Adaptive Antennas and Receivers

Edited By

Melvin M. Weiner

The MITRE Corporation (Retired)
Bedford, Massachusetts, U.S.A.

Taylor & Francis
Taylor & Francis Group
Boca Raton London New York

A CRC title, part of the Taylor & Francis imprint, a member of the
Taylor & Francis Group, the academic division of T&F Informa plc.

Published in 2006 by
CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number-10: 0-8493-3764-X (Hardcover)
International Standard Book Number-13: 978-0-8493-3764-2 (Hardcover)
Library of Congress Card Number 2005043746

Library of Congress Cataloging-in-Publication Data

Adaptive antennas and receivers / edited by Melvin M. Weiner.
p. cm. -- (Electrical engineering and electronics ; 126)
Includes bibliographical references and index.
ISBN 0-8493-3764-X (alk. paper)
1. Adaptive antennas. I. Weiner, Melvin M. II. Series.

TK7871.67.A33A32 2005
621.382'4--dc22 2005043746

Taylor & Francis Group
is the Academic Division of T&F Informa plc.

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

ELECTRICAL AND COMPUTER ENGINEERING

A Series of Reference Books and Textbooks

FOUNDING EDITOR

Marlin O. Thurston
Department of Electrical Engineering
The Ohio State University
Columbus, Ohio

1. Rational Fault Analysis, *edited by Richard Saeks and S. R. Liberty*
2. Nonparametric Methods in Communications, *edited by P. Papantoni-Kazakos and Dimitri Kazakos*
3. Interactive Pattern Recognition, *Yi-tzuu Chien*
4. Solid-State Electronics, *Lawrence E. Murr*
5. Electronic, Magnetic, and Thermal Properties of Solid Materials, *Klaus Schröder*
6. Magnetic-Bubble Memory Technology, *Hsu Chang*
7. Transformer and Inductor Design Handbook, *Colonel Wm. T. McLyman*
8. Electromagnetics: Classical and Modern Theory and Applications, *Samuel Seely and Alexander D. Poularikas*
9. One-Dimensional Digital Signal Processing, *Chi-Tsong Chen*
10. Interconnected Dynamical Systems, *Raymond A. DeCarlo and Richard Saeks*
11. Modern Digital Control Systems, *Raymond G. Jacquot*
12. Hybrid Circuit Design and Manufacture, *Roydn D. Jones*
13. Magnetic Core Selection for Transformers and Inductors: A User's Guide to Practice and Specification, *Colonel Wm. T. McLyman*
14. Static and Rotating Electromagnetic Devices, *Richard H. Engelmann*
15. Energy-Efficient Electric Motors: Selection and Application, *John C. Andreas*
16. Electromagnetic Compossibility, *Heinz M. Schlicke*
17. Electronics: Models, Analysis, and Systems, *James G. Gottling*
18. Digital Filter Design Handbook, *Fred J. Taylor*
19. Multivariable Control: An Introduction, *P. K. Sinha*
20. Flexible Circuits: Design and Applications, *Steve Gurley, with contributions by Carl A. Edstrom, Jr., Ray D. Greenway, and William P. Kelly*

21. Circuit Interruption: Theory and Techniques, *Thomas E. Browne, Jr.*
22. Switch Mode Power Conversion: Basic Theory and Design, *K. Kit Sum*
23. Pattern Recognition: Applications to Large Data-Set Problems, *Sing-Tze Bow*
24. Custom-Specific Integrated Circuits: Design and Fabrication, *Stanley L. Hurst*
25. Digital Circuits: Logic and Design, *Ronald C. Emery*
26. Large-Scale Control Systems: Theories and Techniques, *Magdi S. Mahmoud, Mohamed F. Hassan, and Mohamed G. Darwish*
27. Microprocessor Software Project Management, *Eli T. Fathi and Cedric V. W. Armstrong (Sponsored by Ontario Centre for Microelectronics)*
28. Low Frequency Electromagnetic Design, *Michael P. Perry*
29. Multidimensional Systems: Techniques and Applications, *edited by Spyros G. Tzafestas*
30. AC Motors for High-Performance Applications: Analysis and Control, *Sakae Yamamura*
31. Ceramic Motors for Electronics: Processing, Properties, and Applications, *edited by Relva C. Buchanan*
32. Microcomputer Bus Structures and Bus Interface Design, *Arthur L. Dexter*
33. End User's Guide to Innovative Flexible Circuit Packaging, *Jay J. Miniet*
34. Reliability Engineering for Electronic Design, *Norman B. Fuqua*
35. Design Fundamentals for Low-Voltage Distribution and Control, *Frank W. Kussy and Jack L. Warren*
36. Encapsulation of Electronic Devices and Components, *Edward R. Salmon*
37. Protective Relaying: Principles and Applications, *J. Lewis Blackburn*
38. Testing Active and Passive Electronic Components, *Richard F. Powell*
39. Adaptive Control Systems: Techniques and Applications, *V. V. Chalam*
40. Computer-Aided Analysis of Power Electronic Systems, *Venkatachari Rajagopalan*
41. Integrated Circuit Quality and Reliability, *Eugene R. Hnatek*
42. Systolic Signal Processing Systems, *edited by Earl E. Swartzlander, Jr.*
43. Adaptive Digital Filters and Signal Analysis, *Maurice G. Bellanger*
44. Electronic Ceramics: Properties, Configuration, and Applications, *edited by Lionel M. Levinson*

45. Computer Systems Engineering Management, *Robert S. Alford*
46. Systems Modeling and Computer Simulation, *edited by Naim A. Kheir*
47. Rigid-Flex Printed Wiring Design for Production Readiness, *Walter S. Rigling*
48. Analog Methods for Computer-Aided Circuit Analysis and Diagnosis, *edited by Takao Ozawa*
49. Transformer and Inductor Design Handbook: Second Edition, Revised and Expanded, *Colonel Wm. T. McLyman*
50. Power System Grounding and Transients: An Introduction, *A. P. Sakis Meliopoulos*
51. Signal Processing Handbook, *edited by C. H. Chen*
52. Electronic Product Design for Automated Manufacturing, *H. Richard Stillwell*
53. Dynamic Models and Discrete Event Simulation, *William Delaney and Erminia Vaccari*
54. FET Technology and Application: An Introduction, *Edwin S. Oxner*
55. Digital Speech Processing, Synthesis, and Recognition, *Sadaoki Furui*
56. VLSI RISC Architecture and Organization, *Stephen B. Furber*
57. Surface Mount and Related Technologies, *Gerald Ginsberg*
58. Uninterruptible Power Supplies: Power Conditioners for Critical Equipment, *David C. Griffith*
59. Polyphase Induction Motors: Analysis, Design, and Application, *Paul L. Cochran*
60. Battery Technology Handbook, *edited by H. A. Kiehne*
61. Network Modeling, Simulation, and Analysis, *edited by Ricardo F. Garzia and Mario R. Garzia*
62. Linear Circuits, Systems, and Signal Processing: Advanced Theory and Applications, *edited by Nobuo Nagai*
63. High-Voltage Engineering: Theory and Practice, *edited by M. Khalifa*
64. Large-Scale Systems Control and Decision Making, *edited by Hiroyuki Tamura and Tsuneo Yoshikawa*
65. Industrial Power Distribution and Illuminating Systems, *Kao Chen*
66. Distributed Computer Control for Industrial Automation, *Dobrivoje Popovic and Vijay P. Bhatkar*
67. Computer-Aided Analysis of Active Circuits, *Adrian Ioinovici*
68. Designing with Analog Switches, *Steve Moore*
69. Contamination Effects on Electronic Products, *Carl J. Tautscher*
70. Computer-Operated Systems Control, *Magdi S. Mahmoud*
71. Integrated Microwave Circuits, *edited by Yoshihiro Konishi*

72. Ceramic Materials for Electronics: Processing, Properties, and Applications, Second Edition, Revised and Expanded, *edited by Relva C. Buchanan*
73. Electromagnetic Compatibility: Principles and Applications, *David A. Weston*
74. Intelligent Robotic Systems, *edited by Spyros G. Tzafestas*
75. Switching Phenomena in High-Voltage Circuit Breakers, *edited by Kunio Nakanishi*
76. Advances in Speech Signal Processing, *edited by Sadaoki Furui and M. Mohan Sondhi*
77. Pattern Recognition and Image Preprocessing, *Sing-Tze Bow*
78. Energy-Efficient Electric Motors: Selection and Application, Second Edition, *John C. Andreas*
79. Stochastic Large-Scale Engineering Systems, *edited by Spyros G. Tzafestas and Keigo Watanabe*
80. Two-Dimensional Digital Filters, *Wu-Sheng Lu and Andreas Antoniou*
81. Computer-Aided Analysis and Design of Switch-Mode Power Supplies, *Yim-Shu Lee*
82. Placement and Routing of Electronic Modules, *edited by Michael Pecht*
83. Applied Control: Current Trends and Modern Methodologies, *edited by Spyros G. Tzafestas*
84. Algorithms for Computer-Aided Design of Multivariable Control Systems, *Stanoje Bingulac and Hugh F. VanLandingham*
85. Symmetrical Components for Power Systems Engineering, *J. Lewis Blackburn*
86. Advanced Digital Signal Processing: Theory and Applications, *Glenn Zelniker and Fred J. Taylor*
87. Neural Networks and Simulation Methods, *Jian-Kang Wu*
88. Power Distribution Engineering: Fundamentals and Applications, *James J. Burke*
89. Modern Digital Control Systems: Second Edition, *Raymond G. Jacquot*
90. Adaptive IIR Filtering in Signal Processing and Control, *Phillip A. Regalia*
91. Integrated Circuit Quality and Reliability: Second Edition, Revised and Expanded, *Eugene R. Hnatek*
92. Handbook of Electric Motors, *edited by Richard H. Engelmann and William H. Middendorf*
93. Power-Switching Converters, *Simon S. Ang*
94. Systems Modeling and Computer Simulation: Second Edition, *Naim A. Kheir*
95. EMI Filter Design, *Richard Lee Ozenbaugh*
96. Power Hybrid Circuit Design and Manufacture, *Haim Taraseiskey*

97. Robust Control System Design: Advanced State Space Techniques, *Chia-Chi Tsui*
98. Spatial Electric Load Forecasting, *H. Lee Willis*
99. Permanent Magnet Motor Technology: Design and Applications, *Jacek F. Gieras and Mitchell Wing*
100. High Voltage Circuit Breakers: Design and Applications, *Ruben D. Garzon*
101. Integrating Electrical Heating Elements in Appliance Design, *Thor Hegbom*
102. Magnetic Core Selection for Transformers and Inductors: A User's Guide to Practice and Specification, Second Edition, *Colonel Wm. T. McLyman*
103. Statistical Methods in Control and Signal Processing, *edited by Tohru Katayama and Sueo Sugimoto*
104. Radio Receiver Design, *Robert C. Dixon*
105. Electrical Contacts: Principles and Applications, *edited by Paul G. Slade*
106. Handbook of Electrical Engineering Calculations, *edited by Arun G. Phadke*
107. Reliability Control for Electronic Systems, *Donald J. LaCombe*
108. Embedded Systems Design with 8051 Microcontrollers: Hardware and Software, *Zdravko Karakehayov, Knud Smed Christensen, and Ole Winther*
109. Pilot Protective Relaying, *edited by Walter A. Elmore*
110. High-Voltage Engineering: Theory and Practice, Second Edition, Revised and Expanded, *Mazen Abdel-Salam, Hussein Anis, Ahdab El-Morshedy, and Roshdy Radwan*
111. EMI Filter Design: Second Edition, Revised and Expanded, *Richard Lee Ozenbaugh*
112. Electromagnetic Compatibility: Principles and Applications, Second Edition, Revised and Expanded, *David Weston*
113. Permanent Magnet Motor Technology: Design and Applications, Second Edition, Revised and Expanded, *Jacek F. Gieras and Mitchell Wing*
114. High Voltage Circuit Breakers: Design and Applications, Second Edition, Revised and Expanded, *Ruben D. Garzon*
115. High Reliability Magnetic Devices: Design and Fabrication, *Colonel Wm. T. McLyman*
116. Practical Reliability of Electronic Equipment and Products, *Eugene R. Hnatek*
117. Electromagnetic Modeling by Finite Element Methods, *João Pedro A. Bastos and Nelson Sadowski*
118. Battery Technology Handbook, Second Edition, *edited by H. A. Kiehne*
119. Power Converter Circuits, *William Shepherd and Li Zhang*

120. Handbook of Electric Motors: Second Edition, Revised and Expanded, *edited by Hamid A. Toliyat and Gerald B. Kliman*
121. Transformer and Inductor Design Handbook, *Colonel Wm T. McLyman*
122. Energy Efficient Electric Motors: Selection and Application, Third Edition, Revised and Expanded, *Ali Emadi*
123. Power-Switching Converters, Second Edition, *Simon Ang and Alejandro Oliva*
124. Process Imaging For Automatic Control, *edited by David M. Scott and Hugh McCann*
125. Handbook of Automotive Power Electronics and Motor Drives, *Ali Emadi*
126. Adaptive Antennas and Receivers, *edited by Melvin M. Weiner*

In memory of our parents, Kate and William

Melvin and Donald Weiner

In memory of my wife, Clara

Ronald Fante

Preface

The primary intent of this book is to provide an introduction to state-of-the art research on the modeling, testing, and application of adaptive antennas and receivers. As such, it provides a baseline for engineers, scientists, practitioners, and students in surveillance, communication, navigation, government service, artificial intelligence, computer tomography, neuroscience, and security intrusion industries. This book is based on work performed at Syracuse University and The MITRE Corporation with sponsorship primarily by the U. S. Air Force.

At issue is the detection of target signals in a competing electromagnetic environment which is much larger than the signal *after* conventional signal processing and receiver filtering. The competing electromagnetic environment is *external* system noise (herein designated as "noise") such as clutter residue, interference, atmospheric noise, man-made noise, jammers, external thermal noise (optical systems), *in vivo* surrounding tissue (biological systems), and surrounding material (intrusion detection systems). The environment is statistically characterized by a probability density function (PDF) which may be Gaussian, or more significantly, nonGaussian. For applications with an objective of target detection, the signal is assumed to be from a moving target within the surveillance volume and with a velocity greater than the minimum discernable velocity.

In radars, which look down at the ground to detect targets, the clutter echo power can be 60 to 80 dB larger than the target echo power before signal processing. The target is detected by measuring the difference in returns from one pulse to the next. This method is based on the underlying assumption that the clutter echo power and the radar system are stable between pulses whereas the target signal is not. The degree of stability influences the subclutter visibility (SCV), i.e., the ratio by which the target echo power may be weaker than the coincident clutter echo power and still be detected with specified detection and false-alarm probabilities. The receiving systems of interest comprise an antenna array, digital receiver, signal processor, and threshold detector.[1]

The electromagnetic environment is assumed to be characterized by a "noise" voltage with a PDF that is *temporally* Gaussian but not necessarily *spatially* Gaussian. Conventional signal detection, for a specified false alarm rate or bit error rate, is achieved by measuring the magnitude-squared output of a linear Gaussian receiver compared to a single threshold determined by the variance of the noise voltage averaged over all the cells of the total surveillance volume.

A linear Gaussian receiver is defined as a receiver matched to the frequency spectrum of the signal and assumes a temporally Gaussian PDF "noise" voltage. In conventional signal detection, the probability of signal detection within any cell of the surveillance volume is small if the signal power is small compared to the average noise variance of the total surveillance volume.

This book considers the more general case where the "noise" environment may be spatially nonGaussian. The book is divided into three parts where each part presents a different but sequentially complementary approach for increasing the probability of signal detection within at least some of the cells of the surveillance volume for a nonGaussian or Gaussian "noise" environment. These approaches are: Approach A. *Homogeneous Partitioning of the Surveillance Volume*; Approach B. *Adaptive Antennas*; and Approach C. *Adaptive Receivers*.

Approach A. *Homogeneous Partitioning of the Surveillance Volume*. This approach partitions the surveillance volume into homogeneous, contiguous subdivisions. A homogeneous subdivision is one that can be subdivided into arbitrary subgroups, each of at least 100 contiguous cells, such that all the subgroups contain stochastic spatio-temporal "noise" sharing the same PDF. At least 100 cells/subgroup are necessary for sufficient confidence levels (see Section 4.3). The constant false-alarm rate (CFAR) method reduces to Approach A if the CFAR "reference" cells are within the same homogeneous subdivision as the target cell. When the noise environment is not known *a priori*, then it is necessary to sample the environment, classify and index the homogeneous subdivisions, and exclude those samples that are not homogeneous within a subdivision. If this sampling is not done in a statistically correct manner, then Approach A can yield disappointing results because the estimated PDF is not the actual PDF. Part I *Homogeneous Partitioning of the Surveillance Volume* addresses this issue.

Approach B. *Adaptive Antennas*. This approach, also known as *space-time adaptive processing*, seeks to minimize the competing electromagnetic environment by placing nulls in its principal angle-of-arrival and Doppler frequency (space-time) domains of the surveillance volume. This approach utilizes $k = \mathrm{NM}$ samples of the signals from N subarrays of the antenna over a coherent processing interval containing M pulses to (1) estimate, in the space-time domain, an $NM \times NM$ "noise" covariance matrix of the subarray signals, (2) solve the matrix for up to N unknown "noise" angles of arrival and M unknown "noise" Doppler frequencies, and (3) determine appropriate weighting functions for each subarray that will place nulls in the estimated angle-of-arrival and Doppler frequency domains of the "noise".

Approach B is a form of filtering in those domains. Consequently, the receiver detector threshold can be reduced because the average "noise" voltage variance of the surveillance volume is reduced. The locations and depths of the nulls are determined by the relative locations and strengths of the "noise" sources in the space-time domain and by differences in the actual and estimated "noise" covariance matrices. The results are influenced by the finite number k of

stochastic data samples and the computational efficiency in space-time processing the samples. Part II *Adaptive Antennas* addresses these issues and presents physical models of several applications.

Approach C. *Adaptive Receivers*. For each homogeneous subdivision of the surveillance volume, this approach generally utilizes a nonlinear, nonGaussian receiver whose detection algorithm is matched to the sampled "noise" voltage spatial PDF of that subdivision. When the nonGaussian "noise" waveform is spikier than Gaussian noise, the nonlinear receiver is more effective than a linear receiver in reducing the detecton threshold for a given false alarm rate *provided that the estimated spatial PDF is an accurate representation of the actual PDF*. If the estimated spatial PDF is Gaussian, then the nonlinear receiver reduces to a linear Gaussian receiver. At issue are (1) how to model, simulate, and identify the random processes associated with the *correlated* "noise" data samples and (2) how to determine the nonlinear receiver and its threshold that are best matched to those data samples. Part III *Adaptive Receivers* addresses and illustrates these issues with some applications.

Approach C should not be implemented until Approaches A and B have been implemented. For a prespecified false alarm probability, Approach A or B alone have a better probability of target detection than in their absence. The combination of Approaches A and B has a better probability of target detection than Approach A or B alone. The combination of Approaches A, B, and C has a still better probability of target detection. For this reason, this book often refers to the combination of Approaches A, B, and C as the *weak* signal problem, (*i.e.*, small signal-to-noise ratio case); the combination of Approaches A and B or Approach A or B alone as the *intermediate* signal problem, (*i.e.,* intermediate signal-to-noise ratio case); and the absence of all three approaches as the *strong* signal problem, (*i.e.,* large signal-to-noise ratio case). Approaches A and C are usually more difficult to implement than Approach B alone because "noise" spatial PDF is more difficult to measure than "noise" variance. However, for the weak signal problem, Approaches A and C can be worth the effort as is shown in Part III. All of these approaches have benefited from orders-of-magnitude increases in the speeds of beam scanning and data processing made possible by recent technological advances in integrated circuits and digital computers. However, equally important, are the recent advances in methodology which are reported in this book.

Adaptive antennas originated in the 1950s with classified work by S. Applebaum followed later by P. W. Howells, both of whom published their work about 40 years ago.[2,3] Practical techniques for space-time processing of the sampled data originated with B. Widrow and colleagues approximately a year later.[4]

A nonlinear nonGaussian receiver for weak signal detection, in the presence of "noise" whose PDF is not necessarily Gaussian, originated with D. Middleton approximately 45 years ago.[5] The receiver is designated a "locally optimum detector" (LOD) because, in a Taylor series expansion of the numerator of the likelihood ratio (LR) about a unity operating point, only the second term

(a linear test statistic) is retained and the first term (a constant) is combined as part of the threshold. Thus, for small signal-to-disturbance ratio, a sensitive yet computationally simpler test statistic is obtained, resulting in either a nonlinear receiver for non-Gaussian disturbances or a linear matched filter for Gaussian disturbances with deterministic signals. Unlike an adaptive receiver, Middleton's LOD utilized a fixed detection algorithm and threshold that were determined *a priori* to the detection process.

The feasibility of an adaptive receiver was made possible less than 15 years ago when Aydin Ozturk (Professor, Dept. of Mathematics, Syracuse University) developed an algorithm for identifying and estimating univariate and multivariate distributions based on sample-order statistics.[6,7] At that time, my brother, Donald D. Weiner (Professor, Dept. of Electrical and Computer Engineering, Syracuse University), in collaboration with his doctoral student Muralidhar Rangaswamy and A. Ozturk, conceived the idea of an adaptable receiver which (1) sampled in real time the "noise" environment, (2) utilized the Ozturk algorithm to estimate the "noise" PDF, and (3) utilized the Middleton LOD by matching its detection algorithm and threshold to the estimated "noise" PDF.[8,9]

By 1993, with additional collaboration from Prakash Chakravarthi, Mohamed-Adel Slamani (doctoral students of D. D. Weiner), Hong Wang (Professor, Dept. of Electrical and Computer Engineering, Syracuse University) and Lujing Cai (doctoral student of H. Wang), the core ideas for much of the material in this book had been developed.[10,11] With the exception of Chapters 9 and 10, all of the materials in this book are based on later refinements, elaborations, and applications by D. D. Weiner, his students (Thomas J. Barnard, P. Chakravarthi, Braham Himed, Andrew D. Keckler, James H. Michels, M. Rangaswamy, Rajiv R. Shah, M. A. Slamani, Dennis L. Stadelman), his colleagues at Syracuse University (A. Ozturk, H. Wang), students of H. Wang (L. Cai, Michael C. Wicks), his colleagues at Rome Laboratory (Christopher T. Capraro, Gerard T. Capraro, David Ferris, William J. Baldygo, Vincent Vannicola), his son (William W. Weiner), and Fyzodeen Khan (colleague of T. J. Barnard). Chapter 9 is contributed by George Ploussios (consultant, Cambridge, MA). Chapter 10 consists of reprints of all the refereed journal papers on adaptive antennas individually authored by Ronald L. Fante (Fellow, The MITRE Corporation) or co-authored with colleagues Edward C. Barile, Richard M. Davis, Thomas P. Guella, Jose A. Torres, and John J. Vaccaro.

My interest in the core ideas of this book originated in 1993 from two invited talks at the MITRE Sensor Center.[12,13] The two talks utilized novel mathematical tools (such as the *Ozturk algorithm* and *spherically invariant random vectors*) for more effective implementation of homogeneous partitioning, adaptive antennas, and adaptive receivers. Since that time, the utilization of these tools for those purposes has been reported in the journal literature but not in a book. In July 2003, Marcel Dekker Inc. asked me to recommend a prospective author for a book on *smart* antennas. Smart antennas are nothing more than adaptive antennas (with or without the signal processing associated with adaptive

antennas) which are specifically tailored for the wireless mobile communication industry. Since there were already several books on smart antennas, the publisher agreed instead to accept a proposal from me for the present book.

All of the material for this book is in the public domain. Chapters 1, 7, 9, and 11 were written specifically for this book. The material is from the following sources:

Chap.	Source
1	Contributed by Weiner, M.M.
2	Slamani, M.A., A New Approach to Radar Detection Based on the Partitioning and Statistical Characterization of the Surveillance Volume, University of Massachusetts at Amherst, Final Technical Report, Rome Laboratory, Air Force Material Command, RL-TR-95-164, Vol. 5, Sept. 1995.
3.1	Rangaswamy, M., Michels, J.H., and Himed, B., Statistical analysis of the nonhomogeneity detector for STAP applications, *Digital Signal Processing Journal*, Vol. 14, No. 1, Jan. 2004.
3.2	Rangaswamy, M., Statistical analysis of the nonhomogeneity detector for nongaussian interference backgrounds, *IEEE Trans. Signal Processing*, Vol. 15, Jan./Feb. 2005.
4	Shah, R.R., A New Technique for Distribution Approximation of Random Data, University of Massachusetts at Amherst, Final Technical Report, Rome Laboratory, Air Force Material Command, RL-TR-95-164, Vol. 2, Sept. 1995.
5.1	Ozturk, A., A New Method for Distribution Identification, *J. American Statistical Association*, submitted but not accepted for publication, 1990 (revised 2004). Contributed by A. Ozturk.
5.2	Ozturk, A., A general algorithm for univariate and multivariate goodness-of-fit tests based on graphical representation, *Commun. in Statistics, Part A—Theory and Methods*, Vol. 20, No. 10, pp. 3111–3131, 1991.
6.1	Weiner, W.W., The Ozturk Algorithm: A New Technique for Analyzing Random Data with Applications to the Field of Neuroscience, Math Exam Requirements for Ph.D. in Bioengineering and Neuroscience, Syracuse University, May 9, 1996.
6.2	Slamani, M.A. and Weiner, D.D., Use of image processing to partition a radar surveillance volume into background noise and clutter patches, *Proc. 1993 Conference on Information Sciences and Systems*, Johns Hopkins Univ., Baltimore, Md., March 24–26, 1993.
6.3	Slamani, M.A. and Weiner, D.D., Probabilistic insight into the application of image processing to the mapping of clutter and noise regions in a radar surveillance volume, *Proc. 36th Midwest Symposium of Circuits and Systems*, Detroit, Mi., Aug. 10–18, 1993.

6.4 Slamani, M.A., Ferris, D., and Vannicola, V., A new approach to the analysis of IR images, *Proc. 37th Midwest Symposium in Circuits and Systems*, Lafayette, LA, Aug. 3–5, 1994.

6.5 Slamani, M.A., Weiner, D.D., and Vannicola, V., ASCAPE: An automated approach to the statistical characterization and partitioning of a surveillance volume, *Proc. 6th International Conference on Signal Processing Applications and Technology*, Boston, MA, Oct. 24–26, 1995.

6.6 Keckler, A.D., Stadelman, D.L., Weiner, D.D., and Slamani, M.A., Statistical Characterization of Nonhomogeneous and Nonstationary Backgrounds, Aerosense 1997 Conference on Targets and Backgrounds: Characterization and Representation III, Orlando FL, April 21–24, 1997, *SPIE Proceedings*, Vol. 3062, pp. 31–40, 1997.

6.7 Capraro, C.T., Capraro, G.T., Weiner, D.D., and Wicks, M.C., Knowledge-based map space-time adaptive processing, *Proc. 2001 International Conference on Imaging Science, Systems, and Technology,* Las Vegas, Nev., Vol. 2, pp. 533–538, June 2001.

6.8 Capraro, C.T., Capraro, G.T., Weiner, D.D., Wicks, M.C., and Baldygo, W.J., Improved space-time adaptive processing using knowledge-aided secondary data selection, *Proc. 2004 IEEE Radar Conference*, Philadelphia, PA, April 26–29, 2004.

7 Contributed by Weiner M.M.

8 Cai, L. and Wang, H., Adaptive Implementation of Optimum Space-Time Processing, Chapter 2, Kaman Sciences Corp., Final Technical Report, Rome Laboratory, Air Force Material Command, RL-TR-93-79, May 1993.

9 Contributed by Ploussios, G.

10.1 Fante, R.L., Cancellation of specular and diffuse jammer multipath using a hybrid adaptive array, *IEEE Trans. Aerospace and Electronic Systems*, Vol. 27, No. 5, pp. 823–837, Sept. 1991.

10.2 Barile, E., Fante, R.L., and Torres, J., Some Limitations on the effectiveness of airborne adaptive radar, *IEEE Trans. Aerospace and Electronic Systems*, Vol. 28, No. 4, pp. 1015–1032, Oct. 1992.

10.3 Fante, R.L., Barile, E., and Guilla, T., Clutter covariance smoothing by subaperture averaging, *IEEE Trans. Aerospace and Electronic Systems*, Vol. 30, No. 3, pp. 941–945, July 1994.

10.4 Fante, R.L. and Torres, J., Cancellation of diffuse jammer multipath by an airborne adaptive radar, *IEEE Trans. Aerospace and Electronic Systems*, Vol. 31, No. 2, pp. 805–820, April 1995.

10.5 Fante, R.L., Davis, R.M., and Guella, T.P., Wideband cancellation of multiple mainbeam jammers, *IEEE Trans. Antennas and Propagation*, Vol. 44, No. 10, pp. 1402–1413, Oct. 1996.

10.6 Fante, R.L., Adaptive space-time radar, *J. Franklin Inst.*, Vol. 335B, pp. 1–11, Jan. 1998.

10.7 Fante, R.L., Synthesis of adaptive monopole patterns, *IEEE Trans. Antennas and Propagation*, Vol. 47, No. 5, pp. 773–774, May 1999.

10.8 Fante, R.L., Ground and airborne target detection with bistatic adaptive space-based radar, *IEEE AES Systems Magazine*, pp. 39–44, Oct. 1999.

10.9 Fante, R.L., Adaptive nulling of SAR sidelobe discretes, *IEEE Trans. Aerospace and Electronic systems*, Vol. 35, No. 4, pp. 1212–1218, Oct. 1999.

10.10 Fante, R.L. and Vaccaro, J.J., Wideband cancellation of interference in a GPS receive array, *IEEE Trans. Aerospace and Electronic Systems,* Vol. 36, No. 2, pp. 549–564, April 2000.

10.11 Davis, R.M. and Fante, R.L., A maximum likelihood beamspace processor for improved search and trace, *IEEE Trans. Antennas and Propagation*, Vol. 49, No. 7, pp. 1043–1053, July 2001.

11 Contributed by Weiner, M.M.

12 Rangaswamy, M., Spherically Invariant Random Processes and Radar Clutter Modeling, Simulation, and Distribution Identification, Final Technical Report, Rome Laboratory, Air Force Materiel Command, RL-TR-95-164, Vol. 3, Sept. 1995.

13 Chakravarthi, P., The Problem of Weak Signal Detection, University of Massachusetts at Amhest, Final Technical Report, Rome Laboratory, Air Force Materiel Command, RL-TR-95-164, Vol. 4, Sept. 1995.

14 Barnard, T.J., A Generalization of Spherically Invariant Random Vectors with an Application to Reverberation Reduction in a Correlation Radar, Ph.D. Thesis, Syracuse University, April 1994.

15.1 Barnard, T.J. and Khan, F., Statistical normalization of spherically invariant nonGaussian clutter, *IEEE J. Oceanic Eng.*, 29(2):303–309, April 2004.

15.2 Keckler, A.D., Stadelman, D.L., and Weiner, D.D., NonGaussian clutter modeling and application to radar detection, *Proc. 1997 IEEE National Radar Conference*, Syracuse, N. Y., May 13–15, 1997.

15.3 Stadelman, D.L., Keckler, A.D., and Weiner, D.D., Adaptive Ozturk-based receivers for small signal detection in impulsive nonGaussian clutter, 1999 Conference on Signal and Data Processing of Small Targets, *SPIE Proc.*, Vol. 3809, Denver CO, July 20–22, 1999.

15.4 Stadelman, D.L., Weiner, D.D., and Keckler, A.D., Efficient determination of thresholds *via* importance sampling for monte carlo evaluation of radar performance in nonGaussian clutter, *Proc. 2002 IEEE Radar Conference*, Long Beach, CA, April 22–25, 2002.

15.5 Keckler, A.D. and Weiner, D.D., Generation of rejection method bounds for spherically invariant random vectors, *Proc. 2002 IEEE Radar Conference*, Long Beach, CA, April 22–25, 2002.

15.6 Stadelman, D. and Weiner, D.D., Optimal NonGaussian Processing in Spherically Invariant Interference, Syracuse University, Final Technical Report, Air Force Research Laboratory, AERL-SN-RS-TR-1998–26, March 1998.

15.7 Rangaswamy, M., Michels, J.H., and Weiner, D.D., Multichannel detection for correlated nonGaussian random processes based on innovations, *IEEE Trans. Signal Processing*, Vol. 43, No. 8, pp. 1915–1995, August 1995.

Melvin M. Weiner
Editor

Biography

Melvin M. Weiner received S.B. and S.M. degrees in electrical engineering from Massachusetts Institute of Technology in 1956 (Cooperative Honors Program).

From 1953 to 1956 he was a co-op student at Philco Corp., Philadelphia, working on cathode-ray tubes. From 1956 until the present, he has been engaged in the Boston area in the fields of electromagnetics, physical electronics, optics, lasers, communications, and radar systems. He served as project engineer at Chu Associates (1956 to 1959) on ferrite phase shifters and electronic scanning antennas; senior engineer at EG&G, Inc. (1963 to 1966) on optical detection of high-altitude nuclear explosions; senior staff engineer at Honeywell Radiation Center (1966 to 1967) on infrared surveillance systems and fluxgate magnetometers; staff engineer at AS&E, Inc. (1967 to 1968) on x-ray telescopes, gamma-ray spark chambers, and Cockcroft-Walton voltage multipliers; consulting inventor (1959 to 1971) of the traveling-wave cathode ray tube, nonimaging solar concentrator, first continuous-wave solid-state laser, and first photonic crystal and holey fiber; principal research engineer at AVCO Everett Research Laboratory (1971 to 1978) on optical design and diagnostics of high power CO_2 lasers including co-invention of the off-axis unstable resonator; member of the technical staff at The MITRE Corp. (1978 to 1994) on bistatic radar phenomenology, airborne anti-jam VHF radios, and ground-based over-the-horizon HF radar systems; and author (1994 to present) of two of his three books on electomagnetics, with three additional books in preparation.

He is the author of 36 refereed papers, one book chapter, three books (*Monopole Elements on Circular Ground Planes*, Norwood, MA: Artech House, 1987; *Monopole Antennas*, New York: Marcel Dekker, 2003; *Adaptive Antennas and Receivers*, New York: Marcel Dekker, 2005) and holder of five U. S. patents. Mr. Weiner is a member of the Institute of Electrical and Electronics Engineers, Optical Society of America; Sigma Xi, and Eta Kappa Nu of which he was a national director (1969 to 1971) and founder-chairman of the Motor Vehicle Safety Group (1969 to 1973) contributing to the establishment of the current National Highway Traffic Safety Administration.

Acknowledgments

The 29 contributing authors to this book are from four groups of sources. Chapters 1, 7, and 11 were contributed by the Editor. All the other chapters except Chapter 9 and Chapter 10 were solicited from the Editor's brother D. D. Weiner and D.D. Weiner's doctoral students, colleagues at Syracuse University and U. S. Air Force Rome Laboratory, and son. Chapter 9 was solicited from G. Ploussios, the Editor's former MIT classmate and colleague at Chu Associates and The MITRE Corporation. Chapter 10 was solicited from R. L. Fante, the Editor's former colleague at The MITRE Corporation.

Chapter 2 was supported by U. S. Air Force Rome Laboratory under Contract UM915025/28140. Constructive suggestions and contributions were received from D. D. Weiner, P. Varshney, H. Wang, C. Isik of Syracuse University; L. Slaski, V. Vannicola of Rome Laboratory; V. Lesser of University of Massachusetts; and H. Nawab of Boston University (IPUS Program).

Chapter 3 was supported by the U. S. Air Force Office of Scientific Research under Projects 2304E8, 2304IN, and by in-house research efforts at the U. S. Air Force Research Laboratory.

Chapter 4 was supported by the U. S. Air Force Rome Laboratory under Contract F30602-91-C-0038. Constructive suggestions and contributions were received from D. D. Weiner, L. Slaski, P. Chakravarthi, M. Rangaswamy, and M. Slamani. Support from the Computer Applications and Software Engineering (CASE) Center of Syracuse University is also gratefully acknowledged.

Chapter 5 was supported in part by the National Science Foundation under Grant No. G88135. Support from the Computer Applications and Software Engineering (CASE) Center of Syracuse University is also gratefully acknowledged.

Chapter 6: Section 6.1 received constructive suggestions from R. Smith, D. D. Weiner, S. Chamberlain, E. Relkin, C. Passaglia, M. Slamani, and M. Schecter of Syracuse University. Section 6.2 and Section 6.3 were supported by the U. S. Air Force Rome Laboratory under Contract F30602-91-C-0038. Section 6.4 was supported by the U. S. Air Force Office of Scientific Research under Rome Laboratory Contract AF30602-94-C-0203. Section 6.5 was supported by the U. S. Air Force Rome Laboratory and received constructive suggestions and contributions from J. Michels, M. Rangaswamy, H. Hottelet, C. Zamara, B. Testa, and D. Hui. Section 6.6 was supported in part by the U. S. AFRL Sensors Directorate under Contract F30602-97-C-0065 and received constructive suggestions from G. Genello.

Chapter 8 was supported by the U. S. Air Force Rome Laboratory under Contract F30602-89-C-0082.

Chapter 10: Section 10.1 and Section 10.2 were supported by the U. S. Air Force Electronic Systems Division under Contract F19628-89-C-0001 and received constructive suggestions and contributions from K. Gude, R. Games (Section 10.1); R. Dipietro, B. N. Suresh Babu, T. Guella, D. Lamensdorf, S. Townes (Section 10.2). Sections 10.3 to 10.10 received constructive suggestions and contributions from R. DiPietro, D. Lamensdorf, J. A. Torres (Section 10.3); R. Martin of Westinghouse, E. Barile, D. Lamensdorf, B. N. Suresh Babu, R. DiPietro, T. Guella, R. Williams (Section 10.4); T. Hopkinson, J. Torres, D. Moulin, R. DiPietro, J. Williams, R. Rama Rao (Section 10.10).

Chapter 12 was supported by the U. S. Air Force Rome Laboratory under Contracts SCEE-DACS-P35498-2, UM 91S025/2810, and F30602-91-C-0038. Constructive suggestions and contributions were received from D. D. Weiner, R. Srinivasan, A. Ozturk, L. Slaski, R. Brown, M. Wicks, J. H. Michels, P. Chakravarthi, T. Sarkar, and H. Schwarzlander. Support from the Academic Computing Services of Syracuse University is also gratefully acknowledged.

Chapter 13 was supported by the U. S. Air Force Rome Laboratory under Contracts SCEE-DACS-P35498-2, UM 91S025/28140, and F30602-91-C-0038. Constructive suggestions and contributions were received from D. D. Weiner, R. Srinivasan, A. Ozturk, M. Rangaswamy, and M. Slamani of Syracuse University; and M. Wicks, R. Brown, L. Slaski, and Michels of Rome Laboratory.

Chapter 14 received constructive suggestions and contributions from A. Ozturk, M. Rangaswamy, D. D. Weiner, and the Martin Marietta Corporation.

Chapter 15: Section 15.2 was supported by the U. S. Air Force Rome Laboratory through the Office of Scientific Research Summer Graduate Student Research Program and received constructive suggestions and contributions from J. Michels, M. Rangaswamy, H. Hottelet, C. Zamora, B. Testa, and D. Hui. Section 15.5 was supported in part by the U. S. Air Force Rome Laboratory through the Office of Scientific Research Summer Graduate Student Research Program and received constructive suggestions and contributions from J. Michels, M. Rangaswamy, and D. Stadelman. Section 15.6 was supported by the Advanced Research Projects Agency of the U. S. Dept. of Defense under Contract F30602-94-C-0287 and received constructive suggestions and contributions from H. Hottelet, D. Hui, A. Keckler, E. J. Dudewicz, and F. Sezgin. Section 15.7 was supported by the U. S. Air Force Office of Scientific Research under Project 2304E8 and the U. S. Air Force Rome Laboratory in-house effort under Project 4506, and received constructive suggestions and contributions from R. Vienneau, T. Robbins, R. Srinivasan, and J. Lennon.

Some of the material in this book has been published in refereed journals, conference proceedings, or other public domains protected by copyright. Receipt of waiver of copyright is gratefully acknowledged as follows:

CHAPTER 3

Rangaswamy, M., Michels, J.H., and Himed, B., Statistical analysis of the nonhomogeneity detector for STAP applications, *Digital Signal Processing Journal*, Vol. 14, No. 3, May 2004, pp. 253–267.

Rangaswamy, M., Statistical analysis of the nonhomogeneity detector for nonGaussian interference backgrounds, *IEEE Trans. Signal Processing*, Vol. 53, No. 6, June 2005, pp.2101–2111.

CHAPTER 5

Ozturk, A., A general algorithm for univariate and multivariate goodness-of-fit tests based on graphical representation, *Commun. in Statistics, Part A–Theory and Methods*, Vol. 20, No. 10, 1991, pp. 3111–3131.

CHAPTER 6

Weiner, W.W., The Ozturk Algorithm: A New Technique for Analyzing Random Data with Applications to the Field of Neuroscience, Math Exam Requirements for PhD in Bioengineering and Neuroscience, Syracuse University, May 9, 1996.

Slamani, M.A. and Weiner, D.D., Use of image processing to partition a radar surveillance volume into background noise and clutter patches, *Proc. 1993 Conference on Information Sciences and Systems*, Johns Hopkins University, Baltimore, MD, March 24–26, 1993.

Slamani, M.A. and Weiner, D.D., Probabilistic insight into the application of image processing to the mapping of clutter and noise regions in a radar surveillance volume, *Proc. 36th Midwest Symposium of Circuits and Systems*, Detroit, MI, Aug. 10–18, 1993.

Slamani, M.A., Ferris, D., and Vannicola, V. A new approach to the analysis of IR images, *Proc. 37th Midwest Symposium in Circuits and Systems*, Lafayette, LA, Aug. 3–5, 1994.

Slamani, M.A., Weiner, D.D., and Vannicola, V., ASCAPE: An automated approach to the statistical characterization and partitioning of a surveillance volume, *Proc. 6th International Conference on Signal Processing Applications and Technology*, Boston, MA, Oct. 24–26, 1995.

Keckler, A.D., Stadelman, D.L., Weiner, D.D., and Slamani, M.A., Statistical Characterization of Nonhomogeneous and Nonstationary Backgrounds, Aerosense 1997 Conference on Targets and Backgrounds: Characterization and Representation III, *SPIE Proceedings*, Vol. 3068, Orlando, FL, April 21–24, 1997.

Capraro, C.T., Capraro, G.T., Weiner, D.D., and Wicks, M.C., Knowledge-based map space-time adaptive processing, *Proc. International Conference on Imaging Science, Systems, and Technology*, Las Vegas, Nev., June 2001, Vol. 2, pp. 533–538.

Capraro, C.T., Capraro, G.T., Weiner, D.D., and Wicks, M.C., Improved space-time adaptive processing using knowledge-aided secondary data selection, *Proc. 2004 IEEE Radar Conference*, Philadelphia, April 26–29, 2004.

CHAPTER 10

Fante, R.L., Cancellation of specular and diffuse jammer multipath using a hybrid adaptive array, *IEEE Trans. Aerospace and Electronic Systems*, Vol. 27, No. 5, Sept. 1991, pp. 823–837.

Barile, E., Fante, R.L., and Torres, J. Some limitations on the effectiveness of airborne adaptive radar, *IEEE Trans. Aerospace and Electronic Systems*, Vol. 28, No. 4, Oct. 1992, pp. 1015–1032.

Fante, R.L., Barile, E., and Guilla, T., Clutter covariance smoothing by sub-aperture averaging, *IEEE Trans. Aerospace and Electronic Systems*, Vol. 30, No. 3, July 1994, pp. 941–945.

Fante, R.L. and Torres, J., Cancellation of diffuse jammer multipath by an airborne adaptive radar, *IEEE Trans. Aerospace and Electronic Systems*, Vol. 31, No. 2, April 1995, pp. 805–820.

Fante, R.L., Davis, R.M., and Guella, T.P., Wideband cancellation of multiple mainbeam jammers, *IEEE Trans. Antennas and Propagation*, Vol. 44, No. 10, Oct. 1996, pp. 1402–1413.

Fante, R.L., Adaptive space-time radar, *J. Franklin Inst.*, Vol. 335B, Jan. 1998, pp. 1–11.

Fante, R.L., Synthesis of adaptive monopole patterns, *IEEE Trans. Antennas and Propagation*, Vol. 47, No. 5, May 1999, pp. 773–774.

Fante, R.L., Ground and airborne target detection with bistatic adaptive space-based radar, *IEEE AES Systems Magazine*, Oct. 1999, pp. 39–44.

Fante, R.L., Adaptive nulling of SAR sidelobe discretes, *IEEE Trans. Aerospace and Electronic Systems,* Vol. 35, No. 4, Oct. 1999, pp. 1212–1218.

Fante, R.L. and Vaccaro, J.J., Wideband cancellation of interference in a GPS receive array, *IEEE Trans. Aerospace and Electronic Systems*, Vol. 36, No. 2, April 2000, pp. 549–564.

Davis, R.M. and Fante, R.L., A maximum likelihood beamspace processor for improved search and trace, *IEEE Trans. Antennas & Propagation*, Vol. 49, No. 7, July 2001, pp. 1043–1053.

CHAPTER 15

Barnard, T.J. and Khan, F., Statistical normalization of spherically invariant nonGaussian clutter, *IEEE J. Oceanic Eng.*, Vol. 29, No. 2, April 2004, pp. 303–309.

Keckler, A.D., Stadelman, D.L., and Weiner, D.D., NonGaussian clutter modeling and application to radar target detection, *Proc. 1997 IEEE National Radar Conference*, Syracuse, N Y., May 13–15, 1997.

Stadelman, D.L., Keckler, A.D., and Weiner, D.D., Adaptive Ozturk-based receivers for small signal detection in impulsive nonGaussian clutter, 1999 Conference on signal and Data Processing of Small Targets, *SPIE Proceedings,* Vol. 3809, Denver, CO, July 20–22, 1999.

Stadelman, D.L., Weiner, D.D., and Keckler, A.D., Efficient determination of thresholds *via* importance sampling for monte carlo evaluation of radar performance in nonGaussian clutter, *Proc. 2002 IEEE Radar Conference*, Long Beach, CA, April 22–25, 2002.

Keckler, A.D. and Weiner, D.D. Generation of rejection method bounds for spherically invariant random vectors, *Proc. 2002 IEEE Radar Conference*, Long Beach, CA, April 22–25, 2002.

Rangaswamy, M., Michels, J.H., and Weiner, D.D., Multichannel detection for correlated nonGaussian random processes based on innovations, *IEEE Trans. Signal Processing*, Vol. 43, No. 8, August 1995, pp. 1915–1995.

The acquisition and production stages of this book were skillfully guided by B. J. Clark of Marcel Dekker and Nora Konopka, Theresa Delforn, and Gerry Jaffe of Taylor & Francis, respectively. Layout was magnificently managed by Carol Cooper of Alden Prepress Services UK.

Contributors

W. J. Baldygo
U.S. Air Force Research Lab
Rome, New York

E. C. Barile
Raytheon Missile Defense Center
Woburn, Massachusetts

T. J. Barnard
Lockheed Martin Maritime Systems
Syracuse, New York

L. Cai
Globespan Semiconductor, Inc.
Redbank, New Jersey

C. T. Capraro
Capraro Technologies
Utica, New York

G. T. Capraro
Capraro Technologies
Utica, New York

P. Chakravarthi
Eka Systems
Germantown, Maryland

R. M. Davis
The MITRE Corp.
Bedford, Massachusetts

R. L. Fante
The MITRE Corp.
Bedford, Massachusetts

D. Ferris
U.S. Air Force Research Lab
Rome, New York

T. P. Guella
The MITRE Corp.
Bedford, Massachusetts

B. Himed
U.S. Air Force Research Lab
Rome, New York

A. D. Keckler
Sensis Corp.
DeWitt, New York

F. Khan
Naval Undersea Warfare Center
Newport, Rhode Island

J. H. Michels
U.S. Air Force Research Lab
Rome, New York

A. Ozturk
Ege University,
Bornova, Izmir, Turkey

G. Ploussios
Consultant
Boston, Massachusetts

M. Rangaswamy
U.S. Air Force Research Lab
Hanscom AFB, Massachusetts

R. R. Shah
Juhu S.,
India

M. A. Slamani
ITT Industries
Alexandria, Virginia

D. L. Stadelman
Syracuse Research Corp.
Syracuse, New York

J. A. Torres
The MITRE Corp.
Bedford, Massachusetts

J. J. Vaccaro
The MITRE Corp.
Bedford, Massachusetts

V. Vannicola
Research Associates for Defense Conversion
Rome, New York

H. Wang
Syracuse University
Syracuse, New York

D. D. Weiner
Syracuse University (retired)
Syracuse, New York

M. M. Weiner
The MITRE Corp. (retired)
Bedford, Massachusetts

W. W. Weiner
Rose-Hulman Institute of Technology
Terra Haute, Indiana

M. C. Wicks
U.S. Air Force Research Lab
Rome, New York

Table of Contents

Part I
Homogeneous Partitioning of the Surveillance Volume

Chapter 1 Introduction 3

M. M. Weiner

Chapter 2 A New Approach to Radar Detection Based on the Partitioning and Statistical Characterization of the Surveillance Volume 5

M. A. Slamani

Chapter 3 Statistical Analysis of the Nonhomogeneity Detector (for Excluding Nonhomogeneous Samples from a Subdivision) 175

M. Rangaswamy, J. H. Michels, and B. Himed

Chapter 4 A New Technique for Univariate Distribution Approximation of Random Data 205

R. R. Shah

Chapter 5 Probability Density Distribution Approximation and Goodness-of-Fit Tests of Random Data 259

A. Ozturk

Chapter 6 Applications 295

W. J. Baldygo, C. T. Capraro, G. T. Capraro, D. Ferris, A. D. Keckler, M. A. Slamani, D. L. Stadelman, V. Vannicola, D. D. Weiner, W. W. Weiner, and M. C. Wicks

Part II
Adaptive Antennas

Chapter 7 Introduction 419

M. M. Weiner

Chapter 8 Adaptive Implementation of Optimum Space–Time Processing 421

L. Cai and H. Wang

Chapter 9 A Printed-Circuit Smart Antenna with Hemispherical Coverage for High Data-Rate Wireless Systems 439

G. Ploussios

Chapter 10 Applications 443

E. C. Barile, R. M. Davis, R. L. Fante, T. P. Guella, J. A. Torres, and J. J. Vaccaro

Part III
Adaptive Receivers

Chapter 11 Introduction 603

M. M. Weiner

Chapter 12 Spherically Invariant Random Processes for Radar Clutter Modeling, Simulation, and Distribution Identification 605

M. Rangaswamy

Chapter 13 Weak Signal Detection 707

P. Chakravarthi

Chapter 14 A Generalization of Spherically Invariant Random Vectors (SIRVs) with an Application to Reverberation Reduction in a Correlation Sonar 799

T. J. Barnard

Chapter 15 Applications....913

T. J. Barnard, A. D. Keckler, F. Khan, J. H. Michels, M. Rangaswamy, D. L. Stadelman, and D. D. Weiner

Appendices....1039

Acronyms....1117

References....1133

Computer Programs available at CRC Press Website....1169

Part I

Homogeneous Partitioning of the Surveillance Volume

1 Introduction

M. M. Weiner

Part I *Homogeneous Partitioning of the Surveillance Volume* discusses the implementation of the *first* of three sequentially complementary approaches for increasing the probability of target detection within at least some of the cells of the surveillance volume for a spatially nonGaussian or Gaussian "noise" environment that is temporally Gaussian. This approach, identified in the Preface as *Approach A*, partitions the surveillance volume into homogeneous contiguous subdivisions. A homogeneous subdivision is one that can be subdivided into arbitrary subgroups, each of at least 100 contiguous cells, such that all the subgroups contain stochastic spatio-temporal "noise" sharing the same probability density function (PDF). At least one hundred cells per subgroup are necessary for sufficient confidence levels (see Section 4.3). The constant false-alarm rate (CFAR) method reduces to Approach A if the CFAR "reference" cells are within the same homogeneous subdivision as the target cell. When the noise environment is not known *a priori*, then it is necessary to sample the environment, classify and index the homogeneous subdivisions, and exclude samples that are not homogeneous within a subdivision. If this sampling is not done in a statistically correct manner, then Approach A can yield disappointing results because the estimated PDF is not the actual PDF. Part I addresses this issue.

Chapter 2 discusses the implementation of Approach A to the radar detection problem. In Section 2.1, the simplest but least versatile implementation is discussed for utilization when statistical knowledge of the environment is known *a priori*. Section 2.2 discusses a feedforward expert system for implementation when the statistical environment is *not* known *a priori* but must be estimated from data samples in real time. Section 2.3 introduces a *feedback* expert system Integrated Processing and Understanding of Signals (IPUS) that augments the feedforward system of Section 2.2 by assessing whether correct signal processing and understanding have taken place and then performs additional data sampling and signal processing if required. Section 2.4 discusses the application of a feedback expert system to radar signal processing. The issues associated with clutter-patch mapping (Section 2.5) and indexing (Section 2.6) with a *feedforward* expert system are implemented by IPUS for a *feedback* expert

system in Section 2.7. Conclusions and suggestions for future research are presented in Section 2.8.

Chapter 3 analyzes the integrity of a nonhomogeneous detector (NHD) for excluding nonhomogeneous samples from a candidate subdivision. The cases of Gaussian and nonGaussian interference environments are discussed in Section 3.1 and Section 3.2, respectively.

Given a finite number of correlated samples that are realizations of a stochastic process, as in Section 2.2 to Section 2.7 and Chapter 3, how does one determine the best-fit approximation to the PDF of those samples? Chapter 4 discusses a new technique, the Ozturk algorithm, for achieving this difficult task. After a review of the literature (Section 4.1), the Ozturk algorithm is summarized (Section 4.2), and then evaluated by simulation results (Section 4.3). Conclusions and suggestions for future work are given in Section 4.4.

A more complete discussion of the Ozturk algorithm is given in Chapter 5 by its originator, Aydin Ozturk.

Chapter 6 presents applications of homogeneous partitioning to neuroscience (Section 6.1), radar detection (Section 6.2, Section 6.3, Section 6.7, and Section 6.8), infra-red image processing (Section 6.4 and Section 6.5), and concealed weapon detection (Section 6.6). Section 6.2 and Section 6.3 summarize an image processing mapping procedure, previously discussed in Section 2.4 to Section 2.7, for distinguishing patches dominated by background noise from those dominated by clutter. Section 6.5 presents a formalized process Automatic Statistical Characterization and Partitioning of Environments (ASCAPE) for that purpose. Section 6.7 and Section 6.8 utilize *a priori* knowledge-based terrain maps to achieve homogeneous partitioning.

2 A New Approach to Radar Detection Based on the Partitioning and Statistical Characterization of the Surveillance Volume

M. A. Slamani

CONTENTS

2.0. Introduction 7
2.1. Radar Detection with *a Priori* Statistical Knowledge of the Environment 8
2.1.1. Introduction 8
2.1.2. SIRV 10
2.1.2.1. Definitions 10
2.1.2.2. Properties of SIRVs 11
2.1.3. Locally Optimum Detector 12
2.2. Understanding of Signal and Detection Using a Feedforward Expert System 13
2.2.1. Introduction 13
2.2.2. Classification of the Test Cells 14
2.2.2.1. Mapping of the Space 14
2.2.2.2. Indexing of the Cells 15
2.2.3. Target Detection 16
2.3. Signal Understanding and Detection Using a Feedback Expert System 20
2.3.1. Introduction 20
2.3.2. IPUS Architecture 20
2.3.2.1. Introduction 20

2.3.2.2. Discrepancy Detection 24
2.3.2.3. Diagnosis and Reprocessing 25
2.3.2.4. Interpretation Process 26
2.3.2.5. SOU and Resolving Control Structure 26
2.3.3. Application of IPUS to Radar Signal Understanding 29
2.4. Proposed Radar Signal Processing System Using a Feedback Expert System 29
2.4.1. Data Collection and Preprocessing 29
2.4.2. Mapping 32
2.5. Mapping Procedure 36
2.5.1. Introduction 36
2.5.2. Observations on BN and CL Cells 37
2.5.2.1. Observations on BN Cells 37
2.5.2.2. Observations on CL Cells 38
2.5.3. Mapping Procedure 39
2.5.3.1. Separation of CL Patches from Background Noise 39
2.5.3.2. Detection of CL Patch Edges and Edge Enhancement 44
2.5.3.3. Conclusion 46
2.5.4. Examples of the Mapping Procedure 47
2.5.4.1. Introduction 47
2.5.4.2. Examples 49
2.5.5. Convergence of the Mapping Procedure 70
2.5.5.1. Introduction 70
2.5.5.2. Separation between BN and CL Patches 73
2.5.6. Extension of the Mapping Procedure to Range–Azimuth–Doppler Cells 79
2.5.7. Conclusion 81
2.6. Indexing Procedure 82
2.6.1. Introduction 82
2.6.2. Assessment Stage 83
2.6.2.1. Identification of the BN and CL Patches 83
2.6.2.2. Computation of CL-to-Noise Ratios 85
2.6.2.3. Classification of CL Patches 85
2.6.3. CL Subpatch Investigation Stage 86
2.6.4. PDF Approximation of WSC CL Patches 87
2.6.4.1. Test Cell Selection 88
2.6.4.2. PDF Approximation 89
2.6.4.3. PDF Approximation Metric 91
2.6.4.4. Outliers 93
2.6.4.5. PDF Approximation Strategy 96
2.6.5. Examples 97
2.6.5.1. Example 1 97
2.6.5.2. Example 2 104

2.6.5.3. Example 3 106
2.6.6. Extension of the Indexing Procedure to Range–Azimuth–Doppler Cells 111
2.6.7. Conclusion 113
2.7. Application of IPUS to the Radar Detection Problem 114
2.7.1. Summary of IPUS Concepts 114
2.7.2. Role of IPUS in the Mapping Procedure 115
2.7.2.1. IPUS Stages Included in the Mapping Procedure 115
2.7.2.2. Observations on the Setting of NCC 117
2.7.3. Examples of Mapping 125
2.7.3.1. Example 1 125
2.7.3.2. Example 2 125
2.7.3.3. Example 3 125
2.7.4. Role of IPUS in the Indexing Procedure 126
2.7.4.1. IPUS Stages Included in the Assessment Stage 127
2.7.4.2. IPUS Stages Included in the CL Subpatch Investigation Stage 127
2.7.4.3. Examples 131
2.7.4.4. IPUS Stages Included in the PDF Approximation Stage 133
2.7.5. Examples of Indexing 147
2.7.5.1. Example 1 148
2.7.5.2. Example 2 152
2.7.5.3. Example 3 163
2.7.6. Conclusion 170
2.8. Conclusion and Future Research 172
2.8.1. Conclusion 172
2.8.2. Future Research 173

2.0. INTRODUCTION

In signal processing applications it is common to assume a Gaussian process in the design of optimal signal processors. However, non-Gaussian processes do arise in many situations. For example, measurements reveal that radar clutter may be approximated by either Weibull, K-distributed, Lognormal, or Gaussian distributions depending upon the scenario.[4–10] When the possibility of a non-Gaussian problem is encountered, the question, as to which probability distributions should be utilized in a specific situation for modeling the data, needs to be answerd.

In practice, the underlying probability distributions are not known a priori. Consequently, an assessment must be made by monitoring the environment. Another consideration is that radar detection problems can usually be divided into strong, intermediate, and weak signal cases. Hence, the system that monitors a radar environment must be able to subdivide the surveillance volume into background noise and clutter patches in addition to approximating the underlying

probability distributions for each patch. This is in contrast to current practice where a single robust detector, usually based on the Gaussian assumption, is employed.

The objective of this work is to develop techniques that monitor the environment and select appropriate detector for processing the data.

The main contributions are: (1) an image processing technique is devised which enables partitioning of the surveillance volume into background noise and clutter patches, (2) a new algorithm, developed by Dr. Ozturk while he was a Visiting Professor at Syracuse University,[27–29] is used to identify suitable approximations to the probability density function for each clutter patch, and (3) rules to be used with the expert system, Integrated Processing and Understanding of Signals (IPUS),[20–22] are formulated for monitoring the environment and selecting the appropriate detector for processing the data.

This dissertation is organized as follows: Section 2.1 discusses some of the difficulties that arise in the classical radar detection problem. Their solution is proposed in Section 2.2 which uses an expert system with feed-forward processing. In Section 2.3 an improved solution is presented using feed-back processing. The general radar detection problem is described in Section 2.4 and a mapping procedure is introduced to separate between background noise and cluter patches. In Section 2.5 an image processing technique is developed for the mapping procedure. Next, an indexing procedure is developed in Section 2.6 to enable the invetigation of clutter subpathces and the approximation of probability distributions for each clutter patch. Finally, expert system rules are developed in Section 2.7 to enable the system to control both the mapping and indexing stages. Conclusions and suggestions for future research are given in Section 2.8.

2.1. RADAR DETECTION WITH *A PRIORI* STATISTICAL KNOWLEDGE OF THE ENVIRONMENT

2.1.1. Introduction

The optimal radar detection problem consists of collecting a set of N samples $(r_0, r_1,\ldots, r_{N-1})$ from a given cell in space, processing the data by a Neyman–Pearson receiver which takes the form of a likelihood ratio test (LRT)[1] and deciding for that cell whether or not a target is present. Let $\mathbf{r}$ denote the vector formed by N samples, $\mathbf{r} = (r_0, r_1,\ldots, r_{N-1})^{\mathrm{T}}$, where T denotes "transpose" and the samples are realizations of the random variables R_o, $R_1 \cdots R_{N-1}$, respectively. The LRT compares a statistic λ to a fixed threshold η. The statistic λ is the ratio between the joint probability density function (PDF), $p_{\mathbf{R}}(\mathbf{r}|H_1)$, of the N samples given that a target is present and the joint PDF, $p_{\mathbf{R}}(\mathbf{r}|H_0)$ of N samples, given that no target is present. H_1 and H_0 denote the hypotheses that a target is present and absent, respectively. This ratio is called LR. The threshold η is determined by constraining the probability of false alarm (PFA) to a specified value.

The binary hypotheses (H_1, H_0) are defined in a way such that, under hypothesis H_1, the kth collected sample, r_k, $k = 0,1,\ldots, N-1$, is composed of a target signal sample, s_k, plus an additive disturbance sample, d_k. Under hypothesis H_0, the kth sample, r_k (where $k = 0, 1,\ldots, N-1$), consists of the disturbance sample d_k. Hence,

$$r_k = \begin{cases} s_k + d_k; & H_1 \\ d_k; & H_0 \end{cases} \qquad k = 0, 1, \ldots, N-1. \tag{2.1}$$

In general, the disturbance sample d_k consists of a clutter (CL) sample c_k, plus a BN sample n_k.

The LRT then takes the form

$$\lambda = \frac{p_{\mathbf{R}}((\mathbf{r})|H_1)}{p_{\mathbf{R}}((\mathbf{r})|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{2.2}$$

For $\lambda > \eta$, H_1 is decided otherwise, H_0 is decided.

Assuming that the samples are statistically independent, the joint PDF $p_{\mathbf{R}}((\mathbf{r})|H_i)$; $i = 0, 1$, is nothing but the product of the N marginal PDFs of the samples. Specifically,

$$p_{\mathbf{R}}((\mathbf{r})|H_i) = \prod_{k=0}^{N-1} p_{R_k}(r_k|H_i); \qquad i = 0, 1 \tag{2.3}$$

The LRT is then readily implemented provided the marginal PDFs are known.

In practice, the real data may be correlated in time, making the statistical independence assumption invalid. Unless the joint PDFs of the correlated samples are assumed to be Gaussian, it is not commonly known how to specify the joint PDFs $p_{\mathbf{R}}((\mathbf{r})|H_i)$; $i = 0, 1$. Many engineers invoke the Gaussian assumption even when it is known to be not applicable. It is for this reason the most of the radars today are Gaussian receivers (i.e., these process data using LRT based on the joint Gaussian PDF). When the target signal, s_k, cannot be filtered from the disturbance, d_k, by means of spatial or temporal processing and d_k is much larger than s_k (where $k = 0, 1,\ldots, N-1$) then r_k approximately equals d_k under hypotheses and high precision is needed to evaluate the LRT because $p_{\mathbf{R}}(\mathbf{r}|H_1)$ becomes approximately equal to $p_{\mathbf{R}}(\mathbf{r}|H_0)$. Specifically,

$$\lambda = \frac{p_{\mathbf{R}}(\mathbf{r}|H_1)}{p_{\mathbf{R}}(\mathbf{r}|H_0)} \approx 1 \tag{2.4}$$

and the statistic λ becomes insensitive to the received data.

Recent work reported by Rangaswamy et al.[2] shows that, it is possible to model N correlated nonGaussian random samples as samples, from a spherically invariant random process (SIRP). The vector $\mathbf{R}$ of the N samples is said to be a spherically invariant random vector (SIRV). More details about SIRVs are presented in Section 2.1.2.

In addition, the work done by Chakravarthi et al.[3] shows for the nonGaussian weak signal case (WSC) (i.e., the average power of s_k is much less than that of d_k, $k = 0, 1,\ldots, N-1$) that the use of a locally optimum detector (LOD) provides improved performance over direct application of the Gaussian LRT. LODs are introduced in Section 2.1.3.

2.1.2. SIRV

Many investigators[4–7] have reported experimental measurements for which the CL PDF has an extended tail. The extended tail gives rise to spiky CL and relatively large probabilities of false alarm. The Gaussian model for the CL fails to predict this behavior. NonGaussian models for the univariate (marginal) CL PDF have been proposed. Commonly reported marginal non-Gaussian PDFs for the CL are Weibull,[4] log-normal[8,9] and *K*-distributed.[5,6,10]

Usually, radars process *N* pulses at a time. A complete statistical characterization of the CL requires the specification of the joint PDF of the *N* samples. The theory of SIRPs provides a powerful mechanism for obtaining the joint PDF of the *N* correlated non-Gaussian random variables.

2.1.2.1. Definitions

An SIRV[2,11] is a random vector PDF of which is uniquely determined by the specification of a mean vector, a covariance matrix and a characteristic first order PDF. The PDF of a random vector is defined to be the joint PDF of the components of the random vector.

An SIRP is a random process such that every random vector obtained by sampling this process is an SIRV.

The work of Kingman[12] and Yao[13] gave rise to a representation theorem, valid for all SIRVs which is summarized below. Let $\mathbf{z} = [z_1, z_2,\ldots, z_N]^{\mathrm{T}}$ denote the realization of a real zero mean Gaussian random vector $\boldsymbol{Z}$ with covariance matrix M. Let s denote the realization of a nonnegative random variable $\boldsymbol{S}$ with PDF $f_S(s)$. It is assumed that S is independent of $\mathbf{Z}$. The representation theorem states that $\mathbf{X}$ is an SIRV *if and only if* it can be expressed in the form $\mathbf{X} = \mathbf{Z}S$. In particular the PDF of $\mathbf{X}$ is given by

$$f_X(\mathbf{x}) = (2\pi)^{-N/2}|\mathbf{M}|^{-1/2}\int_0^\infty s^{-N}\exp\left(-\frac{\mathbf{X}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{X}}{2s^2}\right)f_S(s)\mathrm{d}s \tag{2.5}$$

where: $|\mathrm{M}|$ denotes the determinant of the covariance matrix M. Let $p = \mathbf{X}^{\mathrm{T}}\mathrm{M}^{-1}\mathbf{X}$. Then Equation 2.5 can be written as

$$f_X(\mathbf{x}) = (2\pi)^{-N/2}|\mathrm{M}|^{-1/2}h_N(p) \tag{2.6}$$

where:

$$h_N(p) = \int_0^\infty s^{-N}\exp\left(-\frac{p}{2s^2}\right)f_S(s)\mathrm{d}s \tag{2.7}$$

Note that $f_S(s)$ is defined to be the characteristic PDF of the SIRV.

2.1.2.2. Properties of SIRVs

Several properties of SIRVs[2,14–16] are stated below:

- It has been pointed out in Ref. 15 that when $f_S(s) = \delta(s - 1)$, where $\delta(\cdot)$ is the unit impulse function, the resulting $h_N(P)$ is the familiar multivariate Gaussian PDF.
- Differentiation of Equation 2.7 with respect to the argument gives[16]

$$h_{N+2}(w) = -2\frac{\mathrm{d}h_N(w)}{\mathrm{d}w} \tag{2.8}$$

In Equation 2.8 the argument has been replaced by w because the quadratic form p depends on N and therefore cannot be used in the equality. Equation 2.8 provides a mechanism to relate the functional forms of the higher order PDFs with those of the lower orders. Starting with $N = 1$ and using Equation 2.8 repeatedly gives

$$h_{2N+1}(w) = (-2)^N \frac{\mathrm{d}^N h_1(w)}{\mathrm{d}w^N} \tag{2.9}$$

Starting with $N = 2$,

$$h_{2N+2}(w) = (-2)^N \frac{\mathrm{d}^N h_2(w)}{\mathrm{d}w^N} \tag{2.10}$$

It is therefore possible to construct all higher order PDFs for odd values of N and even values of N, starting from $h_1(w)$ and $h_2(w)$, respectively.

- The PDF of an SIRV is a function of a nonnegative quadratic form. The PDF is uniquely determined by specification of a mean vector, a covariance matrix and characteristic first-order PDF.
- An SIRV is invariant under a linear transformation. More precisely, if $\mathbf{X}$ is an SIRV with characteristic PDF $f_S(s)$, then $\mathbf{Y} = \mathrm{A}\mathbf{X} + \mathbf{B}$ is an SIRV with the same characteristic PDF as $\mathbf{X}$, where A is a matrix and $\mathbf{B}$ is a vector.
- $h_1(p)$ is related to the marginal (first order) PDF of X_k. This can be seen easily by letting $N = 1$ in Equation 2.6. More precisely,

$$\mathrm{f}_{X_k}(x_k) = \frac{1}{\sqrt{2\pi}\sigma_k} h_1(p); \quad k = 0, 1, 2, \ldots, N-1 \tag{2.11}$$

  where $p = \mathbf{x}_k^2/\sigma_k^2$ and σ_k^2 denotes the variance of X_k.
- Two of the possible techniques for obtaining the PDF of an SIRV are:
  (1) If the characteristic PDF of the SIRV is readily available, then evaluate $h_N(p)$ directly from Equation 2.7. Complete the characterization by specifying the mean vector and covariance matrix of the SIRV.

(2) If the characteristic PDF of the SIRV is not readily available, but the marginal PDF is known, first obtain $h_1(p)$ from the marginal PDF using Equation 2.11. Then obtain $f_S(s)$ by solving the integral equation in Equation 2.7 with $N = 1$. Next use $f_S(s)$ in Equation 2.7 for desired N to obtain $h_N(p)$. Finally, specify the mean vector and covariance matrix of the SIRV.

In the work that follows SIRPs will be used to model correlated non-Gaussian disturbances (i.e., CL plus noise).

2.1.3. Locally Optimum Detector

Assume that the received target signals have a very small unknown amplitude θ, so that

$$\begin{aligned} &\text{Under } H_1 \quad \mathbf{r} = \theta\mathbf{s} + \mathbf{d} \\ &\text{Under } H_0 \quad \mathbf{r} = \mathbf{d} \end{aligned} \tag{2.12}$$

A Taylor series expansion of the numerator of the LR results in[17]

$$\frac{p_{\mathrm{R}}(\mathbf{r}|H_1)}{p_{\mathrm{R}}(\mathbf{r}|H_0)} = 1 + \theta \frac{\dfrac{\mathrm{d}}{\mathrm{d}\theta} p_{\mathrm{R}}(\mathbf{r}|H_1)\Big|_{\theta=0}}{p_{\mathrm{R}}(\mathbf{r}|H_0)} + \frac{\theta^2}{2} \frac{\dfrac{\mathrm{d}^2}{\mathrm{d}\theta^2} p_{\mathrm{R}}(\mathbf{r}|H_1)\Big|_{\theta=0}}{p_{\mathrm{R}}(\mathbf{r}|H_0)} + \ldots \tag{2.13}$$

The LRT consists of comparing the LR to a threshold η. For the case where θ is very small, it is assumed that the terms involving θ^2, θ^3, … are negligible with high probability. Ignoring these terms, the LRT simplifies to

$$1 + \theta \frac{\dfrac{\mathrm{d}}{\mathrm{d}\theta} p_{\mathrm{R}}(\mathbf{r}|H_1)\Big|_{\theta=0}}{p_{\mathrm{R}}(\mathbf{r}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{2.14}$$

$$\nu = \frac{\dfrac{\mathrm{d}}{\mathrm{d}\theta} p_{\mathrm{R}}(\mathbf{r}|H_1)\Big|_{\theta=0}}{p_{\mathrm{R}}(\mathbf{r}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\eta - 1}{\theta} = \gamma \tag{2.15}$$

The receiver which performs the above test is referred to as the LOD. Because the probability of detection (PD) and the PFA are given by

$$\mathrm{PD} = \int_{\gamma}^{\infty} p_{\nu}(\nu|H_1)\mathrm{d}\nu \tag{2.16}$$

$$\mathrm{PFA} = \int_{\gamma}^{\infty} p_{\nu}(\nu|H_0)\mathrm{d}\nu \tag{2.17}$$

and because $p_{\nu}(\nu|\mathrm{H}_1)$ approaches $p_{\nu}(\nu|H_0)$ as θ goes to zero, it follows that PD = PFA when $\theta = 0$. The power function of a receiver is defined as the curve relating PD and θ. Under a fixed PFA constraint, a typical power function curve is shown in Figure 2.1 for an LOD.

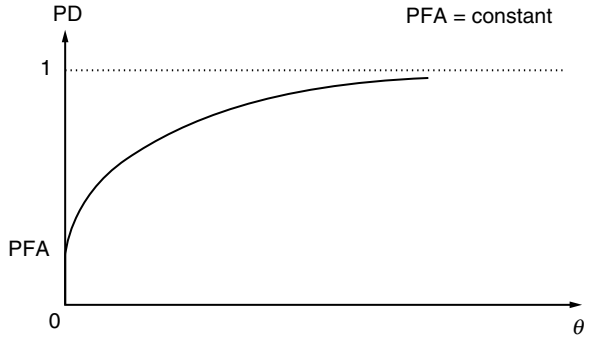

FIGURE 2.1 Power function for a locally optimum detector.

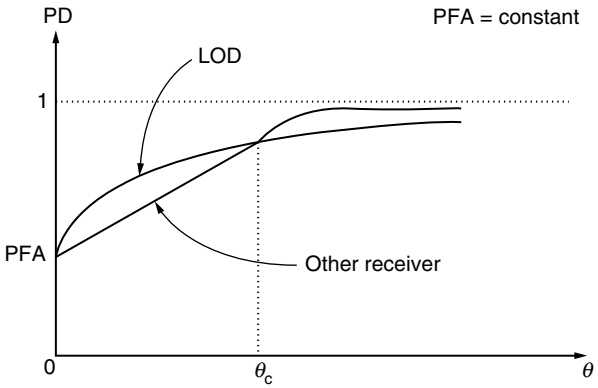

FIGURE 2.2 Illustration of non-optimality of locally optimum detector when $\theta > \theta_c$.

An alternative approach to the weak signal detection problem (i.e., $\theta \ll 1$) is to find that receiver which maximizes the slope of the power function at $\theta = 0$. Solution of this optimization problem results in the same ratio test as given by Equation 2.2. Consequently, the LOD is identical to that receiver which maximizes the slope of power function at $\theta = 0$. As a result, the LOD will maximize PD for θ sufficiently small. However, other receivers may yield larger values of PD when θ becomes large. As shown in Figure 2.2, for $\theta > \theta_c$ another receiver is shown to out-perform the LOD.

2.2. UNDERSTANDING OF SIGNAL AND DETECTION USING A FEEDFORWARD EXPERT SYSTEM

2.2.1. INTRODUCTION

The use of SIRPs in the implementation of LRTs and LODs for radar detection problem allows to derive algorithms for performing strong and weak signal

detection in a nonGaussian environment. Classical detection assumes *a priori* knowledge of joint PDF underlying the received data. In practice, received data can come from a clear region, where background noise (BN) alone is present or from a clutter (CL) region, where returns are due to BN and reflections from such objects as ground, sea, buildings, birds, … etc. When a desired target return is from a clear region and the BN is sufficiently small, the signal-to-noise ratio will be large and the strong signal detector (i.e., the LRT) should be used. However, if a desired target return is from a CL region, two situations can exist. When the desired target can be filtered from the CL by means of space–time processing and the BN is sufficiently small, the signal-to-noise ratio will be large and a strong signal detector should again be used. When the desired target cannot be filtered from the CL by means of space–time processing and the CL return is much larger then the desired target return, a weak signal detector (i.e., LOD) should be used. Use of the LOD in a strong signal situation can result in a severe loss of performance. Hence, it is necessary for the receiver to determine whether a strong or weak signal situation exists.

All of this suggest use of an expert system in the radar detection problem for: (1) monitoring the environment and (2) selecting the appropriate detector for processing the data. This is in contrast to current practice where a single robust detector, usually based on the Gaussian assumption, is employed. In addition depending on statistical changes in the environment over time and space, the expert system enables the receiver to adapt so as to achieve close to optimal performance. The goal of this study is to explore how an expert system can be used to develop an adaptive radar receiver, able to outperform traditional radars with respect to high subCL visibility. The focus of the study deals with detection of weak targets which cannot be filtered from the CL by means of space–time processing. For this purpose, it is convenient to divide the problem into two steps. The first step involves classification of the cells to be tested, while the second has to do with determining whether or not a target is present in the test cell. The two steps are discussed in more details in the following sections.

2.2.2. Classification of the Test Cells

Classification of test cells involves two steps: mapping of the space and indexing of the cells.

2.2.2.1. Mapping of the Space

In this research, the term, space, is used in its most general context. In practice, an effort is made to filter the desired target return from the CL to the extent possible. This is accomplished by performing space–time processing on received data. In particular, given N temporal samples and M spatial samples from a single range ring, spatial and temporal spectra are generated by performing a linear transformation on the NM samples. When random processes are modeled as SIRPs, the space–time processing will not change the nature of random

processes, since a linear transformation of an SIRP results in an SIRP of the same type. For example, if the underlying samples are modeled as Weibull, the transformed samples will also be Weibull. The spatial spectra contain angle of arrival information while the temporal spectra contain Doppler information. Consequently, the space of interest consists of range/angle of arrival/Doppler cells, respectively. It is in this general context that we refer to space and cells.

The mapping of the space can be done in two different ways: first, a CL map of entire space can be generated by assigning a scan of the radar receiver for this purpose. The CL map will indicate *a priori* those locations, which are likely to consist of clear regions (where BN alone is present) and CL patches. Another way, is to do a real-time assessment of a test cell. In other words, the question of whether the test cell is in a clear region or in a CL patch is determined in real-time during the scan.

2.2.2.2. Indexing of the Cells

In the regions where CL patches are present, a sufficient number of reference cells near the test cell must be chosen so that the number of samples from the reference cells (approximately 100 if using Ozturk's algorithm[26,27]) are adequate for approximating the body of the joint PDF that is to be used in characterizing the disturbance of the test cell. It should be recognized that a CL patch could be nonhomogeneous. If so, it may be necessary to subdivide the CL patches into subpatches. The samples from each subpatch would be approximated by a different joint PDF.

An important point is that knowledge of the body of joint PDF describing the disturbance is usually insufficient for determination of threshold of the radar receiver. In this research, our focus is on the weak signal detection problem. As pointed out earlier, the appropriate processor for this problem is the LOD. For the LOD, Chakravathi and Weiner[3,18,23] have conjectured that the body of PDF for the disturbance may be sufficient for determining the tail of the PDF for the LOD statistic. The body of joint PDF describing the disturbance could then be sufficient for determining the threshold of the LOD. This is significant because, only around 100 reference cells are needed when Ozturk's identification algorithm is used.

A set of descriptors need to be allocated for each cell. This allocation is referred to as cell indexing. The first item which needs to be determined is whether the detection problem for the test cell should be classified as strong signal detection, weak signal detection, or an intermediate case which falls between strong and weak signal detection. For the case of strong signal detection, the conventional radar receiver is adequate. For weak signal detection, the LOD should be used. For the intermediate case, it is proposed to use Wang's processor[19] based on the generalized likelihood ratio test (GLRT).

In all the three cases, it is necessary to use reference cells to estimate information needed for implementation of the processor. This assumes that the disturbance is homogeneous such that the reference cells neighboring the

test cell are representative of the test cell. However, a test cell may be near the edge of a CL patch or subpatch. Its location relative to these edges should be known, so that the reference cells, which are truly representative of the test cell, are selected.

For weak signal detection, the LOD differs from one joint PDF to another. Hence, it is necessary to know the type of random processes (e.g., Gaussian, *K*-distributed, Weibull, etc.) associated with the disturbance in the region, where the test cell is located. In practice, Ozturk's algorithm may indicate that, several different PDFs are suitable for approximating a particular disturbance. When this is the case, it may be desirable to implement several LODs and fuse the results. Hence, descriptors should be allocated to each cell for describing one or more joint PDFs in terms of their types, scales, locations, and shape parameters.

2.2.3. Target Detection

As mentioned previously, the type of detector depends upon whether the detection problem has been classified as a strong signal case (SSC), weak signal case (WSC), or intermediate signal case (ISC). Once the type of detector has been selected, parameters for the sufficient statistic of the detector are determined from the cell descriptors. Information from the cell descriptors is used to determine the threshold. When several detectors are employed simultaneously, as could arise in the WSC, a fusion algorithm is used to arrive at a global decision.[31]

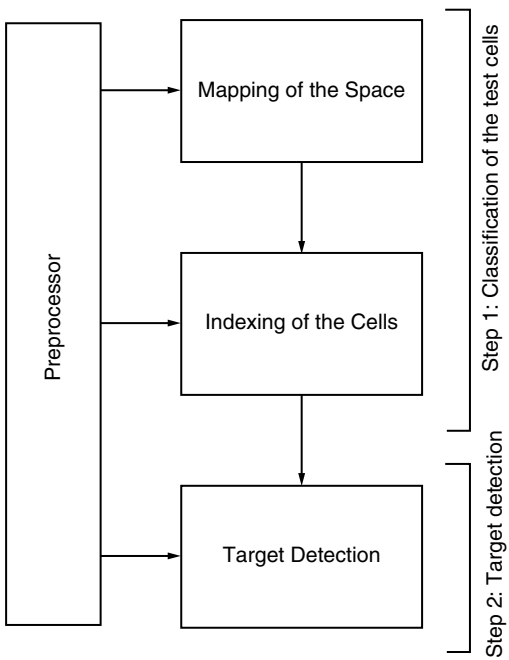

FIGURE 2.3 Signal understanding and detection using a feedforward expert system.

The above discussion provides the basis for signal understanding and detection using a feed forward expert system (SUD/FFES). A block diagram of the SUD/FFES is shown in Figure 2.3.

The preprocessor collects data, performs classical space–time processing, and stores the resulting data. This data is then used by the mapping, indexing, and detection blocks, which implement the two steps discussed previously. A more detailed block diagram is shown in Figure 2.4 where each block is subdivided into signal processing blocks and signal interpretation blocks. Because weak signal detection will be emphasized in this study, the portion dealing with weak signal detection, is shown using solid lines while the remaining portions are enclosed by dashed lines. The solid line portion is referred to as weak signal understanding and detection using a feed forward expert system (WSUD/FFES).

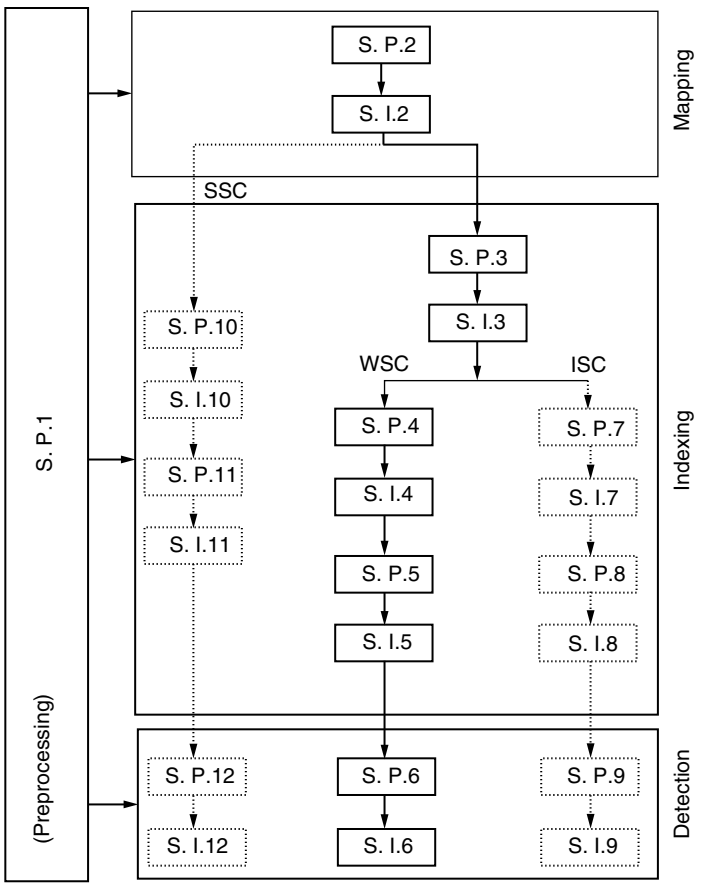

FIGURE 2.4 Signal processing and interpretation blocks in a feedforward expert system.

TABLE 2.1
Functioning of the Signal Processing and Interpretation Blocks in a Feedforward Expert System

	Preprocessing		
S. P. 1	Collects data. Performs space-time processing for each range ring to produce a two dimensional spectrum describing angle of arrival and Doppler frequency information.		
	Mapping		
S. P. 2	Uses spectral analysis and rank ordering or a thresholding technique to distinguish between CL and background noise.		
S. I. 2	Declares cells as BN (therefore SSC) or CL cells.		
	Indexing		
	S. P. 3 Uses rank ordering or thresholding technique to distinguish between WSC and ISC		
	S. I. 3 Declares cells as WSC or ISC		
	SSC	***WSC***	***ISC***
	S. P. 10 For a given test cell assigns a homogeneous distance procedure to every pair of cells within a specified region of the space.	**S. P. 4** For a given test cell assigns a homogeneous distance procedure to every pair of cells within a specified region of the space.	**S. P. 7** For a given test cell assigns a homogeneous distance procedure to every pair of cells within a specified region of the space.
	S. I. 10 Chooses the reference cells for the test cell based upon number of reference cells required and the homogeneous distance procedure.	**S. I. 4** Chooses reference cells for the test cell based upon number of reference cells required and the homogeneous distance procedure.	**S. I. 7** Chooses reference cells for the test cell based upon number of reference cells required and the homogeneous distance procedure.

Continued

TABLE 2.1 Continued

Indexing (cont.)

S. P. 11 Estimates for each test cell the variance needed to determine threshold.	**S. P. 5** Uses Ozturk's algorithm to determine the PDF(s) of the reference cells and the associated parameters (e.g., scale, location, shape, variance…).	**S. P. 8** Estimates signal covariance matrix of the test cell using data from the reference cells.
S. I. 11 Declares the cells from the clear region as SSC, stores the value of the variance for each cell (assuming white Gaussian background noise).	**S. I. 5** Assigns one or more PDF(s) and their parameters to the test cell.	**S. I. 8** Models data as having a jointly Gaussian PDF.

Target Detection

SSC	*WSC*	*ISC*
S. P. 12 Computes threshold and processes data from test cell using conventional radar (LRT).	**S. P. 6** Computes threshold, processes data using one or more LODs, and fuses the results.	**S. P. 9** Computes threshold and processes data using GLRT.
S. I. 12 Target present or absent.	**S. I. 6** Target present or absent.	**S. I. 9** Target present or absent.

A more detailed discussion of each block is provided in Table 2.1, and further more in Sections 2.4–2.7. It may be noted that, this is a feedforward system in the because the mapping, indexing, and detection are done sequentially with no feedback between these blocks.

Since the data collected from the environment is random, it is not possible to carry out an exact probability distribution identification using a reasonably small amount of data. Consequently, it is better to think about the probability distribution identification problem as an approximation problem, where we expect to select PDFs and their parameters, which result in good approximations to the underlying probability distributions.

Because the expert system of Figure 2.3 is feedforward, it does not allow for verification of the interpretations made by any of the signal interpretation blocks of Figure 2.4. Also, reprocessing of the data is not allowed by the feedforward configuration. The system discussed in the next chapter adds to the expert system the ability to reprocess data, if discrepancies are found to exist at the output of any signal interpretation block. The concept of discrepancy analysis is discussed in the next section.

2.3. SIGNAL UNDERSTANDING AND DETECTION USING A FEEDBACK EXPERT SYSTEM

2.3.1. Introduction

To convert the feedforward system of the previous chapter to a feedback system, feedback is introduced within each block and to each previous block. This is shown in Figures 2.5 and 2.6. We refer to such a system as signal understanding and detection using a feedback expert system (SUD/FBES). The expert system analyzes the output of each block and makes an assessment, as to whether correct signal processing and understanding have taken place. The signal interpretation blocks of the feedforward system are augmented to carry out this task. Depending upon the assessment made, additional data and signal processing may be carried out.

The assessment procedure is indicated in Figure 2.5 by a block in the shape of a diamond with one input and two outputs. The diamond input corresponds to the output of the block to be evaluated. If correct signal processing and understanding is believed to have taken place, the block under evaluation is allowed to communicate with the next block directly through the normal feedforward channel. Otherwise, additional data and signal processing are carried out under supervision of the expert system.

The feedback expert system to be used in this study, is called IPUS[20] (Integrated Processing and Understanding of Signals). It is presently under development by V. Lesser and H. Nawab[21] and was successfully applied to an acoustic recognition problem.[22] In this study, it is proposed to apply the IPUS expert system to the radar weak signal detection problem. In the following sections, the theory of IPUS is introduced and examples for application of IPUS to the radar weak signal detection problem are discussed.

2.3.2. IPUS Architecture

2.3.2.1. Introduction

The IPUS architecture has evolved from research on the design of an acoustic recognition system.[22] The goal of such a system is to identify the origins of various sound sources (such as telephones, vacuum cleaners, crying infants, etc.). The complexity of the acoustic recognition problem arises because of two factors:

(a) The need to process a tremendous variety of signal types due to the situation-dependent nature of the input. For example, not only may the input of an acoustic recognition system includes different types of signals, such as narrow-band, impulsive, and harmonic signals, but also may include various combinations of these signals.
(b) The need to change processing goals in a context-dependent way. For example, the goal of a signal understanding system might be to respond

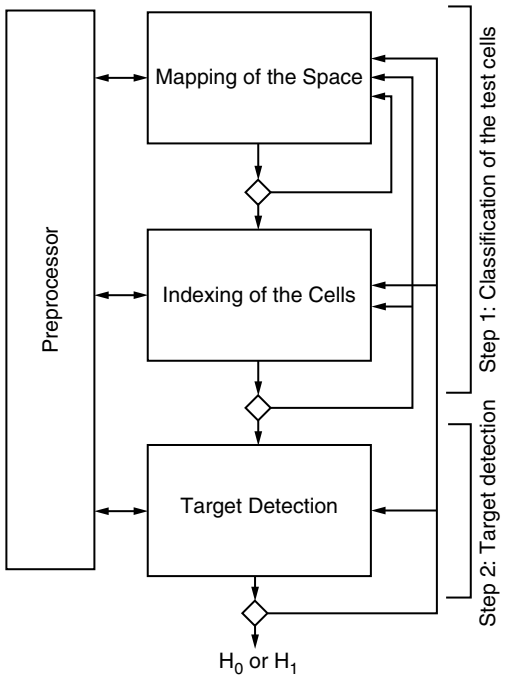

FIGURE 2.5 SUD/FBES.

to either the sounds of an infant or a ringing telephone and ignore other sound sources. If an infant sound is detected, the system would then ignore the telephone and would switch its main goal to determining whether the infant is laughing, crying, or choking.

These two factors also arise in the radar detection problem. Specifically, complexity is encountered because of:

(a) The need to process a tremendous variety of signal types due to the situation-dependent nature of input. For example, the PDFs of random received signals may be Gaussian, Weibull, *K*-distributed, etc., with various values for the scales, locations, and shape parameters.
(b) The need to change processing goals in a context-dependent way. For example, the usual operational mode of a radar involves processing of returns from clear regions which consist of strong signals embedded in a weak Gaussian BN. If a return from a CL region is determined, it must be decided whether either the intermediate or weak signal case (WSC) exists. If the WSC is applicable, one or more LODs need to be selected.

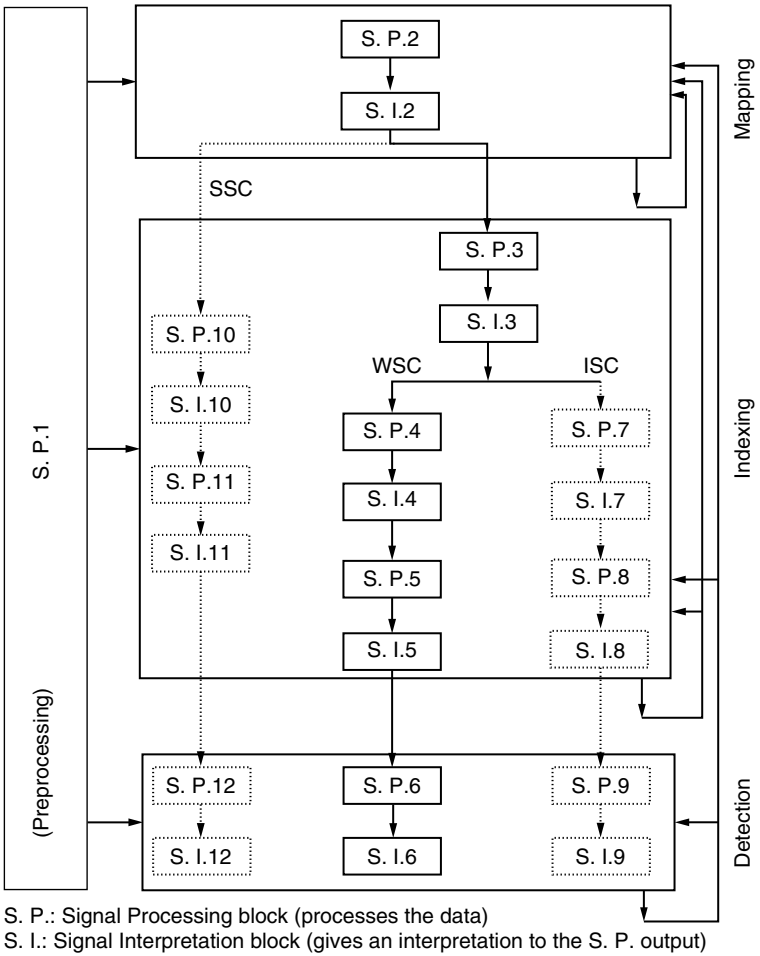

FIGURE 2.6 Signal processing and interpretation blocks in a feedback expert system.

In discussing the IPUS expert system, concepts are illustrated with examples taken from the radar detection problem.

In the radar detection problem considered in this research, it is assumed that, the signal environment is unknown *a priori*, even though mathematical models, for various signals that can arise, are assumed to be known. Once the environment has been determined, application of the appropriate signal detection algorithm (e.g., Gaussian receiver, LODs, GLRT) is straightforward. The difficult problem addressed by the IPUS architecture deals with the use of measured data to identify suitable mathematical models, for approximating the various signals received from the unknown environment.

The algorithms employed in the IPUS architecture to identify mathematical models for approximating the received signals, are referred to as signal processing algorithms (SPAs).

Because of the two factors mentioned previously, it is very difficult or even impossible, to design a single mathematically derived SPA which can be applied to all possible input signals to produce the desired information for each input. To deal with such complexities, the approach taken in the IPUS architecture is for the signal understanding system to have access to a "database" of mathematically derived algorithms. For the radar detection problem, examples are algorithms for the mean estimation, covariance matrix estimation, goodness of fit test and the PDF approximation. This database is indexed by the type of assumptions made about the input signal and the type of output information desired in accordance with the current goals of the signal understanding system.

For example, it may be assumed that the input signal is Gaussian. A goodness of fit test algorithm would be applied to determine whether the data is statistically consistent with the Gaussian assumption. If the Gaussian assumption is not rejected, then the desired output information would be the sample mean vector and sample covariance matrix of the data.

The IPUS architecture utilizes the fact that, signal processing theories often supply a system designer with an SPA which has adjustable control parameters (sampling period of data samples, number and location of reference cells, etc.). SPA denotes a database of SPA "instances," each instance corresponding to a particular set of fixed values for the control parameters. The IPUS architecture is designed to search for appropriate SPA instances to be utilized in particular situations in order to accurately model the unknown environment.

Two basic approaches for carrying out the signal processing are:

(a) Process the incoming signal with all the SPA instances those are potentially relevant to the entire class of possible input signals in the application domain and then choose the output data which has the most consistent interpretation. This approach requires vast amounts of signal-processing output data to be examined by higher level interpretation processes.
(b) Process the incoming signal with one or a small number of the possibly relevant SPA instances, then use some mechanism to recognize whether incorrect processing has taken place. This is followed by determination of the nature of incorrect processing, through a diagnostic reasoning process and finally changing the parameter settings of the SPA with the aim of obtaining an SPA instance appropriate for the processing of input signal. The SPA instance with the adjusted control parameter settings is then used to reprocess the input signal.

In order to select appropriate values for SPA control parameters, the system must consider the current system goals as well as knowledge about certain

characteristics of the particular input signal. This leads to a paradox. Choosing the appropriate control parameter values requires knowledge about the signal. However, this knowledge can only be obtained by first processing the signal with an algorithm with appropriate control parameter setting. Under this situation the following approach is adopted.

The IPUS architecture uses an iterative technique for converging to the appropriate control parameter values. The technique begins by using the best available guess for the SPA control parameters values. The SPA instance output is then analyzed through a discrepancy detection mechanism for indicating the presence of distorted SPA output data. A diagnosis is then performed for mapping the detected discrepancies to distortion hypotheses. A signal reprocessing phase then proposes a new set of values for the control parameters of the SPA with the aim of eliminating the hypothesized distortions. The SPA instance corresponding to new control parameter values is then used to reprocess the input signal. The output from the reprocessing once again undergoes discrepancy detection and if necessary, is followed by diagnosis, signal reprocessing planning, and further reprocessing of the input signal.

The signal data and the interpretation hypotheses derived from that data are stored on a blackboard with hierarchically organized information levels. The hypotheses on the blackboard fall into two basic categories: hypotheses posted to explain the signal data and hypotheses posted to specify expectations about the nature of the signal data. The inferencing on the blackboard is performed by different knowledge sources (KSs) for tasks, such as, discrepancy detection, diagnosis and reprocessing, and data interpretation. These tasks are presented in the following subsections.

2.3.2.2. Discrepancy Detection

Ideally, application of an SPA instance to input data results in undistorted output data. If the control parameters of the SPA instance are not appropriately chosen, distorted output data may result. The key to discrepancy detection is the ability to recognize and classify discrepancies due to distortion introduced by the SPA instance. Three types of discrepancies are possible:

1. The first type of discrepancy is referred to as a *violation*. A violation occurs when the SPA output data implies the presence of a signal which is not a member of the allowable class of input signals. For example, disturbances arising from cells in the clear region are always modeled in this work as Gaussian processes, because of the expectation that BN is Gaussian. Suppose, when output data from an SPA instance implies that the disturbance from a cell in the clear region is non-Gaussian, a violation type of discrepancy would result.
2. The second type of discrepancy is referred to as a *conflict*. A conflict occurs when current SPA output data is inconsistent with expectations arising from interpretations of past data. There are two types of conflicts depending on

whether all, or only a portion, of the current SPA output data is inconsistent. For an example of the first type of conflict, suppose previous SPA output data arose from disturbances in the clear region while current SPA output data is arising from disturbances in a CL patch. A conflict of the first type occurs if all of the current SPA output data, such as an increase in variance and nonGaussianity of the data, conflict with previous interpretations from the clear region. For an example of the second type of conflict, suppose previous SPA output data has resulted in the interpretation that the disturbance is from the clear region. This might be implied by the SPA output data indicating Gaussian statistics, zero mean and a variance level in the range of BN. A conflict of the second type occurs when, even though Gaussian statistics are confirmed by current SPA output data, it also reveals that the mean is no longer zero and the variance level has increased significantly. This could happen if the disturbance is now coming from a Gaussian CL patch where the data is highly correlated.

3. The third type of discrepancy is referred to as a *fault*. A fault can arise in two different ways. First, it occurs when two or more different SPAs, those that are applied to the same data, result in different output interpretations. Second, it occurs when two or more instances of a single SPA (i.e., the same SPA with different parameter values) result in different interpretations when applied to the same data. An example of the first kind of fault would be the situation where SPA # 1, a power level detector, indicates a power level consistent with the BN while SPA # 2, Ozturk's distribution identification algorithm, indicates a nonGaussian distribution. This is a fault because the BN is assumed to be Gaussian. An example of the second kind of fault would be the situation where the use of Ozturk's algorithm based on 50 and 100 samples from the same CL patch results in a different interpretation.

2.3.2.3. Diagnosis and Reprocessing

When the signal being monitored does not satisfy the requirements of the SPA instance, the output of the SPA is distorted resulting in a discrepancy. Once a discrepancy has been detected, a diagnosis procedure is used to identify the distortion which may have led to the discrepancy. Knowing the distortion, either the appropriate parameters of the same SPA can be adjusted or a different SPA can be chosen to reprocess the data. In a sense, the diagnosis procedure maps symptoms (discrepancies) to hypothesized underlying causes (distortions). For example, assume the sample mean of a CL patch is repeatedly being evaluated by processing 50 samples at a time. Although, the first eight trials result in values close to zero, the ninth trial produces a large negative value for the mean. This represents a conflict of the first kind. The diagnosis procedure may surmise that the conflict may be due to the presence of one or more outliers. Consequently, the reprocessing procedure concludes that the data from the ninth trial should be reprocessed after removal of the outliers.

2.3.2.4. Interpretation Process

The interpretation process is a search through a space of sets of interpretation models for modeling signals. When a possible combinatorial explosion in interpretation models does not exist, the interpretation process may be viewed as just a straight-forward classification process. Otherwise, the search must be carried out as a constructive problem solving process. The IPUS architecture employs the constructive problem solving approach, which reduces to the classification approach in absence of a combinatorial explosion.

Constructive problem solving techniques must be used when the set of possible solutions is too large to be enumerated. For example, although the set of PDF types is finite in the radar problem, there are an infinite number of different PDFs possible because of the infinite number of values that can be assumed by the scales, locations, and shape parameters. Consequently, constructive problem solving is needed to approximate the underlying probability distribution of the data.

2.3.2.5. SOU and Resolving Control Structure

At any stage in monitoring the environment, one can never be totally sure that various interpretations are correct. Sources of uncertainty (SOU) always exist with regards to the correct processing of data and the extent of confidence to be placed in the interpretation. The objective of IPUS is to continue to reprocess data, as time permits, so as to reduce the SOUs. This reprocessing is continued until either the level of uncertainty has been reduced to some acceptable level or until all the time allotted for reprocessing has been utilized.

The process of reducing the SOUs can be viewed as a problem solving activity. At each stage of the IPUS architecture, it is necessary to identify SOUs associated with a particular output and to have a strategy for reprocessing the data, so as to efficiently reduce the SOUs associated with a particular output. The Resolving Sources of Uncertainty (RESUN) control[21] structure is used to direct the problem-solving procedure used to gather evidence, in order to resolve particular SOUs in the interpretation hypotheses.

The IPUS architecture and RESUN control structure presently group SOUs into three categories: (1) violation-type discrepancies, (2) conflict-type discrepancies, (3) fault-type discrepancies. The basis for these categories is largely empirical and is continually evolving. These categories are discussed below.

2.3.2.5.1. Violation-Type Discrepancies

A *violation* type discrepancy occurs when signal processing output data violates *a priori* known characteristics of the entire class of possible input signals for the radar problem. For example, BN is assumed to have zero mean. Assume that a signal has been interpreted as white Gaussian and with a very small power level. In addition, the mean has been interpreted as being nonzero. Because of the nonzero mean, the signal cannot be classified as BN. Also, because of the small

power level, the signal cannot be classified as CL. Clearly, a violation exists and the zero-mean condition has been violated. There are various SOUs which could have caused this violation. For example,

(1) there could have been one or more outliers in data,
(2) there may not have been an adequate number of samples processed,
(3) the random samples may not have been representative of the zero-mean signal,
(4) a desired target return could be contained in one or more of the reference cells.

For each of the above SOUs, respectively, one or more procedures exist for their reduction. For example,

(1) extreme values could be excised from the data,
(2) the number of samples processed could be increased,
(3) an entire set of new samples could be processed,
(4) previously detected targets could be checked to determine, whether one or more targets are likely to be present in one or more of the reference cells. If so, these cells should be eliminated as reference cells, for the test cell of interest.

When a violation-type discrepancy is noted, the hypothesis that a violation exists is posted on the interpretation blackboard of the expert system. For each condition being violation and each SOU that could have caused the violation, an additional output data hypothesis is posted which describes the condition being violated and the SOU that could have caused the violation. The SOUs, which fail to produce some particular support evidence for a single input signal or fail to result in any valid explanation for a combination of input signals, are classified as negative evidence SOUs. The hypothesis concerning the violation-type discrepancy and each of the additional hypotheses are then connected by a negative evidence link and are labeled as **VIOLATION-NEGATIVE EVIDENCE SOU**. Further problem solving to reduce the SOU can then be carried out by reprocessing the underlying signal with different SPA instances.

2.3.2.5.2. Conflict-Type Discrepancies

After processing a certain quantity of data, various expectations related to the data to be processed arise. *Conflicts* occur when these expectations are not met. If all the expectations are not met, the conflict is said to be of the first kind. If some expectations are met while others are not, the conflict is said to be of the second type.

An unverified expectation hypothesis exists when there is a conflict of the first type. A **NO-SUPPORT SOU** label is attached to each unverified

expectation hypothesis. For each expectation (i.e., component of the hypothesis) not being met and for each SOU that might be responsible, an additional hypothesis is posted on the interpretation blackboard which describes the unsupported expectations and the corresponding SOUs. The hypothesis concerning the conflict-type discrepancy and each of the additional hypotheses are connected by a link labeled **NO-SUPPORT SOU**. Further problem solving can then attempt to find SPA instances so as to produce output data that can support the expectation.

When there is a conflict of the second type, a partially verified expectation hypothesis is posted on the interpretation blackboard with a **PARTIAL-EVIDENCE SOU** label. As in the previous case, each unsupported expectation (i.e., inference or hypothesis component) is associated with a hypothesis, labeled **NO-SUPPORT SOU**, which describes the unsupported expectation and the corresponding SOU. The hypothesis labeled **PARTIAL-EVIDENCE SOU** can trigger further problem solving, to find support for lower component hypotheses labeled **NO-SUPPORT SOU**.

By way of example, assume that the previous ten test cells were classified as having been in the clear. Our expectation hypothesis, therefore, is that the 11th test cell will also be in the clear. If the 11th test cell fails the Gaussian, zero-mean, and small power level inferences, we have a conflict of the first type. On the other hand, if the Gaussian and zero-mean inferences are found to have support while the small power level inference is unsupported, we have a conflict of the second type. SOUs and corresponding linked hypotheses would be identified with each unsupported inference and additional signal processing would be carried out to reduce the level of SOUs as time permits.

2.3.2.5.3. Fault-Type Discrepancies

Fault-type discrepancies arise when two different SPAs or instances produce different output interpretations. In such a case, a composite hypothesis is created which is a copy of the more reliable or the two output interpretations. A link labeled with a **NEGATIVE-EVIDENCE SOU** connects the less reliable hypothesis to the composite hypothesis. Further problem solving, attempts to remove the negative-evidence SOU by reprocessing the signal using different SPA instances for the less reliable SPA. Two possible outcomes are then possible. Either the negative evidence SOU is eliminated or it is replaced by another negative-evidence SOU. In the latter case, further problem solving can attempt to reprocess the signal with the more reliable SPA but using different SPA instances.

By way of example, assume that a zero-mean inference is to be supported by a sample-mean algorithm and a median detector algorithm. Assume that, only a small number of data samples are available. Under this assumption, the median detector algorithm is likely to be more reliable. Suppose, the median detector algorithm supports the zero-mean inference while the sample mean algorithm does not. RESUN creates the composite hypothesis that the mean is

zero. The composite hypothesis is limited to the hypothesis that the mean is not zero by using a **NEGATIVE-EVIDENCE OUTLIER** label. The data samples are then reprocessed by sample mean algorithm using SPA instances, which delete some of the extreme data values. If the zero-mean inference is supported, the negative evidence SOU is assumed to have been eliminated. On the contrary, the signal is reprocessed by the median detector algorithm, where different sets of outliers may be chosen.

2.3.3. Application of IPUS to Radar Signal Understanding

The IPUS architecture is suitable when a single SPA instance cannot correctly process all input signals which can potentially arise in a signal understanding application. In the radar problem, the variety of probability distributions underlying the data, along with different tasks to be carried out in monitoring the environment (CL mapping and cell indexing), necessitate more than one SPA instance, rendering IPUS suitable for the radar problem.

In this research we emphasize the weak singal case (WSC). As explained previously, the branch of Figure 2.6 corresponding to the WSC is referred to as WSUD/FBES. From the figure, it is clear that tasks of WSUD/FBES have been subdivided into mapping, indexing and detection. Assuming the mapping and indexing to be done properly, application of the LOD is straightforward.

2.4. PROPOSED RADAR SIGNAL PROCESSING SYSTEM USING A FEEDBACK EXPERT SYSTEM

2.4.1. Data Collection and Preprocessing

Assume that a radar transmits a periodic signal composed of a series of predefined RF pulses. Let T be the period of the signal and τ_e denote pulse duration. T is known as the pulse repetition interval (PRI). Assuming rectangular-shaped pulses, the envelope of emitted signals is shown in Figure 2.7 where A denotes pulse amplitude.

For ease of discussion, consider a low pulse repetition frequency (PRF) radar for which there is no range ambiguity. During the time interval $[pT + \tau_e, (p + 1)T]$, of the pth PRI, $p = 1, 2, 3, \ldots$, the radar functions are in the receive

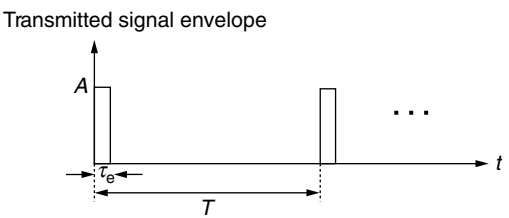

FIGURE 2.7 Envelope of the transmitted signals.

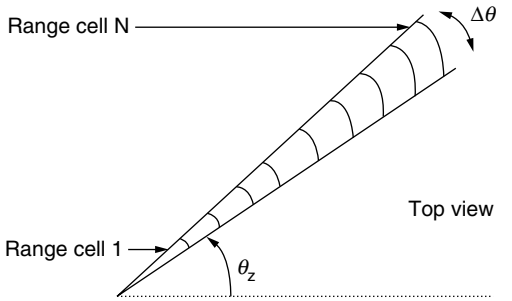

FIGURE 2.8 Range cells in a constant azimuth beam.

mode and collect data, due to the pth pulse from the azimuth angle, θ_z, to which the radar antenna is oriented. The collected data corresponds to returns from different range cells as illustrated in the top view of Figure 2.8 where $\Delta\theta$ denotes the antenna beam width.

The round trip delay, t_d, from a given range R is given by

$$t_d = \frac{2R}{c} \tag{2.18}$$

where, c denotes the velocity of propagation of the electromagnetic wave. For the pth period of the transmitted signal, the time interval $[pT + \tau_e, (p+1)T]$ is subdivided into J range bins of duration τ_e. Each range bin corresponds to a range cell with its range extent denoted by r_e. Figure 2.9 shows the different range cells and their corresponding range bins on the time axis for a fixed azimuth angle θ_z.

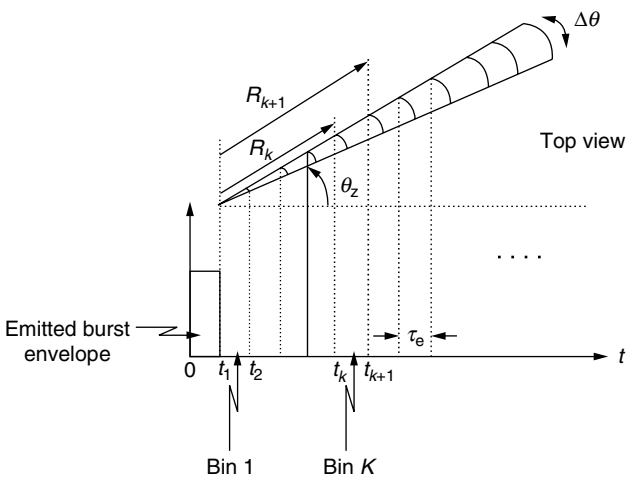

FIGURE 2.9 Range cells and their corresponding time bins.

Let R_k and R_{k+1} denote the ranges to the leading and trailing edges of the kth range cell. Also, let t_k and t_{k+1} denote the corresponding round trip delay times. Then, using Equation 2.18 it can be written

$$t_k = \frac{2R_k}{c} \quad \text{and} \quad t_{k+1} = \frac{2R_{k+1}}{c} \tag{2.19}$$

Taking the difference and utilizing the fact that a range bin is defined to be of time extent, τ_e, results in

$$t_{k+1} - t_k = \tau_e = \frac{2(R_{k+1} - R_k)}{c} \tag{2.20}$$

Hence, each range bin of time extent, τ_e, corresponds to a range cell of range extent, r_e, where

$$r_e = R_{k+1} - R_k = c\frac{\tau_e}{2} \tag{2.21}$$

This process is repeated for different azimuth angles, θ_z. The collection of cells corresponding to different ranges and azimuth angles can be represented in a rectangular format, as illustrated in Figure 2.10. The collection of range cells corresponding to a constant azimuth angle is shown in Figure 2.8. For a fixed range, the collection of azimuth cells is as shown in Figure 2.11. Let the range bins be indexed by k_R, $1 \le k_R \le J$, and the azimuth bins be indexed by k_A, $1 \le k_A \le K$. The resolution cell corresponding to $k_R = j$ and $k_A = k$ is referred to as the jkth range–azimuth (R/A) cell, as illustrated in Figure 2.10.

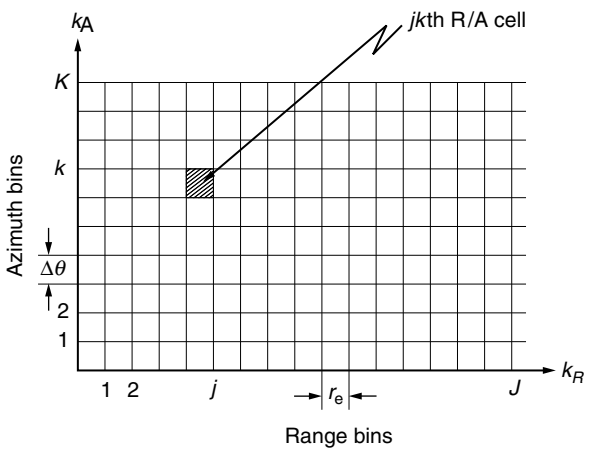

FIGURE 2.10 Rectangular format for representative cells at different ranges and azimuth angles.

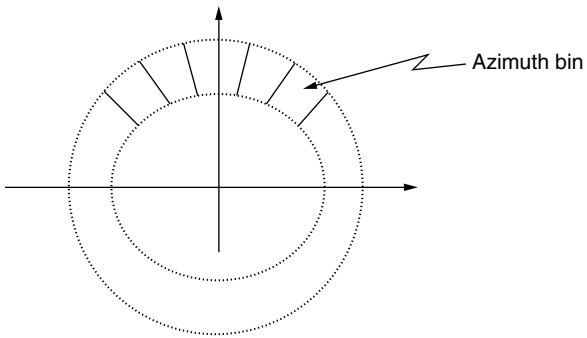

FIGURE 2.11 Collection of azimuth cells for a fixed range.

In practice, the antenna beam dwells on each R/A cell for a time duration, known as the dwell time which is equal to P PRIs. Consequently, P pulse returns from each R/A cell, spaced T seconds apart, can be processed. After decomposing the P received pulses into a sequence of I and Q quadrature components, a Doppler filter bank is implemented by performing a M-point fast Fourier transform (FFT) on each of the P quadrature component samples where $M \geq P$ depending upon whether or not zero padding is used. Let $y_{Ip}(j, k)$ and $y_{Qp}(j, k)$, $p = 1, 2, 3, \ldots, P$, denote the I and Q components from the pth PRI of the jkth R/A cell. Similarly, let $Y_{Im}(j, k)$ and $Y_{Qm}(j, k)$, $m = 1, 2, 3, \ldots, M$, denote the mth FFT coefficient from the jkth R/A cell. In addition to computing the temporal data magnitudes, $y_p(j, k)$, from the I and Q quadrature components for each PRI, the magnitude of each Fourier coefficient, $Y_m(j, k)$, is computed, as shown in Figure 2.12, to form the FFT coefficient magnitudes.

The collection of cells corresponding to different ranges, azimuth angles, and Doppler frequencies can be represented in a 3-D cubic format, as illustrated in Figure 2.13. Let the Doppler bins be indexed by k_{D}, $1 \leq k_{\mathrm{D}} \leq M$. The resolution cell corresponding to $k_{\mathrm{R}} = j$, $k_{\mathrm{A}} = k$, and $k_{\mathrm{D}} = m$ is referred to as the jkmth range–azimuth–Doppler (R/A/D) cell. The total number of R/A/D cells is given by $J \times K \times M$.

The temporal data magnitudes and FFT coefficient magnitudes are used by the mapping processor to enable classification of the R/A/D cells into BN cells and CL cells. The mapping approach is described in the next section for the general radar problem.

2.4.2. Mapping

The purpose of mapping is to declare a R/A/D cell, as either a BN cell or CL cell. As shown in Figure 2.14, two types of data are formed in the preprocessing stage. These are, temporal data magnitudes, $y_p(j, k)$, and FFT coefficient magnitudes, $Y_m(j, k)$. The set of $y_p(j, k)$ magnitudes consists of P data points for each R/A cell in the $J \times K$ R/A plane, while, the set of $Y_m(j, k)$

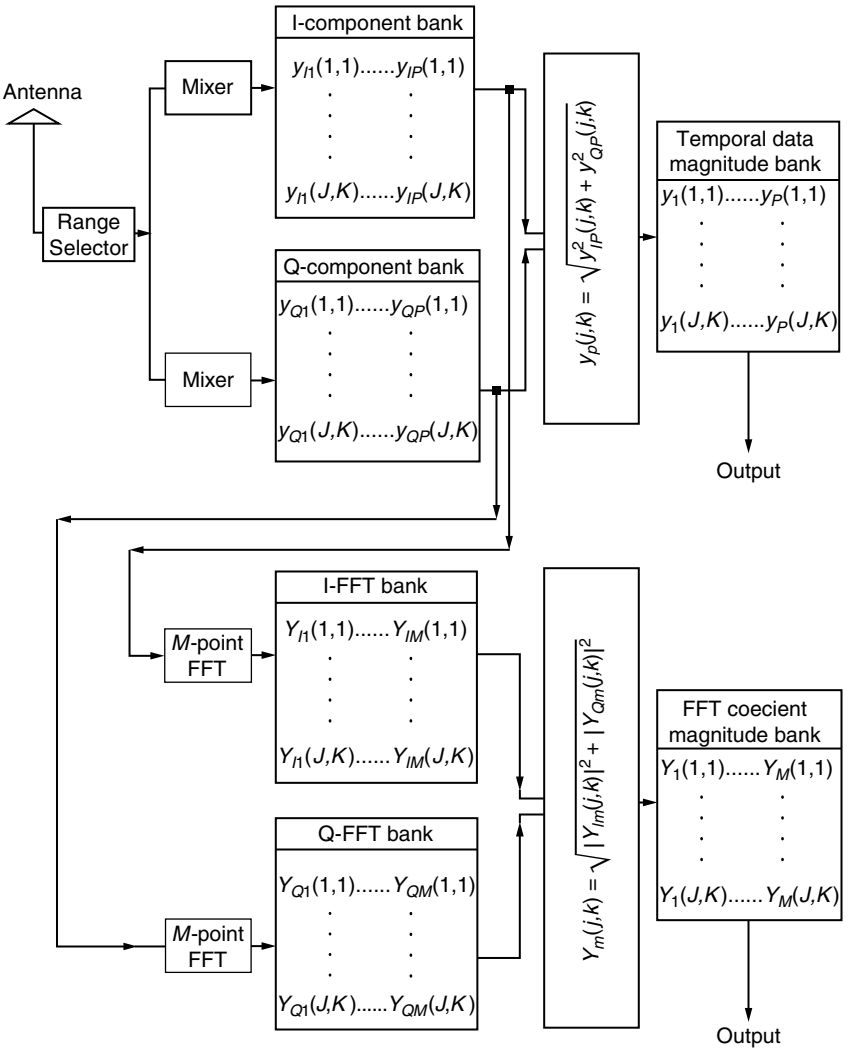

FIGURE 2.12 Block diagram of data processing stage.

magnitudes consists of one data point for each R/A/D cell in the $J \times K \times M$ R/A/D space.

The mapping is done in two stages. First, using $y_p(j, k)$, the R/A plane is mapped into BN and CL (CL) R/A cells. Then, using $Y_m(j, k)$, the 3-D R/A/D space is mapped into BN and CL R/A/D cells. The necessity for performing mapping in two stages is explained next.

In the first mapping stage, an average $P(j, k)$ is formed for each R/A cell in the $J \times K$ R/A plane by averaging the power of the P temporal data magnitudes,

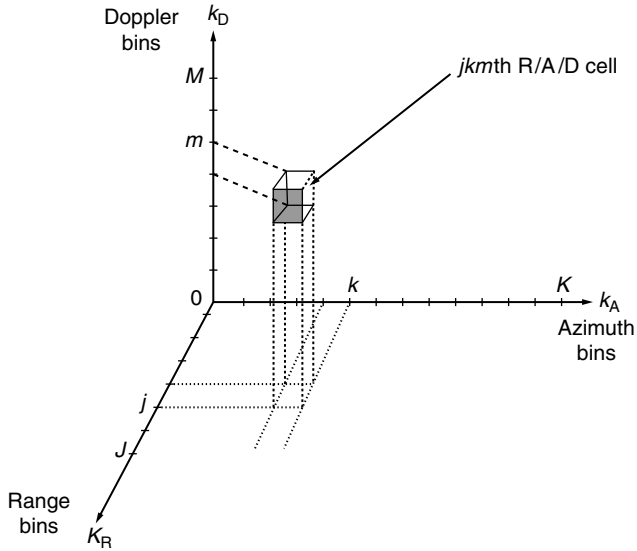

FIGURE 2.13 Cubic format for representing cells at different ranges, azimuth angles, and doppler frequencies.

$y_p(j, k)$ as shown in Equation 2.22,

$$P(j,k) = \frac{1}{P}\sum_{p=1}^{P} y_p^2(j,k); \quad \begin{matrix} j = 1,2,\ldots,J \\ k = 1,2,\ldots,K. \end{matrix} \tag{2.22}$$

The average powers $P(j, k)$ are used to classify BN and CL cells in the R/A plane, as is discussed in the next chapter.

The second stage involves the 3-D R/A/D space. In particular, each cell in the R/A/D space needs to be classified as either BN or CL. The classification performed in the first mapping stage is useful in reducing the effort required in the second stage. If the *jk*th R/A cell is BN, then all of the *jkm* Doppler cells, $m = 1, 2,\ldots, M$, will also be BN. However, if the *jk*th R/A cell is CL, then the *jkm* Doppler cells, $m = 1, 2,\ldots, M$, may or may not be CL depending upon the CL spectrum. Consequently, only the Doppler cells corresponding to R/A cells that have been identified as CL in the first mapping stage need to be examined in the second stage. This reduces considerably the amount of processing needed for the second mapping stage.

Notice that two-way communications is required between the preprocessing and mapping stages. As shown in Figure 2.14, the preprocessing stage generates the quadrature components, the FFT coefficients, and their magnitudes. The first mapping stage classifies the R/A cells as either BN or CL. For those R/A cells classified as BN, the corresponding Doppler cells are also classified as BN by the second mapping stage. For those R/A cells that are

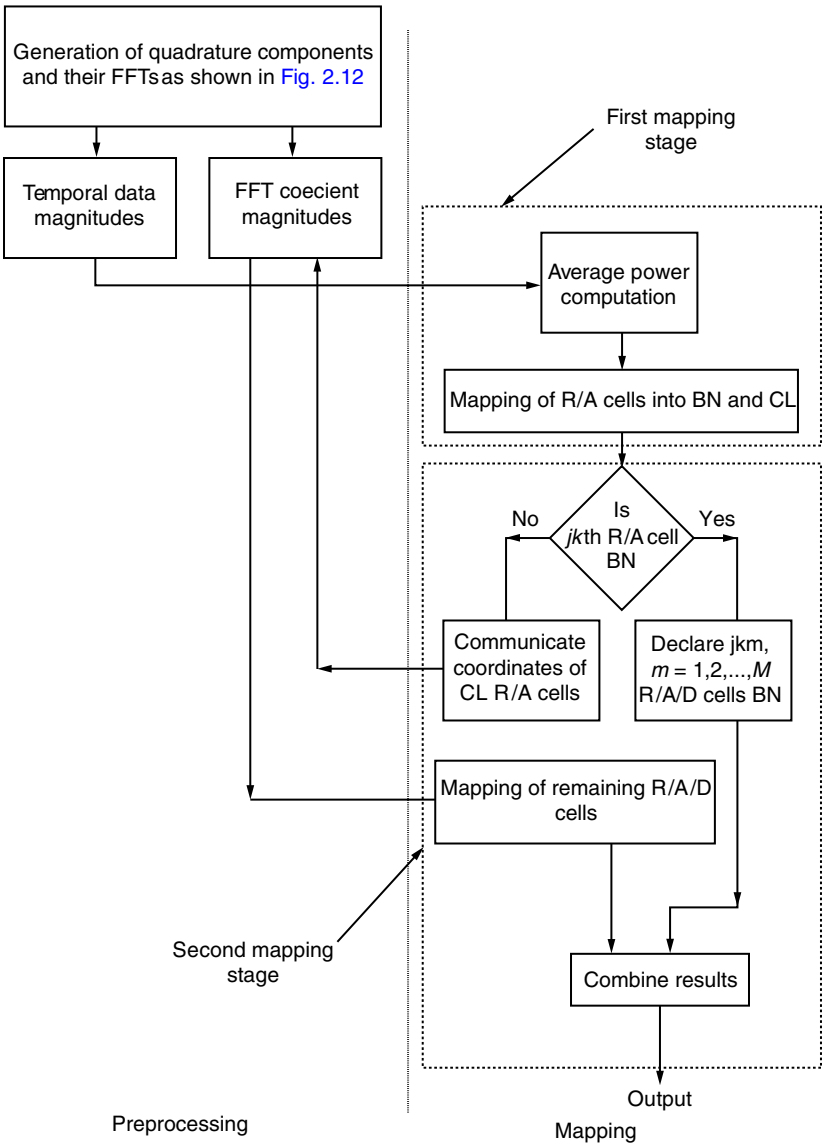

FIGURE 2.14 Block diagram of the preprocessing and mapping of data.

classified as CL, the second mapping stage obtains the FFT coefficient magnitudes. These are then used to classify the remaining R/A/D cells as either BN or CL.

The major focus of this work is to demonstrate how the principles of IPUS can be applied to the radar problem. For that purpose, procedures and rules have been developed that utilize the IPUS capabilities to solve the various

stages of the radar problem. Throughout this work the procedures and rules have been tested with nontrivial radar examples. The first procedure, presented in the next section, is a mapping technique that classifies cells into BN and CL.

2.5. MAPPING PROCEDURE

In this section a mapping procedure for a surveillance volume subdivided into R/A cells is presented along with examples. Also, extension of the mapping procedure to R/A/D cells is discussed in the last section. General procedures are presented without explanations of how to choose parameters. The explanations are provided in Section 2.7.

2.5.1. INTRODUCTION

Assume that $J \times K$ R/A cells are scanned by a radar antenna. Furthermore, for simplicity, assume that the dwell time is equal to the PRI so that only a single pulse is processed from each cell, i.e., $P = 1$. In this case, an FFT is not possible and the block diagram of the preprocessing stage reduces to that shown in Figure 2.15. When an FFT is not possible, the block diagram of the preprocessing and mapping stages, shown in Figure 2.14, simplifies to Figure 2.16. Only $J \times K$ temporal data magnitudes $y(j, k)$ are available to the mapping stage. In this case, only $J \times K$ R/A cells need to be mapped into BN and CL cells and the mapping process of Figure 2.14 reduces to the first mapping stage alone. As explained in Section 2.4.2, an average power $P(j, k)$ is formed for every R/A cell in the $J \times K$ R/A plane. For $P = 1$, Equation 2.22 becomes

$$P(j,k) = y_1^2(j,k); \qquad \begin{matrix} j = 1,2,\ldots,J \\ k = 1,2,\ldots,K. \end{matrix} \tag{2.23}$$

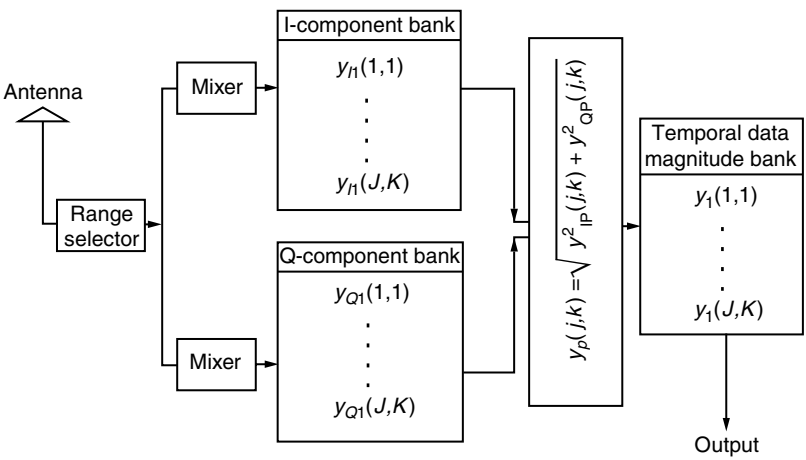

FIGURE 2.15 Block diagram of data preprocessing stage for $P = 1$.

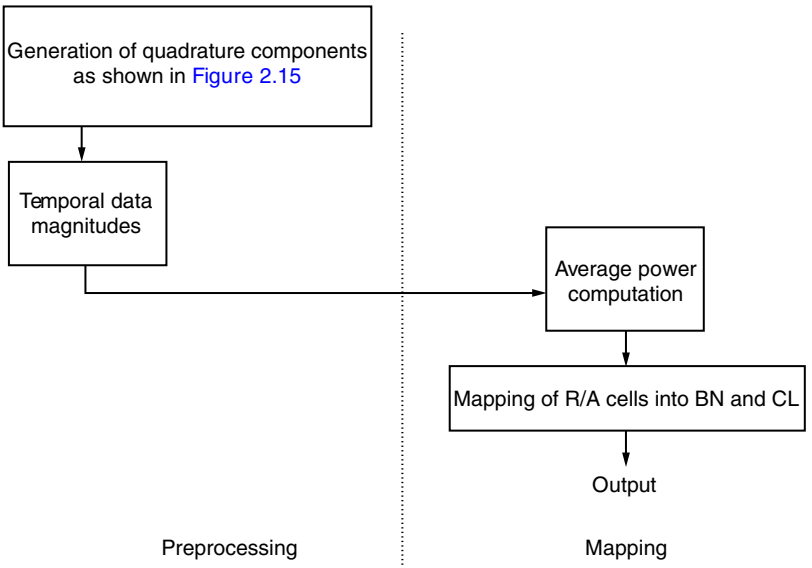

FIGURE 2.16 Block diagram of preprocessing and mapping stages for $P = 1$.

At this point, the R/A plane consists of two different types of regions that need to be identified. These are clear regions, where BN alone is present, and CL patches, where CL and additive BN are present. Let us, first, examine the nature of the BN and CL in order to understand the theory behind the procedure development for mapping.

2.5.2. Observations on BN and CL Cells

Assuming additive BN and CL, the following observations are based on many computer generated examples of a BN region containing CL patches where the CL-to-noise ratio (CNR) is assumed to be greater than 0 dB. The BN envelope is assumed to be Rayleigh while the CL envelope may be either Rayleigh, *K*-distributed, Weibull, or lognormal.

2.5.2.1. Observations on BN Cells

The following observations on BN cells were noted:

- On an average, the BN data values are smaller than the CL data values.
- Large data values may exist in a BN cell that may be higher than some data values of a CL cell.
- Large data values in the BN tend to be isolated points.
- The number of BN data significantly larger than the average is relatively small. Figure 2.17 shows a typical BN data histogram.

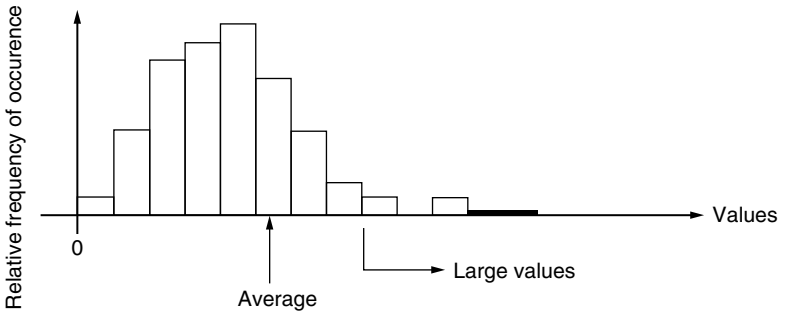

FIGURE 2.17 Example of a BN histogram.

– The relatively small numbers of large BN data are distributed sparsely throughout the surveillance volume.

2.5.2.2. Observations on CL Cells

The following observations on CL cells were noted:

– On an average, CL data values are higher than BN data values.
– A CL region contains additive CL and BN.
– Small data values may exist in a CL cell that may be smaller than some data values of a BN cell.
– The large CL data values tend to be larger than most of the large BN data values assuming positive CNR.
– Whereas the BN data values are distributed over the entire surveillance volume, the CL data values are distributed only over the CL regions.
– Large data values in the CL tend to be clustered.

Figure 2.18 shows an example of the distribution of large values of BN data denoted by (×) and large values of CL data denoted by (o). Notice that the large

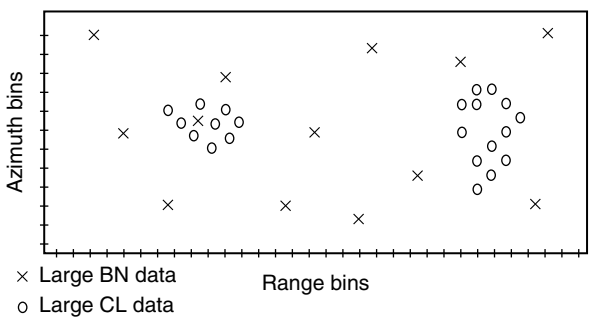

FIGURE 2.18 Concentration of large data in a scanned volume.

BN data values are sparsely distributed over a large area while large values of CL are concentrated in small regions, the CL regions.

2.5.3. Mapping Procedure

Using the fact that CL patches, on an average, have stronger radar returns, the mapping processor begins by setting a threshold which results in a specified fraction of BN cells, as explained in Section 2.7. Image processing is then used to establish the BN and CL patches. If the final image contains a significantly different fraction of BN than originally established by the initial threshold, the process is repeated with a new threshold. The mapping processor iterates until it is satisfied that the final scene is consistent with the latest specified threshold. Finally, CL patch edges are detected using image processing technique.

The mapping procedure consists of two steps. The first step is the identification of CL patches within BN and second is the detection of CL patch edges and their enhancements. These two steps are explained next.

2.5.3.1. Separation of CL Patches from Background Noise

Identification of CL patches within BN is performed by following steps: thresholding, quantization, correction, and assessment.

2.5.3.1.1. Thresholding and Quantization

Identification of CL patches within BN starts by setting a threshold q that results in a specified fraction of BN cells. Then a quantized volume is formed as follows: all R/A cells with average power less than q are given a value of zero and all R/A cells with average power above q are given a value of one (unity). Let $Q(j,k)$ represent the quantized value of the jkth R/A cell. Then,

$$Q(j,k) = \begin{cases} 1 & \text{if } P(j,k) \geq q \\ 0 & \text{if } P(j,k) < q \end{cases} \qquad j = 1,2,\ldots,J \text{ and } k = 1,2,\ldots,K \quad (2.24)$$

where $P(j,k)$, the average power of the jkth R/A cell, is defined in Equation 2.23.

Figure 2.19 shows the computer generated CL patches and BN region which are to be separated by mapping procedure. A, B, C, and D denote the CL patches. Figure 2.20 shows a typical contour plot of the quantized R/A volume. Initially, let R/A cells with a quantized value of one be declared as CL cells and R/A cells with a quantized value of zero be declared as BN cells. By comparing Figures 2.19 and 2.20, note that the quantized version differs from the original. This is due to the fact that even though the average powers of BN cells are expected to fall under the threshold typically, some BN cells have an average power that falls above the threshold. Similarly, while the average powers of CL cells are expected to fall above the threshold, typically, some CL cells have an average power that falls under the threshold. Also, as explained in Section 2.7,

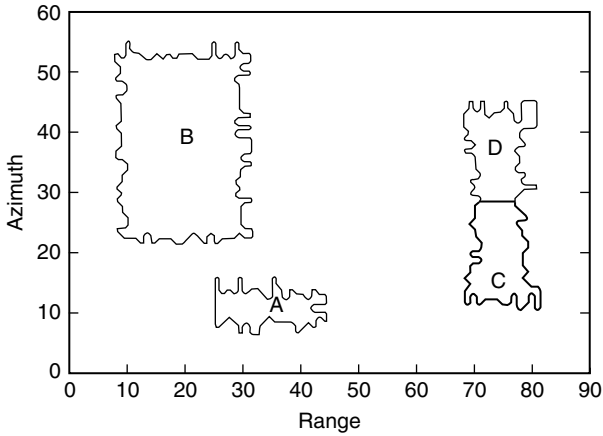

FIGURE 2.19 Example of a generated range–azimuth volume that is to be separated by the mapping procedure.

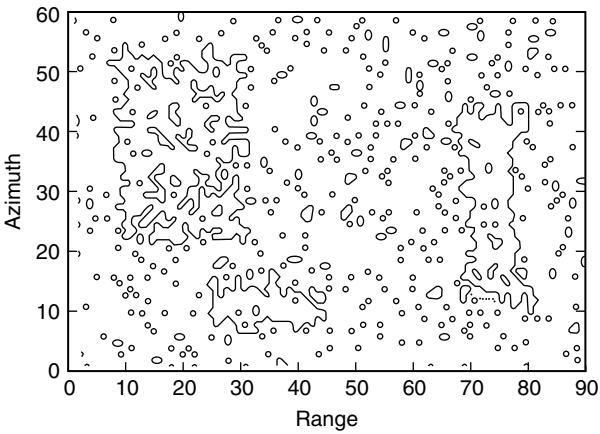

FIGURE 2.20 Example of a quantized range–azimuth volume.

the first setting of the threshold, which is somewhat arbitrary, is likely not to be the best for identifying CL patches within BN.

2.5.3.1.2. Correction

After quantization, the next step is to correct misclassified BN and CL cells. Consider a set of three by three R/A cells. As shown in Figure 2.21, let the center cell be referred to as the test cell and the surrounding cells be referred to as the neighboring cells. Assume that a CL patch cannot be formed by a single cell. In this case, every test cell in the CL patch has at least one neighboring cell that belongs to the same CL patch.

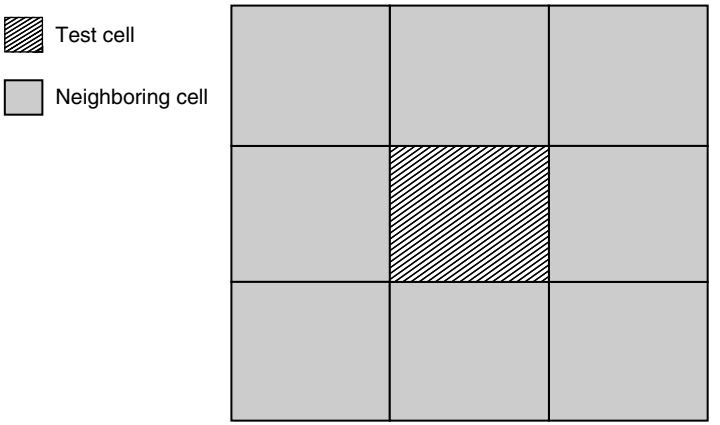

FIGURE 2.21 3 × 3 Range–azimuth cells.

A test cell belonging to a CL patch that has at least one neighboring BN cell is referred to as a CL edge cell. On the other hand, a test cell that belongs to a CL patch, for which none of the neighboring cells are in the BN, is referred to as an inner CL cell.

The proposed correction technique consists of transforming the quantized volume into a "corrected" volume. The transformation consists of the following steps:

1. As explained in Chapter 7, choose the necessary number of CL neighboring cells in the quantized volume (NCQ), for a test cell in the quantized volume to be declared as a CL cell in the corrected volume. NCQ can take one of the following values: 5, 6, 7, 8.
2. For every test cell in the quantized volume count the number of neighboring CL cells. If the number is greater than or equal to NCQ declare the test cell as a CL cell in the corrected volume. Otherwise, declare the test cell as a BN cell in the corrected volume.

When all the cells of quantized volume have been tested, a corrected volume consisting of declared BN or CL R/A cells is obtained. Notice that a cell type is not updated until the entire correction is done. Figure 2.22 shows an example of a corrected volume. Notice that the latter volume has larger homogeneous regions than the quantized one.

Because NCQ is chosen to be relatively large (i.e., NCQ = 5, 6, 7, or 8), BN cells those were incorrectly identified in the quantized volume as CL cells, due to their large power, tend to be reclassified as BN cells. Also, inner CL cells in the quantized volume are recognized as CL cells in the corrected volume. Meanwhile, most of the CL edge cells in the quantized volume are recognized as BN cells in the corrected volume. This results in an over-correction where most

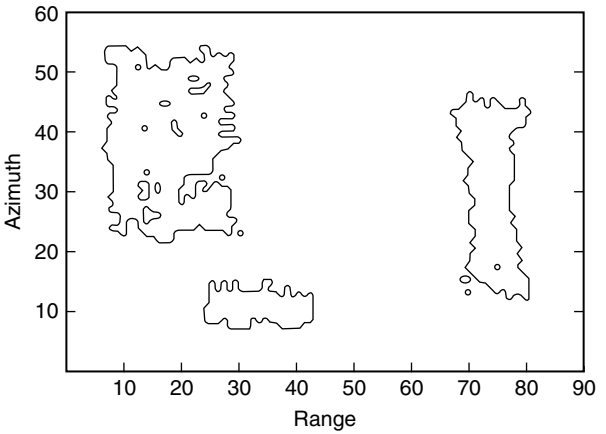

FIGURE 2.22 Example of a "corrected" range–azimuth volume.

of the CL edge cells are identified as BN. As an example, when NCQ = 8, only inner CL cells in the quantized volume are recognized as CL cells in the corrected volume and all CL edge cells in the quantized volume are recognized as BN cells in the corrected volume. In order to recover the edge cells, a second correction stage is needed where the first corrected volume will be transformed into a second corrected volume. Let the first corrected volume be referred to as the "corrected-quantized" volume (CQV) and the second "corrected" volume be referred to as the "corrected–corrected" volume (CCV). The following steps are used to transform the CQV into the CCV:

1. As explained in Section 2.7, choose the necessary number of CL neighboring cells, NCC, for a test cell in the CQV to be declared as a CL cell in the CCV. NCC can take one of the following values: 1, 2, 3, or 4.
2. For every test cell in the CQV count the number of neighboring CL cells. If the number is greater than or equal to NCC declare the test cell as a CL cell in the CCV. Otherwise, declare the test cell as a BN cell in the CCV.

Figure 2.23 shows the CCV obtained by transforming the CQV of Figure 2.22.

2.5.3.1.3. Assessment

Let BNQP, BNCQP, and BNCCP denote the percentage of BN cells in the "quantized," corrected–quantized and corrected–corrected volumes, respectively. BNQP is prespecified so as to determine the threshold for the quantized volume, whereas BNCQP and BNCCP are computed after the CQV and the CCV are obtained.

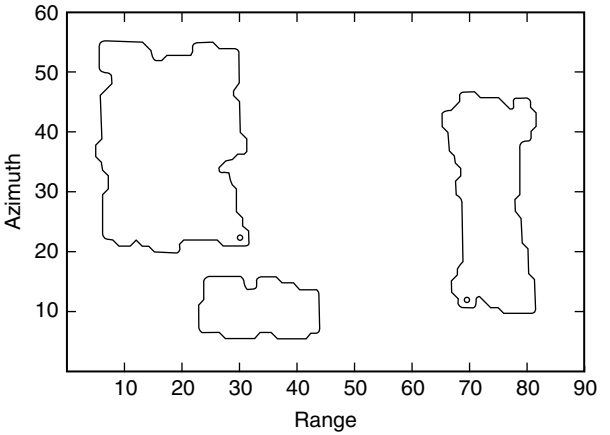

FIGURE 2.23 Example of the CCV corresponding to the CQV of Figure 2.22.

The assessment process consists of comparing BNCQP and BNCCP to BNQP in order to determine whether or not the percentages of the BN cells after correction are consistent with the percentage of BN cells in the quantized volume. When there is no consistency, further quantization, correction and assessment are performed until consistency is obtained.

The thresholding/quantization, first correction, second correction, and assessment stages are used to find the best threshold to separate between BN and CL patches. Once BNQP has been set, a threshold is computed. Then corrections are made to try and build the BN region and CL patches. The correction stages relabel some of the above-threshold cells as BN cells if they are likely to belong to the BN, and some of the below-threshold cells as CL cells if they are likely to belong to a CL patch, based on the choices for NCQ and NCC. Depending on how good or how bad of a choice is the threshold, many or few cells are relabeled, respectively. At the end of the procedure, BNCCP is computed and compared to BNQP. If the values are within a certain range, few cells would have been relabeled, the threshold is accepted and the assessment passes. Otherwise, many cells have been relabeled and the threshold is rejected. The iterative process then continues by setting another threshold through the choice of a new value for BNQP.

Rules for choosing NCQ, NCC, and BNQP and for determining when consistency of the percentages are obtained are explained in Section 2.7.

2.5.3.1.4. Smoothing

Examples have shown in many cases that when the percentages are consistent, CL declared patches may contain isolated BN declared cells. Because small powers can arise in a CL patch as explained in Section 2.5.2, it is most likely that isolated BN cells in CL patches are CL cells. The smoothing process is used to detect these isolated cells and label them adequately by transforming

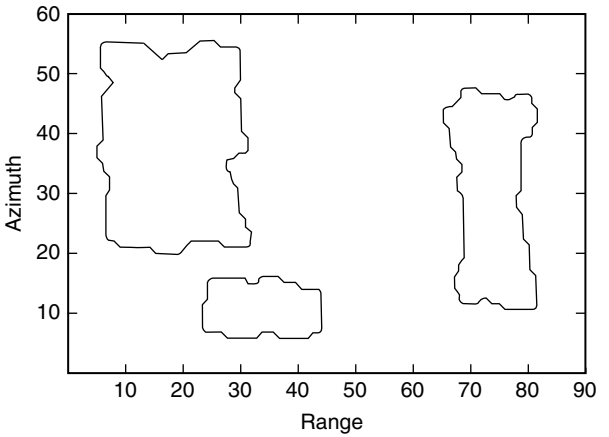

FIGURE 2.24 Example of the smoothed volume corresponding to the corrected-corrected volume (CCV) of Figure 2.23.

the CCV into a smoothed volume (SV). The smoothing technique consists of the following steps:

1. Choose the necessary number of CL neighboring cells, NS, for a BN identified test cell in the CCV to be declared as a CL cell in the SV where NS can take one of the following values: 5, 6, 7, or 8.
2. For every BN identified cell in the CCV count the number of neighboring CL cells. If the number is greater than or equal to NS, declare the test cell as a CL cell in the SV. Otherwise, declare the test cell as a BN cell in the SV.

Figure 2.24 shows an example of the SV corresponding to the CCV of Figure 2.23. Note that the CL patches are smooth and do not contain any isolated points.

2.5.3.2. Detection of CL Patch Edges and Edge Enhancement

2.5.3.2.1. Detection of CL Patch Edges

After smoothing, each cell in the SV has been declared as either a CL or BN cell. The next step is to determine which of the CL cells are located on the edges of the CL patches. This is important for subsequent radar signal processing if reference cells for estimating parameters of a test cell are to be chosen properly.

Identification of CL edge (CLE) cells is done by the use of an image processing technique referred to in the image processing literature as unsharp masking.[24,25] It consists of the following steps:

1. A weighting filter consisting of a three by three array of cells is constructed, as shown in Figure 2.25, where the center cell has a

$w(-1, -1) = -1$	$w(0, -1) = -1$	$w(1, -1) = -1$
$w(-1, 0) = -1$	$w(0, 0) = 8$	$w(1, 0) = -1$
$w(-1, 1) = -1$	$w(0, 1) = -1$	$w(1, 1) = -1$

FIGURE 2.25 Weighting filter.

weight given by $w(0, 0) = 8$ and the neighboring cells have weights given by $w(-1, -1) = w(0, -1) = w(1, -1) = w(-1, 0) = w(1, 0) = w(-1, 1) = w(0, 1) = w(1, 1) = -1$. The center cell is positioned on the test cell. Notice that the weights of the filter cells sum to zero. In particular,

$$\sum_{m=-1}^{1} \sum_{n=-1}^{1} w(m, n) = 0 \tag{2.25}$$

2. Assume the weighting filter is centered at the jkth cell in SV. The cells corresponding to the 3×3 array of the weighting filter have quantized values as illustrated in Figure 2.26. By definition,

$$\mathrm{SQ}(j, k) = \begin{cases} 1 & \text{if the } jk\text{th cell in SV is declared as CL} \\ 0 & \text{if the } jk\text{th cell in SV is declared as BN} \end{cases} \tag{2.26}$$

where $j = 1, 2, \ldots, J$ and $K = 1, 2, \ldots, K$. To avoid filter cells falling outside SV, the coordinates of the jkth cell at which the filter is centered are constrained to $j = 2, 3, \ldots, J - 1$, and $k = 2, 3, \ldots, K - 1$.

3. Form the sum

$$S = \sum_{m=-1}^{1} \sum_{n=-1}^{1} w(m, n)\mathrm{SQ}(j + m, k + n) \tag{2.27}$$

   – If S is equal to zero, all cells have the same assigned value. This can arise only when the test cell is not an edge cell.

SQ(j–1, k–1)	SQ(j,k–1)	SQ(j+1,k–1)
SQ(j–1, k)	SQ(j, k)	SQ(j+1, k)
SQ(j–1, k+1)	SQ(j, k+1)	SQ(j+1, k+1)

FIGURE 2.26 Quantized values of the 3×3 array corresponding to the jkth cell.

- If S is positive, the test cell is an edge cell and is labeled as such.
- If S is negative, the test cell cannot be an edge cell. On the other hand, one or more of the neighboring cells are guaranteed to be an edge cell.

The three situations are illustrated in Figure 2.27, where the black and white cells represent BN and CL cells, respectively. In Figure 2.27(a,b), $S = 0$ because all nine cells are in BN and CL, respectively. Observe that the test cell is not an edge cell. In Figure 2.27(c), $S = 4 > 0$. Note that the test cell is an edge cell. Finally, in Figure 2.27(d), $S = -2 < 0$ and the test cell is not an edge cell. At the end of the edge detection procedure, each cell in the original volume has been labeled as CL, BN or CLE cell. At this point, the mapping is done. The final volume is referred to as the mapped volume (MV).

2.5.3.2.2. Enhancement of CL Patch Edges

The edges detected after smoothing, tend not to follow the irregular edges which may actually exist. Consequently, the edges are further enhanced by examining the power levels of cells, just outside and on the edge cells. If the power levels of these cells exceed the threshold set in the quantization stage, those are declared as edge cells, otherwise, as BN cells.

At the end of edge enhancement procedure, edges are detected and each cell in the original volume is labeled as either CL, BN, or CLE cell.

2.5.3.3. Conclusion

As shown in Figure 2.28, the mapping procedure consists of the following steps: thresholding or quantization, correction, assessment, smoothing, edge

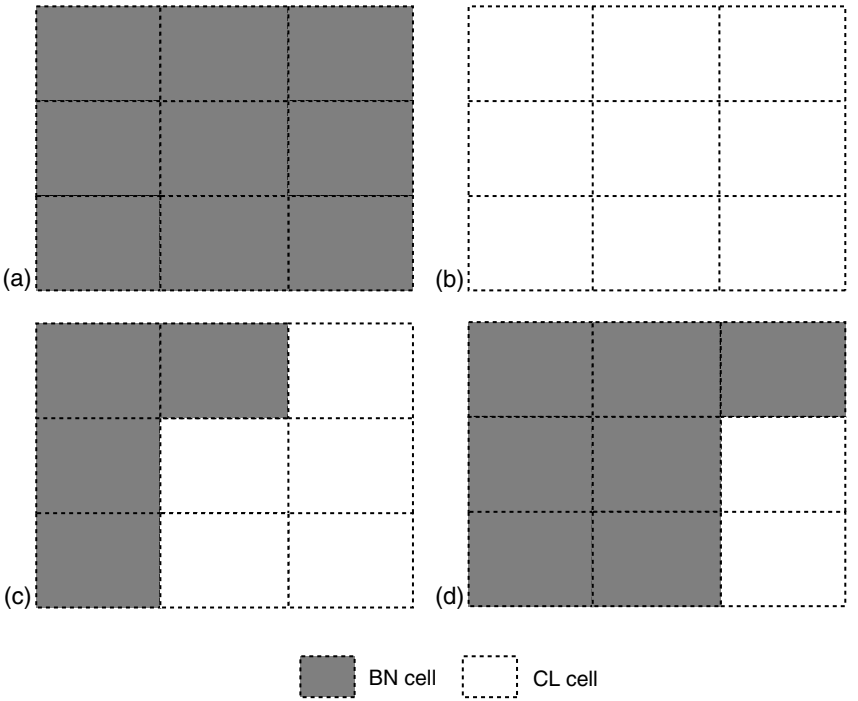

FIGURE 2.27 Example of unsharp masking technique. (a) $S = 0$, (b) $S = 0$, (c) $S > 0$, (d) $S < 0$.

enhancement, and edge detection. As will be explained later on in Section 2.5.7, the feedback loop which connects the assessment procedure to the thresholding or quantization, first correction, and second correction blocks, is controlled by IPUS.

2.5.4. Examples of the Mapping Procedure

Before presenting examples to illustrate how the mapping procedure performs a separation between BN and CL patches, a brief review of the mapping procedure is first given followed by three different examples of the mapping procedure. The generated scene and the distributions of CL patch data are shown for every example. Also, only the results corresponding to the initial and final choices of BNQP are illustrated in each case. The in-between settings of BNQP are listed in tables.

2.5.4.1. Introduction

The mapping procedure begins by selecting a threshold, such that, the percentage of BN cells relative to total number of cells is equal to a specified value, denoted

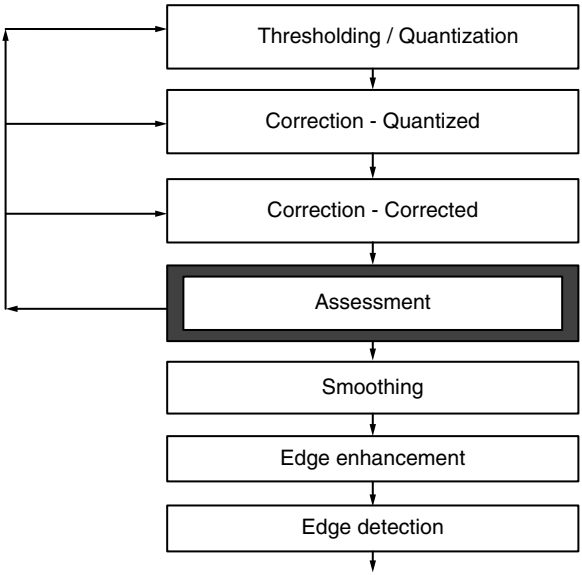

FIGURE 2.28 Block diagram of the mapping procedure.

by BNQP. Two correction stages then ensue. In the first correction stage, each cell in the quantized volume, denoted by QV, is tested by a three by three mask centered on the test cell. The test cell is labeled as BN, only if less than NCQ of the eight neighboring cells are declared as CL in the QV volume, where NCQ is a parameter specified by the user. In the following discussion it will be shown that, the first correction stage tries to restore the right tail of BN PDF, which had been severely distorted by the quantization. After the first correction stage, the corrected volume is denoted by CQV. The second correction stage attempts to correctly reclassify the edges of the CL patches. This is done by testing each cell in CQV, using once again, a three by three mask, centered on the test cell. The test cell is labeled as BN if less than NCC of the eight neighboring cells is declared as CL in the CQV volume where NCC is a parameter specified by user. Typical values for NCQ are 5, 6, 7, 8 while typical values for NCC are 1, 2, 3, 4. In the following discussion, it will be shown that the second correction stage attempts to restore the shapes of the BN and CL PDFs. After the second correction stage, the corrected volume is denoted by CCV. The percentage of BN cells relative to the total number of cells in the CCV volume is denoted by BNCCP. BNCCP is compared to BNQP. If the difference $|BNCCP - BNQP|$ is smaller than a prespecified value, the iteration process ends and some additional processing is done to enhance and label the edges. As explained in Section 2.7, if the difference is not too large, additional iterations are made with new values for NCQ and NCC. If these do not lead to convergence or if the difference is too large, the whole process is repeated by selecting a new threshold. If the difference

is large, the new value for BNQP is chosen to be the previous BNCCP. Otherwise, the new value of BNQP is chosen to be half way in between the previous values of BNQP and BNCCP.

2.5.4.2. Examples

In the following examples, it is assumed that the radar scans over 120° in the azimuthal direction and has a beam width of 2°. Hence, the azimuthal axis is subdivided into 60 bins. Also, the length of a range cell is given by $c\tau/2$ where $c = 3 \times 10^8$ m/s is the speed of the light and τ is the radar pulse width. Assuming a pulse width of $\tau = 1$ μs and a maximum range of $R_{\max} = 13.5$ km, the range axis is divided into $\{R_{\max}/[(c\tau)/2]\} = 90$ bins. Consequently, this subdivision of the entire surveillance volume yields a total of 5400 R/A cells.

In all examples presented in this dissertation CL plus noise data in a given cell are obtained by simply adding the envelopes of the CL and BN values for that cell. Of course, the envelope of the sum of two random processes does not equal to the sum of the envelopes. Nevertheless, it is possible to develop the important concepts of mapping and indexing using any set of random variables. The sum of envelopes approach was chosen in order to simplify the many computer simulations required by this research. Specifically, it was not necessary to first generate random processes whose envelopes are Weibull, Lognormal, and *K*-distributed.

The following three examples illustrate, respectively, the cases of three problems, *viz*., easy, difficult but resolvable, difficult and not resolvable. These examples are discussed throughout the work.

2.5.4.2.1. Example 1

Consider a surveillance volume containing four homogeneous CL patches, denoted by A, B, C, and D. CL patches C and D are contiguous and form a single nonhomogeneous CL patch C/D. Let CNR denote the CL to BN average power ratio. Table 2.2 lists the parameters of each CL patch. BN, which is spread throughout the surveillance volume, is Rayleigh-distributed.

In this example, 66.07% of the total scanned volume is occupied by BN alone. Also, CL patch C contains the same number of cells as CL patch D. Figures 2.29 and 2.30 show the CL patch boundaries and the 3-D envelope plot for the surveillance volume, respectively. The PDFs and histograms of the BN and CL patches are shown in Figure 2.31. Comparing Figure 2.31(a,b), notice that some of the BN data are larger than some of the CL data of CL patches A and B.

The iteration process begins with the threshold set such that 10% of the sorted data are below the threshold, i.e., BNQP = 10%. Although the value of 10% is arbitrary, the initial setting of BNQP should be low. With the setting at 10%, the contour plot of the quantized volume is shown in Figure 2.32. In this figure, the closed solid line contours surround the BN declared regions and the rest is CL. Notice that a lot of BN cells have data exceeding the threshold

TABLE 2.2
CL Patch Parameters (Example 1)

CL	CNR (dB)	Data Distribution	Shape Parameter	Number of Cells
A	10	Rayleigh	None	117
B	20	*K*-distributed	10.0	1423
C	30	Lognormal	0.01	146
D	40	Weibull	5.00	146

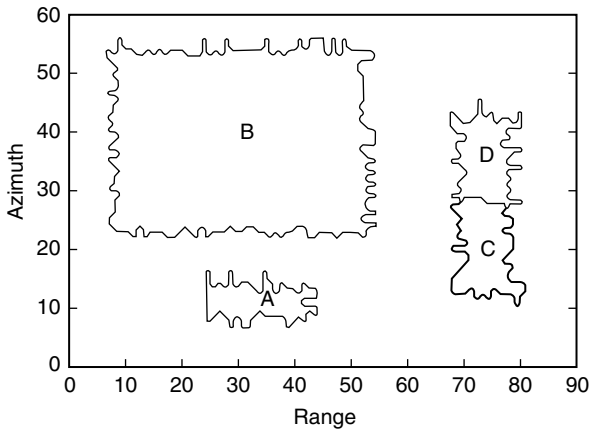

FIGURE 2.29 Boundaries of the CL regions (Example 1).

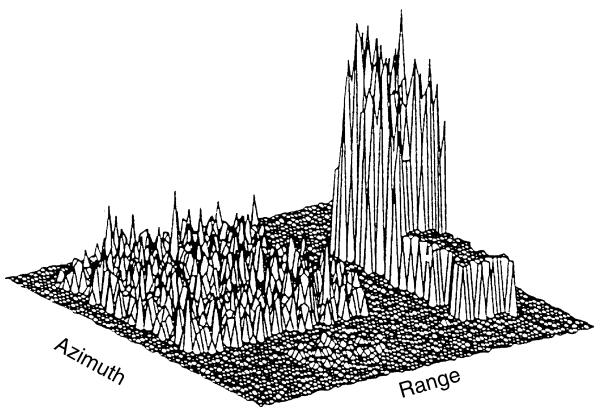

FIGURE 2.30 3-D envelope plot of the surveillance volume (Example 1).

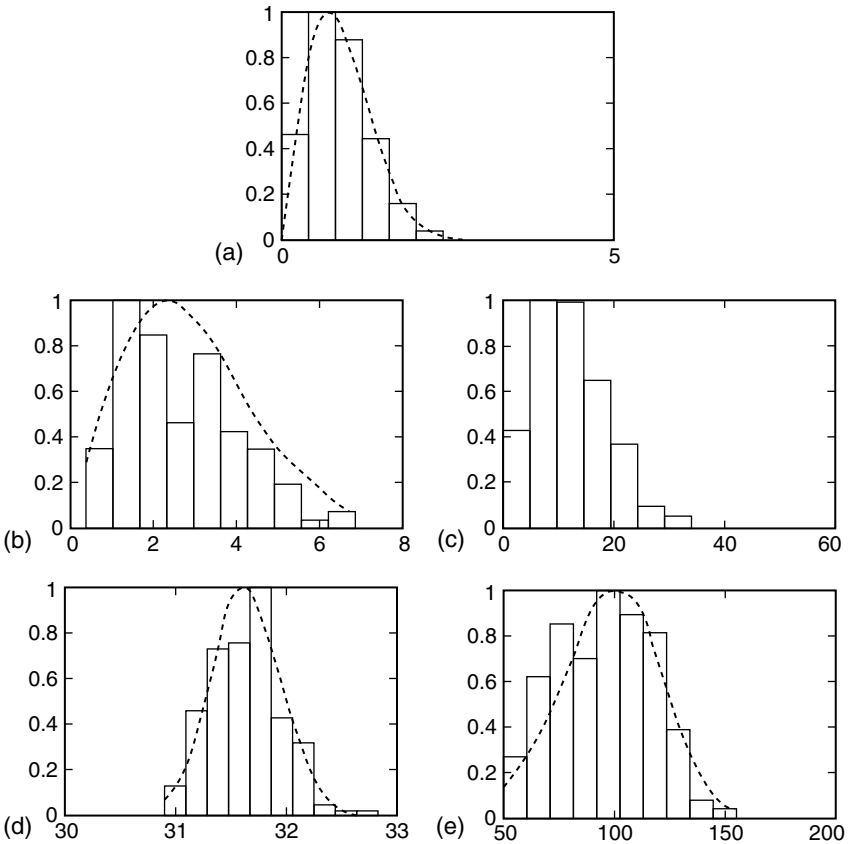

FIGURE 2.31 Probability density functions and histograms of the BN and CL patches (Example 1). (a) Rayleigh-distributed background noise, (b) Rayleigh-distributed CL patch A, (c) *K*-distributed CL patch B, (d) Lognormal-distributed CL patch C, (e) Weibull-distributed CL patch D.

because the threshold was set very low. This results in a lot of small BN regions. For the purpose of comparison, the boundaries of the original CL patches are shown in dotted lines. With, NCQ set to eight and NCC set to one, the contour plots resulting from the first and second corrections, respectively, are shown in Figures 2.33 and 2.34.

As shown in Table 2.3, when BNQP = 10% and (NCQ, NCC) = (8, 1), BNCCP = 22% which results in |BNQP − BNCCP| = 12% being large. A new threshold is then chosen so that BNQP = 22% which is the previous value obtained for BNCCP. The iterative process continues until it is found for BNQP = 63.37%, NCQ = 5, and NCC = 3, that |BNQP − BNCCP| is less than 1%. In particular, for these parameter values |BNQP − BNCCP| = |63.37 − 64.17| < 1%. At this stage the assessment passes and the iterative process stops.

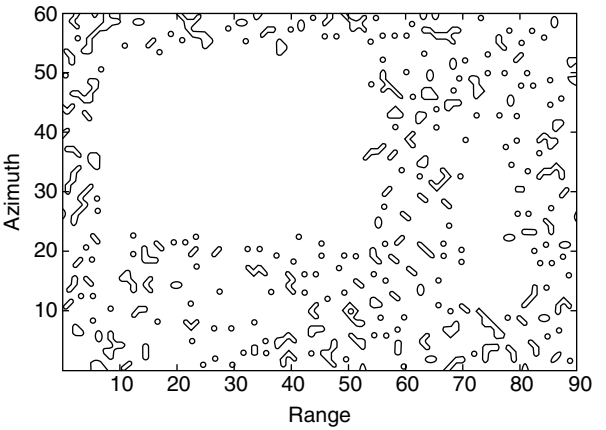

FIGURE 2.32 Contour plot of the quantized volume with BNQP = 10% (Example 1).

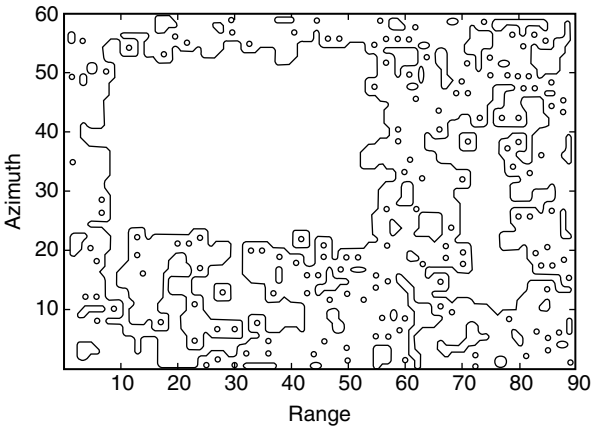

FIGURE 2.33 Contour plot of the first correction volume with NCQ = 8 (Example 1).

Figures 2.35–2.37 show the contour plots obtained after quantization, first correction and second correction, respectively, when BNQP = 63.37%, NCQ = 5, and NCC = 3. Note first from Figure 2.35 that even with this setting of BNQP, where the threshold is very close to its true value, some of the BN data exceed the threshold and form very small CL patches, and some of the CL data in every patch falls below the threshold causing holes in the CL patch. Figure 2.36 shows how most of the erroneously declared BN and CL cells have been correctly reclassified as CL and BN cells, respectively. Next, the edges are augmented as shown in Figure 2.37. However, comparing the resulting edge contour to the ideal contour shows that not all the edges have been recovered even though the CL patches have been identified.

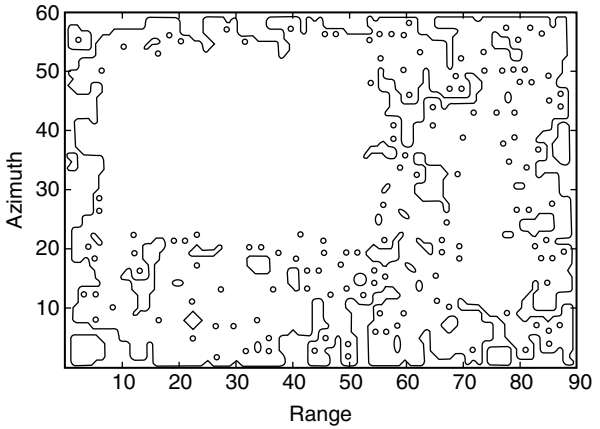

FIGURE 2.34 Contour plot of the second correction volume with NCC = 1 (Example 1).

Next, smoothing, edge enhancement, and edge detection are performed. The results are shown in Figures 2.38–2.40, respectively. First, comparing Figures 2.37 and 2.38, note that smoothing did not result in any change from the mapped volume. This is because the regions of Figure 2.37 do not contain any holes whereas the purpose of smoothing is to remove holes. Comparing the edge enhanced regions in Figure 2.39 (in solid lines) to the ideal one (in dotted lines), one can see how close the two sets of regions have become. In fact, at the end of edge enhancement, 65.54% of the total number of cells are declared BN where as ideally the BN percentage of the generated scene was 66.07%. Note that, only 10

TABLE 2.3
Settings of Percentage of BN Cells in the Quantized Volume (BNQP) (Example 1)

BNQP (%)	Parameter Values	
10.00	NCQ = 8	BNCQP = 51.56
(guess)	NCC = 1	BNCCP = 22.00
22.00	NCQ = 8	BNCQP = 68.28
latest BNCCP	NCC = 1	BNCCP = 50.39
50.39	NCQ = 7	BNCQP = 70.93
latest BNCCP	NCC = 1	BNCCP = 63.37
63.37	NCQ = 5	BNCQP = 67.35
latest BNCCP	NCC = 1	BNCCP = 59.72
63.37 same as	NCQ = 5	BNCQP = 67.35
latest BNQP	NCC = 3	BNCCP = 64.17

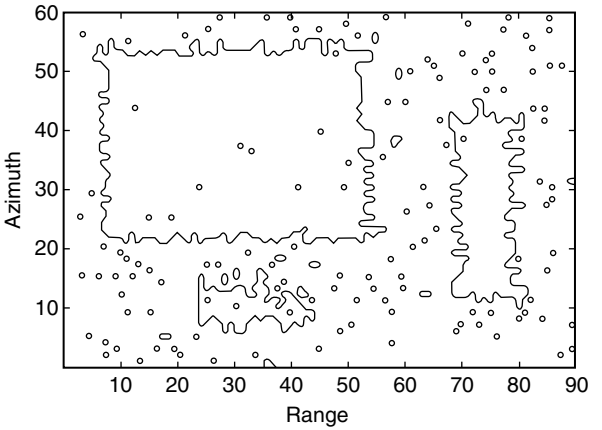

FIGURE 2.35 Contour plot of the quantized volume with BNQP = 63.37% (Example 1).

CL cells were misidentified and associated with the BN. Out of these, six had data values under the threshold (BNQP = 63.37%). Also, 39 BN cells were misidentified and associated with the CL. Of these, 21 had data values above the threshold. These results in the ratios

$$\frac{\text{misidentified CL cells below threshold}}{\text{total number of CL cells}} = 0.22\% \tag{2.28}$$

and

$$\frac{\text{misidentified BN cells below threshold}}{\text{total number of BN cells}} = 0.50\% \tag{2.29}$$

Figure 2.40 shows the edges in the edge detected volume.

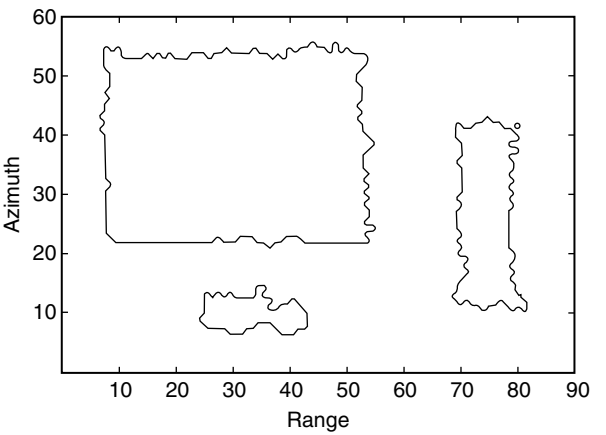

FIGURE 2.36 Contour plot of the first correction volume with NCQ = 5 (Example 1).

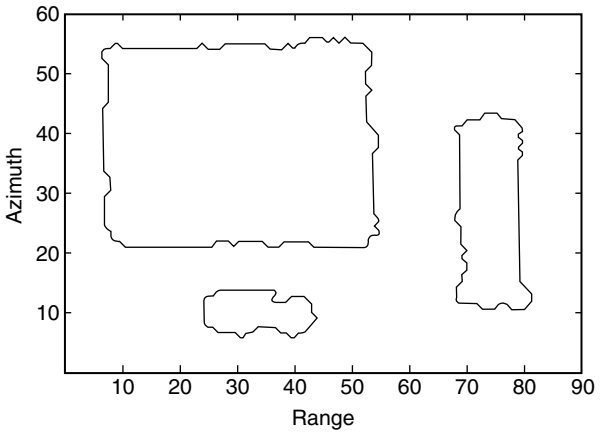

FIGURE 2.37 Contour plot of the second correction volume with NCC = 3 (Example 1).

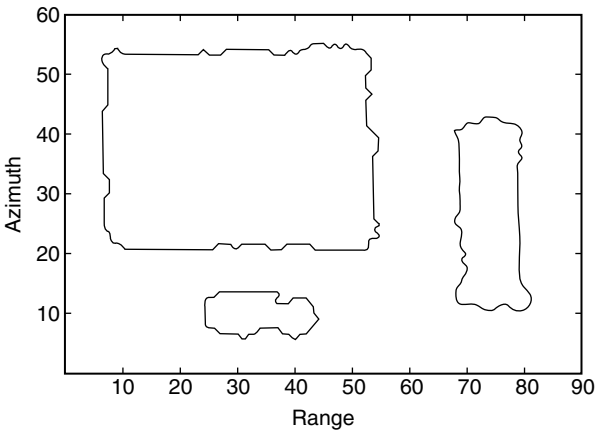

FIGURE 2.38 Contour plot of the smoothed volume with NS = 7 (Example 1).

An assessment of the mapping procedure results in three CL patches with the number of cells for each patch as listed in Table 2.4. By comparing Tables 2.2 and 2.4, you can notice how close the number of cells are.

2.5.4.2.2. Example 2

Consider, a surveillance volume containing four homogeneous CL patches, denoted by A, B, C, and D, as shown in Figure 2.41. CL patches C and D are contiguous and form a single nonhomogeneous CL patch C/D. Table 2.5 lists the parameters of each CL patch. In addition, the BN which is spread throughout the surveillance volume is Rayleigh-distributed.

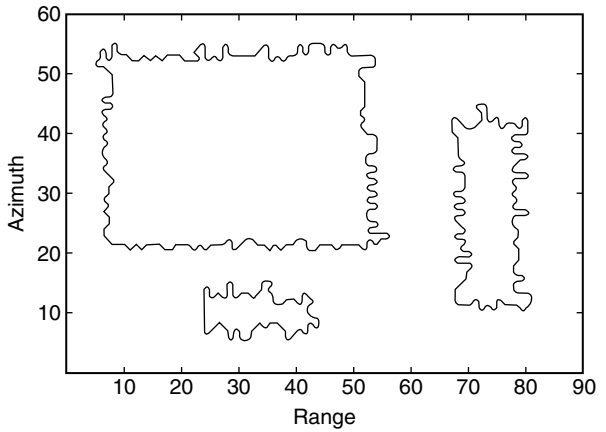

FIGURE 2.39 Contour plot of the edge enhanced volume (Example 1).

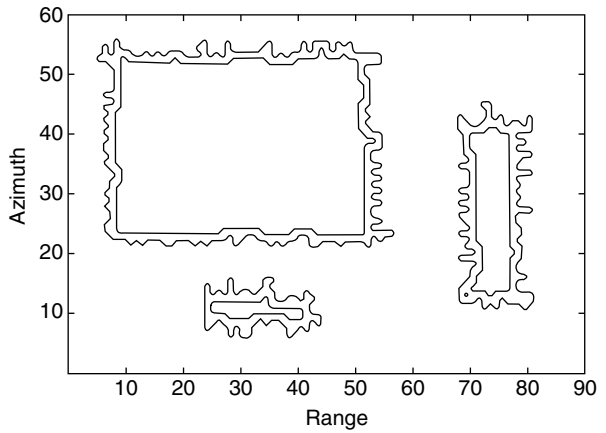

FIGURE 2.40 Contour plot of the edge selected volume (Example 1).

TABLE 2.4
CL Patch Parameters after Assessment (Example 1)

CL	Number of Cells
A	113
C + D	304
B	1444

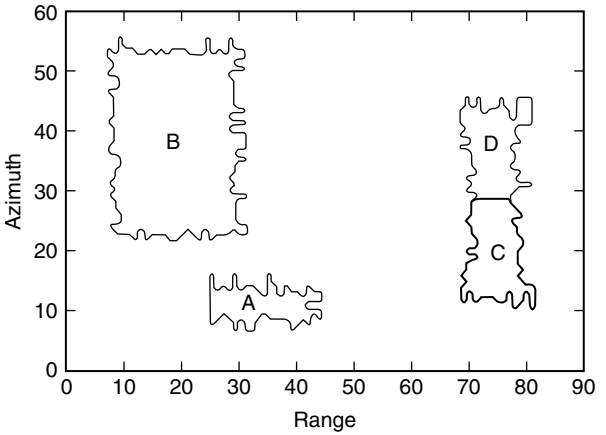

FIGURE 2.41 Boundaries of the CL regions (Example 2).

In this example, 79.80% of the total scanned volume is occupied by BN alone. Also, CL patch C contains approximately the same number of cells as CL patch D. Figures 2.41 and 2.42 show the CL patch boundaries and the 3-D envelope plot for the surveillance volume, respectively. The PDFs and histograms of the BN and CL patches are shown in Figure 2.43. In this example, not only is the CNR for every patch low but the shape parameters of the PDFs from which the CL data have been generated are such that the CL in patches A, B, and C is very spiky, as shown by the long tails in Figure 2.43(b–d). The shape of the histograms for CL patches A, B, and C is such that a lot of CL data overlap the BN data in value making it difficult to separate between the BN and CL patches just by setting a threshold.

The iteration process begins with the threshold, once again, arbitrarily set such that 10% of the sorted data are below the threshold, i.e., BNQP = 10%. With this setting, the contour plot of the quantized volume is shown in Figure 2.44. In this figure, the closed solid line contours surround the BN declared regions and the rest is CL. Notice that a lot of BN cells have data exceeding the threshold because the threshold was set very low. This results in

TABLE 2.5
CL Patch Parameters (Example 2)

CL	CNR (dB)	Data Distribution	Shape Parameter	Number of Cells
A	10	*K*-distributed	1.5	117
B	10	Weibull	1.0	672
C	10	Lognormal	1.0	151
D	10	Rayleigh	None	151

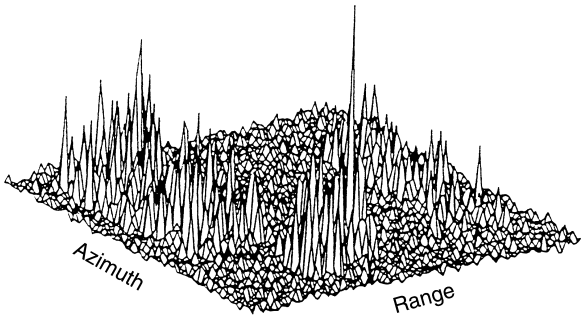

FIGURE 2.42 3-D Envelope plot of the surveillance volume (Example 2).

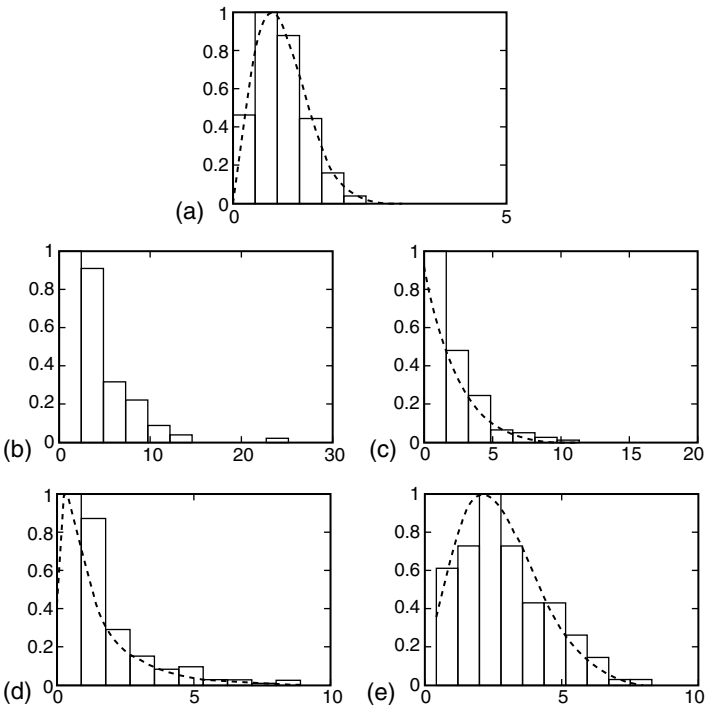

FIGURE 2.43 Probability density functions and histograms of the BN and CL patches (Example 2). (a) Rayleigh-distributed background noise, (b) Rayleigh-distributed CL patch A, (c) *K*-distributed CL patch B, (d) Lognormal-distributed CL patch C, (e) Weibull-distributed CL patch D.

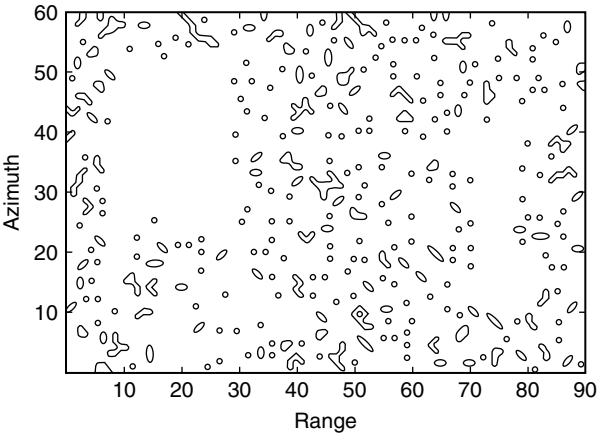

FIGURE 2.44 Contour plot of the quantized volume with BNQP = 10% (Example 2).

a lot of small BN regions. For the purpose of comparison, the boundaries of the original CL patches are shown in dotted lines. With, NCQ set to 8 and NCC set to 1, the contour plots resulting from the first and second corrections, respectively, are shown in Figures 2.45 and 2.46.

As shown in Table 2.6, when BNQP = 10% and (NCQ, NCC) = (8, 1), BNCCP = 20.59% which results in |BNQP − BNCCP| = 10.59% being large. A new threshold is then chosen so that BNQP = 20.59% which is the previous value obtained for BNCCP. The iterative process continues until it is found for BNQP = 75.78%, NCQ = 5, and NCC = 1, that |BNQP − BNCCP| < 1%. In particular, for these parameter values |BNQP − BNCCP| =

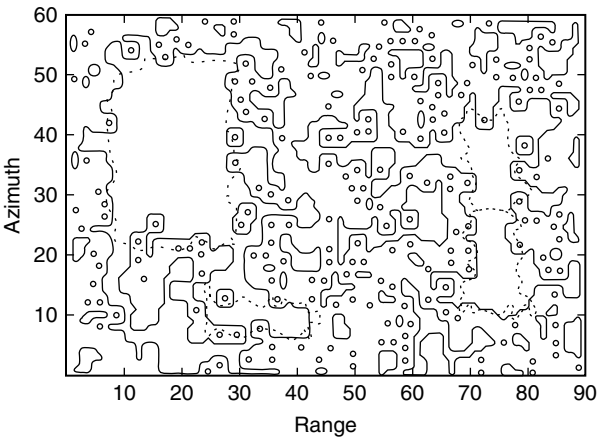

FIGURE 2.45 Contour plot of the first correction volume with NCQ = 8 (Example 2).

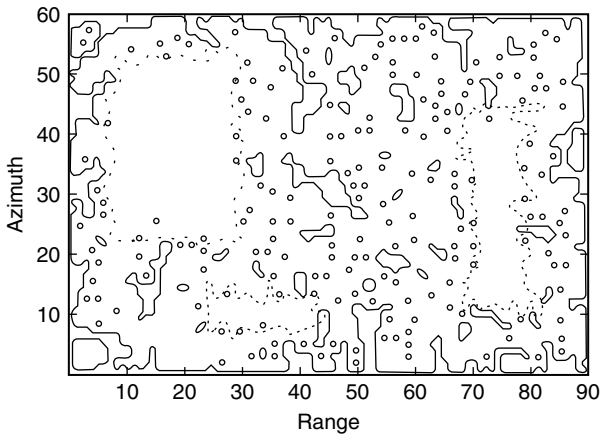

FIGURE 2.46 Contour plot of the second correction volume with NCC = 1 (Example 2).

$|75.78 - 75.31| < 1\%$. At this stage the assessment passes and the iterative process stops.

Figures 2.47–2.49 show the contour plots obtained after quantization, first correction and second correction, respectively, when BNQP = 75.78%, NCQ = 5, and NCC = 1. First, note from Figure 2.47 that even with this setting of BNQP where the threshold is very close to its true value, a lot of BN data

TABLE 2.6
Settings of Percentage of BN Cells in the Quantized Volume (BNQP) (Example 2)

BNQP (%)	Parameter Values	
10.00	NCQ = 8	BNCQP = 56.35
(guess)	NCC = 1	BNCCP = 20.59
20.59	NCQ = 8	BNCQP = 77.43
latest BNCCP	NCC = 1	BNCCP = 48.04
48.04	NCQ = 7	BNCQP = 84.11
latest BNCCP	NCC = 1	BNCCP = 72.30
72.30	NCQ = 5	BNCQP = 82.17
latest BNCCP	NCC = 1	BNCCP = 73.94
72.30 same as	NCQ = 5	BNCQP = 82.17
latest BNQP	NCC = 3	BNCCP = 79.26
75.78	NCQ = 5	BNCQP = 83.50
half way rule	NCC = 1	BNCCP = 75.31

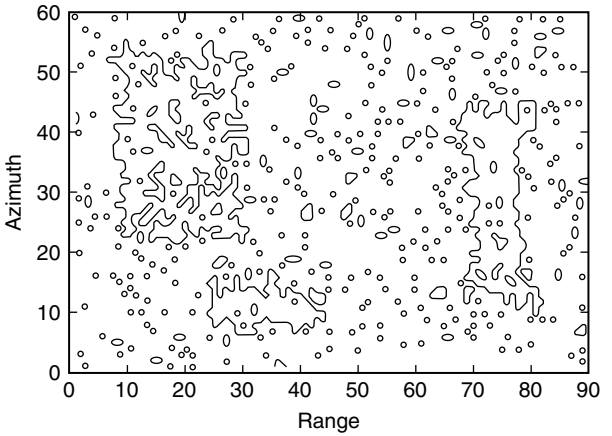

FIGURE 2.47 Contour plot of the quantized volume with BNQP = 75.78% (Example 2).

exceed the threshold and form very small CL patches, and also, a lot of the CL data in every patch falls below the threshold causing holes in the CL patch. Figure 2.48 shows how most of the erroneously declared BN and CL cells have been correctly reclassified as CL and BN cells, respectively. Next, the edges are augmented as shown in Figure 2.49. Comparing the resulting edge contour to the ideal one, it shows that not all the edges have been completely recovered even though the CL patches have been identified. Also, notice in this case, the CL patches are not homogeneous and contain holes in them.

Next, smoothing, edge enhancement and edge detection are performed. The results are shown in Figures 2.50–2.52, respectively. First, on correcting

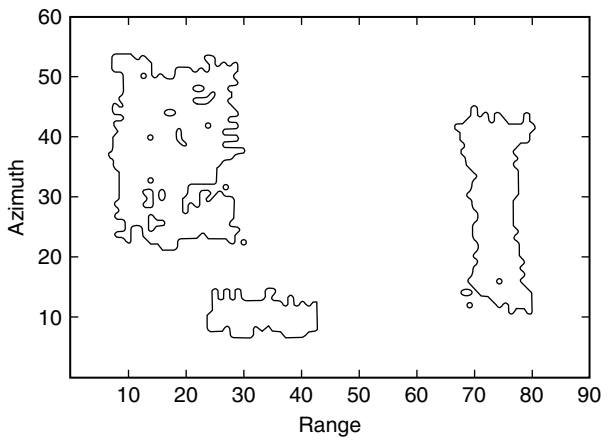

FIGURE 2.48 Contour plot of the first correction volume with NCQ = 5 (Example 2).

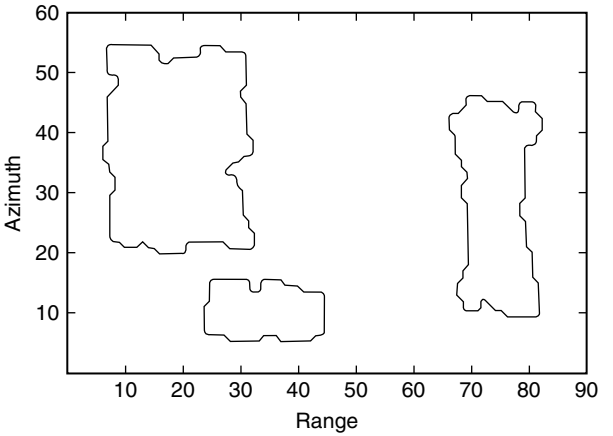

FIGURE 2.49 Contour plot of the second correction volume with NCC = 1 (Example 2).

Figures 2.49 and 2.50, note that the smoothing reclassified the BN declared cells causing holes inside the CL as CL cells. Comparing the edge enhanced regions in Figure 2.51 (in solid lines) with the ideal one (in dotted lines), one can see how close the two sets of regions have become. In fact, at the end of edge enhancement, 80.31% of total number of cells are declared BN where as, ideally the BN percentage of generated scene was 79.80%. Note that, 86 CL cells were misidentified and associated with the BN. Out of these 72 had data values under the threshold (BNQP = 79.59%). Also, 58 BN cells were misidentified and associated with the CL. Of these, 32 had data values above the threshold.

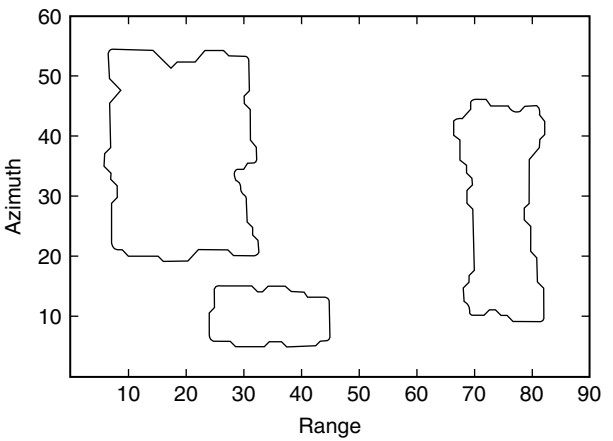

FIGURE 2.50 Contour plot of the smoothed volume with NS = 7 (Example 2).

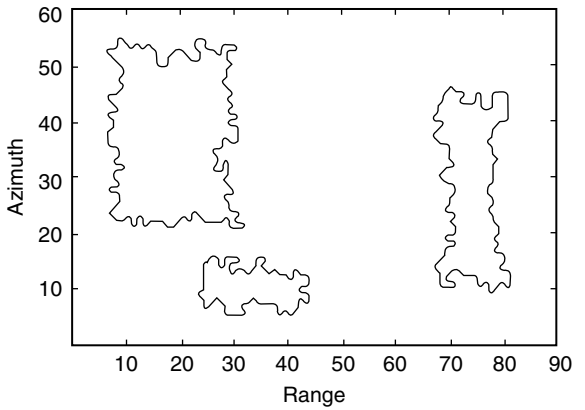

FIGURE 2.51 Contour plot of the edge enhanced volume (Example 2).

These result in

$$\frac{\text{misidentified CL cells above the threshold}}{\text{total number of CL cells}} = 1.28\% \tag{2.30}$$

$$\frac{\text{misidentified BN cells below the threshold}}{\text{total number of BN cells}} = 0.60\% \tag{2.31}$$

Figure 2.52 shows the edges of the edge detected volume.

An assessment of the mapping procedure results in three CL patches with the number of cells for each patch as listed in Table 2.7. By comparing Tables 2.5 and 2.7, you can notice how close the number of cells are.

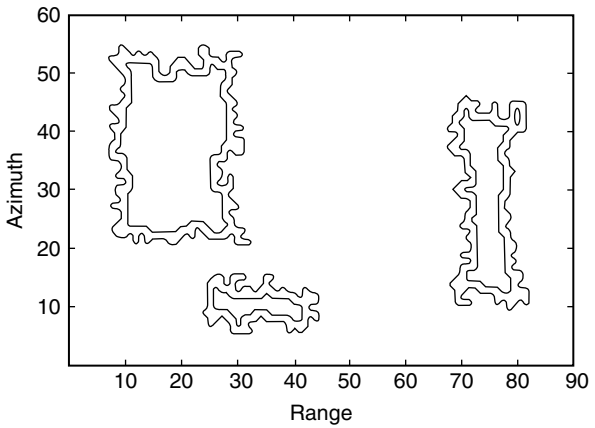

FIGURE 2.52 Contour plot of the edge detected volume (Example 2).

TABLE 2.7
CL Patch Parameters after Assessment (Example 2)

CL	Number of Cells
A	133
C + D	308
B	622

2.5.4.2.3. Example 3

Consider a surveillance volume containing four homogeneous CL patches, denoted by A, B, C, and D, as shown in Figure 2.53. CL patches C and D are contiguous and form a single nonhomogeneous CL patch C/D. Table 2.8 lists the parameters of each CL patch. In addition, the BN is Rayleigh-distributed.

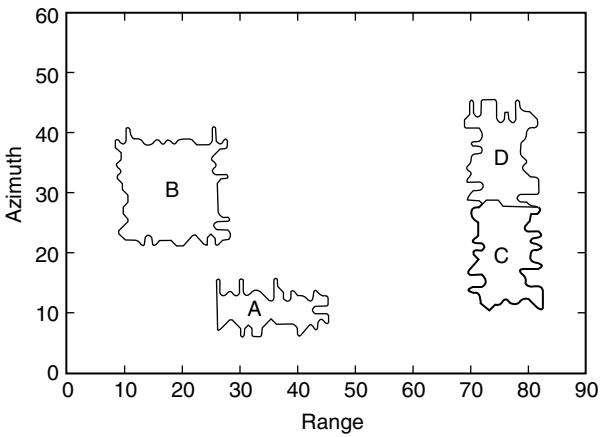

FIGURE 2.53 Boundaries of the CL regions (Example 3).

TABLE 2.8
CL Patch Parameters (Example 3)

CL	CNR (dB)	Data Distribution	Shape Parameter	Number of Cells
A	10	Rayleigh	None	117
B	10	*K*	10.0	298
C	10	Lognormal	0.01	161
D	10	Weibull	10.0	162

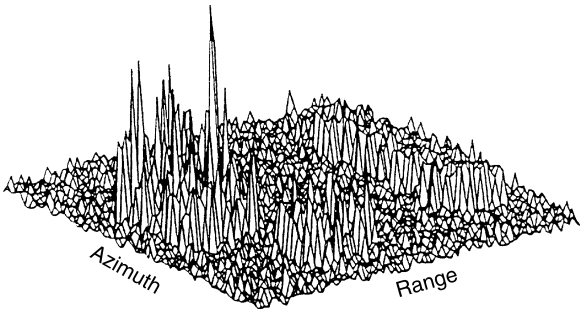

FIGURE 2.54 3-D envelope plot of the surveillance volume (Example 3).

In this example, 86.33% of the total scanned volume is occupied by BN alone. Also, CL patch C contains approximately the same number of cells as CL patch D. Figures 2.53 and 2.54, show the CL patch boundaries and the 3-D plots of the surveillance volume, respectively. Notice in Figure 2.54 how small are some of the CL data compared to the BN data and how large are some of the BN data compared to the CL data. This is due to the fact that the CNR for each patch is low. (Assuming an SNR of 10 dB in order to obtain reasonable detections in the absence of CL, CNR = 10 dB implies equal signal and CL powers. For values of CNR less than 10 dB, the signal would be larger than CL and no WSC would be available.) In this example, the mapping procedure is tested on a situation where CNR for every CL patch is low. The PDFs and histograms of the BN and CL patches are shown in Figure 2.55. Comparing the histograms in Figure 2.55(a–c,e), notice that, some of the BN data are higher than some of the CL data in patches A, B, and D. Also, observe that overlaps between histograms of the BN with CL patches A and B are noticeable.

The iteration process begins with the threshold, once again, it is set in such a way that 10% of the sorted data are below the threshold, i.e., BNQP = 10%. With this setting, the contour plot of the quantized volume is shown in Figure 2.56. In this figure, the closed solid line contours surround the BN declared regions and the rest is CL. Notice that a lot of BN cells have data that exceed the threshold because the threshold was set very low. This results in a lot of small BN regions. For purpose of comparison, the boundaries of the original CL patches are shown in dotted lines. With NCQ set to eight and NCC set to one, the contour plots resulting from the first and second corrections are shown in Figures 2.57 and 2.58, respectively.

As shown in Table 2.9, when BNQP = 10% and (NCQ, NCC) = (8, 1), BNCCP = 20.04% which results in |BNQP − BNCCP| = 10.04% being large. A new threshold is then chosen so that BNQP = 20.04% which was the previous value obtained for BNCCP. The iterative process continues until it is found for BNQP = 84.70%, NCQ = 5, and NCC = 3, that |BNQP − BNCCP| < 1%. For these parameter values, |BNQP − BNCCP| = |84.70 − 85.46| < 1%. At this stage the assessment passes and the iterative process stops.

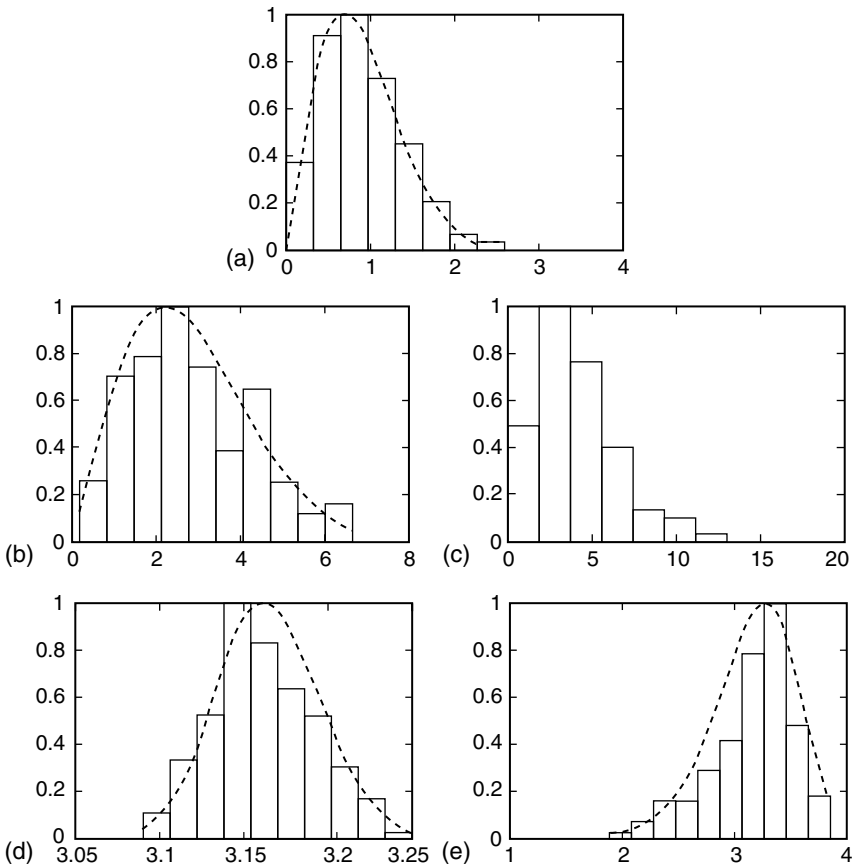

FIGURE 2.55 Probability density functions and histograms of the BN and CL patches (Example 3). (a) Rayleigh-distributed background noise, (b) Rayleigh-distributed CL patch A, (c) *K*-distributed CL patch B, (d) Lognormal-distributed CL patch C, (e) Weibull-distributed CL patch D.

Figures 2.59–2.61 show the contour plots obtained after quantization, first correction and second correction, respectively, when BNQP = 84.70%, NCQ = 5, and NCC = 3. First, note from Figure 2.59 that even with this setting of BNQP, where the threshold is very close to its true value, some of the BN data exceed the threshold and form very small CL patches and some of the CL data in every patch falls below the threshold causing holes in the CL patch. Figure 2.60 shows how most of the erroneously declared BN and CL cells have been correctly reclassified as BN and CL cells, respectively. Next, the edges are augmented as shown in Figure 2.61. Comparison of the resulting contour with the ideal one, shows that even though the CL patches have been identified, the edges are not completely recovered.

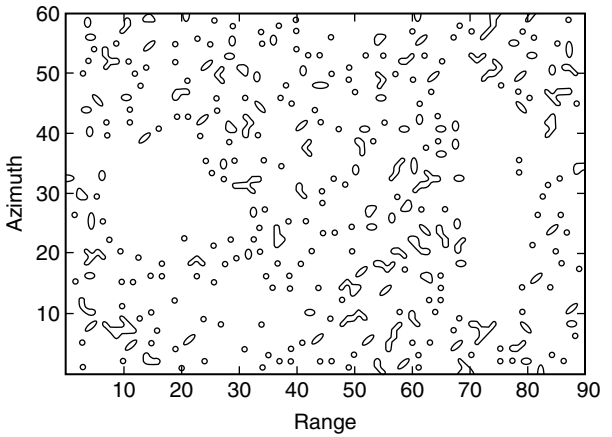

FIGURE 2.56 Contour plot of the quantized volume with BNQP = 10% (Example 3).

Next, smoothing, edge enhancement, and edge detection are performed. The results are shown in Figures 2.62–2.64, respectively. First, comparing Figures 2.61 and 2.62, note that smoothing did not result in any change from the mapped volume. This is because the regions of Figure 2.61 do not contain any holes whereas the purpose of smoothing is to remove holes. Comparing the edge enhanced regions in Figure 2.63 (in solid lines) with the ideal one (in dotted lines), one can see how close the two sets of regions have become. In fact, at the end of edge enhancement, 85.72% of the total number of cells are declared BN where as, ideally the BN percentage of the generated scene was 86.33%. Note that only six CL cells were misidentified and associated with the BN. Out of these, five had data values under the threshold (BNQP = 84.70%). Also, 39 BN

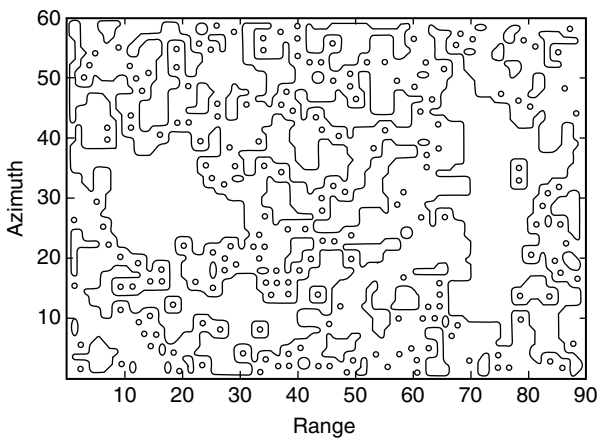

FIGURE 2.57 Contour plot of the first correction volume with NCQ = 8 (Example 3).

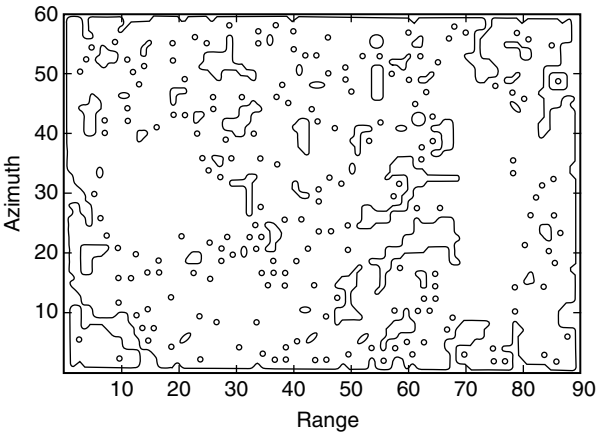

FIGURE 2.58 Contour plot of the second correction volume with NCC = 1 (Example 3).

cells were misidentified and associated with the CL. Of these, 10 had data values above the threshold. These result in

$$\frac{\text{misidentified CL cells above the threshold}}{\text{total number of CL cells}} = 0.13\% \qquad (2.32)$$

TABLE 2.9
Settings of Percentage of BN Cells in the Quantized Volume (BNQP) (Example 3)

BNQP (%)	Parameter Values	
10.00	NCQ = 8	BNCQP = 56.17
(guess)	NCC = 1	BNCCP = 20.04
20.04	NCQ = 8	BNCQP = 77.98
latest BNCCP	NCC = 1	BNCCP = 43.78
43.78	NCQ = 8	BNCQP = 90.78
latest BNCCP	NCC = 1	BNCCP = 82.65
82.65	NCQ = 7	BNCQP = 91.17
latest BNCCP	NCC = 1	BNCCP = 84.70
84.70	NCQ = 5	BNCQP = 87.61
latest BNCCP	NCC = 1	BNCCP = 82.00
84.70 same as	NCQ = 5	BNCQP = 87.61
latest BNQP	NCC = 3	BNCCP = 85.46

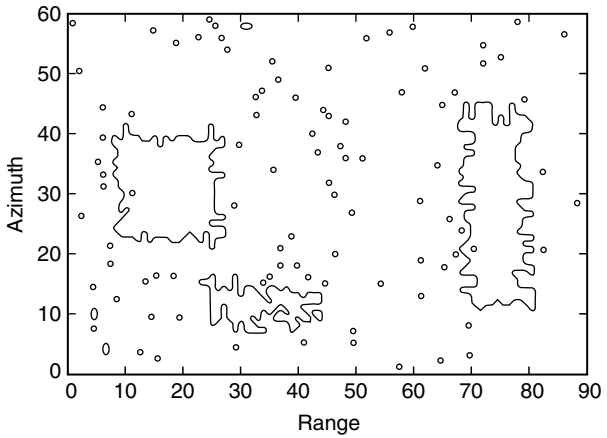

FIGURE 2.59 Contour plot of the quantized volume with BNQP = 84.70% (Example 3).

and

$$\frac{\text{misidentified BN cells below the threshold}}{\text{total number of BN cells}} = 0.62\% \qquad (2.33)$$

Figure 2.64 shows the edges of the edge detected volume.

An assessment of the mapping procedure results in three CL patches with the number of cells for each patch as listed in Table 2.10. By comparing Tables 2.8 and 2.10, you can notice how close the number of cells are.

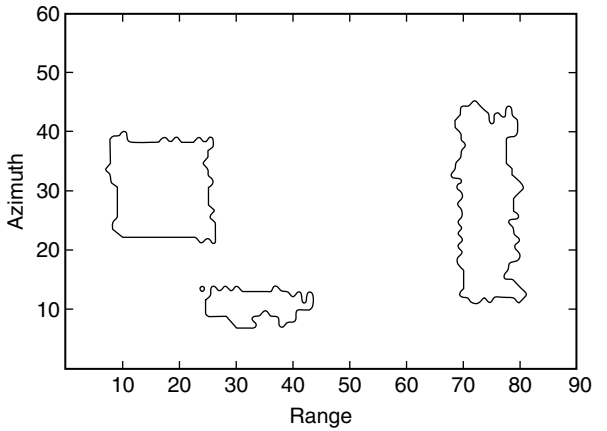

FIGURE 2.60 Contour plot of the first correction volume with NCQ = 5 (Example 3).

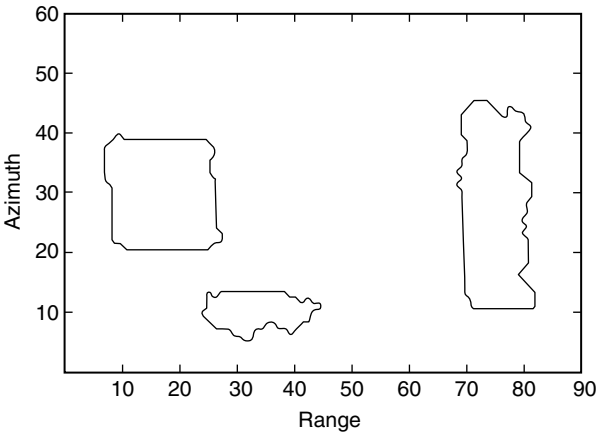

FIGURE 2.61 Contour plot of the second correction volume with NCC = 3 (Example 3).

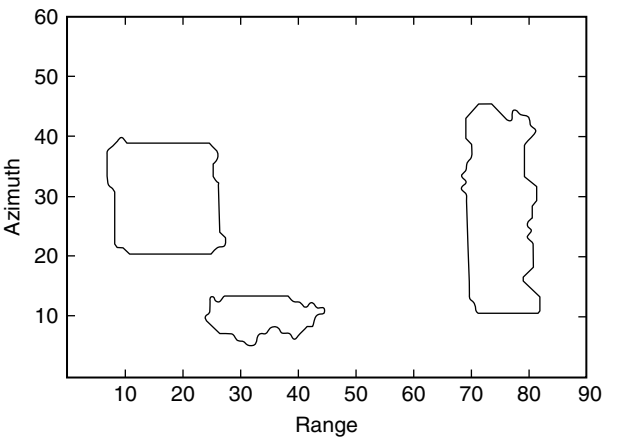

FIGURE 2.62 Contour plot of the smoothed volume with NS = 7 (Example 3).

2.5.5. Convergence of the Mapping Procedure

2.5.5.1. Introduction

Consider an image containing two regions where the PDFs for each region have nicely separated peaks as shown in Figure 2.65(a) and the overall PDF for both regions is as shown in Figure 2.65(b). In practice, a histogram is generated which approximates the overall PDF. Note that the individual PDF of each region is unknown. Because the individual PDFs are adequately separated, the overall histogram will be bimodal and separation between the two regions is readily obtained by placing the threshold $T1$ between the two peaks as shown

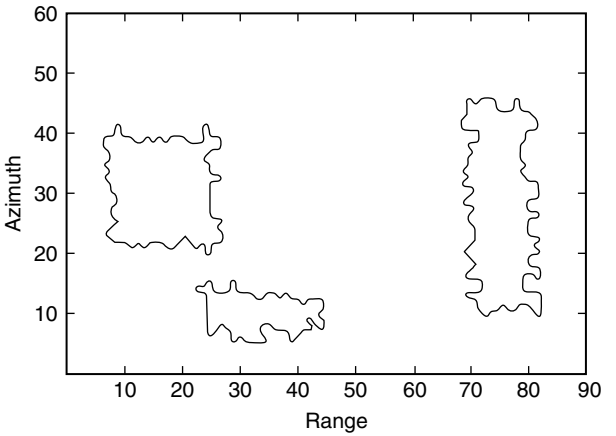

FIGURE 2.63 Contour plot of the edge enhanced volume (Example 3).

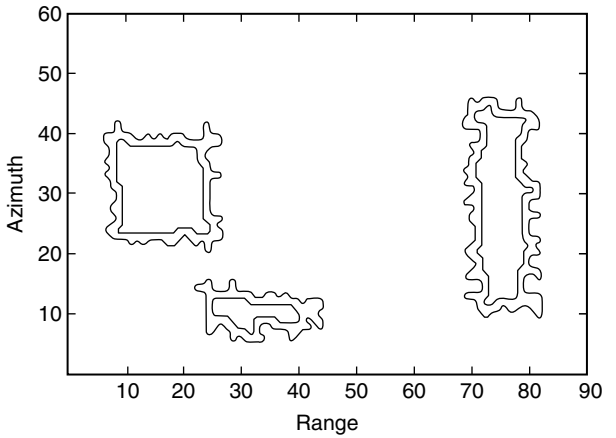

FIGURE 2.64 Contour plot of the edge selected volume (Example 3).

TABLE 2.10
CL Patch Parameters after Assessment (Example 3)

CL	Number of Cells
A	124
C + D	341
B	306

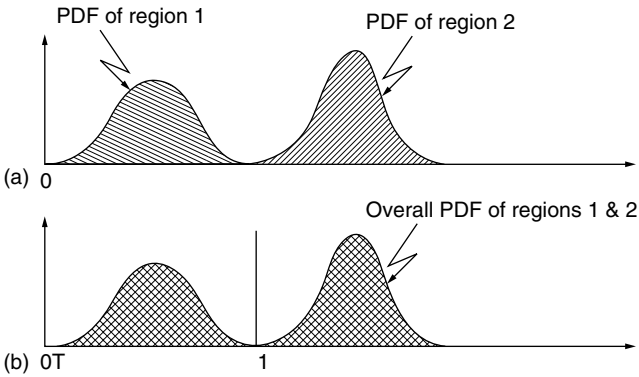

FIGURE 2.65 Nonoverlapping PDFs of two distinct regions. (a) Individual PDFs for each region. (b) Overall PDF for both regions.

in Figure 2.65(b). Cells with data values lower than $T1$ are declared to belonging to region one, while cells with data values higher than $T1$ are declared to belong to region two.

Now consider the slightly overlapping PDFs as shown in Figure 2.66(a,b). Although the overall PDF of the data regions is again bimodal, a noticeable overlap now exists between the tails. Once again, a threshold $T1$ is used to separate between the two regions. However, now a significant number of cells will be misclassified and corrections should be made to the extent possible.

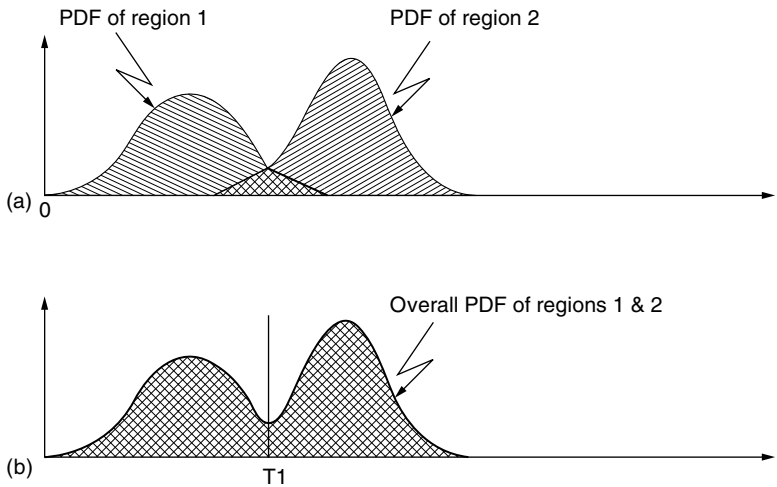

FIGURE 2.66 Overlapping PDFs of two distinct regions with a small overlapping area. (a) Individual PDFs for each region. (b) Overall PDF for both regions.

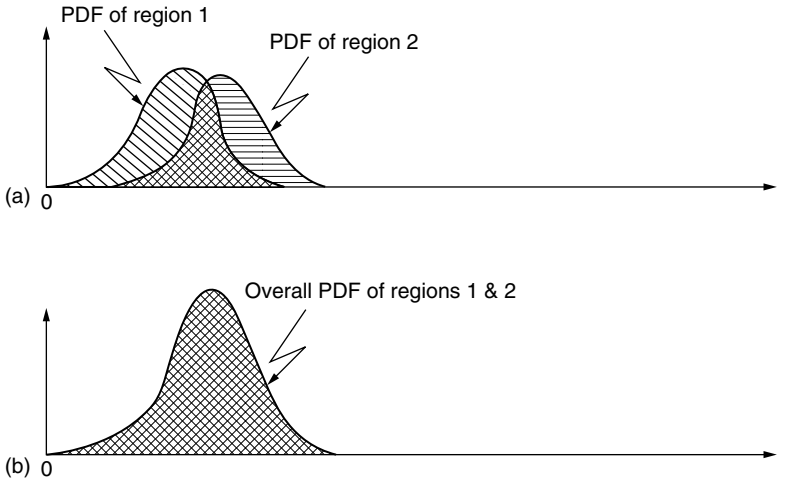

FIGURE 2.67 Overlapping PDFs of two distinct regions with a big overlapping area. (a) Individual PDFs for each region. (b) Overall PDF for both regions.

Figure 2.67 shows a more complicated case where the two regions now have major overlap between the tails. The overall PDF of the data from both regions is now unimodal and it is not possible to choose a threshold which separates the two regions without significant misclassifications.

In this section, it is shown that the mapping procedure described previously can adaptively choose a threshold and correct misclassifications so as to obtain good approximations to the PDFs of each region. The mapping procedure enables the region having the smallest envelopes, on average, to be separated from the remaining regions. By successive application of the mapping procedure, it is possible to first separate out the region with smallest envelopes, followed by the region with next smallest envelopes, and so forth. In the first application of the mapping procedure to a radar surveillance volume, region one consists of the BN while region two consists of the entire set of CL patches.

2.5.5.2. Separation between BN and CL Patches

An important observation relative to selection of the threshold, via specification of BNQP, is made when considering a surveillance volume consisting of only BN. The objective is to determine the range of threshold settings, for which mapping procedure declares the entire surveillance volume as BN. For this purpose, we consider the situation where entire scene consists of BN and the data are generated from a Rayleigh PDF.

For the scenario under consideration, the whole scene is homogeneous BN and the mapping, if done correctly, should identify the entire volume as a single BN region. Table 2.11 shows the values of BNCCP obtained for different settings of BNQP. Notice from the table that, a single region results only when threshold

TABLE 2.11
BN Percentages and Threshold Values Corresponding to the Rayleigh Distributed Background Noise

BNQP (%)	Threshold Value	BNCCP (%)
10	0.32	19.48
20	0.47	42.80
30	0.59	71.61
40	0.72	89.06
50	0.84	96.26
60	0.95	99.85
70	1.10	100.00
80	1.29	100.00
90	1.53	100.00

exceeds 70% of the data. As shown in Figure 2.68, the mapping procedure correctly reclassifies all cells corresponding to values above the threshold as long as the same is sufficiently towards the tail of PDF. In general, it is observed that the mapping procedure works best provided the threshold is adequately positioned towards the tail of PDF.

Another parameter that arises in the mapping procedure is BNCQP which denotes the percentage of BN cells after the first correction relative to the total number of cells in the surveillance volume. To gain insight into the relationships among BNQP, BNCQP, and BNCCP, we return to example number two, previously discussed in Section 2.5.4.2.2, where a lot of CL patch data overlap with the BN data in a manner similar to Figure 2.70. In Table 2.12, different values of these parameters are tabulated as the mapping procedure converges to the end result. If the test cell is to be declared as CL, recall that at least NCQ and NCC neighboring cells are required to be declared as CL in the QV and CQV during the first and second corrections, respectively.

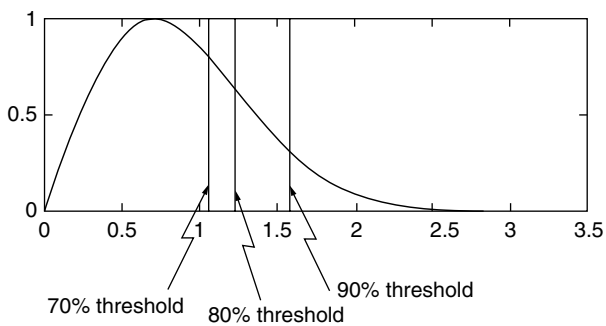

FIGURE 2.68 PDF of the Rayleigh-distributed patch.

TABLE 2.12
BN Percentages (Example 2)

BNQP (%) (NCQ, NCC)	BNCQP (%)	BNCCP (%)	|BNCQP − BNCCP|	|BNCCP − BNQP|
10.00 (8, 1)	56.35	20.59	35.76	10.59
20.59 (8, 1)	77.43	48.04	29.39	27.45
48.04 (7, 1)	84.11	72.30	11.81	24.26
72.30 (5, 1)	82.17	73.94	8.23	1.64
72.30 (5, 3)	82.17	79.26	2.91	6.96
75.78 (5, 1)	83.50	75.31	8.19	0.27

Table 2.13 tracks the mapping procedure during the first correction stage (denoted by QV → CQV), during the second correction stage (denoted by CQV → CCV) and at the end of the two correction stages (denoted by QV → CCV). All percentages given are with respect to the total number of cells in the surveillance volume. Initially, the threshold is set such that BNQP percentage of the total number of cells is below the threshold. The first correction stage requires that at least NCQ of the neighboring cells be above the threshold if the test cell is to be classified as a CL cell. Under the column headed by QV → CQV, $(CL \rightarrow BN)_1$ denotes the percentage of the total number of cells in the surveillance volume that were above the threshold but are reclassified as BN cells during the first correction stage. Similarly, $(BN \rightarrow CL)_1$ denotes the percentage of the total number of cells in the surveillance volume that were below the threshold but are reclassified as CL cells after the first correction stage. Note that the difference, $(CL \rightarrow BN)_1 - (BN \rightarrow CL)_1$, is the net percentage of the total number of cells in the surveillance volume that have been reclassified from CL to BN cells after the first correction stage. Similar statements apply for: (1) the second correction stage to $(CL \rightarrow BN)_2$,

TABLE 2.13
CL-to-BN and BN-to-CL Transitions (Example 2)

BNQP (%) (NCQ, NCC)	QV → CQV		CQV → CCV		QV → CCV	
	$(CL \rightarrow BN)_1$ (%)	$(BN \rightarrow CL)_1$ (%)	$(CL \rightarrow BN)_2$ (%)	$(BN \rightarrow CL)_2$ (%)	$(CL \rightarrow BN)$ (%)	$(BN \rightarrow CL)$ (%)
10.00 (8, 1)	49.83	3.46	3.70	39.46	10.61	0.00
20.59 (8, 1)	59.35	2.48	2.96	32.35	27.48	0.00
48.04 (7, 1)	38.07	1.98	0.61	12.43	27.48	3.20
72.30 (5, 1)	13.93	4.04	0.09	8.31	9.63	7.96
72.30 (5, 3)	13.93	4.04	0.65	3.56	11.31	4.33
75.78 (5, 1)	11.85	4.11	0.04	8.22	7.59	8.04

$(BN \rightarrow CL)_2$, and $(CL \rightarrow BN)_2 - (BN \rightarrow CL)_2$ under the column headed by CQV → CCV and (2) for the combined results of the two correction stages to $(CL \rightarrow BN)$, $(BN \rightarrow CL)$, and $(CL \rightarrow BN) - (BN \rightarrow CL)$ under the column headed by QV → CCV. Note that

$$\begin{aligned}(CL \rightarrow BN) - (BN \rightarrow CL) = &- [(CL \rightarrow BN)_1 - (BN \rightarrow CL)_1] \\ &+ [(CL \rightarrow BN)_2 - (BN \rightarrow CL)_2].\end{aligned} \tag{2.34}$$

Also,

$$\begin{aligned}BNCQP - BNQP &= (CL \rightarrow BN)_1 - (BN \rightarrow CL)_1 \\ BNCCP - BNCQP &= (CL \rightarrow BN)_2 - (BN \rightarrow CL)_2 \\ BNCCP - BNQP &= (CL \rightarrow BN) - (BN \rightarrow CL) \\ &= [BNCQP - BNQP] + [BNCCP - BNCQP]\end{aligned} \tag{2.35}$$

The mapping procedure involves iterations which continue until the difference BNCCP − BNQP is sufficiently small. From Equation 2.35 it is seen that convergence results when

$$(CL \rightarrow BN) \approx (BN \rightarrow CL) \tag{2.36}$$

Consequently, near convergence, the combined effect of the two correction stages should result in the percentage of CL cells reclassified as BN cells being approximately equal to the percentage of BN cells reclassified as CL cells. Alternatively, from Equation 2.35 convergence results when

$$[BNCQP - BNQP] \approx -[BNCCP - BNCQP] \tag{2.37}$$

or equivalently, when

$$[(CL \rightarrow BN)_1 - (BN \rightarrow CL)_1] \approx -[(CL \rightarrow BN)_2 - (BN \rightarrow CL)_2] \tag{2.38}$$

Thus, near convergence, the net percentage of cells which have been reclassified from CL to BN cells during the first correction stage should approximately equal the negative of the net percentage of cells which have been reclassified from CL to BN cells during the second correction stage. These observations are helpful in coming up with rules for determining the next setting of the parameters in the iteration process.

By way of example, when BNQP = 10%, the threshold is such that 10% of total number of cells in surveillance volume fall below the threshold while 90% fall above. The situation is pictured in Figure 2.69(b). With reference to Table 2.13, when NCQ = 8, 49.83% of the total cells in the surveillance volume which were classified as CL cells for being above threshold, are reclassified as BN cells after the first correction stage, whereas 3.46% of the total cells, which were classified as BN cells because they were below the threshold, are reclassified as CL cells. The net percentage of the cells reclassified as BN is 49.83% − 3.46% = 46.37%. For the second correction stage, with NCC = 1,

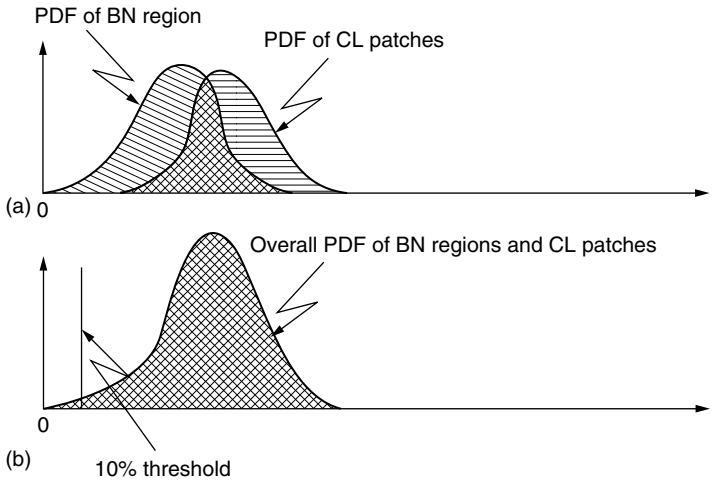

FIGURE 2.69 Overlapping PDFs of BN region and CL patches. (a) Individual PDFs for each region. (b) Overall PDF for both regions.

3.70% of the total cells in the CQV surveillance volume, which were classified as CL cells after the first correction stage, are reclassified as BN cells because these do not have at least one neighboring CL cell. Similarly, 39.46% of the total cells in the CQV surveillance volume, which were classified as BN cells, are reclassified as CL cells because these have one or more neighboring CL cells. The last row of Table 2.13 corresponds to a situation close to convergence. When the threshold is set such that 75.78% of the total number of cells in the surveillance volume are below threshold, the combined effect of the two correction stages results in

$$(\mathrm{CL} \rightarrow \mathrm{BN}) = 7.59\% \approx (\mathrm{BN} \rightarrow \mathrm{CL}) = 8.04\% \tag{2.39}$$

Similarly,

$$[(\mathrm{CL} \rightarrow \mathrm{BN})_1 - (\mathrm{BN} \rightarrow \mathrm{CL})_1] = 8.74\% \approx -[(\mathrm{CL} \rightarrow \mathrm{BN})_2 - (\mathrm{BN} \rightarrow \mathrm{CL})_2] = 8.18\%. \tag{2.40}$$

Equivalently, from Table 2.12 for BNQP = 75.78%,

$$[\mathrm{BNCQP} - \mathrm{BNQP}] = 7.72\% \approx -[\mathrm{BNCCP} - \mathrm{BNCQP}] = 8.19\% \tag{2.41}$$

Insight into the manner by which the PDFs of BN and CL are modified during the correction stages, is obtained by examining pertinent amplitude histograms for various surveillance volumes QV, CQV, and CCV. The overall amplitude histogram for the generated data of the QV volume is shown in Figure 2.70(a). Note that the histogram is unimodal and it is not possible to distinguish between the BN and the CL cells. In fact, by just looking at the histogram one would not suspect that more than one region exists. When the threshold is set at 0.37 such that BNQP = 10%, many of the BN cells are classified as CL due to the low

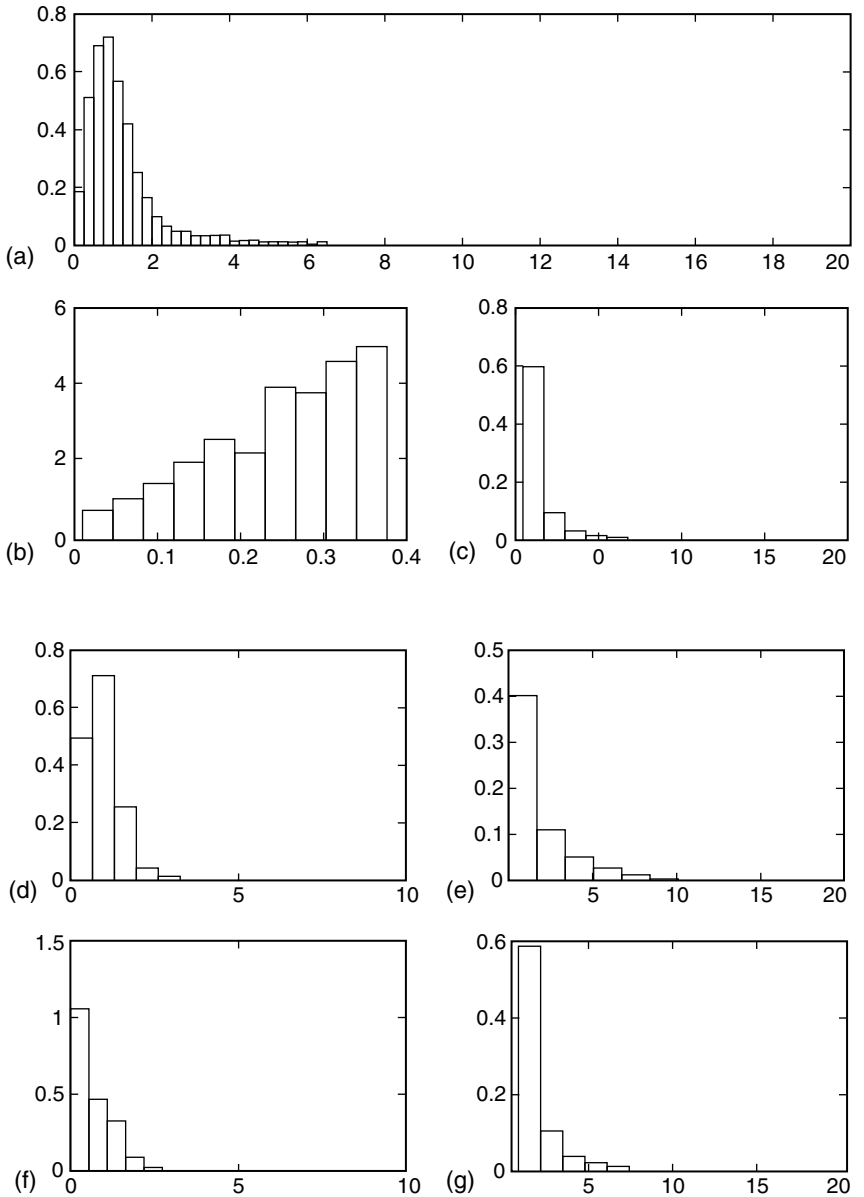

FIGURE 2.70 Region histograms corresponding to BNQP = 10%, NCQ = 8, and NCC = 1. (a) Overall histogram of the generated data. (b) BN histogram at the quantization stage. (c) CL histogram at the quantization stage. (d) BN at the first correction stage. (e) CL histogram at the first correction stage. (f) BN histogram at the second correction stage. (g) CL histogram at the second correction stage.

threshold. The amplitude histograms for the BN and CL cells in the QV volume are shown in Figure 2.70(b,c), respectively. Note that the BN histogram is truncated to a cell amplitude of 0.35.

The amplitude histograms for the CQV volume resulting from the first correction stage are shown in Figure 2.70(d,e). Comparing Figure 2.70(d) with (b), it is seen that many cells with amplitudes above the threshold value of 0.35 have been reclassified as BN. The amplitude histograms for the CCV volume resulting from the second correction stage are shown in Figure 2.56(f,g). Note the further enhancement of the right tail of the BN amplitude histogram. This enhancement is due to the image processing and is in spite of the low threshold value.

During the iteration process the threshold is gradually increased and converges to a value in the vicinity of 1.51 for which BNQP = 75.78%. In Figure 2.71(a), this threshold is shown in the overall histogram for the QV volume. The amplitude histograms for the BN and CL cells in the QV volume are shown in Figure 2.71(b,c), respectively. Note that the amplitudes of the BN cells fall below 1.51 whereas those of the CL cells fall above 1.51. The results of the first and second correction stages and the edge enhancement stage are shown in Figure 2.71(d–i), respectively. To provide a basis for comparison, the actual BN and CL amplitude histograms are shown in Figure 2.71(j,k). The strong similarity between the amplitude histograms of Figure 2.71(h,i) and those of Figure 2.71(j,k) indicates that the mapping procedure has converged satisfactorily. Note how nicely the final histograms of Figure 2.71(h,i) have evolved from the original histograms of Figure 2.70(b,c).

In general, the first correction stage begins to establish the right tail of the BN amplitude histogram and reshapes the CL amplitude histogram by reclassifying mislabeled BN cells. The second correction stage reshapes the bodies and the tails of the BN and CL histograms by recovering the CL edges.

2.5.6. Extension of the Mapping Procedure to Range–Azimuth–Doppler Cells

Assume that the dwell time is $P = M > 1$. In this case, an FFT is possible and the block diagram of the preprocessing and mapping of data is as shown in Figure 2.14. As explained in Section 2.4.2, the mapping consists of two stages. The first mapping stage operates on R/A cells while the second one operates on R/A/D cells.

Once the R/A surveillance volume has been mapped into BN and CL cells, as explained in Section 2.5.3, the second mapping stage starts by declaring as BN cells as R/A/D cells corresponding to R/A cells which were previously declared as BN in the first mapping stage. Meanwhile, the FFT magnitudes of R/A cells which were previously declared as CL in the first mapping stage, are obtained from the preprocessing blocks to enable classification of the remaining R/A/D cells as either BN or CL cells.

Note that a row of range cells, having a fixed azimuth, in the R/A volume corresponds to a R/D plane in the R/A/D volume. Also, a row of azimuth cells, having a fixed range in the R/A volume corresponds to an A/D plane in the R/A/D

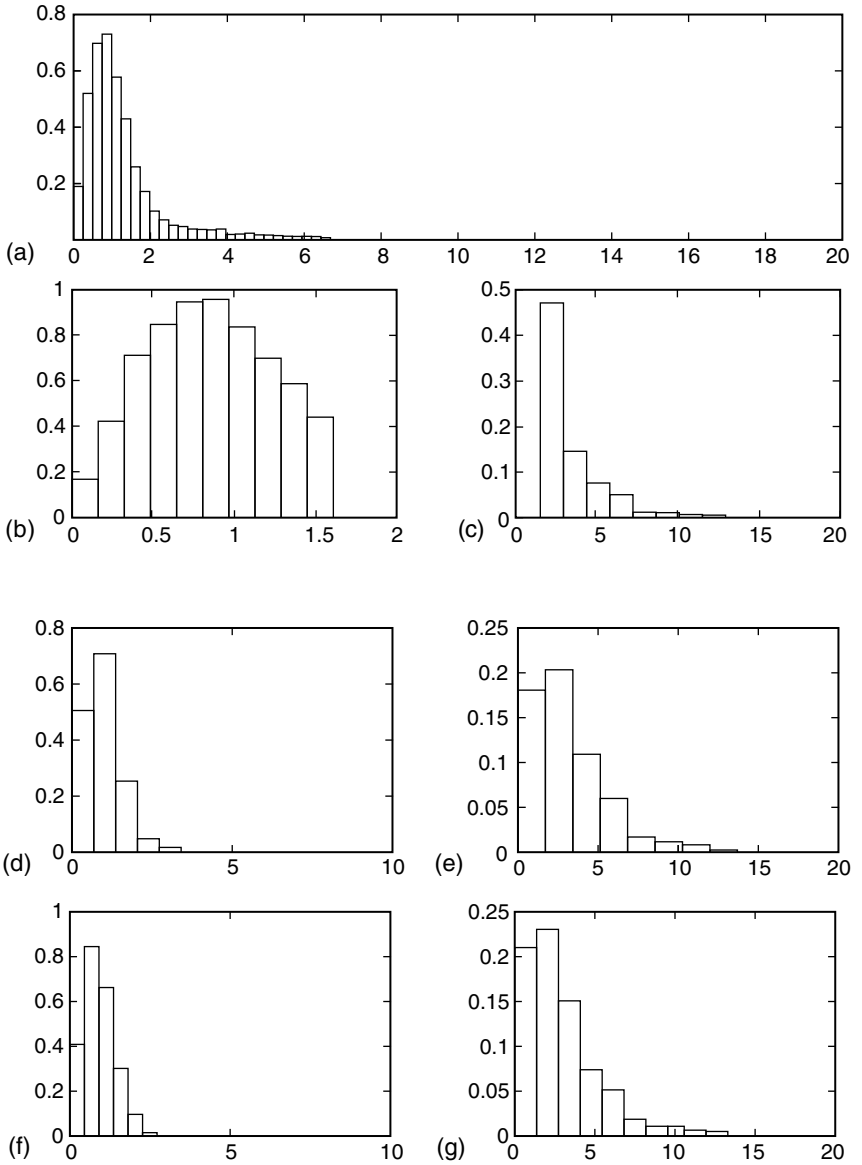

FIGURE 2.71 Region histograms corresponding to BNQP = 75.78%, NCQ = 5, and NCC = 1. (a) Overall histogram of the generated data. (b) BN histogram at the quantization stage. (c) CL histogram at the quantization stage. (d) BN at the first correction stage. (e) CL histogram at the first correction stage. (f) BN histogram at the second correction stage. (g) CL histogram at the second correction stage. (h) BN histogram at the MV. (i) CL histogram at the MV. (j) Actual BN histogram of the generated data. (k) Actual CL histogram at the generated data.

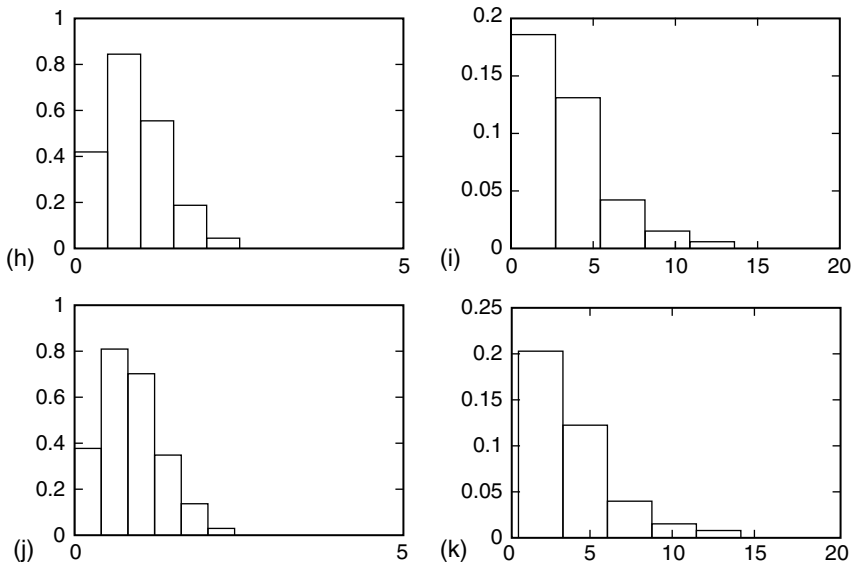

FIGURE 2.71 Continued.

volume. These are illustrated in Figure 2.72(a,b), respectively. Hence, range and azimuth rows of BN declared cells in the R/A volume correspond to R/D and A/D planes, respectively, of BN declared cells in the R/A/D volume and no additional processing is needed for such a plane.

Mapping of the R/A/D volume is done on a plane-by-plane basis. Either a R/A plane (with parameterized azimuth) or an A/D plane (with parameterized range) can be considered. The best choice is one which would involve the least processing. For example, if there are more R/D BN declared planes than A/D BN declared planes in the R/A/D volume, the best choice in this case would be to process R/D planes, and vice versa.

Assume that R/D plane-by-plane processing has been chosen. The mapping procedure becomes a two dimensional problem, and, thresholding/quantization, first correction, second correction, smoothing, edge-enhancement, and edge detection are performed as explained in Section 2.5.3.

When the second mapping stage is done, the R/A/D volume will consist of a R/A/D BN region and R/A/D CL patches in the form of 3-D objects.

2.5.7. Conclusion

In summary, a mapping procedure was presented in Section 2.5 which allowed for distinguishing between BN and CL patches. The procedure was illustrated with examples which showed how the mapping procedure works, even under hard conditions as in the example of Section 2.5.4.2.2, where the histograms of the CL

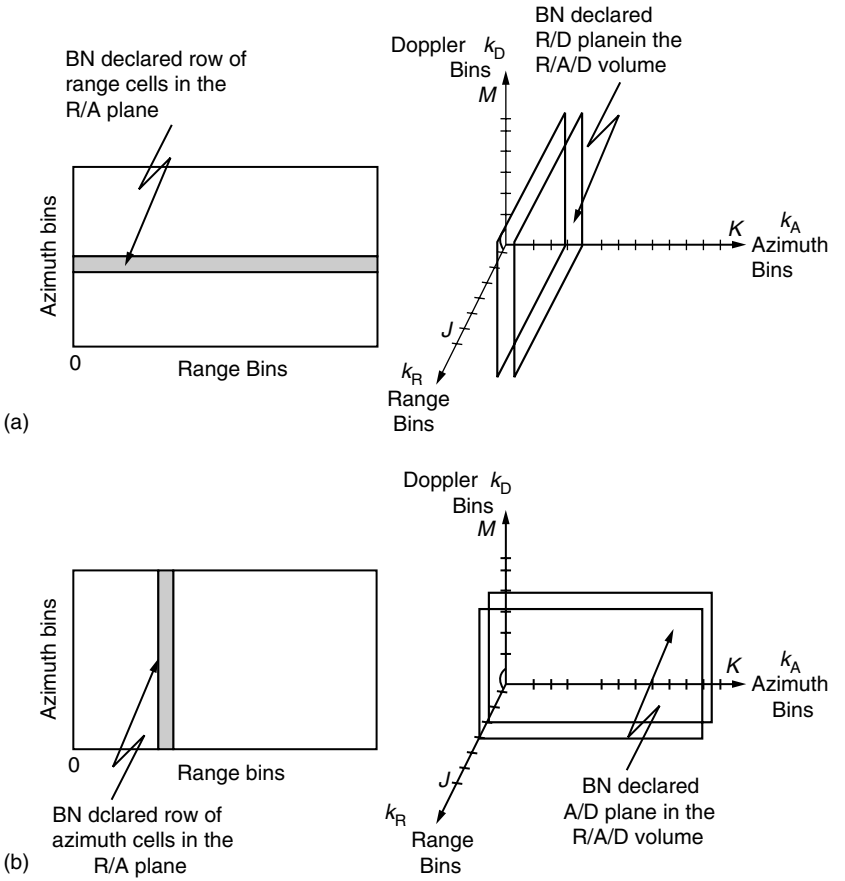

FIGURE 2.72 BN declared rows and their corresponding BN declared planes: (a) BN row of range cells and their BN range–Doppler plane (b) BN row of azimuth cells and BN azimuth–Doppler plane.

patches and BN region overlapped to render the total histogram unimodal such that one would not suspect the presence of more than one distribution.

It is to be noticed that there was no discussion in this chapter about the rules which govern the choice of parameters viz., BNQP, NCQ, and NCC. These rules will be introduced later in Section 2.7 along with the role of IPUS in mapping procedure.

2.6. INDEXING PROCEDURE

2.6.1. INTRODUCTION

As explained in Table 2.1, and shown in examples of Section 2.5.4.2, the mapping procedure subdivides the surveillance volume into BN regions and CL patches labeling the cells in surveillance volume as either BN or CL cells.

In this research it is assumed that BN is white and sufficiently weaker than a desired target return so that the SSC is applicable and detections can be obtained using the conventional Gaussian receiver with little degradation in performance. Consequently, no additional indexing is needed for BN cells.

On the other hand, more processing is needed on CL cells before the detection process can begin. This is because (1) CL patches may be nonhomogeneous containing two or more CL subpatches, (2) in each CL patch, or CL subpatch if it exists, either the SSC, WSC, or ISC will apply, (3) for each WSC CL patch or subpatch, its PDF must be approximated. Also, for each ISC CL patch or subpatch, its covariance matrix must be approximated. All these steps need to be implemented during the indexing stage and are explained in this chapter.

The indexing procedure starts with an assessment stage by which (1) CL patches and the BN region are identified by assigning a unique identification number to all cells within a CL patch or BN region, and (2) CNR and number of cells in each CL patch are determined. The next stage, called the subpatch investigation stage, consists of determining the existence of CL subpatches. If CL subpatches are found, the assessment stage is then carried out, once again, for the entire surveillance volume. The final stage of the indexing procedure is the PDF approximation stage where the PDF for each WSC CL patch is approximated. The assessment, CL subpatch investigation, and PDF approximation stages are explained next.

2.6.2. Assessment Stage

As explained above, assessment is carried out: (1) to assign a unique number to the BN region and each CL patch, (2) to estimate the CNR for each CL patch, and (3) to classify CL patches as either SSC, WSC or ISC regions.

2.6.2.1. Identification of the BN and CL Patches

When the mapping is completed, recall that the mapped volume has a value of zero assigned to BN cells and a value of one assigned to the CL cells. Therefore, nothing more needs to be done for the BN region as all of its cells are already indexed by the number zero. On the other hand, all cells in each of the CL patches are assigned a value of one. Thus, a numbering procedure has to be implemented to enable the computer to distinguish between the various CL patches. The approach taken in this work is to assign every cell in the first patch investigated the number one, every cell in the second patch investigated the number two, and so on until all patches have been indexed with consecutive integers. In this way, all cells in each CL patch are assigned a unique number.

If a cell belongs to a new CL patch, the key to the numbering is the ability to recognize this fact. This is done by considering a mask of five cells as shown in Figure 2.73 where the white cells represent neighboring cells and the shaded one is the test cell to be numbered. Since the surveillance volume has previously been augmented by adding rows of BN to its left, top right, and bottom edges, there is

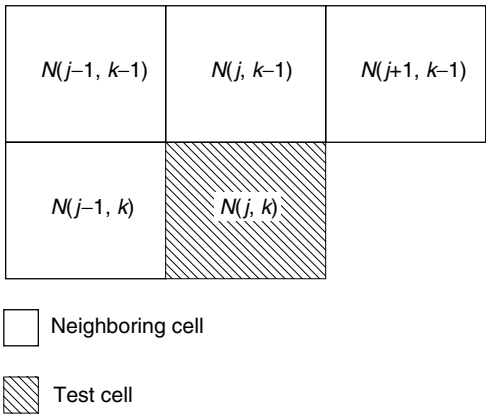

FIGURE 2.73 Mask used in numbering.

no problem positioning this mask for each test cell in the original unaugmented surveillance volume.

Let $M(i, j)$ be the value assigned to the ijth cell in the MV where,

$$M(i,j) = \begin{cases} 0 & \text{if } ij\text{th cell is declared as BN} \\ 1 & \text{if } ij\text{th cell is declared as CL.} \end{cases} \tag{2.42}$$

Assuming that the test cell to be numbered is the jkth cell in the original unaugmented surveillance volume, let the assigned number be denoted by $N(j, k)$. Also, let G denote the unique number assigned to the CL patch previously investigated and H the minimum positive number assigned to neighboring cells. Then, by definition, we have

$$N(j,k) = \begin{cases} 0 & \text{if } M(j,k) = 0 \\ (G+1) & \text{if all neighboring cells are numbered 0} \\ H & \text{if at least one of the neighboring cells is numbered nonzero.} \end{cases} \tag{2.43}$$

The number G is incremented by unity whenever a new CL patch is detected.

Because a CL patch boundary may be sharply shaped, as shown in the example of Figure 2.74, the numbering procedure may end up by assigning two different numbers for different cells of the same CL patch. This anomaly is avoided by further testing the neighboring cells of the cell to be numbered as follows:

1. For the given cell to be numbered, look up the numbers assigned to the set of neighboring cells $(j-1, k-1)$, $(j, k-1)$, $(j+1, k-1)$, and $(j-1, k)$,
2. Take the minimum nonzero number of those in step one,
3. Reassign all nonzero numbered neighboring cells the minimum nonzero number from step two,

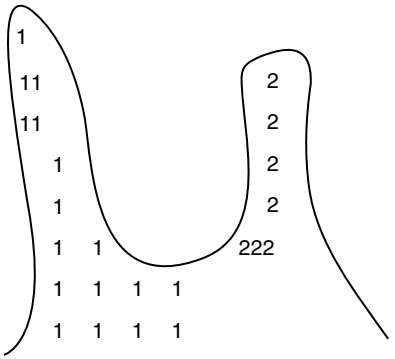

FIGURE 2.74 Example of a sharply-shaped boundary.

4. Revisit all the cells in the surveillance volume that have previously been numbered. If any cell is assigned a nonzero number identical to one of those in step one, reassign that cell the minimum nonzero number of step two.

For example, with respect to Figure 2.74, the above steps have the effect of assigning a value of one to all cells of the CL patch shown.

Once numbering is completed, the BN cells are assigned a value of zero, and every CL patch is assigned a unique positive number.

2.6.2.2. Computation of CL-to-Noise Ratios

The CNR for CL patch k is given by $[(\bar{d}_k - \bar{n})/\bar{n}]^2$, expressed in dB, where $\bar{d}_k$ is the sample average of the CL plus BN values distributed over CL patch k and $\bar{n}$ is the sample average of the BN values distributed over the BN region.

2.6.2.3. Classification of CL Patches

Throughout this work the minimum target SNR is assumed to be 10 dB so that targets are readily detected when they appear in the BN region. The classification of a CL patch as an SSC, WSC, or ISC then depends on its CNR. Ranges of the CNR for the strong, weak, and intermediate signal cases are given in Table 2.14. The bounds for this work were chosen based on experience gained through computer simulations. For example, when CNR = 8 dB, the average signal power is 2 dB larger than the average CL power. Nevertheless, non-Gaussian CL tends to be spiky. Consequently, even when CNR = 8 dB, there are regions where the CL is much larger than the signal. Consequently, 8 dB is chosen as the lower bound on CNR for WSC. Also, when CNR = 5 dB, the average signal power is 5 dB larger than the average CL power. At this level, the CL typically dominates the signal at only a few isolated points. Thus, 5 dB is chosen as the upper bound on CNR for SSC. ISC is defined to fall between the two bounds.

TABLE 2.14
Classification of a CL Patch

Classification	CNR
SSC	CNR $<$ 5 dB
ISC	5 dB $\leq$ CNR $\leq$ 8 dB
WSC	CNR $>$ 8 dB

When a CL patch is classified as SSC, the returns from this patch are processed by the same detector as that used for the BN region. Only the threshold needs to be adjusted properly. When a CL patch is classified as ISC, the GLRT[19] is used to process the associated returns. Finally, when a CL patch is classified as WSC, it is necessary to determine the associated PDF so that the appropriate processor can be selected from the library of weak signal detectors.

2.6.3. CL Subpatch Investigation Stage

Recall that mapping consists of appropriately selecting a threshold to distinguish between BN and CL patches using only the assumption that the BN power, on average, is smaller than the CL power. This same approach may be used, once again to extract that CL subpatch with the lowest average power from a set of contiguous CL subpatches of higher average powers in a given CL patch. In this case, the CL patch containing CL subpatches will be viewed as a volume containing a CL subpatch region with low average power and a set of subpatches with higher average powers occupying the rest of the CL patch area as shown in Figure 2.75.

The mapping procedure, therefore, is used to extract the CL subpatch with the lowest average power from among the remaining CL subpatches in a given CL patch. Because numbering has already labeled each patch with a unique number, it is straightforward for the computer program to select a patch and check for the presence of subpatches in it.

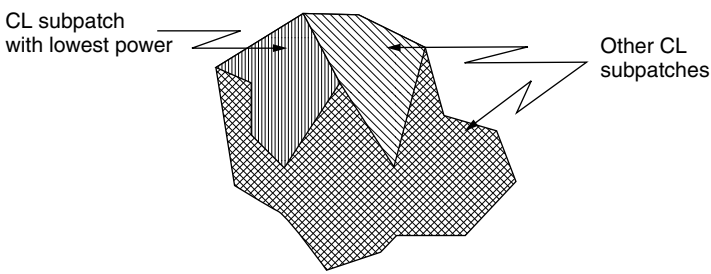

FIGURE 2.75 Example of a CL patch containing CL subpatches.

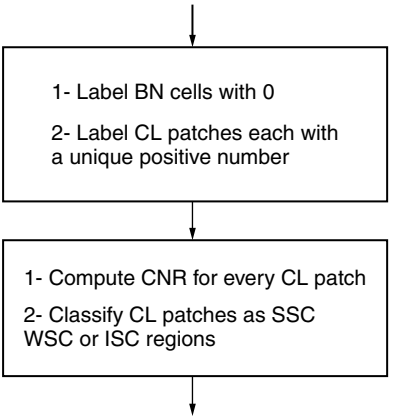

FIGURE 2.76 Assessment stage.

For each CL patch, the mapping procedure is performed iteratively until it is hypothesized, that every subpatch in a given CL patch is homogeneous and cannot be partitioned further. After all CL subpatches have been extracted, the surveillance volume consists of a BN region and CL patches that may or may not be contiguous. Notice that every CL subpatch is now referred to as a CL patch.

If CL subpatches are found to exist, assessment is performed once again to (1) allocate, or reallocate, a number for every CL patch in the surveillance volume, (2) compute the CNR for every patch, and (3) classify the CL patches as SSC, WSC, or ISC regions. Assessment and subpatch investigation are described by the flow charts of Figures 2.76 and 2.77. Figure 2.78 shows the order in which assessment and CL subpatch investigation are performed.

Once WSC regions have been identified and numbered, their respective PDFs are approximated as explained in the next section.

2.6.4. PDF Approximation of WSC CL Patches

The PDF approximation of WSC CL patches follows two steps. During the first step, which is also referred to as test cell selection, a WSC CL patch is selected from among the WSC declared CL patches, a set of test cells is chosen in that CL patch, and reference cells are identified for each test cell which belong to the selected CL patch and are the closest to the test cell. In the second step, referred to as PDF approximation, the data of the reference cells are processed by the Ozturk algorithm so that the PDF of the test cell can be approximated. The process iterates for next WSC CL patch until each WSC CL patch in the surveillance volume has its PDF approximated. Test cell selection, PDF approximation, outlier definition and a PDF approximation strategy are introduced next.

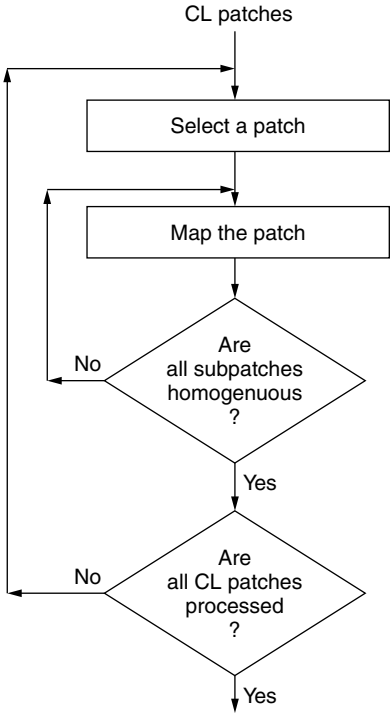

FIGURE 2.77 CL Subpatch investigation stage.

2.6.4.1. Test Cell Selection

As explained above, test cell selection involves three steps:

(i) A WSC CL patch is chosen from among the WSC declared CL patches. This can be done automatically by the computer program since at this stage every CL patch has been labeled with a unique number.
(ii) A set of N_T test cells is then chosen in the WSC CL patch being processed where the value of N_T depends upon the extent to which the patch needs to be characterized. Note that any cell in the CL patch can be a test cell. A possible choice for the test cells is equally spaced test cells that cover the entire area of the CL patch.
(iii) Finally, for each test cell, a set of reference cells is selected. The reference cells must be in the same CL patch as the test cell and should be closest in distance to it to comply with the assumption that the reference cells are representative of the test cells.

In order to select the reference cells for a given test cell, the program starts by centering a three by three mask around the test cell and choosing as reference cells

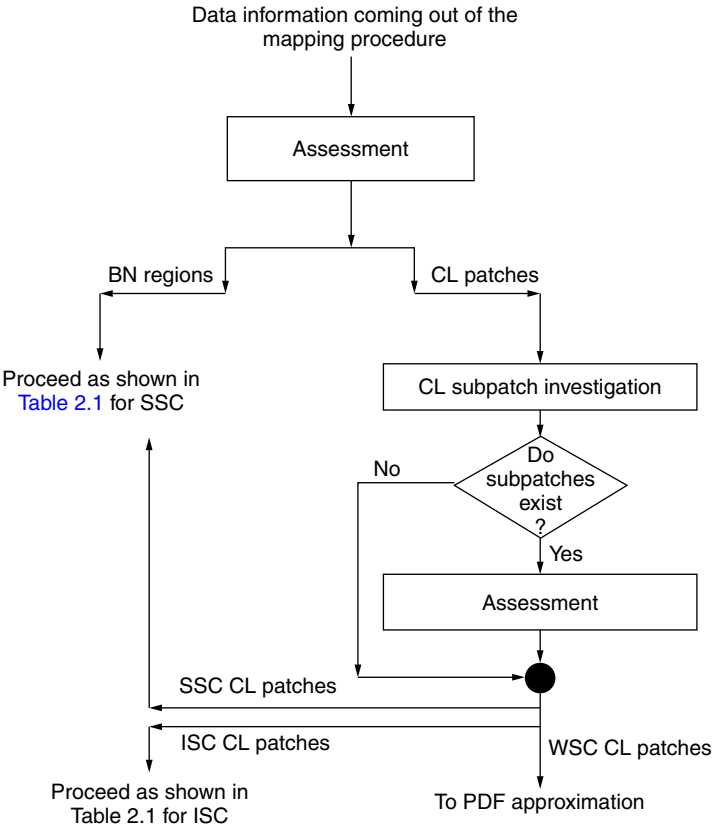

FIGURE 2.78 Assessment and CL subpatch investigation.

those neighboring cells within the mask that are declared to be in the same CL patch as the test cell. If the desired number of reference cells are not obtained, the program increases the size of mask by adding one row and one column to each boundary of the three by three mask. This results in a five by five mask where only the cells in augmented rows and columns need to be examined. The process of adding one row and one column to each boundary of the previous mask continues until desired number of reference cells have been obtained.

2.6.4.2. PDF Approximation

Approximation of the PDF underlying a test cell consists of processing the data in the reference cells. The PDF approximation is performed by the Ozturk algorithm which consists of two modes, which is also referred to as the goodness of fit test mode and the approximation chart mode. These two modes are discussed in the following sections.

2.6.4.2.1. Introduction to Ozturk Algorithm

The Ozturk algorithm[26] is a recent algorithm based on sample order statistics which has been developed[27–29] and reported[30] for univariate distribution approximation. This algorithm has two modes of operation. In the first mode, the algorithm performs a goodness of fit test. The test determines, to a desired confidence level, whether the random data is statistically consistent with a specified probability distribution. In the second mode of operation, the algorithm approximates the PDF underlying the random data. In particular, by analyzing the random data and without any *a priori* knowledge, the algorithm identifies the density function which best approximates the data from a stored library of PDFs. Estimates of the location, scale, and shape parameters of the PDF are provided by the algorithm. The algorithm has been found to work reasonably well for observation sizes as small as 75 to 100. Throughout this work, a number of 100 reference cells are selected for each test cell.

2.6.4.2.2. Goodness of Fit Test

The goodness of fit determines whether or not the set of data samples provided to the algorithm is statistically consistent with a specified distribution, referred to as the null hypothesis. Let N_R denote the number of reference cells. For the null hypothesis, the program utilizes a Monte Carlo simulation of 2000 trials to generate an averaged set of N_R linked vectors in the *uv* plane, as shown in Figure 2.79. Using the standardized sample order statistics of the data, the program then creates a second system of N_R linked vectors in the *uv* plane, as shown in Figure 2.79(a). The terminal points of the linked vectors, as well as

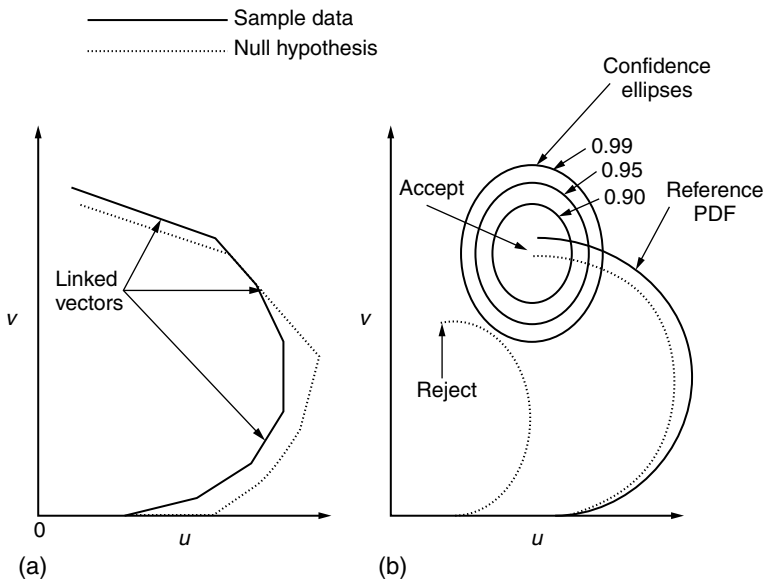

FIGURE 2.79 Goodness of fit test: (a) linked vectors; (b) confidence ellipses.

the shapes of their trajectories, are used in determining whether or not to accept the null hypothesis. The null hypothesis is the distribution against which the sample data is to be tested.

The algorithm provides quantitative information as to how consistent the sample data set is with the null hypothesis distribution by use of confidence contours. In confidence contours, each contour is derived from a specified probability so that the end point falls within the contour, given that the data comes from the null distribution. An example of these contours is shown in Figure 2.79(b) for probabilities of .9, .95, and .99. If the end point of the sample data linked vector locus falls within a contour, then the sample data set is said to be statistically consistent with the null hypothesis at a confidence level based on the probability specified for that contour. If the sample data set is truly consistent with the null hypothesis, the system of sample linked vectors is likely to closely follow that for the system of null linked vectors.

2.6.4.2.3. Approximation Chart Mode

The approximation chart mode is simply an extension of the goodness of fit test mode. Following a similar approach as outlined in the section for the goodness of fit mode, random samples are generated from a library of different univariate probability distributions. In the goodness of fit test mode, the locus of end point was obtained for the null hypothesis and sample size, N_R. For the approximation chart mode we go one step further by obtaining the locus of end point for each distribution from the library of distributions for the given sample size, N_R, and for various choices of the shape parameter(s). Thus, depending on whether it has a shape parameter or not, each distribution is represented by a trajectory or point in the two dimensional *uv* plane. The distributions which are plotted on the distribution approximation chart are: (1) Gaussian, (2) Uniform, (3) Exponential, (4) Laplace, (5) Logistic, (6) Cauchy, (7) Extreme Value, (8) Gumbel type-2, (9) Gamma, (10) Pareto, (11) Weibull, (12) Lognormal, (13) Student-*T*, (14) *K*-distributed, (15) Beta, and (16) Su-Johnson. Figure 2.80 shows an example of the approximation chart. Note that every point in the approximation chart corresponds to a specific distribution. That point closest to the sample data locus end point is chosen as the best approximation to the PDF underlying the random data. This closest point is determined by projecting the sample locus end point to all points on the approximation chart and selecting that point whose perpendicular distance from the sample point is the smallest. Once the PDF underlying the sample data is selected, the shape, location and scale parameters are then approximated.

2.6.4.3. PDF Approximation Metric

In the goodness of fit test, it is hypothesized that whenever the end point of the sample data locus falls within the .99 probability confidence contour of the null distribution, the PDF underlying the data can be approximated by the null

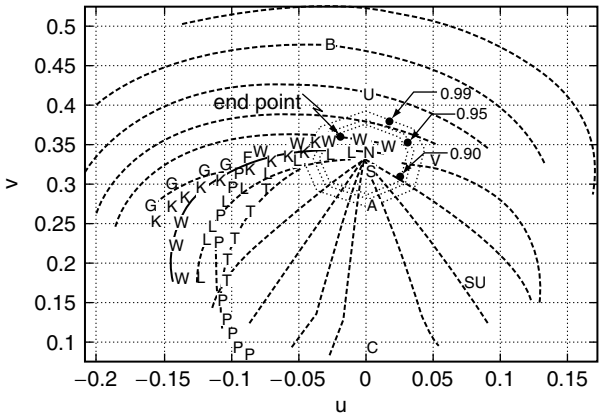

FIGURE 2.80 Approximation chart: N = normal, U = uniform, E = negative exponential, A = Laplace, S = logistic, C = Cauchy, V = extreme value, T = gumbel type-2, G = gamma, P = pareto, W = Weibull, L = lognormal, K = *K*-distributed, B = beta, SU = Su-Johnson. Sample data locus end point ● is shown with 0.99, 0.95, 0.90 confidence contours and candidate PDFs.

distribution. Because the confidence contours are neither circular nor exactly elliptical, analytically determining whether the locus end point falls within the contour is a difficult computational problem. Thus, by definition, the sample data is declared to be from the null distribution provided the locus end point of the sample linked vectors is within the locus end point of the null distribution by a distance that is less than half the length of the minor axis of the .99 confidence contour.

Although the approximation chart can accommodate a wide variety of PDFs, most experimentally measured data for radar CL envelopes are approximated as arising from either Rayleigh, Weibull, Lognormal, or *K*-distributed PDFs, as noted in Section 2.1.2. Hence, for ease of implementation, it is assumed throughout this work that random data in a homogeneous region are generated from and approximated by one of the cited PDFs. Note that because the Weibull PDF becomes the Rayleigh PDF when its shape parameter equals two, only Weibull, Lognormal, or *K* distributions are used. The half-length of the minor axis of the .99 confidence contour is listed in Table 2.15 for the Rayleigh, Weibull, Lognormal, and *K*-distributed PDFs for various values of the shape parameters when $N_R = 100$. For a shape parameter that is not listed in the table, the half-length of the minor axis of the .99 confidence contour is approximated by that of the closest shape parameter for the particular PDF. Because the smallest PDF approximation metric that appears in Table 2.15 is $d_{min} = 0.13 \times 10^{-1}$, the null distribution is declared as the approximating distribution without the need to refer to the table whenever the distance between the end points of the sample and null distribution linked vectors is less than d_{min}.

TABLE 2.15
PDF Approximation Metric for Different PDFs, $N_R = 100$

PDF	Shape Parameter	Half-length of the Minor Axis ($\times 10^{-1}$)
Rayleigh	—	0.47
Weibull	0.1	0.25
	0.5	0.41
	1.0	0.41
	2.0	0.47
	3.0	0.47
	4.0	0.45
	5.0	0.44
	10.0	0.44
Lognormal	0.01	0.48
	0.05	0.45
	0.1	0.43
	0.2	0.42
	0.3	0.42
	0.4	0.46
	1.0	0.41
	5.0	0.28
	10.0	0.13
K-distribution	0.01	0.13
	0.1	0.37
	1.0	0.29
	5.0	0.47
	10.0	0.47
	20.0	0.49
	40.0	0.49
	50.0	0.49

2.6.4.4. Outliers

Even though the mapping procedure does a good job in separating the BN and CL regions one or more outliers may exist in a set of reference cells. For example, outliers may arise due to (1) misidentified BN cells in a CL patch or misidentified CL cells in the BN, (2) cells having data values of low probability of occurrence, and (3) cells containing signals from strong targets.

In correctly approximating the PDF underlying a set of data, outliers can cause a problem by significantly altering the set of linked vectors generated by the Ozturk algorithm. To illustrate this, a set of $N_R = 100$ reference data, referred to as set A, are generated from the Lognormal distribution with shape parameter 0.01. The histogram of this set is plotted in Figure 2.81. Also, another set, referred

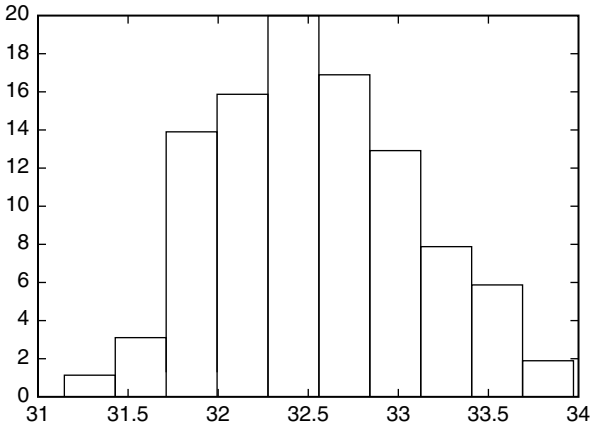

FIGURE 2.81 Histogram of set A.

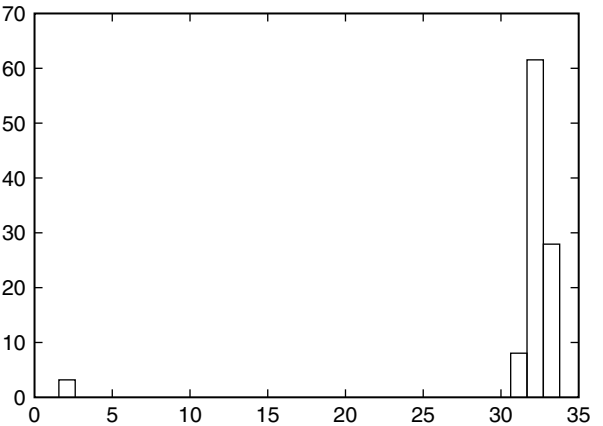

FIGURE 2.82 Histogram of set B.

to as set B, is formed which contains 97 data points from set A and three data points with very small values to constitute the outliers in the set. Figure 2.82 shows the histogram of this second set. Note from the histogram that the three data points have resulted in an isolated bar. The two sets are processed by the Ozturk algorithm and have their locus end points plotted in approximation charts of Figures 2.83 and 2.84. Note how the end point in Figure 2.84 for the set containing outliers (set B) is far removed from the Lognormal PDF from which 97 out of the 100 data points of set B were generated. To understand the cause, Figures 2.85 and 2.86 show the plots of the goodness of fit test for set A and set B, respectively, where the null hypothesis is the Lognormal PDF with shape parameter 0.01. Comparing the two figures, it is noted that the linked vectors in set B are smaller than those of set A, causing the locus end point for set B to fall

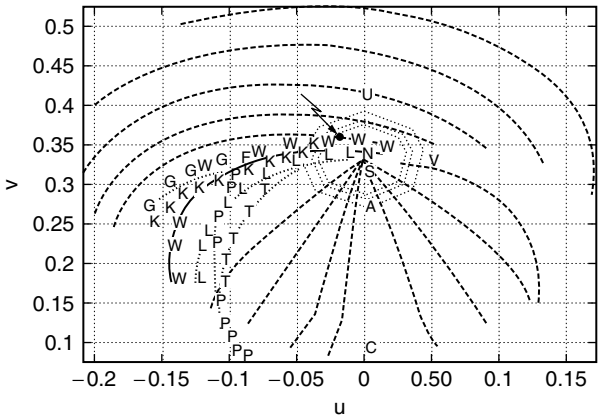

FIGURE 2.83 Approximation chart for the end point of set A.

way below the end point for set A and, therefore, outside the confidence ellipses. This is due to the fact that, the amplitudes of linked vectors are proportional to the magnitude of standardized data, which depend on the mean and standard deviation of the set. The three outliers do not significantly affect the mean of the set but do increase the variance tremendously causing the standardized data and, therefore, the amplitudes of the linked vectors to become smaller. In this example, the mean and standard deviation for set A are equal to 32.54 and 4.69, respectively, while the mean and standard deviation for set B are equal to 31.63 and 18.08, respectively. This example illustrates what can happen when three BN cells with small data values are misidentified and associated with a set of CL cells.

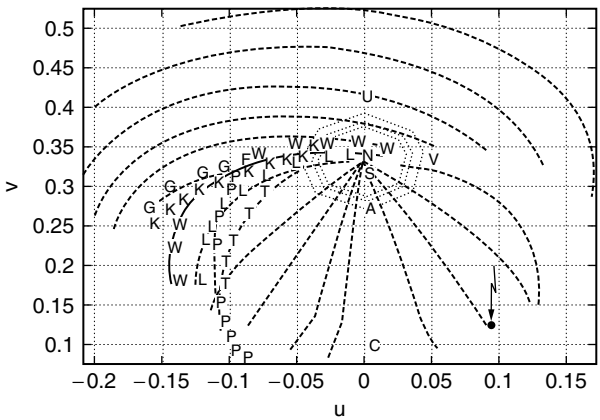

FIGURE 2.84 Approximation chart for the end point of set B.

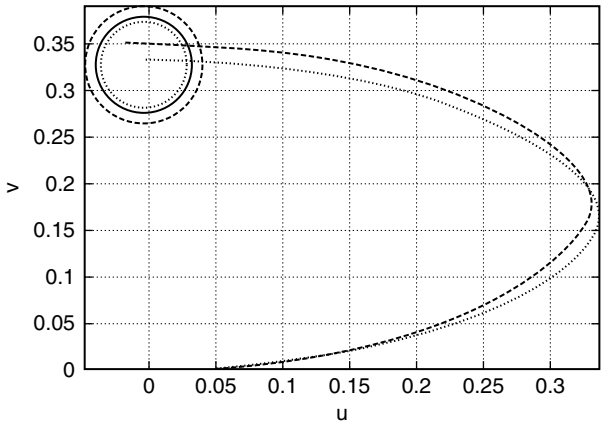

FIGURE 2.85 Goodness of fit test for set A.

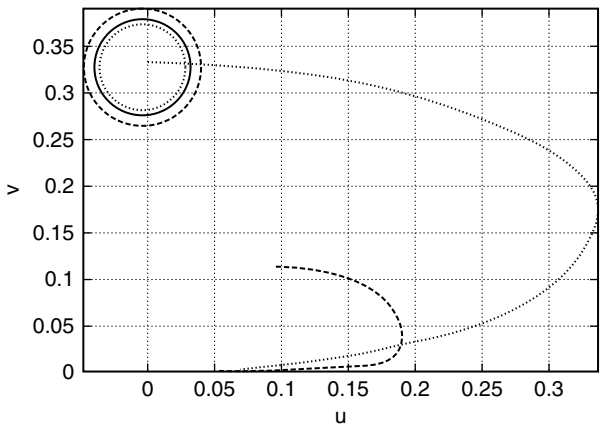

FIGURE 2.86 Goodness of fit test for set B.

One way to identify outliers within a region is to compute the mean m and standard deviation σ within that region and call any cell whose data value falls outside the interval $[m - 2\sigma, m + 2\sigma]$ as an outlier. This method is used to identify outliers within a set of selected reference cells. When this method is applied to set B, whose histogram is shown in Figure 2.82, the three outliers are identified and removed from the set.

2.6.4.5. PDF Approximation Strategy

As mentioned in Section 2.6.4.3, CL data are generated in this work from either the Weibull, Lognormal, or K-distributed distributions. Consequently,

the strategy for approximating the PDF underlying a CL patch consists of the following steps:

1. Select a total of N_T test cells that are evenly spread throughout the CL patch.
2. For each test cell, choose its closest $N_R = 100$ reference cells as described in Section 2.6.4.1.
3. Using the Ozturk algorithm, determine the distance between the locus end point of the data linked vectors and its projection onto the Weibull, Lognormal, and *K*-distributed trajectories.
4. Discard those PDFs for which the distances in step 3 exceed the corresponding distances obtained from Table 2.15.
5. If all possible PDFs are discarded in step 4, excise any outliers that may exist from the data and proceed to step 3.
6. Use IPUS, as discussed in the next chapter, to determine the best PDF(s) to approximate the data.

Steps 1–5 are now illustrated through examples.

2.6.5. Examples

In this section, indexing is performed on the same examples considered in Section 2.5.4 where mapping procedure was carried out. Discussion of each example ends with a table indicating the approximations to be used for the PDFs underlying selected test cells in the CL patches. The quality of the approximations is discussed in Section 2.7 for each of the examples.

2.6.5.1. Example 1

Consider the example of Section 2.5.4.2.1 where the parameters of the generated scene are given in Table 2.2.

2.6.5.1.1. Assessment

The indexing procedure starts by first assessing the mapped volume. This consists of labeling the CL patches, estimating CNR for each patch, and counting the number of cells in each patch. The results are presented in Table 2.16. CL patches labeled one, two and three, are shown in Figure 2.87. By comparing Table 2.2 with Table 2.16, it is seen that reasonable results are obtained for the CNRs and numbers of cells. Because the CL patch labeled as two consists of C + D, its CNR and number of cells, respectively, are approximated by the average of the CNRs and the sum of the number of cells for C and D.

2.6.5.1.2. CL Subpatch Investigation

The next step in the indexing procedure consists of identifying subCL patches within a CL patch. This is done by selecting a CL patch and applying the mapping

TABLE 2.16
Assessment Parameters (Example 1)

CL Patch	CL Patch Label	CNR (dB)	Number of Cells
A	1	9.04	113
C + D	2	36.66	304
B	3	23.31	1444

procedure to it. If one or more subpatches exist, the procedure attempts to first identify the subpatch having the lowest average power. If a subpatch does not exist, the procedure attempts to recognize this situation.

Let PLQP, PLCQP, and PLCCP represent the percentage number of cells of the subpatch with the lowest power among the possible subpatches of a CL patch in the quantized, first-corrected and second-corrected stages, respectively. Tables 2.17–2.19 show the results of the mapping procedure applied, respectively, to CL patches one, two, and three. For CL patches one and three, the mapping procedure results with PLCCP being equal to 100%, reflecting the conclusion that the subpatches with the lowest average power in CL patches one and three occupy 100% of the area in each patch. Therefore, it is concluded that CL patches one and three are homogeneous and do not contain any subpatches. On the other hand, mapping of CL patch two results with PLCCP = 48.75%. This indicates that CL patch two contains at least two subpatches and the subpatch with the lowest power occupies 48.75% of the total area of patch two. This is consistent with the generated surveillance volume, where CL patch C has a smaller power than CL patch, D and C occupies 50% of the total area of CL patch two.

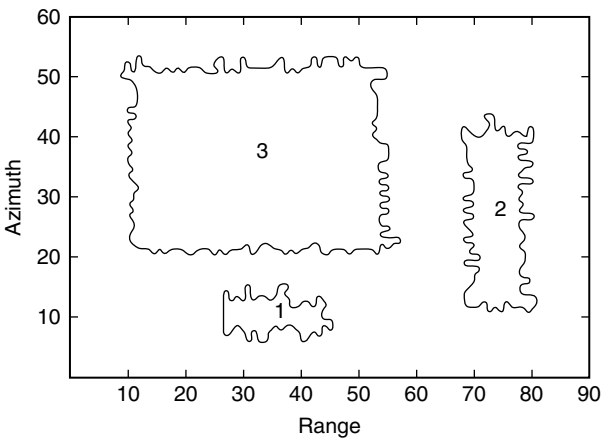

FIGURE 2.87 Contour plot of the mapped volume after numbering (Example 1).

TABLE 2.17
Mapping Procedure Applied to CL Patch 1 (Example 1)

PLQP (%)	Parameter Values	
10.00	NCQ = 8	PLCQP = 70.59
(guess)	NCC = 1	PLCCP = 20.59
20.59	NCQ = 8	PLCQP = 94.12
previous PLCCP	NCC = 1	PLCCP = 82.35
82.35	NCQ = 8	PLCQP = 100.0
previous PLCCP	NCC = 1	PLCCP = 100.0

TABLE 2.18
Mapping Procedure Applied to CL Patch 2 (Example 1)

PLQP (%)	Parameter Values	
10.00	NCQ = 8	PLCQP = 39.38
(guess)	NCC = 1	PLCCP = 7.50
20.00	NCQ = 8	PLCQP = 59.38
previous PLCCP	NCC = 1	PLCCP = 43.12
43.12	NCQ = 7	PLCQP = 55.00
previous PLCCP	NCC = 1	PLCCP = 49.38
49.38	NCQ = 5	PLCQP = 50.62
previous PLCCP	NCC = 1	PLCCP = 46.25
49.38	NCQ = 5	PLCQP = 50.62
previous PLCCP	NCC = 3	PLCCP = 48.75

TABLE 2.19
Mapping Procedure Applied to CL Patch 3 (Example 1)

PLQP (%)	Parameter Values	
10.00	NCQ = 8	PLCQP = 50.25
(guess)	NCC = 1	PLCCP = 12.19
12.19	NCQ = 8	PLCQP = 58.51
previous PLCCP	NCC = 1	PLCCP = 16.61
16.61	NCQ = 8	PLCQP = 73.62
previous PLCCP	NCC = 1	PLCCP = 28.80
28.80	NCQ = 8	PLCQP = 92.57
(guess)	NCC = 1	PLCCP = 66.19
66.19	NCQ = 8	PLCQP = 100.0
previous PLCCP	NCC = 1	PLCCP = 100.0

2.6.5.1.3. Reassessment

Finally, mapping was applied to CL patches two and four with the result that both patches were judged to be homogeneous.

Treating each identified subpatch as a patch, assessment is done once more to relabel all subpatches in the surveillance volume. Results of the assessment are tabulated in Table 2.20. Now the surveillance volume is identified to include four CL patches where CL patches one, two, three, and four coincide with CL patches A, C, B, and D, respectively, of the original scene. Note how close are the number of cells and CNR for every patch when comparing Tables 2.2 and 2.20. Note also that using the values of Table 2.14, every CL patch in the surveillance volume has been identified as a WSC region. Also, using the ranges of the CNR for the strong, weak, and intermediate signal cases, as given in Table 2.14, classification of the CL patches is shown in column five of Table 2.20. Figure 2.88 shows the result of the numbering of the new mapped volume.

TABLE 2.20
Reassessment Parameters (Example 1)

CL Patch	CL Patch Label	CNR (dB)	Number of Cells	Classification
A	1	9.04	113	WSC
B	2	30.63	151	WSC
C	3	23.31	1444	WSC
D	4	40.54	146	WSC

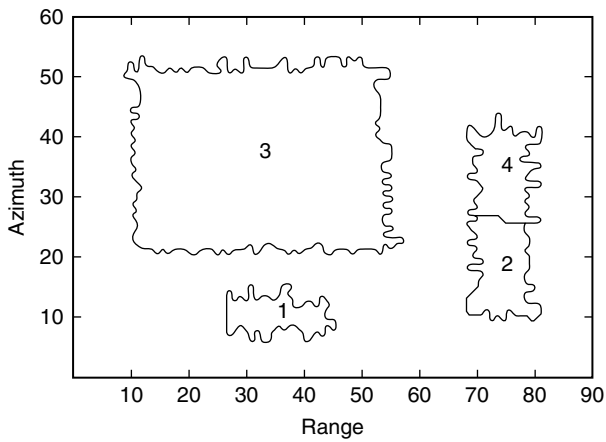

FIGURE 2.88 Contour plot of the MV after renumbering (Example 1).

2.6.5.1.4. Steps 1–5 of the PDF Approximation Strategy

As detailed in Section 2.6.4.5, the first step in the PDF approximation strategy is to select a total of N_T test cells that are evenly spread throughout the CL patch. For each of the four CL patches listed in Table 2.20, let $N_T = 5$. Assume each CL patch is scanned from left to right and from top to bottom. In each CL patch let the first cell scanned be numbered as one. The last cell scanned in each patch is assigned a number equal to the total number of cells in that patch. The test cells selected for CL patches one, two, three, and four are listed in the first column of Tables 2.21–2.24, respectively. Note that the test cells are evenly spread throughout each CL patch.

The second step in the PDF approximation strategy is to choose for each test cell the closest $N_R = 100$ reference cells as detailed in Section 2.6.4.1.

Using the Ozturk algorithm, the third step is to determine the distance between the locus end point of the data linked vectors and its projection onto the Weibull, Lognormal, and *K*-distributed trajectories. The distance, shape, scale, and location parameters for the approximating PDFs are tabulated in columns four, five, six, and seven, respectively, of Tables 2.21–2.24. The (u, v) coordinates of the data locus end point are listed in column eight of the tables. Finally, column three provides a ranking based on the smallest distance between the approximating PDF and the data locus end point.

TABLE 2.21
PDF Approximation for CL Patch 1 (Example 1)

Test Cell No.	PDF	Rank	Distance	Shape	Scale	Location	(u, v)
	W	1	0.72×10^{-2} T	2.12	3.20	0.58	(−0.0298, 0.3503)
1	L	5	0.22×10^{-1} T	0.18	7.48	−4.18	
	K	3	0.17×10^{-1} T	50.0	0.43	0.77	
	W	2	0.90×10^{-2} T	2.14	3.26	0.54	(−0.0293, 0.3320)
28	L	5	0.24×10^{-1} T	0.18	7.76	−4.47	
	K	3	0.18×10^{-1} T	50.0	0.43	0.75	
	W	1	0.74×10^{-2} T	1.94	3.01	0.74	(−0.0376, 0.3500)
57	L	5	0.24×10^{-1} T	0.21	6.74	−3.50	
	K	3	0.13×10^{-1} T	50.0	0.43	0.69	
	W	2	0.13×10^{-1} T	2.06	3.16	0.63	(−0.0313, 0.3561)
85	L	6	0.28×10^{-1} T	0.19	7.38	−4.10	
	K	3	0.21×10^{-1} T	50.0	0.43	0.73	
	W	1	0.23×10^{-2} T	2.01	3.09	0.69	(−0.0330, 0.3455)
113	L	5	0.18×10^{-1} T	0.20	7.02	−3.75	
	K	2	0.11×10^{-1} T	50.0	0.43	0.73	

TABLE 2.22
PDF Approximation for CL Patch 2 (Example 1)

Test Cell No.	PDF	Rank	Distance	Shape	Scale	Location	(*u*, *v*)
1	W	17	0.24×10^{-0} Ψ	5.00	9.00	24.12	(−0.0952, 0.1143)
	L	19	0.25×10^{-0} Ψ	0.76	1.97	30.20	
	K	20	0.26×10^{-0} Ψ	32.0	0.72	28.99	
1	W	2	0.87×10^{-2} T	2.60	1.39	31.30	(−0.0177, 0.3512)
w/o 3	L	4	0.21×10^{-1} T	0.10	5.02	27.50	
Outliers	K	6	0.27×10^{-1} T	50.0	0.16	31.58	
38	W	23	0.29×10^{-0} Ψ	0.73	1.81	30.84	(0.1281, 0.1361)
	L	21	0.27×10^{-0} Ψ	0.81	2.27	29.80	
	K	20	0.26×10^{-0} Ψ	50.0	0.72	28.09	
38	W	1	0.63×10^{-2} T	2.82	1.46	31.30	(−0.0122, 0.3486)
w/o 4	L	4	0.18×10^{-1} T	0.06	7.81	24.77	
Outliers	K	14	0.31×10^{-1} T	50.0	0.15	31.67	
76	W	24	0.30×10^{-0} Ψ	0.74	1.81	30.79	(0.1310, 0.1360)
	L	21	0.27×10^{-0} Ψ	1.10	1.29	30.67	
	K	20	0.26×10^{-0} Ψ	50.0	0.71	28.09	
76	W	1	0.20×10^{-2} T	2.52	1.25	31.43	(−0.0202, 0.3406)
w/o 4	L	3	0.11×10^{-1} T	0.13	3.62	28.89	
Outliers	K	5	0.21×10^{-1} T	50.0	0.14	31.66	
114	W	26	0.31×10^{-0} Ψ	0.52	1.19	31.02	(0.1581, 0.1563)
	L	23	0.29×10^{-0} Ψ	1.84	0.37	31.41	
	K	21	0.27×10^{-0} Ψ	50.0	0.84	27.16	
114	W	12	0.28×10^{-1} T	2.45	1.19	31.41	(−0.0235, 0.3143)
w/o 5	L	3	0.14×10^{-1} T	0.18	2.57	29.86	
Outliers	K	11	0.28×10^{-1} T	50.0	0.14	31.60	
151	W	26	0.31×10^{-0} Ψ	0.52	1.19	31.02	(0.1581, 0.1563)
	L	23	0.29×10^{-0} Ψ	1.85	0.38	31.41	
	K	21	0.27×10^{-0} Ψ	50.0	0.84	27.16	
151	W	10	0.28×10^{-1} T	2.44	1.19	31.48	(−0.0235, 0.3147)
w/o 5	L	3	0.14×10^{-1} T	0.18	2.58	29.85	
Outlier	K	11	0.28×10^{-1} T	50.0	0.14	31.60	

Step 4 is to discard those PDFs for which the distances in step 3 exceed the corresponding distances obtained from Table 2.15. In Tables 2.21–2.24, the results are indicated by either Ψ or T, respectively, depending upon whether the PDFs are discarded or not.

TABLE 2.23
PDF Approximation for CL Patch 3 (Example 1)

Test Cell No.	PDF	Rank	Distance	Shape	Scale	Location	(*u*, *v*)
1	W	5	0.17×10^{-1} T	1.62	13.1	0.82	(−0.0476, 0.3235)
	L	1	0.25×10^{-2} T	0.34	20.4	−9.04	
	K	3	0.75×10^{-2} T	15.6	3.90	−1.19	
361	W	6	0.24×10^{-1} T	1.67	13.7	1.65	(−0.0438, 0.3171)
	L	1	0.49×10^{-2} T	0.33	21.4	−8.75	
	K	4	0.15×10^{-1} T	15.8	3.94	−0.01	
722	W	3	0.19×10^{-1} T	2.67	22.1	−4.93	(−0.0154, 0.3620)
	L	5	0.32×10^{-1} T	0.07	1.15	−101.14	
	K	7	0.35×10^{-1} T	50.0	2.43	−0.14	
1083	W	1	0.33×10^{-2} T	1.63	11.8	3.41	(−0.0508, 0.3433)
	L	5	0.22×10^{-1} T	0.31	19.8	−6.95	
	K	2	0.11×10^{-1} T	35.0	2.34	1.51	
1444	W	1	0.36×10^{-2} T	2.2	16.9	−1.17	(−0.0285, 0.3393)
	L	3	0.11×10^{-1} T	0.19	37.7	−24.57	
	K	4	0.12×10^{-1} T	50.0	2.18	0.26	

TABLE 2.24
PDF Approximation for CL Patch 4 (Example 1)

Test Cell No.	PDF	Rank	Distance	Shape	Scale	Location	(*u*, *v*)
1	W	5	0.44×10^{-1} T	4.30	86.9	16.66	(0.0165, 0.3811)
	L	6	0.52×10^{-1} Ψ	0.01	2625	−2529.01	
	K	20	0.72×10^{-1} Ψ	50.0	6.34	57.92	
37	W	5	0.36×10^{-1} T	3.88	79.5	23.66	(0.0102, 0.3751)
	L	6	0.44×10^{-1} T	0.01	2628	−2532.56	
	K	18	0.63×10^{-1} Ψ	50.0	6.35	57.54	
73	W	2	0.78×10^{-2} T	5.00	96.2	17.69	(0.0119, 0.3630)
	L	5	0.23×10^{-1} T	0.01	2547	−2508.77	
	K	19	0.60×10^{-1} Ψ	50.0	6.16	58.36	
110	W	1	0.52×10^{-2} T	5.00	99.3	5.24	(0.0153, 0.3317)
	L	5	0.16×10^{-1} T	0.01	2616	−2520.22	
	K	18	0.56×10^{-1} Ψ	50.0	6.32	58.84	
146	W	1	0.47×10^{-2} T	4.37	92.2	11.62	(0.0080, 0.3430)
	L	3	0.14×10^{-1} T	0.01	2736	−2640.34	
	K	18	0.48×10^{-1} T	50.0	6.61	56.10	

TABLE 2.25
Assessment Parameters (Example 2)

CL Patch	CL Patch Label	CNR (dB)	Number of Cells
A	1	12.03	133
C + D	2	8.22	308
B	3	8.65	622

In the shaded background rows of Table 2.22, related to CL patch 2, note that all possible PDFs are discarded. Consequently, step 5 is implemented where outliers that exist are excised from each set of reference cells. Steps 3 and 4 are then repeated and the results presented with a white background in Table 2.22. With the outliers removed, note that the distance measure for one or more of the approximating PDFs has become significantly smaller.

2.6.5.2. Example 2

Consider now the example of Section 2.5.4.2.2 where the parameters of the generated scene are given in Table 2.5.

2.6.5.2.1. Assessment

Following the same strategy discussed in Section 2.6.5.1, results of the assessment stage are presented in Table 2.25. CL patches 1, 2 and 3, are shown in Figure 2.88. Comparison of Tables 2.5 and 2.25 shows that reasonable results are obtained for the CNRs and number of cells.

2.6.5.2.2. CL Subpatches Investigation

Tables 2.26–2.28, show the results of the mapping procedure applied, respectively to CL patches one, two, and three. For CL patches one, two, and three,

TABLE 2.26
Mapping Procedure Applied to CL Patch 1 (Example 2)

PLQP (%)	Parameter Values	
10.00	NCQ = 8	PLCQP = 36.73
(guess)	NCC = 1	PLCCP = 8.16
20.00	NCQ = 8	PLCQP = 71.43
modified PLQP	NCC = 1	PLCCP = 28.57
28.57	NCQ = 8	PLCQP = 87.76
previous PLCCP	NCC = 1	PLCCP = 55.10
55.10	NCQ = 8	PLCQP = 100.0
previous PLCCP	NCC = 1	PLCCP = 100.0

TABLE 2.27
Mapping Procedure Applied to CL Patch 2 (Example 2)

PLQP (%)	Parameter Values	
10.00	NCQ = 8	PLCQP = 60.38
(guess)	NCC = 1	PLCCP = 19.50
19.50	NCQ = 8	PLCQP = 81.76
previous PLCCP	NCC = 1	PLCCP = 49.06
49.06	NCQ = 8	PLCQP = 94.97
previous PLCCP	NCC = 1	PLCCP = 81.76
81.76	NCQ = 8	PLCQP = 100.0
previous PLCCP	NCC = 1	PLCCP = 100.0

the mapping procedure results with PLCCP being equal to 100% reflecting the conclusion that the subpatches with the lowest power in CL patches one, two, and three occupy 100% of the area in each patch. Therefore, it is concluded that all of these CL patches are homogeneous and do not contain any subpatches.

In fact, CL patch two consists of subpatches C and D. By using expert system rules, it is shown in Section 2.7 that CL patch two can be further separated.

2.6.5.2.3. Reassessment

Because the CL subpatch identification does not result in finding any subpatches within CL patches one, two and three, the mapped volume is not changed and assessment is not necessary.

2.6.5.2.4. Steps 1–5 of the PDF Approximation Strategy

Tables 2.29 to 2.31 list the results of steps 1–5 of the PDF approximation strategy applied to CL patches one, two, and three, respectively.

TABLE 2.28
Mapping Procedure Applied to CL Patch three (Example 2)

PLQP (%)	Parameter Values	
10.00	NCQ = 8	PLCQP = 63.36
(guess)	NCC = 1	PLCCP = 22.12
22.12	NCQ = 8	PLCQP = 89.63
previous PLCCP	NCC = 1	PLCCP = 62.90
62.90	NCQ = 8	PLCQP = 100.0
previous PLCCP	NCC = 1	PLCCP = 100.0

TABLE 2.29
PDF Approximation for CL Patch 1 (Example 2)

Test Cell No.	PDF	Rank	Distance	Shape	Scale	Location	(*u*, *v*)
1	W	5	0.16×10^{-1}T	1.40	4.54	0.20	(−0.0566, 0.3207)
	L	1	0.31×10^{-2} Ψ	0.40	6.67	−2.93	
	K	2	0.57×10^{-2} Ψ	9.93	1.90	−1.00	
33	W	5	0.13×10^{-1}T	1.37	4.57	0.23	(−0.0604, 0.3232)
	L	3	0.73×10^{-2} T	0.41	6.63	−2.80	
	K	1	0.17×10^{-2} T	8.61	2.09	−1.10	
66	W	6	0.44×10^{-1} Ψ	0.86	2.93	1.48	(−0.0802, 0.2751)
	L	2	0.16×10^{-1} Ψ	0.77	3.14	0.37	
	K	5	0.35×10^{-1} Ψ	1.71	4.36	−0.36	
66	W	4	0.13×10^{-1} T	1.30	3.82	0.68	(−0.0651, 0.32116)
w/o 3	L	3	0.82×10^{-2} T	0.45	5.26	−1.64	
outliers	K	1	0.16×10^{-2} T	6.96	2.03	−0.58	
99	W	5	0.23×10^{-1} T	1.27	4.11	0.62	(−0.0636, 0.3113)
	L	1	0.65×10^{-3} T	0.48	5.35	−1.59	
	K	3	0.11×10^{-1} T	6.23	2.35	−0.79	
133	W	7	0.56×10^{-1} Ψ	0.73	2.39	2.05	(−0.0732, 0.2641)
	L	3	0.29×10^{-1} Ψ	0.83	2.82	−0.85	
	K	6	0.48×10^{-1} Ψ	1.73	4.93	−0.23	
133	W	5	0.18×10^{-1}T	1.36	4.08	0.55	(−0.0586, 0.3181)
w/o 4	L	1	0.21×10^{-2} T	0.42	5.77	−2.06	
outliers	K	2	0.71×10^{-2} T	8.80	1.93	−0.62	

2.6.5.3. Example 3

Finally, consider the example of Section 2.5.4.2.3 where the parameters of the generated scene are given in Table 2.8.

2.6.5.3.1. Assessment

Following the same strategy discussed in Section 2.6.5.1, results of the assessment stage are presented in Table 2.32. CL patches one, two and three, are shown in Figure 2.89. Comparison of Tables 2.8 and 2.32 shows that reasonable results are obtained for the CNRs and number of cells.

2.6.5.3.2. CL Subpatch Investigation

Tables 2.33–2.35 show the results of the mapping procedure applied, respectively, to CL patches one, two, and three. For CL patches one, two, and three, the mapping procedure results with PLCCP being equal to 100% reflecting the conclusion that the subpatches with the lowest power in CL patches one, two, and three occupy 100% of the area in each patch. Therefore, it is concluded that all CL patches are homogeneous and do not contain any subpatches.

TABLE 2.30
PDF Approximation for CL Patch 2 (Example 2)

Test Cell No.	PDF	Rank	Distance	Shape	Scale	Location	(*u*, *v*)
1	W	8	0.57×10^{-1} Ψ	0.73	1.08	1.47	(−0.0725, 0.2622)
	L	3	0.30×10^{-1} T	0.84	1.27	0.93	
	K	6	0.50×10^{-1} Ψ	1.71	2.02	0.41	
77	W	7	0.85×10^{-1} Ψ	0.72	1.09	1.53	(−0.0813, 0.2659)
	L	2	0.21×10^{-1} T	0.85	1.27	0.99	
	K	5	0.41×10^{-1} Ψ	1.47	2.19	0.51	
154	W	2	0.18×10^{-1} T	1.34	2.56	0.91	(−0.0740, 0.3510)
	L	7	0.37×10^{-1} T	0.40	3.98	−1.07	
	K	4	0.29×10^{-1} T	1.92	1.16	0.08	
231	W	3	0.93×10^{-2} T	2.08	4.06	0.24	(−0.0316, 0.3338)
	L	2	0.67×10^{-2} T	0.21	8.63	−4.97	
	K	4	0.93×10^{-2} T	50.0	0.55	0.42	
308	W	1	0.01×10^{-3} T	1.99	3.99	0.25	(−0.0335, 0.3435)
	L	4	0.16×10^{-1} T	0.20	9.10	−5.52	
	K	2	0.94×10^{-2} T	50.0	0.56	0.29	

TABLE 2.31
PDF Approximation for CL Patch 3 (Example 2)

Test Cell No.	PDF	Rank	Distance	Shape	Scale	Location	(*u*, *v*)
1	W	3	0.69×10^{-2} T	0.90	2.06	0.87	(−0.0997, 0.3072)
	L	6	0.19×10^{-1} T	0.69	2.45	−0.12	
	K	1	0.26×10^{-2} T	1.69	2.94	−0.30	
156	W	1	0.13×10^{-2} T	1.13	2.66	0.52	(−0.0830, 0.2397)
	L	5	0.25×10^{-1} T	0.52	3.51	−1.02	
	K	4	0.13×10^{-1} T	4.00	2.08	−0.65	
311	W	4	0.13×10^{-1} T	1.17	2.50	0.88	(−0.0743, 0.3181)
	L	3	0.11×10^{-1} T	0.52	3.21	−0.45	
	K	1	0.48×10^{-3} T	4.47	1.78	−0.12	
466	W	4	0.13×10^{-1} T	1.17	3.12	0.81	(−0.0743, 0.3174)
	L	3	0.11×10^{-1} T	0.52	3.99	−0.84	
	K	1	0.11×10^{-2} T	4.41	2.25	−0.45	
622	W	5	0.24×10^{-1} T	0.76	2.13	1.43	(−0.0965, 0.2867)
	L	1	0.33×10^{-2} T	0.79	2.57	0.31	
	K	3	0.16×10^{-1} T	1.36	4.03	−0.19	

TABLE 2.32
Assessment Parameters (Example 3)

CL Patch	CL Patch Label	CNR (dB)	Number of Cells
A	1	9.34	124
C + D	2	10.63	341
B	3	13.47	306

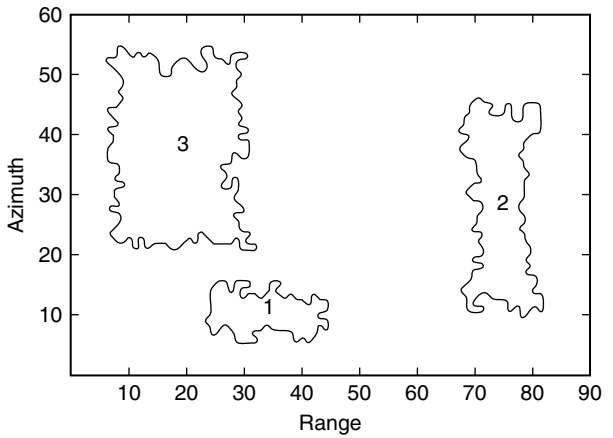

FIGURE 2.89 Contour plot of the mapped volume after numbering (Example 2).

TABLE 2.33
Mapping Procedure Applied to CL Patch 1 (Example 3)

PLQP (%)	Parameter Values	
10.00	NCQ = 8	PLCQP = 48.94
(guess)	NCC = 1	PLCCP = 8.51
20.00	NCQ = 8	PLCQP = 82.98
modified PLQP	NCC = 1	PLCCP = 51.06
51.06	NCQ = 8	PLCQP = 100.0
previous PLCCP	NCC = 1	PLCCP = 100.0

TABLE 2.34
Mapping Procedure Applied to CL Patch 2 (Example 3)

PLQP (%)	Parameter Values	
10.00	NCQ = 8	PLCQP = 42.93
(guess)	NCC = 1	PLCCP = 11.62
11.62	NCQ = 8	PLCQP = 53.54
previous PLCCP	NCC = 1	PLCCP = 13.64
13.64	NCQ = 8	PLCQP = 64.14
previous PLCCP	NCC = 1	PLCCP = 18.69
18.69	NCQ = 8	PLCQP = 77.27
previous PLCCP	NCC = 1	PLCCP = 32.32
32.32	NCQ = 8	PLCQP = 94.44
previous PLCCP	NCC = 1	PLCCP = 69.70
69.70	NCQ = 8	PLCQP = 100.0
previous PLCCP	NCC = 1	PLCCP = 100.0

TABLE 2.35
Mapping Procedure Applied to CL Patch 3 (Example 3)

PLQP (%)	Parameter Values	
10.00	NCQ = 8	PLCQP = 46.53
(guess)	NCC = 1	PLCCP = 14.36
14.36	NCQ = 8	PLCQP = 70.30
previous PLCCP	NCC = 1	PLCCP = 26.24
26.24	NCQ = 8	PLCQP = 92.08
previous PLCCP	NCC = 1	PLCCP = 66.83
66.83	NCQ = 8	PLCQP = 100.0
previous PLCCP	NCC = 1	PLCCP = 100.0

As in example 2, CL patch two consists of subpatches C and D. In Chapter 7 it is shown that expert system rules are unable, in this case, to further subdivide the CL patch. However, it is also shown that the PDFs of subpatches C and D are very similar with identical power levels. Consequently, the inability to subdivide CL patch two is not a serious problem.

2.6.5.3.3. Reassessment

Because the CL subpatch identification does not result in finding any subpatches within CL patches one, two, and three, the mapped volume is not changed and assessment is not necessary (see Figure 2.90).

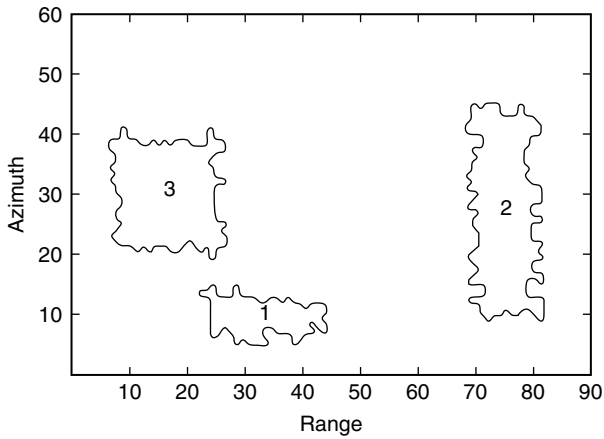

FIGURE 2.90 Contour plot of the mapped volume after numbering (Example 3).

2.6.5.3.4. Steps 1–5 of the PDF Approximation Strategy

Tables 2.36–2.38 list the results of steps 1–5 of the PDF approximation strategy applied to CL patches one, two, and three, respectively.

TABLE 2.36
PDF Approximation for CL Patch 1 (Example 3)

Test Cell No.	PDF	Rank	Distance	Shape	Scale	Location	(*u*, *v*)
1	W	3	0.30×10^{-1} T	2.20	3.90	−0.01	(−0.0250, 0.3733)
	L	7	0.44×10^{-1} Ψ	0.13	12.4	−9.01	
	K	5	0.39×10^{-1} T	50.0	0.49	0.40	
31	W	3	0.27×10^{-1} T	2.32	3.98	−0.07	(−0.0233, 0.3702)
	L	7	0.41×10^{-1} T	0.12	12.9	−9.60	
	K	4	0.36×10^{-1} T	50.0	0.49	0.42	
62	W	2	0.97×10^{-2} T	2.00	3.69	−0.32	(−0.0342, 0.3529)
	L	5	0.25×10^{-1} T	0.20	8.41	−5.00	
	K	3	0.17×10^{-1} T	50.0	0.51	0.54	
93	W	3	0.28×10^{-1} T	2.32	3.98	−0.07	(−0.0238, 0.3709)
	L	7	0.41×10^{-1} T	0.12	13.0	−9.60	
	K	4	0.37×10^{-1} T	50.0	0.49	0.41	
124	W	3	0.28×10^{-1} T	2.10	3.77	0.12	(−0.0276, 0.3727)
	L	7	0.41×10^{-1} T	0.15	10.8	−7.53	
	K	4	0.36×10^{-1} T	50.0	0.49	0.39	

TABLE 2.37
PDF Approximation for CL Patch 2 (Example 3)

Test Cell No.	PDF	Rank	Distance	Shape	Scale	Location	(*u*, *v*)
1	W	15	$0.94 \times 10^{-1}\ \Psi$	5.00	3.06	1.24	(0.0308, 0.2437)
	L	14	$0.94 \times 10^{-1}\ \Psi$	0.01	81.0	−76.96	
	K	20	$0.11 \times 10^{-0}\ \Psi$	50.0	0.19	2.89	
1	W	2	0.13×10^{-1} T	2.39	1.20	3.07	(−0.0227, 0.3554)
w/o 5	L	6	0.26×10^{-1} T	0.13	3.58	0.52	
Outliers	K	4	0.25×10^{-1} T	50.0	0.14	3.24	
85	W	15	$0.11 \times 10^{-0}\ \Psi$	5.00	3.04	1.19	(0.0282, 0.2214)
	L	14	$0.11 \times 10^{-0}\ \Psi$	0.01	80.5	−76.56	
	K	18	$0.13 \times 10^{-0}\ \Psi$	36.5	0.22	2.84	
85	W	1	0.79×10^{-2} T	2.19	1.03	3.14	(−0.0281, 0.3509)
w/o 6	L	5	0.23×10^{-1} T	0.17	2.51	1.51	
Outliers	K	3	0.18×10^{-1} T	50.0	0.13	3.23	
171	W	22	$0.18 \times 10^{-0}\ \Psi$	0.49	0.24	3.68	(0.0365, 0.2186)
	L	14	$0.12 \times 10^{-0}\ \Psi$	0.01	80.0	−76.24	
	K	18	$0.14 \times 10^{-0}\ \Psi$	50.0	0.19	2.78	
171	W	12	0.32×10^{-1} T	2.35	1.16	2.97	(−0.0258, 0.3105)
w/o 5	L	3	0.17×10^{-1} T	0.22	2.05	1.91	
Outliers	K	11	0.30×10^{-1} T	43.8	0.15	3.13	
256	W	15	$0.47 \times 10^{-1}\ \Psi$	4.70	2.45	1.71	(−0.0005, 0.2915)
	L	12	$0.40 \times 10^{-1}\ \Psi$	0.05	9.86	−5.89	
	K	18	$0.55 \times 10^{-1}\ \Psi$	50.0	0.16	2.98	
256	W	1	0.39×10^{-2} T	2.29	1.14	2.98	(−0.0256, 0.3468)
w/o 3	L	5	0.18×10^{-1} T	0.16	2.88	1.07	
Outliers	K	4	0.17×10^{-1} T	50.0	0.14	3.11	
341	W	23	$0.19 \times 10^{-0}\ \Psi$	0.39	0.16	3.72	(0.0469, 0.2188)
	L	14	$0.12 \times 10^{-0}\ \Psi$	0.01	94.1	−90.21	
	K	19	$0.14 \times 10^{-0}\ \Psi$	50.0	0.22	2.57	
341	W	8	0.24×10^{-1} T	2.36	1.21	2.95	(−0.0253, 0.3186)
w/o 7	L	1	0.94×10^{-2} T	0.18	2.56	1.42	
Outlier	K	7	0.24×10^{-1} T	50.0	0.14	3.12	

2.6.6. Extension of the Indexing Procedure to Range–Azimuth–Doppler Cells

It is possible to extend the indexing procedure to the 3-D R/A/D plane. The same steps used in indexing the R/A plane are followed. Namely, (1) an assessment stage is utilized to assign a unique number to the BN region and each 3-D

TABLE 2.38
PDF Approximation for CL Patch 3 (Example 3)

Test Cell No.	PDF	Rank	Distance	Shape	Scale	Location	(*u*, *v*)
1	W	14	0.73×10^{-1} Ψ	0.96	2.90	2.39	(−0.0530, 0.2596)
	L	6	0.47×10^{-1} Ψ	0.72	3.07	1.41	
	K	10	0.62×10^{-1} Ψ	3.03	3.06	0.29	
1	W	12	0.33×10^{-1} T	2.23	5.36	0.03	(−0.0291, 0.3098)
w/o 3	L	3	0.16×10^{-1} T	0.24	8.91	−4.35	
Outliers	K	10	0.29×10^{-1} T	35.3	0.81	0.54	
76	W	14	0.71×10^{-1} Ψ	0.98	3.00	2.39	(−0.0519, 0.2632)
	L	6	0.45×10^{-1} Ψ	0.69	3.26	1.3	
	K	10	0.59×10^{-1} Ψ	3.33	2.97	0.58	
76	W	10	0.29×10^{-1} T	2.18	5.39	0.05	(−0.0300, 0.3135)
w/o 3	L	2	0.13×10^{-1} T	0.24	9.20	−4.61	
Outliers	K	7	0.25×10^{-1} T	38.6	0.79	0.48	
153	W	5	0.17×10^{-1} T	1.31	3.49	1.61	(−0.0626, 0.3176)
	L	1	0.39×10^{-2} T	0.45	4.78	−0.48	
	K	2	0.58×10^{-2} T	7.51	1.81	0.49	
229	W	3	0.36×10^{-1} T	1.50	4.20	1.29	(−0.0673, 0.3742)
	L	7	0.57×10^{-1} Ψ	0.35	6.70	−2.14	
	K	4	0.45×10^{-1} T	42.9	0.82	0.24	
306	W	3	0.25×10^{-1} T	1.50	4.22	1.23	(−0.0645, 0.3623)
	L	6	0.44×10^{-1} T	0.36	6.59	−2.02	
	K	4	0.33×10^{-1} T	31.5	0.95	0.18	

CL patch, compute the CNR for each CL patch, and classify CL patches as to whether they are SSC, WSC, or ISC regions, (2) a CL subpatch investigation stage is utilized to subdivide nonhomogeneous CL patches into contiguous CL subpatches, and (3) a random data analysis stage is utilized to obtain a PDF approximation of WSC CL patches stage.

Even though the indexing procedure for the R/A/D volume follows the same steps as those shown in Figures 2.76–2.78 for the R/A plane, some changes are needed to apply the indexing algorithms of the R/A plane to the R/A/D volume. These changes are:

(i) The mask used in the numbering procedure and shown in Figure 2.73 becomes a 3-D mask which consists of the 13 previously numbered neighboring cells to the test cell. This is illustrated in Figure 2.91 for the *ijk*th test cell.

(ii) For the choice of the reference cells, a three by three by three mask of neighboring cells is initially centered around the test cell instead

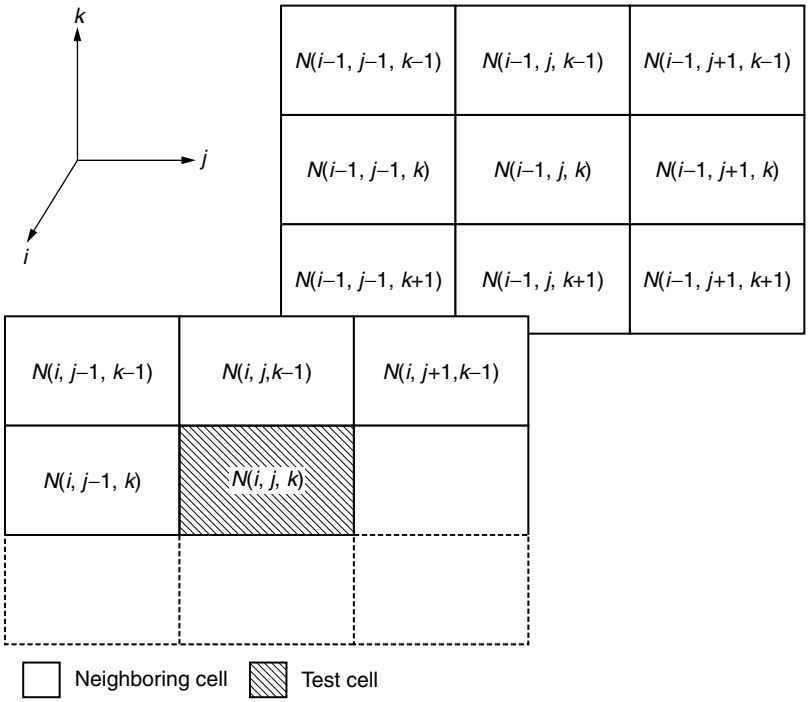

FIGURE 2.91 Mask used in numbering the range–azimuth–Doppler volume.

of the three by three one used for the R/A plane. Then, the mask is augmented by one plane at each of its boundaries resulting in a five by five by five mask. As was done for the R/A case, the process of adding one row and one column to each boundary of the previous mask continues until the desired number of reference cells have been obtained. Note that the process of choosing a specified number of reference cells in the R/A/D volume results in the cells being closer to the test cell than for the R/A plane. For example, in order to choose 100 reference cells in a homogeneous region, a mask of dimension 11×11 is needed in the R/A plane whereas a mask of dimension five by five by five suffices in the R/A/D volume. Thus, the reference cells are a distance of up to five cells away from the test cell in the R/A plane, whereas the distance is only up to two cells away from the test cell in the R/A/D plane.

2.6.7. Conclusion

In summary, an indexing procedure has been presented that allows for numbering each of the BN region and CL patches with a unique number, extracting CL subpatches, and approximating the PDF of the test cells in the CL

patches. The procedure was illustrated with examples that show how the indexing procedure works.

With respect to the indexing procedure presented in Table 2.1, every step of the WSC has been discussed in this chapter. It should be pointed out that the procedures used for choosing test and reference cells in SSC and ISC are same as those used for WSC.

Note that detection is not handled in this work because it is assumed that the appropriate detection algorithms to be used are known once the mapping and indexing are handled correctly.

Up to this point, only the SUD/FFES has been presented. Rules are presented in the next chapter to enable SUD/FBES.

2.7. APPLICATION OF IPUS TO THE RADAR DETECTION PROBLEM

Various aspects of the radar detection problem were illustrated in Figure 2.5. The IPUS architecture decides whether or not a weak signal situation exists by classifying the data to be processed into either SSC, ISC, or WSC. As explained in Table 2.1, the SSC uses the LRT, the ISC uses the GLRT, and the WSC uses the LOD.

Thus far, the mapping and indexing stages for the FFES, shown in Figure 2.4, have been detailed. In this chapter, rules are developed which allow IPUS to supervise (1) the convergence process in the mapping procedure, and (2) the interpretation of the indexing procedure.

Section 2.7.1 summarizes IPUS concepts. In Sections 2.7.2 and 2.7.4, rules are developed for the mapping and indexing procedures, respectively, along with examples which illustrate application of the rules.

2.7.1. Summary of IPUS Concepts

The following items summarize the concepts upon which IPUS is based:

- *Signal processing algorithm (SPA)*: one or more SPAs and the corresponding control parameters need to be defined. The set of values assigned to the control parameters at a given instance is referred to as an SPA instance.
- *Discrepancy detection*: when the signal being monitored by the SPA does not satisfy the requirements of the SPA instance, the output of the SPA is distorted. Recognition of the fact that SPA instance is not properly chosen for the input data stream is referred to as discrepancy detection.
- *Diagnosis procedure*: once a discrepancy has been detected, a diagnosis procedure is used to identify the source of the distortion which may have led to the discrepancy.

– *Reprocessing procedure*: knowing the source of distortion, either the parameters of the same SPA can be readjusted or a different SPA can be chosen to reprocess the data.

In the following sections, (1) SPAs are identified along with their sets of control parameters, (2) rules are developed which enable the detection of discrepancies and identification of sources of distortion, and (3) examples are presented to illustrate the result of reprocessing.

2.7.2. Role of IPUS in the Mapping Procedure

In this section, control by IPUS of the convergence process in the mapping procedure is described.

2.7.2.1. IPUS Stages Included in the Mapping Procedure

2.7.2.1.1. SPA and SPA Instance

It has been shown in the block diagram of Figure 2.28 that part of the mapping procedure consists of a set of four blocks linked by feed-forward and feed-back loops. These blocks are: thresholding or quantization, correction-quantized, correction-corrected, and assessment. Recall that these blocks are used to find the best threshold to separate between BN and CL patches.

The parameters associated with the first three blocks are BNQP, NCQ, and NCC, respectively. The assessment block defines whether or not reprocessing through the feedback loops is needed and, if so, which control parameters should be changed and what values should be assigned to the control parameters. The assessment stage computes BNCQP and BNCCP and compares them to BNQP.

In this application IPUS treats all four blocks as a single SPA. The SPA control parameters are BNQP, NCQ, and NCC. Any one set of the control parameter values is referred to as an SPA instance.

2.7.2.1.2. Observations on Setting of the Control Parameters

In this section, different effects of the control parameters are discussed. Note first that the intervals for the allowable values of the control parameters are given in Section 2.5 and are equal to

$$\begin{aligned} 0\% \le \text{BNQP} \le 100\% \\ 5 \le \text{NCQ} \le 8 \\ 1 \le \text{NCC} \le 4. \end{aligned} \tag{2.44}$$

Recall the BNQP represents the fraction of BN cells in the quantized volume. It is used to determine the threshold q for which all cells with data amplitudes below q are identified as BN and above q are identified as CL in the quantized volume.

Also, NCQ is the minimum number of neighboring cells in the quantized volume required to be identified as CL cells, for a test cell to be declared as a CL cell in the first-corrected volume. Finally, NCC is the minimum number of neighboring cells in the first-corrected volume required to be identified as CL cells, for a test cell to be declared as a CL cell in the second-corrected volume.

BNCQP and BNCCP are computed parameters which represent the BN percentages in the first and second-corrected volumes, respectively.

Define $BNQP_t$ to be the true value for the fraction of BN cells in the generated scene.

As explained in Section 2.5.3, the mapping processor begins by setting a threshold that results in a specified fraction of BN cells equal to BNQP. The mapping processor iterates until the latest scene is consistent with the last specified value of BNQP. When the iteration process ends, it is assumed that

$$BNQP \cong BNQP_t \tag{2.45}$$

2.7.2.1.2.1. Observations on the Setting of BNQP

1. Setting BNQP much smaller than $BNQP_t$: many cells have data amplitudes larger than the threshold resulting in a large number of BN cells being declared as CL cells in the quantized volume.
2. Setting BNQP much larger than $BNQP_t$: CL patch cells may be misclassified due to the fact that some CL patches have data amplitude values below the threshold. This results in many CL cells, being identified as BN cells in the quantized volume.

Conclusion: because (1) the objective of the mapping procedure is to separate between BN and CL patches, (2) the average power of the BN is the lowest among all regions, and (3) the threshold is set adaptively by the assessment stage, the threshold is always set very low at the beginning so that BN information is gained as the process iterates. The threshold, controlled by the assessment stage, is raised adaptively until $BNQP \cong BNQP_t$.

2.7.2.1.2.2. Observations on the Setting of NCQ. Recall that NCQ controls which test cells in the first-corrected volume are to be declared as CL. NCQ is said to be large when its value approaches eight and small when its value approaches five. The following observations relative to NCQ take into consideration that in the initial setting the BNQP is low and then is increased until BNQP approximates the true value $BNQP_t$. Depending on the setting of BNQP with respect to $BNQP_t$, four cases exist:

1. BNQP *much smaller than* $BNQP_t$:
   (a) *Setting* NCQ *small*: in this case, because many CL declared cells exist in the quantized volume due to the low threshold, small NCQ results in the building of a multitude of CL patches which are likely to be so close that they form a single big CL patch in the first-corrected volume.

(b) *Setting* NCQ *high*: here, even though many CL declared cells exist in the quantized volume due to the low threshold, high NCQ results in the building of fewer CL patches than when NCQ is small. This is due to the fact that there must be at least NCQ CL cells neighboring the test cell in the quantized volume, where NCQ is large, in order for the test cell to be declared as a CL cell in the first-corrected volume. In this case, corrections are made and some of the cells previously declared as CL cells in the quantized volume are now declared as BN cells in the first-corrected volume.

2. BNQP *close to* $BNQP_t$: when BNQP is close to its true value, the threshold is high enough to separate between the BN region and CL patches. With either small or large values for NCQ, the CL regions are well approximated. In this case, the choice of NCQ affects the classification of the inner cells of the CL regions. This is because, even though the data amplitudes of CL cells are higher than those of the BN cells, in general, some CL cells with data amplitudes lower than those of the highest BN data values exist and may be lower than the threshold.
   (a) *Setting* NCQ *small*: all test cells in the quantized volume which have at least NCQ neighboring cells are declared as CL cells in the first-corrected volume. Small NCQ helps to correctly classify the inner CL cells. However, note that small NCQ also results in misclassifying BN cells that are surrounded by at least NCQ declared CL cells.
   (b) *Setting* NCQ *high*: every test cell must have a large number of neighboring CL declared cells in the quantized volume for it to be declared as a CL cell in the first-corrected volume. This causes the procedure to misclassify some of the inner CL cells when too many of the neighboring cells have their data amplitudes falling below the threshold. In this case, the identified CL regions are not homogeneous and contain BN declared "holes."

Conclusion: The value of NCQ should be chosen as large as possible at the beginning of the iterative process when the threshold is set very low, to correctly reclassify the maximum number of BN cells misidentified at the thresholding/quantization stage. When the threshold reaches a level, close to its convergence value, NCQ should then be chosen as small to avoid nonhomogeneous CL regions.

2.7.2.2. Observations on the Setting of NCC

Because NCQ truncates the boundaries of the CL regions, NCC is used to augment the edges of the CL declared regions. NCC is said to be large when its value approaches four and small when its value approaches one. In the following

discussion it is assumed that the conclusions previously reached on the settings of BNQP and NCQ are taken into consideration so that BNQP is initially set low to be increased until it approaches its true value $BNQP_t$, while NCQ is initially set to a large value, to be decreased as BNQP approaches its true value. Four cases are then identified:

1. BNQP *much smaller then* $BNQP_t$ *and* NCQ *large*: because NCQ is set large, many CL edge cells are misclassified and associated with the BN region.
   (a) *Setting* NCC *small*: when NCC is set small, many of the edge cells are correctly reclassified from BN cells to CL cells in the second-corrected volume.
   (b) *Setting* NCC *large*: in this case, only a few misclassified CL edge cells are correctly reclassified in the second-corrected volume.
2. BNQP *close to* $BNQP_t$ *and NCQ small*: because NCQ is small, only a few CL edge cells are associated with the BN.
   (a) *Setting* NCC *small*: small NCC causes not only CL edge cells to be recovered but also BN cells to be misclassified in the second-corrected volume.
   (b) *Setting* NCC *large*: in this case, most of the CL edge cells are correctly classified in the second-corrected volume and only few BN cells are misclassified as CL cells.

Conclusion: NCC results in the recovery of CL edge cells and the misclassification of some BN cells close to the CL edge cells. In order to maximize recovery of the CL edge cells and minimize the misclassification of BN cells, NCC should be set small when NCQ is set large in order to recover a lot of CL edge cells that were lost in the first correction. On the other hand, NCC should be set large when NCQ is set small because, in this case, only a few CL edge cells need to be recovered.

2.7.2.2.1. Discrepancy Detection

In this section, rules are developed to enable the detection of discrepancies. The assessment stage of the mapping procedure consists of comparing at each step of the iteration the value for BNCCP with the corresponding BNQP. When BNCCP is not sufficiently close to BNQP, the assessment stage is said to fail. This initiates the discrepancy detection stage. Diagnosis identifies the distortion that may have caused the discrepancy and adjusts one or more of the mapping control parameters for reprocessing of the data.

The strategy behind the iterative process of mapping procedure employs two stages. In the first stage, referred to as the threshold approximation stage, BNQP is varied iteratively by the mapping processor until, as explained later, it is expected that BNQP is within 10% of its true value $BNQP_t$. The second stage, referred to as threshold fine-tuning stage, consists of iteratively varying BNQP

until it converges to within 1% of the last computed value for BNCCP. The two stages are now discussed in detail.

(a) *Threshold approximation stage*: during this stage, two sets of SPA instances are used on the same data of the surveillance volume. For sets BNQP and NCC are the same whereas NCQ is equal to seven for one set and eight for other.

Recall that NCQ is used to recognize the CL patches in the surveillance volume. First consider the situation where BNQP approximately equals $BNQP_t$. Here the threshold is such that it is possible to do a good job of separating between the BN region and CL patches. Steps are then taken to correct misclassified BN and CL data. Note that misclassifications are due to large BN data exceeding the threshold and small CL data falling below the threshold. At this point, setting NCQ to seven and eight, respectively, results in very close values for BNCQP and BNCCP due to the facts that (1) the two masks are very similar (NCQ = 8 requires that eight neighboring cells be declared CL in the quantized volume for a test cell in the first-corrected volume to be declared CL whereas NCQ = 7 requires that seven neighboring cells be declared CL in the quantized volume for a test cell in the first-corrected volume to be declared CL) and (2) only a few cells are misclassified in the quantized volume.

Now consider that BNQP is significantly smaller than $BNQP_t$. In this case many BN cells are misclassified after quantization. Even though masks with NCQ equal to seven and eight are similar, they result in BNCQP and BNCCP being considerably different due to the fact that the large number of misclassified BN cells are so many that they tend to group together. Consequently, changing NCQ from eight to seven simply results in additional BN cells grouping together to form additional CL regions and more edges. Because of this

$$[\text{BNCQP}]_{\text{NCQ}=7} < [\text{BNCQP}]_{\text{NCQ}=8} \tag{2.46}$$

and

$$[\text{BNCCP}]_{\text{NCQ}=7} < [\text{BNCCP}]_{\text{NCQ}=8}$$

Because the second-corrected volume represents the scene where CL patches and their edges are assumed to be properly recovered had the threshold BNQP been chosen properly, BNCCP tries to converge to $BNQP_t$. Thus, it is logical to begin each iteration by assigning to BNQP the latest computed value of $[\text{BNCCP}]_{\text{NCQ}=8}$. The very first value assigned to BNQP is simply a guess. This value should be such that the threshold is low. In all of our examples, the first value of BNQP is chosen equal to 10%.

Using different scenes with different values for $BNQP_t$, it has been determined near convergence that whenever the difference between $\text{BNCCP}|_{\text{NCQ}=7}$ and $[\text{BNCCP}]_{\text{NCQ}=8}$ is within 10%, then BNQP is likely to be within 10% of its true value. This is confirmed in Figure 2.92 where plots of the quantities

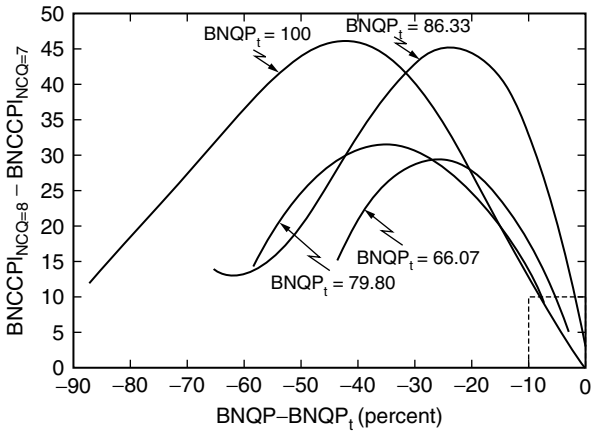

FIGURE 2.92 Plot of $[BNCQP]_{NCQ=8} - [BNCQP]_{NCQ=7}$ vs. $BNQP - BNQP_t$ for different values of $BNQP_t$.

$[BNCCP]_{NCQ} = 8 - [BNCCP]_{NCQ} = 7$ vs. $BNQP - BNQP_t$ are shown for different values of $BNQP_t$. Note that

(i) for different values of $BNQP_t$, all plots are such that near convergence the difference $[BNCCP]_{NCQ=8} - [BNCCP]_{NCQ=7}$ approaches to 0 when $BNQP - BNQP_t$ also approaches 0,
(ii) for different values of $BNQP_t$, when $BNQP - BNQPt > -10\%$, the difference $[BNCCP]_{NCQ=8} - [BNCCP]_{NCQ=7} < 10\%$.

As a result of the above, the threshold approximation stage iterates until it is satisfied that

$$[BNCCP]_{NCQ=8} - [BNCCP]_{NCQ=7} < 10\% \tag{2.47}$$

In summary, as shown in the flow chart of Figure 2.93, a guess for the initial value of BNQP is followed by the execution of the mapping procedure using two different SPA instances. The outputs of the two SPA instances are compared by means of the computed values of BNCCP. If $[BNCCP]_{NCQ=8} - [BNCCP]_{NCQ=7}$ is more than 10%, a discrepancy is detected and it is concluded that the value of BNQP differs from its true value by more than 10%. BNQP is then increased to the latest computed value of $[BNCQP]_{NCQ=8}$. BNQP is varied from one iteration to the next while NCC is kept equal to one. This choice for NCC agrees with the observations made previously where it was concluded that NCC should be set small when NCQ is set large. In this case NCQ has a large value equal to either seven or eight.

The discrepancies that may arise in the threshold approximation stage are due to the fact that two instances of the same SPA result in different interpretations

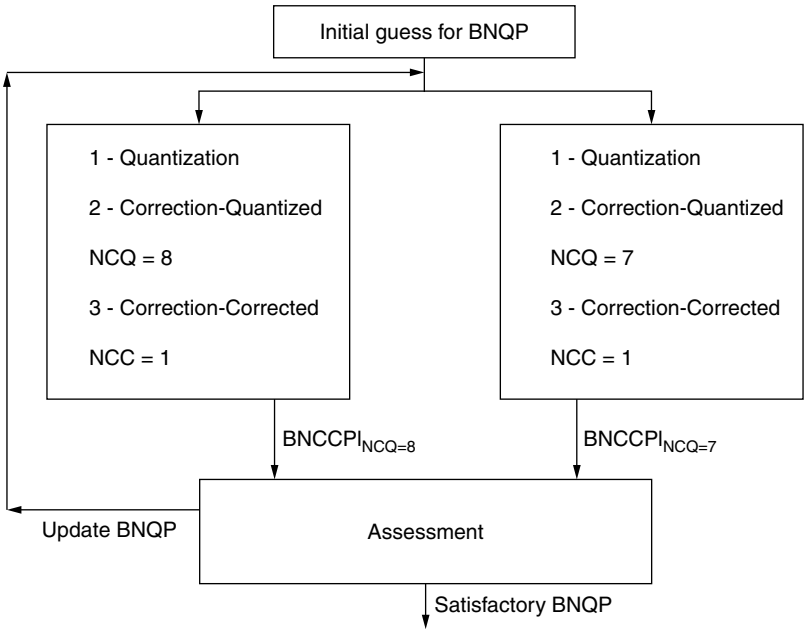

FIGURE 2.93 Threshold approximation stage, discrepancy reduction.

when applied to the same data. As defined in Section 2.3.2.2, such a discrepancy is typified as a fault.

Two fault-type discrepancies are readily identified in the threshold approximation stage. These are as follows:

1. $[\text{BNCCP}]_{\text{NCQ}=8} - [\text{BNCCP}]_{\text{NCQ}=7} > 10\%$: as discussed above, the goal of the threshold approximation stage is to obtain a threshold BNQP that is within 10% of its true value. As shown in Figure 2.92, this is likely only when $[\text{BNCCP}]_{\text{NCQ}=8} - [\text{BNCCP}]_{\text{NCQ}=7} < 10\%$. When the difference between the computed thresholds $[\text{BNCCP}]_{\text{NCQ}=8} - [\text{BNCCP}]_{\text{NCQ}=7}$ is more than 10%, a fault type of discrepancy is detected during the assessment stage. The diagnosis process identifies the fact that $\text{BNQP} - \text{BNQP}_t < (-)10\%$ as the source for the distortion causing the discrepancy. The remedy, in this case, is to increase the value of BNQP during the reprocessing stage to the latest computed value for $[\text{BNCCP}]_{\text{NCQ}=8}$.
2. *Initial* BNQP *set too low*: in some cases, when the initial guess for BNQP is too small, the number of BN cells with data exceeding the threshold is so large that when corrections are made, the second-corrected volume results in many CL declared patches or, in the worst case, a single big CL patch. This results in the values of either $[\text{BNCCP}]_{\text{NCQ}=8}$ or

TABLE 2.39
Fault-Type Discrepancies in the Threshold Approximation Stage

Discrepancy (Fault Type)	Diagnosis (Source of Distortion)	Reprocessing
$[BNCCP]_{NCQ=8} - [BNCCP]_{NCQ=7}$ is more than 10%	$BNQP - BNQP_t > (-)\ 10\%$	Assign to BNQP the latest value of $[BNCCP]_{NCQ=8}$
$[BNCCP]_{NCQ=8} - [BNCCP]_{NCQ=7}$ is less than 10% in the early stages of iteration	Initial value of BNQP is too low	Increase BNQP by 10% from its initial value

$[BNCCP]_{NCQ=7}$ being even smaller than BNQP. In this case, it is possible to obtain a value for the difference $[BNCCP]_{NCQ=8} - [BNCCP]_{NCQ=7}$ that is smaller than 10%. The IPUS control must be suspicious of such a case and declare BNQP-set-too-low as the source for the distortion causing the discrepancy. The remedy, in this case, is to increase the initial value of BNQP during the reprocessing stage. In our examples, we choose to increase BNQP by 10% every time an initial-BNQP-set-too-low fault is obtained.

Table 2.39 summarizes the discrepancies that may occur during the threshold approximation stage.

(b) *Threshold fine-tuning stage*: at the end of the threshold approximation stage BNQP is likely to be within 10% of its true value $BNQP_t$. During the threshold fine-tuning stage, BNQP is varied until BNCCP is within 1% of the corresponding value of BNQP.

During the threshold fine-tuning stage, NCQ is lowered to avoid holes in the CL patches caused by misclassified CL cells whose data values are lower than the threshold. For the same reason BNQP is assigned the latest value of $[BNCCP]_{NCQ=7}$ rather than the latest value of $[BNCCP]_{NCQ=8}$. In addition, the value of NCC is raised to avoid misclassification of BN cells close to the CL edges. These choices for NCQ and NCC agree with observations mentioned in Section 2.7.2.1.2.

The following observations on BNQP, NCQ, and NCC are necessary to understand how these parameters should be automatically set in order for BNCCP to converge to within 1% of BNQP.

1. When BNQP is increased while NCQ and NCC are kept constant, the number of BN cells in the quantized volume is increased and, therefore, BNCQP and BNCCP are likely to increase.
2. When NCQ is increased while BNQP and NCC are kept constant, the requirement on a test cell to be declared as a CL cell in the first-corrected volume becomes more stringent and, therefore, the number

of CL cells in both corrected volumes are likely to decrease. This tends to increase the number of BN cells causing BNCQP and BNCCP to increase.

3. When NCC is increased while BNQP and NCQ are kept constant, the requirement on a test cell to be declared as a CL cell in the second-corrected volume becomes more stringent and, therefore, the number of CL cells in this volume is decreased. Thus, the number of BN cells in the second-corrected volume increases and, consequently, BNCCP increases.

Using the above observations, the following strategy is used by the assessment stage to control the threshold fine-tuning stage,

1. Because BNQP is within 10% of its true value at the beginning of the threshold fine-tuning stage, the threshold is likely to be relatively high. Thus, NCQ should be set to its smallest value of five while, as needed, NCC should be incremented iteratively from its minimum value of one up to its maximum value of four.
2. When the inequality in Equation 2.47 is not satisfied, BNQP should be increased in small steps. Otherwise, the iterative process diverges when the same rule from the threshold approximation stage is used. The approach taken in this work during the threshold fine-tuning stage consists of assigning a value to BNQP that is half way between its latest value and the latest value of BNCCP, i.e.,

$$\text{BNQP} = \frac{[\text{BNQP}]_{\text{latest}} + [\text{BNCCP}]_{\text{latest}}}{2} \tag{2.48}$$

3. The condition set forward for ending the threshold fine-tuning stage is given by

$$|\text{BNQP} - \text{BNCCP}| < 1\% \tag{2.49}$$

Two cases are possible when the inequality in Equation 2.49 is not satisfied: either BNQP $<$ BNCCP or BNQP $>$ BNCCP.

4. When the inequality in Equation 2.49 is not satisfied and BNQP $<$ BNCCP, the control parameters should be varied by the diagnosis procedure such that BNCCP is decreased. In this case, NCQ should be made smaller. If none of the allowable values for NCQ result in the inequality in Equation 2.49 being satisfied, then BNQP is varied according to Equation 2.48.
5. When the inequality in Equation 2.49 is not satisfied and BNQP $>$ BNCCP, the control parameters should be varied by the diagnosis procedure such that BNCCP is increased. In this case, NCC should be made larger. If none of the allowable values for NCC result in the inequality in Equation 2.49 being satisfied, then BNQP is varied according to Equation 2.48.

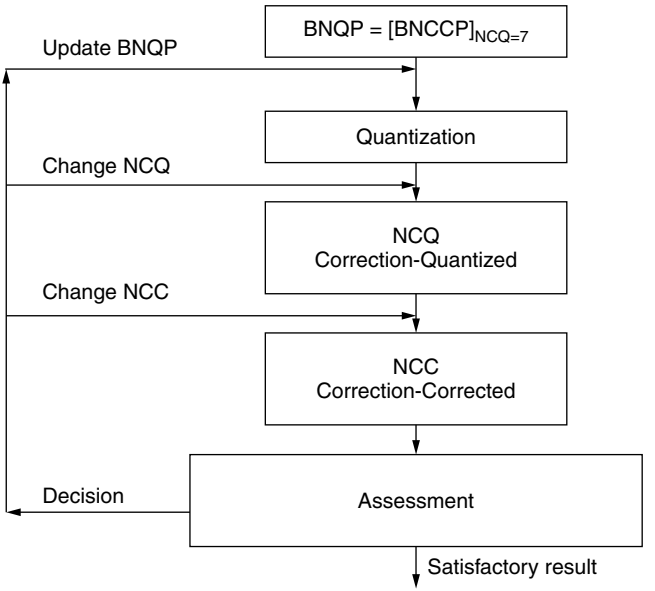

FIGURE 2.94 Threshold fine-tuning stage, discrepancy reduction.

During the threshold approximation stage note that only BNQP was varied. On the other hand, in the threshold fine-tuning stage, any of the parameters BNQP, NCQ, and NCC may be varied.

In summary, as shown in the flow chart of Figure 2.94, the threshold fine-tuning stage begins by assigning to BNQP the latest value of $\text{BNCCP}|_{\text{NCQ}=7}$. Once quantization, first correction, and second correction stages are completed with preselected values for NCQ and NCC, the assessment stage diagnoses the results according to the strategy discussed above, and, depending on the outcome, decides either that reprocessing is necessary with adjusted values for any of the BNQP, NCQ, and NCC parameters, or the threshold fine-tuning stage is completed.

At the end of each iteration of the threshold fine-tuning stage it is expected that the computed value of BNCCP will be within 1% of BNQP. When the inequality in Equation 2.49 is not satisfied, a conflict type of discrepancy is detected based on the inconsistency in the expectation that BNCCP will be within 1% of BNQP. Table 2.40 summarizes the discrepancies that may occur during the threshold fine-tuning stage.

It has been determined through examples that the initial setting of NCQ = 5 is adequate for the threshold fine-tuning stage to converge (i.e., in the examples studied it was never necessary to decrease NCQ).

In the following section, examples are presented which illustrate use of the rules developed for the threshold approximation and threshold fine-tuning stages.

TABLE 2.40
Fault-Type Discrepancies in the Threshold Fine-Tuning Stages

Discrepancy (Conflict Type)	Diagnosis (Source of Distortion)	Reprocessing
$\lvert BNQP - BNCCP\rvert > 1\%$ and $BNQP < BNCCP$	Either NCQ or BNQP are not well adjusted	Decrease NCQ, otherwise update BNQP
$\lvert BNQP - BNCCP\rvert > 1\%$ and $BNQP > BNCCP$	Either NCC or BNQP are not well adjusted	Increase NCC, otherwise update BNQP

2.7.3. Examples of Mapping

In this section, tables describing previous results from the mapping procedure of examples one, two and three in Section 2.5 are modified to demonstrate operation of the rules developed in this chapter for the threshold approximation and threshold fine-tuning stages. In the following examples, define

$$\Delta_1 = [BNCCP]_{NCQ=8} - [BNCCP]_{NCQ=7} \tag{2.50}$$

and

$$\Delta_2 = |BNQP - BNCCP| \tag{2.51}$$

2.7.3.1. Example 1

First, consider the example presented in Section 2.5.4.2.1. Table 2.41 is an expanded version of Table 2.3 in order to include information about the procedures for the threshold approximation and threshold fine-tuning stages. In the first step of the threshold fine-tuning stage note that $\Delta_2 > 1\%$ and $BNCCP < BNQP$. Using the decision rules in Table 2.40, NCC is increased in the next steps.

2.7.3.2. Example 2

Consider now the example presented in Section 2.5.4.2.2. Tables 2.42 is an expanded version of Table 2.3 in order to include information about the procedures for the threshold approximation and threshold fine-tuning stages. In the first step of the threshold fine-tuning stage note that $\Delta_2 > 1\%$ and $BNCCP > BNQP$. Using the decision rules in Table 2.40, BNQP is increased using the half way rule of Equation 2.48.

2.7.3.3. Example 3

Consider now the example presented in Section 2.5.4.2.3. Table 2.43 is an expanded version of Table 2.9 in order to include information about the procedures for the threshold approximation and threshold fine-tuning stages.

TABLE 2.41
Setting of BNQP in the Threshold Approximation and Threshold Fine-Tuning Stages (Example 1)

BNQP (%)	Parameter Values		Δ_i
	Threshold Approximation Stage		
10.00 (guess)	NCQ = 8 NCC = 1 BNCQP = 51.56 BNCCP = 22.00	NCQ = 7 NCC = 1 BNCQP = 25.78 BNCCP = 7.09	$\Delta_1 > 10\%$
22.00 Latest BNCCP	NCQ = 8 NCC = 1 BNCQP = 68.28 BNCCP = 50.39	NCQ = 7 NCC = 1 BNCQP = 54.06 BNCCP = 25.74	$\Delta_1 > 10\%$
50.39 Latest BNCCP	NCQ = 8 NCC = 1 BNCQP = 75.56 BNCCP = 68.46	NCQ = 7 NCC = 1 BNCQP = 70.93 BNCCP = 63.37	$\Delta_1 < 10\%$
	Threshold Fine-Tuning Stage		
63.37 Latest BNCCP	NCQ = 5 NCC = 1 BNCQP = 67.35 BNCCP = 59.72 < BNQP		$\Delta_2 > 1\%$
63.37 Latest BNCCP	NCQ = 5 NCC = 2 BNCQP = 67.35 BNCCP = 62.26 < BNQP		$\Delta_2 > 1\%$
63.37 Same as latest BNQP	NCQ = 5 NCC = 3 BNCQP = 67.35 BNCCP = 64.17		$\Delta_2 < 1\%$

As was the case in example one, note that $\Delta_2 > 1\%$ in the first step of the threshold fine-tuning stage and BNCCP < BNQP. Once again, using the decision rules in Table 2.40, NCC is increased.

2.7.4. Role of IPUS in the Indexing Procedure

In this section, control of the indexing procedure by IPUS is described. Recall that the indexing procedure consists of the assessment, CL subpatch investigation and PDF approximation stages. Control by IPUS of these different stages is discussed next.

TABLE 2.42
Setting BNQP in the Threshold Approximation and Threshold Fine-Tuning Stages (Example 2)

BNQP (%)	Parameter Values		Δ_i
	Threshold Approximation Stage		
10.00 (guess)	NCQ = 8 NCC = 1 BNCQP = 56.35 BNCCP = 20.59	NCQ = 7 NCC = 1 BNCQP = 23.94 BNCCP = 6.33	$\Delta_1 > 10\%$
20.59 Latest BNCCP	NCQ = 8 NCC = 1 BNCQP = 77.43 BNCCP = 48.04	NCQ = 7 NCC = 1 BNCQP = 53.30 BNCCP = 16.94	$\Delta_1 > 10\%$
48.04 Latest BNCCP	NCQ = 8 NCC = 1 BNCQP = 91.83 BNCCP = 81.70	NCQ = 7 NCC = 1 BNCQP = 84.11 BNCCP = 72.30	$\Delta_1 < 10\%$
	Threshold Fine-Tuning Stage		
72.30 Latest BNCCP	NCQ = 5 NCC = 1 BNCQP = 82.17 BNCCP = 79.26 > BNQP		$\Delta_2 > 1\%$
75.78 Half way	NCQ = 5 NCC = 1 BNCQP = 83.50 BNCCP = 75.31		$\Delta_2 < 1\%$

2.7.4.1. IPUS Stages Included in the Assessment Stage

As explained in Section 2.7.2, the assessment stage consists of (1) assigning a unique number to the BN region and each CL patch (2) computing the CNR for each CL patch, and (3) classifying CL patches as either SSC, WSC, or ISC regions. Due to the straightforward implementation of these steps, any control by IPUS of the assessment stage is not discussed further.

2.7.4.2. IPUS Stages Included in the CL Subpatch Investigation Stage

The approach used to extract a CL subpatch from a set of contiguous CL subpatches is the same as that used in the mapping procedure to extract the BN from CL patches. Therefore, control by IPUS is needed in the CL subpatch investigation stage.

TABLE 2.43
Setting of BNQP in the Threshold Approximation and Threshold Fine-Tuning Stages (Example 3)

BNQP (%)	Parameter Values		Δ_i
	Threshold Approximation Stage		
10.00 (guess)	NCQ = 8 NCC = 1 BNCQP = 56.17 BNCCP = 20.04	NCQ = 7 NCC = 1 BNCQP = 24.24 BNCCP = 6.09	$\Delta_1 > 10\%$
20.04 Latest BNCCP	NCQ = 8 NCC = 1 BNCQP = 77.98 BNCCP = 43.78	NCQ = 7 NCC = 1 BNCQP = 51.22 BNCCP = 14.83	$\Delta_1 > 10\%$
43.78 Latest BNCCP	NCQ = 8 NCC = 1 BNCQP = 90.78 BNCCP = 82.65	NCQ = 7 NCC = 1 BNCQP = 84.78 BNCCP = 66.09	$\Delta_1 > 10\%$
82.65 Latest BNCCP	NCQ = 8 NCC = 1 BNCQP = 93.43 BNCCP = 89.39	NCQ = 7 NCC = 1 BNCQP = 91.17 BNCCP = 84.70	$\Delta_1 < 10\%$
	Threshold Fine-Tuning Stage		
84.7 Latest BNCCP	NCQ = 5 NCC = 1 BNCQP = 87.61 BNCCP = 82.00 < BNQP		$\Delta_2 > 1\%$
84.70 Latest BNCCP	NCQ = 5 NCC = 2 BNCQP = 87.61 BNCCP = 83.65 < BNQP		$\Delta_2 > 1\%$
84.70 Same as latest BNQP	NCQ = 5 NCC = 3 BNCQP = 87.61 BNCCP = 85.46		$\Delta_2 < 1\%$

2.7.4.2.1. SPA and SPA Instance

Recall that the CL subpatch investigation stage attempts to extract first, if it exists, the CL subpatch having the lowest average power. This part of the CL subpatch investigation stage consists of a set of four blocks linked by feedforward and feedback loops. These blocks are: thresholding or quantization, correction–quantized, correction–corrected, and assessment.

The parameters associated with first three blocks are PLQP, NCQ, and NCC, respectively. The assessment stage computes PLCQP and PLCCP and compares them to PLQP. IPUS treats all four blocks as a single SPA. The control parameters for the SPA are PLQP, NCQ, and NCC. Any single set of the control parameter values is referred to as an SPA instance.

2.7.4.2.2. Observations on the Setting of the Control Parameters

The observations previously made for the mapping procedure parameters (i.e., BNQP, NCQ, and NCC) also hold for the CL subpatch investigation parameters (i.e., PLQP, NCQ, and NCC).

2.7.4.2.3. Resolution of Discrepancies

As was the case for the mapping procedure, threshold approximation and threshold fine-tuning stages are used in the CL subpatch investigation stage to enable resolution of discrepancies. Because the procedure for extracting a CL subpatch is the same as that used for separating BN and CL patches, discrepancy detection rules for both procedures are the same. The threshold approximation stage block diagram of Figure 2.93 becomes that of Figure 2.95 where BNQP and BNCCP have been replaced by PLQP and PLCCP, respectively. Let $PLQP_t$ represent the true value for PLQP, the threshold approximation stage iterates until it is satisfied that PLQP is within 10% of its true value $PLQP_t$. This is satisfied when

$$[PLCCP]_{NCQ=8} - [PLCCP]_{NCQ=7} < 10\% \tag{2.52}$$

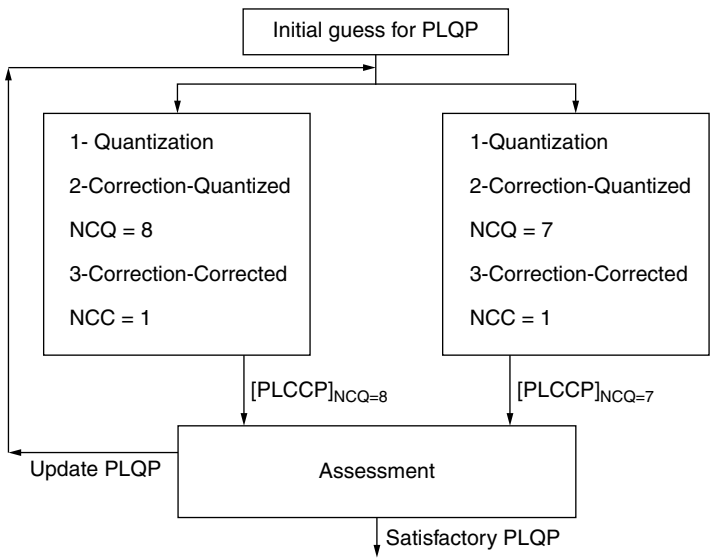

FIGURE 2.95 Threshold approximation stage, resolution of discrepancies.

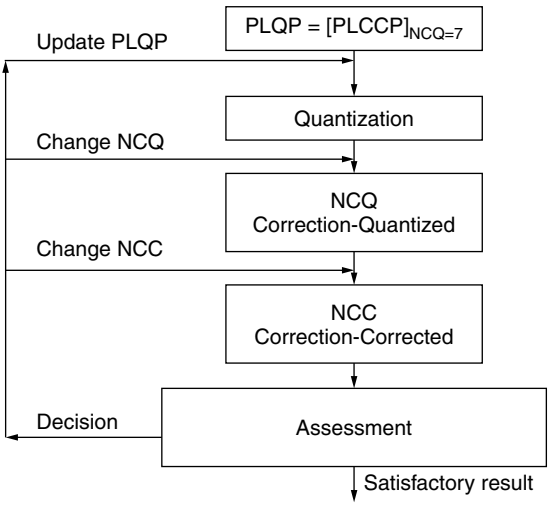

FIGURE 2.96 Threshold fine-tuning stage, resolution of discrepancies.

The threshold fine-tuning stage converges when

$$|\text{PLQP} - \text{PLCCP}| < 1\% \tag{2.53}$$

Figure 2.96 shows the block diagram of the threshold fine-tuning stage.

In addition to the discrepancy rules described in Section 2.7.2.1.3 for the threshold approximation and fine-tuning stages and summarized in Tables 2.39 and 2.40, additional rules are introduced next as a result of some more observations on the behavior of PLQP for the case where a CL patch is homogeneous and does not contain subpatches. These are,

1. It is noted that during the threshold approximation stage it is possible that the inequality in Equation 2.52 will be satisfied with $[\text{PLCCP}]_{\text{NCQ=7 or 8}} = 100\%$. This means that the subpatch with the smallest average power occupies 100% of the CL patch area. Consequently, the CL patch is homogeneous.
2. When the inequality in Equation 2.52 is met with $[\text{PLCCP}]_{\text{NCQ=7}} = 100\%$, there is no need for the threshold fine-tuning stage. This is because the initial value of PLQP in the threshold fine-tuning stage would be equal to 100% and any more processing would also result in PLCCP = 100% regardless of the values chosen for NCQ and NCC.
3. If the threshold fine-tuning stage results in PLCCP = 100% at any iteration, the threshold fine-tuning stage should end because, as in observation 2, any more processing will end with PLCCP = 100% regardless of the values chosen for NCQ and NCC. This, in turn, will make PLQP equal to 100%.

TABLE 2.44
Conflict-Type Discrepancies in the Threshold Approximation Stage

Discrepancy (Conflict Type)	Diagnosis (Source of Distortion)	Reprocessing
$[PLCCP]_{NCQ=8} - [PLCCP]_{NCQ=7}$ is more than 10%	$PLQP - PLQP_t > -10\%$	Assign to PLQP the latest value of $[PLCCP]_{NCQ=8}$
$[PLCCP]_{NCQ=8} - [PLCCP]_{NCQ=7}$ is less than 10% in the early stages of iteration	Initial value of PLQP is too low	Increase PLQP by 10% from its initial value
$[PLCCP]_{NCQ=8} - [PLCCP]_{NCQ=7}$ is less than 10% and $PLCCP|_{NCQ=7or8} = 100\%$	CL patch homogeneous	No threshold fine-tuning is needed

Tables 2.44 and 2.45 summarize the discrepancies that may occur during the threshold approximation and fine-tuning stages for the CL subpatch investigation.

Examples are next presented to illustrate the control by IPUS of the CL subpatch investigation.

2.7.4.3. Examples

In this section, tables describing the CL subpatch investigation of examples one, two and three in Section 2.6 are modified to include the rules developed in this chapter for the threshold approximation and threshold fine-tuning stages. In the following examples, define $\Delta_1 = [PLCCP]_{NCQ=8} - [PLCCP]_{NCQ=7}$ and $\Delta_2 = |PLQP - PLCCP|$.

2.7.4.3.1. Example 1

First consider the example of Section 2.6.5.1. Tables 2.46–2.48 are expanded versions of Tables 2.17–2.19 to include information about the procedures for

TABLE 2.45
Conflict-Type Discrepancies in the Threshold Fine-Tuning Stage

Discrepancy (Conflict Type)	Diagnosis (Source of Distortion)	Reprocessing
$\|PLQP - PLCCP\| > 1\%$ and $PLQP < PLCCP$	Either NCQ or PLQP are not well adjusted	Decrease NCQ, otherwise update PLQP
$\|PLQP - PLCCP\| > 1\%$ and $PLQP > PLCCP$	Either NCC or PLQP are not well adjusted	Increase NCC, otherwise update PLQP
$\|PLQP - PLCCP\| > 1\%$ and $PLCCP = 100\%$	CL patch homogeneous	Stop threshold fine-tuning

TABLE 2.46
Setting of PLQP in the Threshold Approximation and Threshold Fine-Tuning Stages, CL Patch 1 (Example 1)

	CL Patch 1		
PLQP (%)			Δ_i
	Threshold Approximation Stage		
10.00 (guess)	NCQ = 8 NCC = 1 PLCQP = 70.59 PLCCP = 20.59	NCQ = 7 NCC = 1 PLCQP = 23.53 PLCCP = 2.94	$\Delta_1 > 10\%$
20.59 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 94.12 PLCCP = 82.35	NCQ = 7 NCC = 1 PLCQP = 79.41 PLCCP = 38.24	$\Delta_1 > 10\%$
82.35 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 100.0 PLCCP = 100.0	NCQ = 7 NCC = 1 PLCQP = 100.0 PLCCP = 100.0	$\Delta_1 < 10\%$

the threshold approximation and threshold fine-tuning stages. Note in Table 2.46 no threshold fine-tuning is needed for CL patch 1. This is due to the fact that CL patch 1 is homogeneous and does not contain subpatches. Also in Table 2.48 note that the threshold fine-tuning stage needs not be carried out as PLCCP = 100% in the last row before of the threshold approximation stage. However, when the fine-tuning stage is carried out, PLQP converges to 99.33%.

2.7.4.3.2. Example 2

Consider the example of Section 2.6.5.2. Tables 2.49–2.51 are expanded versions of Tables 2.52–2.54 to include information about the procedures for the threshold approximation and threshold fine-tuning stages. In Table 2.50 note that the threshold fine-tuning steps, presented in the shaded rows, need not be carried out.

2.7.4.3.3. Example 3

Finally, consider the example of Section 2.6.5.3. Tables 2.52–2.54 are expanded versions of Tables 2.33–2.35 to include information about the procedures for the threshold approximation and threshold fine-tuning stages. Note from Tables 2.26–2.28 that no threshold fine-tuning is needed for CL patches one, two, and three. Also, note that that the last four shaded rows of Table 2.54 need not be carried out.

TABLE 2.47
Setting of PLQP in the Threshold Approximation and Threshold Fine-Tuning Stages, CL Patch 2 (Example 1)

PLQP (%)	CL Patch 2		Δ_i
	Threshold Approximation Stage		
10.00 (guess)	NCQ = 8 NCC = 1 PLCQP = 39.38 PLCCP = 7.50	NCQ = 7 NCC = 1 PLCQP = 9.37 PLCCP = 0.00	$\Delta_1 < 10\%$
20.00 (new guess)	NCQ = 8 NCC = 1 PLCQP = 59.38 PLCCP = 43.12	NCQ = 7 NCC = 1 PLCQP = 38.12 PLCCP = 8.75	$\Delta_1 > 10\%$
43.12 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 61.88 PLCCP = 51.25	NCQ = 7 NCC = 1 PLCQP = 55.00 PLCCP = 49.38	$\Delta_1 > 10\%$
	Threshold Fine-Tuning Stage		
49.38 Latest PLCCP	NCQ = 5 NCC = 1 PLCQP = 50.62 PLCCP = 46.25		$\Delta_2 > 1\%$
49.38 Same as latest PLQP	NCQ = 5 NCC = 2 PLCQP = 50.62 PLCCP = 46.88		$\Delta_2 > 1\%$
49.38 Same as latest PLQP	NCQ = 5 NCC = 3 PLCQP = 50.62 PLCCP = 48.75		$\Delta_2 < 1\%$

2.7.4.4. IPUS Stages Included in the PDF Approximation Stage

Recall that the objective of the PDF approximation stage is to approximate the PDF underlying a particular CL patch region. A six-step strategy for PDF approximation was presented in Section 2.6.4.5. The first five steps consisted of (1) selecting a total of N_T test cells that are evenly spread throughout the CL patch, (2) choosing for each test cell the closest $N_R = 100$ reference cells, as described in Section 2.6.4.1, (3) Using the Ozturk algorithm to determine

TABLE 2.48
Setting of PLQP in the Threshold Approximation and Threshold Fine-Tuning Stages, CL Patch 3 (Example 1)

PLQP (%)	CL Patch 3		Δ_i
	Threshold Approximation Stage		
10.00 (guess)	NCQ = 8 NCC = 1 PLCQP = 50.25 PLCCP = 12.19	NCQ = 7 NCC = 1 PLCQP = 14.61 PLCCP = 0.33	$\Delta_1 > 10\%$
12.19 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 58.51 PLCCP = 16.61	NCQ = 7 NCC = 1 PLCQP = 20.28 PLCCP = 0.83	$\Delta_1 > 10\%$
16.61 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 73.62 PLCCP = 28.80	NCQ = 7 NCC = 1 PLCQP = 33.97 PLCCP = 2.42	$\Delta_1 > 10\%$
28.80 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 92.57 PLCCP = 66.19	NCQ = 7 NCC = 1 PLCQP = 70.03 PLCCP = 22.79	$\Delta_1 > 10\%$
66.19 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 100.0 PLCCP = 100.0	NCQ = 7 NCC = 1 PLCQP = 99.83 PLCCP = 98.66	$\Delta_1 < 10\%$
	Threshold Fine-Tuning Stage		
98.66 Latest PLCCP	NCQ = 5 NCC = 1 PLCQP = 100.0 PLCCP = 100.0		$\Delta_2 > 1\%$
99.33 Half way Rule	NCQ = 5 NCC = 1 PLCQP = 100.0 PLCCP = 100.0		$\Delta_2 < 1\%$

the distance between the locus end point of the data linked vectors and its projection onto the Weibull, Lognormal, and *K*-distributed trajectories, (4) Discarding those PDFs for which the distances in step 3 exceed the corresponding half length of the minor axis obtained from Table 2.15, and (5) Excising any

TABLE 2.49
Setting of PLQP in the Threshold Approximation and Threshold Fine-Tuning Stages, CL Patch 1 (Example 2)

	CL Patch 1		
PLQP (%)			Δ_i
	Threshold Approximation Stage		
10.00 (guess)	NCQ = 8 NCC = 1 PLCQP = 36.73 PLCCP = 8.16	NCQ = 7 NCC = 1 PLCQP = 8.16 PLCCP = 2.04	$\Delta_1 < 10\%$ $\Delta_1 > 10\%$
20.00 (new guess)	NCQ = 8 NCC = 1 PLCQP = 71.43 PLCCP = 28.57	NCQ = 7 NCC = 1 PLCQP = 55.10 PLCCP = 10.20	
28.57 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 87.76 PLCCP = 55.10	NCQ = 7 NCC = 1 PLCQP = 55.10 PLCCP = 10.20	$\Delta_1 > 10\%$
55.10 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 100.0 PLCCP = 100.0	NCQ = 7 NCC = 1 PLCQP = 95.92 PLCCP = 79.59	$\Delta_1 > 10\%$
79.59 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 100.0 PLCCP = 100.0	NCQ = 7 NCC = 1 PLCQP = 100.0 PLCCP = 100.0	$\Delta_1 < 10\%$

outliers that may exist from the data and proceeding to step 3 when all possible PDFs are discarded in step 4. These steps were illustrated through examples in Section 2.6.5. Step 6 consists of the use of IPUS to determine one or more PDFs to approximate the data in a particular CL patch. Thus, IPUS is needed to complete the PDF approximation stage of the indexing procedure.

In order to be able to identify the SPA, SPA instance, and discrepancies associated with the PDF approximation stage, step 6 of the PDF approximation strategy has to be formulated.

2.7.4.4.1. Step 6 of the PDF Approximation Strategy

When the five first steps of the PDF approximation strategy are completed, the information available to step 6 consists of (1) identification numbers of the test cells, (2) best PDF(s) to approximate the data of the reference cells in every test cell along with their shape, scale, and location parameters, (3) distance from

TABLE 2.50
Setting of PLQP in the Threshold Approximation and Threshold Fine-Tuning Stages, CL Patch 2 (Example 2)

PLQP (%)	CL Patch 2		Δ_i
	Threshold Approximation Stage		
10.00 (guess)	NCQ = 8 NCC = 1 PLCQP = 60.38 PLCCP = 19.50	NCQ = 7 NCC = 1 PLCQP = 16.35 PLCCP = 0.00	$\Delta_1 > 10\%$
19.50 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 81.76 PLCCP = 49.06	NCQ = 7 NCC = 1 PLCQP = 46.54 PLCCP = 8.80	$\Delta_1 > 10\%$
49.06 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 94.97 PLCCP = 81.76	NCQ = 7 NCC = 1 PLCQP = 81.76 PLCCP = 57.23	$\Delta_1 > 10\%$
81.76 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 100.0 PLCCP = 100.0	NCQ = 7 NCC = 1 PLCQP = 99.37 PLCCP = 94.97	$\Delta_1 < 10\%$
	Threshold Fine-Tuning Stage		
94.97 Latest PLCCP	NCQ = 5 NCC = 1 PLCQP = 100.0 PLCCP = 100.0		$\Delta_2 > 1\%$
97.47 Half way Rule	NCQ = 5 NCC = 1 PLCQP = 100.0 PLCCP = 100.0		$\Delta_2 > 1\%$
98.73 Half way rule	NCQ = 5 NCC = 1 PLCQP = 100.0 PLCCP = 100.0		$\Delta_2 > 1\%$
99.36 Half way rule	NCQ = 5 NCC = 1 PLCQP = 100.0 PLCCP = 100.0		$\Delta_2 < 1\%$

TABLE 2.51
Setting of PLQP in the Threshold Approximation and Threshold Fine-Tuning Stages, CL Patch 3 (Example 2)

PLQP (%)	CL Patch 3		Δ_i
	Threshold Approximation Stage		
10.00 (guess)	NCQ = 8 NCC = 1 PLCQP = 63.36 PLCCP = 22.12	NCQ = 7 NCC = 1 PLCQP = 17.97 PLCCP = 0.69	$\Delta_1 > 10\%$
22.12 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 89.63 PLCCP = 62.90	NCQ = 7 NCC = 1 PLCQP = 63.82 PLCCP = 19.12	$\Delta_1 > 10\%$
62.90 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 100.0 PLCCP = 100.0	NCQ = 7 NCC = 1 PLCQP = 99.08 PLCCP = 94.70	$\Delta_1 < 10\%$

the locus end point of every test cell to the approximating PDF(s), and (4) (u, v) coordinates of the locus end point for every test cell. This information has been tabulated for the selected examples of Sections 2.6.5.1–2.6.5.3.

The following strategy is used in step 6:

6.1. Only those kinds of PDFs among the Weibull, Lognormal and *K* distributions, denoted by W, L, and K, respectively, that pass step 4 to all of the test cells in the patch are considered as possible approximating kind of PDF(s) for the CL patch. A violation type of discrepancy occurs when none of the distributions pass step 4 for all of the test cells in the patch. This situation initiates a search for subpatches within the patch.

6.2. The ranking of each possible approximating type of PDF is summed over all of the test cells in the CL patch. That distribution having the lowest sum is chosen as the best approximation kind, for underlying CL patch distribution. For example, consider Table 2.23 which contains the distribution rankings for each test cell. The sum of rankings for W, L, and K are 16, 15, and 20, respectively. Consequently, L is chosen as the best approximating kind.

6.3. A check is made to determine whether, the patch under consideration is homogeneous. If it is, numerical values are determined for the best approximating kind of PDF. If it is not, a fault type of discrepancy

TABLE 2.52
Setting of PLQP in the Threshold Approximation and Threshold Fine-Tuning Stages, CL Patch 1 (Example 3)

PLQP (%)	CL Patch 1		Δ_i
	Threshold Approximation Stage		
10.00 (guess)	NCQ = 8 NCC = 1 PLCQP = 48.94 PLCCP = 8.51	NCQ = 7 NCC = 1 PLCQP = 6.38 PLCCP = 0.00	$\Delta_1 < 10\%$
20.00 (new guess)	NCQ = 8 NCC = 1 PLCQP = 82.98 PLCCP = 51.06	NCQ = 7 NCC = 1 PLCQP = 46.81 PLCCP = 10.64	$\Delta_1 > 10\%$
51.06 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 100.0 PLCCP = 100.0	NCQ = 7 NCC = 1 PLCQP = 95.74 PLCCP = 78.72	$\Delta_1 > 10\%$
Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 100.0 PLCCP = 100.0	NCQ = 7 NCC = 1 PLCQP = 100.0 PLCCP = 100.0	$\Delta_1 < 10\%$

occurs and a subpatch search is initiated. The check for homogeneity proceeds by determining whether all of the test cells in the patch fall within the .99 confidence contour of the approximating PDF for each test cell. However, in contrast to Section 2.6.4.3, the length of the minor axis is used as opposed to the half length. The lengths of the minor axes are tabulated in Table 2.55 for $N_R = 100$ and various shape parameters. The procedure is carried out as follows:

6.3.a. Select the first test cell in the CL patch.

6.3.b. Using the best approximating kind of PDF found in step 6.2, and the shape, parameter associated with the selected test cell, determine the length of the minor axis from Table 2.55.

6.3.c. Using the (u, v) coordinates for the locus end points, compute the distances between the (u, v) coordinates of the selected test cell and every other test cell.

6.3.d. Check the distances computed in 6.3.c. against the minor axis length obtained in 6.3.b. Those test cells within the minor axis length from the selected test cell pass and are indicated by a T. The remaining cells are indicated by an Ψ.

TABLE 2.53
Setting of PLQP in the Threshold Approximation and Threshold Fine-Tuning Stages, CL Patch 2 (Example 3)

PLQP (%)	CL Patch 2		Δ_i
	Threshold Approximation Stage		
10.00 (guess)	NCQ = 8 NCC = 1 PLCQP = 42.93 PLCCP = 11.62	NCQ = 7 NCC = 1 PLCQP = 16.67 PLCCP = 1.51	$\Delta_1 > 10\%$
11.62 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 53.54 PLCCP = 13.64	NCQ = 7 NCC = 1 PLCQP = 19.19 PLCCP = 1.51	$\Delta_1 > 10\%$
13.64 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 64.14 PLCCP = 18.69	NCQ = 7 NCC = 1 PLCQP = 24.24 PLCCP = 2.52	$\Delta_1 > 10\%$
18.69 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 77.27 PLCCP = 32.32	NCQ = 7 NCC = 1 PLCQP = 44.95 PLCCP = 7.57	$\Delta_1 > 10\%$
32.32 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 94.44 PLCCP = 69.70	NCQ = 7 NCC = 1 PLCQP = 78.79 PLCCP = 37.37	$\Delta_1 > 10\%$
69.70 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 100.0 PLCCP = 100.0	NCQ = 7 NCC = 1 PLCQP = 100.0 PLCCP = 100.0	$\Delta_1 < 10\%$

6.3.e. Select the second test cell in the CL patch and repeat 6.3.b. through 6.3.d.

6.3.f. Continue this process until all test cells have been selected.

6.3.g. Record the results in a table referred to as the exclusion table. The CL patch is assumed to be homogeneous only if no Ψs appear in the exclusion table.

6.3.h. When only Ts appear in the exclusion table, the CL patch is considered to be homogeneous. An approximating PDF is obtained by averaging over all of the test cells the shape, scale,

TABLE 2.54
Setting of PLQP in the Threshold Approximation and Threshold Fine-Tuning Stages, CL Patch 3 (Example 3)

PLQP (%)	CL Patch 3		Δ_i
	Threshold Approximation Stage		
10.00 (guess)	NCQ = 8 NCC = 1 PLCQP = 46.53 PLCCP = 14.36	NCQ = 7 NCC = 1 PLCQP = 11.39 PLCCP = 1.98	$\Delta_1 > 10\%$
14.36 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 70.30 PLCCP = 26.24	NCQ = 7 NCC = 1 PLCQP = 31.68 PLCCP = 5.44	$\Delta_1 > 10\%$
26.24 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 92.08 PLCCP = 66.83	NCQ = 7 NCC = 1 PLCQP = 68.81 PLCCP = 24.75	$\Delta_1 > 10\%$
66.83 Latest PLCCP	NCQ = 8 NCC = 1 PLCQP = 100.0 PLCCP = 100.0	NCQ = 7 NCC = 1 PLCQP = 99.01 PLCCP = 93.07	$\Delta_1 < 10\%$
	Threshold Fine-Tuning Stage		
93.07 Latest PLCCP	NCQ = 5 NCC = 1 PLCQP = 100.0 PLCCP = 100.0		$\Delta_2 > 1\%$
96.53 Half way Rule	NCQ = 5 NCC = 1 PLCQP = 100.0 PLCCP = 100.0		$\Delta_2 > 1\%$
98.26 Half way Rule	NCQ = 5 NCC = 1 PLCQP = 100.0 PLCCP = 100.0		$\Delta_2 > 1\%$
99.13 Half way Rule	NCQ = 8 NCC = 1 PLCQP = 100.0 PLCCP = 100.0		$\Delta_2 < 1\%$

TABLE 2.55
Minor Axis Length for Different PDFs, $N_R = 100$

PDF	Shape Parameter	Length of the Minor Axis ($\times 10^{-1}$)
Rayleigh	—	0.94
Weibull	0.1	0.50
	0.5	0.82
	1.0	0.82
	2.0	0.94
	3.0	0.94
	4.0	0.90
	5.0	0.88
	10.0	0.88
Lognormal	0.01	0.96
	0.05	0.90
	0.1	0.86
	0.2	0.84
	0.3	0.84
	0.4	0.92
	1.0	0.82
	5.0	0.56
	10.0	0.26
K-distribution	0.01	0.26
	0.1	0.74
	1.0	0.58
	5.0	0.97
	10.0	0.94
	20.0	0.98
	40.0	0.98
	50.0	0.98

and location parameters of the best approximating kind of PDF obtained in 6.2.

6.3.i. When Ψs appear in the exclusion table, a fault type of discrepancy occurs. This is discussed in a later section dealing with discrepancies.

Steps 6.3.a.–6.3.h. are now illustrated using the data in Table 2.23. As noted in step 6.2., Lognormal is the best PDF approximating kind for the CL patch under consideration. Therefore, using Table 2.55, the lengths of the minor axes corresponding to the shape parameter values of the Lognormal PDF found in Table 2.23 are listed in Table 2.56 for the different test cells. The second column and first row of Table 2.56 indicate the assigned numbers of the test cells.

TABLE 2.56
Distance Between All Pairs of Test Cells for the Example of Table 2.19

Length from Table 2.55	Cell No.	1	361	722	1083	1444
0.084	1	0	0.007	0.050	0.020	0.025
0.084	361	0.007	0	0.053	0.027	0.027
0.084	722	0.050	0.053	0	0.040	0.026
0.084	1083	0.020	0.027	0.040	0	0.023
0.090	1444	0.025	0.027	0.026	0.023	0

TABLE 2.57
Exclusion Table for the Example of Table 2.19

Cell No.	1	361	722	1083	722
1	T	T	T	T	T
361	T	T	T	T	T
722	T	T	T	T	T
1083	T	T	T	T	T
1444	T	T	T	T	T

The remaining entries in the table represent the distance between the locus end points of the tabulated pairs of test cells.

The exclusion table corresponding to Table 2.56 is shown in Table 2.57 and consists Ts and Ψs depending on whether or not the distances in every row of Table 2.56 are smaller of larger than the length shown in the first column of the same row. For this example, note that the entries are all Ts indicating that all test cells are within the length of the minor axes for the appropriate lognormal distribution. It is concluded that the patch under investigation is homogeneous. This example is completed in Section 2.7.5.1.

2.7.4.4.2. SPA and SPA Instance

Steps 1–6 of the PDF approximation stage involve the choice of N_T test cells, the selection of N_R reference cells for each test cell, use of the Ozturk algorithm to find the best approximating PDF(s) for each test cell, and determination of an approximating PDF when the CL patch is homogeneous or implementation of a subpatch search procedure when the CL patch is nonhomogeneous. These steps, which are diagrammed in Figure 2.97, constitute the SPA. A particular setting of the parameters N_T, and N_R constitutes an SPA instance.

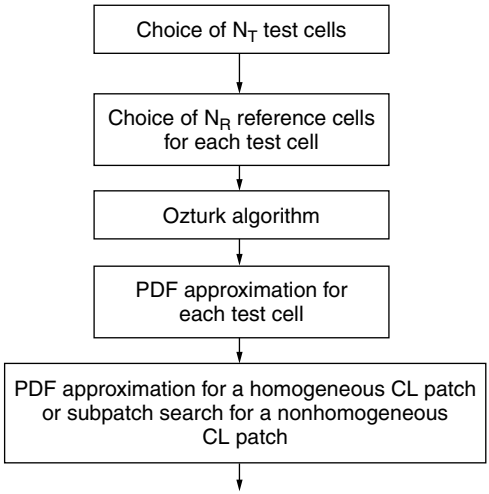

FIGURE 2.97 PDF approximation stage.

2.7.4.4.3. Resolution of Discrepancies

In this section, rules are developed to enable resolution of discrepancies that may occur during the PDF approximation stage. Discrepancies may arise during two different phases of this stage: (1) during the PDF approximation for each test cell and (2) during PDF approximation for a homogeneous CL patch or subpatch search for a nonhomogeneous CL patch. These phases are viewed as two substages and are investigated next.

(1) PDF *approximation for each test cell*: in order to approximate the PDF(s) of a test cell, steps 1–5 of the PDF approximation stage are implemented, as presented in Section 2.6.4.5. As shown in the examples of Section 6.5, the outcome of the Ozturk algorithm phase is rejected whenever none of the allowable distributions (i.e., W, L, K) is found to be a suitable approximation. This decision is based on the fact that the data are known to be generated from either W, L, or K. When the data cannot be approximated by one of the allowable PDFs, a violation type of discrepancy is detected. At this point, the diagnosis process hypothesizes that the presence of outliers may be the source for the distortion causing the discrepancy. The first row of Table 2.58 summarizes resolution of the discrepancy that may occur in the PDF approximation of a test cell.

Even though the data are known to be generated from the W, L, and *K* distributions, these may not be representative. As a result, the allowable distributions may not provide suitable approximations even after removal of the outliers. Because the data have been identified to be belonging to WSC, the Gaussian receiver is likely to produce false alarms. Consequently, the discrepancy is resolved by discarding the data. The second row of Table 2.58 summarizes resolution of this discrepancy.

TABLE 2.58
Discrepancies in the PDF Approximation of a Test Cell Stage

Discrepancy (Violation Type)	Diagnosis (Source of Distortion)	Reprocessing
None of the allowable distributions is found to be a suitable approximation for the test cell	Presence of outliers	Remove outliers
After removal of outliers none of the allowable distributions is found to be a suitable approximation for the test cell	Data is not representative	Discard data

(2) PDF *approximation for a* CL *patch*: in order to approximate the PDF of a CL patch, step 6 of the PDF approximation stage is implemented. As discussed in Section 2.7.4.4.1, the decision about existence of a suitable approximating PDF for a CL patch is based on the outcome of the exclusion table. As stated previously, a discrepancy is detected when one or more Ψs occur in the exclusion table.

When this happens, two cases are possible:

(2a) The CL patch is homogeneous. In this case, the discrepancy is of the conflict type because the expectation that only Ts should be encountered in the exclusion table has not materialized. However, the full length minor axis criterion has yet to be applied to the PDF which would be obtained by averaging over the parameters of all test cells for the best approximating PDF kind. This is necessary because an Ψ which appears in the exclusion table due to two test cells having widely separated (u, v) coordinates may disappear when the (u, v) coordinates of each test cell are compared to those of the average PDF. Should this fail, an exclusion table is generated for the next best ranked PDF kind. If only Ts appear in the exclusion table, the parameters of this PDF kind are averaged over all the test cells and the average PDF is used to approximate the homogeneous CL patch. If one or more Ψs appear in the exclusion table, phase (2a) is repeated until all possible PDF kinds have been exhausted. Should all possibilities fail, it is assumed that the CL patch is not homogeneous and the process initiates phase (2b).

(2b) The CL patch is not homogeneous. In this case, the discrepancy is of the violation type because none of the allowable PDF kinds is able to model the entire CL patch. Therefore, it should be modeled by two or more approximating PDFs. By using the exclusion table to identify groupings and by introducing new test cells between the groupings to generate even larger exclusion tables, the system attempts to determine a suitable number of PDFs for modeling the nonhomogeneous CL patch. This is achieved by examining the patterns of Ψs and Ts which appear, as discussed below.

TABLE 2.59
Exclusion Table for Cells a, b, c, d, and e

Cell No.	a	b	c	d	e
a	T	Ψ	T	Ψ	Ψ
b	Ψ	T	Ψ	T	T
c	T	Ψ	T	Ψ	Ψ
d	Ψ	T	Ψ	T	T
e	Ψ	T	Ψ	T	T

For example, five test cells numbered a, b, c, d, and e might yield the exclusion table shown in Table 2.59, By interchanging rows b and c and columns b and c, Table 2.60 results. From this table, there are seen to be two groupings: cells a, c and cells b, d, e. To determine whether the CL patch may consist of more than two groupings, new test cells, numbered f, g, h, which are located between the groupings are introduced. After rearranging rows and columns, the exclusion table given by Table 2.61 may result. By inspection, it is concluded that cells f, g, h are not common to the previous two groupings. It is concluded that the CL patch is likely to be composed of at least three groupings. On the other hand, the introduction of cells f, g, h might result in the exclusion table shown in Table 2.62. Here, cells f, g, and h are seen to be common to both of the original groupings. Cells f, g, and h can be classified as "border" test cells (i.e., cells having reference cells extending into more than one homogeneous subregion of the CL patch). It is concluded that two subregions are likely to exist within the CL patch. Further refinements can be made by introducing additional new test cells and studying the resulting exclusion tables.

In general, additional groupings are identified when new test cells cannot be associated with existing groupings. Otherwise, new test cells will either be associated with existing groupings or be classified as border cells. When all cells of the CL patch have been used as either test cells or reference cells and when

TABLE 2.60
Exclusion Table when Cells b and c are Interchanged

Cell No.	a	b	c	d	e
a	T	T	Ψ	Ψ	Ψ
c	T	T	Ψ	Ψ	Ψ
b	Ψ	Ψ	T	T	T
d	Ψ	Ψ	T	T	T
e	Ψ	Ψ	T	T	T

TABLE 2.61
Exclusion Table with Additional Cells f, g, and h Not Common to Other Cells

Cell No.	a	c	f	g	h	b	d	e
a	T	T	Ψ	Ψ	Ψ	Ψ	Ψ	Ψ
c	T	T	Ψ	Ψ	Ψ	Ψ	Ψ	Ψ
f	Ψ	Ψ	T	T	T	Ψ	Ψ	Ψ
g	Ψ	Ψ	T	T	T	Ψ	Ψ	Ψ
h	Ψ	Ψ	T	T	T	Ψ	Ψ	Ψ
b	Ψ	Ψ	Ψ	Ψ	Ψ	T	T	T
d	Ψ	Ψ	Ψ	Ψ	Ψ	T	T	T
e	Ψ	Ψ	Ψ	Ψ	Ψ	T	T	T

new test cells result only in border cells, the refinement procedure stops and the system assumes that additional groupings do not exist.

(2c) The next step in the procedure is to identify those cells in the various homogeneous subregions of the CL patch. In general, each test cell in a grouping has associated with it N_R reference cells. It is assumed that all N_R reference cells belong to the same subregion as the corresponding test cell. All the cells identified for a particular subregion are assigned the same number. For example, all the cells in the kth subregion are assigned the number k. At this point each cell in the CL patch will have one or more numbers assigned to it. Those cells with a single assigned number are assumed to belong to the numbered subregion. Those cells with more than one assigned number are assumed to belong to a border region bordering the numbered subregions. For example, a cell with the assigned numbers j and k is assumed to be in a border region bordering the jth and kth subregions. Similarly, a cell with the assigned numbers j, k, and m is assumed to be in a border region bordering the jth, kth, and mth subregions.

TABLE 2.62
Exclusion Table with Additional Cells f, g, and h Common to Other Cells

Cell No.	a	c	f	g	h	b	d	e
a	T	T	T	T	T	Ψ	Ψ	Ψ
c	T	T	T	T	T	Ψ	Ψ	Ψ
f	T	T	T	T	T	T	T	T
g	T	T	T	T	T	T	T	T
h	T	T	T	T	T	T	T	T
b	Ψ	Ψ	T	T	T	T	T	T
d	Ψ	Ψ	T	T	T	T	T	T
e	Ψ	Ψ	T	T	T	T	T	T

TABLE 2.63
Discrepancies in the PDF Approximation of a CL Patch

Discrepancy	Diagnosis (Source of Distortion)	Reprocessing
Not all test cells can be approximated by the best average PDF (conflict type)	CL patch may be homogeneous but cannot be approximated by the best average PDF	Try the next best average PDF
Not all test cells can be approximated by any of the average PDF(s) (violation type)	CL patch is not homogeneous	Separate subpatches and approximate their corresponding PDF(s)

(2d) Once the subregions have been identified, the PDF associated with each subregion is approximated. This is accomplished by following the procedure previously outlined in substeps 6.1–6.3 of step 6 for the PDF approximation strategy. Depending upon the application, the data from the border regions can either be discarded or approximated by one of the PDFs from the bordering subregions.

In summary, two types of discrepancies can arise during the PDF approximation of a CL patch, as shown in Table 2.63. When the CL patch may be homogeneous but cannot be approximated by the best average PDF, a conflict type of discrepancy is noted and the diagnosis suggests that the next best average PDF be tried. On the other hand, when it is concluded from the exclusion table not all cells in a CL patch can be approximated by the same PDF, a violation type of discrepancy is detected indicating that the CL patch is not homogeneous. The diagnosis recommends that the CL patch data should be reprocessed to search for subpatches within the CL patch. Once the subregions have been identified, their PDFs are approximated.

2.7.5. Examples of Indexing

The IPUS concepts presented in Section 2.7.4 are now illustrated by continuing the examples treated previously in Sections 2.6.5.1–2.6.5.3. In this section, the possible kinds of approximating PDFs are ranked as explained in Section 2.7.4.4. For each approximating kind of PDF, the corresponding exclusion table is then built. Conclusions from examination of the exclusion tables are indicated with either an "All pass" or a "Not all pass" label. A not all pass label indicates that the corresponding exclusion table includes at least one Ψ and that the full length minor axis criterion fails with the average PDF for one or more test cells. On the other hand, an All pass label indicates that the corresponding exclusion table does not include any Ψs. Hence, all test cells may be approximated by the corresponding kind of PDF. For the kinds of PDFs that can approximate

the distribution of the CL patch the scale, shape, and location parameters are averaged over all of the test cells in the CL patch to obtain an average PDF. PDFs that cannot approximate a particular CL patch either because of failure to pass the half length criterion or a not all pass label are labeled N/A to indicate not applicable.

In order to evaluate the effectiveness of the proposed approach for the partitioning of the surveillance volume, results obtained for the number of CL patches, number of cells in each CL patch, and estimated CNR for each CL patch are compared to the generated ones. On the other hand, because the CL patch cells contain additive BN data, the PDF underlying a CL patch is not the PDF of the generated CL data alone. In general, analytical expressions for the PDF of the sum of BN and CL data are difficult to obtain. However, by using the Ozturk algorithm on 1000 BN plus CL data points, a very accurate approximation is obtained for the underlying PDF of the CL patch. Examples are now presented to illustrate control by IPUS of the PDF approximation stage.

2.7.5.1. Example 1

Consider Tables 2.21–2.24 which summarize the results of steps one through five of the PDF approximation strategy for the example of Section 2.6.5.1. Note that CL patches one, two, and three can be approximated by any one of the three PDF kinds, whereas, as indicated in Table 2.24, CL patch four can be approximated only by the Weibull PDF. Tables 2.64–2.67 summarize the results for step 6 of the PDF approximation strategy for the four CL patches identified thus far. An exclusion table was generated for the allowable PDF kinds in each CL patch. As shown in Tables 2.64–2.67, all of the PDF kinds are labeled All pass indicating that none of the exclusion tables contained an Ψ mark. Consequently, average values for the parameters of all the allowable PDF kinds are also tabulated in Tables 2.64–2.67.

Table 2.68 uses the results presented in Tables 2.20 and 2.64–2.67 to list the best average PDF for each CL patch. Comparing the results in Tables 2.14 and 2.68, note that (1) regions for all four of the CL patches have been estimated, (2) the sample CNRs for CL patches A, C, and D, as evaluated according to Section 2.6.2.2, are within 1 dB of the generated values while that for CL patch B is within 3.31 dB, (3) the number of cells included in CL patches A, B, C, and D

TABLE 2.64
Average PDFs for CL Patch 1 (Example 1)

Average PDF	Rank	Exclusion Table	Shape	Scale	Location
W	1	All pass	2.05	3.14	0.64
L	3	All pass	0.19	7.27	−4.00
K	2	All pass	50.0	0.43	0.73

TABLE 2.65
Average PDFs for CL Patch 2 (Example 1)

Average PDF	Rank	Exclusion Table	Shape	Scale	Location
W	2	All pass	2.56	1.29	31.38
L	1	All pass	0.13	4.32	28.17
K	3	All pass	50.0	0.15	31.62

TABLE 2.66
Average PDFs for CL Patch 3 (Example 1)

Average PDF	Rank	Exclusion Table	Shape	Scale	Location
W	2	All pass	1.96	15.52	−0.04
L	1	All pass	0.25	20.09	−30.09
K	3	All pass	33.28	2.96	0.08

TABLE 2.67
Average PDFs for CL Patch 4 (Example 1)

Average PDF	Rank	Exclusion Table	Shape	Scale	Location
W	1	All pass	4.51	90.82	14.97
L			N/A		
K			N/A		

TABLE 2.68
Assessment Parameters of the Mapping and Indexing Stages (Example 1)

CL Patch (No.)	CNR (dB)	Number of Cells	PDF (Best)	Shape Parameter	Scale Parameter	Location Parameter
A (1)	9.04	113	W	2.05	3.14	0.64
B (3)	23.31	1444	L	0.25	20.09	−30.09
C (2)	30.63	151	L	0.13	4.32	28.17
D (4)	40.56	146	W	4.51	90.82	14.97

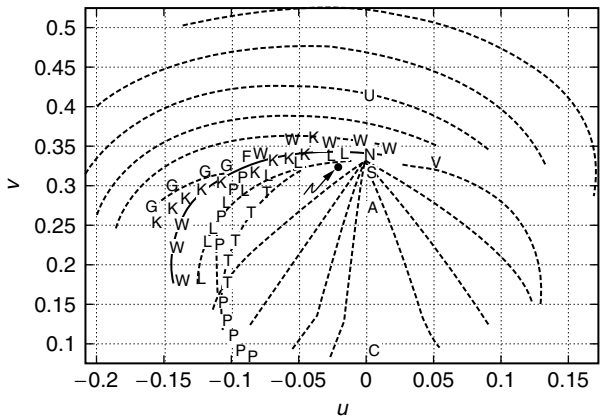

FIGURE 2.98 Location of the locus end point for CL patch A (Example 1).

are close to the number of cells generated. In fact, a separate evaluation indicates that 96.5, 98.3, 97.2, and 99.3% of the generated cells have been correctly classified for CL patches A, B, C, and D, respectively.

With regard to the goodness of the PDF approximations, Figures 2.98, 2.100, 2.101, and 2.103 show the location in the approximation chart, indicated by o, for the accurate approximation of the underlying CL patch PDF (i.e., PDF corresponding to the sum of CL and BN in the patch) (Figure 2.104). These were obtained by generating 1000 points for each CL patch. The results are summarized in Table 2.69. Figures 2.99, 2.101, 2.103, and 2.105 show the PDFs from Table 2.68 (solid lines) obtained by the mapping and indexing stages, superimposed on the corresponding PDFs from Table 2.69 (dashed lines).

The comparison for CL patch A is given in Figure 2.99. Although the best approximating PDF is Weibull while the underlying PDF is Lognormal, the curves are reasonably close except for a small offset in their location. Figure 2.101 shows the comparison for CL patch B. Here the best approximating and underlying PDFs are Lognormal with shape parameters 0.25 and 0.27, respectively. However, the scale and location parameters are not as closely

TABLE 2.69
Accurate Approximation of Underlying CL Patch PDFs (Example 1)

CL Patch	PDF	Distance	Shape	Scale	Location	(*u*, *v*)
A	L	0.61×10^{-2}	0.16	9.33	−5.65	(−0.0213, 0.3225)
B	L	0.35×10^{-2}	0.27	26.25	−13.43	(−0.0416, 0.3282)
C	W	0.22×10^{-2}	2.17	0.87	31.56	(−0.0296, 0.3382)
D	SU–Johnson	0.79×10^{-2}	(40.0, 40.0)	54.19	164.51	(0.0223, 0.3379)

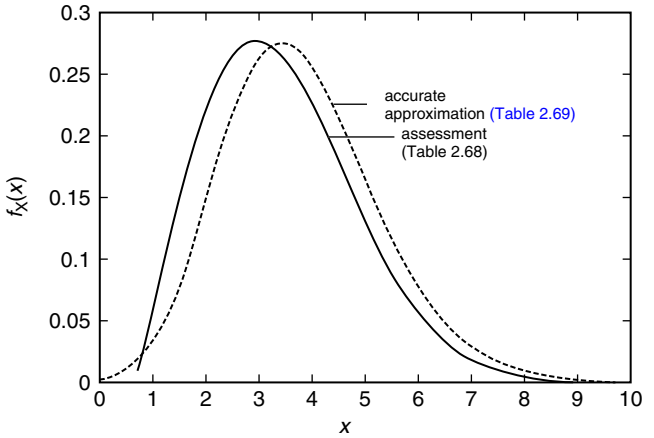

FIGURE 2.99 PDF comparison for CL patch A (Example 1).

matched. This accounts for the discrepancy between the two curves. The results for CL patch C are shown in Figure 2.103. As with CL patch A, the best approximating PDF is Weibull while the underlying PDF is Lognormal. In this case, the two curves differ substantially and a poor approximation has been made. Finally, the comparison for CL patch D is shown in Figure 2.105. Although the best approximating PDF is Weibull while the underlying PDF is SU-Johnson, note that the curves are reasonably close except for a small offset in their height and location.

The best approximating PDFs from Table 2.68 are superimposed on the histograms for CL patches A, B, C, and D in Figures 2.106–2.109, respectively. Whether or not the best approximating PDFs are close to the underlying PDFs, note that the best approximating PDFs nicely overlay the corresponding histograms.

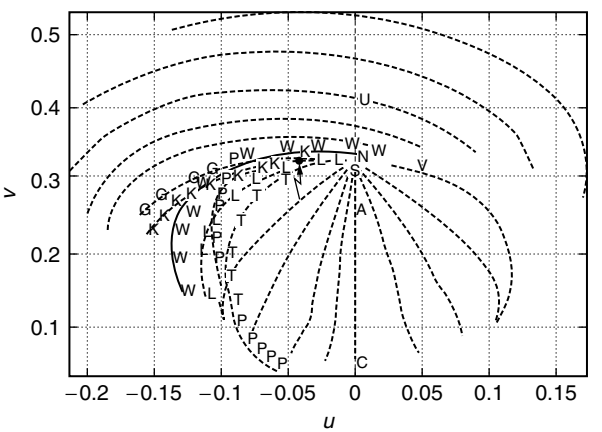

FIGURE 2.100 Location of the locus end point for CL patch B (Example 1).

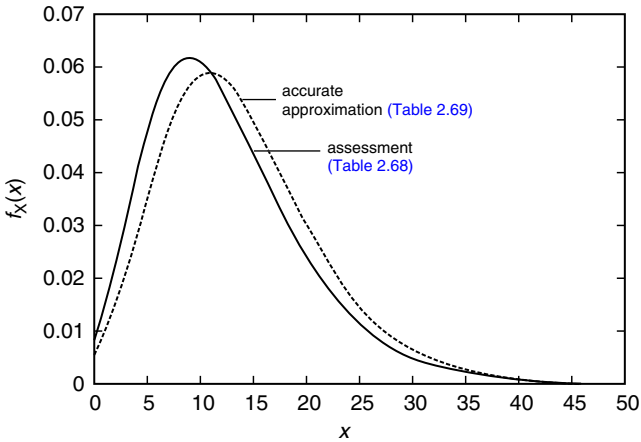

FIGURE 2.101 PDF comparison for CL patch B (Example 1).

Relative to CL patch C, it is seen that the large amount of data with values above 33 are not representative of the underlying PDF shown in Figure 2.102. This may be due to the inclusion in CL patch C of cells originating in CL patch D.

2.7.5.2. Example 2

Consider now Tables 2.29–2.31 which summarize the results of steps 1–5 of the PDF approximation strategy for the example of Section 2.6.5.2. Note that CL patches one and three can be approximated by any one of the three PDF kinds, whereas, as indicated in Table 2.30, CL patch two can be approximated only by the Lognormal PDF. Tables 2.70–2.72 summarize the results for step 6 of the PDF approximation strategy for the three CL patches identified thus far.

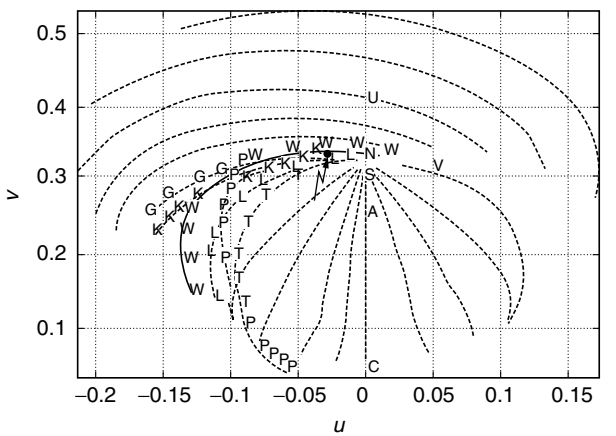

FIGURE 2.102 Location of the locus end point for CL patch C (Example 1).

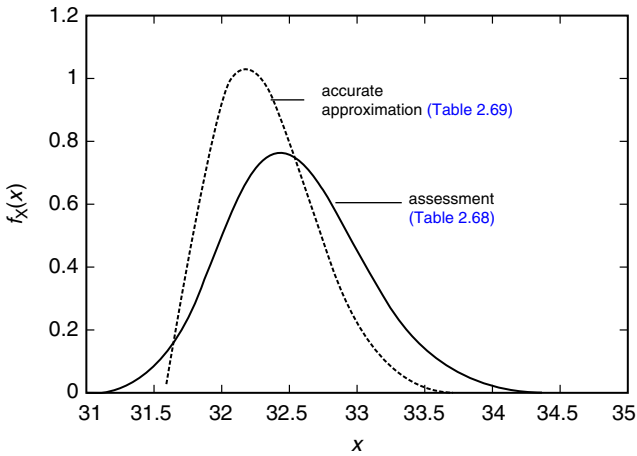

FIGURE 2.103 PDF comparison for CL patch C (Example 1).

An exclusion table was generated for the allowable PDF kinds in each CL patch. As shown in Tables 2.71 and 2.72, the Lognormal and K PDF kinds are labeled not all pass for CL patches 2 and 3, respectively, indicating that the corresponding exclusion tables contained one or more Ψs and that the full length minor axis criterion has failed when applied to the average PDFs obtained by averaging over the parameters of all the test cells for the corresponding PDF kinds. Consequently, a N/A label is posted for each of the Lognormal and K PDF kinds in Tables 2.71 and 2.72, respectively. Average values for the parameters of all the allowable PDF kinds labeled with an all pass are also tabulated in Tables 2.70–2.72.

The exclusion table, corresponding to the lognormal entries in Table 2.30 for CL patch 2, is shown in Table 2.73. The patterns of Ψs and Ts in the table clearly

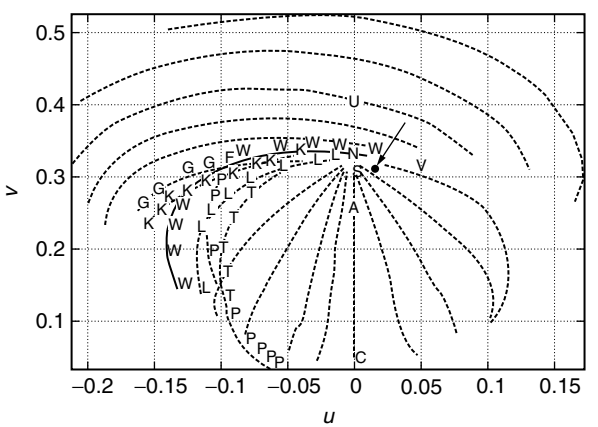

FIGURE 2.104 Location of the locus end point for CL patch D (Example 1).

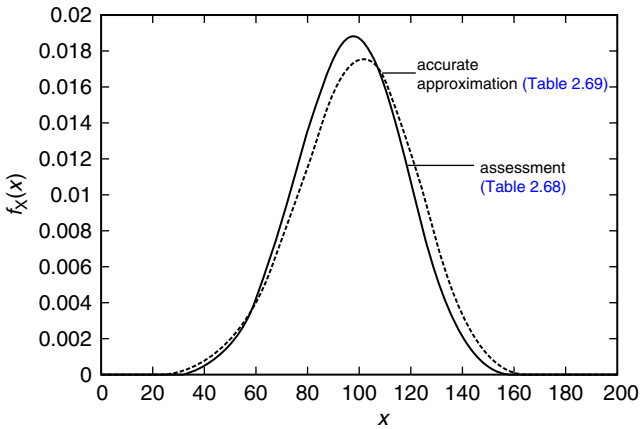

FIGURE 2.105 PDF comparison for CL patch D (Example 1).

defines two different regions. One region consists of test cells numbered one and 77. The other region consists of the remaining test cells numbered 154, 231, and 308. This suggests the presence of at least two homogeneous subpatches in CL patch two. In order to investigate whether more than two subpatches exist, three more test cells are considered which are located in between the two cited regions. The additional test cells are numbered 96, 115, and 134. Steps 1–5 of the PDF approximation strategy for these cells are summarized in Table 2.74. The new exclusion table, shown in Table 2.75 suggests that only two subpatches exist in the CL patch. The overlapping that exists between the two regions is due to the test cells located near the edges separating the two subpatches and whose reference cells extend to both regions. The conclusion that CL patch two consists of two subpatches is recognized by the IPUS program as a discrepancy.

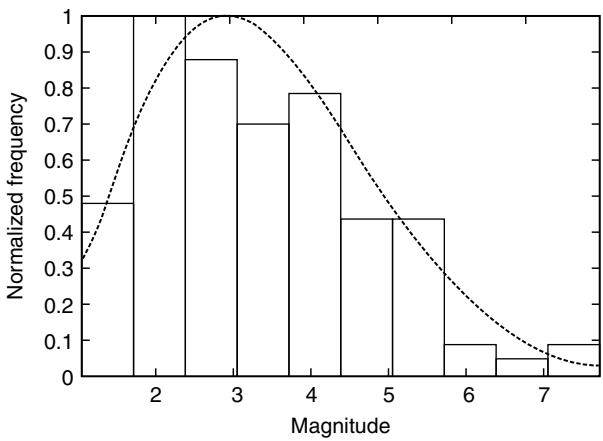

FIGURE 2.106 Histogram and best approximating PDF for CL patch A (Example 1).

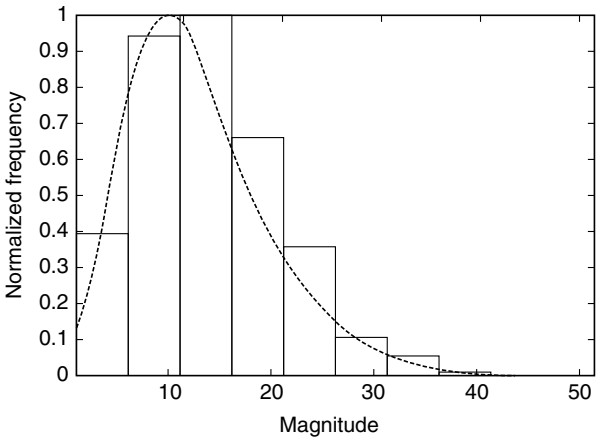

FIGURE 2.107 Histogram and best approximating PDF for CL patch B (Example 1).

Having detected the discrepancy, the diagnosis process in the IPUS control initiates a search for the subpatches. Specifically, (1) test cells numbered one and 77 have their reference cells numbered one, while test cells numbered 154, 231, 308 have their reference cells numbered two, (2) all reference cells numbered with one and two at the same time are declared as cells close to the boundary separating the CL subpatches. Their reference cells defines the boundary region where a test cell in one region has some of its reference cells extending to the other region. When this step is completed, the subpatches and the boundary region are defined as shown in Figure 2.110. Note that the CL patch now consists of three different regions. The lower portion of the CL patch includes part of

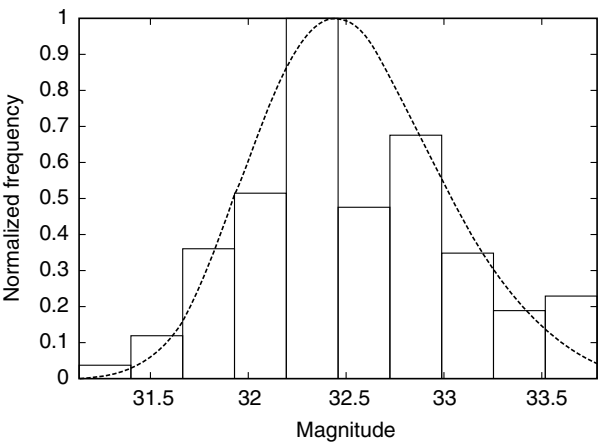

FIGURE 2.108 Histogram and best approximating PDF for CL patch C (Example 1).

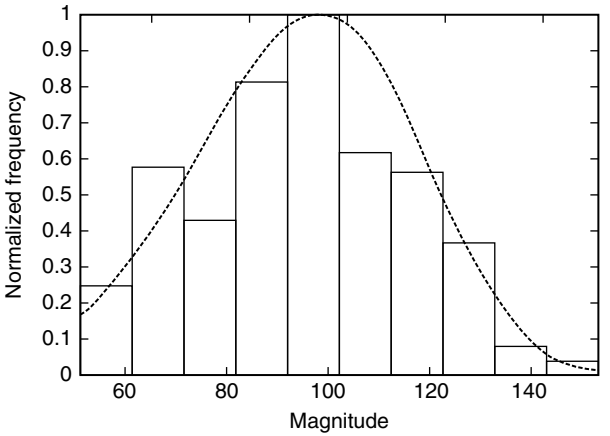

FIGURE 2.109 Histogram and best approximating PDF for CL patch D (Example 1).

TABLE 2.70
Average PDFs for CL Patch 1 (Example 2)

Average PDF	Rank	Exclusion Table	Shape	Scale	Location
W	3	All pass	1.34	4.22	0.46
L	1	All pass	0.43	5.94	−2.20
K	1	All pass	8.11	2.06	−0.82

TABLE 2.71
Average PDFs for CL Patch 2 (Example 2)

Average PDF	Rank	Exclusion Table	Shape	Scale	Location
W	N/A	N/A	N/A	N/A	N/A
L	1	Not all pass	N/A	N/A	N/A
K	N/A	N/A	N/A	N/A	N/A

TABLE 2.72
Average PDFs for CL Patch 3 (Example 2)

Average PDF	Rank	Exclusion Table	Shape	Scale	Location
W	2	All pass	1.03	2.49	0.90
L	3	All pass	0.61	3.15	−0.42
K	1	Not all pass	N/A	N/A	N/A

TABLE 2.73
Exclusion Table for the Example of Table 2.30

Cell No.	1	77	154	231	308
1	T	T	Ψ	Ψ	Ψ
77	T	T	Ψ	Ψ	Ψ
154	Ψ	Ψ	T	T	T
231	Ψ	Ψ	T	T	T
308	Ψ	Ψ	T	T	T

TABLE 2.74
PDF Approximation for Test cells No. 96, 115, and 134, CL Patch 2 (Example 2)

Test Cell No.	PDF	Rank	Distance	Shape	Scale	Location	(*u*, *v*)
96	W	2	0.57×10^{-2} T	0.91	1.56	1.29	(−0.0992, 0.3091)
	L	6	0.21×10^{-1} T	0.68	1.87	0.52	
	K	1	0.41×10^{-2} T	1.75	2.15	0.42	
115	W	3	0.16×10^{-1} T	1.00	1.83	1.25	(−0.1045, 0.3327)
	L	7	0.40×10^{-1} T	0.60	2.34	0.24	
	K	5	0.27×10^{-1} T	2.17	2.10	0.37	
134	W	2	0.16×10^{-2} T	1.04	1.97	1.28	(−0.0888, 0.3224)
	L	5	0.23×10^{-1} T	0.59	2.45	0.27	
	K	4	0.10×10^{-1} T	3.08	1.86	0.31	

TABLE 2.75
Exclusion Table for the Example of Tables 2.30 and 2.74 (Example 2)

Cell No.	1	77	96	115	134	154	231	308
1	T	T	T	T	T	Ψ	Ψ	Ψ
77	T	T	T	T	T	Ψ	Ψ	Ψ
96	T	T	T	T	T	T	T	T
115	T	T	T	T	T	T	T	T
134	T	T	T	T	T	T	T	T
154	Ψ	Ψ	T	T	T	T	T	T
231	Ψ	Ψ	T	T	T	T	T	T
308	Ψ	Ψ	T	T	T	T	T	T

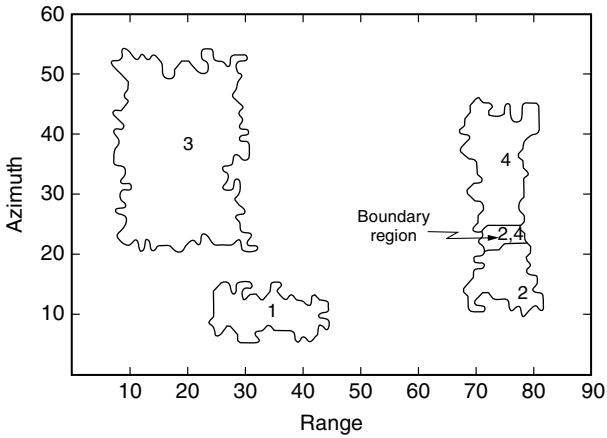

FIGURE 2.110 Boundaries of the scene resulting from step 6 of the PDF approximation stage (Example 2).

generated CL patch C while the upper portion includes part of generated CL patch D. The in-between region contains cells from CL patches C and D. The subpatches are relabeled two and four, as shown in Figure 2.110. It is found that CL patch two contains 99 cells, CL patch four contains 183 cells, and the boundary region contains 26 cells. Also, based on the 99 cells of CL patch two and the 183 cells of CL patch four, the CNRs for CL patches two and four are equal to 10.10 and 8.47 dB, respectively.

The PDF(s) of the CL subpatch regions are next approximated by considering the average PDF(s) in test cells 1 and 77 for CL patch two and in test calls 154, 231, and 308 for CL patch four. In absence of additional information, the test cells located in the boundary region are approximated either by the PDF of CL patch two or the PDF of CL patch four. Tables 2.76 and 2.77 summarize the results for step six of the PDF approximation for the subpatches numbered two and four.

TABLE 2.76
Average PDFs for CL Patch 2 Using Test Cells Numbered 1 and 77 (Example 2)

Average PDF	Rank	Exclusion Table	Shape	Scale	Location
W	N/A	N/A	N/A	N/A	N/A
L	1	All pass	0.84	1.27	0.96
K	N/A	N/A	N/A	N/A	N/A

TABLE 2.77
Average PDFs for CL Patch 4 Using Cells Numbered 154, 231, and 308 (Example 2)

Average PDF	Rank	Exclusion Table	Shape	Scale	Location
W	1	All pass	1.80	3.54	0.47
L	3	All pass	0.27	7.24	−3.85
K	2	All pass	33.97	0.76	0.26

Table 2.78 uses the results presented in Tables 2.25, 2.70, 2.72, 2.76, and 2.77 to list the best average PDF for each CL patch. Note that the L and *K* distributions are listed as the best approximating PDFs for CL patch A because their rankings were identical. Comparing the results in Tables 2.5 and 2.78 note that (1) regions for all four of the CL patches have been estimated, (2) the sample CNRs for CL patches B, C, and D, as evaluated according to Section 2.6.2.2, are within 1.5 dB of the generated values while that for CL patch A is within 2.03 dB, (3) the number of cells included in CL patches A and B are close to the number of cells generated. In fact, a separate evaluation indicates that 98.7%, and 92.5% of the generated cells have been correctly classified for CL patches A, and B, respectively. Also, 63.3% and 99.1% of the generated cells have been correctly classified for CL patches C, and D, respectively.

With regard to the goodness of the PDF approximations, Figures 2.111, 2.113, 2.115, and 2.117 show the location in the approximation chart, indicated by o, for the accurate approximation of the underlying CL patch PDF (i.e., PDF corresponding to the sum of CL and BN in the patch). These were obtained by generating 1000 points for each CL patch. The results are summarized in Table 2.79. Figures 2.112, 2.114, 2.116, and 2.118 show the PDFs from Table 2.78 (solid lines) obtained by the mapping and indexing stages, super-imposed on the corresponding PDFs from Table 2.79 (dashed lines).

TABLE 2.78
Assessment Parameters of the Mapping and Indexing Stages (Example 2)

CL Patch (No.)	CNR (dB)	Number of Cells	PDF (Best)	Shape Parameter	Scale Parameter	Location Parameter
A (1)	12.03	133	L	0.43	5.94	−2.20
			K	8.11	2.06	−0.82
B (3)	8.65	622	W	1.03	2.49	0.90
C (2)	10.10	99	L	0.84	1.27	0.96
D (4)	8.47	183	W	1.80	3.54	0.47

TABLE 2.79
Accurate Approximation of Underlying CL Patch PDFs (Example 2)

CL Patch	PDF	Distance	Shape	Scale	Location	(u, v)
A	W	0.73×10^{-2}	1.31	3.98	1.12	$(-0.0733, 0.3364)$
B	L	0.30×10^{-1}	0.49	3.67	-1.13	$(-0.0840, 0.3297)$
C	SU-Johnson	0.11×10^{-3}	$(1.31, -0.7)$	0.69	2.94	$(-0.0555, 0.2590)$
D	W	0.33×10^{-2}	2.24	3.43	2.48	$(-0.0278, 0.3345)$

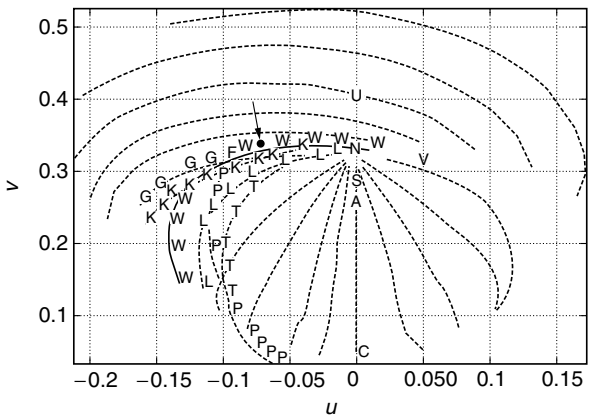

FIGURE 2.111 Location of the locus end point for CL patch A (Example 2).

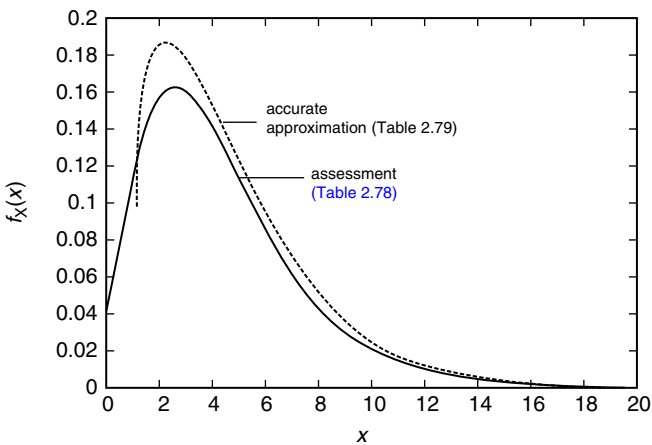

FIGURE 2.112 PDF comparison for CL patch A (Example 2).

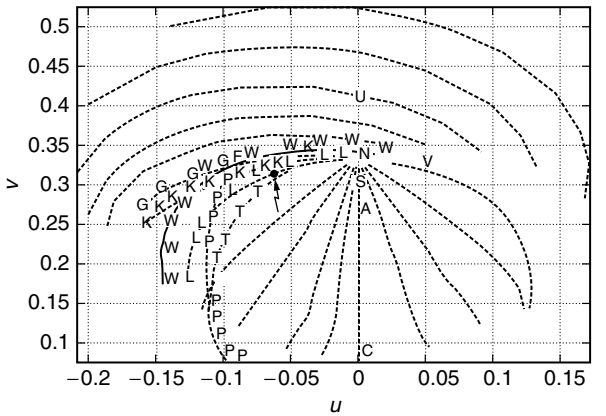

FIGURE 2.113 Location of the locus end point for CL patch B (Example 2).

The comparison for CL patch A is given in Figure 2.112. Although the best approximating PDF is Lognormal while the underlying PDF is Weibull, the curves are reasonably close except for a difference in their heights. Figure 2.114 shows the comparison for CL patch B. In contrast to CL patch A, the best approximating PDF is Weibull while the underlying PDF is Lognormal. Note that the curves are reasonably close. However, a small offset in location is used to compensate for a slight mismatch in shape. The results for CL patch C are shown in Figure 2.116. The best approximating PDF is Lognormal while the underlying PDF is SU-Johnson. In this case, the two curves differ substantially and a poor approximation has been made. Had the SU-Johnson been included in the library of allowable PDF kinds, a better approximating PDF would have been obtained. Finally, the comparison for CL patch D is shown in Figure 2.118. Here the best

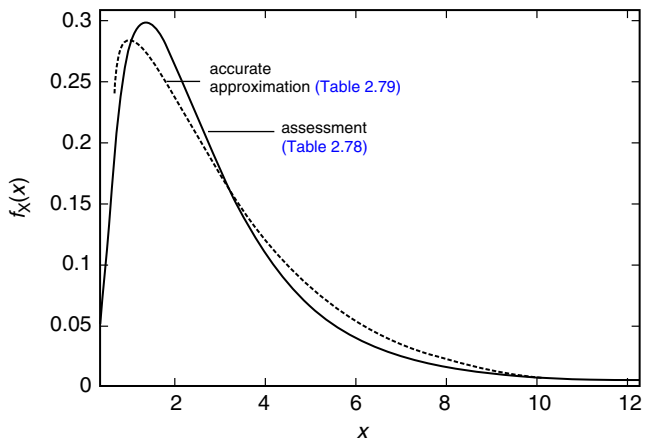

FIGURE 2.114 PDF comparison for CL patch B (Example 2).

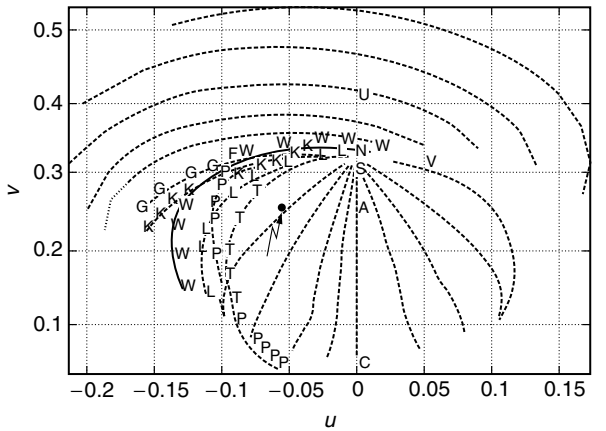

FIGURE 2.115 Location of the locus end point for CL patch C (Example 2).

approximating and underlying PDFs are Weibull with scale parameters 3.54 and 3.43, respectively. However, the shape and location parameters are not close. This accounts for the discrepancy between the two curves.

The best approximating PDFs from Table 2.78 are superimposed on the histograms for CL patches A, B, C, and D in Figures 2.119–2.122, respectively. Whether or not the best approximating PDFs are close to the underlying PDFs, note that the best approximating PDFs nicely overlay the corresponding histograms. Relative to CL patch D, it is seen that the large amount of data with values below 2.5 are not representative of the underlying PDF shown in Figure 2.118. This may be due to inclusion in CL patch D of cells originating in CL patch C, as seen in Figure 2.110.

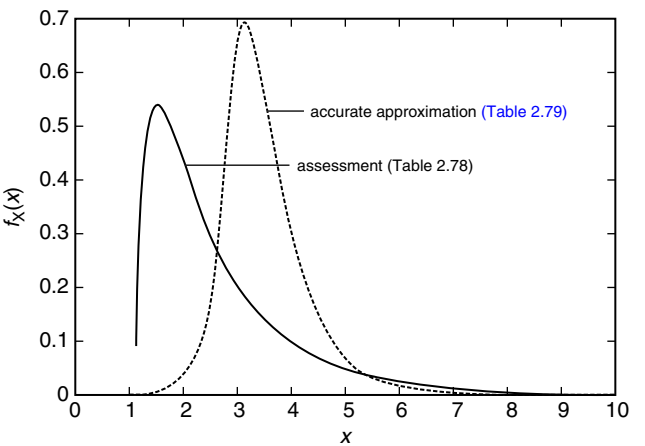

FIGURE 2.116 PDF comparison for CL patch C (Example 2).

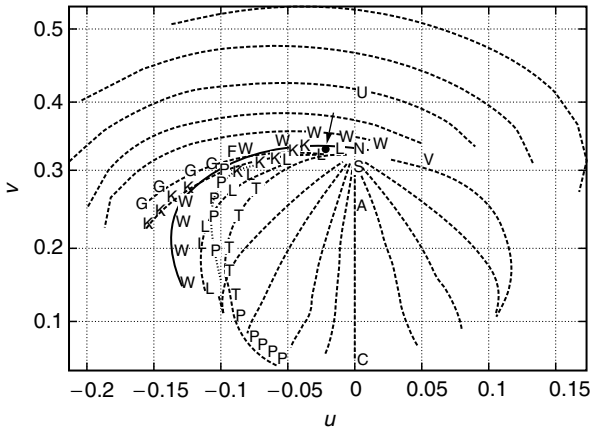

FIGURE 2.117 Location of the locus end point for CL patch D (Example 2).

2.7.5.3. Example 3

Finally, consider Tables 2.29–2.31 which summarize the results of steps 1–5 of the PDF approximation strategy for the example of Section 2.6.5.3. Note that CL patch 2 can be approximated by any one of the three PDF kinds, whereas, as indicated in Tables 2.29 and 2.31, CL patches one and three can be approximated only by the Weibull and K PDFs. Tables 2.80–2.82 summarize the results for step six of the PDF approximation strategy for the three CL patches identified thus far. An exclusion table was generated for the allowable PDF kinds in each CL patch. As Shown in Tables 2.80–2.82, all of the PDF kinds are labeled All pass indicating that none of the exclusion tables contained an Ψ mark. Consequently, average values for the parameters of all the allowable PDF kinds are also tabulated in Tables 2.80–2.82.

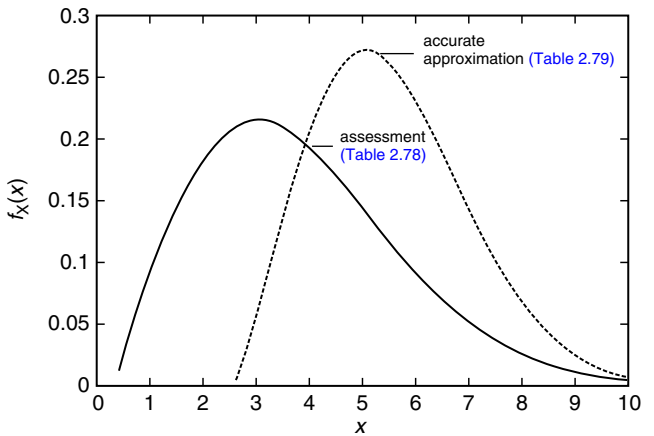

FIGURE 2.118 PDF comparison for CL patch D (Example 2).

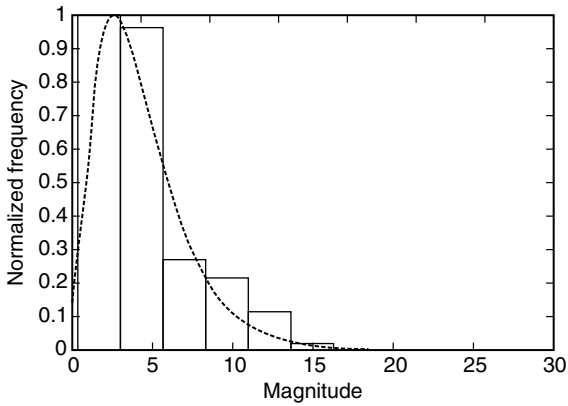

FIGURE 2.119 Histogram and best approximating PDF for CL patch A (Example 2).

Table 2.83 uses the results presented in Tables 2.32 and 2.80–2.82 to list the best average PDF for each CL patch. Comparing the results in Tables 2.8 and 2.83 note that (1) regions for CL patches A and B have been estimated, whereas the regions for the contiguous CL patches C and D have been approximated as a single region, (2) the sample CNRs for CL patch A, as evaluated according to Section 2.6.2.2, is within 1 dB of the generated values while that for CL patch B is within 3.47 dB, (3) the number of cells included in CL patches A and B are close to the number of cells generated. In fact, a separate evaluation indicates that 97.8% and 98.3% of the generated cells have been correctly classified for CL patches A and B, respectively.

For this particular example, the PDF approximation strategy is not able to recognize that more than one subpatch exists in CL patch 2. In order to evaluate the severity of this result, we analyze the PDFs of CL patches C and D. As shown

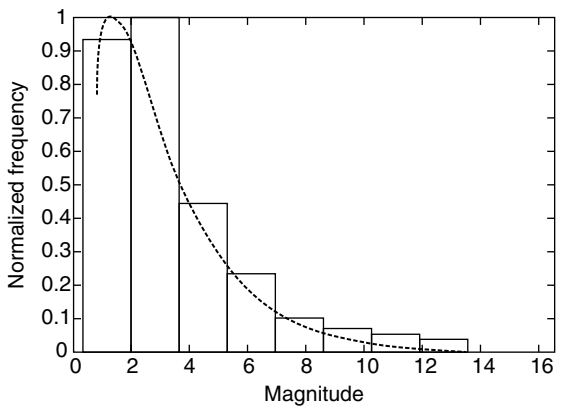

FIGURE 2.120 Histogram and best approximating PDF for CL patch B (Example 2).

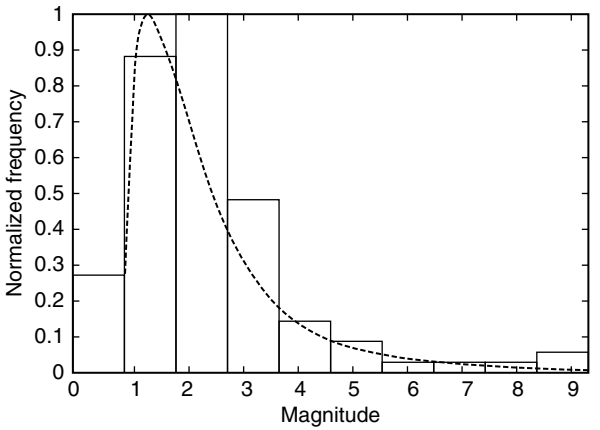

FIGURE 2.121 Histogram and best approximating PDF for CL patch C (Example 2).

in Figure 2.54(d,e), even though the data amplitude histograms of CL patches C and D are different (CL patch C data amplitudes extend only from 3.1 to 3.25 whereas CL patch D data amplitudes extend from 2 to 3.75), they have the same maximum located around 3.15. When BN is added to CL patches C and D, their histograms become as shown in Figure 2.123(a,b), respectively. When BN data is added, the data amplitudes of CL patches C and D extend over longer intervals ranging from 3.1 to 5.4. Note, also, that the shapes of both histograms are very similar. Thus, when BN data is added CL patches C and D can be approximated by the same PDF. It is, therefore, concluded for this example that it is reasonable to approximate CL patches C and D as a single CL patch (numbered 2).

With regard to the goodness of the PDF approximations, Figures 2.124, 2.126, 2.128, and 2.130 show the location in the approximation chart, indicated

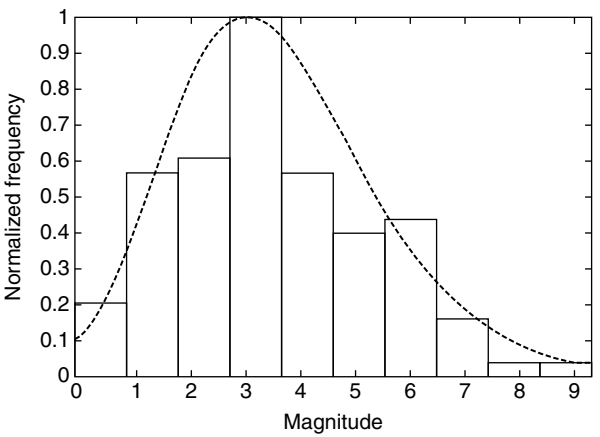

FIGURE 2.122 Histogram and best approximating PDF for CL patch D (Example 2).

TABLE 2.80
Average PDFs for CL Patch 1 (Example 3)

Average PDF	Rank	Exclusion Table	Shape	Scale	Location
W	1	All pass	2.16	3.86	−0.07
L	N/A	N/A	N/A	N/A	N/A
K	2	All pass	50.0	0.49	0.43

TABLE 2.81
Average PDFs for CL Patch 2 (Example 3)

Average PDF	Rank	Exclusion Table	Shape	Scale	Location
W	2	All pass	2.32	1.15	3.02
L	1	All pass	0.17	2.72	1.29
K	3	All pass	48.76	0.14	3.17

TABLE 2.82
Average PDFs for CL Patch 3 (Example 3)

Average PDF	Rank	Exclusion Table	Shape	Scale	Location
W	2	All pass	1.74	4.53	0.84
L	N/A	N/A	N/A	N/A	N/A
K	1	All pass	31.16	1.03	0.39

TABLE 2.83
Assessment Parameters of the Mapping and Indexing Stages (Example 3)

CL Patch (No.)	CNR (dB)	Number of Cells	PDF (Best)	Shape Parameter	Scale Parameter	Location Parameter
A (1)	9.34	124	W	2.16	3.86	−0.07
B (3)	13.47	306	K	31.16	1.03	0.39
C + D (2)	10.63	341	L	0.17	2.72	1.29

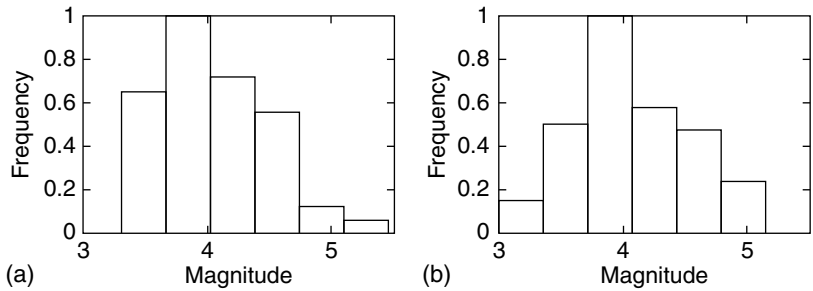

FIGURE 2.123 Data amplitude histograms including additive BN data (Example 3). (a) CL patch C; (b) CL patch D.

TABLE 2.84
Accurate Approximation of Underlying CL Patch PDFs (Example 3)

CL Patch	PDF	Distance ($\times 10^{-2}$)	Shape	Scale	Location	(u, v)
A	W	0.33	2.24	3.43	2.48	(−0.0278, 0.3345)
B	W	0.80	2.0	5.11	0.53	(−0.0353, 0.3485)
C	W	0.19	2.18	1.02	2.93	(−0.0292, 0.3384)
D	SU–Johnson	0.17	(2.97, −0.0094)	1.53	5.74	(−0.0098, 0.3156)

by o, for the accurate approximation of the underlying CL patch PDF (i.e., PDF corresponding to the sum of CL and BN in the patch). These were obtained by generating 1000 points for each CL patch. The results are summarized in Table 2.84. Figures 2.125, 2.127, 2.129, and 2.131 show the PDFs from

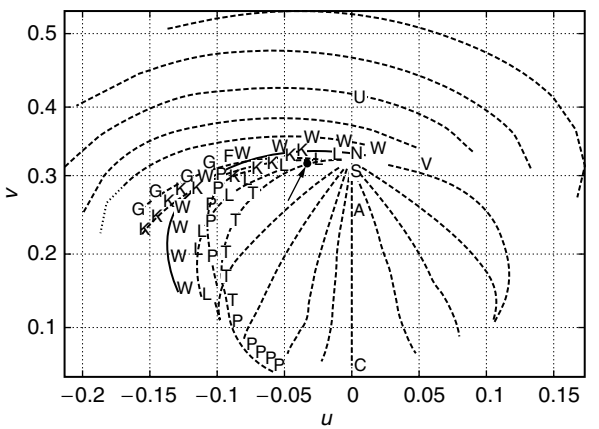

FIGURE 2.124 Location of the locus end point for CL patch A (Example 3).

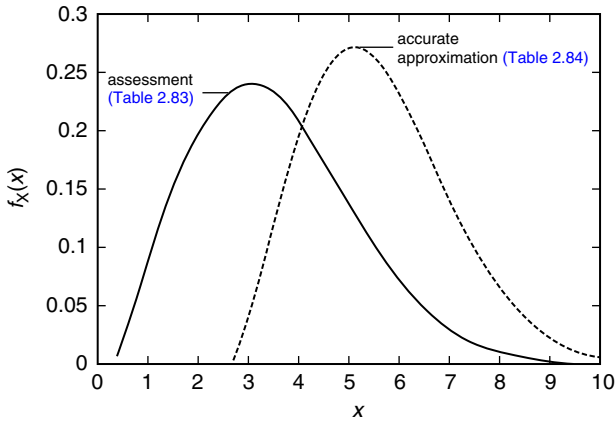

FIGURE 2.125 PDF comparison for CL patch A (Example 3).

Table 2.83 (solid lines) obtained by the mapping and indexing stages, superimposed on the corresponding PDFs from Table 2.84 (dashed lines).

The comparison for CL patch A is given in Figure 2.125. Here the best approximating and underlying PDFs are Weibull with shape parameters 2.16 and 2.24 and scale parameters 3.86 and 3.43, respectively. However, location parameters are not as closely matched. This accounts for the discrepancy between the two curves. Figure 2.127 shows the comparison for CL patch B. Although the best approximating PDF is K while the underlying PDF is Weibull, the curves are reasonably close. The results for CL patch C, defined as part of CL patch 2, are shown in Figure 2.129. The best approximating PDF is Lognormal while

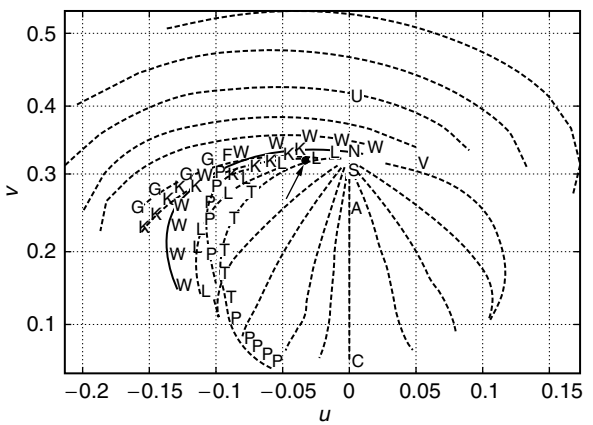

FIGURE 2.126 Location of the locus end point for CL patch B (Example 3).

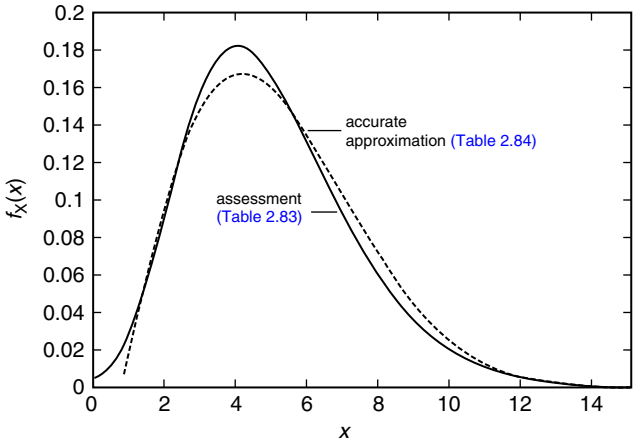

FIGURE 2.127 PDF comparison for CL patch B (Example 3).

the underlying PDF is Weibull. In this case, the two curves are reasonably close except for a small offset in their location and height. Finally, the comparison for CL patch D, defined as part of CL patch 2, is shown in Figure 2.131. The best approximating PDF is Lognormal while the underlying PDF is SU-Johnson. The two curves differ substantially and a poor approximation has been made. As was the case in Example 2, had the SU-Johnson been included in the library of allowable PDF kinds, a better approximating PDF would have been obtained.

The best approximating PDFs from Table 2.83 are superimposed on the histograms for CL patches A, B, C, and D in Figures 2.132–2.135, respectively. Whether or not the best approximating PDFs are close to the underlying PDFs, note that the best approximating PDFs nicely overlay the corresponding histograms.

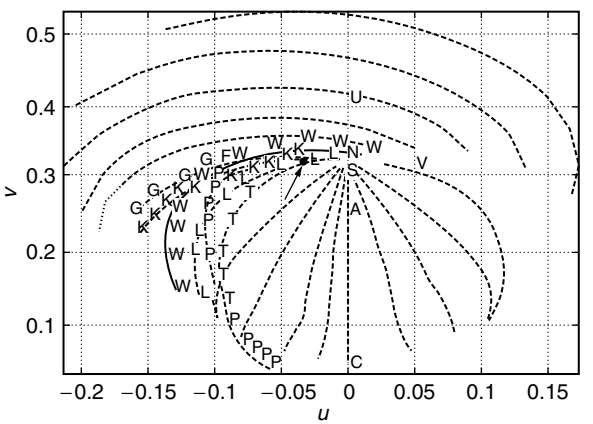

FIGURE 2.128 Location of the locus end point for CL patch C (Example 3).

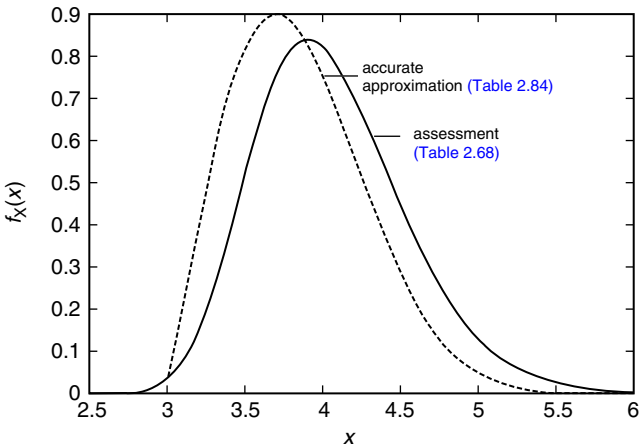

FIGURE 2.129 PDF comparison for CL patch C (Example 3).

2.7.6. Conclusion

In this chapter, expert system rules were developed to control the decisions needed for operation of the mapping and indexing stages. Discrepancies were determined which enable identification of the sources of distortions which in turn enables reprocessing.

Using the results developed in this chapter the three examples begun in Section 2.5.4 were completed. For Example one, in which the CNRs for CL patches A, B, C, and D were 10, 20, 30, and 40 dB, respectively, it was possible to identify and approximate all four CL regions. This was possible because of the significant difference in amplitudes between the CL patches and the BN.

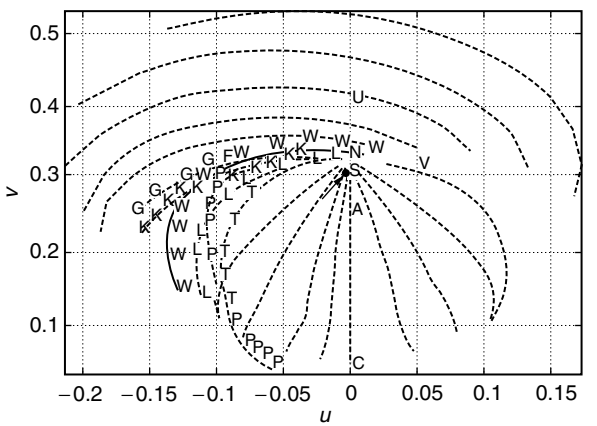

FIGURE 2.130 Location of the locus end point for CL patch D (Example 3).

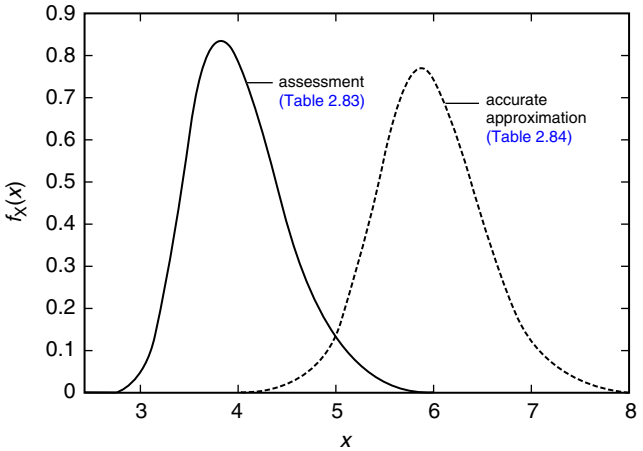

FIGURE 2.131 PDF comparison for CL patch D (Example 3).

Example two had added complexity because the CNRs of the contiguous CL patches C and D were identical. Even so, the indexing procedure was able to recognize that CL patch two consisted of two CL subpatches. In addition, the poor PDF approximation that can result with a limited library of PDFs was demonstrated. Finally, Example three investigated the situation where, even though the underlying PDFs of CL patches C and D were noticeably different, the addition of BN resulted in similar histograms. As a result, CL patches C and D were modeled as a single homogeneous CL patch. This example illustrates one of the pitfalls that can arise with small amounts of random data.

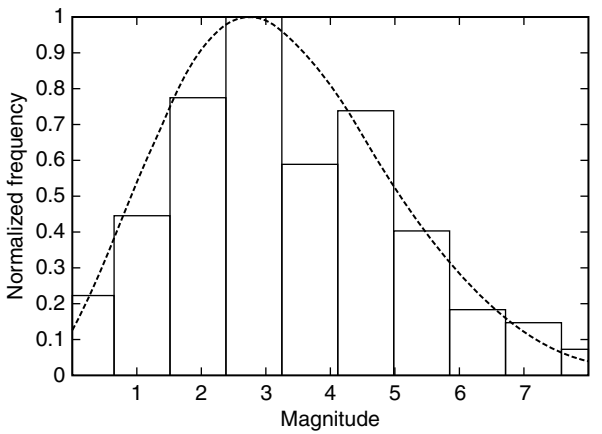

FIGURE 2.132 Histogram and best approximating PDF for CL patch A (Example 3).

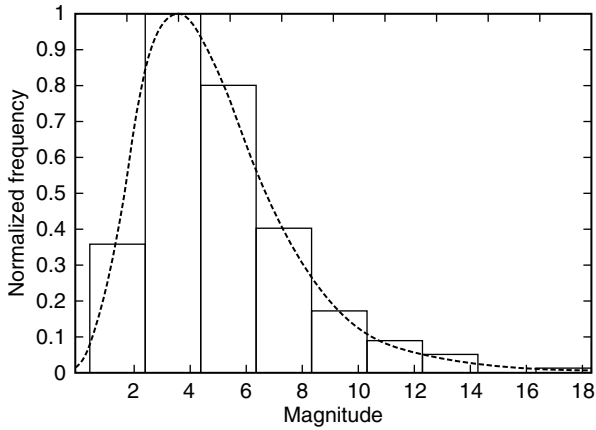

FIGURE 2.133 Histogram and best approximating PDF for CL patch B (Example 3).

2.8. CONCLUSION AND FUTURE RESEARCH

2.8.1. Conclusion

Current radars are typically designed on the assumption that the CL and BN are Gaussian random processes. Such receivers are matched filters which maximize signal-to-noise ratio by filtering to the extent possible the desired target returns from the disturbance. When it is not possible to separate the CL from the target by means of spatial or temporal filtering, the optimal Gaussian receiver performs poorly.

However, it is known that an optimal nonGaussian receiver in a non Gaussian environment can perform significantly better than the Gaussian receiver

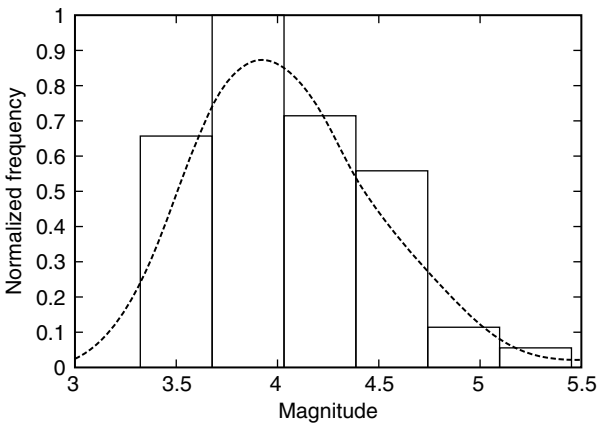

FIGURE 2.134 Histogram and best approximating PDF for CL patch C (Example 3).

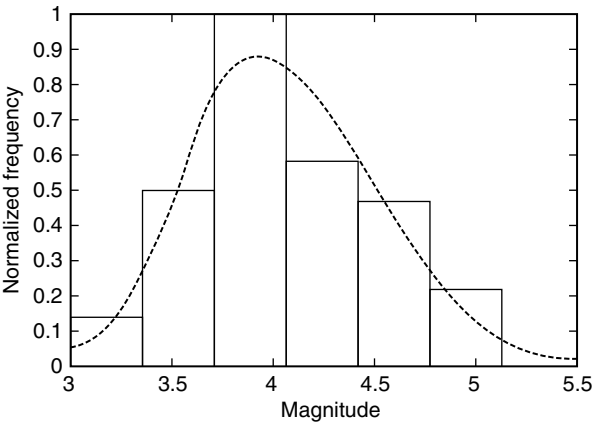

FIGURE 2.135 Histogram and Best Approximating PDF for CL patch D (Example 3).

in situations where the disturbance and target spectra completely overlap. The major problem is that there are an infinite different nonGaussian probability distributions and, in a practical situation, the environment is unknown.

This dissertation demonstrates, in conjuncture with an expert system, that it is possible to partition a surveillance volume into CL and BN regions and to approximate the underlying probability distribution of each region. This is highly significant because real data is likely to be nonhomogeneous and nonstationary. It is for this reason that optimal processors may not work well on real data. The results obtained in this research provide an innovative approach to analyzing and characterizing real data. Several computer generated examples are used to demonstrate the proposed methods.

2.8.2. Future Research

In this dissertation it was shown that the mapping and indexing stages have potential for being able to successfully monitor a random environment. The following problems were suggested by this research and remain to be addressed to in the future:

- Measures should be developed to enable a qualitative performance analysis of the proposed mapping and indexing procedures and the rules to be used by IPUS.
- Criteria to guarantee convergence of the iteration process in the mapping procedure should be developed.
- The IPUS control should be able not only to control individually the mapping and indexing stages but also to interact backward between these stages. Recall from Section 2.6.4.4 that outliers may exist in a set of reference cells that may arise due to (1) misclassified BN cells in a CL patch or misclassified CL cells in the BN, (2) cells having data

values of low probability of occurrence, and (3) cells containing signals from strong targets. Feedback between the indexing and mapping stages should be introduced to account and correct for outliers.

- A strategy should be developed to enable handling of the next set of data from the same environment to update the information of the mapping and indexing stages without processing all the data but rather by doing a series of checks.
- A detection strategy and expert system rules need to be developed for the target detection stage presented in Section 2.2.
- Following the detection stage, a target tracking strategy needs to be developed. Note, for example, that when a target is present in a given cell, its presence in subsequent cells should be consistent with a reasonable track.
- Environmental models based on collected real data and physical considerations should be verified using the Ozturk algorithm in order to gain confidence in these models.
- The Ozturk algorithm currently analyzes univariate random data. It is proposed to extend the Ozturk algorithm to the multivariate random data case.
- The trade off in computational complexity, for generating an average approximating PDF for each CL patch or subpatch vs. using the Ozturk algorithm once for all cells in a region, should be investigated.
- The effectiveness of using various image processing procedures found in the literature[24,25] for the mapping and indexing procedures should be investigated.

3 Statistical Analysis of the Nonhomogeneity Detector (for Excluding Nonhomogeneous Samples from a Subdivision)

CONTENTS

3.1. Gaussian Interference Backgrounds 176
(*M. Rangaswamy, J. H. Michels, and B. Himed*)
3.1.1. Introduction 176
3.1.2. Generalized Inner Product Statistics: Known Covariance Matrix 177
3.1.3. Generalized Inner Product Statistics: Unknown Covariance Matrix 178
3.1.4. Nonhomogeneity Detector 181
3.1.5. Performance Analysis of the Adaptive Matched Filter Test 183
3.1.6. Conclusions 187
3.2. NonGaussian Interference Backgrounds 187
(*M. Rangaswamy*)
3.2.1. Introduction 187
3.2.2. Preliminaries 189
3.2.3. Nonhomogeneity Detector for NonGaussian Interference Scenarios 190
3.2.3.1. Covariance Matrix Estimation 190
3.2.3.2. Maximally Invariant NHD Test Statistic 191
3.2.3.3. PDF and Moments of the NonGaussian NHD Test Statistic 192

3.2.3.4. Goodness-of-Fit Test 193
3.2.4. Performance Analysis..... 194
3.2.5. Conclusion 204

3.1. GAUSSIAN INTERFERENCE BACKGROUNDS

(M. RANGASWAMY, J. H. MICHELS, AND B. HIMED)

3.1.1. Introduction

An important issue in space–time adaptive processing (STAP) for radar target detection is the formation and inversion of the covariance matrix underlying the clutter and interference. Typically, the unknown interference covariance matrix is estimated from a set of independent identically distributed (iid) target-free training data that is representative of the interference statistics in a cell under test. Frequently, the training data is subject to contamination by discrete scatterers or interfering targets. In either event, the training data becomes nonhomogeneous. Consequently, it is not representative of the interference in the test cell. Estimates of the covariance matrix from nonhomogeneous training data result in severely undernulled clutter. Consequently, CFAR and detection performance suffer. Significant performance improvement can be achieved by employing preprocessing to select representative training data.

Consideration has previously been given to the problem of target detection using improved training strategies[1–4] and to the impact of nonhomogeneity on STAP performance.[5–8] The distribution information of a class of multivariate probability density functions (PDF) is succinctly determined through an equivalent univariate PDF of a quadratic form.[9] An application of this result is the nonhomogeneity detector (NHD) based on the generalized inner product (GIP).[1–5]

Nonhomogeneity of the training data arises from a number of factors such as contaminating targets, presence of strong discretes, and nonstationary reflectivity properties of the scattering surface. In these scenarios, the test cell disturbance covariance matrix, $\mathbf{R}_T$, differs significantly from the estimated covariance matrix, $\hat{\mathbf{R}}$ formed using target-free disturbance realizations from adjacent reference cells. If a large number of test cell data realizations are available, the underlying nonhomogeneity is characterized via the eigenvalues of $\hat{\mathbf{R}}^{-1}\mathbf{R}_T$.[10] However, in radar applications, only a single realization of test cell data is usually available. Consequently, the resulting estimate of $\mathbf{R}_T$ is singular. Hence, the empirically formed GIP has been compared with a theoretical mean corresponding to a "known" covariance matrix.[1–5] Large deviations of the GIP mean from the theoretical mean have been ascribed to nonhomogeneity of the training data. Such an approach provides meaningful results in the limit of large training data size. In practice, the amount of training data available for a given application is limited by system considerations such as bandwidth and fast scanning arrays. Furthermore, the inherent temporal and spatial nonstationarity of the interference precludes

the collection of large amounts of training data. Consequently, the approach of Refs.1–5 can be misleading since it ignores finite data effects and the resulting variability in the covariance matrix estimate.[11] Specifically, we note that the empirical GIP mean using an estimated covariance matrix with finite data can be twice as large as the corresponding GIP mean for a known covariance matrix in some instances. Consequently, such a scenario can easily lead to incorrect classification of training data.

The normalized GIP, P', admits a remarkably simple stochastic representation as the ratio of two statistically independent chi-square distributed random variables.[11] Consequently, the normalized GIP follows a central-F distribution.[20–22] The main result of this paper lies in exploiting these facts to construct a formal goodness-of-fit test for selecting homogeneous training data and its application to the performance of the adaptive matched filter (AMF) test.[12–14] Other applications of the F-distribution can be found in Refs. 23–25.

Section 3.1.2 briefly reviews the GIP statistics for the case of a known covariance matrix. In Section 3.1.3 we discuss the GIP statistics for the case of an unknown covariance matrix. Section 3.1.4 introduces the NHD and derives formal goodness-of-fit tests based on the GIP statistics described in Section 3.1.3. The AMF test performance with and without training data contamination using simulated and measured data is presented in Section 3.1.5. The AMF performance is shown to degrade with contaminated training data. It is further shown in Section 3.1.5 that the use of NHD preprocessing enables selection of representative training data. Consequently, use of NHD preprocessing restores the AMF test performance to case where there is no training data contamination.

3.1.2. Generalized Inner Product Statistics: Known Covariance Matrix

Let $\mathbf{X} = [X_1, X_2, \ldots, X_M]^{\mathrm{T}}$ denote a complex random vector with zero mean and known positive definite Hermitian covariance matrix $\mathbf{R}$ where T denotes transpose. The quadratic form given by

$$Q = \mathbf{X}^{\mathrm{H}}\mathbf{R}^{-1}\mathbf{X} \tag{3.1}$$

has the important property[11]

$$E(Q) = M \tag{3.2}$$

where E denotes mean expected value.

This result is important in that it is independent of the PDF underlying $\mathbf{X}$ and is only a function of the dimension of the random vector. If the PDF of $\mathbf{X}$ is known, the corresponding PDF of Q can be readily derived. For Gaussian distributed $\mathbf{X}$, i.e., $\mathbf{X} \sim CN(0, \mathbf{R})$, the PDF of Q is a chi-squared distribution with

M complex degrees of freedom. More precisely the PDF of Q is given by

$$f_Q(q) = \begin{cases} \dfrac{q^{M-1}}{\Gamma(M)}\exp(-q) & q \geq 0 \\ 0 & \text{Otherwise} \end{cases} \tag{3.3}$$

where $\Gamma(\cdot)$ is the Eulero–Gamma function.

The GIP based NHD calculates the quadratic form Q using an estimated covariance matrix (formed from iid target free training data) and compares its mean with M. Deviations from M have been attributed to nonhomogeneities in the training data.[1–5] In practice, the interference covariance matrix is formed from a finite amount of training data. The statistical variability associated with the data could introduce additional errors and thus, deviations of the GIP from M cannot entirely be ascribed to the presence of nonhomogeneities. Consequently, it is useful to work with the statistics of Q formed with an estimated covariance matrix with finite sample support. The GIP PDF and moments are quite different from those of Equation 3.2 and Equation 3.3 for the finite sample support problem.

3.1.3. Generalized Inner Product Statistics: Unknown Covariance Matrix

Let $\mathbf{X} \sim CN(0, \mathbf{R}_T)$ denote the random test data vector and $\mathbf{Z}$ denote a random data matrix, whose columns $\mathbf{Z}_k, k = 1, 2, \ldots, K$ are iid $CN(\mathbf{0}, \mathbf{R})$ target-free training data vectors. For homogeneous (representative) training data, $\mathbf{R}_T = \mathbf{R}$. The sample covariance matrix given by $\hat{\mathbf{R}} = (1/K)ZZ^H$ is the maximum likelihood estimate of the covariance matrix. Let

$$P = \mathbf{X}^H\hat{\mathbf{R}}^{-1}\mathbf{X} \tag{3.4}$$

We derive a canonical stochastic representation for the normalized GIP, $P' = P/K$, in terms of two statistically independent chi-squared distributed random variables in Appendix A. Consequently, we have

$$P = \frac{R_1}{R_2/K} \tag{3.5}$$

where R_1 and R_2 are statistically independent chi-squared distributed random variables with PDFs given by

$$f_{R_1}(r_1) = \begin{cases} \dfrac{r_1^{M-1}}{\Gamma(M)}\exp(-r_1) & r_1 \geq 0 \\ 0 & \text{Otherwise} \end{cases} \tag{3.6}$$

$$f_{R_2}(r_2) = \begin{cases} \dfrac{r_2^{K-M}}{\Gamma(K-M+1)}\exp(-r_2) & r_1 \geq 0 \\ 0 & \text{Otherwise} \end{cases} \tag{3.7}$$

respectively. Consequently, P' follows a central-F distribution[20–22] given by

$$f_{p'}(p') = \begin{cases} \dfrac{(p')^{M-1}}{\beta(M,L)(1+p')^{L+M}} & p' \geq 0 \\ 0 & \text{Otherwise} \end{cases} \tag{3.8}$$

where $\beta(M,L) = \int_0^1 \varsigma^{M-1}(1-\varsigma)^{L-1}\,d\varsigma$ and $L = K - M + 1$.

The statistical equivalence of P' to the ratio of two independent chi-squared distributed random variables is fascinating in that it permits rapid calculation of the moments of P. More importantly, it is extremely useful in Monte Carlo studies involving computer generation of P. For homogeneous training data, the use of Equation 3.5 circumvents the need to explicitly generate the test data vector $\mathbf{X}$ and the training data vectors used for covariance estimation. For large M and perforce K, significant computational savings can be realized from the method of Equation 3.5. It can be readily shown that

$$E(P) = \frac{M}{\left(1 - \dfrac{M}{K}\right)} \tag{3.9}$$

$$\text{Var}(P) = \sigma_P^2 = \frac{M}{\left(1 - \dfrac{M}{K}\right)^2\left[1 - \dfrac{(M+1)}{K}\right]} \tag{3.10}$$

where $E(P)$ and $\text{Var}(P)$ denote the mean and variance of P, respectively. Observe that the moments of P formed from an estimated covariance matrix (sample covariance matrix) with finite sample support deviate significantly from the corresponding moments for the case of a known covariance matrix given by Equation 3.1 and Equation 3.2. For example, with $K = 2M$, there is a 100% deviation of the mean of Equation 3.9 from that of Equation 3.1. Therefore, comparison of an empirically formed GIP with the theoretical mean of Equation 3.1 provides misleading results in that a finite data effect is ascribed to training data nonhomogeneity.

We then study the representation of Equation 3.5 in the limit of large K. For this purpose, we consider the characteristic function of R_2/K given by

$$\phi_{\frac{R_2}{K}}(j\omega) = E\left[\exp\left(-j\omega\frac{r_2}{K}\right)\right] = \frac{1}{\left(1 + \dfrac{j\omega}{K}\right)^{K-M+1}} \tag{3.11}$$

For $K \to \infty$, $\lim_{K\to\infty} \phi_{\frac{R_2}{K}}(j\omega) = \exp(-j\omega)$. Taking the inverse Fourier Transform, we have

$$\lim_{K\to\infty} f_{\frac{R_2}{K}}(r) = \delta(r-1) \tag{3.12}$$

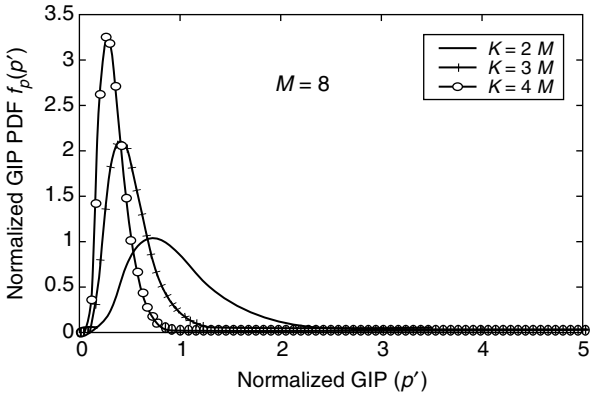

FIGURE 3.1 Normalized generalized inner product PDF.

Hence, for $K \to \infty$, R_2/K becomes unity with probability one. Thus, the GIP for this case is simply R_1 and hence, follows a Chi-squared distribution with M complex degrees of freedom. Consequently, for $K \to \infty$, $E(P) = \sigma_P^2 = M$ corresponding to the known covariance matrix results. Consequently, the GIP statistical representation given by Equation 3.5 provides additional insights on the NHD. The numerator random variable corresponds to the GIP statistics for known covariance matrix. The denominator random variable succinctly embeds the deleterious effects of estimating the covariance matrix with finite sample support. Deviation of the normalized GIP statistics from the PDF of Equation 3.8 can then be attributed to nonhomogeneity of the training data.

Figure 3.1 shows the PDF of P' for several values of K with $M = 8$ for Gaussian interference statistics. Observe that the variance of P' decreases with increasing K. This is anticipated since $\hat{\mathbf{R}} \to \mathbf{R}$ with probability one as $K \to \infty$. Consequently, the statistics of P' incur a dependence on K resulting from the use of finite sample support in estimating the covariance matrix.

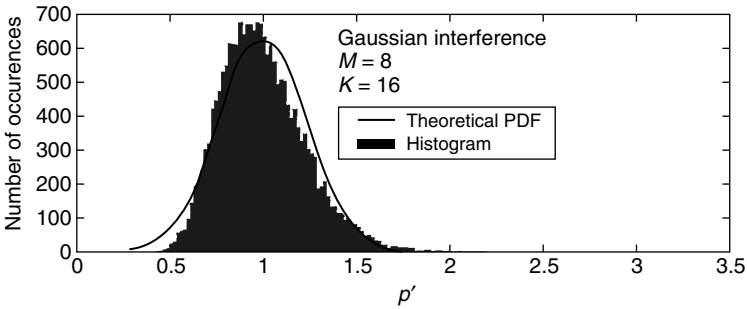

FIGURE 3.2 Empirical and theoretical generalized inner product PDF.

The results presented in Figure 3.2 correspond to the case of homogeneous training data. They show a comparison of the histogram of P' obtained from Monte Carlo realizations using simulated data with the theoretically predicted PDF of P' obtained from Equation 3.8. The results reveal good agreement between the theoretical prediction and the empirically generated values. The mean value of P, 15.957, obtained via 50,000 Monte Carlo realizations compares well with the theoretical predicted value of 16.

3.1.4. Nonhomogeneity Detector

We now present two methods for selecting homogeneous data from a set of training data. The first method exploits the central-F distribution of P' given by Equation 3.8 to construct a formal goodness-of-fit test, while the second method relies upon a comparison of empirically formed P' with the theoretical mean predicted by Equation 3.9 and discarding those realizations for which P' deviates significantly from the theoretical mean. More precisely, the difference between the empirical realizations of P' and the theoretically calculated mean value is calculated for each realization of P'. This difference is then rank ordered and the training data realizations corresponding to the least deviation from the theoretical mean are retained for subsequent use in STAP algorithms. The cumulative distribution function of P' is given by

$$\Pr(P' \leq r) = 1 - \text{betainc}\left(\frac{1}{r+1}, M, L\right) \tag{3.13a}$$

where betainc (x, m, n) is the incomplete beta function defined by

$$\text{betainc}(x, m, n) = \frac{1}{\beta(m, n)} \int_0^x w^{m-1}(1-w)^{n-1} \mathrm{d}w \tag{3.13b}$$

The goodness-of-fit test consists of determining whether realizations of P' formed from a given set of training data are statistically consistent with the PDF of Equation 3.1. For this purpose a suitable type-I error, α, is chosen. More precisely, α is simply the probability of incorrectly rejecting the hypothesis that a given realization of P' is statistically consistent with the PDF of Equation 3.8. Specifically, we seek a threshold, λ such that

$$\alpha = \Pr(P' > \lambda) = 1 - \Pr(P' \leq \lambda) = \text{betainc}\left(\frac{1}{\lambda+1}, M, L\right) \tag{3.14}$$

λ is determined from a numerical inversion of Equation 3.14. The goodness-of-fit test consists of forming realizations of P' from a set of training data and rejecting those training data vectors for which P' exceeds λ. The second method is based on comparing the realizations of P with the theoretically predicted mean of P given by Equation 3.9 and retaining those realizations exhibiting least deviation from the theoretically predicted mean of Equation 3.9. Examples

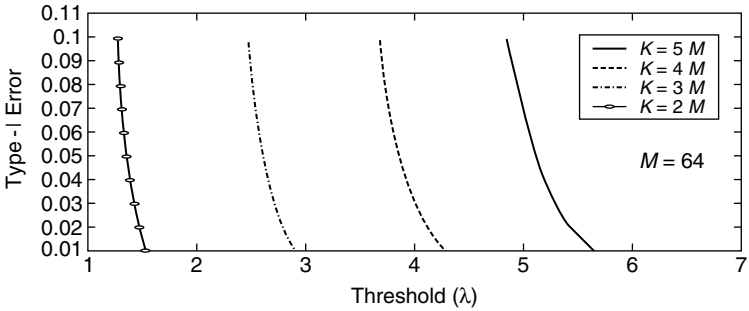

FIGURE 3.3 Type-I error versus threshold.

that illustrate the two approaches are presented. For a given training data set, a moving window approach is used to form realizations of P'. This approach is suboptimal because it does not guarantee statistical independence of the realizations of P'. However, we adopt this approach due to the limited training data support. For the examples presented here, data from the MCARM program[15] corresponding to 16 pulses and 8 channels from acquisition "220" on Flight 5, cycle "e" is used.

Figure 3.3 plots the type-I error (α) versus threshold for $M = 64$. Here different values of K are chosen to illustrate the threshold behaviour. For each value of α, λ is determined from a numerical inversion of Equation 3.14 and we observe an increase in λ for a given K.

The plot in Figure 3.4 shows P' and λ corresponding to $\alpha = 0.1$ as a function of range. A moving window approach is used to obtain P' for each range cell considered. Nonhomogeneity of the training data is seen in those

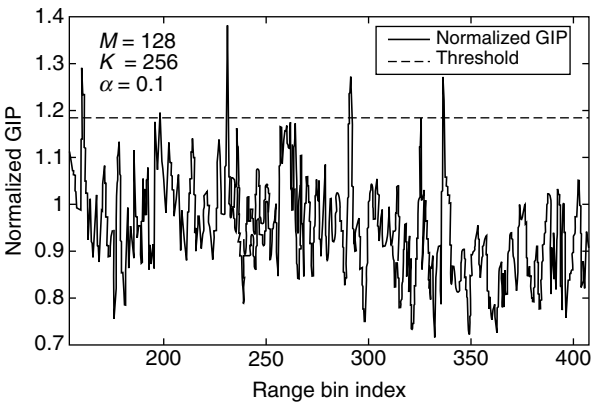

FIGURE 3.4 Normalized generalized inner product versus range, $\alpha = 0.1$.

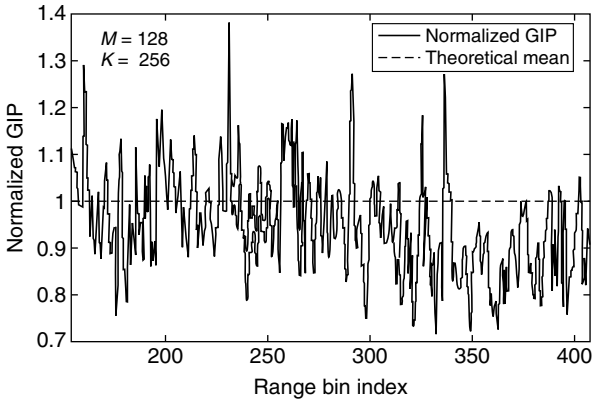

FIGURE 3.5 Normalized GIP versus range, simple normalization of GIP mean.

range cells for which P' exceeds λ. Figure 3.5 plots the normalized GIP as a function of range. The normalized GIP theoretical mean obtained from Equation 3.9 with a simple normalization is also shown. Values of the normalized GIP exceeding the theoretical mean correspond to nonhomogeneous training data realizations.

3.1.5. Performance Analysis of the Adaptive Matched Filter Test

In this section, we consider the performance analysis of the AMF test[12–14] in nonhomogeneous training data. The AMF test is given by

$$\Lambda = \frac{|\mathbf{S}^{\mathrm{H}}\hat{\mathbf{R}}^{-1}\mathbf{X}|^2}{|\mathbf{S}^{\mathrm{H}}\hat{\mathbf{R}}^{-1}\mathbf{S}|} \gtrless \lambda_{\mathrm{AMF}} \tag{3.15}$$

where $\mathbf{S}$ is the random spatio-temporal steering vector, $\mathbf{X}$ is the received random data vector, $\hat{\mathbf{R}}$ is the sample covariance matrix given by $\hat{\mathbf{R}} = (1/K)\sum_i^K \mathbf{Z}_i\mathbf{Z}_i^{\mathrm{H}}$ with $\mathbf{Z}_i$ denoting independent identically distributed training data and λ_{AMF} is a threshold selected to obtain a desired probability of false alarm.

For the case of homogeneous training data, analytical expressions for the probability of false alarm and probability of detection are given by Ref. 13

$$P_{\mathrm{fa}} = \int_0^1 \frac{f_\rho(\rho)\mathrm{d}\rho}{(1+\lambda_{\mathrm{AMF}}\rho)L} \tag{3.16}$$

$$P_{\mathrm{d}} = 1 - \int_0^1 \sum_{k=1}^{L} \left(\frac{L}{k}\right)\rho^k\lambda_{\mathrm{AMF}}^k G_k\left[\frac{\rho b}{(1+\rho\lambda_{\mathrm{AMF}})}\right]\frac{f_\rho(\rho)\mathrm{d}\rho}{(1+\rho\lambda_{\mathrm{AMF}})^L} \tag{3.17}$$

where

$$f_\rho(\rho) = \frac{(1-\rho)^{M-2}\rho^L}{\beta(M-1, L+1)} \qquad L = K - M + 1 \quad (3.18)$$

$$G_k(x) = \exp(-x)\sum_{n=0}^{k-1}\frac{x^n}{n!} \quad (3.19)$$

and "b" is related to the output signal-to-noise ratio (SNR). For $K \to \infty$, the sample covariance matrix tends to the true clutter covariance matrix, $\mathbf{R}$. Consequently, the AMF test converges to the matched filter (optimal receiver in Gaussian disturbance) for large K. The expressions for the matched filter P_{fa} and P_{d} are given by[13]

$$P_{\text{fa}} = \exp(-\lambda_{\text{MF}}) \quad (3.20)$$

$$P_{\text{d}} = \exp(-A)\sum_{k=0}^{\infty}\frac{A^k}{k!}[1 - G_k(\lambda_{\text{MF}})] \quad (3.21)$$

where A is related to the output SNR and λ_{MF} is the matched filter threshold.

Figure 3.6 presents P_{d} versus output signal-to-interference plus noise ratio (SINR). Relevant test parameters are reported in the plot. The matched filter (MF) curve obtained from Equation 3.21 corresponds to the optimal performance in Gaussian clutter. The P_{d} curve for the AMF operating in homogeneous Gaussian clutter follows from Equation 3.17 and exhibits performance to within 3 dB of the MF. The AMF performance operating in nonhomogeneous training data with and without NHD preprocessing is carried out by Monte Carlo simulation. For this example, the training data contained 30 high-amplitude, mainbeam discrete targets located at various range cells and Doppler frequencies. Initial sample support for NHD preprocessing is $6M$. A sliding window approach is used to select

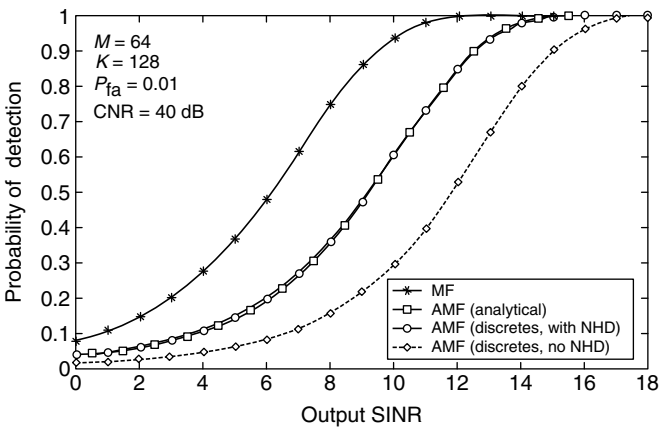

FIGURE 3.6 Performance of the adaptive matched filter "with" and "without" nonhomogeneity detector.

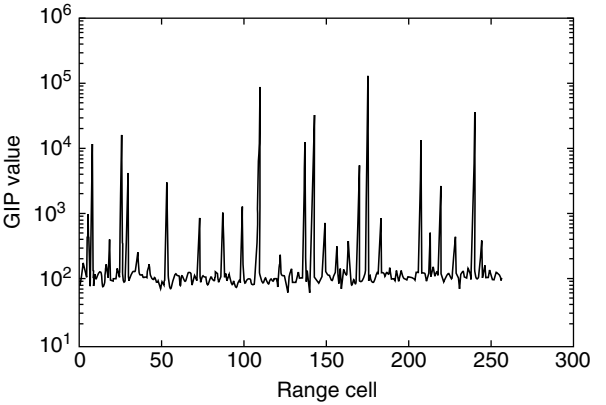

FIGURE 3.7 Generalized inner product versus range.

a subset consisting of $4M$ training data realizations. Each GIP value obtained at a specific range cell is computed using $\hat{\mathbf{R}}$ formed from $2M$ adjacent training data vectors. Previously, we noted the suboptimality of this scheme. In practice, its use is dictated by training data size limitations. In this manner $4M$ GIP values are obtained. The NHD preprocessing used in this example is based on a comparison of the empirical GIP with its theoretical mean value given by Equation 3.9. The training data used in forming $\hat{\mathbf{R}}$ after NHD processing is obtained by sorting the GIP values and retaining $K = 2M$ realizations corresponding to the smallest GIP deviation from the theoretical mean of Equation 3.9. Observe that the AMF performance in nonhomogenous clutter degrades severely. Also note that, for this case, NHD preprocessing restores the AMF performance to its analytical value.

Figure 3.7 shows a plot of the GIP versus range prior to NHD preprocessing for the simulated data used in carrying out the performance analysis of Figure 3.6. Figure 3.8 shows a plot of the sorted absolute value of the difference between

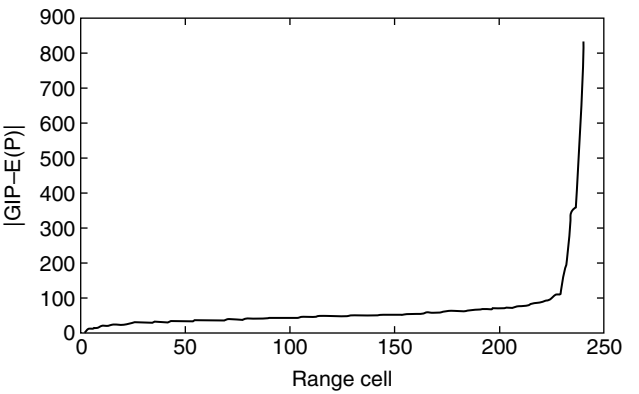

FIGURE 3.8 Absolute value of difference between GIP and theoretical GIP mean versus range.

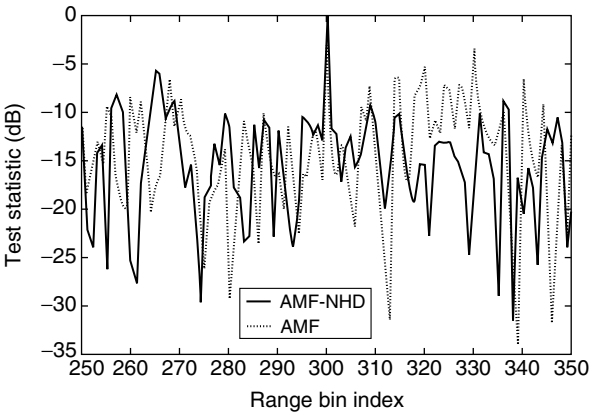

FIGURE 3.9 Test statistic versus range.

the GIP and its theoretical mean versus range after NHD preprocessing for the example in Figure 3.6. Observe the absence of discretes in the first $K = 2M$ range cells.

Figure 3.9 depicts performance using measured data from the MCARM program.[15] For this case, it is not possible to present performance in terms of detection probability versus SINR. This is due to the fact that only one realization of target present data is available. Hence, we present a plot of the detection test statistic versus range. Since the AMF test statistic is an ad hoc estimate of the output SINR, and since the probability of detection is a monotonically increasing function of the output SINR, this is an acceptable performance metric.

Performance of the AMF without NHD processing degrades significantly in nonhomogeneous clutter. Performance improvement is noted when the AMF is employed in nonhomogeneous data with NHD preprocessing. Consequently, the use of NHD affords moderate performance improvement of the AMF test in nonhomogeneous clutter. The performance with measured data is characterized by the ratio Ψ_1, of the test statistic at the test cell to the mean of the test statistics formed from adjacent cells, and also by the ratio Ψ_2, of the test statistic at the test cell to the highest test statistic formed from adjacent cells. Table 3.1 shows these values for the AMF test with and without NHD preprocessing.

TABLE 3.1
Adaptive Matched Filter Performance with Measured Data

Algorithm	ψ_1 (dB)	ψ_2 (dB)
AMF with NHD	13.25	5.68
AMF	11.83	3.38

3.1.6. Conclusions

Section 3.1 has made several significant contributions. First, we provided a statistical characterization of the GIP based NHD developed in Refs. 1–5. We showed that the underlying GIP statistics deviate significantly when the unknown covariance matrix is estimated using finite sample support. A canonical representation for the GIP in terms of two statistically independent Chi-Square distributed random variables and the resulting central-F distribution for the normalized GIP were then used to construct goodness-of-fit tests, whose performance is presented using both simulated and measured data. Application of this method as a preprocessing method for training data selection in the AMF algorithm was presented. Performance of the AMF in contaminated training data degrades significantly. The use of our preprocessing method for training data selection restores the AMF performance to within 3 dB of the optimal MF performance. This fact is illustrated with simulated as well as measured data from the MCARM program. Future work will undertake extensive performance comparisons between covariance based STAP methods such as the AMF and the normalized adaptive matched filter (NAMF)[16] (with NHD preprocessing) and model-based parametric STAP tests such as the parametric adaptive matched filter (PAMF) [17], normalized parametric adaptive matched filter (N-PAMF)[18] and fast adaptive processors[19] in nonhomogeneous interference backgrounds.

3.2. NONGAUSSIAN INTERFERENCE BACKGROUNDS

(M. RANGASWAMY)

3.2.1. Introduction

An important issue in space–time adaptive processing (STAP) for radar target detection is the formation and inversion of the covariance matrix underlying the disturbance. In practice, the unknown interference covariance matrix is estimated from a set of iid target-free training data that is assumed to be representative of the interference statistics in a cell under test. Frequently, the training data is subject to contamination by discrete scatterers or interfering targets. In either event, the training data becomes nonhomogeneous. As a result, it is not representative of the interference in the test cell. Hence, standard estimates of the covariance matrix from nonhomogeneous training data result in severely undernulled clutter. Consequently, CFAR and detection performance suffer. Significant performance improvement can be achieved by employing preprocessing to select representative training data.

Consideration has previously been given to the problem of target detection using improved training strategies[1–5] and to the impact of nonhomogeneity on STAP performance.[5–8] Use of the nonhomogeneity detector (NHD) based on the generalized inner product (GIP) measure for STAP problems involving Gaussian interference scenarios has been addressed[1–4,8,9] and extended significantly to include the effects of finite sample support used for covariance

matrix estimation.[10,11] However, the corresponding problem for nonGaussian interference scenarios has received limited attention. This is due to the fact that tractable models for correlated nonGaussian interference have become available only in recent works.[12–14]

Nonhomogeneity of training data can occur due to environmental factors, such as, the presence of strong discrete scatterers, dense target environments, nonstationary reflectivity properties of the scanned area, and radar system configurations like conformal arrays and bistatic geometries. A variety of robust adaptive signal processing methods to combat specific types of nonhomogeneities have been developed.[15–19]

In Section 3.2, we concern ourselves with the problem of training data nonhomogeneity caused by dense target environments and present the NHD for nonGaussian interference scenarios. More specifically, two p-tuple random vectors $\mathbf{X}_t$ and $\mathbf{X}_s$ having covariance matrices $\mathbf{R}_t$ and $\mathbf{R}_s$, respectively, are defined to be nonhomogeneous if $\mathbf{R}_s^{-1}\mathbf{R}_t \neq v\mathbf{I}$, where $\mathbf{I}$ denotes the $p \times p$ identity matrix and v is an arbitrary positive scale factor. In other words, the random vectors are defined to be nonhomogeneous if they do not share the same covariance structure.

In Section 3.2, we derive the NHD for nonGaussian interference scenarios, which can be modeled by spherically invariant random processes (SIRP), and present a statistical analysis of the resultant NHD test. Section 3.2.2 presents the relevant mathematical preliminaries. In Section 3.2.3 we discuss the issues of covariance matrix estimation using finite data as well as the use of a maximally invariant test statistic for the NHD. Furthermore, we present a statistical analysis of the NHD and show that a formal goodness-of-fit test can be constructed for selecting homogeneous training data. The basis of our NHD strategy is a characterization of the statistics pertaining to homogeneous SIRP clutter scenarios and rejecting realizations departing from these statistics. Performance analysis is discussed in Section 3.2.4. Also included therein is a performance comparison with existing NHD tests. Conclusions and future research directions are outlined in Section 3.2.5.

In general, the problem of nonhomogeneity detection for SIRPs is complicated by the fact that the underlying SIRP covariance matrix and characteristic PDF are unknown. Knowledge of the SIRP characteristic PDF is assumed in this chapter as a first step towards addressing the problem. This information can be gained from estimates of the first order PDF obtained from experimental data using histogram or moment techniques.[20] Significant performance penalty is incurred if this information is unavailable. This fact is illustrated through an example in Section 3.2.4.

The main contributions of this paper are summarized below:

1. Reduce the NHD problem for SIRP interference scenarios to one of testing whether two data sets share a common covariance structure but have different levels by proper use of the maximum likelihood estimate of the covariance matrix.

2. Provide a formal goodness-of-fit test using a scale invariant test statistic.
3. Analytical expressions for the NHD PDF, which enable calculation of the threshold setting for the NHD test.
4. Performance analysis of the NHD test using simulated and measured radar data.
5. Performance comparison with existing NHD tests demonstrating superior performance in both SIRP as well as Gaussian scenarios.

3.2.2. PRELIMINARIES

Let $\mathbf{X} = [X_1 X_2 \ldots X_m]^{\mathrm{T}}$ denote a complex spherically invariant random vector (SIRV) having zero mean, positive definite Hermitian covariance matrix $\mathbf{R}$ and characteristic PDF $f_V(v)$. The PDF of $\mathbf{X}$ is given by[21]

$$f_{\mathbf{X}}(x) = \pi^{-M}|R|^{-1}h_{2M}(q) \tag{3.22}$$

where $|\cdot|$ denotes determinant and

$$q = \mathbf{X}^{\mathrm{H}}\mathbf{R}^{-1}\mathbf{X}, \qquad h_{2M}(w) = \int_0^\infty v^{-2M}\exp\left(-\frac{w}{v^2}\right)f_V(v)\mathrm{d}v \tag{3.23}$$

Every SIRV admits a representation of the form $\mathbf{X} = \mathbf{Z}V$,[22] where $\mathbf{Z}$ has a complex-Gaussian PDF, $CN(0, \mathbf{R})$, and V is a statistically independent random variable with PDF $f_V(v)$. Consequently, the covariance matrix of $\mathbf{X}$ is given by $\mathbf{R}_{\mathbf{X}} = \mathbf{R}E(V^2)$. In practice, $\mathbf{R}$ and $f_V(v)$ are unknown. For the purpose of this paper, we assume knowledge of $f_V(v)$ and treat the problem of nonhomogeneity detection with respect to unknown $\mathbf{R}$. Validity of the SIRP model for clutter encountered in STAP applications has been extensively discussed.[23]

Previous works[1–4,8–11,24] employed the GIP based NHD for Gaussian interference scenarios. The GIP based method relies on the statistics of a quadratic form given by $\hat{Q} = \hat{\mathbf{X}}^{\mathrm{H}}\hat{\mathbf{R}}^{-1}\mathbf{X}$ where $\hat{}$ denotes estimate and H denotes Hermitian. This method can be used as an NHD test statistic in SIRV interference if a perfect estimate of the covariance matrix can be obtained, which calls for an extremely large sample support size (infinite sample support). However, in practice the training data available in a given application is limited by system considerations such as, the bandwidth, fast scanning arrays, and more fundamentally the underlying spatio-temporal nonstationarity of the scenario. Thus, one is almost always forced to work with finite sample support. Consequently, the covariance matrix estimate for this problem can be obtained to within a constant of the sample covariance matrix, which is the maximum likelihood estimate of the covariance matrix underlying the Gaussian component of the SIRV. Typically, this constant is unknown in practice. Hence, proposed goodness-of-fit tests[3,4,9,10] cannot be properly implemented for this problem. On the other hand, proposed implementations of the NHD tests using the sample covariance matrix estimate for $\hat{\mathbf{R}}$ in SIRV scenarios[3,4,9–11,24] lead to incorrect declaration of data nonhomogeneity. This fact

is illustrated in the examples presented in Section 3.2.4. Therefore, we seek a scale invariant test statistic for this problem.

3.2.3. Nonhomogeneity Detector for NonGaussian Interference Scenarios

Let $\mathbf{X} \sim \text{SIRV}[0, \mathbf{R}, f_V(v)]$ denote the complex SIRV test data vector where $\mathbf{R}$ is unknown. Further, let $\mathbf{X}_i$, $i = 1, 2, \ldots K$, denote iid complex $\text{SIRV}[0, \mathbf{R}, f_V(v)]$ training data. The first step in deriving the NHD for SIRVs involves obtaining the maximum likelihood estimate of the underlying covariance matrix.

This estimate is then used in a test statistic that exhibits maximal invariance with respect to the unknown scaling of the estimated covariance matrix. The resulting test statistic takes the form of a NAMF, that has been extensively analyzed in Ref. 25–27 and references therein. As noted previously, the basis of our strategy to detect nonhomogeneity in the data is to first characterize the NHD PDF in homogeneous SIRP clutter scenarios and use this information to construct a formal goodness-of-fit test to reject data realizations that depart from the said PDF.

3.2.3.1. Covariance Matrix Estimation

The unknown covariance matrix is estimated from representative SIRV training data sharing the covariance structure of that of the test cell. Maximum likelihood (ML) estimation of the covariance matrix for SIRVs was first considered in Ref. 28. The work of Ref. 28 showed that covariance matrix estimation for SIRVs can be treated in the framework of a complete–incomplete data problem and pointed out that the maximum likelihood estimate of the covariance matrix is a weighted sample matrix. Since the covariance matrix estimate cannot be obtained in closed form, Refs. 28 and 29 use an iterative method known as the expectation–maximization (EM) algorithm. More precisely, let $\mathbf{x}_i$, $i = 1, 2, \ldots, K$, denote the realizations a given set of independent identically distributed training data sharing the covariance matrix of the random variable test data vector $\mathbf{X}$. Refs. 28 and 29 show that the ML estimate of the covariance matrix is given by

$$\hat{\mathbf{R}} = \frac{1}{K} \sum_{i=1}^{K} c_i x_i x_i^{\mathrm{H}} \tag{3.24}$$

where

$$c_i = -\frac{h'_{2M}(q_i)}{h_{2M}(q_i)}, \qquad h'_{2M}(w) = \frac{\partial h_{2M}(w)}{\partial w} = -h_{2M+2}(w) \tag{3.25}$$

and $q_i = \mathbf{x}_i^{\mathrm{H}} \hat{\mathbf{R}}^{-1} \mathbf{x}_i$, $i = 1, 2, \ldots K$. Since both sides of Equation 3.24 involve $\hat{\mathbf{R}}$ (the right hand side implicitly through c_i), it is not possible to obtain the estimate in closed form. Consequently, Ref. 28 used the EM algorithm to obtain an iterative solution to the problem. We adopt the approach of Ref. 28 for obtaining

the covariance matrix estimate in this work. A derivation of the covariance matrix estimate is contained in Appendix B. We note therein that the EM algorithm yields an estimate that is to within a multiplicative constant of the sample covariance matrix, which is the ML estimate of the covariance matrix underlying the Gaussian component of the SIRV. This fact was verified for all the simulated data examples presented in Section 3.2.4 by examining the eigenvalues of the estimated covariance matrix obtained at the convergence of the EM algorithm. Details pertaining to the initial start and convergence properties of the EM algorithm can be found in Ref. 28. The next step is to use this estimate in a maximally invariant decision statistic for nonhomogeneity detection.

Recognizing the need to know the characteristic SIRV PDF, which may be hard to obtain in some practical applications, the works of Refs. 30 and 31 propose recursive covariance matrix estimators for the class of nonGaussian processes where the random variable V of the SIRP model is treated as a deterministic but unknown parameter. Strictly speaking, the nonGaussian model used in Refs. 30 and 31 departs from the SIRP model due to the treatment of V as a deterministic but unknown scale factor. However, it serves a useful alternative model in some instances.

3.2.3.2. Maximally Invariant NHD Test Statistic

The maximal invariant statistic for different scaling of test and training data is given by[25]

$$\Lambda_{\mathrm{NAMF}} = \frac{|\mathbf{s}^{\mathrm{H}}\hat{\mathbf{R}}^{-1}\mathbf{X}|^2}{[\mathbf{s}^{\mathrm{H}}\hat{\mathbf{R}}^{-1}\mathrm{s}][\mathbf{X}^{\mathrm{H}}\hat{\mathbf{R}}^{-1}\mathbf{X}]} \tag{3.26}$$

where $\mathbf{s} = (1/\sqrt{M})[\,1 \quad 1 \ldots 1\,]^{\mathrm{T}}$. For convenience, we use a simple choice for $\mathbf{s}$ by designating it to be the first column of a normalized discrete Fourier transform (DFT) matrix. However, in most STAP applications, the spatio-temporal steering vector is a function of azimuthal angle and Doppler. Bearing in mind that we are concerned about training data containing contaminating targets, which share the same angle Doppler information as that of a desired target, the spatio-temporal steering vector provides a logical choice for $\mathbf{s}$.

The test statistic of Equation 3.26 has also been proposed as a suboptimal method for adaptive radar target detection in compound-Gaussian clutter.[32] Invariance properties of the test statistic of Equation 3.26 and its geometrical representation have been studied in Ref. 25 and references therein for the case of Gaussian interference statistics using a sample covariance matrix estimate. In SIRP interference, however, each training data vector realization is scaled by a different realization of V. Consequently, maximal invariance of the test statistic of Equation 3.26 afforded by the sample covariance matrix estimate no longer applies. This is due to the fact that the sample covariance matrix is no longer the maximum likelihood estimate of the covariance matrix for SIRV scenarios.[33] However, use of an estimated covariance matrix of the form of Equation 3.24

restores the maximal invariance property of the test statistic of Equation 3.26. The reason is that the resultant covariance matrix estimate is to within a multiplicative constant of the sample covariance matrix. This behavior has been verified for all the simulated data example presented in Section 3.2.4. Consequently, we now have a case where the covariance matrix of the test and training data share the same structure but have different unknown scaling. It has been established in Ref. 25 that Λ_{NAMF} is the invariant test statistic for this problem. Hence, the canonical representation for Λ_{NAMF} in terms of five random variables derived in Ref. 25 applies to this problem in a straightforward manner. However, we emphasize that it is important to properly estimate the SIRV covariance matrix in order to reduce the NHD problem to the case where test and training data covariance matrices differ by an unknown scale factor. This calls for knowledge of the first order SIRV characteristic PDF.

3.2.3.3. PDF and Moments of the NonGaussian NHD Test Statistic

Our comments in the concluding paragraph of Section 3.2.3.2 allow us to use the canonical representation for Λ_{NAMF} contained in Ref. 25 for Gaussian interference scenarios. Consequently, the PDF of the NHD test statistic is readily determined in terms of an equivalent random variable defined by

$$\Lambda_{\text{eq}} = \frac{\Lambda_{\text{NAMF}}}{1 - \Lambda_{\text{NAMF}}} \tag{3.27}$$

It has been shown in Refs. 25, 27 and 34 for Gaussian interference statistics that Λ_{eq} admits a representation in terms of an F-distributed random variable, P, and a beta-distributed loss factor, Γ. Because of the fact that the NHD problem in SIRV interference has been reduced to that of testing whether two data sets share the same covariance structure with differing scale, the results of Refs. 25, 27 and 34 readily extend to the SIRV problem. More precisely, for the case where no target is present in $\mathbf{x}$, Λ_{eq} admits a representation of the form

$$\Lambda_{\text{eq}} = \frac{P}{1 - \Gamma} \tag{3.28}$$

The PDFs of P and Γ are given by

$$f_P(p) = \frac{L}{(1+p)^{L+1}}, \qquad f_\Gamma(\gamma) = \frac{1}{\beta(L+1, M-1)} \gamma^L (1-\gamma)^{M-2} \tag{3.29}$$

where $L = K - M + 1$, and

$$\beta(m, n) = \int_0^1 x^{m-1}(1-x)^{n-1} \mathrm{d}x \tag{3.30}$$

After a little bit of algebra, it follows that the PDF of Λ_{NAMF} with no target present in x is given by

$$f_{\Lambda_{\text{NAMF}}}(r) = \int_0^1 \frac{L(1-\gamma)f_\Gamma(\gamma)(1-r)^{L-1}\mathrm{d}\gamma}{[1-\gamma r]^{L+1}} \tag{3.31}$$

The mean of Λ_{NAMF} is difficult to calculate analytically. Consequently, we work with the mean of Λ_{eq} given by

$$E(\Lambda_{\text{eq}}) = \frac{K}{(K-M)(M-2)} \tag{3.32}$$

to study the convergence properties in the limit of large K. The statistical equivalence, of Λ_{eq} to the ratio of an F-distributed random variable and a beta-distributed loss factor, permits rapid calculation of the moments of Λ_{eq}. Also, it is extremely useful in Monte Carlo studies involving simulation of Λ_{NAMF}. For homogeneous training data, the use of Equation 3.28 circumvents the need to explicitly generate the test data vector $\mathbf{X}$ and the training data vectors used for covariance estimation. For large M and hence large K, significant computational savings can be realized from the method of Equation 3.28. It is instructive to note that the PDF of Λ_{NAMF} depends only on M and K, which are under the control of a system designer, and not on nuisance parameters such as the true covariance matrix underlying the interference scenario. Furthermore, for $K \to \infty$ the mean of Equation 3.32 converges to $E(\Lambda_{\text{eq}}) = (1/(M-2))$, corresponding to the mean of an F-distributed random variable. This is due to the fact that as $K \to \infty$, the estimated covariance matrix approaches the true covariance matrix with probability one and thus the loss factor takes on the value zero with probability one.

3.2.3.4. Goodness-of-Fit Test

Since the PDF and mean of Λ_{NAMF} are known, a formal goodness-of-fit test can be used for nonhomogeneity detection in nonGaussian interference scenarios. The goodness-of-fit test can be formally cast in the form of the following statistical hypothesis test:

H_0: Λ_{NAMF} is statistically consistent with the PDF of Equation 3.31
H_1: Λ_{NAMF} is not statistically consistent with the PDF of Equation 3.31

For this purpose, we need to determine the type-I error given by

$$P_{\text{e}} = P(\Lambda_{\text{NAMF}} > \eta | H_0) = P\left[\Lambda_{\text{eq}} > \frac{\eta}{(1-\eta)} | H_0\right] \tag{3.33}$$

Using Equations 3.27 and 3.29, it follows that the probability of error conditioned on Γ is given by

$$P_e|\Gamma = \frac{1}{[1+(1-\gamma)\eta^*]^L} \tag{3.34}$$

where $\eta^* = \eta/(1-\eta)$. The unconditional type-I error probability is obtained by taking the expectation of Equation 3.34 over Γ and is given by

$$P_e \int_0^1 \frac{f_\Gamma(\gamma)}{[1+(1-\gamma)\eta^*]^L} d\gamma \tag{3.35}$$

In Section 3.2.4 the type-I error is chosen to be 0.01. The threshold, η^*, is determined by a numerical inversion of Equation 3.35. The value of η follows from the relationship $\eta = \eta^*/(1+\eta^*)$. We then form empirical realizations of Λ_{NAMF} from each training data realization using a sliding window approach. In this approach each training data vector is treated as a test cell data vector, whose covariance matrix is estimated from neighboring cell data according to Equation 3.24. We then test for statistical consistency of these realizations of Λ_{NAMF} with the PDF of Equation 3.31. Realizations of Λ_{NAMF} exceeding η correspond to nonhomogeneous training data. A desirable feature of P_e is that it depends only upon K and M and not on nuisance parameters such as the true covariance matrix underlying the interference. Performance analysis of the NHD method is presented in the next section.

3.2.4. Performance Analysis

Performance of the goodness-of-fit test with simulated and measured data is presented here. Figure 3.10 shows the plot of the PDF of Λ_{NAMF} with K as a parameter. Observe that the variance of Λ_{NAMF} decreases with increasing K, as increase of K provides better estimates of the covariance matrix estimate, and when $K \rightarrow \infty$, the estimated covariance matrix approaches the true covariance matrix to within a scale factor with probability one.

Figure 3.11 shows a plot of the Type-I error versus the threshold, η, with K as a parameter. For a given type-I error, the threshold decreases with increasing K, in conformance with the results of Figure 3.10.

For convenience of analysis simulated data examples contained herein use the K-distributed amplitude PDF given by[12,13,21]

$$f_R(r) = \frac{b^{\alpha+1} r^\alpha}{2^{\alpha-1}\Gamma(\alpha)} K_{\alpha-1}(br) \qquad r \geq 0,\ b, \alpha > 0 \tag{3.36}$$

where b and α are the distribution scale and shape parameters, respectively, $K_\nu(\cdot)$ is the modified Bessel function of the second kind of order ν and $\Gamma(\cdot)$ is the Eulero–Gamma function. The K-distribution, which is a member of the class of SIRPs,[12] has been proposed as a model for impulsive clutter resulting from

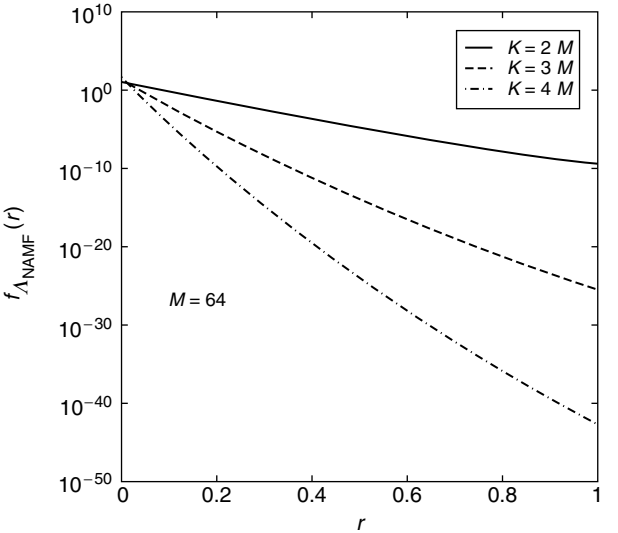

FIGURE 3.10 Λ_{NAMF} PDF for $K = 2M, 3M, 4M$: $M = 64$.

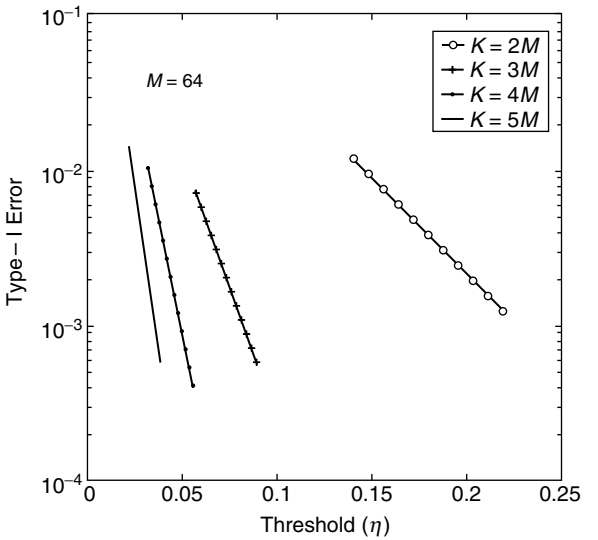

FIGURE 3.11 Λ_{NAMF} type-I error probability vs. threshold: $M = 64$.

terrain and sea scatter.[35,36] Small values of α result in heavy-tails for the PDF of Equation 3.36. The corresponding $fv(\nu)$ and $h_{2M}(\cdot)$ are given by

$$\begin{aligned} f_v(\nu) &= \frac{2b}{\Gamma(\alpha)}(b\nu)^{2\alpha-1}\exp(-b^2\nu^2) \qquad 0 \le \nu \le \infty \\ h_{2M}(w) &= \frac{2b^{2M}}{\Gamma(\alpha)}(b\sqrt{w})^{\alpha-M}K_{\alpha-M}(2b\sqrt{w}) \end{aligned} \tag{3.37}$$

We generate 1024 realizations of a 64-tuple vector from the *K*-distributed SIRP with $\alpha = 0.5$ having a prescribed covariance matrix according to the physical model described in Ref. 37 using the approach of Ref. 13. No targets are added to this data set. Starting from the midpoint (range bin 512), the data set is processed symmetrically on either side using a sliding window. Each cell is treated as a test cell (which may or may not contain contaminating targets). Two guard cells are provided (one on each side of the test cell). One hundred and twenty-eight training data realizations are collected by moving symmetrically on either side of the guard cells for use in covariance matrix estimation. The covariance matrix estimate is obtained using Equation 3.24. Λ_{NAMF} given by Equation 3.26 is then calculated for each test cell using the estimated covariance matrix and compared to a threshold determined from Equation 3.35 for $P_e = 0.01$. Relevant test parameters are reported in the plots.

Figure 3.12 shows the performance of the goodness-of-fit test for simulated homogeneous data from the *K*-distribution[21] with shape parameter 0.5 using

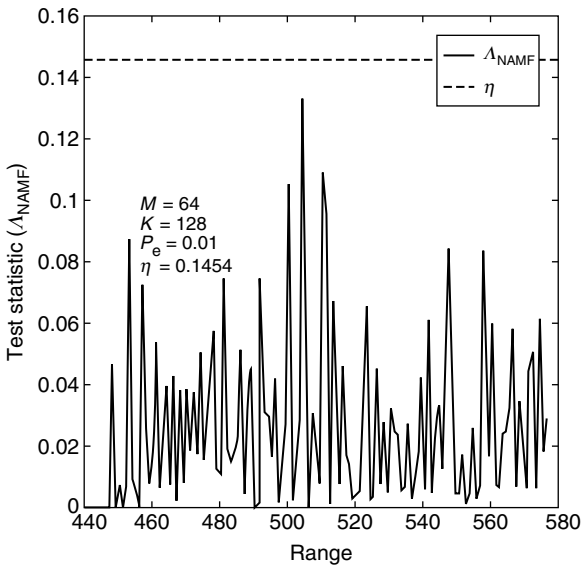

FIGURE 3.12 Λ_{NAMF} versus range bin number for homogeneous *K*-distributed SIRV with $\alpha = 0.5$, $M = 64$, $K = 128$.

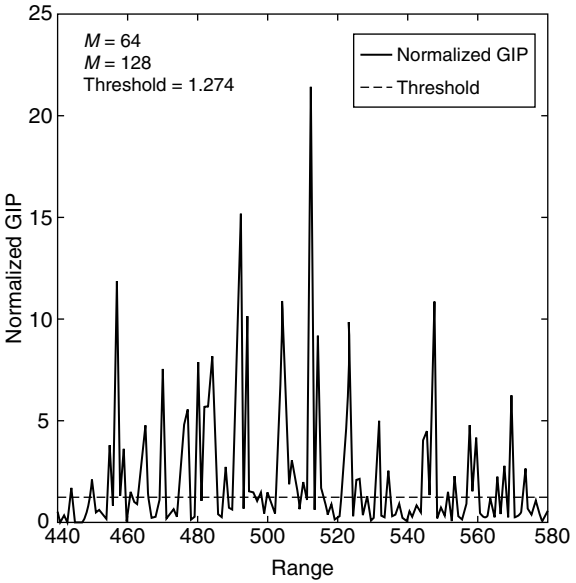

FIGURE 3.13 Normalized GIP ($\mathbf{x}^{\mathrm{H}}\mathbf{S}^{-1}\mathbf{x}/K$) versus range bin number for homogeneous K-distributed SIRV with $\alpha = 0.5$, $M = 64$, $K = 128$ and threshold = 1.274.

the covariance estimate of Equation 3.24. The figure shows a plot of Λ_{NAMF} as a function of range. No realization of Λ_{NAMF} exceeds η, reflecting homogeneity of the data. The experiment was repeated 1000 times and in all cases Λ_{NAMF} did not exceed η confirming the homogeneity of the data.

Figure 3.13 shows the performance of the NHD test proposed in Refs. 10, 11 and 24 based on comparing the normalized GIP, $\mathbf{x}^{\mathrm{H}}\mathbf{S}^{-1}\mathbf{x}/K$, with the threshold setting determined according to Equation 3.25 in Ref. 24. Here $\mathbf{S} = (1/K)\sum_{i=1}^{K}\mathbf{x}_i\mathbf{x}_i^{\mathrm{H}}$ is simply the sample covariance matrix. The data set used here is the same as the data set used for the example in Figure 3.12. The normalized GIP is formed using sliding window processing as described in Refs. 10, 11 and 24. Figure 3.13 shows a plot of the normalized GIP as a function of range. Also plotted is the threshold setting. From the plot it is evident that for almost all range bins the normalized GIP exceeds the threshold, leading to the declaration of nonhomogeneity, when in fact the data is homogeneous.

Figure 3.14 shows the performance of a second goodness-of-fit test proposed in Refs. 10, 11 and 24, which compares the normalized GIP, $\mathbf{x}^{\mathrm{H}}\mathbf{S}^{-1}\mathbf{x}/K$ to a theoretically calculated mean value obtained from Equation 3.6 of Ref. 24. The data used for this example is the same as that used in the example of Figure 3.12. Figure 3.14 shows a plot of the normalized GIP as a function of range. Also shown is the theoretically calculated mean value. Again, we see that for almost all range bins the normalized GIP exceeds the mean value causing an incorrect declaration of data nonhomogeneity.

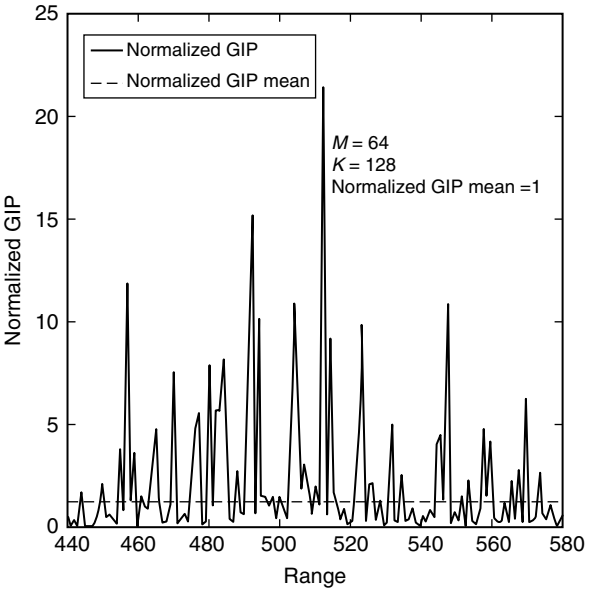

FIGURE 3.14 Normalized GIP ($\mathbf{x}^{\mathrm{H}}\mathbf{S}^{-1}\mathbf{x}/K$) versus range bin number for homogeneous K-distributed SIRV with $\alpha = 0.5$, $M = 64$, $K = 128$ and normalized GIP mean $= 1$.

Figure 3.15 shows the performance of the NHD test proposed in Refs. 3, 4 and 9, which compares the GIP, $\mathbf{x}^{\mathrm{H}}\mathbf{S}^{-1}\mathbf{x}$. to a theoretically specified mean value of M. The data used for this example is the same as that used in the example of Figure 3.12. Figure 3.15 shows a plot of the GIP as a function of range. Also shown is the theoretically specified mean value. Again, we see that for almost all range bins the GIP exceeds the mean value causing an incorrect declaration of data nonhomogeneity. The incorrect declaration of nonhomogeneity is due to the fact that $\mathbf{S}$ is no longer the ML estimate of the covariance matrix for SIRV interference scenarios. Similar results showing an even more severe performance degradation in K-distributed clutter with $\alpha = 0.1$ were obtained. However, these results are not reported here for avoiding tediousness of exposition. The experiments pertaining to Figure 3.13 to Figure 3.15 were repeated 1000 times and all the trials exhibited performance consistent with that reported in those figures.

Figure 3.16 shows the performance of the goodness-of-fit test developed in this paper in K-distributed clutter with shape parameter 0.5. Synthetic targets were injected at range bins 479 and 510 to cause the nonhomogeneity. Nonhomogeneity of the data is evident in those range bins where Λ_{NAMF} exceeds η.

Figure 3.17 shows the performance of the goodness-of-fit test in K-distributed clutter with $\alpha = 0.1$. Synthetic targets were injected at range bins 510 and 552 to cause the nonhomogeneity. Clearly Λ_{NAMF} exceeds η for both of these range bins and thus they are declared to be nonhomogeneous data sets.

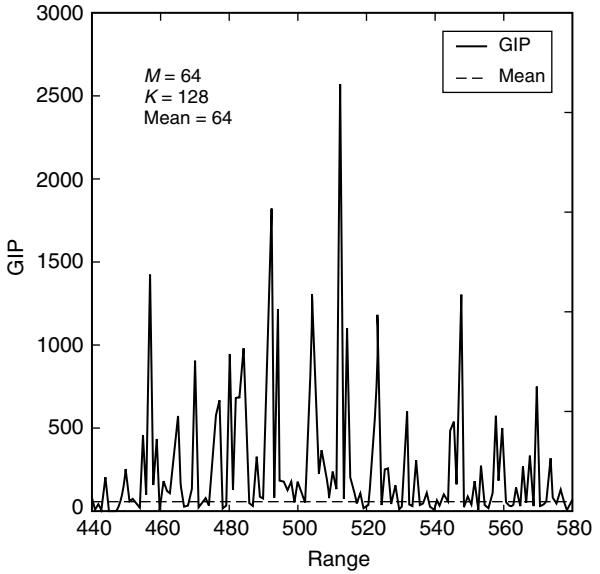

FIGURE 3.15 GIP ($\mathbf{x}^{\mathrm{H}}\mathbf{S}^{-1}\mathbf{x}$) versus range bin number for homogeneous K-distributed SIRV with $\alpha = 0.5$, $M = 64$, $K = 128$ and GIP mean $= 64$.

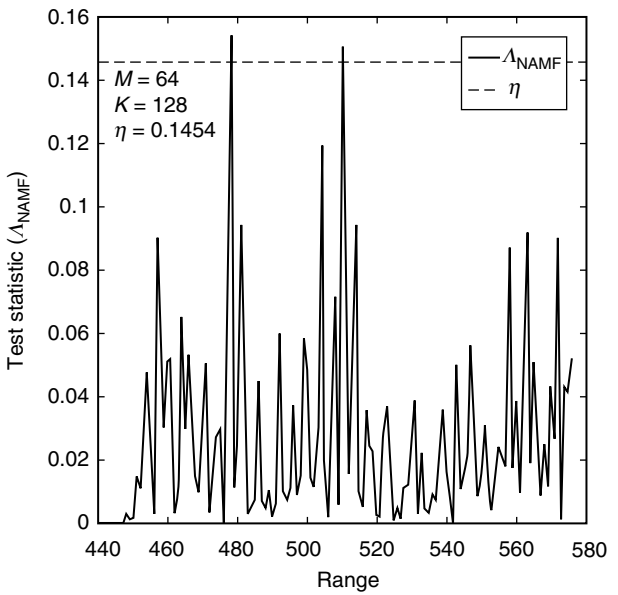

FIGURE 3.16 Λ_{NAMF} versus range bin number for nonhomogeneous K-distributed SIRV with $\alpha = 0.5$, $M = 64$, $K = 128$.

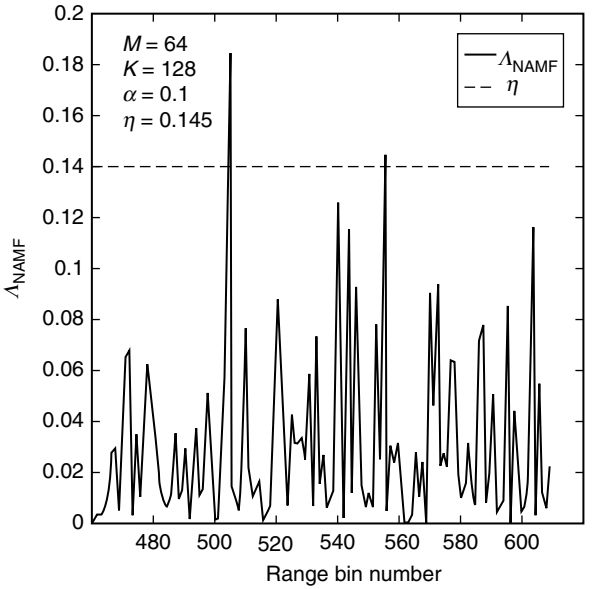

FIGURE 3.17 Λ_{NAMF} versus range bin number for nonhomogeneous K-distributed SIRV with $\alpha = 0.1$, $M = 64$, $K = 128$.

Figure 3.18 shows the results of the goodness-of-fit test using the covariance matrix estimator proposed in Refs. 30 and 31. This estimator does not require knowledge of the first order characteristic PDF of the SIRV and therefore converges faster than the estimator of Equation 3.24, especially for small values of α. The data set used for this example is the same as that used for the example in Figure 3.13. Although a peak in the test statistic is seen at range bin 479, it does not exceed the threshold, while the peak is not seen for range bin 510. Therefore, contaminating targets in range bins 479 and 510 are not detected. Furthermore, the method erroneously reports the presence of a contaminating target at range bin 573. This illustrates the importance of knowing the underlying characteristic PDF to properly estimate the covariance matrix and use it in the NHD statistic.

The results contained in Figure 3.16 to Figure 3.18 were further validated by using 1000 realizations of the experiment and averaging the results over 50 independent trials. In 997 out of the 1000 trials the NHD realizations corresponding to bins 479 and 510 of Figure 3.16 exceeded the threshold. For Figure 3.17, 984 times out of 1000 trials, the realizations corresponding to range bins 510 and 552 exceeded the threshold. The corresponding number for Figure 3.18 was 971. These findings are summarized in Table 3.2.

The examples reported in Figure 3.19 to Figure 3.21 make use of the MCARM data of Ref. 3. The MCARM data consists of measured L-band radar data using a Westinghouse radar mounted on the port-side of a BAC1-11 aircraft. The relevant system parameters are summarized in Table 3.3. The MCARM data is a common test bed for performance analysis and bench-marking of STAP

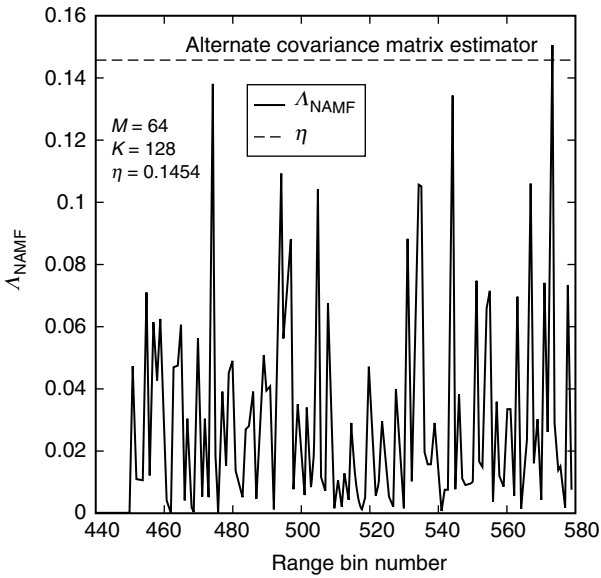

FIGURE 3.18 Λ_{NAMF} versus range bin number for nonhomogeneous K-distributed SIRV with $\alpha = 0.5$, $M = 64$, $K = 128$ and covariance matrix estimate of Refs. 30 and 31.

algorithms and is therefore considered in this paper. Further details pertaining to the MCARM data can be found in Ref. 38. Figure 3.19 shows the results of the goodness-of-fit test for the MCARM data using acquisition 220 on Flight 5, cycle "e" for 8 channels and 16 pulses. The results of Refs. 3, 39 and 40 using the MCARM tend to confirm that the MCARM data is homogeneous for the most part. Statistical analysis of the data indicates that the data is well-approximated by the Gaussian distribution.[3] As a consequence, $c_i = (-h'_{2M}(q_i))/(h_{2M}(q_i)) = 1$ for this case. Hence, the maximum likelihood estimate of the covariance matrix is simply the sample covariance matrix. The test statistic, Λ_{NAMF}, and the threshold, η, are plotted as a function of range. Nonhomogeneity of the data is evident in those bins for which Λ_{NAMF} exceeds η. For the sake of comparison, Figure 3.20 and Figure 3.21 show the performance of the NHD methods of Refs. 10, 11, 24 and

TABLE 3.2
Nonhomogeneity Detector Performance Summary

Figure Number	Range Bin Number	Number of Realizations	Number of Exceedences
3.16	510	1000	997
3.16	479	1000	997
3.17	510	1000	984
3.17	552	1000	984
3.18	573	1000	971

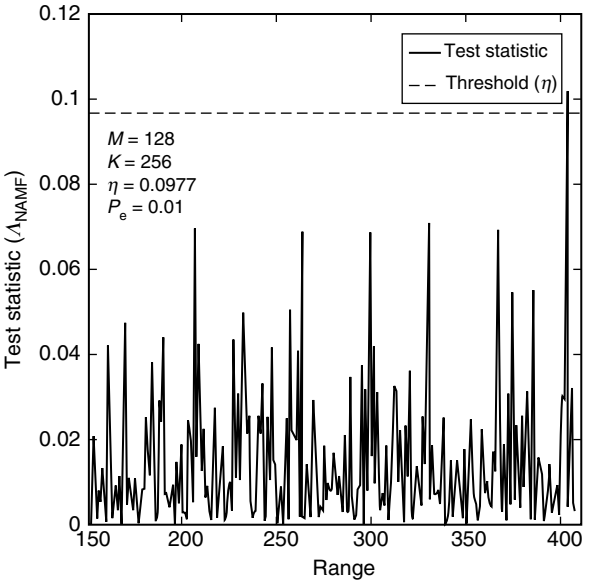

FIGURE 3.19 Λ_{NAMF} versus range bin number using MCARM data: $M = 128$, $K = 256$.

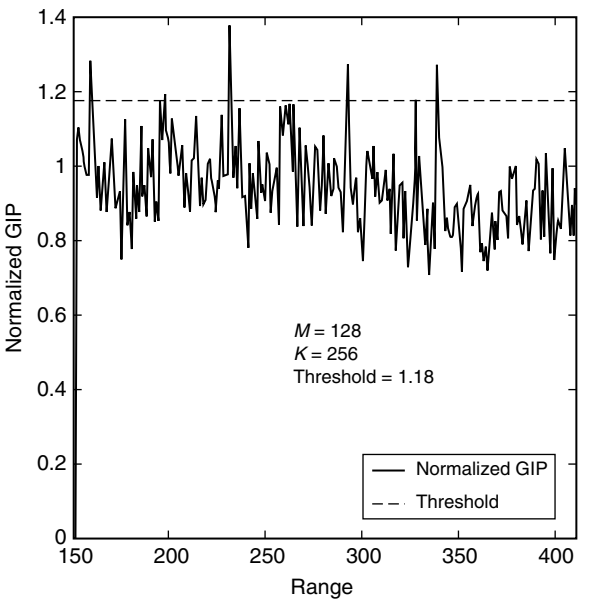

FIGURE 3.20 Normalized GIP ($\mathbf{x}^{\text{H}}\mathbf{S}^{-1}\mathbf{x}/K$) versus range bin number using MCARM data: $M = 128$, $K = 256$.

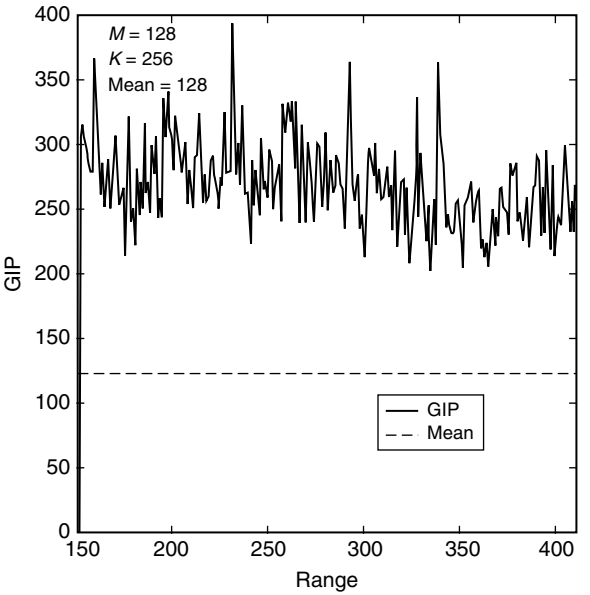

FIGURE 3.21 GIP ($\mathbf{x}^{H}\mathbf{S}^{-1}\mathbf{x}$) versus range bin number using MCARM data: $M = 128$, $K = 256$.

Refs. 3, 4, 9, respectively, using the same MCARM data set used in the example of Figure 3.19. Figure 3.20 shows the results of the NHD test proposed in Refs. 10,11 and 24 based on comparing the normalized GIP, ($\mathbf{x}^{H}\mathbf{S}^{-1}\mathbf{x}/K$) with the threshold setting determined according to Equation 4.2 in Ref. 24. The MCARM data set is processed in the same manner as described in the example of Figure 3.13.

TABLE 3.3
MCARM Data Parameters

Parameter	Value
Transmit frequency	1240 MHz
Transmit beamwidth	6.7° Az., 10.4° El
Waveform	50.4 μs LFM
Peak transmit power	20 kw
Pulse compression ratio	63
Platform altitude	10,000 ft
Platform velocity	100 m/s
Array configuration	11 × 11 planar array
Number of pulses	128
Pulse repetition frequency	2 kHz
Number of unambiguous range bins	630

Figure 3.21 shows the performance of the NHD test proposed in Refs. 3, 4 and 9, which compares the GIP, $\mathbf{x}^{H}\mathbf{S}^{-1}\mathbf{x}$, to a theoretically specified mean value of M using the MCARM data. The MCARM data set is processed in the same manner as described in the example of Figure 3.15. It is seen from Figure 3.20 and Figure 3.21 that a lot more declarations of nonhomogeneity result from the NHD methods of Refs. 10, 11, 24 and Refs. 3, 4, 9, when in fact the MCARM data is homogeneous. However, it may also be noted that the NHD method of Refs. 10,11 and 24 outperforms the GIP based method of Refs. 3, 4 and 9. Thus, the NHD method of this paper outperforms competing techniques for Gaussian interference scenarios as well.

3.2.5. Conclusion

This paper provides a statistical characterization of the NHD for nonGaussian interference scenarios, which can be modeled as a spherically invariant random process. A formal goodness-of-fit test based is derived. Performance analysis of the method is considered in some detail using simulated as well as measured data from the MCARM program. Performance comparison of the method with other NHD techniques is also undertaken. The illustrative examples validate the approach taken and confirm the superior performance of the technique of this paper in both Gaussian and nonGaussian interference scenarios. Future work would include extensive performance analysis using simulated and measured data showing the resulting impact on STAP performance. The performance of several STAP algorithms in Gaussian and nonGaussian interference scenarios has been considered in Ref. 26. Future work will address performance of the methods treated in Ref. 26 in conjunction with NHD processing described herein to combat heterogeneous interference scenarios. Preliminary work (not reported here) in this direction reveals that the estimator of Equation 3.24 is rather slow to converge even for moderate system dimension. A related research direction is the performance comparison of model-based parametric STAP methods (which do not require NHD preprocessing) with sample covariance based STAP methods employing NHD preprocessing in dense target environments. Analysis in this direction is undertaken in Ref. 41.

4 A New Technique for Univariate Distribution Approximation of Random Data

R. R. Shah

CONTENTS

4.1. Literature Review 206
4.1.1. Introduction 206
4.1.2. The Kolmogorov–Smirnov Test 206
4.1.2.1. Example 1 207
4.1.2.2. Example 2 209
4.1.3. The Chi-Square Test 210
4.1.3.1. Example 3 214
4.1.4. Quantile–Quantile Plot 215
4.1.4.1. Example 4 216
4.1.5. Probability–Probability Plot 218
4.1.5.1. Example 5 219
4.2. The Ozturk Algorithm 221
4.2.1. Introduction 221
4.2.2. Definitions 222
4.2.3. The Ozturk Algorithm 223
4.2.3.1. Goodness of Fit Test 223
4.2.3.2. Distribution Approximation 239
4.2.3.3. Parameter Estimation 245
4.3. Simulation Results of the Ozturk Algorithm 248
4.3.1. Goodness of Fit Test Results 249
4.3.1.1. The Univariate Gaussian Case 249
4.3.1.2. The Weibull Case 249
4.3.1.3. The Gamma Case 250
4.3.1.4. The Lognormal Case 250
4.3.2. Distribution Approximation Test Results 252

4.4. Conclusions and Suggestions for Future Work..... 255
4.4.1. Conclusions..... 255
4.4.2. Suggestions for Future Work 256

4.1. LITERATURE REVIEW

4.1.1. Introduction

In the analysis of random data, we encounter situations to check various statistical models or "hypotheses" against the data. The usual situation is to test or check a particular probability distribution for consistency in representing data from a certain experiment. The hypothesis that this distribution is the right one is called *null hypothesis*, often denoted by H_0. This hypothesis may have emerged from long experience associated with an experiment and it is necessary to see whether the hypothesis is still correct when there has been some changes in circumstances which call it into question. Alternatively, the hypothesis may be the result of a theoretical analysis or a logical argument and the theory needs to be verified.

A null hypothesis is ordinarily taken to be quite specific. In particular, location, scale, and shape parameters associated with probability density function (PDF) are specified along with the type of distribution. For example, the PDF of the Weibull distribution changes as its shape parameter is changed. Therefore, Weibull distribution having a different shape parameter from the Weibull distribution of the null hypothesis is assumed to be different. All the other distributions, taken together, define what is referred to as alternative hypothesis, denoted by H_1. Therefore, a question that arises in the analysis of random data is: "If the null hypothesis is not true, what are the suitable approximations to the underlying distribution of the data?"

To answer this question, several tests have been proposed and used, with each of these tests having their own strengths and weaknesses. Some may work well for a particular set of density functions, but poorly for others. We focus our attention on four of the most frequently used tests for analyzing random data. Detailed discussions of these tests follow.

4.1.2. The Kolmogorov–Smirnov Test

This test is based on the idea of a "sample distribution function" — a statistic that is the sample version of population distribution function.

Given a sample $(x_1, x_2, \ldots, x_n)$ of size n, the sample distribution function $F_n(x)$ is the cumulative distribution function (CDF) of a discrete PDF where the random variable X assumes the values $x_1, x_2, \ldots, x_n$ with probability of $1/n$. Consequently, the CDF increases in steps of size $1/n$ at each sample value, rising from zero to the left of the smallest x_i to one at the largest x_i.

4.1.2.1. Example 1

Consider a sample[13] of the following five observations.

$$2.22, -0.83, 0.18, 1.18, 2.05$$

The sample distribution function is easily constructed after the sample values are marked on the x-axis: starting at height zero for the values less than -0.83, the CDF is increased successively by steps of height 1/5 at the ordered sample values -0.83, 0.18, 1.18, 2.05, and 2.22. The result is shown in Figure 4.1.

The Kolmogorov–Smirnov test statistic is defined as the maximum absolute vertical deviation D_n, of sample distribution function, $F_n(x)$, from the CDF, $F_0(x)$, specified by the null hypothesis H_0. If the fit is good, D_n is expected to have a small value. On the other hand, if the underlying distribution has a CDF significantly different from $F_0(x)$, it is expected that the fit will be poor and D_n will be large. Consequently, if values of D_n exceed a prespecified value, called the acceptance limit, H_0 is rejected. Fortunately, the distribution of statistic D_n depends only on the sample size and not on the shape of the distribution being tested. The distribution of D_n has been computed under the assumption that the null hypothesis holds. Results of acceptance limits are given in Table 4.1[12,13] for different sample sizes and for various preselected value of

$$\alpha = \Pr(\text{reject } H_0 | H_0 \text{ true}) \tag{4.1}$$

where α is called the significance level. For large values of n, asymptotic formulae are given for acceptance limits. The test consists of the following steps:

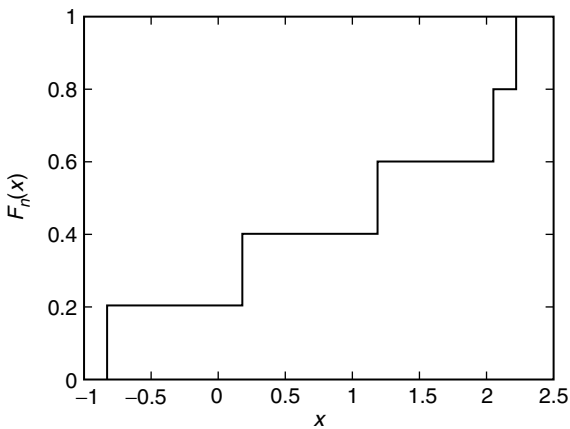

FIGURE 4.1 Sample distribution function (Example 1).

TABLE 4.1
Acceptance Limits for the Kolmogorov–Smirnov Test

	Significance Level				
Sample Size (*n*)	**0.20**	**0.15**	**0.10**	**0.05**	**0.01**
1	0.900	0.925	0.950	0.975	0.995
2	0.684	0.726	0.776	0.842	0.929
3	0.565	0.597	0.642	0.708	0.829
4	0.494	0.525	0.564	0.624	0.734
5	0.446	0.474	0.510	0.563	0.669
6	0.410	0.436	0.470	0.521	0.618
7	0.381	0.405	0.438	0.486	0.577
8	0.358	0.381	0.411	0.457	0.543
9	0.339	0.360	0.388	0.432	0.514
10	0.322	0.342	0.368	0.409	0.486
11	0.307	0.326	0.352	0.391	0.468
12	0.295	0.313	0.338	0.375	0.450
13	0.284	0.302	0.325	0.361	0.433
14	0.274	0.292	0.314	0.349	0.418
15	0.266	0.283	0.304	0.338	0.375
16	0.258	0.274	0.295	0.328	0.391
17	0.250	0.266	0.286	0.318	0.380
18	0.244	0.259	0.278	0.309	0.270
19	0.237	0.252	0.272	0.301	0.361
20	0.231	0.246	0.264	0.294	0.352
25	0.21	0.22	0.24	0.264	0.32
30	0.19	0.20	0.22	0.242	0.29
35	0.18	0.19	0.21	0.23	0.27
40	—	—	—	0.21	0.25
50	—	—	—	0.19	0.23
60	—	—	—	0.17	0.21
70	—	—	—	0.16	0.19
80	—	—	—	0.15	0.18
90	—	—	—	0.14	—
100	—	—	—	0.14	—
Asymptotic formula:	$\frac{1.07}{\sqrt{n}}$	$\frac{1.14}{\sqrt{n}}$	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$

(1) Plot $F_n(x)$ and $F_0(x)$ in the coordinate axes.
(2) By inspection, determine the maximum vertical absolute deviation, given by,

$$D_n = \max_x |F_n(x) - F_0(x)| \tag{4.2}$$

(3) Select a significance level

$$\alpha = \Pr(\text{reject } H_0 | H_0 \text{ true}) \tag{4.3}$$

(4) Accept H_0 if $D_n \leq K$ and reject otherwise.

Note that

$$1 - \alpha = \Pr(\text{accept } H_0 | H_0 \text{ true}) \tag{4.4}$$

Let the CDF of D_n be denoted by $F_{D_n}(x|H_0)$. It follows that

$$\Pr(D_n > K|H_0) = 1 - F_{D_n}(K|H_0) = \alpha \tag{4.5}$$

Consequently, K is the $100(1 - \alpha)$ percentile of $F_{D_n}(x|H_0)$.

4.1.2.2. Example 2

Let the null hypothesis $F_0(x)$ be Gaussian with mean = 32 and standard deviation = 1.8. Consider the following ten observations[13]: 31.0, 31.4, 33.3, 33.4, 33.4, 33.5, 33.7, 34.4, 34.9, 36.2. The corresponding sample distribution function $F_n(x)$, is sketched in Figure 4.2 along with the normal distribution whose mean is 32 and whose standard deviation is 1.8. Assume the significance level is chosen to be $\alpha = .05$. From Figure 4.2 it is determined that the maximum deviation D_n, between the two curves is 0.56. From Table 4.1 the acceptance limit is $K = .409$. Because $D_n > K$, H_0 is rejected.

Although the Kolmogorov–Smirnov test is found to perform quite well even for small sample sizes, it has two principal disadvantages.

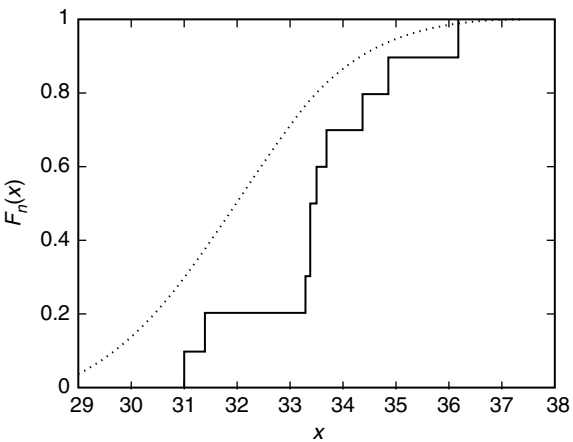

FIGURE 4.2 Sample distribution function (Example 2).

1. To perform the test it is necessary to have *a priori* knowledge about the data in order to be able to specify meaningful null hypotheses.
2. When the null distribution is rejected, no information is provided about which distributions are suitable for approximating the underlying distribution of the data.

4.1.3. The Chi-Square Test

The chi-square (χ^2) test was originally developed for discrete random variables. It is applied to the case of continuous random variables by making a discrete approximation to the continuous probability density function. Because the distribution of the statistic used becomes tractable only as the sample size becomes infinite, the chi-square test should be employed only for large sample sizes.

Consider a null hypothesis with PDF $f_0(x)$ and distribution function $F_0(x)$, as shown in Figure 4.3. Divide the x-axis into k contiguous intervals $E_1, E_2, \ldots, E_k$ from left to right. Note that the ith interval, E_i, consists of the set of points such that

$$a_{i-1} \leq x \leq a_i$$

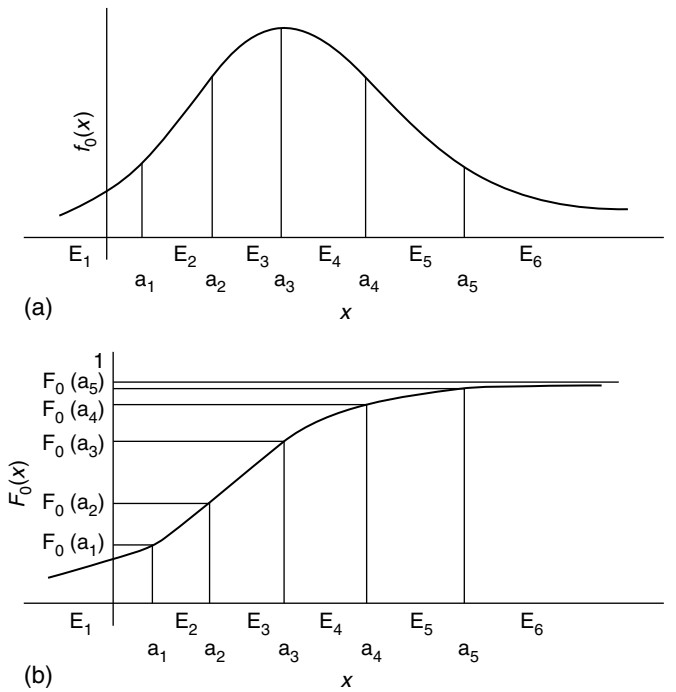

FIGURE 4.3 PDF and distribution function divided into intervals; (a) PDF, (b) distribution function.

where $a_0 = -\infty$ and $a_k = +\infty$. Consequently, a_{i-1} and a_i are the end points of the ith interval, E_i.

Define the probabilities

$$\begin{aligned} p_i &= \Pr(X \in E_i) \\ &= \Pr(a_{i-1} \le X \le a_{i-1}) \\ &= (F_0(a_i) - F_0(a_{i-1})); \qquad i = 1, 2, \ldots, k \end{aligned} \tag{4.6}$$

Observe that p_i is the area under $f_0(x)$ between $x = a_{i-1}$ and $x = a_i$. Also,

$$\sum_{i=1}^{k} p_i = 1 \tag{4.7}$$

Now consider a random experiment consisting of n independent trials. Define

$$f_i = (\text{number of outcomes in } E_i)$$

According to the relative frequency concept,

$$p_i = \lim_{n \to \infty} \frac{f_i}{n} \tag{4.8}$$

Note that

$$\sum_{i=1}^{k} f_i = n \tag{4.9}$$

To test whether the null hypothesis is statistically consistent with the data, the statistic

$$\chi^2 = \sum_{i=1}^{k} \frac{(f_i - np_i)^2}{np_i} \tag{4.10}$$

is evaluated. The null hypothesis is rejected when χ^2 exceeds a critical level M. To determine M, the significance level

$$\alpha = \Pr(\text{reject } H_0 | H_0 \text{ true}) \tag{4.11}$$

is specified. Observe that

$$\alpha = \Pr(\chi^2 > M | H_0) = (1 - F_{\chi^2}(M)) \tag{4.12}$$

where $F_{\chi^2}(\cdot)$ denotes the distribution function of the χ^2 statistic. Since

$$F_{\chi^2}(M) = 1 - \alpha \tag{4.13}$$

M is the $100(1 - \alpha)$ percentile of the χ^2 statistic. M is also referred to as the rejection limit. Values of M are tabulated in Table 4.2.[12,13]

TABLE 4.2
Percentiles of the Chi-Squared Distribution

	p									
Degrees of Freedom	**.01**	**.025**	**.05**	**.10**	**.70**	**.80**	**.90**	**.95**	**.975**	**.99**
1	.000	.001	.004	.016	1.07	1.64	2.71	3.84	5.02	6.63
2	.020	.051	.103	.211	2.41	3.22	4.61	5.99	7.38	9.21
3	.115	.216	.352	.584	3.66	4.64	6.25	7.81	9.35	11.3
4	.297	.484	.711	1.06	4.88	5.99	7.78	9.49	11.1	13.3
5	.554	.831	1.15	1.61	6.06	7.29	9.24	11.1	12.8	15.1
6	.872	1.24	1.64	2.20	7.23	8.56	10.6	12.6	14.4	16.8
7	1.24	1.69	2.17	2.83	8.38	9.80	12.0	14.1	16.0	18.5
8	1.65	2.18	2.73	3.49	9.52	11.0	13.4	15.5	17.5	20.1
9	2.09	2.70	3.33	4.17	10.7	12.2	14.7	16.9	19.0	21.7
10	2.56	3.25	3.94	4.87	11.8	13.4	16.0	18.3	20.5	23.2
11	3.05	3.82	4.57	5.58	12.9	14.6	17.3	19.7	21.9	24.7
12	3.57	4.40	5.23	6.30	14.0	15.8	18.5	21.0	23.3	26.2
13	4.11	5.01	5.89	7.04	15.1	17.0	19.8	22.4	24.7	27.7
14	4.66	5.63	6.57	7.79	16.2	18.2	21.1	23.7	26.1	29.1
15	5.23	6.26	7.26	8.55	17.3	19.3	22.3	25.0	27.5	30.6
16	5.81	6.91	7.96	9.31	18.4	20.5	23.5	26.3	28.8	32.0

17	6.41	7.56	8.67	10.1	19.5	21.6	24.8	27.6	30.2	33.4
18	7.01	8.23	9.39	10.9	20.6	22.8	26.0	28.9	31.5	34.8
19	7.63	8.91	10.1	11.7	21.7	23.9	27.2	30.1	32.9	36.2
20	8.26	9.59	10.9	12.4	22.8	25.0	28.4	31.4	34.2	37.6
21	8.90	10.3	11.6	13.2	23.9	26.2	29.6	32.7	35.5	38.9
22	9.54	11.0	12.3	14.0	24.9	27.3	30.8	33.9	36.8	40.3
23	10.2	11.7	13.1	14.8	26.0	28.4	32.0	35.2	38.1	41.6
24	10.9	12.4	13.8	15.7	27.1	29.6	33.2	36.4	39.4	43.0
25	11.5	13.1	14.6	16.5	28.2	30.7	34.4	37.7	40.6	44.3
26	12.2	13.8	15.4	17.3	29.2	31.8	35.6	38.9	41.9	45.6
27	12.9	14.6	16.2	18.1	30.3	32.9	36.7	40.1	43.2	47.0
28	13.6	15.3	16.9	18.9	31.4	34.0	37.9	41.3	44.5	48.3
29	14.3	16.0	17.7	19.8	32.5	35.1	39.1	42.6	45.7	49.6
30	15.0	16.8	18.5	20.6	33.5	36.2	40.3	43.8	47.0	50.9
40	22.1	24.4	26.5	29.0	44.2	47.3	51.8	55.8	59.3	63.7
50	29.7	32.3	34.8	37.7	54.7	58.2	63.2	67.5	71.4	76.2
60	37.5	40.5	43.2	46.5	65.2	69.0	74.4	79.1	83.3	88.4

The exact distribution of the statistic χ^2 is not simple and depends on the p_i's in the null distribution (H_0) and the distribution under test. It was found that these difficulties markedly disappear as the sample size n becomes large.[12] The distribution of χ^2 under H_0 for large n is approximately one of the family of chi-square distributions, depending on the number of intervals, k, but not on the distribution under test. This family of distributions is characterized by the *number of degrees of freedom*, defined to be $k - 1$. Various percentiles of the chi-square distribution, for selected numbers of degrees of freedom, are given in Table 4.2.

4.1.3.1. Example 3

Table 4.3[13] gives 200 measurements of viscosity. The table gives frequencies f_i corresponding to the ten intervals. Data given in the table can be used to test the hypothesis that the PDF from which they come is normal with mean 32 and standard deviation 1.8. We select the significance level α to be 0.05, so that M equals the 95th percentile of the χ^2 distribution. Table 4.3 shows that k, the number of intervals, is ten. Consequently, the *number of degrees of freedom* given by $k - 1$ equals nine. By extracting the 95th percentile of the χ^2 distribution with nine degrees of freedom from Table 4.2 it is seen that $M = 16.9$.

To illustrate the computations in Table 4.3, the entry for p_4 is obtained from the normal distribution function by means of Equation 4.6. Hence,

$$\Pr(30.05 < X < 31.15) = \Phi\left(\frac{31.15-32}{1.8}\right) - \Phi\left(\frac{30.05-32}{1.8}\right) = 0.1791 \quad (4.14)$$

where $F_0(x) = \Phi[(x-32)/1.8]$ and

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du = \Pr(X \le x) \quad (4.15)$$

TABLE 4.3
Measurements of Viscosity (Example 3)

i	Interval	p_i	$200p_i$	f_i	$(f_i - 200p_i)^2 = z_i$	$z_i/(200p_i)$
1	<27.85	0.0105	2.10	3	0.81	0.3857
2	27.85–28.95	0.0346	6.92	7	0.0064	0.0009
3	28.95–30.05	0.0943	18.86	25	37.65	1.1996
4	30.05–31.15	0.1791	35.82	42	38.10	1.06
5	31.15–32.25	0.2388	47.76	56	67.90	1.422
6	32.25–33.35	0.2161	43.22	30	174.50	4.037
7	33.35–34.45	0.1399	27.98	22	35.80	1.279
8	34.45–35.55	0.0624	12.48	11	2.19	0.175
9	35.55–36.65	0.0195	3.90	3	0.81	0.208
10	>36.65	0.0048	0.96	1	0.0016	0.0017

Values of $\Phi(x)$ are obtained from a table for the standard normal distribution function.

The statistic χ^2 is obtained by summing the entries in the last column with the result $\chi^2 = 10.16$. This does not exceed the 95th percentile of the chi-square distribution with nine "degrees of freedom", i.e., $\chi^2 < M$. The chi-square test, therefore, calls for accepting the null distribution on the basis of given data.

Notice that the chi-square test does not test $F_0(x)$ but only the p_i's. In particular, the natural order of the intervals does not enter the test. Moreover, F_0 is not the only distribution function having the p_i's obtained from F_0. Despite these minor objections, the chi-square test is frequently used in testing a continuous distribution.

Like the Kolmogorov–Smirnov test, the chi-square test suffers from the same disadvantages as mentioned earlier. In addition, it has one more disadvantage, viz., it requires a large sample size to give accurate results.

4.1.4. Quantile–Quantile Plot

A quantile–quantile (Q–Q) plot is a special plot or graphical technique which can be performed to assess the marginal distribution of sample observations. Consider a set of data of size n given by $x_1, x_2, \ldots, x_n$. Let the data be rank-ordered such that $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$. For the jth ordered sample $x_{(j)}$, define

$$p_{(j)} = \frac{j - 1/2}{n} \tag{4.16}$$

where the 1/2 is introduced as a "continuity correction".[14] Let $F_X(x)$ denote the CDF of data. For large enough sample values of n, it then follows that

$$F_X(x_{(j)}) = \Pr(X \leq x_{(j)}) \approx p_{(j)} \tag{4.17}$$

Denote the CDF of the null distribution by $F_0(z)$. The quantile of $F_0(z)$, denoted by $q_{(j)}$, is related to $p_{(j)}$ by

$$F_0(q_{(j)}) = \Pr(Z \leq q_{(j)}) = p_{(j)} \tag{4.18}$$

If the data comes from the same distribution as the null distribution, then

$$x_{(j)} \approx q_{(j)} \tag{4.19}$$

and $x_{(j)}$ can be interpreted as an estimate of the sample quantile.

A Q–Q plot is generated using the following steps:

(1) Collect n data points $x_1, x_2, \ldots, x_n$.
(2) Rank-order the data such that $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$.

(3) Define

$$p_{(j)} = \frac{j - 1/2}{n}; \qquad j = 1, 2, \ldots, n$$

(4) Evaluate the quantile $q_{(j)}$ defined by

$$F_0(q_{(j)}) = p_{(j)}; \qquad j = 1, 2, \ldots, n$$

(5) Plot the pair of points

$$(q_{(j)}, x_{(j)}); \qquad j = 1, 2, \ldots, n$$

When the data comes from the null distribution, the Q–Q plot is likely to approximate a straight line through the origin at 45°.

4.1.4.1. Example 4

A sample[14] of $n = 10$ observations gives the values tabulated in the 2nd column of the Table 4.4. The sample mean and the sample variance are $\hat{m} = .77$ and $\hat{\sigma}^2 = .9414$ respectively. The values of $p_{(j)}$ are computed in the 3rd column. Taking the normal distribution with mean $\hat{m}$ and variance $\hat{\sigma}^2$ as the null distribution, the corresponding quantiles $q_{(j)}$ are evaluated in the 4th column. For example, corresponding to $p_{(9)} = .85$

$$F_0(q_{(9)}) = \Pr(Z \le 1.775) = \int_{-\infty}^{1.775} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(z-0.77)^2/1.8828} dz = 0.85 \qquad (4.20)$$

Consequently, $q_{(9)} = 1.775$.

TABLE 4.4
Observation Table (Example 4)

j	Ordered Observations $x_{(j)}$	Probability Levels $p_{(j)} = (j - 1/2)/n$	Standard Normal Quantiles $q_{(j)}$
1	−1.00	0.05	−0.826
2	−0.10	0.15	−0.235
3	0.16	0.25	0.116
4	0.41	0.35	0.396
5	0.62	0.45	0.649
6	0.80	0.55	0.891
7	1.26	0.65	1.144
8	1.54	0.75	1.424
9	1.71	0.85	1.755
10	2.30	0.95	2.366

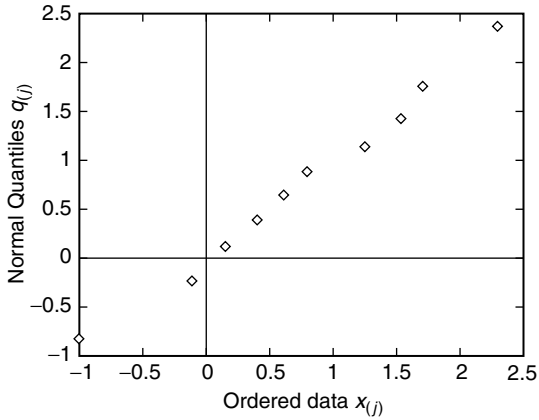

FIGURE 4.4 Q–Q plot (Example 4).

The Q–Q plot for the above data, which is a plot of the ordered data $x_{(j)}$ against the normal quantiles $q_{(j)}$, is shown in Figure 4.4. The pair of points $(q_{(j)}, x_{(j)})$ lie very nearly along a straight line at 45° and it is accepted that these are normally distributed with mean = .77 and variance = .9414.

TABLE 4.5
Critical Points for the Q–Q Plot Correlation Coefficient Test for Normality

	Significance Levels (α)		
Sample Size (n)	**0.01**	**0.05**	**0.10**
5	0.8299	0.8788	0.9032
10	0.8801	0.9198	0.9351
15	0.9126	0.9389	0.9503
20	0.9269	0.9508	0.9604
25	0.9410	0.9591	0.9665
30	0.9479	0.9652	0.9715
35	0.9538	0.9682	0.9740
40	0.9599	0.9726	0.9771
45	0.9632	0.9749	0.9792
50	0.9671	0.9768	0.9809
55	0.9695	0.9787	0.9822
60	0.9720	0.9801	0.9836
75	0.9771	0.9838	0.9866
100	0.9822	0.9873	0.9895
150	0.9879	0.9913	0.9928
200	0.9905	0.9931	0.9942
300	0.9935	0.9953	0.9960

The straightness of the Q–Q plot can be evaluated by calculating the correlation coefficient of the points in the plot. The correlation coefficient for the Q–Q plot is approximated by

$$\hat{r}_{\mathrm{Q}} = \frac{\sum_{j=1}^{n} (x_{(j)} - \hat{m})(q_{(j)} - \hat{q})}{\sqrt{\sum_{j=1}^{n} (x_{(j)} - \hat{m})^2} \sqrt{\sum_{j=1}^{n} (q_{(j)} - \hat{q})^2}} \tag{4.21}$$

where $\hat{q}$ is the sample mean of the quantiles $q_{(j)}$; $j = 1, 2, \ldots, n$. Formally, we select the null hypothesis at a significance level α if $\hat{r}_{\mathrm{Q}}$ exceeds a critical value denoted by M.[14] The values of M have been evaluated for the normal distribution and tabulated in Table 4.5[14] for different sample sizes and significance levels.

For the above example we select $\alpha = .10$. Also, using the information from Table 4.4, it is found that the mean of sample quantiles and standard normal quantiles are, respectively, $\hat{m} = .77$ and $\hat{q} = .768$. Using Equation 4.21, we find that the correlation coefficient, $\hat{r}_{\mathrm{Q}}$, is found to be 0.9943. Referring to Table 4.5, we find that corresponding to $n = 10$ and $\alpha = .10$, the critical point M for the Q–Q plot correlation coefficient test for normality is 0.9351. Because $\hat{r}_{\mathrm{Q}} > .9351$, the hypothesis of normality is accepted.

4.1.5. Probability–Probability Plot

The probability–probability (P–P) plot is another graphical technique that is performed for random data analysis. Just as with the Q–Q plot we consider a set of data of size n given by $x_1, x_2, \ldots, x_n$. The data is rank-ordered such that $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$. Proceeding in line with the Q–Q plot, define for the jth ordered sample, $x_{(j)}$,

$$p_{(j)} = \frac{j - 1/2}{n} \tag{4.22}$$

where the 1/2 is introduced as a "continuity correction".[14] Let $F_X(x)$ denotes the CDF of the data. From the Q–Q plot we know that the $r_{(j)}$'s are the sample quantiles. Denote the CDF of the null distribution by $F_0(z)$. Then $p_{x_{(j)}}$ is defined to be the probability such that

$$F_0(x_{(j)}) = \Pr(Z \leq x_{(j)}) = p_{x_{(j)}} \tag{4.23}$$

If the data comes from the same distribution as the null distribution, it is likely that

$$p_{x_{(j)}} \approx p_{(j)} \tag{4.24}$$

and $p_{(j)}$ can be interpreted as an estimate of the probability $p_{x_{(j)}}$.

A P–P plot is generated using the following steps:

(1) Collect n data points $x_1, x_2, \ldots, x_{n-1}, x_n$.
(2) Rank-order the data such that $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$.
(3) Define

$$p_{(j)} = \frac{j - 1/2}{n}; \qquad j = 1, 2, \ldots, n$$

(4) Evaluate the probability $p_{x_{(j)}}$ defined by

$$F_0(x_{(j)}) = p_{x_{(j)}}; \qquad j = 1, 2, \ldots, n$$

(5) Plot the pair of points

$$(p_{x_{(j)}}, p_{(j)}); \qquad j = 1, 2, \ldots, n$$

When the data comes from the null distribution, the P–P plot is likely to approximate a straight line through the origin at 45°.

4.1.5.1. Example 5

We take the example used with the Q–Q plot and find the P–P plot of the given data. The observations are tabulated in the 2nd column of Table 4.6. Values of $p_{(j)}$ are computed in the 2nd column. The sample mean and the sample variance are $\hat{m} = .77$ and $\hat{\sigma}^2 = .9414$. Finally, taking the normal distribution with mean $\hat{m}$ and variance $\hat{\sigma}^2$ as the null distribution, the corresponding probabilities $p_{x_{(j)}}$ are

TABLE 4.6
Observation Table (Example 5)

j	Ordered Observations $x_{(j)}$	Probability Levels $p_{(j)} = (j - 1/2)/n$	Standard Normal Quantiles $p_{x_{(j)}}$
1	−1.00	0.05	0.0342
2	−0.10	0.15	0.1853
3	0.16	0.25	0.2647
4	0.41	0.35	0.3553
5	0.62	0.45	0.4384
6	0.80	0.55	0.5124
7	1.26	0.65	0.6932
8	1.54	0.75	0.7864
9	1.71	0.85	0.8336
10	2.30	0.95	0.9424

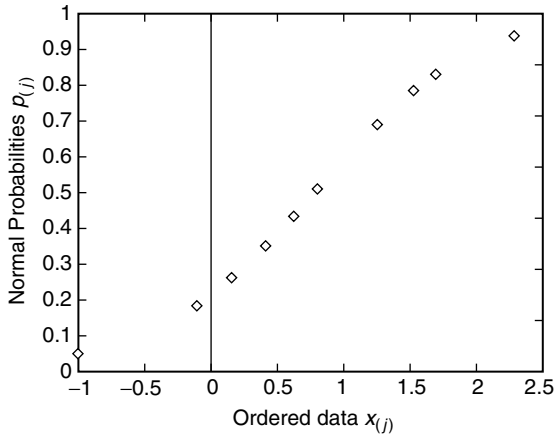

FIGURE 4.5 P–P plot (Example 5).

evaluated in the 4th column. For example, corresponding to $x_{(7)} = 1.26$

$$F_0(x_{(7)}) = \Pr(Z \leq 1.26) = \int_{-\infty}^{1.26} \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-(z-0.77)^2/1.8828} \mathrm{d}z = 0.6932 \quad (4.25)$$

Consequently, $p_{x_{(7)}} = .6932$

The P–P for the above data, which is a plot of values $p_{(j)}$ against the normal probabilities $p_{x_{(j)}}$, is shown in Figure 4.5. The pair of points $(p_{(j)}, p_{x_{(j)}})$ lies very nearly along a straight line at 45° and we accept the notion that these are normally distributed with mean = .77 and variance = .9414.

The straightness of the P–P plot can be evaluated by approximating the correlation coefficient

$$\hat{r}_{\mathrm{P}} = \frac{\sum_{j=1}^{n} (p_{(j)} - \hat{p})(p_{x_{(j)}} - \hat{p}_x)}{\sqrt{\sum_{j=1}^{n} (p_{(j)} - \hat{p})^2} \sqrt{\sum_{j=1}^{n} (p_{x_{(j)}} - \hat{p}_x)^2}} \quad (4.26)$$

where $\hat{p}$ and $\hat{p}_x$ are the sample means of $p_{(j)}$ and $p_{x_{(j)}}$, respectively, with $j = 1, 2, \ldots, n$. Unfortunately, a table for the critical value M for different values of the significance level α, was not found in the literature. If $\hat{r}_{\mathrm{P}}$ is close to unity, $p_{(j)}$ and $p_{x_{(j)}}$ are highly correlated and although a significance level cannot be specified, it is likely that the data can be approximately with null distribution. For this example,

$$\hat{r}_{\mathrm{P}} = 0.9960 \quad (4.27)$$

Because $\hat{r}_{\mathrm{P}} \approx 1$, it is concluded that the data is statistically consistent with the normal distribution having mean $\hat{m} = .77$ and variance $\hat{\sigma}^2 = .9414$.

An attractive property of the Q–Q plot is that it is invariant to a linear transformation. Specifically, the Q–Q plot of a linear function of $x_{(j)}$ is again a straight line at 45°. However, this time the line need not pass through the origin. P–P plots do not have this property. The main drawback of these plots is their weak performances for small and moderate sample sizes. Also, generalization of the Q–Q plot to multivariate distributions is not straightforward. On the other hand P–P plots can be applied to the multivariate situation. Although a statistic exists for evaluating the straightness of the Q–Q plot when the null distribution is standard normal, this statistic is not readily available for other distributions. Consequently, the Q–Q and P–P plots do not readily offer a quantitative goodness of fit test and the decision is mostly made on a subjective basis.

4.2. THE OZTURK ALGORITHM

4.2.1. INTRODUCTION

In testing a null hypothesis for a distributional assumption against an unspecified alternative, no uniformly most powerful or optimal test[2] exists. Because of this, various test procedures have been developed to assess these distributional assumptions. Under certain conditions, (i.e., for a specified null hypothesis, a specified sample size, and a predetermined level of significance) one test procedure may be shown to be more powerful than the other existing procedures. Besides, the power consideration of a given test, computational simplicity, desirable distributional properties of the test statistic and the generality of test procedure are some of the important properties to be considered.

Section 4.1 provided a brief overview of some of these tests. The χ^2 test has been widely used for assessing the distributional assumptions because of its generality and computational simplicity.[2] However, the choice of class intervals for computing the test statistic is arbitrary and the procedure can be used only for large sample sizes. Q–Q plots and P–P plots are among the most widely used graphical procedures for making assessment about the random data. But their performances are weak for small and moderate sample sizes. Also, generalizations of Q–Q plots to the multivariate distributions are not simple.[5,6] As described in Refs. 9, 12, 13, and 17, the Kolmogorov–Smrinov test, which is based on the empirical distribution function of the sample and the null distribution, is also widely used. In fact, comparative studies have shown that, the Kolmogorov–Smirnov statistic has higher power than the χ^2 statistic for many alternatives.[2] There are many other tests such as W test (by Shapiro and Wilk), Anderson's A test,[2] Filliben's correlation coefficient test,[8] etc.

All these tests are goodness of fit tests. Within a certain confidence level, these tests provide information whether a set of random data is statistically consistent with a specified null distribution. However, if the specified distribution is rejected, these tests give no clue about the alternative underlying distribution of data. Thus, we need to have *a priori* knowledge about the random data to be able to use these tests. In practice, a lot of times we have no *a priori* knowledge

about the random signals. For example, the clutter PDF encountered in radar signal processing, is not known *a priori*. Moreover, a lot of these tests require a large number of observations to give accurate results. Obtaining many observations may prove costly in a real-world situation. Consequently, a scheme is necessary, that not only performs the Goodness of Fit test but also approximates the PDF for small number of observations.

A new algorithm based on sample order statistics has been developed in Refs. 1–3 and has been reported in Ref. 10 for univariate distribution approximation. This algorithm has two modes of operation. In the 1st mode, the algorithm performs a goodness of fit test. Specifically, the test determines, to a desired confidence level, whether random data is statistically consistent with a specified probability distribution. In the 2nd mode of operation, the algorithm approximates the PDF underlying random data. In particular, by analyzing the random data and without any *a priori* knowledge, the algorithm identifies, from a stored library of PDFs, the density function which best approximates the data. Estimates of location, scale, and shape parameters of the PDF are provided by the algorithm. The algorithm is typically found to work well for observation sizes of the order of 75 to 100.

In this chapter we present the Ozturk algorithm. It will be demonstrated through examples that the algorithm can be used to test for any distributional assumption (not limited to location-scale family) including univariate and multivariate random variables.

4.2.2. Definitions

Let $F_Y(y)$ denote the PDF of a random variable Y. Consider the linear transformation defined by

$$x = \beta y + \alpha \tag{4.28}$$

The PDF of X is given by

$$f_X(x) = \frac{1}{|\beta|} f_Y\left(\frac{x-\alpha}{\beta}\right) \tag{4.29}$$

where α and β are defined to be the location and scale parameters of $f_X(x)$, respectively. The mean μ_x and the variance σ_x^2 of the random variable x are given by

$$\mu_x = E[x], \qquad \sigma_x^2 = E[(x-\mu_x)^2] \tag{4.30}$$

where E is the expectation operator.

Although the mean and the variance are related to the location and scale parameters, it may be noted that the location parameter is not the mean value and the scale parameter is not the square root of the variance in general. However, for a standard Gaussian PDF $f_Y(y)$, for which the mean is zero and variance is unity, the location parameter is the mean of X and the scale parameter is the standard deviation (square root of the variance) of X.

The coefficient of skewness α_3, and the coefficient of kurtosis α_4, of X are defined to be

$$\alpha_3 = \frac{E[(x-\mu_x)^3]}{\sigma_x^3}, \qquad \alpha_4 = \frac{E[(x-\mu_x)^4]}{\sigma_x^4} \tag{4.31}$$

It is readily shown that α_3 and α_4 are invariant to the values of μ_x and σ_x. For a PDF that is symmetric about the mean, $\alpha_3 = 0$. For the case of the Gaussian distribution, $\alpha_3 = 0$ and $\alpha_4 = 3$.

4.2.3. The Ozturk Algorithm

Any distribution or a family of distributions can be represented by a single point or by a region on an $\alpha_3 - \alpha_4$ plane, respectively, where α_3 is the coefficient of skewness and α_4 is the coefficient of kurtosis (see for example Ref. 15, p. 14). A set of random data can also be represented by a point whose coordinates are given by the sample estimates of α_3 and α_4. Then the best candidate for the underlying true distribution can be identified to be the nearest neighbor distribution on the chart. Although such a chart, based on the coefficient of skewness and kurtosis, provides a useful method of characterizing the distributions, its use is limited because the moments of some distributions do not exist. Other drawbacks of this approach are

1. Estimates of α_3 and α_4 are highly sensitive to extreme observations.
2. Estimates of these moments are highly biased for small sample sizes.[4]
3. The moment estimators are greatly affected by outliers.

This chapter introduces the Ozturk Algorithm which is a general graphical technique that works in two specific modes:

1. In the first mode it performs a formal goodness of fit test for a specified null distribution.
2. In its second mode it provides a graphical representation that gives insight to the distribution which best approximates the data set and provides a way of characterizing the data.

4.2.3.1. Goodness of Fit Test

The goodness of fit test is a complex algorithm that determines whether or not the set of data samples provided to the algorithm are statistically consistent with a specified distribution (the null hypothesis). Using the standard normal distribution with zero mean and unit variance as the *reference distribution*, the standardized sample order statistics are represented by a system of linked vectors. The terminal point of the linked vectors, as well as the shape of their trajectories, are used in determining whether or not to accept the null hypothesis. In its present form the algorithm uses the standard Gaussian distribution as the *reference distribution*. However, any other distribution could be used as the *reference*

distribution. The null hypothesis is the distribution against which the sample data is to be tested. Note that the reference distribution need not be the same as the null distribution.

We begin by introducing several sample order statistics used in the algorithm and thereafter proceed to explain the goodness of fit test procedure

Consider the following three sets of data size n where each corresponding entry in sequential data sets is treated as a random variable:

1. A sample data set
$$X_1, X_2, X_3, \ldots, X_n$$
with mean and standard deviation given by μ_x and σ_x.
2. A null hypothesis data set
$$Z_1, Z_2, Z_3, \ldots, Z_n$$
is generated from any available distribution against which the sample set will be tested. The mean and standard deviation of this data set are defined to be μ_z and σ_z, respectively.
3. A reference distribution data set
$$W_1, W_2, W_3, \ldots, W_n$$
is generated from the standardized Gaussian distribution.

Let $X_{1:n} \leq X_{2:n} \leq \ldots \leq X_{n:n}$ denote the ordered set of samples obtained by ordering X_i; $i = 1, 2, \ldots, n$, where $X_{1:n}$ is the smallest data sample. Similarly, the other two data sets are ordered resulting in the three ordered data sets

$$\begin{aligned} &X_{1:n}, X_{2:n}, X_{3:n}, \ldots, X_{n:n} \\ &Z_{1:n}, Z_{2:n}, Z_{3:n}, \ldots, Z_{n:n} \\ &W_{1:n}, W_{2:n}, W_{3:n}, \ldots, W_{n:n} \end{aligned} \tag{4.32}$$

Define

$$Y_{i:n} = \frac{X_{i:n} - \hat{\mu}_x}{\hat{\sigma}_x}; \qquad i = 1, 2, \ldots, n \tag{4.33}$$

where $\hat{\mu}_x = \sum X_i / n$ is the sample mean and $\hat{\sigma}_x = \sum [(X_i - \mu_x)^2 / (n-1)]^{1/2}$ is the sample standard deviation. These are the standardized order statistics of the sample data. For the null hypothesis, a Monte Carlo simulation consisting of 2000 trials is utilized. The estimate of the expected values of the standardized ith order statistic is defined as

$$\hat{T}_{i:n} = \frac{1}{2000} \sum_{k=1}^{2000} \frac{(Z_{i:n})_k - \hat{\mu}_z}{\hat{\sigma}_z}; \qquad i = 1, 2, \ldots, n \tag{4.34}$$

where $(Z_{i:n})_k$ denotes the ith order statistic from the kth Monte Carlo trial, and $\hat{\mu}_z$ and $\hat{\sigma}_z$ denote the sample mean and sample standard deviation. Also, $\hat{m}_{i:n}$ is defined as the estimate of the expected value of the ith order statistic of

the reference distribution, the standardized Gaussian. Using 2000 Monte Carlo trials.

$$\hat{m}_{i:n} = \frac{1}{2000}\sum_{k=1}^{2000}(W_{i:n})_k; \qquad i = 1, 2, \ldots, n \tag{4.35}$$

where $(W_{i:n})_k$ denotes the ith order statistic from the kth Monte Carlo trial of the reference distribution. When the null hypothesis is the reference distribution, the standardized Gaussian, then

$$\hat{T}_{i:n} \approx \hat{m}_{i:n} \tag{4.36}$$

The goodness of fit test proceeds by joining together two sets of n linked vectors, one for the sample data and one for the null hypothesis. The ith linked vector in each set is characterized by its length and orientation with respect to the horizontal axis. For the sample data, the length of the ith vector a_i, is obtained from the magnitude of the ith standardized order statistic of the data, while its angle or orientation, θ_i is related to $\hat{m}_{i:n}$. More specifically, for the sample data

$$a_i = \frac{|Y_{i:n}|}{n}, \qquad \theta_i = \pi\theta(\hat{m}_{i:n}), \qquad \phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left(-\frac{t^2}{2}\right)\mathrm{d}t \tag{4.37}$$

ϕ_x is the cumulative distribution of the standard Gaussian distribution. We define the sample points Q_k in a two dimensional plane (U, V) by

$$Q_k = (U_k, V_k); \qquad k = 1, 2, \ldots, n \tag{4.38}$$

where $U_0 = V_0 = 0$ and

$$U_k = \frac{1}{k}\sum_{i=1}^{k}|Y_{i:n}|\cos(\theta_i), \quad V_k = \frac{1}{k}\sum_{i=1}^{k}|Y_{i:n}|\sin(\theta_i); \qquad k = 1, 2, \ldots, n \tag{4.39}$$

Similarly, for the null hypothesis the length of the ith vector, b_i, is obtained from the magnitude of the ith standardized order statistic of the null data set. Specifically, for the null data

$$b_i = \frac{|\hat{T}_{i:n}|}{n}, \qquad \theta_i = \pi\phi(m_{i:n}) \tag{4.40}$$

Using the same two dimensional plane, we plot the sample points for the null distribution defined by

$$Q_{0k} = (U_{0k}, V_{0k}); \qquad k = 1, 2, \ldots, n \tag{4.41}$$

where $U_{00} = V_{00} = 0$ and

$$U_{0k} = \frac{1}{k}\sum_{i=1}^{k}|\hat{T}_{i:n}|\cos(\theta_i), \quad V_{0k} = \frac{1}{k}\sum_{i=1}^{k}|\hat{T}_{i:n}|\sin(\theta_i); \qquad k = 1, 2, \ldots, n \tag{4.42}$$

Note that the angle θ remains the same for both sets of linked vectors. However, the magnitude of the linked vector for the sample data is a_i whereas, it is b_i for the null distribution. The angle θ_i is solely dependent on the reference distribution while the magnitude $|Y_{i:n}|$ and $|\hat{T}_{i:n}|$ are solely dependent on sample data and null data sets, respectively. In particular, for the ith sample-linked vector, a_i is dependent on the standardized ith order statistic of the sample data set whereas, for the ith-linked vector of null hypothesis, b_i is dependent on the estimate of the expected value of the ith standardized order statistic of the Monte Carlo simulation of null distribution.

Although $Y_{i:n}$ and $\hat{T}_{i:n}$ are ordered statistics from the smallest to the largest, it may be noted that the magnitude of $Y_{i:n}$ and $\hat{T}_{i:n}$ are not. In fact, with increasing i, $|Y_{i:n}|$ and $|\hat{T}_{i:n}|$ begin large, decrease to approximately zero, and then increase again.

The ith sample and null-linked vectors, respectively, are drawn by joining the points (Q_i, Q_{i-1}) and $(Q_{0i}, Q_{0(i-1)})$. It should be noted that the Q_n and Q_{0n}, given in Equation 4.38 and Equation 4.41 represent the terminal points respectively, of the linked vectors defined above. Figure 4.6 shows the two sets of linked vectors obtained when the sample and null data sets are obtained from the Gaussian distribution with $n = 6$ and $n = 50$. The solid curves in Figure 4.6 show the linked vector for the sample distribution while the dashed curves show the ideal linked vector for the null distribution. When the length n of the data set is large (of the order of 50 points), the linked vector is a smooth arc, as seen in Figure 4.6.

For a typical set of ordered data samples, drawn from null distribution, it is reasonable to expect that the sample-linked vectors would follow the null-linked vectors closely. If the ordered set of samples is not from the null distribution, then the sample-linked vectors are not expected to follow the path of the null-linked vectors closely. Hence, the procedure provides visual information about how well the ordered set of data fit the null distribution. However this is not an *ad hoc* statistical procedure. As shall be seen later on, we do construct test statistics to present a formal way of performing the goodness of fit test to determine whether the data set is statistically consistent with the null hypothesis.

4.2.3.1.1. Properties of the Test Statistic Q_{0n}

An important property of the Q_{0n} statistic is that it is invariant under linear transformation. In particular, we consider the standardization used in Equation 4.33. Let $S_i = cX_i + d$ where c and d are constants. Let μ_s and σ_s denote the mean and standard deviation of the samples, S_i, respectively. Then, it is readily shown that

$$\left|\frac{X_i - \mu_x}{\sigma_x}\right| = \left|\frac{S_i - \mu_s}{\sigma_s}\right|$$

The invariance property follows as a consequence. The advantage of this property is that the PDF of $Q_{0n} = (U_{0n}, V_{0n})$ for a given sample set and reference distribution depends only on the sample size n and is unaffected by the location

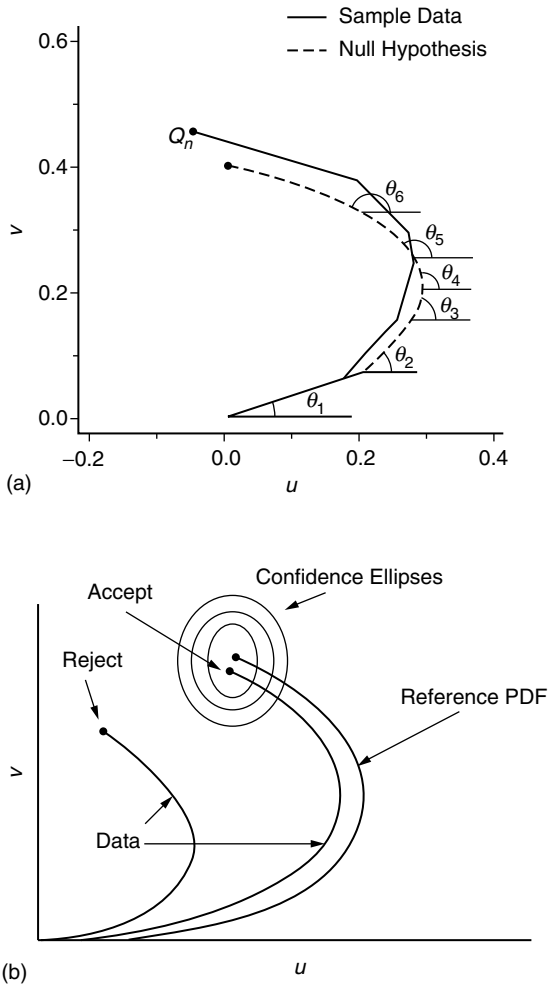

FIGURE 4.6 Linked vectors $Q_K = (U_k, V_k)$; (a) sampled data (solid line) and null hypothesis (dashed line); (b) confidence ellipses.

and scale parameters. The distributional properties of this statistic for testing normality is studied by Ozturk and Dudewicz in Ref. 3.

The exact sampling distribution of Q_{0n} is usually difficult to obtain. However, the empirical distribution of the test statistic Q_{0n} was obtained via Monte Carlo experimentation by Ozturk and Dudewicz in Ref. 3. Using the means, variances, and coefficients of skewness and kurtosis of U_{0n} and V_{0n} based on 50,000 samples for values of n from 3 till 100, they found that the distributions of U_{0n} and V_{0n} approach the normal distribution even for moderate sample sizes. The distributional properties of the statistic $Q_{0n} = (U_{0n}, V_{0n})$ for testing normality

is studied by Ozturk and Dudewicz in Ref. 3. Some of the empirical results obtained by them for the statistic Q_{0n} for standard normal distribution for $3 \leq n \leq 100$ are given below.

$$\begin{aligned} E(U_{0n}) &= 0 \\ E(V_{0n}) &= \mu_v \approx 0.326601 + \frac{0.412921}{n} \\ E(U_{0n}V_{0n}) &\approx 0 \end{aligned} \tag{4.43}$$

$$\begin{aligned} \mathrm{Var}(U_{0n}) &= \sigma_u^2 \approx \frac{0.02123}{n} + \frac{0.01765}{n^2} \\ \mathrm{Var}(V_{0n}) &= \sigma_v^2 \approx \frac{0.04427}{n} - \frac{0.0951}{n^2} \end{aligned} \tag{4.44}$$

Also, it was found empirically, for $n > 10$, that, U_{0n} and V_{0n} are approximately bivariate normal.

An interesting property of this algorithm is that any one of the points Q_{0k}; $k = 1, 2, \ldots, n$, or a selected group of these points can be used as a test statistic to establish a formal test. The algorithm in its present form proposes the general statistic Q_{0n} as the test statistic for testing the null hypothesis.

4.2.3.1.2. Basic Concept of the Confidence Contours

The algorithm provides quantitative information as to how consistent the sample data set is with the null hypothesis distribution by the use of the confidence contours. An example of these contours is shown in Figure 4.6. If the end point of the sample data linked vector curve falls within one or more of these contours, then the sample data set is said to be statistically consistent with the null hypothesis at a confidence level based on the confidence contours. If the sample data set is truly consistent with the null hypothesis, note that the sample-linked vector is likely to closely follow the null-linked vector.

Now, consider the linked vector for the null hypothesis which is based on the standardized expected values of order statistics, Z, for 2000 Monte Carlo simulations. The test statistic Q_{0n}, found by computing the expected value of 2000 end points of the 2000 linked vectors provided by the Monte Carlo simulation, is random. The coordinates of Q_{0n}, U_{0n}, and V_{0n}, may or may not be bivariate Gaussian.

When U_{0n} and V_{0n} are bivariate Gaussian, the confidence contours of null hypothesis are readily determined. A three dimensional bell shaped bivariate Gaussian curve is fitted to the 2000 end points arising from the Monte Carlo simulation. The elliptical contours of this distribution are plotted for various parameters of the significance level α (e.g., 0.01, 0.05, 0.1) where α is defined as the conditional probability that Q_{0n} falls outside the specified ellipse, given that the data comes from the null distribution. $(1 - \alpha)$ is called the confidence level and the corresponding contour is called the $100(1 - \alpha)\%$ confidence contour.

Note that $(1-\alpha)$ is the conditional probability that the Q_{0n} falls inside the specified ellipse, given that the data comes from null distribution.

This could be done for any of the n points of the ordered statistic, Z, along the null-linked vector. Thus, more than one set of confidence contours could be created if there is more than one test statistic. Then, if the sample data is truly consistent with null hypothesis, the sample data-linked vector is likely to pass through a series of confidence contours, determined from the distributions of the test statistics. However, it was found to be unnecessary to clutter-up the graphics with so many contours since it can always be visually detected whether or not the linked vectors are closely following the same trajectory. The option of using more than one test statistic is provided in the algorithm.

Note that the average value of the test statistic, Q_{0n}, of null distribution is at the center of the contours. Thus, the closer the end point of the sample data linked-vector is to the center of confidence contour, the more likely it is that the sample data is coming from null hypothesis. As the significance level decreases, the confidence level increases and the probability that Q_{0n} will fall within the corresponding ellipse will also increase. This results in the fact that the size of the confidence contours increases as the confidence level increases.

For a given sample size, n, the ith angle of any linked vector depends solely on the reference distribution which remains unchanged throughout. Consequently, for a given value of sample size, n, and for a given null hypothesis, values for the magnitude and angle of the points (U_{0k}, V_{0k}) on the null-linked vector, $k = 1, 2, \ldots, n$, may be tabulated. This table, which is dependent on n and the null hypothesis, could be stored and recalled when desired. This can significantly reduce the computational requirements.

4.2.3.1.3. Determining Confidence Contours

As described earlier, the confidence contours are contours from the bivariate PDF of the end point coordinates used to determine the test statistic Q_{0n}. These 2000 end points are obtained from the Monte Carlo simulation. Plotting confidence contours is usually not easy when the joint distribution is not bivariate normal. Further, in order to analytically determine the confidence contours, the joint PDF of U_{0n} and V_{0n} must be known.[4] However, it is difficult to analytically determine this joint PDF. Consequently, a normality transformation is made on the end point coordinates U_{0n} and V_{0n} to obtain statistics $r_{0u} = \psi_1(U_{0n})$ and $r_{0v} = \psi_2(V_{0n})$ where $\psi_1(\cdot)$ and $\psi_2(\cdot)$ are functions operating on U_{0n} and V_{0n}, respectively. A family of distributions called the *Johnson System* is used to perform the transformation on U_{0n} and V_{0n} to obtain a bivariate normal distribution.

The Johnson system of distributions is a flexible family of distributions having four parameters. This system is used to summarize a set of data by means of a mathematical function that will fit the data. The system proposed by Johnson contains three families of distributions which are obtained by

transformations of the form

$$R = \gamma + \eta f_i(G; \lambda, \epsilon); \qquad i = 1, 2, 3 \tag{4.45}$$

where R is a standard normal variable and G is the random variable on which the transformation is performed. γ, η, λ, and ϵ are four parameters of the Johnson system of distributions. In particular, let

$$R_{0u} = \gamma_1 + \eta_1 f_i(U_{0n}; \lambda_1, \epsilon_1)\, R_{0v} = \gamma_2 + \eta_2 f_i(V_{0n}; \lambda_2, \epsilon_2); \quad i = 1,2,3 \tag{4.46}$$

where f_i; $i = 1,2,3$ represent the following three functions suggested by Johnson:

$$f_1(g; \lambda, \epsilon) = \sinh^{-1}\left(\frac{g-\epsilon}{\lambda}\right) \tag{4.47}$$

denotes the S_U distribution.

$$f_2(g; \lambda, \epsilon) = \ln\left(\frac{g-\epsilon}{\lambda+\epsilon-g}\right), \quad \epsilon \le g \le \epsilon + \lambda \tag{4.48}$$

denotes the S_B distribution, and

$$f_3(g; \lambda, \epsilon) = \ln\left(\frac{g-\epsilon}{\lambda}\right), \quad g > \epsilon \tag{4.49}$$

denotes the S_L distribution.

Note that $f_i(g; \lambda, \epsilon)$, $i = 1, 2, \ldots, 3$, are single-valued monotonically increasing functions for the allowed ranges of g. S_L is, in essence, a three parameter distribution since the parameter λ can be eliminated by letting $\gamma^* = \gamma - \eta \ln \lambda$ so that $r = \gamma^* + \eta \ln(g - \epsilon)$. S_B is a distribution bounded on $(\epsilon, \epsilon + \lambda)$ and the S_U is an unbounded distribution. In a plot of the 3rd and 4th order standardized moments where $\sqrt{\alpha_3}$ is plotted vs. α_4, the chosen functions are such that the S_L distributions form a curve dividing the $(\sqrt{\alpha_3}, \alpha_4)$ plane in two regions. The S_B distributions lie in one of the regions and the S_U lie in the other.

While using this system of transformations, the first step is to determine which of the three families should be used for performing the normality transformation. A possible procedure is to compute the sample estimate of the standardized moments, viz., the coefficients of skewness and kurtosis and choose the distribution depending on which of the two regions contains the computed point. However, as described at the beginning of the chapter, this method has major drawback. Consequently, another procedure is used to determine the family of distributions and perform transformations. It is a simple selection rule, which is a function of four percentiles to select one of the three families and to give estimates of the parameters for all the families. It was developed by Slifker J. and Shapiro S. in Ref. 4.

The idea of the selection rule is to try and find a property of the transformation given in Equation 4.45 and use it to select an appropriate member of the Johnson family to approximate a set of data. According to Slifker J. and

Shapiro S., heuristically there must be some relationship concerning the distances in the tails vs. distances in the central portion of the distribution, which could be used to distinguish between bounded and unbounded cases. This led to the following formalization.

Consider any one of the transformations described by Equation 4.45. Choose any fixed value of $r > 0$ of a standard normal variate. Then the points $\pm r$ and $\pm 3r$ divide a horizontal axis into three intervals of equal length given by $(-3r, -r)$, $(-r, r)$, $(r, 3r)$. Let g_{3r}, g_r, g_{-r}, and g_{-3r} be the values corresponding to $3r$, r, $-r$ and $-3r$, under the transformation given in Equation 4.45, respectively. Let

$$m = g_{3r} - g_r \qquad l = g_{-r} - g_{-3r} \qquad p = g_r - g_{-r} \tag{4.50}$$

Because $f_i(g; \lambda, \epsilon)$; $i = 1, 2, \ldots 3$, are single-valued monotonically increasing functions for the allowed ranges of g, it is readily seen that m, l, and p are all greater than zero. For a bounded symmetrical Johnson distribution, it was hypothesized that the distances m and l between each of the outer and inner points would be smaller than the distance p between the two inner points. The converse would be true for the unbounded case. This led to the following more general results:

$$\begin{aligned} &(i)\ \frac{ml}{p^2} > 1 \text{ for any } S_{\mathrm{U}} \text{ distribution;} \\ &(ii)\ \frac{ml}{p^2} < 1 \text{ for any } S_{\mathrm{B}} \text{ distribution;} \\ &(iii)\ \frac{ml}{p^2} = 1 \text{ for any } S_{\mathrm{L}} \text{ distribution.} \end{aligned} \tag{4.51}$$

These properties are proven in Appendix C and can be used to discriminate among the three families.

4.2.3.1.3.1. Selection Procedure. The selection procedure consists of the following steps:

1. Choose a fixed value of $r > 0$. This choice should be motivated by the number of data points. In general, for moderate sized data sets, a value of r less than one should be chosen.[4] Any choice of r greater than one would make it difficult to estimate the percentile of $\pm 3r$. A *typical* choice is to use a value of r close to 0.5 such as 0.524. This would make $3r = 1.572$ and these points correspond to the 70th and the 94.2th percentiles of the standard normal distribution, respectively. However, the larger the number of data points, the larger the value of r that can be selected. In the *Ozturk algorithm* r is chosen to be 0.775449.

2. Determine from a table for the normal distribution the probability $P_a = \Pr(R \leq a)$, where a is taken to be either $3r$, r, $-r$ or $-3r$. For example, if $r = 0.5$ then $P_{0.5} = \Pr(R \leq 0.5) = 0.6915$.
3. Determine integer values of k_a such that

$$P_a \approx \frac{k_a - \frac{1}{2}}{n} \tag{4.52}$$

where

$$k_a = \left[nP_a + \frac{1}{2} \right] \tag{4.53}$$

$[\cdot]$ denotes the closest integer, and $a = 3r, r, -r, -3r$.
4. Obtain n observations of the random variable G, where G is related to the random variable R through Equation 4.45. Order these observations from the smallest to the largest and denote the kth ordered observation by g^k.
5. Let

$$g_a = g^{k_a} \tag{4.54}$$

where $a = 3r, r, -r, -3r$. The connections between g_a, k_a, and P_a are explained in Appendix D.
6. From the values of g_a obtained in step 5, compute the distances m, l, and p according to Equation 4.50.
7. Use the criteria in Equation 4.51 to select the appropriate member of the family of distributions.

Because the g^i's are continuous random variables, the probability is zero that $(ml/p^2) = 1$. Thus, choice of the S_L distribution requires that ml/p^2 fall within some small prespecified tolerance interval around one.

After completion of the selection process, the next step is to estimate parameters of the distribution selected. Estimation of the parameters is accomplished by using formulae given. These allow the estimates to be simply calculated by means of a scientific hand calculator. The formulae for the estimates are given, in terms of the chosen values of r and the computed values of m, l, and p. Derivations of these formulae are provided in Appendix C.

Note that the following formulae express the parameter values as functions of m, l, and p which in turn are functions of g_{3r}, g_r, g_{-r}, and g_{-3r}. In practice, the corresponding parameter estimates are computed based on the ordered sample values, $g^{k_a}; a = 3r, r, -r, -3r$.

(i) Johnson S_U distribution

$$r = \gamma + \eta \sinh^{-1}\left(\frac{q - \epsilon}{\lambda} \right) \tag{4.55}$$

Parameter estimates for Johnson S_U distribution

$$\begin{aligned}
\eta &= \frac{2r}{\cosh^{-1}\left[\frac{1}{2}\left(\frac{m}{p}+\frac{l}{p}\right)\right]}; \quad (\eta > 0)\\
\gamma &= \eta \sinh^{-1}\left[\frac{\frac{l}{p}-\frac{m}{p}}{2\left(\frac{m}{p}\frac{l}{p}-1\right)^{1/2}}\right];\\
\lambda &= \frac{2p\left(\frac{m}{p}\frac{l}{p}-1\right)^{1/2}}{\left(\frac{m}{p}+\frac{l}{p}-2\right)\left(\frac{m}{p}\frac{l}{p}+2\right)^{1/2}}; \quad (\lambda > 0)\\
\epsilon &= \frac{g_r+g_{-r}}{2}+\frac{p\left(\frac{l}{p}-\frac{m}{p}\right)}{2\left(\frac{m}{p}+\frac{l}{p}\right)-2}
\end{aligned} \tag{4.56}$$

(ii) Johnson S_B distribution

$$r = \gamma + \eta \ln\left(\frac{g-\epsilon}{\lambda+\epsilon-g}\right) \tag{4.57}$$

Parameter estimates for Johnson S_B distribution

$$\begin{aligned}
\eta &= \frac{r}{\cosh^{-1}\left(\frac{1}{2}\left[\left(1+\frac{p}{m}\right)\left(1+\frac{p}{l}\right)\right]^{1/2}\right)}; \quad (\eta > 0)\\
\gamma &= \eta \sinh^{-1}\left[\frac{\left(\frac{p}{l}-\frac{p}{m}\right)\left\{\left(1+\frac{p}{m}\right)\left(1+\frac{p}{l}\right)-4\right\}^{1/2}}{2\left(\frac{p}{m}\frac{p}{l}-1\right)}\right];\\
\lambda &= \frac{p\left[\left\{\left(1+\frac{p}{m}\right)\left(1+\frac{p}{l}\right)-2\right\}^2-4\right]^{1/2}}{\frac{p}{m}\frac{p}{l}-1}; \quad (\eta > 0)\\
\epsilon &= \frac{g_r-g_{-r}}{2}-\frac{\lambda}{2}+\frac{p\left(\frac{p}{l}-\frac{p}{m}\right)}{2\left(\frac{p}{m}\frac{p}{l}-1\right)}
\end{aligned} \tag{4.58}$$

(iii) Johnson S_L distribution

$$r = \gamma^* + \eta \ln(g - \epsilon) \tag{4.59}$$

Parameter estimates for Johnson S_L distribution

$$\eta = \frac{2r}{\ln \frac{m}{p}}$$

$$\gamma^* = \eta \ln \left[\frac{\frac{m}{p} - 1}{p\left(\frac{m}{p}\right)^{1/2}} \right] \tag{4.60}$$

$$\epsilon = \frac{g_r + g_{-r}}{2} - \frac{p}{2} \frac{\frac{m}{p} + 1}{\frac{m}{p} - 1}$$

Note that the values of the parameters above are presented in such a way as to emphasize their dependences on the ratios m/p and l/p for the S_U distribution and on p/m and p/l for the S_B distribution. For the S_L distribution, we see from Equation 4.51 that $(l/p) = (m/p)^{-1}$. Thus, the formulae for the S_L distribution parameters are given solely in terms of the single ratio m/p.

Example. We consider a set of data representing the resistor values in a very large-scale integrated (VLSI) circuit. The data and the observed frequencies are shown in Table 4.7. We choose the value of r to be one. Thus, the four values assumed by a are $+3$, 1, -1, and -3. From the table for the normal distribution, the probabilities, $P_a = \Pr(R \leq a)$, for $a = 3$, 1, -1, and -3 are found to be 0.9986, 0.8413, 0.1587, and 0.0014, respectively.

First, consider $a = 3$, for which $P_3 = .9986$. The value of the order number k_3 is given by

$$k_3 = \left[nP_3 + \frac{1}{2} \right] = [(9440)(0.9986) + 0.5] = 9427 \tag{4.61}$$

where $[\cdot]$ denotes the closest integer value. If the raw data were available, we would simply let g_3 equal to the 9427th ordered sample, g^{9427}. However, because the raw data has been grouped into the intervals tabulated in Table 4.7, the value of g^{9427} is unknown. Consequently, interpolation is used to estimate a value for g^{9427}.

Note that the 9427th ordered observation falls in the interval (16.25, 16.75). The probabilities that the resistor values are less than or equal to 16.25 and

TABLE 4.7
Resistor Values

Resistor Interval (in kΩ)	Observed Frequencies
<9.25	—
9.25–9.75	1
9.75–10.25	7
10.25–10.75	18
10.75–11.25	36
11.25–12.25	70
11.75–12.25	115
12.25–12.75	199
12.75–13.25	437
13.25–13.75	929
13.75–14.25	1787
14.25–14.75	2294
14.75–15.25	2082
15.25–15.75	1129
15.75–16.25	275
16.25–16.75	55
16.75–17.25	6
>17.25	—
Total	9440

16.75 kΩ, respectively, are given by

$$\begin{aligned} \Pr(G \leq 16.25) &= \frac{9440 - 6 - 55}{9440} = 0.9935 \\ \Pr(G \leq 16.75) &= \frac{9440 - 6}{9440} = 0.9994 \end{aligned} \tag{4.62}$$

Thus, by the method of interpolation,

$$\frac{16.75 - 16.25}{0.9994 - 0.9935} = \frac{g^{9427} - 16.25}{0.9986 - 0.9935} \tag{4.63}$$

The value of g^{9427} is found to be $16.25 + 0.439 = 16.689$. Setting g_3 equal to g^{9427}, it follows that $g_3 = 16.689$. Values of g_1, g_{-1}, and g_{-3} are found in a similar manner. In summary,

$$g_3 = 16.689, \quad g_1 = 15.242, \quad g_{-1} = 13.581, \quad g_{-3} = 10.409 \tag{4.64}$$

Consequently,

$$\begin{aligned} p &= g_1 - g_{-1} = 1.661 \\ m &= g_3 - g_1 = 1.447 \\ l &= g_{-1} - g_{-3} = 3.172 \end{aligned} \tag{4.65}$$

yielding

$$\frac{ml}{p^2} = 1.664 \tag{4.66}$$

Because the value of (ml/p^2) is found to be significantly greater than 1, it is decided from Equation 4.51 that an S_U distribution is appropriate for transformation. The formulae given in Equation 4.56 are used to obtain the parameter values. Thus,

$$\begin{aligned} \eta &= \frac{2(1)}{\cosh^{-1}\left[\frac{1}{2}(0.871+1.910)\right]} = 2.333 \\ \gamma &= 2.333 \sinh^{-1}\left[\frac{1.910-0.871}{2\sqrt{1.910(0.871)-1}}\right] = 1.402 \\ \lambda &= \frac{2(1.661)\sqrt{(1.910)(0.871)-1}}{(0.871+1.910-2)\sqrt{0.871+1.910+2}} = 1.585 \\ \epsilon &= \frac{15.242+13.581}{2} + \frac{1.661(1.910-0.871)}{2(1.910+0.871-2)} = 15.516 \end{aligned} \tag{4.67}$$

The transformation equation, therefore, becomes

$$r = 1.402 + 2.333 \sinh^{-1}\left(\frac{g-15.516}{1.585}\right) \tag{4.68}$$

where r is a standard normal variable and g is a random variable corresponding to the resistors values.

Once the transformation equations have been obtained for the end point coordinates U_{0n} and V_{0n}, they are applied to the end point data arising from the 2000 Monte Carlo simulations to generate the standard bivariate normal random variables R_{0u} and R_{0v}, respectively. If a type j transformation, $j = 1, 2, 3$, is used, the original data is said to have a Johnson type j-distribution. In practice, U_{0n} and V_{0n} need not have the same distributions (i.e., U_{0n} may be of type i whereas V_{0n} may be of type j and $i \neq j$). An estimate of the correlation coefficient between R_{0u} and R_{0v} is given by

$$\hat{\rho} = \frac{1}{1999} \sum_{i=1}^{2000} \left[\frac{(R_{0u_i} - \hat{\mu}_{r_{0u}})(R_{0v_i} - \hat{\mu}_{r_{0v}})}{\hat{\sigma}_{r_{0u}} \hat{\sigma}_{r_{0v}}} \right] \tag{4.69}$$

where $\hat{\mu}_{r_{0u}}, \hat{\mu}_{r_{0v}}$, and $\hat{\sigma}_{r_{0u}}, \hat{\sigma}_{r_{0v}}$ are the sample means and variances of the 2000 transformed statistics R_{0u} and R_{0v}, respectively.

Because R_{0u} and R_{0v} are bivariate standard normal random variables, their joint PDF can be written as

$$f_{R_{0u},R_{0v}}(r_{0u}, r_{0v}) = \frac{1}{2\pi\sqrt{1-\hat{\rho}^2}} \exp\left(-\frac{t}{2}\right) \tag{4.70}$$

where

$$t = \frac{1}{1-\hat{\rho}^2}(r_{0u}^2 + r_{0v}^2 - 2\hat{\rho} r_{0u} r_{0v}) \tag{4.71}$$

Let $t = t_0$. Then the equation

$$t_0 = \frac{1}{1-\hat{\rho}^2}(r_{0u}^2 + r_{0v}^2 - 2\hat{\rho} r_{0u} r_{0v}) \tag{4.72}$$

is that of an ellipse in the r_{0u}, r_{0v} plane for which

$$f_{R_{0u},R_{0v}}(r_{0u}, r_{0v}) = \frac{1}{2\pi\sqrt{1-\hat{\rho}^2}} \exp\left(-\frac{t_0}{2}\right) \tag{4.73}$$

Points that fall within the ellipse correspond to those points in the r_{0u}, r_{0v} plane for which

$$f_{R_{0u},R_{0v}}(r_{0u}, r_{0v}) > \frac{1}{2\pi\sqrt{1-\hat{\rho}^2}} \exp\left(-\frac{t_0}{2}\right) \tag{4.74}$$

Let α be defined as the probability that r_{0u} and r_{0v} fall outside the ellipse given that the data comes from the null hypothesis. It follows that

$$\alpha = \Pr(T > t_0) \tag{4.75}$$

Note that the bivariate normal distribution is a special case of the spherically invariant random vector (SIRV) where the characteristic PDF is given by Ref. 11.

$$f_s(s) = \delta(s-1) \tag{4.76}$$

The PDF of an N dimensional SIRV involves the same quadratic form t that arises in the N dimensional multivariate Gaussian PDF. For an SIRV, Rangaswamy[11] shows that the PDF of t is

$$f_T(t) = \frac{1}{2^{N/2}\Gamma\left(\frac{N}{2}\right)} t^{[(N/2)-1]} h_N(t); \qquad 0 \le t \le \infty \tag{4.77}$$

where $h_N(t)$ is a monotonically decreasing function given by

$$h_N(t) = \int_0^\infty s^{-N} \exp\left(-\frac{t}{2s^2}\right) f_S(s) \mathrm{d}s \tag{4.78}$$

Substituting Equation 4.76 into Equation 4.78 results in

$$h_N(t) = \int_0^\infty s^{-N} \exp\left(-\frac{t}{2s^2}\right) \delta(s-1) \mathrm{d}s = \exp\left(-\frac{t}{2}\right) \tag{4.79}$$

For the bivariate case, $N = 2$. Consequently, Equation 4.77 reduces to

$$f_T(t) = \frac{1}{2} \exp\left(-\frac{t}{2}\right); \qquad 0 \le t \le \infty \tag{4.80}$$

Hence,

$$\alpha = \Pr(T > t_0) = \int_{t_0}^\infty \frac{1}{2} \exp\left(-\frac{t}{2}\right) \mathrm{d}t = \exp\left(-\frac{t_0}{2}\right) \tag{4.81}$$

Consequently, $t_0 = -2\ln(\alpha)$. Thus Equation 4.72 becomes

$$\frac{1}{1-\hat{\rho}^2}(r_{0u}^2 + r_{0v}^2 - 2\hat{\rho} r_{0u} r_{0v}) = -2\ln(\alpha) \tag{4.82}$$

The contour equation

$$r_{0u}^2 - 2\hat{\rho} r_{0u} r_{0v} + r_{0v}^2 = -2(1-\hat{\rho}^2)\ln(\alpha) \tag{4.83}$$

which is the equation of an ellipse, and is used to determine the $100(1-\alpha)\%$ confidence contour. This is also shown in Refs. 15 and 16. When the statistics R_{0u} and R_{0v} are uncorrelated, the correlation coefficient is zero and Equation 4.83 becomes

$$r_{0u}^2 + r_{0v}^2 = -2\ln(\alpha) \tag{4.84}$$

which is the equation of a circle. Also Equation 4.83 degenerates into a line as the correlation coefficient approaches ± 1.

In the *Ozturk Algorithm*, an *inverse* Johnson Transformation is applied to the points for the confidence ellipses. The locus of resulting points obtained, is then plotted to obtain the corresponding confidence contours in the $U-V$ plane. Consequently, these confidence contours are not necessarily ellipsoidal.

The confidence contours are plotted for a given sample size n. These are then used to make a visual as well as computational test of the null hypothesis. If the terminal point, Q_n, of the sample data, falls inside the contour, the data is declared as being consistent with the null hypothesis with confidence level $(1-\alpha)$. Otherwise the null hypothesis is rejected with a significance level α. Figure 4.7 shows the linked vectors and the confidence contours when the null distribution is standard normal and the sample data size is 100. From the figure it is seen that the sample data is statistically consistent with the null hypothesis at confidence

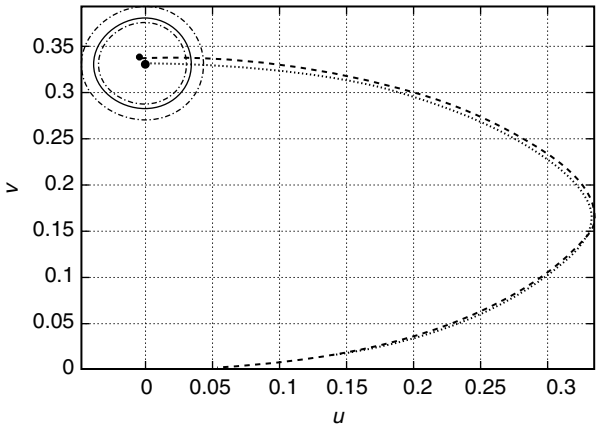

FIGURE 4.7 Goodness of fit chart. Sample data (dashed line) consistent with null hypothesis (dotted line). (90, 95, 99% confidence contours from the innermost to the outermost, respectively.)

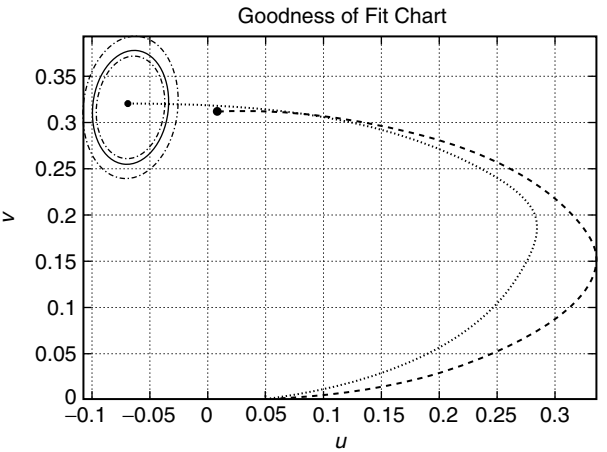

FIGURE 4.8 Goodness of fit chart. Sample data (dashed line) not consistent with null hypothesis (dotted line). (90, 95, 99% confidence contours from the innermost to the outermost, respectively.)

level of 90%. Figure 4.8 shows a case where the sample data is not consistent with the null hypothesis at a significance level of 1%.

4.2.3.2. Distribution Approximation

The *distribution approximation* procedure is simply an extension of the goodness of fit test. Following a similar approach to that outlined in the section for

the goodness of fit test, random samples are generated from a library of different univariate probability distributions. In the goodness of fit test, the statistic $Q_{0n} = (U_{0n}, V_{0n})$, given by Equation 4.42 was obtained for the null hypothesis and for a specified n. For the distribution approximation we go one step further and for each distribution taken from a library of distributions, we obtain the end point statistic Q_n from Equation 4.39 for a given n and for various choices of the shape parameter. Thus, depending on whether it has a shape parameter or not, each distribution is represented by a point or a trajectory in a two dimensional plane whose coordinates are U_n and V_n. Figure 4.9 shows an example of such a representation. The distributions which are plotted on the distribution approximation chart are (1) Gaussian, (2) Uniform, (3) Exponential, (4) Laplace, (5) Logistic, (6) Cauchy, (7) Extreme Value, (8) Gumbel type 2, (9) Gamma, (10) Pareto, (11) Weibull, (12) Lognormal, (13) Student T, (14) K distributed, (15) Beta, and (16) SU Johnson. Tables 4.8 and 4.9 give the standard and the general form respectively, of these distributions.

As mentioned before, the points on the linked vectors for various distributions are computed using Equation 4.39. The magnitude for each point on the linked vectors is computed from values averaged over 2000 Monte Carlo simulations of the ordered statistic, $Y_{i:n}$, obtained from Equation 4.33 while the angles are computed from the *reference distribution* (standard Gaussian). The confidence ellipses are computed only for null hypothesis used in the prior goodness of fit test. Only the end point coordinates Q_n of the linked vectors are provided in the approximation chart. This is due to the fact that the plot would become too cluttered to properly interpret the data if all the lined vectors for these various distributions were provided in the graphics. Also, meaningful information from the linked vectors is contained in the location of their end

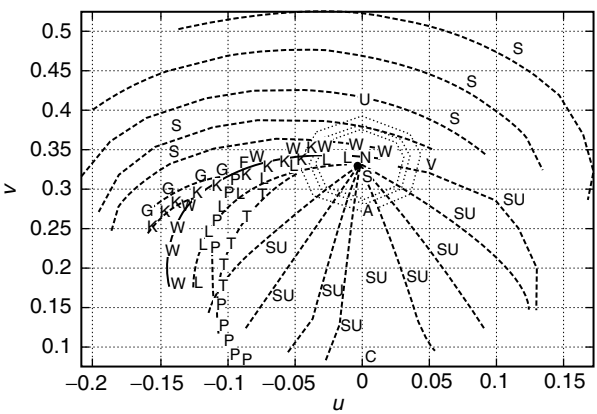

FIGURE 4.9 Distribution approximation chart: N = Normal, U = Uniform, C = Cauchy, L = Lognormal, S = Logistic, A = Laplace, V = Extreme Value, T = T2 Gumbel, G = Gamma, E = Negative Exponential, P = Pareto, K = *K*-Distributed, W = Weibull, B = Beta, SU = SU Johnson.

TABLE 4.8
Standard Forms of the PDF's Used in the Approximation Chart

Distribution	Standard Form $f_Y(y)$	
Gaussian	$(\sqrt{2\pi}\beta)^{-1}\exp\left(-\frac{y^2}{2}\right)$	$-\infty < y < \infty$
Uniform	1	$0 < y < 1$
Exponential	$\exp(-y)$	$0 < y < \infty$
Laplace	$0.5\exp(-\lvert y\rvert)$	$-\infty < y < \infty$
Logistic	$\exp(-y)[1+\exp(-y)]^{-2}$	$-\infty < y < \infty$
Cauchy	$\frac{1}{\pi(1+y^2)}$	$-\infty < y < \infty$
Extreme Value (Type 1)	$\exp(-y)\exp[-\exp(-y)]$	$-\infty < y < \infty$
Gumbel (Type 2)	$\gamma^{y-\gamma-1}\exp(-y^{-\gamma})$	$-\infty < y < \infty$
Gamma	$\frac{1}{\Gamma(\alpha)}\exp(-y)y^{\gamma-1}$	$\alpha < y < \infty$
Pareto	$\frac{\gamma}{y^{\gamma+1}}$	$y > 1,\ \gamma > 0$
Weibull	$\gamma y^{\gamma-1}\exp(-y^{\gamma})$	$y > 0$
Lognormal	$\frac{\gamma}{\sqrt{2xy}}\exp\left[-\frac{\{\gamma\log(y)\}^2}{2}\right]$	$y > 0$
K-Distribution	$\frac{2}{\Gamma(\gamma)}\left(\frac{y}{2}\right)^{\gamma}K_{\gamma-1}y$	$y > 0$
Beta	$\frac{1}{B(\gamma,\delta)}y^{\gamma-1}(1-y)^{\delta-1}$	$0 < y < 1$
Johnson SU	$\frac{\exp\left(\frac{(\sinh^{-1}(y)-\gamma)^2}{2\delta^2}\right)}{\sqrt{2\pi}\delta\sqrt{1+y^2}}$	$-\infty < y < \infty$

points. Therefore, only the end points of all the linked vectors are provided in the approximation chart, along with the confidence ellipse for the selected null distribution.

For each distribution, such as Gaussian, which is uniquely specified by its mean and variance (no shape parameters), there exists a single end point on the approximation chart corresponding to the single unique-linked vector.

For distributions dependent on a single shape parameter, such as Weibull, the end point of the of the linked vector will also be dependent on the shape parameter. Therefore, a sequence of linked vectors must be computed in order to obtain the trajectory on which the end point travels for varying shape parameter. In a sense, the trajectory represents a family of PDFs, having the same distribution but with different shape parameter values. For example, the trajectory for the Weibull distribution is obtained by joining the end points for which the shape parameters are 0.3, 0.4, 0.5, 0.6, 0.8, 1.1, 1.5, 2.0, 3.0, 5.0. As the shape

TABLE 4.9
General Forms of the PDF's Used in the Approximation Chart

Distribution	General Form $f_X(x)$	
Gaussian	$(\sqrt{2\pi}\beta)^{-1}\exp\left(-\frac{(x-\alpha)^2}{2\beta^2}\right)$	$-\infty < x < \infty$
Uniform	$\frac{1}{\beta}$	$\alpha < x < \alpha+\beta$
Exponential	$\frac{1}{\beta}\exp\left(-\frac{(x-\alpha)}{\beta}\right)$	$\alpha < x < \infty$
Laplace	$\frac{0.5}{\beta}\exp\left[-\left\lvert\frac{(x-\alpha)}{\beta}\right\rvert\right]$	$-\infty < x < \infty$
Logistic	$\frac{1}{\beta}\exp\left[-\frac{(x-\alpha)}{\beta}\right]\left[1+\exp\left(-\frac{(x-\alpha)}{\beta}\right)\right]^{-2}$	$-\infty < x < \infty$
Cauchy	$\frac{1}{\pi\beta\left[1+\frac{(x-\alpha)^2}{\beta^2}\right]}$	$-\infty < x < \infty$
Extreme Value (Type 1)	$\frac{1}{\beta}\exp\left[-\frac{(x-\alpha)}{\beta}\right]\exp\left[-\exp\left\{-\frac{(x-\alpha)}{\beta}\right\}\right]$	$-\infty < x < \infty$
Gumbel (Type 2)	$(\gamma/\beta)[(x-\alpha)/\beta]^{-\gamma-1}\exp\left[-\frac{(x-\alpha)^{-\gamma}}{\beta\gamma}\right]$	$\alpha < x < \infty,\ \beta > 0,\ \gamma > 0$
Gamma	$\frac{1}{\beta\Gamma(\alpha)}\exp\left[-\frac{(x-\alpha)}{\beta}\right]\left(\frac{(x-\alpha)}{\beta}\right)^{\gamma-1}$	$\alpha < x < \infty$
Pareto	$\frac{\gamma}{\beta}\frac{1}{\left(\frac{x-\alpha}{\beta}\right)^{\gamma+1}}$	$x > \alpha+\beta,\ \gamma > 0$
Weibull	$\frac{\gamma}{\beta}\left(\frac{x-\alpha}{\beta}\right)^{\gamma-1}\exp\left[-\left(\frac{x-\alpha}{\beta}\right)^{\gamma}\right]$	$x > \alpha$
Lognormal	$\frac{\gamma}{\sqrt{2\pi}\beta\left(\frac{x-\alpha}{\beta}\right)}\exp\left[-\frac{\left\{\gamma\log\left(\frac{x-\alpha}{\beta}\right)\right\}^2}{2}\right]$	$x > \alpha$
K-Distribution	$\frac{2}{\beta\Gamma(\gamma)}\left[\frac{x-\alpha}{2\beta}\right]^{\gamma}K_{\gamma-1}\left[\frac{x-\alpha}{\beta}\right]$	$x > \alpha$
Beta	$\frac{1}{\beta B(\gamma,\delta)}\left(\frac{x-\alpha}{\beta}\right)^{\gamma-1}\left[1-\left(\frac{x-\alpha}{\beta}\right)\right]^{\delta-1}$	$\alpha < x < \alpha+\beta$
Johnson SU	$\frac{1}{\beta}\frac{\exp\left(\frac{\left(\sinh^{-1}\left(\frac{x-\alpha}{\beta}\right)-\gamma\right)^2}{2b^2}\right)}{\sqrt{2\pi}\delta\sqrt{1+\left(\frac{(x-\alpha)}{\beta}\right)^2}}$	$-\infty < x < \infty$

parameter increases, note that the Weibull distribution approaches the Gaussian distribution. This is shown in Figure 4.9. The representation of Figure 4.9 is called an approximation chart.

Similarly, for a distribution dependent on two shape parameters, such as Beta, a sequence of linked vectors must be computed in order to plot the trajectories on which the end point travels for varying shape parameters. This is performed by holding the 1st shape parameter constant and varying the 2nd shape parameter to generate a trajectory, then changing the 1st shape parameter and again holding it constant while varying the 2nd shape parameter, etc. until a family of trajectories is produced which characterizes the distribution.

Thus, an approximation chart as that in Figure 4.9 can be produced. It is apparent that this approximation chart provides a one-to-one graphical representation for each PDF for a given n. Therefore, every point in the approximation chart corresponds to a specific distribution. Thus, if the null hypothesis in the goodness of fit test is rejected, then the distribution which approximates the underlying PDF of the set of random data can be obtained by comparing Q_n, obtained for the samples, with the existing trajectories in the chart. The end point or trajectory closest to the Q_n of the sample data is chosen as an approximation to the PDF underlying the random data. This closest point or trajectory is determined by projecting the sample point Q_n to neighboring points or trajectories on the chart and selecting that point or trajectory whose perpendicular distance from the sample point is the smallest. For example, consider the situation of Figure 4.10. Let $Q_n = (u_n, v_n)$ denote the coordinates of the sample point. Let (x_1, y_1) and (x_2, y_2) denote the coordinates of the points A and B on the trajectory shown in Figure 4.10. The segment of the trajectory between points A and B is assumed to be linear. Let (x_0, y_0) denote the coordinates of the point of intersection of the straight line between A and B and the projection of $Q_n = (u_n, v_n)$ onto this straight line. The equation of the straight line between

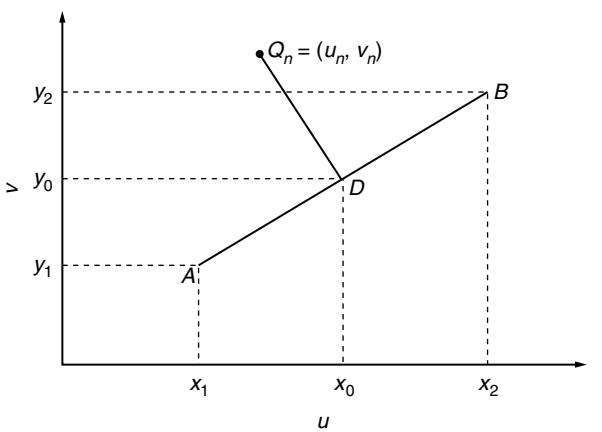

FIGURE 4.10 Computation of distance between sample point and candidate distribution.

the points A and B can be written as

$$y - y_1 = m(x - x_1) \tag{4.85}$$

where

$$m = \frac{y_2 - y_1}{x_2 - x_1}$$

and (x, y) is a point on the line. Also, the equation of the straight line joining (x_0, y_0) and (u_n, v_n) is

$$y - v_n = -\frac{1}{m}(x - u_n) \tag{4.86}$$

where (x, y) is a point on the perpendicular. The coordinates (x_0, y_0) result from letting $x = x_0$ and $y = y_0$ in Equation 4.85 and Equation 4.86. Their solution yields

$$\begin{aligned} x_0 &= \frac{1}{m^2 + 1}[m^2 - my_1 + u_n + mv_n] \\ y_0 &= \frac{1}{m^2 + 1}[y_1 - mx_1 + m_2 v_n + mu_n] \end{aligned} \tag{4.87}$$

Finally, the perpendicular distance from the sample point onto the trajectory between points A and B is

$$D = \sqrt{\frac{1}{m^2 + 1}[m^2\psi_1^2 - 2m\psi_1\psi_2 + \psi_2^2]} \tag{4.88}$$

where

$$\psi_1 = u_n - x_1, \qquad \psi_2 = v_n - y_1 \tag{4.89}$$

The complete distribution approximation algorithm is summarized as follows.

1. Sort the samples $X_1, X_2, \ldots, X_n$ in increasing order.
2. Obtain the standardized order statistic $Y_{i:n}$.
3. Compute U_n and V_n from Equation 4.39 for the library of PDFs mentioned.
4. Obtain an approximation chart based on the sample size n and plot the sample point Q_n on this chart.
5. Compute the distance, D between the sample point Q_n and each of the end points on the chart. Choose the PDF corresponding to the point or trajectory which results in the smallest value for D as an approximation to the PDF of the samples.

The approximation to the underlying PDF of the set of random data can be improved by including as many distributions as possible in the approximation

chart so as to fill as much of the space as possible with candidate distributions. It is emphasized, however, that this procedure does not identify the underlying PDF. It merely gives the best approximation to the distribution underlying the PDF of the data from those available in the chart.

4.2.3.3. Parameter Estimation

Once the probability distribution of the samples is approximated, the next step is to estimate its parameters. The method of distribution approximation discussed in Section 4.2.3.2 lends itself for estimating the parameters of the approximated distribution. We present the estimation procedure for the location, scale, and shape parameters in this section.

4.2.3.3.1. Estimation of Location and Scale Parameters

Let $f(x:\alpha, \beta)$ denote a known distribution which approximates the PDF of the set of random data, where α and β are the location parameter and scale parameter, respectively, of the approximating PDF. Let $X_{i:n}$ denote the ordered statistics of X from a sample of size n. Let $S_{i:n}$ be defined by

$$S_{i:n} = \frac{X_{i:n} - \alpha}{\beta} \tag{4.90}$$

Also, let

$$\mu_{i:n} = E[S_{i:n}] \tag{4.91}$$

Then

$$E[X_{i:n}] = \beta\mu_{i:n} + \alpha \tag{4.92}$$

We consider the following statistics

$$T_1 = \sum_{i=1}^{n} \cos(\theta_i) X_{i:n}, \qquad T_2 = \sum_{i=1}^{n} \sin(\theta_i) X_{i:n} \tag{4.93}$$

where θ_i is the angle defined in Equation 4.37. The expected values of T_1 and T_2 are

$$E[T_1] = \sum_{i=1}^{n} \cos(\theta_i)[\beta\mu_{i:n} + \alpha], \qquad E[T_2] = \sum_{i=1}^{n} \sin(\theta_i)[\beta\mu_{i:n} + \alpha] \tag{4.94}$$

These can be written as

$$E[T_1] = a\alpha + b\beta, \qquad E[T_2] = c\alpha + d\beta \tag{4.95}$$

where

$$a = \sum_{i=1}^{n} \cos(\theta_i) \qquad b = \sum_{i=1}^{n} \mu_{i:n} \cos(\theta_i)$$
$$c = \sum_{i=1}^{n} \sin(\theta_i) \qquad d = \sum_{i=1}^{n} \mu_{i:n} \sin(\theta_i) \tag{4.96}$$

Because the standardized Gaussian distribution is used as the reference distribution for θ_i, it can be shown that $a = 0$.[1] The estimates for β and α are then given by

$$\hat{\beta} = \frac{\widehat{E[T_1]}}{\hat{b}}, \qquad \hat{\alpha} = \frac{\widehat{E[T_2]} - \hat{d}\hat{\beta}}{\hat{c}} \tag{4.97}$$

For sufficiently large n (i.e., $n > 50$), suitable estimates for $E[T_1]$ and $E[T_2]$ are

$$\widehat{E[T_1]} \approx T_1, \qquad \widehat{E[T_2]} \approx T_2 \tag{4.98}$$

Estimates for b and d rely upon an estimates of $\mu_{i:n}\hat{\mu}_{i:n}$ obtained from a Monte Carlo simulation of $S_{i:n}$ where $S_{i:n}$ is generated from the known approximating distribution $f(x; 0, 1)$ having zero location and unity scale parameters. $\hat{\mu}_{i:n}$ is the sample mean of $S_{i:n}$ based upon 2000 Monte Carlo trials. Having $\hat{\mu}_{i:n}$, the estimates for b and d are given by

$$\hat{b} = \sum_{i=1}^{n} \hat{\mu}_{i:n} \cos(\theta_i), \qquad \hat{d} = \sum_{i=1}^{n} \hat{\mu}_{i:n} \sin(\theta_i) \tag{4.99}$$

The scale and location parameters are then estimated by application of Equation 4.97 and Equation 4.98.

4.2.3.3.2. Shape Parameter Estimation

In this section we present the approximate method used for estimating the shape parameter of the approximating PDF. We first consider distributions with only one shape parameter. Let γ denote the shape parameter of the approximating PDF. Since U_n and V_n are location and scale invariant, the point Q_n depends only on the sample size n and the shape parameter γ.

Recall that a point on the trajectories of the approximation chart is obtained by averaging for a specified value of the shape parameter the results from a large number of trials for U_n and V_n. Consequently, for given values of n and γ the coordinates of the corresponding points along the trajectory for a specified distribution, can be characterized by

$$E(U_n) = \phi_1(n, \gamma), \qquad E(V_n) = \phi_2(n, \gamma) \tag{4.100}$$

where the complete trajectory is obtained by repeating large number of trials for U_n and V_n over a suitable range of γ. On a given trial involving the random data,

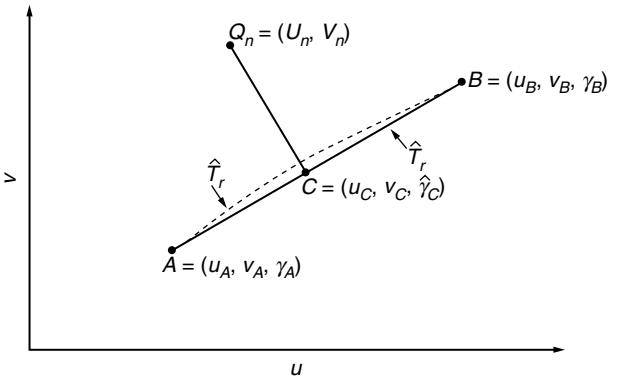

FIGURE 4.11 Shape parameter estimation.

it is likely that the coordinates U_n and V_n obtained for the samples will not coincide with any of the trajectories on the chart. The PDF underlying the random data is approximated by selecting the distribution, corresponding to the point in the trajectory, that falls closest to the sample point Q_n. The situation is illustrated in Figure 4.11. Q_n appears in the figure with coordinates (U_n, V_n). The straight line $\hat{T}_r$ denotes an approximation to a segment of the nearest trajectory which, in general, is a curved segment between points A and B. A is that point on the actual trajectory corresponding to the shape parameter γ_A. Its coordinates are (u_A, v_A). Similarly, B is the point on the actual trajectory corresponding to the shape parameter γ_B. Its coordinates are (u_B, v_B). The slope of the straight line between points A and B is

$$m = \frac{v_B - v_A}{u_B - u_A} \tag{4.101}$$

The equation for the straight line $\hat{T}_r$ is

$$v = v_A + m(u - u_A) \tag{4.102}$$

Point C, with coordinates (u_C, v_C), is the perpendicular projection of Q_n onto $\hat{T}_r$. The straight line linking Q_n and C has a slope equal to $-1/m$ and an equation of the form

$$v = V_n - \frac{1}{m}(u - U_n) \tag{4.103}$$

Since C is a point common to both straight lines, it follows from Equation 4.102 and Equation 4.103 that

$$u_A + m(u_C - u_A) = V_n - \frac{1}{m}(u_C - U_n) \tag{4.104}$$

Solution for u_C results in

$$u_C = \frac{m(V_n - v_A) + mu_A + U_n}{m^2 + 1} \tag{4.105}$$

Let γ_C denote the shape parameter corresponding to the point on the actual trajectory closest to Q_n. An approximation to γ_C is then obtained by linear interpolation on $\hat{T}_r$. The result is

$$\hat{\gamma}_C = \gamma_A + \frac{(\gamma_B - \gamma_A)(u_C - u_A)}{(u_B - u_A)} \tag{4.106}$$

The accuracy of the procedure can be improved by employing a nonlinear interpolation method. It must be emphasized that the location, scale, and shape parameter estimation procedures presented in this section are approximate methods.

The proposed estimation procedure can also be extended to the two shape parameter case. In this case one needs to choose at least three points (u_1, v_1), (u_2, v_2), and (u_3, v_3) and let the shape parameter values corresponding to these three points be γ_1, γ_2, and γ_3, respectively. The points are chosen in such a way that they form the three vertices of a triangle inside which falls the sample point Q_n.[1] Again, by using a linear interpolation in the plane, an approximate solution can be obtained for the parameter estimates.

4.3. SIMULATION RESULTS OF THE OZTURK ALGORITHM

For univariate cases, the power of the Ozturk Algorithm has been studied for various distributions in Refs. 1–3. It was noted in Refs. 1–3 that the power of the algorithm depends on the sample size n, type of the standardized statistic and the null distribution. This algorithm has been found to compare favorably against all the well known tests. Also, the algorithm has been put to use to test its performance against different known distributions. Random data were generated using computer simulations as given in Refs. 18,19. The goodness of fit as well as the distribution approximation test was performed on these data using this algorithm. In this chapter, a brief summary of some of the results obtained is presented.

Data was generated from four different null distributions, viz.,

- Univariate Gaussian
- Weibull (Shape Parameter 1)
- Gamma (Shape Parameter 1)
- Lognormal (Shape Parameter 1)

The goodness of fit test results are tabulated and presented first. The results of the distribution approximation are not easy to tabulate. We shall, therefore, present the result of a single case for the purpose of illustration.

4.3.1. Goodness of Fit Test Results

4.3.1.1. The Univariate Gaussian Case

Data was generated from a Gaussian pseudo random number generator using computer simulations. The data set represented a zero mean and unit variance normal PDF. The following observations were noted.

- It was observed that a sample size of less than 40 is not advisable for the goodness of fit test as it almost always shows that the data is statistically consistent with any null. This is due to the fact that such a small sample size could be used to represent any PDF.
- A sample size between 75 and 100 is found to be good enough to accurately perform the goodness of fit test.
- For a sample size greater than 75 and when the null specified was Gaussian, the goodness of fit test showed that the data was statistically consistent with the null in almost all the cases. For other null hypotheses which are not close to Gaussian in the approximation chart, the goodness of fit test always showed that the data was statistically inconsistent with the nulls. But for null hypotheses which are close to Gaussian in the approximation chart, such as logistic, the goodness of fit test comes up with statistical consistency almost always. This vindicates the fact that the logistic PDF curve is very similar to the Gaussian PDF curve.

Table 4.9 shows the results obtained for this case.

4.3.1.2. The Weibull Case

Data was generated from the Weibull PDF with shape parameter one and the goodness of fit test was performed on it. The following observations were noted.

- The goodness of fit test worked well for a sample size between 75 and 100. For a smaller sample size, the goodness of fit test is not advisable since the results obtained were not accurate.
- When the null specified was Weibull with shape parameter two, the goodness of fit test showed that the data was statistically inconsistent with the null for a sample size greater than 75 in all the cases. This is due to the fact that Weibull (Shape Parameter 1) is far away from Weibull (Shape Parameter two) on the approximation chart.

Table 4.10 shows the results obtained for this case.

TABLE 4.10
Result of the Ozturk Algorithm when Data Generated was Gaussian

Sample Size (*n*)	Data Generated From	Null Distribution	Number of Cases of SC/ Total Number of Cases
5	Gaussian	Gaussian	3/3
25	—	—	3/3
40	—	—	4/7
50	—	—	4/7
75	—	—	7/7
100	—	—	7/8
125	—	—	8/8
150	—	—	8/8
25	—	Uniform	2/5
40	—	—	2/5
50	—	—	1/5
75	—	—	0/5
100	—	—	0/5
150	—	—	0/5
75	—	Exponential	0/8
100	—	—	0/8
150	—	—	0/8
75	—	Laplace	4/8
100	—	—	2/8
150	—	—	1/8
75	—	Logistic	7/8
100	—	—	7/8

SC, Statistical Consistency.

4.3.1.3. The Gamma Case

Data was generated from the Gamma PDF with shape parameter one and the goodness of fit test was performed on it. Observations noted were almost the same as those for the Weibull case. Again a sample size between 75 and 100 was observed to have performed well in this case. The results for this case are tabulated in Table 4.11.

4.3.1.4. The Lognormal Case

A Lognormal pseudo random number generator was used to generate random data representing the Lognormal PDF with a shape parameter of one. Observations noted for the goodness of fit test performance on this data were noted.

TABLE 4.11
Ozturk Algorithm Results when Data Generated was Weibull with Shape Parameter One

Sample Size (*n*)	Data Generated From	Null Distribution	Number of Cases of SC/ Total Number of Cases
5	Weibull (Sh. 1)	Weibull (Sh. 1)	7/8
25	—	—	7/9
40	—	—	4/8
50	—	—	5/8
75	—	—	6/8
100	—	—	8/8
150	—	—	8/8
5	Weibull (Sh. 1)	Weibull (Sh. 2)	7/8
25	—	—	3/8
40	—	—	2/8
50	—	—	0/8
75	—	—	0/9
100	—	—	0/8
150	—	—	0/8

SC, Statistical Consistency; Sh., Shape Parameter.

TABLE 4.12
Ozturk Algorithm Results when Data Generated was Gamma with Shape Parameter One

Sample Size (*n*)	Data generated From	Null Distribution	Number of Cases of SC/ Total Number of Cases
40	Gamma (Sh. 1)	Gamma (Sh. 1)	5/7
50	—	—	7/8
75	—	—	7/9
100	—	—	7/8
150	—	—	8/8
50	Gamma (Sh. 1)	Gamma (Sh. 5)	1/8
75	—	—	0/8
100	—	—	1/8

SC, Statistical Consistency; Sh., Shape Parameter.

- A sample size between 50 and 75 was found to be sufficient to perform the goodness of fit test accurately.
- This sample size gave very good results as far as the distribution approximation was concerned. About 30 times out of 40 the Lognormal PDF showed up in the five closest distributions that could be approximated.

Table 4.12 shows the results of the goodness of fit test for this case.

In general, the goodness of fit test seemed to perform well for a sample size of 100. The confidence contours grow smaller when the sample size is increased. In effect, we could hypothesize that for an infinite sample size the contours would become a point in the two dimensional (U, V) plane. This is intuitively satisfying.

4.3.2. Distribution Approximation Test Results

The distribution approximation test was performed for a number of cases. In fact, it was performed for all the cases in which the goodness of fit test was performed. As mentioned previously, since it is not very easy to tabulate the results of the distribution approximation test for all these cases, results for a single test case are presented below.

Data was generated from standard Gaussian distribution using a Gaussian random number generator. A single test case consisting of 100 data points was considered. Using standard Gaussian as the null distribution, the distribution approximation test was performed on the data set. The 1st results of this test gave the five closest PDF's which the data could approximate. This result is shown in Table 4.13.

TABLE 4.13
Ozturk Algorithm Results when Data Generated was Lognormal with Shape Parameter One

Sample Size (*n*)	Data generated From	Null Distribution	Number of Cases of SC/ Total Number of Cases
40	Lognormal (Sh. 1)	Lognormal (Sh. 1)	5/8
50	—	—	7/8
75	—	—	8/8
100	—	—	8/8
150	—	—	8/8
40	Lognormal (Sh. 1)	Lognormal (Sh. 0.5)	0/9
50	—	—	0/8
75	—	—	0/8
100	—	—	0/8

SC, Statistical Consistency; Sh., Shape Parameter.

TABLE 4.14
Five Closest PDF's Given by Distribution Approximation Test for a Standard Gaussian Data Set

Distribution Number	Distance	Rank
21	0.23728×10^{-07}	1
22	0.89802×10^{-05}	2
20	0.23473×10^{-04}	3
23	0.28047×10^{-04}	4
5	0.47542×10^{-04}	5

Distributions 20, 21, 22, and 23 are all SU-Johnson distributions with different shape parameters, whereas distribution number five is a logistic distribution. Note that the standard Gaussian was the 11th ranked PDF with a distance of 0.47879×10^{-03}. Estimates of the location, scale, and the shape parameters given by the distribution approximation test for these distributions are given in Table 4.14.

The approximation chart for this test case is shown in Figure 4.12. It is obvious from the approximation chart that the PDF's identified for this case are very close to the Gaussian PDF. It is therefore concluded that even though this data set has passed the goodness of fit test with standard Gaussian as the null, they could also be approximated by the set of 5 PDF's identified in Table 4.15.

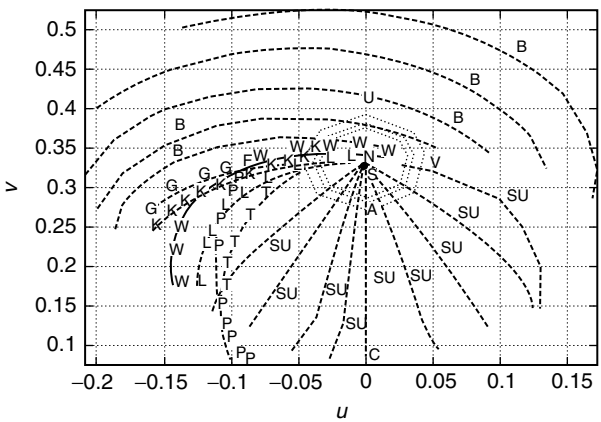

FIGURE 4.12 Distribution approximation chart for a standard Gaussian data set: end point of linked vectors for a standard Gaussian data set, N = Normal, U = Uniform, C = Cauchy, L = Lognormal, S = Logistic, A = Laplace, V = Extreme Value, T = T2 Gumbel, G = Gamma, E = Negative Exponential, P = Pareto, K = *K*-Distributed, W = Weibull, B = Beta, SU = SU Johnson.

TABLE 4.15
Estimates of the Parameters of the Five Closest Distributions Chosen by the Distribution Approximation Test for a Standard Gaussian Data Set

Distribution Number	Location Parameter	Scale Parameter	Shape (1) Parameter	Shape (2) Parameter
21	−0.23435	1.9774	2.4099	−0.2
22	−0.14976	1.9392	2.3697	−0.1
20	−0.39697	1.9747	2.4310	−0.4
23	−0.058871	1.9437	2.3717	0.0
5	−0.06215	0.49795	0.0	0.0

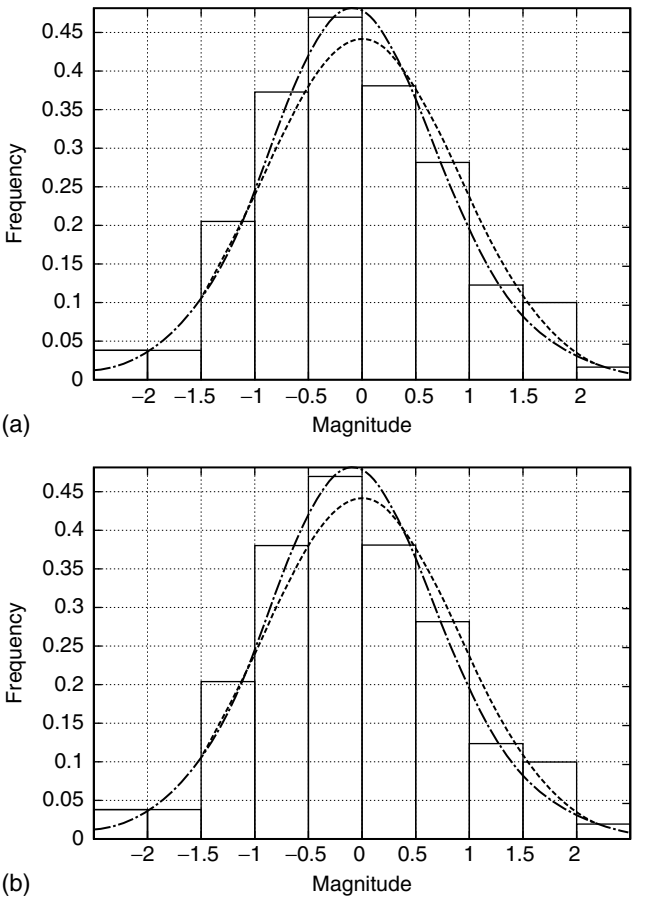

FIGURE 4.13 Histogram fitting (solid line is histogram plotted for the data; dotted line is standard Gaussian); (a) dashed line is PDF number 21 of Table 4.15; (b) dashed line is PDF number 22 of Table 4.15.

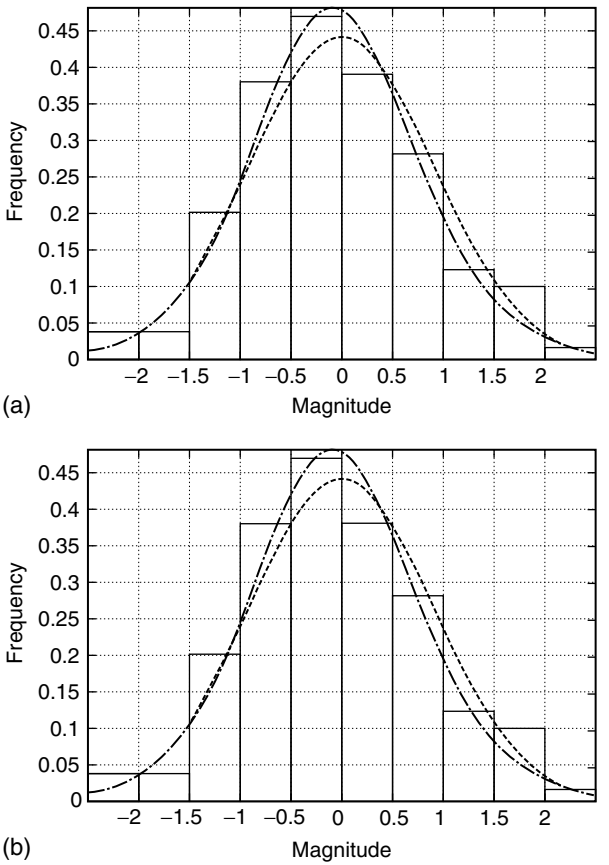

FIGURE 4.14 Histogram fitting (solid line is histogram plotted for the data; dotted line is standard Gaussian); (a) dashed line is PDF number 20 of Table 4.15, (b) dashed line is PDF number 23 of Table 4.15.

In fact these are better approximations than standard Gaussian. This is shown by histogram plots shown in Figures 4.13–4.15. In these plots the histogram of the data is plotted along with the null hypothesis, which is the standard Gaussian, and one of the five distributions, given by the distribution approximation test, on the same coordinate axes. As is obvious from the figures there is very little to choose amongst the five PDF's approximated.

4.4. CONCLUSIONS AND SUGGESTIONS FOR FUTURE WORK

4.4.1. Conclusions

This thesis has discussed various techniques for analyzing random data. Two areas were considered. The first area dealt with goodness of fit tests to determine

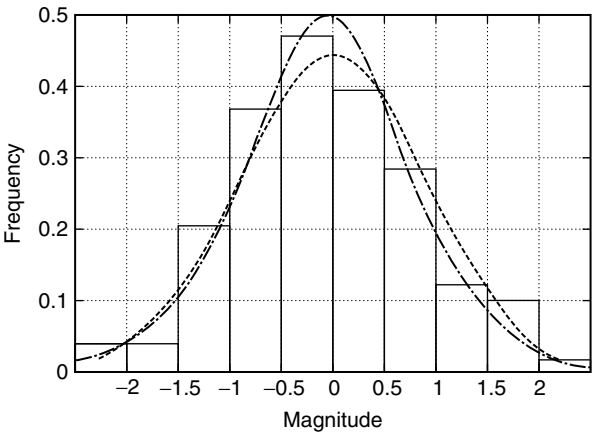

FIGURE 4.15 Histogram fitting (solid line is histogram plotted for the data; dotted line is standard Gaussian; dashed line is PDF number five of Table 4.15).

whether or not a set of random data is statistically consistent with a prespecified probability distribution. After reviewing the Kolmogorov–Smirnov test, chi-square test, Q–Q plots and P–P plots, a new test, called the Ozturk Algorithm, was introduced. This test has an easily understood graphically presentation and works well for sample sizes as small as 100. The second area dealt with approximation of the underlying PDF of random data. Although the other tests were not applicable to this second area, the goodness of fit test of the Ozturk Algorithm was shown to lend itself to generation of a distribution approximation chart from which approximations to the underlying PDF of the random data can be obtained. Again, good results were observed for sample sizes as small as 100. An analysis was provided for generating confidence contours when the random data was nonGaussian. Simulated data was used to evaluate performance of the Ozturk Algorithm and some of the results were presented.

4.4.2. Suggestions for Future Work

Several problems remain to be explored with the Ozturk Algorithm:

1. The Ozturk Algorithm works well for continuous probability density functions. Generalizations of the Ozturk Algorithm to discrete case should be explored.
2. Extension of the Ozturk Algorithm from univariate to multivariate PDF's should be considered. One possibility involves utilization of quadratic forms of the data.[7,11]
3. Rangaswamy[11] demonstrated that multivariate spherically invariant random processes (SIRPs) can be approximated by means of their quadratic forms. However, a probability distribution approximation

chart for SIRPs which could be utilized by the Ozturk Algorithm remains to be generated.

4. The univariate PDF's currently included in the approximation chart are unimodal. Extension to multimodal PDF's should be explored.
5. When the number of data point is much greater than 100, the Ozturk Algorithm requires considerable time to process the data. Ways should be examined for making the algorithm more efficient. This includes parallelization of the algorithm as well as processing the data in groups of 100 and averaging the results.
6. Reduction of the Ozturk Algorithm to chip form should be investigated for real-time applications.

5 Probability Density Distribution Approximation and Goodness-of-Fit Tests of Random Data

A. Ozturk

CONTENTS

5.1. A New Method for Distribution Approximation 260
5.1.1. Introduction 260
5.1.2. Approximation Procedure 261
5.1.3. Performance of the Approximation Procedure 263
5.1.4. Approximation of Multivariate Distributions 264
5.1.5. Parameter Estimation 268
5.1.5.1. Estimating the Location and Scale Parameters 268
5.1.5.2. Estimating the Shape Parameters 270
5.1.6. Distribution Approximation for Mixtures of Distributions 272
5.1.7. Examples 273
5.1.8. Conclusions 275
5.2. A General Algorithm for Univariate and Multivariate Goodness-of-Fit Tests Based on Graphical Representation 276
5.2.1. Introduction 276
5.2.2. The Test Procedure 277
5.2.3. Properties of the Test Statistics 279
5.2.4. Extensions of the Test 282
5.2.5. Examples 283
5.2.6. Empirical Power 286
5.2.7. The Test Algorithm 290
5.2.8. Discussion 292

5.1. A NEW METHOD FOR DISTRIBUTION APPROXIMATION

5.1.1. Introduction

Assessing the distributional assumptions about univariate and multivariate data is a basic concern in statistical applications. A common approach to the problem is to hypothesize a certain probability model and perform an appropriate goodness-of-fit test. Various test procedures, some of which are in graphical nature, have been developed for assessing the distributional assumptions of a sample. For example, probability plots, in particular, quantile–quantile ($Q-Q$) plots are among the most widely used graphical procedures for making assessments about the sample.

A second and perhaps more important aspect of assessing the distributional assumptions of the sample, is to diagnose the true distribution if the null hypothesis is rejected. Most test procedures do not have this diagnostic feature. In $Q-Q$ plots any member of location and scale family of distributions can be displayed and represented by a single curve. Thus by comparing the sample $Q-Q$ plots with theoretical plots, one can obtain some information about the true distribution. However, the $Q-Q$ plots do not offer a formal test. Identification is made on a subjective basis. Also, generalizations of this graphical procedure for diagnosing true distributions are not straightforward when applied to the non location-scale family of distributions and to the multivariate distributions. For a comprehensive review and interpretation of univariate $Q-Q$ plots we refer to Wilk and Gnanedesikan[1] and for the multivariate generalization of $Q-Q$ plots to Easton and McCulloch.[2]

It is well-known that the probability distributions for which the first four moments exist, can be characterized by their coefficients of skewness (α_3) and kurtosis (α_4). Any distribution or family of distributions, can be represented by a single point or by a region on an $\alpha_3-\alpha_4$ plane, respectively (see for example Johnson and Kotz,[3] p. 14). A sample can also be represented by a point whose coordinates are given by the sample estimates of α_3 and α_4. Then the best candidate for the underlying true distribution can be identified to be the nearest neighbor distribution on the chart. Although such a chart based on the coefficients of skewness and kurtosis provide a useful way of characterizing the distributions, its use is limited by the fact that the moments of some distributions do not exist. Another drawback of this approach is that the estimates of α_3 and α_4 are highly sensitive to extreme observations.

In this paper, we introduce a general graphical method for characterizing the univariate and multivariate distributions. The proposed method not only provides a formal goodness-of-fit test based on graphical representation, but also a good basis for displaying the distributions in a two dimensional space. Using normal distribution as a *reference distribution*, the standardized sample order statistics are represented by a system of linked vectors. The terminal point of these linked vectors is used to approximate the true distribution, from which the sample is drawn. A chart is developed for this purpose and it is shown that it can also be used to estimate the parameters of the identified distribution.

The approximation procedure is explained in Section 5.1.2. The performance of the proposed method is discussed in Section 5.1.3. Extension of the proposed procedure to multivariate distributions is discussed in Section 5.1.4. Estimation of the location, scale and the shape parameters are given in Section 5.1.5. Approximation of the mixtures of distributions is explained in Section 5.1.6. Some examples are given in Section 5.1.7 to illustrate the proposed method.

5.1.2. Approximation Procedure

This study was motivated by an attempt to develop a general algorithm for univariate and multivariate goodness-of-fit tests, based on graphical representation. The development of the test procedure for assessing the distributional assumptions is explained in a series of papers by Ozturk[4] and Ozturk and Dudewicz.[5] In this section we first give a brief description of the corresponding test statistic and then explain the distribution approximation procedure. For illustration we will assume that the null distribution is normal. The proposed procedure can be modified by using any member of location-scale family of distributions as a *reference distribution*.

Let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ be an ordered sample from a normal distribution with mean μ and variance σ^2. Let $m_{1:n}, m_{2:n}, \ldots, m_{n:n}$ denote the expected values of the standard normal order statistics. Also let

$$\theta_i = \pi\Phi(m_{i:n}) \tag{5.1}$$

where $\Phi(\cdot)$ is the distribution function of the standard normal distribution and $\pi = 3.1415\ldots$ We define the sample point $Q_n^{(k)}$ in a two dimensional plane by

$$Q_n^{(k)} = (U_n^{(k)}, V_n^{(k)}) \tag{5.2}$$

where

$$U_n^{(k)} = \frac{1}{n}\sum_{i=1}^{n}\{\cos(\theta_i)\}Y_{i:n}^{(k)} \tag{5.3}$$

$$V_n^{(k)} = \frac{1}{n}\sum_{i=1}^{n}\{\sin(\theta_i)\}Y_{i:n}^{(k)} \tag{5.4}$$

and $Y_{i:n}^{(k)}$ is the standardized ith sample order statistic and k stands for type of standardization. Ozturk[4] considered various types of standardized sample order statistics to construct the test statistic $Q_n^{(k)}$. In this study the following standardized statistic will be considered.

$$Y_{i:n}^{(1)} = \frac{|X_{i:n} - \bar{X}|}{S}; \qquad i = 1, 2, \ldots, n \tag{5.5}$$

where $\bar{X} = \sum X_{i:n}/n$ and $S = \{\sum(X_{i:n} - \bar{X})^2/(n-1)\}^{1/2}$. It should be noted that the statistic $Q_n^{(k)}$ given in Equation 5.2 represents the terminal point of the linked vectors with each having a length $|Y_{i:n}^{(k)}|$ and angle with the horizontal axis θ_i.

Under null hypothesis the statistic $Q_n^{(k)} = (U_n^{(k)}, V_n^{(k)})$ has a bivariate distribution with mean $E(Q_n^{(k)}) = \{E(U_n^{(k)}), E(V_n^{(k)})\}$. If the distribution of $Q_n^{(k)}$ was known, then a $100(1 - \alpha)\%$ confidence contour of the point $E(Q_n^{(k)})$ could be obtained to perform a test of hypothesis that the sample comes from a normal distribution. A general algorithm was developed to obtain the empirical distribution of $Q_n^{(k)}$ under null hypothesis.[4]

For a given sample size n, the bivariate distribution of $Q_n^{(k)}$ is approximated by the corresponding empirical distribution.

Following a similar approach, one can also generate random samples from other distributions and obtain the corresponding empirical distribution of the statistic $Q_n^{(k)}$. Thus for each specified alternative distribution and for fixed sample size n, the expected values of the statistic $Q_n^{(k)}$ based on Equation 5.3 and Equation 5.4 can be obtained. It is clear that all the probability distributions for which the statistic $Q_n^{(k)}$ is definable, can be represented on a two dimensional plane by using the standard normal distribution as a *reference distribution*. An example of such representation based on the statistic in Equation 5.5 for which $k = 1$ is given in Figure 5.1. Twelve distributions, namely normal (1), uniform (2), exponential (3), Laplace (4), Logistic (5), Cauchy (6), Extreme value (7), Gumbel-type 2 (8), Gamma (9), Pareto (10), Weibull (11) and Lognormal (12) distributions are considered. Coordinates of the points representing each distribution are obtained through Monte Carlo experiments. By using the IMSL subroutines, 1000 samples of size 50 are generated for each distribution. It can be seen that the location and scale of the family of distributions such as normal, uniform, negative exponential etc., are each represented by a single point since the $Q_n^{(k)}$ statistic, is a location and scale invariant statistic. For a given value of the shape parameter of a distribution, the corresponding point $E(Q_n^{(k)})$ can be obtained similarly. Thus, a curve obtained by joining such points represents the corresponding

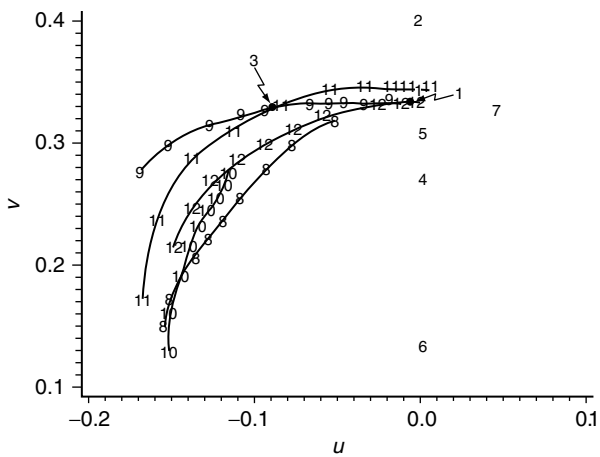

FIGURE 5.1 Identification chart for univariate distributions based on 1000 samples ($n = 50$).

distribution family having one shape parameter. For example the curve for Gamma distribution in Figure 5.1 is obtained by joining the points for which the selected values of the shape parameter were $\gamma = 0.2, 0.3, 0.5, 0.7, 1.0, 2.0, 3.0, 4.0, 6.0, 10.0$. Note that as γ increases the Gamma distribution approaches to the normal distribution.

The chart given in Figure 5.1 provides a useful characterization of various distributions via normal distribution. The relationship between the distributions can be clearly observed.

For a given sample of size n, the point $Q_n^{(k)}$, as an estimate of the true point $E(Q_n^{(k)})$, can be plotted on the chart in Figure 5.1. The statistic $Q_n^{(k)}$ is a function of the linear combinations of the standardized sample order statistics. Hence, it is expected to provide a more consistent estimate of the true point $E(Q_n^{(k)})$ than that of moment estimates.

It may be argued that a given distribution can be represented in a $(U_n^{(k)}, V_n^{(k)})$ plane by a single point. Conversely, a given point in $(U_n^{(k)}, V_n^{(k)})$ region corresponds to a distribution. Thus, if the null hypothesis that the random sample $X_1, X_2, \ldots, X_n$ comes from a normal distribution is rejected, then the underlying true distribution can be diagnosed by comparing the point $Q_n^{(k)} = (U_n^{(k)}, V_n^{(k)})$ with the existing distributions in the chart. Therefore the nearest neighboring distribution can be identified to be the best candidate for true distribution.

The proposed algorithm for distribution identification is summarized as follows

(1) Sort the sample observations $X_1, X_2, \ldots, X_n$ in increasing order.
(2) Obtain the standardized order statistic $Y_{i:n}^{(k)}$.
(3) Compute the coordinates $U_n^{(k)}, V_n^{(k)}$ of the point $Q_n^{(k)}$ from Equation 5.3 and Equation 5.4.
(4) Obtain an identification chart based on the sample of size n and plot the sample point $Q_n^{(k)}$ on this chart.
(5) Compare the point $Q_n^{(k)}$ with the existing distributions in the chart. Identify the nearest neighboring distribution as the best candidate for the true distribution.

It is clear that the accuracy of such procedure may be increased by including as many distributions as possible. Furthermore, confidence ellipses of a point for a given distribution can be obtained empirically. The use of such contours will lead to a decision on the probabilistic basis.

5.1.3. Performance of the Approximation Procedure

In this section we investigate the performance of the proposed identification procedure. It is clear that for identification procedure to be useful, some idea of sampling variation of the statistics $U_n^{(k)}$ and $V_n^{(k)}$ is needed. If the sampling variation of the statistics is too large then the performance of the underlying procedure will be poor.

One way of getting some idea about the sampling variation of the $Q_n^{(k)}$ statistic is to obtain the corresponding confidence ellipses. For $k = 1$ and $n \leqq 100$, Ozturk and Dudewicz[5] obtained the following approximations to the means and the variances of the statistics $Q_n^{(1)}$ and $V_n^{(1)}$ under the normality assumption of the sample

$$\begin{aligned} E\{U_n^{(1)}\} &= 0 \\ E\{V_n^{(1)}\} &= 0.3266 + 0.4129/n \\ \mathrm{Var}\{U_n^{(1)}\} &= 0.02123/n + 0.01765/n^2 \\ \mathrm{Var}\{V_n^{(1)}\} &= 0.04427/n - 0.0951/n^2 \end{aligned}$$

It was also shown empirically that the covariance $\mathrm{Cov}\{U_n^{(1)}, V_n^{(1)}\} = 0$ and the joint distribution of $U_n^{(1)}$ and $V_n^{(1)}$ is approximately bivariate normal for $n > 10$. Based on the above approximations, one can obtain approximate confidence ellipses for various sample sizes.

In a later study, Ozturk[4] used the Johnson system of distributions to obtain approximate confidence ellipses for any point in the identification chart. Examples of such confidence ellipses for sample sizes $n = 50$, 100, and 500 from normal distribution are presented in Figure 5.2. Three confidence ellipses namely 90, 95 and 99% are given for each case. The effect of the sample size on the confidence ellipses can be clearly observed from these figures.

To get a better feeling about the power of the proposed procedure, we need to compare the performance of underlying procedure with those of the Q–Q and moments procedures. However, we did not include the Q–Q procedure in the comparison, since it has a subjective basis. To compare the proposed method with the method of moments, we started with a chart of 19 distributions. A simulation experiment was designed to generate 1000 samples of size $n = 15$, 25, 50, 100 and 500. Then the number, of correct identifications when using each method, is counted for a given distribution and sample size. The corresponding results are given in Table 5.1. It is interesting to see that the proposed method has outperformed the method of moments in most of the cases. The performance of proposed method was no better than its competitor for the normal and uniform distribution when the sample is small.

5.1.4. Approximation of Multivariate Distributions

The proposed method of distribution identification for univariate distributions also can be generalized to identify the multivariate distributions. It should be noted that in many multivariate distributions the marginal distributions are either identical or common, up to location and scale parameters.[6] In such cases the first step would be to assess the assumption that all marginal distributions are common. A simple approach to this problem is to ignore the multivariate nature of data and use univariate procedures to assess the distributional assumptions which may lead to identifying the corresponding multivariate distribution.

However, there is a need for a test which takes into account the general multivariate nature of the observations also.

Let $X' = (X_1, X_2, \ldots, X_n)$ be a random vector having a multivariate normal distribution with mean μ and covariance Σ. A number of methods which utilize the *squared radii*

$$Z_i^2 = (X_i - \bar{X})^i S^{-1} (X_i - \bar{X}) \tag{5.6}$$

where $\bar{X}$ and S are sample estimates of μ and Σ, respectively, have been suggested to test for multivariate normality.[4,7,8]

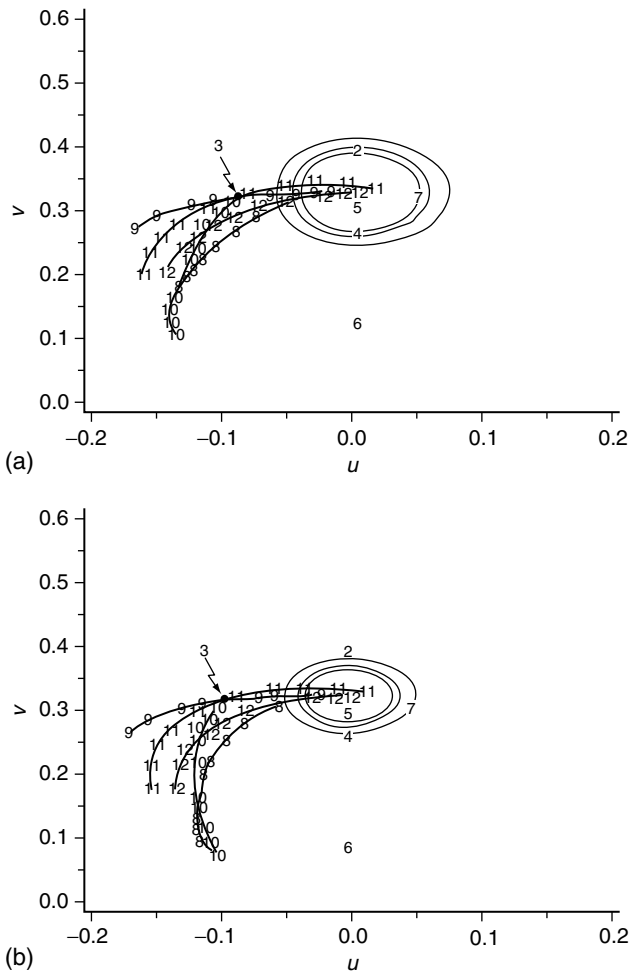

FIGURE 5.2 90, 95, and 99% confidence ellipses of $E\{Q_n^{(k)}\}$ based on a sample of size n from normal distribution: (a) $n = 50$, (b) $n = 100$, (c) $n = 500$.

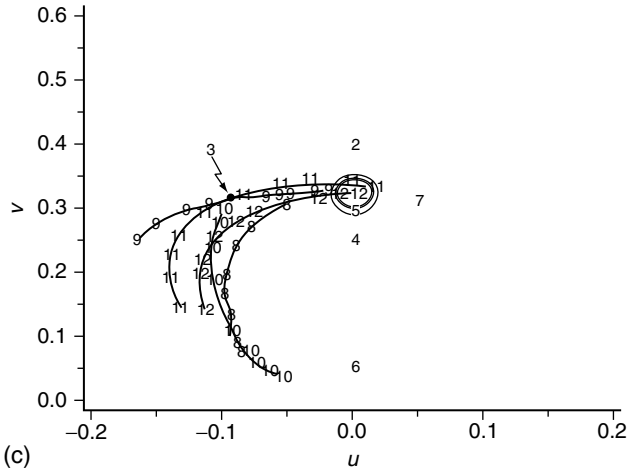

FIGURE 5.2 Continued.

TABLE 5.1
Performance of the Proposed *(Q)* and Moment *(M)* Procedures in Terms of Correctly Identified Samples out of 1000 Samples

	n = 15		*n = 25*		*n = 50*		*n = 100*		*n = 500*	
	Q	***M***	***Q***	***M***	***Q***	***M***	***Q***	***M***	***Q***	***M***
Normal	121	371	217	496	414	652	630	817	905	983
Uniform	609	905	728	941	862	990	956	1000	1000	1000
Exponential	119	49	158	91	270	120	396	175	769	339
Laplace	205	12	420	41	664	99	811	199	983	573
Logistic	60	61	160	91	307	172	515	257	897	680
Cauchy	591	—	716	—	874	—	962	—	1000	—
Extreme value	492	158	592	218	762	286	899	363	1000	617
Gamma (.2)	291	0	364	0	484	101	627	140	928	310
Gamma (.5)	167	0	214	19	297	27	369	37	665	80
Gamma (.7)	130	31	128	37	204	72	256	101	518	178
Gamma (3.)	13	55	33	77	79	90	161	148	412	289
Weibull (.4)	520	0	575	0	590	0	645	0	868	16
Weibull (.6)	69	0	85	0	190	11	285	61	672	160
Weibull (.8)	13	0	30	25	61	85	123	108	455	204
Weibull (1.5)	169	78	258	102	422	132	531	186	642	405
Lognormal (.2)	90	143	128	185	168	260	218	364	224	669
Lognormal (.4)	86	42	125	74	161	104	170	143	72	295
Lognormal (.6)	106	0	142	6	194	32	233	31	130	64
Lognormal (1.)	205	0	257	0	299	0	329	6	447	59

— Moments do not exist.

It is shown that Z_i^2 has a Beta $\{p/2, (n-p-1)/2\}$ distribution.[9] The *squared radii* are location scale and correlation invariant (i.e., the corresponding distribution depends on the number of variables p, and sample size n only). Thus, it is reasonable to develop a procedure of distribution identification on the basis of the *squared radii* Z_i^2. If the true distribution is a multivariate normal, then the corresponding problem is reduced to identifying the univariate *Beta* distribution with parameters $p/2$ and $(n-p-1)/2$, for a given multivariate sample of size n. In this case the univariate $Q_n^{(k)}$ procedure can be used to test the null hypothesis that the underlying distribution comes from a multivariate normal distribution.

If the true distribution is not a multivariate normal, again we propose to utilize the statistic Z_i^2 for distribution identification problems. As in the univariate case the transformed observations Z_i^2 $(i = 1, 2, \ldots, n)$ are used to obtain the statistic $Q_n^{(k)}$. For a given set of p and n values an identification chart similar to ones in Figure 5.1 can be obtained by determining the expected points corresponding to various distributions. Thus the identification procedure would consist of comparing the sample point $Q_n^{(k)}$ based on the *squared radii* with the expected points, and finding the nearest neighboring point of a distribution to estimate the true distribution.

A chart for identifying multivariate distributions is given in Figure 5.3 for $n = 50$, $k = 1$, and $p = 2$. The same set of univariate distributions are used in univariate distribution identification (see Figure 5.1) are also used to generate bivariate samples. Because the $Q_n^{(k)}$ procedure is a location, scale and covariance invariant procedure, the bivariate samples were obtained from the corresponding marginal distributions. For example the bivariate exponential sample is formed by generating two independent samples from the univariate exponential

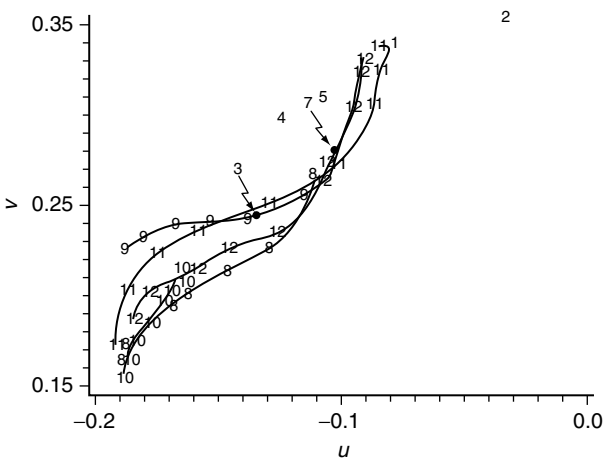

FIGURE 5.3 Identification chart for bivariate distributions based on 1000 samples $(n = 50)$.

distribution. For larger dimensions the same procedure can also be applied to generate multivariate samples. The only difficulty of employing such an algorithm to generate multivariate samples is to determine the joint probability density function for a given covariance structure. However, there are various classes of multivariate distributions whose distribution functions are known and random samples may be generated from these distributions without facing any serious difficulty. For example elliptically contoured distributions have been used widely in Monte Carlo work since they are easy to generate and cover a wide range of distributions.[10]

It is seen that Figure 5.3 displays a useful characterization of the bivariate distributions. For a given sample of size n, one could obtain the transformed values $Z_1^2, Z_2^2, \ldots, Z_n^2$ and compute the corresponding $Q_n^{(k)}$ statistic. The next step for identifying the true distribution will be to plot the point $Q_n^{(k)}$ on the identification chart in Figure 5.3 and determine the nearest neighboring distribution.

It should be pointed out that our proposed distribution approximation procedure for multivariate distributions can be more conveniently used for the distributions for which the joint probability density function is a function of the quadratic form $(\mathbf{X} - \mu)^{\mathrm{T}}\mathbf{V}^{-1}(\mathbf{X} - \mu)$ where μ and $\mathbf{V}$ are the mean vector and covariance matrix, respectively, and T denotes transpose. The elliptically contoured distributions including the multivariate normal distribution are examples of this class of distributions. For these distributions there is a one-to one correspondence between the point $Q_n^{(k)}$ on the identification chart and the corresponding distribution.

5.1.5. Parameter Estimation

Once the true distribution is identified, the next step is to estimate parameters of the underlying distribution by using one of the standard methods of estimation. However, the proposed method of distribution identification lends itself to develop a simple method of estimation for the parameters of the identified distribution. In this section we discuss the estimation of the location and scale parameters based on the $Q_n^{(k)}$ statistic for univariate distributions. An approximate method for estimating the shape parameters will also be explained in this section.

5.1.5.1. Estimating the Location and Scale Parameters

Suppose that the random sample $X_1, X_2, \ldots, X_n$ comes from the distribution function $F(x; \alpha, \beta)$ where α and β are the location and scale parameters respectively. Let $\mu_{1:n}, \mu_{2:n}, \ldots, \mu_{n:n}$ denote the expected values of order statistics from the standard distribution $F_0\{(x - \alpha)/\beta\}$. Also let $\theta_i = \pi\Phi(m_{i:n})$ as defined in Equation 5.1 and $\theta_i' = \pi F_0(\mu_{i:n})$, $i = 1, 2, \ldots, n$. We consider the following statistics

$$T_1 = \sum(\cos\ \theta_i)X_{i:n} \tag{5.7}$$

$$T_2 = \sum(\sin\theta_i)X_{i:n} \tag{5.8}$$

The statistics T_1 and T_2 are linear combinations of the sample order statistics and their expected values may be written as

$$E(T_1) = a\alpha + b\beta \tag{5.9}$$

$$E(T_2) = c\alpha + d\beta \tag{5.10}$$

where $a = \sum\cos\theta_i$, $b = \sum\mu_{i:n}\cos\theta_i$, $c = \sum\sin\theta_i$, $d = \sum\mu_{i:n}\sin\theta_i$. For the normal distribution, it can be shown that $a = 0$. Thus, the statistics

$$\hat{\alpha} = (T_2 - d\hat{\beta})/c \tag{5.11}$$

$$\hat{\beta} = T_1/b \tag{5.12}$$

are unbiased estimators of the parameters α and β, respectively. For symmetric distributions (i.e., distributions for which $\mu_{i:n} = -\mu_{n-i+1}$ $(i = 1, 2, \ldots, n)$) it can be shown that $d = 0$. This result can be applied to Equation 5.11 and Equation 5.12 to give

$$\hat{\alpha} = \sum(\sin\theta_i)X_{i:n}/c = T_2/c \tag{5.13}$$

$$\hat{\beta} = \sum(\cos\theta_i)X_{i:n}/b = T_1/b \tag{5.14}$$

For this special case, it is interesting to note that, up to a constant, the statistic T_1 estimates the location parameter and T_2 estimates the scale parameter.

In this estimation procedure, we used the standard normal distribution as a *reference distribution*. One could also consider using any other member of location-scale family of distributions as a *reference distribution*. In this case the underlying procedure could be modified by substituting $\theta_i' = \theta_i$ in the corresponding equations.

We now illustrate the proposed method for estimating the parameters of the normal distribution. The sample mean $\bar{X}$ is the minimum variance unbiased estimator of the location parameter μ. Estimation of the scale parameter σ has been studied extensively and various estimators have been proposed for this purpose. A review of the underlying estimators are given by Johnson and Kotz.[3]

For the normal distribution $\mu_{i:n} = -m_{i:n}$ where $m_{i:n}$, $(i = 1, 2, \ldots, n)$ are expected standard normal scores. Also, $b = \sum m_{i:n}\cos\theta_i$ and $c = \sum\sin\theta_i$. Substituting these quantities in Equation 5.13 and Equation 5.14 yields

$$\hat{\mu} = \sum(\sin\theta_i)X_{i:n}/c = T_2/c \tag{5.15}$$

$$\hat{\sigma} = \sum(\cos\theta_i)X_{i:n}/b = T_1/b \tag{5.16}$$

where $\hat{\mu}$ and $\hat{\sigma}$ are unbiased estimators of μ and σ. For large n it is well-known that $\Phi(m_{i:n}) \simeq i/(n+1)$ and this result can be used to simplify the above

TABLE 5.2
Efficiencies of $\hat{\mu}$ and $\hat{\sigma}$ Relative to the Least Squares Estimates

n	2	3	4	5	6	7	8	9	10
$\hat{\mu}$	100.0	98.7	97.8	97.1	96.6	96.2	96.0	96.0	95.5
$\hat{\sigma}$	100.0	100.0	99.5	99.0	98.6	98.3	98.0	96.3	97.6

expressions. It follows from Equation 5.15 and Equation 5.16 that the variances of $\hat{\mu}$ and $\hat{\sigma}$ are

$$V(\hat{\mu}) = \sigma^2 \mathbf{k}' \mathbf{V} \mathbf{k}/c^2 \tag{5.17}$$

$$V(\hat{\sigma}) = \sigma^2 \mathbf{l}' \mathbf{V} \mathbf{l}/c^2 \tag{5.18}$$

where $\mathbf{k}' = (\sin\theta_1, \ldots, \sin\theta_n)$ and $\mathbf{l}' = (\cos\theta_1, \ldots, \cos\theta_n)$ and $\mathbf{V}$ is the covariance matrix of the standard normal order statistics.

Relative efficiencies of $\hat{\mu}$ and $\hat{\sigma}$ compared with the corresponding best linear unbiased estimators for $n = 2, 3, \ldots 10$ are shown in Table 5.2. From this table it can be seen that both estimators have quite high (over 95%) efficiencies relative to the best linear estimators. Furthermore, estimator of the scale parameter is always more efficient than those based on mean deviation and range (see Johnson and Kotz[3] for efficiencies of the underlying estimators). In view of the accuracy obtainable with best linear unbiased estimators, it may not be reasonable to consider use of proposed estimators. However the linear coefficients used to compute the best linear unbiased estimator of σ are difficult to obtain. On the other hand if a $Q_n^{(k)}$ test is performed to identify the true distribution, the statistics T_1 and T_2 will already be available. Thus the estimators $\hat{\mu}$ and $\hat{\sigma}$ can easily be computed from Equation 5.15 and Equation 5.16.

5.1.5.2. Estimating the Shape Parameters

The $Q_n^{(k)}$ procedure can also be used to estimate the shape parameters of a given distribution. Let γ and λ denote shape parameters of a distribution. Since the statistics $U_n^{(k)}$ and $V_n^{(k)}$ are location and scale invariant, the expected value of the point $Q_n^{(k)}$ will be some function of sample size n and the shape parameters γ and λ. As explained in previous sections, given any distribution with a given set of parameter values γ_0 and λ_0, one can determine its location on a $(U_n^{(k)}, V_n^{(k)})$ plane.

We first consider the distributions with one shape parameter (say γ) only. Suppose that the statistics $U_n^{(k)}$ and $V_n^{(k)}$ are based on the random sample $X_1, X_2, \ldots, X_n$ from a distribution $F(x; \alpha, \beta, \gamma)$, where, α and β are the location and scale parameters, respectively, and γ is the shape parameter. The expected values of $U_n^{(k)}$ and $V_n^{(k)}$ can be expressed as

$$E(U_n^{(k)}) = \varphi_1(n, \gamma) \tag{5.19}$$

$$E(V_n^{(k)}) = \varphi_2(n, \gamma) \tag{5.20}$$

where φ_1 and φ_2 are some functions of n and γ. For a given sample size n and shape parameter $\gamma = \gamma_0$ the corresponding expected point $EQ_n^{(k)}(\gamma_0) = \{\varphi_1(n, \gamma_0), \varphi_2(n, \gamma_0)\}$ can be determined approximately in the $\{U_n^{(k)}, V_n^{(k)}\}$ plane (see Figure 5.1).

Our proposed method of estimation of the shape parameter γ is based on finding a point $E\{Q_n^{(k)}(\gamma)\}$ such that

$$U_n^{(k)} = \varphi_1(n, \hat{\gamma}) \tag{5.21}$$

$$V_n^{(k)} = \varphi_2(n, \hat{\gamma}) \tag{5.22}$$

where $\hat{\gamma}$ is the sample estimator of γ. However, in a typical situation the sample point may not fall exactly on the curve of one of the distribution families (with one shape parameter). Let the points $E\{Q_{1n}^{(k)}(\gamma_1)\} = (u_1, v_1)$ and $E\{Q_{2n}^{(k)}(\gamma_2)\} = (u_2, v_2)$ denote the expected points based on $\gamma = \gamma_1$ and $\gamma = \gamma_2$ values, respectively. The sample point $Q_n^{(k)} = \{U_n^{(k)}, V_n^{(k)}\}$ lies either on, above or below the curve of the identified distribution. Assuming that the a linear interpolation provides a satisfactory approximation, we proceed to estimate γ approximately by

$$\hat{\gamma} \simeq \gamma_1 + (\gamma_2 - \gamma_1)(x_0 - u_1)/(u_2 - u_1) \tag{5.23}$$

where

$$x_0 = \frac{A(V_n^{(k)} - v_1) + A^2 u_1 + U_n^{(k)}}{A^2 + 1} \tag{5.24}$$

and

$$A = \frac{v_2 - v_1}{u_2 - u_1} \tag{5.25}$$

The accuracy of the estimator $\hat{\gamma}$ is closely related to the distance of the sample point $Q_n^{(k)}$ to the curve of the identified true distribution. However, more accurate results may be obtained by employing a nonlinear interpolation.

The proposed estimation procedure can be extended to the two-shape parameter case also. In this case one needs to choose at least three points $E\{Q_{1n}^{(k)}(\gamma_1, \lambda_1)\}$, $E\{Q_{2n}^{(k)}(\gamma_2, \lambda_2)\}$, and $E\{Q_{3n}^{(k)}(\gamma_3, \lambda_3)\}$ in such a way that the sample point $Q_n^{(k)}$ falls inside a triangle whose vertices are determined by the underlying expected points. Again by using a linear interpolation in the plane, an approximate solution can be obtained for the parameter estimates.

5.1.6. Distribution Approximation for Mixtures of Distributions

In this section we discuss the possibility, under certain assumptions, of using the proposed procedure to estimate both the distribution and its parameters for mixtures of distributions.

Let $F(x;\alpha,\beta)$ denote the distribution function of a random variable X and α and β are the location and scale parameters respectively. The contaminated (mixed) distribution is defined by

$$pF(x;\alpha,\beta)+(1-p)F(x;\alpha+\delta,k\beta) \tag{5.26}$$

where p is the *mixing* parameter ($0 \le p \le 1$). The distribution is location contaminated for $\delta \neq 0$, $k=1$, scale contaminated for $\delta=0$, $k\neq 1$ and location and scale contaminated for $\delta \neq 0$, $k \neq 1$.

The distribution function in Equation 5.26 has five unknown parameters. It can be shown for this distribution that the expected values of the statistics $U_n^{(k)}$ and $V_n^{(k)}$ depend on the parameters p, k, and $\phi=\delta/\beta$. A point on the identification chart usually does not represent a distribution uniquely. Therefore the proposed approximation procedure cannot be used for this general case. However if one of these parameters is known or could be estimated independently then an identification chart based on $U_n^{(k)}$ and $V_n^{(k)}$ statistics could be obtained and the parameters of the identified distribution could be estimated by following a similar procedure as explained in the previous sections.

As an example, consider the location contaminated normal distribution. In this case we assume that two components of mixture of normal distributions have equal variances ($\sigma_1^2=\sigma_2^2=\sigma^2$). The corresponding distribution function may be written as

$$F(x;\mu,\sigma,p,\delta)-p\Phi(x;\mu,\sigma)+(1-p)\Phi(x;\mu+\delta,\sigma) \tag{5.27}$$

where $\Phi(\cdot)$ is the distribution function of normal distribution. The mean and the variance of this distribution are

$$E(X)=\mu+(1-p)\delta \tag{5.28}$$

$$\mathrm{Var}(X)=\sigma^2+p(1-p)\delta^2 \tag{5.29}$$

Preston[11] used the sample skewness and kurtosis coefficients to estimate p and $\phi=\delta/\sigma$. He showed further that one could back-track with given sample mean and variance to obtain the estimates of μ, σ, δ and p. It is obvious for this case that the distribution of the statistics $U_n^{(k)}$ and $V_n^{(k)}$ depends on the parameters p and ϕ. By following a similar approach, one could obtain the estimates of these parameters by using the $U_n^{(k)}$ and $V_n^{(k)}$ statistics. An example of the chart based on location contaminated normal distributions (for $n=50$ and $k=1$) is given in Figure 5.4. The charts in Figure 5.4 shows system of contours each of which is

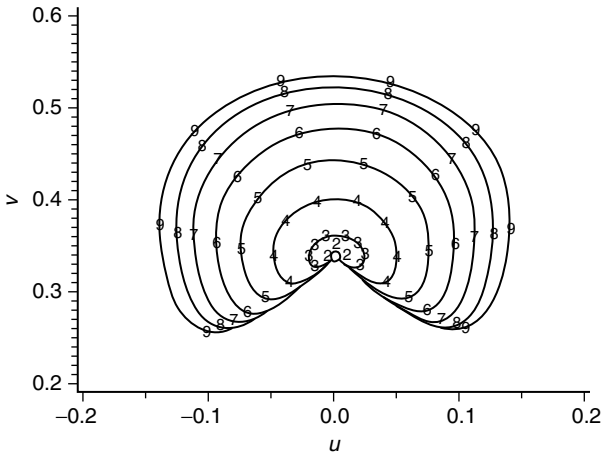

FIGURE 5.4 Identification chart for location contaminated normal distributions ($n = 50$).

obtained by joining the points based on various p values ($p = 0.05, 0.10, \ldots 0.95$) and $\phi = \delta/\sigma$ values ($\phi = 1, 2, 3, 4, 5, 6, 7, 8, 9$).

For a given sample of size n, the corresponding point $Q_n^{(k)}$ in the $\{U_n^{(k)}, V_n^{(k)}\}$ plane can be plotted and the estimates of p and ϕ can be obtained approximately by a linear interpolation. Given the estimates $\hat{p}$ and $\hat{\phi} = \hat{\delta}/\hat{\sigma}$ and the moment estimates of the quantities in Equation 5.28 and Equation 5.29, the estimates of the parameters can be determined as

$$\hat{\sigma} = s/\{1 + \hat{p}(1 - \hat{p})\hat{\phi}^2\}^{1/2} \tag{5.30}$$

$$\hat{\mu} = \bar{X} - (1 - \hat{p})\hat{\phi}\hat{\sigma} \tag{5.31}$$

$$\hat{\delta} = \hat{\phi}\hat{\sigma} \tag{5.32}$$

where $\bar{X}$ and S are the sample mean and standard deviation, respectively.

5.1.7. Examples

Example 1. Consider the following hypothetical sample of 50 observations which were generated from the standard gamma distribution with shape parameter $\gamma = 0.5$.

0.507 0.175 0.388 1.135 1.958 0.156 1.131 0.001 0.229 0.550 0.138 1.571
3.571 0.152 0.276 0.570 0.323 0.009 0.362 0.001 0.220 0.181 0.452 1.640
0.001 0.559 0.230 0.350 0.037 0.275 0.062 1.499 0.048 0.818 0.110 0.010
0.335 1.866 0.001 0.109 0.239 0.328 0.802 0.123 0.021 0.141 0.448 1.977
0.849 0.039

We use the chart in Figure 5.1 to illustrate the identification and parameter estimation procedure. The computed statistics for this sample are $U_{50}^{(1)} = -0.1321$ and $V_{50}^{(1)} = 0.3063$. Plotting this point on the identification chart, we see that the nearest neighbor distribution is the gamma distribution.

The coordinates of the nearest points and the corresponding γ values on the *gamma curve* are

$$Q_1' = (U_1', V_1') = (-0.1513, 0.2987), \quad \gamma_1 = 0.3$$

$$Q_2' = (U_2', V_1') = (-0.1268, 0.3143), \quad \gamma_1 = 0.5$$

From Equation 5.24 and Equation 5.25 we find that $A = (0.3143 - 0.2987)/(-0.1268 + 0.1513) = 0.6367$, and $x_0 = -0.1342$. Substituting these values in Equation 5.23 gives $\hat{\gamma} = 0.3 + (0.2)(-0.1342 + 0.1513)/(-0.1268 + 0.1513) = 0.4397$. Note that the shape parameter is estimated independently from the scale parameter β. To estimate β for this particular case, we use the well-known relationship between the parameters, i.e., $\gamma\beta = \mu$ where μ is the mean of the distribution. If we use the sample mean $\bar{X} = \sum X_i/n$ as an estimate of μ then the scale parameter β may be estimated as $\hat{\beta} = \bar{X}/\hat{\gamma} = 1.229$. The moment and the maximum likelihood estimate (MLE) of γ and β are found to be as follows

Moment estimates	$\tilde{\gamma} = 0.6005$	$\tilde{\beta} = 0.9183$
MLE estimates	$\hat{\gamma}^* = 0.5288$	$\hat{\beta}^* = 1.0219$

If the underlying distribution is assumed to be a three-parameter gamma distribution, then the chart in Figure 5.1 and Equations 5.11 and 5.12 could be used to estimate the location and scale parameters. An obvious advantage of the proposed method of estimation over the method of maximum likelihood is that, the shape parameter is estimated independently and estimation of location and scale parameters do not require iterative computation, which in many cases causes some difficulties.[12]

Example 2. In this example we consider well-known *iris* data of Fisher.[13] Four measurements namely *sepal length*, *sepal width*, *petal length* and *petal width* were taken on each of 50 plants from three varieties (*Iris setosa*, *I. versicolor* and *I. virginica*). This data set has been examined by a number of authors including Ozturk,[4] Royston,[14] Koziol[8] and Small[15] who demonstrated their goodness-of-fit test for multivariate normality. Our objective here is not to perform a formal multivariate normality test but to demonstrate our multivariate distribution identification procedure. However the underlying procedure could be modified by providing a $100(1 - \alpha)\%$ confidence ellipse to test the null hypothesis that each sample comes from a multivariate normal distribution.

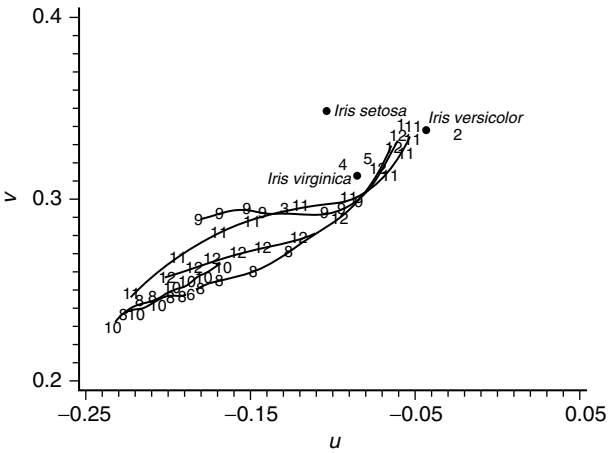

FIGURE 5.5 Identification chart for *Iris* data ($p = 4$, $n = 50$).

For each variety the following statistics (based on type 1 standardization) were computed

$$\begin{aligned} \textit{Iris setosa} \quad & U_{50}^{(1)} = -0.0999, \quad V_{50}^{(1)} = 0.3520 \\ \textit{Iris versicolor} \quad & U_{50}^{(1)} = -0.0453, \quad V_{50}^{(1)} = 0.3407 \\ \textit{Iris virginica} \quad & U_{50}^{(1)} = -0.0878, \quad V_{50}^{(1)} = 0.3136 \end{aligned}$$

Plots of points corresponding to each variate are shown on the identification chart in Figure 5.5. It is seen from the chart that the points representing *I. versicolor* and *I. virginica* are close to the point of multivariate normal distribution (with $p = 4$ and $n = 50$) whereas the point representing *I. setosa* data departs slightly from the underlying point. Evidence for a departure from multivariate normality with *I. setosa* data was reported by Small (1980) at $P = 0.03$ and Ozturk (1991) at $P = 0.07$ while many other procedures did not provide evidence to reject the multivariate normality assumption.

5.1.8. Conclusions

This paper has presented a procedure of distribution approximation for univariate and multivariate distributions through Monte Carlo experiments. In addition, a technique is given for the estimation of parameters of the univariate distributions. Based on the empirical results it is shown that the proposed procedure provides a good basis for approximating the distributions in a nonparametric sense. However, the accuracy of the proposed method among the other factors depends on the number of samples used in Monte Carlo experiments. Therefore there is

a trade-off between the accuracy and the computational time; the accuracy can be increased at the expense of increasing the computational time and *vice versa*.

5.2. A GENERAL ALGORITHM FOR UNIVARIATE AND MULTIVARIATE GOODNESS-OF-FIT TESTS BASED ON GRAPHICAL REPRESENTATION

5.2.1. Introduction

This paper develops a general algorithm for univariate and multivariate goodness-of-fit tests based on graphical representation, using the same principle as explained in a recent report.[1] The derivation of the test statistic Q_n for testing whether a complete sample comes from a specified univariate location-scale family of distributions, is explained in the above reference. Using the normal distribution as an example, it is shown that the proposed procedure provides not only a powerful test for one or more independent complete samples but also an interesting graphical display for identifying the true distribution.

In testing a null hypothesis about a distributional assumption against an unspecified alternative there is generally no uniformly most powerful nor optimal test.[2] Because of this, various test procedures have been developed for assessing the distributional assumptions. Under certain conditions (i.e., for a specified null hypothesis, a specified sample size, and a predetermined level of significance) one test procedure may be shown to be more powerful than the other existing procedures. Besides the power of a given test, computational simplicity, desirable distributional properties of the test statistic and the generality of the test procedure are among the important points to be considered.

The χ^2 goodness-of-fit test has been widely used for assessing the distributional assumptions because of its generality and its computational simplicity. However, the choice of the class intervals for computing the test statistic is arbitrary and the procedure can only be used for moderate or large samples. Tests based on the probability integral transformation have also been widely used to test the null hypothesis for a completely specified distribution. A number of test statistics, including the Kolmogorov–Smirnov D and Anderson–Darling A^2 statistics which are based on the discrepancy between the empirical distribution function of the sample and the null distribution function, have been developed. In the case when the parameters of the null distribution are estimated from the sample, the above procedures are modified to obtain more general tests (for example see Lilliefors[3]). Comparative studies showed that the D statistic has higher power than that of the χ^2 test for many alternatives.[4,5]

The use of probability plotting as a quick means of checking on a hypothesized distribution provides a most valuable aid. For a given random sample, using this informal approach one can visually assess the distributional assumptions.[6] Probability plotting has been used as a basis for developing some formal tests of goodness-of-fit. The W test for normality, suggested by Shapiro

and Wilk[7], has stimulated fresh interest in the subject. Power comparisons based on Monte Carlo studies have shown that the W test provides an omnibus test for normality.[8] A number of modifications of the underlying procedure based on the same principles have been introduced for testing various null distributions.[9,10] Extensions of the W test to other null distributions have not been as successful as in the normal case.[11,12] Stephens[13,14] gave an excellent review of the tests based on the W statistic and the empirical distribution function statistics.

There has been an increasing research movement in developing graphical goodness-of-fit tests. A major force behind this development has been the rapid changes in computer technology. High speed computers having graphical facilities are becoming available to many statisticians. The current state of the computer technology reduced the computational limitations and provided a convenient environment for the development of inventive methods.

In this paper we aim to present a general algorithm based on Monte Carlo simulations for testing the distributional assumptions. The underlying method uses the sample order statistics to display sample patterns graphically. It will be demonstrated through examples that the proposed algorithm can be used to test any distributional assumptions (not limited to the location-scale family) for the univariate–multivariate, one sample and several sample cases.

In the following sections, details of the proposed algorithm are explained. Section 5.2.2 introduces basic notation. Section 5.2.3 presents general properties of the test statistics and Section 5.2.4 explains extensions of the test. Section 5.2.5 is devoted to some examples to illustrate the test procedure for various cases. Section 5.2.6 explains the empirical power of the test and Section 5.2.7 summarizes the general algorithm for the test procedure. Finally, Section 5.2.8 discusses some advantages of the proposed procedure over probability plotting.

5.2.2. The Test Procedure

Let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ be an ordered sample from the distribution function F and suppose that the location-scale distribution $F_0\{(x-\mu)/\sigma\}$ is proposed as the null distribution. In this section we develop a general algorithm for testing the null hypothesis that the sample comes from the hypothesized distribution $F_0\{(x-\mu)/\sigma\}$. The extensions of the procedure to testing for the distributions having unknown shape parameters, and multivariate normal distributions will be given in Section 5.2.4.

A number of procedures used for testing distributional assumptions are based on linear combinations of the standardized sample order statistics. For example, the Shapiro–Wilk W statistic can be written in the following form

$$W^{1/2} = \sum_{i=1}^{n} a_i Y_{i:n} \tag{5.33}$$

where $Y_{i:n}$ is the standardized ith sample order statistic, i.e.,

$$Y_{i:n} = (X_{i:n} - \bar{X})/S \tag{5.34}$$

$\bar{X}$ and S are the sample mean and standard deviation, respectively, and a_i ($i = 1, 2, \ldots n$) are the normalized coefficients of the best linear unbiased estimator of σ. A number of modifications of the W statistic have been introduced by using different sets of coefficients a_i.[15–20] Some other modifications of W are introduced on replacing S, the sample estimate of σ by some other linear estimates.[21]

A number of standardized sample order statistics which can be used to construct test statistics, are given in Table 5.3. Some other location-scale invariant statistics could also be added to the list in the table. The first statistic in the table is the standardized absolute deviation from the sample mean and was used by Ozturk and Dudewicz.[1] The Type 2 statistic is a well-known statistic and has been used widely.[9] Type 3 and Type 4 statistics are obtained by replacing the sample mean by the sample median in Type 1 and Type 2 statistics, respectively. The sample range $\Delta = X_{n:n} - X_{i:n}$ is also considered as an alternative measure of dispersion. Type 5 through Type 8 statistics are obtained from the first four statistics by replacing S by Δ in the corresponding expressions respectively. The Type 9 statistic also provides a simple standardization and may be useful for assessing some distributional assumptions. Note that all the statistics listed in Table 5.3 are location and scale invariant.

We define our statistic, for testing the null hypothesis that the random sample is drawn from a specified location-scale distribution (location and scale parameters are unknown), as follows. Let $Y_{j:n}^{(k)}$ denote the jth standardized sample order statistic based on the *Type* k ($k = 1, 2, \ldots, 9$) standardization as defined in Table 5.3. Also let $m_{i:n}$ denote the expected value of the ith sample order statistic from the distribution function F_0. The ith point ($i = 1, 2, \ldots, n$) on a two dimensional plane is defined by

$$Q_i^{(k)} = (U_i^{(k)}, V_i^{(k)}) \tag{5.35}$$

TABLE 5.3
Various Standardized Sample Order Statistics

Number	Expression	Number	Expression
1	$\lvert X_{i:n} - \bar{X}\rvert/S$	2	$(X_{i:n} - \bar{X})/S$
3	$\lvert X_{i:n} - \tilde{X}\rvert/S$	4	$(X_{i:n} - \tilde{X})/S$
5	$\lvert X_{i:n} - \bar{X}\rvert/\Delta$	6	$(X_{i:n} - \bar{X})/\Delta$
7	$\lvert X_{i:n} - \tilde{X}\rvert/\Delta$	8	$(X_{i:n} - \tilde{X})/\Delta$
9	$(X_{i:n} - X_{i:n})/\Delta$	—	—

where

$$U_i^{(k)} = \frac{1}{n}\sum_{j=1}^{i} \cos\{\pi F_0(m_{j:n})\} Y_{j:n}^{(k)} \tag{5.36}$$

$$V_i^{(k)} = \frac{1}{n}\sum_{j=1}^{i} \sin\{\pi F_0(m_{j:n})\} Y_{j:n}^{(k)} \tag{5.37}$$

For a given set of order statistics $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$, one can plot the corresponding points $Q_1, Q_2, \ldots, Q_n$ on a plane in two dimensions. Furthermore, these points can be joined to obtain linked vectors. These linked vectors are based on the sample order statistics and they reflect a certain pattern under the null hypothesis. Following the same procedure, one can also obtain linked vectors using the expected values of $U_i^{(k)}$ and $V_i^{(k)}$ in place of their sample values.

The proposed test is based on the comparison of the sample and the expected linked vectors. If the null hypothesis is true, we expect that the sample linked vectors follow a pattern which is uniformly close to the null pattern.

Note that the procedure explained above provides a graphical goodness-of-fit test. An interesting property of such a test procedure is that any one of the points $Q_i^{(k)}$, $(i = 1, 2, \ldots, n)$, or a selected group of these points can be used as a test statistic to establish a formal test. In this study we propose the following general statistic

$$Q_n^{(k)} = \{U_n^{(k)}, V_n^{(k)}\} \tag{5.38}$$

for testing the null hypothesis. A $100(1-\alpha)\%$ confidence contour for the point $E(Q_n^{(k)}) = \{E(U_n^{(k)}), E(V_n^{(k)})\}$ can also be constructed. Thus, for a specified k, the above statistic together with a linked vectors chart, is expected to provide a useful test procedure.

The test algorithm may be summarized as follows:

(1) Sort the sample observations in increasing order,
(2) Choose the appropriate type of standardization from Table 5.3 and obtain the sample order statistics,
(3) Obtain the sample and the null-linked vectors,
(4) Construct a $100(1-\alpha)\%$ confidence ellipse for the point $E\{Q_n^{(k)}\}$.
(5) If the sample point $Q_n^{(k)}$ does not fall inside the ellipse then reject the null hypothesis.

5.2.3. Properties of the Test Statistics

In general, the test statistic $Q_n^{(k)}$ is a location and scale invariant statistic. Therefore, it could be used to test a composite null hypothesis that the random sample $X_1, X_2, \ldots, X_n$ comes from a specified distribution. The statistics $U_n^{(k)}$ and $V_n^{(k)}$ are both the linear combinations of standardized sample order statistics $Y_{i:n}^{(k)}$, $(i = 1, 2, \ldots, n)$. It is clear that the joint distributions of $U_n^{(k)}$ and $V_n^{(k)}$ depend on

the null distribution, the type of the standardized statistic k and the sample size n. For a given null distribution, type of statistic, and sample size, the joint distribution of $U_n^{(k)}$ and $V_n^{(k)}$ must be obtained in order to perform a formal test of hypothesis.

The distributional properties of the statistic $Q_n^{(1)} = (U_n^{(1)}, V_n^{(1)})$ for testing normality have been studied by Ozturk and Dudewicz.[1] Based on empirical comparisons, they showed that the joint distribution of the underlying statistics converge rapidly to the bivariate normal distribution and that an approximate $100(1 - \alpha)\%$ confidence ellipse for the point $E\{Q_n^{(1)}\}$ could be obtained for $n > 10$.

The exact sampling distribution of $Q_n^{(k)}$ is usually difficult to obtain. However, using the special properties of some of the null distributions, explicit expressions could be obtained for the means, variances and covariances of $U_n^{(k)}$ and $V_n^{(k)}$. On the other hand, asymptotic joint normality of these statistics could be established for some of the standardized statistics and null distributions.

In this study, we use the empirical distribution of the statistic $Q_n^{(k)}$ in order to construct a general algorithm for testing the distributional assumptions. Means, variances and covariances of the statistics $U_n^{(k)}$ and $V_n^{(k)}$ can be obtained by Monte Carlo experiments. Empirical joint distributions of the underlying statistics could also be obtained by plotting the equi-probability contours.

Examples of the empirical distributions of the statistics $U_n^{(k)}$ and $V_n^{(k)}$ for $k = 1, 2, \ldots, 9$ are given in Figure 5.6. Each distribution is obtained by generating 100,000 random samples of size 50 from the normal distribution. Subroutine DRNNOA in the IMSL (International Mathematical and Statistical Libraries) is used to generate the random samples. It is seen from the figure that the joint distribution of the statistics, $U_n^{(k)}$ and $V_n^{(k)}$, can be approximated by a bivariate normal distribution for $k = 1$, 3 and 6. For $k = 2$, 4, 5, 7, 8 and 9 at least one of the corresponding marginal distributions are symmetric.

Plotting confidence contours is usually not easy when the joint distribution is not bivariate normal. However, if a normality transformation could be made to obtain statistics $Z_U = \psi_1(U_n^{(k)})$ and $Z_V = \psi_2(V_n^{(k)})$ then the corresponding confidence contours could be constructed. In this study we utilize the Johnson system of distributions to transform the distribution of $Q_n^{(k)}$ into a standard bivariate normal distribution.

Let

$$Z_U = \gamma_1 + \delta_1 f_i\{(U_n^{(k)} - \xi_1)/\lambda_1\} \tag{5.39}$$

$$Z_V = \gamma_2 + \delta_2 f_i\{(V_n^{(k)} - \xi_2)/\lambda_2\} \tag{5.40}$$

where $f_1(y) = \log(y)$, $f_2(y) = \log\{y/(1 - y)\}$, $f_3(y) = \sinh^{-1}(y)$ and $f_4(y) = y$. If $U_n^{(k)}$ has a Johnson Type i distribution, and $V_n^{(k)}$ has a Johnson Type j distribution, then the standardized variables Z_U and Z_V have a joint standard bivariate normal distribution with correlation coefficient ρ. Thus the confidence contour, within which $100(1 - \alpha)\%$ of the distribution lies, could be constructed.

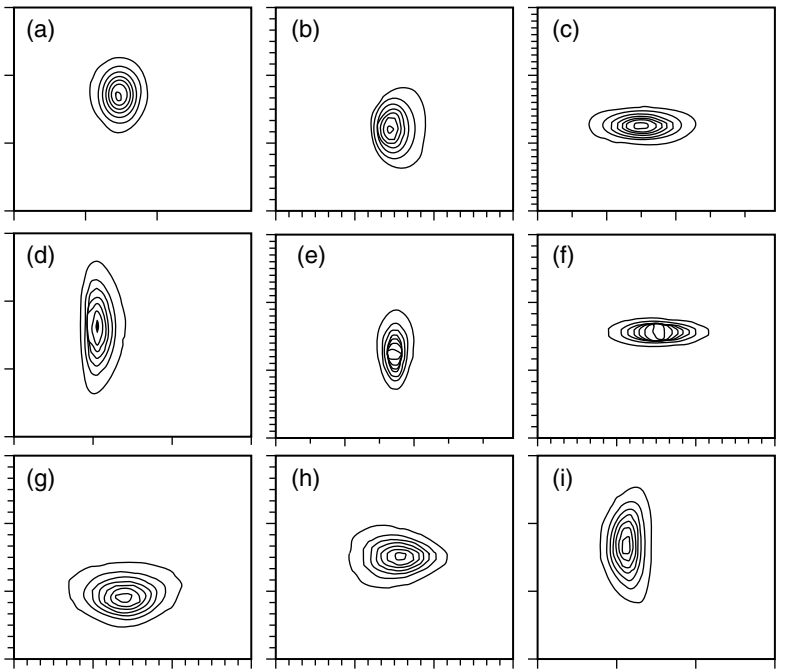

FIGURE 5.6 Empirical distribution of the statistic $Q_n^{(k)}$ for $n = 50$. (a) Type 1, (b) Type 2, (c) Type 3, (d) Type 4, (e) Type 5, (f) Type 6, (g) Type 7, (h) Type 8, (i) Type 9.

The corresponding contour has the equation[31]

$$Z_U^2 - 2\rho Z_U Z_V + Z_V^2 = -2(1-\rho^2)\log(\alpha) \tag{5.41}$$

When the statistics $U_n^{(k)}$ and $V_n^{(k)}$ are uncorrelated, the above equation will have a simpler form. By plotting the contours of equal density of the joint distribution of $U_n^{(k)}$ and $V_n^{(k)}$, $100(1-\alpha)\%$ confidence contours can be obtained.

It is clear that the confidence contours of the statistics $U_n^{(k)}$ and $V_n^{(k)}$ can be obtained if the parameters ρ, γ_i, δ_i, ξ_i and λ_i $(i = 1, 2)$ are known. However, these parameters usually are not known and they must be determined for every null distribution, sample size (n) and type of standardization (k). In this study, we proceed to determine the corresponding parameter values empirically. For this purpose, a number of random samples of size n are generated from the hypothesized distribution. The corresponding $U_n^{(k)}$ and $V_n^{(k)}$ statistics are then computed. Parameters of the distributions in Equation 5.39 and Equation 5.40 are determined by using the estimation method suggested by Slifker and Shapiro.[22] The correlation coefficient ρ is determined from the transformed values Z_U and Z_V of each sample.

5.2.4. Extensions of the Test

The proposed procedure is developed to test the composite hypothesis that a sample comes from a univariate location scale distribution. In this section we explain some extensions of the underlying procedure for more general cases. Examples for each of the corresponding extensions will be given in the next section.

(1) *Testing two or more independent samples*: The generalization of the $Q_n^{(k)}$ test to more than one independent sample is explained by Ozturk and Dudewicz[1] and will be summarized here for completeness. For this case, one constructs the null-linked vectors and the sample linked vectors for each sample. If the level of significance for testing one sample is α then the corresponding overall level for significance α' for testing r independent samples would be $1-(1-r)^r$. Thus $100(1-a')\%$ confidence ellipses must be obtained. The null hypothesis is rejected if at least one of the points $Q_n^{(k)}$ for each sample falls outside the corresponding confidence ellipse.

(2) *Testing for distributions with unknown shape parameters*: Testing for distributions with shape parameters is conceptually different than testing for distributions without shape parameters. In the first case the shape of the distribution depends on the corresponding values of the shape parameter(s) while in the second case the null distribution is transformed to a single form.

If the null distribution has shape parameter(s) then we extend the proposed procedure by using the sample estimates of the underlying parameters. In this case the null and the sample linked vectors, and the corresponding confidence ellipse could be constructed following the same procedure as explained in Section 5.2.2. Then the distribution of the $Q_n^{(k)}$ statistic can be obtained empirically based on the null distribution for which the shape parameter is estimated from the sample.

(3) *Testing for multivariate normality*: A number of test procedures for multivariate normality have been proposed in the literature (see Cox and Small[23]). A common approach for detecting the presence of joint nonnormality has been to perform separate tests of univariate normality on the marginal distributions. For this purpose, our proposed graphical goodness-of-fit procedure could be used to test the marginal normality of the observations on each variable. Distributional assumptions for the other multivariate distributions could also be tested using the same $Q_n^{(k)}$ procedure. However, the derivation of the true significance level for such test procedures is a multivariate problem and usually will be unknown.

Another approach for testing multivariate normality is to develop test statistics which explicitly exploit the multivariate nature of the data.[24] The test

procedures in this category concentrate, either on combinations of univariate normality tests or on the invariant graphical procedures, based on the distribution of the squared radii r_i^2.

In this section we demonstrate the use of the $Q_n^{(k)}$ procedure for testing multivariate normality on the basis of the distribution of the ordered r_i^2 values of the observations.

Let $X_1, X_2, \ldots, X_n$ be n independent $p \times 1$ vectors of observations. Then the squared radius for the ith observation is

$$r_i^2 = (X_i - \bar{X})^{\mathrm{T}} S^{-1} (X_i - \bar{X}) \tag{5.42}$$

where $\bar{X}$ is the mean vector, S is the estimated covariance matrix, and T denotes transpose. Under the null hypothesis that the random sample is drawn from a multivariate normal distribution, the distribution of the r_i^2 is a constant multiple of a Beta distribution with parameters $\frac{1}{2}p$ and $\frac{1}{2}(n-p-1)$.[25] Apparently $r_i^2 = (i = 1, 2, \ldots, n)$ are not uncorrelated. However, for large samples the correlation between the r_i^2 may be ignored.[26] In this case the problem of testing the random sample for multivariate normality is reduced to testing the transformed sample $r_1^2, r_2^2, \ldots, r_n^2$ for the Beta $\{\frac{1}{2}p, \frac{1}{2}(n-p-1)\}$ distribution. Note that the $Q_n^{(k)}$ test not only provides a graphical procedure but also a significance test for multivariate normality.

A widely used graphical procedure for assessing joint normality based on *radius and angles* representation,[27] also provides a useful informal graphical procedure. Following a similar procedure, one can obtain the probability integral transformation of r_i^2, values given in Equation 5.42 and $p-1$ normalized angles $\theta_{ij} (i = 1, 2, \ldots, n;\ j = 1, 2, \ldots, p-1)$. The random variables will be approximately independent and uniformly distributed over (0,1). Thus a formal test for multivariate normality could also be constructed by using the $Q_n^{(k)}$ procedure. For this case, the corresponding problem consists of testing p independent samples for uniformity.

5.2.5. Examples

We consider the well-known *iris* data, first used by Fisher.[28] Four variates namely *sepal length*, *sepal width*, *petal length and petal width* were measured on 50 plants from each variety *Iris setosa*, *I. versicolor* and *I. virginica*. The complete data set is given by Kendall.[29] In this study, we use the data for *I. setosa* only as the data for *I. versicolor* and *I. virginica* are less interesting.

Example 1. We illustrate the univariate $Q_n^{(k)}$ procedure on the variety *I. setosa* first with $k = 1$. Our objective is to perform normality tests for each variety separately. The null and sample linked vectors with 90, 95 and 99% confidence contours are given in Figure 5.7(a). To construct these contours, 2000 random samples of size 50 were generated from the standard normal distribution and the statistics $U_n^{(1)}$ and $V_n^{(1)}$ were computed for each sample. The Johnson system of distributions were fitted to the data sets each of which is based on the statistics

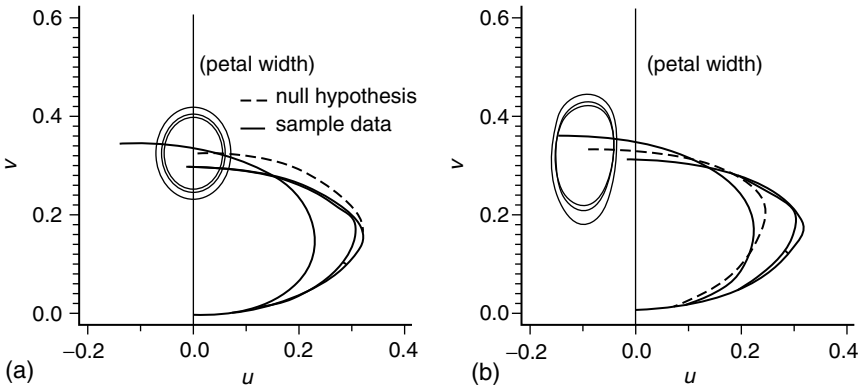

FIGURE 5.7 The $Q_n^{(1)}$ test for *Iris setosa* data: (a) testing for normality, (b) testing for exponentiality.

$U_n^{(1)}$ and $V_n^{(1)}$. Parameters of the Johnson distributions were determined by using the estimation method suggested by Slifker and Shapiro.[22] Linked vector charts for the variates *sepal length*, *sepal width*, *petal length and petal width* are shown in the same figure. The P values corresponding to the $Q_n^{(1)}$ statistics for the variates *sepal length*, *sepal width*, *petal length* and *petal width* are found to be $P_1 = 0.58$, $P_2 = 0.46$, $P_3 = 0.62$, $P_4 = 0.00$, respectively. A *P value* obtained from Equation 5.41 is the probability that the point $Q_n^{(1)} = \{U_n^{(1)}, V_n^{(1)}\}$ falls outside of the $100(1 - P)\%$ confidence contour of the distribution. A small *P value* indicates that the null hypothesis can be rejected at some level of significance.

If the null distribution were exponential the corresponding linked vectors chart could be obtained similarly [Figure 5.7(b)]. For this case the *P values* of the four variates are found to be $P_1 = 0.00$, $P_2 = 0.00$, $P_3 = 0.00$ and $P_4 = 0.02$.

The above tests could be performed also by utilizing the other statistics which are given in Table 5.3. Some examples based on testing the *sepal length* of *I. setosa* for normality are given in Figure 5.8 for $k = 2(1)\,9$. The corresponding P values (following the same order in Table 5.3) are found to be $P_2 = 0.95$, $P_3 = 0.63$, $P_4 = 0.98, P_5 = 0.97, P_6 = 0.88, P_7 = 0.97, P_8 = 0.91, P_9 = 0.81$, respectively.

Example 2. In this example we illustrate the $Q_n^{(k)}$ procedure for testing the *petal width* of *I. setosa* for a gamma distribution with unknown scale and shape parameters. The probability density function of the two parameter gamma distribution is given by

$$f(x; \gamma, \beta) = \frac{1}{\beta^{\gamma}\Gamma(\gamma)} x^{\gamma-1} e^{-x/\beta} \tag{5.43}$$

For a given random sample $X_1, X_2, \ldots, X_n$ the maximum likelihood estimates of the parameters can be obtained. For our data set the maximum likelihood

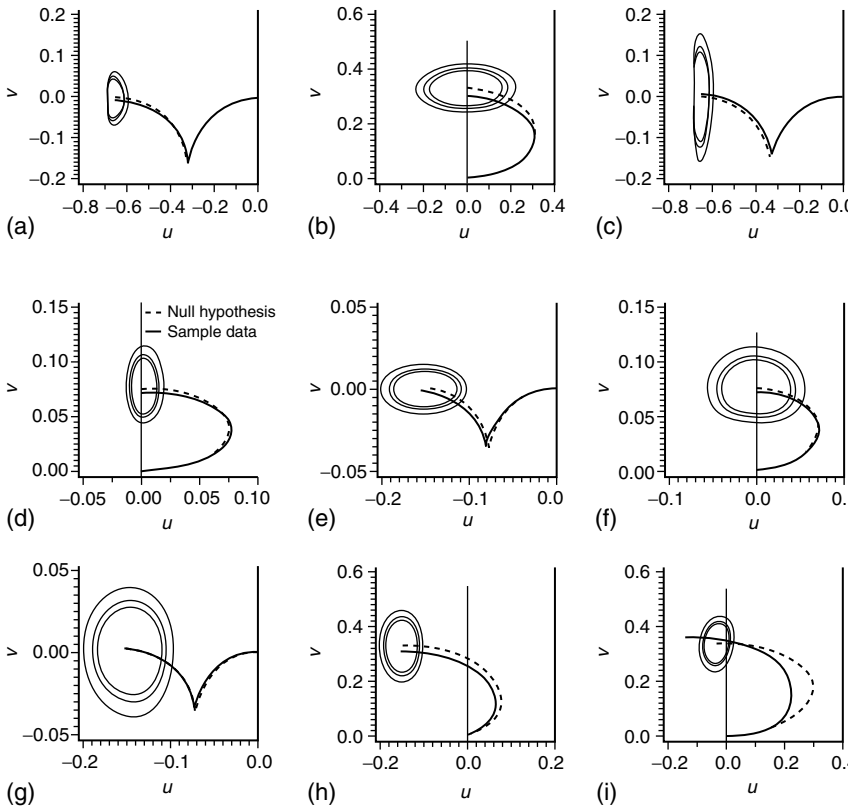

FIGURE 5.8 Various $Q_n^{(k)}$ based on *Iris setosa* data where (a) through (h) is testing for normality and (i) is testing for gamma distribution. (a) Type 2, (b) Type 3, (c) Type 4, (d) Type 5, (e) Type 6, (f) Type 7, (g) Type 8, (h) Type 9 (i) Type 1.

estimates of γ and β are found to be $\hat{\gamma} = 6.246$ and $\hat{\beta} = 0.394$ (see Johnson and Kotz[30,31] for the estimation procedure). Since the $Q_n^{(k)}$ procedure is scale invariant, we need to consider the shape parameter only. To establish a $100\,(1 - \alpha)\%$ confidence contour, we obtain the empirical joint distribution of $Q_n^{(k)}$ based on the standard gamma distribution (i.e., gamma distribution with $\beta = 1$ and $\gamma = 6.246$) by following the same procedure as explained in Section 5.2.3. The linked vectors chart for $k = 1$ is given in Figure 5.8(i). The corresponding *P value* is found to be 0.00.

Example 3. In this example we examine the *I. setosa* data for testing multivariate normality. Using the values of four variates we first compute r_1^2 ($i = 1, 2, \ldots, n$). Testing the null hypothesis of multivariate normality reduces to testing the null hypothesis that the transformed sample $r_1^2, r_2^2, \ldots, r_{50}^2$ comes from a Beta(2, 22.5) distribution. The corresponding linked vector chart is given in

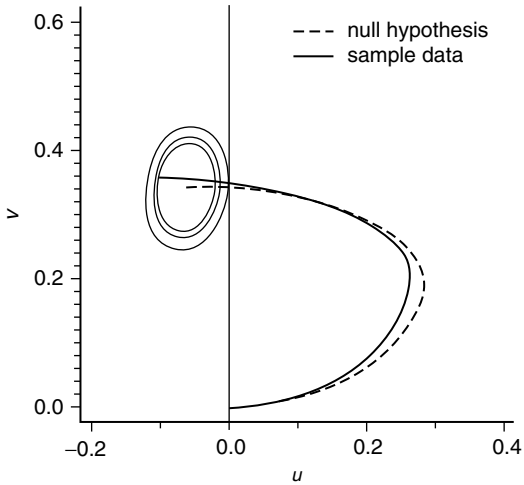

FIGURE 5.9 The $Q_n^{(k)}$ test for multivariate normality ($k = 1$).

Figure 5.9. The *P value* is found to be 0.08. This data set was also examined by Koziol,[32] Small,[33] and Royston.[34] The respective *P values* are given to be 0.20, 0.01 and 0.001. From example 1, it is clear that the source of nonnormality is the highly skewed fourth variate (*petal width*).

5.2.6. Empirical Power

For $k = 1$, Ozturk and Dudewicz[1] have studied the power of the $Q_n^{(k)}$ procedure for testing normality. Using a wide range of alternative distributions, it is shown that the power of the $Q_n^{(1)}$ test is comparable with that of the Shapiro–Wilk test.

It is clear that the power of the $Q_n^{(k)}$ test depends on the sample size (n), type of the standardized statistic (k) and the null distribution. To make a general comparison of power, one needs to consider a wide range of alternative distributions and various significance levels. Such a comparison would require an extensive study and will be pursued in future work.

In this study, we perform a limited comparison in order to have some feeling about the power of the $Q_n^{(k)}$ procedure. Seven distributions, namely normal, uniform, exponential, Laplace, logistic, Cauchy and extreme value were chosen for the comparison of power. From each alternative distribution, 2000 samples of size 25 were generated and the 10% empirical powers were obtained by computing the number of cases that the *P value* was less than 0.10. The results are given in Table 5.4. Alternative distributions are given in the columns and null distributions are given in the rows. The results are obtained for all types of standardizations. For example using the Type 1 statistic ($k = 1$), the empirical power of the $Q_n^{(1)}$ procedure for testing normality is 0.38 when the alternative

TABLE 5.4
10% Power of the $Q_n^{(k)}$ Test Based on 2000 Samples of Size 25

Null Distribution	k	Normal	Uniform	Exponential	Laplace	Logistic	Cauchy	Extreme
Normal	1	0.10	0.38	0.87	0.41	0.16	0.95	0.40
Uniform	1	0.35	0.10	0.89	0.84	0.57	1.00	0.62
Exponential	1	0.84	0.88	0.10	0.86	0.85	0.96	0.99
Laplace	1	0.21	0.77	0.85	0.11	0.14	0.81	0.36
Logistic	1	0.08	0.49	0.86	0.22	0.08	0.90	0.31
Cauchy	1	0.76	0.98	0.82	0.28	0.57	0.11	0.70
Extreme v	1	0.32	0.53	0.99	0.56	0.40	0.94	0.10
Normal	2	0.09	0.39	0.90	0.40	0.21	0.96	0.45
Uniform	2	0.38	0.09	0.95	0.83	0.58	0.99	0.72
Exponential	2	0.91	0.91	0.10	0.91	0.89	0.97	0.99
Laplace	2	0.17	0.83	0.81	0.09	0.11	0.80	0.30
Logistic	2	0.07	0.53	0.86	0.22	0.09	0.90	0.37
Cauchy	2	0.79	0.99	0.34	0.28	0.55	0.09	0.54
Extreme v	2	0.41	0.56	1.00	0.59	0.46	0.93	0.08
Normal	3	0.09	0.39	0.76	0.40	0.14	0.95	0.31
Uniform	3	0.34	0.11	0.82	0.81	0.54	0.99	0.56
Exponential	3	0.63	0.78	0.10	0.68	0.62	0.92	0.94
Laplace	3	0.25	0.80	0.74	0.10	0.16	0.78	0.36
Logistic	3	0.11	0.54	0.74	0.19	0.10	0.89	0.30
Cauchy	3	0.82	0.99	0.82	0.33	0.65	0.09	0.74
Extreme v	3	0.24	0.46	0.95	0.48	0.30	0.93	0.10
Normal	4	0.09	0.51	0.65	0.37	0.17	0.94	0.30
Uniform	4	0.37	0.10	0.89	0.82	0.57	0.99	0.63
Exponential	4	0.47	0.89	0.09	0.23	0.35	0.69	0.59
Laplace	4	0.35	0.88	0.45	0.10	0.21	0.73	0.31
Logistic	4	0.13	0.65	0.50	0.14	0.10	0.86	0.19
Cauchy	4	0.93	1.00	0.70	0.60	0.82	0.10	0.79
Extreme v	4	0.11	0.25	0.73	0.13	0.09	0.82	0.09
Normal	5	0.09	0.60	0.89	0.37	0.16	0.95	0.38
Uniform	5	0.54	0.09	0.88	0.91	0.74	1.00	0.70
Exponential	5	0.85	0.91	0.10	0.84	0.83	0.94	0.99
Laplace	5	0.22	0.88	0.83	0.09	0.13	0.78	0.39
Logistic	5	0.11	0.69	0.87	0.18	0.10	0.89	0.33
Cauchy	5	0.74	0.99	0.83	0.26	0.52	0.10	0.69
Extreme v	5	0.36	0.72	0.99	0.59	0.42	0.94	0.11
Normal	6	0.09	0.62	0.93	0.37	0.20	0.94	0.46
Uniform	6	0.63	0.10	0.93	0.92	0.79	1.00	0.79
Exponential	6	0.90	0.95	0.09	0.87	0.88	0.93	1.00

Continued

TABLE 5.4 Continued

Null Distribution	k	Normal	Uniform	Exponential	Laplace	Logistic	Cauchy	Extreme
Laplace	6	0.17	0.88	0.88	0.10	0.09	0.78	0.36
Logistic	6	0.09	0.72	0.93	0.17	0.09	0.87	0.38
Cauchy	6	0.58	1.00	0.73	0.22	0.38	0.08	0.53
Extreme v	6	0.43	0.78	1.00	0.61	0.49	0.93	0.12
Normal	7	0.10	0.58	0.74	0.37	0.16	0.94	0.28
Uniform	7	0.55	0.11	0.81	0.92	0.72	0.99	0.68
Exponential	7	0.63	0.84	0.10	0.66	0.61	0.92	0.94
Laplace	7	0.27	0.91	0.77	0.10	0.16	0.77	0.38
Logistic	7	0.12	0.73	0.74	0.18	0.08	0.88	0.28
Cauchy	7	0.79	1.00	0.87	0.32	0.59	0.09	0.77
Extreme v	7	0.28	0.65	0.95	0.47	0.32	0.92	0.10
Normal	8	0.09	0.63	0.37	0.28	0.14	0.90	0.17
Uniform	8	0.58	0.08	0.74	0.90	0.74	1.00	0.67
Exponential	8	0.25	0.71	0.09	0.17	0.19	0.73	0.43
Laplace	8	0.30	0.93	0.45	0.08	0.17	0.66	0.28
Logistic	8	0.15	0.79	0.37	0.12	0.10	0.77	0.17
Cauchy	8	0.87	1.00	0.84	0.51	0.73	0.12	0.81
Extreme v	8	0.14	0.62	0.74	0.22	0.14	0.85	0.08
Normal	9	0.12	0.58	0.96	0.43	0.26	0.93	0.44
Uniform	9	0.63	0.10	0.97	0.91	0.80	0.99	0.78
Exponential	9	0.98	0.96	0.10	0.98	0.98	0.98	1.00
Laplace	9	0.19	0.89	0.95	0.10	0.11	0.75	0.20
Logistic	9	0.07	0.69	0.92	0.20	0.10	0.86	0.30
Cauchy	9	0.69	1.00	0.57	0.24	0.46	0.12	0.57
Extreme v	9	0.40	0.71	1.00	0.62	0.48	0.95	0.09

distribution is uniform. Note that when the alternative distribution for testing normality is also normal then the corresponding empirical power is 0.10.

It is interesting to see that among the other factors the type of the standardization has strong influence on the power of the test. For example, when the null distribution is normal and the alternative distribution is uniform, the corresponding powers are 0.38 for $k = 1$ and 0.63 for $k = 8$. Similar results can be observed for many other cases. This suggests that if the alternative distribution can be specified, then the power of the test can be improved by choosing the appropriate type of standardization.

The power of the $Q_n^{(k)}$ test (with $k = 9$) is also compared with that of the Shapiro–Wilk W test. The W test is known to be an *omnibus* test and it is developed particularly for testing normality. Three sample sizes ($n = 15$, 35 and 50) and 35 alternative distributions are used in the power comparison and

TABLE 5.5
10% Power of the $Q_n^{(9)}$ and W Tests (for Normality) Based on 5000 Samples

	$n = 15$		$n = 35$		$n = 50$	
Null Distribution	W	$Q_n^{(9)}$	W	$Q_n^{(9)}$	W	$Q_n^{(9)}$
Uniform	0.26	0.27	0.74	0.80	0.95	0.96
Beta(3, 3)	0.08	0.08	0.13	0.12	0.22	0.16
Beta(4, 4)	0.08	0.07	0.10	0.10	0.15	0.11
Beta(5, 5)	0.08	0.08	0.10	0.08	0.12	0.09
Student t(1)	0.79	0.81	0.98	0.97	0.99	0.99
Student t(2)	0.52	0.52	0.74	0.78	0.82	0.87
Student t(3)	0.32	0.38	0.54	0.58	0.60	0.70
Student t(5)	0.21	0.25	0.30	0.40	0.32	0.45
Student t(10)	0.13	0.16	0.17	0.21	0.16	0.24
Student t(25)	0.12	0.12	0.12	0.13	0.11	0.14
Lambda(.46, .46)	0.10	0.10	0.25	0.25	0.43	0.38
Lambda(.85, .85)	0.20	0.23	0.67	0.72	0.90	0.92
Lambda(.98, .98)	0.25	0.25	0.72	0.80	0.95	0.96
Cauchy	0.80	0.82	0.97	0.97	0.99	0.99
Logistic	0.16	0.17	0.18	0.24	0.20	0.29
Johnson Sb(0, 1.5)	0.08	0.07	0.11	0.09	0.14	0.13
Johnson Sb(0, 2.0)	0.08	0.08	0.09	0.07	0.12	0.08
Johnson Sb(0, 2.5)	0.09	0.09	0.09	0.07	0.10	0.08
Exponential	0.79	0.75	0.99	1.00	1.00	1.00
Triangular	0.07	0.07	0.10	0.10	0.17	0.13
Lognormal	0.88	0.87	0.99	0.99	0.99	1.00
Chi-square(1)	0.97	0.97	1.00	1.00	0.99	1.00
Chi-square(3)	0.62	0.58	0.95	0.96	1.00	0.99
Chi-square(5)	0.44	0.40	0.81	0.76	0.93	0.92
Beta(3, 2)	0.13	0.11	0.26	0.23	0.41	0.33
Beta(5, 1)	0.60	0.58	0.97	0.98	0.99	1.00
Weibull(3)	0.08	0.09	0.10	0.10	0.12	0.11
Weibull(4)	0.09	0.08	0.09	0.08	0.10	0.07
Von Mises(2)	0.15	0.19	0.19	0.24	0.18	0.28
Extreme value	0.32	0.28	0.62	0.55	0.76	0.69
Lambda(.20, .15)	0.10	0.09	0.09	0.08	0.10	0.08
Lambda(.05, .10)	0.15	0.16	0.25	0.22	0.28	0.27
Lambda(.07, .10)	0.13	0.13	0.13	0.15	0.14	0.17
Johnson Sb(1, 2)	0.11	0.09	0.14	0.12	0.20	0.15
Johnson Sb(1, 3)	0.09	0.10	0.10	0.09	0.11	0.10

the corresponding results based on 5000 samples are given in Table 5.5. It is noted that the power of the $Q_n^{(9)}$ procedure is uniformly similar to (and in some cases higher than) that of the W test for the alternative distributions considered in the comparison.

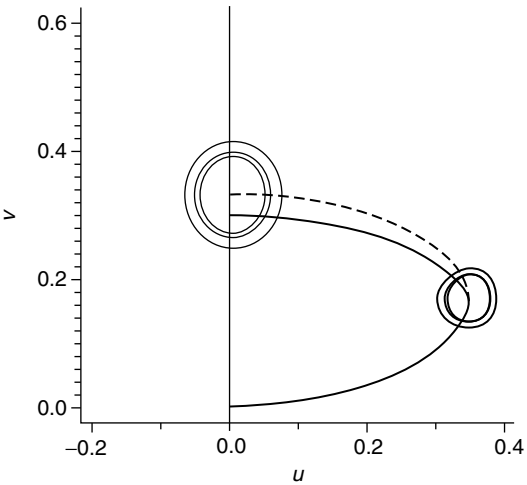

FIGURE 5.10 Linked vector chart with confidence contours at two points.

Another possibility of improving the power for the test is to obtain a second confidence contour at the point $E\{Q^{(k)}_{[n/2]}\}$. An example of such confidence contours is given in Figure 5.10. The sample linked vectors are obtained by using the *Iris* data (*sepal length*). A major difficulty of such an approach is that the level of significance is not determinable because of the strong dependence between the statistics $Q^{(k)}_n$ and $Q^{(k)}_{[n/2]}$.

However, for a given level of overall significance α, the individual level of significance less than α could be determined empirically for each point. In many cases, the use of the second confidence contour is important since the point $Q^{(k)}_n$ can be reached through various patterns in the linked vectors. Therefore, the construction of such a contour and determination of the empirical level of significance is included in the present algorithm.

5.2.7. The Test Algorithm

It is noted that the $Q^{(k)}_n$ procedure provides a general algorithm for testing the distributional assumptions. In this section, we will describe this algorithm. The input to the procedure consists of $n \times 1$ dimensional observations X_i $(i = 1, 2, \ldots, n)$.

At the first stage of the procedure the null hypothesis is specified. If the null distribution has an unknown shape parameter, then this parameter is estimated using an appropriate method (i.e., method of maximum likelihood, least squares, etc.).

The second stage consists of checking whether the random sample is univariate or not. If the sample is from a multivariate distribution then the corresponding squared radii r_i^2 values are computed. In this case there will only

be one transformed sample consisting of squared radii. Another option at this stage is to make a linear transformation to obtain p (approximately) independent samples as explained in Section 5.2.4.

The third stage of the process is the sorting of the observations in the samples. Note that, in the multivariate case sorting is only done for the transformed samples.

At the next stage the overall level of significance is determined. If there is only one sample the corresponding significance level is α. If there are several independent samples (including the one sample multivariate case) then the overall level of significance α' is obtained.

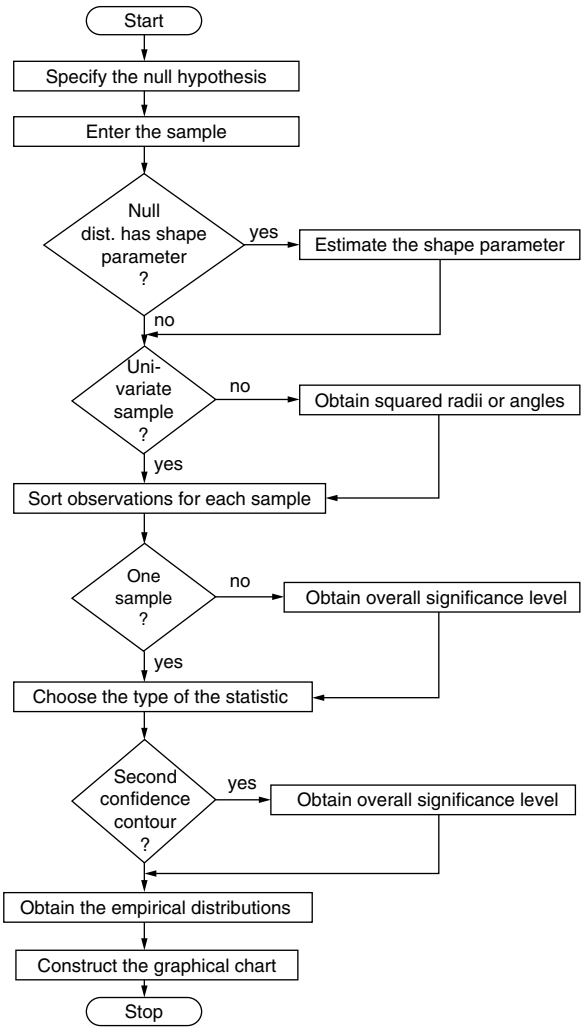

FIGURE 5.11 Flowchart of the test algorithm.

The type of standardization is determined at the fifth stage. If no choice is made the Type 1 statistic is used as default. Once the type of statistic is decided then the corresponding quantities $m_{i:n}$ and $E\{Y_{i:n}^{(k)}\}$ are obtained empirically. The empirical distribution of the statistic $Q_n^{(k)}$ is also determined at this stage.

The final stage of the process is to plot the confidence contours, null and sample linked vectors. If a second confidence contour is required then the corresponding overall level of significance is determined empirically.

A summary flowchart of the steps involved in computer implementation is given in Figure 5.11. A FORTRAN program based on this algorithm has been developed and the resulting graphics are obtained using the corresponding SAS procedures.

5.2.8. Discussion

The aim of this paper has been to develop a general algorithm based on $Q_n^{(k)}$ procedure for testing the distributional assumptions of a random sample. The underlying algorithm provides not only a general algorithm but also a graphical procedure for testing a single sample or several independent samples for specified univariate or multivariate distributions.

An in-depth study of the general power properties of the proposed test has not been attempted but preliminary power comparisons have indicated that it has good power properties. The power of the test may be increased by using a second confidence contour. However much further work is required to establish some criteria for the optimal choice of k for a specified null distribution.

It is evident that both the $Q_n^{(k)}$ procedure and $Q-Q$ type of plots are based on similar principles, that is, comparing the sample order statistics with their expected values. In $Q-Q$ plots the sample order statistics are plotted against the expected order statistics and the resulting plots are compared with a straight line. In the $Q_n^{(k)}$ procedure the sample standardized linked vectors are compared with their expected linked vectors. The length of the sample and the expected vectors are determined by the sample and the expected value of the corresponding standardized order statistics respectively, and the direction (the angle between the ith vector and the horizontal axis) of the vectors are determined by using the null distribution function.

Some advantages of the $Q_n^{(k)}$ procedure over $Q-Q$ plots are summarized as follows

(1) The $Q-Q$ plots do not provide a formal test. Linearity of the plots is justified on a subjective basis. The $Q_n^{(k)}$ procedure displays the sample observations in such a way that the sample vector patterns are compared with the expected patterns. Furthermore a confidence contour for the terminal point of the expected linked vectors (or if required a second confidence contour) is provided so that a formal test can be performed.
(2) It is not easy to appreciate the linearity of the $Q-Q$ plots when the sample is small. For example, Aly and Ozturk[19,20] attempted to overcome

this difficulty by introducing Hodges–Lehmann $Q-Q$ plots. However, the use of $Q-Q$ plots for small samples is still not straightforward. This situation does not create any serious problems for the $Q_n^{(k)}$ procedure.

(3) An attractive property of the $Q_n^{(k)}$ procedure is that if the null hypothesis is rejected then one can estimate the skewness and kurtosis of the true distribution. This means that if the null hypothesis is rejected then the $Q_n^{(k)}$ procedure can be used to identify the true distribution. The $Q-Q$ plots also can be used to characterize the distributions.[35,36] However, identification of a distribution by using the $Q-Q$ plots, is made on the subjective basis only (our research is being continued along this line).

(4) The $Q_n^{(k)}$ procedure provides a general algorithm to test the distributional assumptions for one or several samples from univariate or multivariate distributions. Extensions of $Q-Q$ plots to multivariate cases is not easy. Furthermore the $Q_n^{(k)}$ procedure may be used to assess the distributional assumptions for the non location-scale family of distributions.

(5) The algorithm proposed in this paper does not require knowledge of the expected values $m_{i:n}$ or $m'_{i:n}$. These quantities are obtained empirically. Clearly the accuracy of such an approximation depends on the number of Monte-Carlo samples. Our experience has shown that 1000 samples for a moderate size of sample would provide satisfactory approximations to both $m_{i:n}$ and $m'_{i:n}$ $(i = 1, 2, \ldots, n)$. Thus, determination of the optimal number of Monte Carlo samples will be a trade off between the accuracy of the underlying approximations and the computational time.

Besides the generality and good power properties, the proposed test is graphical in nature. Not only can it be used to enhance scientific work, but also it has a potential use in teaching and research.

6 Applications

CONTENTS

6.1. The Ozturk Algorithm: A New Technique for Analyzing Random Data with Applications to the Field of Neuroscience 297
(*W. W. Weiner*)
6.1.1. Introduction to the Ozturk Algorithm 297
6.1.1.1. Overview 298
6.1.1.2. Sample Simulation 299
6.1.2. Detailed Description of the Ozturk Algorithm 306
6.1.2.1. The Standardized Order Statistic 306
6.1.2.2. The Goodness-of-Fit Test 307
6.1.2.3. Calculation of Linked Vectors in the U-V Plane 308
6.1.2.4. Calculation of Confidence Ellipses 311
6.1.2.5. The Best-Fit Test 312
6.1.2.6. Estimation of Location and Scale Parameters 314
6.1.3. Analysis of Spontaneous Auditory Nerve Activity of Chinchillas 316
6.1.3.1. Analysis of Two Fibers with Different Spontaneous Rates 325
6.1.3.2. Analysis of Pulse-Number Distributions 327
6.1.4. Analysis of Efferent Optic-Nerve Activity in the Horseshoe Crab 333
6.1.4.1. Characterization of Interburst Intervals 335
6.1.4.2. Trends in the Shape Parameter 339
6.1.5. Analysis of the Visual Field of the Horseshoe Crab 341
6.1.5.1. Total Interommatidial Angles 344
6.1.5.2. Horizontal and Vertical Interommatidial Angles 346
6.1.6. Applications of the Ozturk Algorithm in Neuroscience 348
6.2. Use of Image Processing to Partition a Radar Surveillance Volume into Background Noise and Clutter Patches 349
(*M. A. Slamani and D. D. Weiner*)
6.2.1. Introduction 349
6.2.2. Observations about BN and CL 350
6.2.2.1. Observations about BN 351
6.2.2.2. Observations about CL 351
6.2.3. Mapping Procedure 351

6.2.3.1. Separation of CL Patches from BN 351
6.2.3.2. Detection of Clutter Patch Edges 354
6.2.3.3. Enhancement of Clutter Patch Edges 355
6.2.4. Example 355
6.3. Probabilistic Insight into the Application of Image Processing to the Mapping of Clutter and Noise Regions in a Radar Surveillance Volume 359
(*M. A. Slamani and D. D. Weiner*)
6.3.1. Introduction 359
6.3.2. Separation between BN and CL Patches 360
6.3.3. Summary 368
6.4. A New Approach to the Analysis of IR Images 368
(*M. A. Slamani, D. Ferris, and V. Vannicola*)
6.4.1. Introduction 368
6.4.2. ASCAPE 369
6.4.3. Mapping Procedure 371
6.4.3.1. Identification of Lowest Average Power Level (LP) 371
6.4.3.2. Detection of Patch Edges 371
6.4.4. Statistical Procedure 372
6.4.4.1. Introduction to Ozturk Algorithm 372
6.4.4.2. Outliers 374
6.4.4.3. Strategy to SubPatch Investigation Using the Statistical Procedure 374
6.4.5. Expert System Shell IPUS 375
6.4.6. Example: Application of ASCAPE to Real IR Data 376
6.4.7. Conclusion 381
6.5. Automatic Statistical Characterization and Partitioning of Environments (ASCAPE) 382
(*M. A. Slamani, D. D. Weiner, and V. Vannicola*)
6.5.1. Problem Statement 382
6.5.2. ASCAPE Process 385
6.5.3. Application of ASCAPE to Real IR Data 385
6.5.4. Conclusion 386
6.6. Statistical Characterization of Nonhomogeneous and Nonstationary Backgrounds 386
(*A. D. Keckler, D. L. Stadelman, D. D. Weiner, and M. A. Slamani*)
6.6.1. Introduction 386
6.6.2. Application of ASCAPE to Concealed Weapon Detection 387
6.6.3. The SIRV Radar Clutter Model 390
6.6.4. Distribution Approximation Using the Ozturk Algorithm 392
6.6.5. Approximation of SIRVs 395
6.6.6. NonGaussian Receiver Performance 398
6.6.7. Concluding Remarks 400

6.7. Knowledge-Based Map Space Time Adaptive Processing (KBMapSTAP) 400
(*C. T. Capraro, G. T. Capraro, D. D. Weiner, and M. C. Wicks*)
6.7.1. Introduction 400
6.7.2. Clutter Model 401
6.7.3. Representative Secondary Clutter 402
6.7.4. Airborne Radar Data 402
6.7.5. *A Priori* Data 403
6.7.6. Research Problem, Hypothesis, and Preliminary Findings 403
6.7.7. Conclusions and Future Work 407
6.8. Improved STAP Performance Using Knowledge-Aided Secondary Data Selection 408
(*C. T. Capraro, G. T. Capraro, D. D. Weiner, M. C. Wicks, and W. J. Baldygo*)
6.8.1. Introduction 408
6.8.2. Radar and Terrain Data 409
6.8.3. Approach 410
6.8.3.1. STAP Algorithm 410
6.8.3.2. Registration of the Radar with the Earth 411
6.8.3.3. Data Selection 412
6.8.3.4. Corrections for Visibility 412
6.8.3.5. Secondary Data Guard Cells 413
6.8.4. Results 413
6.8.5. Conclusion 415

6.1. THE OZTURK ALGORITHM: A NEW TECHNIQUE FOR ANALYZING RANDOM DATA WITH APPLICATIONS TO THE FIELD OF NEUROSCIENCE

(W. W. WEINER)

6.1.1. Introduction to the Ozturk Algorithm

The purpose of this project is to introduce a new algorithm for analyzing random data. This algorithm is now referred to as the Ozturk Algorithm, named after the person who invented it (Ozturk[1]). In addition to describing this algorithm, this chapter will also present three detailed applications that illustrate the power of this algorithm. In the first part of this chapter, a brief description of the Ozturk Algorithm and its advantages over classical techniques will be discussed. The reader will also be introduced to the basic operation of this algorithm. In the next section of this chapter, a detailed description of how the Ozturk Algorithm works will be given. For those readers who are only interested in applications of the Ozturk Algorithm and not in the mathematical specifics of its operation,

this section may be omitted. The next three sections of this chapter will discuss three applications of the Ozturk Algorithm:

1) Analysis of auditory nerve activity in the chinchilla
2) Analysis of efferent optic nerve activity in the horseshoe crab
3) Analysis of the visual field of the horseshoe crab.

Finally, the last section of this chapter will discuss potential applications of the Ozturk Algorithm in neuroscience.

6.1.1.1. Overview

The Ozturk Algorithm provides a very powerful tool for describing and analyzing data. The algorithm provides two primary modes of operation: it can be used to describe the probability density function (PDF) (from a prespecified library) that best fits a set of data through the use of a graphical solution, or it can be used to perform a goodness-of-fit test between a set of data and a specified density function (null distribution). The technique this algorithm employs also provides the user with considerable flexibility. The library of density functions, which the algorithm uses can be modified or expanded.

The primary advantages of the Ozturk Algorithm are that very few sample points are required to run this algorithm, and that unlike classical statistical techniques, no assumptions about the null hypothesis are required for fitting data. Classical goodness-of-fit tests such as the chi-square test and Q–Q plots require large sample sizes to give accurate results. In contrast, the Ozturk algorithm performs extremely well with as few as 50 to 75 data points. In fact, empirical testing of this algorithm has found that 100 to 150 data points works as well as much larger sample sizes for most applications (Shah[2]). The other main advantage of the Ozturk Algorithm is its ability to fit data without specifying a null hypothesis. Classical tests such as the Kolmogorov–Smirnov test require the user to assume a null distribution. Furthermore, with these classical techniques, should the null distribution be rejected, no insight is gained about which distribution provides a suitable fit to the data.

When using the Ozturk Algorithm, it is important to make sure that one correctly interprets the information provided by the algorithm. The algorithm generates a list of the PDFs which best fit the data. It also gives the value of the parameters for each of these PDFs. One must keep in mind that the critical issue is NOT which distribution the algorithm determines best fits the data, but the fact that an equation is obtained for using to describe the data. In fact, if one plots several of the PDFs that best fit the data, one will find that each of these PDFs looks very similar, regardless of the type of function. As a result, a number of PDFs may accurately describe the data.

Once the Ozturk algorithm determines which PDF best approximates the data, a considerable amount of information about the data set can be calculated. Since all of the density functions in the algorithm's library are well characterized,

information such as mean and variance can be computed. In addition, one can use the PDF to generate a cumulative distribution function or a Fourier Transform. This enables probability and frequency information to be extracted. Furthermore, if one is attempting to model data, the Ozturk Algorithm can be used to provide an appropriate expression for describing a phenomenon.

The Ozturk Algorithm may also be used to run a goodness-of-fit test. In this mode, a null distribution is specified (including shape parameters), and the algorithm provides a graphical solution describing the goodness-of-fit. Three confidence ellipses are plotted, corresponding to confidence levels of 90, 95, and 99%. If a data set falls within one or more of these ellipses, the sample data is said to be statistically consistent with the null hypothesis at a confidence level based on the confidence ellipses. Each confidence ellipse describes the conditional probability that given the null hypothesis is true, the data set should fall within the ellipse with a probability corresponding to the percent of the confidence ellipse. In other words, for the confidence ellipse corresponding to a confidence level of 99% the sample data, if consistent with the null hypothesis, will fall within this ellipse 99% of the time. It should be noted that the size of the confidence ellipses depends on the number of points in the sample data: the greater the number of points, the smaller the ellipses. More will be said about the confidence ellipses in the next section of this chapter.

6.1.1.2. Sample Simulation

The best way to gain an appreciation for the Ozturk Algorithm is to consider a sample simulation. The purpose of this example is to familiarize the reader with the output of the algorithm and to demonstrate how to interpret a goodness-of-fit test and a best-fit test using this algorithm. In this example, a data set consisting of 500 points was analyzed. For the purpose of this example, the nature of the data is unimportant. Also, as will be shown in later sections, this data set is much larger than necessary for generating accurate results with this algorithm.

Before running the algorithm, the number of points in the data set must be entered (in this case 500). Next, the density function that will be used as the null distribution must be specified. If no density function is entered, the normal distribution is used as the null distribution. The algorithm will then list the five density functions that best fit the data. In this particular example, the density functions listed were: (1) Weibull, (2) Gamma, (3) K distribution, (4) Lognormal, and (5) Gumbel (type-2). After this, the algorithm will then calculate the parameters of best fit for a specified density function. As an example, for these data, the respective equations for the best-fit Weibull and Gamma functions were specified as follows:

$$f(x) = \frac{\gamma}{\beta}\left[\frac{x-\alpha}{\beta}\right]^{\gamma-1} e^{-\left[\frac{x-\alpha}{\beta}\right]^{\gamma}} \tag{6.1}$$

where $\alpha = -0.36E - 3$, $\beta = 0.29E - 1$, $\gamma = 1.32$ and

$$g(x) = \frac{1}{\beta\Gamma(\gamma)} e^{-\frac{(x-\alpha)}{\beta}} \left[\frac{x-\alpha}{\beta} \right]^{\gamma-1} \tag{6.2}$$

where $\alpha = -0.28E - 2$, $\beta = 0.15E - 1$, $\gamma = 2.0$, and $\Gamma(x) =$ gamma function.

After the Ozturk Algorithm has generated the above information, a graphical solution for the best-fit and goodness-of-fit tests may be obtained. The interpretation of these graphical solutions will be discussed next.

Tables 4.8 and 4.9 of Section 4.2.3 list the PDFs that the Ozturk Algorithm currently uses. These functions are listed both in standard form and general form, the difference between the two is that the general form incorporates the transformation $y = (x - \alpha)/\beta$. Therefore, as indicated in Tables 4.8 and 4.9, the relationship between the standard form and the general form of the PDFs is

$$g(x) = f(y) \left. \frac{dy}{dx} \right|_{y=\frac{x-\alpha}{\beta}} = \frac{1}{\beta} f\left(\frac{x-\alpha}{\beta} \right)$$

In the general form of the density function, α and β are referred to as the location and scale parameters, respectively. These parameters are similar but not equivalent to the mean and standard deviation (except in the case of the normal distribution). Many of the density functions listed in the appendix also contain one or two shape parameters (γ and δ). The shape parameters, as the name implies, determine the shape of the distribution function within a family of shapes associated with a specified variate. A more detailed description of the location, scale, and shape parameters may be found in (Hastings and Peacock[3]). To help get a feel for how the location, scale and shape parameters affect a density function, the parameters are modified one at a time for the Weibull function described above. These results are illustrated in Figure 6.1. Notice that only the shape parameter alters the general shape of the function.

In Figure 6.1, the bolded line (filled circles) in each plot corresponds to the best-fit Weibull distribution, as specified in the text. In each plot only one variable is changed. The other two variables are kept constant at the value specified for the best Weibull distribution. The value assigned to the modified parameter is given in the legends for each graph ($a =$ location parameter, $b =$ scale parameter, and $g =$ shape parameter).

The Ozturk Algorithm uses a graphical solution to perform the best-fit test. This solution involves mapping each of the distributions that are listed in Tables 4.8 and 4.9 into a plane (defined as the U-V plane). The details of this mapping will be discussed in the next section. The sample data is also mapped into this U-V plane. The algorithm then determines which distribution in the U-V plane is closest to the mapped sample data, and this distribution is specified as the best-fit distribution for the sample data. For the example discussed above, the graphical solution to the best-fit test is shown in Figure 6.2.

This figure shows the mapping of all the distributions in Tables 4.8 and 4.9, as well as the 500 sample data point into the U-V plane. The location of the mapping of the data sample is represented in the U-V plane by a filled circle.

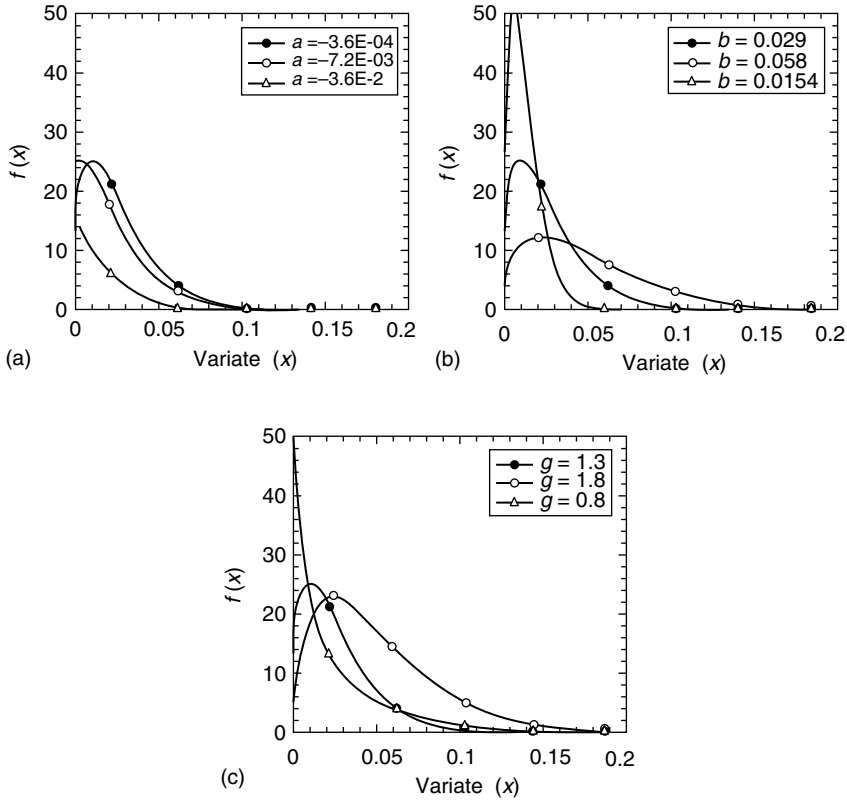

FIGURE 6.1 Effect of location, scale, and shape parameters on Weibull PDF. (a) Variation of location parameter *a*. (b) Variation of scale parameter *b*. (c) Variation of shape parameter *g*.

Notice that the five distributions that are closest to the mapped data point are the Weibull, Gamma, K distribution, Lognormal, and Gumbel (type-2). [N = Normal, U = Uniform, E = Exponential, A = Laplace, L = Logistic, C = Cauchy, V = Extreme Value (type-1), T = Gumbel (type-2), G = Gamma, P = Pareto, W = Weibull, L = Lognormal, K = K-distribution. The upper five dashed lines represent Beta, and the lower nine dashed lines represent SU Johnson.]

Normally, the Ozturk Algorithm uses different colors to represent the distributions in these graphs. Unfortunately, when printing these graphs using a gray-scale printer, the resulting printout becomes much more cluttered looking and difficult to interpret. I apologize for this difficulty, and ask the reader to keep in mind that the actual display of the graphical solution is easier to interpret when in color and on a full-size computer screen.

A few comments about this mapping are necessary. As will be seen in the next section, the density functions that lack any shape parameters (γ and δ) get

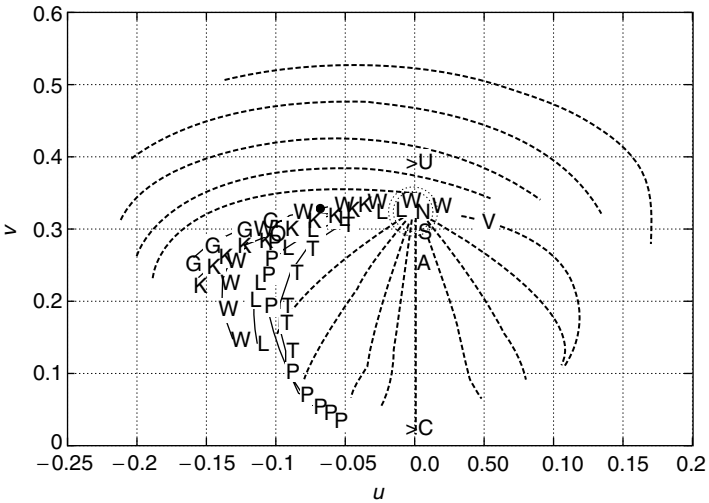

FIGURE 6.2 Distribution approximation chart for graphical solution of the Ozturk algorithm to the best-fit test. • mapped end point of 500 sample data set.

mapped as points in the U-V plane. Examples of these density functions include: Normal, Uniform, Exponential, Laplace, Logistic, Cauchy, and Extreme Value (type-1). The density functions that contain a single shape parameter get mapped as lines in the U-V plane. Each point on the line corresponds to a specific value of the shape parameter. Examples of these density functions include: Gumbel (type-2), Gamma, Pareto, Weibull, Lognormal, and K-distribution. Finally, the density functions that contain two shape parameters get mapped as multiple lines in the U-V plane. Each line corresponds to a fixed value of one shape parameter, with the points on this line corresponding to specific values of the second shape parameter. Examples of these density functions include the Beta and SU Johnson. The sample data set gets mapped as a single point in the U-V plane, and as stated earlier, the distributions that best fit the sample data are the ones located closest to this mapped point. As will be seen in the next section, the exact location of the mapping of distributions into the U-V plane depends upon the number of points in the data sample.

The best way to ensure that a best-fit solution is accurately determined for any data set is to make sure that the U-V plane is filled with known distributions. By so doing, regardless of where in the plane the sample data gets mapped, one is assured that a known distribution will lie close to this point. In fact, it is for this reason that the distributions listed in Table 4.8 and Table 4.9 are used by the Ozturk Algorithm. Examination of Figure 6.2 demonstrates that almost the entire U-V plane is filled with known distributions. The algorithm could be made even more rigorous by adding distributions that cover other locations in the U-V plane.

The graphical solution also enables one to get a feel for how precise the best-fit test is. In this example, the sample data gets mapped right on top of the curve

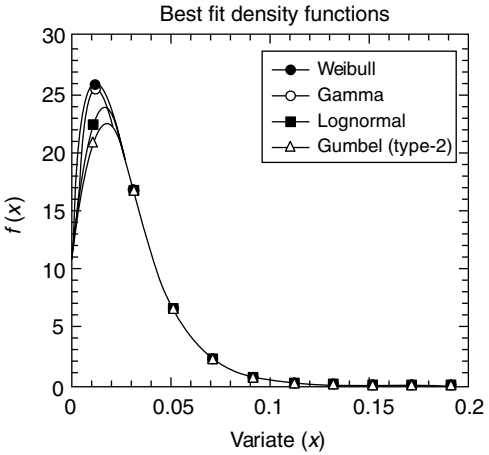

FIGURE 6.3 Four of the five best-fit density functions to the text data set.

corresponding to the Weibull distribution. Therefore, the Weibull distribution provides a perfect fit to the data and is thus the most accurate density function for describing the data. Nevertheless, any of the five best-fit distributions provide a good fit for the data. In fact, these density functions when plotted together all look very similar (Figure 6.3). Therefore, any of the five density functions that the algorithm states best fit the data would very accurately describe the data. With experience, one will find that all of the density functions within a given region of the U-V plane look very similar when plotted. In this respect, one has some flexibility when choosing which best-fit density function to use for modeling data. In Figure 6.3, the K distribution has been omitted due to the complexity of this function. The parameters used to plot these functions were determined by the Ozturk algorithm. The equations describing these functions are listed in Tables 4.8 and 4.9. [Weibull: $\alpha = 0.36\text{E} -3$, $\beta = 0.029$, $\gamma = 1.3$; Gamma: $\alpha = 0.28\text{E} - 2$, $\beta = 0.015$, $\gamma = 2.0$; Gamma: $\alpha = 0.28\text{E} - 2$, $\beta = 0.015$, $\gamma = 2.0$; Lognormal: $\alpha = 0.022$, $\beta = 0.044$, $\gamma = 0.41$; Gumbel (type-2): $\alpha = 0.64$, $\beta = 0.66$, $\gamma = 40$.]

A goodness-of-fit test was also performed on these data. When doing a goodness-of-fit test, the sample data are compared with a specified distribution (the null distribution). After selecting a density function for the null distribution, values for the shape parameters must be given (if applicable). Two different null distributions were used for the goodness-of-fit test: Normal and Weibull. Since the Normal distribution lacks any shape parameters, none were specified. For the Weibull distribution, the shape parameter used for the goodness-of-fit test was 1.3 (the one found from the best-fit test). The results of these goodness-of-fit tests are shown in Figure 6.4. In this test the sample data and null hypothesis distribution get mapped as trajectories in the U-V plane. Based on these tests, a Normal distribution for the sample data can be *rejected* with 99% confidence level while a Weibull distribution can be *accepted* with 99% confidence level.

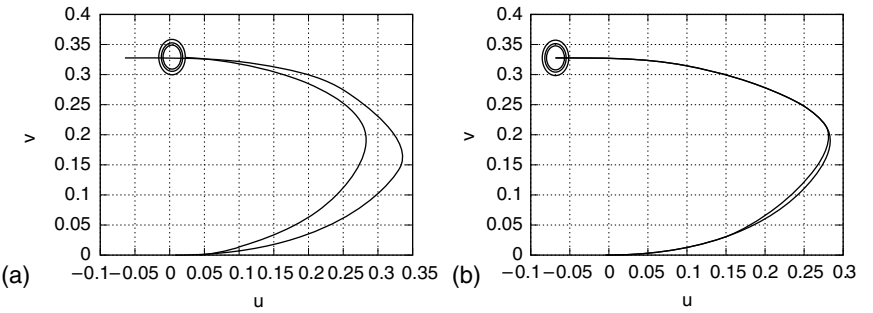

FIGURE 6.4 Comparison of two goodness-of-fit tests for the same sample data. (a) Null hypothesis distribution is normal. (b) Null hypothesis distribution is Weibull.

The outer ellipse corresponds to a 99% confidence level, middle ellipse to a 95% confidence level, and inner ellipse to a 90% confidence level. The larger the ellipse, the more likely the sample distribution will pass through the ellipse.

The details of the goodness-of-fit test will be discussed in the following section. However, the basic premise of the graphical solution is that the sample data and the null distribution get mapped as trajectories in the U-V plane. To create these trajectories, the Ozturk Algorithm first generates the same number of samples from the null distribution as are present in the sample data. Next, both the sample data and the null data are ordered sequentially in ascending order. (It is for this reason that independent random data samples must be used.) Finally, each data point from the sample and null data set is converted to an ordered statistic and mapped in the U-V plane. These mapped points are connected to form the resulting trajectories, as illustrated in Figure 6.4. If the data are completely consistent with the null distribution, these two trajectories will overlap everywhere. On the other hand, if the data are not consistent with the null distribution, these two trajectories will differ considerably.

As can be seen in Figure 6.4, the mapped trajectory of the data and the Normal distribution differ markedly, whereas the trajectory of the data and the Weibull distribution (for a shape parameter of 1.3) almost completely overlap. Therefore, these trajectories provide one with qualitative information regarding the goodness-of-fit test: the more identical the trajectories, the more similar are the data and the null distribution. It is important to remember, however, that for those distributions described by shape parameters, every value of the shape parameter will lead to a differently mapped trajectory.

The graphical solution to the goodness-of-fit test also provides quantitative information. Confidence ellipses are plotted for the end point of the null distribution trajectory. Therefore, the center of the confidence ellipses corresponds to the end point of the mapped null distribution trajectory. Three confidence ellipses are plotted. The largest ellipse corresponds to a confidence level of 99% (0.01 level of significance), the middle ellipse to a confidence level of 95%

(0.05 level of significance), and the smallest ellipse to a confidence level of 90% (0.10 level of significance). The confidence ellipses describe the probability that the end point of the mapped trajectory from the sample data will lie within the ellipses given that the null distribution is true. In other words, given that the null hypothesis is true, the end point of the trajectory from the sample data will be located inside the largest ellipse 99% of the time, the middle ellipse 95% of the time, and the smallest ellipse 90% of the time. Put in another way, if the end point of the mapped trajectory from the sample data lies outside of the large ellipse, the null hypothesis can be rejected with 99% confidence (or a 0.01 level of significance). The level of significance refers to the probability that the null hypothesis is rejected given that it is true. Therefore if the mapped end point from the sample data lies outside of all the confidence ellipses, we can reject the null hypothesis with a level of significance greater than 0.01. This means that we will be wrong in rejecting the null hypothesis less than one percent of the time. In this particular example, it is evident that the Normal distribution can be rejected as the null hypothesis with 99% confidence and that the Weibull distribution is consistent with the null hypothesis with 99% confidence.

It is important to point out that the size of the confidence ellipses is completely determined by the choice of the null distribution and the number of points in the sample data set. This statement is intuitively satisfying. The fewer the number of points, the larger the confidence ellipses will be. When few data points are present, the confidence ellipses will cover a large portion of the U-V plane, and almost all data sets will map inside these ellipses. This fact makes sense because if only a few samples (say, 10 to 20) are used, it is almost impossible to reject the idea that these data are from a specific density function. On the other hand, as the number of data points increases, the confidence ellipses eventually converge to a single point in the U-V plane. Therefore, in theory, a data set with an infinite number of points will be consistent with only one PDF. Similarly, the variability of the null distribution affects the size of the confidence ellipses: the greater the variability of the null distribution, the larger the confidence ellipses.

One last point concerning the confidence ellipses merits mentioning. When the Ozturk Algorithm implements the graphical solution to the best-fit test, the confidence ellipses are also plotted on the distribution approximation chart. This can be seen in Figure 6.4 where the Normal distribution was specified as the null distribution. Notice that the three confidence ellipses are plotted as dotted circles around the "N" in this figure. Also notice that the U-V coordinates for the confidence ellipses in Figures 6.2 and 6.4 agree. Thus, even when doing a best-fit test, information about the goodness-of-fit test is provided by the Ozturk Algorithm.

Now that a basic understanding of the Ozturk Algorithm has been presented, a detailed discussion of the best-fit test and goodness-of-fit test will be provided in the next section. Those readers who are only interested in the applications of the Ozturk Algorithm and not in the mathematical specifics of its operation may omit this section.

6.1.2. Detailed Description of the Ozturk Algorithm

In this section, a detailed explanation of how the Ozturk Algorithm performs the best-fit test and goodness-of-fit test will be given. Initially, the concept of a standardized order statistic will be discussed. Then, the technique used by the Ozturk Algorithm to perform the goodness-of-fit test will be described. This discussion will include how the sample data and null distribution get mapped as trajectories in the U-V plane and how the confidence ellipses are calculated. Finally, the technique used by the algorithm to perform the best-fit test will be discussed. A description of how the parameters in a best-fit test for a particular distribution are calculated will be given. The theoretical information in this section describing the operation of the Ozturk Algorithm has been taken from several sources (Shah,[2] Ozturk,[1] and Ozturk and Dudewicz[4]).

6.1.2.1. The Standardized Order Statistic

The Ozturk Algorithm is appropriate for analyzing any unimodal random data set. Currently, the algorithm is in the process of being expanded to multivariate and multimodal distributions (personal communication). The one assumption the algorithm makes is that the random data are from independent trials, and thus, the order of the data does not matter. Data for which this assumption is not valid should not be used by the algorithm.

The Ozturk Algorithm organizes the data samples in sequential order: $X_1, X_2, X_3, \ldots, X_n$, such that $X_1 < X_2 < X_3, \ldots, < X_n$. The standardized ith order statistic (Y_i) for each sample is defined as $Y_i = X_i - \mu_x/\sigma_x;\ i = 1, 2, 3, \ldots, n$ where

$$\mu_x = \sum_{1}^{n} \frac{X_i}{n}$$

is the sample mean and

$$\sigma_x = \sum_{i}^{n} \sqrt{\frac{(X_i - \mu_x)^2}{n-1}}$$

is the sample standard deviation. One of the major advantages of the standardized order statistic is that it is invariant under linear transformation. The usefulness of this property will become evident shortly.

Table 4.8 and Table 4.9 list all of the PDFs that the Ozturk Algorithm currently uses. These functions are listed both in standard form and general form, where the difference between the two is that the general form incorporates the transformation $y = (x - \alpha)/\beta$. Therefore, as indicated in the tables, the relationship between the standard form and the general form of the PDFs is

$$g(x) = f(y)\frac{\mathrm{d}y}{\mathrm{d}x}\bigg|_{y=\frac{x-\alpha}{\beta}} = \frac{1}{\beta} f\left(\frac{x-\alpha}{\beta}\right)$$

In the general form of the density function, α and β are referred to as the location and scale parameters, respectively. The advantage of using standardized order statistics is that given the linear transformation described above, the standard order statistics of the random variable X and Y are equal. That is

$$\frac{Y_i - \mu_y}{\sigma_y} = \frac{X_i - \mu_x}{\sigma_x}$$

where μ_x, μ_y, σ_x, and σ_y, are the sample means and sample standard deviations as defined previously. The above statement can easily be proven by noting that

$$\mu_y = \frac{E[X]}{\sigma_x} - \frac{\mu_x}{\sigma_x} = 0$$

and

$$\sigma_y = \sqrt{\frac{\text{Var}[X]}{\sigma_x^2}} = 1$$

This property enables the Ozturk Algorithm to perform the goodness-of-fit test and best-fit test using the standard form of the PDFs (a simpler form than the general one). As a result, in these calculations, the location and scale parameters are irrelevant since they do not affect the standardized order statistics. Put in another way, by using the standardized order statistics, the algorithm normalizes the data for any location and scale parameter. The end result is that only the shape parameters and type of density function affect the standardized order statistics. It is for this reason — as will be seen later in this section — that density functions that lack any shape parameters map as points in the U-V plane, whereas those density functions that have either one or two shape parameters map as a line or a series of lines in the U-V plane.

6.1.2.2. The Goodness-of-Fit Test

The Ozturk Algorithm has two modes of operation: a goodness-of-fit mode and a best-fit mode. A detailed description of the goodness-of-fit test will be provided first because once the procedure for this test has been explained, it will be easier to understand the best-fit test.

To perform the goodness-of-fit test, the Ozturk Algorithm uses three sets of data: a reference distribution, a null distribution, and a sample data set. For convenience, in this algorithm, the standard normal distribution is used as the reference distribution. However, there is no reason that another distribution could not be used. Similarly, the null distribution may be any density function that is listed in Tables 4.8 and 4.10. Nevertheless, one should keep in mind that additional distributions could be used as the null distribution if they were programmed into the Ozturk Algorithm. In addition to specifying that density function to use as the null distribution, it is also necessary to define all values

of the shape parameters (if applicable) for that density function. Finally, the sample data may describe any univariate random variable whose samples are independent events. The number of samples in the data set may range from 1 to 9000; however, it is recommended that one use at least 50 data points to obtain accurate results. The number of data points calculated for the reference distribution and the null distribution is the same as the number of data points in the sample data set.

The idea behind the goodness-of-fit test is that all of the points in the sample data set and null distribution are ordered sequentially, and then standardized order statistics are calculated for each point. The algorithm then transforms these statistics into linked vectors, one vector for each point. The magnitude of the vectors is determined by the sample data and null distribution while the angle of the vector is determined solely from the reference distribution. The sum of these vectors leads to the mapped trajectories in the U-V plane: one for the sample data and one for the null distribution (see Figure 6.4). Confidence ellipses are then calculated for the end point of the mapped trajectory of the null distribution.

6.1.2.3. Calculation of Linked Vectors in the U-V Plane

In this section the random variable X will be used when describing the sample data set, the random variable N will be used when describing the null distribution, and the random variable R will be used when describing the reference distribution. The following paragraphs explain the procedure used by the Ozturk Algorithm for computing the linked vectors.

The sample data from which the standardized order statistics are calculated is entered into the algorithm as a text file. In contrast, the average of 2000 Monte Carlo simulations is used to generate data for the reference and null distribution. The number of points generated with each Monte Carlo simulation is the same as the number of samples in the data set to be analyzed. The first calculation made by the algorithm is to order each of the samples as described below:

$$X_1, X_2, X_3, \ldots, X_n \qquad N_{1;k}, N_{2;k}, N_{3;k}, \ldots, N_{n;k} \qquad R_{1;k}, R_{2;k}, R_{3;k}, \ldots, R_{n;k} \tag{6.3}$$

such that

$$\begin{aligned} &X_1 < X_2 < X_3 < \cdots < X_n \\ &N_{1;k} < N_{2;k} < N_{3;k} < \cdots < N_{n;k} \\ &R_{1;k} < R_{2;k} < R_{3;k} < \cdots < R_{n;k} \end{aligned} \tag{6.4}$$

where $N_{i;k}$ refers to the ith order statistic from the kth Monte Carlo trial for the null distribution and $R_{i;k}$ refers to the ith order statistic from the kth Monte Carlo trial for the reference distribution. Once this ordering has been performed, the Ozturk Algorithm calculates standardized order statistics for each distribution.

The equations describing this process are listed below:

$$Y_i = \frac{X_i - \mu_x}{\sigma_x}; \quad i = 1, 2, 3, \ldots, n \tag{6.5}$$

$$M_i = \frac{1}{2000} \sum_{k=1}^{2000} \frac{(N_i)_k - \mu_{n;k}}{\sigma_{n;k}}; \quad i = 1, 2, 3, \ldots, n \tag{6.6}$$

$$S_i = \frac{1}{2000} \sum_{k=1}^{2000} (R_i)_k; \quad i = 1, 2, 3, \ldots, n \tag{6.7}$$

where μ_x and σ_x correspond to the sample mean and sample standard deviation for the sample data set and $\mu_{n;k}$ and $\sigma_{n;k}$ correspond to the sample mean and sample standard deviation for the kth Monte Carlo simulation for the null distribution. Since the reference distribution used by the Ozturk Algorithm is the Standard Normal, $\mu_{r;k} = 0$ and $\sigma_{r;k} = 1$. It is for this reason that these parameters are omitted in the calculation of S_i. Essentially, the equations listed above indicate that the ith standardized order statistics for the null distribution and reference distribution are calculated by taking the average of 2000 ith standardized order statistics that are generated from Monte Carlo simulations.

Once the above calculations have been made, the Ozturk Algorithm computes the length and orientation of each linked vector. Two sets of linked vectors are calculated: one for the sample data and one for the null distribution. The ith linked vector in each set corresponds to the ith ordered sample in the data set. The following equations describe how the magnitude and angle of these linked vectors are determined:

$$a_i = \frac{|Y_i|}{n} \tag{6.8}$$

$$b_i = \frac{|M_i|}{n} \tag{6.9}$$

$$\theta_i = \pi\theta(S_i); \theta(S_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{S_i} e^{\left(\frac{-t^2}{2}\right)} \mathrm{d}t \tag{6.10}$$

where a_i represents the lengths of the ith linked vector for the sample data, b_i represents the length of the ith linked vector for the null distribution, and θ_i represents the orientation (measured from the horizontal axis) of the ith linked vector for both the sample data and the null distribution. Notice from these equations that only the reference distribution determines the angle of the ith linked vector, and that this angle is the same for both the sample data and the null distribution. Similarly, it is the magnitude of the ith standardized order statistic that determines the length of the ith linked vector.

Now that the formulas used to calculate the linked vectors have been given, it is possible to define a U-V plane where the ordered pair $Q_k = (U_k, V_k)$ is

specified as follows:

$$U_0 = V_0 = 0 \tag{6.11}$$

$$U_k = \frac{1}{n}\sum_{i=1}^{k} |Y_i|\cos(\theta_i); \quad k = 1, 2, 3, \ldots, n \tag{6.12}$$

$$V_k = \frac{1}{n}\sum_{i=1}^{k} |\mathrm{Y}_i|\sin(\theta_i); \quad k = 1, 2, 3, \ldots, n \tag{6.13}$$

A similar ordered pair, $P_k = (U_{0k}, V_{0k})$, is likewise defined below:

$$U_{00} = V_{00} = 0 \tag{6.14}$$

$$U_{0k} = \frac{1}{n}\sum_{i=1}^{k} |M_i|\cos(\theta_i); \quad k = 1, 2, 3, \ldots, n \tag{6.15}$$

$$V_{0k} = \frac{1}{n}\sum_{i=1}^{k} |\mathrm{M}_i|\sin(\theta_i); \quad k = 1, 2, 3, \ldots, n \tag{6.16}$$

The first set of equations defines the trajectory of linked vectors associated with the sample data, while the second set of equations defines the trajectory of linked vectors associated with the null distribution. The Ozturk Algorithm graphs the mapped trajectories found in the goodness-of-fit test by connecting the set of points described by Q_k and P_k for $k = 1, 2, 3, \ldots, n$.

A few general comments about the above equations are worth mentioning. One important thing is to remember that the angle of the ith linked vector is purely determined by the reference distribution is the same for both the sample data and the null distribution, and based on how it is defined, starts at 0 and increases to a maximum of 180 degrees (as measured from the horizontal axis). For this reason the general shape of all of the trajectories in the goodness-of-fit test are similar (Figure 6.4). In addition, due to the definition of the standardized order statistic, the magnitude of the ith linked vector for the sample data and null distribution start out large, then decrease to almost zero, and then become large again. The length of these vectors reach a minimum around an angle of 90°. It can also be seen from the above equations that Q_n and P_n represent the endpoint of the linked vectors for the sample data and null distribution, respectively. Finally, the trajectories defined by Q_k and P_k for $k = 1, 2, 3, \ldots, n$ become smoother as the number of samples (n) increases.

Although the Ozturk Algorithm only generates confidence ellipses for the end point of the null distribution trajectory, qualitative information about the goodness-of-fit test is provided by comparing the path of the trajectory for the sample data with that for the null distribution. It is reasonable to expect that the linked vectors from the sample data will closely follow the linked vectors from the null distribution if the sample data are consistent with the null

distribution. As a result, the goodness-of-fit test provides visual information about how well the ordered set of data fit the null distribution (Figure 6.4). Nevertheless, the Ozturk Algorithm also uses statistical procedures for quantifying the goodness-of-fit test through the calculation of confidence ellipses. This procedure is described next.

6.1.2.4. Calculation of Confidence Ellipses

The confidence ellipses provide a quantitative statistical description of how consistent the sample data are with the null distribution. As discussed earlier, if the end point of the trajectory for the sample data (Q_n) falls within one or more of the ellipses, the sample data set is said to be statistically consistent with the null hypothesis at a confidence level based on the confidence ellipses. Each confidence contour is directly associated with a significance level and is defined as the conditional probability that the end point of the trajectory for the sample data (Q_n) falls inside the specified ellipse given that the data comes from the null distribution. Therefore, the 99% confidence ellipse corresponds to a 0.01 level of significance, the 95% confidence ellipse corresponds to a 0.05 level of significance, and the 90% confidence ellipse corresponds to a 0.10 level of significance.

To understand how the Ozturk Algorithm calculates the confidence ellipses, one must remember that the algorithm performs 2000 Monte Carlo simulations of sample size n for the null distribution. Therefore, 2000 end points are calculated, and the average of these values is used as the actual value of the end point (Q_n). As a result, the endpoint (Q_n) corresponds to the ordered pair (U_{0n}, V_{0n}) where both U_{0n} and V_{0n} are random variables consisting of 2000 samples. It is from these random variables that the confidence ellipses are calculated.

The actual nature of the underlying distribution for the random variables U_{0n} and V_{0n} is unknown. However, if these variables were described by a bivariate Gaussian distribution, the calculation of the confidence ellipses would be straightforward. A three dimensional bell shaped Gaussian curve could be fit to the 2000 end points calculated by the Monte Carlo simulation and the corresponding confidence ellipses plotted for the desired significance levels. If a correlation existed between the random variables U_{0n} and V_{0n} the confidence contours would be ellipses, but if no correlation existed, the confidence contours would be circles.

Unfortunately, however, as just mentioned, the underlying distributions of U_{0n} and V_{0n} are not known. Therefore, the Ozturk Algorithm uses a complicated procedure to transform the random variables U_{0n} and V_{0n} into a single random variable that is bivariate Gaussian. After this transformation is made, the confidence ellipses are calculated. Finally, the actual confidence contours are obtained by taking the inverse transform of the confidence ellipses obtained from the bivariate Gaussian random variable. A family of distributions called the Johnson System is used to perform the transformation on U_{0n} and V_{0n} in order to obtain a bivariate Gaussian distribution. The details of this transformation are tedious and beyond the scope of this chapter. A complete description of this process can be found in (Shah[2]).

Two comments about the confidence ellipses are worth mentioning. For one thing, the confidence contours will only be ellipses if U_{0n} and V_{0n} can be described by a bivariate Gaussian distribution without undergoing any transformation. In all other cases, these confidence contours may not look like ellipses due to the transformation process. Hence, the phrase "confidence ellipses" is used loosely in this chapter since many of these contours are not actually ellipses. In addition, the size of the confidence ellipses depends on the sample size (n) and the selection of the null distribution. As the sample size decreases, the variability in P_n increases, and the confidence ellipses become larger. Similarly, if a null distribution is selected that has a high degree of variability, the variability of P_n will also be high, and the confidence ellipses will be relatively larger (for a given sample size). If the sample size used is too small (less than 50 samples), the confidence ellipses become so large that the goodness-of-fit test becomes meaningless; almost any sample data will be consistent with the null hypothesis. Nevertheless, an inspection of the mapped trajectories between the sample data and the null distribution provides a qualitative indicator of how consistent the data and the null distribution are.

6.1.2.5. The Best-Fit Test

The best-fit test uses a distribution approximation technique and is simply an extension of the goodness-of-fit test. A detailed discussion of this procedure can be found in Shah[2] and Ozturk.[1]

In the goodness-of-fit test, it was explained how trajectories for the sample data and null distribution get mapped into the U-V plane. The ordered pairs, Q_n and P_n, correspond to the end points of these trajectories. In the best-fit test, this procedure is extended one step further; the end point for every distribution is calculated and plotted in the U-V plane. In addition, since the end point for a particular distribution depends on the value of the shape parameters, for those distributions where shape parameters exist, end points for several values of these shape parameters are computed, and these end points are mapped as trajectories in the U-V plane. In other words, only one end point is calculated for those distributions that lack any shape parameters (Normal, Uniform, Exponential, Laplace, Logistic, Cauchy, and Extreme Value [type-1], and this end point maps as a point in the U-V plane; several end points are calculated for those distributions containing one shape parameter [Gumbel (type-2)], Gamma, Pareto, Weibull, Lognormal, and K-distribution), and these end points map as a line in the U-V plane; and even more end points are calculated for those distributions containing two shape parameters (Beta and SU Johnson), and these end points map as a series of lines in the U-V plane. As an example, the Ozturk Algorithm calculates end points for the Weibull distribution for the following values of the shape parameter: 0.3, 0.4, 0.5, 0.6, 0.8, 1.1, 1.5, 2.0, 3.0, and 5.0. These end points are then mapped in the U-V plane, connected and correspond to the line labeled Weibull in the distribution approximation chart (Figure 6.2). When a distribution contains two shape parameters, then the first shape parameter is held constant

at some value while the second shape parameter is varied, and then this process is repeated for different values of the first shape parameter. It is important to keep in mind that each end point is computed from 2000 Monte Carlo simulations where the number of samples used in the simulation (n) matches the number of points in the sample data set. Therefore, the appearance of the U-V plane for the best-fit test (Figure 6.2) is dependent upon the number of samples in the data set being analyzed.

Although the above process may seem very tedious and time consuming, the Ozturk Algorithm speeds up this process considerably by tabulating data. In other words, the coordinates of the end points for specific values of the shape parameters and a specific distribution are stored in a data table as a function of sample size. Using this technique, it becomes unnecessary to recalculate these end points with each simulation, which would greatly slow down the processing time. In addition, for those distributions that contain shape parameters, the Ozturk Algorithm only calculates (stores) end points for a fixed number of values for these shape parameters. These points are then connected together to form the lines in the U-V plane found in the distribution approximation chart.

Once a distribution approximation chart is generated for all of the distributions stored in the algorithm's library and for the particular sample size of the data set, the sample data is mapped in the U-V plane. Recall that the coordinates of the end point for the sample data correspond to the point, Q_n. This graphical solution to the best-fit test is referred to by the Ozturk Algorithm as the Distribution Approximation Chart (Figure 6.2). Inspection of this plot allows one to determine which distributions most closely fit the data; those distributions that lie closest to the mapped data point (Q_n) provide the best fit. In addition to this graphical solution, the Ozturk Algorithm also determines which five distributions best fit the data. The algorithm determines the closest point or trajectory by projecting the mapped data point to neighboring points or trajectories on the chart, and selecting the point or trajectory whose perpendicular distance from the sample point is the smallest. If the identified density functions contain shape parameters, the values of these parameters are determined through interpolation. One must remember that the distributions are mapped in the U-V plane by connecting the end points corresponding to selected values of the shape parameters. It should also be noted that the library of density functions has been carefully chosen such that most of the U-V plane is filled with mapped end points or trajectories of end points. In so doing, the algorithm ensures that regardless of where the end point of the sample data is mapped in the U-V plane, there will be a mapped distribution located close to this end point.

After the algorithm identifies the five density functions which best fit the sample data and interpolates the value of the shape parameters (if applicable), then the standard form of the density function is completely characterized (as indicated in Tables 4.8 and 4.9). However the location (α) and scale (β) parameters still need to be computed. Recall that the Ozturk Algorithm makes use

of standardized order statistics that are independent of linear transformations such as the location and scale parameters. Therefore, to compute the location and scale parameters, the process of taking standardized order statistics must be reversed. The process used by the algorithm to compute the value of the location and scale parameters is described next.

6.1.2.6. Estimation of Location and Scale Parameters

The Ozturk Algorithm calculates the value of the location (α) and scale (β) parameters by inverting the procedure used to generate the standardized order statistics. Recall that the random variable X corresponds to the sampled order data while the random variable Y corresponds to the standardized order statistics for the sample data. The relationship between these two variables is that $Y_i = (X_i - \alpha)/\beta$ where Y_i represents the ith standardized order statistic and X_i represents the ith sampled order data. Once the Ozturk Algorithm specifies the distributions that best fit the data along with the corresponding value of the shape parameters (if applicable), then the random variable Y is completely described according to the standard form of the density functions listed in Tables 4.8 and 4.9. In order to obtain a complete description of the random variable X, the location and scale parameters must be computed. To calculate the value of these two parameters, the algorithm defines two new variables as follows:

$$C_1 = \sum_{i=1}^{n} \cos(\theta_i) X_i \tag{6.17}$$

$$C_2 = \sum_{i=1}^{n} \sin(\theta_i) X_i \tag{6.18}$$

where θ_i corresponds to the angle defined by Equation 6.10, X_i refers to the random variable describing the ith ordered sample from the data set, and n corresponds to the number of samples in the data set. Based on the definition of the above two variables and making use of the substitution $E[X_i] = \beta\mu_i + \alpha$, the expected values of C_1 and C_2 are equal to:

$$E[C_1] = \sum_{i=1}^{n} \cos(\theta_i)(\beta\mu_i + \alpha) \tag{6.19}$$

$$E[C_2] = \sum_{i=1}^{n} \sin(\theta_i)(\beta\mu_i + \alpha) \tag{6.20}$$

where μ_i represents the sample mean of Y_i. The value of μ_i is calculated by the Ozturk Algorithm by running 2000 Monte Carlo simulations for the standard form of the best-fit density function described by Y_i and averaging the means of these simulations. In other words, the algorithm uses the standard form of the selected best-fit distribution (as described in Tables 4.8 and 4.9) to generate n samples

from this density function, and takes the average of these samples. This process is repeated for 2000 simulations, and the average of all of these simulations is the value used for μ_i.

The equations describing the expected values of C_1 and C_2 can be rewritten as:

$$E[T_1] = a\alpha + b\beta \tag{6.21}$$

$$E[T_2] = c\alpha + d\beta \tag{6.22}$$

where

$$a = \sum_{i=1}^{n} \cos(\theta_i) \tag{6.23}$$

$$b = \sum_{i=1}^{n} \mu_i \cos(\theta_i) \tag{6.24}$$

$$c = \sum_{i=1}^{n} \sin(\theta_i) \tag{6.25}$$

$$d = \sum_{i=1}^{n} \mu_i \sin(\theta_i) \tag{6.26}$$

Since the standardized Gaussian distribution is used by the Ozturk Algorithm as the reference distribution, θ_i starts at 0 ($i = 1$) and ends up at $\pi(i = n)$. Therefore, $a = 0$, as defined above. Rearranging the preceding equations and letting $a = 0$, it can be shown that:

$$\alpha = \frac{E[T_2] - d\beta}{c} \tag{6.27}$$

$$\beta = \frac{E[T_1]}{b} \tag{6.28}$$

If the sample size is sufficiently large (i.e., $n > 50$), it can be shown that $E[T_1] = T_1$ and $E[T_2] = T_2$ (Shah[2]). In addition, since the algorithm computes the value of μ_i, as described above, and the value of θ_i is also calculated, values for the variables b, c and d are easily obtained. Using these values, the algorithm computes the value of the location and scale parameters. Once these two parameters are calculated, the general form of the density function describing the random variable X is completely defined (as indicated in Tables 4.8 and 4.9).

In this section of the chapter, a detailed description of the technique used by the Ozturk Algorithm to perform a goodness-of-fit test and best-fit test was provided. In the next three sections, three applications of the Ozturk Algorithm in the field of neuroscience will be presented.

6.1.3. Analysis of Spontaneous Auditory Nerve Activity of Chinchillas

The usefulness of the Ozturk Algorithm for analyzing data will first be demonstrated using data from the auditory nerve. All of the data analyzed in this section was provided by Evan Relkin and was recorded from chinchillas. One of the major limitations when recording from a single unit is the ability to maintain a stable recording from an auditory fiber for an extended period of time. Nevertheless, one must make sure that enough data are collected so that an accurate analysis of these data can occur. It would be of considerable advantage to the investigator if spike trains could be recorded for a shorter amount of time for each stimulating condition without compromising the information gained from these data. In this section, it will be shown how the Ozturk Algorithm can be used to gather data more efficiently.

In the first part of this study, spontaneous nerve activity was analyzed. The random variable inputted into the algorithm was the time between successive spikes (i.e., interspike intervals). Since these spike trains are undriven, they are not subjected to adaptation, and hence, they represent random, independent events. One must recall that the Ozturk Algorithm assumes the data set to consist of independent trials. For this reason, driven spike trains are not considered. Two different fibers were analyzed, one with a spontaneous activity of 38 spikes/sec (fiber #1) and one with a spontaneous activity of 74 spikes/sec (fiber #2). According to Liberman,[5] both of these fibers would be classified as high spontaneous rate (SR) fibers. (Liberman classifies those fibers having an SR greater than 17.5 spikes/sec as high SR, between 0.5 and 17.5 spikes/sec as medium SR, and below 0.5 spikes/sec as low SR.) The purpose of this portion of the study was to determine the minimum number of interspike intervals necessary to accurately characterize the interspike interval histogram. In addition, the Ozturk Algorithm was used to determine how the best-fit distributions, for fibers of two different spontaneous activities, compare.

For the fiber with a spontaneous activity of 38 spikes/sec (fiber #1), 2198 different interspike intervals were obtained while for the fiber with a spontaneous activity of 74 spikes/sec (fiber #2) a total of 7053 interspike intervals were collected. All of these intervals were inputted into the Ozturk Algorithm and the five best-fit distributions for each fiber were computed. Tables 6.1 and 6.2 summarize the five best-fit density functions for each fiber, respectively. For both fibers, the K-distribution provides the best fit to the data while the Gamma distribution provides the second best fit. Due to the complexity of the K-distribution (this density function involves a form of the Bessel function that is described by a series) the Gamma distribution will be used as the best-fit approximation for the interspike interval histograms of these fibers. Figure 6.5 shows the graphical solution to the best-fit test for each fiber. From this figure, it can be seen that the Gamma function, with an appropriate value for the shape parameter, provides an excellent fit to both sets of data. Thus, the Gamma distributions as described in Tables 6.1 and 6.2 will be

TABLE 6.1
Parameters of Five Best-Fit Distributions, Listed in Order of Best Fit, for all 2198 Samples of Interspike Intervals of Chinchilla Auditory Nerve Activity Collected from Fiber #1

Best Fit Distributions	Location Parameter (α)	Scale Parameter (β)	Shape 1 Parameter (γ)	Shape 2 Parameter (δ)
K-distribution	−0.95241E-02	0.18754E-01	0.45784E-01	—
Gamma	−0.13320E-02	0.16496E-01	0.16635E-01	—
Weibull	−0.13467E-02	0.26395E-01	0.11726E-01	—
Lognormal	−0.14145E-01	0.35292E-01	0.50601E-00	—
Exponential	−0.46305E-02	0.22387E-01	—	—

assumed to provide the best characterization of the interspike interval histograms for the two fibers.

In Figure 6.5, the location of the mapped data set is represented in the U-V plane by a filled circle. For fiber #1 the five distributions that are located closest to the mapped data point are the K-distribution, Gamma, Weibull, Lognormal, and Exponential. For fiber #2, the five distributions that are closest to the mapped data point are the K-distribution, Gamma, Exponential, Weibull, and Lognormal. [N = Normal, U = Uniform, E = Exponential, A = Laplace, L = Logistic, C = Cauchy, V = Extreme Value (type-1), T = Gumbel (type-2), G = Gamma, P = Pareto, W = Weibull, L = Lognormal, K = K-distribution. The upper five dashed lines represent Beta, and the lower nine dashed lines represent SU Johnson.]

In the next portion of this analysis, subsets of points from the sample data were inputted into the algorithm to determine the minimum number of points necessary to accurately characterize the data. For fiber #1, subsets of interspike

TABLE 6.2
Parameters of Five Best-Fit Distributions, Listed in Order of Best Fit, for all 7053 Samples of Interspike Intervals of Chinchilla Auditory Nerve Activity Collected from Fiber #2

Best fit Distributions	Location Parameter (α)	Scale Parameter (β)	Shape 1 Parameter (γ)	Shape 2 Parameter (δ)
K-Distribution	−0.49604E-02	0.10898E-01	0.41161E-01	—
Gamma	−0.33691E-03	0.93693E-02	0.15965E-01	—
Exponential	0.26819E-02	0.12407E-01	—	—
Weibull	0.10954E-02	0.14334E-01	0.11495E-01	—
Lognormal	−0.72029E-02	0.19000E-01	0.51858E-00	—

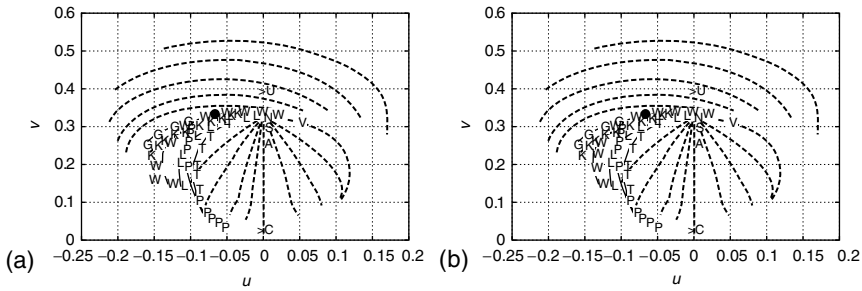

FIGURE 6.5 Distribution approximation charts for two data sets. (a) fiber #1 data set. (b) fiber #2 data set. • mapped end point of entire data set.

intervals containing either 500, 150, 100, 75 or 50 data points were also analyzed by the algorithm. For each of these subsets, four different sets of data were used (i.e., four sets of 500 points, four sets of 150 points, etc.). These four sets were selected such that none of the data points appeared in more than one of the four subsets. The Ozturk Algorithm was then used to determine the best-fit Gamma distribution for each of these groups. (Recall that using all of the data, the best-fit density function for both fibers was the Gamma distribution.)

Figure 6.6 summarizs these results. The graphs on the left show plots of the best-fit Gamma distributions for each of the four subsets of points; the bolded line in each plot represents the best-fit Gamma distribution for all of the data points collected from fiber #1. The graphs on the right show a portion of the graphical solution to the best-fit test as determined by the Ozturk Algorithm. These graphs show where in the U-V plane each of the four subsets of data is mapped. The filled circle in these plots represents the location of the mapped end point for the entire data set. In other words, the graphs on the right show the U-V plane for the distribution approximation chart (see Figure 6.5) without actually showing where each of the distributions listed in Tables 4.8 and 4.9 gets mapped in this plane; only the location of the mapped data point for each of the four subsets, along with the location of the mapped data point for the entire data set are shown. (Recall that the mapped end point for each set of sample data corresponds to the point Q_n, as defined in the previous section.)

In Figure 6.6, the bolded line represents the best-fit Gamma distribution for the entire data set and the symbol "x" represents the corresponding mapped end point in the U-V plane. The parameter values were determined by the Ozturk Algorithm. Figure 6.6(a): open circles, $\alpha = -0.0028$, $\beta = 0.015$, $\gamma = 2.0$; filled circles, $\alpha = 0.00024$, $\beta = 0.016$, $\gamma = 1.8$; open squares, $\alpha = -0.0017$, $\beta = 0.016$, $\gamma = 1.6$; filled squares, $\alpha = -0.0012$, $\beta = 0.018$, $\gamma = 1.6$. Figure 6.6(b): open circles, $\alpha = -0.0033$, $\beta = 0.016$, $\gamma = 1.8$; filled circles, $\alpha = -0.00055$, $\beta = 0.013$, $\gamma = 1.9$; open squares, $\alpha = -0.0019$, $\beta = 0.012$, $\gamma = 1.9$; filled squares, $\alpha = -0.0029$, $\beta = 0.019$, $\gamma = 1.7$. Figure 6.6(c): open circles,

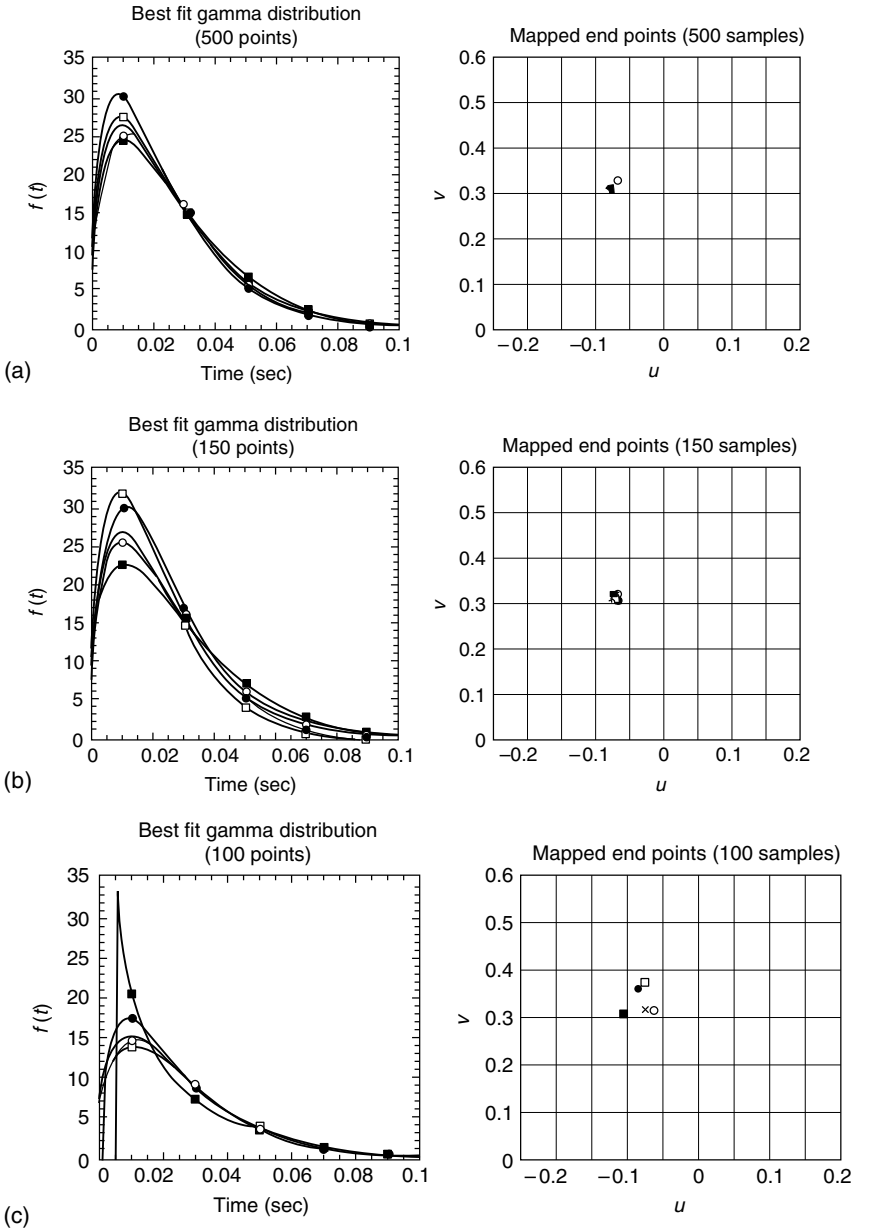

FIGURE 6.6 Best-fit gamma distributions and mapped end points for fiber #1 interspike intervals. Distributions are shown for entire data set of 2198 samples and also for four subsets. (a) 500 samples in subset. (b) 150 samples in subset. (c) 100 samples in subset. (d) 75 samples in subset. (e) 50 samples in subset.

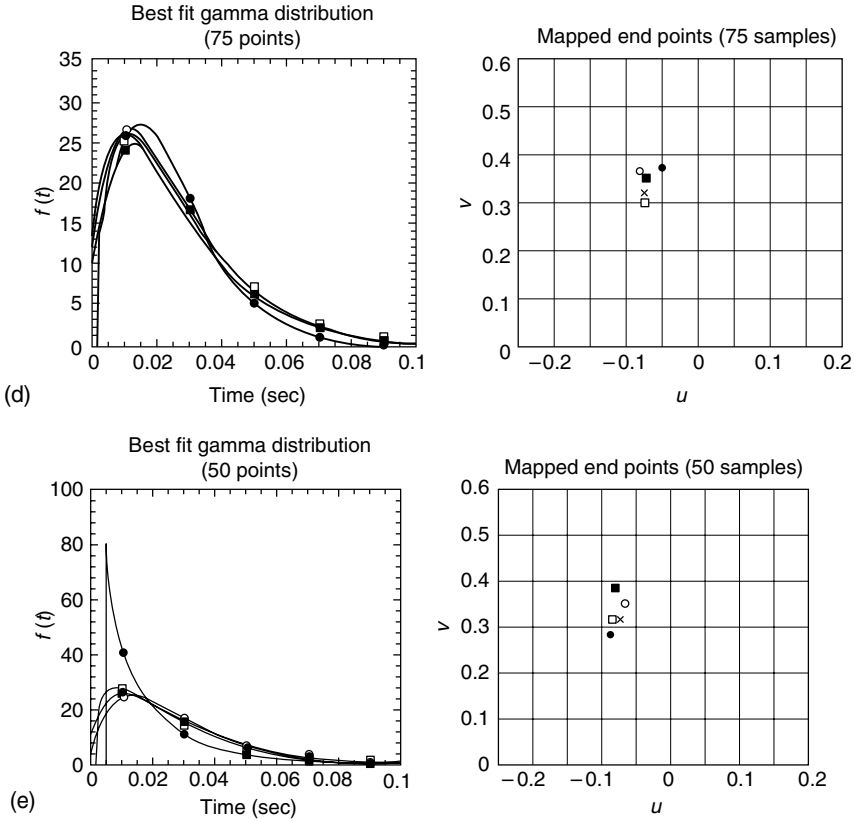

FIGURE 6.6 Continued.

$\alpha = -0.003329$, $\beta = 0.0155$, $\gamma = 2.0$; filled circles, $\alpha = 0.0011$, $\beta = 0.017$, $\gamma = 1.4$; open squares, $\alpha = -0.0021$, $\beta = 0.017$, $\gamma = 1.8$; filled squares, $\alpha = 0.0050$, $\beta = 0.030$, $\gamma = 0.79$. Figure 6.6(d): open circles, $\alpha = -0.0017$, $\beta = 0.014$, $\gamma = 1.9$; filled circles, $\alpha = -0.0070$, $\beta = 0.0090$, $\gamma = 3.4$; open squares, $\alpha = 0.0012$, $\beta = 0.018$, $\gamma = 1.6$; filled squares, $\alpha = -0.0032$, $\beta = 0.014$, $\gamma = 2.2$. Figure 6.6(e): open circles, $\alpha = -0.0013$, $\beta = 0.014$, $\gamma = 2.0$; filled circles, $\alpha = -0.0043$, $\beta = 0.021$, $\gamma = 0.80$; open squares, $\alpha = 0.0019$, $\beta = 0.021$, $\gamma = 1.3$; filled squares, $\alpha = -0.0025$, $\beta = 0.015$, $\gamma = 1.9$.

Inspection of Figure 6.6 qualitatively reveals that the Ozturk Algorithm does a reasonable job approximating the interspike interval histograms with as few as 50 samples. The plots of the mapped end points indicate that the variability in the location of the mapped data point when 150 samples or more are used is extremely small. However, even when as few as 50 samples are used, the location of the mapped subsets are all scattered close to the location of the mapped end point for the entire data set. The graphs of the best-fit Gamma

distributions are even more compelling. For all but two of the data sets (filled squares with 100 samples and filled circles with 50 samples) the best-fit Gamma distributions almost overlay exactly with the best-fit Gamma distribution for the entire data set, and even the two data sets that deviate the most only differ in a small portion of the function; the tails of these distributions all overlap. Therefore, even without any quantitative analysis, these figures indicate that 50 to 100 samples are sufficient for characterizing interspike interval histograms. Nevertheless, a goodness-of-fit test was also performed with the Ozturk Algorithm for each subset of data using the best-fit Gamma distribution for the entire data set as the null distribution, and in *every* case the subset of points was statistically consistent with the null distribution with a 99% confidence level. Figure 6.7 shows the graphical solution for four of these goodness-of-fit tests, two using 500 points and two using 100 points. In all four cases, the trajectory for the sample data closely follows the trajectory for the null distribution. Also notice, as explained in the previous section, that the confidence ellipses are larger when fewer samples are used.

A similar analysis was also done for fiber #2; however, since this data set consists of more than 7000 interspike intervals, subsets containing 1000, 500,

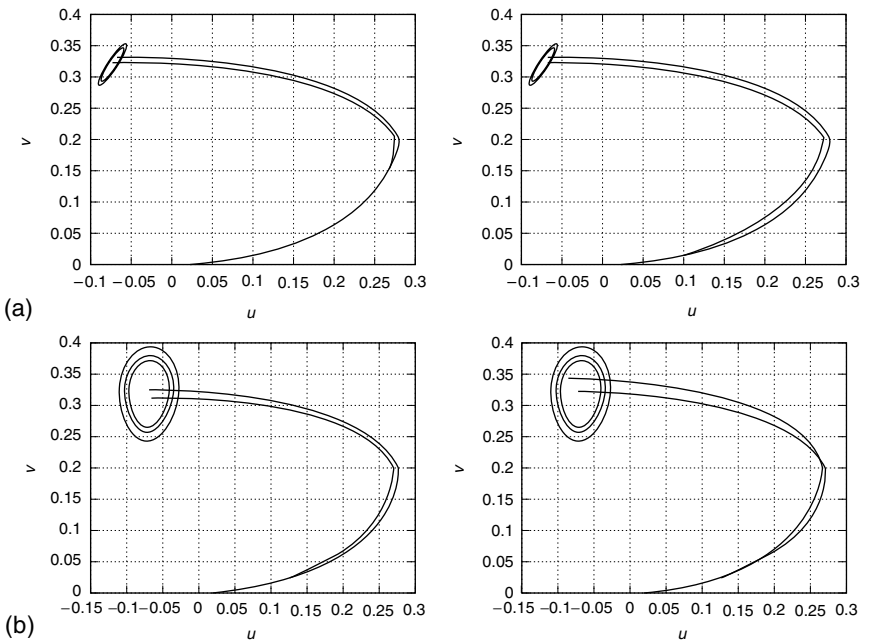

FIGURE 6.7 Goodness-of-fit tests for fiber #1 interspike intervals. Null hypothesis is best-fit gamma distribution for entire data set (i.e., $\gamma = 1.66$). Outer, middle, and inner ellipses correspond to confidence Levels of 99%, 95%, and 90%, respectively. (a) Two subsets of 500 samples. (b) Two subsets of 100 samples.

150, 100, 75 and 50 samples were used. As before, four groups of randomly selected interspike intervals for each sample size were analyzed, except for the 1000 point case where seven data sets were used. The results of these simulations by the Ozturk Algorithm are summarized in Figure 6.8(a) to (f) and are very similar to the results from the first fiber. Once again, these graphs indicate that 50 to 100 interspike intervals are sufficient for characterizing the interspike interval histogram. The Ozturk Algorithm was also used to perform a goodness-of-fit test for each subset where the best-fit Gamma distribution for the entire data set was used as the null distribution, and once again, *every* subset of data was statistically consistent with the null hypothesis with a 99% confidence level.

In Figure 6.8, the bolded line represents the best-fit Gamma distribution for the entire data set and the symbol "x" represents the corresponding mapped end point in the U-V plane. The parameter values were determined by the Ozturk Algorithm. Figure 6.8(a): open circles, $\alpha = -0.00065$, $\beta = 0.0091$, $\gamma = 1.7$; filled circles, $\alpha = -0.00021$, $\beta = 0.0088$, $\gamma = 1.5$; open squares, $\alpha = 0.000079$, $\beta = 0.0087$, $\gamma = 1.6$; filled squares, $\alpha = 0.00046$, $\beta = 0.0085$, $\gamma = 1.5$; open triangles, $\alpha = 0.00023$, $\beta = 0.0086$, $\gamma = 1.4$; filled triangles, $\alpha = -0.00068$, $\beta = 0.0086$, $\gamma = 1.7$; open diamonds, $\alpha = -0.00073$, $\beta = 0.0077$, $\gamma = 1.8$; Figure 6.8(b): open circles, $\alpha = -0.0091$, $\beta = 0.0079$, $\gamma = 1.8$; filled circles, $\alpha = -0.00029$, $\beta = 0.0072$, $\gamma = 1.8$; open squares, $\alpha = -0.00046$, $\beta = 0.0074$, $\gamma = 1.7$; filled squares, $\alpha = -0.00034$, $\beta = 0.0068$, $\gamma = 1.8$. Figure 6.8(c): open circles, $\alpha = -0.0070$, $\beta = 0.0057$, $\gamma = 3.8$; filled circles, $\alpha = 0.0011$, $\beta = 0.017$, $\gamma = 1.4$; open squares, $\alpha = -0.0021$, $\beta = 0.017$, $\gamma = 1.8$; filled squares, $\alpha = 0.0050$, $\beta = 0.030$, $\gamma = 0.79$. Figure 6.8(d): open circles, $\alpha = -0.0011$, $\beta = 0.0087$, $\gamma = 1.3$; filled circles, $\alpha = 0.0039$, $\beta = 0.013$, $\gamma = 0.75$; open squares, $\alpha = 0.0015$, $\beta = 0.010$, $\gamma = 1.3$; filled squares $\alpha = 0.0019$, $\beta = 0.0073$, $\gamma = 2.3$. Figure 6.8(e): open circles, $\alpha = 0.0036$, $\beta = 0.015$, $\gamma = 0.77$; filled circles, $\alpha = 0.0028$, $\beta = 0.018$, $\gamma = 0.88$; open squares, $\alpha = 0.00048$, $\beta = 0.0072$, $\gamma = 1.5$; filled squares, $\alpha = -0.00088$, $\beta = 0.0050$, $\gamma = 2.5$. Figure 6.8(f): open circles, $\alpha = -0.019$, $\beta = 0.0019$, $\gamma = 16$; filled circles, $\alpha = -0.0033$, $\beta = 0.0062$, $\gamma = 2.8$; open squares, $\alpha = 0.0048$, $\beta = 0.0089$, $\gamma = 0.37$; filled squares, $\alpha = -0.0017$, $\beta = 0.0039$, $\gamma = 3.4$.

In this portion of the study, the question addressed was how many samples are required to adequately produce an interspike interval histogram. The above results indicate that 50 to 100 interspike intervals are sufficient for characterizing the histogram. Furthermore, one should keep in mind that once the best-fit density function is determined, a great deal of additional information (e.g., mean and variance) also becomes available. In the next portion of this study, the Ozturk Algorithm was used to compare the best-fit density function for these two fibers. Since the spontaneous activity of these fibers differs by a factor of 2 (38 spikes/s for fiber #1 and 74 spikes/s for fiber #2) it will be interesting to see whether a single family of distributions with specific shape parameters can describe the interspike interval histogram for both fibers.

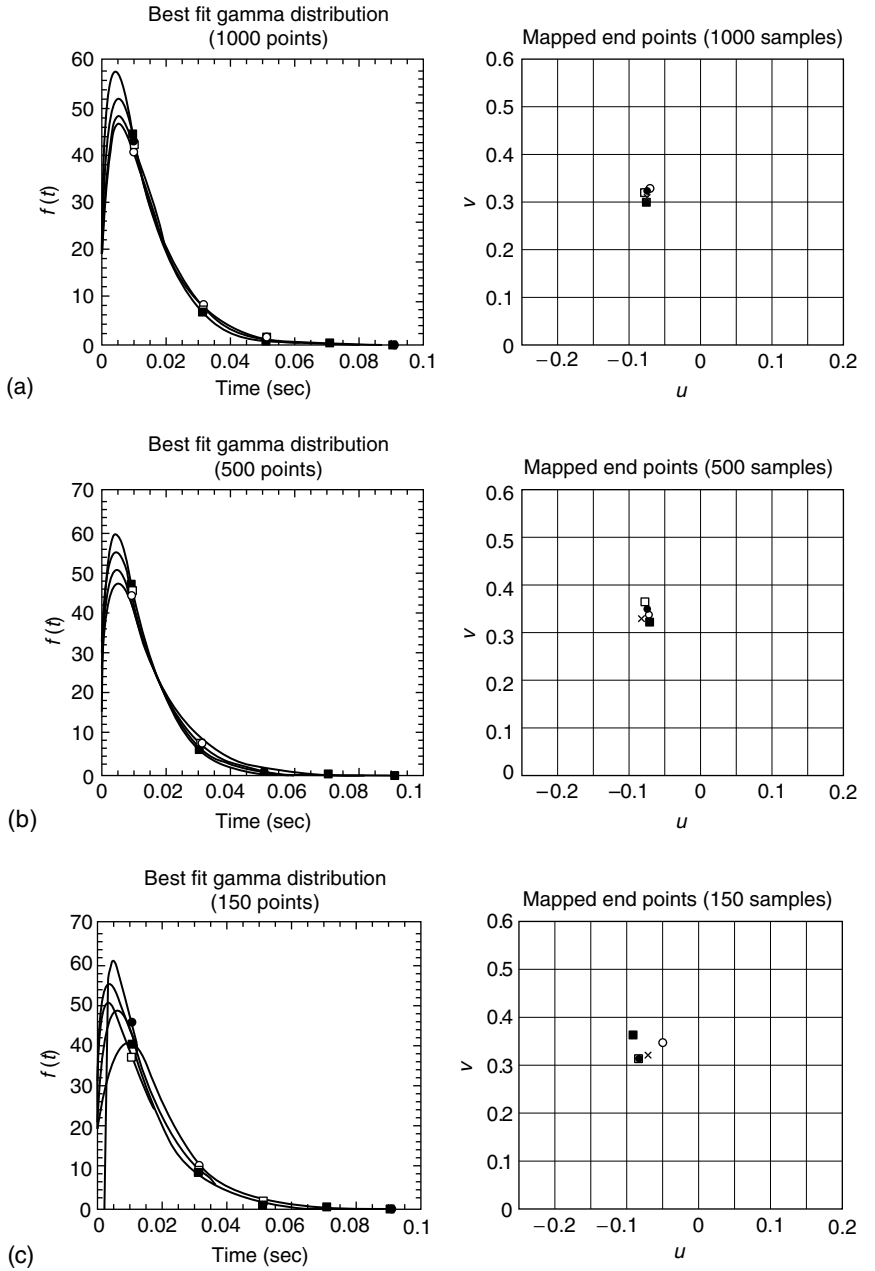

FIGURE 6.8 Best-fit gamma distributions and mapped end points for fiber #2 interspike intervals. Distributions are shown for entire data set of 7053 samples and four subsets. (a) 1000 samples in subset. (b) 500 samples in subset. (c) 150 samples in subset. (d) 100 samples in subset. (e) 75 samples in subset. (f) 50 samples in subset.

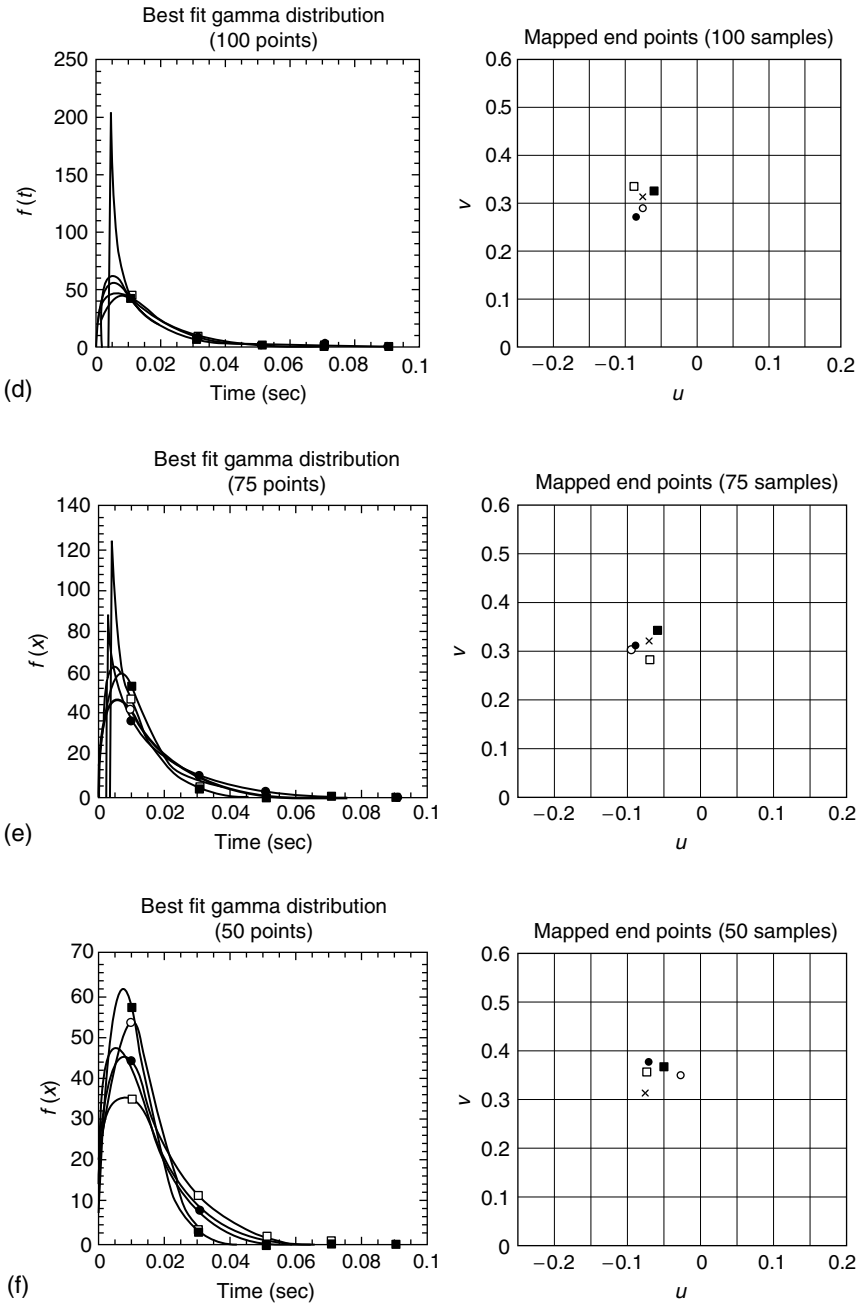

FIGURE 6.8 Continued.

6.1.3.1. Analysis of Two Fibers with Different Spontaneous Rates

In the first part of this study, it was shown that 50 to 100 samples are sufficient for characterizing an interspike interval histogram, and that the best-fit density function is the Gamma distribution. The bottom-right graph in Figure 6.9 demonstrates how well the best-fit Gamma distribution describes binned data for 1000 interspike intervals from fiber #2. However, one interesting finding from this study was that the best-fit density function for both fibers #1 and #2 was the Gamma distribution despite the fact that the spontaneous activity of these fibers differs considerably (38 spikes/s as compared to 74 spikes/s). Therefore, one question that arises is whether a single type of density function with fixed

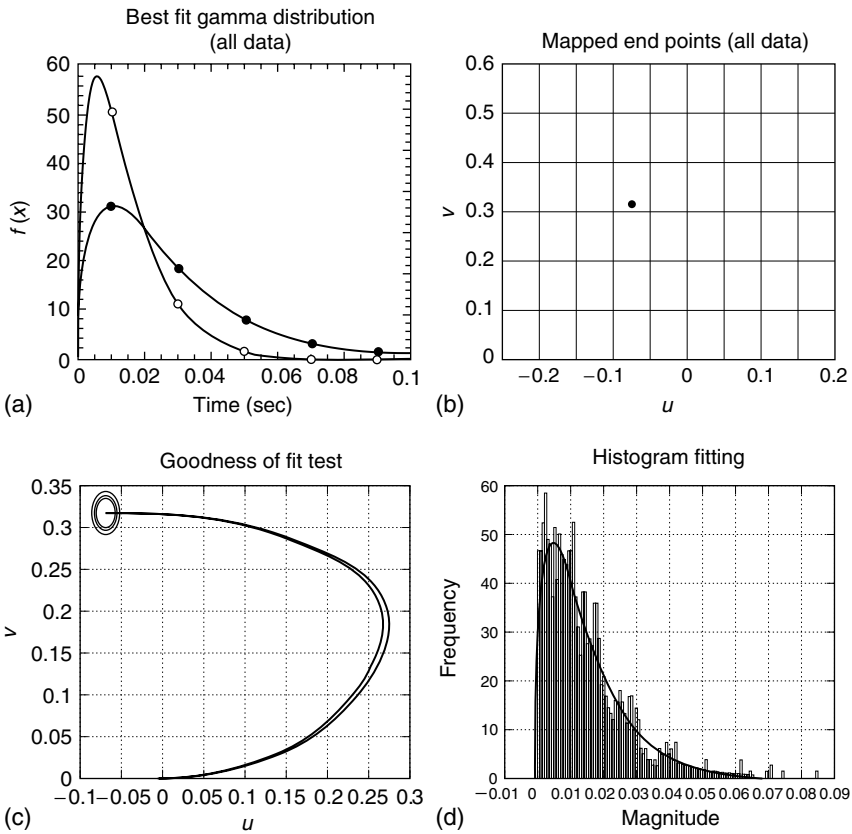

FIGURE 6.9 Comparison of best-fit tests of sampled interspike intervals for two fibers with different spontaneous activity rates. (a) Best-fit gamma distributions for fiber #1 (filled circles) and fiber #2 (open circles). (b) Overlapping mapped end points for fiber #1 (filled circles) and fiber #2 (open circle). (c) Goodness-of-fit test for best-fit gamma distribution of fiber #1. Null hypothesis is best-fit gamma distribution of fiber #2. (d) Best-fit gamma distribution and histogram for 1000 data samples of fiber #2.

shape parameter values can describe the spontaneous activity of any auditory nerve fiber. Using the Ozturk Algorithm, it is possible to address this question.

Figure 6.9 provides some compelling evidence suggesting that the spontaneous activity for these two auditory nerve fibers is well described by a Gamma distribution with a shape parameter of approximately 1.63; the two high-SR fibers only differ in their values of the location (α) and scale parameters (β). Figure 6.9(a) shows a plot of the best-fit Gamma distribution, as listed in Tables 6.1 and 6.2, for fiber #1 (filled circles) and fiber #2 (open circles). The difference between these two functions reflects the fact that fiber #2 has a spontaneous rate approximately twice that of fiber #1. However, Figure 6.9(b) shows that the data set for both of these fibers gets mapped in the same location of the U-V plane. Therefore, even though the spontaneous activity of these two fibers is different, the same type of Gamma distribution provides the best fit for both sets of data; only the values of the location and scale parameters differ. In addition, when a goodness-of-fit test is performed using the best-fit Gamma distribution for fiber #2 as the null distribution and the interspike intervals for fiber #1 as the sample data, the results indicate that the two sets of data agree almost perfectly. Figure 6.9(c) shows the graphical solution of this goodness-of-fit test; the sample data are statistically consistent with the null hypothesis at a 99% confidence level, and even more remarkably, the two mapped trajectories overlay almost entirely. This result provides strong evidence that, at least for these two fibers, a Gamma distribution with a shape parameter of 1.63 characterizes the interspike intervals of spontaneous auditory nerve activity.

To more rigorously test this idea, an additional simulation was performed. Generally, the mathematical function thought to describe spontaneous spike generation in the auditory nerve is a Poisson process with recovery time. Therefore, a computer program was made to simulate spike generation from this process. The computer program was provided by Evan Relkin, and the recovery function used is one that was described by Gaumond et al.[6,7] The simulation was run for 50 seconds using a spontaneous activity rate of 45 spikes/sec (also corresponding to a high spontaneous rate fiber). From this simulation, 2129 interspike intervals were generated. These data were analyzed by the Ozturk Algorithm, and the best-fit density function was determined to be a Gamma distribution. More remarkably, however, was the fact that the location of the mapped end point in the U-V plane was extremely similar to that for the other two fibers. In addition, when a goodness-of-fit test was performed on these data using the best-fit Gamma distribution for fiber #2 as the null distribution, the two mapped trajectories for the sample data and null data matched almost perfectly; the simulated data were statistically consistent with the null distribution at a 99% confidence level. A summary of these results is given in Figure 6.10.

Although the spontaneous activity from different classifications of several more auditory nerve fibers and simulations (i.e., medium and low SR fibers) needs to be tested before definitive conclusions may be reached, the above examples demonstrate the usefulness of the Ozturk Algorithm. For one thing, the algorithm determined that a Gamma distribution with a shape parameter of

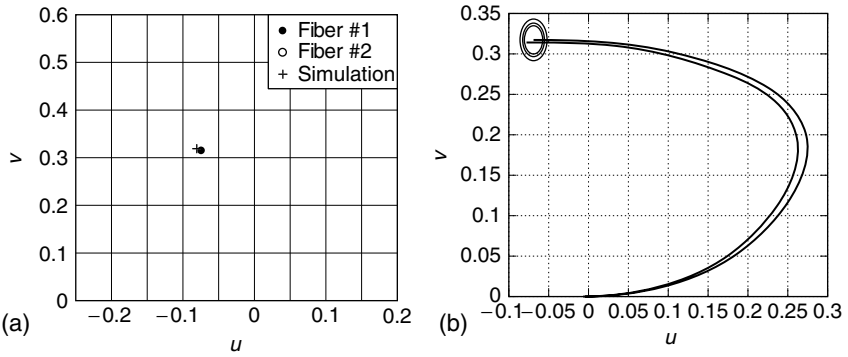

FIGURE 6.10 Computer simulation of 2129 interspike intervals from a poisson process with a Gaumond recovery function. (a) Mapped end points of simulation data and entire data sets of fibers #1 and #2. (b) Goodness-of-fit test for computer simulation. Null hypothesis is fiber #2 best-fit Gamma distribution with parameters $\alpha = 0.0013$, $\beta = 0.016$, $\gamma = 1.5$.

approximately 1.63 provides an excellent fit to interspike interval histograms for high-SR fibers. Secondly, the algorithm provides a method for testing the hypothesis that the spontaneous activity of all auditory nerve fibers can be described by this function. Thirdly, the Ozturk Algorithm allows a statistical verification of Gaumond's prediction that spontaneous activity in auditory fibers can be simulated using a Poisson process with a Gaumond recovery function. Based on the above simulation, his prediction appears to be an accurate one. And lastly, the algorithm enables an even more general test of spontaneous activity to be performed. Perhaps the spontaneous rate of all types of neurons can be characterized by the same function. A similar analysis to the above one performed on different classes of neurons would provide a method for addressing this question.

6.1.3.2. Analysis of Pulse-Number Distributions

The last part of the analysis in this section was done at the request of Evan Relkin. One method for analyzing spikes is to generate pulse-number distributions. In this analysis, the number of events (in this case spikes) are counted for a specific time interval, and these counts are binned for successive sweeps of the same stimulating condition (e.g., Relkin and Pelli[8]).

Dr. Relkin is currently analyzing some data using this technique (personal communication). In his study, a 25 msec interval is used for generating the pulse-number distributions. From these data, means and variances are computed. One problem arising with these calculations is that many more samples are required to accurately compute the variance than the mean; experimental evidence suggests that approximately the square number of samples are needed to compute the

variance with the same accuracy as the mean. To circumvent this difficulty, Dr. Relkin has generated an equation that relates mean and standard deviation for 25 msec intervals: $\sigma = 0.18\mu^{0.36}$ where σ corresponds to the standard deviation and μ corresponds to the mean. Since the Ozturk Algorithm determines the distribution of best-fit for a set of data, and the mean and variance of these distributions are known (see Table 6.3[20–24]), it was believed that the algorithm might provide a more efficient method for determining the variance of the pulse-number distributions.

The same two fibers were used in this study, and, as before, the data were provided by Evan Relkin. The stimulus paradigm was such that each stimulus condition contained 20 sweeps, and each sweep consisted of 200 ms of spontaneous activity. Pulse-number distributions were generated by counting the number of spontaneous spikes occurring in a 25 ms period. Therefore, one stimulus condition, consisting of 20 sweeps, contains 160 intervals of 25 ms. The number of 25 ms intervals used to bin the data was either 960, 500, 250, 200, 150, 100, 75, 50, or 25. For the case where 960 intervals were used, exactly six stimulus conditions of 20 sweeps were required.

The Ozturk Algorithm was used to analyze these data. The random variable inputted into the algorithm was the number of spikes occurring in 25 ms interval. In other words, if 250 intervals were used, then a data set containing 250 samples was entered into the algorithm, the value of each sample representing the number of spikes that occurred in a particular 25 ms interval. It is important to point out that this random variable is discrete. Nevertheless, since the algorithm is being used to predict means and variances, as opposed to curve fitting the data, a discrete random variable is acceptable.

For both fibers, nine different sets of data, each generated from a different number of samples, were analyzed. The Ozturk Algorithm was used to determine the best-fit density function for each data set. The means and variances for these distributions were computed. The formulas used for these calculations are listed in Table 6.3 and correspond to the general form of the density functions in Tables 4.8 and 4.9. One should remember that the actual class of the best-fit distribution is unimportant; it is only used to compute the means and variances. Tables 6.4 and 6.5 summarize the results obtained for fibers #1 and #2, respectively. The standard deviations were calculated by taking the square root of the variance.

When the means and standard deviations generated from the Ozturk Algorithm are compared to the best-fit equation determined by Dr. Relkin, the results are impressive.

Figure 6.11(a) shows these results. Although the variability around the curve is greater for the second fiber than the first one, this finding is not unusual. Dr. Relkin noticed when he was generating his curve fit that the variability around this curve increased considerably as the spontaneous rate of the fiber increased (personal communication). In fact, Dr. Relkin commented that the variability of the data from the Ozturk Algorithm was much less than the variability he normally observes.

TABLE 6.3
Mean and Variance of Standard and General Forms of PDFs in Tables 4.8 and 4.9

Distribution	Mean	Variance
1. Normal	α	β^2
2. Uniform	$\frac{2\alpha+\beta}{2}$	$\frac{\beta^2}{12}$
3. Exponential	$\alpha+\beta$	β^2
4. Laplace	α	$2\beta^2$
5. Logistic	α	$\frac{(\beta\pi)^2}{3}$
6. Cauchy	not defined	not defined
7. Extreme Value (type-1)	$\alpha+\beta\Gamma'(1)$ where $\Gamma'(1)=-0.57721$	$\frac{(\beta\pi)^2}{6}$
8. Gumbel (type-2)	$\Gamma(1-\gamma-1)+\alpha$	$\beta^2\Gamma(1-2\gamma-1)+2\alpha(\beta-1)\Gamma(1-\gamma^{-1})-\Gamma^2(1-\gamma^{-1})$
9. Gamma	$\beta\Gamma+\alpha$	$\beta^2\gamma$
10. Pareto	$\frac{\beta\gamma}{\gamma-1}+\alpha\gamma>1$	$\beta^2\left[\frac{\gamma}{\gamma-2}-\left(\frac{\gamma}{\gamma-1}\right)^2\right],\ \gamma>2$
11. Weibull	$\beta\Gamma\left(\frac{\gamma+1}{\gamma}\right)+\alpha$	$\beta^2\left[\Gamma\left(\frac{\gamma+2}{\gamma}\right)-\left\{\Gamma\left(\frac{\gamma+1}{\gamma}\right)\right\}^2\right]$
12. Lognormal	$\beta\exp^{\left(\frac{\gamma^2}{2}\right)}+\alpha$	$\beta^2\exp^{(\gamma^2)}(\exp^{(\gamma^2)}-1)$
13. K-Distribution	$\frac{1.77\beta\Gamma(\gamma+0.5)}{\Gamma(\gamma)}+\alpha$	$\beta^2\left[4\gamma-3.14\left(\frac{\Gamma(\gamma+0.5)}{\Gamma(\gamma)}\right)^2\right]$
14. Beta	$\frac{\beta\gamma}{\gamma+\delta}+\alpha$	$\frac{\beta^2\gamma\delta}{(\gamma+\delta)^2(\gamma+\delta+1)}$
15. Johnson SU	$[\beta\exp^{\left(\frac{1}{2\gamma^2}\right)}\sin h(\delta)]+\alpha$	$\frac{\beta^2}{2}[\exp^{\left(\frac{1}{\gamma^2}\right)}-1][\exp^{\left(\frac{1}{\gamma^2}\right)}\cos h(2\delta)+1]$

Note: The means and variances listed above correspond to the general form of the PDFs listed in Table 4.9.

TABLE 6.4
Parameters of Best-Fit Distributions for Nine Sets of 25 msec Samples of Interspike Intervals of Chinchilla Auditory Nerve Activity Collected from Fiber #1

Best-Fit Distribution	Number of Samples	Location Parameter (α)	Scale Parameter (β)	Shape 1 Parameter (γ)	Shape 2 Parameter (δ)	Distribution Mean	Distribution Std. Deviation
Beta	960	−0.71743E + 00	0.29501E + 01	0.15966E + 01	0.16E + 01	0.75605	0.72004
Beta	500	−0.64751E + 00	0.28556E + 01	0.14932E + 01	0.16E + 01	0.73099	0.70530
Weibull	250	−0.21597E + 01	0.32E + 01	0.47862E + 01	—	0.77016	0.70162
Weibull	200	−0.11257E + 01	0.21774E + 01	0.26765E + 01	—	0.81068	0.77958
Weibull	150	−0.16792E + 01	0.27022E + 01	0.3746E + 01	—	0.75954	0.73041
Beta	100	−0.28112E + 00	0.21496E + 01	0.66257E + 00	0.8E + 00	0.69269	0.68188
Weibull	75	−0.2314E + 01	0.33584E + 01	0.5E + 01	—	0.76958	0.70625
Johnson S.U.	50	0.89717E + 00	0.75416E + 00	0.12942E + 01	0	0.89717	0.80883
Beta	25	−0.63365E + 00	0.21895E + 01	0.11252E + 01	0.8E + 00	0.64602	0.63089

TABLE 6.5
Parameters of Best-Fit Distributions for Nine Sets of 25 msec Samples of Interspike Intervals of Chinchilla Auditory Nerve Activity Collected from Fiber #2

Best-Fit Distribution	Number of Samples	Location Parameter (α)	Scale Parameter (β)	Shape 1 Parameter (γ)	Shape 2 Parameter (δ)	Distribution Mean	Distribution Std. Deviation
Beta	960	−0.70106E + 00	0.53329E + 01	0.25861E + 01	0.32E + 01	1.6825	1.0178
Beta	500	−0.69648E + 01	0.37508E + 01	0.13357E + 01	0.16E + 01	1.6369	0.94149
Beta	250	0.80749E + 01	0.37225E + 01	0.10714E + 01	0.16E + 01	1.5737	0.95217
Lognormal	200	−0.20146E + 02	0.21955E + 02	0.53507E + 01	—	1.8405	1.1773
Beta	150	0.28177E + 00	0.36410E + 01	0.92656E + 00	0.16E + 01	1.6170	0.93436
Beta	100	0.42671E + 00	0.29290E + 01	0.57095E + 00	0.8E + 00	1.6465	0.93774
Beta	75	0.61713E + 00	0.27104E + 01	0.44209E + 00	0.8E + 00	1.5818	0.86667
Beta	50	−0.22230E + 01	0.65131E + 01	0.47614E + 01	0.32E + 01	1.6722	1.0667
Beta	25	−0.25382E + 00	0.54295E + 01	0.17489E + 01	0.32E + 01	1.6649	1.0641

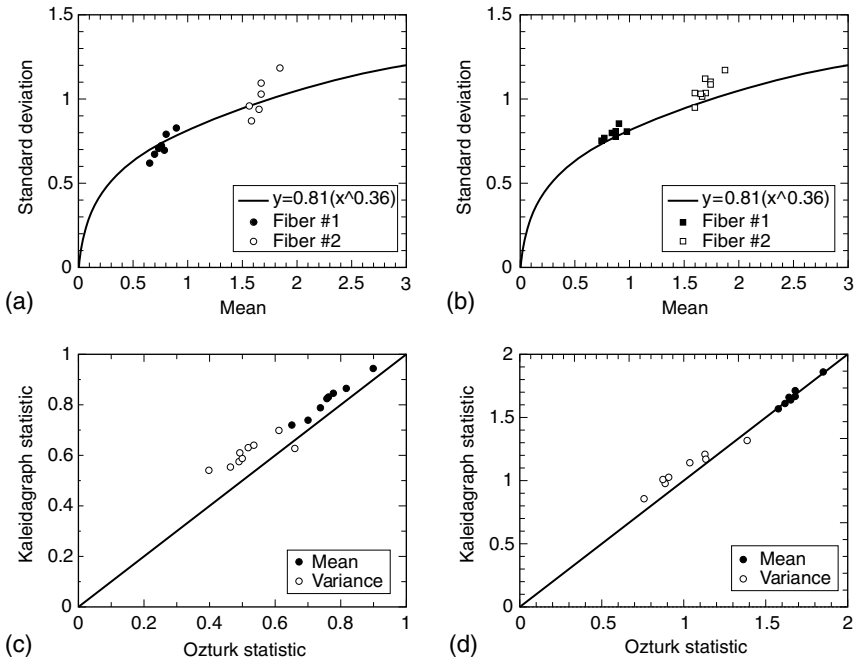

FIGURE 6.11 Comparison of the Ozturk statistics and the sample statistics generated by Kaleidagraph application. (a) Ozturk statistics vs. curve fit determined by Evan Relkin. (b) Kaleidagraph statistics vs. curve fit determined by Evan Relkin. (c) Ozturk statistics vs. Kaleidagraph statistics for fiber #1. The line in this graph passes through the origin and has a slope of one. (d) Ozturk statistics vs. Kaleidagraph statistics for fiber #2. The line in this graph passes through the origin and has a slope of one.

The application, Kaleidagraph, was also used to calculate sample means and standard deviations for all of the data inputted into the Ozturk Algorithm. These data are shown in Figure 6.11(b). It is interesting to note that unlike the statistics generated from the Ozturk Algorithm, all of the sample statistics fall on one side of the curve (with the exception of one data point from fiber #2). In fact, when calculating the deviation of the Ozturk statistics and Kaleidagraph statistics from Evan Relkin's curve fit, it was found that the mean-square deviation of the Ozturk statistics was 9% less for fiber #1 and 18% less for fiber #2 than the deviation for the corresponding Kaleidagraph statistics.

Another way to compare the two methods for calculating statistics is to plot the Ozturk mean and variance versus the sample mean and variance for each set of data.[†] Figure 6.11(c) and (d) shows this comparison. The line drawn in these

[†] The variance has been plotted instead of the standard deviation since the mean and standard deviation for fiber #1 are similar, and thus, the data points would overlay in this plot.

graphs passes through the origin and has unity slope. If the statistics from both methods matched perfectly, one would predict that the data would fall on this line. As can be seen from these graphs, for fiber #1, the Kaleidagraph statistics for both mean and variance are consistently greater than those of the Ozturk statistics; for fiber #2, although the sample mean and Ozturk mean agree well, the sample variance is consistently greater than the Ozturk variance. Therefore, by using the Ozturk Algorithm to determine the distribution of best-fit and calculating the mean and variance of this distribution, one obtains a more accurate prediction of these statistics than by using the sample statistics. In addition, it is important to keep in mind that the Ozturk Algorithm can generate an excellent estimate of the pulse-number distribution's mean and variance with as few as 25 samples.

6.1.4. Analysis of Efferent Optic-Nerve Activity in the Horseshoe Crab

The second application of the Ozturk Algorithm was analyzing efferent optic nerve activity in the horseshoe crab. Numerous studies have been carried out to demonstrate the effects of the efferent activity on visual sensitivity. Some of these studies include Barlow et al.[9], Kass and Barlow[10], Chamberlain and Barlow[11,12], Barlow et al.[13], and Kier and Chamberlain.[14] Two additional studies provide a detailed description of the organization of the efferent activity (Barlow[15] and Kass and Barlow[16]).

Some of the conclusions from these two studies include:

1. Each lateral optic nerve contains a small but separate group of efferent fibers (approximately 10 to 20).
2. Efferent fibers fire in bursts at rates up to two bursts per second.
3. Each efferent fiber fires one spike in a burst.
4. Bursts occur synchronously in the lateral optic nerves.
5. Efferent activity undergoes a circadian rhythm.
6. Coupling among efferent neurons changes during the circadian cycle.

Although several effects of the efferent activity on visual sensitivity are known and some general properties about the organization of this activity have been made, a detailed study of the structure of this information was lacking. For example, Barlow[15] comments, "The repetitive bursts of impulses indicate that the efferent cell bodies in the brain are either coupled together or receive nearly synchronous inputs from the circadian clock. The bursts become less distinct in the early morning hours, suggesting that the synchrony changes during the circadian cycle." In other words, Barlow observed that the efferent activity occurred in bursts, with each fiber firing once per burst, and that these bursts slowed down at dusk (onset of activity) and dawn (offset of activity) and were most rapid in the middle of the night.

To further explore the organization of the efferent activity, Chris Passaglia and I performed an additional study (Passaglia et al.[17]). In this study, we found a

patterned structure to the activity; bursts of spikes occur in clusters, and these clusters of bursts occur in packets. In general, the time between bursts is less than ten seconds, the time between clusters is around ten seconds, and the time between packets is on the order of minutes. Figure 6.12 illustrates these concepts by showing several minutes of efferent activity throughout the night. Each "spike" in these traces correspond to a burst of action potentials. We also observed in our study that the efferent activity was highest in the middle of the night, and that the time between bursts, clusters, and packets all increased at the onset and offset of activity. In other words, the efferent activity started off slowly (long time between bursts, clusters, and packets) increased to its fastest rate during the middle of the night, and then gradually slowed down again toward morning. During the day there is no efferent activity. Figure 6.12 also illustrates this phenomenon. To help learn about the structure of the efferent activity, the Ozturk Algorithm was used to analyze the best-fit distributions for the interburst intervals.

In Figure 6.12, there was no activity recorded before 6 p.m. or after 6 a.m. in this animal, and the activity was most rapid between 11 p.m. and 1 a.m. Each spike in these traces corresponds to a burst of action potentials, and each efferent fiber fires one action potential per burst. The upper-left inset shows an expanded time scale where approximately seven action potentials can be resolved.

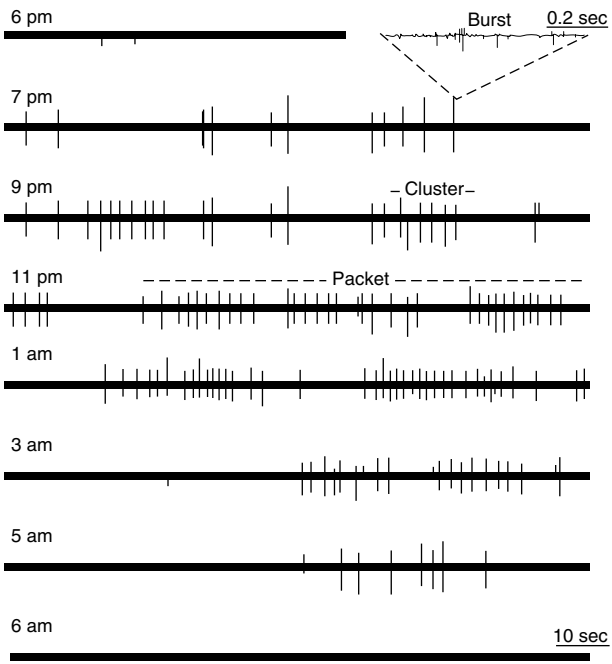

FIGURE 6.12 Eight 3-min traces of horseshoe crab efferent optic nerve activity recorded throughout the night.

The number of action potentials occurring in each burst stays constant throughout an experiment.

6.1.4.1. Characterization of Interburst Intervals

Given the observation that the rate of the efferent activity is modulated throughout the night, an interesting question that arises is whether a single distribution can provide a good fit to the data regardless of the time of the activity. In other words, as the efferent activity is modulated, does the type of distribution which best fits the data change, or rather does the best-fit distribution remain the same and the parameters of this distribution change? Since the Ozturk Algorithm assumes independent samples, only interburst intervals, and not the time between clusters and packets, were analyzed. In addition, data samples were only analyzed over a maximum duration of a one hour interval, the assumption being that the efferent spike trains remain relatively stationary over this interval, and hence, the assumption of independent trials remains valid. Qualitative inspection of the efferent spike trains over one hour intervals supports this assumption. The efferent spike trains that were recorded from these intervals included the time of occurrence of all of the spikes. The data were transformed into interspike intervals by calculating the time between successive spikes. Intervals corresponding to times between clusters and packets were subjectively identified and deleted. Three experiments, done on three different animals, were analyzed. For each of these experiments, the interburst data for successive one hour intervals were analyzed.

For all three experiments, the Ozturk Algorithm was used to determine the best-fit distribution of the interburst intervals for each one hour interval. For each experiment, the coordinates of the mapped end points from the one hour intervals were averaged, and this averaged end point was inputted into the Ozturk Algorithm to determine the best-fit distribution for the averaged data. Interestingly, in all three experiments, the best-fit distribution for the averaged data was the Lognormal distribution. Even more remarkably, the shape parameter value was almost identical in each experiment. Figure 6.13(a) shows the location of the averaged end point in the U-V plane for each experiment. Figure 6.13(b) shows the corresponding plots of the best-fit Lognormal distribution. Notice how similar the three distributions are despite the fact that these experiments were done on three different animals. In Figure 6.13, the parameter values for these distributions are given in Table 6.6.

As mentioned earlier, the rate of efferent activity is modulated throughout the night. Thus, best-fit Lognormal distributions were also obtained for each one hour interval. The best-fit Lognormal distribution of the averaged data was used as the null distribution when performing goodness-of-fit tests. Table 6.6 provides a summary of the number of samples used and the parameter values obtained for the best-fit Lognormal distribution for every condition, which was analyzed by the Ozturk Algorithm. Notice how similar the shape parameter value is for the averaged data in all three experiments.

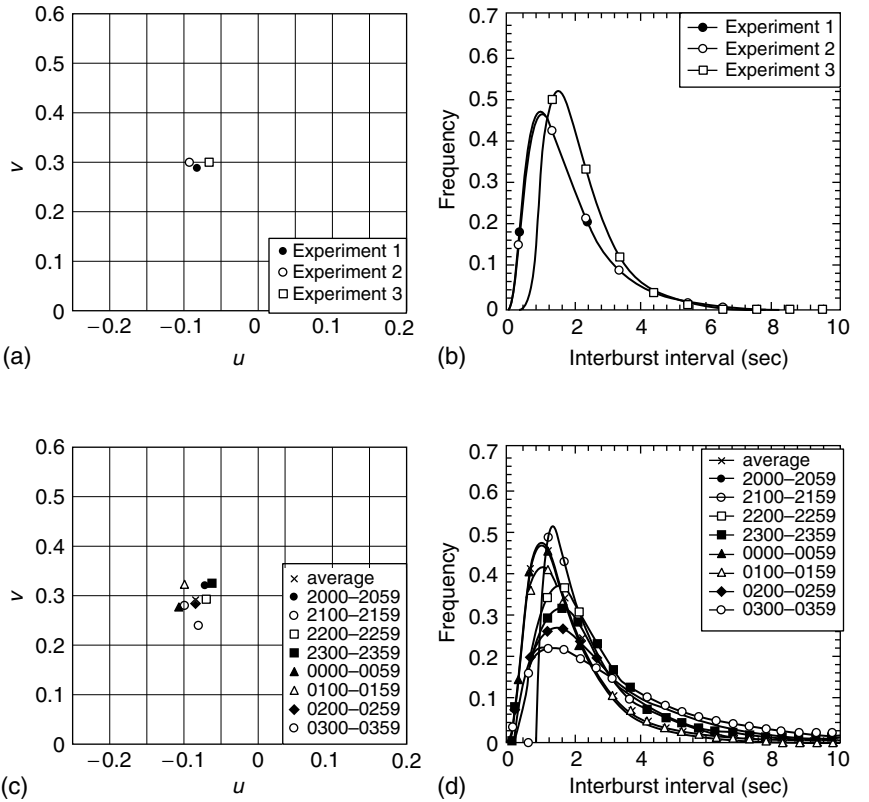

FIGURE 6.13 Mapped end points and corresponding best-fit lognormal distributions for interburst intervals of efferent optic nerve activity in the horseshoe crab. (a) Mapped end points averaged over all hourly intervals for each of three experiments. (b) Lognormal distributions averaged over all hourly intervals for each of three experiments. (c) Mapped end points for each one-hour interval analyzed in experiment 1. (d) Lognormal distributions for each one hour interval analyzed in experiment 1.

The results in Table 6.6 indicate that efferent activity was recorded from 8 p.m. to 4 a.m. in experiment 1, from midnight to 6 a.m. in experiment 2, and from midnight to 7 a.m. in experiment 3. The reason that this activity was not monitored throughout an entire night is due to the complexity of the recording. Remembering that the efferent activity does not begin until dusk, it is futile to begin the surgery until after the clock has turned on the activity. In addition, it can take several hours until efferent activity is isolated — especially given the ratio of afferents to efferents within the optic nerve. Thus, the only way to obtain a recording for an entire night is if the recording is maintained in excess of 24 h. That is, the recording of the entire night is obtained during the *second* night of activity since the onset of activity will be recorded for the second night.

TABLE 6.6
Parameters of Best-Fit Lognormal Distributions, for Each One-Hour Intervals, and for the Average of All One-Hour Intervals, in Three Different Experiments of Efferent Optic Nerve Activity in the Horseshoe Crab

	Time Interval	Number of Samples	Location Parameter	Scale Parameter	Shape Parameter
Experiment 1	Average	277	−0.29873E − 01	0.16047E + 01	0.65588E + 00
	2000–2059	98	−0.17282E + 01	0.40188E + 01	0.50696E + 00
	2100–2159	257	0.88015E + 00	0.12853E + 01	0.10285E + 01
	2200–2259	297	−0.32991E − 01	0.21864E + 01	0.58224E + 00
	2300–2359	381	−0.12828E + 01	0.33888E + 01	0.40119E + 00
	0000–0059	422	0.23578E + 00	0.14394E + 01	0.81168E + 00
	0100–0159	372	−0.27404E + 00	0.19148E + 01	0.6E − 00
	0200–0259	245	−0.26715E + 00	0.27248E + 01	0.67234E + 00
	0300–0359	145	−0.99534E − 01	0.30335E + 01	0.82254E + 00
Experiment 2	Average	480	−0.22409E − 01	0.16548E + 01	0.64258E + 00
	0000–0059	901	0.20427E + 00	0.64307E + 00	0.68058E + 00
	0100–0159	619	0.31337E + 01	0.10708E + 01	0.57936E + 00
	0200–0259	571	0.17248E + 00	0.87607E + 00	0.69301E + 00
	0300–0359	443	0.93756E − 01	0.11927E + 01	0.67886E + 00
	0400–0459	206	−0.98192E + 00	0.29723E + 01	0.63250E + 00
	0500–0559	141	−0.27794E + 00	0.25508E + 01	0.67267E + 00
Experiment 3	Average	741	0.30030E + 00	0.16864E + 01	0.52238E + 00
	0000–0059	1114	0.62426E + 00	0.13595E + 01	0.63071E + 00
	0100–0159	1047	0.59993E + 00	0.15211E + 01	0.55657E + 00
	0200–0259	946	0.80420E + 00	0.15346E + 01	0.68594E + 00
	0300–0359	877	−0.19083E + 01	0.44820E + 01	0.30436E + 00
	0400–0459	659	−0.12900E + 01	0.41819E + 01	0.42255E + 00
	0500–0559	302	0.73848E + 00	0.38726E + 01	0.65947E + 00
	0600–0659	243	−0.38711E + 01	0.70872E + 01	0.44340E + 00

Keeping in mind that the amount of efferent activity begins slowly, increases to a maximum in the middle of the night, and then slows down again, inspection of the number of samples listed in Table 6.6 indicates that only the last half of the efferent activity was recorded for experiments 2 and 3. Nevertheless, there are still enough one hour intervals included in these experiments to study the relationship between the modulation of efferent activity and the best-fit distribution of this activity.

Figure 6.13(c) and (d) show the location of the mapped end points in the U-V plane and the corresponding best-fit Lognormal distributions for each one hour interval from experiment 1. Figure 6.14 shows these same plots for experiments 2 and 3. The average data for each experiment is represented by an "x" in the U-V plane and a bolded line in the Lognormal plots. As mentioned before, the average

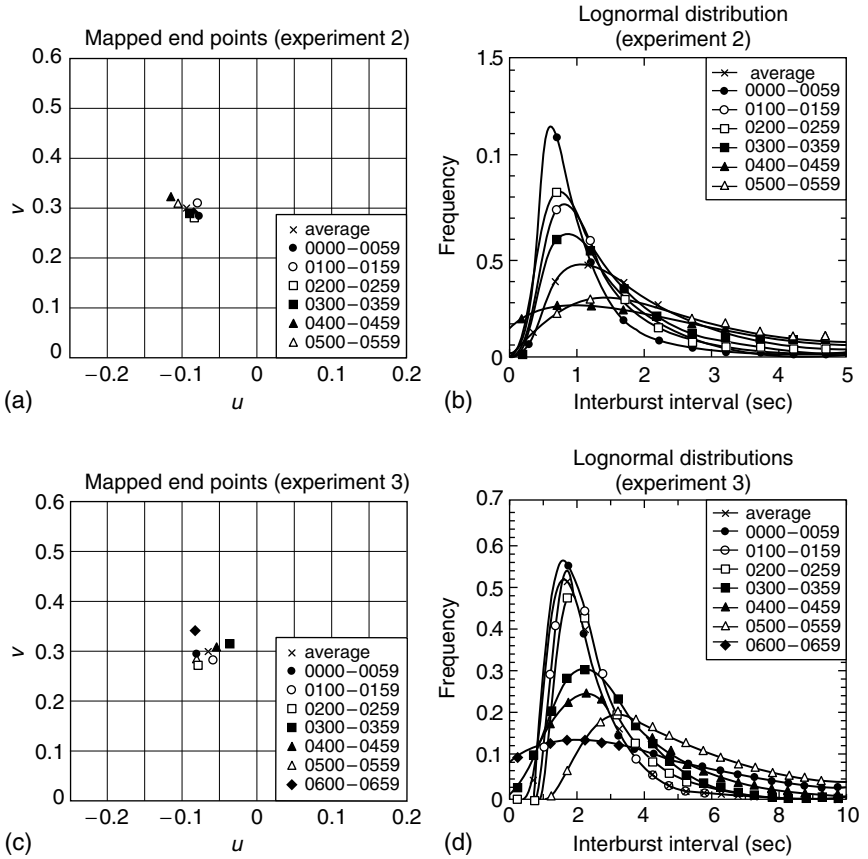

FIGURE 6.14 Mapped end points and corresponding best-fit lognormal distributions for each hourly interburst intervals of efferent optic nerve activity in the horseshoe crab. (a) Mapped end points for experiment 2. (b) Lognormal distributions for experiment 2 (Scales of axes differ for each hourly interval.) (c) Mapped end points for experiment 3. (d) Lognormal distributions for experiment 3. (Scales of axes differ for each hourly interval.)

data were generated by averaging the location of all of the one hour mapped end points in the U-V plane and averaging the number of samples used in each one hour interval. These results were inputted into the Ozturk Algorithm, and the location of the mapped end point and parameter values for the best-fit Lognormal distribution were determined.

Goodness-of-fit: tests were performed on each one hour interval using the average data of the appropriate experiment as the null distribution. In *every* case, the sample data for the one hour intervals were statistically consistent with the null hypothesis at a 99% confidence level (results not shown). This fact indicates that a single Lognormal distribution can sufficiently characterize the interburst

intervals of the efferent activity regardless of the time of night of the activity. Coupled with the fact that the best-fit Lognormal distribution was approximately the same for all three experiments, it appears that interburst intervals of efferent activity can be described with a single Lognormal density function independent of the time of activity or the animal from which the activity was recorded. Figures 6.13 and 6.14 help to illustrate these points. The location of the mapped end points for each one hour interval contain some variability; however, this variability is centered around the average data. Similarly, although the overall amount of activity is modulated throughout the night as indicated in the Lognormal plots, the fact that all of the end points from each one hour interval tend to cluster in one location of the U-V plane indicates that a single distribution function with a relatively constant shape parameter still provides the best fit to these interburst intervals. It is also interesting to note that the variability in the location of the mapped end points is greater between one hour intervals from the same animal than between the location of the mapped end points for the average data from different animals [Figures 6.13(a,c) and 6.14(a,c)]. The fact that the plots of the best-fit density functions for each one hour interval change throughout the night despite the relative constancy of the shape parameter indicates that the location and scale parameters are what encode the modulation of the efferent activity, and not the type of density function. These graphs demonstrate that as the rate of the efferent activity slows down, the best-fit density functions become flatter and more spread out. This flattening can be attributed to a decoupling in the organization of the efferent activity (Figure 6.12). That is, just as the action potentials of each fiber within a burst tend to decouple at the onset and offset of activity, it also appears that the general organization of the bursts, clusters, and packets also reflects this decoupling; the time between bursts, clusters, and packets all become more variable. In Figure 6.14 the parameter values for the Lognormal distributions are given in Table 6.6.

6.1.4.2. Trends in the Shape Parameter

In the previous section, it was shown that a Lognormal distribution with a single shape parameter adequately describes the interburst intervals of the efferent optic nerve activity. In addition, although there is some variability in the value of this shape parameter, all of the interburst intervals can fit into a single Lognormal density function at a 99% confidence level.

One question that still arises, however, is whether any trends exist in the value of the shape parameter. Put another way, does any correlation exist between the value of the shape parameter and the time of night of the efferent activity? To help address this question, the value of the shape parameter was plotted for successive one hour intervals for all three experiments. These results are summarized in Figure 6.15. In each case, a linear curve fit has been used on the data. In all three cases, the slope of this curve fit is close to zero (experiment $1 = 0.011$; experiment $2 = 0.0030$; and experiment $3 = -0.22$). The standard error has been plotted as error bars in all three graphs. These results indicate a

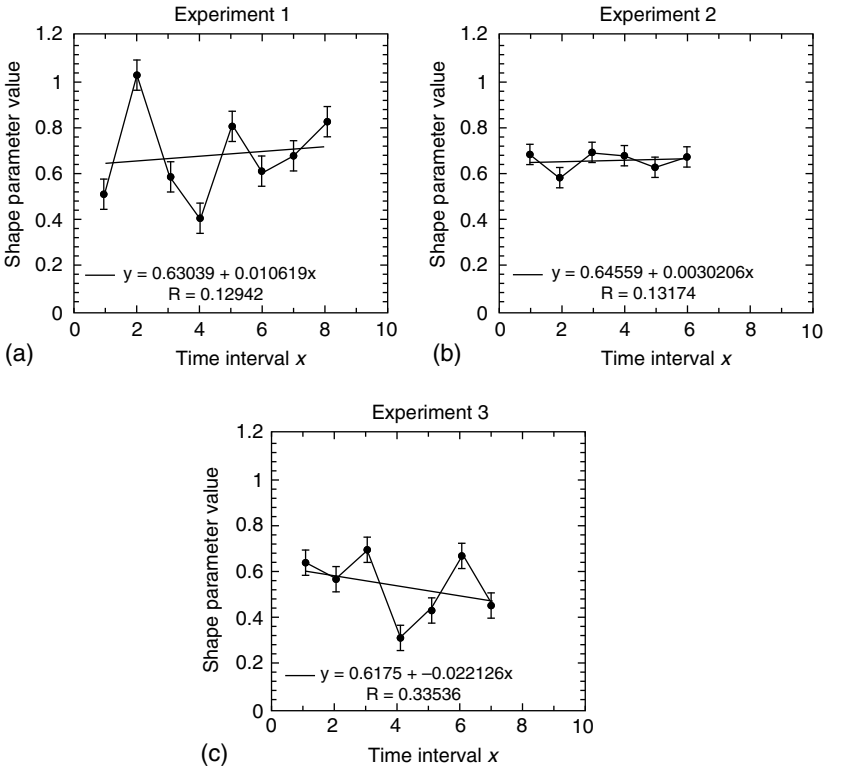

FIGURE 6.15 Trends in shape parameter values for successive hourly time intervals in horseshoe crab experiments. Standard error bars with best linear curve fit of correlation coefficient R are shown. (a) Experiment 1. (b) Experiment 2. (c) Experiment 3.

lack of any trend in the shape parameter as a function of efferent modulation: in experiment 1 the shape parameter tends to increase slightly towards morning, while in experiment 2 the shape parameter remains relatively constant, and in experiment 3 the shape parameter tends to decrease slightly toward morning. Therefore, although the shape parameter does contain some degree of variability, no apparent trends in this variability exist. In Figure 6.15, the successive time intervals and shape parameter values are listed in Table 6.6. Lastly, it is important to consider the significance of these results. For one thing, these results indicate that a Lognormal distribution with a shape parameter around 0.61 should be used to model the interburst intervals of the efferent activity; the value of the location and scale parameters are what reflect the modulation of this activity. In addition, one could perform a similar analysis on the time between clusters and packets. It would be very interesting to see whether these best-fit distributions would also be Lognormal density functions with similar shape parameter values. Thirdly, the fact that a single density function appears to characterize the interburst intervals

despite the modulation of this activity suggests that a single process is involved in generating this activity, and that the modulation of this process leads to the pattern of efferent activity occurring throughout the night. Finally, it is interesting to consider the significance of the Lognormal distribution providing the best fit for the efferent activity. Is it coincidental that this well described density function is the one that happens to best fit the interburst intervals? One particularly noteworthy observation is the relationship between the Lognormal distribution and the central limit theorem (Papoulis[18]). The central limit theorem states that given n independent positive random variables, as n becomes large, the distribution of the product of these random variables will be Lognormal. Whether any relationship exists between this mathematical statement and the physiology remains to be determined. Nevertheless, it is interesting to postulate whether the clock of the horseshoe crab generates efferent activity through a multiplicative operation involving a large number of random variables. Regardless of the answer to this question, the usefulness of the Ozturk Algorithm for analyzing data is apparent from this example.

6.1.5. Analysis of the Visual Field of the Horseshoe Crab

The last application of the Ozturk Algorithm was analyzing the visual field of horseshoe crabs. In a recent study of mine (Weiner and Chamberlain[19]), I measured the extent and resolution of the visual field of these animals. I found that two different eye shapes exist and have named the animals possessing these two eye shapes as "morlocks" and "eloi" (after H. G. Wells' two varieties of humans in his book *The Time Machine*[24]). Morlocks have a relatively smaller and flatter eye with maximum resolution in the anteroventral quadrant of their visual field. In contrast, the lateral eye of eloi is relatively larger and bulgier; these animals have much more uniform resolution in their visual field. Figures 6.16 and 6.17 summarize these findings for morlocks and eloi, respectively. In these figures, contour plots show the distribution of interommatidial angles — the angle formed by the intersection of the optic axes of adjacent ommatidia — across the eye in both the horizontal and vertical direction.

In Figures 6.16(a,b) and 6.17(a,b), the contour lines are labelled in degrees. In Figures 6.16(c) and 6.17(c), the lengths of the vertical and horizontal axes of the ellipse correspond to the magnitude of the respective interommatidial angles. The size of each ellipse provides information about the resolution of the eye. The eccentricity of each ellipse provides information about the relationship between vertical and horizontal interommatidial angles. The "acute zone" in the anteroventral portion of the morlock eyes is clearly visible as the region of small ellipses. In all three maps, anterior is to the left and dorsal is toward the top. The length of the eye along the horizontal axis is slightly exaggerated. Figures 6.16 and 6.17 are based on 82 samples from Weiner and Chamberlain.[19]

In this section of the chapter, the Ozturk Algorithm was used to statistically investigate the differences in distribution of interommatidial angles in the eyes of

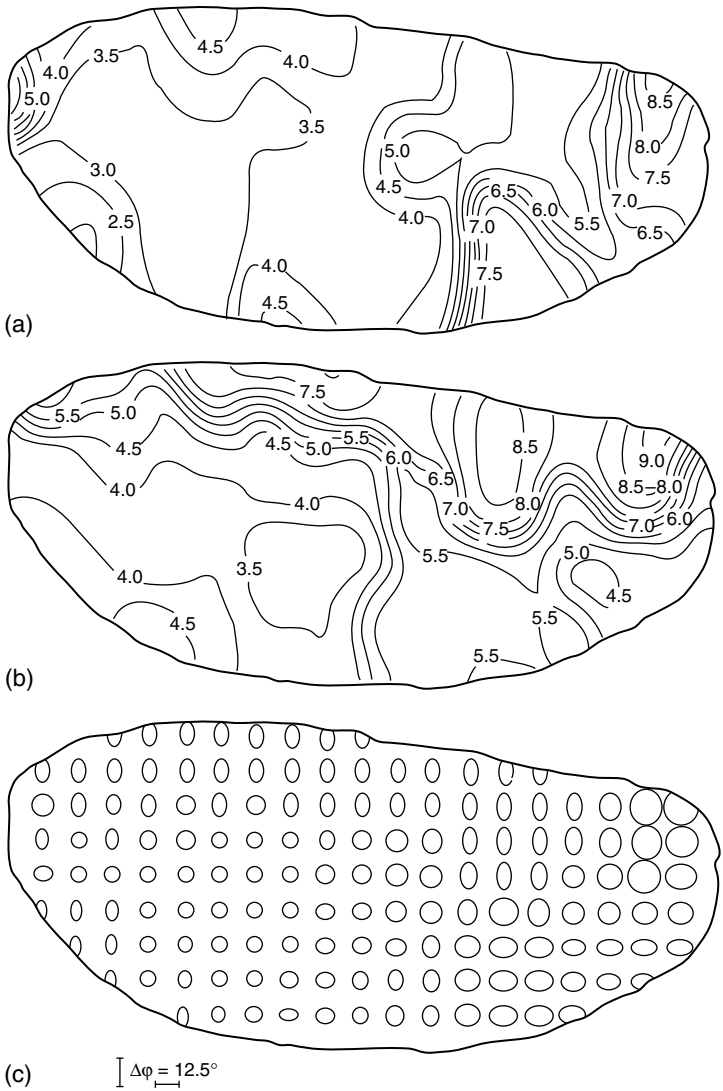

FIGURE 6.16 Distribution of interommatidial angles across the morlock eye of a horseshoe crab. (a) Contour map in the horizontal direction. (b) Contour map in the vertical direction. (c) Elliptical plots.

morlocks and eloi. The random variable used in this analysis was the interommatidial angle. It is important to remember that the Ozturk Algorithm assumes that the random variable represents independent samples. For this particular case, this assumption is not valid since the value of the interommatidial angle is a function of the location in the eye. To help reduce this error,

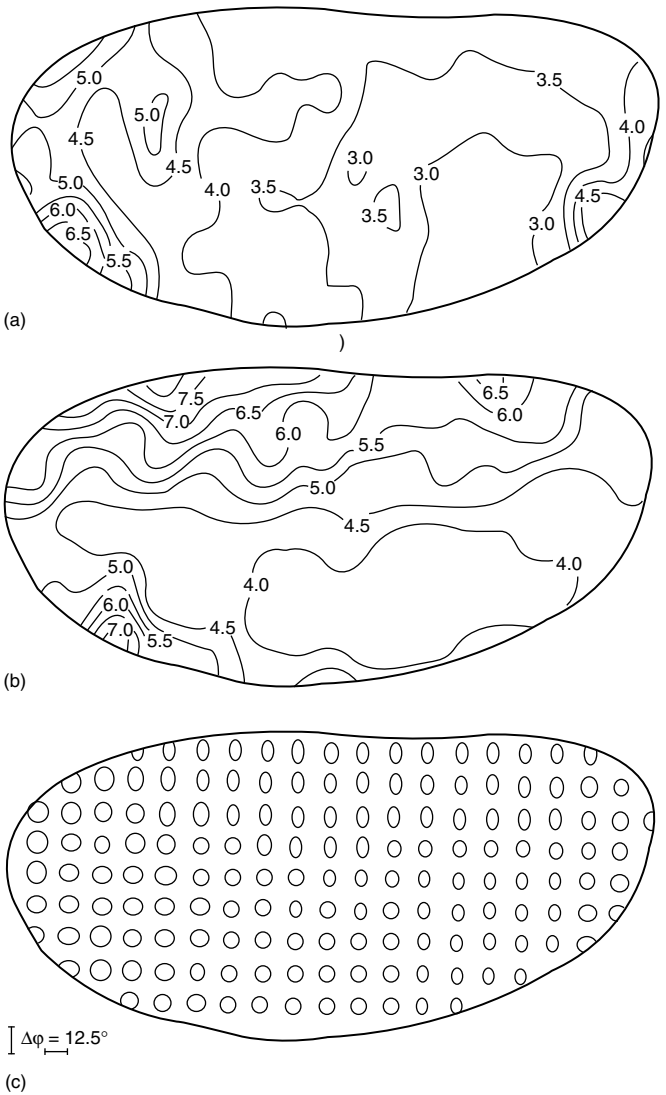

FIGURE 6.17 Distribution of interommatidial angles across the eloi eye of a horseshoe crab. (a) Contour map in the horizontal direction. (b) Contour map in the vertical direction. (c) Elliptical plots.

interommatidial angles were taken from samples uniformly distributed throughout the eye. By so doing, every portion of the eye was sampled equally. In addition, in my previous study it was found that the distribution of interommatidial angles is approximately the same for all animals of a particular eye shape. Therefore, in this analysis only the eyes from two animals were

analyzed: one morlock and one eloi. The two animals selected had approximately the same interocular distance.

Throughout this section, three interommatidial angles will be referred to: horizontal, vertical, and total. The horizontal and vertical angles refer to the interommatidial angle in the anterior-posterior and dorsal–ventral directions, respectively. The total angle refers to the overall interommatidial angle found by combining the horizontal and vertical angles into a single solid angle.

6.1.5.1. Total Interommatidial Angles

In the first portion of this study, the distribution of total interommatidial angles was analyzed using 145 samples. The best-fit distribution for a morlock eye and an eloi eye was determined. Figure 6.18(c,d) shows a histogram of these interommatidial angles with the best-fit density function plotted on the same graph (c): morlock, (d): eloi. These histograms indicate that the total interommatidial angles result in a much more uniform visual field in eloi than

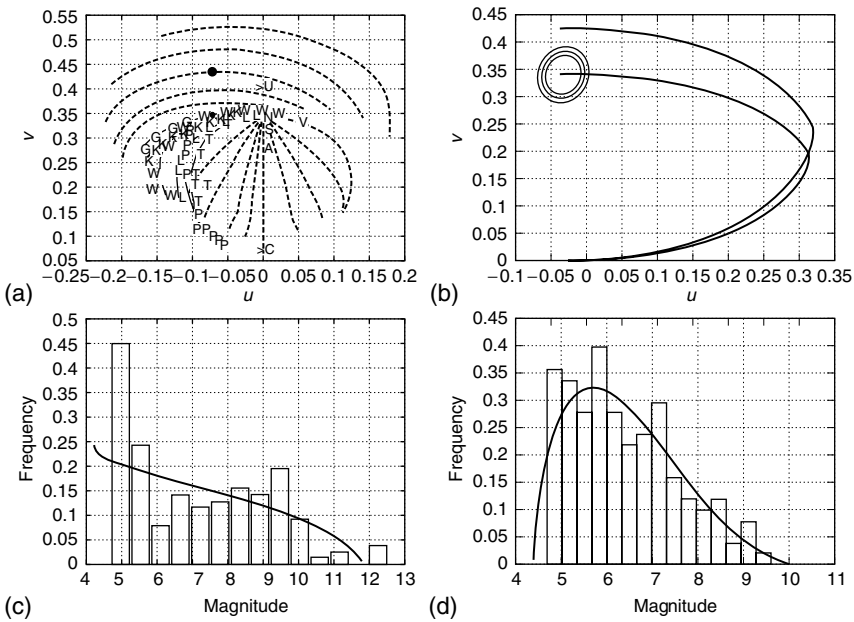

FIGURE 6.18 Statistical analysis of the distribution of total interommatidial angles in morlock and eloi lateral eyes of a horseshoe crab. (a) Distribution approximation chart for a morlock eye. (b) Goodness-of-fit test for a morlock eye. Null hypothesis is best-fit distribution function for an eloi eye. (c) Histogram and best-fit distribution function of total interommatidial angles (degrees) for a morlock eye. (d) Histogram and best-fit distribution function of total interommatidial angles (degrees) for an eloi eye.

in morlock. After all, if the visual field had completely uniform resolution, the interommatidial angles would be the same everywhere and a histogram plot of interommatidial angles would show a single bar at this value. From Figure 6.18(d), it can be seen that most of the interommatidial angles in eloi occur in the range from 4.75 to 7.5 degrees. In contrast, Figure 6.18(c) reveals that morlock eyes have interommatidial angles almost uniformly distributed between 4.75 and 10.25 degrees. Occurrence of a peak in this histogram plot around 5 degrees reflects the presence of an area in the eye where maximum resolution exists. The contour plots of Figure 6.16 indicate that this area corresponds to the antero-ventral quadrant of the visual field. The fact that morlock eyes have a more uniform distribution of total interommatidial angles over a wider range of values than eloi eyes is consistent with the finding that eloi have more uniform resolution in their visual field.

Another way of showing that eloi eyes have a more uniform visual field than morlock eyes is by comparing the variance of the total interommatidial angles. Using the parameters of the best-fit density function, the variance can easily be calculated from the formulas listed in Table 6.3. The best-fit density function and parameter values in Figure 6.18 for each eye are given below. From these values, the variance in total interommatidial angles is 3.65 degrees for the morlock eye and 1.33 degrees for the eloi eye.

In Figure 6.18(a), the location of the mapped sample data in the U-V plane is indicated by a filled circle. In this figure: N = Normal, U = Uniform, E = Exponential, A = Laplace, L = Logistic, C = Cauchy, V = Extreme Value (type-1), T = Gumbel (type-2), G = Gamma, P = Pareto, W = Weibull, L = Lognormal, K = K-Distribution. The upper five dashed lines represent Beta and the lower nine dashed lines represent SU Johnson. In Figure 6.18(b), the ellipses correspond to the confidence ellipses for the goodness-of-fit test where the best-fit density function for the eloi eye was used as the null hypothesis distribution. The total interommatidial angles from the morlock eye were used as the sample data. In Figure 6.18(c), the data were generated from 145 samples. The best-fit distribution is a Beta: $\alpha = 4.6$, $\beta = 6.0$, $\gamma = 0.58$, $\delta = 0.8$. In Figure 6.18(d), the data were generated from 145 samples. The best-fit distribution is Beta: $\alpha = 4.4$, $\beta = 5.9$, $\gamma = 1.6$, $\delta = 3.2$.

The Ozturk Algorithm was also used to determine whether the distribution of total interommatidial angles were statistically consistent for morlock and eloi eyes. In this analysis, a goodness-of-fit test was performed using the distribution of total interommatidial angles for a morlock eye as the sample data and the best-fit density function for the corresponding angles of an eloi eye as the null distribution. The results of this test are shown in the top-right graph of Figure 6.18(b). This test clearly indicates that the two sets of data are statistically inconsistent at a confidence level greater than 99%. This figure also indicates that the mapped trajectory in the U-V plane for the two data sets are very different. Thus, the Ozturk Algorithm provides statistical verification that the visual fields of morlocks and eloi are different.

6.1.5.2. Horizontal and Vertical Interommatidial Angles

In the next portion of this study, the Ozturk Algorithm was used to analyze the statistical distribution of interommatidial angles in the horizontal and vertical directions. The data from the same two eyes, which were used above were also used in this study. Histograms of horizontal and vertical interommatidial angles are plotted for both eyes in Figure 6.19. Inspection of these histograms reveals that the horizontal and vertical interommatidial angles are between 2 and 8.25 in every case except for the vertical direction of morlocks. This fact suggests that it is the wide range of vertical interommatidial angles in morlocks, which contributes to the larger amount of variability in the total interommatidial angles. As mentioned earlier, morlock eyes tend to be flatter and slightly smaller. In fact, an eloi animal with the same interocular distance as a morlock animal will have about 15% more ommatidia in its lateral eye (Weiner and Chamberlain[19]), and most of these additional ommatidia are located in the dorsal–ventral dimension (i.e., the two eyes have approximately the same length in the anterior–posterior

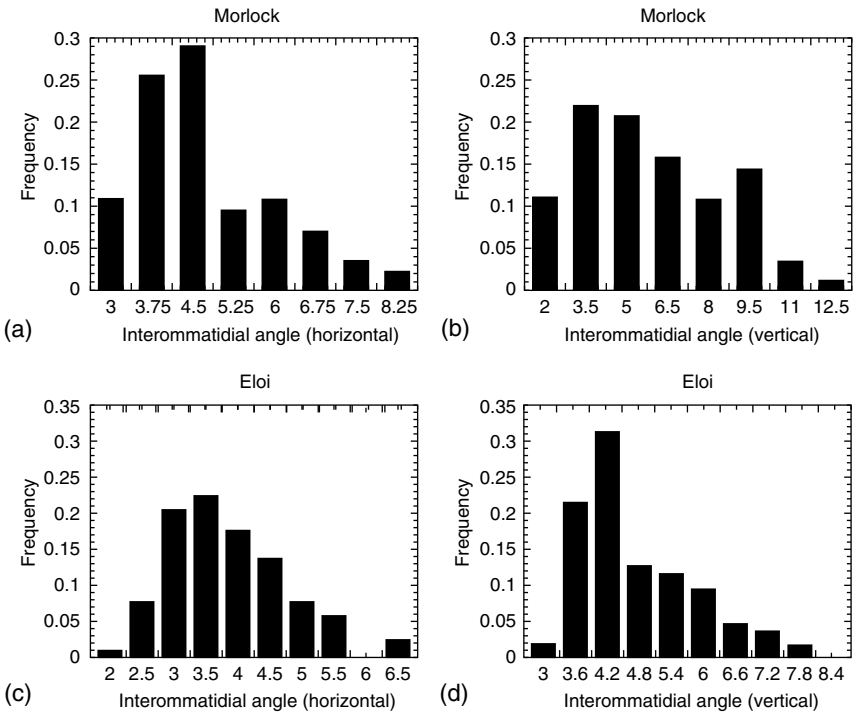

FIGURE 6.19 Histograms of horizontal and vertical interommatidial angles (degrees) for morlock and eloi eyes of a horseshoe crab. (a) Horizontal angles in a morlock eye. (b) Vertical angles in a morlock eye. (c) Horizontal angles in an eloi eye. (d) Vertical angles in an eloi eye.

direction). Therefore, eloi have bulgier eyes with more ommatidia, and it appears that these additional ommatidia improve the resolution of the eye in the dorsal–ventral direction, as compared to morlocks. In contrast, morlocks have poorer resolution in the dorsal–ventral direction; however, they have an area of maximal resolution in the anteroventral quadrant of their eye. Hence, a trade-off between uniform resolution and an acute zone exists.

The Ozturk algorithm was used to determine the best-fit density function for the distribution of horizontal and vertical interommatidial angles in the morlock and eloi eyes; 82 samples were used for the morlock eye while 102 samples were used for the eloi eye. Then, the algorithm was used to perform two goodness-of-fit tests: one for the horizontal angles, and one for the vertical angles. In each test, the interommatidial angles from the morlock eye were used as the sample data while the best-fit density function for the eloi eye was used as the null distribution. The graphical solutions from these tests are shown in Figure 6.20(a) for horizontal angles and Figure 6.20(b) for vertical angles. These tests indicate that the distribution of horizontal and vertical interommatidial angles in morlocks and eloi are statistically inconsistent at a confidence level greater than 99%. The goodness-of-fit tests provide statistical support to the claim that two different eye shapes exist in horseshoe crabs. In my 1994 study, I reached the same conclusion; however, the Ozturk Algorithm provides a statistical tool for quantifying this conclusion.

Finally, one might legitimately point out that the assumption of independent samples, as required by the Ozturk Algorithm, was invalid in this problem. To help alleviate this error, samples were taken, which were uniformly distributed throughout an eye. Nevertheless, the value of these samples is still dependent on where in the eye these samples were located. The significance of this last point comes from the interpretation of the goodness-of-fit test. Should the data from the two eyes have been statistically consistent, the only valid conclusion could have been that the *distribution* of the interommatidial angles in the two eyes is

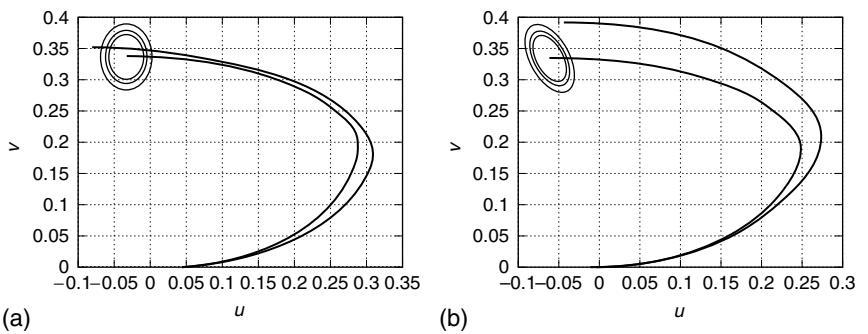

FIGURE 6.20 Goodness-of-fit tests for the distribution function of intercommatidial angles for the morlock eye of a horseshoe crab. (a) Horizontal angles from sample data of morlock eye. Null hypothesis is best-fit distribution from an eloi eye. (b) Vertical angles from sample data of morlock eye. Null hypothesis is best-fit distribution from an eloi eye.

statistically consistent, NOT that the two eyes have the same visual field. Conversely, the fact that the horizontal and vertical interommatidial angles were statistically inconsistent in the two eyes, given that samples were uniformly distributed throughout each eye, ensures that the visual field of the two eyes is different. Therefore, even when the assumption of independent samples is not completely valid, the Ozturk Algorithm can still be effectively used to extract statistical information.

In Figure 6.20, the outer ellipse corresponds to a 99% confidence level and the middle ellipse to a 90% confidence level. The best-fit distributions for the morlock eye are for the horizontal angles, Beta: $\alpha = 3.7$, $\beta = 7.0$, $\gamma = 0.80$, $\delta = 3.3$ and for the vertical angles, Beta: $\alpha = 2.7$, $\beta = 10.2$, $\gamma = 0.96$, $\delta = 1.6$. The best-fit distributions for the eloi eye are for the horizontal angles, K-distribution: $\alpha = 2.4$, $\beta = 0.38$, $\gamma = 24.8$ and for the vertical angles, Weibull: $\alpha = 3.4$, $\beta = 1.8$, $\gamma = 1.5$.

6.1.6. Applications of the Ozturk Algorithm in Neuroscience

In the first section of this chapter, a brief description of the Ozturk Algorithm was given and its advantages over classical techniques were discussed. Then, in the next section of this chapter, a detailed description of the algorithm was provided. Finally, in the last three sections of this chapter, three applications of the Ozturk Algorithm were presented. The three applications that were described were selected for a number of reasons. First, these examples demonstrate how the algorithm can be used to provide statistical information for a wide variety of problems. In addition, the examples discussed represent multiple modalities, as well as both anatomical and physiological processes. Finally, the last two examples were selected because I performed most of the experiments necessary to collect these data. In the final section of this chapter, I will point out several other applications of the Ozturk Algorithm in the field of neuroscience.

The Ozturk Algorithm is appropriate for analyzing data whether one is studying anatomy, physiology, or psychophysics. In the case of anatomy, the algorithm can be used to determine the distribution of a measured dimension for a particular structure. As an example, one could use the algorithm to find the best-fit density function for the diameter of a class of axons. One could also use the algorithm to determine whether this best-fit function is statistically consistent with the distribution of a second class of axons. This type of analysis could be performed on any structure, and the advantage of using the Ozturk Algorithm is that very few samples need to be collected and measured to produce reliable results.

Several applications of the Ozturk Algorithm exist for physiology as well. The algorithm can be used to analyze spike trains and to determine whether two such trains are best fit by the same density function. In addition, membrane potentials can be analyzed and the best-fit density function used to generate a cumulative distribution function. By so doing, the probability of a specified threshold being exceeded could be calculated. In addition, since the algorithm requires so few points to accurately fit data, the amount of sweeps necessary to

collect data might be reduced. This provides an enormous advantage in the cases where stable recordings are problematic.

Even in psychophysics, applications of the Ozturk Algorithm exist. Perhaps the algorithm could be used to more efficiently track the thresholds of subjects. After all, once a density function for an experiment is obtained, the cumulative distribution function can be calculated and any arbitrarily defined threshold determined. It would be very interesting to see whether using the Ozturk Algorithm reduced the amount of trials required when tracking a subject's threshold. This would greatly reduce the time required to perform psychophysical experiments. In addition, the Ozturk Algorithm seems especially useful when one is modeling data; the algorithm can easily be implemented to provide the exact equation of a best-fit density function for a particular process or set of data. Once this equation is determined a wealth of information is available to the researcher such as: mean, variance, frequency content and probabilities.

Undoubtedly, many other applications of the Ozturk Algorithm in the field of neuroscience exist. The purpose of this chapter was to provided a detailed description of the algorithm and present some of its potential uses. Hopefully, as exposure to the Ozturk Algorithm increases, many people will begin to use and benefit from this innovative statistical tool.

6.2. USE OF IMAGE PROCESSING TO PARTITION A RADAR SURVEILLANCE VOLUME INTO BACKGROUND NOISE AND CLUTTER PATCHES

(M. A. SLAMANI AND D. D. WEINER)

6.2.1. Introduction

The use of spherically invariant random processes (SIRPs)[1] in the implementation of likelihood ratio tests (LRTs) and locally optimum detectors (LODs)[2] for the radar problem allows us to derive algorithms for performing both strong and weak signal detection in a nonGaussian environment. Classical detection assumes *a priori* knowledge of the joint PDF underlying the received data. In practice, received data can come from a clear region, where background noise alone is present, or from a clutter region, where returns are due to reflections from such objects as ground, sea, buildings, birds,… etc. When a desired target return is from a clear region and the background noise is sufficiently small, the signal-to-noise ratio will be larger and the strong signal detector (i.e., the LRT) should be used. However, if a desired target return is from a clutter region, two situations can exist. When the desired target can be separated from the clutter by means of space–time processing and the background noise is sufficiently small, the signal to noise ratio will be large and a strong signal detector should again be used. When the desired target cannot be separated from the clutter by means of space–time processing and the clutter return is much larger then the desired target

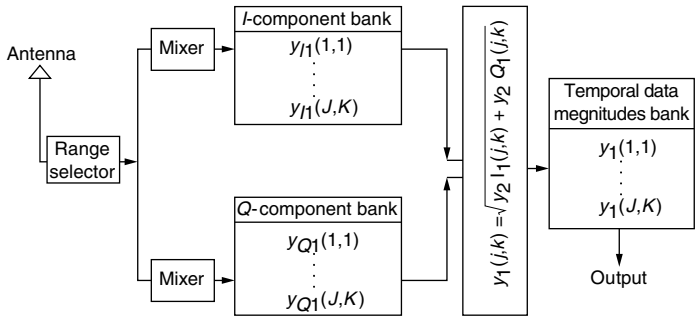

FIGURE 6.21 Block diagram of data preprocessing stage.

return, then a weak signal detector (i.e., LOD) should be used. Use of the LRT (LOD) in a weak (strong) signal situation can result in a severe loss in performance. Hence, it is necessary for the receiver to determine whether a strong or weak signal situation exists.

All of these suggest use of image processing along with an expert system in the radar detection problem for (1) monitoring the environment and (2) selecting the appropriate detector for processing the data. This is in contrast to current practice where a single robust detector, usually based on the Gaussian assumption, is employed. In addition, depending on statistical changes in the environment over time and space, the expert system enables the receiver to adapt so as to achieve close to optimal performance. The goal of this study is to explore how image processing along with an expert system can be used to develop an adaptive radar receiver that is able to outperform traditional radars with respect to high subclutter visibility. The focus of this chapter deals with the partitionning of a radar surveillance volume into background noise and clutter patches. We refer to this as mapping of the surveillance volume.

Assume that $J \times K$ range/azimuth (R/A) cells are scanned by a radar antenna and that the dwell time is equal to the pulse repitition interval (PRI) so that only a single pulse is processed from each cell. The block diagram of the preprocessing stage is shown in Figure 6.21. An average power $P(j,k)$ is formed for every R/A cell in the $J \times K$ R/A plane. In particular,

$$P(j,k) = y_1^2(j,k); \quad j = 1,2,...,J \quad k = 1,2,...,K \tag{6.29}$$

where $y_1(j,k)$ represents the temporal data magnitude of the jkth R/A cell.

At this point, the R/A plane consists of two different types of regions that need to be identified. There are clear regions, where background noise (BN) alone is present, and clutter patches, where both clutter (CL) and additive BN are present.

6.2.2. Observations about BN and CL

The following observations are based on computer generated examples of BN and CL data where the clutter-to-noise ratio (CNR) is assumed to be greater than

0 dB. The BN envelope is assumed to be Rayleigh while the CL envelope may be either Rayleigh, K-distributed, Weibull or Lognormal.[3]

6.2.2.1. Observations about BN

(1) On average, the BN data values are smaller than the CL data values.
(2) Large data values exist in a BN that may be higher than some data values of the CL.
(3) Large data values in the BN tend to be isolated points.
(4) The number of BN data significantly larger than the average is relatively small.
(5) The relatively small number of large BN data is distributed evenly throughout the surveillance volume.

6.2.2.2. Observations about CL

(1) On average, CL data values are higher than BN data values.
(2) A CL region contains additive CL and BN.
(3) Small data values exist in the CL that may be larger than some data values of the BN.
(4) The large CL data values are larger than the largest BN data values assuming positive CNR.
(5) Whereas the BN data values are distributed over the entire surveillance volume, the CL data values are distributed only over the clutter regions.
(6) Small CL data values exist and may be smaller than the large BN data values.
(7) Large data values in the CL tend to be clustered.

6.2.3. Mapping Procedure

Using the fact that clutter patches, on average, have stronger radar returns, the mapping processor begins by setting a threshold that results in a specified fraction of BN cells. Image processing is then used to establish the background noise and clutter patches. If the final image contains a significantly different fraction of BN cells than originally established by the initial threshold, the process is repeated with a new threshold. The mapping processor iterates until it is satisfied that the final scene is consistent with the latest specified threshold. Finally, clutter patch edges are detected using an image processing technique.

The mapping procedure consists of two steps. The first step is the identification of CL patches within BN. The second is the detection of clutter patch edges. These two steps are explained next.

6.2.3.1. Separation of CL Patches from BN

Identification of CL patches within BN is performed by the following steps: thresholding, quantization, correction, and assessment.

6.2.3.1.1. Thresholding and Quantization

Identification of CL patches within BN starts by setting a threshold q that results in a specified fraction of BN cells. Then a quantized volume is formed as follows: all R/A cells with average power less than q are given a value of zero and all R/A cells with average power above q are given a value of one. Let $Q(j,k)$ represent the quantized value of the jkth R/A cell. Then,

$$Q(j,k) = \begin{cases} 1 & \text{if } P(j,k) \geq q \\ 0 & \text{if } P(j,k) < q \end{cases} \qquad j = 1,2,\ldots,J \quad k = 1,2,\ldots,K \tag{6.30}$$

where $P(j,k)$, the average power of the jkth R/A cell, is defined in Equation 6.29.

In general, the quantized version differs from the original. This is due to the fact that even though the average powers of BN cells are expected to fall under the threshold, while the average powers of CL cells are expected to fall above the threshold, on average, some BN cells have an average power that falls above the threshold and some CL cells have an average power that falls under the threshold. Also, the first setting of the threshold, which is somewhat arbitrary, is likely not to be the best for identifying CL patches within BN.

6.2.3.1.2. Correction

Consider a set of 3×3 R/A cells. Let the center cell be referred to as the test cell and the surrounding cells be referred to as the neighboring cells. Assume that a clutter patch cannot be formed by a single cell. In this case, every test cell in the clutter patch has at least one neighboring cell that belongs to the same clutter patch.

A test cell belonging to a clutter patch that has at least one neighboring BN cell is referred to as a CL edge cell. On the other hand, a test cell belonging to a CL patch for which none of the neighboring cells are in the BN is referred to as an inner CL cell.

The proposed correction technique consists of transforming the quantized volume into a "corrected" volume. The transformation consists of the following steps:

1. Choose the necessary number of CL neighboring cells, NCQ, for a test cell in the quantized volume to be declared as a CL cell in the corrected volume. NCQ can take one of the following values: 5, 6, 7, 8.
2. For every test cell in the quantized volume count the number of neighboring CL cells. If the number is greater than or equal to NCQ declare the test cell as a CL cell in the corrected volume. Otherwise, declare the test cell as a BN cell in the corrected volume.

When all the cells of the quantized volume have been tested, a "corrected" volume consisting of declared BN or CL R/A cells is obtained.

Because NCQ is chosen to be relatively large (i.e., NCQ = 5, 6, 7 or 8), BN cells that were incorrectly identified in the quantized volume as CL cells due to

their large power tend to be reclassified as BN cells. Also, inner CL cells in the quantized volume are recognized as CL cells in the "corrected" volume. Meanwhile, most of the CL edge cells in the quantized volume are recognized as BN cells in the "corrected" volume. This results in an overcorrection if most of the CL edge cells are identified as BN. As an example, when NCQ = 8, only inner CL cells in the quantized volume are recognized as CL cells in the "corrected" volume and all CL edge cells in the quantized volume are recognized as BN cells in the "corrected" volume. In order to recover the edge cells, a second correction stage is needed where the first "corrected" volume will be transformed into a second "corrected" volume. Let the first "corrected" volume be referred to as the "corrected-quantized" volume (CQV) and the second "corrected" volume be referred to as the "corrected-corrected" volume (CCV). The following steps are used to transform the CQV into the CCV:

1. Choose the necessary number of CL neighboring cells, NCC, for a test cell in the CQV to be declared as a CL cell in the CCV. NCC can take one of the following values: 1, 2, 3 or 4.
2. For every test cell in the CQV count the number of neighboring CL cells. If the number is greater than or equal to NCC declare the test cell as a CL cell in the CCV. Otherwise declare the test cell as a BN in the CCV.

6.2.3.1.3. Assessment

Let BNQP, BNCQP and BNCCP denote the percentage of BN cells in the quantized, "corrected-quantized" and "corrected-corrected" volumes, respectively. BNQP is prespecified so as to determine the threshold for the quantized volume, whereas BNCQP and BNCCP are computed after the CQV and the CCV are obtained.

The assessment process consists of comparing BNCQP and BNCCP to BNQP in order to determine whether or not the percentages of the BN cells after correction are consistent with the percentage of BN cells in the quantized volume. When there is no consistency, further quantization, correction and assessment are performed until consistency is obtained.

6.2.3.1.4. Smoothing

Examples have shown that when the percentages are consistent, clutter declared patches may contain isolated BN declared cells. Because small BN powers can arise in a CL patch as explained in Section 6.2.2.1, it is most likely that the BN isolated cells in the CL patches are CL cells. The smoothing process is used to detect these isolated cells and label them adequately by transforming the CCV into a smoothed volume (SV). The smoothing technique consists of the following steps:

1. Choose the necessary number of CL neighboring cells NS for a BN identified test cell in the CCV to be declared as a CL cell in the SV where NS can take one of the following values: 5, 6 ,7 or 8.

2. For every BN identified cell in the CCV count the number of neighboring CL cells. If the number is greater than or equal to NS, declare the test cell as a CL cell in the SV. Otherwise declare the test cell as a BN cell in the SV.

6.2.3.2. Detection of Clutter Patch Edges

After smoothing, each cell in the SV has been declared as either a CL or BN cell. The next step is to determine which of the CL cells are located on the edges of the CL patches. This is important for subsequent radar signal processing if reference cells for estimating parameters of a test cell are to be chosen properly.

Identification of CL edge (CLE) cells in done by the use of an image processing technique referred to in the image processing literature as unsharp masking.[4,5] It consists of the following steps:

1. A weighting filter consisting of a 3×3 array of cells is constructed such that the center cell has a weight given by $w(0,0) = 8$ and the neighboring cells have weights given by $w(-1,-1) = w(0,-1) = w(1,-1) = w(-1,0) = w(1,0) = w(1,1) = w(0,1) = w(1,1) = -1$. The center cell is positioned on the test cell. Notice that the weights of the filter cells sum to zero. In particular,

$$\sum_{m=-1}^{1} \sum_{n=-1}^{1} w(m,n) = 0 \tag{6.31}$$

2. Assume the weighting filter is centered at the jkth cell in SV. The cells corresponding to the 3×3 array of the weighting filter have quantized values as illustrated in Figure 6.22. By definition,

$$SQ(j,k) = \begin{cases} 1 & \text{if the } jk\text{th cell in SV is declared as CL} \\ 0 & \text{if the } jk\text{th cell in SV is declared as BN} \end{cases} \tag{6.32}$$

where

$$j = 1, 2, \ldots, J \text{ and } k = 1, 2, \ldots, K$$

$SQ(j-1,k-1)$	$SQ(j,k-1)$	$SQ(j+1,k-1)$
$SQ(j-1,k)$	$SQ(j,k)$	$SQ(j+1,k)$
$SQ(j-1,k+1)$	$SQ(j,k+1)$	$SQ(j+1,k+1)$

FIGURE 6.22 Quantized values of the 3×3 array corresponding to the jkth cell.

To avoid filter cells falling outside SV, the coordinates of the *jk*th cell where the filter is centered are constrained to $j = 2, 3, \ldots, J - 1$, and $k = 2, 3, \ldots, K - 1$.

3. Form the sum

$$S = \sum_{m=-1}^{1} \sum_{n=-1}^{1} w(m,n)SQ(j+m, k+n) \tag{6.33}$$

(1) If S is equal to zero, all cells have the same assigned value. This can arise only when the test cell is not an edge cell. (2) If S is positive, the test cell is an edge cell and is labeled as such. (3) If S is negative, the test cell cannot be an edge cell. On the other hand, one or more of the neighboring cells are guaranteed to be an edge cell.

6.2.3.3. Enhancement of Clutter Patch Edges

The edges deducted after smoothing tend not to follow the irregular edges that may actually exist. Consequently, the edges are further enhanced by examining the power levels of cells just outside the edge cells and the edge cells. If the power levels of these cells exceed the threshold, they are declared as edge cells otherwise they are declared as BN cells.

At the end of the edge enhancement procedure, edges are detected and each cell in the original volume is labeled as either CL, BN or CLE cell. At this point, the mapping is done.

6.2.4. Example

Consider a scanned volume containing four homogeneous clutter patches, denoted by A, B, C, D. Clutter patches C and D are contiguous and form a single nonhomogeneous clutter patch C/D as shown in Figure 6.23. The PDFs and histograms of the background noise and clutter patches are shown in Figure 6.24. The clutter-to-noise ratio for all clutter patches is 10 dB. A 3D-data plot of the scanned volume is shown in Figure 6.25. 86.33% of the total number of cells belong to the background noise. The iteration process began with the threshold arbitrarily being set such that 10% of the returns are below the threshold. After seven iterations, the process converged to a threshold so that 82.31% of the returns are below the threshold. The resulting quantized volume is shown in Figure 6.26. The corrected and smoothed volumes are shown in Figures 6.27 and 6.28, respectively. The edge-enhanced volume is shown in Figure 6.29. Finally, those cells determined to be on the edges of the clutter regions are shown in Figure 6.30. At the end of this process, only 1 CL cell was misidentified and associated with the BN. It was below the threshold. Also, 25 BN cells were misidentified and associated with CL. Of these 15 were above the threshold.

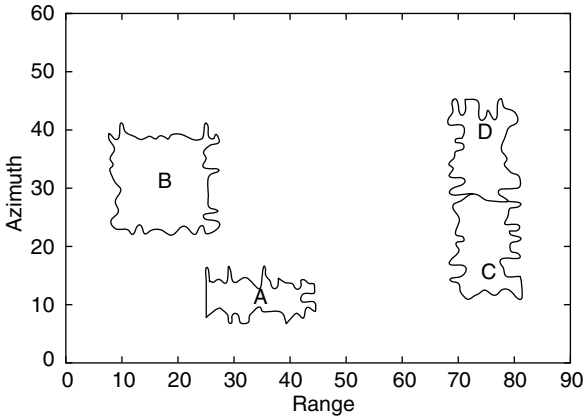

FIGURE 6.23 Contour plot of the ideal volume.

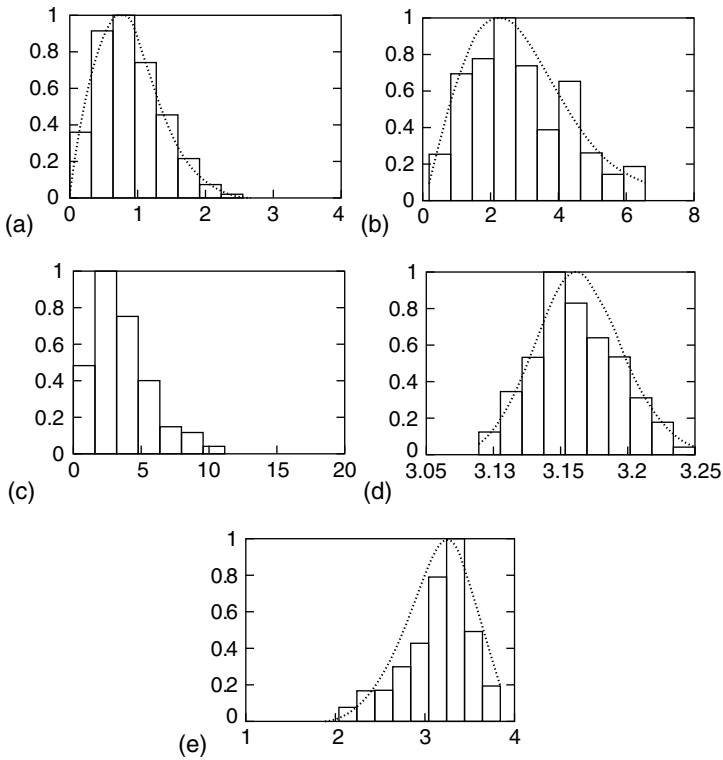

FIGURE 6.24 PDFs and histograms of the background noise and clutter patches. (a) Rayleigh distributed background noise. (b) Rayleigh distributed clutter A. (c) K-distributed clutter B with shape parameter $\rho = 10$. (d) Lognormal distributed clutter C with shape parameter $\rho = 0.01$. (e) Weibull distributed clutter D with shape parameter $\rho = 10$.

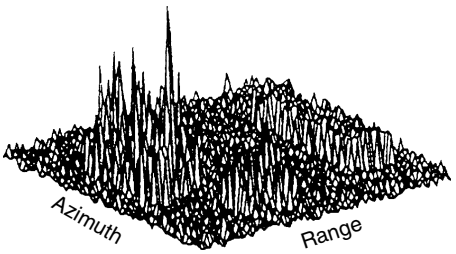

FIGURE 6.25 3D-data plot of the scanned volume.

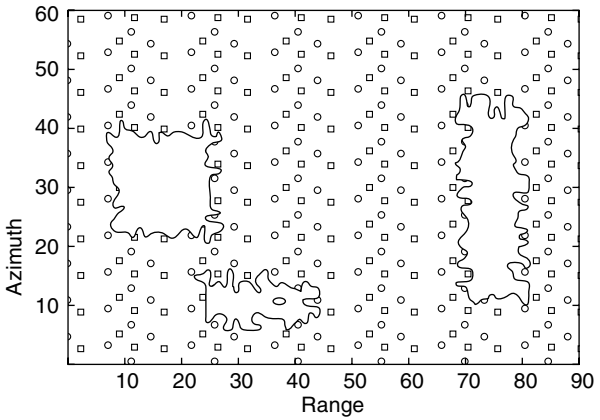

FIGURE 6.26 Contour plot of the quantized volume with threshold set so that 82.31% of the returns are below the threshold.

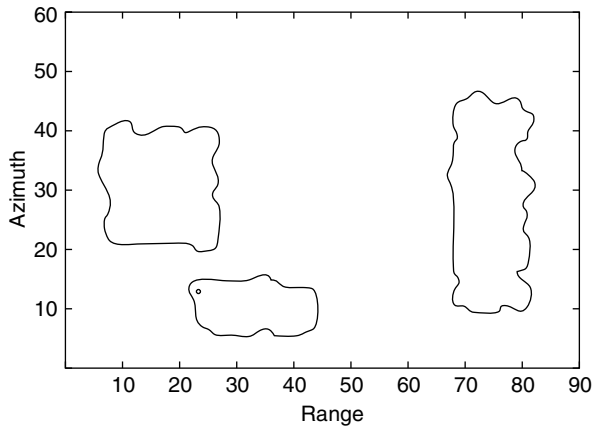

FIGURE 6.27 Contour plot of the corrected volume with NCC = 5.

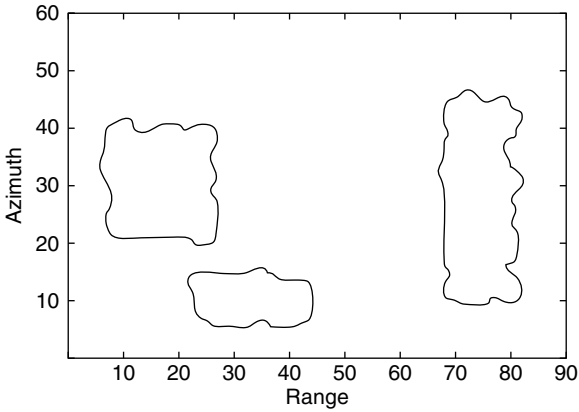

FIGURE 6.28 Contour plot of the smoothed volume with NS = 7.

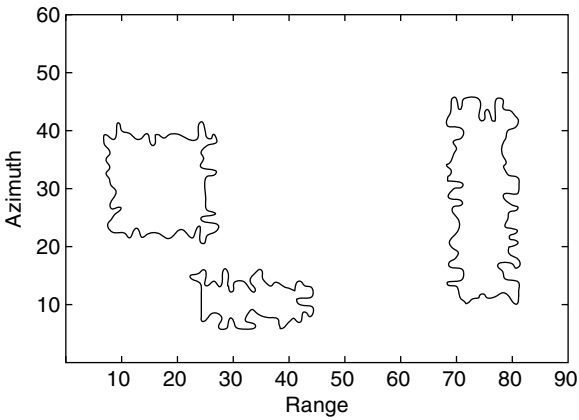

FIGURE 6.29 Contour plot of the edge enhanced volume.

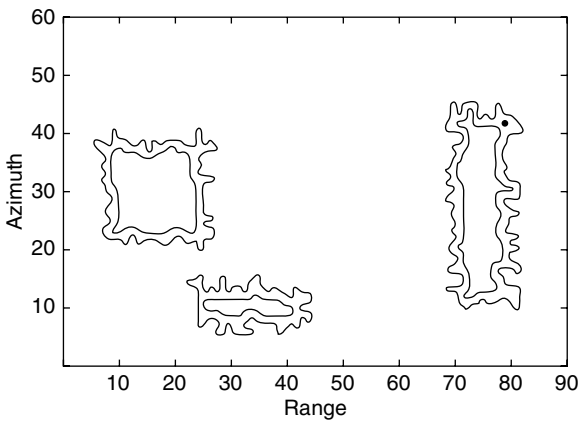

FIGURE 6.30 Contour plot of the CL edge cells.

6.3. PROBABILISTIC INSIGHT INTO THE APPLICATION OF IMAGE PROCESSING TO THE MAPPING OF CLUTTER AND NOISE REGIONS IN A RADAR SURVEILLANCE VOLUME

(M. A. SLAMANI AND D. D. WEINER)

6.3.1. Introduction

In a nonGaussian radar detection problem the choice of a signal processing algorithm depends upon whether the target is imbedded in weak background noise (a strong signal problem) or strong clutter (a weak signal problem).[1] This section gives probabilistic insight into the use of image processing for partitioning a radar surveillance volume into clutter and noise regions. An example is also given to support this insight. The mapping procedure involving quantization, correction, assessment, and smoothing has been previously described.[2]

Consider an image containing two regions where the envelope PDFs for each region are nicely separated as shown in Figure 6.31(a) and the overall PDF for both regions is as shown in Figure 6.31(b). In practice, given the image to analyze, a histogram that approximates the overall PDF is generated. Note that the individual PDF of each region is unknown. However, because the individual PDFs are adequately separated, the overall histogram will be bimodal and separation between the two regions is readily obtained by placing the threshold T1 between the two peaks as shown in Figure 6.31(b). Cells with data values lower than T1 are declared as belonging to region 1, while cells with data values higher than T1 are declared as belonging to region 2.

Now consider the slightly overlapping PDFs as shown in Figure 6.32(a) and (b). Although the overall PDF of the data regions is again bimodal, there is now noticeable overlap between the tails. Once again, a threshold T1 is used to separate between the two regions. However, now a significant number of cells will be misclassified and corrections should be made to the extent possible.

Figure 6.33 shows a more complicated case where the two regions now have major overlap between the tails. The overall PDF of the data from both regions is

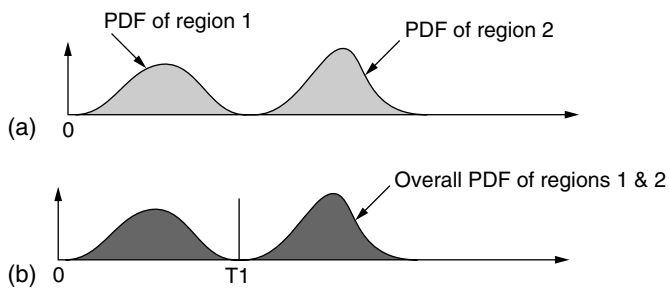

FIGURE 6.31 Nonoverlapping PDFs of two distinct regions. (a) Individual PDFs for each region. (b) Overall PDF for both regions.

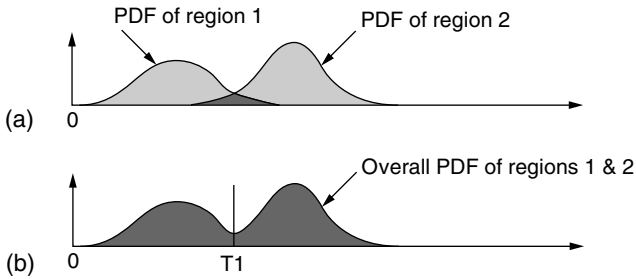

FIGURE 6.32 Overlapping PDFs of two distinct regions with a small overlapping area. (a) Individual PDFs for each region. (b) Overall PDF for both regions.

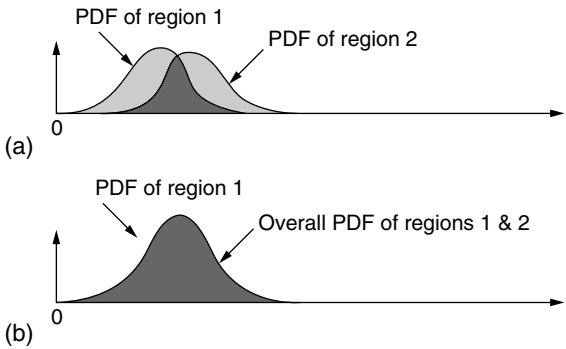

FIGURE 6.33 Overlapping PDFs of two distinct regions with a big overlapping area. (a) Individual PDFs for each region. (b) Overall PDF for both regions.

now unimodal and it is not possible to choose a threshold that separates the two regions without significant misclassifications.

In this chapter, it is shown that the mapping procedure described in Ref. 2 can adaptively choose a threshold and correct misclassifications so as to obtain good representations for the PDFs of each region. The mapping procedure enables the region having the smallest envelopes, on average, to be separated from the remaining regions. By successive application of the mapping procedure, it is possible to first separate out the region with the smallest envelope, followed by the region with the next smallest envelope, and so forth. In the first application of the mapping procedure to a radar surveillance volume, region 1 consists of the background noise (BN) while region 2 consists of the entire set of clutter (CL) patches.

6.3.2. Separation between BN and CL Patches

Before discussing the separation between BN and CL patches, a brief review of the mapping procedure presented in Ref. 2 is first given. The mapping procedure begins by selecting a threshold such that the percentage of BN cells relative to the

total number of cells is equal to a specified value, denoted by BNQP. Two correction stages then ensue. In the first correction stage, each cell in the quantized volume, denoted by QV, is tested by a 3×3 mask centered on the test cell. The test cell is labeled as BN only if less than NCQ of the eight neighboring cells were declared as CL in the QV volume where NCQ is a parameter specified by the user. In the following discussion it will be shown that the first correction stage tries to restore the right tail of the BN PDF, which had been severely distorted by the quantization. After the first correction stage, the corrected volume is denoted by CQV. The second correction stage attempts to correctly reclassify the edges of the CL patches. This is done by testing each cell in CQV using, once again, a 3×3 mask centered on the test cell. The test cell is labeled as BN if less than NCC of the eight neighboring cells were declared as CL in the CQV volume where NCC is a parameter specified by the user. Typical values for NCQ are 5, 6, 7, and 8 while typical values for NCC are 1, 2, 3, and 4. In the following discussion it will be shown that the second correction stage attempts to restore the shapes of both the BN and CL PDFs. After the second correction stage, the corrected volume is denoted by CCV. The percentage of BN cells relative to the total number of cells in the CCV volume is denoted by BNCCP. BNCCP is compared to BNQP. If the difference BNCCP − BNQP is smaller than a prespecified value, the process ends. If the difference is not too large, additional iterations are made with new values for NCQ and NCC. If these do not lead to convergence or if the difference is too large, the whole process is repeated by selecting a new threshold. If the difference is large, the new value for BNQP is chosen to be the previous BNCCP. Otherwise, the new value of BNQP is chosen to be half way in between the previous values of BNQP and BNCCP.

Another parameter that arises in the mapping procedure is BNCQP denoting the percentage of background noise cells after the first correction relative to the total number of cells in the surveillance volume. To gain insight into the relationship between BNQP, BNCQP, and BNCCP, we return to the example discussed in Ref. 2 where some of the cells are BN and the remainders are CL.

TABLE 6.7
Background Noise (BN) Percentages in the Example of Ref. 2

BNQP (%) (NCQ, NCC)	BNCQP (%)	BNCCP (%)	BNCQP − BNCCP	BNCCP − BNQP
10.00 (8,1)	56.17	20.04	36.13	10.04
20.04 (8,1)	77.98	43.78	34.20	23.74
43.78 (8,1)	90.78	82.65	8.13	38.87
82.65 (7,1)	91.17	86.13	5.04	3.48
84.39 (7,1)	91.30	86.26	5.04	1.87
84.39 (5,1)	87.59	81.98	5.61	2.41
84.39 (5,2)	87.59	83.98	3.61	0.41

TABLE 6.8
CL-to-BN and BN-to-CL Transitions in the Example of [2]

	QV → CQV		CQV → CCV		QV → CCV	
BNQP (%) (NCQ, NCC)	$(CL \rightarrow BN)_1$ (%)	$(BN \rightarrow CL)_1$ (%)	$(CL \rightarrow BN)_2$ (%)	$(BN \rightarrow CL)_2$ (%)	(CL → BN) (%)	(BN → CL) (%)
10.00 (8,1)	49.61	3.43	3.63	39.76	10.06	0.00
20.04 (8,1)	60.74	2.78	3.22	37.43	23.76	0.00
43.78 (8,1)	47.44	0.42	0.61	8.741	38.89	0.00
82.65 (7,1)	8.94	0.41	0.04	5.07	4.89	1.38
84.39 (7,1)	7.31	0.39	0.05	5.09	3.24	1.35
84.39 (5,1)	4.61	1.39	0.02	5.63	2.24	4.63
84.39 (5,2)	4.61	1.39	0.05	3.67	2.59	2.96

In Table 6.7, different values of these parameters are tabulated as the mapping procedure converges to the end result. If the test cell is to be declared as CL, recall that NCQ and NCC refer to the minimum number of neighboring cells required to be declared as CL in the QV and CQV during the first and second corrections, respectively. Table 6.8 tracks the mapping procedure during the first correction stage (denoted by QV → CQV), during the second correction stage (denoted by CQV → CCV) and at the end of the two correction stages (denoted by QV → CCV). All percentages given are with respect to the total number of cells in the surveillance volume. Initially, the threshold is set such that BNQP percent of the total number of cells is below the threshold. The first correction stage requires that at least NCQ of the neighboring cells be above the threshold if the test cell is to be classified as a CL cell. Under the column headed by QV → CQV, $(CL \rightarrow BN)_1$ denotes the percentage of the total number of cells in the surveillance volume that were above the threshold but are reclassified as BN cells during the first correction stage.

Similarly, $(BN \rightarrow CL)_1$ denotes the percentage of the total number of cells in the surveillance volume below the threshold but are reclassified as CL cells after the first correction stage. Note that the difference, $(CL \rightarrow BN)_1 - (BN \rightarrow CL)_1$, is the net percentage of the total number of cells in the surveillance volume, which have been reclassified from CL to BN cells after the first correction stage. Similar statements apply for (1) the second correction stage to $(CL \rightarrow BN)_2$, $(BN \rightarrow CL)_2$, and $(CL \rightarrow BN)_2 - (BN \rightarrow CL)_2$ under the column headed by CQV → CCV and (2) for the combined results of the two correction stages to (CL → BN), (BN → CL), and (CL → BN) − (BN → CL) under the column headed by QV → CCV. Note that

$$(CL \rightarrow BN) - (BN \rightarrow CL) = [(CL \rightarrow BN)_1 - (BN \rightarrow CL)_1] + [(CL \rightarrow BN)_2 - (BN \rightarrow CL)_2] \quad (6.34)$$

Also,

$$\begin{aligned} \text{BNCQP} - \text{BNQP} &= (\text{CL} \rightarrow \text{BN})_1 - (\text{BN} \rightarrow \text{CL})_1 \\ \text{BNCCP} - \text{BNCQP} &= (\text{CL} \rightarrow \text{BN})_2 - (\text{BN} \rightarrow \text{CL})_2 \\ \text{BNCCP} - \text{BNQP} &= (\text{CL} \rightarrow \text{BN}) - (\text{BN} \rightarrow \text{CL}) \\ &= [\text{BNCQP} - \text{BNQP}] + [\text{BNCCP} - \text{BNCQP}] \end{aligned} \tag{6.35}$$

The mapping procedure involves iterations that continue until the difference BNCCP − BNQP is sufficiently small. From Equation 6.35, it is seen that convergence results when

$$(\text{CL} \rightarrow \text{BN}) \approx (\text{BN} \rightarrow \text{CL}) \tag{6.36}$$

Consequently, near convergence, the combined effect of the two correction stages should result in the percentage of CL cells reclassified as BN cells being approximately equal to the percentage of BN cells reclassified as CL cells. Alternatively, from Equation 6.35, convergence results when

$$[\text{BNCQP} - \text{BNQP}] \approx -[\text{BNCCP} - \text{BNCQP}] \tag{6.37}$$

or equivalently, when

$$[(\text{CL} \rightarrow \text{BN})_1 - (\text{BN} \rightarrow \text{CL})_1] \approx -[(\text{CL} \rightarrow \text{BN})_2 - (\text{BN} \rightarrow \text{CL})_2] \tag{6.38}$$

Thus, near convergence, the net percentage of cells that have been reclassified from CL to BN cells during the first correction stage should approximately equal the negative of the net percentage of cells that have been reclassified from CL to BN cells during the second correction stage. These observations are helpful in coming up with rules for determining the next setting of the parameters in the iteration process.

By way of example, when BNQP = 10%, the threshold is such that 10% of the total number of cells in the surveillance volume fall below the threshold while 90% fall above. The situation is pictured in Figure 6.34(b). With reference to Table 6.8, when NCQ = 8, 49.61% of the total cells in the surveillance volume, classified as CL cells because they were above the threshold, are reclassified as BN cells after the first correction stage whereas 3.43%, classified as BN cells because they were below the threshold, are reclassified as CL cells. The net percentage of cells reclassified as BN is 49.61% − 3.42% = 46.18%. For the second correction stage, with NCC = 1, 3.63% of the total cells in the CQV surveillance volume, classified as CL cells after the first correction stage, are reclassified as BN cells because they do not have at least one neighboring CL cell. Similarly, 39.76% of the total cells in the CQV surveillance volume classified as BN cells are reclassified as CL cells because they have one or more neighboring CL cells. The last row of Table 6.8 corresponds to a situation close to convergence. With the threshold set such that 84.39% of the total number of cells in the surveillance volume are below the threshold, note that the combined effect

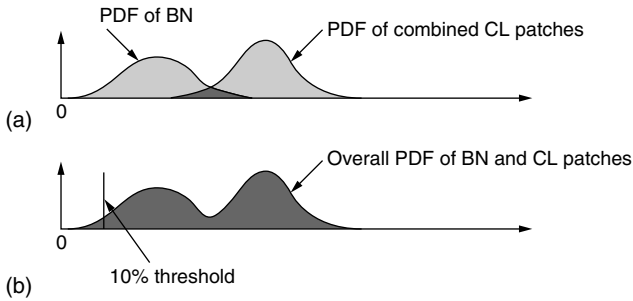

FIGURE 6.34 Overlapping PDFs of background noise (BN) and clutter (CL) regions. (a) Individual PDFs for each region. (b) Overall PDF for both regions.

of the two correction stages has resulted in

$$(\text{CL} \rightarrow \text{BN}) = 2.59\% \approx (\text{BN} \rightarrow \text{CL}) = 2.98\% \tag{6.39}$$

Similarly,

$$\begin{aligned} &[(\text{CL} \rightarrow \text{BN})_1 - (\text{BN} \rightarrow \text{CL})_1] = 3.22\% \\ \approx &-[(\text{CL} \rightarrow \text{BN})_2 - (\text{BN} \rightarrow \text{CL})_2] = 3.62\% \end{aligned} \tag{6.40}$$

Equivalently from Table 6.7,

$$[\text{BNCQP} - \text{BNQP}] = 3.20\% \approx -[\text{BNCCP} - \text{BNCQP}] = 3.61\% \tag{6.41}$$

Insight into the manner by which the PDFs of BN and CL are modified during the correction stages is obtained by examining pertinent amplitude histograms for the various surveillance volumes QV, CQV, and CCV. The overall amplitude histogram for the generated data of the QV volume is shown in Figure 6.35(a). Roughly speaking, the amplitude of the BN cells appears to extend from 0 to 2.5 while those of the CL cells appear to extend from 2.5 to 34. When the threshold is set at 0.35 such that BNQP = 10%, many of the BN cells are classified as CL due to the low threshold. The amplitude histograms for the BN and CL cells in the QV volume are shown in Figure 6.35(b) and (c), respectively. Note that the BN histogram is truncated to an amplitude of 0.35. Also, note that many cells with amplitude below 2.5 are misclassified as CL and are included in the first bar of the CL histogram.

The amplitude histograms for the CQV volume resulting from the first correction stage are shown in Figure 6.35(d) and (e). Comparing Figure 6.35(d) with (b), it is seen that many cells with amplitudes above the threshold value of 0.35 have been reclassified as BN. Also, by comparing Figure 6.35(e) with (c), we can see that the height of the first bar has been reduced from 0.26 to 0.225 indicating that many of the CL cells reclassified as BN came from this bin.

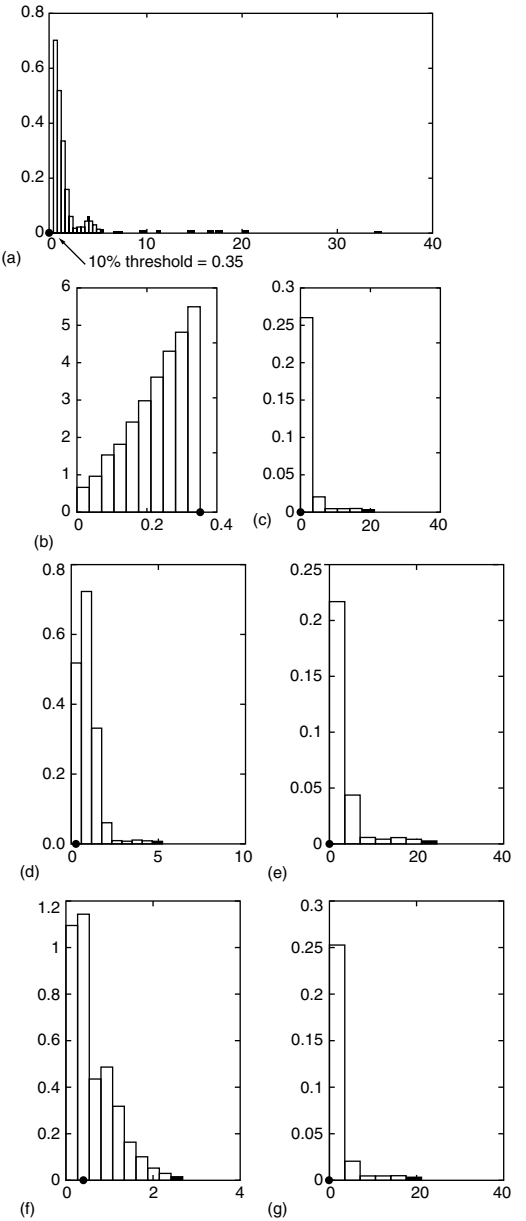

FIGURE 6.35 Regions histograms corresponding to BNQP = 10%, NCQ = 8, and NCC = 1. (a) Overall histogram of the generated data. (b) BN histogram at the quantization stage. (c) CL histogram at the quantization stage. (d) BN histogram at the first correction stage. (e) CL histogram at the first correction stage. (f) BN histogram at the second correction stage. (g) CL histogram at the second correction stage.

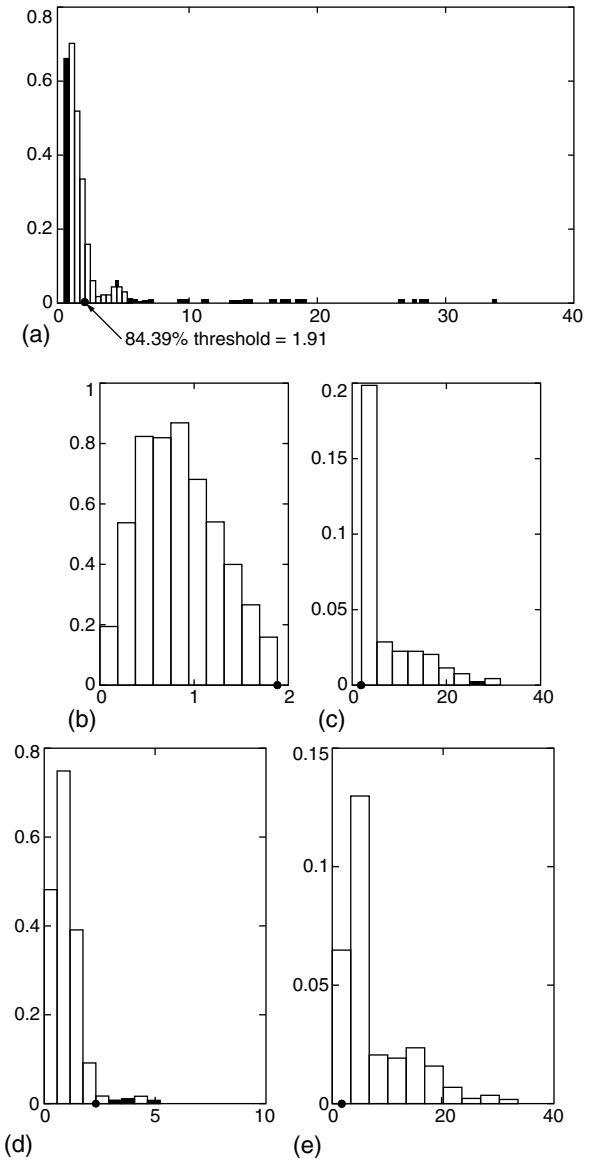

FIGURE 6.36 Regions Histograms Corresponding to BNQP = 84.39%, NCQ = 5, and NCC = 2. (a) Overall histogram of the generated data. (b) BN histogram at the quantization stage. (c) CL histogram at the quantization stage. (d) BN histogram at the first correction stage. (e) CL histogram at the first correction stage. (f) BN histogram at the second correction stage. (g) CL histogram at the second correction stage. (h) BN histogram at the mapped volume. (i) CL histogram at the mapped volume. (j) Actual BN histogram of the generated data. (k) Actual CL histogram of the generated data.

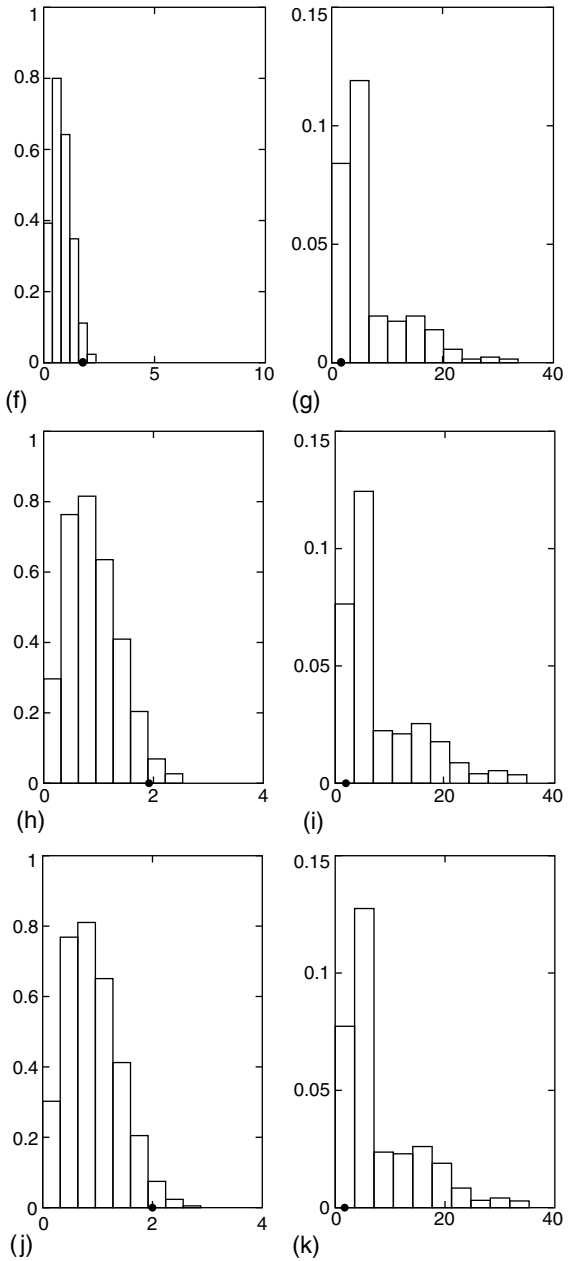

FIGURE 6.36 Continued.

Observe the reconstruction of the right tail of the BN amplitude histogram. The amplitude histograms for the CCV volume resulting from the second correction stage are shown in Figure 6.35(f) and (g). Note the further enhancement of the right tail of the BN amplitude histogram. This enhancement is due to the image processing and is in spite of the low threshold value.

During the iteration process the threshold is gradually increased and converges to a value in the vicinity of 1.91 for which BNQP = 84.39%. In Figure 6.36(a), this threshold is shown in the overall histogram for the QV volume. The amplitude histograms for the BN and CL cells in the QV volume are shown in Figure 6.36(b) and (c), respectively. Note that the amplitude of the BN cells fall below 1.9 whereas those of the CL cells fall above 1.9. The results of the first and second correction stages and the edge enhancement stage are shown in Figures 6.36(d) and (e), Figure 6.36(f) and (g), and Figure 6.36(h) and (i), respectively. To provide a basis for comparison, the actual BN and CL amplitude histograms are shown in Figure 6.36(j) and (k). The strong similarity between the amplitude histograms of Figure 6.36 (h) and (i) and those of Figure 6.36(j) and (k) indicates that the mapping procedure has converged satisfactorily. Note how nicely the final histograms of Figure 6.36(h) and (i) have evolved from the original histograms of Figure 6.35(b) and (c).

In general, the first correction stage begins to establish the right tail of the BN amplitude histogram and reshape the CL amplitude histogram by reclassifying mislabeled BN cells. The second correction stage reshapes both the bodies and the tails of the BN and CL histograms by recovering the CL edges.

6.3.3. Summary

In previous paper, techniques were presented for treating the weak signal problem[1] and for partitioning a radar surveillance volume into BN and CL patches.[2] This section has provided a probabilistic insight into the technique that helps to explain its success.

6.4. A NEW APPROACH TO THE ANALYSIS OF IR IMAGES (M. A. SLAMANI, D. FERRIS, AND V. VANNICOLA)

6.4.1. Introduction

In signal processing applications it is common to assume a Gaussian problem in the design of optimal signal processors. However, nonGaussian processes do arise in many situations. For example, measurements reveal that clutter can be approximated by either Weibull, K-distributed, Lognormal, or Gaussian distributions depending upon the scenario.[1] When the possibility of a nonGaussian problem is encountered, the question as to which probability distributions should be utilized in a specific situation for modeling the data needs to be answered. In practice, the underlying probability distributions are not

known *a priori*. Consequently, an assessment must be made by monitoring the environment to subdivide the surveillance volume into homogeneous patches in addition to approximating the underlying probability distributions for each patch.

The assessment of the environment is performed by an automatic statistical characterization and partitioning of environments (ASCAPE) process, previously used on simulated data.[2,3] ASCAPE uses two separate procedures to determine all homogeneous patches and subpatches in the IR image. The first procedure, referred to as the mapping procedure, is used to separate contiguous homogeneous regions by segregating between their power levels. The second procedure, referred to as the statistical procedure, separates contiguous homogeneous patches by segregating between their probabilistic data distributions. The statistical procedure uses the Ozturk algorithm, a newly developed algorithm for analyzing random data.[4] Furthermore, the statistical procedure identifies suitable approximations to the PDF for each homogeneous patch and determines the location of outliers. Convergence of the procedures is controlled by an expert system shell.

In this work, ASCAPE is introduced in Section 6.4.2. The mapping and statistical procedures are presented in Sections 6.4.3 and 6.4.4, respectively. The expert system shell is discussed in Section 6.4.5. Finally, an example illustrating the different stages of ASCAPE when applied to real data of an IR image is given in Section 6.4.6.

6.4.2. ASCAPE

The ASCAPE process, shown in Figure 6.37, consists of four interactive blocks. The first block is a preprocessing block that performs classical space–time processing on the collected data. Then, based on the mapping procedure, the second block separates contiguous homogeneous patches and subpatches by segregating between their average power level. The next block goes one step further and separates contiguous homogeneous subpatches by segregating between their probabilistic data distributions. This block also approximates the PDF of each homogeneous subpatch and determines the location of outliers in the scene. The final block indexes the scene under investigation by assigning a set of descriptors to every cell in the scene. For each cell, the indexing is used to indicate to which homogeneous patch the cell belongs, whether it is an edge cell or an outlier, which cells can be chosen as reference cells if the cell is to be tested; and which PDF best approximates the data in the cell. Note that the reference cells are cells from the same homogeneous patch and closest to the cell to be tested.

The forward and backward interactions between the different blocks are controlled by an expert system shell referred to as Integrated Processing and Understanding of Signals (IPUS) developed jointly by the University of Massachusetts and Boston University.[5]

When ASCAPE is followed by a detection stage (e.g., target detection in Radar), all information needed by the detector is available for every cell in the

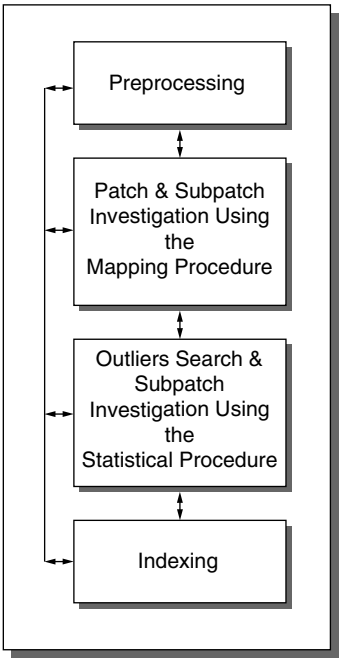

FIGURE 6.37 ASCAPE's Block Diagram.

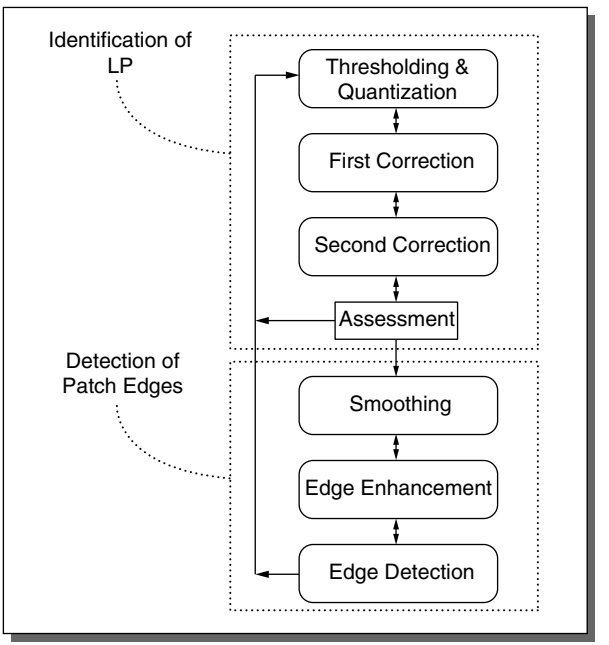

FIGURE 6.38 Mapping Procedure.

scene. Furthermore, given the PDF that can approximate the patch where the test cell is located, the appropriate detector is readily selected. This is in contrast to classical detection approach where a single detector (usually the Gaussian detector) is used in processing the entire scene.

6.4.3. Mapping Procedure

As shown in Figure 6.38 and explained in detail in Refs. 2 and 6. The mapping procedure consists of two stages. In the first stage, the patch with the lowest average power, referred to as the lowest patch (LP), among all remaining patches (RPs) is identified. In the second stage, edges of the LP are enhanced and detected.

These two stages are repeated to identify the next LP and so on. The mapping procedure is repeated continuously until it is not possible to separate anymore between patches, and all patches are declared to be homogeneous. Once it becomes impossible to find any more patches, every patch is processed by the mapping procedure, as discussed above, for detection of subpatches.

The two stages are explained next.

6.4.3.1. Identification of Lowest Average Power Level (LP)

This stage consists of an iterative process for automatically setting a threshold to separate between the patch with the lowest average power level (LP) and the remaining patches (RPs). It is composed of the blocks labeled Quantization and Thresholding, First Correction, Second Correction, and Assessment, as shown in Figure 6.38. Using the fact that the LP, on average, has smaller magnitudes than the RPs, identification of LP within the RPs starts by setting a threshold q that results in a specified fraction of LP cells. Then a quantized volume is formed as follows: all cells with magnitude less than q are given a value of zero and all cells with magnitude above q are given a value of one. Masking techniques are then used in the First Correction and Second Correction blocks, as described in Refs. 3 and 6 to establish the LP and the RPs. If the final scene contains a significantly different fraction of LP cells than originally established by the initial threshold, the Assessment block decides that the process is to be repeated with a new threshold. The mapping processor iterates until the final scene is consistent with the latest specified threshold. It is to be noted that most of the thresholding techniques found in the literature assume that the histogram of the scene data is at least bimodal.[7–9] In practice, data collected from different regions may result in a unimodal histogram making it difficult to select a threshold. The mapping procedure has been shown to be powerful enough to separate between regions even when their histograms overlap significantly.[2,6]

6.4.3.2. Detection of Patch Edges

This stage consists of an edge enhancement and detection process to enable the detection of the edges for the different patches. It is composed of the blocks

labeled Smoothing, Edge Enhancement, and Edge Detection. The smoothing block uses a masking technique to detect isolated cells in the LP and RPs patches and label them adequately. The isolated cells are due to RPs declared cells in LP and LP declared cells in RPs. These originate from cells with large magnitudes in LP and cells with small magnitudes in RPs. The edges obtained after smoothing tend not to follow the irregular edges that may actually exist. Consequently, the edges are further enhanced in the edge enhancement block by examining the magnitudes of the edge cells and cells just outside the edge cells. If the magnitudes of these cells fall below the latest threshold q, they are declared as LP edge cells (LPE). Otherwise they are declared as RPs edge cells (RPsE). At the end of the edge enhancement procedure, edges are detected in the edge detection block using the unsharp masking technique.[8] This is important for subsequent signal processing if reference cells for estimating parameters of a test cell are to be chosen properly.

6.4.4. Statistical Procedure

When no more patches (subpatches) can be found, the mapping procedure ends and is followed by the statistical procedure that is applied to every patch and subpatch declared to be homogeneous by the mapping procedure in order to (1) further separate nonhomogeneous subpatches having very similar power levels but different statistical distributions, (2) locate outliers in the scene, and (3) approximate the PDF of each homogeneous patch and subpatch.

The Ozturk algorithm is used by the statistical procedure to approximate the PDF of each patch and is presented next followed by the definition of outliers and the strategy used in the statistical procedure.

6.4.4.1. Introduction to Ozturk Algorithm

The Ozturk algorithm is based on sample order statistics and is used for univariate distribution approximation.[4,6] This algorithm has two modes of operation. In the first mode, the algorithm performs a goodness-of-fit test. The test determines, to a desired confidence level, whether the random data is statistically consistent with a specified probability distribution. The program utilizes a Monte Carlo simulation of 2000 trials to generate an averaged set of NR linked vectors in the u-v plane, as shown in Figure 6.39(a). Using the standardized sample order statistics of the data, the program then creates a second system of NR linked vectors in the u-v plane. The terminal points of the linked vectors, as well as the shapes of their trajectories, are used in determining whether or not to accept the null hypothesis. The null hypothesis is the distribution against that the sample data is to be tested. The algorithm provides quantitative information as to how consistent the sample data set is with the null hypothesis distribution by use of confidence ellipses where each ellipse is derived from a specified probability that the end point falls within the ellipse, given that the data comes from the null distribution. An example of these ellipses is also

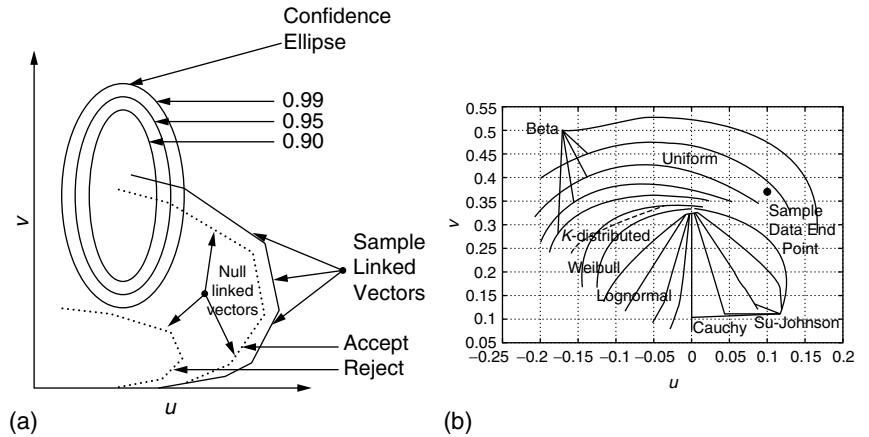

FIGURE 6.39 Ozturk algorithm. (a) Goodness-of-fit test. (b) Approximation chart.

shown in Figure 6.39(a) for probabilities of 0.90, 0.95, and 0.99. If the end point of the sample data linked vector locus falls within an ellipse, then the sample data set is said to be statistically consistent with the null hypothesis. The confidence level is the probability specified for that ellipse. If the sample data set is truly consistent with the null hypothesis, the system of sample linked vectors is likely to closely follow that for the system of null linked vectors. This mode is referred to as the *goodness-of-fit mode*.

In the second mode of operation, referred to as the *approximation chart mode*, and following a similar approach to that outlined in the goodness-of-fit mode, random samples are generated from a library of different univariate probability distributions. In the goodness-of-fit test mode, the locus end point was obtained for the null hypothesis and sample size, NR. For the approximation chart mode we go one step further by obtaining the locus end point for each distribution from the library of distributions for the given sample size, NR, and for various choices of the shape parameters. Thus, depending on whether it has a shape parameter or not, each distribution is represented by a trajectory or point in the two dimensional u-v plane. Figure 6.39(b) shows an example of the approximation chart. Note that every point in the approximation chart corresponds to a specific distribution. The point closest to the sample data locus end point is chosen as the best approximation to the PDF underlying the random data. This closest point is determined by projecting the sample locus end point to all points on the approximation chart and selecting that point whose perpendicular distance from the sample point is the smallest. Once the PDF underlying the sample data is selected, the shape, location, and scale parameters are then approximated. The algorithm has been found to work reasonably well for observation sizes as small as 100. Throughout this work, 100 reference cells are selected for approximating the PDF of each test cell.

6.4.4.2. Outliers

Outliers that may exist in a set of reference cells can alter the statistical distribution of the set of data under examination. Outliers may be due to (1) misidentified LP cells in the RPs or vise-versa, (2) cells having data values of very low probability of occurrence, (3) cells with nonrepresentative data, and (4) cells containing signals from strong targets. One way to identify outliers within a region composed of a set of 100 reference cells is to compute the mean m and the standard deviation s within that region and call as an outlier any cell whose data value falls outside the interval $[m - ks,\ m + ks]$ where k is an empirical parameter (usually between 1.5 and 3) to be set by the user. In our case k has been set equal to two.

6.4.4.3. Strategy to SubPatch Investigation Using the Statistical Procedure

The statistical procedure is applied to every patch and subpatch that has been declared as being homogeneous by the mapping procedure. For each patch and subpatch, a set of test cells evenly spread throughout the patch or subpatch, and their 100 closest reference cells are first selected. Let each set of 100 cells be referred to as a tile. Note that the sets of 100 reference cells are chosen to be disjoint, the closest to and belonging to the same patch as their respective test cells. This results in (1) the sets being shaped as 10×10 square tiles inside a patch and (2) tiles tracking the shape of the edges near the boundary of the patch. As shown in the block diagram of Figure 6.40, the statistical procedure consists of four steps that are performed as follows:

(1) Using the goodness-of-fit mode of the Ozturk algorithm, a First Gaussianity check is performed by the first block on every tile to ensure whether the data in the tile are Gaussian or not. This results in every patch having its tiles labeled as either Gaussian or nonGaussian.
(2) Existing outliers are located in those tiles declared as nonGaussian in (1). This step is performed by the second block.
(3) For every nonGaussian declared tile, cells with outliers are excised from the tile and replaced with the closest cells to the tile whose data are not outliers, and, the Gaussianity check is performed once again as in step (1).
(4) The last block, labeled PDF Approximation and Detection of Subpatches, consists of the following substeps:
    (i) Every set of contiguous nonGaussian tiles is declared as a subpatch.
    (ii) Using the Ozturk Algorithm, the (u,v) coordinates of the locus end point is obtained for every tile declared as nonGaussian in step (3).
    (iii) For every nonGaussian subpatch, as declared in (i), a check is made to ensure whether or not the data of the set of tiles that

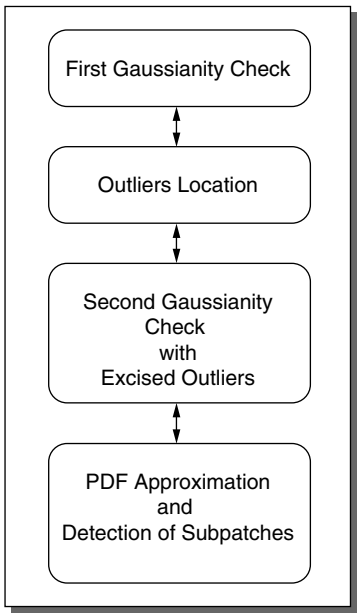

FIGURE 6.40 Statistical procedure.

constitutes a subpatch can fit within a unique confidence ellipse and therefore be approximated by a unique PDF. This is done by computing the average (u_{av}, v_{av}) coordinates from the (u,v) coordinates of all tiles of the same subpatch and obtaining the best approximating PDF and its corresponding confidence ellipse, as described in Section 6.4.4.1. A check is then made to verify whether all (u,v) coordinates of the tiles of the same subpatch are within the confidence ellipse. If not, the tiles are regrouped so that all (u,v) coordinates for each group of tiles can fit within the same ellipse. Each group forms then a subpatch with best approximating PDF defined by the center of the confidence ellipse.

When the statistical procedure ends, each cell in every patch is declared as either Gaussian, nonGaussian, or outlier. In addition, PDFs are approximated for each nonGaussian cell, and existing subpatches, formed by the sets of contiguous tiles whose (u,v) coordinates fit under the same confidence ellipse, are identified.

6.4.5. Expert System Shell IPUS

The expert system shell used to control the different stages of the mapping and indexing procedures is known as Integrated Processing and Understanding

of Signals (IPUS) and was developed jointly by University of Massachusetts and Boston University.[5] The IPUS architecture utilizes the fact that signal processing theories supply the system designer with a signal processing algorithm (SPA) that has adjustable control parameters. Each instance, corresponding to a particular set of fixed values for the control parameters, is referred to as an SPA instance. The IPUS architecture is designed to search for appropriate SPA instances to be used in order to accurately model the unknown environment. The search is initiated by detection of a discrepancy at the output of a given SPA due to the fact that the signal being monitored by the SPA does not satisfy the requirements of the SPA instance. Once a discrepancy has been detected, a diagnosis procedure is used to identify the source of the distortion that may have led to the discrepancy. Then, either parameters of the same SPA are readjusted, or a different SPA is chosen to reprocess the data.

In our case of interest, each block in the different stages and substages of the ASCAPE processor consists of an SPA and SPA instances. Rules have been developed enabling the detection of discrepancies at the output of the SPAs, and identification of different possible distortion sources that would cause the discrepancies.[6] In Figures 6.37, 6.38 and 6.40 note that the arrows connecting different blocks are double headed. This is to allow for ASCAPE, controlled by IPUS, to assess its decisions, correct any discrepancies, and adapt to any changes in the environment being monitored.

6.4.6. Example: Application of ASCAPE to Real IR Data

Consider an IR image of real data collected over lake Michigan. As shown in Figure 6.41, the scene consists of two major regions: lake and land. Furthermore, the three dimensional plot of the data magnitudes in Figure 6.42 shows that the data in the lake region are regular, whereas the data in the land region are irregular and contain large discretes near the boundary close to the lake.

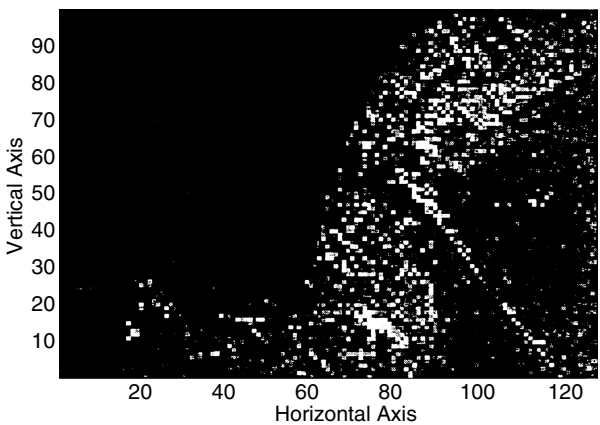

FIGURE 6.41 Original scene (over Lake Michigan).

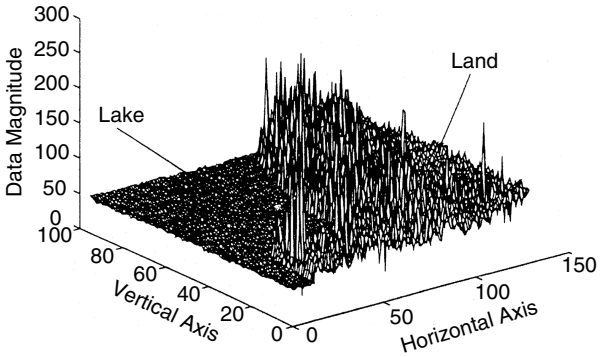

FIGURE 6.42 Three-Dimensional magnitude plot of the original scene.

ASCAPE is applied to the IR scene. First, the preprocessing block identifies the data in the scene, which is uncorrelated as required by the Ozturk algorithm for the univariate PDF approximation. Then, application of the mapping procedure results in the segmentation of the scene into three different patches and their respective boundaries, as shown in Figure 6.43. Parameters of each patch are summarized in Table 6.9. Note that In addition to regions 1 and 2, corresponding to land and lake, respectively, the mapping procedure detects a third region, labeled 3. This region can be sighted in the original scene of Figure 6.41 and is not large in size. In fact, Table 6.9 shows that patch 3 contains only 15 cells as opposed to patches 1 and 2, which contain 6680 and 5337 cells, respectively. Furthermore, the values in Table 6.9 of the variance, mean, an average power of patch 3 are closer to those of the lake than those of the land. This indicates that patch 3 may be a very small body of water. Edges of the different patches are also

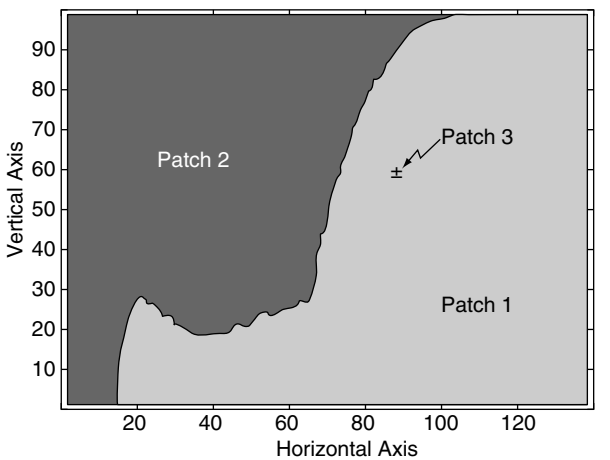

FIGURE 6.43 Scene after the mapping procedure: detection of patches.

TABLE 6.9
Mapping Assessment

	Patch 1	Patch 2	Patch 3
Number of cells	6680	5337	15
Mean	87.61	49.47	34.4
Variance	420.20	4.13	7.41
Average Power (dB)	39.08	33.89	34.40

detected as shown in Figure 6.44. Next, the mapping procedure is used again to investigate the presence of any subpatches in every previously detected patch (i.e., patches labeled 1, 2, and 3). In this case, no subpatches are detected and all three patches are declared to be homogeneous. At this point, it is not possible to separate anymore between existing contiguous homogeneous subpatches by segregating between their power levels.

Next, the statistical procedure is applied to every previously declared homogeneous patch in order to separate further between any existing contiguous subpatches that may have similar power levels but different data distributions. The procedure proceeds as follows for every patch:

(1) Test cells and their respective tiles (sets of 100 reference cells) are selected, spread throughout the patch. Recall that the sets of 100 cells are chosen to be disjoint, the closest to and belonging to the same patch as their respective test cells. This results in sets being shaped as 10×10 square tiles inside the patch and in tiles tracking the shape of the boundary near the boundary of the patch. The tiles are then tested

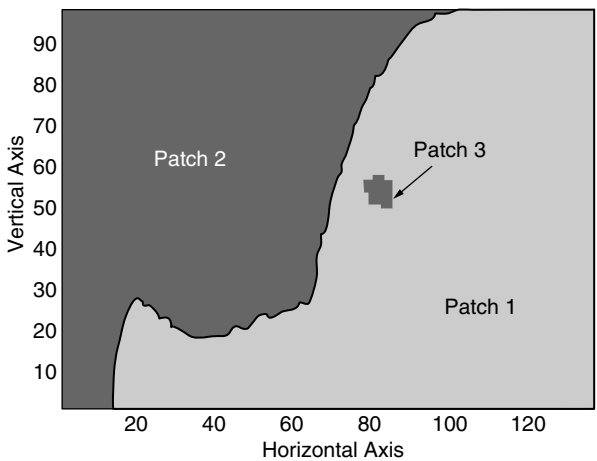

FIGURE 6.44 Scene after the mapping procedure: detection of edges.

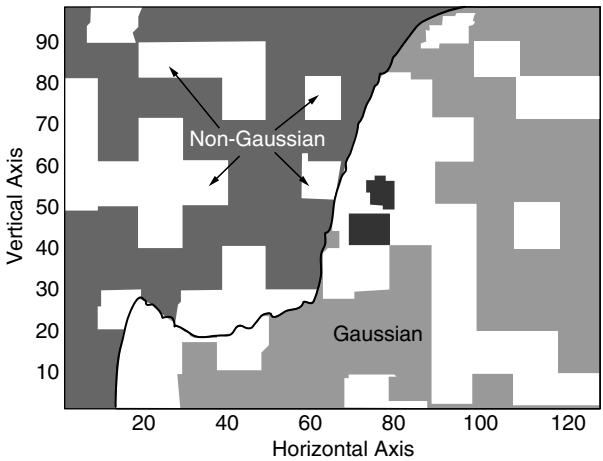

FIGURE 6.45 Scene after the first gaussianity check.

for Gaussianity using the goodness-of-fit test of the Ozturk Algorithm. The result of this step is shown in Figure 6.45. Note that a lot of nonGaussian tiles, represented by the white regions, exist in patches 1 and 2. Note also that patch 3 is not processed by the statistical procedure due to the fact that it has only 15 cells while a minimum of 100 cells are required for the Ozturk algorithm to result in a meaningful approximation for the distribution of the data.

(2) Once the Gaussian and nonGaussian tiles are determined in each patch, cells with outliers are located in the nonGaussian declared tiles and excised.

(3) The nonGaussian previously declared tiles with excised outliers are tested again for Gaussianity. The results are shown in Figure 6.46 where the gray and white regions represent the Gaussian and NonGaussian tiles, respectively, whereas the black pixels represent the location of outliers. The quantitative results are summarized in Table 6.10. Note that:

(i) The area occupied by the nonGaussian tiles has been reduced from 38.19% to 15.94% of the total area of the scene when outliers are excised.

(ii) Even when outliers are excised, nonGaussian tiles exist and thus the nonGaussian problem is important to be considered.

(iii) By comparing Figures 6.41 and 6.46, note that the trajectories of the outliers located in the land region (patch 1) follow paths that can be distinguished in the original scene. These paths may represent highways or roads. Therefore outliers may represent regions in the scene that cannot be considered as patches due to the fact that they do not obey the rules set forward for a patch to be composed of inner cells and edge cells.

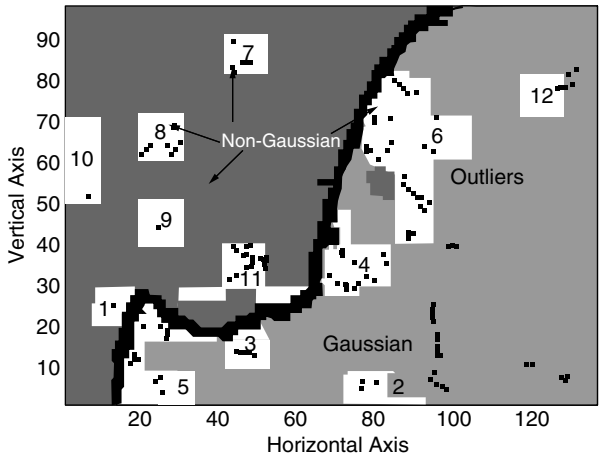

FIGURE 6.46 Scene after the statistical procedure.

(4) Following the instructions presented in step (4) of Section 6.4.4.3, nonGaussian subpatches are determined and numbered. As shown in Figure 6.46, twelve nonGaussian subpatches exist in the scene, six in the lake region and six in the land region. Also, the PDFs of these nonGaussian subpatches are approximated and the results are posted in Table 6.11. Note that the statistical procedure was able to detect more subpatches and therefore subdivide further the regions due to the fact that even though the different regions in each patch have very similar average power, the distribution of the data within each patch differs from one region to another. This is clearly seen in the magnitude plot of the land region in Figure 6.42.

At the end of step (4), processing of the image results in three main patches (labeled 1, 2, and 3) and their respective edges as shown in Figure 6.44. Also, outliers are located in patches 1 and 2. In addition, nonGaussian subpatches and their approximating PDFs are determined within every one of patches 1 and 2.

TABLE 6.10
Percentages of Gaussian, NonGaussian, and Outlier Cells

	Percentage of Gaussian Cells (%)	Percentage of NonGaussian Cells (%)	Percentage of Outliers (%)
Tiles with present outliers, step (1)	61.81	38.19	—
Tiles with excised outliers, step (3)	82.43	15.94	1.63

TABLE 6.11
PDF Types and Parameters of the NonGaussian Subpatches

Sub patch	1	2	3	4	5	6
PDF Type	Beta	Beta	SU-Johnson	Beta	Beta	Beta
Shape 1	6.0	1.0	1.2	1.0	1.0	1.0
Shape 2	1.6	3.2	−0.2	3.2	1.6	3.2
	7	**8**	**9**	**10**	**11**	**12**
PDF Type	Beta	Beta	Beta	Beta	SU-Johnson	Beta
Shape 1	1.0	2.0	2.0	2.0	1.2	1.0
Shape 2	0.8	0.8	1.6	1.6	−0.2	1.6

The statistical procedure is followed by the indexing stage in which every cell in the scene is assigned a set of descriptors to indicate, for each cell: the homogenous patch to which it belongs, whether it is an edge cell or an outlier, which cells can be chosen as reference cells if the cell is to be tested, and which PDF best approximates the data in the cell.

6.4.7. Conclusion

NonGaussian detectors should be utilized in cases that cannot be handled by the classical Gaussian detector (e.g., the nonGaussian regions where the clutter returns are much stronger than those of the targets). In such cases, the appropriate detector must be selected. This is done by monitoring the environment under investigation to (1) segment the scene into homogeneous regions, and (2) approximate the PDFs of each region.

This work has presented: (1) a new adaptive image processing procedure which segments contiguous homogeneous regions with different power levels, (2) a statistical procedure to segment contiguous homogeneous regions with similar power levels but different data distributions, (3) detection and excision of outliers and its significance, and (4) determination of the PDFs of nonGaussian regions. These procedures are part of the new process for the automatic statistical characterization and partitioning of environments (ASCAPE).

Work under investigation and future work include: (1) tailoring/tuning of the edges for the subpatches detected by the statistical procedure, limited thus far by the requirement of 100 reference cells so that the Ozturk Algorithm leads to meaningful results, (2) a performance study of the nonGaussian detectors vs. the Gaussian detector, in different types of environments and under different circumstances, (3) development of more expert system rules to enable ASCAPE to be applied to different types of data (e.g., radar, IR, sonar, medical imaging, etc.), and (4) application of ASCAPE to medical imaging (e.g., detection of tumors in lung) and to other areas.

6.5. AUTOMATIC STATISTICAL CHARACTERIZATION AND PARTITIONING OF ENVIRONMENTS (ASCAPE)

(M. A. SLAMANI, D. D. WEINER, AND V. VANNICOLA)

6.5.1. Problem Statement

In signal processing applications it is common to assume a Gaussian problem in the design of optimal signal processors. However, studies have shown that the Gaussian receiver performs very poorly in strong interference whenever the interference and signal spectra cannot be separated by filtering. For example, consider the spectra shown in Figure 6.47 consisting of 24 Doppler bins with uniformly spaced targets indicated by the small arrows, embedded in background noise and a bell shaped Gaussian interference. The optimal Gaussian based detector (referred to as the joint-domain localized generalized likelihood ratio receiver) is applied to each Doppler bin. The performance of the receiver,[1] shown in Figure 6.48, reveals that the probability of detection (PD) of the receiver is close to unity everywhere except for Doppler bins 11, 12, and 13, in which a strong Gaussian interference-to-noise ratio (INR) of 50 dB exists and PD falls rapidly to the probability of false alarm (PFA).

A question that arises is "Could improved detection have been obtained in bins 11 to 13 had the disturbance been nonGaussian?" the answer is "It is possible to achieve significant improvement in detection performance when the disturbance is nonGaussian". Table 6.12 presents the results for a weak target in K-distributed clutter where the target and clutter spectra are coincident (i.e., not separable).[2] Note that the K-distributed based detector provides significantly improved performance even when the PD of the Gaussian detector approaches the PFA. For example, with a signal-to-clutter ratio (SCR) of SCR = 0 dB and PFA = 10^{-5}, PD equals 0.3 for the K-distributed detector and 10^{-5} for the Gaussian detector. This means that the K-distributed based detector can detect three out of ten targets, whereas the Gaussian detector can detect only

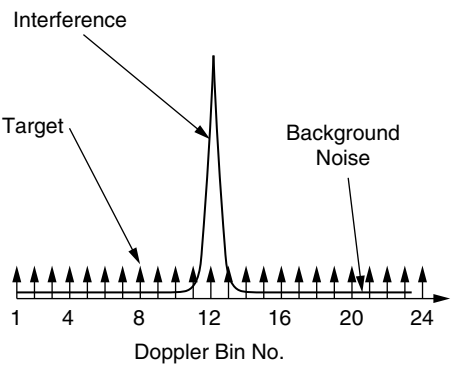

FIGURE 6.47 Targets in NonGaussian interference.

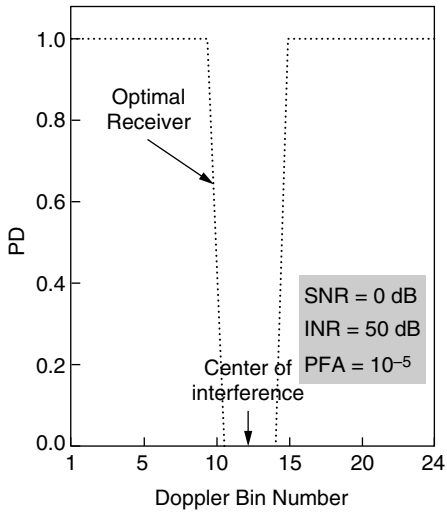

FIGURE 6.48 Performance of the optimal receiver.

one in 100,000 targets on the average. *The important point* is that detections can be made in a nonGaussian environment using a nonGaussian detector even when such performance is not possible with the Gaussian detector. *Unfortunately*, there are problems associated with the utilization of nonGaussian distributions:

(1) There are a multitude of nonGaussian distributions.
(2) Some nonGaussian distributions have shape parameters that result in different shaped PDFs.
(3) For a given set of data, it is difficult to determine which PDF can approximate the data.

Figure 6.49 summarizes the different cases that may arise depending on whether the target is embedded in background noise or in background noise plus

TABLE 6.12
Comparison of NonGaussian and Gaussian Based Detectors

SCR (dB)	PFA	PD K-Distributed Detector	PD Gaussian Detector
0	10^{-2}	0.50	0.06
−10	10^{-2}	0.40	0.02
−20	10^{-2}	0.22	0.01
0	10^{-5}	0.30	10^{-5}
−10	10^{-5}	0.25	10^{-5}
−20	10^{-5}	0.10	10^{-5}

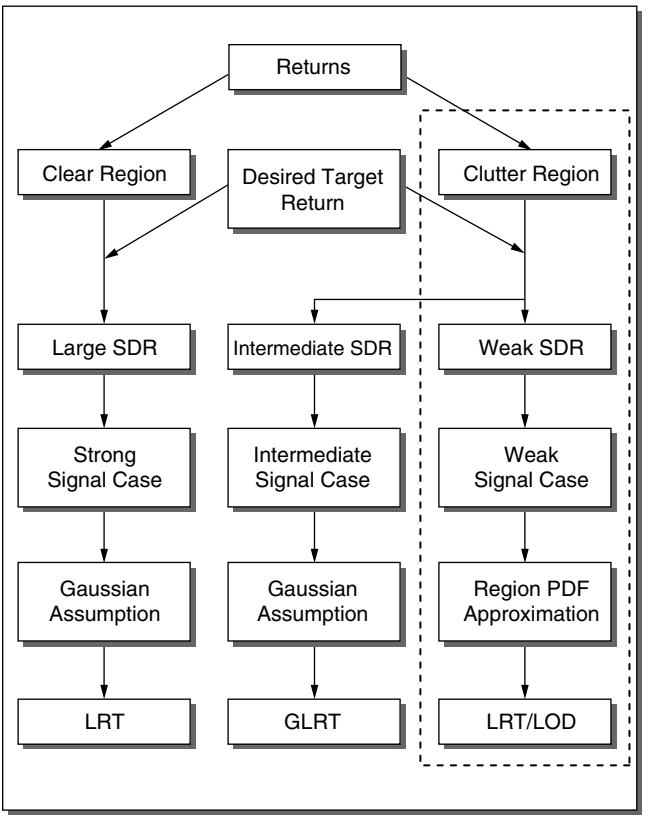

FIGURE 6.49 Different target cases.

disturbance noise such as clutter. This can result either in a strong signal case with a large signal-to-disturbance ratio (SDR), intermediate signal case with a large signal-to-disturbance ratio (SDR), or a weak signal case. Note that the Gaussian assumption is used for the cases where the likelihood-ratio-test (LRT) and generalized likelihood-ratio-test (GLRT)[1] detectors are utilized. For the weak signal case, the PDF of the region has to be approximated in order to enable the use of the appropriate LRT or the suboptimal locally-optimum-detector (LOD).[1] All of this suggests the necessity for continuously monitoring the environment and subdividing the surveillance volume into homogeneous patches in addition to approximating the underlying probability distributions for each patch. This is achieved by the process referred to as the Automated Statistical Characterization And Partitioning of Environments (ASCAPE)[3–5] presented next.

6.5.2. ASCAPE Process

The ASCAPE process, shown in Figure 6.37 of Section 6.4.2, consists of four interactive blocks. The first block is a preprocessing block that performs classical space–time processing on the collected data. Then, based on the mapping procedure, described in Refs. 3 and 4 the second block separates contiguous homogeneous patches and subpatches by segregating between their average power levels. Using the statistical procedure explained in Refs. 4 and 5 the next block goes one step further and separates homogeneous subpatches by segregating between their probabilistic data distributions. Furthermore, this block identifies suitable approximations to the PDF of each homogeneous patch and determines the location of outliers in the scene.[4,5] The statistical procedure uses the Ozturk algorithm, a newly developed algorithm for analyzing random data.[6] The final block indexes the scene under investigation by assigning a set of descriptors to every cell in the scene. For each cell, the indexing is used to indicate to which homogeneous patch the cell belongs, whether it is an edge cell or an outlier, which cells can be chosen as reference cells if the cell is to be tested, and which PDF best approximates the data in the cell. Note that the reference cells are cells that belong to the same homogeneous patch and are closest to the cell to be tested.

The forward and backward interactions between the different blocks are controlled through rules of an expert system shell referred to as Integrated Processing and Understanding of Signals (IPUS) developed jointly by the University of Maussachusetts and Boston University.[7]

When ASCAPE is followed by a detection stage (e.g., target detection in Radar), all information needed by the detector is available for every cell in the scene. Furthermore, given the PDF that can approximate the patch in which the test cell is located, the appropriate detector is readily selected. This is in contrast to the classical detection approach where a single detector (usually the Gaussian detector) is used in processing the entire scene.

6.5.3. Application of ASCAPE to Real IR Data

Consider an image of real IR data collected over lake Michigan. As shown in Figure 6.41 in Section 6.4.6, the scene consists of two major regions: lake and land. Furthermore, the three dimensional plot of the data magnitudes shown in Figure 6.42 indicates that the data in the lake region are regular, whereas the data in the land region are irregular and contain large discretes especially near the boundary close to the lake.

ASCAPE is applied to the IR scene. First, the preprocessing block identifies that the data in the scene is uncorrelated as required by the Ozturk algorithm for the univariate PDF approximation. Then, application of the mapping procedure results in the segmentation of the scene into three different patches and their respective boundaries, as shown in Figure 6.44. Note that in addition to regions 1 and 2, corresponding, respectively, to land and lake, the mapping procedure detects a third region, labeled 3. This region can be observed in the original scene

of Figure 6.41 and is not large in size. Comparison of the mean, variance, and the power of the different regions reveals that patch 3 may be a very small body of water. Edges of the different patches are also detected.

Next, the statistical procedure is applied to every previously declared homogeneous patch in order to separate further between any existing contiguous subpatches that may have similar power levels but different data distributions. As shown in Figure 6.46, the procedure results in the determination of nonGaussian subpatches within each previously declared homogeneous patch and the location of outliers. In addition, for the nonGaussian subpatches, approximate PDFs are determined. Comparing Figures 6.46 and 6.41, note that outliers tend to have a physical significance and might represent string-like patches such as roadways.

The statistical procedure is followed by the indexing stage when every cell in the scene is assigned a set of descriptors to indicate, for each cell, the homogeneous patch to which it belongs, whether it is an edge cell or an outlier, which cells can be chosen as reference cells when parameters in the cell are to be estimated, and which PDF best approximates the data in the cell.

6.5.4. Conclusion

NonGaussian detectors should be utilized in cases that cannot be handled by the classical Gaussian detector (e.g., nonGaussian regions where the clutter and target spectra cannot be separated). In such cases, the appropriate detector must be selected. This is done by using ASCAPE that (1) monitors the environment under investigation, (2) segments the scene into homogeneous patches, and (3) approximates the PDFs of each patch.

Work under investigation and future work include: (1) tailoring/tuning of the edges for subpatches detected by the statistical procedure that is limited by the requirement of 100 reference cells for the Ozturk algorithm, (2) performance improvement of the nonGaussian detectors over the Gaussian detector in different types of environments and under different circumstances, (3) development of more expert system rules to enable ASCAPE to be applied to different types of data (e.g., radar, IR, sonar, medical imaging, etc.), and (4) Application of ASCAPE to medical imaging (e.g., detection of tumors in lung) and to other areas.

6.6. STATISTICAL CHARACTERIZATION OF NONHOMOGENEOUS AND NONSTATIONARY BACKGROUNDS

(A. D. KECKLER, D. L. STADELMAN, D. D. WEINER, AND M. A. SLAMANI)

6.6.1. Introduction

If detection and estimation theory is to be successfully applied to weak signal problems, an adequate statistical characterization of the random background

noise and clutter is required. Backgrounds are commonly assumed to be Gaussian and the literature is replete with detection and estimation algorithms based upon this assumption. Unfortunately, backgrounds encountered in real applications, such as those occurring with tumors embedded in lung tissue, weapons concealed under clothing, and tanks hidden in forests, are likely to be nonhomogeneous, nonstationary, and nonGaussian. Inadequate characterization of the background can lead to severely degraded performance for the weak target problem.

This chapter addresses real-world environments where backgrounds are too complex and unpredictable to be modeled *a priori*. The strategy employed here is to monitor the environment and process the data so as to produce homogeneous partitions, which are statistically characterized in terms of PDFs. For this purpose, a procedure known as ASCAPE (Automated Statistical Characterization and Partitioning of Environments) has been developed.[1,2] ASCAPE identifies partitions in two steps. First, image processing and expert system techniques are used to identify partitions within a scene that can be separated based upon differences in their average intensity levels. Then the Ozturk algorithm,[3] a newly developed algorithm for analyzing random data, is used to divide the partitions into subpatches that can be separated based upon differences in their underlying probability distributions. The ASCAPE procedure is illustrated in Section 6.6.2 in the context of concealed weapon detection.

The Ozturk algorithm, as originally designed, can only approximate the PDF of univariate random data. However, in many applications, such as a coherent radar that jointly processes N looks at a target, it is necessary to characterize the random background in terms of an N-dimensional PDF. Spherically invariant random vectors (SIRVs) are introduced in Section 6.6.3 as a useful approach for modeling correlated, nonGaussian multivariate data. Section 6.6.4 describes the univariate Ozturk algorithm. In Section 6.6.5, the Ozturk algorithm is extended to handle correlated, multivariate random variables classified as SIRVs. Finally, application of the multivariate Ozturk algorithm to a weak signal nonGaussian detection problem is discussed in Section 6.6.6.

6.6.2. Application of ASCAPE to Concealed Weapon Detection

The capabilities of ASCAPE are demonstrated in this section. Consider the scene of Figure 6.50, which shows a person carrying a concealed weapon located on the right rib cage. The scene consists of real millimeter wave data collected from a Millitech Corporation MMW sensor. Figure 6.51 shows the result when the Sobel operator, a well known edge detection algorithm, is applied to the scene in an attempt to detect the weapon. Clearly, the Sobel operator is unable to detect the weapon's edges. This approach fails because the average intensity of the region surrounding the weapon differs only slightly from that of the region where the weapon is located. The image processing and expert system techniques of ASCAPE are then applied to the data of Figure 6.50. The ASCAPE result is shown in Figure 6.52. Even though the differences in average intensities are

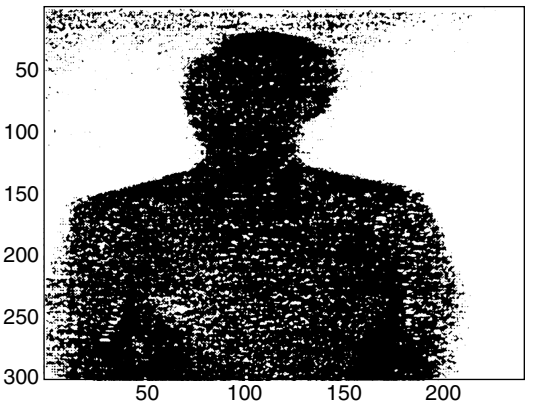

FIGURE 6.50 Original scene.

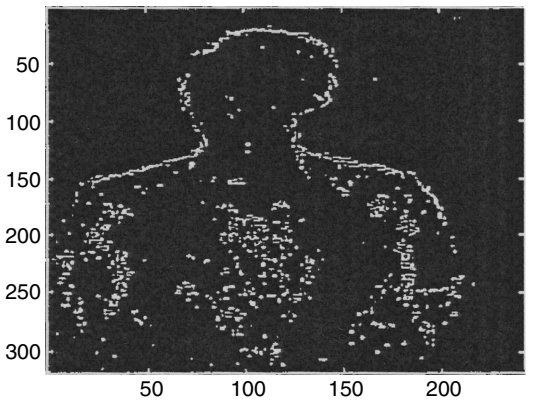

FIGURE 6.51 Results of the sobel operator.

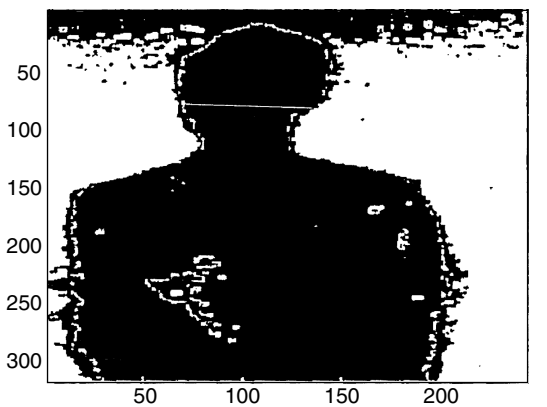

FIGURE 6.52 Composite image.

FIGURE 6.53 Simulated scene.

small, ASCAPE is still able to detect the concealed weapon based on relative average intensities.

Use of the Ozturk algorithm for dividing a partition into subpatches is now illustrated with simulated data. The simulated scene is shown in Figure 6.53 consisting of two regions, denoted as A and B. Region A represents the background and contains 10,000 data samples generated from a Weibull probability distribution with zero mean, unit variance, and a shape parameter value of 0.6. Region B represents a handgun and contains 10,000 data samples generated from a Gaussian distribution with zero mean and unit variance. As a result, regions A and B have the same average intensity. Thus, the complete scene of Figure 6.53 simulates a partition as would be obtained by the first stage of ASCAPE. The histograms of regions A and B are shown in Figure 6.55, along with the combined

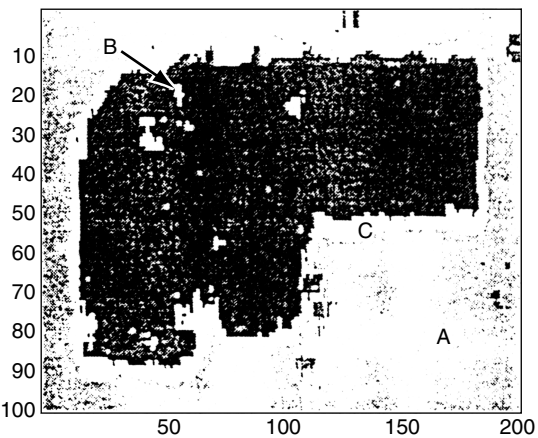

FIGURE 6.54 Mapped scene.

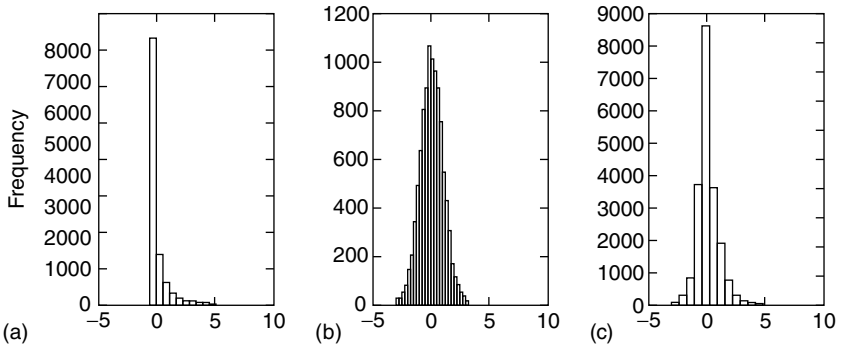

FIGURE 6.55 Histograms. (a) Region A. (b) Region B. (c) Combined data from both region A and region B.

histogram of the entire scene. Examination of the combined histogram in no way suggests the presence of two regions in Figure 6.53.

The Ozturk algorithm is then applied to the data in Figure 6.53. The algorithm uses 100 pixels at a time to statistically characterize the data in terms of PDF approximations. The image is then partitioned into a number of subpatches determined by differences in their probability distributions. The result is shown in Figure 6.54 and clearly indicates that the Ozturk algorithm identifies three regions:

(1) Region A, the background
(2) Region B, the gun
(3) Region C, the edges of the gun.

Additional details concerning the Ozturk algorithm are presented in Section 6.6.4.

6.6.3. The SIRV Radar Clutter Model

Conventional radar receivers are based on the assumption of Gaussian distributed clutter. However, the Weibull and K-distribution are shown to approximate the envelope of some experimentally measured nonGaussian clutter data.[4–8] The detection performance of the Gaussian receiver in this environment is significantly below that of the optimum nonGaussian receiver, especially for weak target returns.

NonGaussian clutter is often observed to be "spiky," as illustrated in Figure 6.56. In such cases, the threshold of the conventional Gaussian receiver must be raised in order to maintain the desired false alarm rate. This results in a reduction of the probability of detection. In contrast, nonGaussian receivers contain nonlinearities that limit large clutter spikes and allow a lower threshold to be used, which improves performance for targets with a low signal-to-clutter

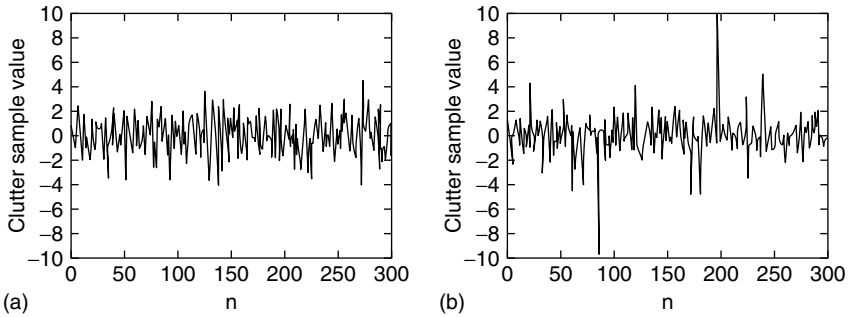

FIGURE 6.56 Comparison of Gaussian data with NonGaussian data of equal variance. (a) Gaussian example. (b) NonGaussian example.

ratio (SCR). Determination of these nonGaussian receivers requires specification of suitable PDFs for the clutter.

The nonhomogeneous and nonstationary clutter environment must be monitored to adapt detection algorithms over the surveillance volume. This is complicated by the need for an efficient technique to accurately approximate a joint clutter PDF that incorporates the pulse-to-pulse correlation. Spherically invariant random vectors (SIRVs), which are used in this chapter, have been shown to be useful for modeling correlated nonGaussian clutter.[9] The class includes many distributions of interest, such as the Gaussian, Weibull, Rician, and K-distributed, among others.[9–12]

A random vector $\mathbf{Y}$ of dimension N and realization $\mathbf{y}$ is defined to be an SIRV if and only if its PDF has the form,

$$f_{\mathbf{Y}}(\mathbf{y}) = (2\pi)^{-\frac{N}{2}} |\Sigma|^{-\frac{1}{2}} h_N(q(\mathbf{y})) \tag{6.42}$$

where Σ is an $N \times N$ nonnegative definite matrix, $q(\mathbf{y})$ is the quadratic form defined by

$$q = q(\mathbf{y}) = (\mathbf{y} - \mathbf{b})^{\mathrm{T}} \Sigma^{-1} (\mathbf{y} - \mathbf{b}) \tag{6.43}$$

$\mathbf{b}$ is the $N \times 1$ mean vector, and $h_{\mathrm{N}}(\cdot)$ is a positive, monotonic decreasing function for all N.[13] Equivalently, an SIRV $\mathbf{Y}$ can be represented by the linear transformation,

$$\mathbf{Y} = A\mathbf{X} + \mathbf{b} \tag{6.44}$$

where $\mathbf{X}$ is a zero-mean SIRV with uncorrelated components represented by

$$\mathbf{X} = S\mathbf{Z} \tag{6.45}$$

$\mathbf{Z}$ is zero-mean Gaussian random vector with independent components, and S is a nonnegative random variable independent of $\mathbf{Z}$. The PDF $f_{\mathrm{S}}(s)$ uniquely determines the type of SIRV and is known as the characteristic PDF of $\mathbf{Y}$.

Since the matrix A is specified independenty of $f_S(s)$, an arbitrary covariance matrix, $\Sigma = AA^T$, can be introduced without altering the type of SIRV.

This representation is used to obtain

$$h_N(q) = \int_0^\infty s^{-N} e^{-\frac{q}{2s^2}} f_s(s)\mathrm{d}s \tag{6.46}$$

and subsequently, the PDF of the quadratic form is

$$f_Q(q) = \frac{1}{2^{\frac{N}{2}}\Gamma\left(\frac{N}{2}\right)} q^{\frac{N}{2}-1} h_N(q) \tag{6.47}$$

Since $h_N(q)$ uniquely determines each type of SIRV, Equation 6.47 indicates that the multivariate approximation problem is reduced to an equivalent univariate problem.

It is not always possible to obtain the characteristic PDF, $f_S(s)$, in closed form. However, an N-dimensional SIRV with uncorrelated elements can be expressed in random variable generalized spherical coordinates R, θ, and ϕ_k for $k = 1, \ldots, N-2$, where the PDF of R is given by

$$f_R(\mathrm{r}) = \frac{\mathrm{r}^{N-1}}{2^{\frac{N}{2}-1}\Gamma\left(\frac{N}{2}\right)} h_N(\mathrm{r}^2) u(\mathrm{r}) \tag{6.48}$$

The angles θ and ϕ_k are statistically independent of the envelope R and do not vary with the type of SIRV. When $f_S(s)$ is unknown, Equation 6.48 is used both to generate SIRVs and to determine $h_N(q)$.[9]

It is desirable to develop a library of SIRVs for use in approximating unknown clutter returns. Table 6.13 contains the characteristic PDFs and $h_N(q)$s of some SIRVs for which analytical expressions are known. For simplicity, the results presented for the Weibull and Chi SIRV are valid only for even N. Additional SIRVs, such as the generalized Rayleigh, generalized Gamma, and Rician, are developed in Ref. 9.

The discrete Gaussian mixture SIRV in Table 6.13 is of special interest. Its PDF is a simple finite weighted sum of Gaussian PDFs. This is useful for approximating many other SIRVs, as well as generating unique distributions.

6.6.4. Distribution Approximation Using the Ozturk Algorithm

It is important to suitably model the clutter PDF to obtain improved detection performance of weak signals in nonGaussian clutter. Ozturk developed a general graphical method for testing whether random samples are statistically consistent with a specified univariate distribution.[3] The Ozturk algorithm is based upon sample order statistics and has two modes of operation. The first mode consists of the goodness-of-fit test. The second mode of the algorithm approximates the PDF of the underlying data by using a test statistic generated from the goodness-of-fit test to select from a library of known PDFs.

TABLE 6.13
Characteristic PDFs and $h_N(Q)$ Functions for Known SIRVs

Marginal PDF	Characteristic PDF $f_S(s)$	$h_N(q)$
Gaussian	$\delta(s-1)$	$e^{-\frac{q}{2}}$
Student-t	$\frac{2b}{\Gamma(\nu)2^{\nu}} b^{2\nu-1} s^{-(2\nu+1)} e^{-\frac{b^2}{2s^2}} u(s)$	$\frac{2^{\frac{N}{2}} b^{2\nu} \Gamma\left(\nu+\frac{N}{2}\right)}{\Gamma(\nu)(b^2+q)^{\frac{N}{2}+\nu}}$
Laplace	$b^2 s e^{-\frac{b^2 s^2}{2}} u(s)$	$b^N (b\sqrt{q})^{1-(N/2)} K_{\frac{N}{2}-1}(b\sqrt{q})$
K-distributed Envelope	$\frac{2b}{2^{\alpha}\Gamma(\alpha)} (bs)^{2\alpha-1} e^{-\frac{b^2 s^2}{2}} u(s)$	$\frac{b^N}{\Gamma(\alpha)} \frac{(b\sqrt{q})^{\alpha-\frac{N}{2}}}{2^{\alpha-1}} K_{\frac{N}{2}-\alpha}(b\sqrt{q})$
Cauchy	$\sqrt{\frac{2}{\pi}} b S^{-2} e^{-\frac{b^2}{2s^2}} u(s)$	$\frac{2^{\frac{N}{2}} b \Gamma\left(\frac{N}{2}+\frac{1}{2}\right)}{\sqrt{\pi}(b^2+q)^{\frac{N}{2}+\frac{1}{2}}}$
Chi Envelope $\nu \leq 1$	$\frac{2^{\nu+1} b^{2\nu}}{\Gamma(\nu)\Gamma(1-\nu)} \frac{s^{2\nu-1}}{(1-2b^2 s^2)^{\nu}} u(s) u\left(\frac{1}{b\sqrt{2}} - s\right)$	$\frac{(2)^{\frac{N}{-2}} b^{2\nu}}{\Gamma(\upsilon)} \sum_{k=1}^{\frac{N}{2}} \binom{\frac{N}{2}-1}{k-1} \mathrm{q}^{\nu-k} b^{N-2k} \frac{\Gamma(k-\nu)}{\Gamma(1-\nu)} e^{-b^2 q}$
Gaussian Mixture $w_k > 0, \sum_{k=1}^{K} w_k = 1$	$\sum_{k=1}^{K} w_k \delta(s - s_{\mathrm{k}})$	$\sum_{k=1}^{K} w_k s_k^{-N} e^{-\frac{q}{2s_k^2}}$
Weibull Envelope $0 < b < 2$		$(-2)^{\frac{N}{2}} e^{-a\sigma^b} q^{\frac{b}{2}} \sum_{k=1}^{N} B_k \frac{(a\sigma^b)^k}{k!} q^{\frac{kb}{2}-\frac{N}{2}}$, $B_k = \sum_{m=1}^{k} (-1)^m \binom{k}{m} \prod_{i=0}^{M-1} \left(\frac{mb}{2} - i\right)$

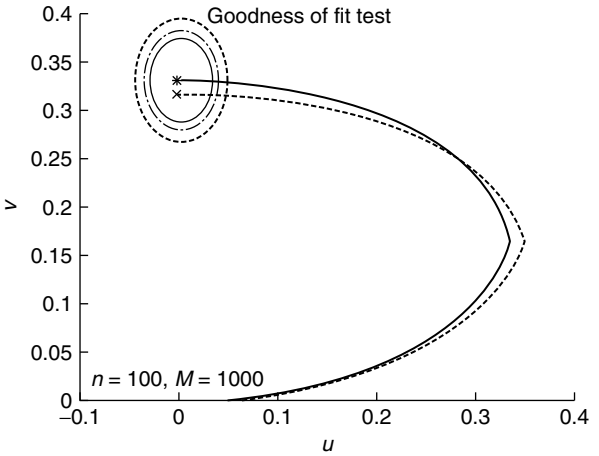

FIGURE 6.57 Linked vectors and 90%, 95%, and 99% confidence intervals for the standard Gaussian distribution.

The goodness-of-fit test is illustrated in Figure 6.57. The solid curve denotes the ideal locus of the null distribution obtained by averaging 1000 Monte Carlo simulations of 100 data samples, where the Gaussian distribution is chosen as the null distribution. The 90%, 95%, and 99% confidence contours are shown. The

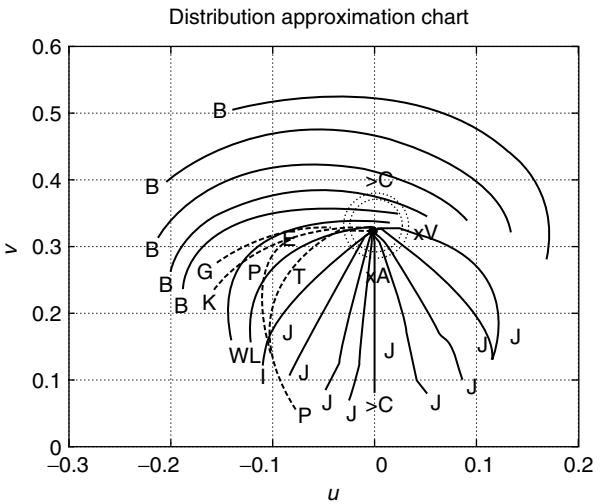

FIGURE 6.58 Ozturk approximation chart for univariate distributions (B = Beta, G = Gamma, K = K-distribution, P = Pareto, L = Lognormal, T = Gumbel, E = Exponential, V = Extreme Value, A = Laplace, S = Logistic, U = Uniform, N = Normal, W = Weibull, C = Cauchy, J = SU Johnson). • End point of 100 data samples corresponding to that of Figure 6.57.

dashed curve shows the locus of test data, which is accepted as being Gaussian distributed with significance 0.1.

An approximation to an unknown distribution is obtained by examining the location of the end point coordinate. An approximation chart is constructed for a library of PDFs by plotting the end point coordinates for each density in the library. A distribution that does not depend upon a shape parameter will appear as a single point on the approximation chart. Distributions that have a single shape parameter, such as the Weibull or K-distributions, will appear as trajectories. Distributions with more than one shape parameter are represented by a family of trajectories. A sample approximation chart for univariate distributions is shown in Figure 6.58 for 100 data samples and 1000 Monte Carlo simulations.

6.6.5. Approximation of SIRVs

The distribution approximation technique described above applies to univariate distributions. It is seen from Equation 6.42 and Equation 6.46 that the characteristic PDF of an SIRV is invariant with respect to the vector dimension N and uniquely determines the SIRV. If the data can be appropriately modeled as an SIRV, then the marginal distribution can be used to uniquely distinguish it from all other SIRVs. Since the marginal distribution of an SIRV is univariate, the procedure discussed in Section 6.6.4 can be applied directly. However, knowledge of the marginal distribution is insufficient to ensure that multivariate data can be modeled as an SIRV.

Multivariate sample data can be rejected as having a particular type of SIRV density if the envelope distribution is not supported by the Ozturk algorithm. In addition, the angle distributions must be checked for consistency. However, the angle distributions are independent of the type of SIRV considered and are useful only for verifying that sample data is not SIRV distributed.

The approximation problem is further complicated since the covariance matrix of the underlying SIRV distribution is usually unknown. The maximum likelihood (ML) estimate of the covariance matrix for a known zero-mean SIRV is given by

$$\hat{\sum}_y = \frac{1}{K}\sum_{k=1}^{K} \frac{h_{N+2}\left(\mathrm{y}_k^{\mathrm{T}}\hat{\Sigma}_y\mathrm{y}_k\right)}{h_N(\mathrm{y}_k^{\mathrm{T}}\hat{\Sigma}_y\mathrm{y}_k)}\mathrm{y}_k\mathrm{y}_k^{\mathrm{T}} \tag{6.49}$$

since Equation 6.49 depends upon $h_N(q)$, the ML estimate of the covariance matrix cannot be used in the approximation problem. Alternatively, a statistic formed using the well known sample covariance matrix is used in this chapter to select the appropriate approximation for the clutter distribution. This statistic is given by

$$\hat{\mathrm{r}} = [(\mathbf{y} - \hat{\mathbf{b}}_y)^{\mathrm{T}}\hat{s}_y^{-1}(\mathbf{y} - \hat{\mathbf{b}}_y)]^{\frac{1}{2}} \tag{6.50}$$

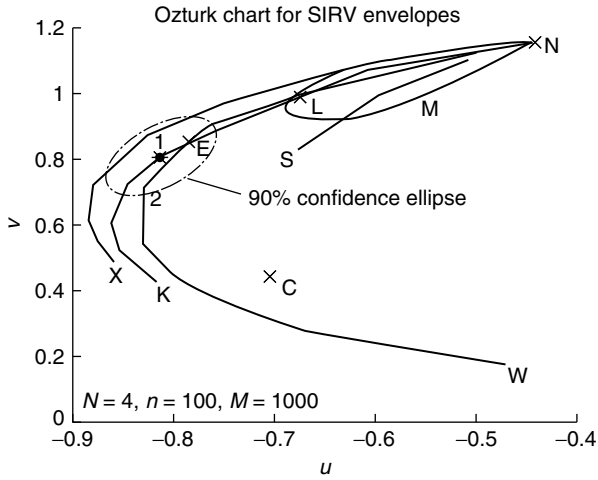

FIGURE 6.59 Approximation chart for SIRV envelope statistic, N = 4 (C = Cauchy, X = Chi envelope, E = Exponential envelope, K = K-distributed envelope, M = Discrete Gaussian Mixture, N = Normal, S = Student-*t*, W = Weibull, L = Laplace).

where $\hat{s}_y$ is the sample covariance matrix, given by

$$\hat{s}_y = \frac{1}{n-1}\sum_{k=1}^{n}(\mathbf{y}_k - \hat{\mathbf{b}}_y)(\mathbf{y}_k - \hat{\mathbf{b}}_y)^{\mathrm{T}} \tag{6.51}$$

and $\hat{\mathbf{b}}_y$ is the sample mean. Approximation charts using the envelope statistic $\hat{\mathrm{R}}$ of Equation 6.50 are shown in Figures 6.59 and 6.60 for vector dimensions $N = 2$ and

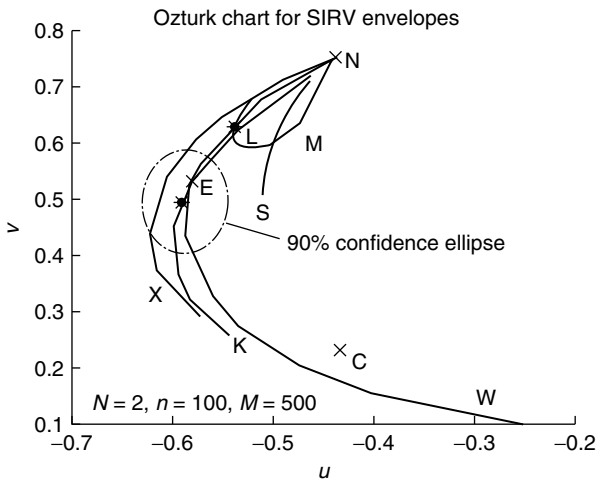

FIGURE 6.60 Approximation chart for SIRV envelope statistic, N = 2 (C = Cauchy, X = Chi envelope, E = Exponential envelope, K = K-distributed envelope, M = Discrete Gaussian Mixture, N = Normal, S = Student-t, W = Weibull, L = Laplace).

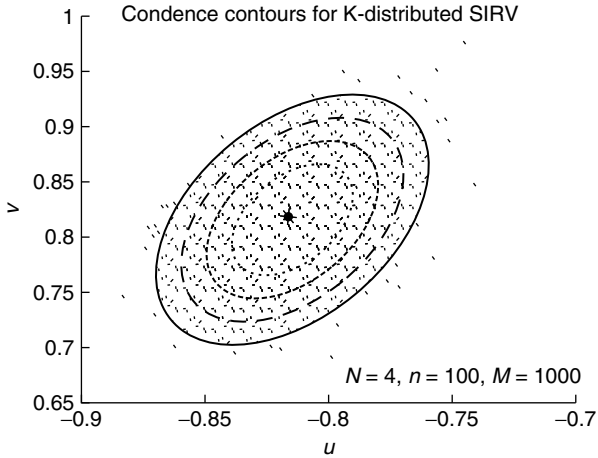

FIGURE 6.61 The 50, 70, 80, and 90% confidence contours for the K-distributed SIRV envelope.

$N = 4$, respectively. The 90% confidence contours for the K-distribution with shape parameter $\alpha_k = 0.4$ are shown on the charts. Surprisingly, the size of the confidence intervals does not significantly increase as the dimension of the SIRV increases. While the sample covariance matrix of Equation 6.51 may be a poor estimate of the actual matrix, the statistic of Equation 6.50 appears to be insensitive to this estimation. Figure 6.61 shows the scatter of locus end points for 1000 simulations of K-distributed data. Each end point is obtained from 100 vectors of four components.

As seen in Figures 6.59 and 6.60, the confidence contours overlap several trajectories on the approximation charts. Therefore, it is possible that any one of

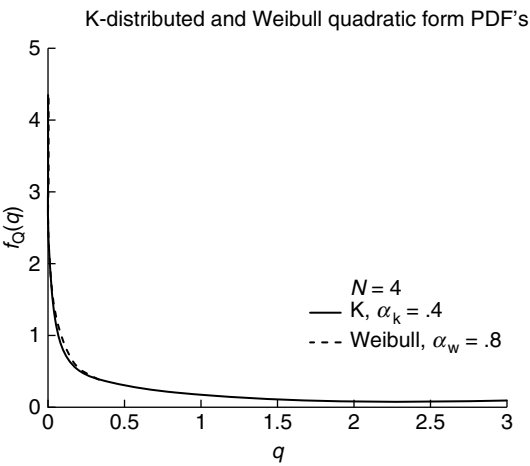

FIGURE 6.62 Comparison of Weibull and K-distributed quadratic form PDFs.

several different types of SIRV distributions may be selected by the Ozturk algorithm to approximate an SIRV distributed sample. Figure 6.62 compares the quadratic from PDF for two distributions that fall within the confidence contour shown in Figure 6.59. The locus end point of a K-distributed SIRV with shape parameter $\alpha_k = 0.4$ is marked with a "1" in Figure 6.59. The locus end point of a Weibull SIRV with shape parameter $\alpha_w = 0.8$ is labeled with a "2". The close fit between these PDFs, even when their locus end points are separated within the confidence contour, suggests similar distributions fall within a particular localized region of the Ozturk chart. Consequently, distributions whose locus end points are contained within a confidence contour are expected to be suitable approximations.

6.6.6. NonGaussian Receiver Performance

The performance of an adaptive detection scheme using the Ozturk PDF approximation algorithm to regularly update the choice of receiver is evaluated by simulating SIRV clutter. The clutter power is assumed to be much greater than the background noise power for the weak signal problem. Consequently, only the clutter PDF is used to model the total disturbance. The clutter is also assumed to have zero mean and a known covariance matrix, $\tilde{\Sigma}$. The amplitude of the desired signal is modeled as an unknown complex random variable, which is constant over each P-pulse coherent processing interval. The phase of the complex amplitude is assumed to have a $U(0, 2\pi)$ PDF. Thus, the form of the ML estimate for the complex amplitude is the same for all SIRVs, and the generalized likelihood ratio test (GLRT) is[11]

$$T_{\mathrm{GLRT}}(\tilde{\mathbf{r}}) = \frac{h_{2P}\left(2\left(\tilde{\mathbf{r}}^{\mathrm{H}}\tilde{\Sigma}^{-1}\tilde{\mathbf{r}} - \dfrac{|\tilde{\mathbf{s}}^{H}\tilde{\Sigma}^{-1}\tilde{\mathbf{r}}|^2}{\tilde{\mathbf{s}}^{\mathrm{H}}\tilde{\Sigma}^{-1}\tilde{\mathbf{s}}}\right)\right)}{h_{2\mathrm{P}}(2\tilde{\mathbf{r}}^{\mathrm{H}}\tilde{\Sigma}^{-1}\tilde{\mathbf{r}})} \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \eta \tag{6.52}$$

where examples of $h_{2\mathrm{P}}(\cdot)$ are listed in Table 6.13. The GLRT of Equation 6.52 is formulated in terms of the complex lowpass elopes of the receive data, $\tilde{\mathbf{r}}$, and known signal pattern, $\tilde{\mathbf{s}}$. Previous investigation has shown there is little or no degradation in performance of the GLRT for the known covariance problem, when compared to the Neyman–Pearson test.[8,14]

Figure 6.63 compares the two-pulse detection performance of the adaptive Ozturk-based receiver to several other receivers in K-distributed clutter. The magnitude of the complex target amplitude is assumed to be Swerling I. The shape parameter of the clutter is chosen as $\alpha_k = 0.4$, which is within the range of values measured for real data.[4] The performance is evaluated for an identity covariance matrix, and may be interpreted as a function of the SCR at the output of a prewhitening filter when the clutter samples are correlated. Detection results are obtained by processing 100,000 vector samples of K-distributed clutter. The solid curve shows the baseline detection performance of the K-distributed GLRT designed for 0.001 probability of false alarm (P_{FA}). The adaptive receiver

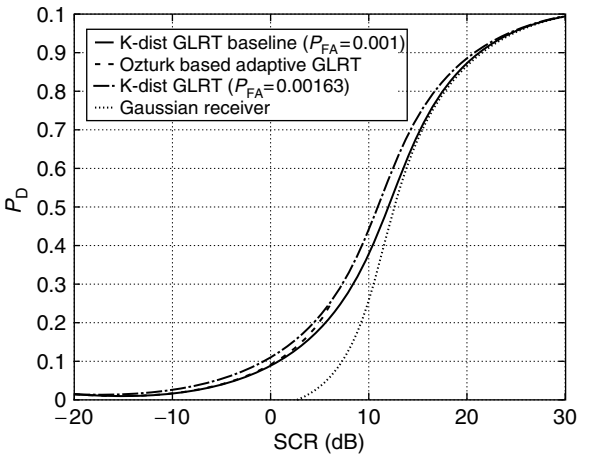

FIGURE 6.63 Adaptive Ozturk-Based receiver comparison for a swerling 1 target magnitude with U(0, 2π) phase.

performance, also indicated in Figure 6.63, is obtained by partitioning the data into 50 intervals of 2000 samples each. The first 100 samples of each interval are processed with the Ozturk algorithm to obtain the data end points marked with a "+" in Figure 6.64. For each data end point, the corresponding 2000 sample interval is processed by a GLRT designed from the PDF associated with the closest library end point. While the known covariance matrix is used in the GLRT implementation, the sample covariance matrix for each 100 samples is used in the Ozturk algorithm, as described in Sections 6.6.4 and 6.6.5.

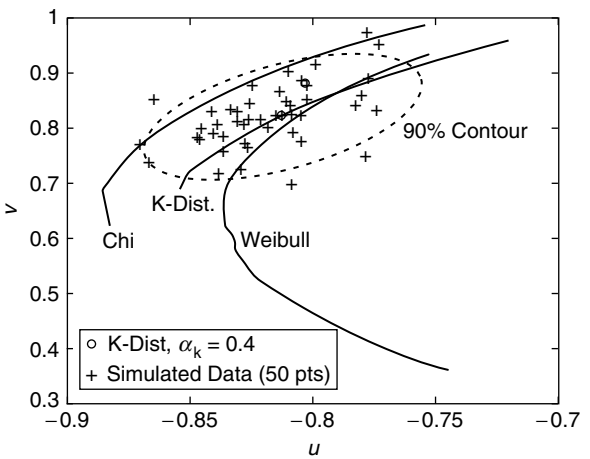

FIGURE 6.64 Ozturk algorithm end point scatter diagram for K-distributed data with $\alpha_k = 0.4$.

Performance of the adaptive receiver closely matches the baseline performance and shows significant improvement over the Gaussian receiver for SCR values below 10 dB. The measured P_{FA} for the adaptive receiver is 0.00163, which is slightly above the design value. This explains why the probability of detection (P_D) for the adaptive receiver exceeds that of the baseline receiver at large SCR values. Baseline K-distributed receiver performance for this higher measured P_{FA} is also included in Figure 6.63 for comparison.

The adaptive receiver processed data associated with all the end points is shown in the scatter diagram of Figure 6.64, including those that fell outside the 90% confidence contour. Nonetheless, the localized P_D for each interval did not vary significantly from the average value reported in Figure 6.63 for a given SCR.

6.6.7. Concluding Remarks

This work provides significant contributions to the partitioning and statistical characterization of complex real-world environments. ASCAPE is shown to be an effective tool for this purpose. New results are presented that allow the Ozturk algorithm to adequately approximate multivariate SIRV PDFs from only 100 sample clutter vectors. A simple radar example is presented for K-distributed clutter with known covariance matrix and 0.001 probability of false alarm. A receiver that adaptively processes the data based on the Ozturk PDF approximation has near optimum performance for this example, thus, demonstrating the successful application of the Ozturk algorithm to weak signal detection. Furthermore, the adaptive receiver has significantly better detection performance than the Gaussian receiver at low SCRs, with only a slight increase in the P_{FA}.

6.7. KNOWLEDGE-BASED MAP SPACE TIME ADAPTIVE PROCESSING (KBMapSTAP)

(C. T. CAPRARO, G. T. CAPRARO, D. D. WEINER, AND M. C. WICKS)

6.7.1. Introduction

Space-time adaptive processing (STAP) is viewed as a potentially effective means for suppressing ground clutter received by an airborne radar. However, a serious problem with any STAP approach involves the accurate estimation of unknown clutter statistics. This problem is further complicated by the fact that airborne radars are likely to encounter nonhomogeneous clutter environments. Previous efforts[1,2] have recognized this problem and have shown the benefits of using *a priori* data to increase performance in nonhomogeneous clutter environments for Constant False Alarm Rate (CFAR) processing and preadaptive filtering with STAP.

This work documents the results of our effort to develop, implement, and test a computer-based algorithm to utilize *a priori* terrain data in order to improve target detection. Our approach was to leverage existing terrain datasets to help selectively choose secondary data for estimating the clutter covariance matrix needed for post-Doppler radar processing. In so doing we will show that performance can be improved. This use of terrain data provides insight into how to build one aspect of the next generation signal processing algorithm and to possibly extend its use to other areas such as tracking and identification.

Section 6.7.2 of the work provides a description of our clutter model. Section 6.7.3 discusses the difficulty in choosing secondary data for the estimation of a clutter covariance matrix in a nonhomogeneous environment and an approach for easing this difficulty with adaptive post-Doppler processing. Section 6.7.4 departs from theory-based discussion and presents a brief description of an airborne radar measurement program used in testing our methodology. Section 6.7.5 describes our *a priori* data approach to estimate the clutter covariance matrix in nonhomogeneous environments. Section 6.7.6 presents our results and Section 6.7.7 presents our conclusions and recommended future work.

6.7.2. Clutter Model

Ward's clutter model[3] is employed to determine whether or not available secondary data may be useful in estimating the clutter covariance matrix of a test cell. Ward approximates a continuous field of clutter by modeling the clutter return from each range ring as the superposition of a large number of independent point scatters or clutter patches evenly distributed in azimuth about the radar. For simplicity, we assume unambiguous range. Then the clutter return at any instant is from a single range ring.

If we divide the range ring into a total of N_c clutter patches, each patch has an angular extent given by $\Delta\theta = 2\pi/N_c$. The response in the nth channel, due to the mth pulse, in the lth range ring, after summing over all k patches is

$$X_{nm\ell} = \sum_{k=1}^{N_c} \alpha_{\ell k} e^{j2\pi(m\bar{\omega}_{\ell k} + n\nu_{\ell k})} \tag{6.53}$$

where: $\bar{\omega}_{\ell k}$ is the normalized Doppler frequency, $\nu_{\ell k}$ is the normalized spatial frequency, and $\alpha_{\ell k}$ is the complex received signal amplitude. From this equation the clutter covariance matrix for the lth range ring can be expressed as

$$\mathrm{M}_\ell = \sum_{k=1}^{N_c} E[|\alpha_{\ell k}|^2] \underline{v}_{\ell k} \underline{v}_{\ell k}^{\mathrm{H}} \tag{6.54}$$

where $E[|\alpha_{\ell k}|^2]$ is the estimation of the mean-square value of the complex amplitude magnitude for each of the N_c clutter patches in the range ring and $v_{\ell k}$ is the space–time steering vector. Since the space–time steering vector can be

specified *a priori* the estimation of the clutter covariance matrix reduces to the estimation of $E[|\alpha_{\ell k}|^2]$. Therefore, it is important to have a good method for estimating this value by properly choosing representative clutter data.

6.7.3. Representative Secondary Clutter

Assume the test cell where a target is to be detected is located in the lth range ring. Since M_1, the clutter covariance matrix of the lth range ring is unknown, the objective is to select secondary data from other range rings in order to estimate M_1. Suppose attention is focused on the (l')th range ring where $l' \neq 1$. The question that arises is, "Is the clutter in the (l')th range ring representative of the clutter in the lth range ring?"

This is true provided each clutter patch in the lth range ring having a specific mean-square complex amplitude magnitude and a specific pair of normalized Doppler and spatial frequencies has a corresponding clutter patch in the (l')th range ring having approximately the same mean-square complex amplitude and approximately the same normalized Doppler and spatial frequencies.

Even though the pairs of normalized Doppler and spatial frequencies remain invariant from one range ring to another, it is unlikely in a nonhomogeneous clutter environment that $E[|\alpha_{l'k'}|^2] = E[|\alpha_{lk}|^2]$ for all N_c pairs of clutter patches in the two range rings. In fact, unless the clutter is entirely homogeneous throughout both range rings, it is unlikely that the clutter in the (l')th range ring will be representative of the clutter in the lth range ring over the entire clutter ridge.

However, the concept of representative secondary clutter data may be meaningful on a selective basis. For example, consider postDoppler adaptive beamforming where nonadaptive Doppler filtering is first performed separately on the M pulses from each array element. In effect, this produces at each array element the output of M Doppler filters that subdivide the normalized Doppler frequency interval into M contiguous Doppler bins. The basic idea is that a Doppler filter, with the capability for very low Doppler sidelobes, rejects the clutter whose Doppler frequencies fall outside of its passband. In this way, the residual clutter along the clutter ridge is localized in terms of its normalized spatial frequencies. Adaptive spatial filtering is subsequently performed to reduce the residual clutter. This is repeated for each of the M Doppler filters. Because the residual clutter in normalized Doppler and spatial frequencies is confined to a localized region along the clutter ridge, it is no longer necessary that the range ring from where secondary data is being collected be equivalent in its entirely to the range ring in which the test cell is located. Now the clutter in only a few patches of each range ring need be equivalent, i.e., those that lie along the same iso-Doppler ridge.

6.7.4. Airborne Radar Data

To assist us in building and testing our methodology for selecting equivalent range rings we used data gathered under a U.S. Air Force program. The AFRL Sensors Directorate Multichannel Airborne Radar Measurements (MCARM)

program was designed to collect multichannel clutter data from an airborne platform with a side looking radar.[4] Northrop Grumman collected MCARM data during flights over the Delmarva Peninsula and the east coast of the United States. A Northrop Grumman owned BAC 1-11 was used as a platform for the L-Band radar data collection system. The radar consisted of 32 subapertures combined into 22 adaptive channel elements. The elements were arranged in a 2×11 array. Data was collected at a variety of pulse repetition frequencies (PRFs) over various terrain including mountains, rural, urban, and land/sea interfaces. There were a total of eleven flights with more than 50 Gb of data collected and additional flights planned. We chose this data because of its varied and heterogeneous clutter environment.

6.7.5. *A PRIORI* DATA

Digital terrain data was obtained from the United States Geological Survey (USGS) to classify the ground environment that the MCARM radar was irradiating. Since the Delmarva Peninsula has little variation in elevation we decided not to incorporate digital elevation data that would provide a measure of the angular reflection back to the antenna. Instead, we chose Land Use and Land Cover (LULC) data that classifies terrain using a grid of 200 by 200 meter cells, and codes that describe the terrain in each cell. There are nine major codes and 38 minor codes that have a more detailed description. The LULC data provides a measure of the amount of radar reflection and absorption from the ground. In order to simplify our approach we only used the major codes and, if deemed necessary, planned on using the minor codes later. An example of LULC major codes are: Urban Areas, Agricultural Land, Water, etc.

6.7.6. RESEARCH PROBLEM, HYPOTHESIS, AND PRELIMINARY FINDINGS

Can postDoppler STAP performance be improved by choosing secondary data based upon *a priori* map data? To determine the answer to this question we compared our results with what we call the standard algorithm or sliding window algorithm. The sliding window algorithm chooses $N/2$ range rings above and below the test ring minus two guard rings, where N is twice the number of independent channels of the MCARM radar, which is 22 (see Figure 6.65). The sliding window algorithm has an implicit assumption that the range rings near the test ring are homogeneous and are representative of the test ring. Our algorithm chooses secondary data by comparing the LULC codes of the Doppler patch that interferes with the test patch in the same range ring and all of the patches that lie on the same iso-Doppler curve of interest. Our assumption is that the major interferer after range and Doppler filtering will be the clutter due to the ground within the same range ring as the test cell.

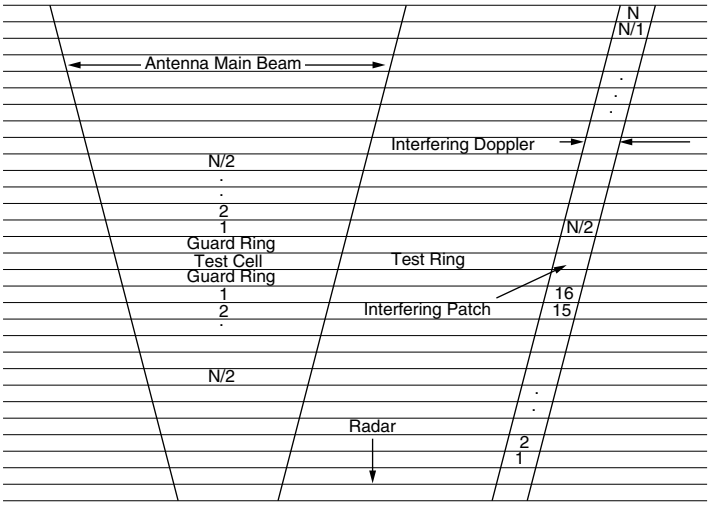

FIGURE 6.65 Sliding window and KBMapSTAP secondary data selection.

It was our hypothesis that our algorithm (KBMapSTAP) would do as good as the sliding window algorithm where the test and surrounding area are homogeneous and KBMapSTAP would do better than the sliding window algorithm for areas where the ground is heterogeneous. To test our hypothesis we injected a target at different range rings with the same radial velocity and power. The only difference in the implementation of the two algorithms was the choice of the secondary range rings. After Doppler processing, we calculated a Modified

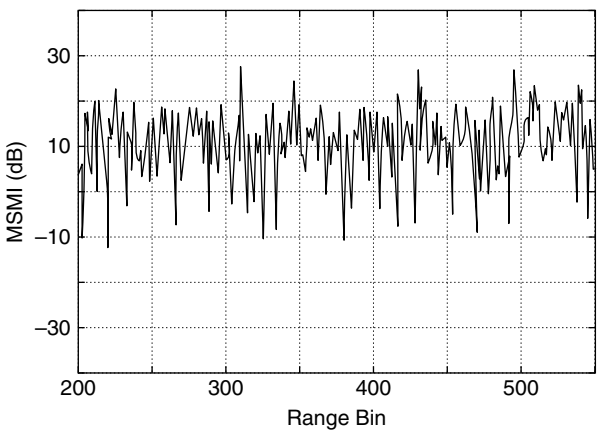

FIGURE 6.66 Sliding window — no injected target (mean MSMI = 12.9, variance MSMI = 28.9).

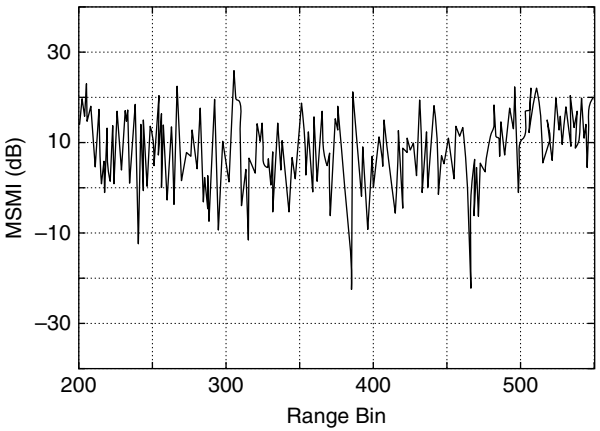

FIGURE 6.67 KBMapSTAP — no injected target (mean MSMI = 11.7, variance MSMI = 26.6).

Sample Matrix Inversion (MSMI) statistic for each range ring of interest.[1,5]

$$\text{MSMI}_i = \frac{|s^{\text{H}}\hat{R}^{-1}x_i|^2}{s^{\text{H}}\hat{R}^{-1}s} \tag{6.55}$$

where s is the space–time steering vector, $\hat{R}$ is the estimate of the clutter covariance matrix, and x_i is the radar return vector for the ith range ring. It can be seen that the MSMI has a thresholding or detection quality similar to a constant false alarm rate (CFAR) property. That is, a MSMI threshold can be chosen y such that those radar returns, x_i, having an MSMI that exceeds y may be considered as potential targets.

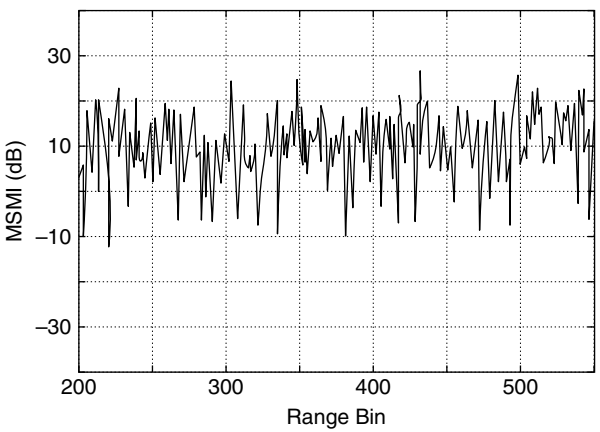

FIGURE 6.68 Sliding window — target injected at range bin 296 (mean MSMI = 12.7, variance MSMI = 28.3).

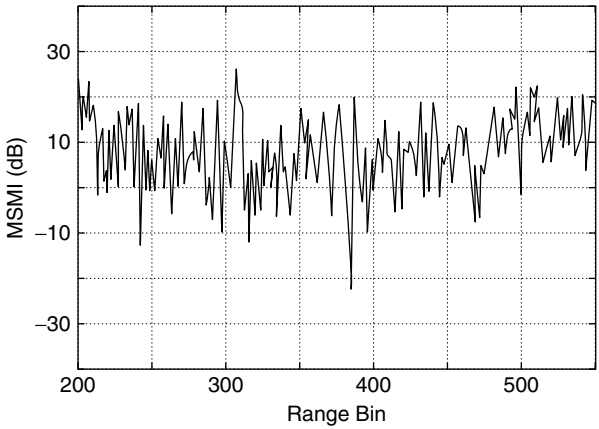

FIGURE 6.69 KBMapSTAP — target injected at range bin 296 (mean MSMI = 12.3, variance MSMI = 35.4).

Figures 6.66 and 6.67 represent the MSMI results for the two algorithms without an injected target. The mean and variance of the results are slightly smaller for KBMapSTAP than for the sliding window algorithm. If a threshold of 20 dB were chosen, then the KBMapSTAP would detect fewer false alarms than the sliding window algorithm.

In heterogeneous environments, KBMapSTAP did consistently better than the sliding window algorithm. For example, Figures 6.68 and 6.69 have a target injected at the same power at range bin 296 and show the MSMI output from each algorithm. If a threshold is chosen at 25 dB we can see that the sliding window algorithm would not detect the target. However, KBMapSTAP would clearly detect it at 5 dB above the threshold.

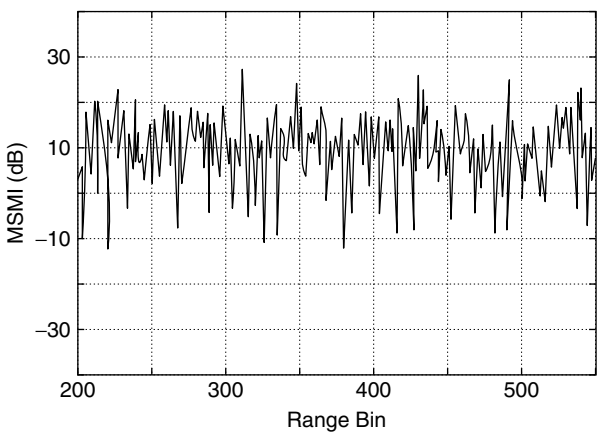

FIGURE 6.70 Sliding window — target injected at range bin 475 (mean MSMI = 12.4, variance MSMI = 28.5).

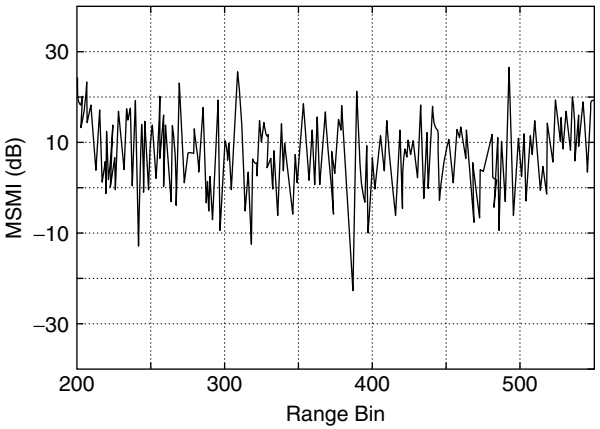

FIGURE 6.71 KBMapSTAP — target injected at range bin 475 (mean MSMI = 11.2, variance MSMI = 26.9).

To test our hypothesis that the KBMapSTAP algorithm would perform the same as the standard algorithm when a target occurred in a homogeneous clutter environment, we injected a target in range bin 475. This range is in water and is surrounded by water such that the major ground clutter is due also to water. Figure 6.70 shows the result of the sliding window algorithm and Figure 6.71, the result of KBMapSTAP. It was conjectured that the KBMapSTAP would do as good as the sliding window algorithm and it did. One could argue however, that it did better considering the lower mean and variance clutter levels.

6.7.7. Conclusions and Future Work

From our limited analysis it can be concluded that the KBMapSTAP algorithm outperforms the standard or sliding window algorithm for heterogeneous clutter environments and performs approximately the same for homogeneous clutter environments. PostDoppler STAP performance can be improved. The data presented here are limited. More analysis and development is required before a quantitative measure of performance can be obtained.

There are some issues that also need to be explored. The data from the USGS database were collected approximately ten years before the radar data was obtained. It is likely that some of the USGS data was not current when the radar data were collected. Techniques to validate map data with the radar need to be explored for those cases where recent map data are not available and when weather and environmental conditions have changed, e.g., snow and flooding.

Map precision is important when the radar's range and angle resolution is significantly different from the map data precision. For our experiment the range resolution of the radar was 120 m and the LULC data points were at a resolution of 200 m by 200 m. Even with this difference in precision the KBMapSTAP

algorithm performed well. A sensitivity analysis should be performed and the clutter patch characterization portion of the algorithm should be modified for varying precision permutations between the radar and the available map data.

The Delmarva area is relatively flat and using LULC data worked well. If however, the terrain is mountainous then the algorithm must include digital elevation model data. This area needs further investigation along with tests to evaluate its performance.

Finally, the LULC data we used did not contain explicit information about man-made features such as railroads, roads, bridges, power lines, etc. The USGS does offer Digital Line Graph data that maps these features. Future work should be done to incorporate these data into the KBMapSTAP algorithm and tests done to measure improvement.

6.8. IMPROVED STAP PERFORMANCE USING KNOWLEDGE-AIDED SECONDARY DATA SELECTION

(C. T. CAPRARO, G. T. CAPRARO, D. D. WEINER, M. C. WICKS, AND W. J. BALDYGO)

6.8.1. Introduction

In order to estimate the clutter covariance matrix needed for STAP, range rings located close to the cell under test are normally chosen as secondary data. If N cells are required for estimation, $N/2$ above the test cell and $N/2$ below the test cell, excluding guard cells, are typically chosen. It is assumed this sliding window (cell averaging symmetric) method of secondary data selection produces cells that are representative of the clutter in the test cell. However, in a nonhomogeneous terrain environment this assumption may not be valid. The amount of secondary data required for proper estimation of the covariance matrix in a stationary environment is predicted to be between two and five times the number of degrees of freedom (DOF) of the radar.[1] As a result, the sample support needed may span hundreds of meters, or even kilometers, depending on the range resolution and the DOF of the radar. Terrain boundaries such as land-water or forest-farmland interfaces are likely to occur. This nonstationarity due to nonhomogeneous terrain can lead to a poor estimation of the clutter covariance matrix and, in turn, poor cancellation of the clutter.

Several authors[2–4] have proposed statistical nonhomogeneity detectors, in both Gaussian and nonGaussian distributed clutter environments, to excise outliers contained within the secondary data. They have demonstrated the deleterious effects of nonhomogeneous secondary data and have shown improvements in STAP performance by filtering the outliers in the selection process.

We propose an approach, in the area of knowledge-aided STAP,[5–8] which uses digital terrain data to aid in choosing representative secondary data. The assumption is that the estimation of the covariance matrix will improve by

choosing secondary data based upon how well its terrain classification compares with the cell under test and, therefore, the STAP algorithm will cancel the interference due to terrain more effectively. This approach is not proposed as a replacement for statistical nonhomogeneity detection. We envision it as a preprocessing step that will enhance the ability of these detectors to filter other types of nonhomogeneities such as targets or discretes.

There is a growing amount of terrain data publicly available at resolutions as small as 10 m. Agencies such as the National Imagery and Mapping Agency,[‡] NIMA (recently renamed National Geospatial-Intelligence Agency, NGA), and the United States Geological Survey,[§] USGS, offer digitized geospatial data containing terrain elevation, classification (urban, agricultural, forested, etc.), linear features (roads, power lines, railroads, etc.), and multispectral imagery. Several software products are also available to aid in converting and viewing the data.[¶]

In this chapter, we provide a description of the measured airborne radar data and terrain data used to demonstrate our approach. We present a method for registering the radar data with the Earth using a more accurate elliptical Earth model. We describe an automated secondary data selection algorithm based on terrain classification. We include a description of the corrections applied to the radar data in order to account for some variations encountered with our approach. We present issues related to range-Doppler spread and propose a solution. Finally, we show results comparing our approach of secondary data selection with the sliding window method.

6.8.2. Radar and Terrain Data

Measured airborne radar data, for this research, was obtained from the AFRL Sensors Directorate's Multi-Channel Airborne Radar Measurements (MCARM) program.[9] The datasets consist of multi-channel clutter data collected by an airborne platform with a side looking radar. The radar was configured with a 2 by 11 channel linear array including sum and delta analog beamformers. MCARM operated at L-Band, in low, medium, and high pulse repetition frequency (PRF) modes. It had a range resolution of approximately 120 m with about 500 range bins of data. Each coherent processing interval (CPI) consisted of 128 pulses and the clutter was typically unambiguous in Doppler. Northrop Grumman collected the data during flights over the Delmarva Peninsula and the East coast of the United States in the mid-1990s. There were eleven flights with an in-scene moving target simulator (MTS) in some of the data collection experiments. The MTS transmitted five Doppler tones (0, -200, -400, -600, and -800 H_z) and was used as the basis for evaluating our results.

[‡] To obtain more information go to http://www.nga.mil.
[§] To obtain more information go to http://www.usgs.gov.
[¶] A versatile and inexpensive product is available from Global Mapper Software, LLC. at http://www.globalmapper.com.

Digital terrain data was obtained from the USGS to classify the ground environment illuminated by MCARM. National Land Cover Data (NLCD) was chosen, which has 21 terrain classifications and a spatial resolution of 30 m.[10] (Each 30 by 30 meter area is given a classification.) The terrain is hierarchically grouped by major classifications, such as, urban areas, barren land, water, etc., and subgrouped into minor classifications, such as, high intensity residential urban areas, low intensity residential urban areas, etc. These data were collected in the 1990s at about the same time as the MCARM experiments. As part of our effort, the NLCD data was converted to an unprojected geographic coordinate system (latitude and longitude) and stored in a relational database for flexible search and retrieval. Other available terrain datasets that provide elevation and linear feature information were not used in this chapter because the Delmarva Peninsula is relatively flat and the NLCD data already contains some information about major roads, bridges, and railroads. However, these additional datasets should be considered especially in mountainous environments or areas where more detailed information about linear features are required.

6.8.3. Approach

6.8.3.1. STAP Algorithm

In order to test our approach, we chose to implement a single-bin post-Doppler STAP algorithm.[11] Although it has been shown in Ref. 12 that heavy Doppler tapering is needed and that multibin postDoppler performance is theoretically better, the single-bin algorithm requires less sample support. The algorithm also nonadaptively suppresses main beam ground clutter and localizes, at the Doppler of interest, competing ground clutter in angle (see Figure 6.72). As a result, the terrain, presumably causing the dominant interference, is confined to a narrow

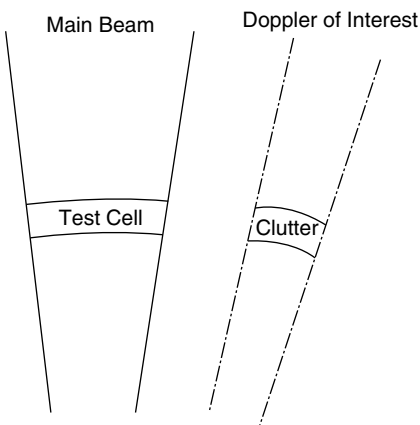

FIGURE 6.72 Location of competing ground clutter with single-bin post-Doppler STAP algorithm.

angular region. Therefore, instead of comparing terrain within the entire range ring under test with other range rings, we compared just the competing range-Doppler cell at the Doppler of interest with cells along the same iso-Doppler.

6.8.3.2. Registration of the Radar with the Earth

Registration of the radar with the Earth was performed to determine the terrain illuminated by the MCARM radar during a CPI. A system of three nonlinear equations was developed to calculate the position of a point on the Earth given a slant range, a Doppler frequently, and an oblate spheroid (elliptical) model of the Earth. Figure 6.73 illustrates the geometry of the problem.

An Earth-Centered Earth-Fixed (ECEF) geographical coordinate system was used. The point $P_r(x_r,y_r,z_r)$ represents the position of the radar while the point on the Earth to be determined is designated as $P_e(x,y,z)$. Also shown in Figure 6.73 is the slant range, R_s, to point P_e and the iso-Doppler of interest (represented as a dashed line).

Assuming the radar is flying slower than the maximum unambiguous Doppler velocity, the intersection of the slant range with the iso-Doppler contour and the Earth's surface occurs at two points, P_e and a mirror point on the iso-Doppler. However, since the radar data was gathered by a side-looking radar we need only determine the location of one of the two points depending on the orientation of the radar platform.

The first equation is related to the slant range and is simply the Euclidian distance between points P_e and P_r. The functional form of the equation is given as

$$F_1(x,y,z) = (x - x_r)^2 + (y - y_r)^2 + (z - z_r)^2 - R_s^2 = 0 \qquad (6.56)$$

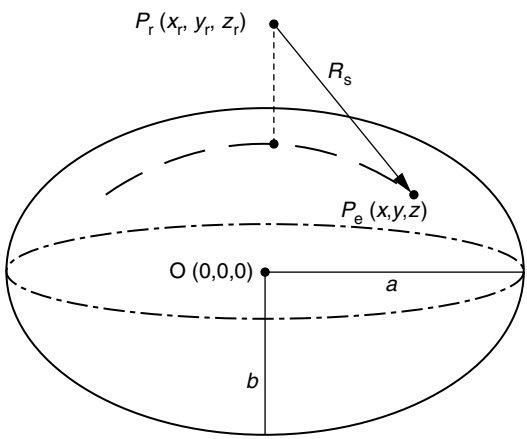

FIGURE 6.73 Registration geometry.

The second equation represents the iso-Doppler contour on the Earth and takes the form,

$$F_2(x, y, z) = (x - x_r)\nu_{rx} + (y - y_r)\nu_{ry} + (z - z_r)\nu_{rz} - \frac{f_d \lambda R_s}{2} = 0 \quad (6.57)$$

where ν_{rx}, ν_{ry}, ν_{rz} are the components of the radar's velocity vector, f_d is the Doppler frequency and λ is the wavelength of the radar. The last equation models the Earth's surface as an oblate spheroid and is defined as,

$$F_3(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{a^2} + \frac{z^2}{b^2} - 1 = 0 \quad (6.58)$$

where a and b are the semi-major and semi-minor radii of the Earth, respectively. Values for these parameters were obtained from the WGS84 (also known as GRS80) world geodetic datum because they were used by the USGS to define the coordinates of the terrain data. However, depending upon where registration is to be performed on the Earth, there may be a more accurate local datum available.

In order to find solutions for x, y, and z, an iterative Newton-Raphson method[13] was used until the method converged to a solution within a certain tolerance. The initial point of the iteration was calculated from a spherical Earth model and was chosen to be near the point of interest, P_e. This helped the Newton–Raphson method converge to P_e instead of its mirror point. A check was done to ensure the result was on the correct side of the radar platform.

Atmospheric propagation effects such as ducting were not modeled in this chapter. In addition, the elevation of the terrain was not included since the area was relatively flat and close to sea level.

6.8.3.3. Data Selection

Once the registration algorithm was complete, it was used to determine the locations of the boundary points along the iso-Doppler of interest, corresponding to each range-Doppler cell. The points were then converted from ECEF coordinates to geographic coordinates and used to query the NCLD terrain database. The query returned a count of each of the 21 terrain classifications contained within the boundary points. The result obtained for each cell was stored in a vector and normalized to account for the variation in area of the range-Doppler cells.

After each of the range-Doppler cell terrain vectors was computed, a comparison was made with the test cell vector using an Euclidian distance measure. The cells whose terrain vector was closest to the test cell's were selected.

6.8.3.4. Corrections for Visibility

One of the advantages of the sliding window method of secondary data selection is that the data, including the test cell, are very close in range. As a result, it is less susceptible to variations in power due to range, clutter reflectivity due to

grazing angle, and Doppler frequency due to array misalignment with the radar's velocity vector.[11,14] Since our approach may choose secondary data that spans a larger range extent, corrections were applied to the radar data to account for these variations. The corrections were also applied for the sliding window method.

6.8.3.5. Secondary Data Guard Cells

As part of our approach, issues concerning range-Doppler spread were addressed. During the analysis of the MCARM data containing the MTS simulated targets, we noticed that a certain amount of range-Doppler spread occurred. This may have been caused by numerous factors. As a result, cells experience signal contamination from neighboring cells. This violates the requirement that secondary data be independent and identically distributed (i.i.d.) when used in estimating the clutter covariance matrix. In order to mitigate this effect, guard cells were placed around the range-Doppler cells selected for secondary data. Excluding these secondary data guard cells (SDGC) is analogous to the standard practice of placing guard cells around the test cell. The number of SDGC used was chosen by the amount of spread measured.

6.8.4. Results

The results presented compare the sliding window method of secondary data selection to our knowledge-aided approach. A single CPI from flight 5, acquisition 151, of the MCARM program, was processed, which contains simulated target signals from the MTS. A modified sample matrix inversion (MSMI) test statistic[15] is plotted versus range bin for each of the results obtained. The ratio of the MTS signal's MSMI value to the range averaged MSMI value is our preferred performance measure (PPM) in this work.

In Figure 6.74, all 22 channels of the MCARM array were used for STAP. A total of 44 secondary data samples was chosen for the estimation of the

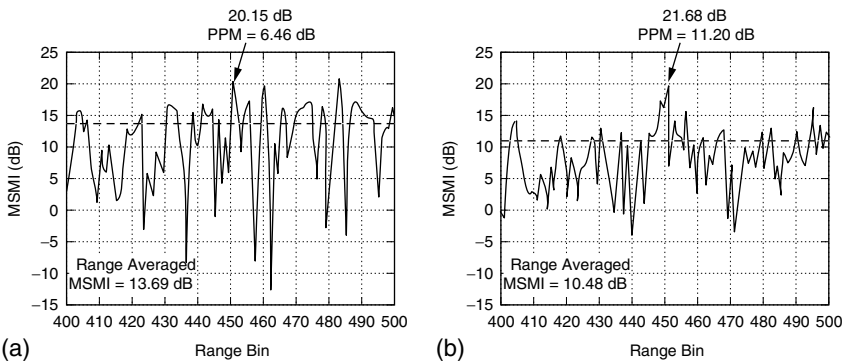

FIGURE 6.74 Using full array. (a) Sliding window method. (b) Knowledge-aided approach.

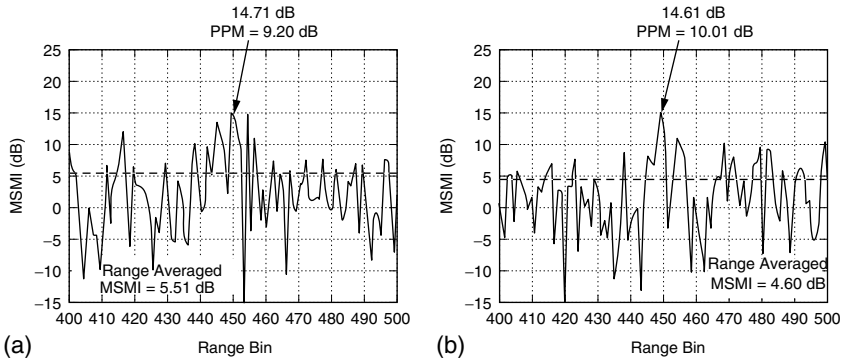

FIGURE 6.75 Using full array and secondary data guard cells. (a) Sliding window method. (b) Knowledge-aided approach.

covariance matrix. Guard cells were placed around the cell under test. However, no secondary data guard cells were excluded. The simulated target is located at range bin 450 (see arrow) and its MSMI value and PPM are given in each figure. The range averaged MSMI value is also given and represented by a dashed line.

The PPM of our knowledge-aided approach, as illustrated in Figure 6.74, was approximately 4.7 dB better than the sliding window method. Notice that the knowledge-aided approach not only raised the MSMI value of the target but it also lowered the range averaged MSMI statistic.

As mentioned above, there was some range-Doppler spread in the radar data. Figure 6.75 shows the results obtained when guard cells are placed around the secondary data as well as the cell under test. It can be seen that the range-averaged MSMI value is significantly lowered, in both cases, by 6 to 8 dB. Furthermore, the PPM of the simulated target, using the sliding window method and SDGC, was almost 3 dB better. However, the knowledge-aided approach did slightly poorer with SDGC. This may be caused by the reduction in sample support due to

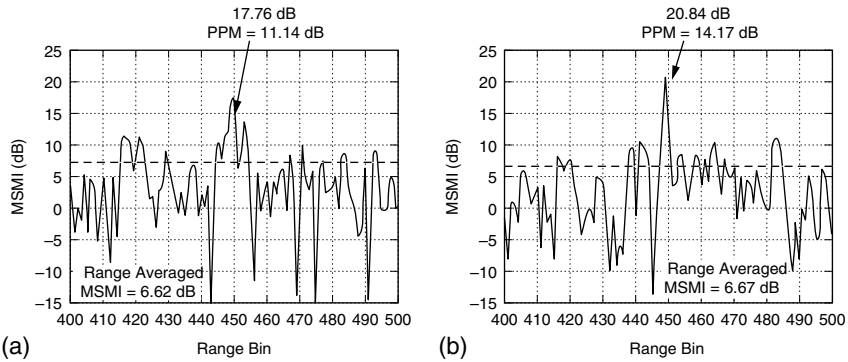

FIGURE 6.76 Using half array and secondary data. (a) Sliding window method. (b) knowledge-aided approach.

elimination of the guard cells from use as secondary data. Since the knowledge-aided approach may choose secondary data farther in range from the test cell than the sliding window method, the added constraint of SDGC may not be beneficial in some cases.

In Figure 6.76, the returns from only the top row of the MCARM array, consisting of 11 of the 22 available channels, were used for STAP to compensate for the reduction in available secondary data because of SDGC. Although this reduces the number of degrees of freedom for the adaptive filter, it also reduces the amount of sample support needed from 44 samples to 22. The results show an increase in performance, for both cases, compared to the previous results, and a best performance with the knowledge-aided approach.

6.8.5. Conclusion

We have presented a knowledge-aided approach utilizing terrain data to select secondary data for STAP. Measured airborne radar data from AFRL Sensors Directorate's MCARM program and digital terrain data from the USGS were used to evaluate this approach. Corrections were applied to the radar data due to variability in factors affecting the clutter returns. Guard cells were placed around secondary data to mitigate the real world effects of range-Doppler spread. A single-bin postDoppler STAP algorithm was chosen and a comparison was performed between the standard sliding window method and the knowledge-aided approach. The results illustrate the benefits of using terrain information, *a priori* data about the radar, and the importance of statistical independence when selecting secondary data for improving STAP performance.

Further study is needed to determine how well terrain classification data correlates with airborne radar clutter statistics. Other types of terrain data should be studied as well in order to explore their potential as an aid for STAP.

Future work will also include integrating this novel approach into the AFRL Signal Processing Evaluation, Analysis and Research (SPEAR) Testbed, configuring and evaluating it with several STAP algorithms and measured GMTI radar datasets. The SPEAR Testbed provides a means of assessing performance against a variety of signal processing metrics to aid in the comparison of multiple competing adaptive signal processing approaches.

Additionally, SPEAR allows detection data to be input into several GMTI tracking algorithms giving the user the ability to capture performance metrics at the tracking stage. This unique ability—to correlate signal processing and tracking metrics across a diverse set of signal processing algorithms, measured and simulated datasets, and knowledge sources—provides a compelling means to demonstrate the effectiveness of this technology to the warfighter.

Part II

Adaptive Antennas

7 Introduction

M. M. Weiner

Part II *Adaptive Antennas* discusses implementation of the *second* of three sequentially complementary approaches for increasing the probability of detection within at least some cells of the surveillance volume for external "noise" which can be either Gaussian or non-Gaussian in the spatial domain but *is* Gaussian in the temporal domain. This approach, identified in the preface as Approach B and also known as *space–time adaptive processing*, seeks to reduce the competing electromagnetic environment by placing nulls in its principal angle-of-arrival and Doppler frequency (space–time) dimensions. This approach utilizes, $k = \mathbf{NM}$ samples of signals from **N** subarrays of the antenna, over a coherent processing interval containing **M** pulses to (1) estimate in the space–time domain, an $\mathbf{NM} \times \mathbf{NM}$ "noise" covariance matrix of the subarray signals, (2) solve the matrix for up to **N** unknown "noise" angles of arrival and **M** unknown "noise" Doppler frequencies, and (3) determine appropriate weighting functions for each subarray which will place nulls in the estimated angle-of-arrival and Doppler frequency domains of the "noise". Approach B is a form of filtering in those domains. Consequently, the receiver detector threshold can be reduced because the average "noise" voltage variance of the surveillance volume is reduced. The locations and depths of the nulls are determined by the relative locations, strengths of the "noise" sources in the space–time domain and by the differences between the actual and estimated "noise" covariance matrices. The results are influenced by the finite number k of stochastic data samples and the computational efficiency in space–time processing of the samples. Part II *Adaptive Antennas* addresses these issues and presents physical models for several applications.

Chapter 8 discusses, the Joint-Domain Localized General Likelihood Ratio (JDL-GLR) algorithm "as an attractive solution to the problem of the joint-domain optimum space–time processor of high order (N_S spatial channels $\times$ N_t pulses in a coherent processing interval) with a fast convergence rate and high computational efficiency, together with such highly desirable features as the embedded constant false-alarm rate (CFAR) and robustness in non-Gaussian clutter/interference".

Chapter 9 presents a practical design of a *smart* antenna for high data-rate wireless systems. An *adaptive* antenna utilizes space–time adaptive processing, to minimize its radiation patterns in the direction or directions of external

electromagnetic "noise". A smart antenna, on the other hand, utilizes *either* space–time adaptive processing *or* up-dated knowledge (assuming such information is available) of the location of interferers to achieve the same objective. A smart antenna is therefore, a more general class of antenna but not necessarily is as complex or effective as an adaptive antenna.

Chapter 10 presents applications of adaptive antennas for cancellation of jammer multipath and main beam and side lobe jamming (Section 10.1, Section 10.4, Section 10.5, Section 10.10); clutter reduction (Section 10.2, Section 10.3, Section 10.6, Section 10.8, 10.9); improved monopulse patterns (Section 10.7); and improved search and track (Section 10.11). Chapter 10 is based on physical models assuming that the clutter is statistically homogeneous. It is assumed that the clutter and jammer covariance matrix estimates (based on a finite number of samples) yield a good approximation to the true covariance matrix of the physical model. Of course, the physical model is only as good as its representation of the real-life surveillance-volume clutter physics. With the exception of Section 10.6, no experimental data is presented in Chapter 10. In real-life situations, often the surveillance volume is not statistically homogeneous, its clutter physics is different from the assumed physics, and the number of available samples are too few. The methodologies of Part I and Chapter 8 address these real-life situations. Therefore, the use of those methodologies, together with physical models of Chapter 10, are expected to yield real-life experimental results which are closer to the optimistic theoretical results, predicted by physical models of Chapter 10.

8 Adaptive Implementation of Optimum Space–Time Processing

L. Cai and H. Wang

CONTENTS

8.1. Introduction 421
8.2. Data Modelling 423
8.3. Difference Among the Performance Potentials of the Cascade and Joint-Domain Processors 425
8.4. The JDL–GLR Algorithm 430
8.4.1. The JDL–GLR Principle 431
8.4.2. The JDL–GLR Detection Performance 433
8.4.3. Detection Performance Comparison 433
8.4.4. Other Features of JDL–GLR 436
8.5. Conclusions and Discussion 436

8.1. INTRODUCTION

For detection of weak targets in strong clutter/interference of complicated angle-Doppler spectrum, it is highly desirable for an airborne radar system to have the optimum or near optimum performance. As the clutter or interference spectrum is unknown to the system and the clutter or interference environment may be varying in time and space, i.e., non stationary and non homogeneous, the signal processor must be adaptive with a sufficiently fast convergence-rate.

Consider a system, which employs N_s spatial channels (subarrays of a phased-arrays) and has N_t pulses in its Coherent Processing Interval (CPI). The optimum processor or the Neyman–Pearson's likelihood ratio test for such a system, is well developed in Ref. 1 under the assumption of Gaussian clutter or interference. This processor, to be referred to as the joint-domain optimum processor in this chapter, has the highest performance potential which can be approached by adaptive algorithms such as the Sample-Matrix-Inversion (SMI),[2] the

Generalized Likelihood Ratio (GLR),[3,4] and the Modified SMI.[5,6] To approach this detection performance potential, however, these algorithms require the training-data set (i.e., the so-called secondary data set) to have at least $2N_sN_t \sim 3N_sN_t$ independent and identically distributed (iid) data vectors. Obviously, such a training-data size requirement is impractical even for moderate N_s and N_t, as the environment in which an airborne surveillance system operates, is usually severely nonstationary and nonhomogeneous. Besides, the computational load can easily become unbearable in practice since it is proportional to $(N_sN_t)^3$. One should also note that, lowering N_s and N_t is not necessarily desirable in practice as the performance potential critically depends on these when the angle-Doppler spectrum of the clutter or interference is complicated.

The more popular approach to space–time processing can be classified as cascade processing with either the beamformer-Doppler processor configuration or the opposite order configuration. In this chapter, the former will be called the *Space–time (S–T) configuration* and the latter the *Time–Space (T–S) configuration*. Obviously, the optimum detection theory can be applied separately to spatial and temporal parts of S–T and T–S configurations, together with various adaptive algorithms available for each part. Of course, the convergence-rate and computation load problems associated with adaptive implementation of the joint-domain optimum processor also appear with the cascade configurations, only to a lesser extent. When the convergence occurs, the performance of an adaptive implementation with the S–T or T–S configuration should approach that of the optimum processor with the same configuration. Cascade processing, especially the S–T configuration, has been so popular in recent years that it seems to replace the joint-domain processor in the airborne surveillance application. Arguments can often be heard about which cascade configuration has higher detection performance potential.

The first objective of this chapter is to show that

(1) None of the two cascade configurations is better than the other, and
(2) the performance potential of cascade configurations can fall far below that of the joint-domain optimum processor. In other words, we show that if one wants to approach the highest performance potential offered by the joint-domain optimum processor, both cascade configurations should be avoided.

As pointed out earlier in this section, it is difficult in practice to approach the performance potential of the joint-domain optimum processor with the straightforward application of adaptive algorithms such as the SMI, Modified SMI, GLR, etc., especially in a severely non stationary and non homogeneous environment, even if the heavy real-time computation could become affordable. Therefore, the second objective of this chapter is to develop a new adaptive algorithm for the joint-domain optimum processor, which should be more data-efficient and computationally efficient than the aforementioned ones. This new

algorithm is an extension of our recent work reported in Refs. 7 and 8 for adaptive Doppler-domain processing.

We will first formulate the data model in Section 8.2. In Section 8.3 we will compare the performance potentials of the cascade and joint-domain processors. The new adaptive algorithm for the joint-domain optimum processor is presented in Section 8.4, together with its performance analysis and comparison. Finally, Section 8.5 summarizes the conclusions with some discussion of related issues.

8.2. DATA MODELLING

Consider a narrowband antenna array with N_s spatial channels (subarrays). Each channel receives N_t data samples corresponding to the return of a train of N_t coherent pulses for a given range cell. Let the column vector x_{tn_s}, $N_t \times 1$, represent the N_t baseband complex (I/Q) data samples of the n_sth channel. The data matrix $\mathbf{X}$, $N_t \times N_s$, is defined by

$$\mathbf{X} = [x_{t1} x_{t2} \ldots x_{tN_s}] = \begin{bmatrix} x_{s1}^{\mathrm{T}} \\ x_{s2}^{\mathrm{T}} \\ \vdots \\ x_{sN_t}^{\mathrm{T}} \end{bmatrix} \tag{8.1}$$

where "T " denotes the transpose and the row vectors of $\mathbf{X}$, $x_{sn_t}^{\mathrm{T}}$, $n_t = 1, 2, \ldots, N_t$, are the "snapshots" obtained along the spatial channels.

Under the signal-absence hypothesis H_0, the data matrix $\mathbf{X}$ consists of clutter or interference and noise components only, i.e.,

$$\mathbf{X} = \mathbf{C} + \mathbf{N} \tag{8.2}$$

where $\mathbf{C}$ and $\mathbf{N}$ represent the clutter or interference and noise, respectively, and are assumed to be independent. Under the signal-presence hypothesis H_1, a target signal component also appears in the data matrix, i.e.,

$$\mathbf{X} = \alpha \mathbf{S} + \mathbf{C} + \mathbf{N} \tag{8.3}$$

where α is an unknown complex constant representing the amplitude of the signal and $\mathbf{S}$ the signal matrix of a known form. We call $\mathbf{X}$ the primary data set as it is from the range cell under the hypothesis test.

For simplicity of discussion only, we assume that the spatial channels are collinear, identical, omni-directional, and equally spaced with spacing d; and that the pulses of the coherent pulse trains are identical with a constant Pulse Repetition Frequency (PRF). Under these assumptions, the $n_t n_s$th entry of the signal matrix $\mathbf{S}$ has the following form

$$s(n_t, n_s) = \exp\left[\mathrm{i}2\pi(n_t - 1)\frac{2v}{\lambda \mathrm{PRF}} + \mathrm{i}2\pi(n_s - 1)\frac{d \sin\theta}{\lambda} \right] \tag{8.4}$$

where v is the radial velocity of the target, θ the direction of arrival of the target-return planewave with respect to the broadside of the array, and λ the radar wavelength. Denoting

$$f_{st} = \frac{2v}{\lambda \text{PRF}} \tag{8.5}$$

as the "normalized Doppler frequency" of the target signal, and

$$f_{ss} = \frac{d \sin \theta}{\lambda} \tag{8.6}$$

as the "spatial frequency," $\mathbf{S}$ can be expressed by

$$\mathbf{S} = \mathbf{s}_s^{\mathrm{T}} \otimes \mathbf{s}_t \tag{8.7}$$

where: $\otimes$ is the Kronecker tensor product, and

$$\mathbf{s}_t = [\,1, \quad \exp(\mathrm{i}2\pi f_{st}), \quad \cdots \quad , \exp(\mathrm{i}2\pi(N_t - 1)f_{st})\,]^{\mathrm{T}} \tag{8.8}$$

and

$$\mathbf{s}_s = [\,1, \quad \exp(\mathrm{i}2\pi f_{ss}), \quad \cdots \quad , \exp(\mathrm{i}2\pi(N_s - 1)f_{ss})\,]^{\mathrm{T}} \tag{8.9}$$

are the signal vectors in time and space domains, respectively. We assume that the parameters PRF, λ, and d have been properly chosen so that f_{st} and f_{ss} are confined within $[-0.5, 0.5]$.

To statistically characterize the clutter or interference and noise components $\mathbf{C}$ and $\mathbf{N}$, we introduce the notation Vec$(\cdot)$ for a matrix operation that stacks the columns of a matrix under each other to form a new column vector. We assume that the $N_t N_s \times 1$ vector Vec$(\mathbf{C} + \mathbf{N})$ has a multivariate complex-Gaussian distribution with zero mean and a covariance matrix $\mathbf{R}$. Under this assumption, x_{tn_s}, $n_s = 1, 2, \ldots, N_s$ and x_{sn_t}, $n_t = 1, 2, \ldots, N_t$ will also be complex zero-mean Gaussian. Let $\mathbf{R}_t$ and $\mathbf{R}_s$ be the covariance matrices of x_{tn_s} and x_{sn_t}, respectively. It is easy to see that $\mathbf{R}_t$ and $\mathbf{R}_s$ are the submatrices of $\mathbf{R}$.

In the cases of unknown clutter or interference statistics, the data from the adjacent range cells, conventionally referred to as the secondary data set, are also needed for estimating the covariance of clutter or interference. Under either hypothesis H_1 or H_0, these consist of the clutter or interference and noise components only, and are denoted by

$$\mathbf{Y}_k = \mathbf{C}_k + \mathbf{N}_k, \qquad N_t \times N_s; \qquad k = 1, 2, \ldots, K \tag{8.10}$$

where K is the number of range cells available. We assume that $\mathbf{Y}_k$, $k = 1,2,\ldots,K$ and $\mathbf{X}$ are independent of each other and bear the same clutter or interference statistics, i.e., Vec$(\mathbf{Y}_k)$ should also have a complex-Gaussian distribution with zero mean and a covariance matrix $\mathbf{R}$.

8.3. DIFFERENCE AMONG THE PERFORMANCE POTENTIALS OF THE CASCADE AND JOINT-DOMAIN PROCESSORS

We will compare the detection performance potentials of the two cascade configurations and the joint-domain processor under the assumption that the clutter- or interference-plus-noise covariance matrix is known. With the known covariance, the S–T configuration is the N_sth order optimum spatial processor followed by the N_tth order optimum temporal (Doppler) processor, the T–S configuration takes the opposite cascade, and the joint-domain processor is the N_sN_tth order optimum processor. Applying the result in Ref. 1 to the above three, we list the optimum weight vectors below for easy reference.

The S–T configuration: we have

$$\mathbf{w}_{s,s-t} = c_{s,s-t}\mathbf{R}_s^{-1}\mathbf{s}_s \tag{8.11}$$

for the spatial domain weight vector, and

$$\mathbf{w}_{t,s-t} = c_{t,s-t}[(\mathbf{w}_{s,s-t}^H \otimes \mathbf{I})\mathbf{R}(\mathbf{w}_{s,s-t} \otimes \mathbf{I})]^{-1}\mathbf{s}_t \tag{8.12}$$

for the temporal domain weight vector, where $c_{s,s-t}$ and $c_{t,s-t}$ are constants. We recall that $\mathbf{R}_s$ and $\mathbf{R}_t$ are the covariance matrices for the rows and columns of $\mathbf{X}$, respectively, and $\mathbf{s}_s$ and $\mathbf{s}_t$ are specified by Equation 8.8 and Equation 8.9. The test statistic is

$$\eta_{s-t} = \mathbf{w}_{t,s-t}^H \mathbf{X} \mathbf{w}_{s,s-t}^* \tag{8.13}$$

The T–S configuration:

$$\mathbf{w}_{t,t-s} = c_{t,t-s}\mathbf{R}_t^{-1}\mathbf{s}_t \tag{8.14}$$

and

$$\mathbf{w}_{s,t-s} = c_{s,t-s}[(\mathbf{I} \otimes \mathbf{w}_{t,t-s}^H)\mathbf{R}(\mathbf{I} \otimes \mathbf{w}_{t,t-s})]^{-1}\mathbf{s}_s \tag{8.15}$$

for the temporal and spatial weight vectors, respectively. The test statistic is

$$\eta_{t-s} = \mathbf{w}_{t,t-s}^H \mathbf{X} \mathbf{w}_{s,t-s}^* \tag{8.16}$$

The joint-domain optimum processor: the whole set of the data is processed all together by an optimum weight vector as

$$\eta_{\mathrm{J}} = \mathbf{w}_J^H \mathrm{Vec}(\mathbf{X}) \tag{8.17}$$

where $\mathbf{w}_{\mathrm{J}}$ is

$$\mathbf{w}_{\mathrm{J}} = c_J \mathbf{R}^{-1} \tag{8.18}$$

with c_{J} being a constant scalar.

One should note that the overall weight vectors for the two cascade configurations can have the following equivalent expressions

$$\mathbf{w}_{s-t} = \mathbf{w}_{s,s-t} \otimes \mathbf{w}_{t,s-t} \tag{8.19}$$

and

$$\mathbf{w}_{t-s} = \mathbf{w}_{s,t-s} \otimes \mathbf{w}_{t,t-s} \tag{8.20}$$

The squared magnitude of the test statistic is compared with a chosen threshold η_0 which is determined by the required probability of false alarm P_{f} as

$$\eta_0 = -\ln P_{\mathrm{f}} \tag{8.21}$$

and the signal presence is claimed if the test statistic surpasses the threshold.

From the result in Ref. 1, the probability of detection of the above three processors has the same form below with their own weight vectors, i.e., $\mathbf{w}_{s-t}$, $\mathbf{w}_{t-s}$, and $\mathbf{w}_{\mathrm{J}}$ to replace $\mathbf{w}$ therein

$$P_{\mathrm{d}} = 1 - \exp(-\gamma) \int_0^{\eta_0} \exp(-t) \mathbf{I}_0(2\sqrt{\gamma t}) \mathrm{d}t \tag{8.22}$$

where:

$$\gamma = |\alpha|^2 \frac{\mathbf{w}^H s s^H \mathbf{w}}{\mathbf{w}^H \mathbf{R} \mathbf{w}} \tag{8.23}$$

and $\mathbf{I}_0(\cdot)$ denotes the zeroth order modified Bessel function of the first kind.

The key to achieving the objective of the comparison easily is to identify a few typical cases, from the vast number of varieties of clutter or interference conditions, which are also simple enough for numerical evaluation. To do so, the following specifics are necessary.

(1) The covariance matrix of the receiver noise is given by

$$\mathrm{E}(\mathrm{Vec}(\mathbf{N})\mathrm{Vec}(\mathbf{N})^H) = \sigma_n^2 \mathbf{I} \tag{8.24}$$

with $\mathbf{I}$ being the $N_t N_s \times N_t N_s$ identity matrix.

(2) The clutter or interference is assumed to have a two-dimension power spectral density of the Gaussian shape centered at $[f_{ct}, f_{cs}]$

$$P_c(f_t, f_s) = \sigma_c^2 \frac{1}{2\pi \sigma_{ft} \sigma_{fs}} \exp\left[-\left(\frac{(f_t - f_{ct})^2}{2\sigma_{f_t}^2} + \frac{(f_s - f_{cs})^2}{2\sigma_{f_s}^2} \right) \right] \tag{8.25}$$

where: f_t and f_s are the normalized Doppler frequency and spatial frequency, respectively, and σ_{f_t} and σ_{f_s} the parameters controlling the spread of the clutter or interference spectrum. The separation between

the signal and the centre of the clutter or interference spectrum is denoted by $\Delta f_t = f_{st} - f_{ct}$ and $\Delta f_s = f_{ss} - f_{cs}$.

(3) The covariance of the clutter or interference corresponding to the above spectrum is then found to be

$$\mathrm{E}(\mathrm{Vec}(\mathbf{C})\mathrm{Vec}(\mathbf{C})^H) = \sigma_c^2 \mathbf{C}_s \otimes \mathbf{C}_t \tag{8.26}$$

where $\mathbf{C}_t$ and $\mathbf{C}_s$ are Toeplitz matrices specified by

$$\mathbf{C}_t = \mathrm{Toeplitz}\{[1,\ \mathrm{e}^{-2(\pi\sigma_{ft})^2 - \mathrm{i}2\pi f_{ct}},\ \cdots,\ \mathrm{e}^{-2(\pi\sigma_{ft}(N_t-1))^2 - \mathrm{i}(N_t-1)2\pi f_{ct}}]\} \tag{8.27}$$

and

$$\mathbf{C}_s = \mathrm{Toeplitz}\{[1,\ \mathrm{e}^{-2(\pi\sigma_{fs})^2 - \mathrm{i}2\pi f_{cs}},\ \cdots,\ \mathrm{e}^{-2(\pi\sigma_{fs}(N_s-1))^2 - \mathrm{i}(N_s-1)2\pi f_{cs}}]\} \tag{8.28}$$

respectively. It is easy to verify that Equation 8.24 and Equation 8.26 will lead to $\mathbf{R}_t = \sigma_c^2 \mathbf{C}_t + \sigma_n^2 \mathbf{I}$ and $\mathbf{R}_s = \sigma_c^2 \mathbf{C}_s + \sigma_n^2 \mathbf{I}$.

We define the clutter/interference-to-noise-ratio (INR) and signal-to-clutter/interference-plus-noise-ratio (SINR) by

$$\mathrm{INR} = \frac{\sigma_c^2}{\sigma_n^2} \tag{8.29}$$

and

$$\mathrm{SINR} = \frac{|\alpha|^2}{(\sigma_n^2 + \sigma_c^2)} \tag{8.30}$$

Three simple cases are identified below in each of which at least one of the cascade configurations suffers severe performance degradation, i.e., significantly departing from the joint-domain optimum.

Case 1. The signal and interference are "well" separated in the angle domain (in the sense that $\Delta f_s > 1/N_s$) but close to each other in the Doppler-domain ($\Delta f_t < 1/N_t$). This situation is shown in the subplot in Figure 8.1. The detection performance vs. SINR for the three processors are plotted in Figure 8.1 with INR = 40 dB and $P_{\mathrm{f}} = 10^{-5}$. The S–T configuration shows almost the same performance potential as the joint-domain optimum in this special case, while the performance loss for the T–S configuration becomes significantly large.

Case 2. The signal and interference are "well" separated in the Doppler-domain but close to each other in the angle domain, as indicated by the subplot in Figure 8.2. The T–S configuration is now close to the joint-domain optimum while the S–T configuration departs significantly.

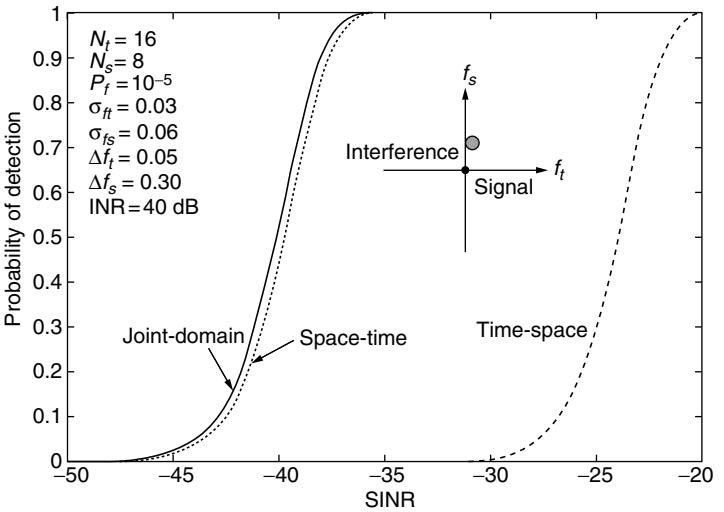

FIGURE 8.1 Performance comparison of the three processing configurations: case 1.

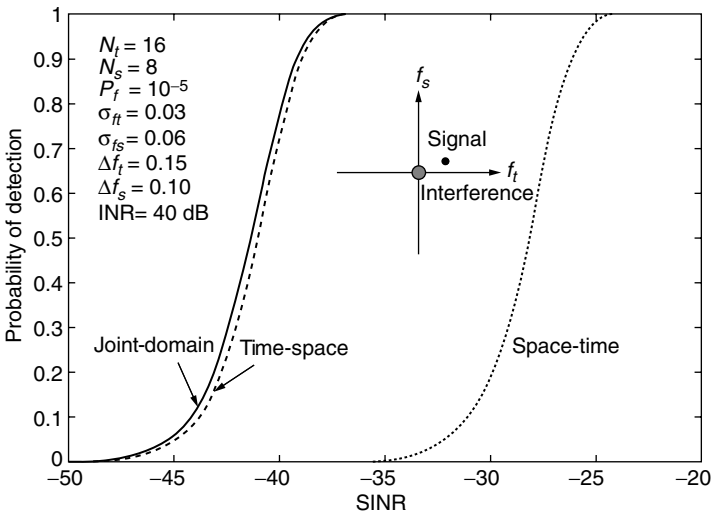

FIGURE 8.2 Performance comparison of the three processing configurations: case 2.

Case 3. The clutter or interference spectrum has two peaks with one close to the signal in the angle domain while the other in the Doppler-domain. In this case *both* cascade configurations fail to approach the joint-domain optimum, as shown in Figure 8.3.

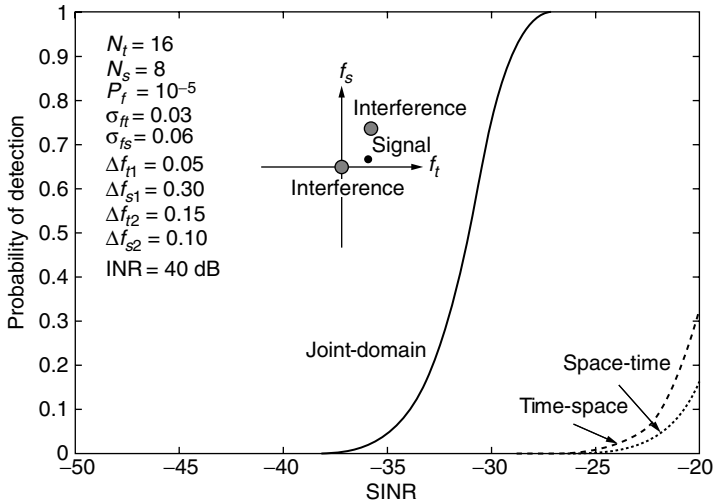

FIGURE 8.3 Performance comparison of the three processing configurations: case 3.

The above three cases are typical in the sense that we can draw from them the following conclusions:

(1) none of the two cascade configurations is better than the other, and
(2) the performance potential of both cascade configurations can fall far below that of the joint-domain optimum processor.

Intuitively, the above conclusions are also well justified. The T–S configuration in Case 1 suppresses the signal as well as the clutter (or interference) as these have little separation in the Doppler frequency domain, so does the S–T configuration in Case 2 in the angle domain. As both Case 1 and Case 2 can appear in practical situations without *a priori* knowledge, preselection of either cascade configuration is thus not appropriate. In Case 3, the signal and clutter (or interference) have little separation in either of the two domains, which results in the failure of both cascade configurations. However, the separation in the joint-domain in Case 3 is still sufficiently large to lead to the success of the joint-domain optimum processor. As an airborne system has to deal with clutter or interference having angle and Doppler spectral spread, it is thus important to make full use of the signal and clutter (or interference) separation, which cannot always be achieved by either of the two cascade configurations.

Although our study so far in this chapter is centered around the detection performance potentials, i.e., under the assumption of known clutter or interference statistics, it is sufficient for us to direct our attention only to the adaptive implementation of the joint-domain optimum processor, since the two cascade configurations have been shown to have limited potentials. This will

be the focus of the remaining part of this chapter. Before we proceed, we should point out that, in addition to the problem of limited potentials, the two cascade configurations may have other serious problems of practical importance which are associated with their adaptive implementations, e.g., the difficulty to achieve a high-quality Constant False Alarm Rate (CFAR). This issue will be briefly discussed later in Section 8.5 to preserve the continuity of our main course.

8.4. THE JDL–GLR ALGORITHM

As pointed out in the introduction, the straightforward application of available adaptive algorithms such as the SMI, Modified SMI, and GLR, etc., has considerable difficulty to approach the joint-domain optimum processor in practice, especially in severely non stationary and non homogeneous environments. Our goal here is to develop an adaptive implementation which is more data-efficient (in the sense of faster convergence or requirement of fewer training-data) as well as more computationally efficient. In addition, it is highly desirable in practice to have the adaptive algorithm possess an embedded CFAR feature and a low sensitivity to the deviation of the clutter or interference distribution from the assumed Gaussian.

To achieve the above goal we will follow the idea of localized adaptive processing as presented in Refs. 7 and 8 for adaptive MTD. Although this idea is similar to that of beam-space processing in Refs. 9–11 under the term of partially adaptive array processing, the work in Refs. 7 and 8 distinguishes itself from the previous study on beam-space processing in the following ways. Refs. 7 and 8 are the first to point out that, for the cases of limited training-data size, the use of localized adaptive processing is almost mandatory, and they have shown that localized adaptive processing can actually *outperform* fully adaptive processing in non stationary and non homogeneous environments. Furthermore, Refs. 7 and 8 are also the first to study localized adaptive processing with the *detection* performance measure, which is of course the primary concern of surveillance systems. In contrast, the previous work on beam-space processing focuses on the *steady state* performance and uses the signal *estimation* performance measure. As the primary concern of this chapter is again detection in severely non stationary and non homogeneous environments, it is natural to follow the work in Refs. 7 and 8. Of course, the extension represents a non trivial task as indicated by the complexity of the joint angle-Doppler-domain.

As discussed in Refs. 7 and 8, the localized processing idea can be applied with a variety of adaptive algorithms such as the SMI, Modified SMI, and GLR. We will again pick up the GLR because it offers the desirable embedded CFAR feature as well as possesses the desirable robustness in nonGaussian clutter or interference.[5,6] Hence, the new algorithm presented in this section will be called the JDL–GLR algorithm, denoting that joint-domain localized (JDL) processing is used in conjunction with the GLR algorithm.

8.4.1. THE JDL–GLR PRINCIPLE

Figure 8.4 illustrates the principle of the JDL–GLR processor we propose. The data in the space–time domain, $\mathbf{X}$, $N_t \times N_s$ is first transformed to the angle-Doppler-domain. This multi-dimensional transform should be invertible to avoid any information loss. It can be done most conveniently via the standard two-dimensional DFT discrete Fourier transform (which is linear and orthogonal), under the assumption made in Section 8.2 for the spatial channels and pulse train. One should note that the gaussianarity assumed for $\mathbf{X}$ will not be affected if the transformation is linear. The angle-Doppler-domain data matrix χ, $N_t \times N_s$, represents the data at the N_t Doppler-bins and N_s angle-bins of the range cell under the hypothesis test. The same transform is also performed on the secondary data $\mathbf{Y}_k$, $k = 1, 2, \ldots, K$, where K is the number of adjacent iid cells, to obtain the angle-Doppler-domain secondary data $\mathbf{y}_k$, $N_t \times N_s$, $k = 1, 2, \ldots, K$.

In practice, only the few angle-bins covering the angle-section centred at the broadside of the array (i.e., around the look direction where most of the transmitted energy is contained) need to be tested, while at most all Doppler-bins should be tested as the target Doppler frequency shift is unknown to the processor. Let N_{s0} be the number of angle-bins of interest. The $N_t \times N_s$ bins to be tested, will be divided into L groups, each of which contains N_{s0} angle-bins and a small number of adjacent Doppler-bins. An example for this grouping is given in Figure 8.5 where $N_t = 24$, $N_s = 12$, and $N_{s0} = 3$. We note that the number of Doppler-bins in each group need not be the same and that some overlap can also be justified. The purpose of dividing along the Doppler axis is to avoid the use of an adaptive processor with large degrees of freedom, which demands

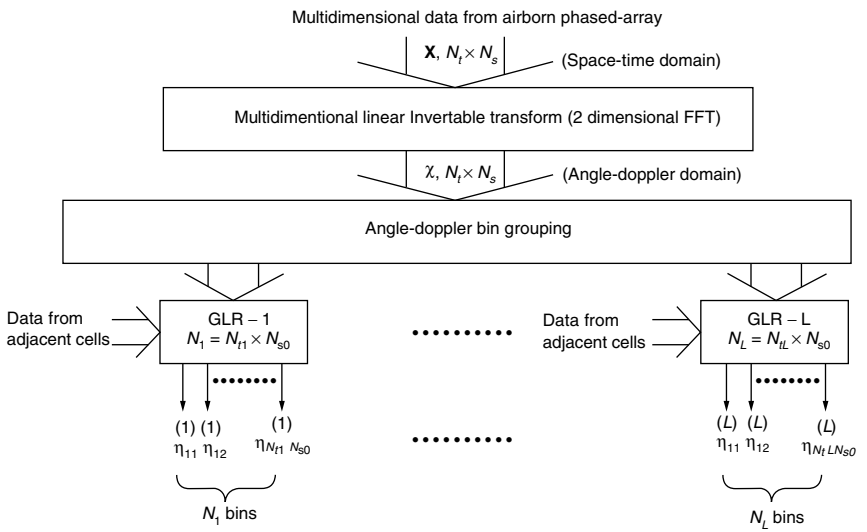

FIGURE 8.4 Block diagram for illustration of the principle of the JDL–GLR processor.

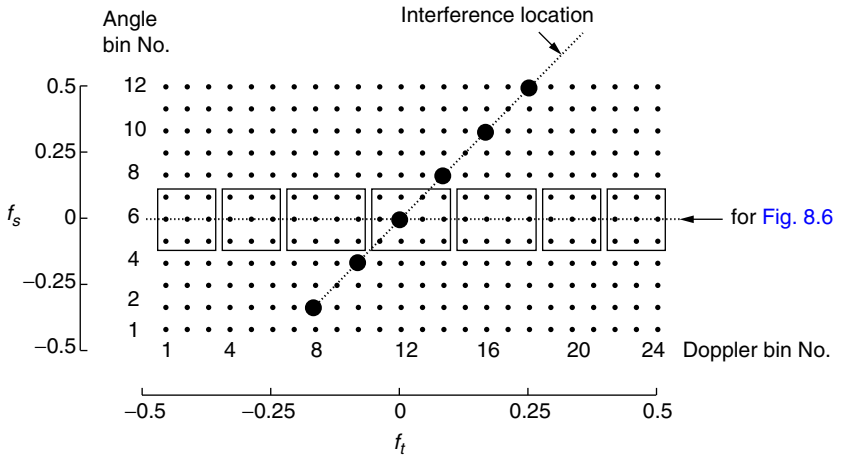

FIGURE 8.5 An JDL–GLR example.

a large training-data set as well as a large amount of computation. This opportunity of "divide-and-conquer" is, of course, made available by the multi-dimensional transformation from the space–time data domain to the angle-Doppler-domain, which decouples the degrees of freedom necessary for handling complicated clutter or interference, from the number of data points to be processed. Based on our experience gained from the work in Refs. 7 and 8, the number of bins in each group is expected to have only minor influence on the detection performance and should be in the range of $2 \times N_{s0} \sim 4 \times N_{s0}$ in general. The angle-Doppler-domain secondary data $\mathbf{y}_k$, $k = 1, 2, \ldots, K$ should be grouped in the same way.

Let N_{tl} be the number of Doppler-bins and $N_l = N_{tl} \times N_{s0}$ the total number of angle-Doppler-bins in the lth group. An N_lth order GLR processor will perform the threshold detection on the N_l bins of the lth group with the test statistic

$$\eta_{nm}^{(l)} = \frac{|\mathrm{Vec}(\mathscr{S}_{nm}^{(l)})^H \hat{\mathbf{R}}_l^{-1} \mathrm{Vec}(\chi_l)|^2}{\mathrm{Vec}(\mathscr{S}_{nm}^{(l)})^H \hat{\mathbf{R}}_l^{-1} \mathrm{Vec}(\mathscr{S}_{nm}^{(l)})[1 + \mathrm{Vec}(\chi_l^H) \hat{\mathbf{R}}_l^{-1} \mathrm{Vec}(\chi_l)]} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_0^{(1)}; \tag{8.31}$$

$$n = 1, 2, \ldots, N_{tl} \qquad m = 1, 2, \ldots, N_{s0}$$

where

$$\hat{\mathbf{R}}_l = \sum_{k=1}^{K} \mathrm{Vec}(\mathbf{y}_{lk}) \mathrm{Vec}(\mathbf{y}_{lk})^H \tag{8.32}$$

and $\mathscr{S}_{nm}^{(l)}$, $N_{tl} \times N_{s0}$, is the signal-steering matrix in the angle-Doppler-domain for the nmth bin of the lth GLR. For a uniform PRF and array spacing, it is easy to see that $\mathscr{S}_{nm}^{(l)}$ has all its entries equal to zero except the nmth one which is $\sqrt{N_t N_s}$.

We note that the threshold $\eta_0^{(l)}$ need not be the same across the L groups as evidenced in Section 8.4.2 below.

8.4.2. The JDL–GLR Detection Performance

The detection performance of the original GLR in Gaussian clutter or interference is given in Refs. 3 and 4 with deterministic modelling and in Ref. 12 with stochastic target modelling. As for the Doppler-domain localized GLR of Refs. 7 and 8, it is straightforward to extend the results in Refs. 3, 4 and 12 to obtain the probabilities of detection and false alarm, P_{d} and P_{f}, of the JDL–GLR with both target models. Below we just list the results for the case of non fluctuating targets with the trivial derivation omitted.

The probability of detection at the nmth bin of the lth GLR is found to be

$$P_{\mathrm{d}}^{(l)}(n,m) = \int_1^0 P_{\mathrm{d}|\rho}^{(l)}(n,m) f_{nm}^{(l)}(\rho)\mathrm{d}\rho \tag{8.33}$$

where

$$P_{\mathrm{d}|\rho}^{(l)}(n,m) = 1 - (1-\eta_0^{(l)})^{K-N_\ell+1} \sum_{k=1}^{K-N_\ell+1} \binom{K-N_t+1}{k} \left(\frac{\eta_0^{(l)}}{1-\eta_0^{(l)}}\right)^k \mathrm{e}^{-\rho\beta_{nm}^{(l)}(1-\eta_0^{(l)})} \times \sum_{m=0}^{K-1} \frac{[\rho\beta_{nm}^{(l)}(1-\eta_0^{(l)})]^m}{m!} \tag{8.34}$$

$$f_{nm}^{(l)}(\rho) = \frac{(K)!}{(K-N_l+1)!(N_l-2)!}\rho^{K-N_\ell+1}(1-\rho)^{N_\ell-2} \tag{8.35}$$

and

$$\beta_{nm}^{(l)} = |\alpha|^2 \mathrm{Vec}(\mathscr{S}_{nm}^{(l)})^H \mathscr{R}_l^{-1} \mathrm{Vec}(\mathscr{S}_{nm}^{(l)}) \tag{8.36}$$

with $\mathscr{R}$ being the covariance matrix of $\mathrm{Vec}(\chi_l)$.

The probability of false alarm for all bins in the lth GLR is given by

$$P_{\mathrm{f}}^{(l)} = (1-\eta_0^{(l)})^{K-N_l+1} \tag{8.37}$$

Obviously the probability of false alarm can be made equal across the L groups by choosing different $\eta_0^{(l)}$, $l = 1, 2, \ldots, L$. Equation 8.37 also clearly indicates that, like the original GLR and the Doppler-domain localized GLR, the JDL–GLR has the "integrated/embedded" CFAR feature as $P_{\mathrm{f}}^{(l)}$, $l = 1, 2, \ldots, L$ do not depend on the covariance of the clutter or interference.

8.4.3. Detection Performance Comparison

Although the convergence-rate advantage of the JDL–GLR can be seen intuitively from the fact that the localized GLR's have much lower degrees of freedom

than a high-order GLR directly applied to the space–time domain data, the numerical example below should demonstrate this advantage clearly.

Consider a system with $N_s = 12$ and $N_t = 24$. The clutter or interference is assumed to have the two-dimensional multipeak Gaussian-shaped power spectrum density (PSD) as shown in Figure 8.6. For convenience of reference we have also indicated the centre locations of this multipeak spectrum in Figure 8.5. The exact expression of this PSD is given by

$$P_c(f_t,f_s) = \sum_{d=1}^{6} \sigma_{cd}^2 \frac{1}{2\pi\sigma_{ft}\sigma_{fs}} \exp\left[-\left(\frac{(f_t - f_{ctd})^2}{2\sigma_{ft}^2} + \frac{(f_s - f_{csd})^2}{2\sigma_{fs}^2} \right) \right] \quad (8.38)$$

where: σ_{cd}^2 is the power of the dth component. Obviously, the total clutter or interference power σ_c^2 is

$$\sigma_c^2 = \sum_{d=1}^{6} \sigma_{cd}^2 \quad (8.39)$$

We set $\sigma_{c1}^2 = \sigma_{c2}^2 = \sigma_{c4}^2 = \sigma_{c5}^2 = \sigma_{c6}^2 = \sigma_{c3}^2/10^{2.5}$, INR $= 50$, and SNR $= 0$ dB which gives SINR $\backsimeq -50$ dB. The thresholds for the processors to be compared are such that every processor has a probability of false alarm $P_f = 10^{-5}$ at each tested bin. We assume that there are $K = 24$ adjacent cells from which the iid secondary data set is obtained.

Consider the following five processors:

(1) the joint-domain optimal,
(2) the JDL–GLR with $L = 7$ localized GLR processors with their coverage shown in Figure 8.5,
(3) the T–S configuration with the optimal processor for each part,

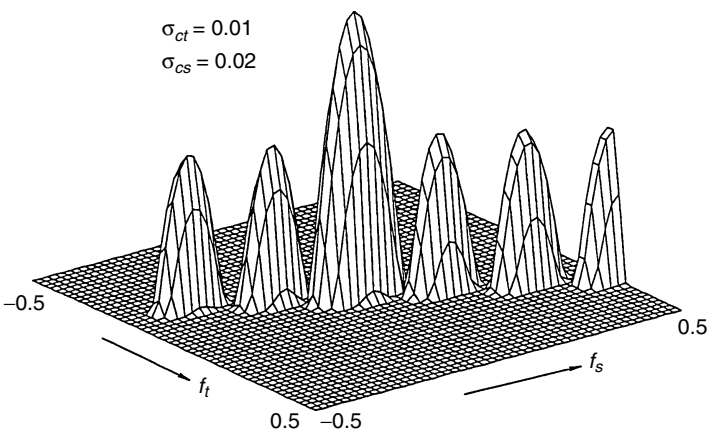

FIGURE 8.6 Two-dimensional power spectral density for the clutter or interference used in the example.

(4) the S–T configuration with the optimal processor for each part, and
(5) a conventional beamformer followed by the optimal temporal processor (i.e., the optimal MTI).

We note that with $N_s = 12$, $N_t = 24$ but $K = 24$ only, any straightforward adaptive implementation of the joint-domain optimal, any adaptive processor with the S–T configuration, or any adaptive processor with the T–S configuration will fail to deliver an acceptable detection performance for this example since $K = 24$ is too small with respect to their degrees of freedom. Therefore, these adaptive processors are excluded from the above list for detailed comparison.

Figure 8.7 shows the probability of detection of the five processors listed in Figure 8.7, at the sixth angle-bin which is the assumed angle of arrival of the target signal. Obviously, the JDL–GLR is the only one that approaches the joint-domain optimal, except at a few bins adjacent to the centre of the strongest clutter or interference spectrum component. The poor performance of the two optimal cascade configurations should not be a surprise from the discussion in Section 8.3. The fact shown in Figure 8.7 that the *ad hoc* processor (BF + opt. MTI) can outperform the two cascade configurations (especially the optimal S–T configuration) is also strong evidence that optimality does not always mean much with a wrong configuration. The optimal S–T configuration gives the poorest performance because its spatial processor, in nulling the clutter or interference, also nulls the target signal. Finally, we comment that a CFAR loss is

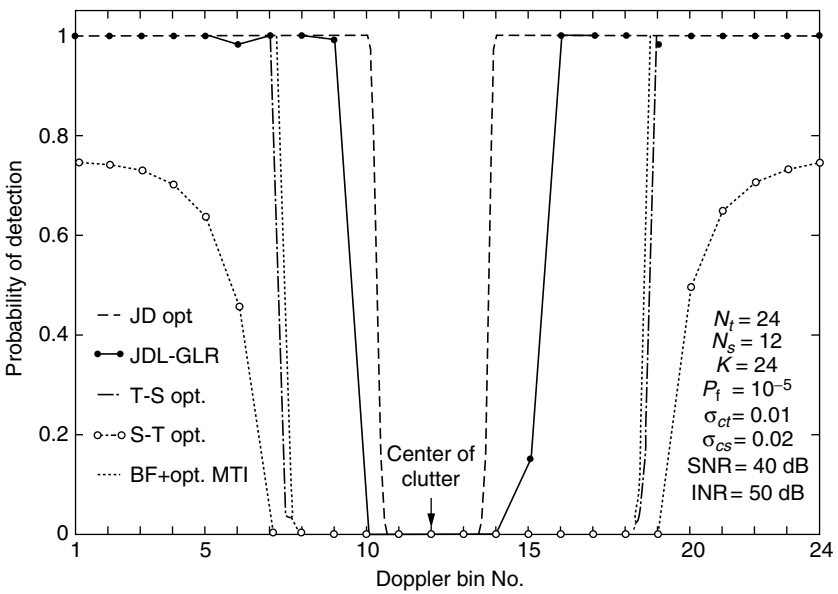

FIGURE 8.7 Detection performance comparison of the five processors.

inevitably associated with any adaptive implementation of the four optimal or partially optimal processors in Figure 8.7, while the embedded CFAR feature of the JDL–GLR makes any other additional CFAR processing unnecessary.

8.4.4. Other Features of JDL–GLR

The computation advantage of the JDL–GLR is clear. Recall that the Nth order GLR has a computation load proportional to N^3. Assume that each localized GLR spans three angle-bins and four Doppler-bins and that $N_t/4$ localized GLRs are required. This leads to a computation load proportional to $(N_t/4)(3\times 4)^3 = 432N_t$ for the JDL–GLR. With a load of $N_t^3N_s^3$ for the straightforward application of the GLR to the space–time domain data, the JDL–GLR will show a computation advantage when $N_t > 4$ and $N_s > 3$. For large N_t and N_s the JDL–GLR offers a computation load reduction by a factor of

$$\gamma = N_t^2N_s^3/432 \tag{8.40}$$

For the example of $N_t = 24$ and $N_s = 12$ in this section, the JDL–GLR's computation load is only 1/2304th of that for the straightforward application of the GLR (or SMI) to the space–time domain data. Like the Doppler-domain localized GLR in Refs. 7,8, the JDL–GLR can further reduce its computation load via deleting the localized GLR processors for the region where the detection performance improvement is unnecessary or impossible. This can be done when some *a priori* information is available about the power concentration of the clutter or interference in the angle-Doppler-domain. Furthermore, the realization of the JDL–GLR benefits from the available parallel processing techniques as its localized GLRs all operate in parallel.

Since the robustness feature in non Gaussian clutter or interference resides with the GLR processor which will not be affected by the linear transformation, the JDL–GLR is expected to maintain its robustness. Computationally intensive simulation is being conducted to confirm this feature and the result will be published separately.[13]

8.5. CONCLUSIONS AND DISCUSSION

This chapter shows:

(1) none of the two cascade configurations is better than the other;
(2) the performance potential of both cascade configurations can fall far below that of the joint-domain optimum processor; and
(3) the JDL–GLR algorithm offers an attractive solution to the problem of approaching the performance potential of the joint-domain optimum processor of a high order $(N_s \times N_t)$ with a fast convergence-rate and high computation efficiency, together with such highly desirable features as the embedded CFAR and robustness in nonGaussian clutter or interference.

We would like to point out that both cascade configurations may have considerable difficulty to achieve a high-quality CFAR in practice when spatial and temporal parts are both adaptive. This is because of the random modulation introduced by the adaptive algorithm for the early part of the cascaded two parts. The problem may become more severe in highly nonstationary and non-homogeneous environments where there is a shortage of a sufficient amount of iid training-data to smooth out the extra random modulation. In contrast, the JDL–GLR presented in this chapter is free of such random modulation and can maintain its CFAR performance with a much smaller amount of iid training-data. Simulation-based comparison, of the CFAR performance of adaptive spatial–temporal processors, can be found in Ref. 13.

9 A Printed-Circuit Smart Antenna with Hemispherical Coverage for High Data-Rate Wireless Systems

G. Ploussios

Whereas an *adaptive* antenna utilizes space–time adaptive processing to minimize its radiation pattern in the direction or directions of external electromagnetic "noise", a *smart* antenna utilizes space–time adaptive processing *or* up-dated knowledge (assuming such information is available) to achieve the same objective. A smart antenna is therefore a more general class of antenna but not necessarily as complex or effective as an adaptive antenna. This chapter describes a recently disclosed[1] smart antenna that is well-suited to high data-rate wireless systems.

This two-port, four-element antenna consists of a unique element configuration and microstrip-feed which provides angle and polarization diversity in an efficient, low-cost, minimum size package (a cylindrical volume $0.1\lambda_0$ high $\times$ $0.6\lambda_0$ in diameter) operating at free-space microwave wavelengths $\lambda_0 \approx 3$ to 30 cm. The antenna has a radiation pattern that is approximately uniform over the entire upper hemisphere and has an operating bandwidth of 10 to 25%. The antenna is well-suited for high data-rate wireless systems on fixed or mobile platforms subject to interference and multipath fading. Conventional designs are of either (a) comparable volume with more restricted angular coverage (very poor low-angle coverage) and limited system bandwidth or (b) comparable performance but more costly and bulky (several times larger in volume), and therefore impractical for most of the applications.

The antenna [Figure 9.1(a and b)] consists of two pairs of elements, the pairs oriented at 90° in azimuth from each other, and each pair (designated a "doublet") with its own port. Each element of the doublet is a quarter-wave resonant bent monopole with a common ground plane oriented in the direction of the horizon.

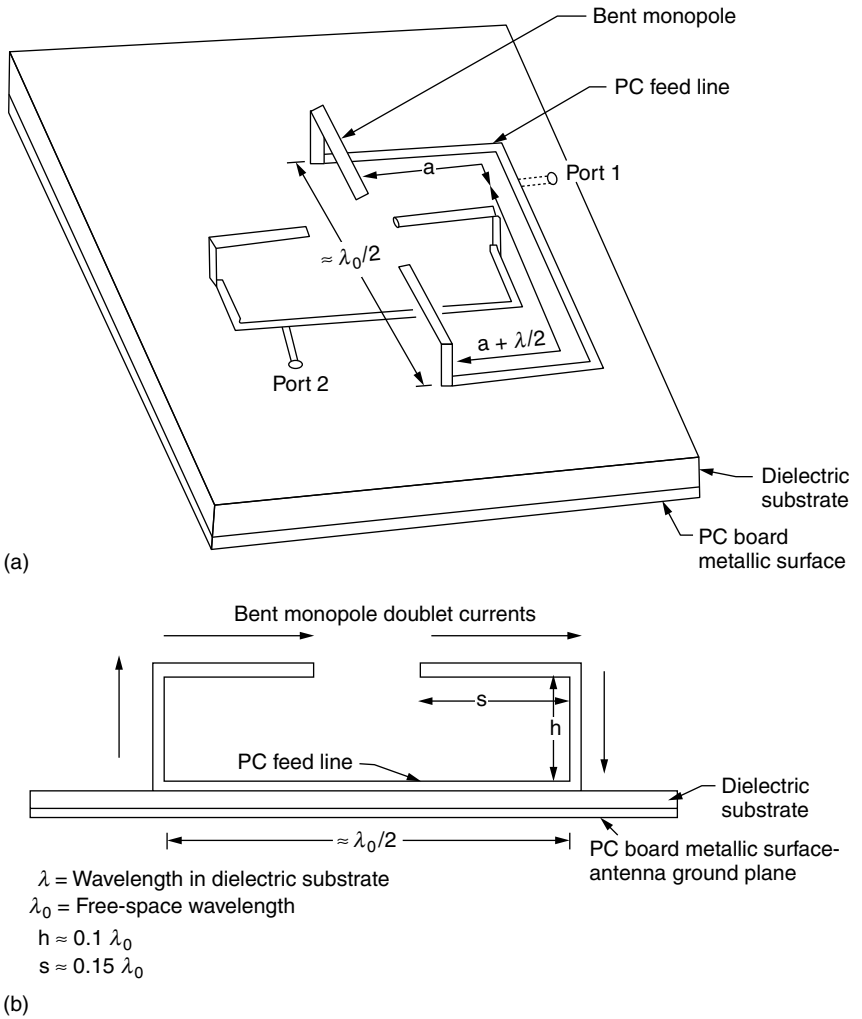

FIGURE 9.1 Smart antenna configuration: (a) bent monopole elements and feed lines, (b) monopole doublet currents.

Each bent element, comprising an electrically short vertical segment ($\approx 0.1\lambda_0$) and a larger horizontal segment ($\approx 0.15\lambda_0$), radiates nearly uniformly over the entire upper hemisphere because each segment has approximately the same peak gain.[2] The two vertical segments of each doublet are separated in space from each other by approximately $0.5\lambda_0$, are each fed approximately 180° out-of-phase (0.5λ in the dielectric substrate) with each other by a microstrip transmission line circuit, and have a combined figure-eight azimuthal pattern with coverage near the horizon but no overhead coverage. The two horizontal segments of each doublet have a figure-eight azimuthal pattern with overhead coverage but reduced

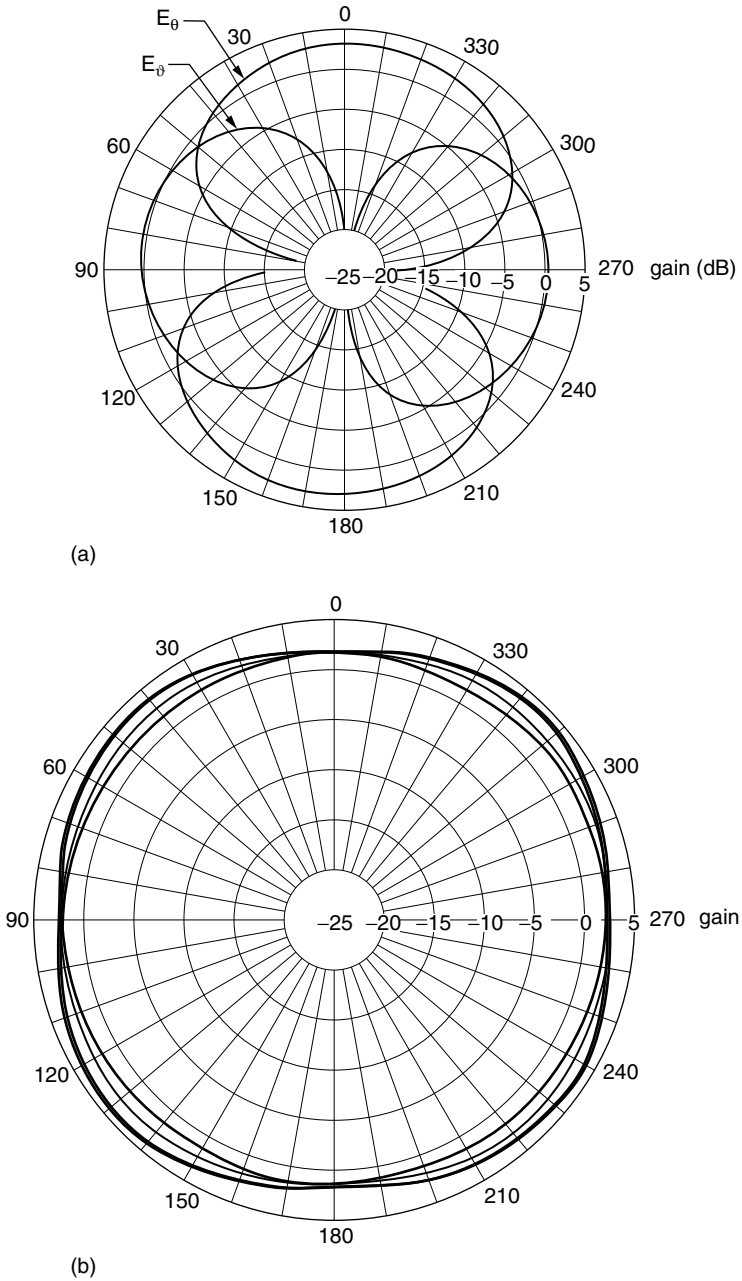

FIGURE 9.2 Single-port azimuthal patterns: (a) E_θ and E_ϕ polarizations (at an elevation angle $\theta = 30°$) at one port of Figure 9.1. The second port patterns are rotated 90°. (b) Circular polarization (at elevation angles $\theta = 10°$, 30°, 60°, 70°, and 90°) formed by combining, with a 90° combiner, the outputs of the two ports of Figure 9.1.

coverage near the horizon. The net result is that each port has figure eight azimuthal patterns for E_θ and E_ν polarization components [Figure 9.2(a)]. The single-port principal plane E_θ and E_ν elevation patterns are very broad, peaking overhead and then decreasing by 3 and 8 dB, respectively, at the horizon. The two ports collectively produce E_θ and E_ν free-space coverage with nearly uniform gain over the entire upper hemisphere. When combined with a 90° combiner, the two ports produce a single-port circularly polarized output with nearly uniform hemispherical coverage [Figure 9.2(b)]. Each of the E_θ and E_ϕ components (not shown) of the circularly polarized output, has similar hemispherical coverage, nominally reduced in power by approximately 3 dB. Alternately, the ports can be (a) selected for maximum ratio of signal-to-interference plus noise, or (b) linearly combined to generate elliptically polarized patterns. In summary, the small smart antenna configuration of Figure 9.1 provides both polarization and angle diversity (including the nulling of interfering signals). The stand-alone antenna configuration of Figure 9.1 can also be arrayed to form a high-gain smart antenna.

10 Applications

CONTENTS

10.1. Cancellation of Specular and Diffuse Jammer Multipath Using a Hybrid Adaptive Array 446
(*R. L. Fante*)
10.1.1. Introduction 446
10.1.2. Why Multipath Requires Additional Degrees of Freedom 446
10.1.3. Generalization 451
10.1.4. Numerical Calculations 456
10.1.5. Summary and Discussion 460
10.2. Some Limitations on the Effectiveness of Airborne Adaptive Radar 461
(*E. C. Barile, R. L. Fante, and J. A. Torres*)
10.2.1. Background 461
10.2.2. Theoretical Introduction 465
10.2.3. Two-Element Displaced Phase Center Antenna 472
10.2.4. Simulation Results 478
10.2.4.1. Internal Clutter Motion 478
10.2.4.2. Aircraft Crabbing 482
10.2.4.3. Near-Field Obstacles 484
10.2.4.4. Antenna Errors (Channel Mismatch) 487
10.2.5. Summary 490
10.3. Clutter Covariance Smoothing by Subaperture Averaging 490
(*R. L. Fante, E. C. Barile, and T. P. Guella*)
10.3.1. Introduction 490
10.3.2. Analysis for an Airborne Radar 492
10.3.3. Summary 496
10.4. Cancellation of Diffuse Jammer Multipath by an Airborne Adaptive Radar 497
(*R. L. Fante and J. A. Torres*)
10.4.1. Introduction 497
10.4.2. Filtered Received Signals 502
10.4.2.1. Received Jammer and Noise Signals 502
10.4.2.2. Interference Covariance Matrix 505
10.4.2.3. Steering-Vector and Received Target Signal 508

10.4.3. Numerical Results 509
10.4.3.1. Introduction 509
10.4.3.2. Tap Spacing 512
10.4.3.3. Total Extent 513
10.4.3.4. Ground Clutter 513
10.4.3.5. Temporal Averaging 514
10.4.3.6. Beam Space 514
10.4.4. Summary and Discussion 517
10.5. Wideband Cancellation of Multiple Mainbeam Jammers 518
(*R. L. Fante, R. M. Davis, and T. P. Guella*)
10.5.1. Introduction 518
10.5.2. Calculation of the Array Performance 520
10.5.3. Simulation Results 523
10.5.3.1. Spatial Span and Location of the Auxiliaries 523
10.5.3.2. Required Number of Auxiliaries and Gain per Auxiliaries 524
10.5.3.3. Signal-to-Interference Ratio after Cancellation 527
10.5.3.4. Simultaneous Nulling of Mainlobe and Sidelobe Jammers 529
10.5.4. Summary and Discussion 530
10.6. Adaptive Space–Time Radar 531
(*R. L. Fante*)
10.6.1. Introduction 531
10.6.2. Understanding the Results in Equation 10.169 and Equation 10.170 533
10.6.3. Sequential Cancellation of Jammers and Clutter 536
10.6.4. Typical Results 538
10.6.5. Additional Considerations 539
10.6.6. Summary 540
10.7. Synthesis of Adaptive Monopulse Patterns 540
(*R. L. Fante*)
10.7.1. Analysis 540
10.7.2. Summary 542
10.8. Ground and Airborne Target Detection with Bistatic Adaptive Space-Based Radar 543
(*R. L. Fante*)
10.8.1. Introduction 543
10.8.2. Analysis 544
10.8.2.1. Sum Beam 544
10.8.2.2. Difference Beam 545
10.8.3. Numerical Studies of Effectiveness 546
10.8.3.1. Sum Beam 548
10.8.3.2. Difference beam 551
10.8.4. Summary 553

10.9. Adaptive Nulling of Synthetic Aperture Radar (SAR) Sidelobe Discretes 553
(*R. L. Fante*)
10.9.1. Introduction 553
10.9.2. Fully Adaptive SAR 554
10.9.3. Overlapped-Subarray SAR 557
10.9.4. Numerical Results 559
10.9.5. Summary 563
10.10. Wideband Cancellation of Interference in a Global Positioning System (GPS) Receive Array 563
(*R. L. Fante and J. J. Vaccaro*)
10.10.1. Introduction 563
10.10.2. Adaptive Filter Weights 564
10.10.2.1. Maximum Signal-to-Interference Ratio 565
10.10.2.2. Minimum Mean Square Error 566
10.10.2.3. Minimum Output Power 567
10.10.3. Signal Distortion Introduced by the Processor 567
10.10.4. Suboptimum Space–Frequency Processing 570
10.10.5. Numerical Simulations 571
10.10.5.1. Introduction 571
10.10.5.2. Effect of Channel Mismatch 574
10.10.5.3. Effect of Steering-Vector Mismatch 576
10.10.5.4. Distortion Introduced by the Adaptive Filter 577
10.10.6. Space–Time vs. Suboptimum Space–Frequency Processing 580
10.10.7. Summary 585
10.11. A Maximum-Likelihood Beamspace Processor for Improved Search and Track 585
(*R. M. Davis and R. L. Fante*)
10.11.1. Introduction 585
10.11.2. Maximum-Likelihood Beamspace Processor (MLBP) 586
10.11.3. Analysis 589
10.11.3.1. The First Stage 589
10.11.3.2. The Second Stage 590
10.11.3.3. Target Detection 592
10.11.4. Numerical Examples 593
10.11.4.1. Improved Clear Environment Search Performance 594
10.11.4.2. Improved Clear Environment Angle Estimation 595
10.11.4.3. Performance against a Single Mainlobe Interferer 596
10.11.5. Summary 600

10.1. CANCELLATION OF SPECULAR AND DIFFUSE JAMMER MULTIPATH USING A HYBRID ADAPTIVE ARRAY

(R. L. FANTE)

10.1.1. Introduction

Most analyses[1,4] of adaptive cancellation of strong jammers consider only the direct signal from the jammer, and ignore any multipath components[5] scattered from the Earth. For a smooth Earth the multipath consists of only a single time-delayed, specularly reflected ray, but for a rough Earth the multipath consists[6] of many time-delayed, diffusely reflected components. The question then arises as to how one can cancel both the direct jammer signal and these multiple reflections. There are a number of choices: one can add more spatial degrees of freedom to the adaptive array, more temporal degrees of freedom, or a combination of both. Additional spatial degrees of freedom can be achieved by using additional auxiliary antenna elements. The additional temporal degrees of freedom can be achieved by using bandwidth partitioning (with a separate adaptive loop in each subband), an adaptive finite impulse response (FIR) filter,[5] or a hybrid system that uses both bandwidth partitioning and adaptive FIR filters. This work is devoted to a study of such hybrid systems. In particular, we study an ideal two-element array that uses bandwidth partitioning in both the main and auxiliary channels, with an Mth-order adaptive FIR filter in each subband of the auxiliary. We then study the ability of this system to cancel specular, moderately diffuse and diffuse multipath, and perform tradeoffs to determine what combinations of bandwidth partitioning and filter order can achieve a specified jammer cancellation level.

10.1.2. Why Multipath Requires Additional Degrees of Freedom

In order to see why jammer multipath is a problem let us first consider the ideal two-element canceler shown in Figure 10.1. Let y_m denote the voltage received at the main antenna terminals at time $t = m\Delta$ and x_m, the voltage received on the auxiliary terminals at that time. Then the residue r_m at the output is

$$r_m = y_m - wx_m \tag{10.1}$$

where the weight w is chosen to minimize the mean square residue. Using this optimum weight it is readily shown that the minimum mean square residue is[1–4]

$$\langle |r|^2 \rangle = \langle |y|^2 \rangle - \frac{|\langle x^* y \rangle|^2}{\langle |x|^2 \rangle} \tag{10.2}$$

where $\langle\ \rangle$ denotes an expectation, and we have removed the subscript m from r_m, x_m, and y_m because the expectations are independent of m for a stationary random process.

We now wish to calculate the residue in Equation 10.2 for the case of a jammer and a single, specularly reflected, multipath ray assuming that both the direct and reflected jammer signals are much stronger than both any desired

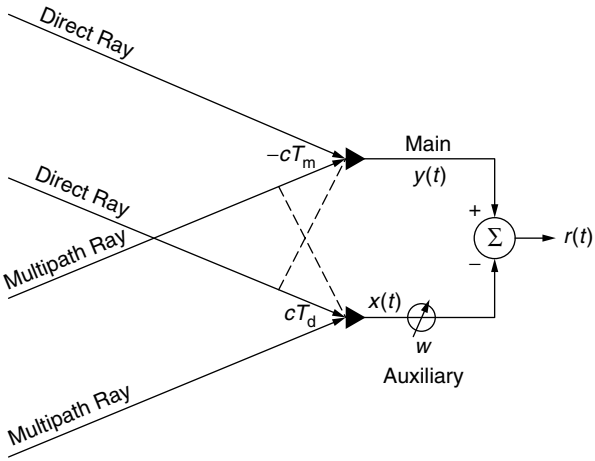

FIGURE 10.1 Ideal two-element adaptive array.

signal and the system noise. Consequently, we ignore both signal and noise in the calculations to follow, because the weights in the adaptive system are then driven by the jammer alone.

Let us denote the direct jammer signal at the main antenna element in Figure 10.1 by $j(t)\exp(\mathrm{i}\omega_0 t)$ where ω_0 is the radian carrier frequency. Also, suppose that the relative strength of the multipath signal at the main antenna is ρ, and its delay relative to the direct jammer signal is t_1. Then, the total jammer signal at the main antenna element is

$$y(t) = j(t) + \rho j(t - t_1)\exp(-\mathrm{i}\omega_0 t_1) \tag{10.3}$$

where a common term, $\exp(\mathrm{i}\omega_0 t)$, has been ignored in Equation 10.3.

Next, let us consider the voltage in the auxiliary channel. If both the jammer and its reflection point on the ground are in the far field of the array in Figure 10.1, we can write the voltage in the auxiliary channel as

$$x(t) = j(t - T_\mathrm{d})\exp(-\mathrm{i}\omega_0 T_\mathrm{d}) + \rho j(t - T_\mathrm{m} - t_1)\exp[-\mathrm{i}\omega_0(T_\mathrm{m} + t_1)] \tag{10.4}$$

where T_d is the time delay between the main and auxiliary elements for the direct ray, and T_m is the delay between the main and auxiliary elements for the multipath ray. We also assume that the jammer power spectral density is uniform and occupies a bandwidth much larger than the bandwidth B_r of the receiver, this latter bandwidth is determined by the bandwidth of the desired signal. Then, because the power spectral density is uniform over the receiver bandwidth B_r, the autocorrelation function $\langle j(t)j^*(t+\tau)\rangle$ is

$$\langle j(t)j^*(t+\tau)\rangle = \mathrm{sinc}(\pi B_\mathrm{r}\tau) \tag{10.5}$$

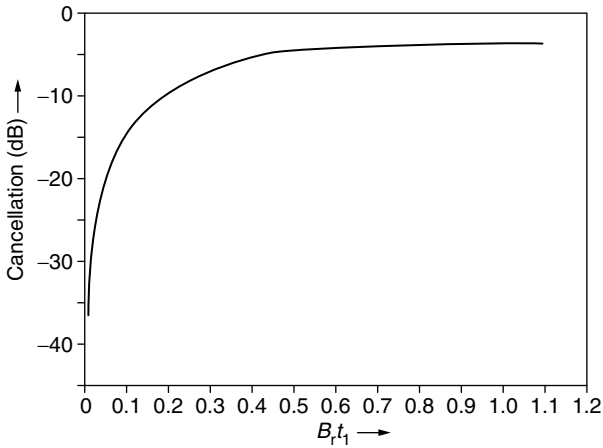

FIGURE 10.2 Cancellation for $\omega_0 t_1 = (2n+1)\pi$, $\omega_0 T_m = (2p+1/2)\pi$, $T_d = 0$, $T_m/t_1 = 0.02$, n and p are integers.

If Equation 10.3 and Equation 10.4 are used in Equation 10.2, with Equation 10.5 used to calculate the required expectations, we obtain an expression for the average residue $\langle |r|^2 \rangle$. This residue, normalized to the power $\langle |y|^2 \rangle$ received in the main channel, has been plotted in Figure 10.2 as a function of bandwidth for the case when $\rho = 0.5$, $T_d = 0$, $\omega_0 t_1 = (2n+1)\pi$, $n =$ integer, $\omega_0 T_m = (2p + 1/2)\pi$, $p =$ integer and $T_m/t_1 = 0.02$. We note that, for a given delay t_1 between the direct and multipath rays, the adaptive canceler in Figure 10.1 is most effective for $B_r t_1 \ll 1$ and ineffective for $B_r t_1 > 1$, because, for the latter case, the direct and multipath signals are then decorrelated and more degrees of freedom are needed to cancel both of them. We find later that diffuse multipath makes matters even worse. Nevertheless, it is clear from Figure 10.2 that if we can make $B_r t_1$ sufficiently small the adaptive canceler does cancel both the direct and multipath jammer signals. This suggests partitioning[7–9] the total band B_r into N subbands of width $B = B_r/N$, as shown in Figure 10.3. This allows us to introduce new degrees of freedom by using an independent weight in each subband, and also provides for channel equalization. In this case, it is demonstrated in Appendix E that the average output power is

$$\int_{-\infty}^{\infty} |r(t)|^2 dt = \sum_{k=0}^{N-1} \langle |r_k|^2 \rangle \tag{10.6}$$

where $\langle |r_k|^2 \rangle$ is the average residue in the kth subband. Likewise, in the absence of an auxiliary channel the average output power is

$$\sum_{k=0}^{N-1} \langle |y_k|^2 \rangle$$

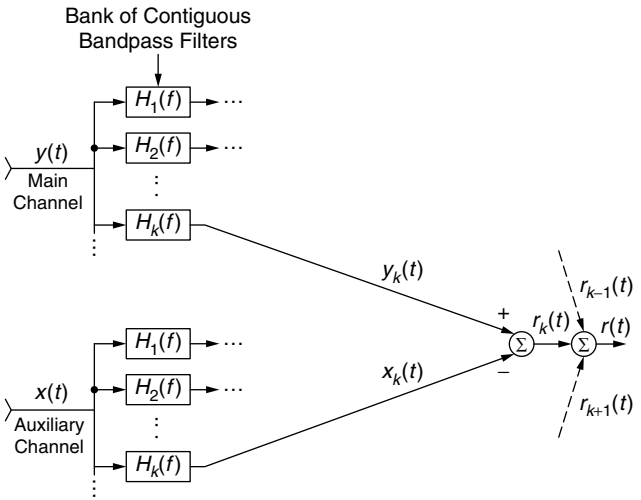

FIGURE 10.3 Partition of B_r into subbands.

so that the normalized residue power, or cancellation ratio, can be defined as

$$C = \frac{\sum_{k=0}^{N-1} \langle |r_k|^2 \rangle}{\sum_{k=0}^{N-1} \langle |y_k|^2 \rangle} \tag{10.7}$$

For specular (diffuse multipath will be studied later) multipath it can be demonstrated that the value of C for given values of Bt_1 and BT_m is nearly independent of N provided N is sufficiently large so that $NBt_1 \gg 1$ and $NTB_m \gg 1$. Therefore, if N is sufficiently large, one can generate universal asymptotic curves for the normalized residue, or equivalently, the cancellation ratio. These asymptotic results are shown in Figure 10.4 for five different multipath reflection coefficients ρ. Although we do not present the details here, analytic approximations to these curves can be developed in the limits when $Bt_1 \ll 1$ and $Bt_1 > 1$. For $Bt_1 \ll 1$ and $BT_m \ll 1$, we can use the Taylor series expansion for $\text{sinc}(\pi B\tau)$ in the expressions for $\langle |r_k|^2 \rangle$ and $\langle |y_k|^2 \rangle$. If we use this, along with the fact that when $NBt_1 \gg 1$ and $NBT_m \gg 1$, the summations over k of $\exp(\pm i2\pi kBt_1)$ and $\exp(\pm i2\pi kBT_m)$ are nearly zero, we find that, provided $T_d = 0$, $\rho < 1$ and $t_1 \gg T_m$

$$C = \frac{2}{3}\left(\frac{\pi\rho}{1+\rho^2}\right)^2 G(\rho)(Bt_1)^2 \tag{10.8a}$$

where

$$G(\rho) = \left[1 - \left(\frac{2\rho}{1+\rho^2}\right)^2\right]^{-1/2} \tag{10.8b}$$

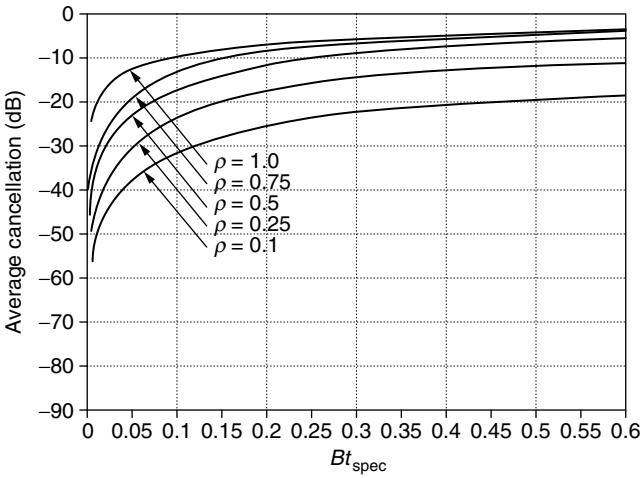

FIGURE 10.4 Average cancellation for specular multipath and one tap ($M = 1$) as subband width B is varied. Average is over 5000 subbands.

It should be noted that C is not necessarily zero when there is no multipath ($\rho = 0$), because we have set $T_d = 0$. In the limit when $\rho = 0$ and $T_d \neq 0$ we find $C = (\pi B T_d)^2/3$.

In the opposite limit when $Bt_1 \gg 1, NBt_1 \gg 1, NBT_m \gg 1$ but $BT_m < 1$ it is easy to show that

$$C = 2\left(\frac{\rho}{1+\rho^2}\right)^2 \tag{10.9}$$

Within their ranges of validity both Equation 10.8a and Equation 10.9 are excellent approximations to the numerical results presented in Figure 10.4.

In order to see how these results are used, let us assume that we need to compensate for specular jammer multipath using bandwidth partitioning alone (later we see the effect of combining adaptive FIR filters with bandwidth partitioning). Suppose the specular-multipath delay t_1 is 10^{-6} s, the receiver bandwidth $B_r = 3 \times 10^6$ Hz, and we wish to obtain a jammer cancellation of at least 40 dB when $\rho = 0.5$. Then, from Figure 10.4, we find that $Bt_1 \simeq 0.6 \times 10^{-2}$ is required to achieve $C = 10^{-4}$. As we see later, diffuse multipath makes matters worse, so in order to allow a safety factor in case the multipath is diffuse, rather than specular, we can decrease Bt_1 by a factor of 2 to obtain $Bt_1 \simeq 0.3 \times 10^{-2}$, or $B = 0.3 \times 10^{-2}/t_1 = 3.0 \times 10^3$ Hz. Therefore, the number of subbands required is $N = B_r/B = 3 \times 10^6/3.0 \times 10^3 = 1000$.

If a DFT (discrete Fourier transform) rather than an analog realization is used for bandwidth partitioning we would use an $N = 2^{10} = 1024$ or $2^{11} = 2048$ point transform for the DFT shown in Figure 10.5. The formal expression for the block-averaged power in this case is given in Appendix F. Note that, as pointed out by Compton,[10] if the block processing shown in Figure 10.5 is replaced by

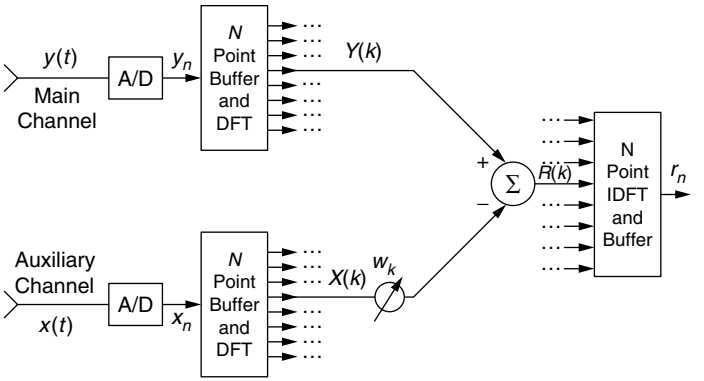

FIGURE 10.5 Bandwidth partitioning using DFTs.

a sliding-window DFT, the inverse discrete Fourier transform (IDFT) shown in Figure 10.5 is unnecessary. Of course, the actual implementation of all DFTs is done using the fast Fourier transform (FFT) algorithm.

In the next section we generalize our results to include diffusely scattered multipath and hybrid cancelers that employ both bandwidth partitioning and adaptive FIR filtering.

10.1.3. Generalization

An analog model for the hybrid canceler is shown in Figure 10.6. A filter bank partitions the main and auxiliary signals into N frequency bins, with an M-tap adaptive FIR filter in each frequency bin. The delay $D = (M-1)\Delta/2$ (for $M =$ odd) is inserted[5] into the main channel to ensure cancellation of all directions of arrival. The canceled signals from each frequency bin are then combined, producing the residue $r(t)$. Although the analog model is easier to understand, in a realistic application the processing is done digitally, with a typical digital realization shown in Figure 10.7. In this case an N-point DFT (calculated with the FFT algorithm) partitions the signal bandwidth B_{r} into N subbands of width $B = B_{\mathrm{r}}/N$. The adaptive processing is then performed in each subband, and the time samples r_{m} are recovered via an IDFT. As discussed previously, the IDFT is unnecessary if sliding-window processing replaces block processing. Also, in practice, the time samples x_{n} may be weighted before Fourier transforming, in order to reduce the frequency sidelobes in each subband.

Let us refer to Figure 10.6, and define $y_k(t)$ as the voltage produced by a jammer plus its multipath in the kth subband of the main channel at time t. Also, define $x_k(t)$ as the voltage in the kth subband of the auxiliary. Then, by generalizing Equation 10.2, it is readily shown that the residual power in the kth subband after cancellation is

$$\langle |r_k|^2 \rangle = \langle |y_k|^2 \rangle - [\mathbf{Z}_k^*]^{\mathrm{T}} [\mathbf{R}_k]^{-1} [\mathbf{Z}_k] \tag{10.10}$$

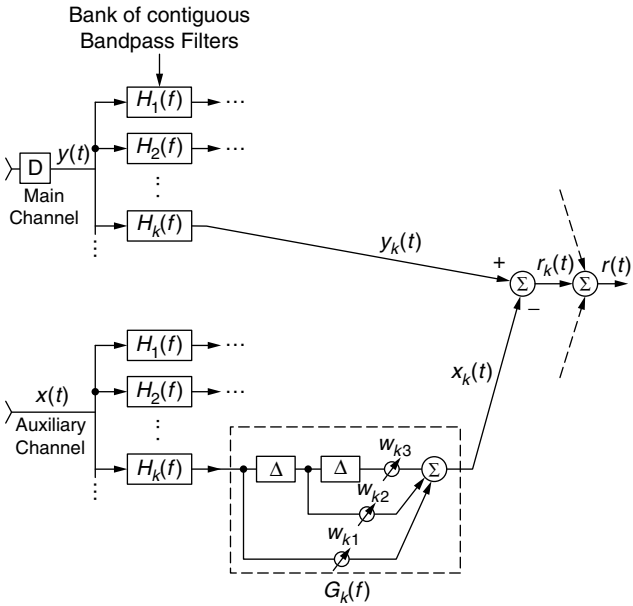

FIGURE 10.6 Analog realization of hybrid canceler.

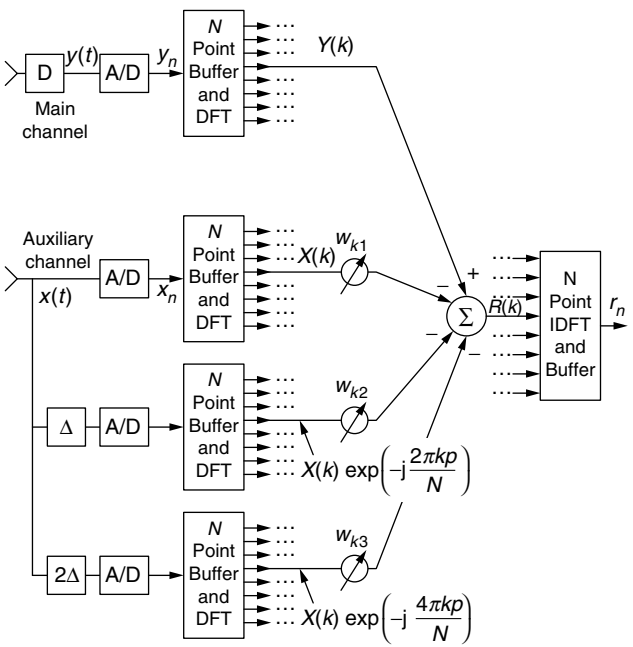

FIGURE 10.7 One possible digital realization of hybrid canceler.

where

$$[\mathbf{R}_k] \equiv \begin{bmatrix} \langle x_k^*(t)x_k(t)\rangle & \langle x_k^*(t)x_k(t-\Delta)\rangle \cdots \\ \langle x_k^*(t-\Delta)x_k(t)\rangle & \langle x_k^*(t-\Delta)x_k(t-\Delta)\rangle \cdots \\ \vdots & \vdots \end{bmatrix} \tag{10.11a}$$

$$[\mathbf{Z}_k] \equiv \begin{bmatrix} \langle x_k^*(t)y_k(t)\rangle \\ \langle x_k^*(t-\Delta)y_k(t)\rangle \\ \vdots \\ \langle x_k^*(t-(M-1)\Delta)y_k(t)\rangle \end{bmatrix} \tag{10.11b}$$

where $\langle\rangle$ denotes an expectation, and we have assumed that there are M taps in the adaptive filter in the kth auxiliary channel. The optimum weights for the kth subband are given by

$$[w_k] = [\mathbf{R}_k]^{-1}[\mathbf{Z}_k] \tag{10.12}$$

We next need to discuss the form of $x_k(t)$ and $y_k(t)$ when the multipath is nonspecular. Diffuse multipath can be modeled by using the glistening surface approach developed by Beckmann and Spizzichino.[6] A typical glistening surface is shown in Figure 10.8, for the flat-Earth approximation. The shaded area in that figure represents the region on the ground producing diffusely scattered jammer multipath. In Ref. 6, an expression is derived for the diffusely scattered power, which becomes particularly simple in the limit when the jammer and the radar are at the same altitude ($h_1 = h_2$). In that limit the total diffuse multipath power received by the radar is

$$P_{\text{diffuse}} = |\rho|^2 \int_{\xi A}^{\pi - \xi A} f(\xi)\mathrm{d}\xi \tag{10.13a}$$

where $f(\xi)$ is plotted in Figure 10.9, ρ is the Fresnel reflection coefficient of the ground, and ξ, ξ_A, and K_β are defined as $\sin^2(\xi/2) = X_1/R$, $\cos^2(\xi/2) = X_2/R$,

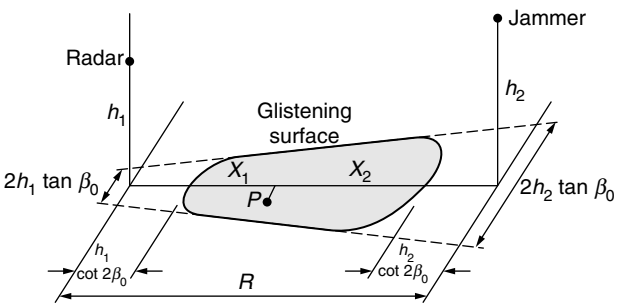

FIGURE 10.8 Glistening surface for diffuse multipath.

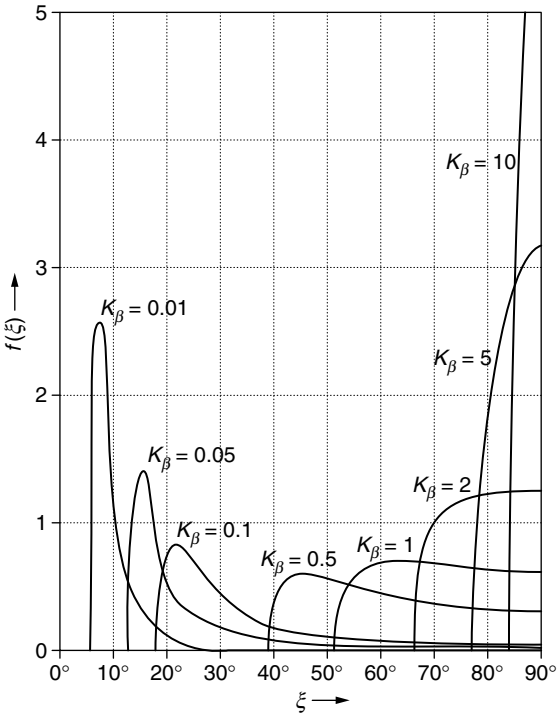

FIGURE 10.9 Function $f(\xi)$. Curves symmetrical with respect to $\xi = 90°$.

$K_\beta = 2h_1/(R\tan\beta_0)$, $\tan\beta_0 = 2\sigma/T$, and $\sin\xi_A \tan\xi_A = K_\beta$. As seen in Figure 10.8, X_1 and X_2 are the projections of an arbitrary point P on the glistening surface, R is the range from the jammer to the radar ($R = X_1 + X_2$), and σ and T are the standard deviation of the surface roughness and its transverse correlation length, respectively. The integral

$$I \equiv \int_{\xi A}^{\pi - \xi A} f(\xi)\mathrm{d}\xi \tag{10.13b}$$

that appears in Equation 10.13a has been evaluated in Ref. 6 and is replotted here in Figure 10.10.

If we now refer back to the definition of K_β, it is evident that $K_\beta \ll 1$ corresponds to a diffuse (rough) surface, $K_\beta \simeq 1$ to a moderately diffuse surface; and $K_\beta \gg 1$ to a specular (smooth) surface. Consequently, from Figure 10.9 it is seen that the multipath from a diffuse surface comes primarily from two regions: one near the jammer and one near the radar. For example, when $K_\beta = 0.05$ the multipath comes mainly from areas centered on $\xi = 15°$ and $\xi = 180° - 15° = 165°$. Likewise, when $K_\beta = 0.5$ (moderately diffuse) the multipath is nearly uniformly distributed from $\xi = 39°$ to $\xi = 180° - 39° = 141°$.

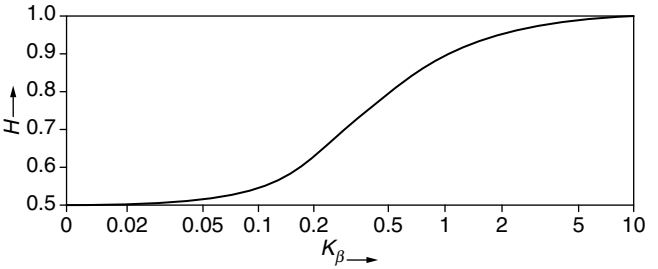

FIGURE 10.10 Integral I in Equation 10.12.

Finally, for $K_\beta \gg 1$ we approach the limit of a single reflection point located at $\xi = 90°$.

The aforementioned multipath contributions can be modeled by a set of Q discrete, independent scatterers with complex amplitudes $A_q \exp(\mathrm{i}\phi_q - \mathrm{i}\omega_0 t_q)$, where the phases ϕ_q are all independent and randomly distributed from 0 to 2π. The amplitudes A_q and the relative time delays t_q of the qth multipath component are chosen in accordance with distribution of $f(\xi)$ in Figure 10.8. Of course, for specular multipath $Q = 1$ and $t_1 = 2(h_1^2 + R^2/4)^{1/2} - R$. The amplitudes A_q and the number of scatterers Q must be chosen to represent the multipath such that

$$\sum_{q=1}^{Q} |A_q|^2 = |\rho|^2 I \tag{10.14}$$

That is, the incoherent addition of all the scatterer powers must equal the total diffusely scattered power given by Equation 10.13a.

When the multipath is represented in accordance with the method described above, the signals $y(t)$ and $x(t)$ in the main and auxiliary channels of either Figure 10.6 or Figure 10.7 may be written as

$$y(t) = j(t - D) + \sum_{q=1}^{Q} A_q j(t - D - t_q)\mathrm{e}^{-\mathrm{i}\omega_0(t_q + D) + \mathrm{i}\phi_q} \tag{10.15}$$

and

$$x(t) = j(t - T_\mathrm{d})\mathrm{e}^{-\mathrm{i}\omega_0 T_\mathrm{d}} + \sum_{q=1}^{Q} A_q j(t - T_{\mathrm{m}q} - t_q)\exp[-\mathrm{i}\omega_0(t_q + T_{\mathrm{m}q}) + \mathrm{i}\phi_q] \tag{10.16}$$

where $T_{\mathrm{m}q}$ is the delay between the main and auxiliary elements for the qth multipath ray, $\omega_0 = 2\pi f_0$, f_0 is the frequency at the center of the desired signal bandwidth B_r, t_q is the time delay at the main element between the direct jammer ray and the qth multipath ray, T_d is the delay of the direct jammer ray between the main and auxiliary channels, $D = (M - 1)\Delta/2$ for $M =$ odd and ϕ_q is a phase randomly distributed between 0 and 2π.

The results in Equation 10.15 and Equation 10.16 are then used to calculate the appropriate matrix elements in Equation 10.10 to Equation 10.12 for each subband, using the result that for the kth subband $\langle j(t)j(t+r)\rangle = \exp(i2\pi kB\tau)\text{sinc}(\pi\beta\tau)$, where $k = 0, \pm 1, \pm 2\ldots$, and k has been ordered such that $k = 0$ corresponds to the subband centered at the middle of the total band B_r. Once both $\langle |y_k|^2\rangle$ and the residue $\langle |r_k|^2\rangle$ have been obtained for each subband, the normalized average power is obtained by summing over all the subbands, as indicated in Equation 10.2. The matrix elements required to compute these quantities are summarized in Appendix G.

10.1.4. Numerical Calculations

Before we proceed to evaluate how well hybrid cancelers perform against various types of jammer multipath, we first need to select an appropriate delay Δ for the adaptive FIR filter in Figures 10.6 and 10.7. When the normalized multipath delay Bt_q is not too close to zero the residue is rather insensitive to the choice of $B\Delta$, as long as $B\Delta$ is not too close to either zero or unity. This point was shown in Ref. 5, Figure 11.19.

However, when $Bt_q \ll 1$ we have found numerically (see Appendix H) that the cancellation is best if $B\Delta$ is of the same order as Bt_q. However, because one does not know apriori what multipath delay to expect, one usually designs for the maximum delay. Consequently, as a compromise we chose $B\Delta = 0.5$, recognizing that this may not be the optimum choice for $Bt_q \ll 1$. Thus, the curves for $M = 3$, 5, and 9 in Figures 10.11 to 10.13 to follow do not necessarily represent the very best we can do with tapped delay lines when $Bt_q \ll 1$. This point is discussed further in Appendix H.

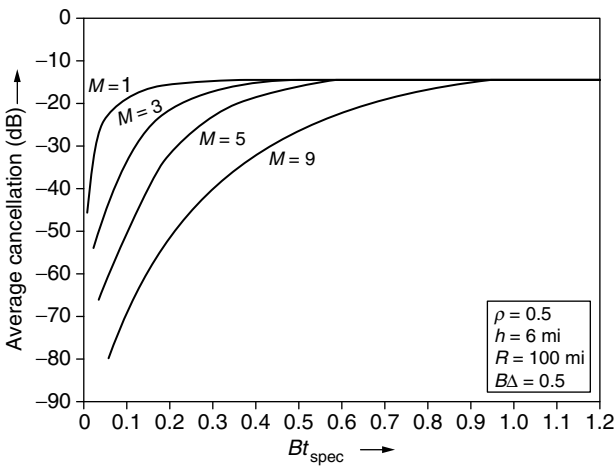

FIGURE 10.11 Average cancellation for diffuse multipath and $\rho = 0.5$.

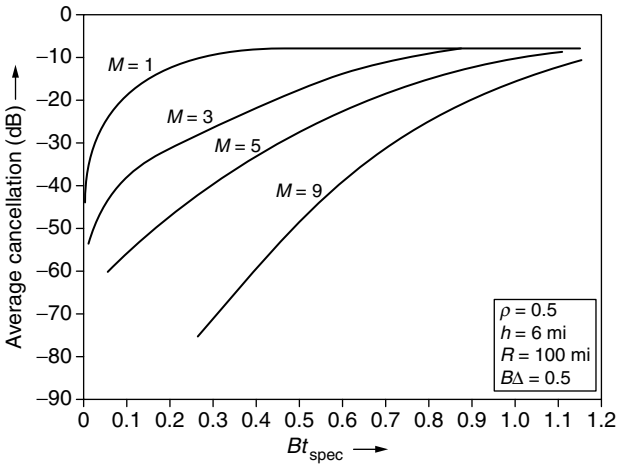

FIGURE 10.12 Average cancellation for moderately diffuse multipath and $\rho = 0.5$.

In order to assess the effect of various types of multipath on performance, we now consider the case when the radar and jammer are separated by 100 statute miles and both are at an altitude of 6 statute miles above ground with a Fresnel reflection coefficient $\rho = 0.5$. For this case, diffuse ($K_\beta = 0.1$), moderately diffuse ($K_\beta = 0.5$) and specular ($K_\beta \gg 1$) multipath are modeled in accordance with the procedure outlined in the preceding section. The sidelobe canceler is modeled as an ideal two-element array with the elements separated by one wavelength at midband. The results are somewhat sensitive to element separation, and we can expect to obtain different residues for other element spacings, although the trends will not change. For all cases, the carrier frequency

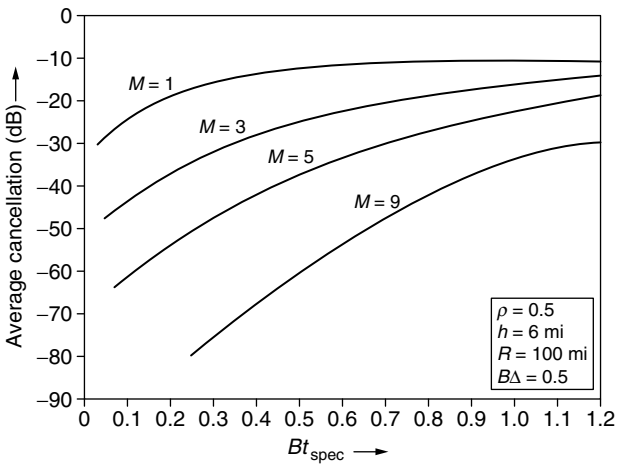

FIGURE 10.13 Average cancellation for specular multipath and $\rho = 0.5$.

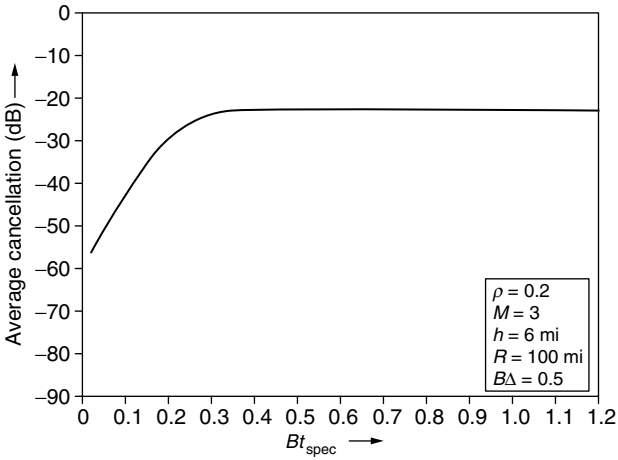

FIGURE 10.14 Average cancellation for diffuse multipath and $\rho = 0.5$.

f_0 was chosen to be 100 B_r. In all results to follow, we have set $B\Delta = 0.5$, in accordance with our discussions in the last paragraph.

Results for the effect of changing B on the normalized jammer residue C after cancellation are presented in Figures 10.11 to 10.13 for diffuse moderately diffuse, and specular multipath, respectively. The effect of decreasing the reflection coefficient ρ is shown in Figure 10.14. All curves have been plotted versus Bt_{spec}, where B is the bandwidth and t_{spec} is the time delay of the specular ray, relative to the direct ray, and is given by

$$t_{spec} = 2[h_1^2 + R^2/4]^{1/2} - R \tag{10.17}$$

In obtaining the curves in Figures 10.11 to 10.14, we have not averaged over thousands of subbands, as would be required for these to be truly asymptotic results, because in order to obtain asymptotic results we require N to be sufficiently large so that NBT_{spec} and $NBT_m \gg 1$. However, for $T_m/t_{spec} = 0.02$ and $Bt_{spec} = 0.01$ this means that $NBT_m = N\,(Bt_{spec})\,(t_m/t_{spec}) = N\,(0.01)\,(0.02) > 1$, so N must be of order 5000 or more. For the $M = 9$ case, this implies 5000 to 10,000 inversions of a 9×9 matrix for each point. Thus, the $M = 1$ curve in Figure 10.13 differs from the corresponding curve in Figure 10.4, because the latter has been averaged over very many subbands. Although the average over hundreds or thousands of subbands give somewhat different results than those presented in Figures 10.11 to 10.13, the trends in both cases are identical, as shown in Appendix H.

We note from Figures 10.11 to 10.13 that, for a given subband width B and number of taps M, diffuse multipath usually represents the worst case (i.e., highest residue), although the asymptote for $Bt_{spec} \gg 1$ is somewhat lower for diffuse multipath than for specular and moderately diffuse multipath because I is smallest

for $K_\beta \ll 1$, as is evident from Figure 10.10. It is evident from Figures 10.11 to 10.13 that for a given multipath condition (e.g., diffuse) one can decrease the residue either by decreasing the subband width B, by increasing the number of taps M, or both. It is also clear, a necessary condition for effective cancellation is that the width B of each subband be sufficiently small so that $Bt_{\text{spec}} \ll 1$.

Let us now apply this result, along with the results in Figures 10.11 to 10.13, to see what combinations of M and N allow us to cancel multipath to a specified level. Suppose that we wish to suppress the jammer signal by 30 dB over a band $B_r = 2$ MHz. Then using Figures 10.11 to 10.13 (in an expanded form) we obtain the results in Figure 10.15, which show the number of subbands N required for different choices of the number of taps M. The values for $M =$ even were obtained using additional curves (not shown). In deriving the curves for $M =$ even, the delay D in the main channel was set equal to zero in the expressions in Appendix G.

Thus, the system designer can achieve the desired cancellation in a number of ways. Which combination of N and M is best will depend on a number of factors, such as how many samples are available, how many computations per output sample are required, etc. A discussion of optimum combinations is given in

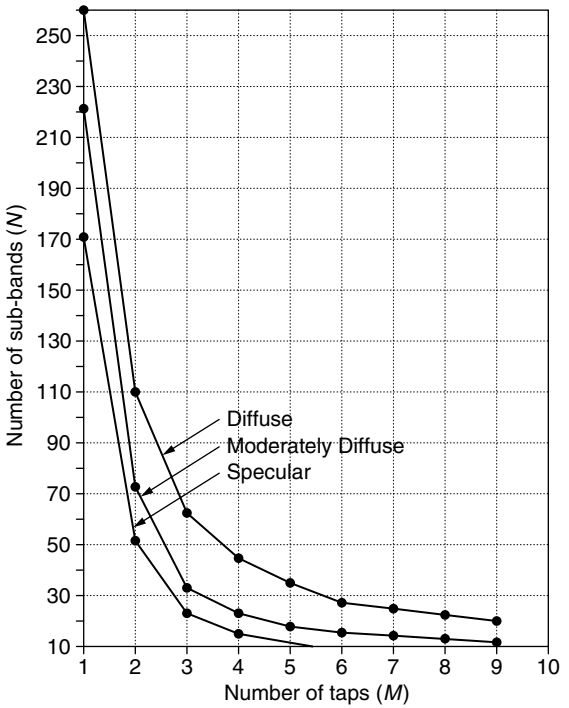

FIGURE 10.15 Number of sub-bands required to achieve 30 dB of Jammer suppression over a 2 MHz band.

Appendix I, where it is indicated that the hybrid system ($M > 1$) is clearly preferable to bandwidth partitioning alone ($M = 1$).

We emphasize that the results we have obtained for N and M (or equivalently, M and B) are for the specific conditions assumed. For lower altitudes, shorter ranges, and reduced Earth reflectivity, the restrictions are less stringent, and one can usually achieve acceptable cancellation with fewer taps per subband and fewer subbands.

10.1.5. Summary and Discussion

We have considered a two-element adaptive canceler that includes both tapped delay lines and bandwidth partitioning, and have studied the ability of this system to cancel a direct jammer signal and its diffuse multipath, assuming the jammer signal is much stronger than any desired signal or system noise.

We did not study the effect of adding more spatial (i.e., more antenna elements) degrees of freedom, although one expects that if more spatial degrees of freedom were available fewer temporal degrees of freedom would have been required. That is, because the phase delay is proportional to $(\omega/c)\sin\theta$, signals at different frequencies, but coming from the same angle θ, are equivalent to signals at the center frequency coming from different angles, and can, therefore, be canceled by putting spatial nulls in the appropriate direction.

An interesting point we should discuss is why our hybrid system (with both tapped delay lines and bandwidth partitioning) gives different results from one with bandwidth partitioning alone, because as Compton[10] has shown, time delays and DFTs are interchangeable, provided the time delay Δ is equal to the intersample period T_s. Let us now show that our system does not satisfy Compton's condition (i.e., $T_s = \Delta$). We found in Appendix D, that although the cancellation is best when the intertap delay Δ is of the order of the multipath delay t_{spec}, acceptable performance is obtained as long as $0.05 < B\Delta < 0.8$, and as a nominal value, for computing Figures 10.11 to 10.15, we chose $B\Delta = 0.5$. The bandwidth B of each subband is B_r/N, where B_r is the total bandwidth and N is the number of subbands, and is equal to the number of points in the DFT. However, the sampling theorem requires that the intersample spacing $T_s = 1/B_r$, so that $B = B_r/N = 1/NT_s$. Therefore, $\Delta = 0.5/B = 0.5\,NT_s$, so that for $N \gg 1$ we have $\Delta \gg T_s$. Consequently, Compton's condition is not satisfied by our system, and this explains why the residue is changed if the tapped delay lines are interchanged with DFTs.

It should also be noted that our results will be unchanged if, in order to remove clutter, a Doppler processor (i.e., another DFT) is placed between the DFT and the weights in each subband of Figure 10.7. The proof of this point is given in Ref. 10.

Finally, we note that the extension of this analysis to a sidelobe canceler with multiple auxiliary antenna elements is straightforward. We simply allow the main channel to have an arbitrary (voltage) gain $G(\theta)$ where θ is the angle relative to boresight, and place a network such as shown in Figure 10.7 behind each

of the S auxiliary antenna elements. If we define $x_k(i, t - p\Delta)$ as the time-delayed voltage of the ith auxiliary element in the kth subband, and then form the vector

$$[\mathbf{A}_k]^{\mathrm{T}} = [x_k(1,t)\ \ x_k(1,t-\Delta)\ \ x_k(1,t-2\Delta)\ldots x_k(2,t)\ldots x_k(S,t)\ldots]$$

we find that the $M \times M$ matrix $[\mathbf{R}_k]$ in Equation 10.11a is replaced by the $MS \times MS$ matrix $\langle[\mathbf{A}_k^*][\mathbf{A}_k]^{\mathrm{T}}\rangle$ and $[\mathbf{Z}_k]$ in Equation 10.11b is replaced by the $MS \times 1$ vector $\langle[\mathbf{A}_k^*]y\rangle$, where $\langle\rangle$ again denotes an expectation. The analysis then proceeds just as before, with the residue in the kth subband calculated using Equation 10.10, etc.

10.2. SOME LIMITATIONS ON THE EFFECTIVENESS OF AIRBORNE ADAPTIVE RADAR

(E. C. BARILE, R. L. FANTE, AND J. A. TORRES)

10.2.1. Background

Unlike a ground-based radar in which nearly all the clutter return is received at or near zero Doppler, the clutter return in an airborne radar has Doppler frequencies spread over a band of width $(4Vf_0/c)$, where V is the platform speed, f_0 is the carrier frequency, and c is the speed of light. This is illustrated in Figure 10.16. An important feature of the Doppler spectrum is that for small depression angles each Doppler f_{d} is uniquely associated with the clutter at an azimuth ϕ satisfying the relation $f_{\mathrm{d}} = (2Vf_0/c)\sin\phi$, so that in sine azimuth–Doppler space all the clutter lies along a single line, as shown in Figure 10.17. In constructing Figure 10.17 we have assumed for ease of presentation that: (i) the pulse

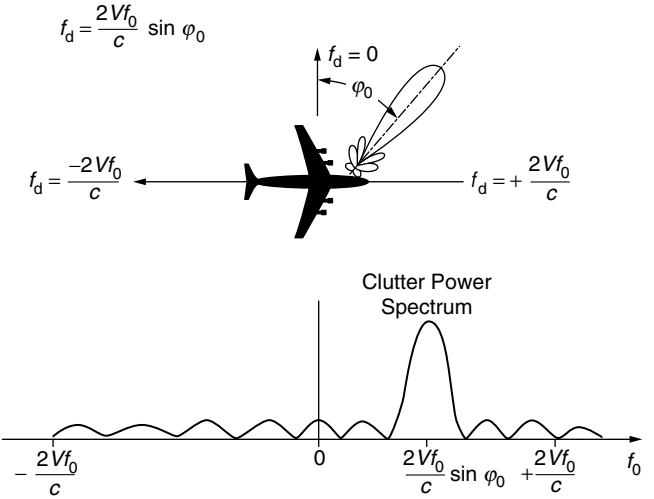

FIGURE 10.16 Doppler spectrum of received clutter for airborne radar.

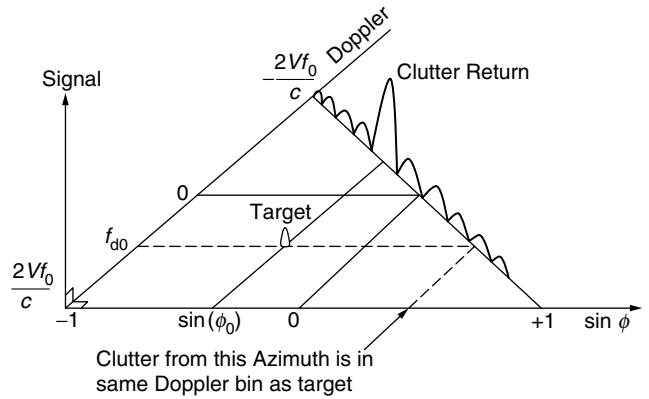

FIGURE 10.17 Clutter return in Doppler–azimuth space.

repetition frequency (PRF) is sufficiently small so that there are no range ambiguities, (ii) the depression angle θ of the range ring under consideration is such that cos θ can be approximated by unity, (iii) the clutter is stationary, and (iv) the transmit and receive patterns have a negligible response on the side of the platform opposite to the side where the main beam is directed. For high PRF radars there are multiple range ambiguities, including some where θ is no longer small. In this case, if the depression angles of the ambiguous range rings are $\theta_1 \cdots \theta_N$, the clutter at a given Doppler f_d comes from azimuth ϕ_n such that $f_d = (2Vf_0/c)\sin \phi_n \cos \theta_n$. If the clutter has internal random motion, or if there is a uniform bias in the relative velocity between the aircraft and the ground (crabbing), or if there are near-field scatterers, then the linear relation between Doppler and sin ϕ cos θ will be disturbed.

A conventional moving target indicator (MTI) is ineffective in canceling airborne clutter because it uses temporal degrees of freedom only, and, hence, produces the filtering action illustrated in Figure 10.18. Thus, MTI cancels the clutter at $\phi = 0°$ but is ineffective at canceling the rest of the clutter. It is evident that, in order to cancel along the diagonal line where the clutter lies, one must add

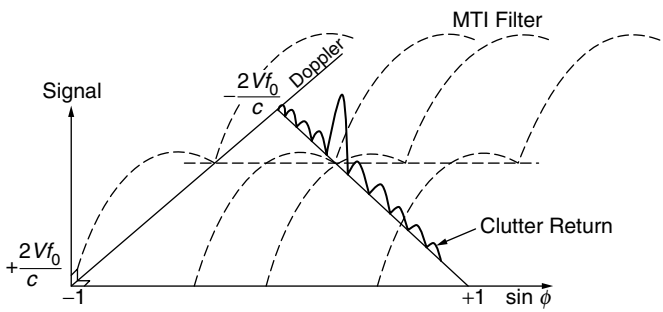

FIGURE 10.18 MTI filter superposed on spectrum of clutter.

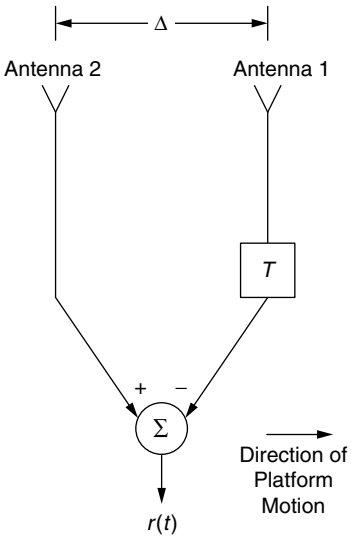

FIGURE 10.19 Two-element DPCA.

some spatial degrees of freedom. The simplest system that combines both spatial and temporal degrees of freedom is the displaced phase center antenna (DPCA). This is illustrated in its simplest form in Figure 10.19.

It is shown elsewhere[1] that if all errors are neglected and the interpulse period T is adjusted so that the element separation $\Delta = 2VT$, then the filtering response of the DPCA is as illustrated in Figure 10.20. Thus, by adding a spatial degree of freedom we have managed to rotate the MTI filter in azimuth–Doppler space so as to put a null on the clutter line.

The problem with nonadaptive DPCA is that it is sensitive to antenna errors, and requires that the platform velocity be known well enough to adjust the

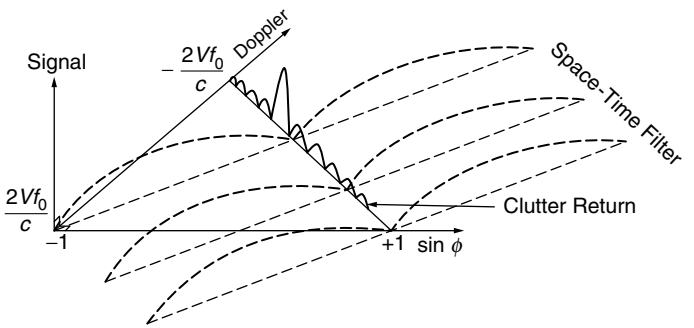

FIGURE 10.20 Space–time filter superposed on spectrum of clutter.

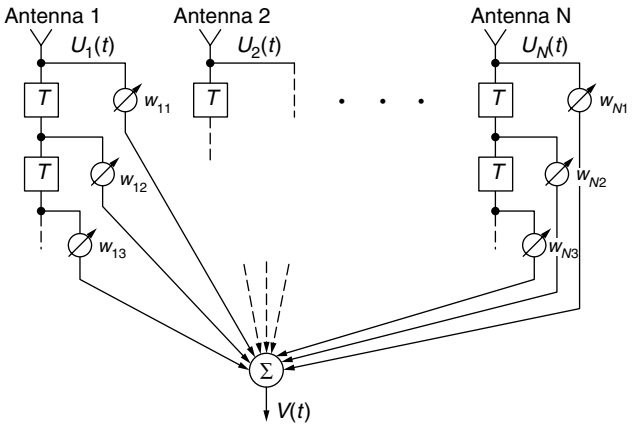

FIGURE 10.21 Generalized space–time processor.

interpulse period to satisfy the condition $T = \Delta/2V$. These difficulties can be overcome by generalizing[2,3] the design to an adaptive system, such as the one shown in Figure 10.21. In this general system the radar is tuned to a target at a specific azimuth ϕ_0 and Doppler f_{d0} and the weights w_{11}, w_{12},… are then adaptively adjusted to maximize the signal-to-clutter-plus-noise ratio. This causes the main beam of the radar to be scanned to ϕ_0, and places a null in its azimuth–Doppler pattern along the clutter line shown in Figure 10.17. Thus, it would appear that one can completely eliminate the clutter. However, in practice this is not the case because of internal clutter motion, crabbing, channel mismatch, and scattering from near-field obstacles, such as the wing on the airborne platform.

Internal clutter motion and crabbing limit cancellation because, as noted earlier, they spread the clutter off the diagonal line in Figure 10.17. Near-field scattering is a problem because, as illustrated in Figure 10.22, in the presence of a near-field obstacle the energy received at a Doppler f_d no longer comes only from

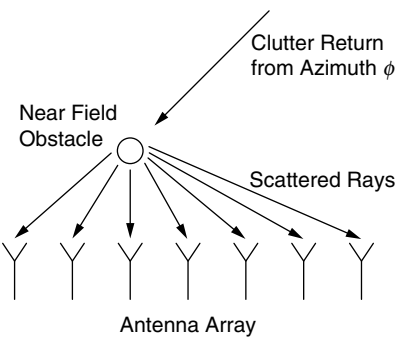

FIGURE 10.22 Scattering by near-field obstacle.

the azimuth ϕ satisfying $f_d = (2Vf_0/c)\sin\phi$, but, rather, from many different azimuths. Thus, near-field obstacles tend to smear the clutter return off the diagonal line in Figure 10.17, and into the entire sine azimuth–Doppler plane. Consequently, we should expect that clutter cancellation in the presence of a near-field obstacle will require more degrees of space–time freedom than are necessary in the absence of any obstacles.

In this chapter we investigate, quantitatively, the limitations placed on space–time cancellation of clutter by internal clutter motion, crabbing, channel mismatch, and near-field obstacles. We first develop very simple analytical models to illustrate these effects, and then proceed to more complex models that are studied numerically.

10.2.2. Theoretical Introduction

A block diagram of a space–time processor is shown in Figure 10.21. The diagram consists of N antennas with L temporal taps having delay T and weights that are adaptively controlled to maximize the signal-to-clutter ratio. If the clutter produces signals $U_1(t), U_2(t)\ldots U_N(t)$ at the terminals of antennas 1 through N, then the output signal $v(t)$ is (ignoring noise)

$$\begin{aligned} v(t) = {} & w_{11}U_1(t) + w_{12}U_2(t) + \cdots w_{1N}U_N(t) + w_{21}U_1(t-T) + w_{22}U_2(t-T) \\ & + \cdots w_{2N}U_N(t-T) + \cdots w_{L1}U_1[t-(L-1)T] + \cdots w_{LN}U_N[t-(L-1)T] \end{aligned} \tag{10.18}$$

The output $v(t)$ may then be subjected to Doppler processing. For example, one could form the samples $v(0), v(T), v(2T)\cdots v[(k-1)T]$, and then take their discrete Fourier transform to yield a frequency decomposition of the space–time output. We do not discuss Doppler processing here.

Equation 10.18 can be rewritten in matrix notation as

$$v = [\mathbf{w}]^{\mathrm{T}}[\mathbf{U}] \tag{10.19}$$

where

$$[\mathbf{w}] = \begin{bmatrix} w_{11} \\ w_{12} \\ \vdots \\ w_{1N} \\ w_{21} \\ \vdots \\ w_{LN} \end{bmatrix}, \quad [\mathbf{U}] = \begin{bmatrix} U_1(t) \\ U_2(t) \\ \vdots \\ U_N(t) \\ U_1(t-T) \\ \vdots \\ U_N(t-(L-1)T) \end{bmatrix}$$

and w_{ij} is the weight for the ith tap on the jth antenna element. The average clutter power can then be written as

$$C = \langle |[\mathbf{w}]^{\mathrm{T}}[\mathbf{U}]|^2 \rangle = [\mathbf{w}^*]^{\mathrm{T}} \langle [\mathbf{U}]^* [\mathbf{U}]^{\mathrm{T}} \rangle [\mathbf{w}] = [\mathbf{w}^*]^{\mathrm{T}} [\mathbf{M}][\mathbf{w}] \tag{10.20}$$

where $\langle\rangle$ denotes an expectation, and $[\mathbf{M}] = \langle [\mathbf{U}^*][\mathbf{U}]^{\mathbf{T}} \rangle$.

When channel noise is included, the covariance matrix $[\mathbf{M}]$ is replaced by

$$[\mathbf{M}'] = [\mathbf{M}] + \sigma^2 [\mathbf{I}] \tag{10.21}$$

where σ^2 is the noise power (assumed to be the same in each channel) and $[\mathbf{I}]$ is the $NL \times NL$ identity matrix.

Next, assume that a point target is present at a given azimuth ϕ_0 and Doppler f_{d0}, and produces an $LN \times 1$ signal vector

$$[\mathbf{s}] = \begin{bmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_N(t - (L-1)T) \end{bmatrix} \tag{10.22}$$

The receive weighting applied to the antenna array to reduce sidelobes is, of course, included in $[\mathbf{s}]$. The corresponding signal power (in the absence of clutter and noise) at the output is given by

$$S = |[\mathbf{w}]^{\mathrm{T}}[\mathbf{s}]|^2 \tag{10.23}$$

Therefore, the signal-to-clutter ratio is

$$\frac{S}{C} = \frac{|[\mathbf{w}]^{\mathrm{T}}[\mathbf{s}]|^2}{[\mathbf{w}^*]^T [\mathbf{M}'][\mathbf{w}]} \tag{10.24}$$

It can be shown (see Refs. 2,4) that the weights that maximize the signal-to-noise-plus-clutter ratio are given by

$$[\mathbf{w}] = \mu [\mathbf{M}']^{-1} [\mathbf{s}^*] \tag{10.25}$$

where μ is a constant, and the corresponding maximum signal-to-noise-plus-clutter ratio is

$$\left(\frac{S}{N + C} \right)_{\max} = [\mathbf{s}]^{\mathrm{T}} [\mathbf{M}']^{-1} [\mathbf{s}^*] \tag{10.26}$$

Next, we recognize that, because $[\mathbf{M}] \equiv \langle [\mathbf{U}^*][\mathbf{U}]^{\mathrm{T}} \rangle$, $[\mathbf{M}']$ is a Hermitian, positive-definite matrix. Consequently, $[\mathbf{M}']^{-1}$ is also Hermitian and positive definite. Furthermore, because $[\mathbf{M}']$ is Hermitian and positive definite, all eigenvalues of $[\mathbf{M}']$ are real and positive, and all its eigenvectors are orthogonal. The matrices $[\mathbf{M}']$ and $[\mathbf{M}']^{-1}$ can then be expanded in terms of the eigenvectors $[\mathbf{e}_k]$

and eigenvalues λ_k of $[\mathbf{M}]$ as

$$[\mathbf{M}'] = \sum_k (\lambda_k + \sigma^2)[\mathbf{e}_k][\mathbf{e}_k]^+ \tag{10.27a}$$

$$[\mathbf{M}']^{-1} = \sum_k \frac{1}{\lambda_k + \sigma^2}[\mathbf{e}_k][\mathbf{e}_k]^+ \tag{10.27b}$$

where $+$ denotes a conjugate transpose. It is evident that if $\mathbf{e}_k$ is an eigenvector of $[\mathbf{M}]$ it is also an eigenvector of $[\mathbf{M}']$, because if

$$[\mathbf{M}][\mathbf{e}_k] = \lambda_k[\mathbf{e}_k] \tag{10.27c}$$

then

$$[\mathbf{M}'][\mathbf{e}_k] = [\mathbf{M}][\mathbf{e}_k] + \sigma^2[\mathbf{I}][\mathbf{e}_k] = (\lambda_k + \sigma^2)[\mathbf{e}_k] \tag{10.27d}$$

If we substitute Equation 10.27b into Equation 10.25, it is immediately evident that the maximum signal-to-noise-plus-clutter ratio is

$$\left(\frac{S}{C+N}\right)_{\max} = \sum_k \frac{|[\mathbf{s}]^{\mathrm{T}}[\mathbf{e}_k]|^2}{\lambda_k + \sigma^2} \tag{10.28}$$

The corresponding weight vector is

$$[\mathbf{w}] = \mu \sum_k \frac{([\mathbf{e}_k]^{\mathrm{T}}[\mathbf{s}])^*}{\lambda_k + \sigma^2}[\mathbf{e}_k] \tag{10.29}$$

We can use Equation 10.29 to determine the azimuth–Doppler response of the space–time processor. Suppose $[\mathbf{s}_u]$ is the signal vector produced by a signal at an azimuth ϕ and Doppler f_{d}, which is different from the signal vector $[\mathbf{s}]$ to which the filter is tuned. Then, using Equation 10.23 and Equation 10.29 we find that the response to $[s_u]$ is

$$P(f_{\mathrm{d}}, \phi) = |[\mathbf{w}]^{\mathrm{T}}[\mathbf{s}_u]|^2 = \mu^2 \left| \sum_k \frac{([\mathbf{e}_k]^{\mathrm{T}}[\mathbf{s}])^*([\mathbf{e}_k]^{\mathrm{T}}[\mathbf{s}_u])}{\lambda_k + \sigma^2} \right|^2 \tag{10.30}$$

By fixing f_{d} and varying ϕ we can obtain the azimuth response for a fixed Doppler. Likewise, by fixing ϕ and varying f_{d} we obtain the Doppler response in a given Azimuth cut.

Based on numerical studies by Klemm,[5] one expects roughly $L + N$ distinct eigenvalues associated with the clutter and $LN - (L + N)$ much smaller eigenvalues that can be associated with the noise. (In the next section we calculate these large and small eigenvalues for DPCA.) If, for the moment, we ignore the noise σ^2, it can be seen from Equation 10.28 and Equation 10.30 that, because λ_k appears in the denominator, both the signal-to-noise-plus-clutter ratio and the radiation pattern are dominated by the behavior of the smallest eigenvalues. Which of these small eigenvalues will actually dominate will depend on which one has the largest value of $|[\mathbf{s}]^{\mathrm{T}}[\mathbf{e}_k]|^2/\lambda_k$. Let us suppose this

happens to be the pth eigenvalue. Then, Equation 10.28 and Equation 10.30 can be approximated by

$$\left(\frac{S}{C+N}\right)_{\max} = \frac{|[\mathbf{s}]^{\mathrm{T}}[\mathbf{e}_p]|^2}{\lambda_p} \tag{10.31}$$

$$P(f_{\mathrm{d}}, \phi) \approx \mu^2 \frac{|[\mathbf{e}_p]^{\mathrm{T}}[\mathbf{s}]|^2}{\lambda_p^2} |[\mathbf{e}_p]^{\mathrm{T}}[\mathbf{s}]|^2 \tag{10.32}$$

Numerical studies of the radiation pattern for cases when the noise is negligible (so that the pattern is determined by the structure of the pth eigenvector $[\mathbf{e}_p]$) have shown that the pattern has many undesirable features. These include splitting and distortion of the main beam and some high sidelobes. This problem can be cured by increasing the system noise σ^2, so that $\sigma^2 \gg \lambda_p$. When this is done a single eigenvector no longer dominates the radiation pattern, and in place of Equation 10.32 we get

$$P(f_{\mathrm{d}}, \phi) \approx \frac{\mu^2}{\sigma^4} \left| \sum_k{}' ([\mathbf{e}_k]^{\mathrm{T}}[\mathbf{s}])^* ([\mathbf{e}_k]^{\mathrm{T}}[\mathbf{s}_u]) \right|^2 \tag{10.33}$$

where the prime on the summation indicates that the summation is over all the small eigenvalues (i.e., over all the eigenvalues associated with the noise). This has the effect of smoothing the radiation pattern, and yields well-behaved patterns. This point is illustrated later in Section 10.2.4.

One additional general point that should be considered before we proceed to specialized studies is the effect of tapering on the receive pattern in sine azimuth–Doppler space. In order to study this effect, consider the voltage receive pattern given by

$$V(f_{\mathrm{d}}, \phi) = [\mathbf{s}_u]^{\mathrm{T}}[\mathbf{w}] = [\mathbf{w}]^{\mathrm{T}}[\mathbf{s}_u] \tag{10.34}$$

where the power pattern $P(f_d, \phi)$ in Equation 10.30 is defined as $P(f_{\mathrm{d}}, \phi) = |V(f_{\mathrm{d}}, \phi)|^2$. If we substitute for $[\mathbf{w}]$ from Equation 10.25, and then add and subtract a term we find

$$V(f_{\mathrm{d}}, \phi) = [\mathbf{s}_u]^{\mathrm{T}}[\mathbf{M}]^{-1}[\mathbf{s}^*] = [\mathbf{s}_u]^{\mathrm{T}}[\mathbf{s}^*] - [\mathbf{s}_u]^{\mathrm{T}}([\mathbf{I}] - [\mathbf{M}]^{-1})[\mathbf{s}^*] \tag{10.35}$$

where $[\mathbf{I}]$ is the identity matrix. The first term on the right-hand side of Equation 10.32 is the ambient azimuth–Doppler pattern of the unadapted array (i.e., all weights set equal to unity) and the second term represents a narrow ridge-beam in sine azimuth–Doppler space that subtracts out the clutter return shown in Figure 10.17. The resulting radiation pattern is shown in Figure 10.23. Unfortunately, when the steering vector is not tapered, as is the case for the results shown in Figure 10.23, the sidelobe response in azimuth and Doppler is poor, and it is therefore desirable to taper the steering vector. Suppose the steering vector $[\mathbf{s}]$ is multiplied by a real $NL \times NL$ diagonal matrix $[\mathbf{D}]$ that

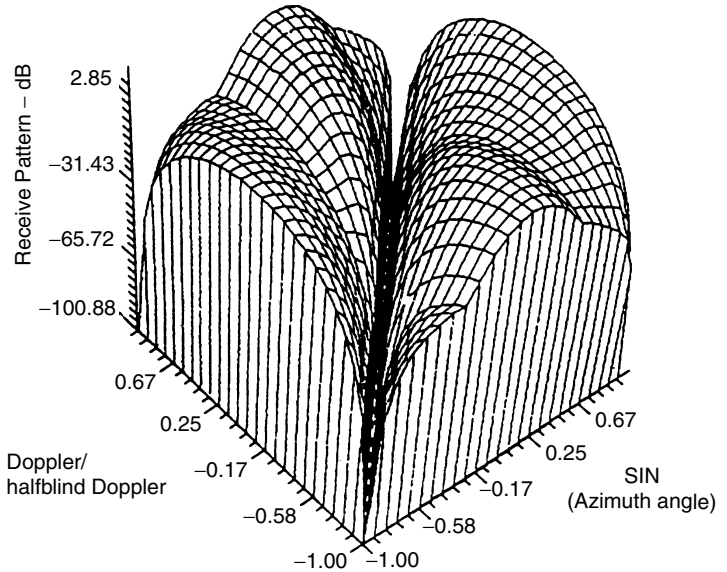

FIGURE 10.23 Adapted azimuth–Doppler pattern for eight-element linear array that processes four pulses.

contains the tapering. Then, the new weight vector is $[\mathbf{w}_2] = [\mathbf{M}]^{-1}[\mathbf{D}][\mathbf{s}^*]$, and the corresponding receive pattern is

$$V_2(f_\mathrm{d}, \phi) = [\mathbf{s}_u]^\mathrm{T}[\mathbf{D}][\mathbf{s}^*] - [\mathbf{s}_u]^\mathrm{T}([\mathbf{I}] - [\mathbf{M}]^{-1})[\mathbf{D}][\mathbf{s}^*] \tag{10.36}$$

The first term on the right-hand side of Equation 10.36 represents the unadapted, tapered radiation pattern, whereas the second term represents a tapered clutter-cancellation beam. Consequently, the space–time processor now produces a radiation pattern that is equal to the unadapted, tapered pattern (with the desired low sidelobes), minus a beam which is nonzero only along the clutter direction and nearly zero elsewhere in sin azimuth–Doppler space. A typical adapted azimuth-pattern-cut is shown in Figure 10.24. By substituting the expression for $[\mathbf{w}_2]$ into Equation 10.24 it is found that tapering produces a slight loss in the output signal-to-clutter-plus-noise ratio below the optimum result given by Equation 10.26. Typical losses are shown in Figure 10.25.

The foregoing analysis assumed the signal was not included in the computation of the covariance matrix. If the signal is included it can lead to deleterious effects, as we now demonstrate. When the signal is included the covariance matrix becomes $[\mathbf{Q}] = [\mathbf{M}] + [\mathbf{s}^*][\mathbf{s}]^\mathrm{T}$ so that the weight vector now is $[\mathbf{w}_3] = ([\mathbf{M}] + [\mathbf{s}^*][\mathbf{s}]^\mathrm{T})^{-1}[\mathbf{D}][\mathbf{s}^*]$. If we apply the matrix inversion lemma we find

$$[\mathbf{w}_3] = ([\mathbf{M}]^{-1}[\mathbf{D}] - \tau\gamma[\mathbf{M}]^{-1})[\mathbf{s}^*] \tag{10.37}$$

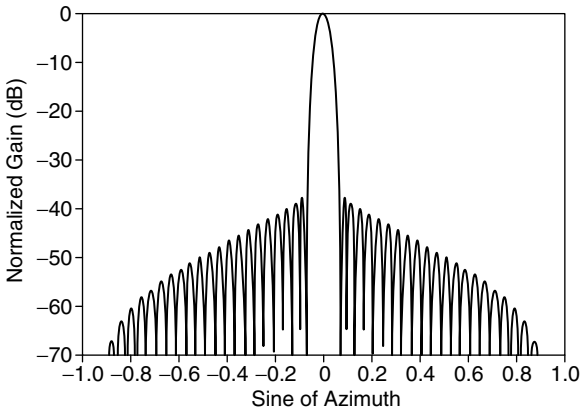

FIGURE 10.24 Adapted azimuth response for Doppler cut corresponding to one half the blind speed. Steering vector tapered but signal not included in weight calculation.

where $\tau^{-1} = [\mathbf{s}]^{\mathrm{T}}[\mathbf{M}]^{-1}[\mathbf{s}^*] + 1$ and $\gamma = [\mathbf{s}]^{\mathrm{T}}[\mathbf{M}]^{-1}[\mathbf{D}][\mathbf{s}^*]$. The voltage radiation pattern now becomes

$$V_3(f_{\mathrm{d}}, \phi) = [\mathbf{s}_u]^{\mathrm{T}}[\mathbf{D}][\mathbf{s}^*] - [\mathbf{s}_u]^{\mathrm{T}}([\mathbf{I}] - [\mathbf{M}]^{-1})[\mathbf{D}][\mathbf{s}^*] - \alpha[\mathbf{s}_u]^{\mathrm{T}}[\mathbf{M}]^{-1}[\mathbf{s}^*] \quad (10.38)$$

where $\alpha = \gamma\tau$. Now recall the definitions of the untapered, adapted pattern $V(f_{\mathrm{d}},\phi)$ as given by Equation 10.35, and the tapered, adapted pattern $V_2(f_{\mathrm{d}},\phi)$ given by Equation 10.36, that includes tapering but excludes the signal from the covariance

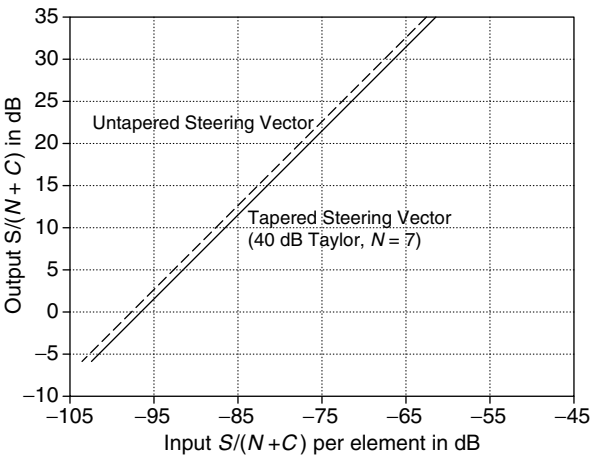

FIGURE 10.25 Effect of tapering vector on output signal-to-noise-plus-clutter ratio of 50-element array when signal is not included in weight calculation.

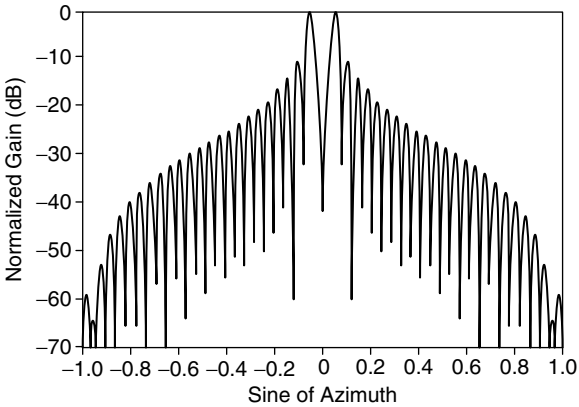

FIGURE 10.26 Adapted azimuth response for Doppler cut corresponding to one half the blind speed. Steering vector tapered and signal included in weight calculation.

computation. Then, it is evident that $V_3(f_d, \phi)$ in Equation 10.38 can be rewritten as

$$V_3(f_d,\phi) = V_2(f_d,\phi) - \alpha V(f_d,\phi) \tag{10.39}$$

That is, the radiation pattern is equal to the tapered, adapted pattern one obtains with the signal excluded from the covariance, minus a coefficient α times the adapted pattern one gets without any tapering at all. Because both beams are pointed to the same point (f_{d0},ϕ_0) in azimuth–Doppler space this can lead to a resultant pattern with some very undesirable features, such as cancellation in the pointing direction, as is evident from Figure 10.26. This produces a corresponding loss in the output signal-to-clutter ratio, as can be seen from Figure 10.27. Thus,

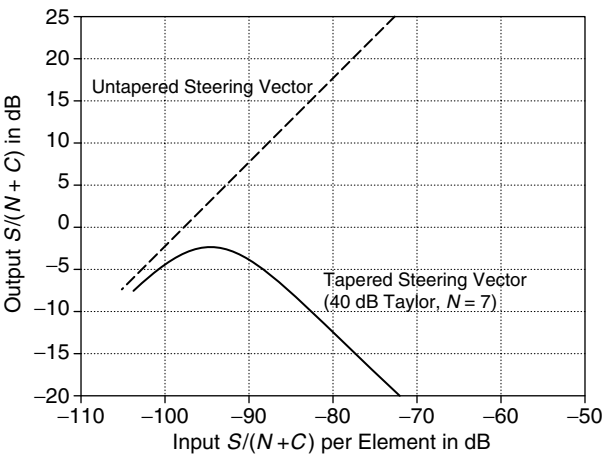

FIGURE 10.27 Effect of tapering vector on output signal-to-noise-plus-clutter ratio of 50-element linear array when signal is included in weight calculation.

one must be careful to exclude the signal from the covariance matrix computation, unless the signal is weak. Although we have not proven it here, it can be readily shown that inclusion of the signal in the computation of the covariance matrix is not a problem in the limit when there is no tapering applied. This can be seen by setting [**D**]=[**I**] in Equation 10.37 and then substituting the result into Equation 10.24. This conclusion does not require that the received signal vector exactly equal the assumed steering vector, because the degradation is gradual. The analysis parallels the development in Ref. 5, Section 2.2.

Now that we have completed the general theoretical introductions we proceed to specifics. In the next section we study the eigenstructure of the simplest possible space–time processor, in order to demonstrate analytically how various effects limit the possible improvement. Then in Section 10.2.4 we numerically study more complex space–time processors.

10.2.3. Two-Element Displaced Phase Center Antenna

In order to illustrate in a simple manner how near-field obstacles (such as an airplane wing), internal clutter motion, and aircraft crabbing limit the performance of an adaptive radar, we now consider the simplest limiting case: a two-element displaced phase center[6–9] antenna (DPCA) that uses the processing shown in Figure 10.19. In this case only w_{11} and w_{22} are nonzero in Equation 10.18, and the covariance matrix [**M**] becomes

$$[\mathbf{M}] = \begin{bmatrix} \langle U_1^*(t)U_1(t)\rangle & \langle U_1^*(t)U_2(t-T)\rangle \\ \langle U_1^*(t-T)U_1(t)\rangle & \langle U_2^*(t-T)U_2(t-T)\rangle \end{bmatrix} \tag{10.40}$$

We must now derive the voltages produced by clutter on the two antennas when a near-field obstacle is present, along with internal clutter motion and aircraft crabbing. In order to do this consider the geometry in Figure 10.28, where we show a single clutter-scatterer of amplitude A_q in the far field of a two-element array, and a near-field obstacle close to the array. For convenience, both the clutter-scatterer and the near-field obstacle have been shown in the plane of the array, but the generalization is straightforward. (For example, when the clutter is not in the plane of the array, $\sin\phi_q$ is replaced by $\sin\phi_q \cos\theta_q$ where θ_q is the depression angle of the scatterer.) The total clutter return will then be obtained by summing over all the individual clutter scatterers, and then averaging. In order to account for the desired platform speed, the internal clutter motion and the aircraft crabbing (which is aircraft motion normal to the antenna array) the clutter-scatterer is assumed to have an x-directed component of velocity $V_x + \delta v_{qx}$ and a y-directed component $V_y + \delta v_{qx}$, where V_x is the aircraft speed in the (desired) direction along the array, δv_{qx} is the x-component of the internal clutter motion of the qth clutter scatterer, V_y is the velocity component produced by crabbing, and δv_{qy} is the y-component of the internal clutter motion of the qth clutter-scatterer. All speeds are measured relative to the antenna array. We further assume that the near-field obstacle scatters isotropically, with a bistatic radar cross section σ_0,

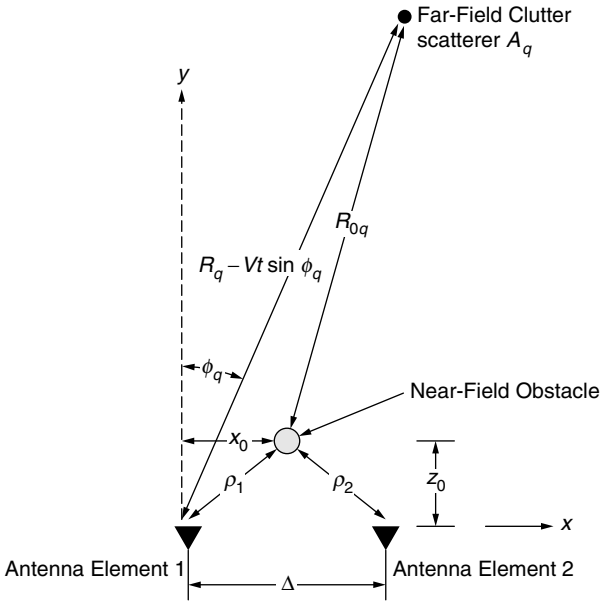

FIGURE 10.28 Obstacle in near field of two-element array.

and that only element 1 of the array transmits, but both elements receive. Then, upon referring to the geometry of Figure 10.29, it is evident that the voltage produced on each receive antenna element by a single far-field clutter-scatterer at angular position ϕ_q is the sum of the four components shown. Upon adding these four components it is straightforward to show that the voltage produced on antenna 1 due to a far-field clutter-scatterer at ϕ_q is

$$U_{1q}(t) = A_q B_q^2 \exp\{\mathrm{i}2k[R_q - (V_x + \delta v_{qx})t \sin \phi_q - (V_y + \delta v_{qy})t \cos \phi_q]\} \quad (10.41)$$

where

$$B_q = 1 + \Gamma_1 \exp[\mathrm{i}k(\rho_1 - \tau_q)] \quad (10.42)$$

and the magnitude of the complex scattering coefficient A_q of the qth clutter-scatterer is proportional to $\sigma_q^{1/2} R_q^{-2}$, where σ_q is the scatterer cross section. Also, $k = \omega/c$, $\tau_q = x_0 \sin \phi_q + z_0 \cos \phi_q$ and $\Gamma_1 = (\sigma_0/4\pi\rho_1^2)^{1/2}$ where σ_0 is the bistatic radar cross section and ρ_1 is the distance between the-near-field obstacle and antenna element 1 as shown in Figure 10.28. In deriving Equation 10.41 and Equation 10.42 it has been assumed that $\rho_1 \ll R_{0q}$, $\rho_1 \ll R_q$, so that R_{0q} can be replaced by R_q in all nonphase terms.

Likewise, the voltage on element 2 due to the qth clutter-scatterer is

$$\begin{aligned} U_{2q}(t) = A_q B_q D_q \exp\{\mathrm{i}2k[R_q - (V_x + \delta v_{qx})t \sin \phi_q \\ - (V_y + \delta v_{qy})t \cos \phi_q] - \mathrm{i}k\Delta \sin \phi_q\} \end{aligned} \quad (10.43)$$

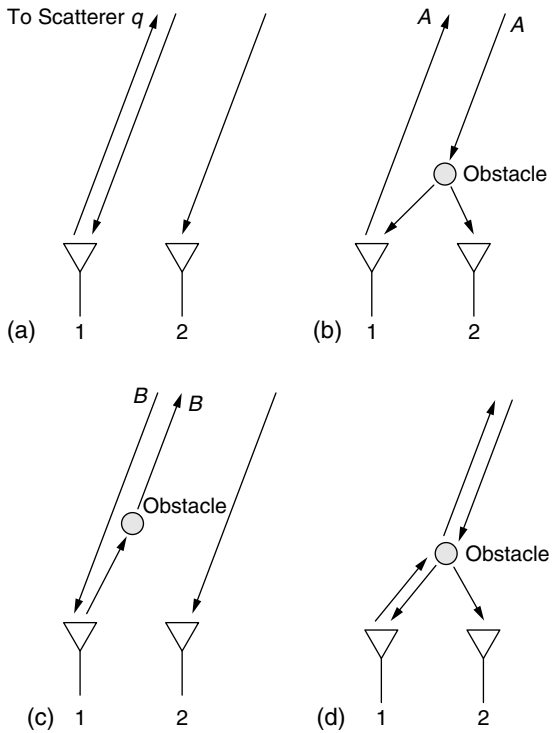

FIGURE 10.29 Four scattering paths for near-field obstacle. (a) Direct path. (b) One bounce path *A*. (c) One bounce path *B*. (d) Two bounce path.

where

$$D_q = 1 + \Gamma_2 \exp[ik(\rho_2 - \tau_q + \Delta \sin \phi_q)] \tag{10.44}$$

and $\Gamma_2 = (\sigma_0/4\pi\rho_2^2)^{1/2}$ where σ_0 is the bistatic radar cross section and ρ_2 is the distance between the near-field obstacle and antenna element 2 as shown in Figure 10.28. The total field on elements 1 and 2 in Figure 10.28 and Figure 10.29 is are obtained by summing over all clutter-scatterers. That is,

$$U_1(t) = \sum_q U_{1q}(t) \tag{10.45}$$

$$U_2(t) = \sum_q U_{2q}(t) \tag{10.46}$$

We can now use Equation 10.41 to Equation 10.46 to calculate the elements of the covariance matrix in Equation 10.40. If we assume that all clutter scatters are independent, and randomly located, then we have

$$\langle A_p A_q^* \rangle = \gamma_q \delta_{pq} \tag{10.47}$$

where $\delta_{pq} = 1$ if $p = q$, and $\delta_{pq} = 0$ otherwise. If we use Equation 10.41 to Equation 10.47 we find that the components of the covariance matrix are

$$M_{11} = \langle |U_1|^2 \rangle = \sum_q \gamma_q |B_q|^4 \tag{10.48}$$

$$M_{22} = \langle |U_2|^2 \rangle = \sum_q \gamma_q |B_q|^2 |D_q|^2 \tag{10.49}$$

$$M_{12} = \langle U_1^*(t) U_2(t-T) \rangle = \sum_q \gamma_q |B_q|^2 B_q^* D_q \langle \exp(\mathrm{i}\beta_q) \rangle \tag{10.50}$$

where

$$\beta_q = k[(2V_x T - \Delta)\sin\phi_q + 2k\delta v_{qx} T \sin\phi_q + 2k(V_y + \delta v_y)T\cos\phi_q] \tag{10.51}$$

and the remaining expectation indicated in Equation 10.50 must be taken over the internal clutter motion. Also, we have not presented M_{21} because $M_{21} = M_{12}^*$, where the asterisk denotes a complex conjugate. Now assume a perfect velocity match so that the processor adjusts the pulse repetition rate to give

$$2V_x T = \Delta \tag{10.52}$$

Also assume the internal clutter motion and crabbing are sufficiently small so that $2kT\delta v_{qx} \ll 1$ and $2k(V_y + \delta v_{qy})T \ll 1$. In this limit we can expand $\exp(\mathrm{i}\beta_q)$ in Equation 10.50 in a Taylor series and then perform the expectation over the internal clutter motion. If we assume that

$$\langle \delta v_{qx} \rangle = \langle \delta v_{qy} \rangle = 0 \tag{10.53}$$

$$\langle \delta v_{qx}^2 \rangle = \langle \delta v_{qy}^2 \rangle = \frac{\sigma_v^2}{2} \tag{10.54}$$

where σ_v^2 is the variance of the internal clutter speed, we find that M_{12} can be approximated as

$$M_{12} \approx \sum_q \gamma_q |B_q|^2 B_q^* D_q [1 - (kT\sigma_v)^2 + \mathrm{i}2kV_y T\cos\phi_q - 2(kTV_y\cos\phi_q)^2] \tag{10.55}$$

We can simplify M_{11}, M_{12}, and M_{22} even further if we assume that the near-field obstacle is a weak scatterer, and hence $\Gamma_1 \ll 1, \Gamma_2 \ll 1$. In this limit we can then ignore higher order terms in Γ_1 and Γ_2, so that

$$M_{11} \approx C_0(1 + 4\Gamma_1^2) \tag{10.56}$$

$$M_{22} \approx C_0(1 + \Gamma_1^2 + \Gamma_2^2 + 2\Gamma_1\Gamma_2 G) \tag{10.57}$$

$$M_{12} \approx C_0(1 - \epsilon + 2\Gamma_1^2 + 2\Gamma_1\Gamma_2 H) \tag{10.58}$$

where C_0 is the clutter power received on a single element in the absence of space–time processing and is defined as

$$C_0 = \sum_q \gamma_q \tag{10.59}$$

Also,

$$\epsilon = \frac{\pi^2}{4}\left(\frac{\sigma_v^2 + V_y^2}{V_x^2}\right) \tag{10.60}$$

$$G = \frac{1}{C_0}\sum_q \gamma_q \cos k(\rho_2 - \rho_1 + \Delta \sin \phi_q) \tag{10.61}$$

$$H = \frac{1}{C_0}\sum_q \gamma_q \exp[\mathrm{i}k(\rho_2 - \rho_1 + \Delta \sin \phi_q)] \tag{10.62}$$

In deriving the results in Equation 10.58 it has been assumed that the clutter is approximately uniformly distributed in angle so that

$$\sum_q \gamma_q \cos \phi_q \approx 0 \tag{10.63}$$

$$\sum_q \gamma_q \cos^2 \phi_q \approx \frac{1}{2}\sum_q \gamma_q \tag{10.64}$$

It was further assumed that the interelement spacing Δ was one half wavelength, so that Equation 10.52 gives $V_x T = \lambda/4$, where $\lambda =$ wavelength.

We can use the simplified expressions in Equation 10.56 to Equation 10.58 to obtain the eigenvalues of the covariance matrix. Recall that the eigenvalues λ_1, λ_2 are solutions of the equation $(M_{11} - \lambda)(M_{22} - \lambda) - (M_{12})^2 = 0$, and can, therefore, be expressed as

$$\lambda_{1,2} = \frac{1}{2}(M_{11} + M_{22}) \pm \frac{1}{2}[(M_{11} + M_{22})^2 + 4|M_{12}|^2]^{1/2} \tag{10.65}$$

If the expressions of M_{11}, M_{22}, and M_{12} are substituted into Equation 10.65 we find that the two eigenvalues are

$$\lambda_1 \approx 2C_0 \tag{10.66}$$

$$\lambda_2 \approx 2C_0\left(\frac{\Gamma_1^2}{2} + \frac{\Gamma_2^2}{2} - \Gamma_1\Gamma_2 G + \epsilon\right) \tag{10.67}$$

and the corresponding normalized eigenvectors are

$$[\mathbf{e}_1] \approx \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}; \qquad [\mathbf{e}_2] \approx \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix} \tag{10.68}$$

The steering vector $[\mathbf{s}]$ in Equation 10.22, for a target with a radial velocity V_t is

$$[\mathbf{s}] = s_0\begin{bmatrix} 1 \\ \exp\left(\mathrm{i}\dfrac{4\pi V_t T}{\lambda}\right) \end{bmatrix} \tag{10.69}$$

Therefore, if we use Equation 10.66 to Equation 10.69 in Equation 10.28 we find that

$$\left(\frac{S}{C+N}\right)_{\max} = \frac{S_0}{C_0}\left[\frac{\cos^2\left(\frac{2\pi V_t T}{\lambda}\right)}{1+\frac{N}{C_0}} + \frac{\sin^2\left(\frac{2\pi V_t T}{\lambda}\right)}{\frac{1}{4}(\Gamma_1^2+\Gamma_2^2-2\Gamma_1\Gamma_2 G)+\frac{\pi^2}{8}\varphi_c^2+\frac{\pi^2}{8}\frac{\sigma_v^2}{V_x^2}+\frac{N}{2C_0}}\right] \tag{10.70}$$

where $N=\sigma^2=$ noise power, S_0/C_0 is the signal-to-clutter ratio in the absence of space–time processing and φ_c is the crabbing angle defined as $\tan\varphi_c \approx \varphi_c = V_y/V_x$. From Equation 10.70 it is evident that, because $\Gamma_1 \ll 1, \Gamma_2 \ll 1, \varphi_c \ll 1,\ \varphi_c \ll 1, \sigma_v/V_x \ll 1$ and $(N/2C_0) \ll 1$, the second term on the right-hand side of Equation 10.70 dominates, as long as the target is not at a blind speed (i.e., $V_t = n\lambda/2T$, where $n=$ integer). Therefore, if we ignore the first term on the right-hand side of Equation 10.70 and assume the target is at one-half the blind speed we get for the improvement factor

$$\frac{\left(\frac{S}{C+N}\right)_{\max}}{\left(\frac{S_0}{C_0+N}\right)} = \frac{4}{\Gamma_1^2+\Gamma_2^2-2\Gamma_1\Gamma_2 G+\frac{\pi^2}{2}\varphi_c^2+\frac{\pi^2}{2}\left(\frac{\sigma_v^2}{V_x^2}\right)+\frac{2N}{C_0}} \tag{10.71}$$

where we have approximated S_0/C_0 by $S_0/(C_0+N)$, because the system noise N is much less than the clutter power C_0.

Equation 10.71 shows how each factor limits the possible improvement factor. The term, $\Gamma_1^2+\Gamma_2^2-2\Gamma_1\Gamma_2 G$, which is proportional to the radar cross section of the near-field obstacle, shows how a near-field scatterer limits improvement. The term $(\pi\varphi_c)^2/2$ shows how crabbing limits improvement and the term $(\pi\sigma/V_x)^2/2$ shows how internal clutter motion limits improvement. Of course, in the absence of crabbing, near-field obstacles and internal clutter motion, Equation 10.71 reduces to

$$\left(\frac{S}{C+N}\right)_{\max} = \frac{2S_0}{N} \tag{10.72}$$

so that the improvements is limited only by system noise N. The factor of 2 in Equation 10.72 arises because there are two antenna elements, and the signals add coherently in the output giving a signal power $|2s_0|^2 = 4S_0$, whereas the noise powers add incoherently, giving a total noise power of $2N$.

It should be noted that, although the foregoing analysis was performed for the two-element, two-pulse limit, numerical studies of the N-element, L-pulse system demonstrate the same quadratic dependence of the improvement factor on Γ^2,

$(\sigma_v/V_x)^2$, etc., as exhibited in Equation 10.71 in the limit when $\sigma_v/V_x \to 0$, etc. The only differences are that the coefficients multiplying $(\sigma_v/V_x)^2$, φ_c^2, etc., are different.

10.2.4. Simulation Results

10.2.4.1. Internal Clutter Motion

In this section we separately study each of the effects that limit the clutter cancellation possible with space–time processing. Let us first study the effect of internal clutter motion. Assume that the antenna array is linear and that the clutter is modeled as a distribution of randomly located scatterers having a Gaussian distribution of radial velocity, with a variance σ_v. The number of clutter scatterers is chosen to be much larger than the number of degrees of freedom of the adaptive radar. The other parameters assumed for the simulation are summarized in Table 10.1, and the covariance matrix is calculated with the target excluded, in order to avoid the problems, discussed in Section 10.2.2, that arise when the steering vector is tapered. In performing the calculations the contributions from all ambiguous range rings are included.

In Figure 10.30 we show the loss in the clutter cancellation in a two-pulse processor produced by the effect of internal clutter motion. Note that, as predicted for a two-element array in Section 10.2.3, the loss in improvement is still approximately proportional to σ_v^{-2} for small values of σ_v. The internal clutter motion also produces a severe distortion in the radiation pattern as can be seen from Figure 10.31, where we show an azimuth cut in the plane where the Doppler is one half the blind speed. In the absence of internal clutter motion the adapted pattern is approximately equal to the nonadapted pattern (where the weights are given by the tapered steering vector), except with a null along the clutter line in sine azimuth–Doppler space. As noted in Section 10.2.2, this distortion occurs because the radiation pattern is dominated by the eigenbeams associated with the smallest eigenvalues. A plot of these eigenvalues is shown in Figure 10.32, where the eigenvalues have been normalized so that λ_1 represents the largest and λ_{100}

TABLE 10.1
Simulation Scenario Parameters

Parameter	Value
Number of elements in azimuth	50
Taper in azimuth on transmit	40 dB Taylor, $\bar{N} = 7$
Carrier frequency	*L*-Band
Pulse repetition frequency	3000 Hz
Platform velocity	350 knots
Platform altitude	30,000 ft
Target range	200 nmi
Azimuth electronic scan angle	Broadside

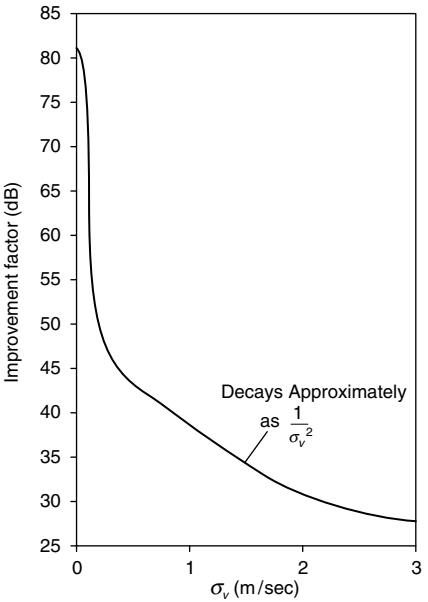

FIGURE 10.30 Effect of internal clutter motion of $[S/(C+N)]$ after processing to quiescent $[S/(C+N)]$, for 50-element array that processes two pulses. Quiescent $[S/(C+N)] = -55.1$ dB.

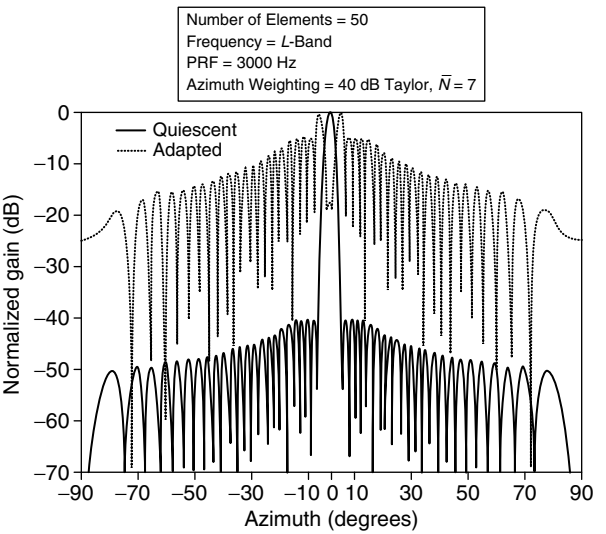

FIGURE 10.31 Quiescent and two-pulse STP antenna patterns with $\sigma_v = 0.4$ m/s.

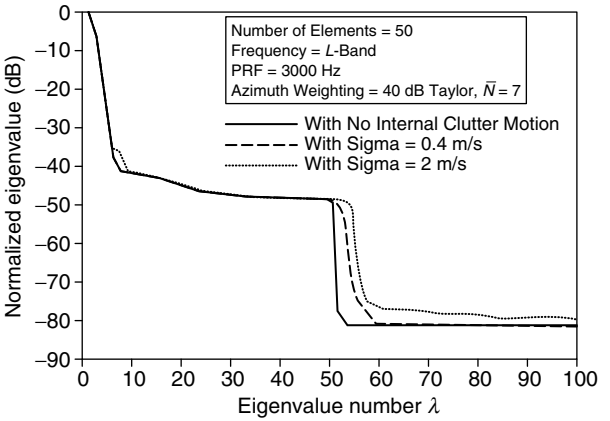

FIGURE 10.32 Effects of ICM on two-pulse STP eigenvalues.

the smallest. Observe how the internal clutter motion modifies the sharp drop-off at λ_{52} between the large eigenvalues (associated with clutter) and the smallest eigenvalues (associated with noise). It is found that these smallest eigenvalues λ_{52} through λ_{100} contribute negligibly to the space–time processing gain, as is evident from Figure 10.33. Thus, the eigenbeams associated with these eigenvalues are contributing nothing towards producing a notch on the clutter, and all they are doing is producing pattern distortion and high azimuth–Doppler sidelobes. One way this problem can be cured is to artificially raise the noise level so that $\sigma^2 = \lambda_{51}$, in accordance with our discussion following Equation 10.33. This nearly eliminates the pattern distortion, as can be seen from Figure 10.34, and even produces a slight increase in $S/(C+N)$, because as seen in Figure 10.33 the eigenbeams associated with λ_{52} to λ_{100} actually cause $S/(C+N)$ to decrease.

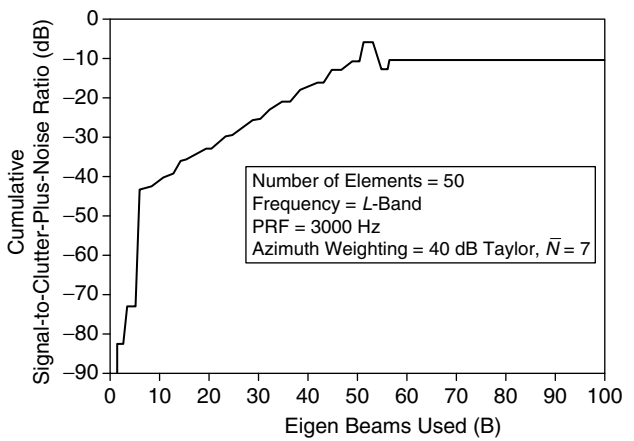

FIGURE 10.33 Cumulative signal-to-clutter-plus-noise ratio for two-pulse STP, with $\sigma_v = 0.4$ m/s and tapered steering vector.

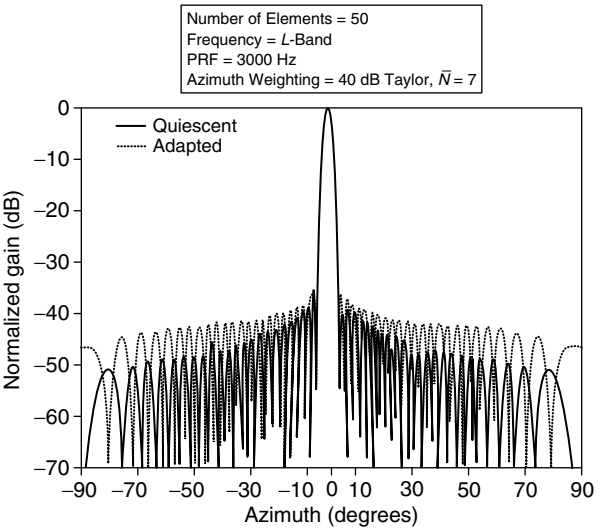

FIGURE 10.34 Quiescent and two-pulse STP antenna patterns with $\sigma_v = 0.4$ m/s and eigenvalue compensation.

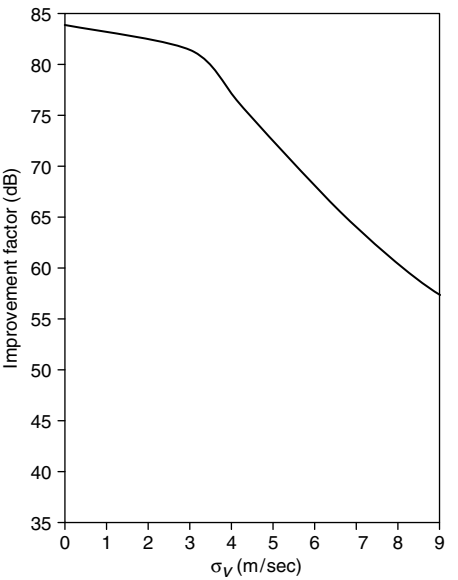

FIGURE 10.35 Effect of internal clutter motion on ratio of $[S/(C+N)]$ after processing to quiescent $[S/(C+N)]$, for 50-element array that processes four pulses. Quiescent $[S/(C+N)] = -55.1$ dB.

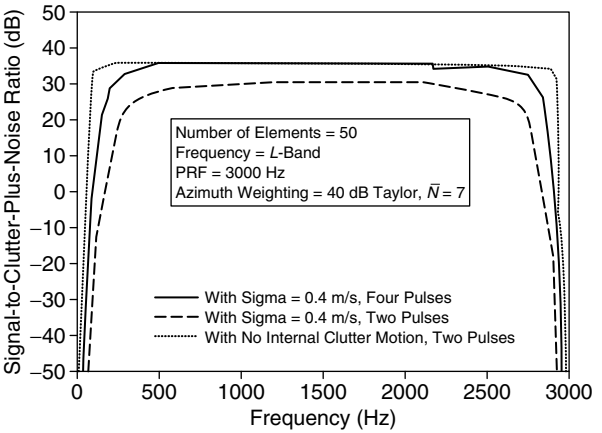

FIGURE 10.36 Effects of ICM on signal-to-clutter-pulse-noise ratio in doppler for two- and four-pulse STP.

However, a better approach to overcoming the effects of internal clutter motion is to increase the number of pulses processed. That is, internal clutter motion causes the clutter to spread off of the clutter line in the Doppler direction in Figure 10.17. Therefore, we would expect that processing more pulses would produce a better Doppler filter, and this is indeed what happens. Using four pulses not only allows us to recover the quiescent radiation pattern, but also significantly increases the value of $S/(C+N)$ over the two-pulse canceler, as can be seen from Figure 10.35. The four-pulse processor also provides an improved response at the output of the Doppler processor, as demonstrated in Figure 10.36.

10.2.4.2. Aircraft Crabbing

We now study the effect of aircraft crabbing, assuming no internal clutter motion. By recalling the analysis leading to Equation 10.65 we expect that aircraft

TABLE 10.2
Effect of Aircraft Crabbing on Improvement Factor of 50-Element Linear Array

Number of Pulses Processed	Improvement Factor (dB)		
	$\varphi_c = 0°$	$\varphi_c = 5°$	$\varphi_c = 10°$
2	80.7	46.5	41.0
4	82.7	82.7	82.7

Note: Improvement factor $= \dfrac{\left(\dfrac{S}{C+N}\right)_{\text{adapted}}}{\left(\dfrac{S}{C+N}\right)_{\text{quiescent}}}$.

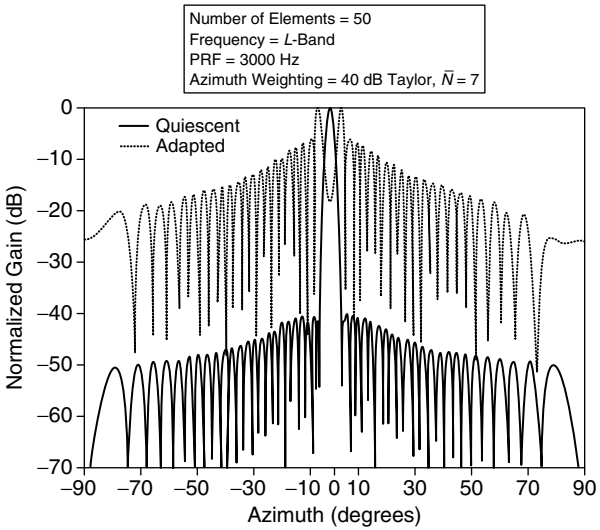

FIGURE 10.37 Quiescent and two-pulse STP antenna patterns with $\varphi_c = 5°$.

crabbing will produce the same sort of effects as internal clutter motion, because it again causes the clutter to move off of the line in sine azimuth–Doppler space shown in Figure 10.17, although in a different fashion than internal clutter motion. That is, internal clutter motion smears the clutter off of the straight line in the sine azimuth–Doppler plane in Figure 10.17, whereas crabbing causes the clutter to lie on a curve in accordance with the relation

$$f_{\mathrm{d}} = \frac{2Vf_0}{c}[\cos\,\phi_c\,\sin\,\phi + \sin\,\varphi_c(1 - \sin^2\,\phi)^{1/2}] \tag{10.73}$$

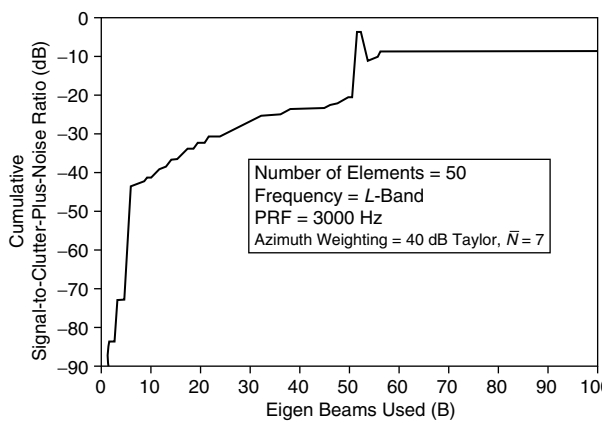

FIGURE 10.38 Cumulative signal-to-clutter-plus-noise ratio for two-pulse STP, with $\varphi_c = 5°$. and tapered steering vector.

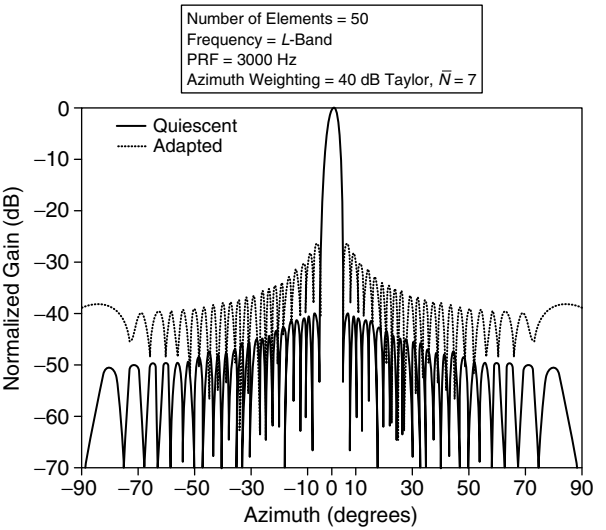

FIGURE 10.39 Quiescent and two-pulse STP antenna patterns with $\varphi_c = 5°$ and eigenvalue compensation.

where f_d is the Doppler frequency, ϕ is the azimuth and φ_c is the crabbing angle. In Table 10.2 we show how crabbing decreases the output signal-to-clutter-plus-noise ratio of a two-pulse processor for the parameters given in Table 10.1, and in Figure 10.37 we show how crabbing distorts the desired radiation pattern. Again the simple cure is found by studying how the eigenbeams contribute to the pattern distortion and the signal-to-clutter-plus-noise ratio. Consider the calculations shown in Figure 10.38. It is evident from this plot that the eigenbeams from 52 through 100 contribute nothing to canceling the clutter, and actually produce a slight loss in the signal-to-clutter-plus-noise ratio. Thus, we can again use eigenvalues compensation (by artificially increasing the noise so that $\sigma^2 = \lambda_{s1}$) to obtain the improved radiation pattern shown in Figure 10.39, as well as the slight improvement in $S/(C+N)$ shown in Table 10.2. However, just as we found for internal clutter motion, a better cure is to process more pulses. If we use a four-pulse processor we recover the desired radiation pattern and, as seen from Table 10.2, completely recover the signal-to-clutter-plus-noise ratio obtained in the absence of crabbing.

10.2.4.3. Near-Field Obstacles

Perhaps the most stressing factor for a space–time processor is a near-field obstacle, because, as noted earlier, it tends to spread to clutter over the entire sine azimuth–Doppler plane. In order to study this problem we considered an obstacle with a bistatic radar cross section $\sigma_0 = 1\ \text{m}^2$ located a distance of ten wavelengths along the normal to the center of a linear array. For these calculations the transmit array was uniformly weighted, as was the steering

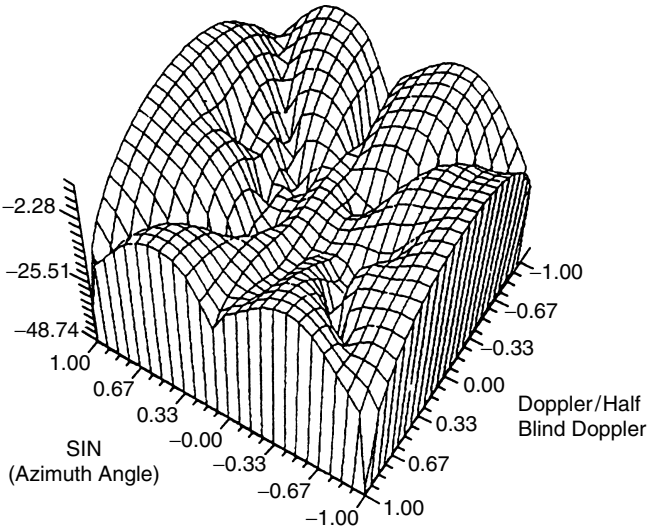

FIGURE 10.40 Adapted azimuth-Doppler pattern when $1\ m^2$ obstacle is located broadside to eight-element array that processes four pulses.

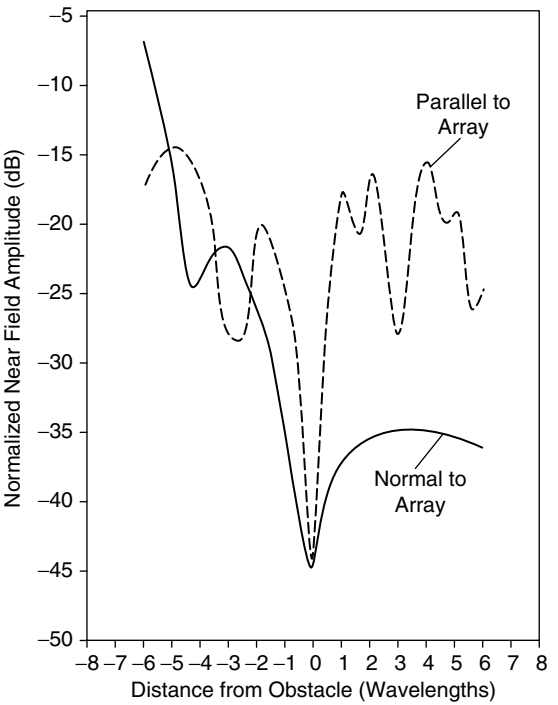

FIGURE 10.41 Adapted field distribution near obstacle located 10λ broadside to 32-element array that uses three-pulse processor, $\sigma_0 = 1\ m^2$.

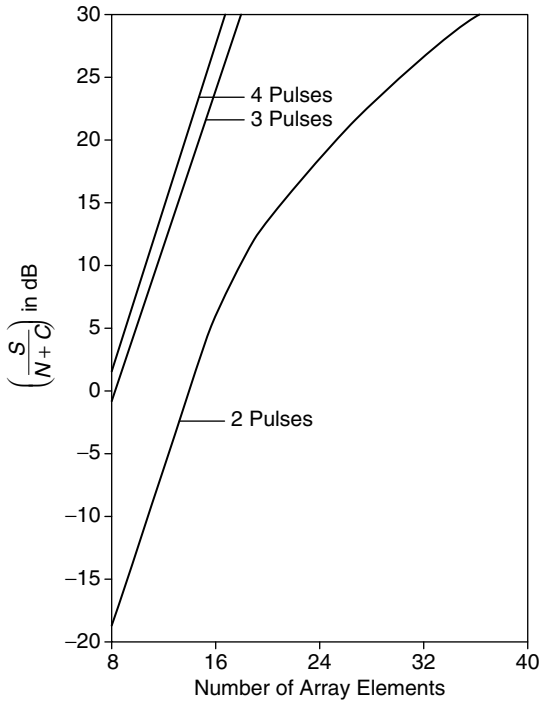

FIGURE 10.42 Adapted signal-to-clutter-plus-noise ratio if 1 m^2 obstacle is located 10λ broadside to array.

vector, and it was assumed that there was no internal clutter motion aircraft crabbing. An adapted radiation pattern in sine azimuth–Doppler space for this case is shown in Figure 10.40. By comparing this pattern with Figure 10.23 we see that, because the near-field obstacle has spread the clutter essentially everywhere in sine azimuth–Doppler space, there is no longer a ridge null (as in Figure 10.23), but rather a series of depressions produced as the processor is trying to cancel clutter that is everywhere, and it simply does not have enough degrees of freedom to do it. A study of the near-field structure shows that what the processor attempts to do is to place a spatial null on the near-field obstacle. This is quite evident from Figure 10.41 where we show the three pulse average of the near field in cuts through the obstacle in planes normal and parallel to the array. This point is further evident from Figures 10.42 and 10.43. In Figure 10.42 we show that increasing the number of elements in the array is quite effective in cancelling the effects of the near-field obstacle, because it allows a deeper null to be placed on the obstacle. However, as seen from Figure 10.43, increasing the number of pulses processed is less effective because it has little effect on the spatial null placed on the object and only reduces the amount of clutter with the same Doppler as the target.

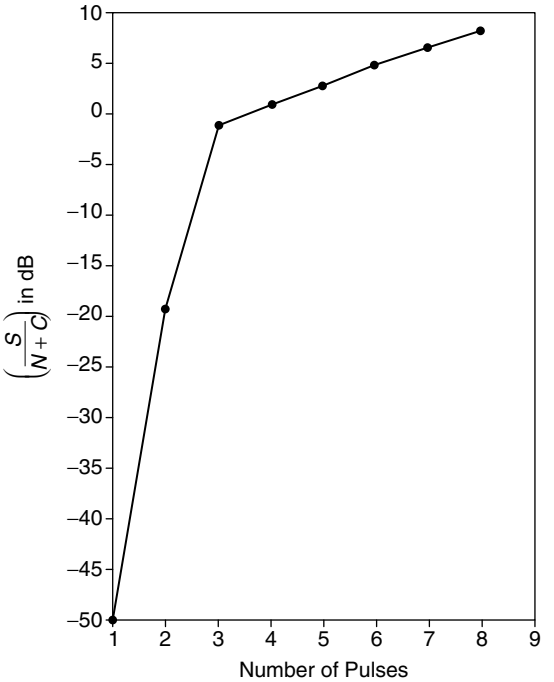

FIGURE 10.43 Adapted signal-to-clutter-plus-noise ratio if 1 m^2 obstacle is located 10λ broadside to eight-element array.

Calculations were also performed for the case of two near-field obstacles each with bistatic cross sections of 1 m^2, with one located ten wavelengths along a normal to the center of the array (i.e., $x = 0$, $z = 10\lambda$) and a second also located ten wavelengths normal to the center of the array but displaced by two wavelengths in the transverse direction (i.e., $x = 2\lambda$, $z = 10\lambda$). In Figure 10.44 we show that now the processor attempts to place a null on each obstacle, and in Figure 10.45 we show that the presence of the second obstacle degrades the signal-to-noise-plus-clutter ratio.

10.2.4.4. Antenna Errors (Channel Mismatch)

Antenna errors can be expected to limit the effectiveness of adaptive radar for several reasons. First, they produce a spatial ripple across the received wavefront that differs from the expected planar wavefront embedded within the steering vector. Consequently, the actual received signal from a target has some features that are different from the assumed steering vector and, hence, the adaptive radar actually attempts to cancel part of the received signal from a true target. This produces a slight degradation in the output signal-to-noise-plus-clutter ratio below the theoretical optimum.

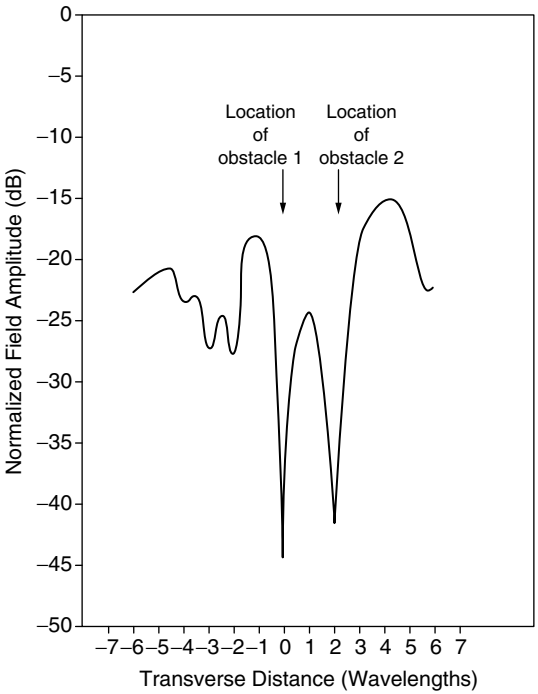

FIGURE 10.44 Adapted field distribution in plane $z = 10\lambda$ for $\sigma_0 = 1$ m^2 obstacles at $x = 0$, $z = 10\lambda$ and $x = 2\lambda$, $z = 10\lambda$. Processor has 32 elements and uses two pulses.

A more stressing effect is the dispersion of errors across the bandwidth of the receiver. Thus, because each receive channel has a different variation (called "channel mismatch") of the phase and amplitude errors across the frequency band, a single receive weight cannot possibly cancel clutter across the entire bandwidth. Mathematically, its effect is to multiply those members of the covariance matrix corresponding to different receive elements by $1 - \sigma_\epsilon^2; -\sigma_\rho^2$, where σ_ϵ^2 is the variance of the phase errors and σ_ρ^2 is the variance of the amplitude errors. The members of the covariance matrix corresponding to the same receive element are unaffected.

An example of the effect of channel mismatch on improvement factor is shown in Figure 10.46. We note that the dispersive errors do not affect the improvement factor as long as σ^{2}/N is less than the noise-to-clutter ratio, but the improvement factor degrades as σ^{-2} thereafter, where $\sigma^2 = \sigma_\epsilon^2 + \sigma_\rho^2$. Also, as expected from classical smoothing theory, increasing the number of elements in the array gives an improvement proportional to the number of elements, N. Thus, as seen from Figure 10.46, increasing the number of elements by a factor of 4 (from 25 to 100) gives a 6 dB increase in the normalized improvement factor.

Because antenna errors are constant from pulse-to-pulse, one would not expect the normalized improvement factor to be significantly changed by varying the number of pulses processed, and this is indeed the case.

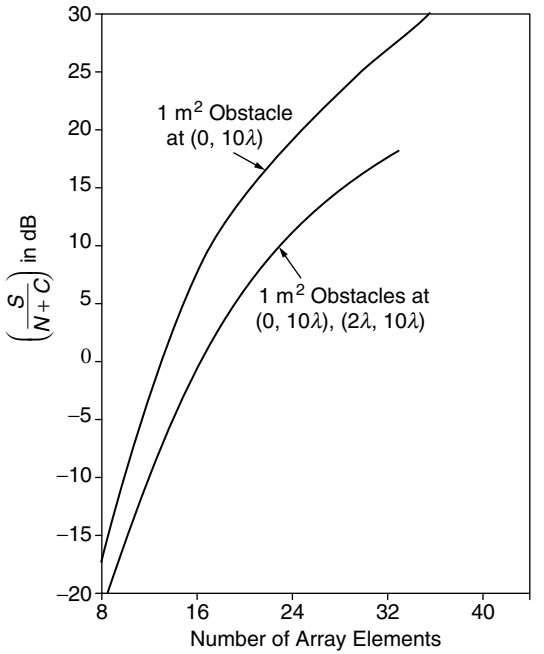

FIGURE 10.45 Adapted signal-to-noise-plus-clutter ratio for two-pulse processor with near-field obstacles.

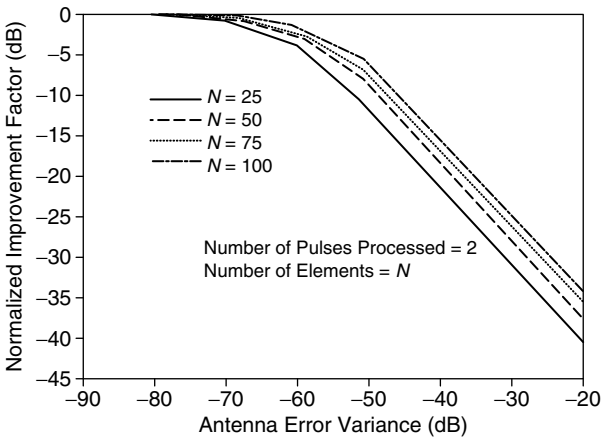

FIGURE 10.46 Effect of channel mismatch on normalized improvement factor of untapered linear array. Quiescent noise-to clutter ratio was approx. -85 dB for $N = 25$, -88 dB for $N = 50$, -90 dB for $N = 75$, and -92 dB for $N = 100$.

10.2.5. Summary

It has been shown that internal clutter motion, aircraft crabbing, and scattering from near-field obstacles can limit the effectiveness of space–time processing in eliminating airborne clutter. An analytical expression (Equation 10.71) was developed to show how each of these effects produces a deterioration in the signal-to-clutter-plus-noise ratio achievable. Furthermore, by studying the spectral decomposition of the covariance matrix it was found that the effects of both internal clutter motion and crabbing could either be compensated by artificially adding noise or by processing more pulses. The latter approach is the method of choice, when possible, especially in a low signal-to-clutter-plus-noise environment.

A near-field obstacle limits performance in a different fashion, and produces a spread of the clutter into all of sine azimuth–Doppler space. It was shown that the space–time processor attempts to compensate for this effect by placing a near-field null on the obstacle. Thus, adding more spatial degrees of freedom (i.e., more elements) is much more effective in eliminating this effect than is adding more temporal degrees of freedom (i.e., processing more pulses).

Channel mismatch also limits effectiveness, and this can be alleviated by controlling the dispersive errors more tightly and by increasing the number of receive elements.

Some people have argued that because the processor attempts to place nulls on near-field obstacles, it might be better to apply a near-field constraint, *a priori*. In Appendix J it is demonstrated that this is a suboptimum solution, and it is always better to allow the processor to determine the near-field structure, because it does it in the fashion that maximizes the signal-to-noise-plus-clutter ratio.

10.3. CLUTTER COVARIANCE SMOOTHING BY SUBAPERTURE AVERAGING

(R. L. FANTE, E. C. BARILE, AND T. P. GUELLA)

10.3.1. Introduction

The detection of small targets with an airborne radar is complicated by the presence of significant sidelobe clutter in the same Doppler bin as the target. The conventional cure is to decrease the sidelobe level, but for targets with very small radar cross section the sidelobe level required may not be achievable in practice. However, one can still eliminate the offending clutter, while also eliminating any jammers present, by using space–time processing to adaptively place nulls on the undesirable sidelobe clutter (and jammers). It has been shown previously[1,2] that an airborne space–time adaptive radar (consisting of an antenna array with an adaptive tapped delay line behind each element) can cancel ground clutter and interferers, while simultaneously preserving a target return, by applying a weight

vector $\mathbf{w}$ to the space–time array, where

$$\mathbf{w} = C\Phi^{-1}\mathbf{s}^* \tag{10.74}$$

In Equation 10.74, C is a constant, s is the target steering vector, and Φ is the covariance matrix of the interference and has a typical component $\Phi_{nm}(r,p) = \langle V_n(rT)V_m^*(pT)\rangle$, where $V_n(rT)$ is the interference (clutter-plus-jammers-plus-noise) voltage received on the nth element for the rth pulse, T is the interpulse period, and $\langle\rangle$ denotes an expectation. One of the key considerations in applying Equation 10.74 is the estimation of Φ. The usual approach is to obtain an estimate using the samples $V_n(t = rT + \mu(\Delta R/c))$ of the voltage in the μth range bin for the rth pulse, where ΔR is the range resolution and c is the speed of light. Then the estimate of $\Phi_{nm}(r,p)$ is

$$\hat{\Phi}_{nm}(r,p) \approx \frac{1}{M}\sum_{\mu=1}^{M} V_n(rT + \mu\Delta R/c)V_m^*(pT + \mu\Delta R/c) \tag{10.75}$$

where M is the number of independent range bins. As shown by Reed et al.[3] this provides an acceptable estimate of $\Phi_{nm}(r,p)$ as long as $M \geq 2N_{\mathrm{DOF}}$, where N_{DOF} is the number of space–time degrees of freedom. However, the maximum number of independent range bins is limited by the ambiguous range of the radar. That is, the range resolution of the radar is $c/2B$, where B is the bandwidth and the ambiguous range is $c/2(\mathrm{PRF})$, PRF being the pulse repetition frequency, so that the number of independent range bins is the number of resolution cells in the ambiguous range. This gives $M = B/\mathrm{PRF}$ independent range bins, or samples. Thus, as long as $B/\mathrm{PRF} > 2N_{\mathrm{DOF}}$, there are sufficient samples to estimate Φ, but it is clear that for high PRF radars it may be difficult to obtain a sufficient number of independent samples by range bin averaging. For this case an adjunct is needed to provide additional smoothing of the covariance. This can be achieved by subaperture averaging,[4–6] which is sometimes used for other purposes. The idea behind subaperture averaging, or spatial smoothing, is the recognition that the estimate of Φ_{nm} for the nth and mth elements is effectively repeated in $\Phi_{n+\nu,m+\nu}$, where ν is an integer, as long as the array elements are uniformly spaced, and the incident field is spatially statistically stationary. Thus, if we have an N-element array and are willing to suffer the loss in resolution entailed in using only N_{A} elements at a time, then (unless $\hat{\Phi}_{nm}$ and $\hat{\Phi}_{n+\nu,m+\nu}$ are highly correlated) we can smooth the covariance by forming

$$\bar{\Phi}_{nm}(r,p) = \frac{1}{K+1}\sum_{\nu=0}^{K} \hat{\Phi}_{n+\nu,m+\nu}(r,P) \tag{10.76}$$

where $K + 1 = N - N_{\mathrm{A}} + 1$ is the number of subarrays. This procedure is illustrated graphically in Figure 10.47 for the case where $N = 10$ and $N_{\mathrm{A}} = 6$. Although it is clear that, in principle, the average in Equation 10.76 should reduce the variance of $\Phi_{nm}(r,p)$, we now need to quantitatively assess its effectiveness.

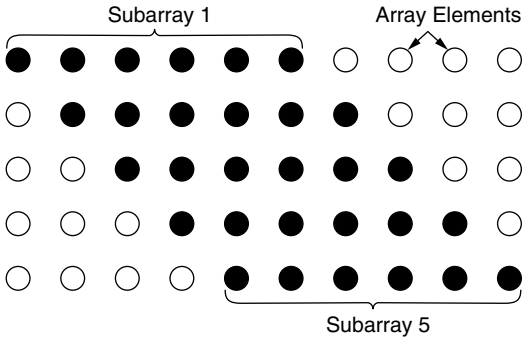

FIGURE 10.47 Ten-element array ($N = 10$) subarrayed to form five ($K + 1 = 5$) overlapped subarrays of six elements ($N_A = 6$).

If we substitute Equation 10.75 and Equation 10.76 and then define

$$\tilde{\Phi}_{nm}(\mu) = \frac{1}{K+1} \sum_{\nu=0}^{K} V_{n+\nu}(rT + \mu\Delta R/c) \times V^*_{m+\nu}(pT + \mu\Delta R/c) \quad (10.77)$$

we can rewrite Equation 10.76 as

$$\bar{\Phi}_{nm}(r,p) = \frac{1}{M} \sum_{\mu=0}^{M} \hat{\Phi}_{nm}(\mu) \quad (10.78)$$

where the arguments r and p have been suppressed in the definition of $\tilde{\Phi}_{nm}$. The mean square difference between $\bar{\Phi}_{nm}$ and the true covariance Φ_{nm} is

$$\sigma^2_{nm}(K+1) = \left\langle \left| \frac{1}{M} \sum_{\mu=1}^{M} \tilde{\Phi}_{nm}(\mu) - \langle \Phi_{nm} \rangle \right|^2 \right\rangle \quad (10.79)$$

If we observe from Equation 10.77 that $\langle \tilde{\Phi}_{nm} \rangle = \Phi_{nm}$, and assume that the values of $\tilde{\Phi}_{nm}(\mu)$ for different range bins are independent (because the range bins do not overlap and, therefore, the scatterers in different bins are different) and statistically stationary, we can rewrite Equation 10.79 as

$$\sigma^2_{nm}(K+1) = \frac{1}{M} [\langle |\bar{\Phi}_{nm}|^2 - |\Phi_{nm}|^2 \rangle] \quad (10.80)$$

In deriving Equation 10.80 from Equation 10.79 it was necessary in evaluating $\langle \tilde{\Phi}_{nm}(\mu) \tilde{\Phi}^*_{nm}(\nu) \rangle$ to separately evaluate the M terms for which $\mu = \nu$, and the $M(M-1)$ terms for which $\mu \neq \nu$. In the next section we evaluate the result in Equation 10.80 for airborne clutter.

10.3.2. Analysis for an Airborne Radar

In order to evaluate σ^2_{nm}, let us consider the case when the adaptive radar is mounted on an airborne platform moving with a constant speed v_A, and the interference is due to ground clutter only (the case of a finite number of

jammers can be treated as a special case of our analysis). Suppose the reference element on the radar transmits a pulse train

$$V_0(t) = \sum_{r'=-\infty}^{\infty} U(t - r'T)e^{i\omega t} \tag{10.81}$$

where ω is the radian signal frequency, and $U(t) = 1$ for $0 \le t \le \tau$ and $U(t) = 0$ otherwise. Then, if there are K_0 clutter scatterers with the kth one located at range R_k and azimuth θ_k, as shown in Figure 10.48, it is readily shown that the received voltage on the nth array element after down conversion is

$$V_n(t) = \sum_{r'} \sum_{k} a_k U\left[\alpha_k t - r'T - 2\frac{R_k}{c} - \frac{n\Delta}{c}\sin\theta_k\right] \times \exp\left\{i\omega\left[(\alpha_k - 1)t - \frac{2R_k}{c} - \frac{n\Delta}{c}\sin\theta_k\right]\right\}. \tag{10.82}$$

where a_k is the strength (including range effects) of the kth clutter scatterer, $\alpha_k = 1 - 2(v_A/c)\sin\theta_k$, Δ is the separation between adjacent array elements, and the summation on k is over all clutter scatterers. We now need to obtain the voltage received on the nth element in the μth range bin for the rth pulse. If the μth range bin is centered on range $R_\mu = \mu\Delta R/c$, we then evaluate Equation 10.82 at $t = rT + 2R\mu/c$. Then, because the pulses do not overlap, only the term $r' = r$ is nonzero in the summation on r'. Also, we assume that both $\nu_A T$ and the array length are small in comparison with the range resolution, $c\tau/2$. In this case we find from Equation 10.82 that the voltage on the nth element in the μth range bin on the rth pulse is

$$V_n(r, \mu) = \sum_{k} a_k U(\gamma_{k\mu})\exp[-i\phi_k - i\varphi_{nr}\sin\theta_k] \tag{10.83}$$

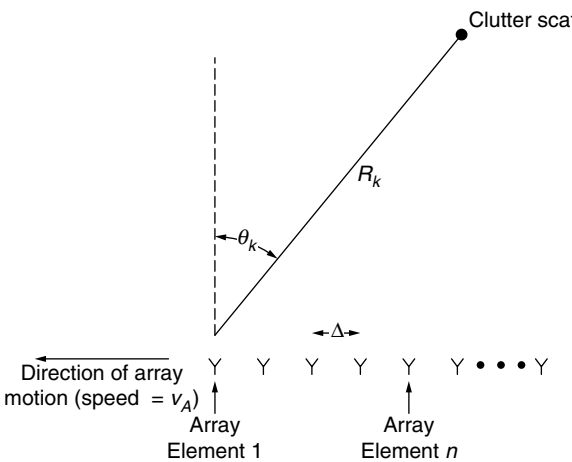

FIGURE 10.48 Location of kth clutter scatterer relative to antenna array.

where

$$\gamma_{k\mu} = \frac{2}{c}(R_k - R_\mu)$$

$$\phi_k = \frac{2\omega R_k}{c}$$

$$\varphi_{nr} = \frac{\omega}{c}\left(n\Delta + 2rv_A T + 2\frac{v_A}{c}R_\mu\right)$$

and we have abbreviated $V_n(rT + \mu\Delta R/c)$ by $V_n(r, \mu)$. Therefore, for the μth range ring, we have

$$V_n(r,\mu)V_m^*(p,\mu) = \sum_k \sum_l a_k a_l^* U(\gamma_{k\mu})U^*(\gamma_{l\mu})\cdot\exp[\mathrm{i}(\phi_l - \phi_k) + \mathrm{i}(\varphi_{mp}\sin\theta_l - \varphi_{nr}\sin\theta_k)] \tag{10.84}$$

If Equation 10.84 is used in Equation 10.77, and the summation over ν is performed, we obtain

$$\tilde{\Phi}_{nm}(\mu) = \sum_k \sum_l a_k a_l^* U(\gamma_{k\mu})U^*(\gamma_{l\mu})\ \exp[\mathrm{i}(\phi_l - \phi_k) + \mathrm{i}(\varphi_{mp}\sin\theta_l - \varphi_{nr}\sin\theta_k) + \mathrm{i}K(\rho_k - \rho_l)/2]\frac{\sin\left(\frac{K+1}{2}\right)(\rho_k - \rho_l)}{(K+1)\sin\frac{1}{2}(\rho_k - \rho_l)} \tag{10.85}$$

where $\rho_k = \omega\Delta\sin\theta_k/c = 2\pi(\Delta/\lambda)\sin\theta_k$ and $K = N - N_A$.

The location (R_k, θ_k) of the kth clutter scatterer can be treated as a random variable, with a probability density $p_k(R_k, \theta_k)$. Let us assume that all scatterer locations are statistically independent so that the joint probability density function $p(R_1, \theta_1; R_2, \theta_2; \cdots; R_N, \theta_N) = p_1(R_1, \theta_1)p_2(R_2, \theta_2)\cdots p_N(R_N, \theta_N)$; we further assume that $p_k(R_k, \theta_k) = p_R(R_k)p_\theta(\theta_k)$. Then, if all scatterer ranges are equally likely, we have $p_R(R_k) =$ constant, and the phase $\phi_k = 2\omega R_k/c$ can therefore be treated as a random variable that is uniformly distributed between 0 and 2π, with a probability density $p_\phi(\phi_k) = 1/(2\pi)$ for $0 \le \phi_k \le 2\pi$. Consequently,

$$\langle\exp[\mathrm{i}(\phi_k - \phi_l)]\rangle = \left(\frac{1}{2\pi}\right)^2\int_0^{2\pi}\mathrm{d}\phi_k\int_0^{2\pi}\mathrm{d}\phi_l\,\exp[\mathrm{i}(\phi_k - \phi_l)] = \delta_{kl} \tag{10.86}$$

and

$$\langle\exp[\mathrm{i}(\phi_k - \phi_l + \phi_r - \phi_s)]\rangle = \delta_{kl}\delta_{rs} + \delta_{ks}\delta_{lr} - \delta_{klrs} \tag{10.87}$$

where $\delta_{kl} = 1$ if $k = l$ and $\delta_{kl} = 0$, otherwise. Also, $\delta_{klrs} = 1$ if $k = l = r = s$ and $\delta_{klrs} = 0$, otherwise. Term δ_{klrs} is needed in Equation 10.77 because otherwise the $k = l = r = s$ term is counted twice. Equation 10.86 is used to average

Equation 10.85 over ϕ, we obtain

$$\langle \tilde{\Phi}_{nm} \rangle_\theta = \sum_k{}' |a_k|^2 \exp[\mathrm{i}(\varphi_{mp} - \varphi_{nr})\sin\theta_k] \tag{10.88}$$

where the prime indicates that the summation is only over clutter scatterers in the μth range ring and $\langle\rangle_\theta$ denotes an expectation on ϕ_k with all θ_k held constant. Likewise, using Equation 10.88 we can show that

$$\langle |\tilde{\Phi}_{nm}| \rangle_\theta = \left| \sum_k{}' |a_k|^2 \exp[\mathrm{i}(\varphi_{mp} - \varphi_{nr})\sin\theta_k] \right|^2 + \sum_k{}' \sum_{\substack{l \\ k \neq l}}{}' |a_k|^2 |a_l|^2 \left[\frac{\sin\left(\frac{K+1}{2}\right)(\rho_k - \rho_l)}{(K+1)\sin\frac{1}{2}(\rho_k - \rho_l)} \right]^2 \tag{10.89}$$

Finally, if Equation 10.88 and Equation 10.89 are substituted into Equation 10.80 we find

$$\sigma^2_{nm}(K+1) = \frac{1}{M} \sum_k{}' \sum_{\substack{l \\ k \neq l}}{}' |a_k|^2 |a_l|^2 \times \left\langle \left[\frac{\sin\left(\frac{K+1}{2}\right)(\rho_k - \rho_l)}{(K+1)\sin\frac{1}{2}(\rho_k - \rho_l)} \right]^2 \right\rangle \tag{10.90}$$

where $\rho_k = (\omega\Delta/c)\sin\theta_k = (2\pi\Delta/\lambda)\sin\theta_k$, and the expectation $\langle\rangle$ in Equation 10.90 is an expectation over the scatterer polar angles θ_k. The expectation in Equation 10.90 has been performed by Monte Carlo methods and the results are presented in Figure 10.49 for the case when the scatterer amplitudes a_k are all equal. The results shown are the averages over 100 realizations, each containing 512 randomly located scatterers per range ring. The normalized variance shown on the ordinate of Figure 10.49 has been defined as $\sigma^2_{nm}(K+1)/\sigma^2_{nm}(1)$ where

$$\sigma^2_{nm}(1) = \frac{1}{M} \sum_k{}' \sum_{\substack{l \\ k \neq l}}{}' |a_k|^2 |a_l|^2 \tag{10.91}$$

From Figure 10.49 we see that the normalized variance decreases approximately as $(K+1)^{-1}$ when the number of subapertures is increased. Thus, subaperture averaging does indeed lead to significant smoothing of the covariance. In fact, if we recognize that the variance per range ring is given by

$$\sigma^2_{nm} = \sum_k{}' \sum_{\substack{l \\ k \neq l}}{}' |a_k|^2 |a_l|^2 \tag{10.92}$$

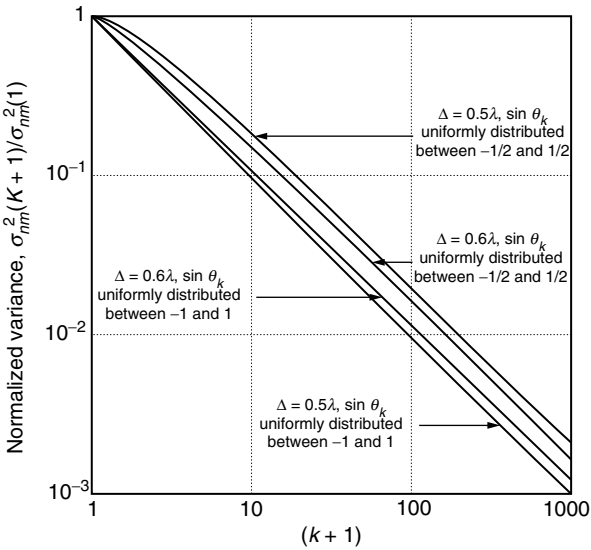

FIGURE 10.49 Normalized variance of (n, m) component of covariance matrix after subaperture averaging.

then it is evident from Equation 10.90, Equation 10.91, and Figure 10.76 that after averaging over both range bins and subapertures the variance is

$$\sigma_{nm}^2(K+1) \approx \frac{1}{M(K+1)}\tilde{\sigma}_{nm}^2 \tag{10.93}$$

Consequently, the variance per range bin has been reduced by the product of the number of range bins and subapertures.

We have also performed the calculations for the case of J jammers. In that case we again obtain the result in Equation 10.90, except that the sums are now over all jammers (rather than all clutter scatterers in a given range ring) and there is no expectation on the polar angles θ_k (because a different realization would still have the jammer at the same angle. Only the phase angles would differ). The numerical results show the same trend observed in Figure 10.49, and are not repeated here.

10.3.3. Summary

It has been demonstrated that the covariance matrix estimate is indeed smoothed by using subaperture averaging. In particular, we have shown that if we have $K+1$ subapertures ($K=0$ corresponds to $N_A = N$ so that there is no spatial smoothing), and M independent range bins, then the equivalent number of independent samples is approximately $(K+1)M$. Therefore, the maximum number N_{DOF} of space–time degrees of freedom that the adaptive radar can have is

$$N_{DOF} \approx \frac{(K+1)M}{2} \tag{10.94a}$$

If we recall that the maximum value of M is B/PRF, we can rewrite this condition as

$$N_{\mathrm{DOF}} = \frac{(K+1)B}{2\mathrm{PRF}} \tag{10.94b}$$

For example, suppose $B = 0.7$ MHz, PRF $= 20$ kHz, and we have $(K+1) = 4$ overlapped subapertures (this implies using three additional array elements because $K + 1 = N - N_{\mathrm{A}} + 1$, so that $N - N_{\mathrm{A}} = 3$). Then, from Equation 10.94b we see that the maximum number of space–time degrees of freedom $N_{\mathrm{DOF}} = 70$. Therefore, subaperture averaging allows us to have sufficient samples to estimate the covariance matrix for a 35 element adaptive space–time array that processes two pulses (the array would actually have 38 elements, but the final three elements are only used to smooth the covariance). This would not be otherwise possible for the high PRF chosen.

10.4. CANCELLATION OF DIFFUSE JAMMER MULTIPATH BY AN AIRBORNE ADAPTIVE RADAR

(R. L. FANTE AND J. A. TORRES)

10.4.1. Introduction

Airborne surveillance radars need to be able to detect targets in an environment that can include the case when hostile airborne and ground-based jammers are present. The surveillance radar can overcome the interference from a ground-based jammer located in the radar sidelobes either by employing very low sidelobes or placing an adaptive spatial null at the jammer location. However, airborne jammers present a special problem because of multipath (i.e., reflection of the jammer interference off the Earth into the radar), even when the jammer is located in the sidelobes of the radar. In regions where the Earth is very smooth (e.g., tranquil sea) this multipath appears at the same azimuth as the direct jammer interference. Both components can usually be canceled with a single spatial null, or by using very low radar sidelobes. However, in regions where the surface of the Earth is diffuse (rough ground, high sea states) the jammer interference is usually scattered off a large portion of the terrain, including that portion subtended by the mainbeam of the radar. In this case, the integrated, diffuse jammer multipath entering the radar through the azimuths subtended by its mainbeam can be considerably larger than that entering through the radar sidelobes. Thus, diffuse multipath creates a unique mainbeam jamming problem, because the jammer interference is spatially distributed in azimuth throughout the mainbeam. Placing a spatial null at a single azimuth in the mainbeam would be ineffective.[†] Furthermore, one cannot place nulls everywhere in the mainbeam

[†] It has been assumed that because of size limitations, the airborne jammer is not highly directional, and the surveillance radar elevation beamwidth is sufficiently large that a portion illuminates the ground; in practical applications these assumptions are usually valid.

because then the target signal would be canceled. Thus, conventional spatial nulling is not a solution, and a different approach is required.

The time-domain approach to cancel such jammer interference is to obtain a replica of the interference, pass it through a delay line that compensates for any time differences between the interference and its replica, and then subtract the two. This idea can be generalized to the cancellation of jammer multipath by considering the geometry in Figure 10.50, where one places an auxiliary beam adjacent to the mainbeam, in order to try to obtain a delayed replica of the interference received through the mainbeam. For a sidelobe jammer, the jammer-to-noise ratio in the auxiliary beam is expected to be of the same order as that in the mainbeam. Then, if the interference received in the auxiliary beam is passed through a tapped delay line that compensates for time differences between the main and auxiliary interferences, the mainbeam interference can be partially or fully canceled. The degree of cancellation depends on the degree of correlation existing between the mainbeam and auxiliary beam interference. A high degree of correlation (which yields good cancellation of mainbeam interference) requires that the product of bandwidth and all time delays between channels be small in comparison with unity.[‡] Note that because the target signal is present in the mainbeam, but only weakly present in the auxiliary beams (because it is in the spatial sidelobes of the auxiliary), the target signal is not canceled when the time-shifted auxiliary channel voltage is subtracted from the main channel voltage.

Although the system discussed above may be reasonably effective in canceling mainbeam diffuse jammer multipath, one might expect that more orthogonal, auxiliary beams are necessary if one is also to cancel the diffuse jammer multipath that enters the radar through its sidelobes. This implies that the maximum cancellation occurs when all possible auxiliary beams are used. However, once the number of beams equals the number of elements, then it does not matter whether the cancellation is performed in element space or beam space because the two are equivalent, as demonstrated in Appendix K.

This discussion suggests that the element space architecture required to cancel diffuse jammer multipath is of the form shown in Figure 10.51. The appropriate tap spacing T has been studied previously[1–3] and the optimum selection is such that $BT \approx 0.3$ to 0.7, where B is the radar bandwidth (it is assumed that the jammer bandwidth exceeds the radar bandwidth). The number of delays $(M-1)$ is chosen so that the range of delays expected is encompassed by the total delay $(M-1)T$ of the tapped delay line. Of course, as noted

[‡] That is, if one wishes to cancel an interference voltage $j(t)$ with a delayed replica $j(t-\tau)$ then the mean residue power after cancellation is $\langle |j(t)-j(t-\tau)|^2 \rangle = 2J[1-\rho(\tau)]$ where $J = \langle |j(t)|^2 \rangle$, $\langle\rangle$ denotes an expectation and $\rho(\tau)$ is the normalized autocorrelation function. For a Gaussian autocorrelation, we have $\rho(\tau) = \exp[-(\tau/\tau_c)^2]$, where the correlation time τ_c is equal to B^{-1}, and B is the bandwidth. Thus, in order for the residue power to be small, we require $\rho(\tau) \approx 1$, which, in turn, requires $\tau/\tau_c \ll 1$, or $B\tau \ll 1$.

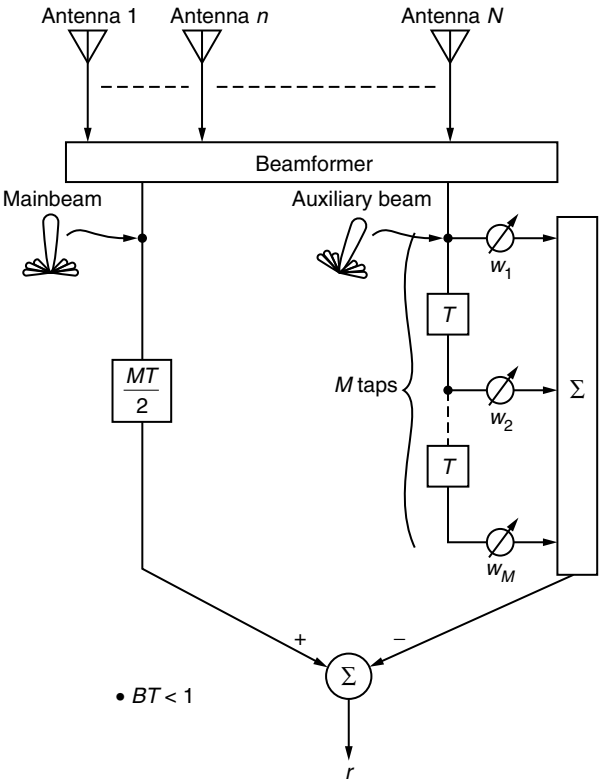

FIGURE 10.50 Generic two-beam adaptive architecture for canceling.

previously, the product of the differential multipath delay and the jammer bandwidth must also be less than unity.

The radar must also cancel conventional airborne clutter and, as illustrated in Ref. 4, this requires an adaptive tapped delay line with the taps (i.e., pulses) spaced by T_p, the inverse of the pulse-repetition frequency (PRF). Thus, the generic space–time processor architecture for canceling both diffuse jammer multipath and clutter requires antenna elements, pulses, and taps as shown in Figure 10.52.

Let us now briefly discuss how the weights in Figure 10.52 are calculated so as to maximize the signal-to-jammer-plus-clutter-plus-noise ratio. Suppose that the voltage present at the terminals of the nth antenna in Figure 10.52 is $U_n(t)$. Then, referring to Figure 10.52, the output voltage r_n in that channel is the weighted sum from M taps and L pulses, and can be written as

$$r_n(t) = \sum_{m=1}^{M} \sum_{l=1}^{L} w_{lm}(n) U_n[t - (l-1)T_p - (m-1)T] \qquad (10.95)$$

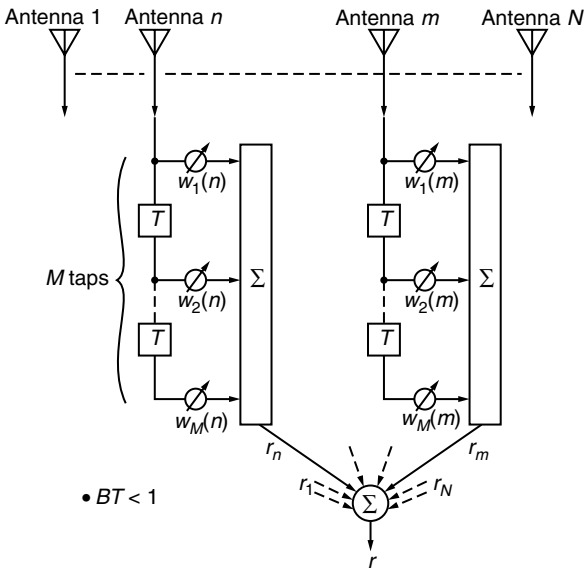

FIGURE 10.51 General fully adaptive element space architecture for canceling diffuse jammer multipath.

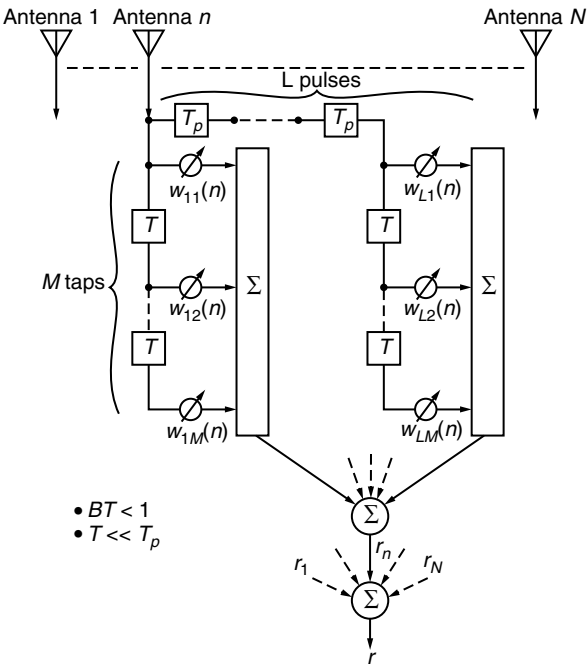

FIGURE 10.52 General fully adaptive element space architecture for canceling diffuse jammer multipath and clutter.

The output voltage r is the sum of all the antenna voltages, and is given by

$$r(t) = \sum_{n=1}^{N} \sum_{m=1}^{M} \sum_{l=1}^{L} w_{lm} U_n[t - (l-1)T_p - (m-1)T] \tag{10.96}$$

Equation 10.96 can be written in vector notation by defining the N M L-by-1 vectors

$$\mathbf{u} = \begin{bmatrix} U_1(t) \\ U_1(t-T) \\ \vdots \\ U_1[t-(L-1)T_p-(M-1)T] \\ U_2(t) \\ U_2(t-T) \\ \vdots \\ U_2[t-(L-1)T_p-(M-1)T] \\ U_3(t) \\ \vdots \\ U_N[t-(L-1)T_p-(M-1)T] \end{bmatrix}, \qquad \mathbf{w} = \begin{bmatrix} w_{11}(1) \\ w_{12}(1) \\ \vdots \\ w_{LM}(1) \\ w_{11}(2) \\ w_{12}(2) \\ \vdots \\ w_{LM}(2) \\ w_{11}(3) \\ \vdots \\ w_{LM}(N) \end{bmatrix} \tag{10.97}$$

Then, in terms of these vectors, we have

$$r - \mathbf{w}^{\mathrm{T}}\mathbf{u} \tag{10.98}$$

where the superscript T denotes a transpose.

It is then readily shown[4–7] that the weight vector that maximizes the signal-to-jammer-plus-clutter-plus-noise ratio is

$$\mathbf{w} = \mu \Phi^{-1} \mathbf{s}^* \tag{10.99}$$

where μ is a constant, Φ is the ideal interference (jammer-plus-clutter-plus noise) covariance matrix defined as

$$\Phi = \langle \mathbf{u}^* \mathbf{u}^{\mathrm{T}} \rangle \tag{10.100}$$

where $\langle\rangle$ denotes an expectation, and $\mathbf{s}$ is the steering vector. The steering vector represents the voltage present at each of the M taps and L pulses in Figure 10.52

when the desired target is present. That is, the desired target produces a vector

$$\mathbf{s} = \begin{bmatrix} s_1(t) \\ \vdots \\ s_1[t-(L-1)T_p-(M-1)T] \\ s_2(t) \\ \vdots \\ s_N[t-(L-1)T_p-(M-1)T] \end{bmatrix} \tag{10.101}$$

When the adaptive weights, as given by Equation 10.99, are applied it is readily shown[4,5] that the signal-to-interference ratio after adaptation is given by

$$\left(\frac{S}{I}\right)_a = \mathbf{s}^{\mathrm{T}} \Phi^{-1} \mathbf{s}^* \tag{10.102}$$

In the next section we compute the ideal covariance matrix and steering vector for the case when jammers, jammer multipath, clutter, and noise are all present, and when the processing is such that the received signal is down-converted (demodulated to baseband) and passed through a matched filter before being subjected to the processing indicated in Figure 10.52. Then in Section 10.4.3, we present numerical results to demonstrate that the proposed architecture is effective in canceling the direct-jammer interference, jammer multipath, and conventional clutter without simultaneously canceling the desired signal. In Section 10.4.4, we summarize our findings, discuss their limitations, and propose areas for future research.

10.4.2. Filtered Received Signals

In this section we calculate the received signals for the jammer, noise, and target. The received jammer voltage consists of a direct-path contribution and multiple, multipath components scattered off the ground. These signals include down-conversion to baseband and matched filtering. The ideal interference covariance matrix and steering vector are also computed.

10.4.2.1. Received Jammer and Noise Signals

Let us refer to Figure 10.53, and assume the ground is subdivided into K subareas, with the kth region having an effective area A_k and bistatic radar cross section per unit area σ_k. We also assume the jammer gain in the direction of the radar is G_0, its gain in the direction of the kth surface patch is G_k, and that the jammer radiates a voltage $j(t)$, with the impedance normalized such that the jammer power P_J is equal to $|j|^2$. Then, if we also assume that each receive element of the radar antenna is isotropic, the total interference voltage on the nth receive

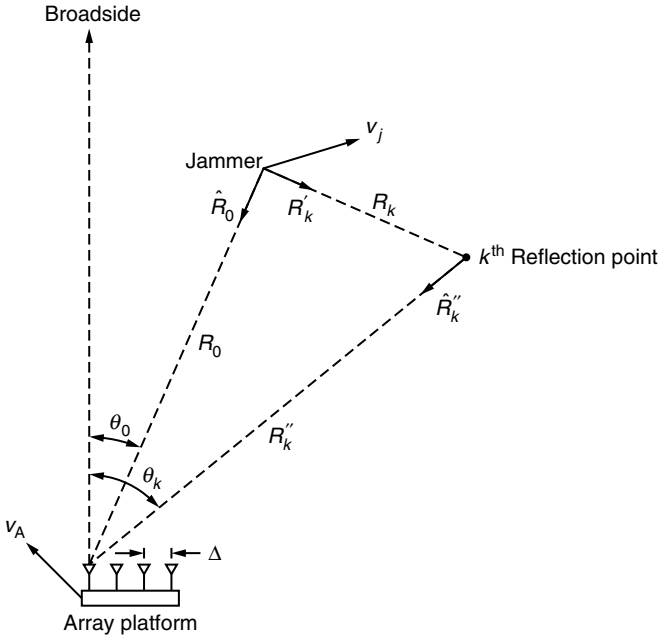

FIGURE 10.53 Model for jammer multipath.

element due to the direct jammer, its multipath, and noise can be written as

$$\tilde{V}_n(t)=\frac{\sqrt{G_0\lambda}}{4\pi R_0}j(\beta_0 t-\tau_{0n})+\frac{\lambda}{(4\pi)^{3/2}}\sum_{k=1}^{K}\frac{\sqrt{\sigma_k A_k G_k}}{R_k' R_k''}j(\beta_k t-\tau_{kn})+\tilde{x}_n(t) \qquad (10.103)$$

where λ is the wavelength, R_0, R_k', and R_k'' are shown in Figure 10.53, $\tilde{x}_n$ is the noise on the nth receive element and

$$\tau_{0n}=\frac{R_0}{c}-(n-1)T_0 \qquad (10.104)$$

$$\tau_{kn}=\frac{1}{c}(R_k'-R_k'')-(n-1)nT_k; \quad k\neq 0 \qquad (10.105)$$

$$T_0=\frac{\Delta}{c}\sin\theta_0\cos\psi_0 \qquad (10.106)$$

$$T_k=\frac{\Delta}{c}\sin\theta_k\cos\psi_k; \quad k\neq 0 \qquad (10.107)$$

$$\beta_0=1-\frac{1}{c}(v_A-v_J)\hat{R}_0 \qquad (10.108)$$

$$\beta_k=1-\frac{1}{c}(v_A\cdot\hat{R}_k''-v_J\cdot\hat{R}_k'); \quad k\neq 0 \qquad (10.109)$$

where v_A is the velocity of the radar platform, v_J is the jammer velocity, Δ is the separation between elements in the receive array, ψ_0 and ψ_k are, respectively, the depression angles of the jammer and the kth reflection point on the ground, and c is the speed of light. The first term in Equation 10.103 is the direct jammer signal, the second term represents the jammer multipath, and the last term is the noise. In deriving Equation 10.103 we have ignored conventional ground clutter, which is added later in our analysis.

If we now define the new voltage

$$V_n(t) = \frac{4\pi R_0}{\lambda (G_0)^{1/2}} \tilde{V}_n(t) \tag{10.110}$$

and a new noise

$$x_n(t) = \frac{4\pi R_0}{\lambda (G_0)^{1/2}} \tilde{x}_n(t) \tag{10.111}$$

we can rewrite Equation 10.103 as

$$V_n(t) = j(\beta_0 t - \tau_{0n}) + \sum_{k=1}^{K} a_k\, j(\beta_k t - \tau_{kn}) + x_n(t) \tag{10.112}$$

where

$$a_k = \left(\frac{\sigma_k A_k}{4\pi}\right)^{1/2} \left(\frac{G_k}{G_0}\right)^{1/2} \frac{R_0}{R_k' R_k''} \tag{10.113}$$

Equation 10.112 is the starting point for our analysis.

We now assume that the received interference signal $V_n(t)$ is filtered and then down-converted (these operations can also be done in the reverse order) by multiplying by $\exp(-\mathrm{i}2\pi f_c t)$, where f_c is the carrier frequency. Then, if $h(t)$ is the impulse response of the filter, the output voltage is

$$y_n(t) = \mathrm{e}^{-\mathrm{i}2\pi f_c t} \int_{-\infty}^{\infty} h(t - \tau) V_n(\tau) \mathrm{d}\tau \tag{10.114}$$

If $V_n(\tau)$ is substituted from Equation 10.112, this becomes

$$y_n(t) = \mathrm{e}^{-\mathrm{i}2\pi f_c t} \sum_{k=0}^{K} a_k \int_{-\infty}^{\infty} h(t - \tau) j(\beta_k \tau - \tau_{kn}) \mathrm{d}\tau$$

$$+ \mathrm{e}^{-\mathrm{i}2\pi f_c t} \int_{-\infty}^{\infty} h(t - \tau) x_n(\tau) \mathrm{d}\tau \tag{10.115}$$

where $a_0 \equiv 1$.

In practice, $h(t)$ is a filter matched to the transmitted signal. If the transmitted signal is $z(t)$, then

$$h(t) = Az^*(t_0 - t) \tag{10.116}$$

where t_0 is the time when the output is a maximum and A is a constant. For a chirp (linear FM) pulse, we find $A = (B/t_s)^{1/2}$ and $z(t)$ has the form

$$z(t) = \exp\left[\mathrm{i}2\pi f_c t + \mathrm{i}\frac{\pi B t^2}{t_s} \right] \cdot \mathrm{rect}(t/t_s) \tag{10.117}$$

where B is the signal or radar bandwidth, rect(x) is equal to unity for $0 \le x \le 1$ and is zero otherwise, and t_s is the temporal duration of the uncompressed pulse. Consequently, for a chirp signal

$$h(t) = \left(\frac{B}{t_s}\right)^{1/2} \exp\left[-\mathrm{i}2\pi f_c(t_0 - t) - \mathrm{i}\frac{\pi B(t_0 - t)^2}{t_s} \right] \mathrm{rect}\left(\frac{t_0 - t}{t_s}\right) \tag{10.118}$$

10.4.2.2. Interference Covariance Matrix

By using Equation 10.115 and Equation 10.118 we now compute the interference covariance matrix for the space–time processor discussed in the last section. The general term of interest is the jammer-plus-noise covariance matrix given by

$$Y_{nm}(t_1, t_2, t) = \langle y_n(t + t_1) y_m^*(t + t_2) \rangle \tag{10.119}$$

where the brackets $\langle\rangle$ denote an expectation. If we assume that $\langle j \rangle = \langle x \rangle = 0$, and that the jammer and noise voltages are statistically independent, we can substitute Equation 10.115 and Equation 10.118 into Equation 10.119 to obtain

$$\begin{aligned} Y_{nm}(t_1,t_2,t) = \Gamma_{12} \sum_{k=0}^{K} \sum_{l=0}^{K} a_k a_l^* \int_{-\infty}^{\infty} \mathrm{d}\tau \int_{-\infty}^{\infty} \mathrm{d}\tau \; & rh(t+t_1-\tau) \cdot h^*(t+t_2\tau') \\ \times \langle j(\beta_k \tau - \tau_{kn}) j^*(\beta_l \tau' - \tau_{\mathrm{ln}}) \rangle + \Gamma_{12} & \int_{-\infty}^{\infty} \mathrm{d}\tau \int_{-\infty}^{\infty} h(t+t_1-\tau) \\ \times h^*(t+t_2-\tau') \langle x_n(\tau) x_m^*(\tau') \rangle & \end{aligned} \tag{10.120}$$

where $a_0 \equiv 1$, and

$$\Gamma_{12} = \exp[-\mathrm{i}2\pi f_c(t_1 - t_2)] \tag{10.121}$$

If the receiver noise is a stationary white random process, then

$$\langle x_n(\tau) x_m^*(\tau') \rangle = S_{xx}(f_c) \delta(\tau - \tau') \delta_{nm} \tag{10.122}$$

where $\delta(\cdots)$ is the Dirac delta function, δ_{nm} is the Kronecker delta, $S_{xx}(f)$ is the power spectrum of the noise, and we have assumed that the noise voltages on the different antennas are independent. Likewise, if the jamming signal is also

a stationary white random process

$$\langle j(\beta_k\tau-\tau_{kn})j^*(\beta_l\tau'-\tau_{lm})\rangle = S_{jj}(f_c)\delta(\beta_k\tau-\beta_l\tau'-\tau_{km}+\tau_{lm}) = \frac{S_{jj}(f_c)}{|\beta_l|}\delta\left(\frac{\beta_k}{\beta_l}\tau-\tau'-\frac{\tau_{kn}}{\beta_l}+\frac{\tau_{lm}}{\beta_l}\right) \quad (10.123)$$

where $S_{jj}(f_c)$ is the power spectral density of the jammer. Consequently, if we substitute Equation 10.122 and Equation 10.123 into Equation 10.120, we obtain

$$Y_{nm}(t_1,t_2,t) = \Gamma_{12}S_{jj}(f_c)\sum_{k=0}^{K}\sum_{l=0}^{K}\frac{a_k a_l^*}{|\beta_l|}\int_{-\infty}^{\infty}d\tau\cdot h(t+t_1-\tau)h^*(t+t_2-\alpha_{kl}\tau-\gamma) + \delta_{nm}\Gamma_{12}S_{xx}(f_c)\int_{-\infty}^{\infty}d\tau\cdot h(t+t_1-\tau)h^*(t+t_2-\tau') \quad (10.124)$$

where

$$\alpha_{kl}=\frac{\beta_k}{\beta_l} \quad (10.125)$$

and

$$\gamma\equiv\gamma(k,l,n,m)=\frac{\tau_{lm}-\tau_{kn}}{\beta_l} \quad (10.126)$$

We can express Equation 10.124 in terms of the transmitted signal $z(t)$ by using Equation 10.116 with $A=(B/t_s)^{1/2}$. The result is

$$Y_{nm}(t_1,t_2,t) = \Gamma_{12}S_{jj}(f_c)\left(\frac{B}{t_s}\right)\sum_{k=0}^{K}\sum_{l=0}^{K}\frac{a_k a_l^*}{|\beta_l|}\int_{-\infty}^{\infty}d\tau\cdot z^*(t_0-t-t_1+\tau) \times z(t_0-t-t_2+\alpha_{kl}\tau+\gamma)+\delta_{nm}\Gamma_{12}S_{xx}(f_c)\left(\frac{B}{t_s}\right)\int_{-\infty}^{\infty}d\tau \times z^*(t_0-t-t_1+\tau)z(t_0-t+t_2+\tau) \quad (10.127)$$

where $z(t)$ is given by Equation 10.117. Letting

$$I_{kl}(n,m,t_1,t_2,t) = \int_{-\infty}^{\infty}d\tau\cdot z^*(t_0-t-t_1+\tau)\cdot z(t-t-t_2+\alpha_{kl}\tau+\gamma) \quad (10.128)$$

$$I'(t_1,t_2) = \int_{-\infty}^{\infty}dr\cdot z^*(t_0-t-t_1+\tau)\cdot z(t_0-t-t_2+\tau) \quad (10.129)$$

we obtain

$$Y_{nm}(t_1,t_2,t) = \Gamma_{12}S_{jj}(f_c)\left(\frac{B}{t_s}\right)\sum_{k=0}^{K}\sum_{l=0}^{K}\frac{a_k a_l^*}{|\beta_l|}I_{kl}(n,m,t_1,t_2,t) \times \delta_{nm}\Gamma_{12}S_{xx}(f_c)\left(\frac{B}{t_s}\right)I'(t_1,t_2) \quad (10.130)$$

The integrals in Equation 10.128 and Equation 10.129 have been evaluated in Appendix L for the case when $z(t)$ is given by Equation 10.117. Note that because of Doppler effects the covariance of the jammer multipath is temporally nonstationary (i.e., I_{kl} depends on time t). As we later see, this means that if the jammer multipath is to be canceled, the weights applied must also vary with time. As we noted earlier, cancellation of diffuse multipath requires that there be correlation between the scattered interference from scatterers in the mainbeam and those in the sidelobes of the radar. This requires that for such scatterers I_{kl} do not vanish and, in particular, that I_{kl}/t_s be of order unity. Based on this requirement it is shown in Appendix L that the necessary conditions for effective cancellation of jammer multipath are:

(1) The differential delay between scatterers must be small in comparison with the uncompressed pulse length.
(2) The product of the differential Doppler frequency and the uncompressed pulse length must be small in comparison with unity.
(3) The product of the radar bandwidth and the differential delay between scatterers must be small compared with unity.

It should also be noted that if multiple, independent jammers are present then the first term in Equation 10.130 is replaced by

$$\Gamma_{12}\left(\frac{B}{t_s}\right)\sum_{v=0}^{J} S_{jj}^{v}(f_c)\sum_{k=0}^{K}\sum_{l=0}^{K}\frac{a_{kv}a_{lv}^{*}}{|\beta_l|}I_{klv}(n,m,t_1,t_2,t)$$

where $S_{jj}^{v}(f_c)$ is the power spectrum of the vth jammer, J is the number of independent jammers, and a_{kv}, I_{klv} are the scattering amplitudes and integrals associated with the vth jammer, respectively.

The covariance due to conventional ground clutter was derived previously.[4] If we denote the components of the clutter covariance matrix by $C_{nm}(t_1,t_2)$ and assume that the clutter return is statistically independent of both the noise and the direct and scattered jammer signals, we then obtain for the (n, m) component of the total interference covariance

$$\Phi_{nm}(t_1,t_2,t) = Y_{nm}(t_1,t_2,t) + C_{nm}(t_1,t_2) \tag{10.131}$$

The times t_1 and t_2 can be expressed in terms of the pulse and tap indices. For reasons that will be obvious when we treat the steering vector, it is convenient to choose the time origin at the center (or as close to the center as possible) of the first tapped delay line. Thus, we can write $t_1 = p_1T_p + l_1T$ and $t_2 = pT_p + l_2T$, where we define $p_1 = (p_1' - 1)$, and $p_2 = (p_2' - 1)$, T_p is the pulse spacing, and T is the tap spacing. The indices p_1' and p_2' are the pulse indices and each ranges from 1 to L. Also, if the number of taps M is odd, we define $l_1 = l_1' - (M+1)/2$, and $l_2 = l_2' - (M+1)/2$, where l_1' and l_2' are the tap indices and each ranges from 1 to M. If the number of taps is even, then

$l_1 = l_1' - (M+2)/2$, and $l_2 = l_2' - (M+2)/2$, where l' and l_2' again each range from 1 to M.

10.4.2.3. Steering-Vector and Received Target Signal

In this section we calculate the steering vector after matched filtering and down-conversion. Suppose the transmitted pulse train is

$$z_x(t) = \sum_{p=1}^{P} u[t - (p-1)T_p]\mathrm{e}^{\mathrm{i}2\pi f_c t} \tag{10.132}$$

where T_p is the interpulse period, f_c is the carrier frequency and, for a linear FM (chirp)

$$u(t) = \mathrm{rect}\left(\frac{t}{t_s}\right)\exp\left(\mathrm{i}\frac{\pi B}{t_s}t^2\right) \tag{10.133}$$

The function rect(x) is equal to unity if $0 \le x \le 1$ and is zero otherwise. Also, B is the radar bandwidth and t_s is the temporal duration of the uncompressed pulse.

If the target is at range R_s, azimuth angel θ_s, and has a velocity $\mathbf{v}_s$, then the signal received on the nth element of the antenna array is

$$\begin{aligned} z_n(t) = &\sum_{p=1}^{P} u[(1-b_s)t - \tau_s - (n-1)T_s - (p-1)T_p] \\ &\times \exp\{\mathrm{i}2\pi f_\mathrm{c}[(1-b_s)t - \tau_s - (n-1)T_s]\} \end{aligned} \tag{10.134}$$

where:

$$T_s = \frac{\Delta}{c}\sin\theta_s\cos\psi_s \tag{10.135}$$

$$\tau_s = \frac{2R_s}{c} \tag{10.136}$$

$$b_s = -\frac{2}{c}(\mathbf{v}_\mathrm{A} - \mathbf{v}_S)\cdot\hat{\mathbf{R}}_s \tag{10.137}$$

Δ is the element spacing, $\mathbf{v}_\mathrm{A}$ is the aircraft velocity, $\hat{\mathbf{R}}_s$ is a unit vector joining the radar to the target, and ψ_s is the target depression angle. If the signal is passed through a matched filter and down-converted, we have

$$s_n(t) = \mathrm{e}^{-\mathrm{i}2\pi f_\mathrm{c}(t-t_0)}\int_{-\infty}^{\infty} h(t-\tau)z_n(\tau)\mathrm{d}\tau \tag{10.138}$$

where $h(\tau)$ is given by Equation 10.118. By using Equation 10.133, we can express $h(\tau)$ in Equation 10.119 in terms of $u(\tau)$. We find

$$h(\tau) = \left(\frac{B}{t_s}\right)^{1/2}\mathrm{e}^{-\mathrm{i}2\pi f_\mathrm{c}(t_0-\tau)}u^*(t_0-\tau) \tag{10.139}$$

Now substitute Equation 10.140 and Equation 10.139 into Equation 10.138. After some simple manipulations, we obtain

$$s_n(t) = \left(\frac{B}{t_s}\right)^{1/2} \sum_{p=1}^{P} \int_{-\infty}^{\infty} \mathrm{d}\tau \cdot u[(1-b_s)\tau - \tau_s - (n-1)T_s - (p-1)T_p] \times u^*(t_0 - t + \tau)\mathrm{e}^{-\mathrm{i}2\pi f_c(b_s\tau+\tau_s+(n-1)T_s)}. \quad (10.140)$$

Next let us evaluate Equation 10.140 at $t = t_0 + \tau_s + l_1T + p_1T_p$, and then make the transformation $\eta = \tau - \tau_s - l_1T - p_1T_p$. Then, because only the $p = p_1$ term remains in the summation in Equation 10.140. We get

$$\begin{aligned} &s_n(t_0 + \tau_s + l_1T + p_1T_p) \\ &= \left(\frac{B}{t_s}\right)^{1/2} \mathrm{e}^{-\mathrm{i}\phi} \int_{-\infty}^{\infty} \mathrm{d}\eta \cdot u[(1-b_s)(\eta + l_1T) - (n-1)T_s \\ &\quad - b_s(\tau_s + p_1T_p)]u^*(\eta)\mathrm{e}^{-\mathrm{i}2\pi b_s f_c \eta} \end{aligned} \quad (10.141)$$

where

$$\phi = 2\pi f_c[(b_s+1)\tau_s + b_s(l_1T + p_1T_p) + (n-1)T_s] \quad (10.142)$$

If we assume that $Bt_s \gg 1, |l_1T| \ll t_s, \pi BD/c \ll 1, f_c t_s \gg p_1$ and $f_c/B \gg p_1$, where D is the total linear dimension of the array, it can be shown that

$$s_n(p_1, l_1) \approx (Bt_s)^{1/2}\mathrm{e}^{-\mathrm{i}\phi_2}\ \mathrm{sinc}(\pi B l_1 T) \quad (10.143)$$

where

$$\phi_2 = 2\pi f_c(1+b_s)\tau_s + 2\pi f_c b_s p_1 T_p + 2\pi f_c(n-1)T_s - \pi B l_1 T \quad (10.144)$$

and the argument $(t_0 + \tau_s + l_1T + p_1T_p)$ has been replaced by (p_1, l_1) in Equation 10.143. Equation 10.143 is the expression used for the steering vector in the calculations to follow. The reader is reminded that l_1 and p_1 are related to the tap index l' and pulse index p'_1 by the relations $l_1 = l'_1 - (M+1)/2$ where M is odd, $l_1 = l'_1 - (M+2)/2$ when M is even and $p_1 = (p'_1 - 1)$, where the tap index l'_1 ranges from 1 to M, and the pulse index p'_1 ranges from 1 to L. Thus, from Equation 10.143 it is evident that the steering vector is centered at tap $l'_1 = (M+1)/2$ when M is odd and on $l'_1 = (M+2)/2$ when M is even.

10.4.3. Numerical Results

10.4.3.1. Introduction

In this section, we describe the model for the location of the diffuse jammer multipath scatterers. We then illustrate how the adaptive processing architectures discussed in Section 10.4.2 can provide cancellation of the diffuse multipath jamming, and in particular, the scattering entering through the radar mainbeam.

Ideally, one would like to model all of the terrain contributing to the diffuse jammer multipath by dividing the entire illuminated area into a grid of coherence regions, and randomly placing a scatterer within each region. However, due to the computational limitations of the simulation, only a subset of the multipath region is modeled. In Appendix L, we show that good cancellation of jammer interference from scatterers located in the radar mainbeam requires that they be highly correlated with jammer interference from scatterers located in the radar sidelobes. Specifically, the differential path delay between the mainbeam and sidelobe scatterers must be small in comparison with the reciprocal of the radar bandwidth B, and the differential Doppler frequency must be small in comparison with the inverse of the uncompressed pulse length. Therefore, we begin by defining a profile of points located on the ground such that the differential path delay between any two points on the profile is zero. Next a set of profiles are chosen to cover the entire extent of the multipath region. The profile separation ΔR is chosen to be approximately equal to the correlation length of the diffuse terrain, where $\Delta R/c$ is of the order of $1/B$. The total extent R of differential path delays covered equals the product of the profile separation ΔR and the number of profiles less one. Finally, a number of scatterers are randomly distributed throughout the region defined by the intersection of the mainbeam and first few sidelobes of the radar with the multipath region defined by the total extent of differential path delays, as illustrated in Figure 10.54. Note that when the sidelobe scatterers are located near the mainbeam, the differential Doppler frequencies between these scatterers and the mainbeam scatterers are also kept small.

For the illustrative examples of this section, we choose eighty to one hundred scatterers; half randomly located in the mainbeam of a 16 element, uniformly illuminated array from -2 to 2° in azimuth, and half located in the sidelobes from

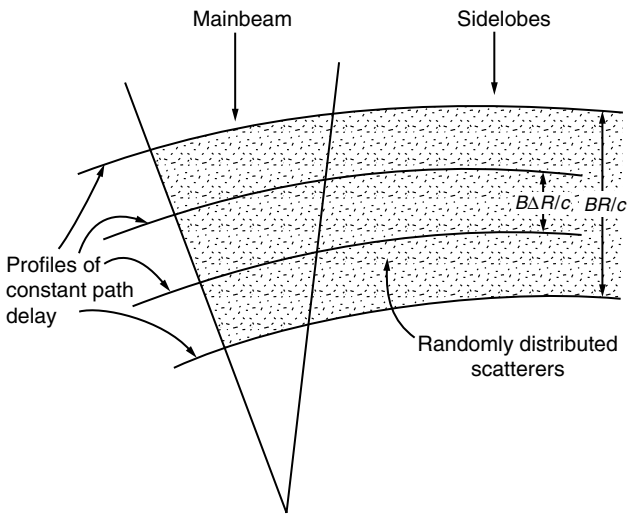

FIGURE 10.54 Model for location of diffuse jammer multipath scatterers.

TABLE 10.3
Simulation Parameters

Number of antenna elements = 16
Frequency = UHF
Pulse repetition frequency = 1000 Hz
Target azimuth = 0°
Jammer azimuth = 40°
Radar bandwidth $B = 1$ MHz
Tap spacing = $0.5/B$
Quiescent multipath JNR = 32.2 dB (direct jamming excluded)

9 to 12°. The frequency is UHF, and PRF is 1000 Hz, the bandwidth B is 1 MHz, and the queiscent jammer-plus-noise-to-noise ratio (JNR) is equal to 32.2 dB. Unless otherwise stated, the examples will not include the contributions from the direct jammer located in the radar sidelobes at 40° in azimuth. The simulation parameters are summarized in Table 10.3. The ability of the adaptive processor to cancel the diffuse multipath jamming is measured as the adaptive signal-to-interference ratio (SIR) at the processor output, in comparison with the adapted SIR achievable in the noise-only case. Specifically, performance is measured as a loss in adapted SIR, where the best achievable value is 0 dB.

Figure 10.55 illustrates the cancellation (assuming instantaneous updating of the weights) of diffuse multipath using the fully adaptive element space architecture shown in Figure 10.51. For this example, the tap spacing is $T = 0.5/B$, and the multipath region is modeled by 80 scatterers covering a total extent BR/c of 5.6. Increasing the number of temporal taps from 1 to 13 significantly

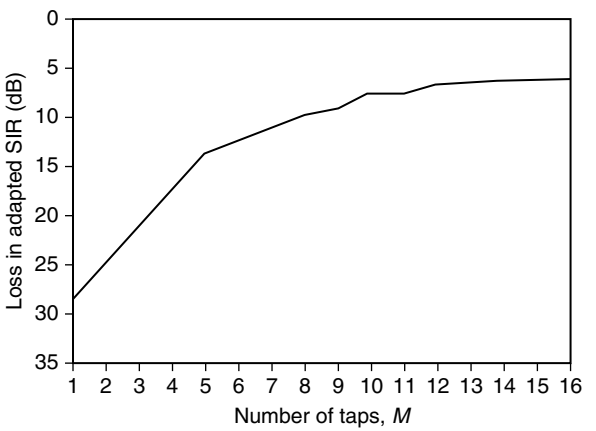

FIGURE 10.55 Cancellation of diffuse jammer multipath using temporal taps for $BT = 0.5$ and $BR/c = 5.6$.

improves the performance of the adaptive processor by 21.5 dB (i.e., reduces the loss in adapted SIR from 28.4 to 6.9 dB); increasing the number of taps from 13 to 16 only improves the performance by an additional 0.6 dB. This behavior is as expected, because the number of taps needed to encompass a total delay $\Delta\tau = R/c$ is given by

$$(M - 1)T = \Delta\tau \tag{10.145}$$

For $T = 0.5/B$, we then find

$$M = 2B\Delta\tau + 1 \tag{10.146}$$

For $BR/c = B\Delta\tau = 5.6$, Equation 10.146 gives $M = 2(5.6) + 1 = 12.2$. Thus, we would expect 12 or 13 taps would be needed to compensate for the multipath, and there is only marginal gain in adding additional taps beyond this number. Note that even for $M \rightarrow \infty$, the performance does not approach the noise-only value (0 dB), because the Doppler and delay match between the sidelobe and mainlobe scatterers is never perfect.

10.4.3.2. Tap Spacing

Figure 10.56 illustrates the performance of the fully adaptive element space processor, using the same multipath scenario modeled in the previous example, when the number of taps and tap spacing are varied. The optimal tap spacing for canceling the diffuse multipath depends on the number of taps used in the architecture. However, for a given number of taps, performance varies less than 3 dB when the tap spacing T varies from $0.3/B$ to $0.7/B$. Therefore, for the remainder of this work, the tap spacing is set to $T = 0.5/B$.

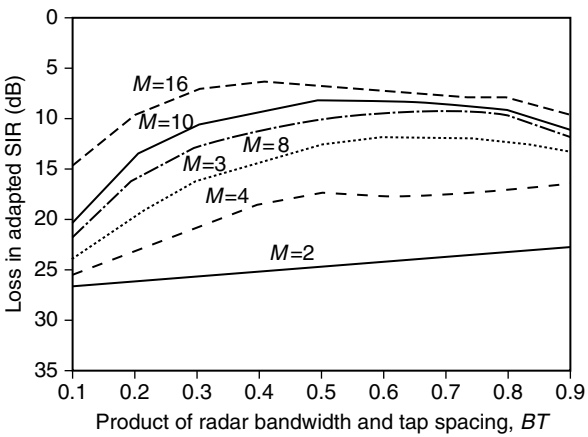

FIGURE 10.56 Effect of tap spacing on the cancellation of diffuse jammer multipath using M temporal taps for $BR/c = 5.6$.

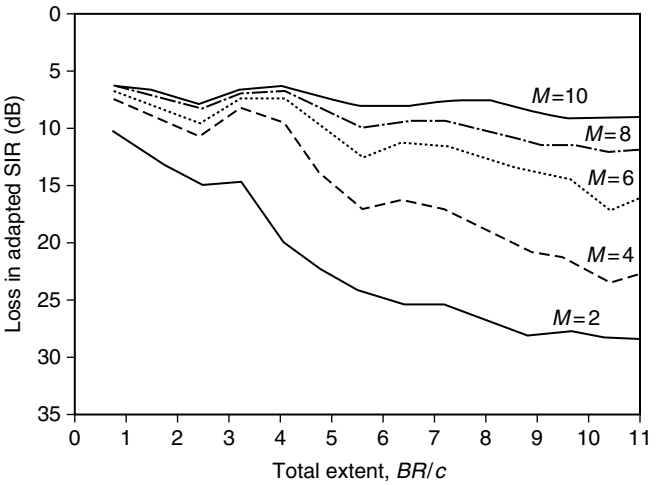

FIGURE 10.57 Effect of total extent on the cancellation of diffuse jammer multipath using M temporal taps for $BT = 0.5$.

10.4.3.3. Total Extent

In this section, we discuss the effect of the total extent of differential path delays on the performance of the element space processor. Figure 10.57 illustrates the loss in adapted SIR as a function of total extent when the number of temporal taps is varied. In general, performance degrades as the total extent of the diffuse multipath region increases. However, the losses in adapted SIR can be mitigated by processing more temporal taps until the required value given by Equation 10.146 is reached. We should comment that the reader may find the oscillations in performance puzzling. These oscillations have been noted previously[1–3] and are produced by the interactions between the scatterers.

10.4.3.4. Ground Clutter

Here, we illustrate how the fully-adaptive element space architecture augmented with PRI taps (i.e., pulses), as shown in Figure 10.52, can cancel both diffuse jammer multipath and ground clutter. For illustrative purposes, the ground clutter has no internal motion, so that two-pulse adaptive processing without temporal taps is sufficient to cancel the clutter when no jammer multipath interference is present.[4,5] For this example, the quiescent clutter-plus-noise-to-noise ratio (CNR) is 43.7 dB, and the multipath region is modeled by 90 scatterers covering a total extent of $BR/c = 6.4$. Figure 10.58 illustrates that without ground clutter (i.e., jammer multipath and noise only), performance improves by nearly 21 dB when the number of taps increases from one to 10. When the ground clutter and a single PRI delay (i.e., two pulses) are added, performance is reduced by less than 1 dB.

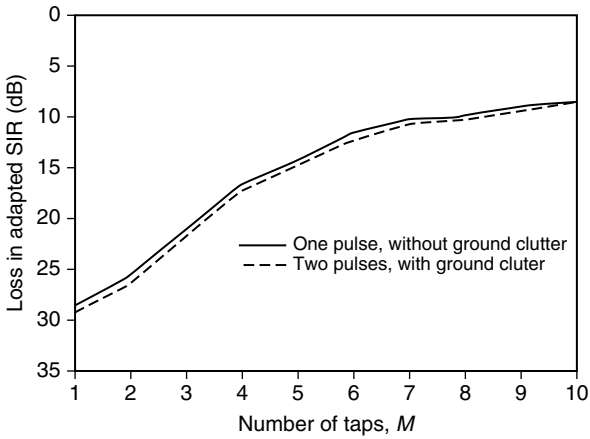

FIGURE 10.58 Cancellation of ground clutter and diffuse jammer multipath for $BR/c = 6.4$ and $BT = 0.5$.

10.4.3.5. Temporal Averaging

It should be stressed that the diffuse jammer multipath covariance matrix is time varying, requiring that the adapted weights be updated at roughly the rate of the highest differential Doppler frequency present. Because this can become computationally burdensome, it may be desirable to reduce the update rate and apply a set of average weights to all the data within a temporal interval. These average weights are calculated as

$$w = \mu \bar{\mathbf{\Phi}}^{-1} \mathbf{s}^* \tag{10.147}$$

where $\bar{\mathbf{\Phi}}$ is the interference covariance matrix averaged over a specified temporal interval and $\mathbf{s}$ is the steering vector. Performance degradations due to the temporal averaging are caused by the decorrelation of the jamming interference from different scatterers. Figure 10.59 illustrates the loss in adapted SIR as a function of the averaging time-interval, when the number of taps is varied. Note that the loss in adapted SIR is an average value over the averaging time-interval, and an averaging time of zero implies that the weights are updated instantaneously. We see that at UHF the performance degrades by less than 3 dB for averaging times up to two PRI (i.e., 2 ms). However, the degradation is larger at *S*-band because, as can be seen from Equation L-7 in Appendix L, the nonstationary portion of the covariance depends on the Doppler frequency difference, which increases as the carrier frequency is increased. In this example, the platform and jammer speeds were 350 nmi/hr.

10.4.3.6. Beam Space

The computational complexity of the fully adaptive element space processor is a significant consideration in practical implementations. The complexity can be

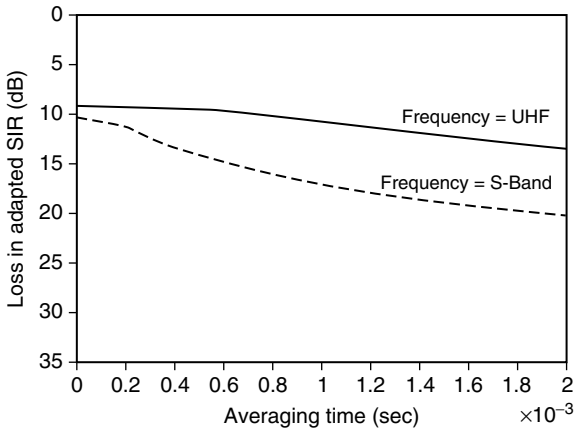

FIGURE 10.59 Effect of temporal averaging on the cancellation of diffuse jammer multipath using temporal taps for $BT = 0.5$ and $BR/c = 6.4$.

reduced, while maintaining performance, by using the beam space architecture. This approach uses beamforming of the array element data to form distinct orthogonal beams, where the number of beam is typically much less than the number of array elements. The spatial information is adaptively filtered using a tapped-delay line at the output of each beam. Appendix K shows that the performances of the element and beam space architectures are equivalent in the limit when the number of elements and beams are the same.

Figure 10.60 shows the performance of the beam space processor versus the number of beams, when the number of temporal taps is varied. The multipath

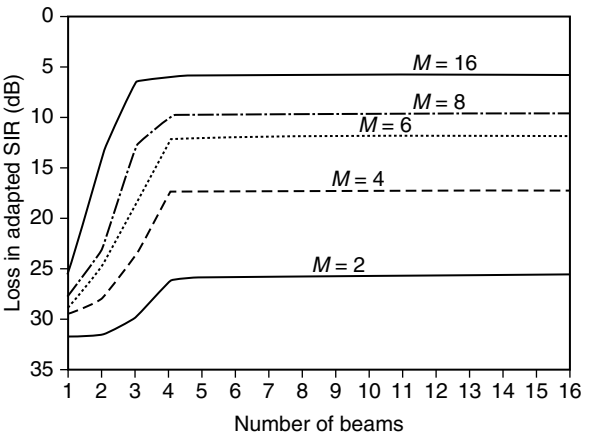

FIGURE 10.60 Effect of temporal averaging on the cancellation of diffuse jammer multipath using M temporal taps (without direct jammer included) for $BT = 0.5$ and $BR/c = 7.2$.

TABLE 10.4
Order of Beam Locations

Beam Number	Beam Angle in Azimuth (Degrees)
1	0.0
2	7.2
3	14.5
4	−7.2
5	22.0
6	−14.5
7	30.0
8	38.7
9	−22.0
10	48.6
11	−30.0
12	61.0
13	−38.7
14	−90.0
15	−48.6
16	−61.0

region is modeled by 100 scatterers covering a total extent of $BR/c = 7.2$. Note that we first form a set of orthogonal beams uniformly spaced in sine of azimuth space, where the number of beams equals the number of array elements. The beam corresponding to the target azimuth (i.e., 0°) is needed to maintain the target signal. Therefore, we order this beam as number one. The remaining beams are then selected based on descending order of received power. This results in the beam ordering of Table 10.4.

Figure 10.60 illustrates that performance degrades by less than 2 dB when three to four beams are used instead of all 16 beams or elements. This can be explained by analyzing the spatial information contained within several of the beams, as shown in Figure 10.61. The beams located at 7.2 and 14.5° contain the spatial information about the sidelobe scatterers, while the beams located at 0 and −7.2 contain spatial information about the mainbeam scatterers.

Figure 10.62 shows the performance of the beam space processor versus the number of beams, when the number of temporal taps is varied, and the contributions from the direct path jamming are now included. The multipath region modeled is identical to the one modeled in the previous example, and the quiescent JNR for the direct path jamming is 14.7 dB. We see that an additional beam is required to cancel the direct path jamming. In particular, the beam located at 38.7° is required since the jammer is located at 40°.

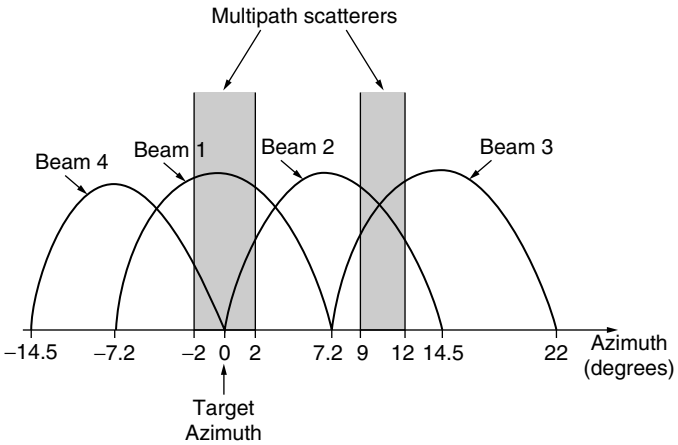

FIGURE 10.61 Azimuth location of the beams and multipath scatterers.

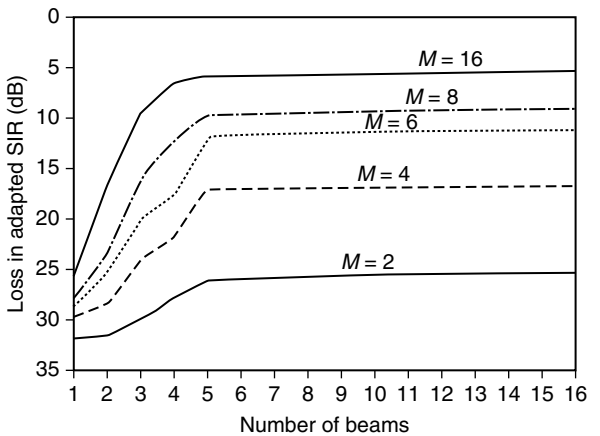

FIGURE 10.62 Effect of beam space processing on the cancellation of diffuse jammer multipath using M temporal taps (with direct jammer included) for $BT = 0.5$ and $BR/c = 7.2$.

10.4.4. Summary and Discussion

We have modeled the interference received by an airborne adaptive radar due to wideband noise, ground clutter, direct-path, sidelobe jammers, and diffuse jammer multipath. It has been demonstrated that, by using the architecture shown in Figure 10.52, it is possible to cancel all of the aforementioned interference, except for the noise. Much of our attention was devoted to the diffuse jammer multipath that is received primarily through the mainbeam of the radar and, hence, cannot be canceled by simply placing adaptive spatial nulls in the radiation pattern. It was found that the diffuse jammer multipath from scatterers

located in the radar mainbeam can be canceled with multipath from highly correlated scatterers located outside the mainbeam. The necessary conditions for highly correlated scatterers (i.e., good cancellation) are as follows:

$$|\Delta\tau_{lk}| \ll t_s \tag{10.148a}$$

$$|\Delta f_{lk}| t_s \ll 1 \tag{10.148b}$$

$$|B\Delta\tau_{lk}| \ll 1 \tag{10.148c}$$

where $\Delta\tau_{lk}$ and Δf_{lk} are, respectively, the differential time delay and Doppler frequency difference for the same antenna between a scatterer in the mainbeam and the scatterer outside the mainbeam; t_s is the uncompressed pulse length; and B is the bandwidth of the radar. When the above conditions are satisfied the diffuse jammer multipath can be canceled (along with ground clutter and direct-path, sidelobe jammers) using the architecture shown in Figure 10.52. The optimum tap spacing T was shown to be approximately $0.5/B$. Therefore, if the differential range extent of the multipath is R, then the total differential delay is R/c, and we expect that

$$M \approx 2BR/c + 1 \tag{10.149}$$

taps are required for effective multipath cancellation, as is clear from Figures 10.57 and 10.58.

It was also shown that, although the diffuse jammer multipath is a temporally nonstationary process, there is only a mild degradation if an average weight is calculated from a temporally averaged interference covariance matrix, in accordance with Equation 10.147, and applied to an entire PRI of data. Finally, we demonstrated that effective cancellation can be performed in the beam space, as well as in the element space, with a considerable reduction in processing requirements. An additional reduction in processing complexity is achieved by using suboptimum frequency domain processing. This approach will be discussed in a future section, and appears quite promising. The foregoing analysis was performed using an ideal interference covariance matrix. The effect of using a finite number of samples and how the samples can be obtained are deferred to a later section.

10.5. WIDEBAND CANCELLATION OF MULTIPLE MAINBEAM JAMMERS

(R. L. FANTE, R. M. DAVIS, AND T. P. GUELLA)

10.5.1. Introduction

MAINBEAM jammers offer a serious challenge to a radar system. The usual solution is to place a strobe in the jammer direction, but this effectively blanks out two beamwidths of azimuth scan. Furthermore, if there is an array of jammers with

an azimuthal spacing of several radar beamwidths, detection can be denied over an entire azimuth sector. Thus, it is desirable to devise a method of compensating for mainbeam jammers. The concept we will employ is to use auxiliary antennas that are separated[1] from the main antenna by sufficiently large distances (i.e., separations that are very much greater than the linear dimension of the main antenna) so that the array can place narrow nulls on the mainbeam jammers while still maintaining peak gain on a closely spaced target. The conventional approach is to divide the main aperture into subapertures and either to implement a fully adaptive array of subarrays or to form multiple adjacent beams and use them in a beamspace canceller. Each of these approaches breaks down when the spacing between the jammer and the quiet target becomes too small (less than about a quarter of a 3-dB beamwidth) because the desired target falls into the jammer notch. A second problem exists in the presence of multiple jammers. The system lacks the resolution to place independent nulls on closely spaced jammers. Physics demands that the effective aperture be increased to realize the requisite improvement in resolution and drives us to the architecture considered.

For narrowband jammers, the problem is relatively simple: one can place N narrowband nulls at angles $\theta_1 \cdots \theta_N$ by simply applying appropriate weights to the outputs of N sufficiently spaced auxiliary antennas, and then adding these to the main array output. However, in the wideband limit, this approach is not sufficient because the location of each null varies with frequency. To make the nulls at $\theta_1, \theta_2, \ldots, \theta_N$ wideband, one needs to place an adaptive finite impulse response (FIR) filter behind each of the auxiliary antennas.[2–6] This leads us to the architecture in Figure 10.63. The main array is modeled as M subarrays that span a total linear dimension D_M in the azimuth plane. The time delays $\tau_1, \tau_2, \ldots, \tau_M$ are required to steer the subarrays to the desired azimuthal scan angle $\theta = \theta_0$. Jammer cancellation is provided by N auxiliary arrays separated from the main array by distances $L_1, L_2, \ldots, L_N$, are these are used to cancel up to N mainbeam jammers. All auxiliaries are assumed to be identical, each consisting of P subarrays with the time delays $\tau, {}'_1, \tau'_2, \ldots, \tau'_p$ used to steer them to $\theta = \theta_0$. Behind each auxiliary is a bulk delay $T_n = L_n \sin \theta_0 / c$, where c is the speed of light and an adaptive tapped delay line with $K + 1$ taps with corresponding (adaptive) weights $w_0, w_1, \ldots, w_K$. These FIR filters are designed to maintain nulls on each jammer over the entire operating band B. Previous studies[2–4] have shown that the intertap delay T should be approximately ϵ/B, where $0.2 \leq \epsilon \leq 0.8$. The required number K of delays per auxiliary is determined by the maximum jammer delay that is likely to be encountered. Suppose jammers can lie anywhere within the 3-dB beamwidth $\Delta\theta_3$ of the mainbeam. Then the maximum delay across the array is $(L_N/c)\sin(\theta_0 \pm \Delta\theta_3/2)$. The bulk delay $T_N = (L_N/c) \sin \theta_0$, shown in Figure 10.63, partially compensates for this, but the difference must be compensated by the tapped delay line. This differential delay is

$$\Delta\tau = \frac{L_N}{c}\left[\sin\left(\theta_0 \pm \frac{\Delta\theta_3}{2}\right) - \sin\theta_0\right] \approx \pm\frac{L_N \Delta\theta_3}{2c}\cos\theta_0 \qquad (10.150)$$

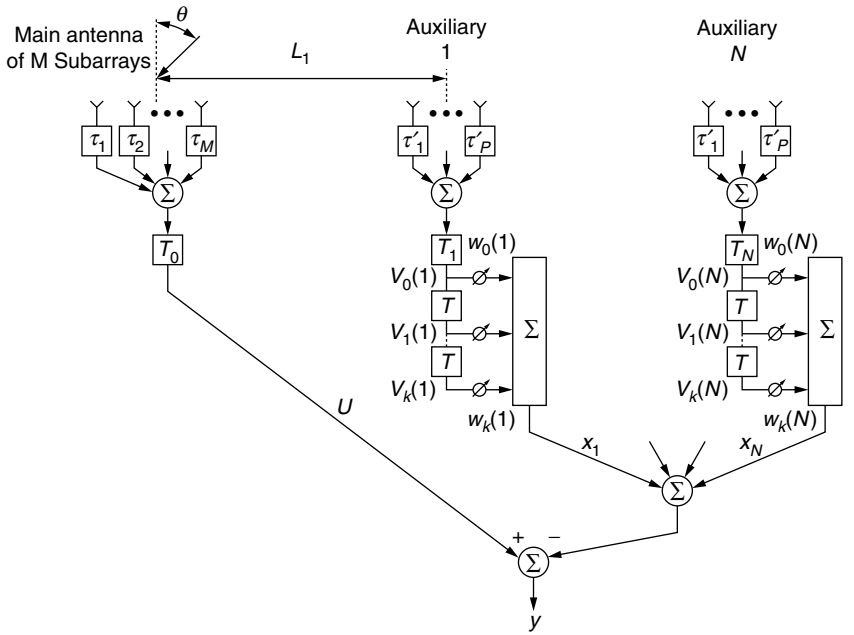

FIGURE 10.63 Main beam plus N auxiliary antennas. The delay $T_0 = KT/2$ and the delays T_n are $(L_n/c)\sin\theta_0$ for $n = 1, 2, \ldots, N$.

If there are K delays per tapped delay line and half are used to compensate for the positive differential delays and half the negative ones, then it is evident that we must set $KT/2 = \Delta\tau$. This gives

$$K = \frac{L_N \Delta\theta_3 \cos\theta_0}{cT} = \frac{L_N B \Delta\theta_3 \cos\theta_0}{\epsilon c} \tag{10.151}$$

In the next section, we will theoretically model the performance of the system shown in Figure 10.63, for the case when there are N_J broad-band noise jammers present in the mainbeam. We will then present the results of a set of simulations that will not only illustrate the performance improvement possible, but also provide answers on how can chooses the gain per auxiliary, the spatial span of the auxiliaries, etc.

10.5.2. Calculation of the Array Performance

Let us first calculate the residual interference after adaptation. We assume that there are N_J broad-band noise jammers located in the Fraunhofer zone (i.e., far field) of the array at azimuths $\theta_1, \theta_2, \ldots, \theta_{NJ}$, that the voltage gain of the main array is $g_M(\theta, f)$, and that all the auxiliary arrays are identical and have a voltage gain $g_A(\theta, f)$ where f is the frequency. Then, upon referring to Figure 10.63, it is evident that the Fourier transform $\tilde{U}(f)$ of the main channel voltage $U(t)$

produced by the N_J jammers is

$$\tilde{U}(f) = \sum_{q=1}^{N_J} \tilde{J}_q(f) g_M(\theta_q, f) \exp(-i2\pi f T_0) + \tilde{Z}(f) \tag{10.152}$$

where $\tilde{J}_q(f)$ and $\tilde{Z}(f)$ are the Fourier transforms of the qth jammer voltage and main antenna noises, respectively, and $T_0 = KT/2$. Likewise, the Fourier transform of the output $x_n(t)$ of the nth auxiliary antenna is

$$\tilde{X}_n(f) = \sum_{q=1}^{N_J} \tilde{J}_q(f) g_A(\theta_q, f) H_n(f) \exp(i2\pi f \tau_{nq}) + \tilde{Z}_n(f) H_n(f) \tag{10.153}$$

where $\tau_{nq} = L_n(\sin\theta_q - \sin\theta_0)/c$, L_n is the separation between the centers of the main array and the nth auxiliary, $\tilde{Z}_n(f)$ is the Fourier transform of the noise at the input to the adaptive FIR filter in the nth channel, and $H_n(f)$ is the frequency response of the FIR filter in the nth channel, defined as

$$H_n(f) = \frac{1}{\sqrt{K+1}} \sum_{k=0}^{K} w_k(n) \exp(-i2\pi k f T) \tag{10.154}$$

The weights $w_k(n)$ in the adaptive FIR filter are adjusted to minimize the output interference power[7–13] and vanish if the jammers vanish. Their calculation for either narrowband or wideband interference is presented in Appendix M. By combining Equation 10.152 and Equation 10.153, we find that the Fourier transform $\tilde{Y}(f)$ of the output voltage $y(t)$ in Figure 10.63 is given by

$$\tilde{Y}(f) = e^{-i2\pi f T_0} \sum_{q=1}^{N_J} J_q(f) \Gamma(\theta_q, f) + \tilde{Z}(f) - \frac{1}{\sqrt{N}} \sum_{n=1}^{N} H_n(f) \tilde{Z}_n(f) \tag{10.155}$$

where N is the number of auxiliary antennas

$$\Gamma(\theta_q, f) = g_M(\theta_q, f) - \frac{g_A(\theta_q, f)}{\sqrt{N}} \sum_{n=1}^{N} \tilde{H}_n(f) \exp(i2\pi f \tau_{nq}) \tag{10.156}$$

and

$$\tilde{H}_n(f) = H_n(f) \exp(i\pi f K T) \tag{10.157}$$

Now let us calculate the output power spectral density, defined as $\langle |\tilde{Y}(f)|^2 \rangle$, where $\langle \cdots \rangle$ denotes an expectation, or ensemble average. If we assume that the noises in different subarrays are independent, and that the voltages received from different jammers are independent, zero-mean, random variables, we find that the power spectrum of the residue after adaptation is

$$\langle |\tilde{Y}(f)|^2 \rangle = \sum_{q=1}^{N_J} S_{Jq}(f) |\Gamma(\theta_q, f)|^2 + S_z(f) \left[1 + \frac{1}{N} \sum_{n=1}^{N} |H_n(f)|^2 \right] \tag{10.158}$$

where $S_z(f)$ is the power spectrum of the noise and $S_{J_q}(f)$ is the power spectrum of the interference received from the qth jammer. The first term on the right-hand side of Equation 10.158 is the uncancelled jammer residue, and the second is the noise residue, sometimes called the "noise carryover."

The voltage $y(t)$ is typically passed through a matched filter that is designed to optimize the signal-to-interference ratio. To discuss this filter, suppose a target is present on radar boresight (i.e., $\theta = \theta_0$). Then the Fourier transform $\tilde{Y}_s(f)$ of the output signal voltage $y_s(t)$ produced by the target is given by

$$\tilde{Y}_s(f) = \mathrm{e}^{-\mathrm{i}2\pi f T_0} \Gamma(\theta_0, f)\tilde{S}(f) \tag{10.159}$$

where $\tilde{S}(f)$ is the Fourier transform of the received target return and $\Gamma(\theta_0, f)$ is given by Equation 10.156 with $\theta_q = \theta_0$. The first term on the right-hand side of $\Gamma(\theta_0, f)$ in Equation 10.156 corresponds to the true target response. The second term may produce spurious targets in different range bins if the bandwidth B is sufficiently large that $D(\sin\theta_q - \sin\theta_0) > c/2B$ where D is the total span of the auxiliary array and θ_q is the location of the qth jammer. A technique to reject the false targets is described in Appendix N. The matched filter that produces the maximum signal-to-interference ratio at time $t = 0$ is

$$\hat{H}_{\mathrm{opt}}(f) = \frac{\tilde{Y}_s^*(f)}{\langle |\tilde{Y}(f)|^2 \rangle} \tag{10.160}$$

and this gives a signal-to-interference ratio

$$\left(\frac{S}{I}\right)_{\mathrm{opt}} = \int_{-\infty}^{\infty} \mathrm{d}f \, \frac{|\tilde{Y}_s(f)|^2}{\langle |\tilde{Y}(f)|^2 \rangle} \tag{10.161}$$

For convenience, we have included both the noise and the jammer interference in the interference term.

In trying to synthesize the optimum matched filter, one needs to know the output-power spectral density $\langle |\tilde{Y}(f)|^2 \rangle$ of the interference, but this is unknown because the locations $\theta_1, \theta_2, \ldots, \theta_Q$ of the jammers are unknown. However, as will be seen later, in the absence of errors the adaptive FIR filters $H_n(f)$ typically reduce the first term on the right-hand side of Equation 10.158 to a level that is well below the noise residue, which is given by the second term on the right-hand side of Equation 10.158. Thus it is acceptable to approximate $\hat{H}_{\mathrm{opt}}$ by

$$\hat{H}_{\mathrm{subopt}} \equiv \frac{\tilde{Y}_s^*(f)}{S_z(f)\left[1 + \dfrac{1}{N}\displaystyle\sum_{n=1}^{N} |H_n(f)|^2\right]} \tag{10.162}$$

This approximation typically gave results for (S/I) that were within less than 0.1 dB of the results computed using the exact expression for $\hat{H}_{\mathrm{opt}}$ given in Equation 10.160. Thus, the results for S/I that will be presented in the next section

are valid for either $\hat{H}_{\text{opt}}$ or $\hat{H}_{\text{subopt}}$. However, when errors are included, the results using $\hat{H}_{\text{opt}}$ and $\hat{H}_{\text{subopt}}$ will be different.

In all calculations to follow it will be assumed that the power spectral densities of both the jammers and noise are white and, therefore, constant over the radar bandwidth B.

10.5.3. Simulation Results

There are a number of significant questions we must now answer. For a given jammer scenario these are:

(1) Where should the auxiliaries be placed and what is the required spatial span of the auxiliary array?
(2) How many auxiliaries should be used and what is the minimum acceptable gain per auxiliary?
(3) What is the minimum target-jammer separation for which the signal-to-interference ratio is acceptable?
(4) With a well-designed system, how much signal-to-interference improvement is possible?

Let us now answer each of these questions.

10.5.3.1. Spatial Span and Location of the Auxiliaries

If there are multiple jammers present and the minimum expected angular separation between jammers is $\Delta\theta$, then one would expect that the spatial span D of the auxiliary array must be of order $C\lambda_0/\Delta\theta$, where λ_0 is the midband wavelength and C is a constant of order unity. To verify this, we considered a number of configurations in which the main antenna and N auxiliary antennas (N was varied between 3 and 12) were located along a straight line. We then studied the cases when all the auxiliaries were located on the same side of the main antenna (as in Figure 10.63) and on both sides, so that the spatial centroid of the auxiliaries lies fairly close to the location of the main antenna. Computer experiments were run for the case when the main antenna and all the auxiliaries had a 30 dB Taylor taper and we found that for up to $N - 1$ jammers present, the cancellation performance was best if the auxiliaries were placed such that their spatial centroid was relatively close (within a few main antenna diameters) to the main antenna array, and the total spatial span D of the auxiliary antennas satisfies the condition

$$\frac{D}{D_M} \geq 0.4 \frac{\Delta\theta_3}{\Delta\theta} \tag{10.163}$$

where $\Delta\theta_3$ is the 3 dB beamwidth of the main antenna and D_M is the diameter of the main antenna. For the Taylor taper used, $\Delta\theta_3 = 1.12\ \lambda_0/D_M$ (at broadside) so

that we can rewrite Equation 10.163 as

$$D > 0.45 \frac{\lambda_0}{\Delta\theta} \tag{10.164}$$

Therefore, the coefficient $C = 0.45$.

If the interjammer spacing is slightly less than $\Delta\theta$, and $N_J < N$ the canceller will not be able to resolve the jammers and the residue will increase. If, however, the interjammer spacing is much less than $\Delta\theta$, adjacent jammers will begin to look like one jammer and the residue will decrease. In results (not shown), it was observed that the interjammer spacing, at which two jammers looked like one jammer, was greater when the phase centers of the main and auxiliary arrays were close, which is another reason for making the spatial centroids of the two arrays close together.

10.5.3.2. Required Number of Auxiliaries and Gain per Auxiliaries

To decide how many auxiliaries are necessary and the minimum gain per auxiliary that is required when N_J jammers are present in the mainbeam, we have considered a geometry where up to 12 auxiliaries of arbitrary gain G_A per auxiliary are located, relative to the main array, along a line at the distances L_1, L_2, L_3, ... given in Table 10.5.

(In practice, to compensate for jammers distributed in elevation, the auxiliaries would not be located along a line but rather on a circular or semicircular arc so that the distances L_1, L_2, ... would actually be projections of the actual locations onto a line normal to the scan direction. The latter observation is one of the reasons why we located more auxiliaries near the ends of the line than the center.) We will consider N_J mainbeam jammers at azimuths

TABLE 10.5
Normalized Auxiliary Locations Relative to Main Array

Auxiliary Number i	L_i/D_M
1	43.57
2	−44.07
3	22.62
4	−20.82
5	32.94
6	−32.33
7	10.87
8	−10.41
9	38.00
10	−38.49
11	29.18
12	−34.70

randomly located within an azimuthal range $\Delta\theta_J$. The value of $\Delta\theta$ is chosen so that the average angular jammer separation $\Delta\theta_J/N_J$ is at least several times larger than the resolution $\Delta\theta$ which, from Equation 10.163, is given by $\Delta\theta/\Delta\theta_3 = 0.4D_M/D$.

Before we present results for the residue after adaptation, let us discuss what we expect. In Appendix O, we show that as long as the number of auxiliaries is greater than or equal to the number of jammers and the jammers are well separated, we expect to cancel nearly all of the jammer interferences so that virtually the entire residue is the noise residue given by

$$\langle |y|^2 \rangle \approx \sigma^2 \left[1 + \frac{\beta G_M N_J}{N G_A} \right] \tag{10.165}$$

where β is determined empirically and is given by

$$\beta = 1 + \frac{N_J - 1}{N - (N_J - 1)} \tag{10.166}$$

σ^2 is the variance of the main-channel noise, G_M is the main-array gain, G_A is the gain of each auxiliary, N_J is the number of jammers, and N is the number of auxiliaries. The quantity NG_A is the total auxiliary gain. Because we expect the adapted residue to depend on $G_M N_J/NG_A$ we have plotted the numerical results as a function of this quantity. In Figures 10.64 and 10.65, we show the normalized narrowband and wideband residues $\langle |y|^2 \rangle/\sigma^2$ obtained by integrating

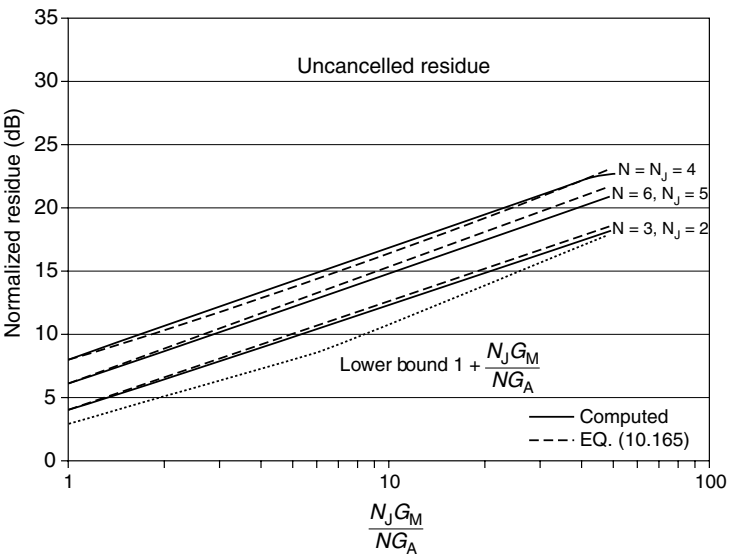

FIGURE 10.64 Residue after adaptation averaged over 25 Monte Carlos of jammer locations within $\Delta\theta_J = \Delta\theta_3/10$ when $B = 0.001\ f_0$, $\theta_0 = 20°$, and auxiliary locations in Table 10.5.

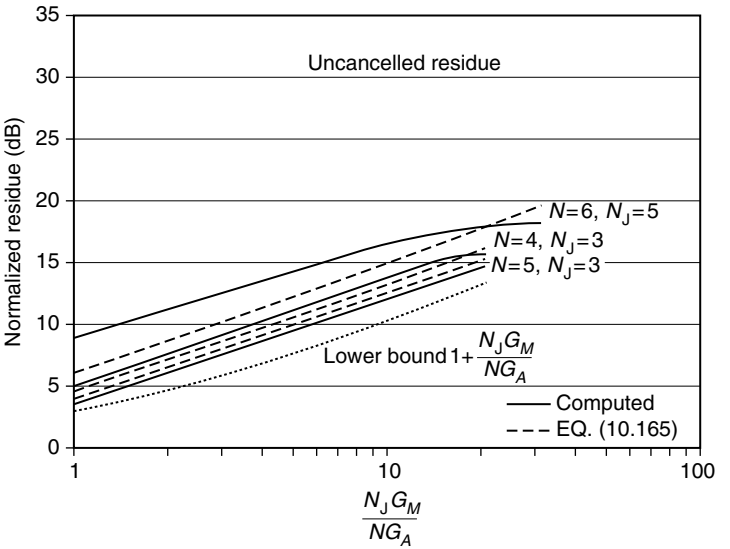

FIGURE 10.65 Residue after adaptation averaged over 25 Monte Carlos of jammer locations within $\Delta\theta_J = \Delta\theta_3/10$ when $B = 0.1\ f_0$, $\theta_0 = 20°$, and auxiliary locations in Table 10.5.

Equation 10.158 over frequency for differing numbers of auxiliaries and jammers. Each of the curves presented is the average of 25 Monte Carlos of the jammer locations within an angular sector $\Delta\theta_J = \Delta\theta_3/10$. In these simulations the array is scanned to $\theta_0 = 20°$ and, for the wideband case ($B = 0.1\, f_0$), we used 40 taps per auxiliary with an intertap delay $T = 0.3/B$.

The theoretical prediction of Equation 10.165 is also shown on these figures. In the narrowband limit we have found that the agreement with the simulated results is excellent for all values of $N \geq N_J$ except where $N = N_J = 2$ (this case will be discussed later). In the wideband limit, the agreement is always excellent for $N \geq N_J + 2$; but for large N_J the agreement degrades somewhat for $N = N_J + 1$ and $N = N_J$. This occurs because in the wideband case the adapted jammer residue is no longer negligible (as was assumed in the derivation of Equation 10.165) unless $N \geq N_J + 2$. That is, in the wideband limit we require $N \geq N_J + 2$ to ensure that the jammer residue is negligible and all that remains is noise carryover.

It should be noted that we have not shown the results for $N = N_J = 2$, because this is a pathological case. When $N = 2$, the array has grating lobes (that are not suppressed as they are for larger N), and the residue becomes quite large whenever both jammers are in grating lobes. For this case the residue was much larger than predicted by Equation 10.165. Because of this grating lobe phenomenon, two auxiliaries should never be used.

10.5.3.3. Signal-to-Interference Ratio after Cancellation

Thus far, we have completely ignored what happens to the desired target signal. We now wish to calculate the actual signal-to-interference ratio Equation 10.161 after adaptation and matched filtering by the filters defined in Equation 10.160. Although the optimum filter defined in Equation 10.160 was used, we should note that using the filter defined in Equation 10.162 gave results for *S/I* within 0.01 dB of the optimum.

Let us consider an antenna system consisting of a main array with 264 columns and three auxiliary arrays of 88 columns each so that the total auxiliary gain NG_A equals the main antenna gain G_M with the auxiliaries located at position 1, 2, and 3 in Table 10.5, relative to the main array. We further assume that the array is scanned to $\theta_0 = 10°$, that a target is located on antenna boresight (i.e., at $\theta = \theta_0$), and that there are two mainbeam jammers present. In performing the calculations to follow, it was assumed that the magnitude squared $|S(f)|^2$ of the signal Fourier transform is constant for $f_0 - B/2$ to $f_0 + B/2$ and zero outside this band. In Figures 10.66 and 10.67, we show the narrowband and wideband signal-to-interference ratios after adaptation and matched filtering. For these cases, the azimuth of jammer one is held constant at 0.09 beamwidths off boresight, and the azimuth of jammer two is varied. We see that there are some fades in *S/I*, but these are filled in as the bandwidth is increased.

Figure 10.68 shows typical wideband antenna patterns before and after cancellation for the case of three auxiliaries, each one third the length of the main antenna, operating against two jammers. Although the peak of the adapted pattern is nearly 7 dB above the peak of the quiescent pattern, the difference would have been about 3 dB if we had normalized the adapted pattern by the total noise residue. Note that the increase in the beamwidth of the adapted beam

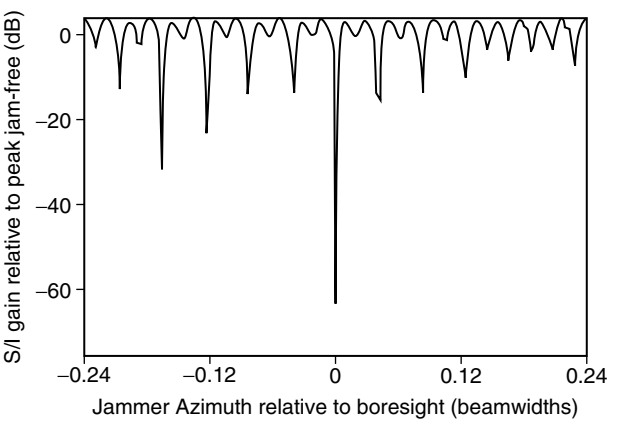

FIGURE 10.66 Signal-to-interference (relative to peak jam-free) when jammer 1 fixed at 0.04 deg off boresight and Jammer 2 Azimuth varied. Three auxiliary antennas: $B = 0.001\, f_0$, $\theta_0 = 10°$, $NG_A = G_M$.

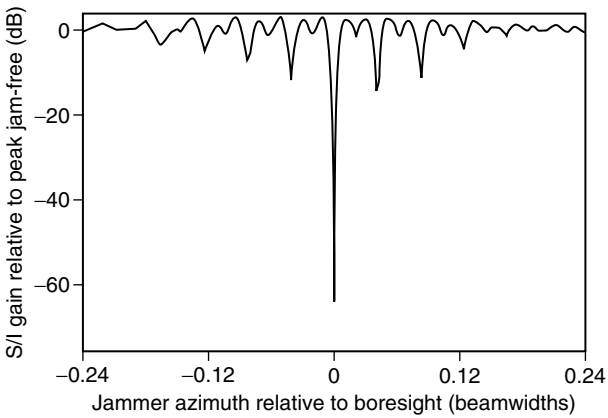

FIGURE 10.67 Signal-to-interference (relative to peak jam-free) when jammer 1 fixed at 0.04 deg off boresight and jammer 2 azimuth varied. Three auxiliary antennas: $B = 0.1\,f_0$, $\theta_0 = 10°$, $NG_A = G_M$.

occurs because each auxiliary is only one third the length of the main antenna. This broadening of the receive beam could be eliminated by using auxiliaries that are the same size as the main array. The auxiliaries need only be receive only and, in some cases, could be reflectors.

It is interesting to note that even in the narrowband case the offboresight fades are relatively narrow. By studying these results, it is easy to show that the probability of a jammer located randomly within the 3 dB beamwidth causing a fade in (S/I) greater than 10 dB is only 0.031. For the wideband case, this drops to

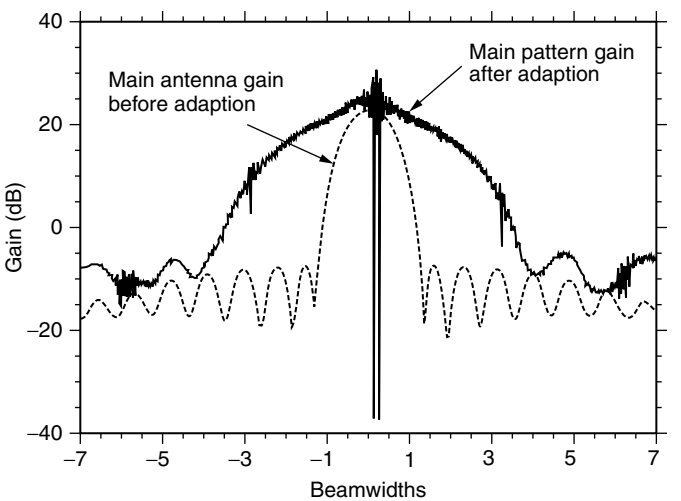

FIGURE 10.68 Pattern gains before and after cancellation for two jammers and three auxiliary antennas: $B = 0.1\,f_0$, $NG_A = G_M$, jammer 1 at 0.04°, jammer 2 at 0.1°.

0.018. Results were also generated (not shown) using six auxiliaries with 44 columns each so that the total auxiliary gain NG_A was still equal to the main array gain. When we increased the number of auxiliaries to six, there were virtually no fades in S/I that exceed 10 dB, except very near boresight.

10.5.3.4. Simultaneous Nulling of Mainlobe and Sidelobe Jammers

An in-depth investigation of the problem of simultaneous nulling of mainlobe and sidelobe jammers was not performed under this effort, but we did perform a limited number of simulations. We had expected to have to add only N_{SL} low-gain auxiliaries to the main antenna to null N_{SL} narrowband sidelobe jammers. However, the bulk delay inserted in series with the high-gain auxiliaries (see Figure 10.63) to time-align the mainlobe jammers does not time-align sidelobe jammers sufficiently to prevent time-delay decorrelation. The problem can be solved by adding low-gain auxiliaries within or next to each high-gain auxiliary, and bulk delaying them with the same delay added to their companion high-gain auxiliaries. Sidelobe jammers received by each high-gain auxiliary will then be highly correlated with sidelobe jammers received in their companion low-gain auxiliaries. The latter approach enables both sidelobe and mainlobe jammers to be nulled by a single large canceller comprised of both multiple high- and low-gain auxiliaries. That is, N_J mainlobe and N_{SL} sidelobe jammers can be simultaneously cancelled by employing N_J high-gain auxiliaries with N_{SL} low-gain auxiliaries added to the main array and to each high-gain auxiliary.

A typical result is illustrated in Table 10.6, where we show the residue after adaptation for the case of one mainlobe jammer and either one or two sidelobe jammers. The main array had 288 elements and the high-gain auxiliary had 88 elements and was located 103.51 m from the main array. Each low-gain auxiliary consisted of one element. The array was scanned to 10°, and the mainlobe jammer was located at 10.04° and had a jammer-to-noise ratio of 32 dB in the main array. Bulk delays of (L_n/c) sin 10° were inserted in series with the outputs of each high-

TABLE 10.6
Cancellation Performance Against One Mainlobe Jammer and One and Two Sidelobe Jammers: Cancellation Bandwidth = 10 MHz. Residue Before Adaptation = 32.4 dB

# ML Jammers	# SL Jammers	# High-Gain Auxs	# Low-Gain Aux at Main	# Low-Gain Aux at High-Gain Aux	Residue-to-Noise Ratio (dB)
1	1	1	1	0	27.7
1	1	1	1	1	6.3
1	2	1	1	1	22.1
1	2	1	2	1	20.5
1	2	1	2	2	7.5

gain auxiliary and its companion low-gain auxiliaries. Sidelobe jammer one was located at 17.5° and had a 22 dB jammer-to-noise ratio in the main array and 27 dB in the high-gain auxiliary. Sidelobe jammer two was located at 23.5° and had a 22 dB jammer-to-noise ratio in the main array and 24 dB in the high-gain auxiliary. The cancellation bandwidth was chosen as 10 MHz and no tapped delay lines were used. From Table 10.6 we see that if the main and high-gain auxiliary each have N_{SL} low-gain auxiliaries the cancellation performance approaches the theoretical performance of 6 dB residue as predicted by Equation 10.165. (The residue would have only been 3 dB had we chosen the auxiliaries to be the same size as the main array. For smaller bandwidths the residues are even closer to the theoretical value.) However, unless we have at least N_{SL} low-gain auxiliaries in or very close to the main array and each high-gain auxiliary, there is a severe performance degradation. Thus, the total number of spatial degrees of freedom required is

$$N_{DOF} = N_J + (N_J + 1)N_{SL} \tag{10.167}$$

where N_J = number of mainbeam jammers and N_{SL} = number of sidelobe jammers. That is, we need N_J high-gain auxiliaries and $(1 + N_J)\, N_{SL}$ low-gain auxiliaries.

It should be noted that we have not discussed the special case when one or more jammers lies in the sidelobes of the main array but within the mainbeam of the high-gain auxiliary. In this case, the performance degrades but the degradation is gradual (about a 6 dB increase in the residue after adaptation). This degradation can be avoided by choosing the gain of each auxiliary equal to that of the main antenna.

10.5.4. Summary and Discussion

We have demonstrated that wideband or narrowband mainbeam jammers can be cancelled using multiple auxiliary antennas each with an adaptive tapped delay line. In the absence of errors, the jammers residue can be cancelled down to below the system noise. The residue after nulling is due mainly to thermal noise carried over from the auxiliaries (Equation 10.165). For a given jammer scenario, we have provided design formulas for the required spatial span of the auxiliaries (Equation 10.163), the number of taps per auxiliary (Equation 10.151), and the gain G_A per auxiliary (Equation 10.165). Although the high-gain auxiliaries will be costly, they can be receive-only and, in some cases, could be reflectors.

The problem of simultaneous nulling of mainlobe and sidelobe jammers was not the primary focus of this effort. However, it was noted that the bulk delays used to time align the main and high-gain auxiliaries in the narrowband case will not always be sufficient to time-align sidelobe jammers. One solutions is to add multiple low-gain auxiliaries to the main and high-gain auxiliaries to null multiple sidelobe jammers.

Finally, we have also demonstrated (results not shown) that the difference beams in each of the auxiliaries can be used to cancel jammers in the difference beams in the main channel while preserving to first order the monopulse slope of the main array. Jammer cancellation performance and parameter sensitivities are comparable to that obtained using the sum beams.

10.6. ADAPTIVE SPACE–TIME RADAR

(R. L. FANTE)

10.6.1. Introduction

Unlike a ground-based radar, in which all of the clutter is received at or near zero-Doppler, the clutter return in an airborne radar has Doppler frequencies spread over a band of width $4V/\lambda$, where V is the platform speed and λ is the radar wavelength. Therefore, conventional moving target indicator (MTI) is ineffective in canceling this clutter without also canceling desired targets. Another feature of airborne radar is that there is usually clutter in the same Doppler cell as a target, but it usually arrives from a different azimuth, as illustrated in Figure 10.69 for the case of frozen clutter (i.e. no internal motion). This suggests that adding spatial degrees of freedom to the conventionally-used temporal degrees of freedom should allow us to place a null along the azimuth of the clutter that competes with the target, thus allowing the clutter to be canceled, without canceling the desired target return. In constructing Figure 10.69 we assumed that the pulse repetition rate is such that there are no Doppler ambiguities. A typical space–time processor that combines spatial and temporal degrees of freedom is shown in Figure 10.70. The processor is formed by placing a tapped delay line at the output of each of N antennas of the array, with the taps spaced by one pulse repetition interval (PRI) T. This processor

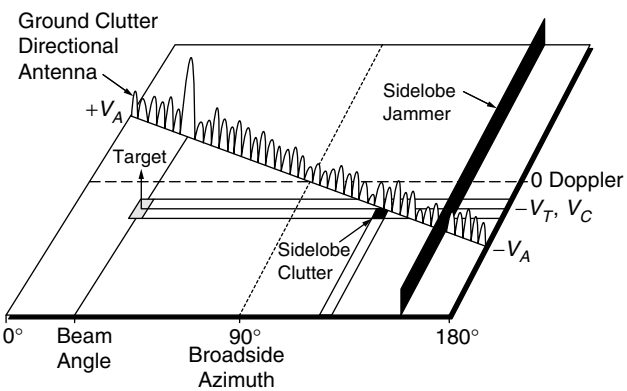

FIGURE 10.69 Azimuth–Doppler plot of clutter and one sidelobe jammer interference for airborne platform.

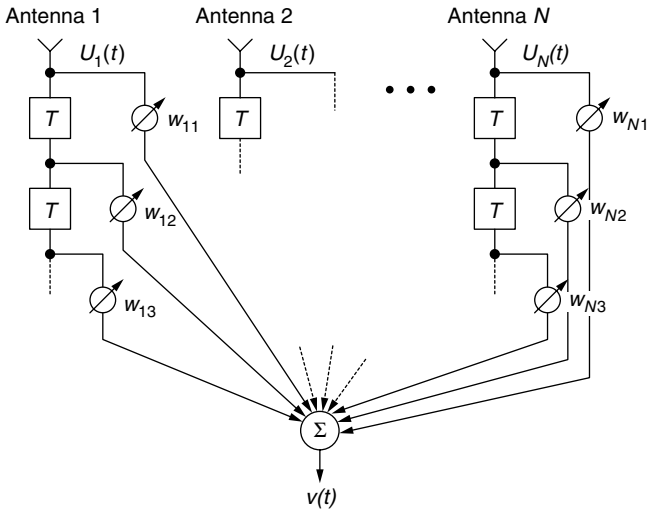

FIGURE 10.70 Generic space–time array.

has a number of degrees of freedom equal to the number N of antennas times the number K of pulses (taps).

The adaptive weights in the processor shown in Figure 10.70 are chosen to maximize the signal-to-noise-plus-clutter-plus-jammer ratio $S/(N + C + J)$ for a specified desired signal (i.e. specified azimuth and Doppler). That is, the adaptive space–time array is a matched filter designed to detect a desired radar signal while rejecting all interference and signals that do not have the properties of the desired signal. The quantity $S/(N + C + J)$ at the output of the array in Figure 10.70 is

$$\frac{S}{N + C + J} = \frac{|\mathbf{w}^{\mathrm{T}}\mathbf{s}|^2}{\langle |\mathbf{w}^{\mathrm{T}}\mathbf{U}|^2 \rangle} \tag{10.168}$$

and this is maximized[1] if the weight vector $\mathbf{w} = [w_{11}, w_{12}, \ldots, w_{1N}, w_{21}, \ldots, w_{2N}, \ldots]^{\mathrm{T}}$ is given by

$$\mathbf{w} = \mu \mathbf{M}^{-1}\mathbf{s}^* \tag{10.169}$$

where μ is a constant, $\mathbf{M}$ is the covariance matrix defined as $\mathbf{M} = \langle \mathbf{U}\mathbf{U}^{+} \rangle$, $\langle\rangle$ denotes an expectation, $\mathbf{U}$ is the received voltage vector given by

$$\mathbf{U} = [U_1(t) \ldots U_N(t),\ U_1(t - T) \ldots U_N(t - T), \ldots]^{\mathrm{T}}$$

and s is the steering vector describing the voltage that would be received on each element of the array if a target at the desired azimuth and with the desired Doppler were present. In the above equations the superscript T denotes a transpose, '*' denotes complex conjugate and + denotes conjugate transpose. When w is given by Equation 10.168 it can be shown by substituting Equation 10.169 into

Equation 10.168 that $S/(N+C+J)$ after adaptation is given by

$$\frac{S}{N+C+J} = \mathbf{s}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{s}^{*} \tag{10.170}$$

In an ideal world where the covariance could be exactly estimated, there is no internal clutter motion, channel match is perfect, the jammers are narrow-band, and there is no jammer multipath, it can be shown that

$$\frac{S}{N+C+J} \rightarrow \frac{|\mathbf{s}|^2}{\sigma^2} \tag{10.171}$$

where $|\mathbf{s}|^2$ is the signal power and σ^2 is the thermal noise power. Thus, under ideal conditions the adaptive processor can cancel all clutter and interference, and all that remains after adaptation is thermal noise.

Although the architecture shown in Figure 10.70 represents an element space configuration, a beam space architecture is equivalent. For example, if a beamformer is placed behind the elements in Figure 10.70, the $N \times 1$ voltage vector $\mathbf{b}$ at the beamformer output can be written as $\mathbf{b} = \mathbf{GU}$, where $\mathbf{G}$ is the $N \times N$ beamformer matrix. Then, if we note that the new steering vector after beamforming is $\mathbf{s}_{\mathrm{b}}^{\mathrm{T}} = \mathbf{GS}$ and that

$$\langle \mathbf{bb}^{+} \rangle = \mathbf{G}\langle \mathbf{UU}^{+} \rangle \mathbf{G}^{+} = \mathbf{GMG}^{+}$$

we can readily show that

$$\left(\frac{S}{N+C+J} \right)_{\text{beam space}} = \mathbf{s}_{b}^{\mathrm{T}} \langle \mathbf{bb}^{+} \rangle^{-1} \mathbf{s}_{\mathrm{b}}^{*} = \mathbf{s}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{s}^{*} \tag{10.172}$$

so that upon comparing Equation 10.172 with Equation 10.173 we see the output signal-to-interference ratio is the same whether an element space or beamspace architecture is used. However, sometimes beamspace may be preferable to element space, because it allows for the development of suboptimum architectures more readily, as will be discussed later.

10.6.2. Understanding the Results in Equation 10.169 and Equation 10.170

In practice one must estimate the covariance matrix using a finite[§] number of samples, but for purposes of understanding we will assume that the number of samples is so large that the ideal covariance matrix is available. For the architecture in Figure 10.70 this is an $NK \times NK$ square matrix $\mathbf{M}$ that can be

[§] For an acceptable covariance estimate the number of samples must be at least $2NK$ and preferably $4NK$, where NK is the total number of adaptive degrees of freedom.

decomposed into its $NK \times 1$ eigenvectors $\mathbf{e}_k$ and eigenvalues λ_k as[2]

$$M = \sum_{k=1}^{NK} \lambda_k \mathbf{e}_k \mathbf{e}_k^+ \tag{10.173}$$

Likewise

$$M^{-1} = \sum_{k=1}^{NK} \frac{\mathbf{e}_k \mathbf{e}_k^+}{\lambda_k} \tag{10.174}$$

Because $\mathbf{M}$ is Hermitian the eigenvalues are real and the eigenvectors are orthogonal. Under ideal conditions (no internal clutter motion, narrowband radar, etc.) there are $N + K$ large eigenvalues associated with the mainbeam and sidelobe clutter [for ambiguous radars there are even more.[3,4] A typical eigenvalue strength plot for clutter only (no jammers) is shown in Figure 10.71.

Strong jammers produce additional large eigenvalues. If there are N_J independent, resolved (i.e., spaced in angle by greater than the 3 dB beamwidth of the antenna) jammers, there are N_J large jammer eigenvalues. However, unresolved jammers and correlated jammers share a single eigenvalue, as indicated in Figures 10.72 and 10.73. Thus, the covariance matrix $\mathbf{M}$ typically has $R = N + K + N_J$ large eigenvalues and $NK - R$ small eigenvalues with magnitudes equal to the variance σ^2 of the system noise. Consequently, Equation 10.176 can be rewritten as

$$\mathbf{M}^{-1} \simeq \sum_{k=1}^{R} \frac{\mathbf{e}_k \mathbf{e}_k^+}{\lambda_k} + \frac{1}{\sigma^2} \sum_{k=R+1}^{NK} \mathbf{e}_k \mathbf{e}_k^+ \tag{10.175}$$

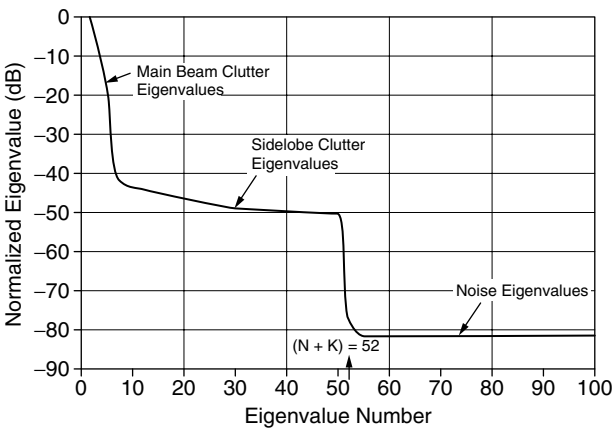

FIGURE 10.71 Clutter eigenvalue spread for $N = 50$, $K = 2$ (no jammers).

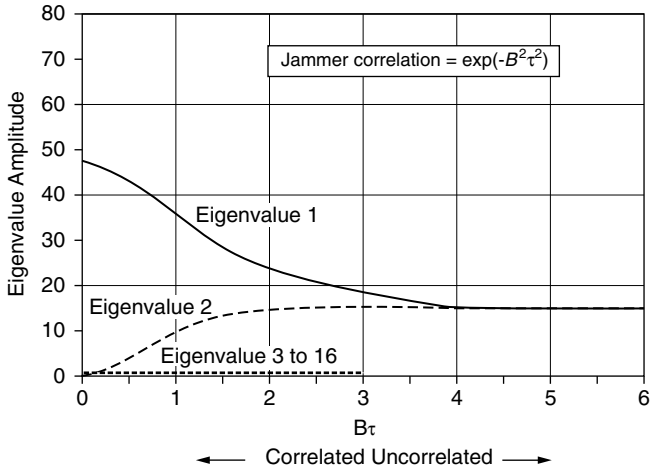

FIGURE 10.72 Effect of jammer correlation on eigenvalues for two resolved jammers when $N = 16$, $K = 1$.

Next, we recognize that the $NK \times NK$ identity matrix $\mathbf{I}$ is defined as[2]

$$I = \sum_{k=1}^{NK} \mathbf{e}_k \mathbf{e}_k^+ = \sum_{k=1}^{R} \mathbf{e}_k \mathbf{e}_k^+ + \sum_{k=R+1}^{NK} \mathbf{e}_k \mathbf{e}_k^+ \tag{10.176}$$

where the first term on the right-hand side of Equation 10.176 defines the interference (clutter plus jammers) subspace and the second defines the noise subspace. Thus, we can rewrite Equation 10.176 as a sum of two projections

$$\mathrm{I} = P_{11} + P_{\perp} \tag{10.177}$$

where P_{11} is the projection into the interference subspace and $P_{\perp}$ is the projection orthogonal to the interference subspace. Therefore, Equation 10.175 can be

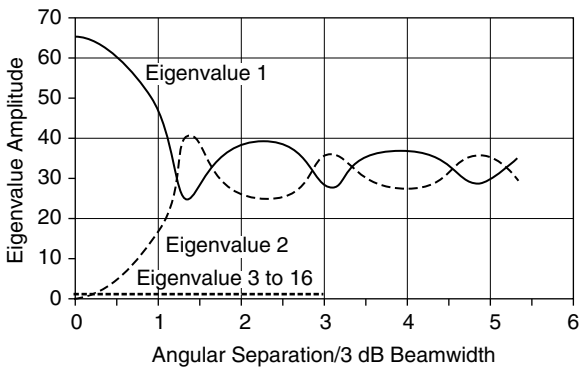

FIGURE 10.73 Effect of jammer separation θ_{sep} on eigenvalues for $N = 16$, $K = 1$, $B\tau = 2$.

rewritten as

$$\mathbf{M}^{-1} = \sum_{k=1}^{R} \frac{\mathbf{e}_k \mathbf{e}_k^+}{\lambda_k} + \frac{1}{\sigma^2} P_\perp \tag{10.178}$$

If Equation 10.178 is now used in Equation 10.169 we see that the weight vector w is

$$w = \mu \sum_{k=1}^{R} \frac{\mathbf{e}_k \mathbf{e}_k^+ \mathbf{s}^*}{\lambda_k} + \frac{\mu}{\sigma^2} P_\perp \mathbf{s}^* \tag{10.179}$$

Because, as can be seen from Figure 10.71, $|\lambda_k| \gg \sigma^2$ we can neglect the first term on the right-hand side of the Equation 10.179 so that

$$\mathbf{w} \simeq \frac{\mu}{\sigma^2} P_\perp \mathbf{s}^* \tag{10.180}$$

This shows that the desired weight vector is the projection of the steering vector s onto the subspace orthogonal to the interference subspace, and means that as long as the steering vector has a component orthogonal to the interference subspace it is possible to cancel the interference while preserving the desired signal (or at least most of it).

It should also be noted that because $|\lambda_k| \gg \sigma^2$ we can use Equation 10.181 to write Equation 10.175 as

$$\mathbf{M}^{-1} \simeq \frac{1}{\sigma^2} \sum_{k=1}^{NK} \mathbf{e}_k \mathbf{e}_k^+ - \frac{1}{\sigma^2} \sum_{k=1}^{R} \mathbf{e}_k \mathbf{e}_k^+ = \frac{1}{\sigma^2} \left(I - \sum_{k=1}^{R} \mathbf{e}_k \mathbf{e}_k^+ \right) \tag{10.181}$$

This is a useful approach for inverting the covariance matrix when there are only a few dominant eigenvalues, and has been exploited in Ref. 6.

By using Equation 10.181 in Equation 10.170 it is easy to see how well the limit in Equation 10.171 is approached.

10.6.3. Sequential Cancellation of Jammers and Clutter

In practice, when both jammers and clutter are simultaneously present it is difficult to cancel both because of interactions in the covariance matrix estimate. Thus, it is desirable to cancel the jammers before tackling the clutter. One way to do this is to use the beam space architecture, place spatial nulls in each beam on any jammers in that beam, and then combine those resulting beams to cancel the ground clutter. Let us analyze here how one forms $L \leq N$ beams with nulls on any jammers present. In particular, we need to derive an $L \times N$ beamforming matrix $\mathbf{\Gamma}$ that produces L beams while simultaneously minimizing jammer power. Let $\mathbf{J}$ be an $N \times 1$ vector consisting of the voltages on the N antenna elements produced by any jammers present. These jammers then produce a voltage vector at beam ports given by $\mathbf{b} = \mathbf{\Gamma J}$. Also, let h_1 be an $N \times 1$ vector of the voltages produced on the N antenna elements by an incoming plane wave at an angle θ_1 relative to the array. Likewise, define h_2 as

the $N \times 1$ vector after a wave at angle θ_2, etc. Then, the condition for forming L beams pointed at θ_1, θ_2, ... is

$$\Gamma_k h_k = C_k \tag{10.182}$$

for $k = 1, 2, \ldots L$, where $\mathbf{\Gamma}_k$ is the $1 \times N$ vector formed by the kth row of $\mathbf{\Gamma}$ and the C_k are specified constraints. Now we desire to minimize the interference power $\langle \mathbf{b}^+ \mathbf{b} \rangle$ subject to the constraints in Equation 10.182. The solution is

$$\mathbf{\Gamma} = \boldsymbol{\beta} \mathbf{D} \mathbf{G} \mathbf{R}^{-1} \tag{10.183}$$

where

$$\mathbf{R} = \langle JJ^+ \rangle \tag{10.184}$$

is the jammer plus noise covariance matrix, $\mathbf{G}$ is the $L \times N$ beamforming matrix in the absence of jammers, $\mathbf{D}$ is a diagonal matrix with the kth diagonal element given by $C_k/(h_k^+ \mathbf{R}^{-1} h_k)$ and β is a constant. Note that if the jammers are absent the jammer plus noise covariance matrix reduces to the noise covariance matrix $\sigma^2 \mathbf{I}_N$, where $\mathbf{I}_N$ is the $N \times N$ identify matrix, so that, as expected, the beamforming matrix Γ reduces to the jammer-free beamforming matrix $\mathbf{G}$.

One may next wonder how one estimates the jammer covariance matrix when both jammers and clutter are present. This can be done by referring to the angle-Doppler plot shown in Figure 10.69, and recognizing that the jammers spread through all Doppler bins. Thus, if one could choose an angle-Doppler region that is well removed from the peak of the clutter ridge it should be possible to estimate the jammer covariance with little clutter interaction. Of course, for radars that are unambiguous[¶] in Doppler this is a trivial task because there are then regions in Doppler space where the clutter is completely absent, but the jammers are present.

Once the jammers have been removed the clutter can then be cancelled either in the beam-pulse domain (also called pre-Doppler) as in Figure 10.74 or in the beam-Doppler domain (also called post-Doppler) as in Figure 10.75. The architecture in Figure 10.75 is what is known as first-order factored processing where the beams in each Doppler frequency bin are independently used to cancel clutter. That is, if $\mathbf{B}(p)$ is an $N \times 1$ vector of the outputs of the N beam ports for Doppler frequency bin p and $\mathbf{S}(p)$ is the $N \times 1$ steering vector[‖] for the Doppler frequency bin p, then the weight vector $\mathbf{W}_p$ used to combine the beam outputs in

[¶] Remember that the clutter only occupies the frequency region from $\pm 2V_s/\lambda$ so that if the pulse repetition frequency exceeds this value there is a clutter-free region in the Doppler domain.

[‖] $S(p)$ is simply the discrete Fourier transform of the time domain steering vector $\mathbf{s}$.

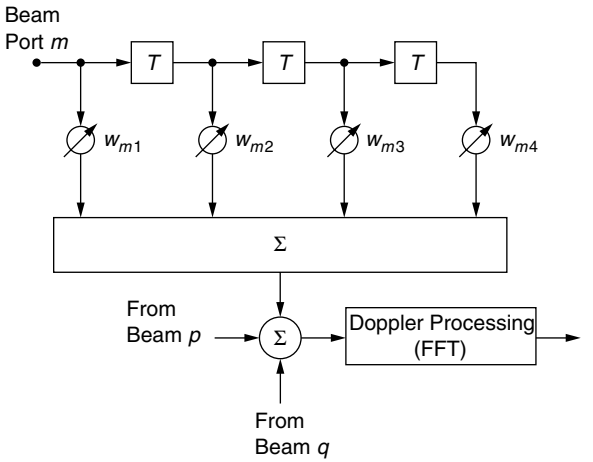

FIGURE 10.74 Beam-pulse space processor.

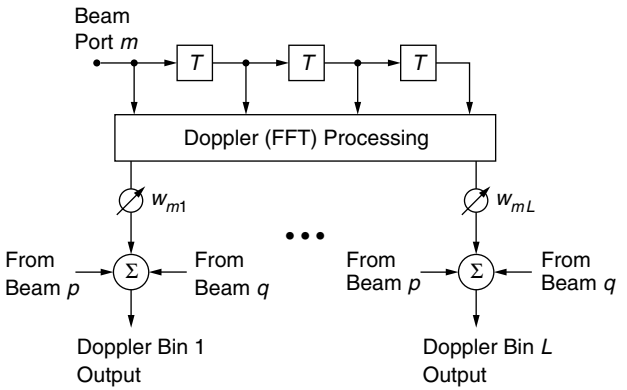

FIGURE 10.75 Beam-doppler space processor (first order factored).

Doppler frequency bin p is

$$\mathbf{W}_p = \mu\langle \mathbf{B}(p)\mathbf{B}^{+}(p)\rangle^{-1}\mathbf{S}^{*}(p) \tag{10.185}$$

where μ is a constant.

In practice, however, it has been found that combining two Doppler bins produces superior cancellation performance. The reader is referred to Ref. 5 for details.

10.6.4. Typical Results

There are several experimental programs to test the adaptive cancellation methods discussed in the last few sections. One is called Mountaintop (Lincoln

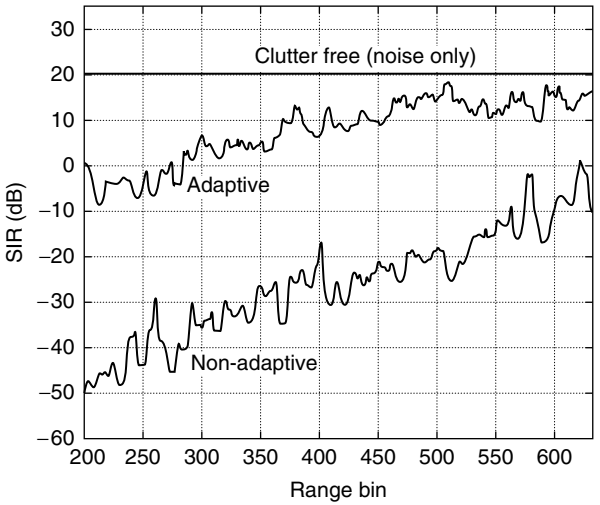

FIGURE 10.76 Improvement in signal to interference ratio for a particular doppler bin and receive azimuth using first order factored processing. The clutter free SNR = 20 dB.

Laboratory) and is designed to evaluate space–time processing using emulated motion of the radar platform. This program has demonstrated excellent cancellation at UHF of emulated airborne clutter at the White Sands Missile Range. Another is called MCARM (Northrup Grumman & Rome Laboratory) and consists of actual BAC-111 flights of a 1.25 GHz adaptive array with 22 spatial degrees of freedom. In Figure 10.76 we show the MCARM signal-to-clutter plus noise ratio (there were no jammers present) before and after adaptation when the processing architecture in Figure 10.75 is used.[7] Note that up to 40 dB of clutter cancellation has been achieved.

The clutter is not cancelled completely for a number of reasons including the fact that the first order factored processing (as in Figure 10.75) is suboptimal and that the clutter is spatially inhomogeneous, so the covariance matrix estimate is not accurate for all ranges.

Additional results are available, but will be omitted for conciseness.

10.6.5. Additional Considerations

In this brief exposition we have ignored many of the implementation problems involved in the actual performance of a space–time adaptive array. Some of these are:

(i) How does one obtain sufficient samples to estimate the covariance matrix and what are the effects of inhomogeneous clutter on its value?
(ii) What are the effects of channel mismatch, radar-aircraft crabbing and scatter from obstacles such as airplane wings that are in the near-field of the array?

(iii) Is the computational load reasonable?
(iv) How does jammer multipath affect the interference cancellation?

All of these points have been investigated to some degree. Jammer multipath is especially serious because the radiation of an airborne sidelobe jammer can be scattered off the ground into the mainbeam (or its elevation sidelobes) of the radar thus giving the appearance of multiple mainbeam jammers. Because the interference is spread in all azimuths across the main beam it cannot be canceled, without simultaneously canceling the target signal, using spatial nulls. Thus, time domain cancellation techniques are required,[8] but these are beyond the scope of this chapter.

10.6.6. Summary

We have given a brief exposition of how a space–time adaptive canceller can remove the deleterious effects of both ground clutter and jammers, thus enabling an airborne radar to detect weak targets. In particular we have shown how the processor accepts target signals, but rejects anything that does not look like a target, as specified by the target's Azimuth and Doppler frequency. Thus, in principle, all interference except system noise can be eliminated.

10.7. SYNTHESIS OF ADAPTIVE MONOPULSE PATTERNS

(R. L. FANTE)

10.7.1. Analysis

SPACE-TIME adaptive processing (STAP) is an effective method used by airborne radars for adaptively canceling clutter and jammers while simultaneously detecting targets. However, while it is straightforward to form adapted sum (Σ) and difference (Δ) beams, the adapted monopulse pattern Δ/Σ may have a highly distorted slope, rendering it ineffective for angular location.[1] In this letter, we present an approach to obtain controlled monopulse patterns for an adaptive radar. The classical STAP architecture is shown in Figure 10.77.[2,3] Our procedure is to first form the adapted sum beam using the classic weight vector[4]

$$\mathbf{w} = \frac{\mathbf{\Phi}^{-1}\mathbf{s}^*}{\mathbf{s}^{\mathrm{T}}\mathbf{\Phi}^{-1}\mathbf{s}^*} \tag{10.186}$$

where for K antennas and M time taps $\mathbf{w}^{\mathrm{T}} = [w_{11}\cdots w_{1M}\cdots w_{N1}\cdots w_{NM}]$, $\mathbf{\Phi}$ is the $NM \times NM$ interference covariance matrix and s is the steering vector for the signal defined as $\mathbf{s}^{\mathrm{T}} = [s_{11}\cdots s_{1M}\cdots s_{NM}]$. For a linear array and a target at azimuth θ_{o} and Doppler frequency f_{o}, the components of $\mathbf{s}$ are $s_{nm} = \exp[\mathrm{i}kx_n\sin\theta_{\mathrm{o}} - \mathrm{i}2\pi rnf_{\mathrm{o}}T]$, where x_n is the location of antenna n and k is the wave number.

The difference beam $\Delta(\theta, f_{\mathrm{o}})$ is now formed such that the received interference is minimized subject to the constraints that $\Delta(\theta_{\mathrm{o}}, f_{\mathrm{o}}) = 0$ and

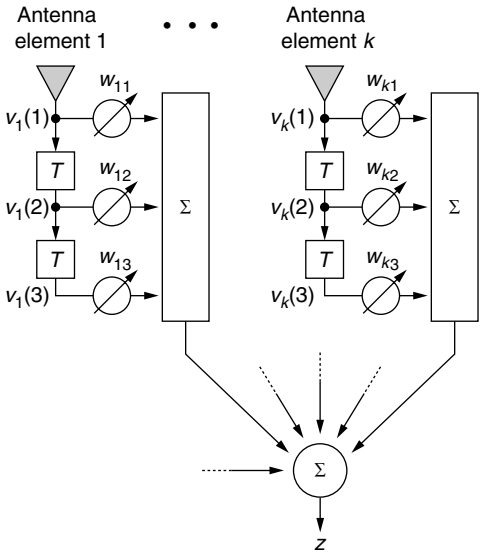

FIGURE 10.77 Space–time adaptive processor.

the ratio Δ/Σ maintains a constant slope at Doppler frequency f as specified by

$$\frac{\Delta(\theta_o \pm \Delta\theta, f_o)}{\sum(\theta_o \pm \Delta\theta, f_o)} = \pm k_s \Delta\theta \tag{10.187}$$

where k_s is a slope constant. If we define a difference-beam weight vector $\mathbf{w}_\Delta$ and recognize that the adapted difference pattern in the azimuth–Doppler domain (θ, f) is $\mathbf{w}_\Delta^T \mathbf{g} = \mathbf{g}^T \mathbf{w}_\Delta$, where $\mathbf{g}(\theta, f)$ is an $NM \times 1$ vector with components $g_{nm} = \exp(ikx_n \sin\theta - i2\pi mfT)$, we see that the above three constraints can be written in matrix notation as

$$\mathbf{H}^T \mathbf{w}_\Delta = \rho \tag{10.188}$$

where[#]

$$\mathbf{H}^T = \begin{bmatrix} \mathbf{g}^T(\theta_o + \Delta\theta, f_o) \\ \mathbf{g}^T(\theta_o, f_o) \\ \mathbf{g}^T(\theta_o - \Delta\theta, f_o) \end{bmatrix} \tag{10.189}$$

$$\boldsymbol{\rho} = k_s \begin{bmatrix} \mathbf{w}^T \mathbf{g}(\theta_o + \Delta\theta, f_o) \\ 0 \\ -\mathbf{w}^T \mathbf{g}(\theta_o - \Delta\theta, f_o) \end{bmatrix} \Delta\theta \tag{10.190}$$

[#] In order to ensure that no anomalies occur we actually used $|w^T g|$ in (Equation 10.190).

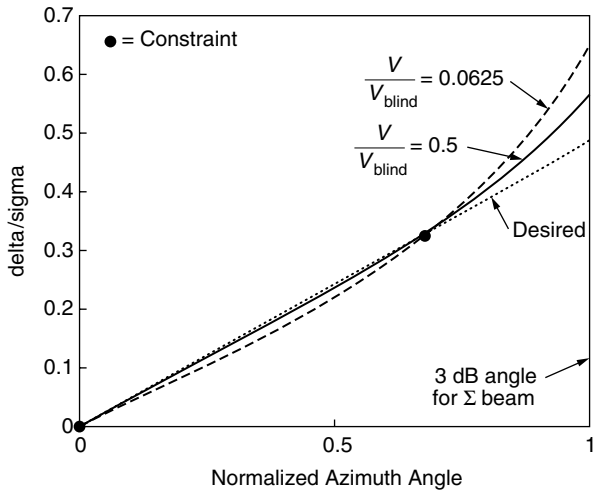

FIGURE 10.78 Positive angle portion of adapted monopulse pattern for three constraints.

The weight vector $\mathbf{w}_\Delta$ that minimizes the difference beam interference

$$\mathbf{w}_\Delta^{\mathrm{H}} \boldsymbol{\Phi} \mathbf{w}_\Delta$$

subject to the constraint in Equation 10.188 is[5]

$$\mathbf{w}_\Delta = \boldsymbol{\Phi}^{-1}\mathbf{H}^*(\mathbf{H}^{\mathrm{T}}\boldsymbol{\Phi}^{-1}\mathbf{H}^*)^{-1}\rho \tag{10.191}$$

In order to illustrate the results, we consider a 13 element linear array with 14 temporal taps per element, designed to detect low-speed targets in heavy ground clutter (clutter-to-noise ratio = 65 dB per element). In this case, the weight vector in Equation 10.186 produces a sum beam having an interference-plus-noise power after adaptation, which is close to the noise floor for all target speeds V such that $0.05 < V/V_{\mathrm{b}} < 0.95$, where V_{b} is the radar blind speed. The weight vector given in Equation 10.191 produces a difference beam with an adapted, interference-plus-noise power close to the noise floor for all target speeds. The adapted monopulse pattern Δ/Σ is shown in Figure 10.78 for two different target[††] speeds. Note that the monopulse slope is nearly linear over the entire 3 dB width of the sum beam, as required. If the constraint in Equation 10.188 is not applied, the adapted monopulse pattern is highly distorted.

10.7.2. Summary

We have developed a procedure to synthesize sum and difference patterns for space–time adaptive arrays in such a way that a specified monopulse response can

[††] The processor produces a different weight vector for each target speed (target Doppler) to which it is tuned.

be achieved. The approach is quite general and has been applied to more scenarios than presented here, including the case of an adaptive array with spatial degrees of freedom only (i.e., $M = 1$). Additional details are available from the author.

10.8. GROUND AND AIRBORNE TARGET DETECTION WITH BISTATIC ADAPTIVE SPACE-BASED RADAR

(R. L. FANTE)

10.8.1. Introduction

A space-based, bistatic surveillance radar has recently been proposed for detecting ground-based moving targets. The concept consists of a geostationary transmit satellite and multiple moving receive satellites that detect moving targets immersed in ground clutter. Because the receiving satellites are in motion, the ground clutter at an angle ϕ relative to the satellite velocity vector appears to have a speed $V\cos\phi$ where V is the satellite speed. Therefore, the ground clutter cannot be removed by simple Doppler processing, as is done with ground-based radars. If the receive aperture has a dimension D along the direction of travel so that the azimuthal extent of the main beam is $\Delta\phi = \lambda/D$ where λ is the wavelength, then the velocity spread of the main beam clutter is $\lambda V/D$. For $\lambda = 0.03$ m, $V = 7500$ m/s and $D = 10$ m the main beam clutter spreads across $\pm$ 23 m/s; consequently, all ground targets in speed range are of -23 to $+23$ m/s are buried within the main beam clutter. One can detect these low-speed targets by increasing the aperture length, and hence, decreasing the width of the main beam. But this may be impractical because one needs apertures that are at least 100 m long to have any chance of detecting targets with radial velocities of 2 m/s. Thus, we desire to consider other approaches that will allow detection, of the low-speed targets, with smaller apertures.

The first approach that suggests itself is DPCA (Displaced Phase Center Antenna) which attempts to compensate for receiver motion by maintaining the phase center of the receive antenna at a fixed location. However, this has several disadvantages, including a relatively wide notch around the zero speed (i.e., the notch in the response around zero speed is fairly wide so that low-speed targets are either not detected or detected poorly). This limitation can be overcome by using a generalization of DPCA called Space–Time Adaptive Processing. In space–time processing an adaptive finite impulse response filter is able to preserve the response at the Azimuth and Doppler bin occupied by the desired target while simultaneously placing nulls on the clutter at all azimuths that produce interference in the same Doppler bin as the target.

In the next section we will develop a theory for a space–time adaptive array using a low-cost partially filled receive aperture. We will demonstrate that this array can not only form a sum beam to detect slowly-moving targets immersed in ground clutter, but can also form an adaptive monopulse channel capable of locating them.

10.8.2. Analysis

10.8.2.1. Sum Beam

Consider a receive array consisting of N subarrays, with each subarray having K columns, as shown in Figure 10.79.

This array is assumed to move at a constant speed V in the direction indicated in Figure 10.79.

In order to cancel clutter that is in motion relative to the receiver (i.e., we can either consider the clutter as stationary and the receiver moving or the receiver as stationary and the clutter moving past it at a speed V) we need to add temporal degrees of freedom behind each subarray, as indicated in Figure 10.80.

We now wish to weight each of the tap voltages in all subarrays and combine them to minimize the clutter while maintaining the receiver gain at some desired target azimuth θ_0 and a desired Doppler frequency f_0.

That is, we set up a bank of values (θ_0, f_0), and for each set we minimize the clutter. In order to perform this minimization let us first define the voltage vector

$$\mathbf{v}^{\mathrm{T}} = [v_1(t),\ v_1(t-T),\ v_1(t-2T), \ldots v_2(t),\ v_2(t-T)\ldots] \tag{10.192}$$

where $v_{\mathrm{n}}(t - pT)$ is the voltage received on tap p of subarray n. Then if we weight and sum the voltages on all time taps in all subarrays we can write that sum as

$$\mathbf{r} = \mathbf{w}^{\mathrm{T}}\mathbf{v} = \mathbf{v}^{\mathrm{T}}\mathbf{w} \tag{10.193}$$

where $\mathbf{w}$ is an $NQ \times 1$ weight vector defined as

$$\mathbf{w}^{\mathrm{T}} = [W_{10}\ W_{11}\ W_{12} \ldots W_{20}\ W_{21}\ldots] \tag{10.194}$$

N is the number of adaptive subarrays, Q is the number of temporal taps per subarray and W_{nk} is the weight applied to the voltage on time tap k of subarray n.

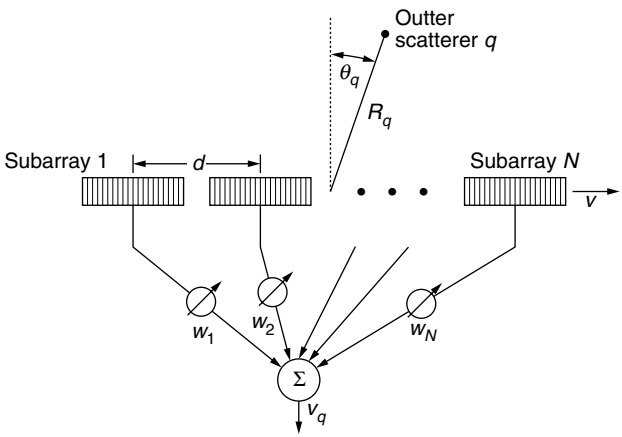

FIGURE 10.79 Antenna array that employs spatial degrees of freedom only.

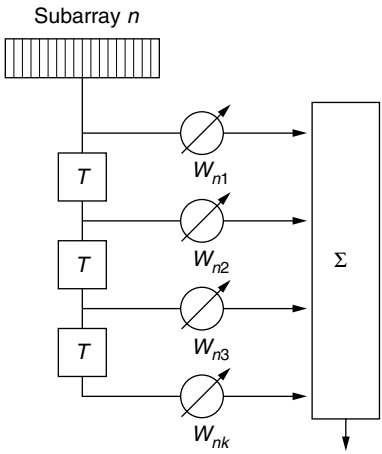

FIGURE 10.80 Adaptive FIR filter behind each array.

We wish to minimize the average clutter power

$$|\mathbf{r}|^2 = \mathbf{w}^{\mathrm{H}}\langle \mathbf{v}^* \mathbf{v}^{\mathrm{T}} \rangle \mathbf{w} = \mathbf{w}^{\mathrm{H}} \boldsymbol{\Phi} \mathbf{w} \tag{10.195}$$

while maintaining the receive pattern gain in the desired target direction and Doppler frequency. In Equation 10.195 the brackets $\langle\rangle$ denote an expectation. A target at the desired azimuth θ_0 and Doppler frequency f_0 produces a voltage on time tap p of subarray n given by $\mathbf{s}_n(p)$. We wish to find the weight vector that minimizes $\langle |\mathbf{r}|^2 \rangle$ while maintaining $\mathbf{w}^{\mathbf{T}}\mathbf{s} =$ constant, where s is an $NQ \times 1$ vector, defined as

$$\mathbf{s}^{\mathrm{T}} = [s_1(0),\ s_1(1),\ \ldots s_2(0),\ s_2(1),\ \ldots] \tag{10.196}$$

The solution is[1,2]

$$\mathbf{w} = \frac{\boldsymbol{\Phi}^{-1}\mathbf{s}^*}{\mathbf{s}^{\mathrm{T}}\boldsymbol{\Phi}^{-1}\mathbf{s}^*} \tag{10.197}$$

10.8.2.2. Difference Beam

We next wish to form a difference beam with the following properties:

1. The received clutter power is minimized.
2. The difference beam has a zero at $\theta = \theta_0$ and maintains a desired slope such that

$$\frac{V_\Delta(\theta_0 \pm \Delta\theta,\ f_0)}{|V_\epsilon(\theta_0 \pm \Delta\theta,\ f_0)|} = \pm k_s \Delta\theta \tag{10.198}$$

where $V_\Delta(\theta)$ is the difference beam voltage after adaptation and $V_\Sigma(\theta)$ is the sum beam voltage pattern after adaptation. The difference beam pattern is given by

$$v_\Delta(\theta, f) = \mathbf{w}_\Delta^{\mathrm{T}} \mathbf{g}(\theta, f) \tag{10.199}$$

where $\mathbf{w}_\Delta$ is the $NQ \times 1$ difference beam weight vector that we will determine later. Sometimes the two constraints in Equation 10.198 are insufficient to maintain the desired slope over the entire half-power width of the sum beam, and if necessary additional constraints can be imposed. For example, in addition to Equation 10.198 we can impose

$$\frac{V_\Delta(\theta_0 \pm \Delta\theta',\ f_0)}{|V_\sum(\theta_0 \pm \Delta\theta',\ f_0)|} = \pm k_s \Delta\theta' \tag{10.200}$$

where $\Delta\theta'$ is different from $\Delta\theta$.

If we use only the two constraints indicated in Equation 10.198 we can write the constraint in the form

$$\mathbf{H}^{\mathrm{T}} \mathbf{w}_\Delta = \rho \tag{10.201}$$

where H^{T} is a $2 \times NQ$ matrix of the form

$$\mathbf{H}^{\mathrm{T}} = \begin{bmatrix} \mathbf{g}^{\mathrm{T}}(\theta_0 + \Delta\theta,\ f_0) \\ \mathbf{g}^{\mathrm{T}}(\theta_0 - \Delta\theta,\ f_0) \end{bmatrix} \tag{10.202}$$

ρ is a 2×1 vector of the form

$$\boldsymbol{\rho} = k_s \begin{bmatrix} \mathbf{w}^{\mathrm{T}} \mathbf{g}(\theta_0 + \Delta\theta,\ f_0) \\ -\mathbf{w}^{\mathrm{T}} \mathbf{g}(\theta_0 - \Delta\theta,\ f_0) \end{bmatrix} \Delta\theta \tag{10.203}$$

and $\mathbf{w}$ is the sum-beam weight vector given in Equation 10.197.

If we use four constraints then the constraint equation still is of the form in Equation 10.201, but now $\mathbf{H}^{\mathrm{T}}$ is a $4 \times NQ$ matrix and $\boldsymbol{\rho}$ is a 4×1 vector. In either case the solution is[1,2]

$$\mathbf{w}_\Delta = \mathbf{R}^{-1} \mathbf{H}^* (\mathbf{H}^{\mathrm{T}} \mathbf{R}^{-1} \mathbf{H}^*)^{-1} \boldsymbol{\rho} \tag{10.204}$$

10.8.3. Numerical Studies of Effectiveness

In this section we will investigate how well a fast-moving radar can detect and locate low speed targets immersed in heavy clutter. However, before we proceed to evaluate target detection we must first investigate a design constraint on the length and maximum allowable radar platform speed. For a stationary transmitter the blind speed for ground targets viewed by the moving receiver is $V_{\mathrm{blind}} = \lambda / T$, where λ is the wavelength and T is the interpulse period. Thus, for a selected blind speed the pulse repetition frequency $1/T$ is determined. For

example, if the wavelength is 0.03 m and the target blind speed is 100 m/sec then $T = 0.0003$ s and the pulse repetition frequency is 3333 Hz. Now suppose the receiver is a linear array of length D moving at a speed V. We would expect that there is a limit on the array speed V that requires successive received pulses fall on the array. That is, we would expect that the temporal pulse processing will not be effective unless $VT < D$. In order to prove this point we first studied the effect of receiver speed V on the effectiveness of a 21 element linear array in detecting a target with a speed of 50 m/sec (one half the blind speed) that is immersed in very heavy clutter (clutter-to-noise ratio per element of 72 dB). In Figure 10.81 we show the loss in signal-to-interference ratio after adaptation, relative to the clutter-free signal-to-noise level as a function of VT/D. Note that as long as $VT < 0.5D$ there is virtually no loss in signal-to-interference ratio, but the loss rises dramatically for $VT \geq D$. Thus, in practice it is necessary to choose D sufficiently large so that

$$\frac{VT}{D} \leq 0.5 \tag{10.205}$$

The parameter VT/D also defines the fraction of the blind speed interval that is occupied by the main beam clutter. The 3 dB beamwidth of the sum beam of the receiver is approximately λ/D. If the receiver is moving at a speed V and the clutter is stationary, then the speed of the clutter at the azimuths corresponding to the ± 3 dB points of the receive beam is $\pm \lambda V/2D$. For a bistatic radar with a stationary transmitter the blind speed is λ / T, so that the

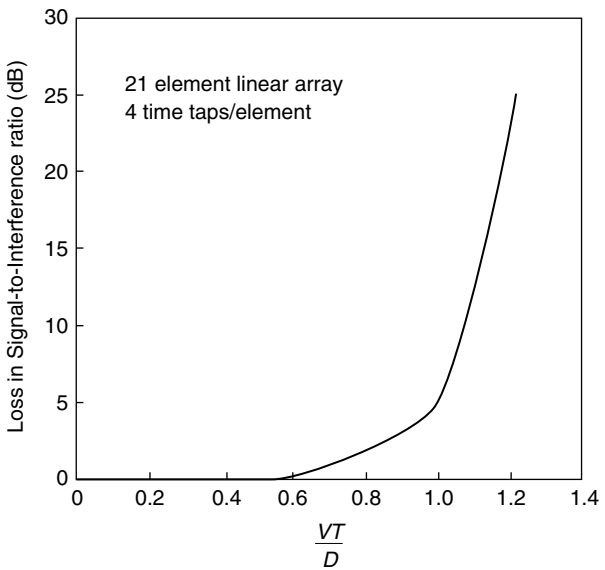

FIGURE 10.81 Effect of platform speed on the signal-to-interference ratio after adaptation.

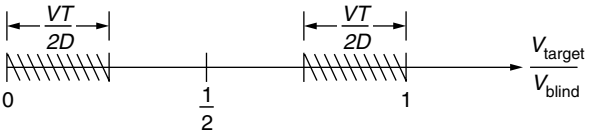

FIGURE 10.82 Fraction of the blind speed interval occupied by main beam clutter.

total fraction of the blind speed occupied by the main beam clutter is

$$\frac{\Delta V_{\text{main}}}{V_{\text{blind}}} = \frac{\lambda V}{D} \cdot \frac{T}{\lambda} = \frac{VT}{D} \tag{10.206}$$

This is illustrated conceptually in Figure 10.82. Because we will often be forced to have values of VT/D close to one half, this means that the 3 dB main beam clutter usually fills nearly one half of the blind speed interval.

10.8.3.1. Sum Beam

We are now ready to present some numerical results for the adapted sum beam. Let us first illustrate how increasing the number of pulses improves detection. We again consider a 21 element linear array attempting to detect a target located broadside to the array in stationary (no internal motion) clutter that is strong enough to produce a clutter-to-noise ratio per element of 72 dB. The target is assumed to produce a signal-to-noise ratio of 0 dB/element, so that before adaptation the clutter is more than 70 dB stronger than the signal. In Figures 10.83 and 10.84 we show the ability of adaptive processing to cancel the clutter for the cases when $VT = 0.1D$ and $VT = 0.5D$, respectively. Note that as more pulses are processed the signal-to-interference ratio after adaptation can be maintained at a high level even for very small target speeds.

Now that we have illustrated the cancellation with several examples, let us proceed to discuss the space-based application. We assume that the transmitter is a stationary satellite but the receiver is a satellite moving at a speed of 7500 m/s, and we desire to detect ground targets with speeds from 0 to 100 m/s. The frequency is X-Band (10 GHz), so the wavelength $\lambda = 0.03$ m. Therefore, we see that for a blind speed of 100 m/s we require $T = 0.0003$ s, or a pulse repetition frequency of 3333 Hz. In order to keep the cost reasonable we consider a receive array that is thinned, as shown in Figure 10.85. For this array each subarray is adaptive, but the individual columns are not. The total length of the array is 6.57 m, and the array is approximately 16% filled. For the parameters chosen $VT/D = 0.34$, satisfying the requirement in Equation 10.205.

We now study the ability of this adaptive array to detect targets in heavy clutter. The clutter is assumed to be uniformly distributed in angle and illuminated by a transmit beam that is ten receive beamwidths wide, and produces a clutter to noise ratio of 77 dB at each receive column. Each subarray of the receive array has an adaptive tapped delay line containing 14 taps so that the total

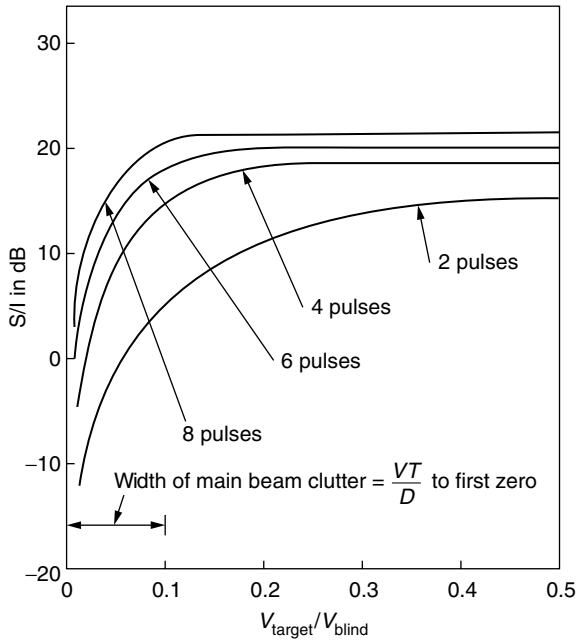

FIGURE 10.83 Effect of processing more pulses on the performance of a 21-element linear array with $VT/D = 0.1$ and element spacing $= \lambda/2$.

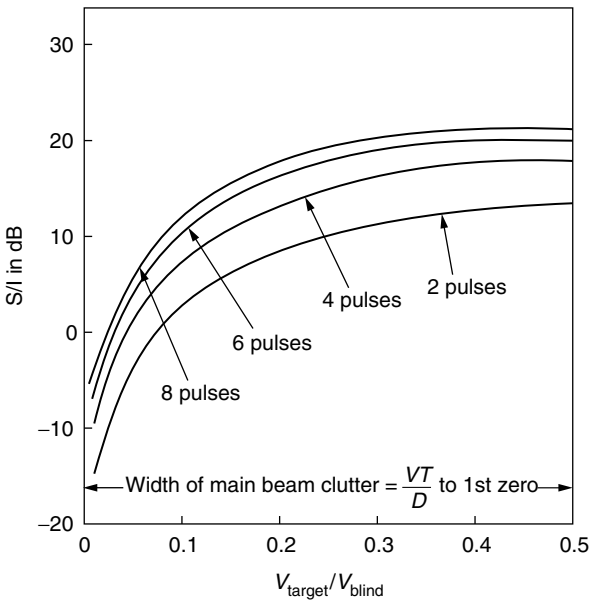

FIGURE 10.84 Effect of processing more pulses on the performance of a 21-element linear array with $VT/D = 0.5$ and element spacing $= \lambda/2$.

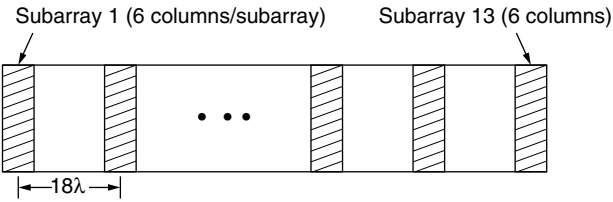

FIGURE 10.85 Thinned receive array.

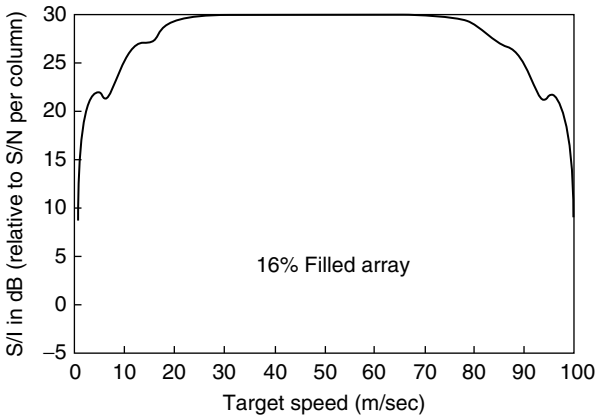

FIGURE 10.86 Signal-to-interference ratio gain for the thinned array in Figure 10.85 when 14 pulses are processed.

number of adaptive degrees of freedom is 14 taps × 13 subarrays = 182. In Figure 10.86 we show the signal-to-interference ratio after adaptation (relative to the signal-to-noise ratio per column in the absence of clutter) as a function of target speed, for a target located broadside to the array. For targets at one half the blind speed, the processor is able to cancel the clutter down to approximately the noise floor, and thus achieve the full gain

$$G = 10 \log_{10}(N_S \cdot N_C \cdot N_T)$$

where N_S is the number of subarrays, N_C is the number of columns per subarray and N_T is the number of temporal taps (pulses). Thus, $G = 10\log_{10}(13.6.14) \cong 30$ dB. The results in Figure 10.86 also indicate that this receive array should be able to detect targets with speeds as low as 2 m/s and as high as 98 m/s without incurring more than 10 dB penalty relative to the maximum gain. A typical adapted sum beam pattern is shown in Figure 10.87 for a filter tuned[‡‡] to 17.5 m/sec.

[‡‡] That is, the results shown are the adapted patterns when the processor is separately tuned to each target speed.

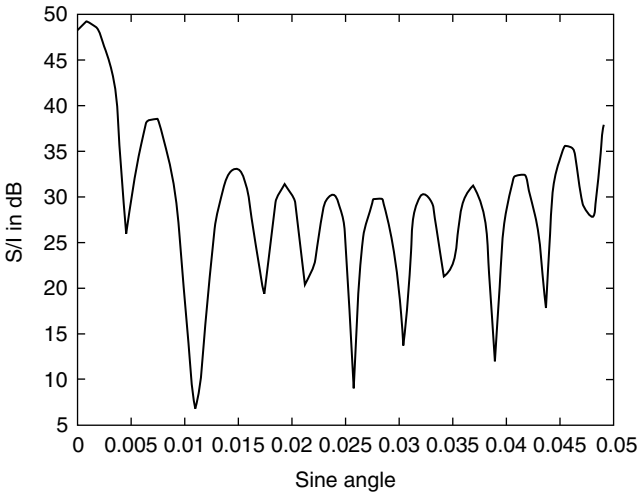

FIGURE 10.87 Adapted sum beam for filter tunes to 17.5 m/s.

This pattern illustrates a feature that can occur with an adaptive processor: there may be a distortion of the receive beam and the peak of the beam may shift slightly from the desired direction (which, in our case, is zero degrees). A shift may occur whenever the clutter that is competing in the same Doppler bin as the target, lies within the main beam. That is, for a target with speed V_{target} the clutter that competes in the same Doppler bin lies in an angular region centered on $V\sin\theta = -V_{\text{target}}$. Thus, for $\sin\theta \cong \theta$ we can rewrite this condition as

$$\frac{\theta}{\theta_{3\text{dB}}} = -\frac{V_{\text{target}}}{V_{\text{blind}}}\left(\frac{D}{VT}\right) \tag{10.207}$$

where θ is the azimuth of the clutter that competes in the same Doppler bin with the target and $\theta_{3\text{dB}}$ is the 3 dB beamwidth of the receive beam. For our array $D/VT = 2.94$ so that the competing clutter lies within the main beam whenever $|V_{\text{target}}| < 0.34V_{\text{blind}}$. This explains why the adapted main beam has shifted in Figure 10.87. The processor is trying to move the competing clutter out of the main beam.

10.8.3.2. Difference Beam

Next we consider the difference beam. We study the same geometry as in the previous section, and attempt to synthesize a desired monopulse slope after adaptation using the two constraints given in Equation 10.198. Although the maximum value of k_s is approximately $k_s = 1.81/\theta_{3\text{dB}}$, we will choose $k_s = 1.81/\theta_{3\text{dB}}$ and apply the constraints at $\Delta\theta = 0.033\theta_{3\text{dB}}$. In Figure 10.88 we show the positive-angle portion of the adapted monopulse pattern for four different target speeds. That is, the plots shown are the monopulse patterns when the adaptive

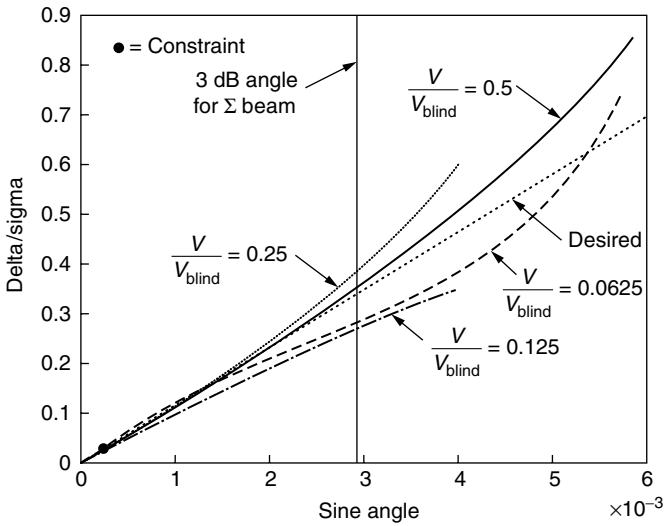

FIGURE 10.88 Adapted monopulse patterns for constraints at $\pm 0.333\theta_{3\text{dB}}$.

filter is tuned to each of the target speeds indicated. Note that the monopulse patterns are all monotonic, but their slope does vary from the desired slope for angles well-removed from the angle where the constraint is set. We note in passing that the clutter plus noise in the adapted difference beam never rises more than about 3 dB above the noise floor for any target speed.

It is disappointing that the monopulse slope in Figure 10.88 is not the same for all target speeds. One might expect to achieve more uniform slopes by applying the constraints at a larger value of $\Delta\theta$. However, this does not work as well as hoped, and when we used $\Delta\theta = 0.2\theta_{3\text{dB}}$ rather than $\Delta\theta = 0.033\theta_{3\text{dB}}$ we actually ended up introducing a bias (i.e., the zero in the monopulse response no longer occurs at $\theta = 0$). This suggests maintaining the constraints at $\pm 0.033\ \theta_{3\text{dB}}$ to force the response to pass through zero at $\theta = 0$ and then adding two new constraints (a total of four constraints) at $\pm\alpha\theta_{3\text{dB}}$ where $\alpha \gg 0.033$.

In Figure 10.89 we show the monopulse response for two different target speeds when the constraints are applied at $\pm 0.033\ \theta_{3\text{dB}}$ and $\pm 0.33\theta_{3\text{dB}}$. Note that the responses are quite close to both each other and the desired response, and it is only near the 3 dB point of the sum beam that there is any significant deviation.

It should be noted that when four rather than two constraints are used there are fewer degrees of freedom available to cancel the clutter in the difference beam. Thus, for targets at one half the blind speed the interference-plus-noise to noise ratio rises from 1.25 to 4.6 dB. However, for the low speed targets (which compete with clutter near the null in the difference beam) the use of four rather than two constraints results in very little difference in the interference-pulse-noise to noise ratio after adaptation.

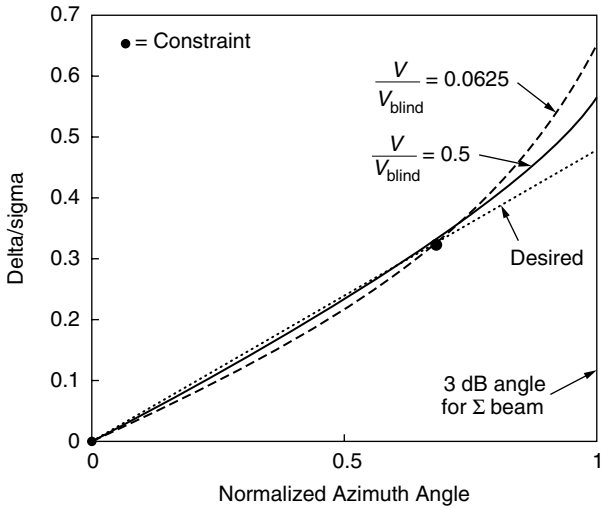

FIGURE 10.89 Adapted monopulse patterns for constraints at $\pm 0.333\theta_{3\text{dB}}$ and $+0.33\theta_{3\text{dB}}$.

10.8.4. SUMMARY

We have demonstrated that by using space–time adaptive, bistatic, spaceborne arrays it is possible to detect slowly-moving targets immersed in strong ground clutter. It is also possible to design an adaptive monopulse that can locate those targets in cross range. If the angular accuracy provided by the adaptive monopulse (approximately 0.03°) is not sufficient one can perhaps construct a long interferometer consisting of two or more of the adaptive arrays discussed in this report. This will be the subject of a future section.

10.9. ADAPTIVE NULLING OF SYNTHETIC APERTURE RADAR (SAR) SIDELOBE DISCRETES

(R. L. FANTE)

10.9.1. INTRODUCTION

Synthetic aperture radar (SAR) uses the motion of the aircraft on which the radar is mounted to achieve a large effective crossrange aperture, thus, achieving high resolution images.[1–5] However, when imaging targets immersed in ground clutter, the image may be affected by the presence of strong discrete scatterers, such as mountains, in the SAR response sidelobes. That is, although the SAR sidelobes are usually low enough so that the image is unaffected by sidelobe clutter, there are occasionally discretes strong enough to break through and corrupt the target image. It is possible to compensate for these strong discrete

scatterers by using adaptive SAR. In fact, it was shown recently[6,7] that the entire SAR image can be formed adaptively without the use of Fourier transforms, as is conventionally done. Here, we pursue a hybrid approach in which the range portion of the SAR image is formed conventionally, but the crossrange portion is formed adaptively in such a way that nulls are placed at the crossrange location of any strong clutter discretes. The main contribution of this work is to demonstrate that by using an array of overlapped, adaptive subarrays rather than a full, adaptive nonoverlapped array to form the crossrange image, a considerable reduction in complexity (i.e., the dimension of the covariance matrix is reduced from an $N \times N$ matrix to an $M \times M$ matrix, where N is the total number of samples in the full array and M is the number of samples per subarray) is possible, without any significant loss in performance.

10.9.2. Fully Adaptive SAR

Let us begin with a brief description of SAR, and then discuss how adaptive processing can be used to cancel the clutter discretes. More detailed discussions of SAR are contained in Refs. 1–5.

Consider radar on an aircraft emitting a series of pulses

$$v(t) = P(t - kT)\exp[-\mathrm{i}2\pi f_c(t - kT)] \tag{10.208}$$

where t = time, k = integer ($k = 1, 2 \ldots N$), T = interpulse period, f_c is the carrier frequency and, typically, $P(t)$ is a chirp pulse, given by[1–3]

$$p(t) = \exp\left(-\mathrm{i}\frac{\pi B t^2}{T_0}\right)\mathrm{rect}\left(\frac{t}{T_0}\right) \tag{10.209}$$

and T_0 is the pulse length and B is the bandwidth. Now suppose this pulse train illuminates a point scatterer at range $R(t)$, and the received signals are recorded in a two-dimensional array $\mathbf{g}(t', t_k)$, where $= t' = t - kT$ is known as the "fast time" and $t_k = kT$ is known as the "slow time." Then, in terms of these variables the received signal, after down conversion, can be written as

$$\mathbf{g}(t', t_k) = AP\left[t' - \frac{2R(t_k)}{c}\right]\exp\left[\mathrm{i}\frac{4\pi f_c}{c}R(t_k)\right] \tag{10.210}$$

where A is a constant that depends on the scatterer strength and range (the range is assumed to vary negligibly during the pulse length T_0) and c is the speed of light.

In order to discuss range migration, let us take the Fourier transform of Equation 10.210 over the fast-time variable t'.

We obtain

$$G(f, t_k) = A\tilde{P}(f)\exp\left[\mathrm{i}\frac{4\pi}{c}(f + f_c)R(t_k)\right] \tag{10.211}$$

where $\tilde{P}(f)$ is the Fourier transform of $P(t)$ and f is the fast frequency. Now let us expand the range to the scatterer in a Taylor series about $t_k = 0$,

$$R(t_k) = R(0) + \dot{R}(0)t_k + \frac{\ddot{R}(0)}{2}t_k^2 + \cdots \tag{10.212}$$

If the result is substituted into Equation 10.211 we see that there is a term in the exponent equal to $4\pi f\, \dot{R}(0)t_k/c$ that couples the fast frequency f and slow time t_k, and may result in a migration of the scatterer out of the range resolution cell during the coherent integration interval $T_c = NT$, where N is the number of pulses that must be integrated to achieve the desired crossrange resolution with the SAR. If $\Delta R = \dot{R}(0)T_c$ is the total range migration during the coherent integration interval, we see that because $|f| \leq B/2$ where $B =$ bandwidth, the term $\exp(\mathrm{i}4\pi f\dot{R}(0)t_k/c)$ can be approximated by unity if $2\pi B\Delta R/c \ll 1$, where B is the fast-frequency bandwidth. If we recall that the range resolution of a radar is $\delta R = c/2B$ we see that $2\pi B\Delta R/c \ll 1$ simply implies $\pi\Delta R \ll \delta R$. When range migration can be ignored, Equation 10.211 can be approximated by

$$G(f, t_k) = A\tilde{P}(f)\exp\left\{\mathrm{i}\frac{4\pi}{c}[(f + f_c)R(0) + f_c\dot{R}(0)t_k + f_c\ddot{R}(0)t_k^2/2]\right\} \tag{10.213}$$

We have shown elsewhere[8] that when migration cannot be ignored it is possible to correct for it by rescaling the data. The discussion here then applies to the rescaled data.

If we now apply the matched range-filter $\tilde{P}^*(f)$, and then take an inverse Fourier transform over the fast-frequency variable, the data produced by this point scatterer in the range slow-time domain is

$$D(\rho,\ t_k) = H(\rho)\exp\left[\mathrm{i}\frac{4\pi}{c}f_c(\dot{R}(0)t_k + \ddot{R}(0)t_k^2/2\right] \tag{10.214}$$

where ρ is the range and

$$H(\rho) = A\int_{-\infty}^{\infty} \mathrm{d}f|\tilde{P}(f)^2|\exp\left[-\mathrm{i}\frac{4\pi f}{c}(\rho - R(0))\right] \tag{10.215}$$

An unimportant phase term has been lumped into A. The scatterer under consideration can be imaged in the range-Doppler domain $(\rho,\ f_\mathrm{d})$ if we first focus it by multiplying the data $D(\rho, t_k) \equiv D(\rho, k)$ by the quadratic focusing function

$$\exp(-\mathrm{i}2\pi f_c\ddot{R}(0)k^2T^2/c)$$

and then take an inverse Fourier transform to get

$$I(\rho, f_\mathrm{d}) = \sum_{k=0}^{N-1} D(\rho, k)\exp(-\mathrm{i}2\pi f_\mathrm{d}kT - \mathrm{i}\beta_0 k^2) \tag{10.216}$$

where $\beta_0 = 2\pi f_c\ddot{R}(0)T^2/c = 2\pi f_c V^2T^2/R(0)c$, V is the radar-platform speed and we have used $\ddot{R}(0) = V^2/R(0)$. This will produce an image at range $\rho = R(0)$ and Doppler $f_\mathrm{d} = 2f_c\dot{R}(0)/c$.

Thus far, we have illustrated how an individual scatterer is imaged. However, the scene being imaged contains multiple scattering centers, so that the data received in the range-slow time domain is actually a summation over the contributions from all the scattering centers. Let us denote the data from all scatterers by $d(\rho, k)$. Our goal is to image a particular scatterer with range-Doppler coordinates (ρ_0, f), while rejecting the interference produced by all the other scatterers, especially those very strong scatterers we have called "discretes."

As can be seen for Equation 10.214 a point scatterer at range ρ_0 and Doppler frequency $f_0 = 2f_c\dot{R}_0/c$ produces a sequence of slow-time samples given by

$$s_0(k) = \exp[\mathrm{i}2\pi f_0 kT + \mathrm{i}\beta_0 k^2] \tag{10.217}$$

where $\beta_0 = 2\pi f_c V^2 T^2/\rho_0 c$, and the range term $H(\rho_0)$ has been suppressed. Equation 10.217 represents the ideal desired response from a selected scatterer in the absence of all interference. Because the platform speed V and the range ρ_0 to the scatterer are known, we can therefore estimate β_0, and then correct for the defocusing by multiplying Equation 10.217 by $\exp(-\mathrm{i}\beta_0 k^2)$ to obtain the modified scatterer response (of course we do not have access to the received signal itself; $\tilde{s}$ is merely the ideal response)

$$\tilde{s}(k) = \exp[\mathrm{i}2\pi f_0 kT] \tag{10.218}$$

Likewise all of the actual data can be modified by the same correction to obtain

$$\tilde{d}(k) = d(\rho,\ k)\exp(-\mathrm{i}\beta_0 k^2) \tag{10.219}$$

where the range dependence has been suppressed in $\tilde{d}$. We now wish to apply a set of adaptive weights $w(k)$ to the modified slow-time data $\tilde{d}(k)$ such that the response from the scatterer (target) to be imaged at coordinates (ρ_0, f_0) is preserved while the contributions from all other scatterers are minimized. If we define the $N \times 1$ vectors $\mathbf{W}$, $\tilde{\mathbf{S}}$, and $\tilde{\mathbf{D}}$ as

$$\mathbf{W}^{\mathrm{T}} = [w(0)\ldots w(N-1)] \tag{10.220a}$$

$$\tilde{\mathbf{S}}^{\mathrm{T}} = [\tilde{s}(0)\ldots\tilde{s}(N-1)] \tag{10.220b}$$

$$\tilde{\mathbf{D}}^{\mathrm{T}} = [\tilde{d}(0)\ldots\tilde{d}(N-1)] \tag{10.220c}$$

we then desire to minimize the mean interference power

$$I = \langle \mathbf{W}^{\mathrm{H}}\tilde{\mathbf{D}}^{*}\tilde{\mathbf{D}}^{\mathrm{T}}\mathbf{W}\rangle = \mathbf{W}^{\mathrm{H}}\langle\tilde{\mathbf{D}}^{*}\tilde{\mathbf{D}}^{\mathrm{T}}\rangle\mathbf{W} \tag{10.221}$$

where preserving the desired scatterer signal through the constraint

$$\mathbf{W}^{\mathrm{T}}\tilde{\mathbf{S}} = 1 \tag{10.222}$$

The weight vector that accomplishes this objective is well known,[§§] and is

[§§] The solution is known as the minimum variance, distortionless look beamformer. It is a special case of the linearly constrained variance beamformer.

given by[9–15]

$$\mathbf{W} = \frac{\mathbf{\Phi}^{-1}\tilde{\mathbf{S}}^*}{\tilde{\mathbf{S}}^{\mathrm{T}}\mathbf{\Phi}^{-1}\tilde{\mathbf{S}}^*} \tag{10.223}$$

where $\mathbf{\Phi}$ is an $N \times N$ covariance matrix defined as

$$\mathbf{\Phi} = \langle \tilde{\mathbf{D}}^* \tilde{\mathbf{D}}^{\mathrm{T}} \rangle \tag{10.224}$$

and $\langle\rangle$ denotes an expectation. Although in principle the aforementioned approach accomplishes the objective, it is impractical, because it is difficult to obtain sufficient independent samples (even using forward-backward subaperture averaging) to estimate $\mathbf{\Phi}$, and even if $\mathbf{\Phi}$ could be estimated, its inversion is computationally burdensome because for a high-resolution image the number of slow-time samples N may be many thousand. However, as we now show there is a simple way around these difficulties.

10.9.3. Overlapped-Subarray SAR

Let us construct an SAR array consisting of Q overlapped adaptive subarrays, and instead of placing nulls (on interfering scatterers) on an array level, as was done in the last section, let us place nulls in each subarray. That is, the total SAR response is the product of the subarray pattern and the array pattern, so that if the subarray pattern has a crossrange null (on a strong discrete), this null also appears in the total pattern. The arrays are overlapped in order to reduce the magnitude of the grating lobes that always appear when nonoverlapped subarrays are employed.

We can form a set of overlapped subarrays (shown in Figure 10.90) by setting

$$k = m + pK \tag{10.225}$$

where $0 \le m \le M - 1$, $0 \le p \le Q - 1$, K is the number of samples overlapped, and M is the number of slow-time samples per subarray, so that the total length of the array is $M + (Q - 1)K$ slow-time samples. In order for the overlapped subarrays to achieve the same crossrange resolution as the nonoverlapped array we require that

$$M + (Q - 1)K = N \tag{10.226}$$

This gives a crossrange angular resolution

$$\delta\theta = \frac{\lambda}{2(N - 1)VT} \tag{10.227}$$

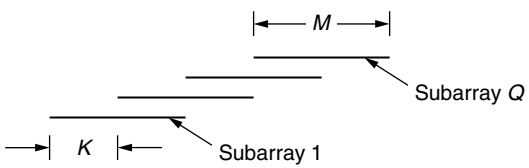

FIGURE 10.90 Full SAR array consists of Q overlapped subarrays of M points per subarray.

where λ is the carrier wavelength, V is the platform speed, and T is the interval between adjacent time samples. If Equation 10.225 is now used in Equation 10.218 we see that the slow-time signal, after quadratic compensation, produced on time sample m within subarray p by a scatterer at Doppler frequency f_0 is

$$\tilde{s}(m,p) = \exp[\mathrm{i}2\pi f_0 T(m + pK)] \tag{10.228}$$

We stress that $\tilde{s}$ is the desired signal that would be produced by the point selected for imaging in the absence of all interference.

Now let us apply a deterministic weight $w_\mathrm{d}(p)$ across the subarrays, in order to control the overall sidelobe response. Then, the desired signal after summing over subarrays is

$$\tilde{s}(m) = \sum_{p=0}^{Q-1} w_\mathrm{d}(p)\exp[\mathrm{i}2\pi f_0 T(m + pK)] \tag{10.229}$$

The correlated data $\tilde{d}(k) = \tilde{d}(m + pK)$ in Equation 10.219 is also weighted and summed over subarrays giving

$$\hat{d}(m) = \sum_{p=0}^{Q-1} w_\mathrm{d}(p)\tilde{d}(m + pK) = \sum_{p=0}^{Q-1} w_\mathrm{d}(p)d(m + pK)\exp(-\mathrm{i}\beta_0 k^2) \tag{10.230}$$

We now wish to apply a set of adaptive weights $w_s(m)$ to the modified slow-time data $\hat{d}(m)$ such that the response $\hat{s}(m)$ from the scatterer at Doppler f_0 is preserved, while the contributions from all other scatterers are minimized. Define the $M \times 1$ vectors

$$\mathbf{W}_s^\mathrm{T} = [w_s(0)\ldots w_s(M-1)] \tag{10.231a}$$

$$\hat{\mathbf{S}}^\mathrm{T} = [\hat{s}(0)\ldots\hat{s}(M-1)] \tag{10.231b}$$

$$\hat{\mathbf{D}}^\mathrm{T} = [\hat{d}(0)\ldots\hat{d}(M-1)] \tag{10.231c}$$

Then, paralleling the discussion in Equation 10.221 to Equation 10.224 we find that the output response $\mathbf{\Gamma} = \mathbf{W}_s^\mathrm{T}\hat{\mathbf{D}}$ has the desired properties if

$$\mathbf{W}_s = \frac{\mathbf{\Gamma}^{-1}\hat{\mathbf{S}}^*}{\hat{\mathbf{S}}^\mathrm{T}\mathbf{\Gamma}^{-1}\hat{\mathbf{S}}^*} \tag{10.232}$$

where $\mathbf{\Gamma}$ is an $M \times M$ matrix, defined as

$$\mathbf{\Gamma} = \langle\hat{\mathbf{D}}^*\hat{\mathbf{D}}\rangle^\mathrm{T} \tag{10.233}$$

The signal-to-interference ratio after adaptation is

$$\frac{\mathbf{S}}{\mathbf{I}} = \hat{\mathbf{S}}^\mathrm{T}\mathbf{\Gamma}^{-1}\hat{\mathbf{S}}^* \tag{10.234}$$

A typical component of Γ is

$$\Gamma_{mn} = \sum_{p=0}^{Q-1} \sum_{q=0}^{Q-1} w_{\mathrm{d}}^{*}(p) w_{\mathrm{d}}(q) \langle \tilde{d}^{*}(m + pK) \tilde{d}(r + qK) \rangle \tag{10.235}$$

We have now achieved a significant computational savings, because we need only estimate and invert an $M \times M$ matrix (where M is of order 10) rather than an $N \times N$ matrix, where N is of order 1000. Furthermore, as is evident from Equation 10.235, by using subarrays one automatically achieves looks at each scatterer from Q different angles (that is, each subaperture in Figure 10.90 views a given scatterer from a different angle) which tends to smooth the covariance estimate. Of course, all of these looks are not independent, because the looks need to be spaced in angle by $\lambda/2D$, where D is the crossrange dimension of the discrete clutter region, in order to be independent. Nevertheless, smoothing does occur. The remainder of the covariance smoothing is achieved through an average over range. That is, although we have suppressed the range index, the corrected data really is $\hat{d}(\rho,\ m + pK)$. Thus the average $\langle d^{*}(m + pK) \tilde{d}(r + qK) \rangle$ that appears in Equation 10.235 is estimated by summing the product over range bins ρ, and then dividing by the number of range bins. As demonstrated elsewhere,[16] one typically needs between $2M$ and $4M$ independent samples to achieve an acceptable estimate of the covariance.

Once the weight vector $\mathbf{W}_s$ has been determined, the response of the SAR to Doppler frequencies f_{d} other than the value f_0 to which the filter is tuned is

$$C(f_{\mathrm{d}}) = \mathbf{W}_s^{\mathrm{T}} h \tag{10.236}$$

where $\mathbf{h}$ is an $M \times 1$ vector, with components

$$h(m) = \sum_{p=0}^{Q-1} w_{\mathrm{d}}(p) \exp[\mathrm{i}2\pi f_{\mathrm{d}} T(m + pK)] \tag{10.237}$$

10.9.4. Numerical Results

In order to evaluate the performance achieved by using the adaptive, overlapped subarrays let us model the strong ground clutter as a set of N_{s} statistically independent point scatterers. In this model the scatterer labeled by the index n produces a voltage $d(k) = A_n \exp(\mathrm{i}2\pi f_n kT + \mathrm{i}\beta_n k^2 T^2)$, and the phase ϕ_n of each complex amplitude A_n is an independent random variable uniformly distributed in the interval $(0,\ 2\pi)$. Using this in Equation 10.235 then gives a theoretical covariance matrix $\mathbf{\Gamma}$ with components

$$\Gamma_{mk} = \sum_{n=1}^{N_{\mathrm{s}}} |A_n|^2 \sum_{p=0}^{Q-1} \sum_{q=0}^{Q-1} w_{\mathrm{d}}^{*}(p) w_{\mathrm{d}}(q) \exp\{ -\mathrm{i}2\pi f_n (m + pK - k - qK)$$

$$- \mathrm{i}\Delta\beta_n [(m + pK)^2 - (k + qK)^2] \} + y_{mn} \tag{10.238}$$

where A_n is the strength of the scatterer n, including the reduction (inherent in Equation 10.215) because the range ρ_n of scatterer n is different from range ρ_0 of the scatterer being imaged, f_n is the Doppler frequency of scatterer n and $\Delta\beta_n = \beta_n - \beta_0$ is its differential quadratic defocusing coefficient. The term y_{mk} is the receiver noise, defined as

$$y_{mk} = \sum_{p=0}^{Q-1} \sum_{q=0}^{Q-1} w_{\mathrm{d}}^*(p) w_{\mathrm{d}}(q) \langle x^*(m + pK) x(k + qK) \rangle \tag{10.239}$$

where $x(\mu)$ is the noise measured on slow-time sample μ. If all slow-time receiver noise samples are statistically independent, so that $\langle x^*(v)x(\mu)\rangle = \sigma^2 \delta_{\mu v}$ then it is readily shown that for real weights and subapertures overlapped by K samples

$$\begin{aligned} y_{mk} &= \sigma^2 \sum_{q=0}^{Q-1} w_{\mathrm{d}}^2(q), \quad k = m \\ &= \sigma^2 \sum_{q=1}^{Q-1} w_{\mathrm{d}}(q) w_{\mathrm{d}}(q-1); \quad k = m \pm K \\ &= \sigma^2 \sum_{q=2}^{Q-1} w_{\mathrm{d}}(q) w_{\mathrm{d}}(q-2); \quad k = m \pm 2K \end{aligned} \tag{10.240}$$

etc. For all other values of m and k we have $y_{mk} = 0$. Thus, because of subarray overlap the noise covariance matrix is nondiagonal. Because of the presence of this noise, the clutter-plus-noise covariance $\mathbf{\Gamma}$ is always nonsingular, even for completely correlated clutter scatterers. However, if the noise is too small the condition number of the matrix $\mathbf{\Gamma}$ may be very large. Thus, in some cases it may be necessary to artificially add noise to the diagonal of $\mathbf{\Gamma}$ to improve the condition number. This may result in a slight loss in the adapted signal-to-interference ratio.

Using the expression for $\mathbf{\Gamma}$ in Equation 10.238 we have evaluated how well the adaptive processor cancels undesired clutter discretes, while simultaneously imaging a target. One measure is the signal-to-interference ratio after adaptation, relative to the signal-to-noise ratio if the clutter discretes were absent. As a test, we chose a strong clutter discrete that produces an interference that is 30 dB above the noise level, and then varied its location f_{d} in Doppler relative to f_0. We chose an SAR array that consisted of 16 subarrays, each containing eight slow-time samples, and overlapped[##] by fifty percent, ($K = 4$). We then calculated the

[##] We have found that as long as the number M of samples in each subarray is much larger than the number of interfering discretes, the 50% overlapped adaptive subarray interference cancellation is within 1 dB of the interference cancellation produced by the fully adaptive array. However, there is nothing magic about our choice of 50% overlap. In fact, a 60% or 70% overlap gives the same interference cancellation with a smaller grating lobe (but at the expense of somewhat larger computational complexity).

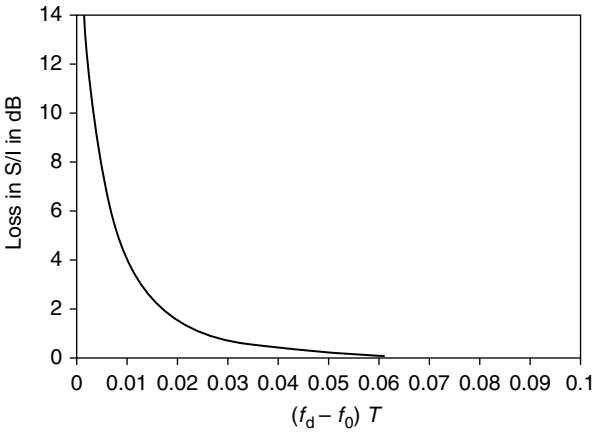

FIGURE 10.91 Loss in signal-to-interference ratio relative to clutter-free limit as function of clutter relative Doppler frequency.

signal-to-clutter-plus-noise ratio after adaptation relative to the signal-to-noise ratio in the absence of the clutter discrete. The results are shown in Figure 10.91, and indicate that as long as $F_dT \equiv (f_d - f_0)T$ is not too small, the clutter discrete is cancelled down to the noise floor.

The adapted SAR response functions (called "point spread functions" in optics) can be calculated by using Equation 10.236. We considered both the case when the deterministic weighting w_d across the array was uniform and Hamming. The response pattern for Hamming weighting and a single strong discrete at $F_dT = (f_d - f_0)T = 0.1$ is shown in Figure 10.92. We note that the processor has placed a null at $F_dT = 0.1$, and that the grating lobe at $F_dT = 0.25$ is down by

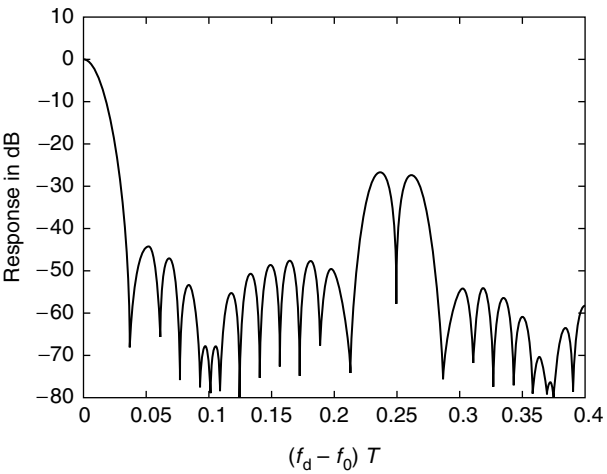

FIGURE 10.92 Adapted SAR response for strong discrete clutter at $(f_d - f_0)T = 0.1$.

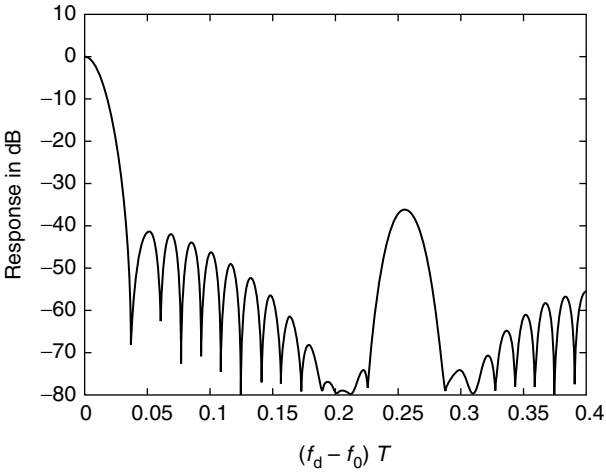

FIGURE 10.93 Adapted SAR response for strong discrete clutter at $(f_d - f_0)T = 0.2$.

nearly 30 dB from the peak response. Except for the null region the adapted pattern differs negligibly from the clutter-free pattern. In Figure 10.93, we show the response for Hamming weighting when the clutter discrete is moved to $F_dT = 0.2$. Now the processor has created a null region at $F_dT = 0.2$, but the pattern is otherwise undisturbed.

Adding additional discrete scatterers at other Doppler frequencies does *not* alter results presented in Figure 10.91. For example, when we sequentially placed additional strong ($C/N = 30$ dB each) clutter scatterers randomly within the region from $(f_d - f_0)T = 0.1$ to $(f_d - f_0)T = 0.5$ we found virtually no dependence of the adapted signal-to-interference ratio on the number of strong discretes. The result for six strong discretes differed only by 0.1 dB from the results for one discrete. This is because in the narrowband case we are considering that the contributions from all the strong discretes add coherently, so that the processor simply places a null on the composite direction of arrival.

We also examined whether there is any degradation because the clutter discretes are unfocused, as evidenced by the term proportional to $\Delta\beta_n = \beta_n - \beta_0$ in the covariance in Equation 10.238. We have found that the adapted signal-to-interference ratio is quite insensitive to discrete defocusing. In particular, we varied $\Delta\beta_n$ from[IIII] 0 to 0.1 and found only a 0.1 dB degradation in the adapted signal-to-interference ratio, even at $\Delta\beta = 0.1$. The adapted response function for this case is

[IIII] Recall that when sampling at the Nyquist rate we have $T = \rho_c/V$, where ρ_c is the crossrange resolution. Thus, $\beta_0 = 2\pi\rho_c^2/\rho_0\lambda$ where ρ_0 is the target range and λ is the wavelength. For a strong discrete at range ρ_1 we have $\Delta\beta = 2\pi\rho_c(\rho_1 - \rho_0)/\lambda\rho_0\rho_1$. If the crossrange resolution is 3 m, the wavelength is 0.2 m, $\rho_0 = 30$ km and $\rho_1 = 35$ km, we find that $\Delta\beta = 0.0013$. Thus, $\Delta\beta = 0.1$ is probably larger than is likely to occur in practice.

nearly identical to that for $\Delta\beta = 0$ except that, because of defocusing, there is no longer a distinct null at the Doppler frequency of the clutter discrete.

10.9.5. Summary

We have developed a computationally simple, hybrid processing approach that removes any image degradation caused by strong, discrete clutter in the sidelobes of a SAR image. The approach involves forming the range image conventionally, but forming the crossrange image in a partially adaptive fashion by using an array of overlapped, adaptive subarrays. The adaptive algorithm used to obtain the subarray weights is known as the minimum-variance, distortionless look beamformer, and has the advantages that the target response is always preserved. Thus, the algorithm does not require that the target signal be small in comparison with the sidelobe clutter, and as shown elsewhere[17] the algorithm is unaffected by the inclusion of the target signal is the covariance matrix, so that the algorithm works well even if the strong discrete scatterers are absent.

It is important, when using the algorithm developed, to always examine the condition number of the matrix $\mathbf{\Gamma}$, because $\mathbf{\Gamma}$ is generally rank deficient. As long as the clutter-to-noise ratio is not extremely large and sufficient samples are used to estimate $\mathbf{\Gamma}$, matrix conditioning will not be an issue. However, in instances where the clutter-to-noise ratio is extremely large, it may be necessary to artificially add noise to the diagonal of $\mathbf{\Gamma}$.

10.10. WIDEBAND CANCELLATION OF INTERFERENCE IN A GLOBAL POSITIONING SYSTEM (GPS) RECEIVE ARRAY

(R. L. FANTE AND J. J. VACCARO)

10.10.1. Introduction

It is desirable that GPS receivers operate efficiently, even in the presence of[22] interference and interference multipath. In order to counter this problem an adaptive antenna array is required because the location of the interference will not be known *a priori*. Conventional adaptive arrays that employ spatial degrees of freedom to place nulls in the direction of interferers perform well over very narrow bandwidths, but may be inadequate for broader band operation, especially when multipath is present. In order to form broadband nulls both spatial and temporal adaptive degrees of freedom are required.[1–3] A generic space–time processor that has the potential to cancel both interferers and their multipath is shown in Figure 10.94. This architecture is a special case of the three-dimensional space–time processor (3D STAP) proposed for radar systems to cancel ground clutter, jammers and jammer multipath.[3–7] The adaptive FIR (finite impulse response) filters are able to tailor the response so as to null interference plus multipath over broader bandwidths, provided the time delays T are less than $1/B$, where B is the operating bandwidth, and the length $(P-1)T$ of

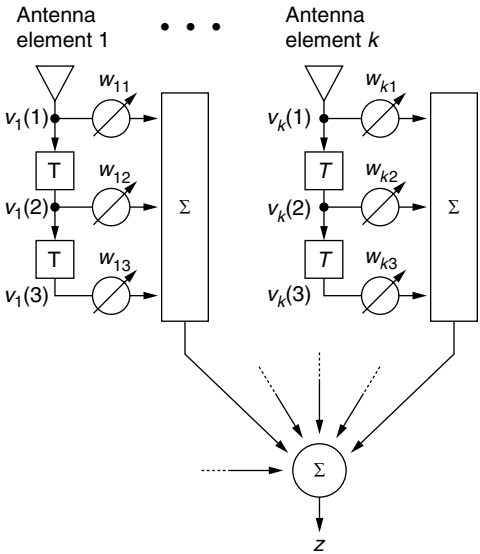

FIGURE 10.94 Adaptive space–time processor with three taps per antenna. Frequency transfer function of filter is $H(f)$.

each adaptive FIR filter is sufficient to encompass the differential multipath delays. The processor works by choosing the adaptive weights w_{kq} so as to preserve the desired GPS signal while simultaneously minimizing all interference. That is, we use the fact that we have information on the angular location of each GPS satellite and the nature of the signal it transmits, to tailor the space–time filter to receive this signal while rejecting others. In the next section we illustrate how this is done for hypothetical antenna arrays.

10.10.2. Adaptive Filter Weights

We now wish to choose the weights w_{kq} shown in Figure 10.94 so as to receive the signal from each[##] GPS satellite while at the same time rejecting all other signals. In order to see how this is done refer to Figure 10.94. The interferers plus their multipath produce a voltage $y_k(t)$ on antenna k, and this voltage is then received and downconverted, producing a voltage $v_k(t)$ at the input to the FIR filter, and a voltage $v_k(t - qT)$ on time tap q. We then desire to choose a set of weights w_{kq} so that the filter output

$$z(t) = \sum_{k=1}^{K} \sum_{q=0}^{P-1} w_{kq} v_k(t - qT) \tag{10.241}$$

[##] That is, a different set of weights is obtained for each GPS satellite.

preserves the GPS signal while canceling the interference and its multipath. In order to compactly formulate this problem let us define the $KP \times 1$ vector $\mathbf{V}$ as

$$\mathbf{V}^{\mathrm{T}} = [\nu_1(0)\ldots\nu_1(P-1)\ldots\nu_K(0)\ldots\nu_K(P-1)] \tag{10.242}$$

where $\nu_k(q) = \nu_k(t - qT)$ is the voltage on time tap q of antenna k produced by the interferers, interference multipath and noise. Also define a $KP \times 1$ weight vector $\mathbf{W}$ as

$$\mathbf{W}^{\mathrm{T}} = [W_{1,0}\ldots W_{1,P-1}\ldots W_{K,0}\ldots W_{K,P-1}] \tag{10.243}$$

Then Equation 10.241 can be rewritten as

$$z = \mathbf{W}^{\mathrm{T}}\mathbf{V} \tag{10.244}$$

The average interference power at the filter output is

$$\langle |z|^2 \rangle = \mathbf{W}^{\mathrm{H}}\langle \mathbf{V}^*\mathbf{V}^{\mathrm{T}} \rangle \mathbf{W} = \mathbf{W}^{\mathrm{H}}\mathbf{R}\mathbf{W} \tag{10.245}$$

where $\langle\rangle$ denotes an expectation and $\mathbf{R}$ is the interference-plus-noise covariance matrix. This covariance matrix is derived in Appendix P and includes the effects of system bandwidth, multipath, and channel mismatch. In practice, $\mathbf{R}$ is estimated[8] by averaging approximately $4KP$ independent time samples of the received interference. Because the GPS signal strength is at least 20 dB below the ambient noise floor, there is no need to be concerned about obtaining signal-free samples.

When the desired GPS signal is incident and signal multipath is absent,[†††] a voltage $s_k(q) \equiv s_k(t - qT)$ is produced on time tap q of antenna k. If we define a $KP \times 1$ vector $\mathbf{S}$ as

$$\mathbf{S}^{\mathrm{T}} = [s_1(0)\ldots s_1(P-1)\ldots s_K(0)\ldots s_K(P-1)] \tag{10.246}$$

we see that the signal voltage at the filter output is

$$z_s = \mathbf{W}^{\mathrm{T}}\mathbf{S} \tag{10.247}$$

We now wish to determine the weight vector W. Because S is a stochastic vector (typically the GPS signal is a Gold code or a PN (pseudonoise) sequence) the classical solution,[8–13]

$$\mathbf{W} = \mathbf{R}^{-1}\mathbf{S}^*$$

used for deterministic signals is not appropriate, so we now consider some different approaches for calculating $\mathbf{W}$.

10.10.2.1. Maximum Signal-to-Interference Ratio

The average signal power at the filter output can be obtained from Equation 10.247 as

$$\langle |z_s|^2 = \mathbf{W}^{\mathrm{H}}\langle \mathbf{S}^*\mathbf{S}^{\mathrm{T}} \rangle \mathbf{W} = \mathbf{W}^{\mathrm{H}}\mathbf{\Gamma}\,\mathbf{W} \tag{10.248}$$

We now wish to determine the weight W that maximizes the average signal-to-interference ratio

$$\mathbf{SI} = \frac{\mathbf{W}^{\mathrm{H}}\boldsymbol{\Gamma}\,\mathbf{W}}{\mathbf{W}^{\mathrm{H}}\mathbf{R}\mathbf{W}} \tag{10.249}$$

The desired weight $\mathbf{W}$ is readily shown[14] to be the eigenvector that corresponds to the maximum eigenvalue λ of

$$\boldsymbol{\Gamma}\,\mathbf{W} = \lambda\mathbf{R}\mathbf{W} \tag{10.250}$$

with the maximum signal-to-interference ratio (ignoring signal multipath) equal to the largest eigenvalue. Note that the process is repeated for each GPS satellite to be used in the solution for the platform position vector.

In practice, signal multipath is not always negligible, so the actual signal is $S_a = S + S_m$, where S_m is the multipath contribution. Then, the ideal signal covariance matrix must be replaced by the actual covariance matrix

$$\boldsymbol{\Gamma}_a = \langle(\mathbf{S}^* + \mathbf{S}_m^*)(\mathbf{S} + \mathbf{S}_m)^{\mathrm{T}}\rangle \tag{10.251}$$

and the actual signal-to-interference ratio after adaptation is

$$\mathbf{SI}_a = \frac{\mathbf{W}^{\mathrm{H}}\boldsymbol{\Gamma}_a\mathbf{W}}{\mathbf{W}^{\mathrm{H}}\mathbf{R}\mathbf{W}} \tag{10.252}$$

Typically, the signal multipath causes the actual signal-to-interference ratio to be a few tenths of a dB below the ideal value. Of course, the interference covariance matrix R always contains the interference multipath.

10.10.2.2. Minimum Mean Square Error

A second approach is to choose the weight vector to minimize the mean square difference between the desired temporal signal sequence s_{d} and the total received voltage $\mathbf{W}^{\mathrm{T}}\mathbf{Z}$, where $\mathbf{Z} = \mathbf{V} + \mathbf{S}$ is the total voltage vector consisting of signal plus all interference. That is, we choose $\mathbf{W}$ to minimize

$$\boldsymbol{\epsilon} = \langle|s_{\mathrm{d}} - \mathbf{W}^{\mathrm{T}}\mathbf{Z}|^2\rangle = \langle s_{\mathrm{d}}s_{\mathrm{d}}^*\rangle - \mathbf{W}^{\mathrm{T}}\langle\mathbf{Z}s_{\mathrm{d}}^*\rangle - \mathbf{W}^{\mathrm{H}}\langle\mathbf{Z}^* s_{\mathrm{d}}\rangle + \mathbf{W}^{\boldsymbol{H}}\langle\mathbf{Z}^*\mathbf{Z}^{\mathrm{T}}\rangle\mathbf{W} \tag{10.253}$$

If we differentiate $\boldsymbol{\epsilon}$ with respect to $\mathbf{W}^{\boldsymbol{H}}$, set the result equal to zero and use the information that the GPS signal is independent of and much weaker than the interference-plus-noise we find

$$\mathbf{W} = \mathbf{R}^{-1}\mathbf{g}_{\mathrm{s}} \tag{10.254}$$

where

$$g_{\mathrm{s}} = \langle\mathbf{S}^* s_{\mathrm{d}}\rangle \tag{10.255}$$

is the first column of the signal covariance matrix $\mathbf{\Gamma}$. The signal-to-interference ratio after adaptation is then obtained by using Equation 10.254 in Equation 10.252.

The weight vectors in Equation 10.250 and Equation 10.254 are not equal, but we have shown elsewhere (using very many typical portions of PN sequences), that the result in Equation 10.254 gives signal-to-interference ratios that are only a few tenths of a dB less than those given using the weight in Equation 10.250. Thus, either approach is acceptable, but in the simulations presented in Section 10.10.5, Equation 10.254 has been used.

10.10.2.3. Minimum Output Power

Because the GPS signal is usually far below the receiver noise level, a simple power minimization is often quite useful. In this method we simply constrain the weight on the middle tap of antenna 1 (see Figure 10.94), and then minimize the output power without attempting to preserve the gain in the signal direction. This method has the disadvantage of allowing for possible signal fades, but the advantage of not requiring the user to know the expected direction of arrival of the incoming GPS signal. This approach was not pursued in this work.

In our analysis in Sections 10.10.2.1 and 10.10.2.2 it was assumed that the angular location of each GPS satellite in view is approximately known, and that a separate weight vector is calculated to either maximize the signal-to-interference ratio or minimize the mean square error for each satellite. However, because the visible GPS satellites usually lie above the horizon and interference very often comes from at or near the horizon, it is possible to maximize the average of SI over some solid angle Ω_0 (or minimize the average of ϵ), where Ω_0 includes nearly the entire upper hemisphere except the region near the horizon. In this case Γ in Equation 10.249 is replaced by $\bar{\Gamma}$ and g_s in Equation 10.255 is replaced by $\bar{g}_s$, where

$$\bar{\Gamma} = \frac{1}{\Omega_0} \int\int_{\Omega_0} d\Omega \langle \mathbf{S}^* \mathbf{S}^T \rangle$$

$$\bar{g}_s = \frac{1}{\Omega_0} \int\int_{\Omega_0} d\Omega \langle \mathbf{S}^* s_d \rangle$$

and $d\Omega$ is the element of solid angle. The advantages of this procedure are that (a) one no longer needs to know the angular locations of each GPS satellite, and (b) a single weight vector is used for all satellites. However, the disadvantage is that we can no longer be assured of achieving full gain at the location of each satellite. In Ref. 22 we present a detailed tradeoff of individual constraints versus hemisphere-averaged constraints.

10.10.3. Signal Distortion Introduced by the Processor

As we see later the processor described in the last section can cancel the interference introduced by ensembles of wideband and narrowband interferers,

but because the architecture in Figure 10.94 does not have a uniform frequency response across the operating band, it is possible that it will introduce a distortion of the desired GPS signal. In order to study this problem, including steering vector mismatch, suppose the GPS satellite is located at polar angle (θ_s, ϕ_s) and we assume it is at a different angle $(\hat{\theta}_s, \hat{\phi}_s)$. Then, it is readily seen that the response of the filter in Figure 10.94 to that satellite is

$$H(f,\ \theta_s,\ \phi_s;\hat{\theta}_s,\ \hat{\phi}_s) = \sum_{k=1}^{K}\sum_{q=0}^{P-1} \tilde{w}_{kq}\exp\left[\mathrm{i}2\pi f\left(\frac{\xi_k}{c} + qT\right)\right] \tag{10.256}$$

where f is the frequency, $\tilde{w}_{kq} = w_{kq}(\hat{\theta}_s,\ \hat{\phi}_s)\exp(\mathrm{i}2\pi f_0\xi_k/c)$, f_0 is the carrier frequency,

$$\xi_k = x_k \sin\theta_s \cos\phi_s + y_k \sin\theta_s \sin\phi_s + z_k \cos\theta_s \tag{10.257}$$

$(x_k,\ y_k,\ z_k)$ is the position of antenna k and $w_{kq}(\hat{\theta}_s,\ \hat{\phi}_s)$ is the weight vector computed for assumed satellite position $(\hat{\theta}_s,\ \hat{\phi}_s)$. Because $(\hat{\theta}_s,\ \hat{\phi}_s)$ and the filter weights w_{kq} are all known (after adaptation), this filter function is readily constructed. If the GPS satellite emits a signal with a Fourier transform $S(f)$, then, in the absence of signal multipath, the signal output after passing through the adaptive filter is $H(f)S(f)$, which can be written in the time domain as

$$r(t) = \int_{-\infty}^{\infty} \mathrm{d}f\ H(f)S(f)\exp(\mathrm{i}2\pi ft) \tag{10.258}$$

where, for notational convenience we have suppressed the angular dependence of H. The GPS receiver estimates time delay by using[15,16] the cross correlation[‡‡‡] of this received signal with the known signal

$$r_0(t) = \int_{-\infty}^{\infty} \mathrm{d}f' S(f')\exp(\mathrm{i}2\pi f' t) \tag{10.259}$$

If the signal is a stationary random process, then

$$\langle S(f)S^*(f')\rangle = P(f)\delta(f - f') \tag{10.260}$$

where $P(f)$ is the power spectrum of the signal. If Equation 10.260 is used, along with Equation 10.258 and Equation 10.259, we see that the cross correlation is

$$C(\tau) = \langle r(t)r_0^*(t+\tau)\rangle = \int_{-\infty}^{\infty} \mathrm{d}f\ P(f)H(f)\exp(-\mathrm{i}2\pi f\tau) \tag{10.261}$$

Typically, $P(f)$ is a positive, symmetric function of f, so that in the absence of $H(f)$ the correlation peak occurs at $\tau = 0$. However, $H(f)$ can introduce both a broadening of the correlation peak and a shift of that peak from the correct value

[‡‡‡] This is referred to in GPS literature as the "code tracking loop."

at $\tau = 0$. This shift (or bias) is introduced by the phase $\phi(f)$ of $H(f)$. In addition, the presence of $H(f)$ could lead to a phase shift at the value of τ where the correlation peak occurs, and this could affect the high-precision, differential GPS systems that use carrier-phase tracking. Thus, in a later section it is necessary to carefully examine the effect of the adaptive antijam filter $H(f)$ on the cross correlation function. Of course, if necessary, both the bias error and any phase shifts, but not the broadening effect, can be corrected by following the adaptive FIR filter with the filter[§§§] $G(f) = H^*(f)$. Then, the cross correlation becomes

$$C'(r) = \int_{-\infty}^{\infty} \mathrm{d}f\ P(f)\ |H(f)|^2\ \exp(-\mathrm{i}2\pi f\tau) \tag{10.262}$$

Because $P(f)|H(f)|^2$ is real, the peak of $|C(\tau)|$ now lies at the correct location ($\tau = 0$) and, clearly, the phase of $C'(0)$ is zero, as desired. However, the correlation peak can potentially be broadened. Thus, there will be no bias in the pseudorange[###] estimate, but the standard deviation of the error in the estimated platform location may be increased. In a later section we show that this compensation filter $G(f)$ is rarely needed. Unfortunately the compensation filter $H^*(f)$ does not cure the range estimation problems caused by signal multipath. Signal multipath causes[17] the correct GPS signal transform $S(f)$ to be replaced by $S(f)(1 + \Delta(f))$, where Δ is of the form

$$\Delta(f) = \sum_{r=1}^{R_\mathrm{m}} a_\mathrm{r}\ \exp(-\mathrm{i}2\pi f\tau_\mathrm{r}) \tag{10.263}$$

and R_m is the number of multipath sources, and a_r, τ_r are the strength and delay of each multipath scatterer, respectively. Note that because the GPS satellite is at a different location than the interferers, the signal multipath delays are different from the interferer multipath delays. In the presence of signal multipath, Equation 10.262 is replaced by

$$C'(\tau) = \int_{-\infty}^{\infty} \mathrm{d}f\ P(f)\ |H(f)|^2\ [1 + \Delta(f)]\ \exp(-\mathrm{i}2\pi f\tau) \tag{10.264}$$

Unless $\Delta(f)$ is quite small, it is evident from Equation 10.264 that the signal multipath can introduce both a bias and a broadening of the correlation peak.

[§§§] It should be recognized that $G(f) = H^*(f)$ is not the only option for compensation. In fact the filter $G(f) = H^*(f)/|H(f)|^2$ is preferable because the cross correlation $C'(\tau)$ after compensation is then undistorted. However, whereas $H(f)$ is a simple FIR filter, $G(f)$ is not.

[###] Because the GPS satellite and receiver clocks are not identical, a GPS system does not measure true range, but rather pseudo-range. If the same clock were used then the differential delay would yield true range.

10.10.4. Suboptimum Space–Frequency Processing

We now wish to explore a suboptimum approach[18] that may possibly reduce the computational complexity without greatly sacrificing performance. One method that readily suggests itself is to process in the frequency rather than the time domain. That is, perhaps we can split the operating band B into M subbands of bandwidth B/M, and then in each subband calculate weights for each antenna that cancels the interference by placing spatial nulls on interferers within that bin while preserving the signal. If the discrete Fourier transform (implemented as a fast Fourier transform (FFT)) really did behave as a "brick wall" filter bank that fully isolated each frequency bin from all others, we expect that this procedure would rival the full time-domain processing discussed earlier. However, because of leakage between bins, even with well-designed windows, it is necessary to consider what is happening in some adjacent bins when attempting to minimize interference in a particular frequency bin. We now discuss a method for accounting for this spillover among bins.

Before we proceed to discuss this approach let us first transform the data into the frequency domain as

$$\mathbf{X}_n(m) = \sum_{k=0}^{M-1} b(k)\, x_n(k) \exp\left(\mathrm{i}\frac{2\pi mk}{M}\right) \tag{10.265}$$

where $b(k)$ is an arbitrary windowing function (e.g., Hamming), $x_n(k)$ is the voltage on time tap k of antenna n and $X_n(m)$ is the voltage in frequency bin m of antenna n. We combine these components into an $KM \times 1$ frequency-domain voltage vector $\mathbf{X}$ as

$$\mathbf{X}^T = [X_1(0)\ldots X_K(0),\ \ X_1(1)\ldots X_K(1)\ldots X_1(m)\ldots X_K(m)\ldots] \tag{10.266}$$

where K is the number of antennas. Next, we define a $KM \times 1$ weight vector $\mathbf{W}_Q(m)$ for order Q processing as

$$\mathbf{W}_1(m)^T = [0\ldots 0 w_1(m)\ldots w_K(m) 0\ldots 0] \tag{10.267}$$

$$\mathbf{W}_2(m)^T = [0\ldots 0 w_1(m)\ldots w_K(m) w_1(m+1)\ldots w_K(m+1) 0\ldots 0] \tag{10.268}$$

$$\mathbf{W}_3(m)^T = [0\ldots 0 w_1(m-1)\ldots w_K(m-1) w_1(m)\ldots w_K(m) w_1(m+1) \times \ldots w_K(m+1) 0\ldots 0] \tag{10.269}$$

Then, for an arbitrary order Q the output in frequency bin m after applying the weights to the data is

$$V_Q(m) = \mathbf{W}_Q^T(m)\mathbf{X} \tag{10.270}$$

Thus, for first order processing, defined by $Q = 1$, the output after weighting in frequency bin m is given by

$$V_1(m) = \sum_{n=1}^{K} w_n(m) X_n(m) \tag{10.271}$$

which clearly shows that bin m is processed without regard for the voltage in the other bins. Second-order processing, defined by $Q = 2$, gives

$$V_2(m) = \sum_{n=1}^{K} [w_n(m)X_n(m) + w_n(m+1)X_n(m+1)] \tag{10.272}$$

and for third-order ($Q = 3$) processing we have

$$V_3(m) = \sum_{n=1}^{K} [w_n(m-1)X_n(m-1) + w_n(m)X_n(m) + w_n(m+1)X_n(m+1)] \tag{10.273}$$

etc. Thus, as we increase the order Q we increasingly account for the data in other bins when calculating the output in a given bin m. Consequently, we expect that as the order is increased we can improve our ability to cancel the interference in each bin while preserving the desired signal there. In the limit when the filter order Q equals the number of taps, it can be shown[19] that space–frequency and space–time processing are equivalent.

For filter order Q, the unknown weights in Equation 10.271 to Equation 10.273 are computed by minimizing the mean square difference between the actual voltage $V_Q(m)$ in bin m and the Fourier transform $S_d(m)$ of the desired signal. The procedure parallels the approach of DiPietro,[18] and is omitted here (complete details are contained in Ref. 20). In a later section we compare suboptimum space–frequency processing with full space–time processing to see if space–frequency processing really offers any computational advantages.

10.10.5. Numerical Simulations

10.10.5.1. Introduction

We now wish to evaluate the cancellation performance of the full order space–time processor described in Section 10.10.2 and of the suboptimum space–frequency processor described in Section 10.10.4.

We assume *a priori* that the receiver in Figure 10.94 contains an analog-to-digital converter (A/D) with sufficient dynamic range to encompass the interference-to-noise ratio. We also assume that each receiver has a response which is essentially linear over the range of operation. If one specifies that, in the presence of a jammer nonlinearity, variations from receiver-to-receiver produce less than a 1 dB increase in output interference, this leads to the requirement[20] that

$$|\beta_n - \beta_m| < \frac{0.47}{J\left(\frac{J}{N}\right)^{1/2}} \tag{10.274}$$

where J is the interferer power, J/N is the interference-to-noise ratio and β_n, β_m are the cubic nonlinearity coefficients of receivers n and m, respectively.

For our first set of simulations we consider a seven-element planar array with elements uniformly spaced around a circle of radius equal to one half wavelength, assume that the operating frequency is 1.575 GHz (L1) and that the operating bandwidth is 20 MHz. In the absence of multipath the adaptive tapped delay lines shown in Figure 10.94 are probably unnecessary except for very strong interferers, because the maximum delay across the array in Figure 10.96 is only 0.67 nsec. Thus, the product of the bandwidth and this delay is only $0.67 \times 10^{-9} \times 20 \times 10^{6} = 0.013$. However, when the array is placed on a platform the maximum differential multipath delay can be of the order of 100 nsec and the delay-bandwidth product then is of order 2, so that the tapped delay lines are clearly necessary. In fact, $B\tau$ products as small as a few tenths produce significant degradation if only spatial nulling is used.

Our simulation accounts for antenna cross-polarization, mutual coupling between antennas and mismatch between the different antenna responses. We have found that for $P > 1$ mutual coupling has virtually no effect[‖‖] on the results unless the mutual coupling coefficient is greater than -15 dB. For nearly all arrays of interest, the mutual coupling is well below this value. In the results to be presented the mutual coupling option was ignored in order to save computer time. The antenna mismatch, however, is an extremely important parameter, and we have expressed the antenna mismatch in terms of a quantity known as the "cancellation ratio," which is minus the logarithm of the normalized residue power when two antennas are illuminated by the same interferer and their outputs subtracted. A high cancellation ratio means that the channels are well matched and a low cancellation ratio implies poorly matched receivers. A derivation of the cancellation ratio in terms of the amplitude and phase mismatch across the operating band is presented in Appendix P.

The multipath is modeled as a set of point scatterers with an arbitrary strength and position. The strength is defined relative to the strength of the direct ray from the interferer. Thus, a strength of -40 dB means that the multipath produces a voltage on each element, that is, 0.01 of the voltage produced by the direct signal from the interferer. Additional details on the multipath model are given in Appendix P.

Let us first examine how multipath affects interferer cancellation. We consider the processor in Figure 10.94 and assume that the array is illuminated by an array of interferers located randomly in azimuth near the horizon (x–y plane). The GPS satellite is assumed to be located 10° off array boresight and the multipath is modeled by four point scatterers located at (d, d), $(d, -d)$, $(-d, d)$, $(-d, -d)$. Thus, for interferers near the horizon the maximum differential delay $\tau_{\max}$ between the direct ray and the multipath is $\tau_{\max} = 2\sqrt{2}d/c$, where c is

[‖‖] For $P = 1$ (no time taps) mutual coupling significantly affects the results if the coupling coefficient exceeds approximately -30 dB.

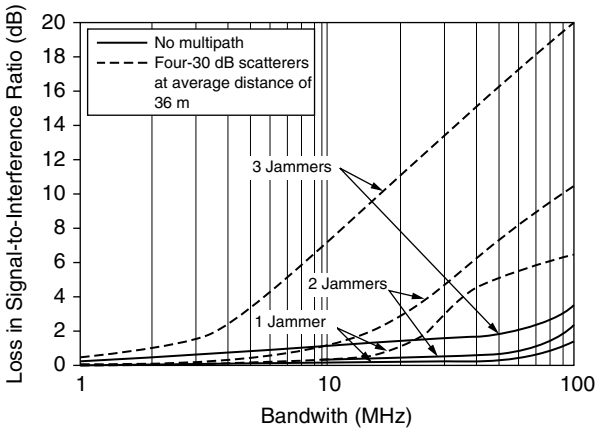

FIGURE 10.95 Effect of multipath on signal-to-interference loss for array that uses spatial degrees of freedom only.

the speed of light. Therefore, in order to compensate for delays of this order we require that the number of time taps P per antenna must satisfy (see Appendix Q)

$$(P-1)T \geq \tau_{\max} \tag{10.275}$$

For a tap spacing $T = 1/B$ and $d = 30$ m this condition becomes $P > 1 + 0.28B$ (MHz), where B(MHz) is the bandwidth in MHz. Thus, if only one tap is used (i.e., spatial degrees of freedom only) we expect performance to degrade significantly if bandwidth is much larger than a few MHz. In Figure 10.95 we show the loss, relative to the interference-free environment, in the adapted signal-to-interference ratio for the cases when the antenna array is illuminated by one, two, and three strong ($J/N = 60$ dB/antenna/interferer) interferers, and multipath is present and absent. Note that although there is little degradation in system performance in the absence of multipath, as expected, when multipath is present the performance begins to degrade badly for bandwidths exceeding a few MHz. This clearly indicates the need for the adaptive FIR filter behind each antenna, even when the antenna array is perfectly equalized (cancellation ratio $= \infty$).

We now illustrate for $B = 20$ MHz how adding temporal taps cures the degradation introduced by the multipath. In order to stress the processor we increase the interference-to-noise ratio from 60 to 70 dB per element per interferer, but again assume the array is well equalized (cancellation ratio = 120 dB). Also, now the four multipath scatterers are placed randomly[###] at an average distance of 12 m from the center of the antenna array, the tap spacing is now

[###] The scatterers are placed randomly (uniform probability density) within a circular annulus in the same plane as the antenna array. The inner radius of the annulus is $\bar{R} - \Delta R/2$ and its outer radius is $\bar{R} + \Delta R/2$ where $\bar{R} = 12$m and ΔR = 2 m.

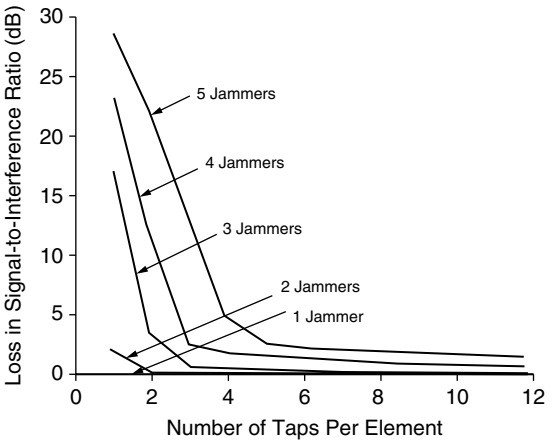

FIGURE 10.96 Illustration that increasing number of time taps decreases loss in signal-to-interference ratio.

$0.84/B$, and the results shown in Figure 10.96 are the averages over ten realizations of the multipath scatterer locations. From Figure 10.96 we see that, as expected, adding temporal taps does indeed reduce the loss in signal to interference caused by the interference plus multipath. The maximum differential multipath delay is approximately 140 nsec so that for $T = 0.84/B$ and $B = 20$ MHz, Equation 10.275 predicts that $Q \geq$ four-taps are required. As can be seen from Figure 10.96 the signal-to-interference loss is always modest if four or more taps are employed.

10.10.5.2. Effect of Channel Mismatch

Although the results is Figure 10.96 are encouraging we must recognize that they are for a perfectly equalized array. We must now explore how imperfect equalization affects performance. If the antenna pair is illuminated by an interferer that produces an interference-to-noise ratio of 70 dB per element then a cancellation ratio of 20 dB implies that the output interference-to-noise ratio after subtraction is 50 dB. However, an adaptive tapped delay line (i.e., FIR filter) of the type shown in Figure 10.94 can dramatically improve the channel equalization. We illustrate this in Figure 10.97 where we show the output interference-plus-noise to noise ratio when a pair of antennas operating over a 20 MHz band are illuminated by a broadband interferer that produces an interference-to-noise ratio of 60 dB/element. The results shown are the averages over ten Monte Carlo simulations of the receiver amplitude and phase ripples, for the case when each receiver has three (random) ripples across the 20 MHz passband. Note that, as expected, the adaptive array self-equalizes if sufficient temporal taps (as was the case previously, the tap spacing is $T = 0.84/B$) are employed.

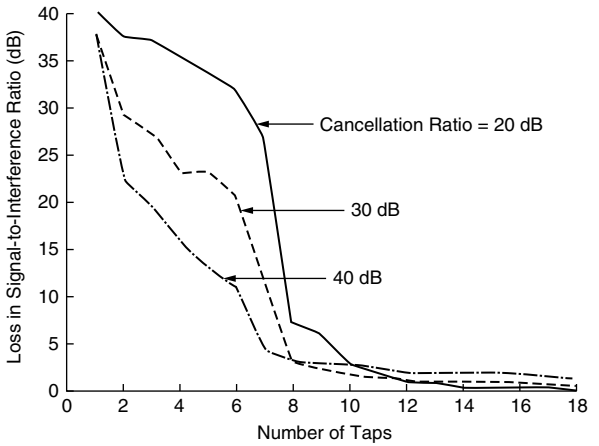

FIGURE 10.97 Illustration of self-equalization by two-element adaptive array.

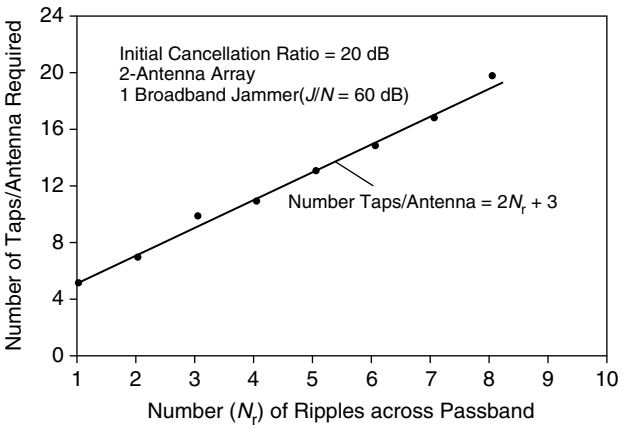

FIGURE 10.98 Self-equalization of two-antenna adaptive array.

We have performed multiple, additional Monte Carlo simulations of the self-equalization[††††] capability of an adaptive array, and some of the results are summarized in Figures 10.98 and 10.99. In Figure 10.98 we consider a two-element adaptive array illuminated by a single broadband interferer, and find that the antenna can self-equalize if each antenna contains $2N_r + 3$ taps, where N_r is the number of ripples across the band B in each receiver response function. Note that this is about half the number of taps calculated by Monzingo and Miller,[2]

[††††] We define "self-equalization" as cancellation of the interference to within 3 dB of the receiver noise floor.

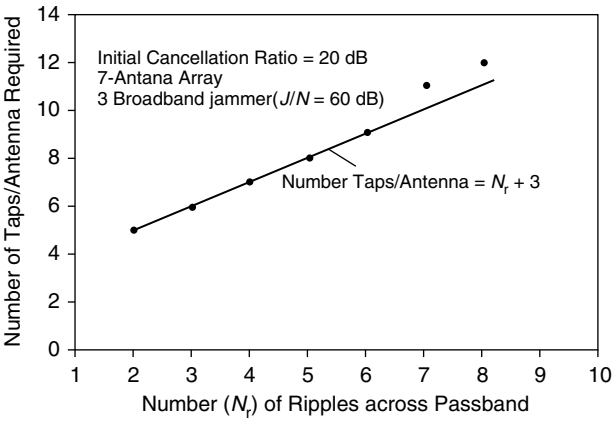

FIGURE 10.99 Self-equalization of seven-antenna adaptive array.

because they used an adaptive FIR filter behind only one antenna and we have one behind both antennas.

In Figure 10.99 we consider the seven-element circular array illuminated by three broadband interferers located randomly along the horizon ($0 <$ azimuth $< 2\pi$). For this case only $N_r + 3$ taps per antenna are required. Fewer taps per antenna are required (as compared with the two-element adaptive array) because the seven-element adaptive array has the freedom to choose the four or five channels that are best matched to each other, and ignore channels with response functions that are very different from the others (the two-element array does not have this luxury).

One should (if possible) choose a receiver technology that has the minimum number of random phase and amplitude ripples across the band B. If the number N_r of ripples is five or fewer it may be feasible to self-equalize. Otherwise a fixed-weight equalization filter is required in each channel.

10.10.5.3. Effect of Steering-Vector Mismatch

Because one does not always know the precise angular location of the GPS satellite, it is important to study the sensitivity of the adaptive filter to errors in satellite location. Therefore, we have performed multiple simulations of the loss in output signal-to-interference ratio as a function of the difference between the true satellite location (θ_s, ϕ_s) and our estimate ($\hat{\theta}_s$, $\hat{\phi}_s$) of its position, for the case of one to three strong jammers (60 dB/jammer/element) located randomly on the horizon (randomly placed over $0 <$ azimuth $< 2\pi$) and a GPS satellite located randomly in the upper hemisphere, but not within 15° of the horizon. The mean loss due to steering vector mismatch is summarized in Table 10.7. Observe that as long as the steering-vector mismatch is less than about 15° the loss is tolerable.

TABLE 10.7
Loss due to Steering-Vector Mismatch

Angular Error in GPS Direction (degrees)	Loss in S/I (dB)
0	0
5	0.1
10	0.6
15	1.4
20	2.4
25	3.7

10.10.5.4. Distortion Introduced by the Adaptive Filter

In Section 10.10.3 we discussed the potential effects on the crosscorrelation function introduced by the adaptive FIR filter in Figure 10.94. There are a number of issues to be considered. These are:

(1) Does the adaptive filter produce a bias (time shift) of the cross correlation function peak that is different for different GPS satellite locations (a constant bias that is the same for all satellites does not affect the estimate of position using the GPS system)?
(2) Is there significant broadening of the cross correlation, leading to an increase in pseudo-range error?
(3) Does the adaptive filter introduce a significant phase error that will affect the estimate of the carrier phase used by high-precision differential GPS systems?

We have provided answers to these questions through multiple simulations, using Equation 10.261 to calculate the cross correlation function for the case when the GPS signal is a PN sequence that has the power spectrum

$$P(f) = \frac{P_0}{B} \sin c^2\left(\frac{\pi f}{B}\right) \tag{10.276}$$

If Equation 10.256 and Equation 10.276 are substituted into Equation 10.261, the time origin shifted to the center tap, and the integration performed, we find

$$C(\tau) = P_0 \sum_{k=1}^{K} \sum_{q=0}^{P-1} \tilde{w}_{kq} \operatorname{tri}(\alpha) \tag{10.277}$$

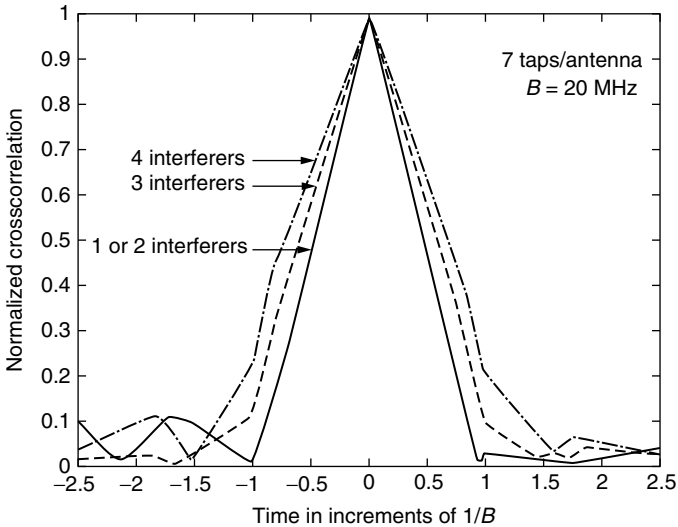

FIGURE 10.100 Uncompensated FIR filter does not significantly affect the cross correlation function.

where $\text{tri}(x) = 1 - |x|$ for $|x| < 1, \text{tri}(x) = 0$ for $|x| > 1$

$$\alpha = B\left[\tau + \frac{\xi_k}{c} + \left[q - \frac{(P-1)}{2}\right]T\right] \tag{10.278}$$

and $\tau = 0$ corresponds to the delay on the center tap of each delay line.

In Figure 10.100 we show a typical result for the normalized cross correlation function $C(\tau)/C(0)$ for the cases when there are one to four, broadband, strong (interference to noise ratio = 60 dB/element/interferer) interferers located randomly in azimuth along the horizon. Also present are four moderate strength (-30 dB) multipath scatterers located randomly near the plane of the antenna at an average range of 14.2 m. From Figure 10.100 we note the following points:

(1) There is no discernable displacement of the cross correlation peak from its correct location at $\tau = 0$. (In fact, in over 100 different simulations the largest shift we observed was only $0.05/B$).
(2) The cross correlation function has nearly the same shape in the presence of one or two interferers as it does in the interference-free case.[‡‡‡‡]

[‡‡‡‡] In the absence of interference the normalized crosscorrelation is $1 - B|\tau|$ for $|\tau| < B^{-1}$, and zero for $|\tau| > B^{-1}$.

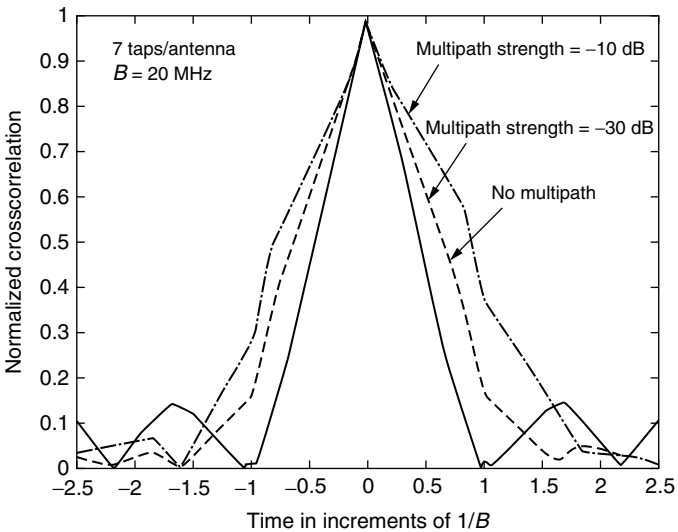

FIGURE 10.101 Cross correlation function is broadened in presence of very strong multipath.

(3) Even in the presence of four strong, broadband interferers there is not a drastic broadening of the cross correlation function. Hence, the increase in pseudo-range error is modest.

It is also important to question whether the presence of stronger multipath would modify the above conclusions. Therefore, we varied the strength and location of the multipath scatterers. In Figure 10.101 we show results for one of the worst cases encountered. Again the crosscorrelation peak is at the correct location, but for very strong multipath (-10 dB) there is a significant broadening of the cross correlation peak. In general, however, multipath this strong is not usual because -10 dB multipath corresponds to a scatterer with a bistatic radar cross section of 253 m^2 at a range of 14.2 m. Thus, we conclude that the adaptive FIR filter in Figure 10.94 does not shift or radically broaden the cross correlation function $C(\tau)$. Consequently there is no range bias error and only a very modest increase in pseudo-range error. The compensation filter $G(f) = H^*(f)$ discussed in Equation 10.262 is unnecessary.

We also studied the carrier phase error, which is equal to the phase of $C(0)$. Typical values of the carrier phase error are shown in Figure 10.102 as a function of steering-vector mismatch, i.e., the difference between the true GPS satellite location (θ_s, ϕ_s) and its estimated position ($\hat{\theta}_s$, $\hat{\phi}_s$). When the steering vector is correctly estimated (mismatch $= 0$) the carrier phase error is a fraction of a degree. Even for rather crude estimates (errors up to 20°) of the steering vector the carrier phase error is only a few degrees. Thus, it appears that the adaptive

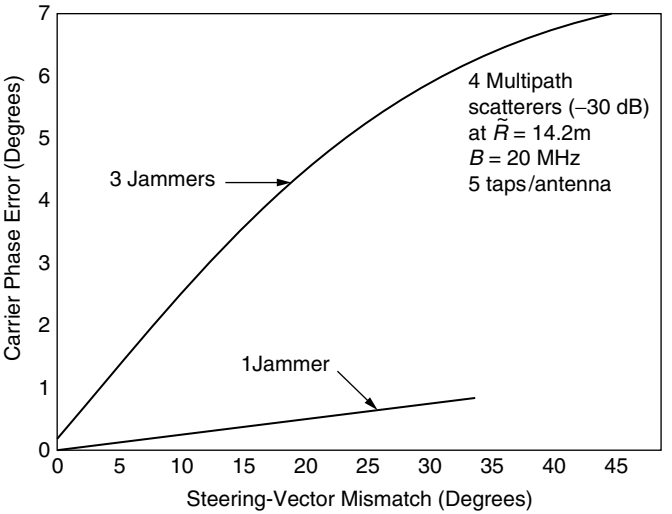

FIGURE 10.102 Steering vector mismatch (degrees).

FIR filter will not have a deleterious effect on differential GPS systems employing carrier phase tracking for high-precision position location.

10.10.6. Space–Time vs. Suboptimum Space-Frequency Processing

We would now like to compare the performance of the null space–time processor with that of the suboptimum space–frequency processor, using the metric of computational complexity. That is, for equal computational complexity, we wish to find which method produces a lower residual interference after adaptation. The number of real operations per second to form and invert the covariance matrix, using an LU factorization and forward-back substitution, for the full space–time processor is[§§§§] 17.33 $(PK)^3 F_u$, where P is the number of time taps and F_u is the weight update rate. The number of operations per second to apply the weights in Equation 10.241 is $(8KP - 2)F_s$, where F_s is the sampling rate, so that the total number of operations per second is the sum of these two quantities.

For suboptimum, space–frequency processing there are 17.33 $(KQ)^3 MF_u$ operations per second to form and invert the covariance matrix, where M is the number of frequency bins and Q is the processing order. For nonoverlapped FFTs we require $(K + 1)(2 + 5 \log_2 M) F_s$ operations per second to apply real weights

[§§§§] This assumes that 4 PK samples are used to estimate the covariance matrix, and that the covariance domain is used. If the QR decomposition is done directly on sampled data then 17.33 $(PK)^3 F_u$ is replaced by 32 $(PK)^3 F_u$.

(e.g., Blackman) to the data block (of M samples), do the FFTs for each antenna and then perform an inverse FFT (IFFT). In order to reduce windowing losses, the FFTs are usually overlapped by at least fifty percent. If the FFTs are overlapped by a factor O_v (i.e., $O_v = 1$ means no overlap, $O_v = 2$ means 50% overlap) then there are $O_v(K+1)\cdot(2+5\log M)F_s$ operations per second to do the FFTs and IFFT. Also, there are $(8QK-2)O_vF_s$ operations per second to apply the weights. Thus, the total number of operations per second is

$$17.33(KQ)^3MF_u + O_vF_s[8KQ - 2 + (K+1)(2+5\log_2 M)]$$

There are now two cases to be considered. In a highly dynamic environment the covariance matrix update rate will dominate the computations count. However, in a less dynamic environment the computation of the FFTs and application of the weights will dominate. Typically, if the weight update rate is less than 1 kHz the weight application will dominate the computations count, so that for space–time processing the number of operations per second is approximately $(8KP-2)F_s$, and for order Q space–frequency processing the number of operations per second[#####] is approximately $O_vF_s[8KQ - 2 + (K+1)\cdot(2+5\times\log_2 M)]$. In the results to follow we assume that the update rate is slower than 1 kHz and that a 50% overlap ($O_v = 2$) is used.

In order to now compare performance versus processing requirements, we consider the scenario where the seven-element circular array is operating over a bandwidth B of 20 MHz and illuminated by zero, one, two and three broadband (bandwidth greater than or equal to 20 MHz) interferers plus multiple narrowband (i.e., bandwidth $\rightarrow$ 0) interferers. Each interferer has an interference-to-noise ratio of 70 dB, and there are four multipath scatterers present with a strength of -30 dB and an average range of 15 m. The narrowband interferers are randomly located in azimuth at elevations close to the horizon, and are also randomly placed within each frequency bin. In Figures 10.103 and 10.104 we show the residual interference after adaptation for the case of sixteen bin, suboptimum, space–frequency processing using a Blackman-Harris window. The cancellation ratio for all cases is 120 dB. The results shown are the averages over three realizations of the random placements of the narrowband interferers within the frequency bins. Also shown is space–time processing (STAP) using four time taps. For each curve, we show the relative processing complexity normalized to the processing complexity for four-tap space–time processing. From Figure 10.103 and 10.104 it is evident that four-tap space–time processing outperforms both order 1 and order 2

[#####] We can also consider memory requirements. A space–time processor with K antennas and P taps requires each processing chain to collect 4 KP samples to estimate the covariance matrix, so that the K antenna system has a memory requirement of $4K^2P$ complex. An order 1 space–frequency processor requires each antenna processing chain to collect 4 KM samples (i.e., 4 K samples per frequency bin) so that the total memory requirement is $4K^2M$. Thus, the space–frequency processor requires M/P more storage than the space–time processor. Thus, memory requirements favor space–time over space–frequency processing. Ultimately, this data requirement will limit adaptation rates for space–frequency processing by the same M/P ratio relative to space–time processing.

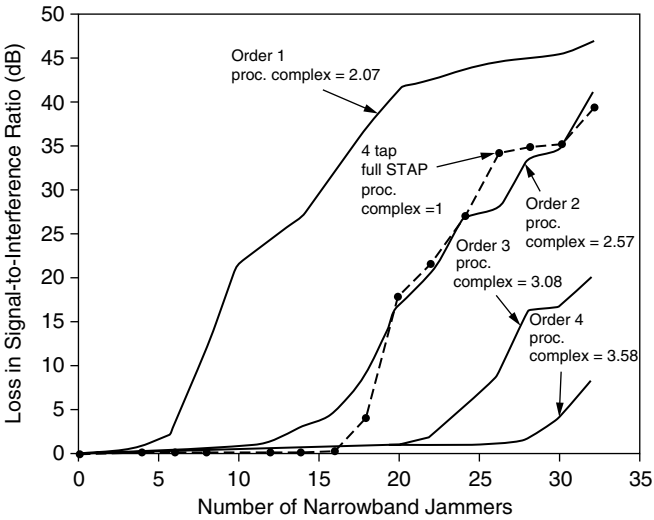

FIGURE 10.103 Comparison of 16 bin space–frequency processing with four-tap space–frequency processing. One wideband interferer.

suboptimum, 16 bin, space–frequency processing and requires less than one half the computations. For example, in Table 10.8 we compare four-tap STAP with 16 bin, order 2, space–frequency processing, using the performance measure that there is less than a 3 dB increase in the interference-plus-noise to noise ratio after

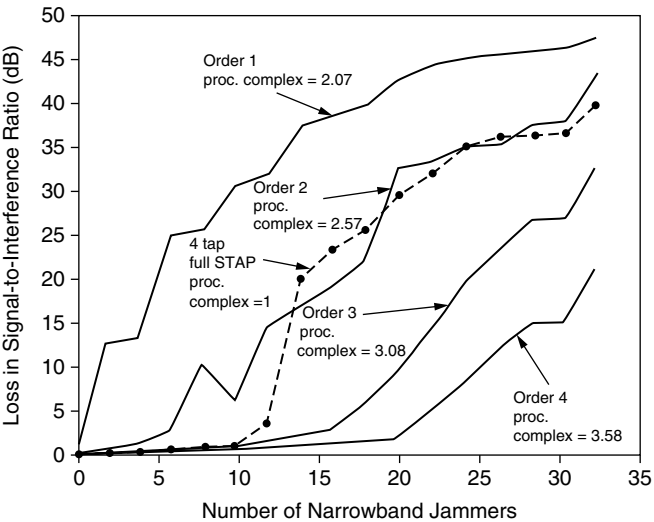

FIGURE 10.104 Comparison of 16 bin space–frequency processing with four-tap space–time processing. Two wideband interferers.

TABLE 10.8
Number of Interferers Canceled without Increasing Noise Floor by More than 3 dB

Processing Type	Relative Processing Complexity	Number of Interferers Cancelled			
Space–Frequency	2.57	0 wideband	1 wideband	2 wideband	3 wideband
16 bin, order 2		23 narrowband	14 narrowband	6 narrowband	1 narrowband
Four-Tap	1	0 wideband	1 wideband	2 wideband	3 wideband
Space–Time		26 narrowband	17 narrowband	12 narrowband	5 narrowband

adaptation. Clearly, space–time is the winner when only 16 frequency bins are used in the space–frequency processor.

The conclusion does not change if a larger FFT is used. A 64 bin, order 1 space–frequency processor requires 2.79 times the number of computations as a four-tap space–time processor, so let us compare these two approaches. First, consider the scenario where the seven-element array is illuminated by three, randomly located, wideband interferers plus a varying number of narrowband interferers. As before, the cancellation ratio is 120 dB. The interference-plus-noise to noise ratio after adaptation for both processors is shown in Figure 10.105, which indicates that the four-tap space–time processor performs as well as the first order 64 bin space–frequency processor with a Blackman-Harris window.

It is important to recognize, however, that the aforementioned conclusions apply to the case when the weight update rate is less than 1 kHz. When the update rate is much faster than 1 kHz the first-order space–frequency processor with 128 or more bins[|||||||] is often superior to the space–time processor (e.g., a 128 point FFT processor gives equal performance using fewer operations per second than the space–time processor when the update rate exceeds 10 kHz).

It is interesting to physically interpret some of the results in Figures 10.103–10.105. Consider the P tap space–time processor. In view of our discussion in Section 10.10.4 that space–time and full-order space–frequency processing are equivalent, we can view this space–time processor as a full order, P bin, space–frequency processor. If the processor has K antennas and P bins, there are $KP - 1$ degrees of freedom available. The desired GPS signal is broadband, spreading across all P frequency bins. Thus, in each bin one spatial

[|||||||] One drawback with large-size FFTs is that the bandwidth of each frequency bin becomes very small (each bin has a bandwidth of B/M where B is the operating bandwidth and M is the FFT size). Thus, in order to estimate the $K \times K$ covariance matrix one requires $4K$ independent samples, which requires a time interval of a $4\ KM/B$ sec. For $K = 7$, $M = 256$, $B = 20$ MHz this requires a time interval of 0.358 msec. Typically, in a dynamic environment the weights will need to be updated once per millisecond, so $M = 256$ is probably an upper limit.

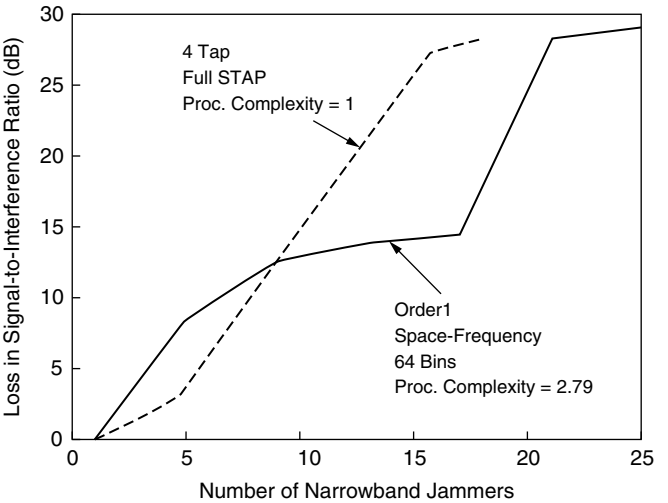

FIGURE 10.105 Comparison of 64 bin space–frequency processing with four-tap space–time processing. Three wideband interferers.

degree of freedom is required to protect the signal. Each broadband interferer also requires one spatial degree of freedom in each bin in order to place a null on that interferer. This leaves a total of

$$N_0 = KP - 1 - P - NP = (K - N - 1)P - 1 \tag{10.279}$$

degrees of freedom to cancel narrowband interferers, where N is the number of broadband interferers. For $K = 7$ antennas, $P = 4$ bins, and $N = 1$ broadband interferer, we see from Equation 10.279 that, in theory, the processor can cancel up to $N_0 = (7 - 2)4 - 1 = 19$ narrowband interferers. From Figure 10.103 we see that, because of frequency overlap and because multipath is also present, one cannot quite cancel 19 narrowband interferers down to the noise floor, but it is possible to cancel 16 or 17 narrowband interferers down to near the noise floor. Likewise, for $N = 2$ broadband interferers, Equation 10.279 predicts that 15 narrowband interferers can also be canceled, whereas the simulation in Figure 10.104 indicates that 11 or 12 are actually canceled down to near the noise floor.

The low-order ($Q = 1, 2$) suboptimum space–frequency processor does not perform as well as hoped because of leakage from bin to bin, even when a Blackman-Harris window (which has extremely low sidelobes) is used. The problem is that, although the window reduces the sidelobes, it spreads the mainlobe frequency response over multiple bins. Consequently, a narrowband interferer in frequency bin m has strong leakage into bins $m \pm 1, m \pm 2$ so that spatial nulls are required in all five bins to null the single narrowband interferer. Thus, the available degrees of freedom are rapidly consumed. Other windows, such as the Hamming window, fare no better. Although, not presented here, we also studied the use of polyphase filters to alleviate this leakage problem, but had limited success.

10.10.7. Summary

We have demonstrated that by using a space–time processor the interference from multiple, strong interferers plus their multipath can be canceled down close to the noise floor without producing serious loss or distortion of a GPS signal. We also found that the processor demonstrates a remarkable capability for self-equalization. An approximate relationship for the number of broadband (N) and narrowband (N_0) interferers that can be canceled is

$$N_0 + NP \simeq (K - 1)P - 1 \tag{10.280}$$

where P is the number of time taps used behind each antenna in Figure 10.94 and K is the number of antennas. The result above assumes that the wideband interferers occupy the entire operating band B and that the narrowband interferers are randomly located across B. The number of time taps must satisfy

$$(P - 1)T > \tau_{\max} \tag{10.281}$$

where T is the time-tap spacing, which must be less than $1/B$, and $\tau_{\max}$ is the maximum differential time delay between the direct interferer ray and the multipath.

Previously, we speculated that suboptimum space–frequency adaptive processing would work nearly as well as space–time processing, but would require fewer computations. We have found that this speculation is untrue unless the weights be updated very rapidly. For moderate update rates the space–time processor outperformed the suboptimum space–frequency processor of equal computational complexity. However, we also studied the case when the weights needed to be updated very rapidly (e.g., once per 100 μs) so that the weight update begins to dominate the computations count. In this case the space–time and space–frequency (with FFT size > 64) processors both gave roughly equal performance for equal computations count. Thus, space–time is not always a clear winner, and if the update rate is sufficiently rapid (e.g., roughly 10 kHz or faster) the first-order space–frequency processor with 128 or 256 bins outperforms the space–time processor.

Although all results presented here employed ideal covariance matrices, we also performed some simulations using estimated covariance matrices (using four times the number of degrees of freedom samples of the voltages for the estimate) and obtained essentially the same results.

10.11. A MAXIMUM-LIKELIHOOD BEAMSPACE PROCESSOR FOR IMPROVED SEARCH AND TRACK

(R. M. DAVIS AND R. L. FANTE)

10.11.1. Introduction

MODERN radars utilize phased-array antennas and analog beamformers. The search function is performed using a single beam to interrogate a grid of spatial

positions. If a target is detected, its angle of arrival (AOA) is measured using two orthogonal difference beams. Simultaneous ratioing of the difference-beam voltages to the sum-beam voltage provides an instantaneous estimate of target location. The process is referred to as monopulse.

The technology currently exists to do digital beamforming (DBF). Direct RF digitization at the element level is possible in low-frequency systems. At higher frequencies analog beamforming can be used to form subarrays whose outputs can be downconverted and digitized. The authors believe that with DBF comes the possibility of using multiple highly overlapped receive beams together with maximum-likelihood (ML) processing to improve both search and track performance. We have developed a maximum-likelihood beamspace processor (MLBP) that implements the new approach and is the subject of this paper. The processor and its advantages are described in Section 10.11.2. The processor is analyzed in Section 10.11.3 and numerical results are presented in Section 10.11.4. The work is summarized in Section 10.11.5.

10.11.2. Maximum-Likelihood Beamspace Processor (MLBP)

The MLBP architecture (Figure 10.106) entails dividing the antenna aperture, which may consist of a large array containing thousands of elements, into a small number of subarrays (M), and digitizing their outputs. A typical value of M would be 16. The digitized signals are then passed through two stages of processing. In the first stage, the subarray voltages are multiplied by N sets of complex weights and summed to form N highly overlapped beams. A typical value of N would be four. The centers of the beams are all located within the 3 dB contour of the transmit beam. Their exact location is determined *a priori*. Interference is removed from each of the beams by choosing the subarray weights to minimize the output mean square power subject to a beam-pointing constraint, or maximize the output signal-to-interference ratio (SIR). Our procedure is a variation of an approach discussed by Baranowski and Ward.[1] In particular, our first stage beams are formed adaptively and take the form of highly overlapped, fixed sum beams. Our choice of

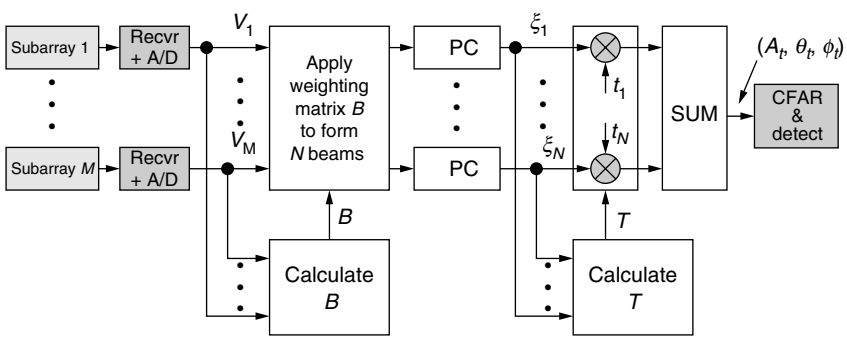

FIGURE 10.106 Maximum-likelihood beamspace processor (MLBP).

architecture was driven by our desire to both remove interference and eliminate beam shape loss. The first level of processing is equivalent to what has historically been referred to as a fully-adaptive array of subarrays. Applebaum was one of the first to demonstrate the virtue of the adaptive array architecture.[2] Many techniques exist for choosing the adaptive weights used to form the beams. The weights are chosen in such a way as to constrain the gain in the pointing direction of each of the N beams. The constraint preserves the gain in the pointing direction and reduces beam shape distortion. Pulse compression (PC) is done after the first stage processing to prevent the first stage adaptive arrays from attempting to cancel target returns. Target (often referred to as signal) cancellation can be prevented if the signal to noise ratio in each of the beams is much less than unity prior to adaptive nulling (see Appendix S for additional discussion). Performing pulse compression at the beam level also reduces computational complexity (only four channels must be compressed versus 16 had we done it at the subarray level).

During the second stage of processing, the N beams are themselves adaptively weighted and combined to form a single output beam. The beam weights in the second stage are chosen to maximize the probability of detecting a target return. The processing used to choose the second set of weights is referred to as maximum likelihood (ML). Kelly, Reed, and Root[3] were among the first to demonstrate the virtues of ML. The ML process entails performing a search over all possible AOA directions of the target return and picking the direction that yields a set of beam weights with the highest probability of declaring a target to be present. The search can be constrained to a region slightly larger than the 3 dB contour of the transmit beam. An independent search must be performed in every range cell within the range sweep. Thus, there will be a separate set of optimum beam weights for each range cell. We assume, herein, that the target geometry and range resolution of the radar is such that there will only be one target per cell.

The authors will show that under the assumption of Gaussian interference, the second stage processing is equivalent to operating a fully adaptive array of beams where the weights are chosen to maximize SIR in the composite output beam. The observation emphasizes the commonality of the ML and adaptive array processing paradigms.

Two impediments to performing the maximum-likelihood search in the past have been the computational burden and the possibility of getting stuck in local minima. Advances in high speed computing, together with the idea of limiting the maximum-likelihood search to only a few beams, make the computation problem tractable. Performing the search in beam space with highly overlapped beams also appears to reduce the probability of getting stuck in local minima. The MLBP architecture offers many advantages, notably:

(1) The ML search eliminates the need to form azimuth and elevation difference beams to measure target AOA, thereby, reducing the required RF beamforming, receiver, and analog-to-digital conversion hardware by up to a factor of three.

(2) Using multiple highly overlapped beams on receive, instead of a single beam in conjunction with ML, nearly eliminates beam shape loss on receive and increases the volume of space, which can be searched for a given number of transmissions. The MLBP can search a given volume of space using approximately 40% fewer pulses compared to a traditional single-beam system. This provides the radar with additional time and energy to perform other functions.
(3) The MLBP architecture supports improved angle accuracy compared to monopulse. The improvement varies from a few percent when the target is at the peak of the transmit beam to approximately 30% when the target is at the 3 dB (half power) point on the transmit beam for a system using four highly overlapped beams.
(4) The MLBP supports angle estimation over a larger angular volume than monopulse. Monopulse systems can usually only be relied upon when the target is within the 3 dB contour. The MLBP, however, degrades gracefully as the target moves outside the 3 dB contour and down the skirt.
(5) The fully adaptive array of subarrays followed by the fully adaptive array of beams supports mainlobe and sidelobe interference suppression while minimizing beam shape distortion and signal cancellation.

A few comments regarding the first advantage are in order. The antenna designer, while wanting to do digital beamforming and monopulse processing, is faced with a difficult decision. The problem is how to simultaneously apply aperture weighting (tapering) to sum and difference channels to achieve low sidelobes. One option is to form a large number of small subarrays, digitize their output voltages, and then split each one three ways and apply the appropriate weighting (Taylor or Bayliss) digitally. The large number of small subarrays is required to appropriately sample the tapering function. The first option requires a large number of receivers and analog-to-digital converters (ADCs). The second option is to split the output of each individual element three ways, apply tapering near the array face, and duplicate the subarray beamforming three times. The second option supports digitization of a smaller number of larger subarrays, but trades the smaller number of receivers and ADCs for additional RF beamforming stages (to form the larger subarrays), weighting circuitry, and possibly analog time delay units to support wideband modes. A small number of large subarrays can be used since there is no sampling problem associated with applying the taper weights. The MLBP supports the best of both options. Since we do not form difference beams, we can apply taper weights to the elements within each subarray and use a small number of receivers and ADCs by digitizing a small number of large subarrays. Although we do not consider wideband operation in this paper, we expect that fewer time delay steering units may be required with the MLBP compared to the second option, because the multiple overlapped beams cover a larger volume of space than a single beam.

10.11.3. Analysis

The MLBP depicted in Figure 10.106 uses two stages of sequential processing on receive. In the first stage, the subarray voltages are adaptively weighted and summed to form N highly overlapped beams. The weights are chosen to maximize SIR in each beam separately. The second stage entails adaptively weighting and summing the beam voltages to form a single composite output beam. A search process is used to find the adaptive beam weights that maximize the probability of a target being present in the output beam. Interference that is not cancelled during the first stage of processing, but is correlated beam to beam, can be nulled in the second stage. In the analysis to follow the narrowband approximation (signal bandwidth $\ll$ carrier frequency) is implicit.

10.11.3.1. The First Stage

Suppose we are given the voltages out of M subarrays and we wish to form N interference-free beams centered on the directions $(\theta_1, \phi_1), (\theta_2, \phi_2), \ldots, (\theta_N, \phi_N)$, in spherical coordinate system, where θ is the polar angle and ϕ is the azimuthal angle. If $\mathbf{v}(k) = [v_1(k), v_2(k), \ldots v_M(k)]^{\mathrm{T}}$ is a vector of the transposed voltages (the superscript T denotes transpose) at the outputs of the subarrays at discrete time k, we wish to multiply this set of voltages by an $N \times M$ beamforming matrix

$$\mathbf{B} = \begin{bmatrix} \mathbf{w}_1^H \\ \mathbf{w}_2^H \\ \vdots \\ \mathbf{w}_N^H \end{bmatrix} \tag{10.282}$$

that will remove interference while simultaneously preserving the gain in the N beam pointing directions. It is well known[2,6] that the weight vector $\mathbf{w}_n = [w_{n1} \cdots w_{nM}]^{\mathrm{T}}$ given by

$$\mathbf{w}_n = \frac{\mathbf{R}^{-1}\mathbf{S}_n}{\mathbf{S}_n^H \mathbf{R}^{-1}\mathbf{S}_n} \tag{10.283}$$

maximizes SIR in the beam pointing direction (θ_n, ϕ_n), thus, producing the desired result. In Equation 10.283, $\mathbf{R}$ is the $M \times M$ covariance matrix of subarray voltages defined as $\mathbf{R} = \langle \mathbf{v}\mathbf{v}^H \rangle$, where $\langle\rangle$ denotes expectation. $\mathbf{R}$ can be approximated as

$$\mathbf{R} = \frac{1}{K} \sum_{k=1}^{K} \mathbf{v}(k)\mathbf{v}^H(k) \tag{10.284}$$

where $K \geq 4M$ and the superscript H denotes Hermitian (conjugate transpose[7,8]). Also $\mathbf{S}_n = [s_{n1}, s_{n2}, \cdots, s_{nM}]^{\mathrm{T}}$ is a $M \times 1$ vector that defines the steering direction

(θ_n, ϕ_n) of the nth beam. The numerator of the weight equation is often referred to as a Wiener whitening filter, while the denominator implements a point constraint by scaling the adaptive pattern to provide the desired gain in the beam steering direction. A typical component of $\mathbf{S}_n$ is

$$s_{nm} = g_m(\theta_n, \phi_n)\exp[jk(x_m \sin\theta_n\cos\phi_n + y_m \sin\theta_n\sin\phi_n + z_m \cos\theta_n)] \tag{10.285}$$

where $g_m(\theta_n, \phi_n)$ is the voltage gain of the mth subarray in the direction (θ_n, ϕ_n), k equals $2\pi/\lambda$, λ is the wavelength, and (x_m, y_m, z_m) is the location of the center of the mth subarray in a Cartesian coordinate system. The voltages in each of the output beams (see Figure 10.106) from the first stage of processing are denoted by $\xi_1, \xi_2, \ldots\xi_N$. If we define $\boldsymbol{\xi} = |\xi_1, \xi_2, \ldots\xi_N|^{\mathrm{T}}$, then the first stage beamformer output voltage vector is

$$\boldsymbol{\xi} = \mathbf{Bv} \tag{10.286}$$

10.11.3.2. The Second Stage

Once the beams have been formed and most of the interference removed, PC is performed on the beam voltages. Since there will be fewer beams than subarrays, performing PC at the beam level is computationally efficient. As previously noted, performing PC at the beam level, rather than the subarray level, also ensures that the first stage adaptive array will not attempt to cancel target returns. Prior to pulse compression, target returns in the subarrays will be well below thermal noise and the adaptive array will, to first order, ignore them (see Appendix S). There may be some applications involving high pulse repetition frequency (PRF) radars where the presence of clutter will force PC to be performed at the subarray level. If a clutter-free region does not exist, then PC followed by Doppler processing may have to be performed at the subarray level in order to generate clutter-free Doppler cells to sample the interference.

Although, the amplitude and direction of a target return is unknown, we can postulate a unit-amplitude target in range bin n in direction (θ, ϕ). This produces an $M \times 1$ voltage vector $\mathbf{h} = [h_1 h_2 \cdots h_M]^{\mathrm{T}}$ at the output of the subarrays, where

$$h_m = g_m(\theta, \phi)\exp[jk(x_m \sin\theta\cos\phi + y_m \sin\theta\sin\phi + z_m \cos\theta)] \tag{10.287}$$

This hypothesis also produces a voltage vector $\mathbf{G} = \mathbf{Bh}$ in beamspace. If the target is nonfluctuating[####] and has unknown amplitude and phase denoted by the complex quantity A, the voltage vector in beamspace produced by the postulated

[####] For a fluctuating target we must estimate $A_1, A_2 \ldots A_K$, in addition, to (θ, ϕ), where A_k is the amplitude at time sample k. The probability function to be maximized becomes $p(\xi_1 \ldots \xi_K | A_1 \ldots A_K, \theta, \phi)$. The analysis is more difficult, but straightforward provided the time samples are sufficiently spaced to make the noise samples independent.

target is

$$\boldsymbol{\xi}_{\mathbf{t}} = A\mathbf{B}\mathbf{h} = A\mathbf{G} \tag{10.288}$$

If we let the vector $\boldsymbol{\rho}$ denote the component of $\boldsymbol{\xi}$ containing uncancelled interference and thermal noise in the second stage beams, then the first stage output vector $\boldsymbol{\xi}$ can be written as

$$\boldsymbol{\xi} = \boldsymbol{\rho} + \boldsymbol{\xi}_{\mathbf{t}} = \boldsymbol{\rho} + A\mathbf{G} \tag{10.289}$$

Assuming that $\boldsymbol{\rho}$ is a circular, normal random vector, it can be shown[4,5,9] that the probability density function for $\boldsymbol{\rho}$ is

$$p(\boldsymbol{\rho}) = \pi^{-N}[\det(\mathbf{C})]^{-1} \exp(-\boldsymbol{\rho}^H \mathbf{C}^{-1}\boldsymbol{\rho}) \tag{10.290}$$

where $\mathbf{C}$ is the $N \times N$ covariance matrix of beam voltages defined by

$$\mathbf{C} = \langle \boldsymbol{\rho}\boldsymbol{\rho}^H \rangle = \mathbf{B}\mathbf{R}\mathbf{B}^{\mathrm{H}} \tag{10.291}$$

Solving Equation 10.289 for $\boldsymbol{\rho}$ and substituting into Equation 10.290 gives the conditional probability density

$$p(\xi|A, \theta_t, \phi_t) = \pi^{-N}[\det(\mathbf{C})]^{-1} \cdot \exp[-(\boldsymbol{\xi} - A\mathbf{G})^{\mathrm{H}} \mathbf{C}^{-1}(\boldsymbol{\xi} - A\mathbf{G})] \tag{10.292}$$

The maximum-likelihood target parameter estimator $(\hat{A},\ \hat{\theta}_t,\ \hat{\phi}_t)$ occurs when Equation 10.292 is a maximum or when

$$\Gamma = (\boldsymbol{\xi} - A\mathbf{G})^{\mathrm{H}} \mathbf{C}^{-1}(\boldsymbol{\xi} - A\mathbf{G}) \tag{10.293}$$

is a minimum. We can solve for the estimator $\hat{A}$, by differentiating Γ with respect to A and setting the result to zero. Performing the differentiation we find

$$\hat{A} = \mathbf{T}^{\mathrm{H}} \xi \tag{10.294}$$

where the beam weighting vector T is given by

$$\mathbf{T} = \frac{\mathbf{C}^{-1}\mathrm{G}}{\mathbf{G}^{\mathrm{H}}\mathbf{C}^{-1}\mathbf{G}} \tag{10.295}$$

We can show that the estimator given by Equation 10.294 is unbiased by substituting Equation 10.289 into Equation 10.294 and taking the expected value. Since $\langle \boldsymbol{\rho} \rangle = 0$, then $\langle \hat{A} \rangle = A$.

At this point in the analysis we make an important observation. $\mathbf{T}$ is functionally identical to the weight vector that maximizes SIR in the output beam of a fully adaptive array. Let us compare $\mathbf{T}$ in Equation 10.295 with $\mathbf{w}_n$ in Equation 10.283. In Equation 10.295 the covariance matrix $\mathbf{C}$ of beam voltages replaces the covariance matrix R of subarray voltages in Equation 10.283 and $\mathbf{G}$, the steering vector that coheres the beam voltages of form a composite beam in the direction of the target AOA, replaces $\mathbf{S}_n$, the steering vector that coheres the subarray voltages in the predetermined direction of the nth beam. The numerator of $\mathbf{T}$ is again a Weiner whitening filter where $\mathbf{C}^{-1/2}\boldsymbol{\xi}$ is the whitened data and

$\mathbf{C}^{-1/2}\mathbf{G}$ is the whitened steering vector in beamspace. The denominator is a scalar that acts to constrain the gain of the composite beam in the target direction. Based upon this observation, we could stop the analysis at this point and simply search for the target AOA that maximizes SIR in the output beam. Instead, we will continue the analysis to more firmly establish the equivalency between adaptive array and maximum-likelihood theory.

Substituting Equation 10.294 back into Equation 10.293 gives

$$\Gamma = \boldsymbol{\xi}^{\mathrm{H}}\mathbf{C}^{-1}\boldsymbol{\xi} - (\mathbf{G}^{\mathrm{H}}\mathbf{C}^{-1}\mathbf{G})|\mathbf{T}^{\mathrm{H}}\boldsymbol{\xi}|^2 \tag{10.296}$$

Because $\boldsymbol{\xi}^{\mathrm{H}}\mathbf{C}^{-1}\boldsymbol{\xi}, \mathbf{G}^{\mathrm{H}}\mathbf{C}^{-1}\mathbf{G}$ and $|\mathbf{T}^{\mathrm{H}}\boldsymbol{\xi}|^2$ are positive definite quantities, it is evident that Γ is minimized when $(\mathbf{G}^{\mathrm{H}}\mathbf{C}^{-1}\mathbf{G})|\mathbf{T}^{\mathrm{H}}\boldsymbol{\xi}|^2$ is maximized. That is, the estimate of target location is that direction (θ, ϕ) for which $(\mathbf{G}^{\mathrm{H}}\mathbf{C}^{-1}\mathbf{G})|\mathbf{T}^{\mathrm{H}}\boldsymbol{\xi}|^2$ is a maximum. Let us now discuss the physical meaning of this quantity.

Referring to Figure 10.106 we see that the voltage in the composite output beam after weighting is $\mathbf{T}^{\mathrm{H}}\boldsymbol{\xi}$. In a range bin where a target is present $\mathbf{T}^{H}\boldsymbol{\xi}$ yields an estimate of the target amplitude $\hat{A}$ (from Equation 10.294). However, in range bins where there is no target the mean output interference power is

$$I = \langle|\mathbf{T}^{\mathrm{H}}\boldsymbol{\rho}|^2\rangle = \mathbf{T}\mathbf{C}\mathbf{T}^{\mathrm{H}} = \frac{1}{\mathbf{G}^{\mathrm{H}}\mathbf{C}^{-1}\mathbf{G}} \tag{10.297}$$

Taking the ratio of the power in the range bin with the target present to the power in a bin where there is no target yields an estimate of SIR in the output beam. This is

$$\mathrm{SIR}_{\mathrm{est}} = (\mathbf{G}^{\mathrm{H}}\mathbf{C}^{-1}\mathbf{G})|\mathbf{T}^{\mathrm{H}}\boldsymbol{\xi}|^2 \tag{10.298}$$

It should be noted that Equation 10.298 is actually the ratio of signal-plus-interference power to interference power, but we identify it as SIR because we assume that the tandem adaptive arrays effectively null the interference, so that the signal is much larger than the residual uncancelled interference in the output beam. We can now see that maximizing $(\mathbf{G}^{\mathrm{H}}\mathbf{C}^{-1}\mathbf{G})|\mathbf{T}^{\mathrm{H}}\boldsymbol{\xi}|^2$ is equivalent to searching over (θ, ϕ) to find the value $(\hat{\theta}, \hat{\phi})$ that maximizes SIR in the output beam. The value of SIR at $(\hat{\theta}, \hat{\phi})$ is readily calculated to be $\boldsymbol{\xi}^{\mathrm{H}}\mathbf{C}^{-1}\boldsymbol{\xi}$. After estimating the target AOA, $(\hat{\theta}, \hat{\phi})$, the amplitude and phase of that target can be estimated using Equation 10.294

$$\hat{A} = \mathbf{T}^{\mathrm{H}}(\hat{\theta}, \hat{\phi})\boldsymbol{\xi} \tag{10.299}$$

Although Equation 10.298 looks mathematically formidable, we show in Appendix T that it reduces to a familiar result in the absence of any interferers.

10.11.3.3. Target Detection

Target detection is achieved by comparing the value of $\hat{A}$ in each range cell with a threshold τ determined using conventional constant false alarm rate (CFAR) processing. If $\hat{A} \geq \tau$, a "target" is declared and if $\hat{\mathrm{A}} < \tau$, "no target" is declared.

10.11.4. Numerical Examples

A MATLAB computer simulation was developed to verify the analysis and predict performance of the MLBP. A 1024 element square array divided into 16 square subarrays was used in the simulation. The elements were spaced $\lambda/2$ apart and the array was uniformly weighted on both transmit and receive. The phase shifters within each subarray were set to steer the subarrays to the pointing direction of the transmit beam. Target returns, interferers, and thermal noise were modeled as tones (zero bandwidth). Thermal noise voltages were complex with their real and imaginary components chosen from a unit variance, zero mean, Gaussian distribution. Target returns were assumed to be nonfluctuating. Angle errors were normalized to the 3 dB boresight beamwidth (θ_3) equal to 0.886 times λ/D, where the diameter of the array D was 16λ.

The number of beams in the second stage of the processor was varied from two to eight. The beams were offset from the subarray steering direction (u_0, t_0) by some fraction of θ_3. The direction cosines with the x- and y-axis in the plane of the array are defined in Figure 10.107. Figure 10.107 depicts a typical beam geometry for the baseline case of four-beams. MATLAB function "FMINS" with an error bound of 10^{-4} was used to perform the search over all possible target AOA (θ, ϕ). FMINS uses the Nelder-Mead simplex direct search method. Nelder-Mead is an excellent engineering tool, but is inefficient. The mean and variance of the final error were found to be relatively insensitive to both the starting point of the search and the error used to determine the stopping point.

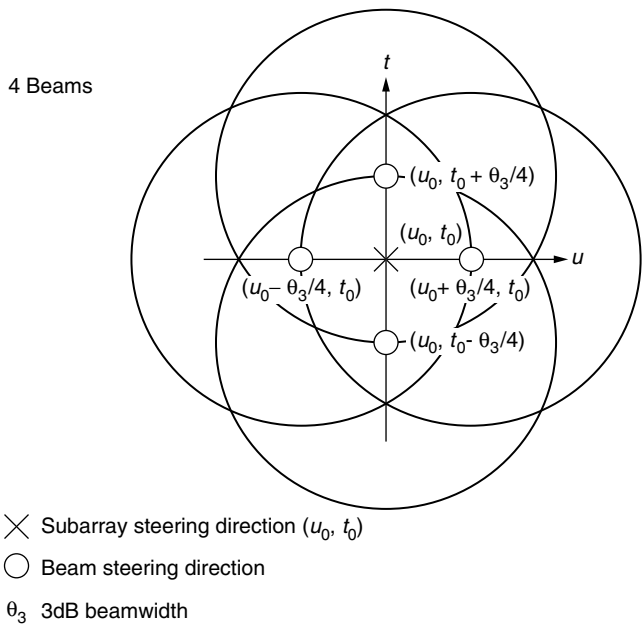

FIGURE 10.107 Beam geometry in direction cosine space.

Since Nelder-Mead is not a gradient-based technique, it may be less likely to get stuck in a local minima than other approaches. We did not notice a problem with local minima in any of the studied cases. Studied cases included one and two mainlobe interferers. We conjecture that the use of highly overlapped beams, each having a constraint in the pointing direction, limits beamshape distortion and the presence of multiple deep nulls.

The number of floating point operations per second (flops) required to implement the MLBP depends upon the algorithm used to search for the AOA. Even though Nelder-Mead is inefficient, we calculated the throughput using it to obtain a benchmark. The number of nonrecurring floating point operations per beam is due mainly to formation and inversion of the sample covariance matrix of subarray voltages and to beamforming (multiplying subarray voltages by the beam weights for each range cell). The recurring computation per beam is equal to the number of operations needed to calculate SIR (see Equation 10.298) for each hypothesized AOA, times the number of hypothesized AOAs per range cell, times the number of range cells. Assuming 500 range cells, 16 subarrays, and 4 beams, it was found that 37 hypothesized AOAs at 1.7 K flops per iteration were typically needed. The total nonrecurring flop count was roughly 200 K flops per beam compared to 31.5 megaflops per beam for the recurring. If the radar puts out 100 beams per second the total throughput (100 times the sum of nonrecurring plus recurring flops) would be 3.1 gigaflops. Note that the flop count is dominated by the recurring count and other search algorithms exist that are more efficient than Nelder-Mead (see, for example, Ref. 10).

10.11.4.1. Improved Clear Environment Search Performance

Figure 10.108 shows computer generated comparisons of the two-way (transmit plus receive) loss in signal-to-noise ratio (SNR) between a conventional

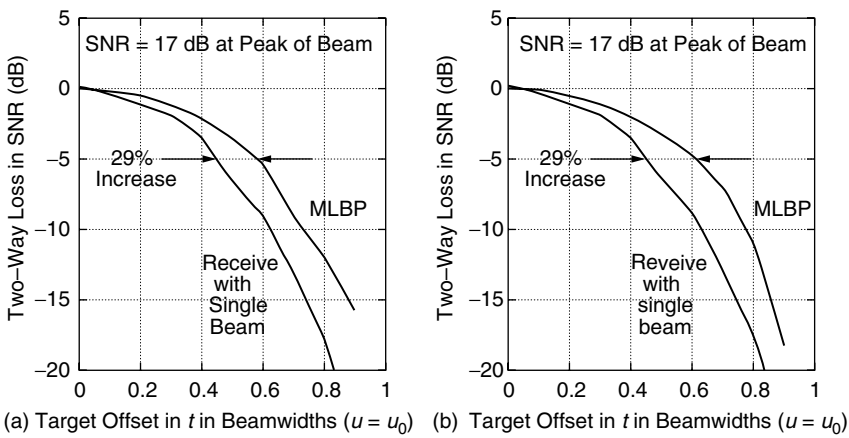

FIGURE 10.108 Comparison of two-way beam shape loss between MLBP and conventional receiver. (a) Three beams. (b) Eight beams.

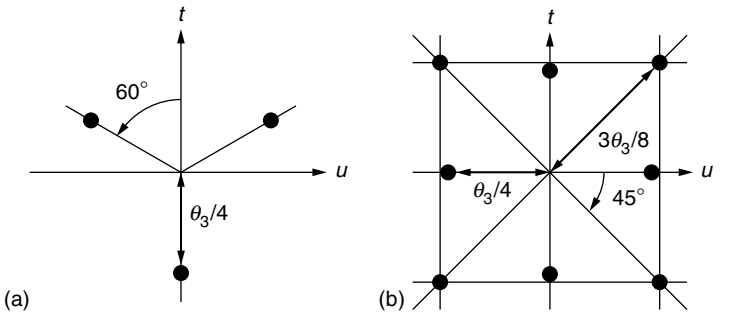

FIGURE 10.109 Beam geometry in direction cosine space. (a) Three beams. (b) Eight beams.

architecture that uses a single receive beam and MLBP architectures using three and eight receive beams. The figure was generated by moving a single target return out from the peak of the transmit beam (also equal to the subarray steering direction (u_0, t_0) along the t-axis ($u_t = u_{target} = u_0$). No interference except receiver noise was present. The curves represent an average of 200 Monte Carlo runs using different receiver noise realizations on the subarray voltages. The figure shows that if search beams are laid down in a line and overlapped at the 5 dB (two-way) loss point, the coverage in one-dimension can be increased by about 29 to 35% as the number of receive beams is increased from three to eight using the MLBP architecture compared to the conventional approach. The improvement is due to the elimination of receive beamshape loss. Figure 10.109 shows beam geometry for the three and eight beam cases.

10.11.4.2. Improved Clear Environment Angle Estimation

Figure 10.110 presents a comparison of the angle-estimation capability of a conventional-monopulse processor with that of the MLBP for the three and eight beam architectures. The figure was generated by moving a single target return out from the peak of the transmit beam (also equal to the subarray steering direction (u_0, t_0) along the t-axis ($u_t = u_{target} = u_0$). The transmit loss was taken into account in the processing and the curves represent an average of 200 Monte Carlo realizations with different receiver noises. The results are compared to a normalized monopulse error (σ) in a thermal noise environment. The one-sigma error was assumed to equal the 3 dB beamwidth divided by 1.6 times the square root of twice the rms SNR.[11]

Figure 10.110 shows that the performances, of both the conventional (monopulse) processor and the MLBP, approach the monopulse error for high SNR (target near subarray pointing direction). At low SNR (as the target moves off the peak and down the skirt of the transmit beam), the MLBP produces significantly smaller one-sigma angle errors than the monopulse processor. The improvement when the target is located at the 3 dB point, for example, varied

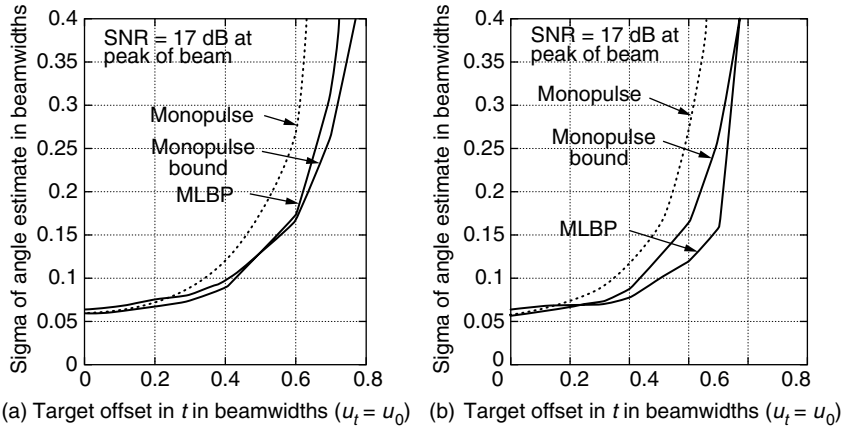

FIGURE 10.110 Comparison of angle accuracy between MLBP and conventional monopulse receiver in clear environment. (a) Three beams. (b) Eight beams.

from about 16% for the three-beam architecture to 33% for the eight-beam architecture. Using more than eight beams did not materially improve performance over the eight-beam system. In fact, a four-beam system did almost as well as the eight-beam system in the clear environment. A four-beam MLBP system was interfaced to a six-state Kalman filter tracking simulation and the resulting track errors were compared with a monopulse system that used the same Kalman-filter tracker.[12] The performance of the two systems (results not shown) was nearly the same for on-axis tracking, but the track errors of the MLBP system were nominally 50% lower than those of the monopulse system for off-axis tracking. In an operational radar, the beam will normally be pointed at one of the largest objects[†††††] in the beam and most targets will be tracked off-axis. It is also noteworthy that while monopulse systems can fail catastrophically as the target return approaches the sum beam null, the MLBP should degrade gracefully and angle measurements made outside the 3 dB transmit contour should be reliable. In summary, the improvement in tracking performance compared to monopulse is due mainly to the fact that the ML search for the target AOA results in the composite output beam being pointed directly at each tracked target.

10.11.4.3. Performance against a Single Mainlobe Interferer

Figure 10.111 shows typical computer simulated performance of a four-beam MLBP operating against a single mainlobe interferer. A target was located a distance of $\theta_3/8$ from the subarray steering direction along a cut that made a 30° angle with the t-axis. A single mainlobe interferer was stepped from the

[†††††] For example, when tracking a reentering missile complex, the radar might track the booster (tank) and the separating reentry vehicles would be tracked off axis.

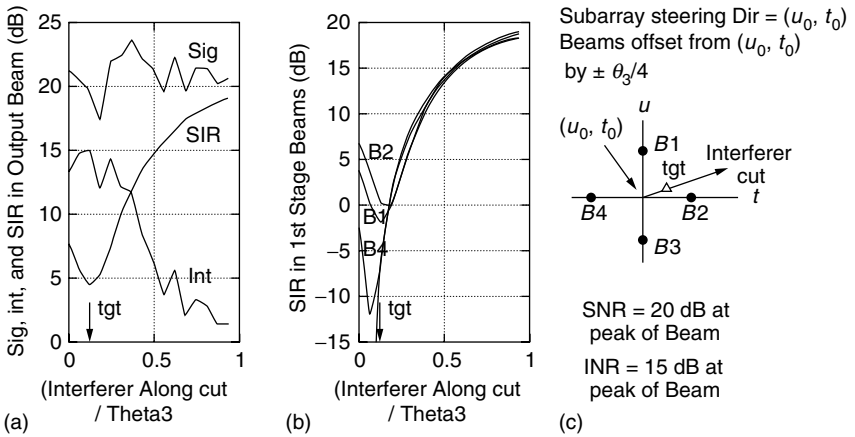

FIGURE 10.111 Four-beam MLRP operation against a single-mainlobe interferer. (a) Signal, interference, SIR. (b) Signal-to-interference ratio (SIR). (c) Four beam geometry.

subarray-steering direction along the cut, through the target's location, and out to a distance of one beamwidth. The target SNR and the INR (interferer-to-noise ratio) were set equal to 20 and 15 dB, respectively, when received at the peak of a beam pointed in the subarray steering direction. A nonfluctuating jammer waveform (constant amplitude and phase) was used when generating the figure. Although, the figure represents only a single realization of receiver noise, nearly identical SIR performance (not shown) was obtained using a complex stochastic waveform and averaging the signal and interference over 100 Monte Carlo runs. The left insert in Figure 10.111 plots the signal, interference, and SIR in the output beam, while the right insert shows SIR in each of the four first stage beams. Note that the SIR in the composite beam in Figure 10.111 is usually larger than any of the SIRs in the four first stage beams. The observation emphasizes the fact that the second-stage processing does not merely pick the first-stage beam having the highest SIR ratio. The first-stage beams are adaptively weighted and summed to form a composite adapted output beam.

A significant improvement in SIR was observed in all computer runs involving a single interferer except when it was very close to the target return—and even then the degradation in SIR was always less than 2 dB of the value in the absence of any adaptivity (which was 5 dB in Figure 10.111). The degradation observed for small target-interferer separations was able to be eliminated by adding more beams. One might have expected more degradation than observed when the interferer approached the target. We hypothesized that the gain constraints placed in the pointing direction of each of the first-stage beams were limiting their ability to null the interferer. This prompted us to examine more closely why the fully adaptive arrays of subarrays operating on the first-stage beams did not result in more interferer nulling and more nulling of the target return along with it.

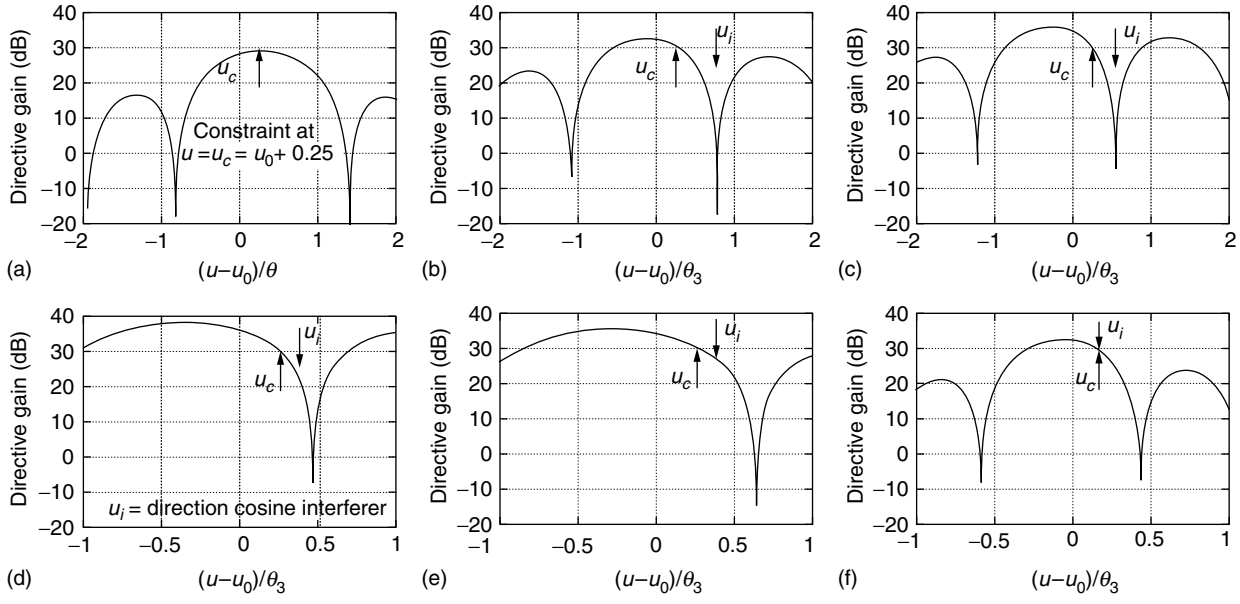

FIGURE 10.112 Adapted antenna patterns vs. direction cosine separation $u_i - u_c$ of interferer from constraint point (INR = 15 dB at peak, $\theta_3 = 3$ dB beamwidth). (a) No interferer. (b) $u_i - u_c = \theta_3/2$. (c) $u_i - u_c = \theta_3/4$. (d) $u_i - u_c = \theta_3/8$. (e) $u_i - u_c = \theta_3/32$. (f) $u_i - u_c = \theta_3/64$.

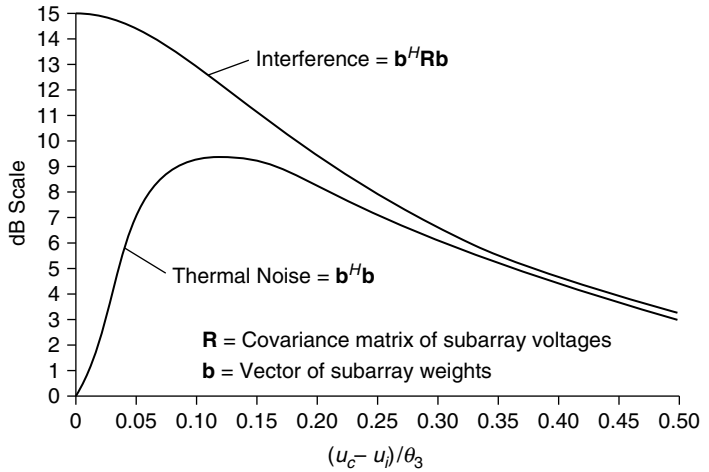

FIGURE 10.113 Total interference and thermal noise in an adapted beam vs. direction cosine separation $u_i - u_C$ of interferer from constraint point ($\theta_3 = 3$ dB beamwidth).

Figure 10.112 shows adapted antenna patterns for the first stage beam that is pointed to $(u_0 + \theta_3/4t_0)$ as a single interferer located at u_i is moved toward the peak of the beam along the u-axis. The location of the gain constraint is denoted by u_c. The interference-to-noise ratio (INR) was set to 15 dB at the peak of the beam and the fluctuation noise was removed from the interferer. Each insert represents a single realization (no Monte Carloing). The inserts demonstrate that when the interferer is brought closer than about $\theta_3/4$ to the constraint point, the array does not have enough resolution to maintain the constraint and place the null on the interferer. In fact, as the separation between the interferer and the constraint point is changed from $\theta_3/8$ to zero, the null is observed to move back out to where it was in the absence of the interferer. Thus, as anticipated, we find that the point constraints placed at the peak of the first-stage beams prevent significant interferer nulling over a relatively large area.

Although the patterns in Figure 10.112 can be used to estimate the exact amount of interferer and target nulling, they cannot be used to determine the overall interference after nulling. The total interference is due to the sum of the uncancelled interferer and the thermal noise in the beam. As the interferer approaches the constraint point, the adaptive subarray weights used to form the first-stage beam become large. Multiplication of the weights by the thermal noise causes an increase in the noise figure in the beam. Figure 10.113 plots the norm of the subarray weights (the thermal noise) and the total interference in the first-stage beam versus separation of the interferer from the constraint point. Note that the weights increase until the separation reaches about $\theta_3/8$ and then decrease as the separation goes from about $\theta_3/8$ to zero. Increasing INR will allow the null to move slightly closer to the constraint point before reversing direction.

The reversal point appears to be the point at which the norm of the weights (and the output noise figure) is a maximum.

10.11.5. Summary

The authors have described a maximum-likelihood beamspace processor for use with an antenna architecture that performs digitization at the subarray level. The MLBP uses two stages of processing. The first stage entails operating N fully adaptive arrays of subarrays in parallel to form N highly overlapped beams. The second stage entails operating a single fully adaptive array of beams to form a single composite output beam pointed in the direction of the target return. The pointing direction of the output beam is determined by performing a search over a constrained region of possible target AOAs and declaring the angular location that produces the highest SIR in the output beam to be the target location. The processing must be repeated for each range cell. A typical system might use 16 subarrays and four-output beams. The architecture supports improved radar performance by eliminating beamshape loss on receive and enabling on-axis tracking of all targets.

Simulation results predict that a four-beam MLBP will be able to search roughly 70% more volume for the same number of transmissions compared to a conventional approach using a single receive beam. The improvement is due to the elimination of receive-beamshape loss. The improvement in angle accuracy compared to monopulse varied from a few percent for a target located at the peak of the transmit beam, to roughly 30% for a target located at the 3 dB point for a four-beam MLBP system. The improvement in angle accuracy translated into roughly a 50% reduction in Kalman filter tracking errors for a target whose track averaged about a third of a beamwidth off-axis.

The new architecture supports nulling of both mainlobe and sidelobe interference. Most nulling occurs in the first stage processing, but residual—interference that is correlated beam to beam—can be further nulled in the second stage. Significant improvement in SIR was observed using a four-beam MLBP operating against a single-mainlobe interferer except for very small target-interferer separations. The small degradation observed for small separations, compared to using no adaptivity, was able to be eliminated by adding more beams. The fact that more degradation was not observed for small separations was shown to be due to the reason that the gain constraints placed in the first-stage beams preclude interferer nulling over a relatively large region.

We believe that future radars will abandon monopulse in favor of maximum-likelihood target location and that a beamspace architecture using highly overlapped beams supports practical and efficient implementation.

Part III

Adaptive Receivers

Part III

Adaptive Receivers

11 Introduction

M. M. Weiner

Part III *Adaptive Receivers* discusses implementation of the *third* of three sequentially complementary approaches for increasing the probability of detection, within at least some cells of the surveillance volume, in presence of external "noise" which can be either Gaussian or nonGaussian in the spatial domain but Gaussian in the temporal domain. For each homogeneous subdivision of the surveillance volume, this approach, identified in the Preface as Approach C, generally utilizes a nonlinear, nonGaussian receiver whose detection algorithm is matched to the sampled "noise" voltage spatial probability density function (PDF) of that subdivision. When the nonGaussian "noise" waveform is spikier than Gaussian noise, the nonlinear receiver is more effective than a linear receiver in reducing the detection threshold for a given false alarm rate, *provided that the estimated spatial PDF is an accurate representation of the actual PDF*. If the estimated spatial PDF is Gaussian, then the nonlinear receiver reduces to a linear Gaussian receiver. The issues are (1) how to model, simulate, and identify the random processes associated with the *correlated* "noise" data samples and (2) how to determine the nonlinear receiver and its threshold which are best matched to those data samples. Part III *Adaptive Receivers* addresses these issues and gives several applications.

Chapter 12 addresses the modeling, simulation, and distribution approximation of correlated spatially nonGaussian radar clutter. After introducing the problem (Section 12.1), the properties of spherically invariant random processes are presented (Section 12.2); several examples of multivariate probability density functions arising from SIRPs are given (Section 12.3); two new canonical computer simulation procedures are developed (Section 12.4) and a new algorithm is introduced to address the distribution approximation of the clutter, using relatively small sample sizes (Section 12.5). A spherically invariant random vector (SIRV), containing N components is defined in Section 12.2 by Equation 12.1. When the components are expressed in an N dimensional spherical coordinate system, the definition and the very name *spherically invariant* follow from the necessary and sufficient condition, that the PDF of the magnitude of a white zero-mean SIRV is dependent only upon the magnitude of the vector and is statistically independent of the remaining $(N-1)$ angular coordinates of the vector (cf. Equation 12.20). SIRVs are useful in approximating

the PDF of sampled nonGaussian and Gaussian random processes (provided the process is an SIRP, i.e., every random vector of the process is a SIRV regardless of sampling instants or the number of samples).

Chapter 13 addresses performance of the Locally Optimum Detector (LOD) in radar weak signal detection for finite sample sizes, where the nonGaussian radar clutter is modeled as a SIRP. A canonical form is established for the LOD that is a product of the Gaussian linear receiver and a zero memory nonlinearity (Section 13.2). A new method for threshold estimation based on extreme value theory is derived which reduces by order of magnitudes the computation and sample sizes required to set the threshold (Section 13.3). The performance is evaluated by computer simulations for finite sample sizes for multivariate Student-T and K-distributed disturbances (Section 13.4 and Section 13.5).

Chapter 14 generalizes SIRVs to include the case of independent vectors with application to reverberation reduction in a correlation sonar.

Chapter 15 presents applications to: statistical normalization of spherically invariant nonGaussian clutter on sonar displays (Section 15.1); nonGaussian clutter modeling for radar target detection (Section 15.2); small signal detection in impulsive nonGaussian clutter (Section 15.3); importance sampling for reducing the number of Monte Carlo simulation bounds to obtain threshold estimates (Section 15.4); rejection method bounds for efficient Monte Carlo computer simulation of SIRVs (Section 15.5); optimal nonGaussian processing using a new canonical form for the optimal nonlinear receiver (Section 15.6); and multichannel detection for correlated nonGaussian processing using an innovations-based detector, viz., a tracking filter or Kalman filter, which utilizes the information in the last signal pulse to predict the next signal pulse (Section 15.7).

12 Spherically Invariant Random Processes for Radar Clutter Modeling, Simulation, and Distribution Identification

M. Rangaswamy

CONTENTS

12.1. Introduction 606
12.2. Background 608
12.2.1. Introduction 608
12.2.2. Definitions 609
12.2.3. Characterization of SIRVs 610
12.2.4. Determining the PDF of a SIRV 615
12.2.5. Properties of SIRVs 618
12.2.5.1. PDF Characterization 618
12.2.5.2. Closure Under Linear Transformation 618
12.2.5.3. Minimum Mean Square Error Estimation 618
12.2.5.4. Distributions of Sums of SIRVs 621
12.2.5.5. Markov Property for SIRPs 622
12.2.5.6. Kalman Filter for SIRPs 624
12.2.5.7. Statistical Independence 625
12.2.5.8. Ergodicity of SIRPs 625
12.2.6. Conclusion 626
12.3. Radar Clutter Modelling Using Spherically Invariant Random Processes 626
12.3.1. Introduction 626
12.3.2. Problem Statement 628

12.3.3. Techniques for Determining the SIRV PDF 630
12.3.3.1. SIRVs with Known Characteristic PDF 630
12.3.3.2. SIRVs with Unknown Characteristic PDFs 631
12.3.3.3. Hankel Transform Approach 632
12.3.4. Examples of Complex SIRVs 634
12.3.4.1. Examples Based on the Characteristic PDF 634
12.3.4.2. Examples Based on Marginal Envelope PDF 641
12.3.4.3. Examples Using the Marginal Characteristic Function 651
12.3.5. Significance of the Quadratic Form of the SIRV PDF 656
12.3.6. Conclusion 664
12.4. Computer Generation of Simulated Radar Clutter Characterized as SIRPs 665
12.4.1. Introduction 665
12.4.2. Preliminaries 666
12.4.3. Two Canonical Simulation Procedures for Generating SIRVs 670
12.4.4. Performance Assessment of the Simulation Schemes 675
12.4.5. Conclusions 677
12.5. Distribution Approximation to Radar Clutter Characterized by SIRPs 680
12.5.1. Introduction 680
12.5.2. Definitions 682
12.5.3. Goodness of Fit Test 682
12.5.4. Distribution Approximation 689
12.5.5. Parameter Estimation 694
12.5.5.1. Estimation of Location and Scale Parameters 695
12.5.5.2. Shape Parameter Estimation 696
12.5.6. Assessing the Distributional Properties of SIRVs 697
12.5.7. Distribution Identification of SIRVs 700
12.5.8. Alternative Method for Parameter Estimation 704
12.5.9. Conclusions 705
12.6. Conclusions 705
12.6.1. General Remarks 705
12.6.2. Suggestions for Future Research 706

12.1. INTRODUCTION

The problem of weak signal detection in nonGaussian noise is of interest to engineers in many disciplines such as, radar, sonar, digital communications, and radio astronomy. In this research, we are interested in the detection of weak radar targets in a strong clutter background. When a signal is transmitted by a radar, the resulting received signal consists of returns from a target (desired signal) and

returns from such objects as buildings, trees, water, land, and weather, depending on the environment. Any return other than that from the target is an unwanted signal and is defined as *radar clutter*.

The fundamental issues need to be addressed are:

(1) Specification of suitable statistical models for the clutter background.
(2) Development of efficient computer simulation procedures for generating samples characterized by the various statistical models.
(3) Development of an approximation procedure for fitting one or more statistical models to a set of experimental data whose distribution is not own *a priori*.

With respect to item (1), a complete statistical clutter model should provide us with a closed form analytical expression for the PDF of the clutter. This should include information about the correlation of the clutter, thereby enabling us to take advantage of this correlation in order to cancel the clutter. We are interested in processing either N correlated complex samples or $2N$ correlated quadrature components obtained by sampling a complex random process that can have a non-Rayleigh envelope PDF with a phase uniformly distributed over the interval $(0, 2\pi)$. In addition, it is assumed that the envelope and the phase are statistically independent. We can think of a vector $\mathbf{Y} = [Y_{c1}, \ldots Y_{cN}, Y_{s1} \ldots Y_{sN}]^T$ to represent the collection of the complex samples obtained from the complex process. The subscripts c and s denote the in-phase and out-of-phase quadrature components, respectively. The issue we need to address is the specification of the joint PDF of the N complex samples or the $2N$ quadrature components. In other words, we need to specify the PDF of the vector $\mathbf{Y}$. We require that the multivariate PDFs be specified in closed form to facilitate their use in the derivation of optimal radar signal processors.

For each PDF of $\mathbf{Y}$ developed, we need to obtain an efficient computer simulation scheme for generating random vectors having this PDF. Computer simulation is necessary because generally it will not be possible to evaluate analytically the performance of optimal nonGaussian radar receivers. In such cases, performance must be evaluated by computer simulation.

More often than not, the background clutter is not known *a priori*. Also, the parameters of the clutter are unknown. Therefore, we need to develop estimation procedures to approximate the clutter PDF and its parameters. We need to address the problem of distribution approximation and parameter estimation. This needs to have an extensive list of possible multivariate PDFs so that a good fit can be found for a given environment. In practice, only a small number of samples may be available. Therefore, there is a need to develop efficient procedures to handle small sizes.

Because the clutter is apt to change in space and time, the radar signal processor must be adaptive to meet the changing conditions. The approach proposed here for the characterization, computer generation, and distribution

approximation of the clutter is based on the theory of spherically invariant random processes (SIRPs) that lends itself to adaptive receivers.

This dissertation addresses the above issues and is organized as follows. In Section 12.2 we present a review of the literature as it pertains to the modeling of radar clutter as a SIRP. Section 12.3 presents techniques for obtaining the joint PDF of N complex, nonGaussian, random variables, assuming that the clutter can be characterized as a SIRP. The need for a library of multivariate nonGaussian PDFs is discussed. Several examples illustrating the various techniques for specifying the multivariate nonGaussian PDF are provided. Finally, a key result for identifying the multivariate nonGaussian PDFs arising from SIRPs is presented in Section 12.3.

Section 12.4 deals with the problem of computer generation of correlated nonGaussian radar clutters that can be characterized as SIRPs. Two canonical simulation procedures are presented. A graphical goodness of fit test procedure is presented, to validate the simulation procedures.

In Section 12.5 we concern ourselves with the distribution approximation of radar clutter characterized by SIRPs. A new graphical scheme based on a key result presented in Section 12.3 is used to address the distribution approximation problem. This procedure reduces the multivariate distribution approximation problem to an equivalent univariate distribution approximation problem, resulting in considerable computational simplicity. Finally, a new technique for shape parameter estimation is suggested based on the approximation procedure. The chief advantage of this scheme is that relatively few samples are needed for the distribution approximation problem.

Conclusions and suggestions for future research are presented in Section 12.6.

12.2. BACKGROUND

12.2.1. Introduction

We present an overview of the literature as it pertains to the modeling of radar clutter by SIRPs. In addition, relevant mathematical preliminaries are presented in this chapter. When a radar transmits a signal, the received echo may consist of returns from one or more targets, buildings, trees, water, land, and weather depending on the environment. The target returns contribute to the desired signal while the other returns contribute the clutter. Many investigators[1–4] have reported experimental measurements for which the clutter PDF has an extended tail. The extended tail gives rise to relatively large probabilities of false alarm. The Gaussian model for the clutter fails to predict this behavior. Two approaches have been used to explain the nonGaussian behavior. One of them is based on the fact that application of the central limit theorem (CLT) is not appropriate. The other approach is based on the nonstationary reflectivity properties of the scanned areas. In any event, nonGaussian models for the univariate (marginal) clutter PDF have been proposed. Commonly reported marginal nonGaussian PDFs for the clutter are Weibull,[1] Lognormal,[5,6] and K-distributions.[2,3,7] Second order

statistics for these models have been reported in terms of autocorrelation functions or power spectral densities.[4,8]

The Weibull[1] and Lognormal[2] models for radar clutter are primarily based on empirical studies, while the *K*-distribution has been shown to have physical significance[2,9] as the observed statistical properties can be related to the electromagnetic and geometric factors pertaining to the scattering surface. Computer simulation schemes for Weibull and Lognormal clutter based on the univariate PDFs and correlation functions have been developed in Ref. 10 and, Ref. 11 respectively. Extension of the Weibull, Lognormal, and *K*-distributed clutter models for coherent radar processing have been developed in, Refs. 5, 12 and, Ref. 13 respectively.

Statistical characterization of the clutter is necessary in order to obtain the optimal radar signal processor. Usually, radars process N pulses at a time. A complete statistical characterization of the clutter requires the specification of the joint PDF of the N samples. When the pulse returns are statistically independent, the joint PDF is simply the product of the marginal PDFs. However, the clutter can be highly correlated. In fact, the correlation between samples is useful in canceling the clutter. Consequently, it is desirable to include the correlation information in the multivariate PDF. For nonGaussian processes this can be done in more than one way. The theory of SIRP provides a powerful mechanism for obtaining the joint PDF of the N correlated nonGaussian random variables. Applications for the theory of SIRPs can be found in the problem of random flights,[14] signal detection and estimation problems in communication theory,[15,16] speech signal processing,[17,18] and radar clutter modeling and simulation.[13,19–22] The following sections provide a brief overview of literature on the theory of SIRPs.

12.2.2. Definitions

In this section we present certain definitions and mathematical preliminaries pertaining to the theory of SIRPs.

Consider a random process characterized by the random vector $\mathbf{Y} = [Y_1, Y_2, \ldots, Y_N]^{\mathrm{T}}$, $N \geq 1$, whose realization is $\mathbf{y} = [y_1, y_2, \ldots, y_N]^{\mathrm{T}}$ where T denotes matrix transpose. The vector $\mathbf{Y}$ is said to be a SIRV if its PDF has the form[14]

$$f_{\mathbf{Y}}(\mathbf{y}) = K \cdot |\mathbf{\Sigma}|^{-1/2} h_N(p) \tag{12.1}$$

where k is a normalization constant chosen so that the volume under the curve of the PDF is unity, $\mathbf{\Sigma}$ is a $N \times N$ nonnegative definite matrix, $h_N(p)$ is a one-dimensional, nonnegative, real-valued monotonically decreasing function with argument $p = [\mathbf{y} - \mathbf{b}]^{\mathrm{T}} \cdot \mathbf{\Sigma}^{-1} \cdot [\mathbf{y} - \mathbf{b}]$, $\mathbf{b}$ is a N by 1 vector, $|\mathbf{\Sigma}|$ is the determinant of the matrix $\mathbf{\Sigma}$, and $\mathbf{\Sigma}^{-1}$ is the inverse of the matrix $\mathbf{\Sigma}$. This paper follows the space-saving conventions, often adopted by the signal processing community, of not using brackets to denote a matrix quantity and of expressing a column vector as the transpose of a row vector. Note that the argument p is a scalar quadratic form. Note also that the PDF of a SIRV is elliptically symmetric (i.e., constant

contours of $f_{\mathbf{Y}}(\mathbf{y})$ are composed of ellipses). If every random vector obtained by sampling a random process $y(t)$ is a SIRV, regardless of the sampling instants or the number of samples, then the process $y(t)$ is defined to be a SIRP.

Kingman[14] introduced the definition of spherically symmetric random vectors (SSRV). In particular, a random vector $\mathbf{X} = [X_1, X_2, \ldots X_N]^{\mathrm{T}}$ is said to be spherically symmetric provided its PDF has the form

$$f_x(\mathbf{x}) - kh_N[(x_1^2 + x_2^2 + \cdots + x_N^2)^{\frac{1}{2}}] = kh_N(\mathbf{x}^{\mathrm{T}}\mathbf{x}) \tag{12.2}$$

where $h_N(\cdot)$ is an arbitrary, nonnegative, monotonically decreasing radial function of dimension N, and k is a normalization constant chosen so that the volume under the curve of the PDF is unity. The subscript N is used to emphasize that we are dealing with N random variables. Throughout the manuscript, it is assumed that the PDF of a random vector is the joint PDF of its components. Equivalently, if $\boldsymbol{\omega} = [\omega_1, \omega_2, \ldots, \omega_N]^{\mathrm{T}}$, the characteristic function of the SSRV $\mathbf{X}$ defined by $\Phi_{\mathbf{X}}(\boldsymbol{\omega}) = E[\exp(j\boldsymbol{\omega}^{\mathrm{T}}\mathbf{X})]$, has the form

$$\Phi_{\mathbf{X}}(\boldsymbol{\omega}) - g_N[(\omega_1^2 + \omega_2^2 + \ldots + \omega_N^2)^{\frac{1}{2}}] \tag{12.3}$$

where $g_N(\cdot)$ is a nonnegative conjugate symmetric function that is magnitude integrable. A SSRV is a special case of a SIRV, arising from Equation 12.1 when $\mathbf{b} = 0$ and $\boldsymbol{\Sigma} = \mathbf{I}$ where $\mathbf{I}$ is the identity matrix. In Appendix U.2, we prove that the characteristic function of a SSRV is also spherically symmetric.

12.2.3. Characterization of SIRVs

In this section we present some important theorems that help us to characterize the PDF of a SIRV. The work of Yao,[15] Kingman,[23] and Wise[24] gave rise to a representation theorem for SSRVs. The representation theorem can be stated as follows.

Theorem 1. *A random vector $\boldsymbol{X} = [X_1, X_2, \ldots X_N]^{T}$ is a SSRV for any N, if and only if there exists a nonnegative random variable T (not to be confused with the symbol T used to denote the transpose of a matrix, as in this theorem) such that the random variables X_i, $(i = 1, 2, \ldots, N)$ conditioned on $T = t$ are independent, identically distributed, Gaussian random variables with zero mean and variance equal to $2t$.*

Proof. Necessity: By definition, the characteristic function of $\mathbf{X}$ is

$$\Phi_{\mathbf{X}}(\boldsymbol{\omega}) = E[\exp(j\boldsymbol{\omega}^{\mathrm{T}}\mathbf{X})] = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \exp(j\boldsymbol{\omega}^{\mathrm{T}}\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) \mathrm{d}\mathbf{x} \tag{12.4}$$

The PDF of the random variable T is introduced by noting that

$$f_{\mathbf{X}}(\mathbf{x}) = \int_{-\infty}^{\infty} f_{\mathbf{X},T}(\mathbf{x}, t)\mathrm{d}t = \int_{-\infty}^{\infty} f_{\mathbf{X}|T}(\mathbf{x}|t) f_T(t)\mathrm{d}t \tag{12.5}$$

Substituting into the expression for the characteristic function and interchanging the order of integration we obtain

$$\Phi_{\mathbf{X}}(\boldsymbol{\omega}) = \int_{-\infty}^{\infty} \Phi_{\mathbf{X}|T}(\boldsymbol{\omega}, t) f_T(t)\mathrm{d}t \tag{12.6}$$

where

$$\Phi_{\mathbf{X}|T}(\boldsymbol{\omega}, t) = \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \exp(j\boldsymbol{\omega}^{\mathrm{T}}\mathbf{x}) f_{\mathbf{X}|T}(\mathbf{x}|t)\mathrm{d}\mathbf{x} \tag{12.7}$$

Since $\mathbf{X}$ is a SSRV for any N, its characteristic function has the form of Equation 12.3. This requires that the functional form of $\Phi_{\mathbf{X}|T}(\boldsymbol{\omega}, t)$ remain unchanged for all N. Furthermore, $\Phi_{\mathbf{X}|T}(\boldsymbol{\omega}, t)$ must also be a function of $(\omega_1^2 + \omega_2^2 + \cdots + \omega_N^2)$ for any choice of N. The only characteristic function satisfying these conditions[23] is

$$\Phi_{\mathbf{X}|T}(\boldsymbol{\omega}, t) = \exp[-t(\omega_1^2 + \omega_2^2 + \cdots + \omega_N^2)] \tag{12.8}$$

where the conditional PDF of $\mathbf{X}$, given $T = t$, is recognized to be multivariate Gaussian, with X_i, $(i = 1, 2, \ldots, N)$ being statistically independent identically distributed, zero mean Gaussian random variables with variance $2t$. Because the variance equals $2t$, T must be a nonnegative random variable. Necessity follows. Note that the theorem does not give any physical significance for T. Neither does it reveal how to determine $f_T(t)$.

Sufficiency: To prove sufficiency, we need to show that every product of a Gaussian random vector $\mathbf{Z}$ having zero mean and identity covariance matrix and a random variable $S' = \sqrt{2T}$ with PDF $f_{S'}(s')$ results in a PDF of the form of Equation 12.2.

In particular, consider the product $\mathbf{X} = \mathbf{Z}S'$. The PDF of $\mathbf{X}$ conditioned on S' is then given by

$$f_{\mathbf{X}|S'}(\mathbf{x}|s') = (2\pi)^{-\frac{N}{2}} |s'|^{-N} \exp\left(-\frac{p'}{2|s'|^2}\right) \tag{12.9}$$

where $p' = \mathbf{x}^{\mathrm{T}}\mathbf{x}$. From the theorem on total probability, the PDF of $\mathbf{X}$ can be written as

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-\frac{N}{2}} \int_{-\infty}^{\infty} |s'|^{-N} \exp\left(-\frac{p'}{2|s'|^2}\right) f_{S'}(s')\mathrm{d}s' \tag{12.10}$$

For convenience, we can write the PDF of S' as

$$f_{S'}(s') = f_1(s') + f_2(s') \tag{12.11}$$

where

$$f_1(s') = \begin{cases} f_1(s') & s' \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{12.12}$$

and

$$f_2(s') = \begin{cases} f(s') & s' \leq 0 \\ 0 & \text{otherwise} \end{cases} \tag{12.13}$$

Then, Equation 12.10 can be expressed as

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-\frac{N}{2}} \int_{-\infty}^{0} |s'|^{-N} \exp\left(-\frac{p'}{2|s'|^2}\right) f_2(-s')\mathrm{d}s'$$
$$+ (2\pi)^{-\frac{N}{2}} \int_{0}^{\infty} |s'|^{-N} \exp\left(-\frac{p'}{2|s'|^2}\right) f_1(s')\mathrm{d}s' \tag{12.14}$$

Making the change of variable $-s' = \zeta$ in the first integral of Equation 12.14, we have

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-\frac{N}{2}} \int_{0}^{\infty} (s')^{-N} \exp\left(-\frac{p'}{2(s')^2}\right) f_{S'}(s')\mathrm{d}s' \tag{12.15}$$

Thus, it is clear that regardless of whether S' is positive or negative, the PDF of $\mathbf{X}$ has the form of Equation 12.15. Henceforth, we always consider the product of $\mathbf{Z}$ and a nonnegative random variable S in our analysis.
Comparing Equation 12.15 and Equation 12.2, we can write $k = (2\pi)^{-\frac{N}{2}}$ and

$$h_N(p') = \int_{0}^{\infty} (s')^{-N} \exp\left(-\frac{p'}{2(s')^2}\right) f_{S'}(s')\mathrm{d}s' \tag{12.16}$$

Note that $h_N(p)$ given by Equation 12.16 is a nonnegative, monotonically decreasing function of p, for all N. Therefore, the PDF of Equation 12.15 is entirely equivalent to that of Equation 12.2. This establishes the theorem. Thus, the PDF of a SSRV is uniquely determined by the specification of a Gaussian random vector having zero mean and identity covariance matrix, and a first order PDF $f_S(s)$ called the characteristic PDF.

The following theorem in Ref. 25 states that a SIRV is related to a SSRV by a linear transformation.

Theorem 2. *If $\mathbf{X}$ is a SSRV, with characteristic PDF $f_S(s)$, then the deterministic linear transformation*

$$\mathbf{Y} = \mathbf{AX} + \mathbf{b} \tag{12.17}$$

results in $\mathbf{Y}$ *being a SIRV having the same characteristic PDF. It is required that* $\mathbf{A}$ *be a matrix such that* $\mathbf{A}\mathbf{A}^T$ *is nonsingular and* $\mathbf{b}$ *be an* N *by 1 vector.*

Proof. Since $\mathbf{X}$ is a SSRV, we can express $\mathbf{X}$ as $\mathbf{X} = \mathbf{Z}S$, where $\mathbf{Z}$ is a Gaussian random vector having zero mean and identity covariance matrix and S is a nonnegative random variable. Hence,

$$\mathbf{Y} = \mathbf{A}\mathbf{Z}S + \mathbf{b} \tag{12.18}$$

Conditioned on S, the PDF of $\mathbf{Y}$ is Gaussian, with mean vector equal to $\mathbf{b}$ and covariance matrix equal to $\mathbf{A}\mathbf{A}^{\mathrm{T}}s^2$. The PDF of $\mathbf{Y}$ conditioned on S is given by

$$f_{\mathbf{Y}|S}(\mathbf{y}|s) = (2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}}s^{-N}\exp\left(-\frac{p}{2s^2}\right) \tag{12.19}$$

where $p = (\mathbf{y}-\mathbf{b})^T\mathbf{\Sigma}^{-1}(\mathbf{y}-\mathbf{b})$ and $|\mathbf{\Sigma}|$ denotes the determinant of the covariance matrix $\mathbf{\Sigma} = \mathbf{A}\mathbf{A}^{\mathrm{T}}$. Implicit herein is the assumption that S has unit mean square value. Using the theorem on total probability, the PDF of $\mathbf{Y}$ can be written as

$$f_{\mathbf{Y}}(\mathbf{y}) = (2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}}h_N(p) \tag{12.20}$$

where

$$h_N(p) = \int_0^\infty s^{-N}\exp\left(-\frac{p}{2s^2}\right)f_S(s)\mathrm{d}s \tag{12.21}$$

The PDF of $\mathbf{Y}$ is of the form of Equation 12.1. Therefore $\mathbf{Y}$ is a SIRV.

The PDF of a SIRV is uniquely determined by the specification of a mean vector, a covariance matrix, and a first order PDF called the characteristic PDF. Theorem 1 for SSRVs generalizes for SIRVs in a straightforward manner. The only difference is that, conditioned on the nonnegative random variable T, the $\{\mathbf{Y}_k : (k = 1,2,\ldots N)\}$ are no longer statistically independent. Instead, the PDF of $\mathbf{Y}$ conditioned on T is a multivariate Gaussian PDF. By the same argument used for SSRVs, a SIRV can be written as a product of a Gaussian random vector and a nonnegative random variable with the only difference that the mean of the Gaussian random vector need not be zero and its covariance matrix is not the identity matrix. As a corollary of Theorem 2,[15] it can be readily shown that every linear transformation on a SIRV results in another SIRV having the same characteristic PDF. As a special case, when $f_S(s) = \delta(s-1)$ where $\delta(\cdot)$ is the unit impulse function, $h_N(p) = \exp(-p/2)$ and the corresponding SIRV PDF given by Equation 12.20 is the multivariate Gaussian PDF. Therefore, the multivariate Gaussian PDF is a special case of the SIRV PDF.

The following theorem[16] provides an interesting property of SSRVs when represented in generalized spherical coordinates.

Theorem 3. *A random vector* $\mathbf{X} = [X_1, \ldots X_N]^{\mathrm{T}}$ *is a SSRV if and only if there exist* N *random variables* $R \in (0, \infty)$, $\Theta \in (0, 2\pi)$ *and* $\Phi_k \in (0, \pi)$, $(k = 1, \ldots N-2)$ *such that when the components of* $\mathbf{X}$ *are expressed in the generalized spherical coordinates*

$$\begin{aligned} X_1 &= R\cos(\Phi_1) \\ X_k &= R\cos(\Phi_k)\prod_{i=1}^{k-1}\sin(\Phi_i) \qquad 1 < k \le N-2 \\ X_{N-1} &= R\cos(\Theta)\prod_{i=1}^{N-2}\sin(\Phi_i) \\ X_N &= R\sin(\Theta)\prod_{i=1}^{N-2}\sin(\Phi_i) \end{aligned} \tag{12.22}$$

then the random variables R Θ *and* Φ_k *are mutually statistically independent and have PDFs of the form*

$$\begin{aligned} f_R(r) &= \frac{r^{N-1}}{2^{\frac{N}{2}-1}\Gamma\left(\frac{N}{2}\right)}h_N(r^2)u(r) \\ f_{\phi_k}(\phi_k) &= \frac{\Gamma\left(\frac{N-k+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{N-k}{2}\right)}\sin^{N-1-k}(\phi_k)[u(\phi_k) - u(\phi_k - \pi)] \\ f_\Theta(\theta) &= (2\pi)^{-1}[u(\theta) - u(\theta - 2\pi)] \end{aligned} \tag{12.23}$$

where $\Gamma(\cdot)$ *is the Eulero Gamma function and* $u(\cdot)$ *is the unit step function.*

Proof. Since the random vector $\mathbf{X}$ is a SSRV, its PDF is of the form of Equation 12.2 with $h_N(p')$ being given by Equation 12.16. The Jacobian of the transformation given by Equation 12.22 is obtained in Ref. 26 as

$$J = \left(R^{N-1}\prod_{k=1}^{N-2}\sin^{N-1-k}(\phi_k)\right)^{-1} \tag{12.24}$$

Using Equation 12.2 and Equation 12.24 and noting that $R^2 = \sum_{k=1}^{N} X_k^2$, the joint PDF of R, Θ and Φ_k $(k = 1, 2, \ldots N-2)$ becomes

$$f_{R,\Theta,\Phi_1\ldots\Phi_{N-2}}(r, \theta, \phi_1 \ldots \phi_{N-2}) = \frac{r^{N-1}}{(2\pi)^{\frac{N}{2}}}h_N(r^2)\prod_{k=1}^{N-2}\sin^{N-1-k}(\phi_k) \tag{12.25}$$

Since the joint PDF in Equation 12.25 can be written as a product of the marginal PDFs given in Equation 12.23, the variables R, Θ and Φ_k, are mutually statistically independent with the prescribed PDFs. In order to prove the sufficient part of the property, we start with the marginal PDFs of R, Θ and Φ_k given by Equation 12.23 and, under the assumption of statistical independence, obtain

the joint PDF of Equation 12.25. Using the inverse of the Jacobian given by Equation 12.24, results in the PDF of $\mathbf{X}$ as given by Equation 12.22.

12.2.4. Determining The PDF of a SIRV

In this section we shall present schemes for determining the PDF of a SIRV. We recognize that the PDF of a SIRV is uniquely determined by the specification of a mean vector, a covariance matrix, and a characteristic first order PDF; and that the SIRV PDF has the form of Equation 12.20. Several techniques are available in the literature for specifying $h_N(p)$. The simplest technique is to use Equation 12.21. However, this procedure requires the knowledge of the characteristic PDF $f_S(s)$. Therefore, when $f_S(s)$. is not known in closed form or it is difficult to evaluate the integral in Equation 12.21, alternate methods for specifying $h_N(p)$ must be examined.

To study the behavior of $h_N(p)$, it is convenient to replace p, which is a quadratic form depending on N, by the dummy scalar variable q. We then write

$$h_N(q) = \int_0^\infty s^{-N} \exp\left(-\frac{q}{2s^2}\right) f_S(s)\mathrm{d}s \tag{12.26}$$

When both sides of Equation 12.26 are differentiated with respect to q, we obtain

$$\frac{\mathrm{d}h_N(q)}{\mathrm{d}q} = -\frac{1}{2}\int_0^\infty s^{-N-2} \exp\left(-\frac{q}{2s^2}\right) f_S(s)\mathrm{d}s \tag{12.27}$$

The right hand side of Equation 12.27 is related to $h_{N+2}(q)$ by the factor of $-\frac{1}{2}$. Thus, we have an interesting result pointed out in Ref. 19 that

$$h_{N+2}(q) = (-2)\frac{\mathrm{d}h_N(q)}{\mathrm{d}q} \tag{12.28}$$

As

$$f_{\mathbf{Y}}(\mathbf{y}) = (2\pi)^{-\frac{N+2}{2}} |\mathbf{\Sigma}|^{-\frac{1}{2}} h_{N+2}(p) \tag{12.29}$$

when $\mathbf{Y}$ is of dimension $N+2$, if follows that $h_N(q)$ must be a monotonically decreasing function for all N. Equation 12.28 provides a mechanism for relating higher order PDFs with those of lower order, for a given SIRV. More precisely, starting with $N=1$ and $N=2$, and using Equation 12.28 repeatedly, gives the following pair of recurrence relations.

$$\begin{aligned} h_{2N+1}(q) &= (-2)^N \frac{\mathrm{d}^N h_1(q)}{\mathrm{d}q^N} \\ h_{2N+2}(q) &= (-2)^N \frac{\mathrm{d}^N h_2(q)}{\mathrm{d}q^N} \end{aligned} \tag{12.30}$$

Therefore, starting from $h_1(q)$ and $h_2(q)$ all PDFs of odd and even order, respectively, can be generated by the use of Equation 12.30. However, since $h_N(\cdot)$

is defined to be a nonnegative monotonically decreasing function for all N, $h_1(\cdot)$ and $h_2(\cdot)$ must belong to a class of functions that are nonnegative and monotonically decreasing. Consequently, their successive derivatives will alternate between negative and positive functions that are monotonically increasing and decreasing, respectively. Given $h_N(q)$, the Nth order SIRV PDF is given by

$$f_{\mathbf{Y}}(\mathbf{y}) = (2\pi)^{-\frac{N}{2}} \left|\boldsymbol{\Sigma}\right|^{-\frac{1}{2}} h_N(p) \tag{12.31}$$

where $h_N(p)$ is nothing more than $h_N(p)$ with q replaced by p.

Another approach for specifying $h_N(p)$, which begins with the univariate characteristic function, has been proposed in Refs. 15, 16, and 27, requiring that the univariate characteristic function be a real even function whose magnitude is integrable. Also, it is assumed that the components of the SIRV are identically distributed. Under these conditions, it has been shown that

$$h_N(p) = (\sqrt{p})^{1-\frac{N}{2}} \int_0^\infty \omega^{\frac{N}{2}} \phi(\boldsymbol{\omega}) J_{\frac{N-2}{2}}(\boldsymbol{\omega}\sqrt{p}) \mathrm{d}\boldsymbol{\omega} \tag{12.32}$$

where $\phi(\boldsymbol{\omega})$ is the univariate characteristic function and $J_\alpha(\eta)$ is the Bessel function of order α. Equation 12.32 has an elegant proof by induction, which is presented here. From Equation 12.20 it follows that $h_1(p)$ is related to the first order SIRV PDF of the ith component. More explicitly, we can write

$$f_{Y_i}(\mathrm{y}_i) = (\sqrt{2\pi}\sigma)^{-1} h_1(p_i) \qquad i = 1, 2, \ldots, N \tag{12.33}$$

where $p_i = (y_1^2/\sigma^2)$ and σ^2 is the common variance of the random variables Y_i $(i = 1, 2, \ldots, N)$. For convenience, assume that σ^2 is unity. The univariate characteristic function is then given by

$$\phi_i(\boldsymbol{\omega}) = \int_{-\infty}^{\infty} f_{Y_i}(y_i) \exp(j\boldsymbol{\omega} y_i) \mathrm{d}y_i \tag{12.34}$$

Using the inverse Fourier transform and noting that $y_i = \sqrt{p_i} h_1(p_i)$ can be expressed in terms of the characteristic function as

$$h_1(p_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \phi_i(\boldsymbol{\omega}) \exp(-j\boldsymbol{\omega}\sqrt{p}) \mathrm{d}\boldsymbol{\omega} \tag{12.35}$$

Since $\phi_i(\boldsymbol{\omega})$ is the same for all i, the subscript i in Equation 12.35 can be dropped. In addition, because $\phi(\boldsymbol{\omega})$ is an even function, we can rewrite Equation 12.35 as

$$h_1(p) = \sqrt{\frac{2}{\pi}} \int_0^\infty \phi(\boldsymbol{\omega}) \cos(\boldsymbol{\omega}\sqrt{p}) \mathrm{d}\boldsymbol{\omega} \tag{12.36}$$

Recognizing that $\cos(x) = \sqrt{\frac{\pi x}{2}} J_{-\frac{1}{2}}(x)$, and replacing p by the dummy variable q,

we have

$$h_1(q) = (\sqrt{q})^{\frac{1}{2}} \int_0^\infty \boldsymbol{\omega}^{\frac{1}{2}} \phi(\boldsymbol{\omega}) J_{-\frac{1}{2}}(\boldsymbol{\omega}\sqrt{q}) \mathrm{d}\boldsymbol{\omega} \tag{12.37}$$

Since the derivation makes use of Equation 12.28 it is necessary to consider odd and even values of N separately. For odd values of N, Equation 12.3 can be written as

$$h_{2N-1}(q) = (q)^{\frac{3}{2}-N} \int_0^\infty \boldsymbol{\omega}^{N-\frac{1}{2}} \phi(\boldsymbol{\omega}) J_{\frac{2N-3}{2}}(\boldsymbol{\omega}\sqrt{q}) \mathrm{d}\boldsymbol{\omega} \tag{12.38}$$

Equation 12.38 is now shown to hold for all N by means of induction. With $N = 1$, Equation 12.38 reduces to Equation 12.37. It remains to show that Equation 12.38 is valid when N is replaced by $N + 1$. Differentiating both sides of Equation 12.38 with respect to q, we obtain

$$\frac{\mathrm{d}h_{2N-1}(q)}{\mathrm{d}q} = \int_0^\infty \boldsymbol{\omega}^{N-\frac{1}{2}} \phi(\boldsymbol{\omega}) \frac{\mathrm{d}}{\mathrm{d}q}\left[(\sqrt{q})^{\frac{3}{2}-N} J_{\frac{2N-3}{2}}(\boldsymbol{\omega}\sqrt{q})\right] \mathrm{d}\boldsymbol{\omega} \tag{12.39}$$

First, focus on the term $\frac{\mathrm{d}}{\mathrm{d}q}\left[(\sqrt{q})^{\frac{3}{2}-N} J_{\frac{2N-3}{2}}(\boldsymbol{\omega}\sqrt{q})\right]$. Since this involves the derivative of a product, we can write

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}q}\left[(\sqrt{q})^{\frac{3}{2}-N} J_{\frac{2N-3}{2}}(\boldsymbol{\omega}\sqrt{q})\right] \\ &= \frac{1}{2}\left(\frac{3}{2} - N\right)(\sqrt{q})^{-\frac{1}{2}-N} J_{\frac{2N-3}{2}}(\boldsymbol{\omega}\sqrt{q}) + (\sqrt{q})^{\frac{3}{2}-N} \frac{\mathrm{d}}{\mathrm{d}q}\left[J_{\frac{2N-3}{2}}(\boldsymbol{\omega}\sqrt{q})\right] \end{aligned} \tag{12.40}$$

Using the identity[28]

$$\frac{\mathrm{d}J_\alpha(\eta)}{\mathrm{d}\eta} = \frac{\alpha}{\eta} j_\alpha(\eta) - J_{\alpha+1}(\eta) \tag{12.41}$$

we have

$$\frac{\mathrm{d}}{\mathrm{d}q}\left[J_{\frac{2N-3}{2}}(\boldsymbol{\omega}\sqrt{q})\right] = \frac{\boldsymbol{\omega}}{2}(\sqrt{q})^{-1}\left[\frac{2N-3}{2\boldsymbol{\omega}\sqrt{q}} J_{\frac{2N-3}{2}}(\boldsymbol{\omega}\sqrt{q}) - J_{\frac{2N-1}{2}}(\boldsymbol{\omega}\sqrt{q})\right] \tag{12.42}$$

Substituting Equation 12.42 into Equation 12.40 gives

$$\frac{\mathrm{d}}{\mathrm{d}q}\left[(\sqrt{q})^{\frac{3}{2}-N} J_{\frac{2N-3}{2}}(\boldsymbol{\omega}\sqrt{q})\right] = -\frac{\boldsymbol{\omega}}{2}(\sqrt{q})^{\frac{1}{2}-N} J_{\frac{2N-1}{2}}(\boldsymbol{\omega}\sqrt{q}) \tag{12.43}$$

Consequently, Equation 12.39 reduces to

$$\frac{\mathrm{d}h_{2N-1}(q)}{\mathrm{d}q} = -\frac{1}{2}(\sqrt{q})^{\frac{1}{2}-N} \int_0^\infty \boldsymbol{\omega}^{N+\frac{1}{2}} \phi(\boldsymbol{\omega}) J_{\frac{2N-1}{2}}(\boldsymbol{\omega}\sqrt{q}) \mathrm{d}\boldsymbol{\omega} \tag{12.44}$$

However, from Equation 12.28 we know that $h_{2N+1}(q) = (-2)\frac{\mathrm{d}h_{2N-1}(q)}{\mathrm{d}q}$. Hence, we have from Equation 12.44

$$h_{2N+1}(q) = (\sqrt{q})^{\frac{1}{2}-N} \int_0^\infty \boldsymbol{\omega}^{N+\frac{1}{2}} \phi(\boldsymbol{\omega}) J_{\frac{2N-1}{2}}(\boldsymbol{\omega}\sqrt{q}) \mathrm{d}\boldsymbol{\omega} \tag{12.45}$$

Because Equation 12.45 is identical to Equation 12.38 with N replaced by $N + 1$, it has been shown by induction that Equation 12.38 is valid for all N. For ease of derivation, it was assumed that the components of $\mathbf{Y}$ have identical variances. However, since the functional form of $h_N(p)$ is invariant to the choice of p, it follows that Equation 12.32 is valid for all odd values of N.

In a similar manner, starting with $h_2(p)$, it can be shown that

$$h_{2N+2}(p) = (\sqrt{p})^{-N} \int_0^\infty \boldsymbol{\omega}^{N+1} \phi(\boldsymbol{\omega}) J_N(\boldsymbol{\omega}\sqrt{p}) \mathrm{d}\boldsymbol{\omega} \tag{12.46}$$

for all N. Note that Equation 12.46 is identical to Equation 12.32 with N replaced by $2N + 2$. The proof of this result is presented in Section 12.3. Thus, in general, for any N (odd or even), we can write $h_N(p)$ as in Equation 12.32.

12.2.5. Properties of SIRVs

In this section we present certain important properties of SIRVs.

12.2.5.1. PDF Characterization

The multivariate PDF of a SIRV as given by Equation 12.20 and Equation 12.21 is uniquely determined by the specification of a mean vector $\mathbf{b}$, a covariance matrix $\boldsymbol{\Sigma}$, and a characteristic first order PDF $f_S(s)$. The PDF of the SIRV is a nonnegative, real valued monotonically decreasing function, $h_N(\cdot)$, of a nonnegative quadratic form multiplied by a constant. The type of SIRV is determined by the form of $h_N(\cdot)$ or, equivalently, the choice of $f_S(s)$. Higher order PDFs can be obtained by the use of Equation 12.32 whereas lower order PDFs can be obtained in the usual manner by integrating out the unwanted variables. We discuss this procedure in Appendix U.3. The PDFs of all orders are of the same type. The marginal PDFs are used to classify the type of SIRV.

12.2.5.2. Closure Under Linear Transformation

As shown in Theorem 2 of Section 12.2.3, every linear transformation of the form of Equation 12.17 on a SIRV results in another SIRV having the same characteristic PDF.

This feature is called the closure property of SIRVs.[15,16]

12.2.5.3. Minimum Mean Square Error Estimation

In minimum mean square error estimation (MMSE) problems, given a set of data, SIRVs are found to result in linear estimators.[15,27,29] An interesting proof of this property is presented here. Let $\mathbf{Y} = [\mathbf{Y}_1^{\mathrm{T}} \mathbf{Y}_2^{\mathrm{T}}]^{\mathrm{T}}$ where $\mathbf{Y}_1 = [Y_1, Y_2, \ldots Y_m]^{\mathrm{T}}$ and $\mathbf{Y}_2 = [Y_{m+1}, Y_{m+2}, \ldots Y_N]^{\mathrm{T}}$ denote the partitions of $\mathbf{Y}$. It has been pointed out in Ref. 30 that the minimum mean square error estimate of the random vector $\mathbf{Y}_2$ given the observations from the random vector $\mathbf{Y}_1$, is given by

$$\hat{\mathbf{Y}}_2 = E[\mathbf{Y}_2 | \mathbf{Y}_1] \tag{12.47}$$

where $E[\mathbf{Y}_2|\mathbf{Y}_1]$ denotes the conditional mean, i.e., the expected value of $\mathbf{Y}_2$ given $\mathbf{Y}_1$. Assume that $\mathbf{Y}$ is a SIRV of dimension N with characteristic PDF $f_S(s)$. Also, for convenience, it is assumed that the mean of $\mathbf{Y}$ is zero. The covariance matrix of $\mathbf{Y}$ denoted by $\boldsymbol{\Sigma}$ can be portioned as

$$\boldsymbol{\Sigma} = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix} \tag{12.48}$$

where $\mathbf{C}_{11}$ denotes the covariance matrix of $\mathbf{Y}_1$, $\mathbf{C}_{12}$ denotes the cross covariance matrix of the vectors $\mathbf{Y}_1$ and $\mathbf{Y}_2$, $\mathbf{C}_{21}$ is the transpose of $\mathbf{C}_{21}$, and $\mathbf{C}_{22}$ denotes the covariance matrix of the vector $\mathbf{Y}_2$. The PDF of $\mathbf{Y}_2$ given $\mathbf{Y}_1$ is expressed as

$$f_{\mathbf{Y}_2|\mathbf{Y}_1}(\mathbf{y}_2|\mathbf{y}_1) = \frac{f_{\mathbf{Y}}(\mathbf{y})}{f_{\mathbf{Y}_1}(\mathbf{y}_1)} \tag{12.49}$$

Recall from Equation 12.20 and Equation 12.21 that

$$f_{\mathbf{Y}}(\mathbf{y}) = (2\pi)^{-\frac{N}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} h_N(p) \tag{12.50}$$

where

$$h_N(p) = \int_0^\infty s^{-N} \exp\left(-\frac{p}{2s^2}\right) f_S(s) \mathrm{d}s \tag{12.51}$$

and $p = \mathbf{y}^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \mathrm{y}$. Note that the inverse covariance matrix can be partitioned as[26]

$$\boldsymbol{\Sigma}^{-1} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \tag{12.52}$$

where:

$$\begin{aligned} \mathbf{A} &= (\mathbf{C}_{11} - \mathbf{C}_{12}\mathbf{C}_{22}^{-1}\mathbf{C}_{21})^{-1} \\ \mathbf{B} &= -\mathbf{A}\mathbf{C}_{12}\mathbf{C}_{22}^{-1} \\ \mathbf{C} &= -\mathbf{D}\mathbf{C}_{21}\mathbf{C}_{11}^{-1} \\ \mathbf{D} &= (\mathbf{C}_{22} - \mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{C}_{12})^{-1} \end{aligned} \tag{12.53}$$

Expanding the quadratic form, we have

$$p = \mathbf{y}_1^{\mathrm{T}}\mathbf{A}\mathbf{y}_1 + \mathbf{y}_1^{\mathrm{T}}\mathbf{B}\mathbf{y}_2 + \mathbf{y}_2^{\mathrm{T}}\mathbf{C}\mathbf{y}_1 + \mathbf{y}_2^{\mathrm{T}}\mathbf{D}\mathbf{y}_2 \tag{12.54}$$

Adding and subtracting $\mathbf{y}_1^{\mathrm{T}}\mathbf{C}_{11}^{-1}\mathbf{y}_1$ to the right hand side of Equation 12.54 gives

$$p = \mathbf{y}_1^{\mathrm{T}}(\mathbf{A} - \mathbf{C}_{11}^{-1})\mathrm{y}_1 + \mathbf{y}_1^{\mathrm{T}}\mathbf{C}_{11}^{-1}\mathbf{y}_1 + \mathbf{y}_1^{\mathbf{T}}\mathbf{B}\mathbf{y}_2 + \mathbf{y}_2^{\mathrm{T}}\mathbf{C}\mathbf{y}_1 + \mathbf{y}_2^{\mathrm{T}}\mathbf{D}\mathbf{y}_2 \tag{12.55}$$

Note that

$$\mathbf{A} - \mathbf{C}_{11}^{-1} = -\mathbf{B}\mathbf{C}_{21}\mathbf{C}_{11}^{-1} \tag{12.56}$$

Hence,

$$p = \mathbf{y}_1^{\mathrm{T}}\mathbf{C}_{11}^{-1}\mathbf{y}_1 - \mathbf{y}_1^{\mathrm{T}}\mathbf{B}\mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{y}_1 + \mathbf{y}_1^{\mathrm{T}}\mathbf{B}\mathbf{y}_2 + \mathbf{y}_2^{\mathrm{T}}\mathbf{C}\mathbf{y}_1 + \mathbf{y}_2^{\mathrm{T}}\mathbf{D}\mathbf{y}_2 \tag{12.57}$$

However, it can be shown that

$$\begin{aligned} \mathbf{y}_2^{\mathrm{T}}\mathbf{C}\mathbf{y}_1 &= -\mathbf{y}_2^{\mathrm{T}}\mathbf{D}\mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{y}_1 \\ \mathbf{y}_1^{\mathrm{T}}\mathbf{B}\mathbf{y}_2 &= -\mathbf{y}_1^{\mathrm{T}}\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{D}\mathbf{y}_2 \\ -\mathbf{y}_1^{\mathrm{T}}\mathbf{B}\mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{y}_1 &= \mathbf{y}_1^{\mathrm{T}}\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{D}\mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{y}_1 \end{aligned} \tag{12.58}$$

Making these substitutions in the expression for p, it follows that

$$\begin{aligned} p = \mathbf{y}_1^{\mathrm{T}}\mathbf{C}_{11}^{-1}\mathbf{y}_1 + \mathbf{y}_2^{\mathrm{T}}\mathbf{D}\mathbf{y}_2 - \mathbf{y}_2^{\mathrm{T}}\mathbf{D}\mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{y}_1 - \mathbf{y}_1^{\mathrm{T}}\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{D}\mathbf{y}_2 \\ + \mathbf{y}_1^{\mathrm{T}}\mathbf{C}_{11}^{-1}\mathbf{C}_{12}\mathbf{D}\mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{y}_1 \end{aligned} \tag{12.59}$$

This can be rewritten as

$$p = \mathbf{y}_1^{\mathrm{T}}\mathbf{C}_{11}^{-1}\mathbf{y}_1 + (\mathbf{y}_2 - \mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{y}_1)^{\mathrm{T}}\mathbf{D}(\mathbf{y}_2 - \mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{y}_1) \tag{12.60}$$

For simplicity, we define

$$\begin{aligned} p_1 &= \mathbf{y}_1^{\mathrm{T}}\mathbf{C}_{11}^{-1}\mathbf{y}_1 \\ p_2 &= (\mathbf{y}_2 - \mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{y}_1)^{\mathrm{T}}\mathbf{D}(\mathbf{y}_2 - \mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{y}_1) \end{aligned} \tag{12.61}$$

Then,

$$p = p_1 + p_2 \tag{12.62}$$

From Equation 12.62 and Equation 12.49 to Equation 12.51, we have

$$f_{\mathbf{Y}_2|\mathbf{Y}_1}(\mathbf{y}_2|\mathbf{y}_1) = \frac{k}{f_{Y_1}(y_1)}\int_0^\infty s^{-N}\exp\left(-\frac{p_1+p_2}{2s^2}\right)f_s(s)\mathrm{d}s \tag{12.63}$$

where $k = (2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}}$ Next, consider

$$E[\mathbf{Y}_2|\mathbf{Y}_1] = \frac{k}{f_{\mathbf{Y}_1}(\mathbf{y}_1)}\int_0^\infty s^{-N}\exp\left(-\frac{p_1}{2s^2}\right)\int_{\mathbf{Y}_2}\mathbf{y}_2\exp\left(-\frac{p_2}{2s^2}\right)\mathrm{d}\mathbf{y}_2 f_S(s)\mathrm{d}s \tag{12.64}$$

Noting that

$$\int_{\mathbf{Y}_2}\mathbf{y}_2\exp\left(-\frac{p_2}{2s^2}\right)\mathrm{d}\mathbf{y}_2 = (2\pi)^{\frac{N-m}{2}}|\mathbf{D}|^{-\frac{1}{2}}s^{N-m}[\mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{y}_1] \tag{12.65}$$

gives

$$E[\mathbf{Y}_2|\mathbf{Y}_1] = \frac{k_1}{f_{\mathbf{Y}_1}(\mathbf{y}_1)}\int_0^\infty s^{-m}\exp\left(-\frac{p_1}{2s^2}\right)f_S(s)\mathrm{d}s \tag{12.66}$$

where $k_1 = (2\pi)^{-\frac{m}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}}|\mathbf{D}|^{-\frac{1}{2}}[\mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{y}_1]$. When a matrix is partitioned as in Equation 12.52, it is known that[31]

$$|\mathbf{\Sigma}| = |\mathrm{C}_{11}||\mathbf{C}_{22} - \mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{C}_{12}| \tag{12.67}$$

Since

$$\mathbf{D} = (\mathbf{C}_{22} - \mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{C}_{12})^{-1} \tag{12.68}$$

it follows that

$$|\boldsymbol{\Sigma}| = |\mathbf{C}_{11}||\mathbf{D}^{-1}| \tag{12.69}$$

Thus,

$$|\boldsymbol{\Sigma}^{-1}| = |\mathbf{C}_{11}|^{-1}|\mathbf{D}| \tag{12.70}$$

Hence $k_1 = (2\pi)^{-\frac{m}{2}}|\mathbf{C}_{11}|^{-\frac{1}{2}}[\mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{y}_1]$. Finally, since

$$f_{\mathbf{Y}_1}(\mathbf{y}_1) = (2\pi)^{-\frac{m}{2}}|\mathbf{C}_{11}|^{-\frac{1}{2}}\int_0^\infty s^{-m}\exp\left(-\frac{p_1}{2s^2}\right)f_S(s)\mathrm{d}s \tag{12.71}$$

$$\hat{\mathbf{Y}}_2 = E[\mathbf{Y}_2|\mathbf{Y}_1] = [\mathbf{C}_{21}\mathbf{C}_{11}^{-1}\mathbf{y}_1] \tag{12.72}$$

It is seen that the MMSE estimate of $\mathbf{Y}_2$, given the data $\mathbf{Y}_1$, is a linear function of $\mathbf{Y}_1$.

If the random vectors $\mathbf{Y}_1$ and $\mathbf{Y}_2$ have nonzero means denoted by $\mathbf{b}_1$ and $\mathbf{b}_2$, respectively, then Equation 12.72 takes the form

$$E[\mathbf{Y}_2|\mathbf{Y}_1] = \mathbf{b}_2 + \mathbf{C}_{21}\mathbf{C}_{11}^{-1}(\mathbf{y}_1 - \mathbf{b}_1) \tag{12.73}$$

As a consequence of this property, when the random vectors $\mathbf{Y}_1$ and $\mathbf{Y}_2$ are uncorrelated so that $\mathbf{C}_{21} = 0$, then we have

$$E[\mathbf{Y}_2|\mathbf{Y}_1] = \mathrm{b}_2 = E[\mathbf{Y}_2] \tag{12.74}$$

This property is referred to as semi independence in Refs. 15, 27, and 32. However, for all SIRVs except the Gaussian, this result does not imply that

$$f_{\mathbf{Y}_2|\mathbf{Y}_1}(\mathbf{y}_2|\mathbf{y}_1) = f_{\mathbf{Y}_2}(\mathbf{y}_2) \tag{12.75}$$

This emphasizes the property that although uncorrelatedness guarantees statistical independence for Gaussian random vectors, it is not a general property of SIRVs.

12.2.5.4. Distributions of Sums of SIRVs

While it is true that the sum of two jointly Gaussian random vectors is also Gaussian, the same is not true for SIRVs in general. This result holds for two SIRVs, when they are statistically independent, having zero mean and when the covariance matrix of the first is within a multiplicative constant of the covariance matrix of the second.[15,16] More precisely, let $\mathbf{Y}_1 = [Y_{11},\ Y_{12},\ldots Y_{1N}]^{\mathrm{T}}$ and $\mathbf{Y}_2 = [Y_{21},\ Y_{22},\ldots Y_{2N}]^{\mathrm{T}}$ denote two independent zero mean SIRVs. The covariance matrix and characteristic PDF of $\mathbf{Y}_1$ are denoted by $\boldsymbol{\Sigma}_1$ and $f_{S_1}(s_1)$, respectively. The corresponding quantities for $\mathbf{Y}_2$ are denoted by $\boldsymbol{\Sigma}_2$ and $f_{S_2}(s_2)$.

We are interested in obtaining the distribution of the sum given by

$$\mathbf{Y} = \mathbf{Y}_1 + \mathbf{Y}_2 \tag{12.76}$$

The characteristic function of $\mathbf{Y}$ is given by

$$E[\exp(j\boldsymbol{\omega}^{\mathrm{T}}\mathbf{Y})] = g_1\left(\boldsymbol{\omega}^{\mathrm{T}}\boldsymbol{\Sigma}_1\boldsymbol{\omega}\right)g_2\left(\boldsymbol{\omega}^{\mathrm{T}}\sum\nolimits_2\boldsymbol{\omega}\right) \tag{12.77}$$

where $g_1(\cdot)$ and $g_2(\cdot)$ are the characteristic functions of $\mathbf{Y}_1$ and $\mathbf{Y}_2$, respectively. If $\mathbf{Y}$ is a zero mean SIRV, then its characteristic function has the form

$$E[\exp(j\boldsymbol{\omega}^{\mathrm{T}}\mathbf{Y})] = g\left(\boldsymbol{\omega}^{\mathrm{T}}\sum\boldsymbol{\omega}\right) \tag{12.78}$$

In order to write Equation 12.77 as a function of a single quadratic form, $\boldsymbol{\Sigma}_2$ must be within a multiplicative constant of $\boldsymbol{\Sigma}_1$.

12.2.5.5. Markov Property for SIRPs

An interesting property of SIRPs is that a zero mean wide sense stationary SIRP is Markov if and only if its autocorrelation function has the form

$$R(t_1, t_2) = \exp(-a|(t_1 - t_2|) \tag{12.79}$$

This result is well known for the special case of a zero mean wide sense stationary Gaussian random process. To demonstrate the more general result we consider N samples from a zero mean wide sense stationary SIRP $y(t)$. Let $\mathbf{Y} = [Y_1, Y_2, \ldots, Y_N]^{\mathrm{T}}$ denote the vector of successive samples obtained from the SIRP.

Given that $y(t)$ is a zero mean wide sense stationary Markov SIRP, we first show that its autocorrelation function must have the form of Equation 12.79. Let Y_1, Y_2, and Y_3 denote the random variables obtained by sampling $y(t)$ at time instants t_1, t_2, and t_3 such that $t_1 \leq t_2 \leq t_3$. Since $y(t)$ is a Markov process, the joint PDF of Y_1, Y_2, and Y_3 can be expressed as

$$f_{Y_1,Y_2,Y_3}(y_1, y_2, y_3) = f_{Y_1}(y_1)f_{Y_2|Y_1}(y_2|y_1)f_{Y_3|Y_2}(y_3|y_2) \tag{12.80}$$

The autocorrelation function $R(t_3, t_1) = E[Y_3Y_1]$ is given by

$$R(t_3, t_1) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y_3y_1f_{Y_1,Y_2,Y_3}(y_1, y_2, y_3)\mathrm{d}y_1\ \mathrm{d}y_2\ \mathrm{d}y_3 \tag{12.81}$$

Also,

$$R(t_2, t_2) = E[Y_2^2] = \int_{-\infty}^{\infty} y_2^2 f_{Y_2}(y_2)\mathrm{d}y_2 \tag{12.82}$$

Hence,

$$\begin{aligned} R(t_3, t_1)R(t_2, t_2) = &\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y_3y_1f_{Y_1,Y_2,Y_3}(y_1, y_2, y_3) \\ &\times \mathrm{d}y_1\ \mathrm{d}y_2\ \mathrm{d}y_3y_2^2f_{Y_2}(y_2)\mathrm{d}y_2 \end{aligned} \tag{12.83}$$

Using Equation 12.80 we can rewrite the above equation as

$$R(t_3,t_1)R(t_2,t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y_3 y_2 f_{Y_3,Y_2,}(y_3,y_2)\mathrm{d}y_3\ \mathrm{d}y_2 \times \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y_2 y_1 f_{Y_2,Y_1,}(y_2,y_1)\mathrm{d}y_2\ \mathrm{d}y_1 \tag{12.84}$$

Consequently,

$$R(t_3,t_1)R(t_2,t_2) = R(t_3,t_2)R(t_2,t_1) \tag{12.85}$$

The only nontrivial autocorrelation function satisfying this property is given by Equation 12.79.

Since $y(t)$ is a zero mean SIRP, it follows that $E[\mathbf{Y}] = 0$. Letting $b = \exp(-a)$, we can write the covariance matrix of $\mathbf{Y}$ as

$$\mathbf{\Sigma} = \begin{bmatrix} 1 & b & \ldots & b^{N-1} \\ b & 1 & \ldots & b^{N-2} \\ b^2 & b & \ldots & b^{N-3} \\ \ldots & \ldots & \ldots & \ldots \\ b^{N-1} & b^{N-2} & \ldots & 1 \end{bmatrix} \tag{12.86}$$

under the assumption that $[t_1, t_2, \ldots, t_N] = [1, 2, \ldots, N]$. We then make use of Equation 12.73 to obtain

$$E[Y_N | Y_{N-1}, Y_{N-2}, \ldots, Y_1] = [b^{N-1} b^{N-2}, \ldots, b]\mathbf{\Sigma}_{\mathbf{y}'}^{-1}\mathbf{Y}' \tag{12.87}$$

where $\mathbf{Y}' = [Y_1, Y_2, \ldots, Y_{N-1}]^{\mathrm{T}}$ and

$$\mathbf{\Sigma}_{\mathbf{y}'} = \begin{bmatrix} 1 & b & \ldots & b^{N-2} \\ b & 1 & \ldots & b^{N-3} \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ b^{N-2} & b^{N-3} & \ldots & 1 \end{bmatrix} \tag{12.88}$$

Recognizing that

$$\mathbf{\Sigma}_{\mathbf{y}'}^{-1} = \frac{1}{1-b^2}\begin{bmatrix} 1 & -b & 0 & \ldots & 0 & \ldots \\ -b & 1+b^2 & -b & 0 & \ldots & 0 \\ 0 & -b & 1+b^2 & \ldots & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & -b & 1+b^2 & -b \\ 0 & \ldots & \ldots & \ldots & -b & 1 \end{bmatrix} \tag{12.89}$$

Therefore, we can rewrite Equation 12.87 as

$$E[Y_N|Y_{N-1}, Y_{N-2}, ..., Y_1] = bY_{N-1} \tag{12.90}$$

From Equation 12.73, we also obtain

$$E[Y_N|Y_{N-1}] = bY_{N-1} \tag{12.91}$$

Clearly $E[Y_N|Y_{N-1}] = E[Y_N|Y_{N-1}, Y_{N-2}, ..., Y_1]$. Since this must be true for all choices of $Y_1, Y_2, ..., Y_{N-1}$, it follows that $f_{Y_N|Y_{N-1},Y_{N-2}...,Y_1}(y_N|y_{N-1}, y_{N-2}, ..., y_1) = f_{y_N|y_{N1}}(y_N|y_{N1})$. Hence, y(t) is Markov.

12.2.5.6. Kalman Filter for SIRPs

It has been shown by Chu in Ref. 29 that the Kalman filter for SIRPs is identical to the corresponding filter for a Gaussian random process. The model considered in Ref. 29 is given by

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{F}_k\mathbf{x}_k + \mathbf{G}_k\mathbf{w}_k \qquad k = 0, 1, ..., N-1 \\ \mathbf{y}_k &= \mathbf{H}_k\mathbf{x}_k + \mathbf{v}_k \qquad k = 0, 1, ..., N-1 \end{aligned} \tag{12.92}$$

where $\mathbf{x}_k$ denotes the state vector of the underlying process, $\mathbf{w}_k$ is its excitation vector, $\mathbf{y}_k$ denotes the observation vector, and $\mathbf{v}_k$ is the measurement noise. It is assumed that $\mathbf{x}_k$, $\mathbf{w}_k$, and $\mathbf{v}_k$ are jointly SIRP with a common characteristic PDF $f_S(s)$. Also, let

$$\begin{aligned} &E|\mathbf{x}_k| = \overline{\mathbf{x}_k} \qquad k = 0, 1, ..., N-1 \\ &E[(\mathbf{x}_k - \overline{\mathbf{x}_k})(\mathbf{x}_k - \overline{\mathbf{x}_k})^{\mathrm{T}}] = \mathbf{M}_k \\ &E[\mathbf{w}_k] = E[\mathbf{v}_k] = 0 \\ &E[(\mathbf{x}_k - \overline{\mathbf{x}_k})\mathbf{w}_k^{\mathrm{T}}] = E(\mathbf{x}_k - \overline{\mathbf{x}_k})\mathbf{v}_k^{\mathrm{T}}] = E[\mathbf{w}_k\mathbf{v}_k^{\mathrm{T}}] = 0 \\ &E[\mathbf{w}_k|\mathbf{w}_{km}] = \mathbf{Q}_k\delta_{l,m} \\ &E[\mathbf{v}_k|\mathbf{v}_{km}] = \mathbf{R}_k\delta_{l,m} \end{aligned} \tag{12.93}$$

where $\mathbf{w}_{km}$ and $\mathbf{v}_{km}$ are the mth components of $\mathbf{w}_k$ and $\mathbf{v}_k$, respectively, and $\delta_{l,m}$ is the Kronecker delta function. Hence, $\mathbf{x}_k$, $\mathbf{w}_k$, and $\mathbf{v}_k$ are mutually uncorrelated while $\mathbf{w}_k$ and $\mathbf{v}_k$ are each white with zero mean.

The innovations vectors is defined as

$$\breve{\mathbf{y}}_{k|k-1} = \mathbf{y}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1} \tag{12.94}$$

where $\hat{\mathbf{x}}_{k|k-1}$ is the MMSE estimate of $\mathbf{x}_k$ given the observation vectors up to $k-1$. The covariance matrix of the innovations can be shown to be

$$\mathrm{Cov}(\breve{\mathbf{y}}_{k|k-1}) = \mathbf{S}_{k|k-1} = (\mathbf{H}_k \mathbf{M}_k \mathbf{H}_k^{\mathrm{T}} + \mathbf{R}_k) \tag{12.95}$$

It can be readily shown that $\mathbf{x}_k$ and $\mathbf{y}_k$ are jointly SIRP. Therefore, the MMSE estimate of $\mathbf{x}_k$ given the observation vectors up to $k-1$ is a linear function of $\mathbf{y}_m$, $m = 1, 2, \ldots, k-1$, as shown by Equation 12.73. Hence, the Kalman filter equations for SIRPs are identical to those for the Gaussian case. The Kalman gain denoted by $\mathbf{K}_{k|k}$ is expressed as

$$\mathbf{K}_{k|k} = \mathbf{M}_k \mathbf{H}_k^{\mathrm{T}} \mathbf{S}_{k|k-1}^{-1} \tag{12.96}$$

The measurement update $\hat{\mathbf{x}}_{k|k}$ is given by

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_{k|k} \hat{\mathbf{y}}_{k|k-1} = (\mathbf{I} - K_{k|k}) \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_{k|k} \mathbf{y}_k \tag{12.97}$$

The covariance matrix of the error in the update can be written as

$$\mathbf{C}_k = \mathbf{M}_k - \mathbf{M}_k \mathbf{H}_k^{\mathrm{T}} (\mathbf{H}_k \mathbf{M}_k \mathbf{H}_k^{\mathrm{T}} + \mathbf{R}_k)^{-1} \mathbf{H}_k \mathbf{M}_k \tag{12.98}$$

The prediction is then given by

$$\hat{\mathbf{x}}_{k+1|k} = \mathbf{F}_k \hat{\mathbf{x}}_{k|k} \tag{12.99}$$

Finally, the covariance matrix of the prediction is expressed as

$$\mathbf{M}_{k+1} = \mathbf{F}_k \mathbf{C}_k \mathbf{F}_k^{\mathrm{T}} + \mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k^{\mathrm{T}} \tag{12.100}$$

When systems driven by nonGaussian noise are encountered in practice, under the assumption of joint SIRP, these equations provide an efficient computation formula for the Kalman filter.

12.2.5.7. Statistical Independence

We point out that the only case where the components of a SSRV are statistically independent occurs when the SSRV is Gaussian. This property is proved in Appendix U.1.

12.2.5.8. Ergodicity of SIRPs

It has been pointed out in Ref. 27 that an ergodic SIRP is necessarily Gaussian. The proof of the nonergodicity of SIRPs (except Gaussian) can be easily obtained using the representation theorem[15] for SIRPs, which states that an SIRP is a univariate randomization of the Gaussian random process. More precisely, if $y(t)$ is an SIRP, then it can be expressed as $y(t) = Sz(t)$, where S is a nonnegative random variable and $z(t)$ is a Gaussian random process. Clearly, if $z(t)$ is stationary, then $y(t)$ will also be stationary. However, different realizations of S result in different scale factors for the sample functions of $y(t)$. Therefore, time averages will differ from one sample function to another, and, in general, will not equal the corresponding ensemble average. Consequently, $y(t)$ cannot be ergodic. When S is a nonrandom constant, $y(t)$ is a Gaussian random process. Then $y(t)$

will be ergodic provided $z(t)$ is also ergodic. It is concluded that only Gaussian SIRPs can be ergodic.

12.2.6. Conclusion

In this chapter, we have presented an overview of the literature on the modelling of radar clutter and the theory of SIRPs. It is clear from this chapter that the PDF of a SIRV is uniquely determined by the specification of a mean vector, a covariance matrix, and a characteristic first order PDF. It is also seen that many interesting properties of Gaussian random processes extend readily to SIRPs. A major difference with nonGaussian SIRPs is their nonergodic behaviour. Consequently, time averages do not result in corresponding ensemble averages. However, if ensemble averages are used instead of time averages, then nonergodicity is not a serious problem. In the following chapters, we shall present the application of SIRPs for nonGaussian radar clutter modelling, simulation, and distribution approximation.

12.3. RADAR CLUTTER MODELLING USING SPHERICALLY INVARIANT RANDOM PROCESSES

12.3.1. Introduction

In this chapter we consider the use of the theory of SIRPs for modelling correlated nonGaussian radar clutter. It has been pointed out in Section 12.2 that radar clutter can be nonGaussian and that radars process N pulses at a time. Furthermore, the clutter can be highly correlated. Therefore, by clutter modelling we mean the specification of the joint PDF of the N correlated clutter samples. Since we are dealing with correlated clutter, the joint PDF cannot be constructed by simply taking the product of the marginal PDFs. This chapter presents a mathematically elegant and tractable approach for specifying the joint PDF of N clutter samples. In addition, we discuss the characterization of Gaussian and nonGaussian correlated random vectors, the need for a library of multivariate PDFs for modelling correlated nonGaussian clutter, several techniques for establishing this library, reduction of the problems of and, finally, reduction of the problems of distribution identification of multivariate correlated nonGaussian random vectors, to that of equivalent identification of a univariate distribution of nonnegative guadratic form.

Specifically, the problem of modelling a random vector obtained by sampling a stochastic process $y(t)$ at N time instants is of interest to us. The stochastic process may be real or complex. In addition, there is no restriction on the number of samples obtained or the sampling time instants. In order to completely characterize the random vector we need to specify the joint PDF of the N samples (real or complex) or, equivalently, specify the joint characteristic function. This problem is very well treated when the underlying stochastic process sin Gaussian. The joint PDF in this case can be written as $(2\pi)^{-\frac{N}{2}}|\mathbf{\Sigma}|^{-\frac{1}{2}}\exp(-p/2)$, where p is a nonnegative quadratic form given by $p = [\mathbf{y} - \boldsymbol{\mu}]^{\mathbf{T}}\mathbf{\Sigma}^{-1}[\mathbf{y} - \boldsymbol{\mu}]$. Here $\boldsymbol{\mu}$ and $\mathbf{\Sigma}$

denote the mean vector and covariance matrix, respectively, of the Gaussian random vector $\mathbf{Y}$ whose components are the N samples of $y(t)$. However, if $y(t)$ is not a Gaussian random process, there is no unique specification for the joint PDF of the N samples except when the samples are statistically independent.

When processing real world data, neither the Gaussian nature of the underlying stochastic process nor the statistical independence of the samples is guaranteed. In fact, it is likely that the samples may be correlated. Hence, we need to obtain multivariate nonGaussian PDFs when can model the correlation between samples. In practice, radar clutter can vary from one application to another. Therefore, we need to have available a library of possible multivariate nonGaussian PDFs so that an appropriate PDF can be chosen to approximate the data for each clutter scenario.

The theory of SIRPs provides us with elegant and mathematically tractable techniques to construct multivariate nonGaussian PDFs. SIRPs are generalizations of the familiar Gaussian random process. The PDF of every random vector obtained by sampling a SIRP is uniquely determined by the specification of a mean vector, a covariance matrix, and a characteristic first order PDF. In addition, the PDF of a random vector obtained by sampling a SIRP is a function of a nonnegative quadratic form. However, the PDF does not necessarily involve an exponential dependence on the quadratic form, as in the Gaussian case. Such a random vector is called a SIRV.

There are two kinds of models for nonGaussian radar clutter. One is called the endogenous model, where the desired nonGaussian process with prescribed envelope PDF and correlation function is realized by using a zero memory nonlinear transformation on a Gaussian process having a prespecified correlation function. In this approach it is not possible to independently control the envelope PDF and the correlation properties of the nonGaussian process. In addition, not all nonlinearities give rise to a nonnegative definite covariance matrix at their outputs.[33,34] The second model is called an exogenous product model.[13] In this model, the desired nonGaussian clutter is generated by the product of a Gaussian random process and an independent nonGaussian process, which may be highly correlated. In this scheme, the desired envelope PDF and the correlation properties can be controlled independently. The exogenous model can be thought of as slowly time variant nonGaussian process modulating a Gaussian random process. The SIRP is a special case of the exogenous model, arising when the modulating process does not change rapidly during the observation interval and can be approximated as a random variable. This is due to the fact that the representation theorem for SIRPs allows us to explicitly write the nonGaussian process as a product of a Gaussian process and a nonnegative random variable. By assuming statistical independence between the modulating random variable and the Gaussian process, it is possible to independently control the nonGaussian enveloped PDF and its correlation properties. The SIRP is the only known case of the exogenous multiplicative model allowing the specification of the Nth order PDF.

Section 12.3.2 outlines the problem of interest. In Section 12.3.3 we present several techniques to obtain SIRVs. Examples based on various techniques

described in Section 12.3.3 are used to obtain a library of SIRV PDFs in Section 12.3.4. Finally, in Section 12.3.5, we present a key result which characterizes SIRVs by using the quadratic form appearing in their PDFs.

12.3.2. Problem Statement

We assume we are dealing with coherent radar clutter. By coherent radar clutter, we mean that the clutter is processed in terms of its in-phase and out-of-phase quadrature components. Pre-detection radar clutter, being a bandpass random process, admits a representation of the form

$$y(t) = \mathrm{Re}\{\tilde{y}(t)\exp(j\omega_0 t)\} \tag{12.101}$$

where $\tilde{y}(t) = y_c(t) + jy_s(t)$ denotes the complex envelope of the clutter process, ω_0 is a known carrier frequency, $y_c(t)$ and $y_s(t)$ denote the in-phase and out-of-phase quadrature components of the complex process $\tilde{y}(t)$, respectively. Equation 12.101 can be rewritten as

$$y(t) = y_c(t)\cos(\omega_0 t) - y_s(t)\sin(\omega_0 t) \tag{12.102}$$

We are interested in specifying the joint PDF of N samples obtained by sampling the process $y(t)$. Since it is always more convenient to work with the associated low pass process, we consider the equivalent problem of specifying the PDF of N complex samples obtained from the complex process $\tilde{y}(t)$. The PDF of a complex random variable is defined to be the joint PDF of its in-phase and out-of-phase quadrature components. Therefore, it follows that the joint PDF of N complex random variables is the joint PDF of the $2N$ in-phase and out-of-phase quadrature components. While dealing with complex random variables, it is sometimes more convenient to work with their envelope and phase. The envelope R_i and phase Θ_i of a complex random variable $\tilde{Y} = Y_{ci} + \mathrm{j}Y_{si}$ are defined by

$$\begin{aligned} R_i &= \sqrt{Y_{ci}^2 + Y_{si}^2} \\ \Theta_i &= \arctan\left(\frac{Y_{si}}{Y_{ci}}\right) \end{aligned} \tag{12.103}$$

We consider the problem of specifying the PDF of a random vector $\mathbf{Y}^{\mathbf{T}} = [\mathbf{Y_c^T}:\mathbf{Y_s^T}]$ obtained by sampling the random process $\tilde{y}(t)$, where $\mathbf{Y_c} = [Y_{c1}, Y_{c2}, \ldots, Y_{cN}]^{\mathrm{T}}$ and $\mathbf{Y_s} = [Y_{s1}, Y_{s2}, \ldots, Y_{sN}]^{\mathrm{T}}$. The subscripts c and s denote the in-phase and out-of-phase quadrature components, respectively. We assume that the process $y(t)$ is a wide sense stationary random process. The necessary and sufficient conditions for $y(t)$ to be a wide sense stationary random process[30] are:

(A) The quadrature components have zero mean.
(B) The envelope of the pairwise quadrature components is statistically independent of the phase and the phase is uniformly distributed over

the interval $(0, 2\pi)$. This results in the pair wise quadrature components being identically distributed and their joint PDF being circularly symmetric. This also results in the orthogonality of the pairwise quadrature components at each sampling instant.

(C) The autocovariance function and crosscovariance function of the quadrature process of the complex process $\tilde{y}(t) = y_c(t) + jy_s(t)$ satisfy the condition given by

$$\begin{aligned} K_{cc}(\tau) &= K_{ss}(\tau) \\ K_{sc}(\tau) &= -K_{sc}(\tau) \end{aligned} \tag{12.104}$$

where:

$$\begin{aligned} K_{cc}(\tau) &= E\{X_c(t)X_c(t-\tau)\} \\ K_{ss}(\tau) &= E\{X_s(t)X_s(t-\tau)\} \\ K_{cs}(\tau) &= E\{X_c(t)X_s(t-\tau)\} \\ K_{sc}(\tau) &= E\{X_s(t)X_c(t-\tau)\} \end{aligned} \tag{12.105}$$

Any choice of autocovariance and crosscovariance functions is allowed as long as requirement (C) is satisfied and the resulting covariance matrix of $\mathbf{Y}$ is nonnegative definite.

Due to requirement (A), $E\{\mathbf{Y_c}\} = E\{\mathbf{Y_s}\} = 0$. If follows that $E\{\mathbf{Y}\} = 0$. As a consequence of requirements (B) and (C), the covariance matrix of $\mathbf{Y}$, given by

$$\mathbf{\Sigma} = \begin{bmatrix} \mathbf{\Sigma_{ss}} & | & \mathbf{\Sigma_{ss}} \\ — & | & — \\ \mathbf{\Sigma_{ss}} & | & \mathbf{\Sigma_{ss}} \end{bmatrix} \tag{12.106}$$

must satisfy the conditions:

$$\begin{aligned} \mathbf{\Sigma_{cc}} &= \mathbf{\Sigma_{ss}} \\ \mathbf{\Sigma_{cs}} &= -\mathbf{\Sigma_{sc}} \end{aligned} \tag{12.107}$$

with the elements of the main diagonal of the matrices $\mathbf{\Sigma_{cs}}$ and $\mathbf{\Sigma_{sc}}$ being equal to zero. Note that $\mathbf{\Sigma_{cc}} = E\{\mathbf{Y_c}\mathbf{Y_c^T}\}$, $\mathbf{\Sigma_{cs}} = E\{\mathbf{Y_c}\mathbf{Y_s^T}\}$, $\mathbf{\Sigma_{sc}} = E\{\mathbf{Y_s}\mathbf{Y_c^T}\}$, and $\mathbf{\Sigma_{ss}} = E\{\mathbf{Y_s}\mathbf{Y_s^T}\}$. Finally, we point out, regardless of the value of N, we always have an even order PDF when dealing with quadrature components. We are now in a position to proceed with the characterization of $\mathbf{Y}$ as a SIRV.

For a SIRV, it is pointed out that the PDF of a given order automatically implies all lower order PDFs. For example, if N random variables are jointly Gaussian, it is well known that the ith order PDF, $i = 1, 2, \ldots, N-1$, is Gaussian. This property of SIRVs is called internal consistency. The requirements (A) to (C) arising from the wide sense stationarity requirements of the process $y(t)$ are called external consistency conditions. Requirements (A) to (C) are not inherent to the SIRP and do not hold when the SIRP is not wide sense stationary.

12.3.3. Techniques for Determining the SIRV PDF

In this section, several techniques are presented for obtaining $h_{2N}(p)$ defined in Section 12.2. For convenience, temporal wide sense stationarity of the underlying bandpass process is assumed. However, the functional form of $h_{2N}(\cdot)$ is unaffected whether or not the random process is temporally wide sense stationary. Hence, it is allowable to let $p = (\mathbf{y} - \mathbf{b})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{y} - \mathbf{b})$ in the final result, in general, where $\mathbf{b}$ is any mean nonzero vector and $\boldsymbol{\Sigma}$ is any nonnegative definite matrix.

Recall from Section 12.2 that the PDF of a SIRV $\mathbf{Y}^{\mathrm{T}} = [\mathbf{Y}_{\mathrm{c}}^{\mathrm{T}}:\mathbf{Y}_{\mathrm{s}}^{\mathrm{T}}]$, with $\mathbf{Y}_{\mathrm{c}}$ and $\mathbf{Y}_{\mathrm{s}}$ as defined in Section 12.3.2 is given by

$$f_{\mathbf{Y}}(\mathbf{y}) = (2\pi)^{-N}|\boldsymbol{\Sigma}|^{-\frac{1}{2}}h_{2N}(p) \tag{12.108}$$

Assuming temporal wide sense stationarity, $p = \mathbf{y}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{y}$ where $\boldsymbol{\Sigma}$ is given by Equation 12.106. The mean vector of $\mathbf{Y}$ is zero due to requirement (A) in Section 12.3.2. The covariance matrix $\boldsymbol{\Sigma}$ having the form of Equation 12.106 and satisfying the requirements of Equation 12.107 is readily determined when the autocorrelation function of the process is specified. Given $\boldsymbol{\Sigma}$, several techniques for obtaining $h_{2N}(p)$ are presented in this section.

The representation theorem for SIRVs allows us to express $\mathbf{Y}$ as a product of a Gaussian random vector $\mathbf{Z}$, having the same dimensions as $\mathbf{Y}$ and a nonnegative random variable S. For the problem of radar clutter modelling, since it is desirable to control the nonGaussian nature of $\mathbf{Y}$ and its correlation properties independently, we assume that the random variable S is statistically independent of $\mathbf{Z}$. In addition, the covariance matrix of the SIRV can be made equal to the covariance matrix of the Gaussian random vector by requiring $E(S^2)$ to be unity. Finally, it is pointed out that the mean of $\mathbf{Z}$ is necessarily zero.

A physical interpretation can be given to $\mathbf{Z}$ and S. Consider a surveillance volume subdivided into contiguous range-Doppler-azimuth cells. Assuming a large enough cell size such that many scatterers are located in each cell, the N pulse returns from a given cell can be modelled as the Gaussian vector $\mathbf{Z}$ according to the central limit theorem. Also assume that the average clutter power remains constant over the N pulse returns in a coherent processing interval. However, the average clutter power is allowed to vary independently from cell to cell since different sets of scatterers are located in each cell. The variation of the average clutter power from cell to cell is modelled by the square of the nonnegative random variable S.

12.3.3.1. SIRVs with Known Characteristic PDF

We consider specification of the PDF of the SIRV $\mathbf{Y}$ when its characteristic PDF is known is closed form. We have pointed out in the previous section that the mean vector of $\mathbf{Y}$ is zero. Also, we have discussed the specification of the covariance matrix of $\mathbf{Y}$. Now, we shall focus on the specification of $h_{2N}(p)$.

As a consequence of the representation theorem, we can write.

$$h_{2N}(p) = \int_0^\infty s^{-2N} \exp\left(-\frac{p}{2s^2}\right) f_S(s)\, ds \tag{12.109}$$

Equation 12.109 enables us to specify $h_{2N}(p)$ when the characteristic PDF $f_S(s)$ is known is closed form. However, in some cases, even though an analytical expression is known for the characteristic PDF, it may be difficult to evaluate the integral in Equation 12.109 in closed form. In such instances, an alternate method for specifying $h_{2N}(p)$ must be examined. The method presented in the next section is useful for these cases.

12.3.3.2. SIRVs with Unknown Characteristic PDFs

When the characteristic PDF of the SIRV is unknown or when the integral in Equation 12.109 is difficult to evaluate, we propose an alternate method to obtain $h_{2N}(p)$. Recall that we are dealing with an even order PDF. Therefore, we can use Equation 12.30 starting with $h_2(q)$ to obtain $h_{2N}(q)$. It is worthwhile pointing out that $h_{2N}(\cdot)$ is related to the first order envelope PDF. From requirement (B) of Section 12.3.2, the joint PDF of the ith in-phase and out-of-phase quadrature components can be expressed as

$$f_{Y_{ci},Y_{si}}(y_{ci}, y_{si}) = (2\pi)^{-1}\sigma^{-2} h_2(p) \qquad i = 1, 2, \ldots, N \tag{12.110}$$

where $p = (y_{ci}^2 + y_{si}^2)/\sigma^2$ and σ^2 denotes the common variance of the in-phase and out-of-phase quadrature components. The envelope and phase corresponding to the ith quadrature components is given by

$$\begin{aligned} R_i &= \sqrt{Y_{ci}^2 + Y_{si}^2} \\ \Theta_i &= \arctan\frac{Y_{si}}{Y_{ci}} \end{aligned} \tag{12.111}$$

Due to the assumption of wide sense stationarity, we can drop the subscript i in Equation 12.111. The Jacobian of the transformation given by Equation 12.111 is $J = R^{-1}$, where J denotes the Jacobian. Using the Jacobian in Equation 12.110 results in the joint PDF of R and Θ being given by

$$f_{R,\Theta}(r, \theta) = \frac{r}{2\pi\sigma^2} h_2\left(\frac{r^2}{\sigma^2}\right) \tag{12.112}$$

Clearly, the joint PDF in Equation 12.112 can be factored as a product of the marginal PDFs of the random variables R and Θ. Consequently, the random variables R and Θ are statistically independent with PDFs given by

$$\begin{aligned} f_R(r) &= \frac{r}{\sigma^2} h_2\left(\frac{r^2}{\sigma^2}\right) \qquad 0 \le r \le \infty \\ f_\Theta(\theta) &= (2\pi)^{-1} \qquad 0 \le \theta \le 2\pi \end{aligned} \tag{12.113}$$

Equation 12.113 relates the envelope PDF to $h_2(\cdot)$. Hence, we can write

$$h_2\left(\frac{r^2}{\sigma^2}\right) = \frac{\sigma^2}{r} f_R(r) \tag{12.114}$$

Thus, Equation 12.114 provides a mechanism to obtain $h_2(q)$. Starting from $h_2(q)$, we then use Equation 12.30 to obtain $h_{2N}(q)$. Since not all nonGaussian envelope PDFs are admissible for characterization as SIRVs, we must check that $h_2(q)$ and its derivatives satisfy the monotonicity conditions stated in Section 12.2. Finally, $h_{2N}(p)$ is obtained by simply replacing q by $p = (\mathbf{y} - \mathbf{b})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mathbf{b})$ in $h_{2N}(q)$.

12.3.3.3. Hankel Transform Approach

In this section we present an approach based on the Hankel transform for specifying $h_{2N}(p)$. Recall that the joint PDF of the ith in-phase and out-of-phase quadrature components of $\mathbf{Y}$ is given by Equation 12.110. For convenience, it is assumed that σ^2 is unity. Dropping the subscript i from Equation 12.110, the joint characteristic function of Y_{ci} and Y_{si} is expressed as

$$\phi_{Y_c,Y_s}(\omega_1,\omega_2) = (2\pi)^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(j\omega_1 y_c + j\omega_2 y_s) h_2(y_c^2 + y_s^2) \mathrm{d}y_c \, \mathrm{d}y_s \tag{12.115}$$

Introducing the transformations

$$\begin{aligned} R &= \sqrt{Y_c^2 + Y_s^2} \\ \Theta &= \arctan \frac{Y_s}{Y_c} \\ \omega &= \sqrt{\omega_1^2 + \omega_2^2} \\ \alpha &= \arctan \frac{\omega_2}{\omega_1} \end{aligned} \tag{12.116}$$

we can rewrite Equation 12.115 as

$$\begin{aligned} \phi_{Y_c,Y_s}(\omega_1,\omega_2) = (2\pi)^{-1} \int_0^{\infty} \int_{-\infty}^{2\pi} \exp[j\omega r\{\cos(\theta)\cos(\alpha) \\ + \sin(\theta)\sin(\alpha)\}] r h_2(r^2) \mathrm{d}r \, \mathrm{d}\theta \end{aligned} \tag{12.117}$$

Noting that $\cos(A - B) = \cos(A)\cos(B) + \sin(A)\sin(B)$, we can rewrite Equation 12.117 as

$$\phi_{Y_c,Y_s}(\omega_1,\omega_2) = (2\pi)^{-1} \int_0^{\infty} \int_0^{2\pi} \exp[j\omega r \cos(\theta - \alpha)] r h_2(r^2) \mathrm{d}r \, \mathrm{d}\theta \tag{12.118}$$

Interchanging the order of integration in Equation 12.118, and recognizing that[35]

$$J_0(x) = \frac{1}{2\pi}\int_0^{2\pi} \exp[jx\cos(\beta-\gamma)]\mathrm{d}\beta \tag{12.119}$$

where $J_0(x)$ is the Bessel function of order zero, we have

$$\phi_{Y_c,Y_s}(\omega_1,\omega_2) = \int_0^{\infty} rh_2(r^2)J_0(\omega r)\mathrm{d}r \tag{12.120}$$

From Equation 12.120, it is clear that the joint characteristic function of Y_c and Y_s is a function of $\omega = \sqrt{\omega_1^2+\omega_2^2}$. Hence, it is a circularly symmetric characteristic function. Denoting this function by $\psi(\omega)$, we can write

$$\psi(\omega) = \int_0^{\infty} rh_2(r^2)J_0(\omega r)\mathrm{d}r \tag{12.121}$$

Equation 12.121 is recognized as the Hankel transform of order zero of $h_2(r^2)$. Using the inverse Hankel transform, we obtain

$$h_2(r^2) = \int_0^{\infty} \omega\psi(\omega)J_0(\omega r)\mathrm{d}\omega \tag{12.122}$$

Introducing the dummy variable ω, we can write

$$h_2(q) = \int_0^{\infty} \omega\psi(\omega)J_0(\omega\sqrt{q})\mathrm{d}\omega \tag{12.123}$$

We can then use Equation 12.130 to obtain $h_{2N}(q)$. More explicitly, we can write

$$h_{2N}(q) = (-2)^{N-1}\int_0^{\infty} \omega\psi(\omega)\frac{\mathrm{d}^{N-1}}{\mathrm{d}q^{N-1}}[J_0(\omega\sqrt{q})]\mathrm{d}\omega \tag{12.124}$$

Using the identity[35]

$$\frac{\mathrm{d}J_0(\eta)}{\mathrm{d}\eta} = -J_1(\eta) \tag{12.125}$$

we have

$$\frac{\mathrm{d}J_0(\omega\sqrt{q})}{\mathrm{d}q} = -\frac{\omega}{2}\omega^{-\frac{1}{2}}J_1(\omega\sqrt{q}) \tag{12.126}$$

Use of the recurrence relation[35]

$$\frac{\mathrm{d}}{\mathrm{d}\eta}[\eta^{-\alpha}J_\alpha(\eta)] = -\eta^{-\alpha}J_{\alpha+1}(\eta) \tag{12.127}$$

results in

$$\frac{d^2}{dq^2}[J_0(\omega\sqrt{q})] = \frac{\omega^2}{4}(\sqrt{q})^{-2}J_2(\omega\sqrt{q}) \tag{12.128}$$

Repeated use of Equation 12.127 gives

$$\frac{d^{N-1}}{dq^{N-1}}[J_0(\omega\sqrt{q})] = (-1)^{N-1}\frac{\omega^{N-1}}{2^{N-1}}(\sqrt{q})^{-N+1}J_{N-1}(\omega\sqrt{q}) \tag{12.129}$$

Substituting Equation 12.129 into Equation 12.124 gives

$$h_{2N}(q) = (\sqrt{q})^{1-N}\int_0^\infty \omega^N \psi(\omega)J_{N-1}(\omega\sqrt{q})d\omega \tag{12.130}$$

Finally, $h_{2N}(p)$ is obtained from Equation 12.130 by replacing q by $p = (\mathbf{y}-\mathbf{b})^T\boldsymbol{\Sigma}^{-1}(\mathbf{y}-\mathbf{b})$. This completes the proof of Equation 12.32 for even values of N, which was previously deferred. The integral in Equation 12.130 is recognized as the Hankel transform of order $N-1$ of $\psi(\omega)$. A number of Hankel transforms have been provided in Ref. 36 and these will be made use of in the examples presented in Section 12.3.4.

12.3.4. Examples of Complex SIRVs

This section presents examples based on the approaches discussed in Section 12.3.3 and is divided into three parts. In Section 12.3.4.1, we present examples that assume the knowledge of the characteristic PDF. In Section 12.3.4.2, the marginal envelope PDF is assumed to be known whereas in Section 12.3.4.3, knowledge of the marginal characteristic function is assumed. Finally, at the end of Section 12.3.4.3 we point out some univariate PDFs that cannot be generalized to SIRV characterization. We consider the problem of determining the PDF of the random vector $\mathbf{Y}^T = [\mathbf{Y}_c^T : \mathbf{Y}_s^T]$ specified in Section 12.3.2. It is assumed that the mean vector of $\mathbf{Y}$ and its covariance matrix $\boldsymbol{\Sigma}$ are known. Consequently, specification of the PDF of $\mathbf{Y}$ of the form of Equation 12.108 reduces to determination of $h_{2N}(p)$.

12.3.4.1. Examples Based on the Characteristic PDF

12.3.4.1.1. Gaussian Distribution

The Gaussian marginal PDF for the quadrature components having mean b_k and variance σ_k^2 is

$$f_{Y_k}(y_k) = \frac{1}{\sqrt{(2\pi)\sigma_k}}\exp\left(-\frac{(y_k-b_k)^2}{2\sigma_k^2}\right) \qquad -\infty \le y_k \le \infty \tag{12.131}$$

The characteristic PDF for this example is given by

$$f_S(s) = \delta(s-1) \tag{12.132}$$

where $\delta(\cdot)$ is the unit impulse function. Using Equation 12.21, it is seen that the resulting $h_N(p)$ is given by

$$h_N(p) = \exp\left(-\frac{p}{2}\right) \tag{12.133}$$

where $p = (\mathbf{y} - \mathbf{b})^T \mathbf{\Sigma}^{-1} (\mathbf{y} - \mathbf{b})$. The corresponding PDF for any N is given by Equation 12.20. For $N = 1$, this result reduces to Equation 12.131. When $\mathbf{Y}$ is made up of quadrature components, we obtain the corresponding $h_{2N}(p)$ by simply replacing N by $2N$ in Equation 12.133. Whenever a characteristic PDF can be made to approach a unit impulse function displaced to the right of the origin by appropriate choice of its parameters, it follows that the corresponding SIRV PDF will approach the Gaussian PDF.

12.3.4.1.2. K-Distribution

The K-distributed envelope PDF, by definition, is given by

$$f_R(r) = \frac{2b}{\Gamma(\alpha)} \left(\frac{br}{2}\right)^{\alpha} K_{\alpha-1}(br) u(r) \tag{12.134}$$

where α is the shape parameter of the distribution, b denotes the scale parameter of the distribution, $K_N(t)$ is the Nth order modified Bessel function of the second kind, and $u(r)$ is the unit step function. The K-distributed enveloped PDF is commonly used for modelling radar clutter PDFs that have extended tails.[2, 9, 19, 20] In particular, the PDF becomes heavy tailed as α approaches zero. Plots of Equation 12.134 for several values of α are shown in Figure 12.1 to Figure 12.4.

The K-distributed envelope PDF arises when we consider the product of a Rayleigh distributed random variable R' and an independent Chi-distributed random variable V. More precisely, we consider the product $R = R'V$, with R'

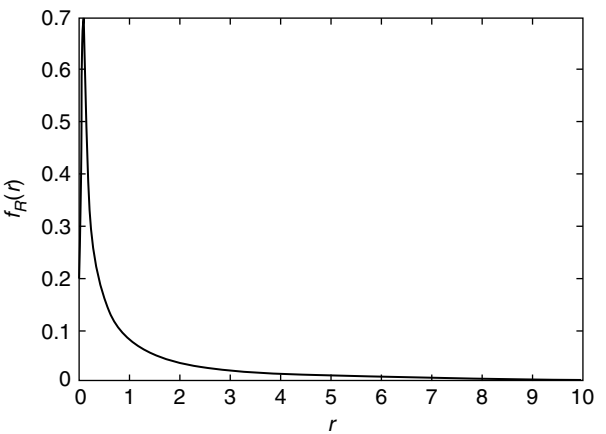

FIGURE 12.1 K-Distribution, $b = 0.31$, $\alpha = 0.05$.

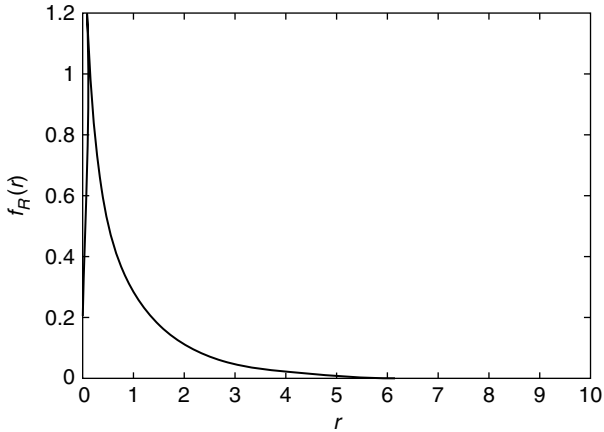

FIGURE 12.2 K-Distribution, $b = 0.77$, $\alpha = 0.3$.

and V being statistically independent. Their PDFs are given by

$$f_{R'}(r') = r' \exp\left(-\frac{[r']^2}{2}\right) \qquad 0 \le r' \le \infty \tag{12.135}$$

and

$$f_V(v) = \frac{2b}{\Gamma(\alpha)2^{\alpha}}(bv)^{2\alpha-1}\exp\left(-\frac{b^2v^2}{2}\right) \qquad 0 \le v < \infty \tag{12.136}$$

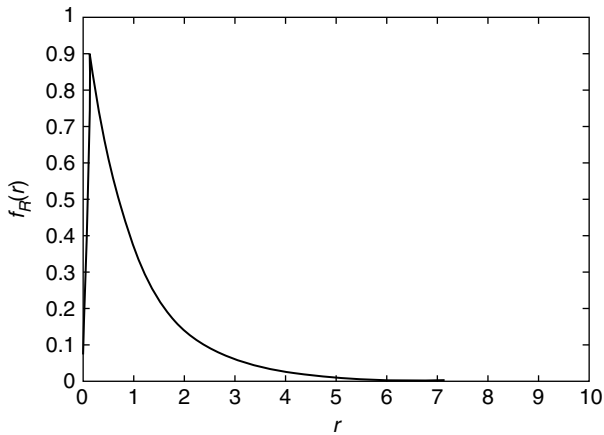

FIGURE 12.3 K-Distribution, $b = 1.0$, $\alpha = 0.5$.

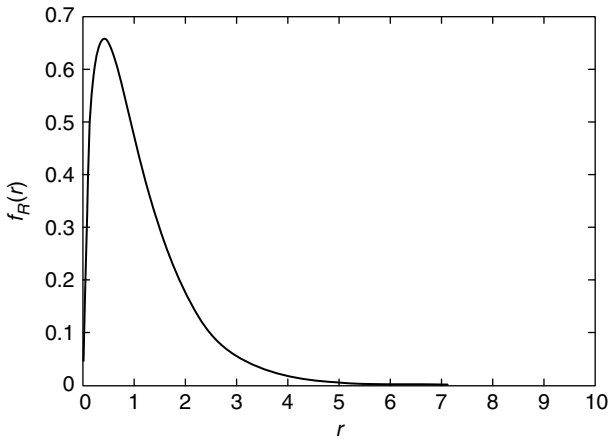

FIGURE 12.4 *K*-Distribution, $b = 1.4$, $\alpha = 0.99$.

respectively. Consequently, the PDF of R is given by

$$f_R(r) = \int_0^\infty f_{R|V}(r|v) f_V(v) \mathrm{d}v$$

$$= \int_0^\infty \frac{r}{v^2} \exp\left(-\frac{r^2}{2v^2}\right) \frac{2b}{\Gamma(\alpha)2^\alpha} (bv)^{2\alpha-1} \exp\left(-\frac{b^2v^2}{2}\right) \tag{12.137}$$

From Ref. 35 we have

$$K_v(xz) = \frac{z^v}{2} \int_0^\infty \exp\left[-\frac{x}{2}\left(t + \frac{z^2}{t}\right)\right] t^{-v-1} \mathrm{d}t \quad \left[|\arg z| < \frac{\pi}{4}\right],\ z > 0 \tag{12.138}$$

Letting $v^2 = t$ in Equation 12.137 and using the result of Equation 12.138, the PDF of Equation 12.134 follows.

The quadrature components corresponding to the Rayleigh envelope PDF are independent identically distributed zero mean Gaussian random variables having unit variance. The PDF of the quadrature components corresponding to R' is expressed as

$$f_{Z_c}(z) = f_{Z_s}(z) = (2\pi)^{-\frac{1}{2}} \exp\left(-\frac{z^2}{2}\right) \tag{12.139}$$

where Z_c and Z_s denote the in-phase and out-of-phase quadrature components, respectively. The quadrature components arising from the *K*-distributed envelope PDF, denoted by Y_c and Y_s, respectively, can be expressed as

$$\begin{aligned} Y_c &= Z_c V \\ Y_s &= Z_s V \end{aligned} \tag{12.140}$$

Note that $|\tilde{Y}| = |\tilde{Z}|V$ and $\Theta_{\tilde{Y}} = \Theta_{\tilde{z}}$. Consequently, the PDF of Y_c is given by

$$f_{Y_c}(y_c) = \frac{b^{2\alpha}}{\sqrt{2\pi}\Gamma(\alpha)2^{\alpha-1}} \int_0^\infty v^{2\alpha-2}\exp\left[-\frac{1}{2}\left(\frac{y_c^2}{v^2} + b^2v^2\right)\right]dv \qquad (12.141)$$

Making the change of variables $t = b^2v^2$ and $z^2 = b^2y_c^2$, and using Equation 12.138, the PDF of Y_c is expressed as

$$f_{Y_c}(y_c) = \frac{2b}{\Gamma(\alpha)\sqrt{2\pi}2^\alpha}|by_c|^{\alpha-\frac{1}{2}}K_{\frac{1}{2}-\alpha}(b|y_c|) \qquad -\infty < y_c < \infty \qquad (12.142)$$

where the absolute value denoted by $|\cdot|$ is used on account of the requirement that $z > 0$. In a similar manner, it can be shown that the PDF of Y_s has the same functional form as Equation 12.142. The PDF of Equation 12.142 is called the Generalized Laplace PDF.[16]

The characteristic PDF for the *K*-distributed SIRV is

$$f_S(s) = \frac{2b}{\Gamma(\alpha)2^\alpha}(bs)^{2\alpha-1}\exp\left(-\frac{b^2s^2}{2}\right)u(s) \qquad (12.143)$$

Using Equation 12.21 and Equation 12.138,

$$h_N(p) = \int_0^\infty s^{-N}\exp\left(-\frac{p}{2s^2}\right)\frac{2b}{\Gamma(\alpha)2^\alpha}(bs)^{2\alpha-1}\exp\left(-\frac{b^2s^2}{2}\right)ds \qquad (12.144)$$

Making the change of variables $t = b^2s^2$ and $z^2 = b^2p$, the resulting $h_N(p)$ is given by

$$h_N(p) = \frac{b^N}{\Gamma(\alpha)}\frac{(b\sqrt{p})^{\alpha-\frac{N}{2}}}{2^{\alpha-1}}K_{\frac{N}{2}-\alpha}(b\sqrt{p}) \qquad (12.145)$$

The corresponding SIRV PDF for any N is given by using Equation 12.20. For the case when $N = 1$, this reduces to Equation 12.142. When dealing with quadrature components, we use Equation 12.145 with N replaced by $2N$

12.3.4.1.3. Student-t Distribution

The Student-*t* distribution for the quadrature components is given by

$$f_{Y_k}(y_k) = \frac{\Gamma(v+\frac{1}{2})}{b\sqrt{\pi}\Gamma(v)}\left(1 + \frac{y_k^2}{b^2}\right)^{-v-\frac{1}{2}} \qquad (-\infty < y_k < \infty),\ \ v > 0 \qquad (12.146)$$

where b is the scale parameter, v is the shape parameter $\Gamma(v)$ is the Eulero–Gamma function and $k = c, s$. Plots of the Student-*t* distribution are shown for several values of v in Figures 12.5 to 12.8. The characteristic PDF for this example is

$$f_S(s) = \frac{2}{\Gamma(v)}\left(\frac{1}{2}\right)^v b^{2v-1}(s^{-1})^{2v+1}\exp\left(-\frac{b^2}{2s^2}\right)u(s) \qquad (12.147)$$

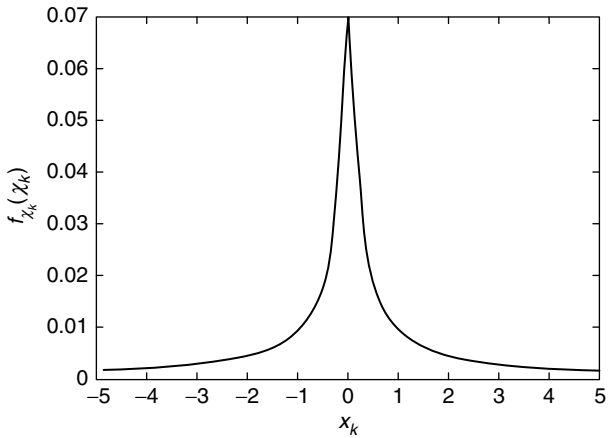

FIGURE 12.5 Student-t distribution, $b = 0.14$, $\upsilon = 0.01$.

Use of Equation 12.21 results in $h_N(p)$ being given by

$$h_N(p) = \frac{2^{\frac{N}{2}} b^{2v} \Gamma(v + \frac{N}{2})}{\Gamma(v)(b^2 + p)^{\frac{N}{2}+v}} \tag{12.148}$$

The corresponding SIRV PDF for any N is given by Equation 12.20. For $N = 1$, this result reduces to Equation 12.146. When dealing with quadrature components, we make use of Equation 12.148 with N replaced by $2N$.

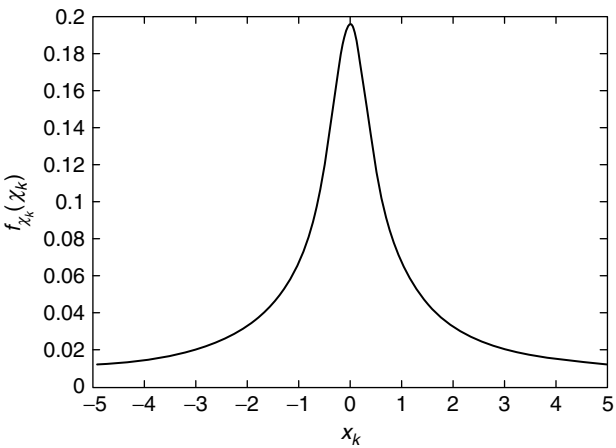

FIGURE 12.6 Student-t distribution, $b = 0.45$, $\upsilon = 0.1$.

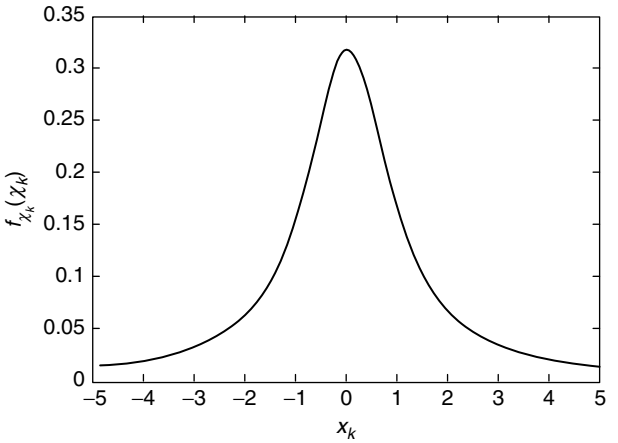

FIGURE 12.7 Student-t distribution, $b = 1.0$, $\upsilon = 0.5$.

12.3.4.1.4. Mixture of Gaussian PDFs

An interesting nonGaussian marginal PDF that is admissible as a SIRV is the mixture of Gaussian PDFs. We consider the PDF given by

$$f_{Y_k}(y_k) \sum a_i (2\pi k_i^2)^{-\frac{1}{2}} \exp\left(- \frac{(y_k - b_k)^2}{2k_i^2} \right) \tag{12.149}$$

for the quadrature components of $\mathbf{Y}$. The characteristic PDF for this example is given by

$$f_S(s) = \sum a_i \delta(s - k_i) \tag{12.150}$$

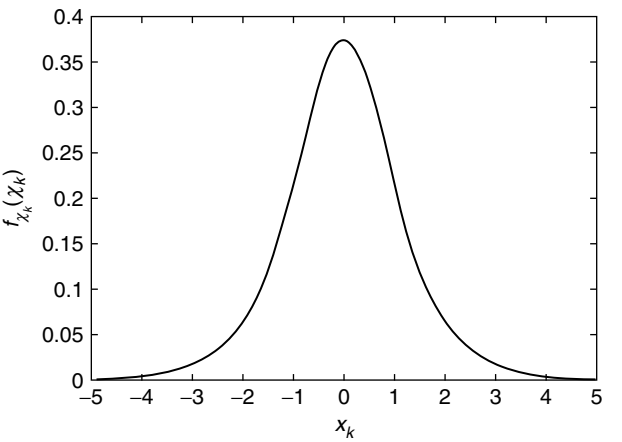

FIGURE 12.8 Student-t distribution, $b = 2.23$, $\upsilon = 2.5$.

Note that S is a discrete random variable, with a_i denoting the probability $P(S = k_i)$. Also, it is required that

$$\begin{aligned} a_i &\geq 0 \qquad i = 1, 2, \ldots \\ \sum_i a_i &= 1 \end{aligned} \tag{12.151}$$

Using Equation 12.21, it is seen that

$$h_N(p) = \sum_i k_i^{-N} a_i \exp\left(-\frac{p}{2k_i^2}\right) \tag{12.152}$$

The corresponding SIRV PDF for any N is given by Equation 12.20. For $N = 1$, this result reduces to Equation 12.49. When dealing with quadrature components, we make use of the result of Equation 12.152 with N replaced by $2N$. Note that the a_i's can be assigned any desired discrete distribution.

12.3.4.2. Examples Based on Marginal Envelope PDF

We shall report here on some new SIRV PDFs obtained starting from the marginal envelope PDF. Note in general, that the characteristic PDF for all the examples considered here are not available in closed form. Since σ^2 is the common variance of the in-phase and out-of-phase quadrature components, σ^2 is equal to $\frac{1}{2}E(R^2)$. In addition, recall that the binomial coefficient is defined by

$$\binom{l}{i} = \frac{l!}{i!(l-i)!} \tag{12.153}$$

In all the examples in this section, we start with $h_2(q)$ and obtain $h_{2N}(q)$ by the process of successive differentiation. The corresponding $h_{2N}(p)$ for each example is obtained by replacing q by p in $h_{2N}(q)$. In all the examples presented in this section note that the envelope PDFs reduce to the Rayleigh envelope PDF for appropriately chosen parameters.

12.3.4.2.1. Chi Envelope PDF

We consider the Chi distributed envelope PDF given by

$$f_R(r) = \frac{2b}{\Gamma(v)}(br)^{2v-1}\exp(-b^2r^2) \qquad 0 \leq r \leq \infty \tag{12.154}$$

where b denotes the scale parameter and v denotes the shape parameter. Plots of the Chi envelope PDF are shown in Figures 12.8 to 12.10 for several values of v. Using Equation 12.114, we can write

$$h_2(q) = \frac{2}{\Gamma(v)}(b\sigma)^{2v}q^{v-1}\exp(-b^2\sigma^2 q) \tag{12.155}$$

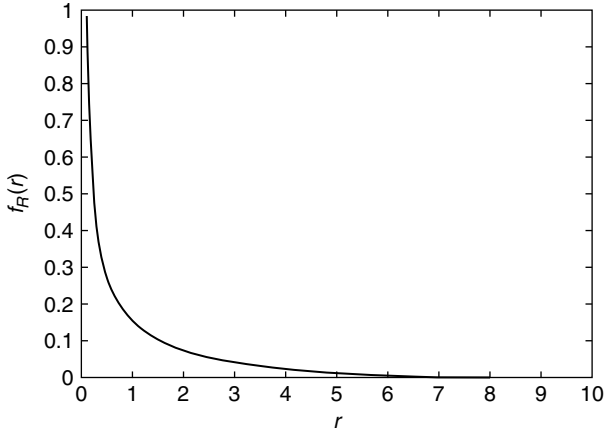

FIGURE 12.9 Chi envelope PDF, $b = 0.22$, $v = 0.1$.

Using Equation 12.30, we have

$$h_{2N}(q) = (-2)^{N-1} \frac{\mathrm{d}^{N-1} h_2(q)}{\mathrm{d}q^{N-1}}$$

$$= \frac{(-2)^{N-1}}{\Gamma(v)} 2(b\sigma)^{2v} \frac{\mathrm{d}^{N-1}}{\mathrm{d}q^{N-1}} [q^{v-1} \exp(-b^2\sigma^2 q)] \tag{12.156}$$

Recall Leibnitz's theorem for the nth derivative of a product,[35] which states that

$$\frac{\mathrm{d}^n(uv)}{\mathrm{d}x^n} = \sum_{k=0}^{n} \binom{n}{k} \frac{\mathrm{d}^k u}{\mathrm{d}x^k} \frac{\mathrm{d}^{n-k} v}{\mathrm{d}x^{n-k}} \tag{12.157}$$

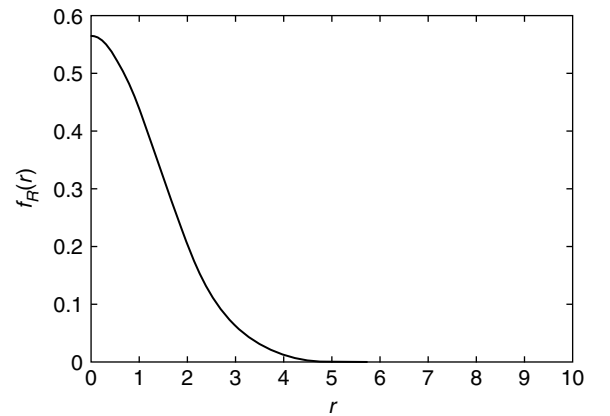

FIGURE 12.10 Chi envelope PDF, $b = 0.5$, $v = 0.5$.

where u and v are functions of x. Noting that

$$\frac{\mathrm{d}^k(q^{v-1})}{\mathrm{d}q^k} = \frac{\Gamma(v)}{\Gamma(v-k)} q^{v-k-1} \tag{12.158}$$

it follows that

$$h_{2N}(q) = (-2)^{N-1} A \sum_{k=1}^{N} G_k q^{v-k} \exp(-Bq) \tag{12.159}$$

where:

$$\begin{aligned} G_k &= \binom{N-1}{k-1} (-1)^{N-k} B^{N-k} \frac{\Gamma(v)}{\Gamma(v-k+1)} \\ A &= \frac{2}{\Gamma(v)} (b\sigma)^{2v} \\ B &= b^2\sigma^2 \end{aligned} \tag{12.160}$$

An important condition is that the SIRV PDF is valid only for $v \leq 1$. This is due to the fact that $h_2(p)$ and its derivatives are monotonically decreasing functions only in the range of values of v mentioned above. Finally, for $v = 1$, note that the Chi envelope PDF reduces to the Rayleigh envelope PDF. The corresponding SIRV PDF then becomes Gaussian.

12.3.4.2.2. Weibull Envelope PDF

The Weibull distributed envelope PDF is given by

$$f_R(r) = abr^{b-1} \exp(-ar^b) \qquad 0 \leq r \leq \infty \tag{12.161}$$

where a is the scale parameter and b is the shape parameter. Plots of the Weibull distribution for several values of b are shown in Figures 12.11 to 12.13. Using Equation 12.114, we have

$$h_2(q) = ab\sigma^b q^{\frac{b}{2}-1} \exp(-a\sigma^b q^{\frac{b}{2}}) = (-2)\frac{\mathrm{d}}{\mathrm{d}q}[\exp(-Aq^{\frac{b}{2}})] \tag{12.162}$$

where $A = a\sigma^b$. From Equation 12.30, we have

$$h_{2N}(q) = (-2)^N \frac{\mathrm{d}^N}{\mathrm{d}q^N}[\exp(-Aq^{\frac{b}{2}})] \tag{12.163}$$

The rule for obtaining the Nth derivative of a composite function is[35]: If $f(x) = F(y)$ and $y = \varphi(x)$, then

$$\frac{\mathrm{d}^N}{\mathrm{d}x^N}\{f(x)\} = \sum_{k=1}^{N} \frac{U_k}{k!} \frac{\mathrm{d}^k}{\mathrm{d}y^k}[F(y)] \tag{12.164}$$

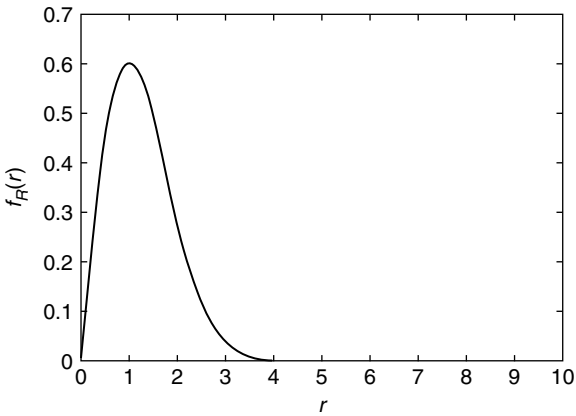

FIGURE 12.11 Chi envelope PDF, $b = 0.7$, $v = 1.0$.

where

$$U_k = \sum_{m=1}^{k} (-1)^{k-m} \binom{k}{m} y^{k-m} \frac{\mathrm{d}^N y^m}{\mathrm{d}x^N} \tag{12.165}$$

Making the association $x = q$ and $y = -Aq^{\frac{b}{2}}$, we have

$$h_{2N}(q) = \sum_{k=1}^{N} C_k q^{\frac{kb}{2}-N} \exp(-Aq^{\frac{b}{2}}) \tag{12.166}$$

where

$$C_k = \sum_{m=1}^{k} (-1)^{m+N} 2^N \frac{A^k}{k!} \binom{k}{m} \frac{\Gamma(1+\frac{mb}{2})}{\Gamma(1+\frac{mb}{2}-N)} \tag{12.167}$$

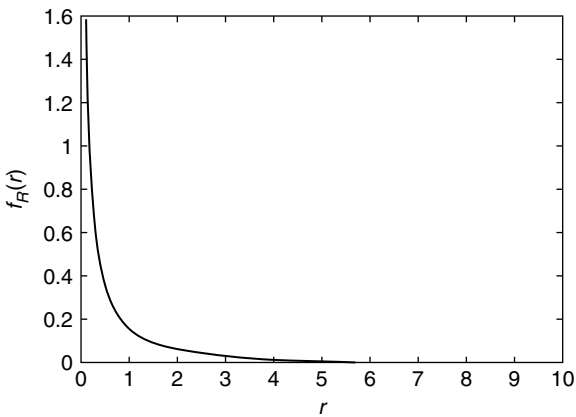

FIGURE 12.12 Weibull distribution envelope PDF, $b = 0.5$, $\alpha = 1.86$.

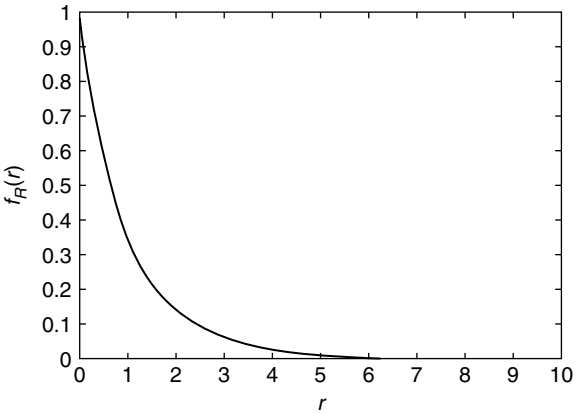

FIGURE 12.13 Weibull distribution envelope PDF, $b = 1.0$, $\alpha = 1$.

The Weibull envelope PDF is admissible for characterization as a SIRV for values of b less than or equal to 2. This is due to the fact that $h_2(q)$ and its derivatives fail to satisfy the monotonicity condition for other values of b. However, this is not a serious restriction for the point of view of radar clutter modeling because the Weibull envelope PDF is of interest in modeling large tailed clutter. Such a situation arises only when $0 < b \leq 2$. The Weibull envelope PDF reduces to the Rayleigh envelope PDF when $b = 2$. The corresponding SIRV PDF then becomes Gaussian. Another case of interest arises when $b = 1$. In this case the Weibull envelope PDF corresponds to the Exponential distributed envelope PDF (Figure 12.14).

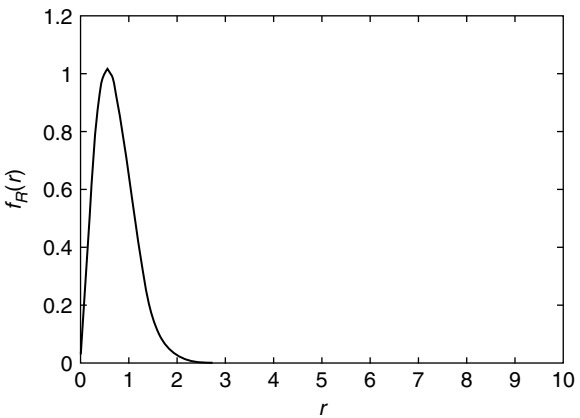

FIGURE 12.14 Exponential distributed envelope, $b = 2.0$, $\alpha = 0.5$.

12.3.4.2.3. Generalized Rayleigh Envelope PDF

The next PDF considered is for the Generalized Rayleigh envelope given by

$$f_{\mathrm{R}}(r) = \frac{\alpha r}{\beta^2 \Gamma\left(\frac{2}{\alpha}\right)} \exp\left[-\left(\frac{r}{\beta}\right)^{\alpha}\right] \qquad 0 \le r \le \infty \tag{12.168}$$

where α is the shape parameter and β is the scale parameter. Plots of the Generalized Rayleigh distribution are shown for several values of α in Figures 12.15 to 12.17.

Proceeding as in the previous example, we find that

$$h_2(q) = A \exp(-Bq^{\frac{\alpha}{2}}) \tag{12.169}$$

where:

$$\begin{aligned} A &= \frac{\sigma^2 \alpha}{\beta^2 \Gamma\left(\frac{2}{\alpha}\right)} \\ B &= \beta^{-\alpha} \sigma^{\alpha} \end{aligned} \tag{12.170}$$

Using Equation 12.25, Equation 12.163, and Equation 12.164, we have

$$h_{2N}(q) = \sum_{k=1}^{N-1} D_k q^{\frac{k\alpha}{2} - N + 1} \exp(-Bq^{\frac{\alpha}{2}}) \tag{12.171}$$

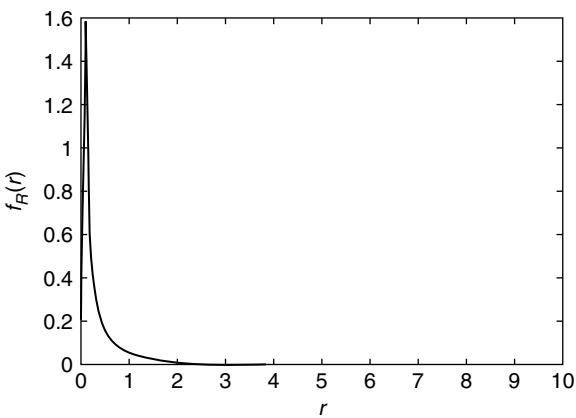

FIGURE 12.15 Generalized Rayleigh distributed envelope PDF, $\alpha = 0.1$, $\beta = 3.45 \times 10^{-15}$.

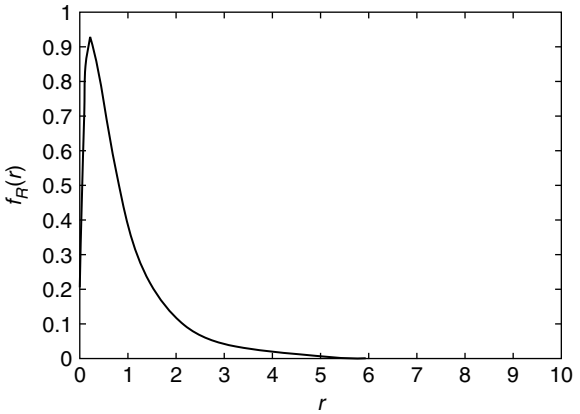

FIGURE 12.16 Generalized Rayleigh distributed envelope PDF, $\alpha = 0.5$, $\beta = 0.048$.

where

$$D_k = \sum_{m=1}^{k} (-1)^{m+N-1} 2^{N-1} \frac{B^k}{k!} \binom{k}{m} \frac{\Gamma\left(1 + \frac{m\alpha}{2}\right)}{\Gamma\left(2 + \frac{m\alpha}{2} - N\right)} \tag{12.172}$$

Note that the SIRV PDF is valid only in the range ($0 \leq \alpha \leq 2$). This is because of the fact that the monotonicity conditions for the derivatives of $h_2(p)$ are satisfied only for the specified range of α. The Generalized Rayleigh envelope PDF reduces to the Rayleigh envelope PDF when $\alpha = 2$.

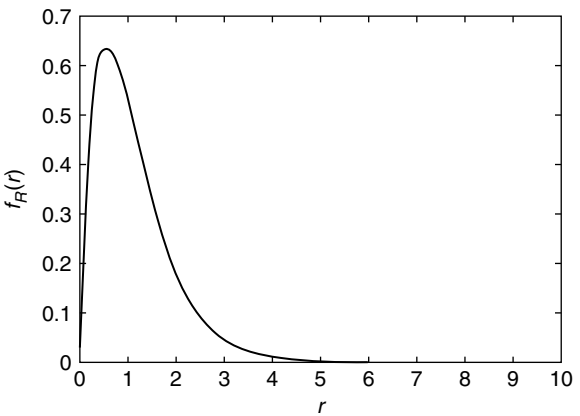

FIGURE 12.17 Generalized Rayleigh distributed envelope PDF, $\alpha = 1$, $\beta = 0.577$.

12.3.4.2.4. Rician Envelope PDF

There are two possible ways in which the Rician envelope PDF occurs. One possibility arises through a complex zero mean random process with correlated quadrature components that are Gaussian. The other is through a nonzero mean complex Gaussian process. The former case is considered here, since the SIRV PDF can be obtained by differentiation of $h_2(q)$. For this case, the envelope PDF is given by (Figure 12.18)

$$f_R(r) = \frac{r}{\sqrt{1-\rho^2}} \exp\left[-\frac{r^2}{2(1-\rho^2)} \right] I_0\left[\frac{\rho r^2}{2(1-\rho^2)} \right] \qquad (12.173)$$
$$(0 \le r \le \infty) \quad (0 < \rho \le 1)$$

where $I_0(x)$ is the modified Bessel's function of the first kind of order zero. Plots of the Rician envelope PDF for several values of ρ are shown in Figures 12.19 to 12.21. Let

$$A = \frac{\sigma^2}{2(1-\rho^2)} \qquad (12.174)$$

Using Equation 12.114 we have

$$h_2(q) = \frac{\sigma^2}{\sqrt{1-\rho^2}} \exp(-Aq) I_0(\rho A q) \qquad (12.175)$$

From Equation 12.30

$$h_{2N}(q) = (-2)^{N-1} \frac{\mathrm{d}^{N-1} h_2(q)}{dq^{N-1}} \qquad (12.176)$$

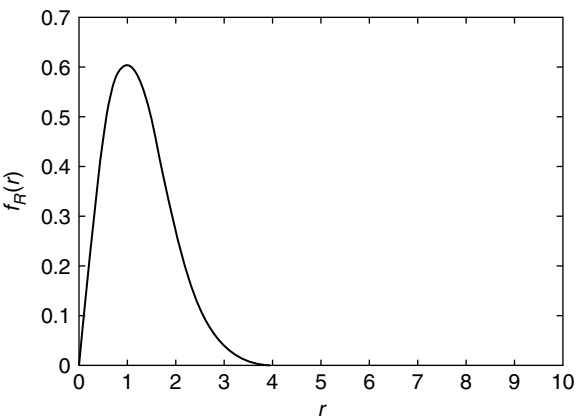

FIGURE 12.18 Generalized Rayleigh distributed envelope PDF, $\alpha = 2$, $\beta = 1.414$.

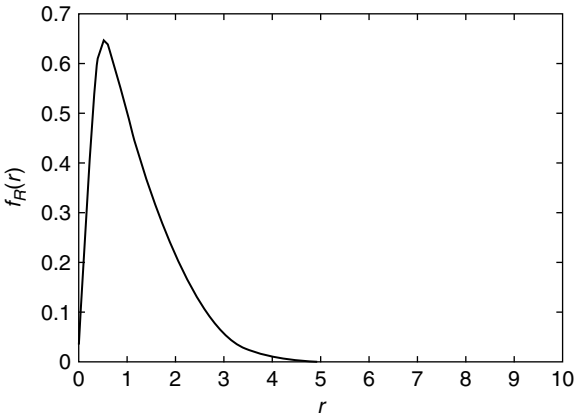

FIGURE 12.19 Rician envelope PDF, $\rho = 0.25$.

We then use Equation 12.157 and the identities[35]

$$I_n(x) = \frac{1}{2\pi}\int_0^{2\pi} \cos(n\theta)\exp[x\cos(\theta)]\mathrm{d}\theta$$
$$\cos^k(\theta) = \frac{1}{2^k}\sum_{m=0}^{k}\binom{k}{m}\cos[(k-2m)\theta] \tag{12.177}$$

to obtain

$$h_{2N}(q) = \frac{\sigma^{2N}}{(1-\rho^2)^{N-\frac{1}{2}}}\sum_{k=0}^{N-1}\binom{N-1}{k}(-1)^k\left(\frac{\rho}{2}\right)^k \xi_k \exp(-Aq) \tag{12.178}$$

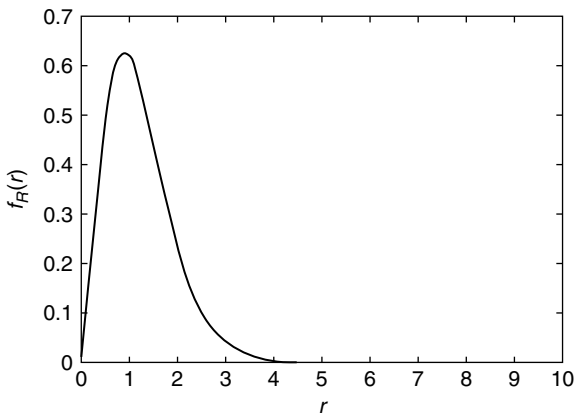

FIGURE 12.20 Rician envelope PDF, $\rho = 0.5$.

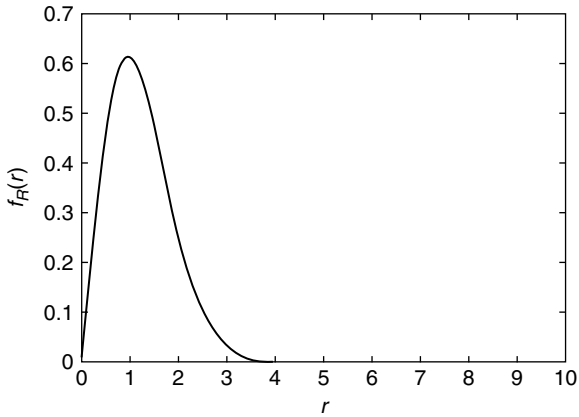

FIGURE 12.21 Rician envelope PDF, $\rho = 0.9$.

where

$$\xi_k = \sum_{m=0}^{k} \binom{k}{m} I_{k-2m}(\rho A q) \tag{12.179}$$

For $\rho = 0$, note that the Rician envelope PDF corresponds to the Rayleigh envelope PDF.

12.3.4.2.5. Generalized Gamma Envelope PDF

In a recent effort[37] the Generalized Gamma envelope PDF has been proposed as a candidate for univariate nonGaussian PDFs for modeling radar clutter. The Generalized Gamma envelope PDF is given by

$$f_{\mathrm{R}}(r) = \frac{ac}{\Gamma(\alpha)}(ar)^{c\alpha-1}\exp(-ar^c), \qquad 0 \le r \le \infty, \;\; a, c, \alpha > 0 \tag{12.180}$$

where a is the scale parameter and α and c are shape parameters of the PDF. Note that the Generalized Gamma envelope PDF reduces to:

1. The Weibull envelope PDF when $\alpha = 1$.
2. The Gamma envelope PDF when $c = 1$.
3. The Exponential envelope PDF when $c = \alpha = 1$.
4. The Chi envelope PDF when $c = 2$.
5. The Rayleigh envelope PDF when $c = 2$ and $\alpha = 1$.

We show that this PDF is admissible as a SIRV. Using Equation 12.114, we can write

$$h_2(q) = Aq^{\frac{c\alpha}{2}-1}\exp(-Bq^{\frac{c}{2}}) \tag{12.181}$$

where $A = \dfrac{(a\sigma)^{c\alpha}c}{\Gamma(\alpha)}$ and $B = a\sigma^c$. Using Equation 12.30, we have

$$\begin{aligned} h_{2N}(q) &= (-2)^{N-1}\frac{\mathrm{d}^{N-1}h_2(q)}{\mathrm{d}q^{N-1}} \\ &= (-2)^{N-1}A\frac{\mathrm{d}^{N-1}}{\mathrm{d}q^{N-1}}[q^{\frac{c\alpha}{2}-1}\exp(-Bq^{\frac{c}{2}})] \end{aligned} \tag{12.182}$$

Using Equation 12.157 and Equation 12.158, we can rewrite Equation 12.182 as

$$\begin{aligned} h_{2N}(q) = (-2)^{N-1}A\sum_{k=0}^{N-1}\binom{N-1}{k}\frac{\mathrm{d}^k}{\mathrm{d}q^k} \\ \times[\exp(-Bq^{\frac{c}{2}})]\frac{\Gamma\left(\dfrac{c\alpha}{2}\right)}{\Gamma\left(\dfrac{c\alpha}{2}-N+k+1\right)}q^{\frac{c\alpha}{2}-N+k} \end{aligned} \tag{12.183}$$

The kth derivative of the exponential term is readily obtained by using Equation 12.164 to Equation 12.167, Hence, we have

$$h_{2N}(q) = \sum_{k=0}^{N-1}F_k q^{\frac{c\alpha}{2}-N}\exp(-Bq^{\frac{c}{2}}) \tag{12.184}$$

where

$$\begin{aligned} F_k = (-2)^{N-1}A\binom{N-1}{k}\frac{\Gamma\left(\dfrac{c\alpha}{2}\right)}{\Gamma\left(\dfrac{c\alpha}{2}-N+k+1\right)} \\ \times\sum_{m=1}^{k}\sum_{l=1}^{m}(-1)^{m+l-1}\frac{B^m}{m!}\frac{\Gamma\left(\dfrac{lc}{2}+1\right)}{\Gamma\left(\dfrac{lc}{2}-k+1\right)}\omega^{\frac{mc}{2}} \end{aligned} \tag{12.185}$$

The Generalized Gamma envelope PDF is admissible for characterization as a SIRV for values of $c\alpha$ less than or equal to 2. This is due to the fact that $h_2(q)$ and its derivatives fail to satisfy the monotonicity condition for other values of $c\alpha$. Interestingly enough, only these values of $c\alpha$ give rise to extended tails for the PDF. Hence, the monotonicity conditions do not impose a serious restriction.

12.3.4.3. Examples Using the Marginal Characteristic Function

Successful use of the marginal characteristic function approach requires the knowledge of various Hankel transforms. For each example, the particular

transform used is cited by equation and page number as it appears in Ref. 36. To illustrate the procedure followed, a detailed derivation is presented in the first example. However, in the remaining examples, we simply list the univariate characteristic function of the quadrature components, the corresponding marginal PDF, and the resulting $h_{2N}(q)$. Finally, $h_{2N}(p)$ is obtained by replacing q with p in the expressions for $h_{2N}(q)$.

12.3.4.3.1. Gaussian Distribution

First, we consider the characteristic function given by

$$\psi(\omega) = \exp\left(-\frac{\omega^2}{2}\right) \tag{12.186}$$

The corresponding marginal PDF of the quadrature components is

$$f_{Y_k}(y_k) = \frac{1}{\sqrt{(2\pi)}}\exp\left(-\frac{y_k^2}{2}\right) \qquad -\infty \le y_k \le \infty \tag{12.187}$$

Equation 12.187 is the PDF of a zero mean unit variance Gaussian random variable. Substitution of Equation 12.186 into Equation 12.130 yields

$$h_{2N}(q) = (\sqrt{q})^{1-N}\int_0^\infty \omega^N \exp\left(-\frac{\omega^2}{2}\right)J_{N-1}(\omega\sqrt{q})\mathrm{d}\omega \tag{12.188}$$

From Ref. 36, Equation (10), p. 29, we have the Hankel transform

$$\int_0^\infty x^{v+\frac{1}{2}}\exp(-ax^2)J_v(xy)\sqrt{xy}\,\mathrm{d}x = \frac{y^{v+\frac{1}{2}}}{(2a)^{v+1}}\exp\left(-\frac{y^2}{4a}\right) \tag{12.189}$$

By making the association that $a = 0.5$, $v = N - 1$, $x = \omega$, and $\mathrm{y} = \sqrt{q}$, the above result becomes

$$\int_0^\infty \omega^N\exp\left(-\frac{\omega^2}{2}\right)J_{N-1}(\omega\sqrt{q})\sqrt{q}\,\mathrm{d}\omega = \sqrt{q}^{N-1+\frac{1}{2}}\exp\left(-\frac{q}{2}\right) \tag{12.190}$$

If follows that

$$h_{2N}(q) = \exp\left(-\frac{q}{2}\right) \tag{12.191}$$

From Equation 12.108, it is seen that the resulting SIRV PDF is the familiar multivariate Gaussian PDF, given by

$$f_{\mathbf{Y}}(\mathbf{y}) = (2\pi)^{-N}|\mathbf{\Sigma}|^{-\frac{1}{2}}\exp\left(-\frac{p}{2}\right) \tag{12.192}$$

12.3.4.3.2. K-Distribution

The marginal characteristic function given by

$$\psi(\omega) = \left(1 + \frac{\omega^2}{b^2}\right)^{-\alpha} \tag{12.193}$$

corresponds to the *K*-distributed envelope whose PDF is

$$f_R(r) = \frac{2b}{\Gamma(\alpha)}\left(\frac{br}{2}\right)^{\alpha} K_{\alpha-1}(br)u(r) \tag{12.194}$$

where α is the shape parameter of the distribution, b denotes its scale parameter, $K_N(t)$ is the Nth order modified Bessel function of the second kind and $u(r)$ is the unit step function. The pertinent Hankel transform for this example is found as Ref. 36, Equation (20), p. 24:

$$\int_0^{\infty} x^{v+\frac{1}{2}}(x^2 + a^2)^{-u-1} J_v(xy)\sqrt{xy}\,\mathrm{d}x = \frac{a^{v-u} y^{u+\frac{1}{2}} K_{v-u}(ay)}{2^u \Gamma(u+1)} \tag{12.195}$$

The resulting $h_{2N}(q)$ is

$$h_{2N}(q) = \frac{b^{2N}}{\Gamma(\alpha)} \frac{(b\sqrt{q})^{\alpha-N}}{2^{\alpha-1}} K_{N-\alpha}(b\sqrt{q}) \tag{12.196}$$

As a special case, when α is equal to unity, Equation 12.193 is the characteristic function of the Laplace distribution for the quadrature components whose PDF is given by

$$f_{Y_k}(y_k) = \frac{b}{2}\exp(-b|y_k|) \quad -\infty \le y_k \le \infty \tag{12.197}$$

where $|y_k|$ denotes the absolute value of y_k and b denotes the scale parameter. The corresponding $h_{2N}(q)$ is given by

$$h_{2N}(q) = b^{2N}(b\sqrt{q})^{1-N} K_{N-1}(b\sqrt{q}) \tag{12.198}$$

Another interesting case of the K-distribution arises when $\alpha = 0.5$. This corresponds to the exponential distribution for the marginal envelope PDF. Therefore, the *K*-distributed envelope PDF with $\alpha = 0.5$ is identical to the Weibull distributed envelope with b = 1. Although the characteristic PDF of the Weibull SIRV is unknown in general, the characteristic PDF of the Weibull SIRV for $b = 1$ is obtained when $\alpha = 0.5$ in Equation 12.143. Finally, we point out that the *K*-distributed envelope reduces to the Rayleigh envelope PDF when α tends to ∞.

12.3.4.3.3. Student-t Distribution

The characteristic function for the Student-*t* distribution with scale parameter b and shape parameter ν is given by

$$\psi(\omega) = \frac{K_\nu(b\omega)(b\omega)^\nu}{2^{\nu-1}\Gamma(\nu)} \tag{12.199}$$

Note the functional similarity with the envelope PDF given by Equation 12.194. The Student-*t* distribution is referred to as the generalized Cauchy distribution in Ref. 38 because the marginal PDF of the quadrature components is given by

$$f_{Y_k}(y_k) = \frac{\Gamma\left(\nu + \frac{1}{2}\right)}{b\sqrt{\pi}\Gamma(\nu)}\left(1 + \frac{y_k^2}{b^2}\right)^{-\nu-\frac{1}{2}} \qquad -\infty \le x_k \le \infty, \nu > 0 \tag{12.200}$$

where $\Gamma(\nu)$ is the Eulero-Gamma function. The relevant Hankel transform (in Ref. 36, Equation 3, p. 63) is

$$\int_0^\infty x^{u+v+\frac{1}{2}} K_u(ax) J_v(xy)\sqrt{xy}\,\mathrm{d}x = \frac{2^{v+u}a^u\Gamma(u+v+1)y^{v+\frac{1}{2}}}{(y^2+a^2)^{u+v+1}} \tag{12.201}$$

Using Equation 12.130, $h_{2N}(q)$ is expressed as

$$h_{2N}(q) = \frac{2^N b^{2\nu}\Gamma(\nu+N)}{\Gamma(\nu)(b^2+q)^{N+\nu}} \tag{12.202}$$

The Cauchy PDF for the quadrature components arises when ν is set equal to $1/2$ in Equation 12.200 and is given by

$$f_{Y_k}(y_k) = \frac{b}{\pi(b^2+y_k^2)} \qquad -\infty \le x_k \le \infty \tag{12.203}$$

where b is the scale parameter. The corresponding $h_{2N}(q)$ is

$$h_{2N}(q) = \frac{2^N b\Gamma\left(\frac{1}{2}+N\right)}{\sqrt{\pi}(b^2+q)^{N+\frac{1}{2}}} \tag{12.204}$$

Note that the Cauchy PDF does not have finite variance. However, this PDF is useful in modeling impulsive noise.[39] Finally, we point out that when $b = \sqrt{2\nu}$ and ν tends ∞ in Equation 12.200, the Student-*t* distribution reduces to the Gaussian distribution.

12.3.4.3.4. Rician Envelope PDF 2

We consider the Rician envelope PDF, arising from a nonzero mean complex Gaussian process, given by

$$f_R(r) = \frac{r}{\alpha^2}\exp\left[-\frac{(r^2+a^2)}{2\alpha^2}\right]I_0\left(\frac{ar}{\alpha^2}\right) \tag{12.205}$$

Plots of the Rician envelope PDF are shown in Figures 12.20 to 12.24 for several values of a and $\alpha = 1$. Note that this PDF approaches the Rayleigh PDF as a tends to zero. For convenience, we assume that $\sigma^2 = (1/2)E(R^2) = 1$. Using Equation 12.114, we have

$$h_2(r^2) = A\exp\left(-\frac{r^2}{2\alpha^2}\right)I_0\left(\frac{ar}{\alpha^2}\right) \tag{12.206}$$

where:

$$A = \frac{\exp\left(-\dfrac{\alpha^2}{2\alpha^2}\right)}{\alpha^2}$$

Noting that[35]

$$\int_0^\infty x\exp(-\alpha x^2)I_\nu(\beta x)J_\nu(\gamma x)\mathrm{d}x = \frac{1}{2\alpha}\exp\left(\frac{\beta^2-\gamma^2}{4\alpha}\right)J_\nu\left(\frac{\beta\gamma}{2\alpha}\right)$$
$$\mathrm{Re}\{\alpha\} > 0,\quad \mathrm{Re}\{\nu\} > -1, \tag{12.207}$$

Equation 12.121 results in the characteristic function

$$\psi(\omega) = \exp\left(-\frac{\omega^2\alpha^2}{2}\right)J_0\left(\frac{\omega a}{\alpha^2}\right) \tag{12.208}$$

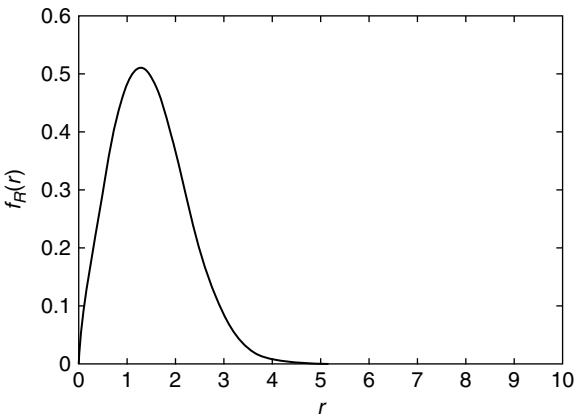

FIGURE 12.22 Rician envelope PDF, $a = 0.25$, $\alpha = 1$.

Recognizing that[35]

$$\int_0^\infty x^{\lambda-1}\exp(-\alpha x^2)J_\mu(\beta x)J_\nu(\gamma x)\mathrm{d}x$$
$$= \frac{\beta^\mu \gamma^\nu \alpha^{-\frac{(\mu+\nu+\lambda+2)}{2}}}{2^{\mu+\nu+1}\Gamma(\nu+1)} \sum_{m=0}^{\infty} \frac{\Gamma\left(m+\frac{\nu}{2}+\frac{mu}{2}+\frac{\lambda}{2}+1\right)}{m!\Gamma(m+\mu+1)}$$
$$\times\left(-\frac{\beta^2}{4\alpha}\right)^m F\left(-m;-\mu-m;\ \nu+1;\frac{\gamma^2}{\beta^2}\right)$$
$$\mathrm{Re}\{\alpha\}>0,\ \ \mathrm{Re}\{\mu+\nu+\lambda\}>-2,\ \ \beta>0,\ \ \gamma>0 \tag{12.209}$$

where $F(-m;-\mu-m;\ \nu+1;(\gamma^2/\beta^2))$ is the four parameter hypergeometric function, it follows from Equation 12.130 that

$$h_{2N}(q) = \frac{\alpha^{2N+2}}{2^{2N+1}\Gamma(N)} \sum_{m=0}^{\infty} \frac{\Gamma(m+N+1)}{m!\Gamma(m+1)}$$
$$\times\left(\frac{-a^2}{2\alpha^6}\right)^m F\left(-m,-m;N;\frac{q\alpha^4}{a^2}\right) \tag{12.210}$$

Since $h_{2N}(q)$ for this example involves an infinite series of hypergeometric functions, its form is mathematically intractable. Therefore, the corresponding multivariate SIRV PDF does note lend itself for use in practical applications.

We point out here that the lognormal envelope PDF and the Johnson (unbounded) distribution are not admissible for extension to SIRVs. This is due to the fact that $h_2(q)$ obtained for each of these distributions fails to satisfy the monotonicity conditions stated in Section 12.3.3. Table 12.1, presents a list of marginal PDFs suitable for extension to SIRVs. Table 12.2 tabulates $h_{2N}(p)$ for those marginal PDFs treated as envelope PDFs while Table 12.3 gives those $h_{2N}(p)$ obtained from the associated marginal characteristic function.

Plots of Equation 12.108 with $N = 1$ for the various SIRV PDFs are shown in Figures 12.25–12.33. In all the plots, the covariance matrix used is given by

$$\mathbf{\Sigma} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \tag{12.211}$$

Observe that each PDF is unimodal. However, the width and height of the peak along with the behavior of the extreme values (i.e., the tails) differ significantly.

12.3.5. Significance of the Quadratic Form of the SIRV PDF

Thus far, our discussion has focused on techniques that can be used to obtain the PDF of a SIRV starting from either the first order PDF or the first order characteristic function. Given random data, we are also interested in the problem

TABLE 12.1
Marginal PDFs Suitable for Extension to SIRVs

Marginal PDF	$f_X(x)$
Chi	$\frac{2b}{\Gamma(\nu)}(bx)^{2\nu-1}\exp(-b^2x^2)$
Weibull	$abx^{b-1}\exp(-ax^b)$
Generalized Rayleigh	$\frac{\alpha x}{\beta^2\Gamma(\frac{2}{\alpha})}\exp[-(\frac{x}{\beta})^\alpha]$
Generalized Gamma	$f_R(r)=\frac{ac}{\Gamma(\alpha)}(ar)^{c\alpha-1}\exp(-ar^c)$
Rician	$\frac{\chi}{\sqrt{1-\rho^2}}\exp[-\frac{\chi^2}{2(1-\rho^2)}]I_0[\frac{\rho x^2}{2(1-\rho^2)}]$
Gaussian	$\sqrt{2\pi}^{-1}\exp(-\frac{x^2}{2})$
Laplace	$\frac{b}{2}\exp(-b\lvert x_k\rvert)$
Cauchy	$\frac{b}{\pi(b^2+x_k^2)}$
K-distribution	$\frac{2b}{\Gamma(\alpha)}(\frac{bx}{2})^\alpha K_{\alpha-1}(bx)u(x)$
Student-*t*	$\frac{\Gamma(\nu+\frac{1}{2})}{b\sqrt{\pi}\Gamma(\nu)}(1+\frac{x_k^2}{b^2})^{-\nu-\frac{1}{2}}$

of approximating the distribution of the underlying data. The problem of multivariate distribution identification is of interest in radar signal detection. Since the background clutter is not known a priori, there is a need to identify the underlying clutter PDF based on measurements obtained from a given environment. Since the radar processes N pulses at a time, knowledge of the joint PDF of the N samples is necessary in order to obtain the optimal radar signal processor for the given clutter background. We present an important theorem here for enabling us to address the distribution identification of a SIRV.

Theorem 4. *The PDF of the quadratic form appearing in Equation 12.20 is given by*

$$f_P(p)=\frac{1}{2^{\frac{N}{2}}\Gamma\left(\frac{N}{2}\right)}p^{\frac{N}{2}-1}h_N(p)\qquad 0\le p\le\infty \tag{12.212}$$

TABLE 12.2
SIRVs Obtained from the Marginal Envelope PDF

Marginal PDF	$h_{2N}(p)$
Chi	$(-2)^{N-1}A\sum_{k=1}^{N} G_k p^{\nu-k}\exp(-Bp)$
	$G_k = \binom{N-1}{k-1}(-1)^{N-k}B^{N-k}\dfrac{\Gamma(\nu)}{\Gamma(\nu-k+1)}$
	$A = \dfrac{2}{\Gamma(\nu)}(b\sigma)^{2\nu}$
	$B = b^2\sigma^2$
	$\nu \le 1$
Weibull	$\sum_{k=1}^{N} C_k p^{(kb/2)-N}\exp\left(-Ap^{b/2}\right)$
	$A = a\sigma^b$
	$C_k = \sum_{m=1}^{k}(-1)^{m+N}2^N\dfrac{A^k}{k!}\binom{k}{m}\dfrac{\Gamma(1+\frac{mb}{2})}{\Gamma(1+\frac{mb}{2}-N)}$
	$b \le 2$
Generalized Rayleigh	$\sum_{k=1}^{N-1} D_k p^{(k\alpha/2)-N+1}\exp\left(-Bp^{(\alpha/2)}\right)$
	$A = \dfrac{\sigma^2\alpha}{\beta^2\Gamma\left(\frac{2}{\alpha}\right)}$
	$B = \beta^{-\alpha}\sigma^{\alpha}$
	$D_k = \sum_{m=1}^{k}(-1)^{m+N-1}2^{N-1}\dfrac{B^k}{k!}\binom{k}{m}\dfrac{\Gamma(1+\frac{m\alpha}{2})}{\Gamma(2+\frac{m\alpha}{2}-N)}$
	$\alpha \le 2$
Generalized Gamma	$h_{2N}(p) = \sum_{k=0}^{N-1} F_k p^{(c\alpha/2)-N}\exp\left(-Bp^{(c/2)}\right)$
	$F_k = (-2)^{N-1}A\binom{N-1}{k}\dfrac{\Gamma\left(\frac{c\alpha}{2}\right)}{\Gamma\left(\frac{c\alpha}{2}-N+k+1\right)}$
	$\times\sum_{m=1}^{k}\sum_{l=1}^{m}(-1)^{m+l-1}\dfrac{B^m}{m!}\dfrac{\Gamma\left(\frac{\ell c}{2}+1\right)}{\Gamma\left(\frac{\ell c}{2}-k+1\right)}p^{\frac{mc}{2}}$
	$c\alpha \le 2$
Rician	$\dfrac{\sigma^{2N}}{(1-\rho^2)^{N-(1/2)}}\sum_{k=0}^{N-1}\binom{N-1}{k}(-1)^k\left(\dfrac{\rho}{2}\right)^k\xi_k\exp(-A)$
	$\xi_k = \sum_{m=0}^{k}\binom{k}{m}I_{k-2m}(\rho A),\ A = \dfrac{p\sigma^2}{2(1-\rho^2)}$

TABLE 12.3
SIRVs Obtained from the Marginal Characteristic Function

Marginal PDF	$h_{2N}(p)$
Gaussian	$\exp(-\frac{P}{2})$
Laplace	$b^{2N}(b\sqrt{p})^{1-N}K_{N-1}(b\sqrt{p})$
Cauchy	$\frac{2^N b\Gamma(\frac{1}{2}+N)}{\sqrt{\pi}(b^2+p)^{N+(1/2)}}$
K-distribution	$\frac{b^{2N}}{\Gamma(\alpha)}\frac{(b\sqrt{p})^{\alpha-N}}{2^{\alpha-1}}K_{N-\alpha}(b\sqrt{p})$
Student-t	$\frac{2^N b^{2\nu}\Gamma(\nu+N)}{\Gamma(\nu)(b^2+p)^{N+\nu}}$

Proof. First, we consider a spherically symmetric random vector (SSRV) $\mathbf{X} = [X_1, X_2, \ldots, X_N]$. Because a SSRV is a special case of the SIRV, the representation theorem can be used to express $\mathbf{X}$ as

$$\mathbf{X} = \mathbf{Z}S \tag{12.213}$$

where $\mathbf{Z}$ is a Gaussian random vector having zero mean and identity covariance matrix and S is a nonnegative random variable with PDF $f_S(s)$. Consider the random variable

$$P' = \mathbf{X}^{\mathrm{T}}\mathbf{X} \tag{12.214}$$

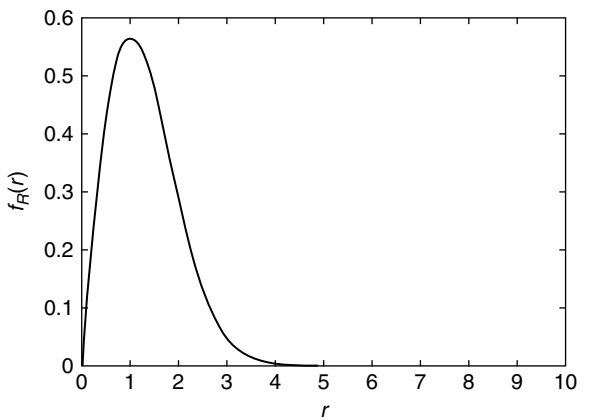

FIGURE 12.23 Rician envelope PDF, $a = 0.5$, $\alpha = 1$.

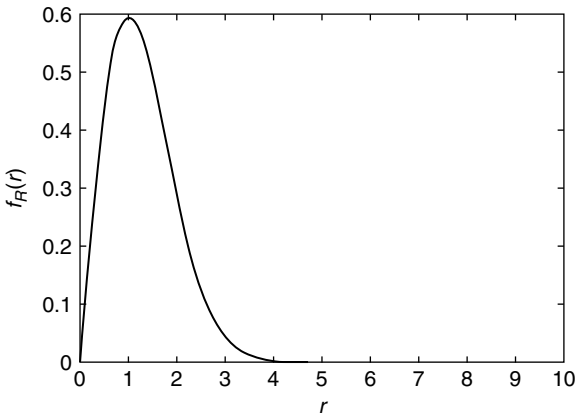

FIGURE 12.24 Rician envelope PDF, $a = 0.9$, $\alpha = 1$.

Using Equation 12.213 in Equation 12.214 gives

$$P' = \mathbf{Z}^{\mathrm{T}}\mathbf{Z}S^2 \tag{12.215}$$

Since $\mathbf{Z}^{\mathrm{T}}\mathbf{Z} = \sum_{i=1}^{N} Z_i^2$ is the sum of the squares of independent identically distributed Gaussian random variables having zero mean and unit variance, the PDF of $V = \mathbf{Z}^{\mathrm{T}}\mathbf{Z}$ is a Chi Square distribution with N degrees of freedom. Consequently,

$$f_V(v) = \frac{v^{\frac{N}{2}-1}}{2^{\frac{N}{2}}\Gamma\left(\frac{N}{2}\right)}\exp\left(-\frac{v}{2}\right) \qquad v \geq 0 \tag{12.216}$$

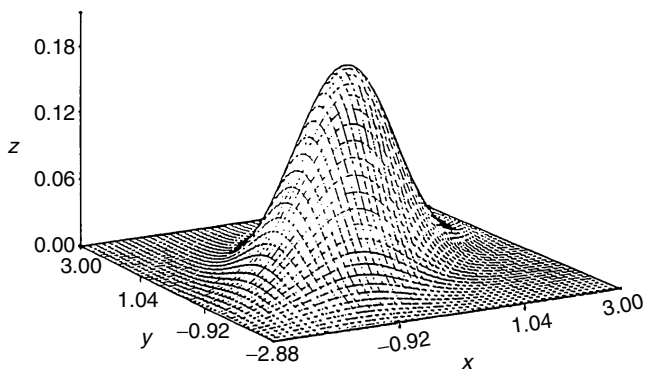

FIGURE 12.25 Gaussian distribution, zero mean, unit variance.

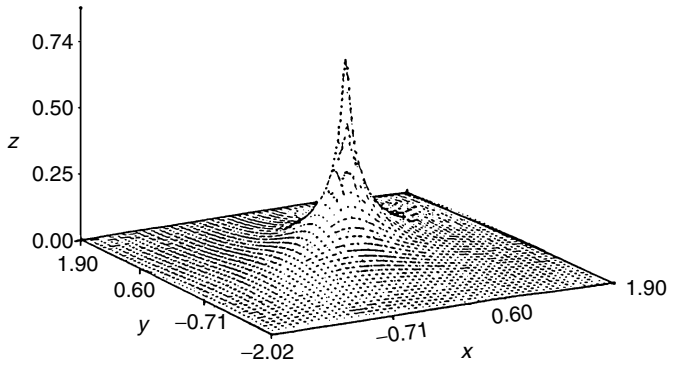

FIGURE 12.26 Laplace distribution, $b = 1$.

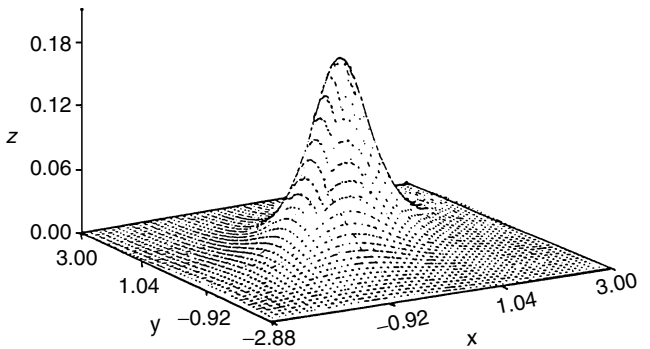

FIGURE 12.27 Cauchy distribution, $b = 1$.

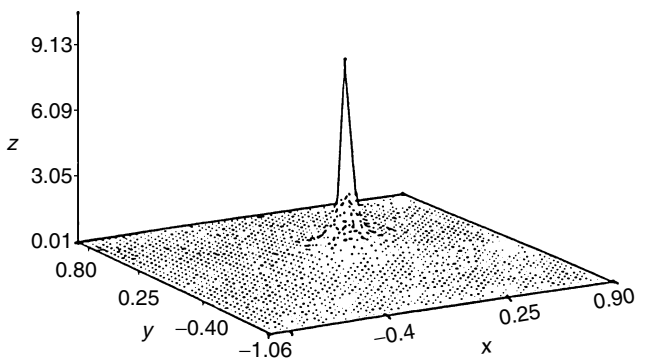

FIGURE 12.28 K-Distribution, $b = 1$, $\alpha = 0.5$.

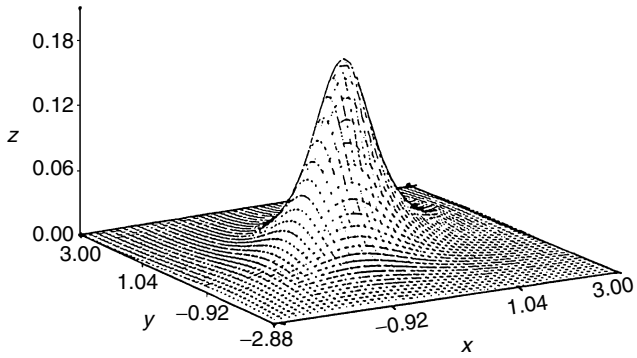

FIGURE 12.29 Student-t distribution, $b = 1$, $\upsilon = 1.5$.

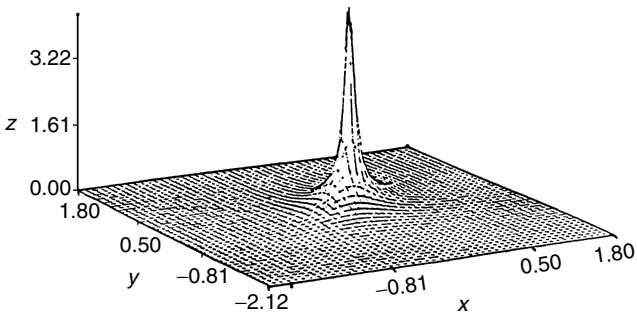

FIGURE 12.30 Chi-distribution, $b = 1$, $\upsilon = 1$.

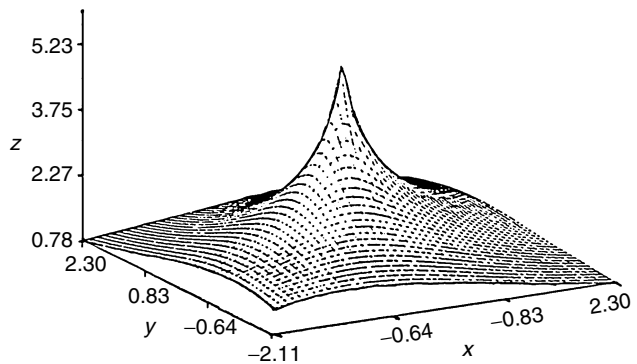

FIGURE 12.31 Generalized Rayleigh PDF, $\alpha = 0.5$, $\beta = 0.05$.

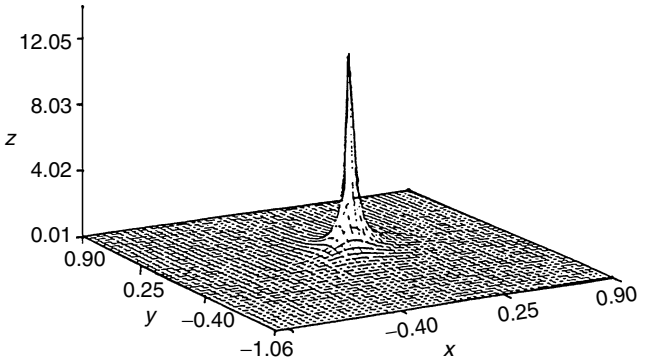

FIGURE 12.32 Weibull distribution, $a = 1$, $b = 1.0$.

The transformation $P' = VS^2$ then results in

$$f_{P'|S}(p'|s) = \frac{(p')^{\frac{N}{2}-1}}{2^{\frac{N}{2}}\Gamma\left(\frac{N}{2}\right)} s^{-N} \exp\left(-\frac{p'}{2s^2}\right) \tag{12.217}$$

From the theorem of total probability, we have

$$f_{P'}(p') = \int_0^\infty \frac{(p')^{\frac{N}{2}-1}}{2^{\frac{N}{2}}\Gamma\left(\frac{N}{2}\right)} s^{-N} \exp\left(-\frac{p'}{2s^2}\right) f_S(s)\mathrm{d}s \tag{12.218}$$

Recall that

$$h_N(p') = \int_0^\infty s^{-N} \exp\left(-\frac{p'}{2s^2}\right) f_S(s)\mathrm{d}s \tag{12.219}$$

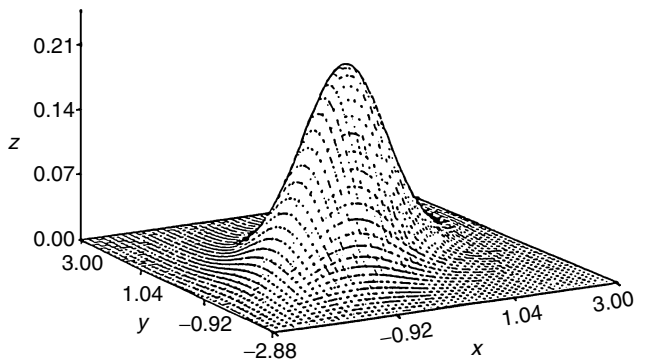

FIGURE 12.33 Rician PDF, $\rho = 0.5$.

Consequently, the PDF of P' is expressed as

$$f_{P'}(p') = \frac{(p')^{\frac{N}{2}-1}}{2^{\frac{N}{2}}\Gamma(\frac{N}{2})} h_N(p') \tag{12.220}$$

Recall that a SIRV $\mathbf{Y} = [Y_1, Y_2, ..., Y_N]^{\mathrm{T}}$ having a mean vector $\mathbf{b}$ and covariance matrix $\mathbf{\Sigma}$ is related to SSRV $\mathbf{X}$ by the linear transformation

$$\mathbf{Y} = \mathbf{AX} + \mathbf{b} \tag{12.221}$$

where $\mathbf{\Sigma} = \mathbf{AA}^{\mathrm{T}}$. Then, the quadratic form appearing in Equation 12.214 can be expressed as

$$P = (\mathbf{Y} - \mathbf{b})^{\mathrm{T}}\mathbf{\Sigma}^{-1}(\mathbf{Y} - \mathbf{b}) \tag{12.222}$$

However, Equation 12.222 is also the quadratic form appearing in Equation 12.20 as the PDF of $\mathbf{Y}$. Since $P = P'$, the PDF of the quadratic form P associated with $\mathbf{Y}$ is

$$f_P(p) = \frac{(p)^{\frac{N}{2}-1}}{2^{\frac{N}{2}}\Gamma(\frac{N}{2})} h_N(p) \tag{12.223}$$

This establishes the theorem. Thus, a SIRV is uniquely characterized by the quadratic form appearing in its PDF. Knowledge of the quadratic form PDF is sufficient to identify the SIRV PDF. This is an important result since it allows us to reduce the multivariate distribution identification problem to the equivalent problem of univariate distribution identification of the quadratic form. It is emphasized that the PDF of P is invariant to the choice of $\boldsymbol{\mu}$ and $\mathbf{\Sigma}$. We point out that the invariance of the PDF of the quadratic form arises from the fact that a SSRV arises from a uniform distribution over an N dimensional hypersphere of radius R. The radius of the hypersphere remains unchanged regardless of whether we consider a SIRV or a SSRV. Only the azimuthal angles and radial angle change depending on whether the random vector is a SSRV or a SIRV. In context of the radar problem, we are dealing with N complex samples or $2N$ quadrature components. The results presented in this section are applicable when N is replaced by $2N$.

12.3.6. Conclusion

In this chapter we have pointed out a method to obtain the PDF of correlated nonGaussian random vectors arising in the problem of radar clutter modelling. The theory of SIRPs has been used to develop the multivariate PDFs. Various techniques have been presented to obtain SIRV PDFs. Several examples are provided to illustrate these techniques. The admissibility of the Chi envelope PDF, Weibull envelope PDF, Generalized Rayleigh envelope PDF, Rician envelope PDF, and Generalized Gamma envelope PDF as SIRVs has been pointed out for the first time. Finally, we have obtained the PDF of the quadratic

form of a SIRV and we have shown that this PDF remains unchanged regardless of whether we are dealing with a SSRV or a SIRV. We have also established that the quadratic form contains all the information that is required in order to identify the SIRV PDF. As a consequence of this result, the problem of a SIRV (multivariate) distribution identification has been reduced to the equivalent identification of the univariate distribution of the nonnegative quadratic form.

12.4. COMPUTER GENERATION OF SIMULATED RADAR CLUTTER CHARACTERIZED AS SIRPs

12.4.1. Introduction

This investigation is motivated by a desire to simulate correlated nonGaussian radar clutter. Various investigators have reported experimental results where nonGaussian marginal PDFs have been used to model the clutter. Usually, radars process N samples at a time. Statistical characterization of the clutter requires the specification of the joint PDF of the N samples. In addition, the clutter may be highly correlated. Hence, the joint PDF must take into account the correlation between samples. Statistical characterization of the clutter is necessary if an optimal radar signal processor is to be obtained. For use of the well known likelihood ratio test, it is necessary to have closed form expressions for the joint PDF of the N clutter samples in order to obtain the optimal radar signal processor. In most cases, it is difficult to evaluate the performance of the optimal radar signal processor analytically when the clutter samples are correlated and nonGaussian. Then computer simulation may be necessary. Therefore, there is a need to develop efficient procedures to facilitate computer simulation of the clutter. A library of multivariate nonGaussian PDFs has been developed in Section 12.3, using the theory of SIRPs and SIRVs. In view of the large number of parameters that are free to be specified, the library of multivariate nonGaussian PDFs can be used to approximate many different radar clutter scenarios. In this section we concern ourselves with the development of computer simulation procedures for the library of nonGaussian PDFs obtained in Section 12.3 so that the performance of any radar signal processor can be evaluated for a variety of different clutter scenarios. Another issue addressed in this section is performance assessment of the simulation procedures.

It has been pointed out in Section 12.3 that the quadratic form appearing in the PDF of the SIRV contains all the information necessary to identify the PDF of the underlying SIRV. We make use of this result in order to assess the performance of the simulation procedures. Some interesting simulation techniques have been proposed for SIRVs in Refs. 18 and 20. The technique suggested in Ref. 18 makes use of Meijer's G-functions. These functions are generalizations of Hypergeometric functions not lending themselves to the development of simple elegant simulation procedures. The technique suggested in Ref. 20 requires transformations from rectangular to spherical coordinates

and then back again. Secondly, this simulation procedure involves the use of the inverse distribution function approach for a rather complicated distribution function. The approach developed in this chapter is simpler to implement than those proposed in Refs. 18 and 20. In addition, a new approach is proposed for assessing the effectiveness of the simulation procedure.

The problem of computer generation of correlated nonGaussian radar clutter is equivalent to the problem of generating random variables with a jointly specified marginal PDF and covariance matrix. While the problem of generating random sequences with either a specified PDF or prescribed covariance function has been well treated,[34] the joint problem has received limited attention. In general, it has been possible to control either the PDF or the correlation function but not both simultaneously. Previous attempts[5,8,10–12] to address the problem of generating random sequences with jointly specified marginal PDF and covariance function have not been successful because the procedures proposed therein made use of zero memory nonlinear (ZMNL) transformations on a correlated Gaussian sequence to obtain the desired nonGaussian sequence. Consequently, the covariance matrix of the nonGaussian sequence was related to that of the Gaussian sequence in a rather complicated manner. Hence, given a certain covariance matrix for the nonGaussian sequence, it was not possible to determine the corresponding covariance matrix of the Gaussian sequence. Furthermore, not all nonlinear transformations gave rise to nonnegative definite covariance matrices at their outputs.[33,34] Thus, using ZMNL transformations on a correlated Gaussian sequence does not offer a practical solution to the joint problem. The techniques presented in this chapter successfully overcome the drawbacks of the previous efforts. This is due to the fact that SIRPs belong to the class of exogenous product models for radar clutter, which allows for independent control of the marginal PDF and correlation function.

In Section 12.4.2, we review some definitions and background information pertaining to the theory of SIRPs. Section 12.4.3 presents two canonical simulation procedures for generating SIRVs. Performance assessment of the simulation procedures is discussed in Section 12.4.4. Finally, conclusions are presented in Section 12.4.5.

12.4.2. Preliminaries

We begin by restating the definitions for a SIRV and a SIRP. A SIRV is a random vector (real or complex) whose PDF is uniquely determined by the specification of a mean vector, a covariance matrix, and a characteristic first order PDF. Equivalently, the PDF of a SIRV can also be referred to as an elliptically contoured distribution. A SIRP is a random process (real or complex) such that every random vector obtained by sampling this process is a SIRV. The work of Yao[15] gave rise to a representation theorem that can be stated as follows (see Theorem 1):

If a random vector is a SIRV, then there exists a nonnegative random variable S such that the PDF of the random vector conditioned on S is a multivariate Gaussian PDF.

We consider the product given by $\mathbf{X} = \mathbf{Z}S$, where $\mathbf{X} = [X_1 \ldots X_N]^{\mathrm{T}}$ denotes the SIRV, $\mathbf{Z} = [Z_1 \ldots Z_N]^{\mathrm{T}}$ is a Gaussian random vector with zero mean and covariance matrix $\mathbf{M}$, and S is a nonnegative random variable with PDF $f_S(s)$. Since it is desirable to independently control the correlation properties and the nonGaussian envelope PDF, $\mathbf{Z}$ and S are assumed to be statistically independent. The PDF of $\mathbf{X}$ conditioned on S is (see Equation 12.14)

$$f_{\mathbf{X}|S}(\mathbf{x}|s) = (2\pi)^{-\frac{N}{2}} |\mathbf{M}|^{-\frac{1}{2}} s^{-N} \exp\left(-\frac{p}{2s^2}\right) \tag{12.224}$$

where p is a nonnegative quadratic form given by $p = \mathbf{x}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{x}$ and $|\mathbf{M}|$ denotes the determinant of the covariance matrix $\mathbf{M}$. The PDF of $\mathbf{X}$ is given by (see Equation 12.15 and Equation 12.16)

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-\frac{N}{2}} |\mathbf{M}|^{-\frac{1}{2}} h_N(p) \tag{12.225}$$

where

$$h_N(p) = \int_0^\infty s^{-N} \exp\left(-\frac{p}{2s^2}\right) f_S(s) \mathrm{d}s \tag{12.226}$$

The PDF of the random variable S is called the characteristic PDF of the SIRV. Therefore, it is apparent that the PDF of a SIRV is completely determined by the specification of a mean vector, a covariance matrix, and a characteristic first order PDF. In addition, the PDF of the SIRV is a function of a nonnegative quadratic form. However, except for the Gaussian case, dependence on the quadratic form is more complicated than the simple exponential. Therefore, an SIRP can be regarded as a generalization of the familiar Gaussian random process. We point out that the covariance matrix of the SIRV is given by $\mathbf{\Sigma} = \mathbf{M}E(S^2)$ where $E(S^2)$ is the mean square value of the random variable S. It is seen that the covariance matrix of the SIRV normalized by the mean square value of S is the covariance matrix of the Gaussian random vector. Note that it is possible to set the covariance matrix of the SIRV equal to that of the Gaussian random vector by requiring $E(S^2)$ to be equal to unity. The desired nonGaussian PDF can be obtained by choosing $f_S(s)$ appropriately. Thus, the SIRV formulation for radar clutter modelling affords independent control over the nonGaussian PDF of the clutter of its correlation properties. Several techniques are available in Section 12.3 for obtaining $h_N(p)$. Note that the Gaussian random vector is a special case of a SIRV and is obtained when $f_S(s) = \delta(s - 1)$ where $\delta(t)$ is the unit impulse function. An interesting interpretation of the representation theorem is that every SIRV is the modulation of a Gaussian random vector by a nonnegative random variable.

Many of the attractive properties of Gaussian random vectors also apply to SIRVs. The most relevant property of SIRVs for the purpose of computer simulation is the closure property under linear transformation[15] stated below (see Theorem 2, Section 12.2.3):

> If $\mathbf{X}$ is a SIRV with characteristic PDF $f_S(s)$, then
>
> $$\mathbf{Y} = \mathbf{AX} + \mathbf{b} \tag{12.227}$$
>
> is also a SIRV with the same characteristic PDF. It is assumed that $\mathbf{A}$ is a nonsingular matrix and $\mathbf{b}$ is a known vector having the same dimension as $\mathbf{X}$.

Theorem 2 provides us with a powerful technique for simulating SIRVs. A white SIRV is defined as one that has a diagonal covariance matrix. In other words, the components of the white SIRV are uncorrelated but not necessarily independent. We can start with a zero mean white SIRV $\mathbf{X}$ having identity covariance matrix and perform the linear transformation given by Equation 12.227 to obtain a SIRV $\mathbf{Y}$ having a nonzero mean and desired covariance matrix $\mathbf{\Sigma}$. The matrix $\mathbf{A}$ and the vector $\mathbf{b}$ are given by

$$\begin{aligned} \mathbf{A} &= \mathbf{E}\mathbf{D}^{\frac{1}{2}} \\ \mathbf{b} &= \boldsymbol{\mu}_y \end{aligned} \tag{12.228}$$

where $\mathbf{E}$ is the matrix of normalized eigen-vectors of the covariance matrix $\mathbf{\Sigma}$, $\mathbf{D}$ is the diagonal matrix of eigen-values of $\mathbf{\Sigma}$ and $\boldsymbol{\mu}_y$ is the desired nonzero mean vector.

In many instances it is not possible to obtain $f_S(s)$ for a SIRV in closed form, even though its existence is guaranteed. In such cases, an alternate approach must be used in order to characterize the SIRV. The following theorem can be used to completely characterize a white SIRV having zero mean and identity covariance matrix (see Theorem 3, Section 12.2.3):

> A random vector $\mathbf{X} = [X_1 \ldots X_N]^{\mathrm{T}}$ is a zero mean white SIRV having identity covariance matrix if and only if there exist random variables $R \in (0, \infty)$, $\Theta \in (0, 2\pi)$ and $\Phi k \in (0, \pi)$, $(k = 1, \ldots N-2)$ such that when the components of $\mathbf{X}$ are expressed in the generalized spherical coordinates
>
> $$\begin{aligned} X_1 &= R\cos(\Phi_1) \\ X_k &= R\cos(\Phi_k)\prod_{i=1}^{k-1}\sin(\Phi_i) \qquad 1 < k \le N-2 \\ X_{N-1} &= R\cos(\Theta)\prod_{i=1}^{N-2}\sin(\Phi_i) \\ X_N &= R\sin(\Theta)\prod_{i=1}^{N-2}\sin(\Phi_i) \end{aligned} \tag{12.229}$$

then the random variables R, Θ and Φ_k are mutually statistically independent and have PDFs of the form

$$\begin{aligned} f_R(r) &= \frac{r^{N-1}}{2^{\frac{N}{2}-1}\Gamma\left(\frac{N}{2}\right)} h_N(r^2)u(r) \\ f_{\Phi_k}(\phi_k) &= \frac{\Gamma\left(\frac{N-k+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{N-k}{2}\right)} \sin^{N-1-k}(\phi_k)[u(\phi_k) - u(\phi_k - \pi)] \\ f_\Theta(\theta) &= (2\pi)^{-1}[u(\theta) - u(\theta - 2\pi)] \end{aligned} \tag{12.230}$$

where $\Gamma(\nu)$ is the Eulero Gamma function and $u(t)$ is the unit step function.

As a consequence of Theorem 3, any SIRV with zero mean and identity covariance matrix can be represented in generalized spherical coordinates that are mutually and statistically independent regardless of the SIRV considered. Also, note that the PDFs of Θ and Φk $(k = 1, \ldots N - 2)$ are functionally independent of the white SIRV considered. Only the PDF of R changes from one white SIRV to another. Note that $R^2 = \sum_{k=1}^{N} X_k^2 = \mathbf{X}^{\mathrm{T}}\mathbf{X}$. Hence R is the norm of the SIRV.

Another important feature of the SIRV is that the quadratic form appearing in its PDF contains all the information necessary to identify the PDF. It follows that knowledge of the PDF of the quadratic form of the SIRV is sufficient to identify the PDF of the corresponding SIRV[21] (see Theorem 4):

The PDF of the quadratic form appearing in Equation 12.225 is given by

$$f_P(p) = \frac{1}{2^{\frac{N}{2}}\Gamma\left(\frac{N}{2}\right)} p^{\frac{N}{2}-1} h_N(p) \qquad 0 \le p \le \infty \tag{12.231}$$

and remains unchanged regardless of whether or not the SIRV is white.

The theorems reviewed in this section will be made use of in the proposed simulation approach, discussed in Section 12.4.3, and in assessing the performance of the simulation procedure, discussed in Section 12.4.4.

In the context of the problem of radar clutter modeling and simulation, the bandpass process $Y(t) = \mathrm{Re}[\tilde{Y}(t)\exp(j\omega_0 t)]$ can be expressed in terms of the corresponding complex, wide sense stationary random processes $\tilde{Y}(t)$. More precisely, we obtain N complex samples by sampling the complex random process $\tilde{Y}(t) = Y_c(t) + jY_s(t)$, where the subscripts c and s denote the in-phase and out-of-phase quadrature components, respectively. This is equivalent to working with a real vector of $2N$ quadrature components, and this is the approach taken in this chapter. Therefore, the results presented in this section are applied to the problem of radar clutter modeling with N replaced by $2N$. For ease of reference, the library of nonGaussian SIRV PDFs obtained in Section 12.3 is repeated here. However, $h_{2N}(p)$ for the SIRVs with known characteristic PDF

TABLE 12.4
$h_{2N}(p)$ for SIRVs with Known Characteristic PDF

Marginal PDF	$h_{2N}(p)$
Laplace	$b^{2N}(b\sqrt{p})^{1-N}K_{N-1}(b\sqrt{p})$
Cauchy	$\dfrac{2^N b\Gamma(\frac{1}{2}+N)}{\sqrt{\pi}(b^2+p)^{N+(1/2)}}$
K-distribution	$\dfrac{b^{2N}}{\Gamma(\alpha)}\dfrac{(b\sqrt{p})^{\alpha-N}}{2^{\alpha-1}}K_{N-\alpha}(b\sqrt{p})$
Student-t	$\dfrac{2^N b^{2\nu}\Gamma(\nu+N)}{\Gamma(\nu)(b^2+p)^{N+\nu}}$

are listed in Table 12.4. The corresponding characteristic PDFs are listed in Table 12.5. Table 12.6 lists $h_{2N}(p)$ for those SIRVs whose characteristic PDF is unknown.

12.4.3. Two Canonical Simulation Procedures for Generating SIRVs

In this section, we concern ourselves with two simulation procedures for generating the SIRVs listed in Tables 12.4 and 12.5. The first simulation procedure to be discussed is applicable when the characteristic PDF, $f_S(s)$, is known. For each of the PDFs listed in Table 12.4, the characteristic PDF $f_S(s)$ is tabulated in Table 12.6, where $E(S^2) = 1$. Since the representation theorem results in the covariance matrix of the SIRV being given by $\mathbf{\Sigma} = \mathbf{M}E(S^2)$, the choice of $E(S^2) = 1$ makes $\mathbf{\Sigma}$ identical to $\mathbf{M}$, the covariance matrix of the Gaussian random vector $\mathbf{Z}$. However, as shown in Table 12.7, the PDFs for the distributions in Table 12.4, as commonly expressed, do not have unit mean square value. In order to obtain the random variable S, with unit mean square value, and the corresponding PDF $f_S(s)$, we generate the random variable V having PDF $f_V(v)$ and mean square value $E(V^2) = a^2$, and perform the scaling $S = \frac{V}{a}$ to obtain the desired S. The simulation procedure for these SIRV PDFs is fairly simple and is stated below:

(1) Generate a sample vector of a white zero mean Gaussian random vector $\mathbf{Z}$, having identity covariance matrix.
(2) Then generate a sample value of the random variable V from the PDF $f_V(v)$. Denote the mean square value of V by a^2.
(3) Normalize the sample value of the random variable V by a to obtain a sample value of the modulating random variable S. In other words generate $S = \frac{V}{a}$.

TABLE 12.5
$h_{2N}(p)$ for SIRVs with Unknown Characteristic PDFs

Marginal PDF	$h_{2N}(p)$
Chi	$(-2)^{N-1}A\sum_{k=1}^{N} G_k p^{\nu-k}\exp(-Bp)$
	$G_k = \binom{N-1}{k-1}(-1)^{k-1}B^{k-1}\dfrac{\Gamma(\nu)}{\Gamma(\nu-k+1)}$
	$A = \dfrac{2}{\Gamma(\nu)}(b\sigma)^{2\nu}$
	$B = b^2\sigma^2$
	$\nu \le 1$
Weibull	$\sum_{k=1}^{N} C_k p^{(kb/2)-N}\exp\left(-Ap^{(b/2)}\right)$
	$A = a\sigma^b$
	$C_k = \sum_{m=1}^{k}(-1)^{m+N}2^N\dfrac{A^k}{k!}\binom{k}{m}\dfrac{\Gamma\left(1+\dfrac{mb}{2}\right)}{\Gamma\left(1+\dfrac{mb}{2}-N\right)}$
	$b \le 2$
Gen. Rayleigh	$\sum_{k=1}^{N-1} D_k p^{(k\alpha/2)-N+1}\exp\left(-Bp^{(\alpha/2)}\right)$
	$A = \dfrac{\sigma^2\alpha}{\beta^2\Gamma\left(\dfrac{2}{\alpha}\right)}$
	$B = \beta^{-\alpha}\sigma^{\alpha}$
	$D_k = \sum_{m=1}^{k}(-1)^{m+N-1}2^{N-1}\dfrac{B^k}{k!}\binom{k}{m}\dfrac{\Gamma\left(1+\dfrac{m\alpha}{2}\right)}{\Gamma\left(2+\dfrac{m\alpha}{2}-N\right)}$
	$\alpha \le 2$
Rician	$\dfrac{\sigma^{2N}}{(1-\rho^2)^{N-(1/2)}}\sum_{k=0}^{N-1}\binom{N-1}{k}(-1)^k\left(\dfrac{\rho}{2}\right)^k\xi_k\exp(-A)$
	$\xi_k = \sum_{m=0}^{k}\binom{k}{m}I_{k-2m}(\rho A),\ A = \dfrac{p\sigma^2}{2(1-\rho^2)}$

TABLE 12.6
Characteristic PDF for SIRVs Listed in Table 12.4, $[E(S^2) = 1]$

Marginal PDF	$f_S(s)$
Laplace	$ab^2 s \exp\left(-\frac{a^2 b^2 s^2}{2}\right) u(s)$
Cauchy	$a^2 b^2 s^{-3} \exp\left(-\frac{b^2}{2a^2 s^2}\right) u(s)$
K-distribution	$\frac{2ab}{\Gamma(\alpha)2^\alpha}(bas)^{2\alpha-1}\exp\left(-\frac{b^2 a^2 s^2}{2}\right) u(s)$
Student-*t*	$\frac{2ab}{\Gamma(\nu)2^\nu} b^{2\nu-1}(as)^{-(2\nu+1)}\exp\left(-\frac{b^2}{2a^2 s^2}\right) u(s)$

(4) Generate the product corresponding to $\mathbf{X} = \mathbf{Z}S$. At this step, we have a sample vector of a white SIRV having zero mean and identity covariance matrix.
(5) Finally, perform the linear transformation given by Equation 12.228 to obtain a sample vector of the SIRV $\mathbf{Y}$ with desired mean and covariance matrix.

Figure 12.34 shows the simulation procedure presented above.

The subroutine RNNOR in IMSL was used for generating the sample vectors of the Gaussian random vector $\mathbf{Z}$. Interestingly enough, the PDFs listed

TABLE 12.7
Related PDF $f_V(v)$

Marginal PDF	$f_V(v)$	$a^2 = E(V)^2$
Laplace	$b^2 v \exp\left(-\frac{b^2 v^2}{2}\right) u(v)$	$\frac{2}{b^2}$
Cauchy	$b^2 v^{-3} \exp\left(-\frac{b^2}{2v^2}\right) u(v)$	∞
K-distribution	$\frac{2ab}{\Gamma(\alpha)2^\alpha}(bv)^{2\alpha-1}\exp\left(-\frac{b^2 v^2}{2}\right) u(v)$	$\frac{2\alpha}{b^2}$
Student-*t*	$\frac{2b}{\Gamma(\nu)2^\nu} b^{2\nu-1} v^{-(2\nu+1)} \exp\left(-\frac{b^2}{2v^2}\right) u(v)$	$\frac{b^2}{2(\nu-1)}$

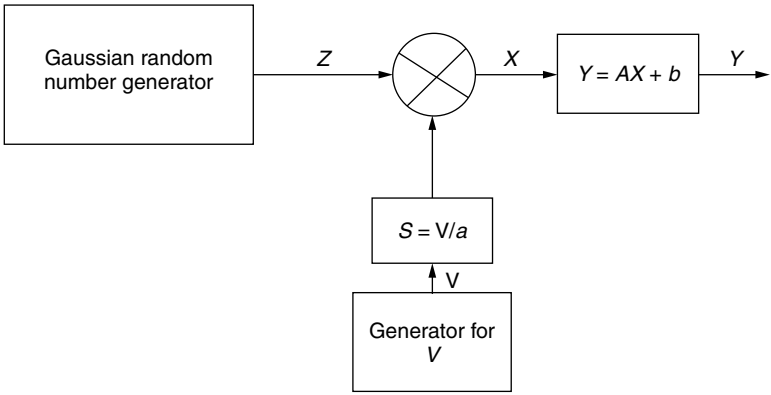

FIGURE 12.34 Simulation scheme for SIRVs with known characteristic PDF.

in Table 12.7 can be related to the PDF of the gamma distribution as discussed below. The PDF $f_V(v)$ for the K-distributed SIRV is a Chi PDF. We first address the random variable generation for the Chi PDF and then provide the transformations for obtaining the random variables for the other PDFs listed in Table 12.7.

Consider the standard Gamma distribution given by

$$f_T(t) = \frac{t^{\alpha-1}}{\Gamma(\alpha)}\exp(-t) \qquad t > 0 \tag{12.232}$$

where α denotes the shape parameter and $\Gamma(\alpha)$ is the Eulero-Gamma function. The random variable V for the Chi PDF is obtained by the transformation $V = \frac{\sqrt{2T}}{b}$. Samples of the random variable T are readily generated by using the IMSL subroutine RNGAM. The procedure for generating the Chi distributed random samples needed for the K-distributed SIRV is summarized below.

1. Generate samples of the random variable T for the standard Gamma distribution of Equation 12.232 by using the IMSL subroutine RNGAM.
2. Perform the transformation $V = \frac{\sqrt{2T}}{b}$.

The PDF $f_V(v)$ for the Laplace SIRV is a Rayleigh PDF and is obtained from that of the K-distributed SIRV by letting $\alpha = 1$. The random variable V for the PDF $f_V(v)$ listed in Table 12.7 for the Student-t SIRV is obtained from the standard Gamma PDF of Equation 12.232 by the transformation $V = \frac{b}{\sqrt{2T}}$ and letting $\alpha = \nu$. Finally, the PDF $f_V(v)$ for the Cauchy SIRV is obtained from that of the Student-t SIRV by letting $\nu = 1$. The procedure for generating the random samples needed for the Student-t SIRV is summarized below.

1. Generate samples of the random variable T for the standard Gamma distribution of Equation 12.232 by using the IMSL subroutine RNGAM.
2. Perform the transformation $V = \frac{b}{\sqrt{2T}}$.

We now concern ourselves with the second simulation procedure that is applicable when the characteristic PDF is unknown, as is the case for SIRVs listed in Table 12.5. This alternate approach makes use of Theorem 3. In particular, this procedure requires the capability to generate the independent random variables, R, Θ, and Φk $(k = 1, 2, \ldots, N-2)$. Generation of the random variables Θ and Φk $(k = 1, 2, \ldots, N-2)$ is extremely difficult from a computational standpoint. This problem is overcome as follows.

Recall that the PDFs of Θ and Φk $(k = 1, 2, \ldots, N-2)$ remain unchanged regardless of the white SIRV considered. Only the PDF of R changes from one white SIRV to another. Furthermore, since a Gaussian random vector is a member of the family of SIRVs, a white Gaussian random vector having zero mean and identity covariance matrix admits a representation of the form of Equation 12.229. It follows that

$$\frac{\mathbf{X}_k}{R} = \frac{\mathbf{Z}_k}{R_{\mathrm{G}}} \qquad k = 1, 2, \ldots, N \tag{12.233}$$

where R is the norm of the desired white SIRV and R_{G} is the norm of the zero mean white Gaussian random vector. Consequently, the components of the desired white SIRV are obtained as

$$\mathbf{X}_k = \frac{\mathbf{Z}_k}{R_{\mathrm{G}}} R \qquad k = 1, 2, \ldots, N \tag{12.234}$$

The simulation procedure is stated below:

(1) Generate a sample vector of the white, zero mean Gaussian random vector $\mathbf{Z}$ having identity covariance matrix.
(2) Compute the norm $R_{\mathrm{G}} = \|\mathbf{Z}\| = \sqrt{\mathbf{Z}^{\mathrm{T}}\mathbf{Z}}$ of the sample vector $\mathbf{Z}$.
(3) Generate a sample of the norm $R = \|\mathbf{X}\| = \sqrt{\mathbf{X}^{\mathrm{T}}\mathbf{X}}$ of the white SIRV from the PDF of R given by Equation 12.230.
(4) Generate a sample vector of the white SIRV $\mathbf{X}$ by computing $\mathbf{X} = \frac{\mathbf{Z}}{R_{\mathrm{G}}} R$.
(5) Finally, perform the linear transformation given by Equation 12.228 to obtain a sample of the SIRV $\mathbf{Y}$ with desired mean and covariance matrix.

The simulation procedure is shown schematically in Figure 12.35. Note that this simulation procedure avoids the explicit generation of the variables Θ and Φ_k $(k = 1, \ldots, N-2)$. The generation procedure for a white Gaussian random vector is well known. Therefore, we need to concern ourselves only with the

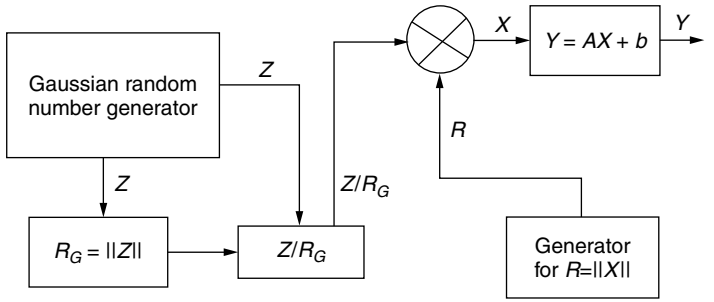

FIGURE 12.35 Simulation scheme for SIRVs with unknown characteristic PDF.

development of a suitable generation scheme for samples of the norm R of the white SIRV $\mathbf{X}$. Generation of the samples of R is not trivial. This is due to the fact that the PDF of R is generally not in a simple functional form. Consequently, it may not be possible to conveniently evaluate analytically the distribution function and its inverse. Hence, generation methods based on the inverse distribution function do not offer a practical solution to this problem. Therefore, in this chapter we generate samples of R by making use of the approach called the "Rejection Method". The rejection method can be used to generate samples of random variables whose cumulative distribution functions are not known, but whose PDFs are known explicitly.[40] The rejection procedure assumes knowledge of the maximum value of the PDF of R for a given SIRV PDF and an estimate for the finite range of the PDF of R so that the area under the PDF curve is close to unity. These quantities are denoted by c and b, respectively. We discuss the rejection procedure in detail in Appendix V. The Rejection method is summarized below:

(1) Generate a uniform random variate U_1 on the interval $(0, b)$.
(2) Generate another uniform variate U_2 on the interval $(0, c)$.
(3) If $U_2 \leq f_R(U_1)$, then $R = U_1$. Otherwise, reject U_1 and return to step 1.

Note that the simulation procedures of Figures 12.34 and 12.35 are canonical as their forms remain unchanged from the simulation of one SIRV to another. Even though the scheme of Figure 12.35 can be used even when $f_S(s)$ is known, the scheme of Figure 12.34 is preferred when S can be generated easily. The linear transformation of Equation 12.228 is a filtering operation. In both schemes, pre-modulation filtering is equivalent to post-modulation filtering. This results from the fact that the representation theorem is valid whether or not the SIRV $\mathbf{X}$ and the Gaussian random vector $\mathbf{Z}$ are white.

12.4.4. Performance Assessment of the Simulation Schemes

In this section we concern ourselves with the performance assessment of the simulation procedures developed in Section 12.4.3. We point out

that the simulation procedures developed in Section 12.4.3 are exact in the sense they are derived without approximation from theory. Hence, departures from the exact SIRVs will depend for the most part on the nonideality of the uniform random number generators and on the number of samples used. Empirical assessment of the simulation procedures is necessary for practical applications.

One possible approach for assessing the distributional properties of the simulated data is to perform a hypothesis test on the marginal distributions of the components of the SIRV. More precisely, the problem is stated as follows.

H_0: The hypothesis that the simulated data is from the desired distribution
H_1: The hypothesis that the simulated data is not from the desired distribution.

For a fixed Type-1 error probability (i.e., the probability that H_1 is accepted given that H_0 is true) each marginal distribution can be checked by employing one of the commonly used goodness of fit procedures. Since the components of the random vectors are not statistically independent, we are now confronted with the problem of developing a goodness of fit test for the multivariate data. In general, it is very difficult to obtain the overall significance level of the test (i.e., the probability that H_0 is accepted given that H_0 is true) for the multivariate goodness of fit testing procedure.

However, an attractive feature of SIRVs is that the quadratic form p appearing in the SIRV PDF contains all the information necessary for identifying the PDF of the SIRV. In other words, knowledge of the PDF of the quadratic form is sufficient to determine the underlying SIRV PDF. Furthermore, the quadratic form PDF remains unchanged regardless of whether the SIRV is white or colored. The PDF of the quadratic form appearing in the SIRV PDF is given by Equation 12.231. For the radar problem where we deal with N complex samples or $2N$ quadrature components, note that we make use of Equation 12.231 with N replaced by $2N$. Hence, we base our goodness of fit test procedure for the generated SIRVs on the PDF of the quadratic form p. Note that we have now reduced the multivariate problem to an equivalent univariate problem involving the goodness of fit test for the PDF of the quadratic form.

In the examples presented in this section, we generated $m = 1000$ realizations of the random vector $\mathbf{Y}$ with $N = 2$ complex samples and from these computed one thousand samples of the quadratic form ρ for each of the non-Gaussia SIRVs whose PDFs are listed in Tables 12.4 and 12.6. In each case, we used the corresponding theoretical PDF of the quadratic form given by Equation 12.231 to test the distribution of the generated quadratic form. The frequency histograms for the generated data and the corresponding theoretical PDFs are shown in Figures 12.36–12.43. In addition, a Chi-Square test was performed on the generated data with the Type-1 error fixed at 0.05 and the null hypothesis was not rejected in each case. The frequency histograms provide a good idea about the true distributions for large sample

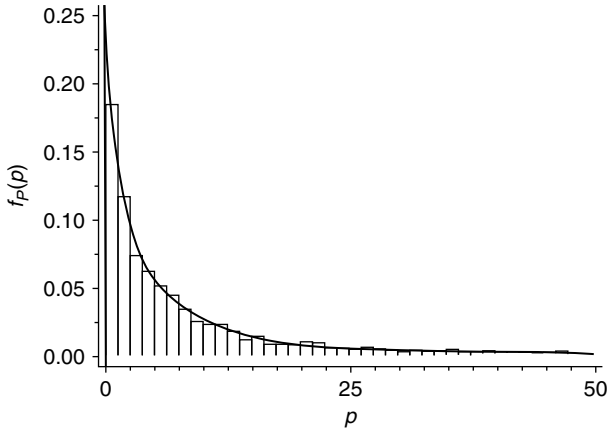

FIGURE 12.36 Theoretical and empirical quadratic form PDFs for Laplace SIRV.

sizes. Observe that the empirical PDFs are very close to the theoretical PDFs. Note that the procedure used in this section to assess the distributional assumptions of the random samples from the SIRV PDFs is a formal goodness of fit test. Similar procedures have been proposed to test for multivariate normality in Refs. 41 and 42.

12.4.5. Conclusions

In this chapter, we have presented two schemes that can be used in practice to simulate correlated nonGaussian radar clutter, when the clutter can be modeled as a SIRP. We pointed out that the simulation schemes developed are canonical

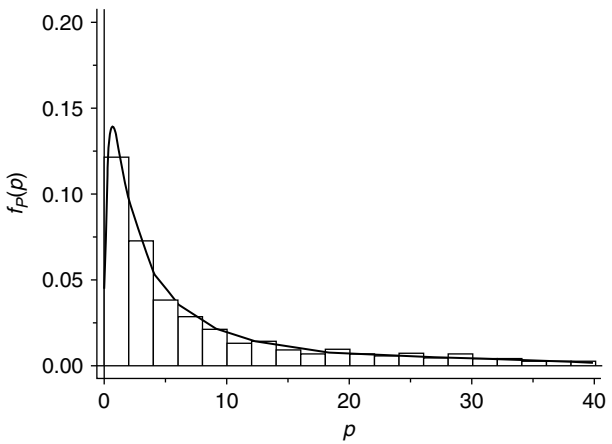

FIGURE 12.37 Theoretical and empirical quadratic form PDFs for Cauchy SIRV.

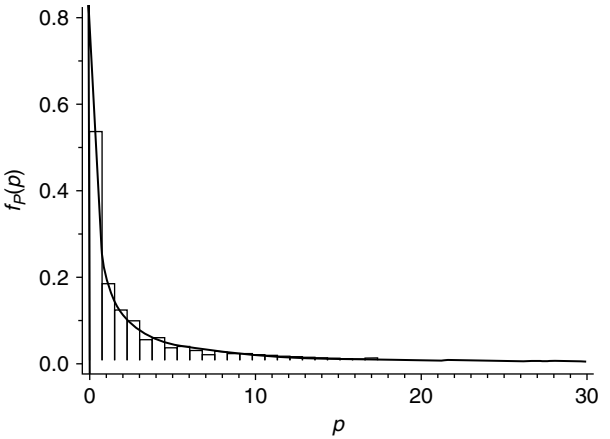

FIGURE 12.38 Theoretical and empirical quadratic form PDFs for *K*-distributed SIRV.

schemes and do not change form the simulation of one SIRV to another. A new approach, based on the PDF of the quadratic form appearing in the SIRV PDF, was used to perform a goodness of fit test in order to assess performance of the proposed simulation schemes. Performance assessment based on this scheme showed excellent agreement between the theoretical and empirical PDFs of the quadratic form. Finally, it was pointed out that use of this technique reduced the goodness of fit test from a multivariate testing procedure to a univariate testing procedure resulting in tremendous processing simplicity. Therefore, this procedure lends itself very well to practical applications.

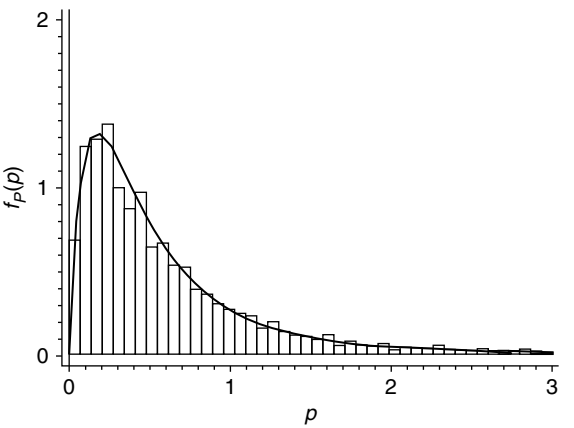

FIGURE 12.39 Theoretical and empirical quadratic form PDFs for student SIRV.

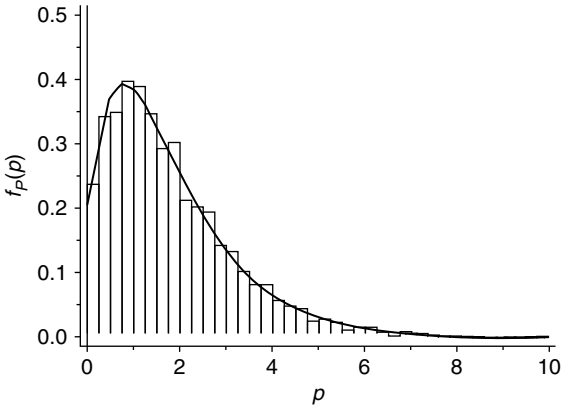

FIGURE 12.40 Theoretical and empirical quadratic form PDFs for Chi distributed SIRV.

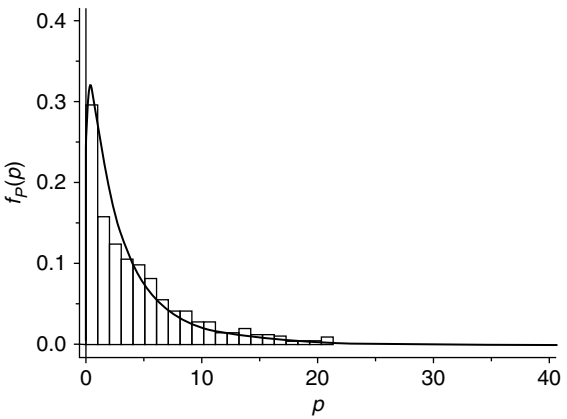

FIGURE 12.41 Theoretical and empirical quadratic form PDFs for generalized Rayleigh SIRV.

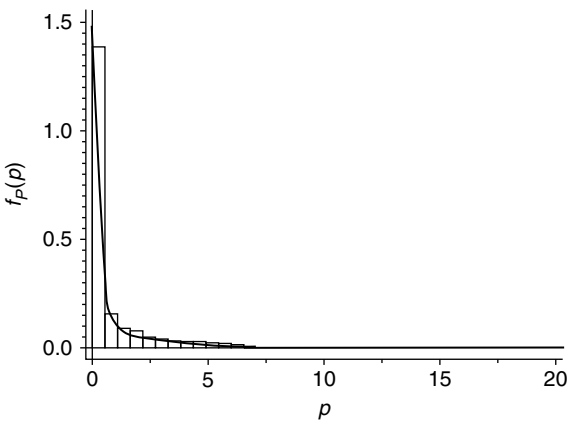

FIGURE 12.42 Theoretical and empirical quadratic form PDFs for Weibull SIRV.

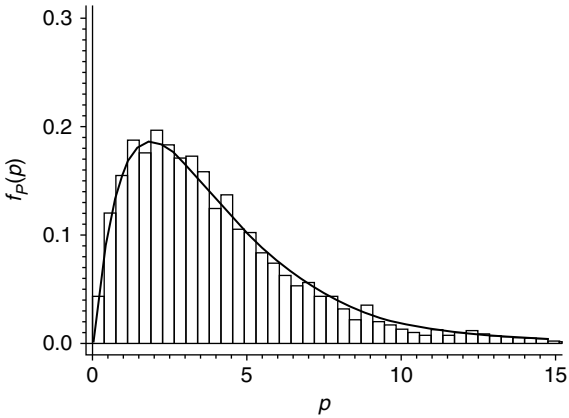

FIGURE 12.43 Theoretical and empirical quadratic form PDFs for Rician SIRV.

12.5. DISTRIBUTION APPROXIMATION TO RADAR CLUTTER CHARACTERIZED BY SIRPs

12.5.1. Introduction

This investigation is motivated by a desire to characterize correlated nonGaussian radar clutter by approximating the underlying PDF of the clutter. Various investigators have reported experimental results where nonGaussian marginal PDFs have been used to model the clutter. Usually, radars process N samples at a time. Statistical characterization of the clutter requires the specification of the joint PDF of the N samples. In addition, the clutter may be highly correlated. Hence, the joint PDF must take into account the correlation between samples. Statistical characterization of the clutter is necessary if an optimal radar signal processor is to be obtained. For use of the well known likelihood ratio test, it is desirable to have a closed form expression for the joint PDF of the N clutter samples in order to obtain the optimal radar signal processor. The joint PDF of the N clutter samples can be easily specified when the clutter is Gaussian. However, when the clutter is nonGaussian and is correlated, many different joint PDFs of the clutter samples can result in the same set of marginal (univariate) distributions having a specified nonGaussian character. The multivariate nonGaussian PDF is uniquely determined from the marginal distribution only when the samples are statistically independent.

Specification of the multivariate PDF is generally a nontrivial problem with no simple best solution.[43] The theory of SIRPs provides a powerful mechanism to obtain the joint PDF of the N correlated, nonGaussian clutter samples. Many of the tractable properties of the Gaussian random process also apply to SIRPs. Typically, background clutter is not known *a priori*. Hence, while dealing with real world data, there is a need to approximate the clutter PDF from a set of

measurements, which in turn needs a library of multivariate nonGaussian PDFs. Such a library has been obtained in Section 12.3 based on the theory of SIRPs. The multivariate PDFs thus obtained are functions of nonnegative quadratic forms. Therefore, these PDFs are sometimes referred to as elliptically symmetric distributions. The multivariate Gaussian PDF belongs to the family of SIRPs. The multivariate Pearson types II and VII are examples of elliptically symmetric multivariate nonGaussian PDFs. SIRPs have received considerable attention over the past two decades since many of the elegant and mathematically tractable properties of the multivariate Gaussian distribution generalize to this class of distributions. Applications of SIRPs can be found in the random flight problem,[14] signal detection,[16] speech signal modeling,[17] and radar clutter modeling.[19,21]

In practice, the clutter PDF encountered in radar signal processing is not known *a priori*. Consequently, a scheme that approximates the clutter PDF based on a set of measured data is necessary. Currently, available tests such as, the Kolmogorov–Smirnov test and the Chi-Square test, address the problem of goodness-of-fit for random data. In particular, these tests provide information about whether a set of random data is statistically consistent with a specified distribution, to within a certain confidence level. However, if the specified distribution is rejected, these tests cannot be used for approximating the underlying PDF of the random data. Moreover, these tests require large sample sizes for reliable results.

In practice, only a small number of samples may be available. Therefore, the scheme used should be efficient for small sample sizes. A new algorithm based on sample order statistics has been developed in Ref. 41 for univariate distribution identification. This algorithm has two modes of operation. In the first mode, the algorithm performs a goodness-of-fit test. Specifically, the test determines, to a desired confidence level, whether random data is statistically consistent with a specified probability distribution. In the second mode of operation, the algorithm approximates the PDF underlying the random data. In particular, by analyzing the random data and without any *a priori* knowledge, the algorithm identifies from a stored library of PDFs that density function which best approximates the data. Estimates of the scale, location, and shape parameters of the PDF are provided by the algorithm. The algorithm typically works well with samples sizes that may be as small as 50 or 100 samples. An extension of this algorithm for the multivariate Gaussian PDF has been considered in Refs. 41 and 44.

In this Section 12.5, using certain properties of SIRPs, we adopt the algorithm developed in Ref. 41 to identify the underlying distribution of a given set of data. In particular, we first show that the multivariate distribution approximation problem for SIRPs is reduced to an equivalent univariate distribution approximation problem. The new algorithm developed by Ozturk in Ref. 41 is used for the univariate approximation problem. Section 12.5.2 presents definitions. Sections 12.5.3 to 12.5.5 summarize the algorithm developed in Ref. 41 for approximating the univariate PDF of a set of random data. In Section 12.5.6 we present a procedure for the goodness of fit test for PDFs arising from SIRPs. The proposed distribution identification algorithm is discussed in Section 12.5.7. Section 12.5.8

proposes a method to estimate the shape parameter based on the procedure developed in Section 12.5.7. Finally, conclusions are presented in Section 12.5.9.

12.5.2. Definitions

Let $f_Y(y)$ denote the PDF of a random variable Y. Consider the linear transformation defined by

$$x = \beta y + \alpha \tag{12.235}$$

The PDF of X is given by

$$f_X(x) = \frac{1}{|\beta|} f_Y\left(\frac{x-\alpha}{\beta}\right) \tag{12.236}$$

where α and β are defined to be the location and scale parameters of $f_X(x)$, respectively. The mean μ_x and variance σ_x of the random variable X are given by

$$\begin{aligned} \mu_x &= E[X] \\ \sigma_x^2 &= E[(X-\mu_x)^2] \end{aligned} \tag{12.237}$$

Although the mean and the variance are related to the location and scale parameters, note that the location parameter is not the mean value and the scale parameter is not the square root of the variance, in general. However, for a standardized Gaussian PDF $f_Y(y)$ with zero mean and unit variance, the location parameter is the mean of X and the scale parameter is the standard deviation (square root of the variance) of X.

The coefficient of skewness, α_3, and the coefficient of kurtosis, α_4, of X, are defined to be

$$\begin{aligned} \alpha_3 &= \frac{E[(X-\mu_x)^3]}{\sigma_x^3} \\ \alpha_4 &= \frac{E[(X-\mu_x)^4]}{\sigma_x^4} \end{aligned} \tag{12.238}$$

It is readily shown that α_3 and α_4 are invariant to the values of μ_x and σ_x. For any PDF that is symmetric about the mean, $\alpha_3 = 0$. For the case of the Gaussian distribution, $\alpha_3 = 0$ and $\alpha_4 = 3$.

12.5.3. Goodness of Fit Test

In this section, we introduce a general graphical method for testing whether a set of random data is statistically consistent with a specified univariate distribution. The proposed method not only yields a formal goodness-of-fit test but also provides a graphical representation that gives insight into how well the random data is representative of the specified distribution (null hypothesis). Using the standard normal distribution with zero mean and unit variance as a reference distribution, the standardized sample order statistics are represented by a system

of linked vectors. The terminal point of these linked vectors, as well as the shape of their trajectories, are used in determining whether or not to accept the null hypothesis.

In this section we first give a brief description of the corresponding test statistic and then explain the goodness of fit test procedure. For illustration purposes, we consider the null distribution to be Gaussian. However, the proposed procedure works for any null hypothesis.

Let X_i; $i = 1, 2, \ldots, n$ denote the ith sample from a Gaussian distribution with mean μ and variance σ^2. Let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ denote the ordered samples obtained by ordering X_i; $i = 1, 2, \ldots, n$. We define

$$Y_i = \frac{|X_i - \bar{X}|}{S} \qquad i = 1, 2, \ldots, n \tag{12.239}$$

where $\bar{X} = \sum X_i/n$ is the sample mean and $S = \{\sum (X_i - \bar{X})^2/(n-1)\}^{1/2}$ is the sample standard deviation. The standardized order statistics are denoted by $Y_{i:n}$ $i = 1, 2, \ldots, n$ and are obtained by ordering the Y_i; $i = 1, 2, \ldots, n$. It follows that

$$Y_{i:n} = \frac{|X_{i:n} - \bar{X}|}{S} \qquad i = 1, 2, \ldots, n \tag{12.240}$$

The ith linked vector is characterized by its length and orientation with respect to the horizontal axis. Let $G_{i:n}$ denote the order statistic from the standard normal reference distribution. Also, let $m_{i:n} = E[G_{i:n}]$. The length of the ith vector, a_i, is obtained from the magnitude of the ith standardized sample order statistic, while its orientation θ_i is related to $m_{i:n}$. More specifically, by definition,

$$\begin{aligned} a_i &= \frac{Y_{i:n}}{n} \\ \theta_i &= \pi\, \Phi(m_{i:n}) \end{aligned} \tag{12.241}$$

where $\Phi(x) = \left(\sqrt{2\pi}\right)^{-1} \int_{-\infty}^{x} \exp\left(-\frac{t^2}{2}\right) dt$ is the cumulative distribution function of the standard Gaussian distribution. We define the sample points Q_k in a two dimensional plane by

$$Q_k = (U_k, V_k) \qquad k = 1, 2, \ldots, n \tag{12.242}$$

where $U_0 = V_0 = 0$ and

$$\begin{aligned} U_k &= \frac{1}{k} \sum_{i=1}^{k} \{\cos(\theta_i)\} Y_{i:n} \qquad k = 1, 2, \ldots, m \\ V_k &= \frac{1}{k} \sum_{i=1}^{k} \{\sin(\theta_i)\} Y_{i:n} \qquad k = 1, 2, \ldots, m \end{aligned} \tag{12.243}$$

The sample linked vectors are obtained by joining the points Q_k. Note that $Q_0 = (0, 0)$. It should also be noted that the statistic Q_n given in Equation 12.242 represents the terminal point of the linked vectors defined above. Figure 12.44

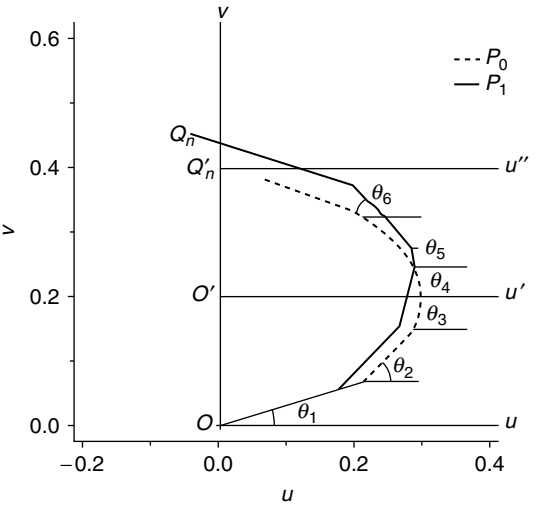

FIGURE 12.44 Linked vector chart: dashed lines P_0 = null linked vectors, solid lines P_1 = sample linked vectors.

shows the linked vectors obtained for the Gaussian distribution with $n = 6$. Since U_n and V_n are random variables for a given n, the corresponding linked vectors must be obtained by averaging the results of Monte Carlo trials. In this case the linked vectors were obtained by averaging the results of 50,000 Monte Carlo trials. The solid curve in Figure 12.44 shows the linked vectors for the sample distribution while the dashed curve shows the ideal linked vectors for the null distribution. The magnitudes and angles of the linked vectors are obtained from Equation 12.241. Note that the angles are independent of the data. Only the magnitudes of the linked vectors change from one trial to another.

For a typical set of ordered samples (i.e., ordered samples drawn from the null distribution) it is reasonable to expect that the sample linked vectors would follow the null pattern closely. If the ordered set of samples is not from the null distribution, the sample linked vectors are not expected to follow the null pattern closely. Hence, the procedure provides visual information about how well the ordered set of samples fit the null distribution.

An important property of the Q_n statistic is that it is invariant under linear transformation. In particular, we consider the standardization used in Equation 12.239. Let $Z_i = cX_i + d$, where c and d are known constants. Let S' denote the sample standard deviation of the samples Z_i. Then, it is readily shown that $|X_i - \bar{X}|/S = |Z_i - \bar{Z}|/S'$. The invariance property follows as a consequence. The advantage of this property is that the PDF of $Q_n = (U_n, V_n)$, for a given sample and reference distribution, depends only on the sample size n and is unaffected by the location and scale parameters. Since it is difficult to determine the joint PDF of U_n and V_n analytically, it is necessary to obtain empirical results.

Assuming that the conditions under the central limit theorem are satisfied, the marginal PDFs of U_n and V_n can be approximated as Gaussian, in the limit of large n. In addition, it is assumed that the joint PDF of U_n and V_n is approximately bivariate Gaussian. Consequently, all that is needed to determine the bivariate PDF is the specification of $E(U_n)$, $E(V_n)$, $E(U_nV_n)$, $\mathrm{Var}(U_n)$, and $\mathrm{Var}(V_n)$. Drawing samples from the Gaussian distribution, it has been shown empirically in Ref. 41 that for $3 \leq n \leq 100$

$$\begin{aligned} E(U_n) &= 0 \\ E(V_n) &= \mu_v \approx 0.326601 + \frac{0.412921}{n} \\ E(U_nV_n) &= 0 \\ \mathrm{Var}(U_n) &= \sigma_u^2 \approx \frac{0.02123}{n} + \frac{0.01765}{n^2} \\ \mathrm{Var}(V_n) &= \sigma_v^2 \approx \frac{0.04427}{n} - \frac{0.0951}{n^2}. \end{aligned} \tag{12.244}$$

Since U_n and V_n are approximately bivariate Gaussian for large or moderate sample sizes, their joint PDF can be written as

$$f_{U_n,V_n}(u_n, v_n) = (2\pi)^{-1}(\sigma_u\sigma_v)^{-1}\exp\left(-\frac{t}{2}\right) \tag{12.245}$$

where

$$t = \frac{u_n^2}{\sigma_u^2} + \frac{(v_n - \mu_v)^2}{\sigma_v^2} \tag{12.246}$$

Let $t = t_0$. Then the equation

$$t_0 = \frac{u_n^2}{\sigma_u^2} + \frac{(v_n - \mu_v)^2}{\sigma_v^2} \tag{12.247}$$

is that of an ellipse in the u_n, v_n plane, for which

$$f_{U_n,V_n}(u_n, v_n) = (2\pi)^{-1}(\sigma_u\sigma_v)^{-1}\exp\left(-\frac{t_0}{2}\right) \tag{12.248}$$

Points that fall within the ellipse correspond to those points in the u_n, v_n plane satisfying

$$f_{U_n,V_n}(u_n, v_n) > (2\pi)^{-1}(\sigma_u\sigma_v)^{-1}\exp\left(-\frac{t_0}{2}\right) \tag{12.249}$$

Let

$$\alpha = P(T > t_0) = P\ [(u_n, v_n) \text{ outside the ellipse given in Eq. (12.247)}] \tag{12.250}$$

It is well known that the PDF of the random variable T defined by Equation 12.246 has a Chi-Square distribution with two degrees of freedom[45] and is

given by

$$f_T(t) = 0.5 \exp\left(-\frac{t}{2}\right) \tag{12.251}$$

Hence,

$$\alpha = 1 - \exp\left(-\frac{t_0}{2}\right) \tag{12.252}$$

Consequently, $t_0 = -2\ln(1-\alpha)$. Thus, Equation 12.247 becomes

$$\frac{u_n^2}{\sigma_u^2} + \frac{(v_n - \mu_v)^2}{\sigma_v^2} = -2\ln(1-\alpha) \tag{12.253}$$

α is known as the significance level of the test. It is the probability that Q_n falls outside the ellipse specified by Equation 12.253 given that the data is coming from a Gaussian distribution. The expression $(1-\alpha)$ is known as the confidence level and the corresponding ellipse is known as the confidence ellipse.

Equation 12.247 can be written in the standardized form

$$1 = \frac{u_n^2}{\sigma_u^2 t_0} + \frac{(v_n - \mu_v)^2}{\sigma_v^2 t_0} \tag{12.254}$$

where the lengths of the major and minor axes are given by $2\max\left[\sigma_u\sqrt{t_0}, \sigma_v\sqrt{t_0}\right]$ and $2\min\left[\sigma_u\sqrt{t_0}, \sigma_v\sqrt{t_0}\right]$, respectively. From Equation 12.252, observe that smaller values of α correspond to larger values of t_0. Consequently, the confidence ellipses become larger as the confidence level is increased.

For a given sample size n ($n \leq 100$) approximate values of μ_v, σ_u^2, and σ_v^2 can be obtained from Equation 12.244. The confidence ellipse of Equation 12.253 can then be used to make a visual as well as computational test of the null hypothesis. If the terminal sample point falls inside the ellipse, then the data is declared as being consistent with the Gaussian distribution with confidence level $1-\alpha$. Otherwise the null hypothesis is rejected with a significance level α.

A major difficulty in determining the joint PDF of U_n and V_n is that the coefficients of skewness and kurtosis of U_n and V_n (see Table 12.8) indicate that the Gaussian approximation for the bivariate PDF may not be satisfactory for $n < 100$. The empirical bivariate PDF of U and V were obtained by using 50,000 Monte-Carlo trials for $n = 3, 10, 20, 30, 50$, and 100. The corresponding constant probability contours of the joint PDF of U_n and V_n are shown in Figure 12.45. The same procedure is used even when the null distribution is different from the Gaussian distribution. However, note that the standard Gaussian distribution is always used as the reference distribution for determining the angles of the linked vectors.

TABLE 12.8
Some Monte Carlo Moments of *U* and *V*

	U				*V*			
n	**Mean**	**Variance**	**Skew.**	**Curtos.**	**Mean**	**Variance**	**Skew.**	**Curtos.**
3	0.00037	0.00846	−0.0036	1.7594	0.47067	0.00168	−0.3174	1.8720
4	0.00052	0.00585	−0.0157	3.1854	0.43458	0.00448	0.3599	2.5413
5	−0.00030	0.00472	0.0186	2.6987	0.41276	0.00475	0.0141	2.2890
6	0.00007	0.00391	−0.0055	2.8623	0.39793	0.00459	−0.0243	2.5650
7	0.00021	0.00333	−0.0033	2.9035	0.38763	0.00426	−0.0552	2.6668
8	−0.00025	0.00290	−0.0025	2.9141	0.38000	0.00399	−0.0694	2.7219
9	0.00008	0.00257	0.0050	2.9201	0.37362	0.00369	−0.0551	2.7671
10	−0.00051	0.00229	0.0041	2.9428	0.36865	0.00349	−0.0528	2.7722
11	0.00032	0.00206	0.0110	2.9936	0.36516	0.00322	−0.0384	2.7936
12	0.00012	0.00189	−0.0213	2.9812	0.36154	0.00302	−0.0496	2.8342
13	0.00005	0.00174	0.0088	2.9905	0.35858	0.00286	−0.0362	2.8251
14	−0.00012	0.00160	−0.0023	3.0230	0.35633	0.00266	−0.0466	2.8730
15	0.00003	0.00151	−0.0133	3.0185	0.35457	0.00253	−0.0370	2.8489
16	0.00008	0.00139	−0.0193	3.0052	0.35271	0.00242	−0.0297	2.8659
17	−0.00005	0.00132	0.0073	3.0413	0.35108	0.00227	−0.0304	2.8528
18	0.00004	0.00124	−0.0018	3.0004	0.34973	0.00218	−0.0088	2.8815
19	0.00041	0.00117	−0.0070	3.0083	0.34829	0.00207	−0.0297	2.8922
20	0.00012	0.00110	0.0065	3.0402	0.34728	0.00197	−0.0344	2.9201
21	−0.00011	0.00105	−0.0013	2.9875	0.34632	0.00189	−0.0368	2.9358
22	0.00008	0.00101	−0.0043	3.0316	0.34525	0.00183	−0.0097	2.4972
23	0.00000	0.00097	−0.0122	2.9970	0.34431	0.00176	−0.0111	2.9226
24	−0.00003	0.00091	−0.0141	3.0537	0.34360	0.00170	−0.0114	2.8925
25	−0.00006	0.00088	−0.0112	3.0061	0.34273	0.00161	−0.0260	2.9097
26	0.00013	0.00084	−0.0153	3.0136	0.34225	0.00156	−0.0387	2.9512
27	0.00003	0.00081	0.0056	3.0316	0.34167	0.00158	−0.0165	2.9120
28	0.00010	0.00078	−0.0030	3.0137	0.34120	0.00146	−0.0155	2.9277
29	−0.00016	0.00076	0.0052	3.0226	0.34056	0.00143	−0.0282	2.9290
30	−0.00006	0.00073	0.0016	3.0383	0.34009	0.00138	−0.0131	2.9298
31	−0.00007	0.00070	−0.0157	3.0120	0.33980	0.00132	−0.0308	2.9065
32	0.00008	0.00069	−0.0055	3.0323	0.33904	0.00129	−0.0083	2.9174
33	0.00009	0.00066	−0.0051	3.0171	0.33901	0.00126	−0.0215	2.9402
34	−0.00012	0.00064	−0.0050	3.0454	0.33827	0.00122	−0.0056	2.9411
35	0.00002	0.00062	0.0023	2.9721	0.33815	0.00119	−0.0126	2.9340
36	0.00008	0.00060	−0.0048	2.9961	0.33789	0.00115	−0.0139	2.9395
37	0.00003	0.00058	−0.0144	3.0520	0.33752	0.00114	0.0082	2.9454
38	−0.00007	0.00057	−0.0002	3.0222	0.33739	0.00111	−0.0187	2.9512
39	0.00000	0.00056	0.0191	3.0391	0.33712	0.00107	−0.0130	2.9595
40	−0.00007	0.00055	0.0161	3.0269	0.33682	0.00104	−0.0267	2.9823
41	−0.00001	0.00052	−0.0189	3.0157	0.33649	0.00103	−0.0085	2.9511
42	0.00010	0.00052	0.0076	2.9818	0.33602	0.00100	−0.0074	2.9569

Continued

TABLE 12.8 Continued

	U				*V*			
n	**Mean**	**Variance**	**Skew.**	**Curtos.**	**Mean**	**Variance**	**Skew.**	**Curtos.**
43	−0.00006	0.00050	0.0036	3.0233	0.33578	0.00098	−0.0099	2.9548
44	0.00001	0.00050	0.0044	3.0025	0.33583	0.00096	−0.0161	2.9515
45	−0.00022	0.00048	0.0025	3.0340	0.33568	0.00093	−0.0128	2.9680
46	0.00004	0.00047	0.0058	3.0439	0.33548	0.00092	−0.0241	2.9734
47	0.00011	0.00045	−0.0022	2.9621	0.33532	0.00089	−0.0068	2.9415
48	0.00004	0.00045	0.0015	3.0596	0.33520	0.00087	−0.0077	2.9795
49	−0.00005	0.00044	−0.0015	3.0063	0.33401	0.00086	−0.0270	2.9255
50	0.00006	0.00044	0.0104	2.9943	0.33473	0.00084	−0.0322	2.9442
51	−0.00007	0.00042	−0.0014	3.0078	0.33448	0.00083	0.0030	2.9332
52	−0.00011	0.00041	0.0086	2.9636	0.33420	0.00081	−0.0207	3.0032
53	0.00000	0.00041	0.0046	3.0510	0.33435	0.00080	−0.0141	2.9716
54	−0.00008	0.00040	−0.0050	2.9991	0.33417	0.00079	0.0080	2.9615
55	0.00006	0.00040	0.0117	3.0644	0.33415	0.00078	0.0028	2.9750
56	−0.00003	0.00038	−0.0100	3.0360	0.33404	0.00076	−0.0178	2.9792
57	−0.00017	0.00038	0.0024	3.0245	0.33380	0.00075	−0.0012	2.9731
58	0.00003	0.00037	0.0091	2.9809	0.33353	0.00073	−0.0038	2.9540
59	0.00006	0.00037	−0.0080	3.0141	0.33371	0.00072	−0.0128	2.9767
60	0.00006	0.00036	0.0189	2.9986	0.33362	0.00071	−0.0061	2.9756
61	−0.00004	0.00035	0.0015	2.9801	0.33339	0.00070	0.0001	2.9630
62	−0.00011	0.00035	0.0092	2.9575	0.33332	0.00069	−0.0189	2.9607
63	−0.00004	0.00034	−0.0160	3.0105	0.33315	0.00069	−0.0124	2.9839
64	0.00007	0.00033	−0.0101	3.0385	0.33300	0.00068	−0.0098	2.9753
65	−0.00001	0.00033	0.0030	3.0394	0.33277	0.00066	0.0018	2.9691
66	0.00013	0.00033	0.0282	3.0242	0.33273	0.00065	−0.0206	2.9699
67	0.00005	0.00032	0.0199	3.0353	0.33273	0.00064	−0.0151	2.9870
68	0.00007	0.00032	0.0138	3.0174	0.33265	0.00063	−0.0028	2.9478
69	0.00001	0.00031	0.0091	3.0329	0.33263	0.00062	−0.0022	2.9994
70	−0.00001	0.00030	−0.0072	3.0419	0.33230	0.00062	0.0111	2.9338
71	0.00018	0.00030	0.0129	3.0295	0.33248	0.00061	−0.0105	2.9610
72	−0.00004	0.00030	0.0067	3.0035	0.33231	0.00060	−0.0231	2.9837
73	−0.00010	0.00030	−0.0108	3.0114	0.33204	0.00059	−0.0071	2.9926
74	0.00002	0.00029	0.0082	3.0396	0.33210	0.00058	−0.0096	2.9811
75	−0.00001	0.00028	−0.0138	3.0576	0.33221	0.00058	−0.0013	2.9766
76	0.00001	0.00028	0.0150	2.9840	0.33202	0.00056	0.0057	2.9676
77	0.00002	0.00028	−0.0006	3.0450	0.33204	0.00055	0.0064	2.9976
78	0.00001	0.00028	0.0187	2.9942	0.33183	0.00056	−0.0099	2.9492
79	−0.00001	0.00027	0.0142	3.0058	0.33195	0.00054	−0.0021	2.9621
80	0.00009	0.00027	−0.0015	2.9686	0.33184	0.00054	−0.0041	2.9766
81	0.00000	0.00027	0.0094	3.0087	0.33175	0.00053	−0.0045	2.9790
82	−0.00002	0.00026	−0.0045	3.0379	0.33182	0.00052	−0.0199	2.9482
83	0.00012	0.00026	0.0130	2.9860	0.33164	0.00052	−0.0102	2.9827

Continued

TABLE 12.8 Continued

	U				*V*			
n	**Mean**	**Variance**	**Skew.**	**Curtos.**	**Mean**	**Variance**	**Skew.**	**Curtos.**
84	0.00004	0.00026	0.0038	3.0150	0.33167	0.00051	−0.0053	2.9966
85	0.00001	0.00025	−0.0090	3.0179	0.33158	0.00051	0.0110	2.9789
86	−0.00012	0.00025	−0.0082	2.9790	0.33144	0.00050	−0.0045	2.9701
87	−0.00001	0.00025	−0.0125	3.0148	0.33136	0.00049	0.0042	2.9658
88	0.00001	0.00024	−0.0044	2.9644	0.33126	0.00049	0.0028	2.9725
89	0.00003	0.00024	−0.0060	3.0055	0.33123	0.00048	−0.0245	2.9970
90	−0.00004	0.00024	−0.0149	2.9977	0.33110	0.00048	−0.0047	2.9775
91	−0.00011	0.00023	−0.0031	2.9945	0.33113	0.00047	0.0022	2.9718
92	0.00003	0.00023	0.0052	3.0332	0.33115	0.00047	0.0019	2.9709
93	−0.00001	0.00023	−0.0034	2.9954	0.33136	0.00047	0.0138	2.9983
94	0.00016	0.00023	0.0202	3.0274	0.33115	0.00046	−0.0221	3.0140
95	0.00004	0.00022	−0.0046	3.0133	0.33100	0.00046	−0.0097	2.9855
96	−0.00002	0.00023	−0.0026	3.0332	0.33086	0.00045	−0.0102	2.9731
97	0.00014	0.00022	0.0147	2.9938	0.33083	0.00045	−0.0136	2.9566
98	−0.00002	0.00022	−0.0049	3.0553	0.33081	0.00044	−0.0041	2.9902
99	0.00002	0.00022	0.0138	3.0012	0.33091	0.00044	0.0032	2.9974
100	0.00002	0.00021	0.0132	3.0208	0.33086	0.00043	0.0007	2.9950

12.5.4. Distribution Approximation

In this section we present a graphical procedure for approximating the underlying PDF of a set of random data based on the goodness-of-fit test procedure discussed in Section 12.5.3.

Following a similar approach to that outlined in Section 12.5.3, random samples are generated from many different univariate probability distributions. For each specified distribution and for a given n, the statistic $Q_n = (U_n, V_n)$ given by Equation 12.243 is obtained for various choices of the shape parameter. Thus, each distribution is represented by a trajectory in the two dimensional plane whose coordinates are U_n and V_n. Figure 12.46 shows an example of such a representation. Twelve distributions namely Gaussian (1), Uniform (2), Exponential (3), Laplace (4), Logistic (5), Cauchy (6), Extreme Value (7), Gumbel type-2 (8), Gamma (9), Pareto (10), Weibull (11), and Lognormal (12) are represented in this chart. Tables 12.9 and 12.10 show the standard form and the general form, respectively, of the PDFs represented in the identification chart. The value of Q_n at each point of the trajectories is obtained by Monte-Carlo experiments using the standard Gaussian distribution as the reference distribution for determining the angles θ_i. The results are based on averaging 1000 trials of $n = 50$ samples from each distribution. The samples from each distribution are obtained by using the IMSL subroutines for specified values of the shape parameter. Since the procedure is location and scale invariant, the trajectory

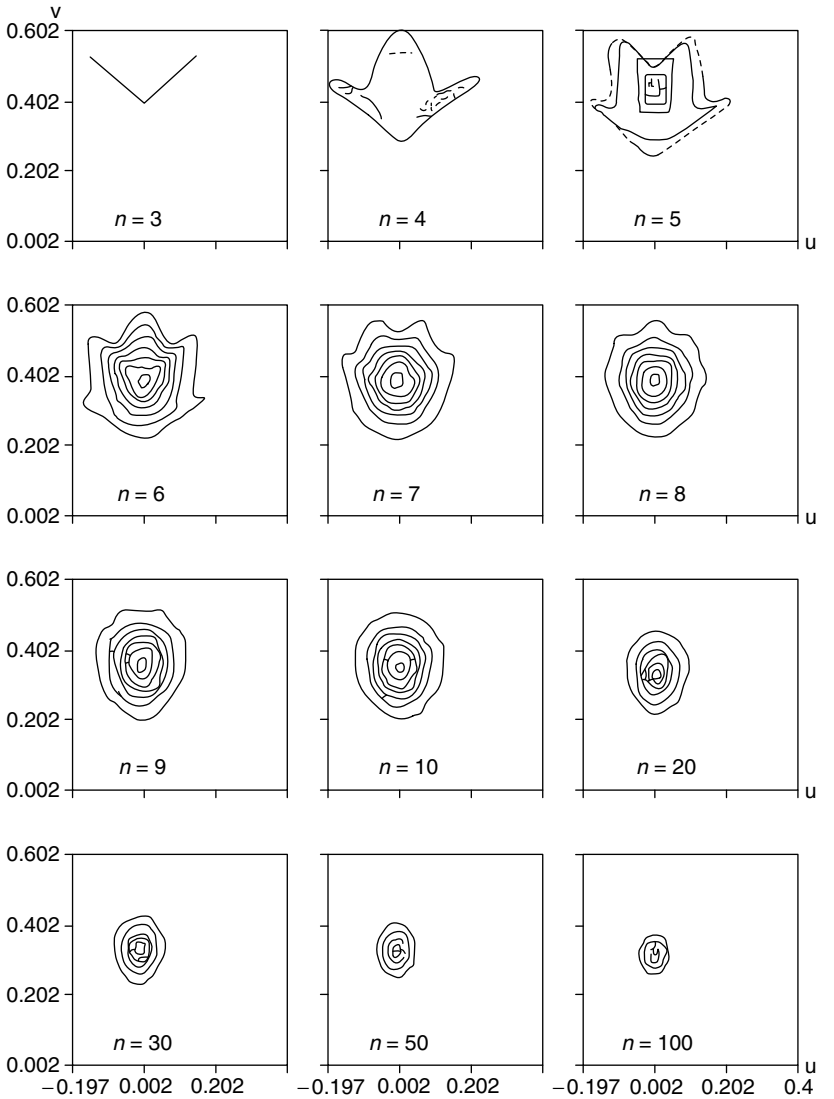

FIGURE 12.45 Empirical distribution of Q_n for several values of n.

reduces to a single point for those PDFs which do not have shape parameters but are characterized only in terms of their location and scale parameters. By way of example, the Gaussian, Laplace, Exponential, Uniform, and Cauchy PDFs are represented by single points in the $U_n - V_n$ plane. However, those PDFs having shape parameters are represented by trajectories. For a given value of the shape parameter, a single point is obtained in the $U_n - V_n$ plane. By varying the shape parameter, isolated points are determined along the trajectory. The trajectory for

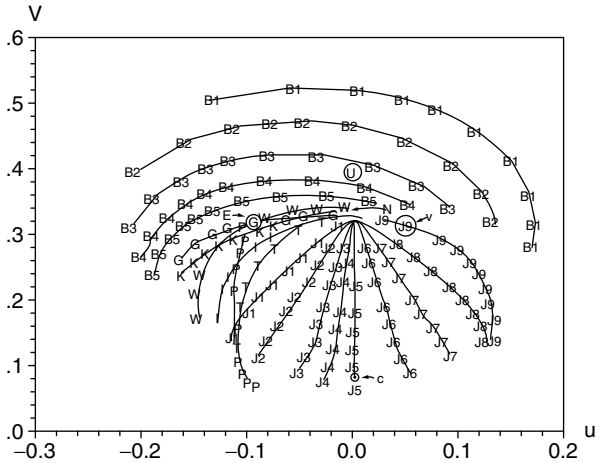

FIGURE 12.46 Identification chart for univariate distributions based on 1000 Samples ($n = 50$). B1 − B5 = Beta, J1 − J9 = SU Johnson, G = Gamma, W = Weibull, K = K-distribution, P = Pareto, I = Lognormal, T = Gumbel, E = Ex ponential, V = Extreme Value, A = Laplace, L = Logistic, U = Uniform, C = Cauchy.

TABLE 12.9
Table of Standard Forms of Univariate PDFs Used for Identification Chart of Figure 12.46

Distribution	Standard Form $f_Y(y)$	
Gaussian	$\left(\sqrt{2\pi}\right)^{-1}\exp\left(-\frac{y^2}{2}\right)$	$-\infty < y < \infty$
Uniform	1	$0 < y < 1$
Exponential	$\exp(-y)$	$(0 < y < \infty)$
Laplace	$0.5\exp(-\lvert y\rvert)$	$-\infty < y < \infty$
Logistic	$\exp(-y)[1+\exp(-y)]^{-2}$	$-\infty < y < \infty$
Cauchy	$\frac{1}{\pi(1+y^2)}$	$-\infty < y < \infty$
Extreme value (Type 1)	$\exp(-y)\exp[-\exp(-y)]$	$-\infty < y < \infty$
Gumbel (Type 2)	$\gamma y^{-\gamma-1}\exp(-y^{-\gamma})$	$-\infty < y < \infty$
Gamma	$\frac{1}{\Gamma(\alpha)}\exp(-y)y^{\gamma-1}$	$0 < y < \infty$
Pareto	$\frac{\gamma}{y^{\gamma+1}}$	$y > 1,\ \gamma > 0$
Weibull	$\gamma y^{\gamma-1}\exp(-y^{\gamma})$	$y > 0$
Lognormal	$\frac{\gamma}{\sqrt{2\pi}y}\exp\left[-\frac{\{\gamma\log(y)\}^2}{2}\right]$	$y > 0$

TABLE 12.10
Table of General Form of Univariate PDFs Used for Identification Chart of Figure 12.46

Distribution	General Form $f_x(\chi)$	
Gaussian	$\left(\sqrt{2\pi}\beta\right)^{-1}\exp\left(-\frac{(x-\alpha)^2}{2\beta^2}\right)$	$-\infty < x < \infty$
Uniform	$\frac{1}{\beta}$	$\alpha < x < \alpha+\beta$
Exponential	$\frac{1}{\beta}\exp\left(-\frac{(x-\alpha)}{\beta}\right)$	$\alpha < x < \infty$
Laplace	$\frac{0.5}{\beta}\exp\left[-\left\lvert\frac{(x-\alpha)}{\beta}\right\rvert\right]$	$-\infty < x < \infty$
Logistic	$\frac{1}{\beta}\exp\left(-\frac{(x-\alpha)}{\beta}\right)\left[1+\exp\left(-\frac{\{x-\alpha\}}{\beta}\right)\right]^{-2}$	$-\infty < x < \infty$
Cauchy	$\frac{1}{\pi\beta\left[1+\frac{(x-\alpha)^2}{\beta^2}\right]}$	$-\infty < x < \infty$
Extreme value (Type 1)	$\frac{1}{\beta}\exp\left[-\frac{(x-\alpha)}{\beta}\right]\exp\left[-\exp\left\{-\frac{(x-\alpha)}{\beta}\right\}\right]$	$-\infty < x < \infty$
Gumbel (Type 2)	$(\gamma/\beta)[(x-\alpha)/\beta]^{-\gamma-1}\exp\left[-\frac{(x-\alpha)^{-\gamma}}{\beta_\gamma}\right]$	$\alpha < X < \infty,\ \beta > 0,\ \gamma > 0$
Gamma	$\frac{1}{\beta\Gamma(\alpha)}\exp\left[-\frac{(x-\alpha)}{\beta}\right]\left(\frac{x-\alpha}{\beta}\right)^{\gamma-1}$	$\alpha < x < \infty$
Pareto	$\frac{\gamma}{\beta}\frac{1}{\left(\frac{x-\alpha}{\beta}\right)^{\gamma+1}}$	$x > \alpha+\beta,\ \gamma > 0$
Weibull	$\frac{\gamma}{\beta}\left(\frac{x-\alpha}{\beta}\right)^{\gamma-1}\exp\left[-\left(\frac{x-\alpha}{\beta}\right)^{\gamma}\right]$	$x > \alpha$
Lognormal	$\frac{\gamma}{\sqrt{2\pi}\beta\left(\frac{x-\alpha}{\beta}\right)}\exp\left[-\frac{\left\{\gamma\log\left(\frac{x-\alpha}{\beta}\right)\right\}^2}{2}\right]$	$x > \alpha$

the PDF is obtained by joining these points. In a sense the trajectory represents a family of PDFs having the same distribution but with different shape parameters values. For example, the trajectory corresponding to the Gamma distribution in Figure 12.46 is obtained by joining the points for which the shape parameters are 0.2, 0.3, 0.5, 0.7, 1.0, 2.0, 3.0, 4.0, 6.0, and 10.0. As the shape parameter increases, note that the Gamma distribution approaches the Gaussian distribution. The representation of Figure 12.46 is called an identification chart. Some distributions, such as, the β distribution and the SU-Johnson system of

distributions, have two shape parameters. For these cases, the trajectories are obtained by holding one shape parameter fixed while the other is varied. For these distributions, several different trajectories are generated in order to cover as much of the $U_n - V_n$ plane as possible. For certain choices of the shape parameters, two or more PDFs become identical. When this occurs, their trajectories intersect on the identification chart.

It is apparent that the identification chart of Figure 12.46 provides a one-to-one graphical representation for each PDF for a given n. Therefore, every point in the identification chart corresponds to a specific distribution. Thus, if the null hypothesis in the goodness-of-fit test discussed in Section 12.5.3 is rejected, then the distribution that approximates the underlying PDF of the set of random data can be found by comparing Q_n obtained for the samples with the existing trajectories in the chart. The closet point or trajectory to the sample Q_n is chosen as an approximation to the PDF underlying the random data. The closet point or trajectory to the sample point is determined by projecting the sample point Q_n to neighboring points or trajectories on the chart and considering that point or trajectory whose perpendicular distance from the sample point is the smallest. Consider the situation of Figure 12.47. Let $Q_n = (u', v')$ denote the coordinates of the sample point. Let x_1, y_1 and x_2, y_2 denote the coordinates of the points A and B on the trajectory shown in Figure 12.47. It is assumed that the segment of the trajectory between the points A and B is linear. Let x_0, y_0 denote the coordinates of the point of intersection of the straight line between A and B and the projection of $Q_n = (u', v')$ onto this straight line. The equation of the straight line between the points A and B can be written as

$$y - y_1 = m(x - x_1) \tag{12.255}$$

where $m = \dfrac{y_2 - y_1}{(x_2 - x_1)}$. Also, the equation of the straight line joining x_0, y_0 and (u', v') is

$$y - v' = -\frac{1}{m}(x - u') \tag{12.256}$$

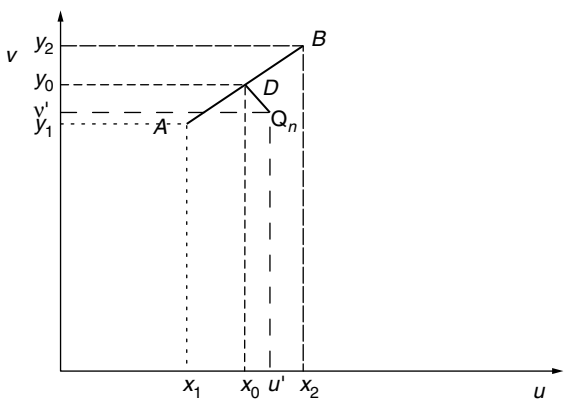

FIGURE 12.47 Distance computation.

The coordinates x_0, y_0 result from the solution of Equation 12.255 and Equation 12.256 and are given by

$$\begin{aligned} x_0 &= \frac{1}{m^2+1}[m^2 x_1 - m y_1 + u' + m v'] \\ y_0 &= \frac{1}{m^2+1}[y_1 - m x_1 + m^2 v' + m u'] \end{aligned} \tag{12.257}$$

Finally, the perpendicular distance from the sample point onto the trajectory between the points A and B is

$$D = \sqrt{\frac{1}{(m^2+1)}[m^2\zeta_1^2 - 2m\,\zeta_1\zeta_2 + \zeta_2^2]} \tag{12.258}$$

where

$$\begin{aligned} \zeta_1 &= u' - x_1 \\ \zeta_2 &= v' - y_1 \end{aligned} \tag{12.259}$$

The complete approximation algorithm is summarized as follows:

1. Sort of samples $X_1, X_2, \ldots, X_n$ in increasing order.
2. Obtain the standardized order statistic $Y_{i:n}$.
3. Compute U_n and V_n from Equation 12.243.
4. Obtain an identification chart based on the sample size n and plot the sample point Q_n on this chart.
5. Compute D using the sample point Q_n and the existing distributions on the chart. Choose the PDF corresponding to the point or trajectory, which results in the smallest value of D as an approximation to the PDF of the samples.

The approximation to the underlying PDF of the set of random data can be improved by including as many distributions as possible in the identification chart so as to fill as much of the space as possible with candidate distributions. However, it is emphasized that this procedure does not identify the underlying PDF. Rather it identifies a suitable approximation to the underlying PDF.

12.5.5. Parameter Estimation

Once the probability distribution of the samples is approximated, the next step is to estimate its parameters. The method discussed in Section 12.5.4 lends itself for estimating the parameters of the approximated distribution. We present the estimation procedure for the location, scale, and shape parameters in this section.

12.5.5.1. Estimation of Location and Scale Parameters

Let $f(x;\alpha,\beta,)$ denote a known distribution that approximates the PDF of the set of random data, where α and β are the location parameter and scale parameter, respectively, of the approximating PDF. Let $X_{i:n}$ denote the ordered statistics of X from a sample of size n. A standardized ordered statistic is defined by

$$W_{i:n} = \frac{X_{i:n} - \alpha}{\beta} \tag{12.260}$$

Let

$$\mu_{i:n} = E[W_{i:n}] \tag{12.261}$$

Then

$$E[X_{i:n}] = \beta\mu_{i:n} + \alpha \tag{12.262}$$

We consider the following statistics

$$T_1 = \sum_i^n \cos(\theta_i)X_{i:n} \qquad T_2 = \sum_i^n \sin(\theta_i)X_{i:n} \tag{12.263}$$

where θ_i is the angle defined in Equation 12.241. The expected values of T_1 and T_2 are

$$E[T_1] = \sum_i^n \cos(\theta_i)[\beta\mu_{i:n} + \alpha] \qquad E[T_2] = \sum_i^n sin(\theta_i)[\beta\mu_{i:n} + \alpha] \tag{12.264}$$

These can be written as

$$E(T_1) = a\alpha + b\beta \qquad E(T_2) = c\alpha + d\beta \tag{12.265}$$

where

$$\begin{aligned} a &= \sum_i^n \cos(\theta_i) \quad b = \sum_i^n \mu_{i:n}\cos(\theta_i) \\ c &= \sum_i^n \sin(\theta_i) \quad d = \sum_i^n \mu_{i:n}\sin(\theta_i) \end{aligned} \tag{12.266}$$

Because the standardized Gaussian distribution is used as the reference distribution for θ_i, it follows that $a = 0$.[41] The estimates for β and α are then given by

$$\hat{\beta} = \frac{\hat{E}[T_1]}{\hat{b}} \qquad \hat{\alpha} = \frac{\hat{E}[T_2] - \hat{d}\hat{\beta}}{c} \tag{12.267}$$

where the symbol Λ is used to denote an estimate. For n sufficiently large (i.e., $n > 50$), suitable estimates for $E[T_1]$ and $E[T_2]$ are

$$\hat{E}[T_1] = T_1 \qquad \hat{E}[T_2] = T_2 \tag{12.268}$$

Estimates for b and d rely upon an estimate of $\mu_{i:n}$. The parameter $\hat{\mu}_{i:n}$ is obtained from a Monte Carlo simulation of $W_{i:n}$ where $W_{i:n}$ is generated from the known approximating distribution $f(x; 0, 1)$ having zero location and unity scale parameters. $\hat{\mu}_{i:n}$ is the sample mean of $W_{i:n}$ based upon 1000 Monte Carlo trials. Having $\hat{\mu}_{i:n}$, the estimates for b and d are given by

$$\begin{aligned} \hat{b} &= \sum_{i}^{n} \mu_{i:n} \cos(\theta_i) \\ \hat{d} &= \sum_{i}^{n} \mu_{i:n} \sin(\theta_i) \end{aligned} \tag{12.269}$$

The scale and location parameters are then estimated by application of Equation 12.267.

12.5.5.2. Shape Parameter Estimation

In this section we present an approximate method for estimating the shape parameter of the approximating PDF. This procedure can be used only when the approximating PDF has a single shape parameter to be determined. Let γ denote the shape parameter of the approximating PDF. Since U_n and V_n are location and scale invariant, the point Q_n depends only on the sample size n and the shape parameter γ.

Recall that the trajectories on the identification chart are obtained by averaging the results of a large number of trials for U_n and V_n. Consequently, for a given value of n, the coordinates of the points along the trajectory for a specified distribution and can be characterized by

$$\begin{aligned} E(U_n) &= \varphi_1(n, \gamma) \\ E(V_n) &= \varphi_2(n, \gamma) \end{aligned} \tag{12.270}$$

where the complete trajectory is obtained by repeating the large number of trials over a suitable range of γ. On a given trial involving the random data it is likely that the coordinates U_n and V_n for the samples will not coincide with any of the trajectories on the chart. The random data is approximated by the distribution that falls closest to the sample point Q_n. The situation is illustrated in Figure 12.48. $Tr1$ and $Tr2$ denote the trajectories for two different candidate distributions denoted by PDF1 and PDF2, respectively. Let x_0 denote the point on $Tr1$ closest to Q_n. Assume that the linear segment of $Tr1$ on which x_0 falls was drawn between the points (u_1, v_1) and (u_2, v_2). Let the shape parameter values corresponding to these points be denoted by γ_1 and γ_2, respectively. Then the value of the shape parameter corresponding to the sample point Q_n is obtained by

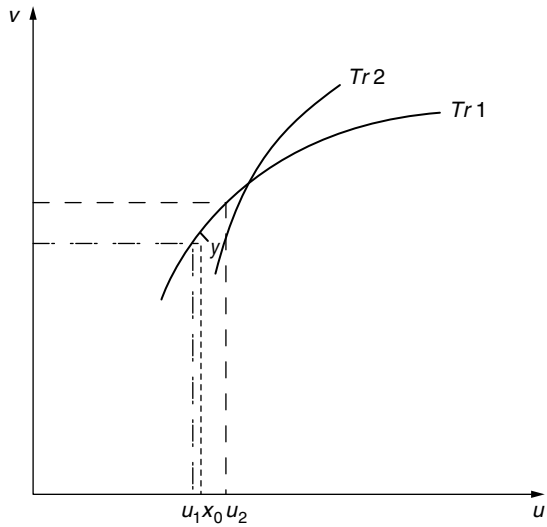

FIGURE 12.48 Shape parameter estimation.

linear interpolation and is given by

$$\hat{\gamma} \approx \gamma_1 + \frac{(\gamma_2 - \gamma_1)(x_0 - u_1)}{(u_2 - u_1)} \tag{12.271}$$

where

$$\begin{aligned} x_0 &= \frac{\{A(V_n - v_1) + A^2 u_1 + U_n\}}{(A^2 + 1)} \\ A &= \frac{(v_2 - v_1)}{(u_2 - u_1)} \end{aligned} \tag{12.272}$$

The accuracy of the procedure can be improved by employing a nonlinear interpolation method. It must be emphasized that the location, scale, and shape parameter estimation procedures presented in this section are approximate procedures.

12.5.6. Assessing the Distributional Properties of SIRVs

A random vector $\mathbf{Y} = [Y_1, Y_2, \ldots, Y_N]^{\mathrm{T}}$ is a SIRV if its PDF has the form

$$f_{\mathbf{Y}}(\mathbf{y}) = (2\pi)^{-\frac{N}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} h_N(p) \tag{12.273}$$

where $p = (\mathbf{y} - \mu)^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{y} - \mu)$ is a nonnegative quadratic form, μ and $\boldsymbol{\Sigma}$ are the mean vector and covariance matrix, respectively, of $\mathbf{Y}$ and $h_N(p)$ is a nonnegative, monotonically decreasing, and real valued function for all N.

Recall from Section 12.5.3 that the PDF of the quadratic form appearing in Equation 12.273 is given by

$$f_P(p) = \frac{p^{\frac{N}{2}} - 1}{2^{\frac{N}{2}}\Gamma\left(\frac{N}{2}\right)} h_N(p)u(p) \tag{12.274}$$

where $\Gamma(\alpha)$ is the Eulero-Gamma function and $u(p)$ is the unit step function. It was also pointed out in Section 12.5.3 that the PDF of the quadratic form is invariant to the choice of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. For example, in the multivariate Gaussian case, the PDF of the quadratic form is the well known Chi-square distribution with N degrees of freedom. Therefore, for a given N, the SIRV is uniquely characterized by the quadratic form. In order to identify the PDF of the underlying SIRV it is sufficient to identify the PDF of the quadratic form. This attractive property of SIRVs enables us to study various distributional aspects of the corresponding multivariate samples. When a radar uses coherent processing, the joint PDF of the $2N$ quadrature components is of interest. Equation 12.273 and Equation 12.274 are then applicable with N replaced by $2N$.

In modeling real world data, the first step is to determine the most appropriate PDF that approximates the data. In the univariate case, the fit and assessment of the goodness of fit for various distributions has been studies extensively and several methods are available for this purpose. However, limited success has been achieved for the multivariate situation. Although a number of multivariate distributions have been developed, the multivariate Gaussian distribution has been the focus of much of the techniques for multivariate analysis.[46]

Assessment of the distributional assumptions for multivariate data is a non trivial problem. Several techniques have been proposed to assess multivariate Gaussianity. In a recent paper Ozturk and Romeu[44] give a review of the methods for testing multivariate Gaussianity. Many of the methods can be modified or generalized to develop goodness of fit methods for SIRVs. If a random vector $\mathbf{Y}$ is a SIRV, then the corresponding marginal distributions must be identical (up to location and scale parameters). Based on this property, one can use the standard univariate goodness of fit testing procedures to assess the degree of similarity of the marginal distributions of the multivariate data. However, such an approach does not provide a way to assess the joint distribution of the components of the multivariate sample.

Since SIRVs can be uniquely characterized in terms of the quadratic form P, Equation 12.274 provides an important property for developing goodness of fit test procedures for SIRVs. Specifically, if the PDF of P can be identified, then the corresponding PDF of the SIRV can also be identified. In fact, many tests for assessment of multivariate Gaussianity are based on the use of this quadratic form.[47] By use of this technique, note that the multivariate distribution approximation problem is reduced to a corresponding univariate distribution approximation of the quadratic form. Any of the classical goodness of fit testing procedures like the Kolmogorov–Smirnov and Chi-Square tests can be used to address the problem of distribution identification of the quadratic form. However, deficiencies of these tests, namely (1) the requirement of large sample sizes for

specifying the parameters of the distributions and (2) weak decisiveness, necessitate use of alternate procedures that are more efficient.

A general algorithm was developed in Ref. 41 to test for univariate and multivariate normality and has been summarized in Sections 12.5.3 to 12.5.5. In this section we propose the use of this algorithm for performing the goodness of fit test for SIRVs. The procedure is summarized here for completeness. Let $\mathbf{Y} = [Y_1, Y_2, \ldots, Y_N]^{\mathrm{T}}$ denote a vector of observations. For each observation vector of size N, we compute the corresponding quadratic form where the maximum likelihood (ML) estimates of the mean vector of $\mathbf{Y}$ and its covariance matrix are used. For the Gaussian case, it is well known that these ML estimates are the sample mean and the sample covariance matrix, respectively.[48] In Appendix W, it is shown that the same results hold for SIRVs.[49] Our goal is to test whether $\hat{P}_i$, $(i = 1, 2, \ldots, n)$ are samples from a certain distribution $F(p; \alpha, \beta, \gamma)$ where α, β are the location and scale parameters, respectively, and γ is the shape parameter.

Let $\hat{P}_{i:n} \leq \hat{P}_{2:n} \leq \cdots \leq \hat{P}_{n:n}$ denote the ordered observations of the quadratic form $\hat{P}_i$ $(i = 1, 2, \ldots, n)$. We define the standardized ith sample order statistic as

$$R_{i:n} = \frac{(\hat{P}_{i:n} - \bar{P})}{S_P} \tag{12.275}$$

where $\bar{P}$ and S_P are the sample mean and sample standard deviation, respectively. Corresponding to the ith sample order statistics $R_{i:n}$ through $R_{i:n}$, the point $Q_i = (U_i, V_i)$ is defined where

$$U_i = \frac{1}{i}\sum_{j=1}^{i} \cos\{\pi\, \Phi(m_{j:n})\}|R_{j:n}|, \quad V_i = \frac{1}{i}\sum_{j=1}^{i} \sin\{\pi\, \Phi(m_{j:n})\}|R_{j:n}| \tag{12.276}$$

where: $\Phi(\cdot)$ and $m_{j:n}$ were defined in Section 12.5.3.

For a given set of n multivariate samples, the points $Q_i = (i = 1, 2, \ldots, n)$ are plotted and joined to obtain a linked vector chart. The linked vectors under the null hypothesis are obtained by averaging the results of 50,000 Monte Carlo trials from the PDF of the quadratic form given in Equation 12.274. The proposed test is based on comparing the sample and the null linked vectors. If the null hypothesis is true, then we expect that the sample linked vectors will follow the null linked vectors closely.

Finally, a formal goodness of fit test is performed using the terminal point of the null linked vectors (i.e., $Q_n = (U_n, V_n)$). Using the central limit theorem, as outlined in Section 12.5.3, confidence ellipses centered at Q_n for the null linked vectors are obtained. If the terminal point of the sample linked vectors does not fall inside the 100 $(1 - \alpha)$% confidence ellipse, then the corresponding null hypothesis is rejected at the α level of significance. Note that the Q_n test provides an interesting graphical representation of the data. An example of such a graphical representation is given in Figure 12.49 for testing a multivariate Gaussian distribution with $n = 50$ and $N = 4$.

It should be noted that the Q_n statistic is location and scale invariant. In other words it is independent of the location and scale parameters. However, it depends

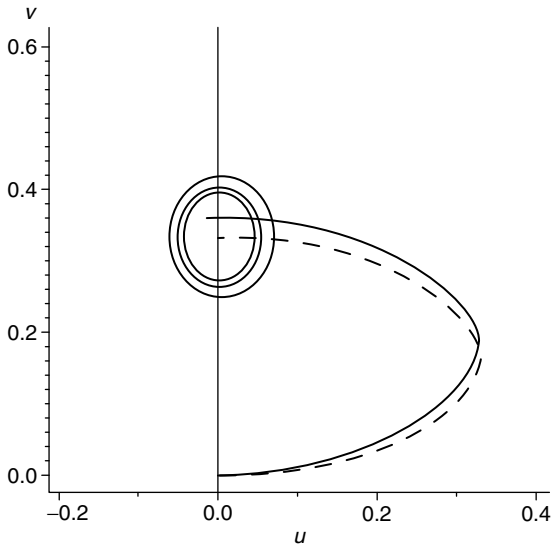

FIGURE 12.49 Goodness of fit test for SIRVs using the Q_n procedure. 90, 95, and 99% contours for the Gaussian distribution. Broken line = null distribution pattern.

on the shape parameter of the null distribution. Assessment of the distributional assumptions of distributions that have shape parameters is conceptually different from the assessment for those that do not have shape parameters. In the former case, we test whether the sample comes from a family of distributions while in the latter case, we test for a simple distribution. One possibility for dealing with this problem is to specify the value of the shape parameter and perform the test in the usual way. If the shape parameter cannot be specified, then an adaptive approach using a sample estimate of the shape parameter must be employed.

Advantages of using the Q_n procedure are explained in Ref. 41. Usually the classical goodness of fit tests end up with either rejecting or accepting the null hypothesis. An attractive property of the Q_n procedure is that it provides some information about the true distributions if the null hypothesis is rejected. Using this property an algorithm for characterizing and identifying the distributions can be developed. The next section explains these ideas.

12.5.7. Distribution Identification of SIRVs

Following the same procedure described in Section 12.5.4, where the reference distribution was Gaussian, an identification chart can be generated for each of quadratic form PDFs of the SIRVs listed in Tables 12.9 and 12.10. Recall from Section 12.5.3 that the PDF of the quadratic form is invariant to the choice of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. Hence, for simplicity, the trajectories for the PDFs of the quadratic forms of the SIRVs listed in Tables 12.9 and 12.10 are obtained by generating the SIRVs having zero mean and identity covariance matrix. Each point on

a trajectory is obtained by averaging the results of 2000 Monte Carlo trials of size $n = 100$. As before, PDFs which do not have shape parameters are represented by a single point in the $U - V$ plane, while those with shape parameters generate a trajectory in the $U - V$ plane by changing the shape parameter.

Assuming coherent radar processing, Tables 12.11 and 12.12 provide a library of $h_{2N}(p)$ for various multivariate SIRV PDFs. An example of the identification chart is given in Figure 12.50 for $N = 4$ and $n = 2000$ where the expected values of $Q_n = (U_n, V_n)$ is plotted for various distribution. The Gaussian distribution was used as the reference distribution. The SIRVs listed in Tables 12.9 and 12.10 are included in the chart and labeled by number. It is noted that the multivariate Gaussian (1), Laplace (2), and Cauchy (3) distributions are represented by single points on the chart; while the multivariate *K*-distribution (8), Chi (9), Generalized Rayleigh (10), Weibull (11), and Rician (12) are represented by trajectories. The Student-*t* distribution (4, 5, 6, 7) with degrees of freedom 3, 5, 10, and 15, respectively, is also shown in the chart. The trajectories for each distribution were obtained by joining ten points resulting from the use of the distributions with parameter values listed in Table 12.13. Each point in the chart is obtained by simulating 2000 samples from the corresponding distributions. The methods developed by Rangaswamy et al.[22,50] were used to generate the multivariate samples in Figure 12.50.

The identification chart provides an interesting display for identifying and characterizing the distributions. Also, relationships, between the various distributions are clearly seen. For example, as their parameters are varied, certain distributions approach the multivariate Gaussian distribution. Also, for appropriately chosen parameters, the multivariate Weibull distribution and the Generalized Rayleigh distribution can be seen to coincide. For a given *N*-variate sample of size *n*, the statistic Q_n based on the sample quadratic forms can be computed and plotted on the identification chart. Then the nearest distribution to the sample point is identified to be the best candidate for the underlying true distribution of the data. An example of such an identification is shown in Figure 12.50 where a well known data set (i.e., *Iris Setosa*[51]) is used to obtain a value for Q_n and is denoted by the point *S*. The *I. Setosa* data consists of four measurements taken from 50 plants. It is seen from Figure 12.50 that the best candidate for approximating the data is the multivariate Chi(9) distribution.

We point out that there are other methods which can be used for the distribution identification problem. A commonly used technique is the $Q - Q$ plot. To identify the underlying distribution the sample quantiles are plotted against the expected quantiles of a reference distribution. Then the resulting shape of the plotted curve is taken as a basis for identifying the corresponding candidate for the true distributions. However, the identification is made on a subjective basis. Even then the procedure is not very easy. Another well known approach of distinguishing between distribution is to characterize them via their skewness (α_3) and kurtosis (α_4) coefficients. In this case, all the distributions are represented by points on the $\alpha_3 - \alpha_4$ plane and the sample data point is compared

TABLE 12.11
SIRVs Obtained from the Marginal Envelope PDF

Marginal PDF	$h_{2N}(p)$
Chi	$(-2)^{N-1}A\sum_{k=1}^{N} G_k p^{\nu-k}\exp(-Bp)$
	$G_k = \binom{N-1}{k-1}(-1)^{k-1}B^{k-1}\dfrac{\Gamma(\nu)}{\Gamma(\nu-k+1)}$
	$A = \dfrac{2}{\Gamma(\nu)}(b\sigma)^{2\nu}$
	$B = b^2\sigma^2$
	$\nu \le 1$
Weibull	$\sum_{k=1}^{N} C_k p^{(kb/2)-N}\exp\left(-Ap^{(b/2)}\right)$
	$A = a\sigma^b$
	$C_k = \sum_{m=1}^{k}(-1)^{m+N}2^N\dfrac{A^k}{k!}\binom{k}{m}\dfrac{\Gamma\left(1+\dfrac{mb}{2}\right)}{\Gamma\left(1+\dfrac{mb}{2}-N\right)}$
	$b \le 2$
Gen. Rayleigh	$\sum_{k=1}^{N} D_k p^{(k\alpha/2)-N+1}\exp\left(-Bp^{(\alpha/2)}\right)$
	$A = \dfrac{\sigma^2\alpha}{\beta^2\Gamma\left(\dfrac{2}{\alpha}\right)}$
	$B = \beta^{-\alpha}\sigma^{\alpha}$
	$D_k = \sum_{m=1}^{k}(-1)^{m+N-1}2^{N-1}\dfrac{B^k}{k!}\binom{k}{m}\dfrac{\Gamma\left(1+\dfrac{m\alpha}{2}\right)}{\Gamma\left(2+\dfrac{m\alpha}{2}-N\right)}$
	$\alpha \le 2$
Rician	$\dfrac{\sigma^{2N}}{(1-\rho^2)^{N-(1/2)}}\sum_{k=0}^{N-1}\binom{N-1}{k}(-1)^k\left(\dfrac{\rho}{2}\right)^k\xi_k\exp(-A)$
	$\xi_k = \sum_{m=0}^{k}\binom{k}{m}I_{k-2m}(\rho A), \qquad A = \dfrac{p\sigma^2}{2(1-\rho^2)}$

with the points representing the theoretical distributions in the same way as in the Q_n procedure. However, estimates of α_3 and α_4 are known to be highly sensitive to extreme observations and therefore, large sample sizes are necessary to perform the identification for a given degree of accuracy.

TABLE 12.12
SIRVs Obtained from the Marginal Characteristic Function

Marginal PDF	$h_{2N}(p)$
Gaussian	$\exp\left(-\frac{P}{2}\right)$
Laplace	$b^{2N}(b\sqrt{p})^{1-N}K_{N-1}(b\sqrt{p})$
Cauchy	$\frac{2^N b\Gamma\left(\frac{1}{2}+N\right)}{\sqrt{\pi}(b^2+p)^{N+(1/2)}}$
K-distribution	$\frac{b^{2N}}{\Gamma(\alpha)}\frac{(b\sqrt{p})^{\alpha-N}}{2^{\alpha-1}}K_{N-\alpha}(b\sqrt{p})$
Student-t	$\frac{2^N b^{2\nu}\Gamma(\nu+N)}{\Gamma(\nu)(b^2+p)^{N+\nu}}$

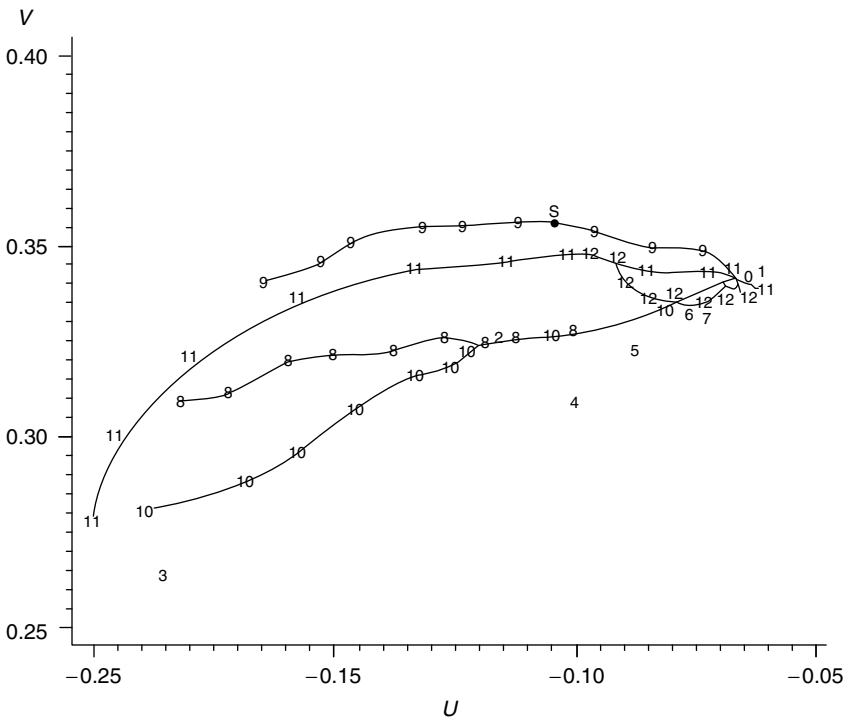

FIGURE 12.50 Identification chart for SIRVs ($n = 2000$, $N = 4$). 1 = Gaussian, 2 = Laplace, 3 = Cauchy, 4 to 7 = Student-t, 8 = K-Distribution, 9 = Chi, 10 = Generalized Rayleigh, 11 = Weibull, 12 = Rician. S is *Iris Setosa* data point.

TABLE 12.13
Shape Parameters of the SIRVs Used for the Identification Chart

K-distribution	0.1, 0.2, 0.3, 0.4, 0.5, 0.7, 0.9, 1.1, 1.5, 1.9
Chi	0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.5, 0.6, 0.75, 0.95
Gen. Rayleigh	0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 1.0, 1.5, 2.0
Weibull	0.3, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0
Rician	0.15, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.85, 0.9

12.5.8. Alternative Method for Parameter Estimation

It is well known that the ML estimate of the covariance matrix of a Gaussian random vector is the sample covariance matrix. Interestingly enough, it has been shown in Ref. 49 that the ML estimate of the covariance matrix $\boldsymbol{\Sigma}$ for SIRVs is the same sample covariance matrix used in the Gaussian case. From Equation 12.274, it is clear that the expected value of the quadratic form can be expressed as

$$E[P] = \varphi(N, \gamma) \tag{12.277}$$

where γ is the shape parameter of the distribution. For those SIRVs where $\varphi(\cdot)$ can be evaluated in closed form and is invertible, the sample mean of P, denoted by $\bar{P}$ can be used to estimate the shape parameter according to

$$\hat{\gamma} = \varphi^{-1}\{\bar{P}, N\} \tag{12.278}$$

where $\bar{P} = \frac{1}{n}\sum_{i=1}^{n} P_i$. For example, in case of the *K*-distribution, we have $E[P] = 2\alpha N$ where α is the shape parameter of the *K*-distribution. Clearly, the shape parameter is given by $\hat{\alpha} = \frac{\bar{P}}{2N}$. Unfortunately, it is not always possible to obtain an invertible closed form expression for $\varphi(\ldots)$. The shape parameter estimation procedure suggested here is not suitable in such a case. An alternate method for the parameter estimation problem is then needed.

In this section we propose to use the Q_n statistic to obtain an approximate estimator for the shape parameter. The underlying procedure is explained in Ref. 41 and is summarized here. Let the points (U_1, V_1) and (U_2, V_2) denote the coordinates of Q_n corresponding to the parameters γ_1 and γ_2, respectively, of a given SIRV. Suppose these points are the nearest points, on the trajectory for the identical distribution to the sample point $Q_n = (U_n, V_n)$ corresponding to the data. Then by using linear interpolation, an approximate estimator of γ is given by

$$\hat{\gamma} \approx \gamma_1 + \frac{(\gamma_2 - \gamma_1)(x_0 - U_1)}{(U_2 - U_1)} \tag{12.279}$$

where

$$x_0 = \frac{\{A(V_n - V_1) + A^2 U_1 + U_n\}}{(A^2 + 1)}, \quad A = \frac{(V_2 - V_1)}{(U_2 - U_1)} \tag{12.280}$$

The accuracy of the proposed estimator for γ depends on the distance between sample point Q_n and the corresponding curve. If necessary, the approximation can be improved by using nonlinear interpolation methods.

12.5.9. Conclusions

In this chapter we have addressed the problem of distribution identification of radar clutter under the assumption that the clutter can be characterized as a SIRP. First and foremost, we have shown that the multivariate distribution identification problem for SIRPs can be reduced to an equivalent univariate distribution identification problem of a nonnegative quadratic form, resulting in considerable processing simplicity. A new algorithm, providing a graphical representation for the goodness of fit test and the distribution identification, has been used. This algorithm, while conceptually simple, it extremely efficient while dealing with small sample sizes. Therefore, it is suitable for use in a variety of practical applications. Finally, based on this algorithm, a new approach has been proposed for estimating the shape parameter of SIRPs.

12.6. CONCLUSIONS

12.6.1. General Remarks

We present conclusions and suggestions for future work in this chapter. We have addressed the problem of modeling, simulation, and distribution identification (multivariate) of correlated nonGaussian radar clutter that can be characterized as a SIRP. The SIRP model for the clutter belongs to the class of exogenous product models, where the clutter process can be decomposed as a product of two independent random processes. One of the processes is Gaussian, while the other is a highly correlated nonGaussian process modulating the Gaussian process. The SIRP model arises as a special case when the modulating process is a nonnegative random variable. This in turn, imposes the requirement that the modulating random process should have a decorrelation time much larger than that of the Gaussian process, so that the modulating process is approximately constant in a given time observation interval.

For example, consider a high resolution airborne radar operating in a maritime environment at low grazing angles. The overall sea clutter return is composed of returns from the capillary waves and the gravity waves. The capillary waves correspond to a rapidly time variant process. It has been pointed out in Ref. 19 that the returns from the capillary waves can be approximated as being jointly Gaussian. Therefore, the returns from the capillary waves can be modeled by a Gaussian random process. The gravity waves correspond to a slowly time variant phenomenon. Furthermore, it has been shown in Ref. 3, 7, and 19 that the decorrelation time of the slowly time variant process is much larger than that of the Gaussian process. Consequently, the slowly time variant process can be approximated by a nonnegative random variable. Hence, the SIRP model

is applicable in this case. In fact, if the nonnegative random variable has a Chi-distribution, the overall sea clutter returns are characterized by a K-distribution. We have pointed out that the K-distribution is a member of the family of SIRPs. Therefore, for this case, the SIRP characterization enables us to determine the optimal radar signal processor. The validity of other SIRPs as models for radar clutter must be determined through an experimental effort.

This dissertation has made the following significant original contributions:

1. Application of the theory of SIRP to obtain a library of multivariate PDFs of correlated nonGaussian random vectors.
2. Derivation of the result that a SIRV is uniquely characterized through the knowledge of the PDF of a quadratic form.
3. Development of elegant and powerful simulation procedures for computer generation of SIRVs.
4. Reduction of the distribution identification of SIRVs from a multivariate problem to an equivalent univariate distribution identification problem.

As a result of these contributions, the problem of modeling, simulation, and distribution identification for SIRVs has resulted in tremendous computational simplicity. Consequently, the schemes developed here are suitable for use in several practical applications.

12.6.2. Suggestions for Future Research

It has been pointed out in this work that many of the attractive properties of Gaussian random processes also apply to SIRPs. Consequently, the use of SIRPs provides a convenient vehicle for solving several signal detection and estimation problems involving correlated nonGaussian processes. In particular, the following issues may be addressed as extensions of this work:

1. Use of experimental data to determine the applicability of the SIRP model for modeling clutter in radar, sonar, and image processing.
2. Application of the Kalman filter for SIRPs.
3. Use of SIRPs for radar ambiguity function analysis.
4. Application of SIRPs for canceling interference in digital communications.
5. Use of SIRPs in innovations based multichannel detection and estimation.
6. Use of SIRPs in linear predictive coding for speech processing.
7. Use of neural networks for identifying SIRPs.
8. Information theoretic considerations, such as, channel capacity and rate distortion theory, for SIRPs.
9. Use of SIRPs in parameter estimation involving the log likelihood function.

13 Weak Signal Detection

P. Chakravarthi

CONTENTS

13.1. Introduction 708
13.1.1. Weak Signal Problem 708
13.1.2. NonGaussian Correlated Data 710
13.1.3. Thesis Organization 711
13.2. The Locally Optimum Detector (LOD) 712
13.2.1. Literature Review 712
13.2.2. Spherically Invariant Random Processes (SIRP) 715
13.2.3. The Derivation of the Locally Optimum Detector 716
13.2.4. The Series Approach 717
13.2.4.1. The Known Signal Case 717
13.2.4.2. The Random Signal Case 720
13.2.5. The Lagrangian Approach 721
13.2.5.1. The Known Signal Case 721
13.2.5.2. The Random Signal Case 723
13.2.6. Special Cases 726
13.2.6.1. The Known Signal Problem 726
13.2.6.2. The Random Signal Problem 729
13.3. Determining Thresholds for the LOD 734
13.3.1. Literature Review 734
13.3.1.1. Classical Methods for Evaluating Thresholds 734
13.3.2. Extreme Value Theory 735
13.3.3. The Radar Problem 736
13.3.4. Methods for Estimating Thresholds 737
13.3.4.1. Estimates Based on Raw Data 737
13.3.4.2. Estimates Motivated by the Extreme Value Theory 738
13.3.5. The Generalized Pareto Distribution 739
13.3.5.1. Methods for Estimating the Parameters of the GPD 742
13.3.5.2. Estimation of Thresholds 748
13.3.6. Numerical Results 749

13.3.6.1. Characterization of Tail Shape for Known Distributions 749
13.3.6.2. Empirical Properties of the Estimators for Known Distributions 749
13.3.6.3. Effect of the Choice of λ on the Threshold Estimates 756
13.3.7. Examples 758
13.3.7.1. Known Distribution Case 758
13.3.7.2. An Unknown Distribution Case 759
13.4. Performance of the LOD for Multivariate Student-*T* and *K*-Distributed Disturbances 764
13.4.1. The Multivariate Student-*T* Distribution 764
13.4.1.1. The Locally Optimum Detector 766
13.4.1.2. Computer Simulation of Performance 768
13.4.1.3. Results of the Computer Simulation 771
13.4.2. The Multivariate *K*-Distribution 776
13.4.2.1. The Locally Optimum Detector 778
13.4.2.2. Computer Simulation of Performance 780
13.4.2.3. Conclusions 783
13.4.3. Determining LOD Threshold with Real Data 785
13.5. Performance of the Amplitude Dependent LOD 788
13.5.1. The Amplitude Dependent LOD for the Multivariate *K*-Distributed Disturbance 789
13.5.1.1. Results of Computer Simulation 790
13.5.2. The Amplitude Dependent LOD for the Student-*T* Distributed Disturbance 794
13.5.2.1. Conclusions 796
13.6. Conclusions 797
13.6.1. Summary 797
13.6.2. Suggestion for Future Research 798

13.1. INTRODUCTION

13.1.1. Weak Signal Problem

In radar applications it is found that the received target signal is contaminated with clutter and thermal noise. The received signal due to undesired reflections from land, sea, atmosphere etc. is called *clutter*. The thermal noise generated by the receiver hardware is typically modeled as a Gaussian random process. This kind of noise is always present. Depending upon the situation, the clutter may or may not be modeled as a Gaussian random process. Also, the power associated with the background clutter may be orders of magnitude larger than the receiver thermal noise or the desired signal power.

In modern radars, temporal and spatial processing are used to separate the target from the clutter. For example, the received signal from a target having a

radial velocity with respect to the radar will experience a Doppler shift. If the target spectrum appears in the tail of the clutter spectrum, then conventional frequency domain techniques can be used to extract the target from the clutter. Similarly, if the spatial spectrum of the target does not overlap that of the clutter, performance will be limited by the background noise rather than the clutter. In this research, use is also made of temporal and spatial processing. However, we are interested in the case where the target temporal and spatial spectra cannot be separated from the strong clutter. By definition, this is referred to as the weak signal detection problem. Given a range–Doppler–azimuth cell in which a target is to be detected, it is assumed that the signal is larger than the background noise but much smaller than the clutter.

In the weak signal problem the performance is limited by the clutter even after temporal and spatial processing. Therefore, it becomes very important to identify the clutter plus noise probability density function (PDF). This density function is the *N*th order joint density function of the received radar samples $r_1, r_2, \ldots, r_N$ in the absence of a target signal. The received waveform can be modeled as a random process. Since we will be sampling this process at N time instants, we need to know the *N*th order joint PDF of the N random variables. In this research effort the performance measures of radar receivers are analyzed, given the *N*th order PDF associated with the random process.

In the hypothesis testing problem, where we have to decide whether the target is present or absent, two kinds of errors can occur:

(1) A false alarm occurs when it is decided that the target is present even though it is not.
(2) A miss occurs when it is decided that the target is not present even though it is.

In many radar problems the chosen criterion is to fix the probability of false alarm at a certain value and then to maximize the probability of detection. In statistical decision theory the *Likelihood Ratio Test* (LRT) is optimum for these kinds of problems. The LRT evaluates the likelihood ratio as the ratio of the *N*th order joint PDF under the alternative hypothesis H_1 (signal present) to the *N*th order joint PDF under the null hypothesis H_0 (signal not present). This ratio is then compared to a certain threshold to make a decision. Under the constraint of a fixed false alarm probability, the Neyman–Pearson receiver obtained on the basis of the LRT is the optimum receiver.

Utilizing an underscore (_) to denote a vector quantity, the components of the received realization vector $\underline{r}$ of the random vector $\underline{R}$ can be written mathematically as

$$H_1 : r_i = s_i + d_i \tag{13.1}$$

$$H_0 : r_i = d_i \qquad i = 1, 2, \ldots, N \tag{13.2}$$

where s_i and d_i represent the ith sample of the desired signal return and the additive disturbance, respectively. Also, let $f_{\underline{R}}(\underline{r}|H_1), f_{\underline{R}}(\underline{r}|H_0), f_{\underline{D}}(\underline{d})$ denote the Nth order PDFs of $\underline{R}$ under H_1, of $\underline{R}$ under H_0 and of the disturbance, respectively. In general, the disturbance may be composed of clutter plus noise. Since it is not possible to separate the clutter and noise components of the disturbance when the disturbance is measured, we focus on the disturbance itself. As the signal becomes very weak (i.e., as the signal to clutter plus noise ratio [SCNR] approaches zero), the numerator and the denominator of the LRT tend to become identical. This is due to the fact that

$$f_{\underline{R}}(\underline{r}|H_1) \approx f_{\underline{R}}(\underline{r}|H_0) = f_{\underline{D}}(\underline{d}) \tag{13.3}$$

This will result in the likelihood ratio being approximately equal to unity independent of the received signal. Thus, if T_s denotes the likelihood ratio,

$$P_D = \int_{\eta}^{\infty} f_{T_s}(T_s|H_1)\mathrm{d}t_s \approx P_F = \int_{\eta}^{\infty} f_{T_s}(T_s|H_0)\mathrm{d}t_s \tag{13.4}$$

where P_D and P_F represent the detection and false alarm probabilities, respectively. Therefore, the LRT performs poorly in the limit as the signal strength tends to zero.

Even though the problem of weak signal detection in radar applications is of great interest, most of the literature by various researchers has been devoted to strong signals in a clutter plus noise background. Optimal and/or very good suboptimal schemes have been proposed to achieve the desired level of performance. Only a relatively small fraction of the literature is devoted to the design of practical schemes for the detection of weak signals. In this study we present a general theory for developing practical detector structures for weak signal problems. Also, computer simulations are used to evaluate performance when the disturbance can be approximated by the multivariate Student-T and K-distributions. In such problems the concept of the Locally Optimum Detector (LOD) is used to come up with the decision rule, which is also a ratio test. For a deterministic signal, a statistic is obtained by taking the ratio of the derivative with respect to the signal strength of the Nth order joint PDF under H_1 to the Nth order joint PDF under H_0. The limit of this ratio as the signal strength tends to zero is evaluated to obtain the test statistic for the decision rule. In the random signal case the test statistic is a ratio, in the limit as the signal strength tends to zero, of the second derivative with respect to the signal strength of the Nth order joint PDF under H_1 to the Nth order joint PDF under H_0. This approach is valid when it is known that the SCNR ratio is very small but the actual value of SCNR is unknown. Thus, the LOD turns out to be a Uniformly Most Powerful (UMP) test for the class of problems where the SCNR is in the neighborhood of zero.

13.1.2. NonGaussian Correlated Data

Previously, general analytic expressions for the various applicable Nth order joint nonGaussian PDFs that allow for correlation between the variables were

unavailable. As a result, researchers in the past assumed independence between the samples. By assuming independence between the samples, they were able to get the Nth order PDF as a product of the marginal density functions. If we carry out the locally optimum test using the Nth order density function based upon independence and evaluate its performance, it is found that an unreasonably large number of samples is needed for acceptable performance. This arises because independent samples imply a white spectrum. Consequently, space-time processing cannot be used to filter the target from the clutter. Based on the concept of Spherically Invariant Random Processes (SIRP), analytical expressions for some Nth order joint nonGaussian PDFs that allow for correlation between the variables are now available. Use of the multivariate expressions for the nonGaussian PDFs and the theory of LODs enables receiver structures for weak signal detection to be derived.

13.1.3. Thesis Organization

The literature review on weak signal detection and the derivation of the LOD are presented in Section 13.2. It is shown that the LOD determines whether a target is present or absent by comparing a statistic, computed from the data, to a set threshold. Both deterministic and random target signals are considered. The receiver structures are specialized to the case for which the clutter plus noise can be approximated by an SIRP.

Since the clutter is assumed to be nonGaussian, the LOD receiver structure turns out to be nonlinear. As a result, system performance must be determined by means of computer simulation. The threshold is conventionally determined through a Monte Carlo procedure. Unfortunately, the number of trials is inversely proportional to the false alarm probability, P_F. For example, when $P_F = 10^{-6}$, a minimum of ten million trials need to be generated. To avoid carrying out so many trials, a new technique, based on extreme value theory, is presented in Section 13.3. It is demonstrated that fairly accurate thresholds can be determined for false alarm probabilities as small as 10^{-7} with as few as 10,000 to 30,000 trials.

Assuming that the clutter plus noise can be approximated by either the multivariate Student-T or K-distributions, the LOD is developed in Section 13.4 for the weak signal detection problem. The system performance is evaluated by means of computer simulation for each distribution. It is shown that the performance improvement for the LOD is significant compared to the linear receiver when the clutter plus noise is approximated by the Student-T distribution. However, the performance improvement compared to the linear receiver is not quite as significant when the clutter plus noise is approximated by the K-distribution.

To enhance the performance of the LOD for weak signal detection when the clutter plus noise is approximated as multivariate K-distributed, a new technique called the amplitude dependent locally optimum detector (ALOD) is presented in Section 13.5. This test, however, is not a UMP test. Based on this test, it is

demonstrated that significant performance improvements can be obtained compared to the linear receiver, even when the clutter plus noise is multivariate *K*-distributed. Summary and conclusions are presented in Section 13.6.

13.2. THE LOCALLY OPTIMUM DETECTOR (LOD)

13.2.1. Literature Review

The concept of the LOD was first established by Neyman and Pearson in 1930.[1,2] Subsequently this was applied to statistical communication and signal processing by several researchers.

David Middleton's work[3,4] on the LOD is based on expanding the LRT in terms of a power series expansion and truncating the series to a first order approximation. In the limit as the signal tends to zero, the canonical structure of the LOD is established with very weak restrictions on the statistical properties of signal and noise. The analysis applies equally well to nonGaussian as well as Gaussian, and nonstationary as well as stationary processes, for stochastic as well as deterministic signals, continuous as well as discrete time signals, and for combinations of signal and noise that need not be additive. In fact, the general character of the results is independent of the particular nature of the signal and noise, although specific noise distributions determine the specific detector structures. Middleton shows that the LOD is a threshold detector with very strong optimality features in the limit of an infinitely large number of samples. However, in our research, we are interested in applications where the number of samples may not be too large.

David Middleton[5] has also extended the problem of threshold or weak signal detection to vector fields involving highly nonGaussian electromagnetic interference and signals that are both narrowband. The emphasis is on a canonical formulation. However, the performance measures presented are obtained on the basis of asymptotically locally optimum algorithms. In Ref. 6 Middleton has also analyzed the performance of the LODs in correlated interference. He points out that the correlation function involves only the second order statistics whereas higher order statistics should be considered for the nonGaussian case. Consequently, the correlated noise model leads to a suboptimal receiver except when the underlying noise distribution, is Gaussian. But when the sample to sample correlation is strong, the resulting algorithms and performance measures can provide noticeable improvements over models that employ independent noise sampling assumptions.

Other researchers in this area, such as, J. H. Miller and John Thomas[7] and Saleem Kassam,[8] have obtained performance of the LOD under the asymptotic condition of an infinitely large number of samples. These researchers have modeled the noise samples as independent, identically distributed random variables. This enables them to have a closed form expression for the *N*th order PDF of multivariate nonGaussian noise. Applying the LOD test they have arrived at the decision statistic. Using the central limit theorem, the test statistic is shown

to approach Gaussian in the limit of very large sample size. Then the performance measures are evaluated. Song and Kassam[9,10] have also derived the LODs for both known signals and random signals in a generalized observation model. In this model additive, multiplicative, and signal dependent noise models are considered. They show that the detectors derived under this model are interesting generalizations of the LODs derived in the additive noise model case. They also analyze the performance of the detector for a finite sample size case. But, the underlying noise distribution is assumed to be bivariate Gaussian.

For a variety of detection problems, Jack Capon[11] concludes that implementation of the LOD is either less or no more complicated than the Neyman–Pearson detector. First, he proposes the LOD for weak signal applications and proceeds to evaluate its performance by comparing it with the Neyman–Pearson detector. The comparison is based on the concept of Asymptotically Relative Efficiency (ARE). ARE is defined as the ratio of sample sizes required for two different detectors to achieve the same error probability and for the same signal to noise ratio, as the signal to noise ratio tends to zero and the sample sizes tend to infinity. On the basis of this comparison, it is shown that the LOD is asymptotically as efficient as the Neyman–Pearson detector. Conte, Izzo, Longo, and Paura[12] have also considered the problem of weak signal detection using LODs for arbitrarily large sample sizes to show that significant improvements are achieved compared to the linear detector. However, when they implement their algorithm for finite sample sizes, they conclude that the promises of the asymptotic theory cannot be achieved using even moderately large sample sizes. Hence, they propose a scheme which is a hybrid of the asymptotically optimum detector and the linear detector. Asymptotically effective nonparametric algorithms for detecting weak signals in nonGaussian interference have also been considered by Valeyev and Aspisov.[13] They conclude that the effectiveness of such algorithms approaches that of the optimum asymptotically, in the limit of large sample size. Raveendra and Srinivasan[14] have derived the locally optimum receiver structure for the coherent detection of continuous phase frequency shift keying (CPFSK) is nonGaussian noise channels. They evaluate the performance of the receiver consisting of a zero memory nonlinearity followed by a correlator for a number of noise models. The measure of performance is the asymptotic relative efficiency. However, they point out two important drawbacks in the analysis of performance through the ARE. The first one is, while large sample sizes are desirable for weak signal detection, increasing the sample size actually makes the LOD suboptimal partly due to the fact that there is an increased effect of higher order terms in the expansion of the likelihood ratio. Secondly, under increasing sampling rates, the assumption of independent samples, used in the derivation of the weak signal detector under nonGaussian conditions becomes invalid. In a Naval Underwater Systems Center report, Raymond Ingram and R. Houle[15] analyze the performance of the optimum and several suboptimum receivers for weak signal detection of known signals in additive, white, nonGaussian noise. He concludes that the implementation of the optimum or suboptimum nonlinear receivers yield significant improvements in performance relative to the receiver which is

optimum in Gaussian noise. However, the receiver structure is more complicated than the linear receiver.

The structure of LODs has also been characterized in terms of locally optimum estimators and correlators.[16] These characterizations are canonical structures involving estimators–correlators. It is shown that if the one step signal predictor is recursive and the noise is white (possibly nonGaussian or non-stationary), there is a canonical structure admitting recursive computation. The motivation to get a recursive structure is to simplify implementations and enhance adaptability. The problem of signal design has been considered for the case of locally optimum detection by Johnson and Orsak.[17] They show for the weak signal problem as we have defined it in Section 13.1 (Introduction), that the detection performance depends on signal energy in proportion to the Fisher information for location. In other words, when the spectra of the signal and disturbance overlap completely, significant improvements do not result from signal design. Another useful result they point out is that, among all distributions having zero mean and the same (finite) variance, the distribution having the smallest Fisher information is the Gaussian. Because of this result, it is concluded that detecting a small signal in Gaussian noise is the most difficult situation possible for an optimal detector. An increase in the signal energy yields the smallest possible performance improvement. Johnson and Orsak also come with explicit expressions to quantify the "small signal regimes" depending on the amplitude distribution of the noise.

Arthur Spaulding[18] compares the performance of the Locally Optimum Binary Detector (LOBD) with that of the linear receiver and a hard limiter. The performance analysis is done via specific examples and through computer Monte Carlo simulation. Under the assumption of independent, identically distributed samples, he concludes that the LOBD approaches its optimum performance only under the limit of large sample sizes and small signal to noise ratios. Also, he shows by way of an example that one cannot always be assured of obtaining great improvements over the linear receiver by using nonlinear processing. This implies that even if the underlying PDF of the interference has a much larger tail than that of the Gaussian, it does not guarantee a much improved performance over the linear processor. The improvement in performance depends on the particular nature of the underlying PDF. Michael Bouvet[19] obtains the LOD by expanding the likelihood ratio and truncating the expansion under the weak signal assumption. He expands the likelihood ratio in two ways: one with respect to the signal and the other with respect to the observation. He then establishes the equivalence between the two different forms of expansion. However, he points out the limitations of these expansions by stating that the results are valid only if the neglected terms are actually negligible with respect to the retained terms. Since the received observations are random, this cannot always be guaranteed.

Shishkov and Penev[20] have considered the case of correlated interference and background white noise, but have restricted themselves to multivariate Gaussian interference. For the known signal case, when the underlying noise distribution is

Gaussian, the optimal detector obtained from the LRT is the same as the weak signal detector obtained from the locally optimum test. Modestino and Ningo[21] were among the earliest researchers to consider weak signal detection arising from bandpass processes. They have modeled the received signal as statistically independent complex samples and then obtained the joint density function of the inphase and quadrature components. Under the assumption that the clutter density function is circularly symmetric, they transform the joint density function to an equivalent one involving the envelope and phase. This model still does not include the correlation from one envelope sample to another and hence large sample sizes are required for good performance. Martinez et al.[22] have considered the case where the noise has a multivariate Laplace distribution, where any nonnegative definite matrix can be used to model the correlation between the random variables. Based on this model they go ahead and derive the LOD which, as expected, is nonlinear. They compare the performance of this detector to the one developed by assuming the noise to be independent and identically distributed, and to the matched filter. They compare the performance of these receivers through the means of ARE and do not analyze the receiver performance for small sample sizes, which is the case of practical interest.

Sangston and Gerlach[23] have used the concept of the SIRP to model multivariate nonGaussian PDFs. They derive the LOD based on this noise model. It turns out that the LOD structure is the matched filter in conjunction with a nonlinearity. They are able to establish the canonical nature of this result for the class of joint density functions arising from SIRPs. They propose an equivalent structure for the LOD. This takes the form of a receiver where the matched filter output is compared to a nonlinear adaptive threshold. However, they do not explore the issue in terms of performance analysis.

13.2.2. Spherically Invariant Random Processes (SIRP)

In general, the radar receiver processes N complex samples (or $2N$ quadrature components) from each resolution cell. To develop an optimal receiver, it is necessary to have a closed form analytical expression for the joint PDF of the received samples. When the N samples are statistically independent, the joint PDF is simply the product of the marginal PDFs. However, clutter samples are likely to be correlated. Because this correlation is useful in canceling the clutter, it is important that the correlation be modeled. Unfortunately, when the received samples are correlated and nonGaussian, there are no unique analytical expressions for their joint PDFs. A search of the mathematical and signal processing literature reveals that the SIRP provides a powerful mechanism for obtaining the PDF of N correlated nonGaussian random variables. SIRPs for clutter modeling and simulation can be found in, Ref. 24 where Conte and Longo model complex clutter as a spherically invariant random process. They point out that the SIRP is ergodic only if the underlying clutter process is Gaussian. This means that time averages cannot be used to approximate ensemble averages. Conte et al.[25] also propose specific computer simulation procedures for

generating clutter realizations from an SIRP with desired correlation properties when the underlying distribution is multivariate Weibull or *K*-distributed. Rangaswamy et al.[26] have developed a library of multivariate correlated nonGaussian PDFs for characterizing various clutter scenarios through the theory of SIRPs. A significant result in this chapter is the proof that the multivariate SIRP PDF approximation problem can be reduced to an equivalent univariate PDF approximation problem. Rangaswamy et al.[27] develop two canonical computer simulation procedures for the generation of any correlated nonGaussian clutter that can be modeled as a spherically invariant random process.

Application of the theory of SIRPs to the problem of signal detection and estimation can be found in Refs. 23 and 28. Yao[28] derives the form of the unit threshold likelihood ratio receiver for the detection of a known deterministic signal in additive SIRP noise. He shows that the optimum receiver is the linear receiver or the matched filter when the threshold is set to unity, a threshold commonly arising in communication systems. This result is very significant because it tells us that nonlinear processing will not improve performance when the threshold is set to unity even though the disturbance is nonGaussian. However, when the threshold is not unity, then the optimal nonGaussian receiver is a nonlinear receiver. Pentini et al.[29] consider the problem of detecting a known target and a Swerling zero target embedded in coherent *K*-distributed clutter. The detectors are derived based on the LRT where the multivariate joint density functions used in the test are obtained from the theory of SIRPs. The receiver performance is then evaluated for the strong signal case. The false alarm probabilities used in obtaining the receiver performance are 10^{-3} and 10^{-4} so as to reduce the number of Monte Carlo trials needed to set thresholds. For a Swerling zero target model, they obtain a probability of detection equal to 0.1 for a false alarm probability of 10^{-3} using four complex samples and a signal to clutter ratio (SCR) of -10 dB. However, they do not explore performance for lower values of SCRs.

13.2.3. The Derivation of the Locally Optimum Detector

The usual criterion in radar problems is to maximize the probability of detection under a fixed false alarm probability constraint. This receiver is called the Neyman–Pearson receiver. The receiver implements the LRT and compares it against a threshold whose value is designed to give the desired false alarm probability. In particular, consider the received vector $\underline{R}^{\mathrm{T}} = [R_1, R_2, \ldots, R_N]$. Introduce the two hypotheses H_0 and H_1 as described below:

$$H_0 : r_i = c_i + n_i \tag{13.5}$$

$$H_1 : r_i = \theta s_i + c_i + n_i \qquad i = 1, 2, \ldots, N \tag{13.6}$$

Thus, H_0 is the hypothesis that the received signal consists solely of clutter plus noise while target signal is assumed to be present under the hypothesis H_1.

Let the joint PDF of $R_1, R_2, \ldots, R_N$ under hypothesis H_k ($k = 0, 1$) be denoted by $f_{\underline{R}}(\underline{r}|H_k)$. The Neyman–Pearson receiver performs the LRT

$$t_s(r) = \frac{f_{\underline{R}}(\underline{r}|H_1)}{f_{\underline{R}}(\underline{r}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{13.7}$$

where η is specified to satisfy the false alarm constraint

$$P_F = \int_{\eta}^{\infty} f_{T_s}(t_s|H_0)\mathrm{d}t_s \tag{13.8}$$

and $f_{T_s}(t_s|H_k)$ is the conditional PDF of the test statistic T_s given hypothesis H_k.

However, when the signal strength is very small relative to the clutter plus noise, the joint density function of the received random variables under H_1 approaches that under H_0. Then the numerator and the denominator of the LRT become approximately equal, leading to numerical difficulties in discriminating between the two hypotheses. The Neyman–Pearson test is of course optimum. However, the form of the LRT can be rearranged to yield a test statistic that is more sensitive to perturbations in the received data. This gives rise to the concept of the LOD. In this chapter the concept of the LOD is developed in detail using two approaches. The first approach is based on a power series expansion of the LRT and the second approach derives the LOD by an optimization using the principle of Lagrangian multipliers. It is shown that both approaches yield identical detector structures, though starting from different theoretical points of view. As the signal strength becomes weaker, the LOD becomes optimum even though its performance for a fixed sample size may not be as good as desired.

13.2.4. The Series Approach

13.2.4.1. The Known Signal Case

Let the additive clutter component $\underline{C} = [C_1, C_2, \ldots, C_N]^T$ be stationary and independent of the stationary white Gaussian background noise $\underline{N} = [N_1, N_2, \ldots, N_N]^T$. The noise variance σ_n^2 is assumed to be several orders of magnitude below the clutter variance σ_c^2 which is taken to be unity without loss of generality. The signal is assumed to be of the form $\theta\underline{s}$, where $\underline{s}$ is known. The components of $\underline{s}$ are chosen to have $|s_i|^2 = 1$ so that the positive parameter θ is a measure of the SCR defined by

$$\mathrm{SCR} = \frac{\theta^2|s_i|^2}{\sigma_c^2} = \theta^2 \tag{13.9}$$

Because the clutter and noise are statistically independent with the noise assumed to have zero mean, the covariance matrix of the disturbance vector

$\underline{D} = \underline{C} + \underline{N}$, denoted by $\mathbf{M}_D$, is equal to the covariance matrix $\mathbf{M}_C$ of the clutter plus the covariance matrix $\mathbf{M}_N$ of the noise. Since the noise is white and stationary, the covariance matrix of the noise is of the form $\mathbf{M}_N = \sigma_n^2\mathbf{I}$, where $\mathbf{I}$ is the identity matrix. When the clutter is highly correlated, the covariance matrix $\mathbf{M}_C$ tends to be ill-conditioned. However, $\mathbf{M}_D$ will not be ill-conditioned because, by adding the small value σ_n^2 to the diagonal elements of $\mathbf{M}_C$, the smallest eigenvalue of $\mathbf{M}_D$ is guaranteed to be no smaller than σ_n^2. Also, addition of $\mathbf{M}_N$ to $\mathbf{M}_C$ ensures that the disturbance spectrum will limit performance even in those frequency intervals where the clutter spectrum is negligible.

With this assumption the LRT takes the form

$$t_s = \frac{f_{\underline{R}}(\underline{r}|H_1)}{f_{\underline{R}}(\underline{r}|H_0)} = \frac{f_{\underline{D}}(\underline{r} - \theta\underline{s})}{f_{\underline{D}}(\underline{r})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{13.10}$$

As mentioned previously, when $\theta \ll 1$, the signal $\theta\underline{s}$ represents a small perturbation in the received vector under hypothesis H_1. Hence, $f_{\underline{R}}(\underline{r}|H_1)$ approximately equals $f_{\underline{R}}(\underline{r}|H_0)$. As a result, T_s is relatively insensitive to $\theta\underline{s}$. One approach at deriving a weak signal detector is to expand the numerator of the LRT in a Taylor series.

For this purpose, let $\underline{y} = \underline{r} - \theta\underline{s}$. Then

$$f_{\underline{R}}(\underline{r}|H_1) = f_{\underline{D}}(\underline{y}) \tag{13.11}$$

Expanding $f_D(\underline{y})$ in a Taylor series about the received vector $\underline{r}$, we obtain

$$\begin{aligned} f_{\underline{D}}(\underline{y}) = f_{\underline{D}}(\underline{r}) &+ \sum_{k_1=1}^{N} (y_{k_1} - r_{k_1}) \frac{\partial f_{\underline{D}}(\underline{y})}{\partial y_{k_1}}\bigg|_{\underline{y}=\underline{r}} \\ &+ \frac{1}{2!}\sum_{k_1=1}^{N}\sum_{k_2=1}^{N} (y_{k_1} - r_k)(y_{k_2} - r_{k_2}) \frac{\partial^2 f_{\underline{D}}(\underline{y})}{\partial y_{k_1}\partial y_{k_2}}\bigg|_{\underline{y}=\underline{r}} + \cdots \\ &+ \frac{1}{n!}\sum_{k_1=1}^{N}\sum_{k_2=1}^{N}\cdots\sum_{k_n=1}^{N} (y_{k_1} - r_{k_1})(y_{k_2} - r_{k_2})\ldots(y_{k_n} - r_{k_n}) \frac{\partial^n f_{\underline{D}}(\underline{y})}{\partial y_{k_1}\partial y_{k_2}\ldots\partial y_{k_n}}\bigg|_{\underline{y}=\underline{r}} \\ &+ \cdots \end{aligned} \tag{13.12}$$

This can be expressed in vector form by introducing the operator

$$(\underline{y} - \underline{r})^T\nabla_y = \sum_{k=1}^{N} (y_k - r_k)\frac{\partial}{\partial y_k} \tag{13.13}$$

where the subscript y on ∇ indicates partial differentiation with respect to the

components of $\underline{y}$. The expansion of $f_{\underline{D}}(\underline{y})$ about the point $\underline{y} = \underline{r}$ then becomes

$$\begin{aligned} f_{\underline{D}}(\underline{y}) &= f_{\underline{D}}(\underline{r}) + [(\underline{y} - \underline{r})^{\mathrm{T}} \nabla_y] f_{\underline{D}}(\underline{y})|_{\underline{y}=\underline{r}} \\ &\quad + \frac{1}{2!}[(\underline{y} - \underline{r})^{\mathrm{T}} \nabla_y]^2 f_{\underline{D}}(\underline{y})|_{\underline{y}=\underline{r}} + \cdots \\ &\quad + \frac{1}{n!}[(\underline{y} - \underline{r})^{\mathrm{T}} \nabla_y]^n f_{\underline{D}}(\underline{y})|_{\underline{y}=\underline{r}} + \cdots \\ &= f_{\underline{D}}(\underline{r}) + \sum_{n=1}^{\infty} \frac{1}{n!}[(\underline{y} - \underline{r})^{\mathrm{T}} \nabla_y]^n f_{\underline{D}}(\underline{y})|_{\underline{y}=\underline{r}} \end{aligned} \tag{13.14}$$

Recall that $\underline{y} = \underline{r} - \theta \underline{s}$, where θ and $\underline{s}$ are constants. Note that $\underline{y} - \underline{r} = -\theta \underline{s}$ and $\partial / \partial_{yk} = \partial / \partial_{rk}$. Then

$$(\underline{y} - \underline{r})^{\mathrm{T}} \nabla_y = \sum_{k=1}^{N} (-\theta s_k) \frac{\partial}{\partial r_k} = -\theta \underline{s}^{\mathrm{T}} \nabla_r \tag{13.15}$$

where the subscript r on ∇ indicates partial differentiation with respect to the components of $\underline{r}$. It follows that the expansion may be written as

$$f_{\underline{D}}(\underline{r} - \theta \underline{s}) = f_{\underline{D}}(\underline{r}) + \sum_{n=1}^{\infty} \frac{(-1^n)}{n!} \theta^n [\underline{s}^{\mathrm{T}} \nabla_r]^n f_{\underline{D}}(\underline{r}) \tag{13.16}$$

In order for the above expansion to be meaningful, it is necessary that all the derivatives in the above expansion exist.

Thus, using the above expansion of $f_{\underline{D}}(\underline{r} - \theta \underline{s})$, the Taylor series expansion of the likelihood ratio about the received vector $\underline{r}$ in Equation 13.10 can be written as

$$T_s(\underline{r}) = 1 + \left[\sum_{n=1}^{\infty} \frac{(-1)^n \theta^n}{n! f_{\underline{D}}(\underline{r})} (s^{\mathrm{T}} \nabla_r)^n \right] f_{\underline{D}}(\underline{r}) \tag{13.17}$$

The first term, being a constant, can be combined with the threshold without loss of optimality. The LOD is obtained by retaining only the term corresponding to $n = 1$ in the infinite summation. For $\theta \ll 1$, it is assumed that the remaining terms in the summation are negligible. On the other hand, because $\underline{r}$ is governed by a random vector and the partial derivatives of the PDF evaluated at $\underline{r}$ may be large, the remaining terms may actually not be negligible. However, it is assumed that this occurs with small probability. The resulting detector structure can then be expressed as

$$T_{\mathrm{LOD}}(\underline{r}) = -\frac{(s^{\mathrm{T}} \nabla_r) f_{\underline{D}}(\underline{r})}{f_{\underline{D}}(\underline{r})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_k \tag{13.18}$$

where η_k is chosen so as to achieve the desired false alarm probability.

13.2.4.2. The Random Signal Case

When the signal is random, $f_{\underline{R}}(\underline{r}|H_1)$ is obtained by integrating the joint density function $f_{\underline{R},\underline{S}}(\underline{r},\underline{s}|H_1)$ over all possible values of $\underline{s}$. Hence,

$$f_{\underline{R}}(\underline{r}|H_1) = \int_{-\infty}^{\infty} f_{\underline{R},\underline{S}}(\underline{r},\underline{s}|H_1)\underline{\mathrm{d}s} = \int_{-\infty}^{\infty} f_{\underline{R}|\underline{S}=\underline{s}}(\underline{r}|\underline{s},|H_1) f_{\underline{S}}(\underline{s})\underline{\mathrm{d}s}$$

$$= E_{\mathrm{s}}[f_{\underline{R}|\underline{S}=\underline{s}}(\underline{r}|\underline{s},H_1)] \tag{13.19}$$

where E_{s} denotes the expectation operation carried out with respect to the random vector $\underline{S}$. Because the denominator of T_{s} in Equation 13.10 is independent of $\underline{s}$, the Taylor series expansion of the likelihood ratio can now be written as

$$T_{\mathrm{s}}(\underline{r}) = 1 + \left[\sum_{n=1}^{\infty} \frac{(-1)^n \theta^n}{n! f_{\underline{D}}(\underline{r})} E_{\mathrm{s}}[(\underline{s}^{\mathrm{T}}\nabla_r)^n] f_{\underline{D}}(\underline{r})\right] \tag{13.20}$$

Once again, as in the known signal case, the unity term appearing in the test statistic can be put into the threshold. If we make the assumption that the expected value of the signal vector is $\underline{0}$, then the $n = 1$ term in the infinite series of Equation 13.20 goes to zero. Thus, for the random signal case, where the signal vector has zero mean, the LOD is defined to be the second term ($n = 2$) in the infinite series. As in the deterministic signal case, θ is assumed to be small enough such that the remaining terms of the series are negligible with high probability. Consequently, the LOD for the random signal case is given by

$$T_{\mathrm{s}2}(\underline{r}) = \frac{\theta^2}{2 f_{\underline{D}}(\underline{r})} E_{\mathrm{s}}[(\underline{s}^{\mathrm{T}}\nabla_r)^2]\ f_{\underline{D}}(\underline{r}) \underset{H_0}{\overset{H_1}{\gtrless}} \eta' \tag{13.21}$$

where $T_{\mathrm{s}2}$ represents the second order term in the Taylor series expansion of T_{s}. The above equation can be rewritten as

$$T_{\mathrm{s}2}(\underline{r}) = \frac{\theta^2}{2 f_{\underline{D}}(\underline{r})} E_{\mathrm{s}}[\nabla_r^{\mathrm{T}} \underline{s}\underline{s}^{\mathrm{T}} \nabla_r] f_{\underline{D}}(\underline{r}) \underset{H_0}{\overset{H_1}{\gtrless}} \eta' \tag{13.22}$$

where, as before, η' is chosen to achieve the specified false alarm probability. Lumping the constant $\theta^2/2$ with the threshold and recognizing that

$$E_{\mathrm{s}}[(\underline{s}^{\mathrm{T}}\nabla_r)^2] = E_{\mathrm{s}}[\nabla_r^{\mathrm{T}} \underline{s}\underline{s}^{\mathrm{T}} \nabla_r] = \nabla_r^{\mathrm{T}} P \nabla_r \tag{13.23}$$

where **P** is the covariance matrix of the signal vector, the detector structure for the locally optimal test becomes

$$T_{\mathrm{LOD}}(\underline{r}) = \frac{\nabla_r^{\mathrm{T}} \mathbf{P} \nabla_r [f_{\underline{\mathrm{D}}}(\underline{r})]}{f_{\underline{\mathrm{D}}}(\underline{r})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{\mathrm{u}} \tag{13.24}$$

13.2.5. The Lagrangian Approach

Consider again the hypotheses testing problem defined in Equation 13.5 and Equation 13.6. Let us define a nonrandomized decision rule $\phi(\underline{r})$ such that

$$\phi(\underline{r}) = \begin{cases} 1; & H_1 \text{ true (target present)} \\ 0; & H_0 \text{ true (target absent)} \end{cases} \tag{13.25}$$

This amounts to partioning the decision space into two regions, S_1 and S_0. A target is declared if the vector $\underline{r}$ is present in the region S_1. If it falls in the region S_0, then the decision is made that the target is absent. The probability of detection equals the probability that the nonrandomized decision rule equals unity, given that hypothesis H_1 is indeed true. This probability will, in general, be a function of θ, the signal-to-clutter ratio. Denoting $\beta(\theta)$ as the probability of detection we have

$$P_{\mathrm{D}} = \beta(\theta) = p[\phi(\underline{r}) = 1 | H_1] = \int_{-\infty}^{\infty} \phi(\underline{r}) f_{\underline{\mathrm{R}}}(\underline{r} | H_1) \underline{\mathrm{d}r} \tag{13.26}$$

$\beta(\theta)$ is also called the power function of the test. The false alarm probability is given by

$$P_{\mathrm{F}} = p[\phi(\underline{r}) = 1 | H_0] = \int_{-\infty}^{\infty} \phi(\underline{r}) f_{\underline{\mathrm{R}}}(\underline{r} | H_0) \underline{\mathrm{d}r} = P_{\mathrm{F}} \tag{13.27}$$

The optimization problem to be discussed in the next section imposes the constraint that the false alarm probability be equal to P_{F}. P_{F} is also defined to be the significance level of the test.

13.2.5.1. The Known Signal Case

As discussed earlier, in the limit as the signal strength tends to zero, the probability of detection becomes approximately equal to the probability of false alarm. Therefore, instead of maximizing the probability of detection, one approach is to maximize the slope of the power function $\beta(\theta)$ curve at the point $\theta = 0$. The function to be maximized and the constraint are given in the following

two equations. Maximize

$$\left.\frac{\partial \beta(\theta)}{\partial \theta^{-}}\right|_{\theta=0} = \left[\frac{\partial}{\partial \theta}\int_{-\infty}^{\infty} \phi(\underline{r}) f_{\underline{\mathrm{R}}}(\underline{r}|H_1)\underline{\mathrm{d}r}\right]_{\theta=0} \tag{13.28}$$

subject to the constraint

$$\int_{-\infty}^{\infty} \phi(\underline{r}) f_{\underline{\mathrm{R}}}(\underline{r}|H_0)\underline{\mathrm{d}r} = \alpha \tag{13.29}$$

We also require that the test be UMP in the sense that $\phi(\underline{r})$ be independent of θ for small neighborhoods in the vicinity of $\theta = 0$. Notice that there is a derivative with respect to θ outside the integral in Equation 13.28. If the function $f_{\underline{\mathrm{R}}}(\underline{r}|H_1)$ is a well behaved function of θ such that its derivative exists at all points, the derivative can be moved inside the integral resulting in

$$\begin{aligned}&\frac{\partial}{\partial \theta}\int_{-\infty}^{\infty} \phi(\underline{r}) f_{\underline{\mathrm{R}}}(\underline{r}|H_1)\underline{\mathrm{d}r} \\ &\quad = \int_{-\infty}^{\infty} \frac{\partial \phi(\underline{r})}{\partial \theta} f_{\underline{\mathrm{R}}}(\underline{r}|H_1)\underline{\mathrm{d}r} + \int_{-\infty}^{\infty} \phi(\underline{r}) \frac{\partial f_{\underline{\mathrm{R}}}(\underline{r}|H_1)}{\partial \theta}\underline{\mathrm{d}r}\end{aligned} \tag{13.30}$$

Because of the UMP requirement, $\partial \phi(\underline{r})/\partial \theta = 0$ and the first integral on the right side of Equation 13.30 integrates to zero. It follows that

$$\frac{\partial}{\partial \theta}\int_{-\infty}^{\infty} \phi(\underline{r}) f_{\underline{\mathrm{R}}}(\underline{r}|H_1)\underline{\mathrm{d}r} = \int_{-\infty}^{\infty} \phi(\underline{r}) \frac{\partial f_{\underline{\mathrm{R}}}(\underline{r}|H_1)}{\partial \theta}\underline{\mathrm{d}r} \tag{13.31}$$

Given the function $(\partial \beta(\theta)/\partial \theta)|_{\theta=0}$ to be maximized along with the false alarm probability constraint, the functional form of the maximization problem using the Lagrange multiplier approach is

$$\max\left[\int_{-\infty}^{\infty} \phi(\underline{r}) \frac{\partial f_{\underline{\mathrm{R}}}(\underline{r}|H_1)}{\partial \theta}\underline{\mathrm{d}r}\bigg|_{\theta=0} + \eta_k\left(\alpha - \int_{-\infty}^{\infty} \phi(\underline{r}) f_{\underline{\mathrm{R}}}(\underline{r}|H_0)\underline{\mathrm{d}r}\right)\right] \tag{13.32}$$

where η_k is the Lagrange multiplier. Expression (13.32) can be rewritten as

$$\max\left[\int_{-\infty}^{\infty} \phi(\underline{r}) \frac{\partial f_{\underline{\mathrm{R}}}(\underline{r}|H_1)}{\partial \theta} - \eta_k f_{\underline{\mathrm{R}}}(\underline{r}|H_0)\underline{\mathrm{d}r}\bigg|_{\theta=0}\right] + \eta_k \alpha \tag{13.33}$$

To maximize the above integral, the decision regions should be chosen such that the integrand is always positive. In other words, the decision regions are chosen such that

$$\left.\frac{\partial f_{\underline{\mathrm{R}}}(\underline{r}|H_1)}{\partial \theta}\right|_{\theta=0} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_k f_{\underline{\mathrm{R}}}(\underline{r}|H_0) \tag{13.34}$$

As was pointed out in the previous section, $f_R(\underline{r}|H_1)$ is identical to $f_D(\underline{r} - \theta\underline{s})$. Therefore, the decision rule becomes

$$\left.\frac{\partial f_D(\underline{r} - \theta\underline{s})}{\partial\theta}\right|_{\theta=0} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_k f_D(\underline{r}) \tag{13.35}$$

The LOD is defined to be the detector implementing the ratio test

$$\frac{\left.\dfrac{\partial f_D(\underline{r} - \theta\underline{s})}{\partial\theta}\right|_{\theta=0}}{f_D(\underline{r})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_k \tag{13.36}$$

The Lagrange multiplier η_k is chosen to satisfy the false alarm constraint. Note that

$$f_D(\underline{r} - \theta\underline{s}) = f_D(r_1 - \theta s_1, r_2 - \theta s_2, \ldots, r_N - \theta s_N) \tag{13.37}$$

As a result,

$$\begin{aligned}\frac{\partial f_D(\underline{r} - \theta\underline{s})}{\partial\theta} &= \frac{\partial f_D(\underline{r} - \theta\underline{s})}{\partial(r_1 - \theta s_1)}\frac{\partial(r_1 - \theta s_1)}{\partial\theta} + \frac{\partial f_D(\underline{r} - \theta\underline{s})}{\partial(r_2 - \theta s_2)}\frac{\partial(r_2 - \theta s_2)}{\partial\theta} \\ &\quad + \cdots + \frac{\partial f_D(\underline{r} - \theta\underline{s})}{\partial(r_N - \theta s_N)}\frac{\partial(r_N - \theta s_N)}{\partial\theta} \\ &= \sum_{k=1}^{N}\frac{\partial f_D(\underline{r} - \theta\underline{s})}{\partial(r_k - \theta s_k)}(-s_k)\end{aligned} \tag{13.38}$$

Consequently,

$$\left.\frac{\partial f_D(\underline{r} - \theta\underline{s})}{\partial\theta}\right|_{\theta=0} = -\sum_{k=1}^{N}\frac{\partial f_D(\underline{r})}{\partial r_k}s_k = -(\underline{s}^T\nabla_r)f_D(\underline{r}) \tag{13.39}$$

Thus, the LOD can also be written as

$$T_{LOD}(\underline{r}) = -\frac{(\underline{s}^T\nabla_r)f_D(\underline{r})}{f_D(\underline{r})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_k \tag{13.40}$$

It can be seen that this detector is identical to the one in Equation 13.18 obtained through the series approach.

13.2.5.2. The Random Signal Case

Consider a random signal S and let its joint PDF be denoted by $f_{\underline{S}}(\underline{s})$. Also, without loss of generality, we can make the assumption that the signal vector has zero mean and each component of the vector has unit variance. Given the

signal vector $\underline{S}$ the joint density function on the received vector under hypothesis H_1 is

$$f_{\underline{R}}(\underline{r}|\underline{s}, H_1) = f_{\underline{D}}(\underline{r} - \theta\underline{s}) \tag{13.41}$$

The power function for the locally optimum test was given in the previous section in Equation 13.26. However, in the random signal case the unconditional density function $f_R(\underline{r}|H_1)$ is obtained by integrating out the random vector $\underline{S}$ from the joint PDF $f_{RS}(\underline{r}, \underline{s}|H_1) = f_R(\underline{r}, |\underline{s}, H_1) f_S(\underline{s})$. Use of Equation 13.41 results in

$$\beta(\theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(\underline{r}) f_{\underline{D}}(\underline{r} - \theta\underline{s}) f_{\underline{S}}(\underline{s}) \mathrm{d}\underline{r}\, \mathrm{d}\underline{s} \tag{13.42}$$

The false alarm constraint is once again given by

$$\int_{-\infty}^{\infty} \phi(\underline{r}) f_{\underline{R}}(\underline{r}|H_0) \mathrm{d}\underline{r} = \alpha \tag{13.43}$$

As before, we wish to maximize $(\partial\beta(\theta)/\partial\theta)|_{\theta=0}$. If the function $f_D(\underline{r} - \theta\underline{s})$ is a well behaved function such that its derivative exists at all points, then

$$\frac{\partial\beta(\theta)}{\partial\theta} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(\underline{r}) \frac{\partial f_{\underline{D}}(\underline{r} - \theta\underline{s})}{\partial\theta} f_{\underline{S}}(\underline{s}) \mathrm{d}\underline{r}\, \mathrm{d}\underline{s} \tag{13.44}$$

It follows from Equation 13.39 that

$$\left.\frac{\partial\beta(\theta)}{\partial\theta}\right|_{\theta=0} = -\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(\underline{r}) \left[\sum_{k=1}^{N} \frac{\partial f_{\underline{D}}(\underline{r})}{\partial r_k} s_k\right] f_{\underline{S}}(\underline{s}) \mathrm{d}\underline{r}\, \mathrm{d}\underline{s} \tag{13.45}$$

Because of the zero mean assumption

$$\int_{-\infty}^{\infty} s_k f_{\underline{S}}(\underline{s}) \mathrm{d}\underline{s} = 0 \tag{13.46}$$

We conclude that

$$\left.\frac{\partial\beta(\theta)}{\partial\theta}\right|_{\theta=0} = 0 \tag{13.47}$$

independent of the choice of $\phi(\underline{r})$. Therefore, to achieve the maximum increase of the power function in the vicinity of the origin, we maximize $(\partial^2\beta(\theta)/\partial\theta^2)|_{\theta=0}$. As before, assuming that the integration and differentiation can be interchanged,

$$\frac{\partial^2\beta(\theta)}{\partial\theta^2} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \phi(\underline{r}) \frac{\partial^2 f_{\underline{D}}(\underline{r} - \theta\underline{s})}{\partial\theta^2} f_{\underline{S}}(\underline{s}) \mathrm{d}\underline{r}\, \mathrm{d}\underline{s} \tag{13.48}$$

However from Equation 13.38

$$\begin{aligned}\frac{\partial^2 f_{\underline{D}}(\underline{r}-\theta\underline{s})}{\partial\theta^2} &= \frac{\partial}{\partial\theta}\sum_{k=1}^{N}\frac{\partial f_{\underline{D}}(\underline{r}-\theta\underline{s})}{\partial(r_k-\theta s_k)}(-s_k)\\ &= \sum_{j=1}^{N}\sum_{k=1}^{N}\frac{\partial^2 f_{\underline{D}}(\underline{r}-\theta\underline{s})}{\partial(r_j-\theta s_j)\partial(r_k-\theta s_k)}\frac{\partial(r_j-\theta s_j)}{\partial\theta}(-s_k)\\ &= \sum_{j=1}^{N}\sum_{k=1}^{N}\frac{\partial^2 f_{\underline{D}}(\underline{r}-\theta\underline{s})}{\partial(r_j-\theta s_j)\partial(r_k-\theta s_k)}s_j s_k \end{aligned} \tag{13.49}$$

Hence,

$$\left.\frac{\partial^2 f_{\underline{D}}(\underline{r}-\theta\underline{s})}{\partial\theta^2}\right|_{\theta=0} = \sum_{j=1}^{N}\sum_{k=1}^{N}\frac{\partial^2 f_{\underline{D}}(\underline{r})}{\partial r_j\partial r_k}s_j s_k = (\nabla_r^{\mathrm{T}}\underline{s}\,\underline{s}^{\mathrm{T}}\nabla_r)f_{\underline{D}}(r) \tag{13.50}$$

Then the second derivative of the power function at the origin takes the form

$$\begin{aligned}\left.\frac{\partial^2\beta(\theta)}{\partial\theta^2}\right|_{\theta=0} &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\phi(\underline{r})(\nabla_r^{\mathrm{T}}\underline{s}\,\underline{s}^{\mathrm{T}}\nabla_r)f_{\underline{D}}(\underline{r})f_{\underline{S}}(\underline{s})\mathrm{d}\underline{r}\,\mathrm{d}\underline{s}\\ &= \int_{-\infty}^{\infty}\phi(\underline{r})E_{\mathrm{s}}(\nabla_r^{\mathrm{T}}\underline{s}\,\underline{s}^{\mathrm{T}}\nabla_r)f_{\underline{D}}(\underline{r})\mathrm{d}\underline{r}\end{aligned} \tag{13.51}$$

Using the approach of Lagrange multipliers to maximize the function in Equation 13.51 along with the constraint Equation 13.43, the optimization problem can be written as

$$\max\left[\int_{-\infty}^{\infty}\phi(\underline{r})E_{\mathrm{s}}[\nabla_r^{\mathrm{T}}\underline{s}\,\underline{s}^{\mathrm{T}}\nabla_r]f_{\underline{D}}(\underline{r})\mathrm{d}\underline{r} + \eta_{\mathrm{u}}\left(\alpha - \int_{-\infty}^{\infty}\phi(\underline{r})f_{\underline{D}}(\underline{r})\mathrm{d}\underline{r}\right)\right] \tag{13.52}$$

The above expression can be rewritten as

$$\max\left[\int_{-\infty}^{\infty}\phi(\underline{r})[E_{\mathrm{s}}[\nabla_r^{\mathrm{T}}\underline{s}\,\underline{s}^{\mathrm{T}}\nabla_r]f_{\underline{D}}(\underline{r}) - \eta_{\mathrm{u}}f_{\underline{D}}(\underline{r})]\mathrm{d}\underline{r}\right] + \eta_{\mathrm{u}}\alpha \tag{13.53}$$

To maximize the integral the decision regions have to be chosen such that the integrand is always nonnegative. The resulting decision regions yield the inequalities

$$E_{\mathrm{s}}[\nabla_r^{\mathrm{T}}\underline{s}\,\underline{s}^{\mathrm{T}}\nabla_r]f_{\underline{D}}(\underline{r}) \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{\mathrm{u}}f_{\underline{D}}(\underline{r}) \tag{13.54}$$

If the covariance matrix of the signal vector is denoted by **P**, then the LOD can be written as

$$T_{\mathrm{LOD}}(\underline{r}) = \frac{(\nabla_r^{\mathrm{T}} \mathbf{P} \nabla_r) f_{\underline{\mathrm{D}}}(\underline{r})}{f_{\underline{\mathrm{D}}}(\underline{r})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{\mathrm{u}} \tag{13.55}$$

which is identical to Equation 13.24.

As a general rule for deriving locally optimum tests, note that we maximize at the origin the first nonvanishing derivative of the power function. For the known and the purely random signal cases the first nonvanishing derivative is the first and the second derivative, respectively.

13.2.6. Special Cases

In this section, LOD structures will be derived for three special cases. In the first it is assumed that the N random variables in the disturbance vector $\underline{D}$ are statistically independent. With this assumption, the joint PDF of the N random variables is the product of the marginal density functions of the individual random variables. In the second, the N random variables are modeled as arising from an SIRP. This model enables us to write the joint PDF of the random variables analytically, accounting for the correlation between the random variables. In the last case, the N correlated random variables are assumed to be jointly Gaussian. The LOD structures are derived for all cases. It turns out in all three cases that the detector can be expressed in a canonical form. This canonical expression is derived for both the known and the random signal problems.

13.2.6.1. The Known Signal Problem

13.2.6.1.1. Disturbance Modeled as Independent Random Variables

From Equation 13.36, the LOD structure in the known signal case is given as

$$\frac{\left.\dfrac{\partial f_{\underline{\mathrm{D}}}(\underline{r} - \theta \underline{s})}{\partial \theta}\right|_{\theta=0}}{f_{\underline{\mathrm{D}}}(\underline{r})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_k \tag{13.56}$$

Let the N random variables in the vector $\underline{D}$ be independent where the PDF of the ith random variable is $f_{\mathrm{D}_i}(d_i)$. Therefore, the conditional joint density functions of the N received random variables are given by

$$f_{\mathrm{R}_1,\mathrm{R}_2,\ldots,\mathrm{R}_N}(r_1, r_2, \ldots, r_N | H_0) = \prod_{i=1}^{N} f_{\mathrm{D}_i}(r_i) \tag{13.57}$$

$$f_{\mathrm{R}_1,\mathrm{R}_2,\ldots,\mathrm{R}_N}(r_1, r_2, \ldots, r_N | H_1) = \prod_{i=1}^{N} f_{\mathrm{D}_i}(r_i - \theta s_i) \tag{13.58}$$

The numerator in the ratio test of Equation 13.56 is evaluated as

$$\left.\frac{\partial f_{\underline{D}}(\underline{r}-\theta\underline{s})}{\partial\theta}\right|_{\theta=0} = \frac{\partial}{\partial\theta}\left[\prod_{i=1}^{N} f_{D_i}(r_i-\theta s_i)\right]\Bigg|_{\theta=0}$$
$$= \sum_{i=1}^{N}\left\{(-s_i)\frac{df_{D_i}(r_i)}{dr_i}\prod_{\substack{j=1\\ j\neq i}}^{N} f_{D_i}(r_j)\right\} \tag{13.59}$$

Thus, from Equation 13.56 the LOD statistic for independent random variables is given by

$$T_{\mathrm{LOD}}(r_1, r_2, \ldots, r_N) = -\sum_{i=1}^{N} s_i \frac{f'_{D_i}(r_i)}{f_{D_i}(r_i)} \tag{13.60}$$

where $f'_{D_i}(r_i)$ denotes the derivative of $f_{D_i}(r_i)$ with respect to r_i. The above equation for the LOD statistic is the canonical form obtained when the random variables are independent. For different density functions, $f_{D_i}(r_i)$, the detector will be different, although its structure remains the same. The canonical form of the detector is shown in Figure 13.1.

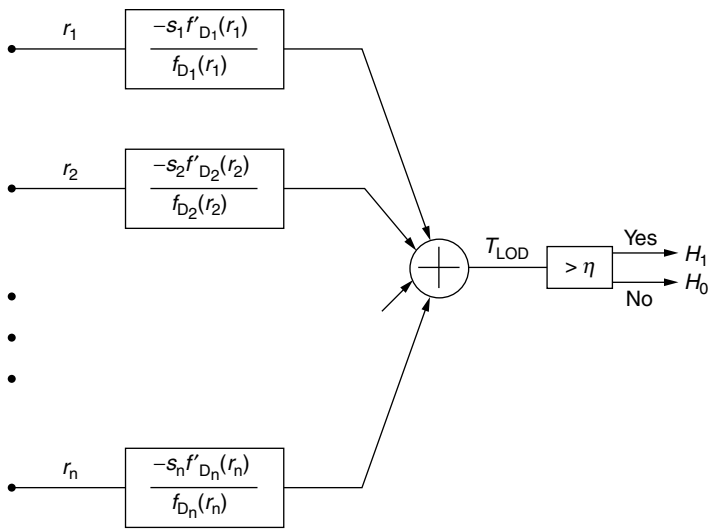

FIGURE 13.1 Canonical form of LOD assuming known signal and independent random variables.

13.2.6.1.2. Disturbance Modeled as an SIRV

When the random variables of the disturbance are drawn from a zero mean SIRP distribution, the joint PDF can be written as

$$f_{\underline{D}}(\underline{d}) = \frac{1}{2\pi^{N/2}|\mathbf{M}|^{1/2}} h_N(p) \tag{13.61}$$

where $p = \underline{d}^{\mathrm{T}}\mathbf{M}^{-1}\underline{d}$, $\mathbf{M}$ is the covariance matrix for the N random variables and $h_N(p)$ is a positive valued, nonlinear function of p. The numerator of the ratio test in Equation 13.56 is then given by

$$\left.\frac{\partial f_{\underline{D}}(\underline{r}-\theta\underline{s})}{\partial\theta}\right|_{\theta=0} = \left.\frac{\partial}{\partial\theta}\left\{\frac{1}{2\pi^{N/2}|\mathbf{M}|^{1/2}} h_N(p)\right\}\right|_{\theta=0} = \frac{1}{2\pi^{N/2}|\mathbf{M}|^{1/2}} \left.\frac{\partial}{\partial\theta}\{h_N(p)\}\right|_{\theta=0} \tag{13.62}$$

where the quadratic form p equals $(\underline{r}-\theta\underline{s})^{\mathrm{T}}\mathbf{M}^{-1}(\underline{r}-\theta\underline{s})$ since $\underline{d} = \underline{r}-\theta\underline{s}$. From the chain rule for differentiation we have

$$\frac{\partial}{\partial\theta}(h_N(p)) = \frac{\partial}{\partial p}(h_N(p))\frac{\partial p}{\partial\theta} \tag{13.63}$$

From the expression for p

$$\left.\frac{\partial p}{\partial\theta}\right|_{\theta=0} = -2(\underline{s}^{\mathrm{T}}\mathbf{M}^{-1}\underline{r}) \tag{13.64}$$

Making use of Equation 13.62 to Equation 13.64 the LOD statistic in Equation 13.56 becomes

$$T_{\mathrm{LOD}}(\underline{r}) = -2(\underline{s}^{\mathrm{T}}\mathbf{M}^{-1}\underline{r})\frac{h_N'(p)}{h_N(p)} \tag{13.65}$$

where $h_N'(p)$ denotes the derivative of the function $h_N(p)$ with respect to the argument p. The LOD statistic in Equation 13.65 represents the canonical structure when the disturbance is modeled as an SIRV. The nonlinear function $h_N(p)$ depends on the particular joint density function used to model the disturbance. The canonical structure for the detector is shown in Figure 13.2. Note that the detector multiplies the output of a matched filter with the output of a nonlinearity. Just as with a Gaussian receiver, the matched filter maximizes the signal to disturbance ratio even though the received signal is nonGaussian (i.e., derivation of the matched filter to maximize signal to noise ratio does not depend on the Gaussian assumption). For nonGaussian problems, matched filtering alone is suboptimum. For SIRPs the optimal receiver requires nonlinear processing as well as matched filtering.

13.2.6.1.3. Random Variables Arising from the Gaussian Distribution

The SIRP class of disturbance reduces to the Gaussian distribution when

$$h_N(p) = \mathrm{e}^{-P/2} \tag{13.66}$$

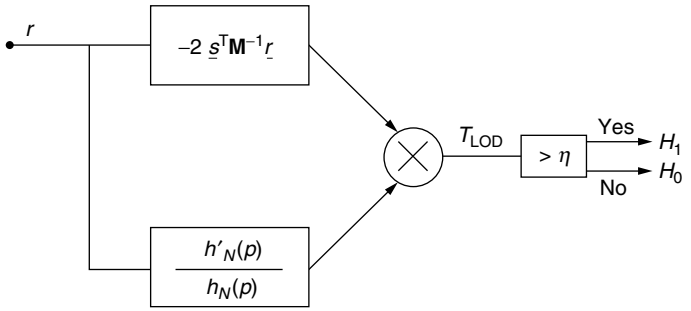

FIGURE 13.2 Canonical form of LOD assuming known signal and random variables arising from an SIRP.

It follows that

$$\frac{h'_N(p)}{h_N(p)} = -\frac{1}{2} \tag{13.67}$$

With reference to Equation 13.65, the LOD statistic becomes

$$T_{\mathrm{LOD}}(\underline{r}) = -\underline{s}^{\mathrm{T}}\mathbf{M}^{-1}\underline{r} \tag{13.68}$$

Interestingly enough, this is identical to the statistic of the LRT for the known signal Gaussian problem.[30] Hence, for the known signal Gaussian problem, the strong and the weak signal detectors are identical. Note that there is no nonlinearity involved with the weak signal detector for this case. To put it another way, the general nonlinear SIRP weak signal detector of Figure 13.2 reduces to the linear receiver or matched filter known to be optimum for the Gaussian problem.

13.2.6.2. The Random Signal Problem

13.2.6.2.1. Independent Disturbance Random Variables

The LOD is given by Equation 13.55 when the signal is random. Repeating Equation 13.55 the LOD structure is

$$T_{\mathrm{LOD}}(\underline{r}) = \frac{(\nabla_r^{\mathrm{T}}\mathbf{P}\nabla_r)f_{\underline{\mathrm{D}}}(\underline{r})}{f_{\underline{\mathrm{D}}}(\underline{r})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{\mathrm{u}} \tag{13.69}$$

P is the random signal covariance matrix. In this section the components of the disturbance vector $\underline{\mathrm{D}}$ are assumed to be statistically independent. The analysis is further simplified when the signal random variables are also assumed to be uncorrelated. The covariance matrix **P** then becomes diagonal. Let the diagonal elements of the matrix **P** be represented by $\sigma_i^2, i = 1, 2, \ldots, N$. Because the disturbance random variables are independent, the joint density function $f_{\mathrm{D}}(\underline{r})$ is again given by the product of the marginal density functions of the

individual random variables. Specifically,

$$f_{\underline{D}}(\underline{r}) = \prod_{i=1}^{N} f_{D_i}(r_i) \tag{13.70}$$

Also, when P is diagonal,

$$\nabla_r^{\mathrm{T}} P \nabla_r = \sum_{i=1}^{N} \sigma_i^2 \frac{\partial^2}{\partial r_i^2} \tag{13.71}$$

Using Equation 13.69 to Equation 13.71 and following the same steps as in the known signal case, the LOD statistic can be derived as

$$T_{\mathrm{LOD}}(\underline{r}) = \sum_{i=1}^{N} \sigma_i^2 \frac{f''_{D_i}(r_i)}{f_{D_i}(r_i)} \tag{13.72}$$

where the double prime indicates second derivative with respect to the argument. The canonical structure derived above is shown in Figure 13.3.

13.2.6.2.2. Disturbance Random Variables from an SIRP Distribution

When the disturbance vector is modeled as having an SIRP distribution, the joint PDF is given by Equation 13.61. The LOD structure for the random signal case is given by Equation 13.69. Since the constant terms in the joint density function cancel out in the numerator and denominator of the ratio test in Equation 13.69,

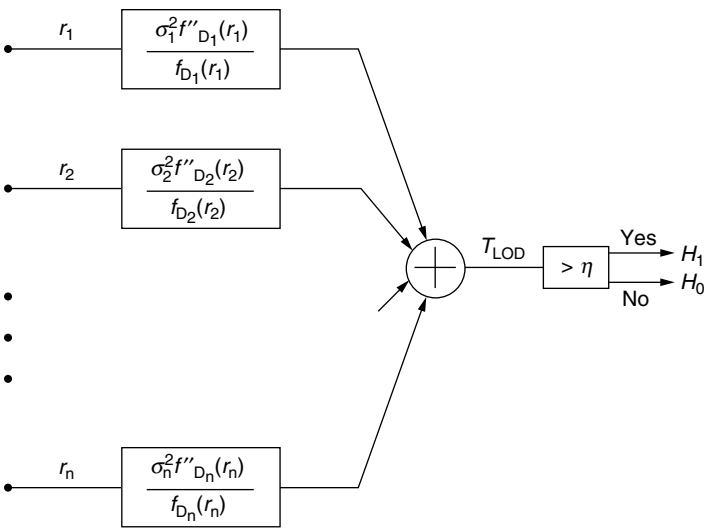

FIGURE 13.3 Canonical form of LOD assuming random signal and independent random variables.

the LOD statistic is obtained by evaluating

$$T_{\mathrm{LOD}}(\underline{r}) = \frac{(\nabla_r^{\mathrm{T}} \mathbf{P} \nabla_r) h_N(p)}{h_N(p)} \tag{13.73}$$

The numerator of Equation 13.73 can be expanded as a sum of terms involving partial derivatives. The result is simplified considerably when the covariance matrix **P** of the signal vector is diagonal. When **P** is chosen to be the Identity matrix (i.e., **P** is diagonal and the variance of each element of the signal vector is unity), the LOD statistic is given by

$$T_{\mathrm{LOD}}(\underline{r}) = \frac{(\nabla_r^{\mathrm{T}} \nabla_r) h_N(p)}{h_N(p)} \tag{13.74}$$

The inner product involving the ∇ vector can be written as

$$\nabla_r^{\mathrm{T}} \nabla_r = \sum_{i=1}^{N} \frac{\partial^2}{\partial r_i^2} \tag{13.75}$$

Application of Equation 13.75, to the numerator of Equation 13.74, results in

$$\sum_{i=1}^{N} \frac{\partial^2 h_N(p)}{\partial r_i^2} = \sum_{i=1}^{N} h_N'(p) \frac{\partial^2 p}{\partial r_i^2} + h_N''(p) \left(\frac{\partial p}{\partial r_i} \right)^2 \tag{13.76}$$

where the prime indicates differentiation with respect to p. Using Equation 13.76 and dividing by $h_N(p)$ the LOD statistic becomes

$$T_{\mathrm{LOD}}(\underline{r}) = \frac{1}{h_N(p)} \left[\sum_{i=1}^{N} \left(h_N'(p) \frac{\partial^2 p}{\partial r_i^2} + h_N''(p) \left(\frac{\partial p}{\partial r_i} \right)^2 \right) \right] \tag{13.77}$$

The quadratic form p can be written as

$$p = \sum_{k=1}^{N} \sum_{l=1}^{N} r_k M_{kl}^{-1} r_l \tag{13.78}$$

where M_{kl}^{-1} represents the kth row and lth column entry of the matrix $\mathbf{M}^{-1}$. From Equation 13.78 $\partial p / \partial r_i$, $(\partial p / \partial r_i)^2$ and $\partial^2 p / \partial r_i^2$ can be calculated. In particular, we have

$$\frac{\partial p}{\partial r_i} = \frac{\partial p}{\partial r_i} \sum_{k=1}^{N} \sum_{l=1}^{N} r_k M_{kl}^{-1} r_l = \sum_{k=1}^{N} r_k M_{ki}^{-1} + \sum_{l=1}^{N} r_l M_{il}^{-1} \tag{13.79}$$

Because of the symmetric nature of the matrices $\mathbf{M}$ and $\mathbf{M}^{-1}$, $M_{ki}^{-1} = M_{ik}^{-1}$. It follows that the square of the above equation is then given by

$$\left(\frac{\partial p}{\partial r_i}\right)^2 = \left(\sum_{k=1}^{N} r_k M_{ki}^{-1} + \sum_{l=1}^{N} r_l M_{il}^{-1}\right)^2 = \left(2\sum_{k=1}^{N} r_k M_{ki}^{-1}\right)^2$$

$$= 4\sum_{k=1}^{N}\sum_{l=1}^{N} r_k M_{ki}^{-1} M_{li}^{-1} r_l = 4\sum_{k=1}^{N}\sum_{l=1}^{N} r_k M_{ki}^{-1} M_{il}^{-1} r_l \tag{13.80}$$

Utilizing Equation 13.79,

$$\frac{\partial^2 p}{\partial r_i^2} = \frac{\partial}{\partial r_i}\left(\sum_{k=1}^{N} r_k M_{ki}^{-1} + \sum_{l=1}^{N} r_l M_{il}^{-1}\right) = 2M_{ii}^{-1} \tag{13.81}$$

With reference to Equation 13.77 and Equation 13.79 to Equation 13.81, define

$$T_{\mathrm{LOD}}^{(1)}(\underline{r}) = \frac{1}{h_N(p)}\left[\sum_{i=1}^{N} h_N'(p)\frac{\partial^2 p}{\partial r_i^2}\right] = 2\frac{h_N'(p)}{h_N(p)}\sum_{i=1}^{N} M_{ii}^{-1} \tag{13.82}$$

$$T_{\mathrm{LOD}}^{(2)}(\underline{r}) = \frac{1}{h_N(p)}\sum_{i=1}^{N} h_N''(p)\left(\frac{\partial p}{\partial r_i}\right)^2 = 4\frac{h_N''(p)}{h_N(p)}\sum_{i=1}^{N}\sum_{k=1}^{N}\sum_{l=1}^{N} r_k M_{ki}^{-1} M_{il}^{-1} r_l$$

$$= \frac{4h_N''(p)}{h_N(p)}\underline{r}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{M}^{-1}\underline{r} \tag{13.83}$$

The LOD statistic that results from Equation 13.82 and Equation 13.83 is written as

$$T_{\mathrm{LOD}}(\underline{r}) = T_{\mathrm{LOD}}^{(1)}(\underline{r}) + T_{\mathrm{LOD}}^{(2)}(\underline{r})$$

$$= \frac{1}{h_N(p)}[2h_N'(p)tr(\mathbf{M}^{-1}) + 4h_N''(p)\underline{r}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{M}^{-1}\underline{r}] \tag{13.84}$$

where $tr(\mathbf{M}^{-1})$ is the sum of the all the diagonal elements of the matrix $\mathbf{M}^{-1}$. The canonical structure of the receiver is shown in Figure 13.4.

13.2.6.2.3. Disturbance Random Variables Arising from the Gaussian Distribution

As pointed out in Section 13.2.6.1.3, the SIRP disturbance reduces to the Gaussian distribution when

$$h_N(p) = \mathrm{e}^{-p/2} \tag{13.85}$$

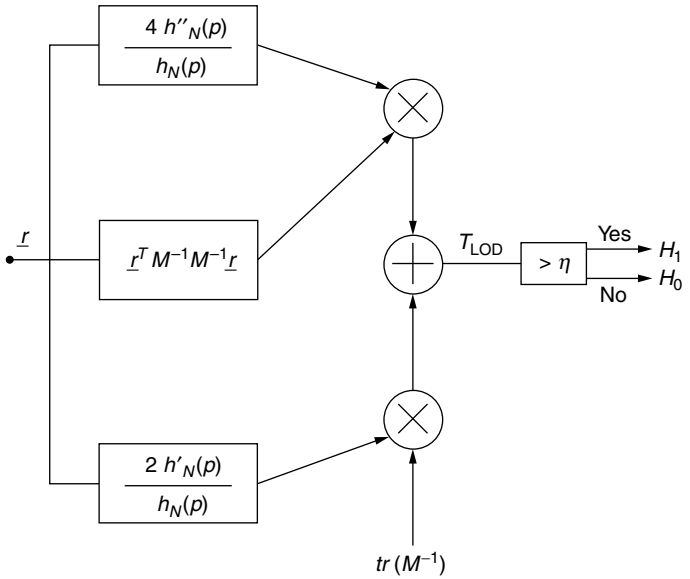

FIGURE 13.4 Canonical form of LOD assuming random signal and random disturbance arising from an SIRP.

From the above equation it follows that

$$\frac{h'_N(p)}{h_N(p)} = -\frac{1}{2} \qquad \frac{h''_N(p)}{h_N(p)} = \frac{1}{4} \tag{13.86}$$

Consequently Equation 13.82 reduces to

$$T^{(1)}_{\mathrm{LOD}}(\underline{r}) = -tr(\mathbf{M}^{-1}) \tag{13.87}$$

whereas Equation 13.83 becomes

$$T^{(2)}_{\mathrm{LOD}}(\underline{r}) = \underline{r}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{M}^{-1}\underline{r} \tag{13.88}$$

Note that $T^{(1)}_{\mathrm{LOD}}(\underline{r})$ is a constant that can be combined with the threshold. As a result, the LOD statistic for the random signal Gaussian problem is given by

$$T_{\mathrm{LOD}}(\underline{r}) = \underline{r}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{M}^{-1}\underline{r} \tag{13.89}$$

Unlike the known signal Gaussian problem, the weak signal LOD statistic does not equal the statistic of the likelihood ratio for the random signal Gaussian problem,[30] which is

$$T_{\mathrm{LR}}(\underline{r}) = \underline{r}^{\mathrm{T}}[\mathbf{M}^{-1} - (\mathbf{M}+\mathbf{P})^{-1}]\underline{r} \tag{13.90}$$

The two statistics becomes equivalent only when

$$\mathbf{M} = \mathbf{P} = \mathbf{I} \tag{13.91}$$

Although the strong and weak signal detectors have different test statistics, the receiver structures are both quadratic in nature.

13.3. DETERMINING THRESHOLDS FOR THE LOD

13.3.1. Literature Review

In this dissertation, multivariate density functions for modeling nonGaussian PDFs are assumed to be known. They are obtained using the theory of SIRPs. Once the multivariate density functions are known, we derive a decision rule using the theory of LODs, that is applicable when the signal to be detected is weak compared to the additive disturbance. The procedure for obtaining the decision rule is explained in detail in Section 13.2.

However, the detector that is obtained on the basis of the theory of LODs is typically nonlinear as the underlying processes are nonGaussian. When the test statistic is nonlinear, it is not possible to evaluate the performance of the detector analytically. Consequently, we have to resort to computer simulations to analyze the performance. There are two steps involved in computer simulations to analyze performance. The first step is to evaluate the threshold so as to obtain the desired false alarm probability. The second step is to evaluate the detection probability once the threshold is set, corresponding to the desired false alarm probability.

13.3.1.1. Classical Methods for Evaluating Thresholds

Monte Carlo methods have typically been used for this purpose. A large number of trials M are generated under the hypothesis that the received signal consists of the disturbance alone. The detector outputs, T_{si} $i = 1, 2, \ldots, M$ corresponding to the generated disturbance vectors, are recorded. Based on the output of the detector, thresholds can be set to obtain the desired false alarm probability. But, in order to establish the threshold for a specified P_{F}, it is necessary to accurately know the behavior of the tail of the test statistic. Unfortunately, the number of trials required for the Monte Carlo technique is very large, as is evident from the rule of thumb

$$M \geq \frac{10}{P_{\mathrm{F}}} \tag{13.92}$$

Hence, if $P_{\mathrm{F}} = 10^{-5}$, one million trials should be generated. Clearly, this is not a very desirable situation. Thus, for a reasonable sample size M, estimation of thresholds corresponding to small false alarm probabilities cannot be made when these methods are used. For Monte Carlo simulations the construction of approximate confidence intervals for the threshold estimates based on various estimators are discussed by Hosking and Wallis.[31]

Some other approaches that do not make use of raw data or their smoothed versions have been suggested by various authors. For example, Harrel and Davis[32] suggested using linear combinations of sample order statistic. Their approach appears to be applicable for estimation of thresholds in the central region of the distribution. However, the underlying threshold estimator is not in a simple computational form and does not offer any additional advantage over the Monte Carlo method in terms of threshold estimation corresponding to small false alarm probabilities.

It has recently been shown[33] that the PDFs of the test statistic can be determined experimentally using a relatively small number of samples (e.g., 50 to 100 samples give accurate fits depending on the distribution). Because the number of samples required by Ozturk's technique is small, it is unlikely that samples will be from the extreme tails of the PDFs. Consequently, the accurate fit mentioned above applies to the main body of the density function.

A number of statisticians have developed methods for estimating the extreme tail of the distributions using the asymptotic properties of extreme order statistics. Assuming an unknown PDF $f_X(x)$, then, for large X, Hill[34] proposed using $\Lambda(x) = 1 - cx^{-\alpha}$ as a limiting form of the distribution function to infer the tail behavior. A similar approach is also given by Pickands.[35] Weissman[36] proposed a different approach based on the joint limiting distribution of the largest k order statistics. His approach is based on the fact that "the largest k order statistics have a joint distribution that is approximately the same as the largest k order statistics from a location and scale exponential distribution." Weissman obtained a simple expression for the estimate of the thresholds corresponding to various false alarm probabilities of the distributions. Based on the empirical comparisons, Boos[37] reported that Weissman's estimators have lower mean squared error than those of standard methods when the tails are exactly exponential. When the tails are not exactly exponential the estimators become highly biased. The mean squared errors of the estimators strongly depend on the choice for k. Although the method is nonparametric in nature, the optimal choice of k requires the knowledge of the parent distribution.

The use of stable distributions to model data having large tailed distributions has attracted considerable attention.[38–41] The independent and identically distributed random variables $Y_1, Y_2, \ldots, Y_n$ are said to have a stable distribution if $Y_1 + Y_2 + \cdots + Y_n$ have the same distribution as the individual random variables. With the introduction of additional parameters, control of the mean, variance, and the skewness of the distribution is possible. A major difficulty with stable distributions is that they usually cannot be expressed in closed form. Also, estimation of parameters is not computationally easy.[42]

13.3.2. Extreme Value Theory

Guida et al.[43] compared the performance analysis of some extrapolative estimators of probability tails with application to radar systems design. They show that the estimates based on the extreme value theory yield clearly superior

accuracy, while achieving a substantial savings in sample size compared to the classical Monte Carlo techniques. However, their method suffers from two major drawbacks. First, they assumed that the underlying unknown distribution is always exponential in nature. This assumption can be restrictive in certain situations. The other drawback is that the samples are partitioned into many smaller sets of samples and the maximum from each set is drawn for estimation purposes. They provide no optimum rule for determining the number of sets to be used in partitioning the original sample even though the accuracy of the estimation depends strongly on the original sample size and the number of sets.

Pickands[35] first suggested that the Generalized Pareto distribution (GPD) can be used to model to extreme tails of PDFs. The GPD is a two parameter distribution, with a scale and a shape parameter. Modeling the extreme tail then corresponds to estimating the two parameters of the GPD. The estimation methods for the GPD have been reviewed by Hosking and Wallis.[31] They considered the method of moments, probability-weighted moments, and the maximum likelihood method for estimating the parameters and the thresholds. Based on computer simulation experiments, they showed that the probability weighed moment method is more reliable than the maximum likelihood method for sample sizes less than 500.

13.3.3. The Radar Problem

The hypothesis testing problem for deciding whether or not a target is present is given by Equation 13.5 to Equation 13.6 in Section 13.2. For weak signal applications, it was shown that the LOD is useful for arriving at a decision rule. For the known signal case, the LOD structure is given by Equation 13.36. Since the test statistic is typically a nonlinear function when $f_D(\underline{r}|H_0)$ and $f_D(\underline{r}|H_1)$ are multivariate nonGaussian density functions, it is not possible, in general, to analytically evaluate in closed form the threshold η for a specified false alarm probability. Given the PDFs of the test statistic, T_s, under hypotheses H_1 and H_0, the detection and false alarm probabilities are

$$P_D = \int_{\eta}^{\infty} f_{T_s}(t_s|H_1)\mathrm{d}t_s \tag{13.93}$$

$$P_F = \int_{\eta}^{\infty} f_{T_s}(t_s|H_0)\mathrm{d}t_s \tag{13.94}$$

P_D and P_F are represented by the shaded areas shown in Figure 13.5. As indicated in the figure, P_F is typically much smaller than P_D.

In practice, the density function of T_s is not known in advance. For example, depending upon various conditions such as terrain, weather etc., the clutter may be best modeled by Gaussian, K-distributed, Weibull or some other probability

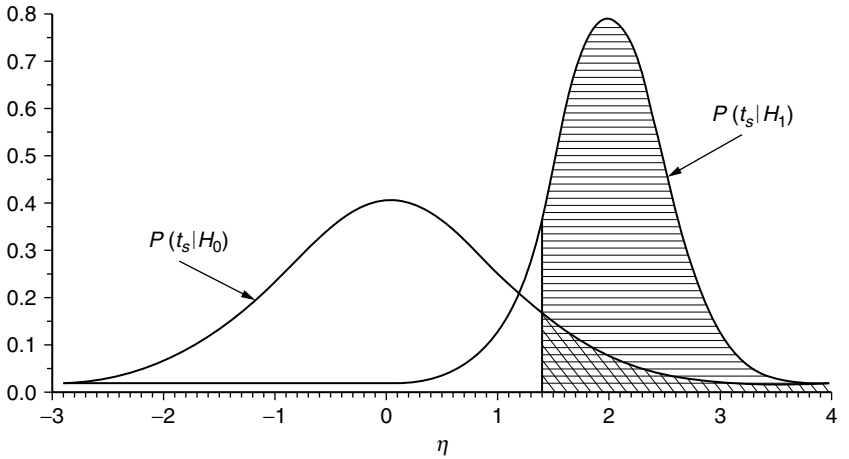

FIGURE 13.5 Shaded areas indicating P_F and P_D.

distribution. In this section a new approach is developed for experimentally determining the extreme tail of $f_{T_s}(t_s|H_0)$, where the number of samples required is several orders of magnitude smaller than that suggested by Equation 13.92. Once the tail of $f_{T_s}(t_s|H_0)$, has been estimated, the threshold can be determined by use of Equation 13.94.

13.3.4. Methods for Estimating Thresholds

13.3.4.1. Estimates Based on Raw Data

In this section we consider some commonly used threshold estimates. These estimates are called raw estimates and are already included in some statistical package programs (e.g., the UNIVARIATE procedure in the SAS[44] package).

Let $X_1 \le X_2 \le \cdots \le X_n$ denote the sample order statistics from a distribution function $F(x)$. Let p denote the desired false alarm probability. Also, let $n(1-p) = j + g$ where j is the integer part of $n(1-p)$. We denote the threshold estimate based on the kth procedure to be described below by $\eta_p^{(k)}$. Four different threshold estimates are given as follows:

$$\eta_\alpha^{(1)} = (1-g)X_j + gX_{j+1} \tag{13.95}$$

$$\eta_\alpha^{(2)} = X_k, \text{ where } k \text{ is the integer part of } [n(1-\alpha) + 1/2] \tag{13.96}$$

$$\eta_\alpha^{(3)} = (1-\delta)X_j + \delta X_{j+1}, \quad \delta = 0 \text{ if } g = 0; \delta = 1 \text{ if } g > 0 \tag{13.97}$$

$$\eta_\alpha^{(4)} = \delta X_{j+1} + (1-\delta)(X_j + X_{j+1})/2, \quad \delta = 0 \text{ if } g = 0; \delta = 1 \text{ if } g > 0 \tag{13.98}$$

It is known that all of the above methods are asymptotically equivalent. Thus, if a large sample size is used (where for example M is determined from

Equation 13.92), the choice of the *best* method is no longer critical. However, in an empirical study,[37] it has been shown that $\eta_{\alpha}^{(4)}$ outperformed the other estimators when $g = 0$. It is noted that the methods based on the above estimators are restricted by the condition $1 \leq n(1-\alpha) \leq n-1$. This implies that the smallest value of the false alarm probability α cannot be lower than $1/n$. Consequently, the threshold corresponding to the smallest false alarm probability, which can be estimated by these procedures depends on the sample size. Thus, for a reasonable size of n, estimation of thresholds for small false alarm probabilities cannot be made when these methods are used.

13.3.4.2. Estimates Motivated by the Extreme Value Theory

Extreme value distributions are obtained as limiting distributions of largest (or smallest) values of sample order statistics. Assuming independent trials, if $X_1 \leq X_2 \leq \cdots \leq X_n$ are order statistics from a common distribution function $F(x)$, then the cumulative distribution function of the largest order statistic is given by

$$G_n(x) = P(X_n \leq x) = [F(x)]^n \tag{13.99}$$

It is clear, as $n \rightarrow \infty$, that the limiting value of $G_n(x)$ approaches zero if $F(x)$ is less than 1, and unity if $F(x)$ is equal to 1 for a specified value of x. A standardized limiting distribution of X_n may be obtained by introducing the linear transformation, $a_n X_n + b_n$, where a_n and b_n are finite constants depending on the sample size n.

In Appendix X, using the theory of limiting distributions,[45] it is shown that if there exist sequences a_n and b_n such that

$$\lim_{n\rightarrow\infty} P\left(\frac{X_n - b_n}{a_n} \leq x\right) = \lim_{n\rightarrow\infty} F^n(a_n x + b_n) = \lim_{n\rightarrow\infty} G_n(a_n x + b_n) \rightarrow \Lambda(x) \tag{13.100}$$

then the solution of Equation 13.100 yields all the possible forms for the distribution function $G_n(x)$ in the limit as $n \rightarrow \infty$. The solutions to the above equation are derived in Appendix X and are rewritten here:

$$\Lambda(x) = \exp(-e^{-x}) \qquad x \geq 0 \tag{13.101}$$

$$\Lambda(x) = \exp(-x^{-k}) \qquad x \geq 0, \;\; k > 0 \tag{13.102}$$

$$\Lambda(x) = \exp(-(-x)^{k}) \qquad x \leq 0, \;\; k > 0 \tag{13.103}$$

In the limit, as n gets large, these are the three types of distribution functions to which the largest order statistic, drawn from almost any smooth and continuous distribution function, converge. By differentiating the three functions, we obtain

analytical expressions for the limiting forms of the PDFs. However, because of the differentiation, it should be recognized that these expressions may not be good approximations to the density functions. In practice, extreme value theory should always be applied to a distribution function, or equivalently, the area under the density function. For $x \geq 0$, differentiation of Equation 13.101 and Equation 13.102 result in

$$1. \quad \frac{d\Lambda(x)}{dx} \approx H(x) = e^{-x} \tag{13.104}$$

$$2. \quad \frac{d\Lambda(x)}{dx} \approx H(x) = kx^{-(k+1)} \quad k \geq 0 \tag{13.105}$$

The first equation above is the well known exponential distribution and the second equation is related to the Pareto distribution. The details that lead to these analytical expressions are shown in Appendix X.

It remains to be explained how the distribution of the largest order statistic is related to the tails of the underlying PDF from which the samples are drawn. The relationship is based on the observation that inferences from short sequences are likely to be unreliable. In particular, instead of observing k sets of n samples and taking the largest order statistic from each of the k sets, it is better to observe a single set of nk samples and use the largest k samples from this set.[46] The k largest order statistics from a vector of nk observations constitute the tail of the underlying distribution especially when n is very large. Therefore, the limiting distribution of the largest order statistic closely approximates the tail of the underlying PDF for large n.

13.3.5. The Generalized Pareto Distribution

The GPD is defined for $x > 0$ by the distribution function

$$G(x) = 1 - (1 + \gamma x/\sigma)^{-1/\gamma}, \quad -\infty < \gamma < \infty, \sigma > 0, \gamma x > -\sigma \tag{13.106}$$

This distribution has a simple closed form and includes a range of distributions depending upon the choice of γ and σ. For example, the exponential distribution results for $\gamma = 0$ and the uniform distribution is obtained when $\gamma = -1$. The GPD defined in Equation 13.106 is valid for all $x > 0$ while Equation 13.104 and Equation 13.105 are valid only for large x.

The PDF corresponding to the GPD is given by

$$g(x) = \frac{d}{dx}\left[1 - \left(1 + \frac{\gamma x}{\sigma}\right)^{-1/\gamma}\right] = \frac{1}{\sigma}\left(1 + \frac{\gamma x}{\sigma}\right)^{-(1/\gamma)-1} \tag{13.107}$$

If we let $\gamma \rightarrow 0$ in the above equation, note that

$$\lim_{\gamma \to 0} = \frac{1}{\sigma}\left(1 + \frac{\gamma x}{\sigma}\right)^{-(1/\gamma)-1} = \frac{1}{\sigma}e^{-x/\sigma} \tag{13.108}$$

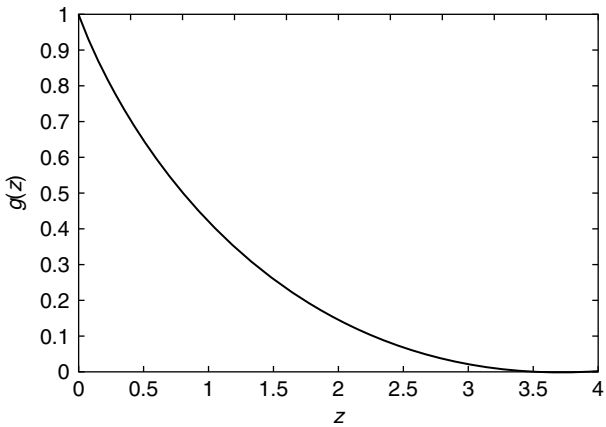

FIGURE 13.6 Generalized Pareto PDF, $\gamma = -1$.

Also, if we let x be large in Equation 13.107, note that

$$\frac{1}{\sigma}\left(1 + \frac{\gamma x}{\sigma}\right)^{-(1/\gamma)-1} \approx \frac{1}{\sigma}\left(\frac{\gamma}{\sigma}\right)^{-(1/\gamma)-1} x^{-(1/\gamma)-1} \tag{13.109}$$

Equation 13.108 and Equation 13.109 are of the same form as Equation 13.104 and Equation 13.105, respectively. Thus, the GPD can be used to approximate both types of tail behavior exhibited by the right tail. Typical plots of the Generalized Pareto PDF for $\gamma < 0$ and $\gamma > 0$ are shown in Figure 13.6 and Figure 13.7, respectively.

We wish to set thresholds for specified false alarm probabilities when the underlying density functions are unknown. To set very small false alarm

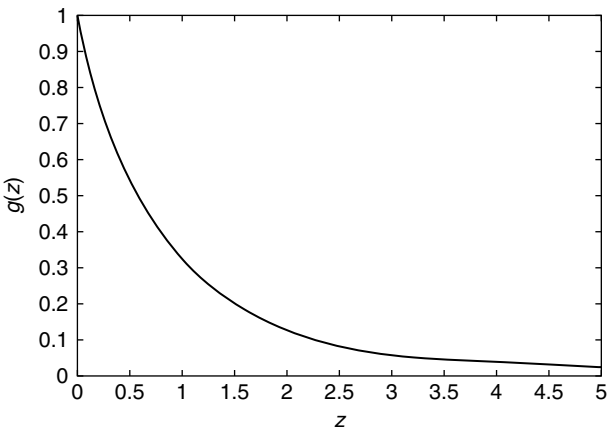

FIGURE 13.7 Generalized Pareto PDF, $\gamma = 1$.

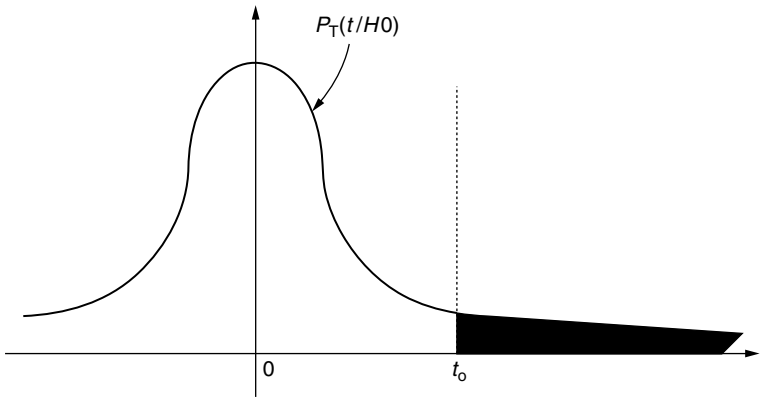

FIGURE 13.8 PDF of test statistic with tail region defined for $t \geq t_0$.

probabilities, the tail of the PDF $f_{T_s}(t_s|H_0)$ has to be accurately modeled. Figure 13.8 represents a typical PDF of the test statistic with the tail region of the PDF being defined as that to the right of $t = t_0$. Figure 13.9 shows the tail translated to the origin. The choice for t_0 is somewhat arbitrary. For example, t_0 can be chosen such that the area in the shaded region equals 0.1, 0.05, or 0.01. It is the portion of the PDF to the right of t_0 that we are interested in modeling by the GPD. In particular, the tail region of the PDF is translated to the origin and modeled as a GPD. Once the estimates of σ and γ have been obtained, the GPD is multiplied by the area of the shaded region and translated back to the point t_0. In this way, the area under the PDF of the test statistic is maintained at unity.

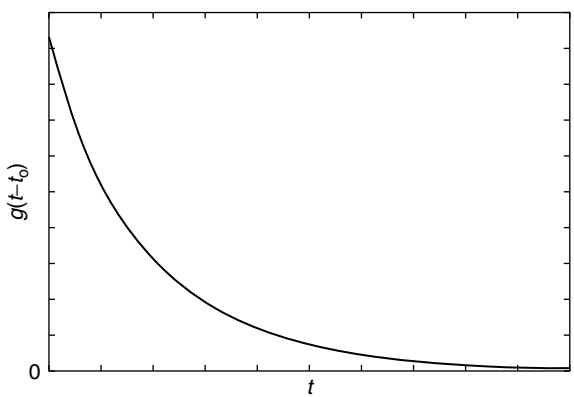

FIGURE 13.9 Tail of the test statistic shifted to origin.

13.3.5.1. Methods for Estimating the Parameters of the GPD

Suppose that the sample ordered statistics $X_1 \leq X_2 \leq \cdots \leq X_n$ are drawn from the distribution function $F(x)$. To estimate the right tail of this distribution it is necessary to determine a value (say, x_0) and then use those sample observations greater than x_0 to obtain the quantity $z = x - x_0$. Once the tail observations have been chosen, the GPD can be fitted to these observations by using standard methods of parameter estimation. Note that the portion of the observations used from a complete set of samples depends on the choice of x_0. One approach to selecting x_0 is to make a histogram of the data and choose x_0 to be near the point of inflection of the histogram. DuMouchel[47] proposed choosing x_0 to be the value such that $\int_{x_0}^{\infty} f_X(x)\mathrm{d}x = 0.1$. Such an approach is less subjective and appears to be satisfactory for many applications. However, it is noted by DuMouchel that "using an even smaller fraction of observations would restrict profitable use of the statistic to much larger sizes. On the other hand, to use more than the upper one tenth of a sample would seem to allow too much dependence on the central part of the distribution." In other words, if a smaller fraction is used, we need larger sample sizes to get an adequate number of samples for estimation and if a larger fraction is used, the body of the distribution may influence estimation of the tail.

Let x_0 be such that $1 - F(x_0) = \int_{x_0}^{\infty} f_X(x)\mathrm{d}x = \lambda$. The distribution function to be used in approximating the tail can be written as

$$\tilde{F}(x) = (1 - \lambda) + \lambda G(x - x_0) = 1 - \lambda\left[1 + \frac{\gamma}{\sigma}(x - x_0)\right]^{-1/\gamma}, \; x > x_0 \quad (13.110)$$

where $G(x)$ is given in Equation 13.106. Assuming that the tail of a given distribution can be adequately approximated by Equation 13.110, the estimation problem of the distribution in the tail region is reduced to estimation of the parameters of the GPD.

In this chapter we consider three methods for the parameters estimation of the GPD. The three methods are maximum likelihood estimation, the method of probability weighted moments, and the ordered sample least squares (OSLS) approach. The first two methods, applied to the GPD, are discussed by Hosking and Wallis.[31] The OSLS approach is a new technique developed in this work. The performance of the three estimation procedures and compared on the basis of estimation bias and mean square error.

13.3.5.1.1. Maximum Likelihood Estimation

The PDF corresponding to the GPD from Equation 13.107, with x replaced by z, is

$$g(z) = \frac{1}{\sigma}\left(1 + \frac{\gamma z}{\sigma}\right)^{-(1/\gamma)-1} \quad (13.111)$$

Given a sample vector $[z_1, z_2, \ldots, z_m]$ from the GPD the joint density function, then $L_{\underline{Z}}(\underline{z})$ of the m samples, assuming independence, is given by

$$L_{\underline{Z}}(\underline{z}) = \frac{1}{\sigma^m} \prod_{i=1}^{m} \left(1 + \frac{\gamma z_i}{\sigma}\right)^{-(1/\gamma)-1} \tag{13.112}$$

To theoretically obtain the maximum likelihood estimates of σ and γ, the logarithm of the joint density function in Equation 13.112 is differentiated with respect to σ and γ, respectively, and the derivatives are set to zero. Let the m largest observations from the unknown distribution whose tail is being modeled by the GPD be placed in the vector $[x_{n-m+1}, x_{n-m+2}, \ldots, x_n]$. Translation of the tail region to the origin results in the vector $[x_{n-m+1} - x_0, x_{n-m+2} - x_0, \ldots, x_n - x_0] = [z_1, z_2, \ldots, z_m]$. Letting $\tau = \gamma/\sigma$ in Equation 13.112 and differentiating the logarithm of the joint density function with respect to σ we get

$$\begin{aligned} -\frac{\mathrm{d}}{\mathrm{d}\sigma} \log L_{\underline{Z}}(\underline{z}) &= \frac{\mathrm{d}}{\mathrm{d}\sigma}\left[m \log(\sigma) + (1 + (\tau\sigma)^{-1}) \sum_{i=1}^{m} \log(1 + \tau z_i) \right] \\ &= \frac{m}{\sigma} + \left(1 - \frac{1}{\sigma^2 \tau}\right) \sum_{i=1}^{m} \log(1 + \gamma z_i/\sigma) \end{aligned} \tag{13.113}$$

By setting Equation 13.113 to zero and solving for σ we obtain

$$\sigma(\tau) = \sum_{i=1}^{m} \log(1 + \tau z_i)/(m\tau) \tag{13.114}$$

The expression for σ is now substituted into Equation 13.113, so as to obtain a function of τ alone. $\hat{\tau}$ is derived by differentiating the expression

$$m \log \sigma(\tau) + (1 + [\sigma(\tau)\tau]^{-1}) \sum_{i=1}^{m} \log(1 + \tau z_i) \tag{13.115}$$

with respect to τ and setting the derivative equal to zero with the constraint that $\tau z_i > -1$. However, the differentiation leads to a nonlinear equation whose analytical solution is not known. This difficulty is circumvented by minimizing Equation 13.115 numerically with respect to τ. The numerical minimization was performed using the Nelder–Mead algorithm.[48] Once the estimate for τ has been obtained, then $\hat{\sigma}$ is obtained form Equation 13.114 and λ is estimated by $\hat{\gamma} = \hat{\sigma}\hat{\tau}$.

13.3.5.1.2. Probability Weighted Moments

The probability weighted moments of a continuous random variable Z with distribution function G are the quantities

$$M_{p,r,s} = E[Z^p G^r(Z)(1 - G(Z))^s] \tag{13.116}$$

where E is the expectation operator of p, r and s are real numbers. For the GPD it is convenient to choose $p = 1$ and $r = 0$, respectively. Then the probability weighted moments are

$$M_{1,0,s} = E[Z(1 - G(Z))^s] \tag{13.117}$$

For the GPD there are two parameters to be estimated, σ, and γ. Substituting $s = 0$ in Equation 13.117, we get

$$\begin{aligned} \varepsilon_0 = M_{1,0,0} &= E[Z] \\ &= \int_0^\infty \frac{Z}{\sigma}\left(1 + \frac{\gamma Z}{\sigma}\right)^{-(1/\gamma)-1} \mathrm{d}Z \end{aligned} \tag{13.118}$$

Letting $1 + (\gamma Z/\sigma) = Y$, Equation 13.118 results in

$$\begin{aligned} \varepsilon_0 &= \frac{\sigma}{\gamma^2}\int_1^\infty (Y-1)Y^{-(1/\gamma)-1}\,\mathrm{d}Y \\ &= \frac{\sigma}{\gamma^2}\left[\frac{Y^{-(1/\gamma)-1}}{-\frac{1}{\gamma}+1} - \frac{Y^{-1/\gamma}}{-\frac{1}{\gamma}}\right]_1^\infty \\ &= \frac{\sigma}{1-\gamma} \end{aligned} \tag{13.119}$$

Letting $s = 1$ in Equation 13.117 we obtain

$$\begin{aligned} \varepsilon_1 = M_{1,0,1} &= E[Z(1 - G(Z))] \\ &= \int_0^\infty \frac{Z}{\sigma}\left(1 + \frac{\gamma Z}{\sigma}\right)^{-1/\gamma}\left(1 + \frac{\gamma Z}{\sigma}\right)^{-(1/\gamma)-1} \mathrm{d}Z \end{aligned} \tag{13.120}$$

Letting $1 + (\gamma Z/\sigma) = Y$, as before, Equation 13.120 results in

$$\varepsilon_1 = \frac{\sigma}{\gamma^2}\int_1^\infty (Y-1)Y^{-(2/\gamma)-1}\mathrm{d}Y = \frac{\sigma}{\gamma^2}\left[\frac{Y^{-(2/\gamma)-1}}{-\frac{2}{\gamma}+1} - \frac{Y^{-2/\gamma}}{-\frac{2}{\gamma}}\right]_1^\infty$$

$$= \frac{\sigma}{2(2-\gamma)} \tag{13.121}$$

The values of ε_0 and ε_1 are obtained from Equation 13.119 and Equation 13.121, respectively, for given values of σ and γ. Since there are two equations in two unknowns, σ and γ can be obtained as functions of ε_0 and ε_1, respectively. Solving for σ and γ we obtain

$$\hat{\sigma} = 2\varepsilon_0\varepsilon_1/(\varepsilon_0 - 2\varepsilon_1) \tag{13.122}$$

$$\hat{\gamma} = 2 - \varepsilon_0/(\varepsilon_0 - 2\varepsilon_1) \tag{13.123}$$

where ε_0 and ε_1 are estimated from the data by the estimators $\varepsilon_0 = \sum_{i=1}^{m} z_i/m$ and $\varepsilon_1 = \sum_{i=1}^{m} (m-i)z_i/\{m(m-1)\}$.[31] Once the value of ε_0 ε_1 are obtained, the estimates of σ and γ are obtained by making use of Equation 13.122 and 13.123. Note that the method of probability weighted moments involves computationally simple expressions for the estimates.

13.3.5.1.3. The Ordered Sample Least Squares Method — A New Approach

The procedure used in maximum likelihood estimation is based on minimizing Equation 13.115. The probability weighted moment estimates are obtained by equation the sample based averages with the theoretical values of the quantity $E[Z(1-G(Z))^s], s = 0, 1,$ where $Z = X - x_0$. A third approach is the OSLS method based on the principle of minimizing the squared distance between the ordered sample and the expected value of the ordered sample. Computer simulations reveal that this can be a more suitable approach for estimating the parameters.

In Appendix X the method for evaluating the mean and the variance of the rth order statistic from a sample of size n is presented. For the GPD the mean and the variance of the rth order statistic can be derived since the probability distribution function is known in closed from. Let x be replaced by z in Equation 13.106 and let $G(z) = u$. Solution for z results in

$$z = G^{-1}(u) = \frac{\sigma}{\gamma}[(1-u)^{-\gamma} - 1] \tag{13.124}$$

Making use of the above equation and Equation X. 62 in Appendix X, the expected value of the rth order statistic Z_r is

$$E(Z_r) = \frac{\sigma}{\gamma}\frac{n!}{(r-1)!(n-r)!}\left[\int_0^1 ((1-u)^{-\gamma} - 1)u^{r-1}(1-u)^{n-r}\mathrm{d}u\right] \tag{13.125}$$

The integral in the above equation can be broken into two parts as follows.

$$E(Z_r) = \frac{\sigma}{\gamma}\frac{n!}{(r-1)!(n-r)!} \times\left[\int_0^1 (1-u)^{-\gamma}u^{r-1}(1-u)^{n-r}\mathrm{d}u - \int_0^1 u^{r-1}(1-u)^{n-r}\mathrm{d}u\right] \tag{13.126}$$

From results presented in Gradshtyn and Ryzhik[49], the expression for $E(Z_r)$ becomes

$$\begin{aligned} E(Z_r) &= \frac{\sigma}{\gamma}\frac{n!}{(r-1)!(n-r)!} \\ &\quad \times\left[\frac{(r-1)!(n-r-\gamma)!}{(n-\gamma)!} - \frac{(r-1)!(n-r)!}{n!}\right] \\ &= \frac{\sigma}{\gamma}\left[\frac{n!(n-r-\gamma)!}{(n-r)!(n-\gamma)!} - 1\right] \\ &= \frac{\sigma}{\gamma}\left[\frac{\Gamma(n+1)\Gamma(n-r-\gamma+1)}{\Gamma(n-r+1)\Gamma(n-\gamma+1)} - 1\right] \end{aligned} \tag{13.127}$$

To calculate the variance of Z_r, we first calculate $E(Z_r^2)$. Making use of Equation 13.124 and Equation X. 65 in Appendix X.3, the expected value of Z_r^2 is

$$E(Z_r^2) = \frac{\sigma^2}{\gamma^2}\frac{n!}{(r-1)!(n-r)!}\left[\int_0^1 ((1-u)^{-\gamma}-1)^2 u^{r-1}(1-u)^{n-r}\mathrm{d}u\right] \tag{13.128}$$

Expanding the square in the integrand of the above equation gives

$$\begin{aligned} E(Z_r^2) &= \frac{\sigma^2}{\gamma^2}\frac{n!}{(r-1)!(n-r)!} \\ &\quad \times\left[\int_0^1 ((1-u)^{-2\gamma} - 2(1-u)^{-\gamma} + 1)u^{r-1}(1-u)^{n-r}\mathrm{d}u\right] \end{aligned} \tag{13.129}$$

Making use of results from Ref. 49 the above integral evaluates to

$$\begin{aligned} E(Z_r^2) &= \frac{\sigma^2}{\gamma^2}\frac{n!}{(n-r)!}\left[\frac{(n-r-2\gamma)!}{(n-2\gamma)!} - \frac{2(n-r-\gamma)!}{(n-\gamma)!} + 1\right] \\ &= \frac{\sigma^2}{\gamma^2}\frac{\Gamma(n+1)}{\Gamma(n-r+1)} \\ &\quad \times\left[\frac{\Gamma(n-r-2\gamma+1)}{\Gamma(n-2\gamma+1)} - \frac{2\Gamma(n-r-\gamma+1)}{\Gamma(n-\gamma+1)} + 1\right] \end{aligned} \tag{13.130}$$

From Equation 13.127 and Equation 13.130 and the relation $\mathrm{Var}(Z_r) = E(Z_r^2) - E^2(Z_r)$, we have

$$\begin{aligned}\mathrm{Var}(Z_r) = {} & \frac{\sigma^2}{\gamma^2}\frac{\Gamma(n+1)}{\Gamma(n-r+1)} \\ & \times\left[\frac{\Gamma(n-r-2\gamma+1)}{\Gamma(n-2\gamma+1)} - \frac{2\Gamma(n-r-\gamma)+1}{\Gamma(n-\gamma+1)} + 1\right] \\ & - \left\{\frac{\sigma}{\gamma}\left[\frac{\Gamma(n+1)\Gamma(n-r-\gamma+1)}{\Gamma(n-r+1)\Gamma(n-\gamma+1)} - 1\right]\right\} \end{aligned} \tag{13.131}$$

Simplifying the above equation results in

$$\begin{aligned}&\mathrm{Var}(Z_r) \\ &= \frac{\sigma^2}{\gamma^2}\left[\frac{\Gamma(n+1)}{\Gamma(n-r+1)}\frac{\Gamma(n-r-2\gamma+1)}{\Gamma(n-2\gamma+1)} - \frac{\Gamma^2(n+1)}{\Gamma^2(n-r+1)}\frac{\Gamma^2(n-r-\gamma+1)}{\Gamma^2(n-r+1)}\right]\end{aligned} \tag{13.132}$$

Letting $Q_r(\gamma) = \dfrac{\Gamma(n+1)}{\Gamma(n-r+1)}\dfrac{\Gamma(n-r-\gamma+1)}{\Gamma(n-\gamma+1)}$ results in

$$E(Z_r) = \mu_r = \frac{\sigma}{\gamma}\{Q_r(\gamma) - 1\} \tag{13.133}$$

$$\mathrm{Var}(Z_r) = \sigma_r^2 = \frac{\sigma^2}{\gamma^2}\{Q_r(2\gamma) - (Q_r(\gamma))^2\} \tag{13.134}$$

A computationally simpler expression can be found for $Q_r(\gamma)$ by making use of the properties of the gamma function. Dividing $Q_r(\gamma)$ by $Q_{r-1}(\gamma)$ we get

$$\frac{Q_r(\gamma)}{Q_{r-1}(\gamma)} = \frac{\dfrac{\Gamma(n+1)}{\Gamma(n-r+1)}\dfrac{\Gamma(n-r-\gamma+1)}{\Gamma(n-\gamma+1)}}{\dfrac{\Gamma(n+1)}{\Gamma(n-r+2)}\dfrac{\Gamma(n-r-\gamma+2)}{\Gamma(n-\gamma+1)}} = \frac{n-r+1}{n-r-\gamma+1} \tag{13.135}$$

Equation 13.135 reduces to

$$Q_r(\gamma) = \prod_{i=1}^{r}\frac{n-i+1}{n-i+1-\gamma} \tag{13.136}$$

To find the least squares estimates of the parameters we write the following nonlinear model for the rth sample order statistic

$$Z_r = E(z_r) + e_r, \qquad r = 1, 2, \ldots, m \tag{13.137}$$

where the error term e_r has a distribution with mean 0 and variance σ_r^2. Since the order statistics are not independent, the errors are also not independent. Because of the nonlinear structure of the model in Equation 13.137 and correlated error,

the method of least squares estimation does not offer a straightforward solution to the estimation problem. Even so, in this study we proceed to sue the OSLS procedure to estimate the parameters.

In Equation 13.133, we note that the scale parameter σ appears linearly whereas the shape parameter γ does not. The least squares estimates are obtained by minimizing the quantity

$$S = \sum_{r=1}^{m} e_r^2 = \sum_{r=1}^{m} [Z_r - \frac{\sigma}{\gamma}(Q_r(\gamma) - 1)]^2 \tag{13.138}$$

Since σ appears linearity in the above expression, minimization can be achieved analytically. Differentiating Equation 13.138 with respect of σ and setting the derivative equal to zero results in

$$2\sum_{r=1}^{m}\left(Z_r - \frac{\sigma}{\gamma}(Q_r(\gamma) - 1 - 1)\right)\left(-\frac{1}{\gamma}(Q_r(\gamma) - 1 - 1)\right) = 0 \tag{13.139}$$

The solution for σ from the above equation is

$$\hat{\sigma}(\gamma) = \gamma \frac{\sum_{r=1}^{m} Z_r(Q_r(\gamma) - 1)}{\sum_{r=1}^{m} (Q_r(\gamma) - 1)^2} \tag{13.140}$$

The expression for $\hat{\sigma}$ is substituted in Equation 13.138 and the resulting expression is minimized with respect to γ. After the substitution the resulting expression is nonlinear and minimization cannot be performed analytically. Using the Nelder–Mead algorithm,[48] the minimization is done numerically. Once the estimate of $\hat{\gamma}$ is obtained, $\hat{\sigma}$ is obtained from Equation 13.140.

Recall that the GPD is being used to approximate the tail of the underlying distribution. Hence, the ordered statistics Z_r, $r = 1, 2, \ldots, m$, from the GPD actually correspond to the ordered statistics $X_{n-m+1} - x_0$, $X_{n-m+2} - x_0, \ldots, X_n - x_0$ from the underlying distribution.

The least squares procedure results in a computationally convenient algorithm. It is emphasized that the minimization of S is carried out only with respect to the single parameter γ. Furthermore, the underlying criterion is based on minimizing the distance between the empirical values and the expected values of the ordered samples. Some numerical comparisons are given in Section 13.3.6.

13.3.5.2. Estimation of Thresholds

The GPD that is estimated from the data is used to approximate the tail of the unknown, underlying distribution. We now show that the threshold is related to the approximating distribution function in a direct manner. With reference to Equation 13.110, let $\hat{\eta}_\alpha$ denote the estimate of the threshold corresponding to

a false alarm probability α. We then have

$$\tilde{F}(\hat{\eta}_\alpha) = 1 - \alpha = 1 - \lambda\left[1 + \frac{\gamma}{\sigma}(\hat{\eta}_\alpha - x_0)\right]^{-1/\gamma} \tag{13.141}$$

Solving for $\hat{\eta}_\alpha$,

$$\hat{\eta}_\alpha = x_0 + \sigma(q^{-\gamma} - 1)/\gamma \tag{13.142}$$

where $\lambda = 1 - F(x_0)$, $q = (1 - \alpha)/\lambda$ and $x_0 = F^{-1}(1 - \lambda)$. For many applications DuMouchel[47] suggests that $\lambda = 0.1$ be used. As will be discussed in the subsequent sections, the optimal value of λ depends on the threshold being estimated. Since the distribution function $F(x)$ is not known, x_0 cannot be determined for a given value of α. Therefore, following common practice, the sample order statistic X_{n-m}, where $m = [\lambda n]$ and [.] denotes the integer part operator, is used as an estimate of x_0.

13.3.6. Numerical Results

13.3.6.1. Characterization of Tail Shape for Known Distributions

We first discuss a method for estimating the parameters of the GPD when the underlying distribution is known. Choose x_0 such that $1 - F(x_0) = 0.1$. Then define the points $p_i \quad i = 1, 2, \ldots, 1000$ by

$$p_i = 0.90005 + 0.0001\,(i - 1) \tag{13.143}$$

Analytically evaluate $x_i = F^{-1}(p_i)$ from the known distribution. Using the 1000 values of x_i, the maximum likelihood estimation, the OSLS, and the probability weighted moments procedures were applied to determine the corresponding γ values for various distributions. The results are given in Table 13.1. The number in parentheses for the Weibull and Lognormal distributions is the value of the shape parameter. For the remaining distributions the number denotes the degrees of freedom. Since σ is a scale parameter, the shape parameter γ best describes the tail shape. For the exponential and the uniform distributions the value of γ can be obtained analytically. The parameter γ is 0 for the exponential distribution and is -1 for the uniform distribution. Since the size of the tail decreases with decreasing γ, the relationship between the tail behavior and the corresponding values of the shape parameter γ can be clearly inferred from this table.

13.3.6.2. Empirical Properties of the Estimators for Known Distributions

Seven distributions with widely differing tail behaviors were chosen in order to investigate the adequacy of the approximation of extreme tails by the GPD and to compare the three estimation procedures. The gamma distribution and Weibull distribution with shape parameter of value 3 have tails that decay faster than those of the exponential PDF. The tails of the chi-square

TABLE 13.1
Tail Parameter γ Describing the Upper Ten Percent of Various Distributions

Distribution	OSLS	ML	PWM
Gaussian	−0.144	−0.151	−0.174
Weibull (3)	−0.163	−0.168	−0.194
Weibull (.67)	0.108	0.129	0.137
Weibull (.5)	0.201	0.265	0.263
Student-T (3)	0.290	0.260	0.261
Student-T (5)	0.132	0.099	0.090
Student-T (8)	0.031	0.006	−0.010
Lognormal (1)	0.232	0.259	0.258
Chi-square (1)	0.030	0.034	0.044
Chi-square (4)	−0.024	−0.033	−0.034
Chi-square (8)	−0.047	−0.058	−0.064

OSLS = Ordered Sample Least Square, ML = Maximum Likelihood, PWM = Probability Weighted Moments.

distribution with 4 degrees of freedom and the Student-T distribution with 8 degrees of freedom are approximately the same as those of the exponential PDF. Finally, the Student-T distribution with 4 degrees of freedom and the Lognormal distribution with shape parameter of value 1 have tails that decay slower than those of the exponential PDF.

Let η and $\hat{\eta}$ denote the true and estimated thresholds, respectively. A Monte Carlo experiment was performed to investigate the normalized bias, $\hat{\eta} - \eta/\eta$ and the normalized mean square error $((\hat{\eta} - \eta)/\eta)^2$ of the proposed threshold estimates. The four sample sizes given by $m = 25$, 50, 100, and 1000 were considered. Each set of samples was obtained by generating n observations and taking the largest $m = 0.1n$ observations. For example, a set of samples of size 25 was obtained by selecting the largest 25 observations from a collection of 250 samples. For each of the four different values of m, $k = 200{,}000/m$ trials were performed for each of the seven distributions. The median of the normalized bias values was computed for each distribution and estimation procedure. The results for $P_{\mathrm{F}} = 10^{-k}$, $k = 2, 3, \ldots, 7$ are given in Table 13.2. Similarly the median of the positive square root of the normalized mean square error are presented in Table 13.3. The results in the two tables differ because the sign of $(\hat{\eta} - \eta)/\eta$ is lost in the normalized root mean square values computed in Table 13.3. Extremely poor estimates for η were obtained in a few of the trials. These poor estimates could severely influence an arithmetic mean of the estimates. To avoid this problem, median values were used in place of arithmetic means.

The empirical results in Table 13.2 indicate that the newly proposed OSLS estimator generally has a smaller normalized bias than the other estimators for

TABLE 13.2
Median of the Normalized Bias Values for Different Percentiles

P_F		10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
			(a) $n\lambda = 25$				
Normal	OSLS	−0.0112	0.0043	−0.0040	−0.0276	−0.0571	−0.0872
Normal	ML	−0.0034	0.0187	0.0328	0.0358	0.0281	0.0137
Normal	PWM	−0.0084	−0.0208	−0.0560	−0.1015	−0.1464	−0.1924
Weibull (3)	OSLS	−0.0048	0.0013	−0.0041	−0.0202	−0.0418	−0.0619
Weibull (3)	ML	0.0039	0.0481	0.0938	0.1374	0.1776	0.2137
Weibull (3)	PWM	−0.0037	−0.0106	−0.0333	−0.0635	−0.0919	−0.1216
t (4)	OSLS	−0.0424	−0.0792	−0.1658	−0.2727	−0.3872	−0.4922
t (4)	ML	−0.0166	−0.1115	−0.2526	−0.4045	−0.5416	−0.6541
t (4)	PWM	−0.0218	−0.0929	−0.2160	−0.3498	−0.4761	−0.5881
t (8)	OSLS	−0.0221	−0.0186	−0.0572	−0.1164	−0.1975	−0.2879
t (8)	ML	−0.0104	−0.0468	−0.1169	−0.2077	−0.3055	−0.4033
t (8)	PWM	−0.0129	−0.0452	−0.1095	−0.2039	−0.3063	−0.4115
Chi-sq (4)	OSLS	−0.0209	−0.0039	0.0241	0.0333	−0.0088	−0.0104
Chi-sq (4)	ML	−0.0037	0.0943	0.2518	0.4571	0.6185	0.8810
Chi-sq (4)	PWM	−0.0144	−0.0205	−0.0334	−0.0576	−0.1254	−0.1624
Lognormal	OSLS	−0.0835	−0.0982	−0.0634	0.0016	0.1007	0.2567
Lognormal	ML	−0.0058	0.1836	0.5932	1.2736	2.4832	4.4947
Lognormal	PWM	−0.0543	−0.0878	−0.0931	−0.0728	−0.0228	0.0639
Pareto (−0.25)	OSLS	−0.0092	0.0208	0.0423	0.0631	0.0780	0.0874
Pareto (−0.25)	ML	−0.0030	0.0523	0.1190	0.1868	0.2479	0.2969
Pareto (−0.25)	PWM	−0.0077	0.0052	0.0121	0.0199	0.0237	0.0278
			(b) $n\lambda = 50$				
Normal	OSLS	0.0036	0.0073	−0.0068	−0.0354	−0.0676	−0.1022
Normal	ML	0.0042	0.0323	0.0467	0.0578	0.0528	0.0380
Normal	PWM	−0.0012	−0.0118	−0.0459	−0.0861	−0.1318	−0.1742
Weibull (3)	OSLS	−0.0022	−0.0007	−0.0133	−0.0337	−0.0571	−0.0838
Weibull (3)	ML	0.0056	0.0500	0.0991	0.1436	0.1847	0.2199
Weibull (3)	PWM	−0.0014	−0.0105	−0.0342	−0.0629	−0.0937	−0.1256
t (4)	OSLS	−0.0147	−0.0646	−0.1800	−0.3209	−0.4501	−0.5063
t (4)	ML	−0.0068	−0.0867	−0.2264	−0.3736	−0.5120	−0.6291
t (4)	PWM	−0.0078	−0.0622	−0.1662	−0.2973	−0.4233	−0.5391
t (8)	OSLS	−0.0062	−0.0222	−0.0841	−0.1723	−0.2694	−0.3703
t (8)	ML	−0.0031	−0.0502	−0.1352	−0.2385	−0.3460	−0.4517
t (8)	PWM	−0.0032	−0.0336	−0.1064	−0.2041	−0.3051	−0.4046
Chi-sq (4)	OSLS	−0.0092	−0.0004	0.0051	0.0060	−0.0498	−0.0686
Chi-sq (4)	ML	0.0115	0.1134	0.2755	0.4775	0.6368	0.9150
Chi-sq (4)	PWM	−0.0041	−0.0087	−0.0191	−0.0407	−0.1123	−0.1488
Lognormal	OSLS	−0.0544	−0.0594	−0.0272	0.0458	0.1573	0.3274
Lognormal	ML	0.0092	0.2177	0.6336	1.3811	2.6197	4.7101
Lognormal	PWM	−0.0302	−0.0391	−0.0185	0.0413	0.1480	0.2977

Continued

TABLE 13.2 Continued

P_F		10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
Pareto (−0.25)	OSLS	−0.0052	0.0100	0.0214	0.0326	0.0404	0.0448
Pareto (−0.25)	ML	0.0005	0.0463	0.1011	0.1560	0.2003	0.2357
Pareto (−0.25)	PWM	−0.0050	−0.0018	−0.0012	−0.0019	−0.0023	−0.0012
			(c) $n\lambda = 100$				
Normal	OSLS	0.0017	−0.0016	−0.0253	−0.0637	−0.1040	−0.1464
Normal	ML	0.0068	0.0263	0.0306	0.0229	0.0063	−0.0185
Normal	PWM	0.0018	−0.0181	−0.0549	−0.1022	−0.1524	−0.1986
Weibull (3)	OSLS	0.0005	−0.0017	−0.0164	−0.0376	−0.0624	−0.0888
Weibull (3)	ML	0.0037	0.0270	0.0564	0.0840	0.1003	0.1158
Weibull (3)	PWM	0.0004	−0.0095	−0.0320	−0.0607	−0.0918	−0.1220
t (4)	OSLS	−0.0064	−0.0441	−0.1421	−0.2680	−0.3922	−0.5031
t (4)	ML	−0.0004	−0.0564	−0.1650	−0.2907	−0.4174	−0.5354
t (4)	PWM	−0.0003	−0.0478	−0.1403	−0.2636	−0.3809	−0.4949
t (8)	OSLS	−0.0024	−0.0134	−0.0751	−0.1606	−0.2578	−0.3548
t (8)	ML	0.0011	−0.0342	−0.1145	−0.2123	−0.3157	−0.4216
t (8)	PWM	0.0013	−0.0271	−0.0955	−0.1888	−0.2892	−0.3916
Chi-sq (4)	OSLS	−0.0032	−0.0028	−0.0077	−0.0198	−0.0841	−0.1111
Chi-sq (4)	ML	0.0175	0.1189	0.2655	0.4581	0.5917	0.8298
Chi-sq (4)	PWM	−0.0004	−0.0089	−0.0238	−0.0448	−0.1143	−0.1520
Lognormal	OSLS	−0.0159	−0.0542	−0.0876	−0.1089	−0.0940	−0.0617
Lognormal	ML	−0.0111	−0.0251	−0.0068	0.0536	0.1499	0.3104
Lognormal	PWM	−0.0165	−0.0210	0.0141	0.0924	0.2315	0.3965
Pareto (−0.25)	OSLS	−0.0023	0.0109	0.0255	0.0350	0.0419	0.0471
Pareto (−0.25)	ML	0.0033	0.0544	0.1170	0.1739	0.2215	0.2611
Pareto (−0.25)	PWM	−0.0014	0.0004	0.0052	0.0084	0.0112	0.0129
			(d) $n\lambda = 1000$				
Normal	OSLS	0.0035	−0.0013	−0.0244	−0.0613	−0.1010	−0.1432
Normal	ML	0.0059	0.0017	−0.0259	−0.0626	−0.1075	−0.1476
Normal	PWM	0.0028	−0.0192	−0.0586	−0.1064	−0.1560	−0.2016
Weibull (3)	OSLS	0.0013	−0.0023	−0.0175	−0.0381	−0.0627	−0.0885
Weibull (3)	ML	0.0020	−0.0018	−0.0159	−0.0386	−0.0641	0.0909
Weibull (3)	PWM	0.0010	−0.0092	−0.0297	−0.0578	−0.0880	−0.1192
t (4)	OSLS	0.0058	−0.0044	−0.0605	−0.1574	−0.2690	−0.3715
t (4)	ML	0.0141	−0.0137	−0.1018	−0.2167	−0.3326	−0.4406
t (4)	PWM	0.0141	−0.0176	−0.1104	−0.2277	−0.3479	−0.4598
t (8)	OSLS	0.0033	−0.0021	−0.0452	−0.1167	−0.2001	−0.2896
t (8)	ML	0.0070	−0.0117	−0.0664	−0.1464	−0.2404	−0.3382
t (8)	PWM	0.0045	−0.0219	−0.0896	−0.1825	−0.2857	−0.3862
Chi-sq (4)	OSLS	0.0003	0.0012	−0.0057	−0.0167	−0.0826	−0.1107
Chi-sq (4)	ML	0.0012	−0.0021	−0.0152	−0.0354	−0.1026	−0.1349
Chi-sq (4)	PWM	0.0006	−0.0011	−0.0080	−0.0263	−0.0934	−0.1211

Continued

TABLE 13.2 Continued

P_F		10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
Lognormal	OSLS	−0.0038	−0.0221	−0.0259	0.0055	0.0646	0.1638
Lognormal	ML	−0.0098	0.0063	0.0616	0.1767	0.3495	0.5999
Lognormal	PWM	−0.0128	−0.0004	0.0567	0.1683	0.3400	0.5771
Pareto (−0.25)	OSLS	0.0002	0.0002	0.0012	0.0007	0.0003	0.0000
Pareto (−0.25)	ML	−0.0002	−0.0010	−0.0044	−0.0061	−0.0081	−0.0094
Pareto (−0.25)	PWM	0.0003	−0.0011	−0.0007	−0.0006	−0.0035	−0.0038

OSLS = Ordered Sample Least Square, ML = Maximum Likelihood, PWM = Probability Weighted Moments.

small or moderate sample sizes. Overall the second smallest normalized bias is achieved by the probability weighted moments method. The maximum likelihood estimator has the largest normalized bias when $P_F \geq 10^{-5}$, especially for the long tailed distributions. The normalized bias all three estimators decrease as the sample size increases. When the parent distribution is GPD, all three estimators perform very well. Even so, the ordered sample least square estimator outperforms the others. The relatively strong performance for the GPD is explained as follows. The extreme value theory is based on the premise that tails of smooth continuous distributions tend towards the GPD. For the GPD, this premise is exactly satisfied. Hence, the corresponding performance is noticeably better than for other distributions.

The results for the median of the normalized root mean square error are surprising. The maximum likelihood estimator is known to be asymptotically efficient. This is always true when the samples are drawn for the underlying distribution (in our case from the GPD). This property of the maximum likelihood estimator can be observed in Table 13.3 when $m = 1000$ but not for smaller sample sizes. Although the OSLS method has a smaller normalized root mean square error in many cases, there is no clear winner with respect to this criterion.

From the empirical results based on a limited number of distributions and sample sizes, it is not easy to make a strong recommendation as to which method to use in practice. However, it terms of the normalized bias, the ordered samples least squares estimator appears to perform better than the other estimators in estimating the large thresholds when $P_F \leq 10^{-6}$. In any event, it is seen that the extreme value theory can be used successfully to determine threshold values, when the false alarm probability is very small.

Two practical advantages of estimation based on extreme value theory are: (1) When there is a constraint on the number of samples, the thresholds obtained from extreme value theory are expected to be closer to the true thresholds than those obtained by conventional Monte Carlo techniques. However, in both

TABLE 13.3
Median RMS Errors for Various Percentiles

P_F		10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
			(a) $n\lambda = 25$				
Normal	OSLS	0.0558	0.1127	0.2022	0.2825	0.3507	0.4044
Normal	ML	0.0558	0.0909	0.1459	0.2057	0.2588	0.3070
Normal	PWM	0.0559	0.1215	0.2121	0.2920	0.3586	0.4117
Weibull (3)	OSLS	0.0257	0.0577	0.1089	0.1580	0.2031	0.2415
Weibull (3)	ML	0.0258	0.0531	0.0950	0.1378	0.1780	0.2139
Weibull (3)	PWM	0.0256	0.0624	0.1149	0.1659	0.2110	0.2495
t (4)	OSLS	0.1069	0.2261	0.4160	0.5989	0.7397	0.8405
t (4)	ML	0.1051	0.2353	0.4157	0.5812	0.7127	0.8097
t (4)	PWM	0.1019	0.2329	0.4213	0.5956	0.7368	0.8344
t (8)	OSLS	0.0781	0.1666	0.3073	0.4455	0.5701	0.6730
t (8)	ML	0.0779	0.1493	0.2554	0.3648	0.4689	0.5649
t (8)	PWM	0.0775	0.1752	0.3180	0.4544	0.5783	0.6787
Chi-sq (4)	OSLS	0.0610	0.1313	0.2441	0.3592	0.4650	0.5455
Chi-sq (4)	ML	0.0721	0.2179	0.4459	0.7901	1.1783	1.7789
Chi-sq (4)	PWM	0.0592	0.1384	0.2500	0.3622	0.4666	0.5446
Lognormal	OSLS	0.1335	0.2452	0.4362	0.6271	0.7785	0.8785
Lognormal	ML	0.1439	0.4007	0.7303	1.4149	2.7312	5.0774
Lognormal	PWM	0.1260	0.2582	0.4463	0.6281	0.7737	0.8705
Pareto (−0.25)	OSLS	0.0409	0.0787	0.1348	0.1752	0.2017	0.219
Pareto (−0.25)	ML	0.0402	0.0763	0.1419	0.2075	0.2640	0.3127
Pareto (−0.25)	PWM	0.0411	0.0866	0.1430	0.1817	0.2084	0.2240
			(b) $n\lambda = 50$				
Normal	OSLS	0.0401	0.0772	0.1391	0.1981	0.2548	0.3042
Normal	ML	0.0394	0.0689	0.1122	0.1559	0.1959	0.2328
Normal	PWM	0.0399	0.0865	0.1530	0.2192	0.2759	0.3273
Weibull (3)	OSLS	0.0180	0.0393	0.0743	0.1135	0.1511	0.1854
Weibull (3)	ML	0.0185	0.0509	0.0997	0.1447	0.1859	0.2214
Weibull (3)	PWM	0.0180	0.0442	0.0852	0.1263	0.1661	0.2017
t (4)	OSLS	0.0779	0.1826	0.3506	0.5179	0.6633	0.7724
t (4)	ML	0.0768	0.1910	0.3602	0.5244	0.6688	0.7762
t (4)	PWM	0.0760	0.1778	0.3332	0.4899	0.6303	0.7386
t (8)	OSLS	0.0561	0.1228	0.2316	0.3503	0.4666	0.5698
t (8)	ML	0.0553	0.1219	0.2226	0.3385	0.4504	0.5529
t (8)	PWM	0.0554	0.1306	0.2405	0.3613	0.4793	0.5807
Chi-sq (4)	OSLS	0.0431	0.0890	0.1678	0.2509	0.3351	0.4109
Chi-sq (4)	ML	0.0489	0.1661	0.3386	0.5487	0.7664	1.1112
Chi-sq (4)	PWM	0.0426	0.0939	0.1747	0.2584	0.3431	0.4185
Lognormal	OSLS	0.0975	0.1834	0.3439	0.5155	0.6660	0.7990
Lognormal	ML	0.0993	0.3381	0.6769	1.3921	2.6297	4.7240

Continued

TABLE 13.3 Continued

P_F		10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
Lognormal	PWM	0.0864	0.1954	0.3510	0.5143	0.6621	0.8012
Pareto (−0.25)	OSLS	0.0289	0.0534	0.0890	0.1162	0.1346	0.1486
Pareto (−0.25)	ML	0.0284	0.0602	0.1149	0.1675	0.2084	0.2417
Pareto (−0.25)	PWM	0.0293	0.0616	0.1032	0.1320	0.1533	0.1666
			(c) $n\lambda = 100$				
Normal	OSLS	0.0284	0.0522	0.0964	0.1414	0.1863	0.2305
Normal	ML	0.0290	0.0517	0.0840	0.1123	0.1433	0.1689
Normal	PWM	0.0281	0.0584	0.1090	0.1635	0.2134	0.2585
Weibull (3)	OSLS	0.0128	0.0273	0.0529	0.0811	0.1101	0.1378
Weibull (3)	ML	0.0131	0.0389	0.0790	0.1202	0.1570	0.1868
Weibull (3)	PWM	0.0126	0.0312	0.0622	0.0942	0.1284	0.1590
t (4)	OSLS	0.0550	0.1400	0.2801	0.4336	0.5739	0.6909
t (4)	ML	0.0525	0.1377	0.2716	0.4165	0.5497	0.6627
t (4)	PWM	0.0527	0.1334	0.2619	0.4046	0.5323	0.6469
t (8)	OSLS	0.0386	0.0914	0.1770	0.2761	0.3758	0.4732
t (8)	ML	0.0388	0.0896	0.1735	0.2710	0.3734	0.4763
t (8)	PWM	0.0384	0.0869	0.1727	0.2750	0.3777	0.4817
Chi-sq (4)	OSLS	0.0287	0.0649	0.1264	0.1932	0.2699	0.3373
Chi-sq (4)	ML	0.0350	0.1437	0.2959	0.4688	0.6092	0.8592
Chi-sq (4)	PWM	0.0283	0.0686	0.1289	0.1948	0.2730	0.3383
Lognormal	OSLS	0.0683	0.1527	0.2794	0.4174	0.5299	0.6290
Lognormal	ML	0.0652	0.1515	0.2690	0.4039	0.5465	0.6769
Lognormal	PWM	0.0647	0.1417	0.2519	0.3805	0.5218	0.6710
Pareto (−0.25)	OSLS	0.0201	0.0372	0.0637	0.0845	0.0997	0.1110
Pareto (−0.25)	ML	0.0197	0.0569	0.1192	0.1746	0.2221	0.2613
Pareto (−0.25)	PWM	0.0201	0.0434	0.0718	0.0952	0.1108	0.1220
			(d) $n\lambda = 1000$				
Normal	OSLS	0.0077	0.0182	0.0373	0.0643	0.1017	0.1440
Normal	ML	0.0087	0.0160	0.0362	0.0632	0.1075	0.1476
Normal	PWM	0.0081	0.0247	0.0586	0.1064	0.1560	0.2016
Weibull (3)	OSLS	0.0037	0.0086	0.0194	0.0393	0.0630	0.0890
Weibull (3)	ML	0.0040	0.0078	0.0191	0.0397	0.0649	0.0909
Weibull (3)	PWM	0.0036	0.0108	0.0300	0.0578	0.0880	0.1192
t (4)	OSLS	0.0203	0.0534	0.1383	0.2476	0.3717	0.4763
t (4)	ML	0.0213	0.0447	0.1083	0.2168	0.3326	0.4406
t (4)	PWM	0.0213	0.0499	0.1207	0.2306	0.3479	0.4598
t (8)	OSLS	0.0135	0.0298	0.0726	0.1376	0.2121	0.3018
t (8)	ML	0.0129	0.0272	0.0750	0.1518	0.2436	0.3406
t (8)	PWM	0.0129	0.0349	0.0939	0.1830	0.2863	0.3863
Chi-sq (4)	OSLS	0.0104	0.0207	0.0362	0.0588	0.1094	0.1408
Chi-sq (4)	ML	0.0099	0.0192	0.0363	0.0589	0.1095	0.1429

Continued

TABLE 13.3 Continued

P_F		10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-7}
Chi-sq (4)	PWM	0.0100	0.0211	0.0400	0.0602	0.1103	0.1433
Lognormal	OSLS	0.0206	0.0528	0.1222	0.1836	0.2429	0.3276
Lognormal	ML	0.0195	0.0434	0.0984	0.2012	0.3581	0.5999
Lognormal	PWM	0.0201	0.0410	0.0927	0.1919	0.3445	0.5770
Pareto (−0.25)	OSLS	0.0061	0.0101	0.0158	0.0213	0.0247	0.0278
Pareto (−0.25)	ML	0.0063	0.0092	0.0154	0.0198	0.0243	0.0268
Pareto (−0.25)	PWM	0.0065	0.0126	0.0222	0.0306	0.0375	0.0428

OSLS = Ordered Sample Least Square, ML = Maximum Likelihood, PWM = Probability Weighted Moments.

techniques an increase in sample size offers greater accuracy in estimating thresholds. (2) Because the estimate of the tail of the underlying distribution is in closed form, estimation can be made for thresholds corresponding to extremely small false alarm probabilities independent of the sample size. In experiments with fixed amounts of data, this is an important advantage.

13.3.6.3. Effect of the Choice of λ on the Threshold Estimates

As was mentioned previously, only those samples exceeding x_0 are used in estimating the GPD parameters. The value of x_0 is determined by λ. The results presented in Table 13.2 and Table 13.3 were obtained by means of Monte Carlo experiments where $\lambda = 0.1$ was used independent of the value of P_F for which the threshold was being estimated. When the false alarm probability was extremely small, the bias and root mean square errors were quite large for some distributions. The smaller the value of λ, the better will be the GPD approximation over the extreme tail being approximated. When λ is chosen too large, a better fit is found for that portion of the distribution closer to the center at the expense of lesser accuracy in the extreme tail. Of course, there is a tradeoff between the choice of λ and the number of data samples available for determining the parameters of the GPD.

In our application the major objective is to approximate the extreme tails corresponding to thresholds of 10^{-5} or smaller. Consequently, we explored the implications of selecting values less than 0.1 for λ. To accomplish this, we obtained the theoretical values of x_i for the standard Normal and Lognormal distributions corresponding to $F^{-1}(p_i)$ where $p_i = (i - 0.5/n)\ \ i = 1, 2, \ldots, n$, and $n = 1000$ and 10,000, respectively. These two distributions are chosen because they represent extremes: The Normal distribution is light tailed (decays faster) while the Lognormal is a heavy tailed (decays slower) distribution.

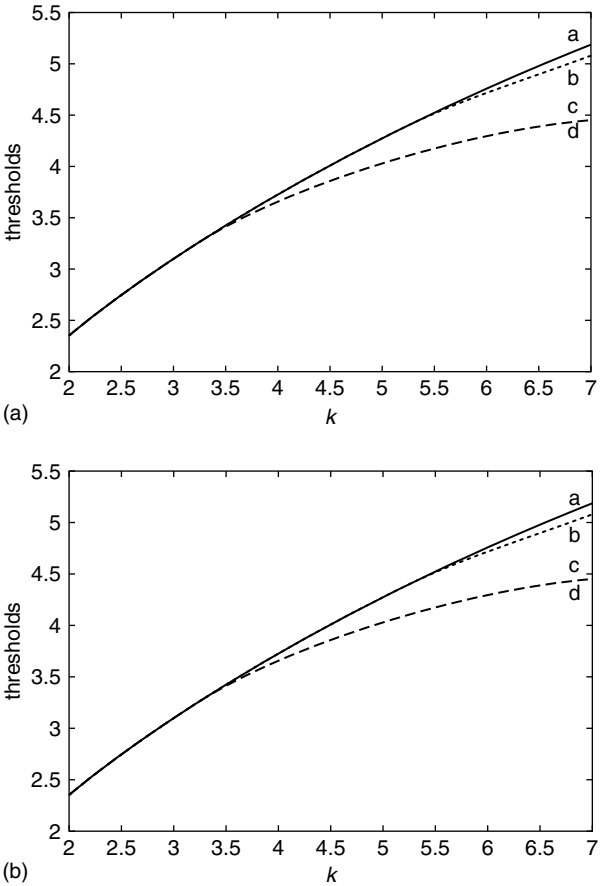

FIGURE 13.10 Normal distribution. Data points correspond to $k = 2, 3, 7$. a: True, b: $\lambda = 0.01$, c: $\lambda = 0.05$, d: $\lambda = 0.10$. (a) $n = 10000$ thresholds for $P_F = 10^{-k}$. (b) $n = 1000$ Thresholds for $P_F = 10^{-k}$.

The number of the x_i samples used to determine the parameters of the GPD is given by λn. The parameters were estimated using the OSLS procedure for values of λ equal to 0.1, 0.05, and 0.01. The resulting GPDs were then used to determine the thresholds for false alarm probabilities given by $P_F = 10^{-k}$ where $k = 2, 3,\ldots,7$. These results are presented in Figure 13.10 and Figure 13.11, where both the theoretical and approximated thresholds are plotted as a function of k for (A) Normal distribution ($n = 10{,}000$), (B) Normal distribution ($n = 1000$), (C) Lognormal distribution ($n = 10{,}000$), (D) Lognormal distribution ($n = 1000$). For $k \geq 5$, it is seen that $\lambda = 0.01$ (curve b) appears to be the best choice for approximating the thresholds. The best results were obtained with $n = 10{,}000$. However, good results were obtained with $n = 1000$.

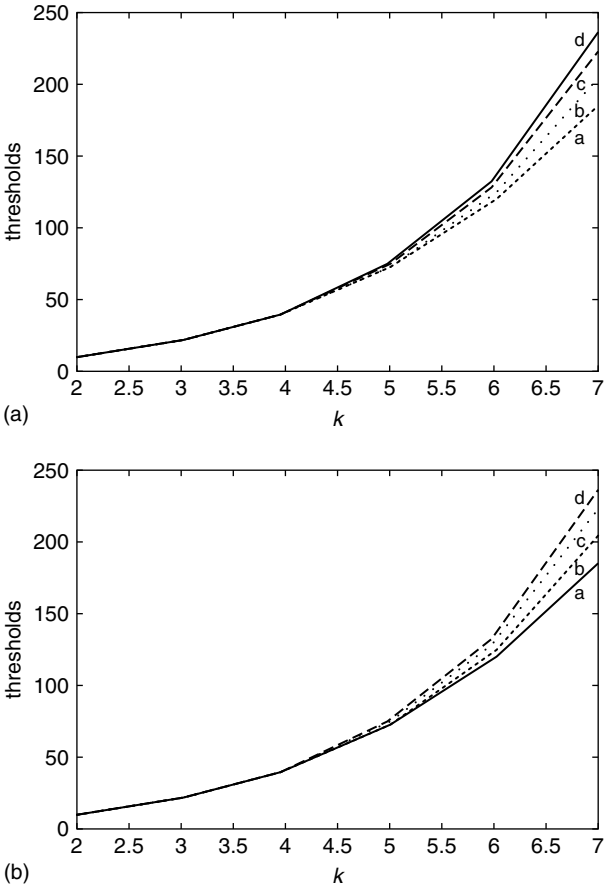

FIGURE 13.11 Lognormal distribution. Data points correspond to $k = 2, 3, 7$. a: True, b: $\lambda = 0.01$, c: $\lambda = 0.05$, d: $\lambda = 0.10$. (a) $n = 10,000$ thresholds for $P_F = 10^{-k}$. (b) $n = 1000$ Thresholds for $P_F = 10^{-k}$.

13.3.7. Examples

13.3.7.1. Known Distribution Case

To evaluate the accuracy of the threshold value estimates, 10,000 random samples were generated from the Gaussian and Lognormal distributions and the upper tails of these two distributions were modeled as Generalized Pareto. In Section 13.3.4.1 and Section 13.3.4.3, theoretical values given by $x_i = F^{-1}(p_i)$ were used to estimate the tail. In this section randomly generated samples are used in place of the theoretical values. Choosing $\lambda = 0.01$, the theoretical thresholds of the Gaussian distribution for $P_F = 10^{-k}, k = 2, 3, \ldots, 7$ are 2.326, 3.090, 3.719, 4.265, 4.753, and 5.199, respectively. The thresholds estimated,

based on one set of random samples are 2.315, 3.223, 3.847, 4.370, 4.855, and 5.292. For the Lognormal distribution the theoretical thresholds corresponding to $P_F = 10^{-k}$, $k = 2, 3, \ldots, 7$ are 10.240, 21.982, 41.224, 71.157, 115.981, and 181.152. Once again, using $\lambda = 0.01$, the thresholds estimated based on one set of random samples are 10.449, 22.862, 42.473, 69.216, 112.229, and 183.495. Note that the estimated results are very close to the true thresholds. We note here that these results were obtained on the basis of one set of observations from the two known distributions, corresponding to a particular seed value. For a different set of samples the estimates will be different depending on the tail behavior of that set of samples. But, unless the samples are really not a true representative of the distribution from which they are drawn, we expect that the estimates based on different samples should give threshold values yielding false alarm probabilities close to the design value.

13.3.7.2. An Unknown Distribution Case

In the previous section the underlying distributions were known to us and the estimates based on the extreme value theory were encouraging for both light and heavy tail behavior. In this example, we take a nonGaussian problem where the underlying distribution is unknown.

The two hypotheses characterizing the detection problem are given in Equation 13.1 and Equation 13.2. We consider the weak signal case for which the clutter is much stronger than the background noise. The LOD[8] has been shown to be suitable for the weak signal detection problem. Under hypothesis H_1, the signal is denoted by θs_i, where θ is a measure of the signal strength. For a deterministic signal and a given set of observations $\underline{r} = [r_1, r_2 \ldots, r_N]^T$ the LOD performs the ratio test

$$T_{\mathrm{LOD}}(\underline{r}) = \frac{\left.\dfrac{\partial P_{\underline{\mathrm{R}}|H_1}(\underline{r}|H_1)}{\partial \theta}\right|_{\theta=0}}{\partial P_{\underline{\mathrm{R}}|H_0}(\underline{r}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_k \tag{13.144}$$

where $P_{\underline{\mathrm{R}}|H_i}(\underline{r}|H_i)$ is the joint PDF of $r_1, r_2, \ldots r_N$ under hypothesis $H_i : i = 0, 1$.

Martinez, Swaszek, and Thomas[22] studied the locally optimal detection problem for nonGaussian distributions and considered the bivariate Laplace distribution as an example. In this section we illustrate the procedure for determining the thresholds of a LOD based on $N = 2$ and the received samples under H_0 having the bivariate Laplace distribution given by

$$f_{\underline{\mathrm{R}}}(r_1, r_2) = \frac{1}{2\pi|\mathbf{M}|^{1/2}} K_0[(2\underline{r}^T\mathbf{M}^{-1}\underline{r})^{1/2}] \tag{13.145}$$

when $\mathbf{M}$ is the covariance matrix for the two samples, $|\mathbf{M}|$ denotes its determinant, $\underline{r}^T\mathbf{M}^{-1}\underline{r}$ is equal to $(r_1^2 - 2\rho r_1 r_2 + r_2^2)/(1 - \rho^2)$, ρ is the correlation coefficient between R_1 and R_2, and $K_0(\cdot)$ is the modified Bessel function of the

second kind of zero order. The resulting LOD statistic is[22]

$$T_{\mathrm{LOD}}(r_1, r_2)\left(\frac{2}{\underline{r}^{\mathrm{T}}\mathbf{M}^{-1}\underline{r}}\right)^{1/2}\frac{K_1[(2\underline{r}^{\mathrm{T}}\mathbf{M}^{-1}\underline{r})^{1/2}]}{K_0[(2\underline{r}^{\mathrm{T}}\mathbf{M}^{-1}\underline{r})^{1/2}]} \times \underline{s}^{\mathrm{T}}\mathbf{M}^{-1}\underline{r} \qquad (13.146)$$

where $\underline{s} = (s_1, s_2)^{\mathrm{T}}$, $\underline{s}^{\mathrm{T}}\mathbf{M}^{-1}\underline{r} = (r_1 - \rho r_2)s_1 + (r_2 - \rho r_1)s_2$ and $K_1(.)$ is the modified Bessel function of the second kind of first order. s_1 and s_2 are the known signal levels. In this example we take $s_1 = 1$ and $s_2 = -1$. Because of the complexity of $T_{\mathrm{LOD}}(.)$, it is not possible to determine a closed form expression for its PDF.

In many applications in radar, thresholds have to be set to achieve desired false alarm probabilities based on a sample size with orders of magnitude less than $10/P_{\mathrm{F}}$. As will be pointed out later, the statistic in Equation 13.146 represents a worst case situation as our simulations indicate that the variance of the test statistic is extremely large. To investigate the reliability of the thresholds estimated based on extreme value theory with smaller sample sizes, 10,000 pairs of observations (r_1, r_2) were generated form the bivariate Laplace distribution given in Equation 13.145, with $\rho = 0.90$. The values of $T_{\mathrm{LOD}}(r_1, r_2)$ were computed for each pair and sorted in increasing order. Corresponding to $\lambda = 0.01$, the largest 100 values of the underlying statistic (the top one per cent) were selected to fit the GPD. This experiment was repeated 250 times. The threshold corresponding to a certain false alarm probability P_{F} of the distribution of the statistic $T_{\mathrm{LOD}}(r_1, r_2)$ is estimated from Equation 13.142 as $\tilde{\eta}_\alpha = x_0 + \frac{\hat{\sigma}}{\hat{\gamma}}[(\frac{1-\alpha}{0.01})^{-\hat{\gamma}} - 1]$ where x_0 is the $9900th$ largest value of the statistic. Thresholds were estimated for false alarm probabilities $P_{\mathrm{F}} = 10^{-k}$, $k = 2, \ldots, 7$ for each repetition of the experiment. Histograms of these threshold values are shown in Figure 13.12 to Figure 13.17, for the different P_{F}s. To give a better appreciation for the range of values, the bins are not necessarily of equal width. The histograms give an indication of the spread in the threshold values depending

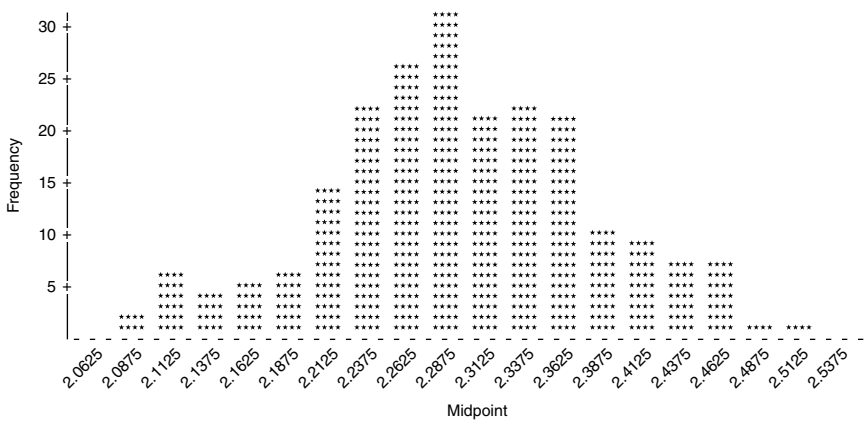

FIGURE 13.12 Histogram of threshold values, $P_{\mathrm{F}} = 10^{-2}$.

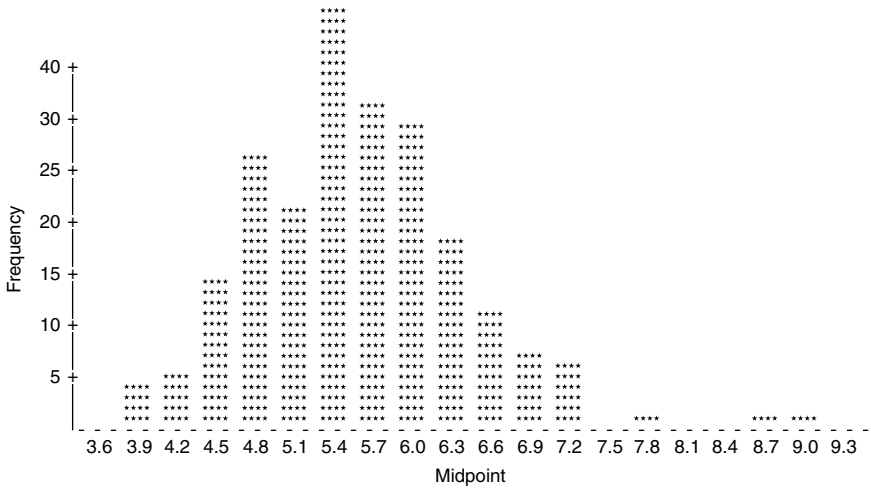

FIGURE 13.13 Histogram of threshold values, $P_F = 10^{-3}$.

on the particular samples collected. From the histograms corresponding to false alarm probabilities of 10^{-2}, 10^{-3}, and 10^{-4} we can see that the threshold estimates obtained on the basis of even one set of samples is likely to approximately yield the desired P_F. Since the underlying distribution of $T_{LOD}(.)$ is unknown, one measure of the accuracy of the estimate is the extent to which most of the estimates fall in one bin of the histogram. Also, we can see that there is negligible overlap between the estimated threshold values in the histograms for the three different values of P_F. This supports the claim that the estimated

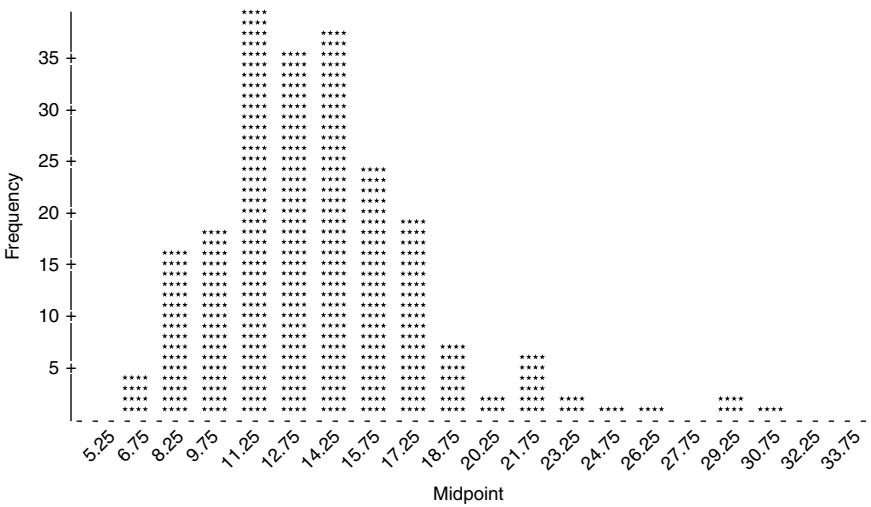

FIGURE 13.14 Histogram of threshold values, $P_F = 10^{-4}$.

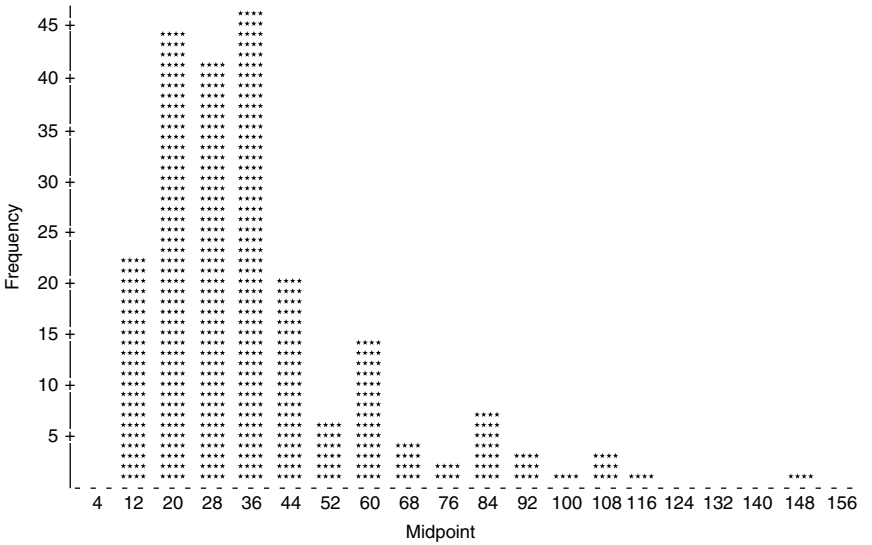

FIGURE 13.15 Histogram of threshold values, $P_F = 10^{-5}$.

threshold is likely to yield a false alarm probability of the same order as the desired P_F. There is a higher overlap in the thresholds of the histograms for $P_F = 10^{-5}$, 10^{-6}, and 10^{-7}. Also, there is much higher spread in the threshold values estimated. Based on the excellent results obtained for the same choices of P_F in the known cases of the previous section, these results are surprising. However, it is explained as follows. The γ values of the GPD estimated for the different repetitions of this experiment lie in the range $0.45 - 0.55$. This represents an extremely heavy tailed distribution. From Table 13.1 we can see that the Lognormal distribution, which is quite a heavy tailed distribution, has $\gamma = 0.232$. The heavy tailed nature of the detector statistic can also be observed

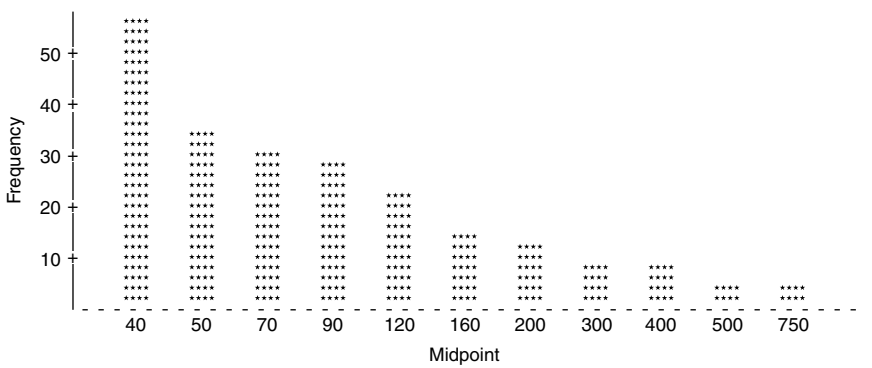

FIGURE 13.16 Histogram of threshold values, $P_F = 10^{-6}$.

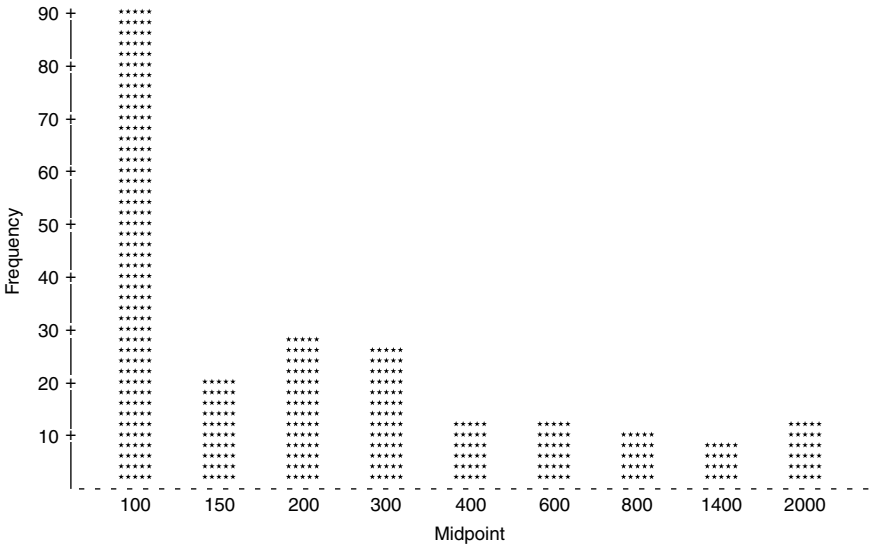

FIGURE 13.17 Histogram of threshold values, $P_F = 10^{-7}$.

by comparing the large threshold values seen in the histograms with the corresponding thresholds of the Gaussian and the Lognormal distributions. The variance of the GPD is given by

$$\begin{aligned} \mathrm{Var}(X) &= \frac{\sigma^2}{(1-\gamma^2)(1-2\gamma)} \qquad \gamma < 0.5 \\ &= \infty \qquad\qquad\qquad\qquad \gamma \geq 0.5 \end{aligned} \tag{13.147}$$

Thus, the bivariate Laplace results in a very highly fluctuating statistic with an extremely large variance. As such, it represents a "worst case" situation for empirically determining the threshold. A much larger sample size is needed to obtain reliable threshold estimates because of the exceedingly large tail of the underlying distribution.

In general, an indication of how heavy the true tail may be for an unknown distribution is given by the estimate of γ for the GPD. When an extremely heavy tail is indicated, another strategy for estimating the thresholds when P_F is very small is to choose the median value of the thresholds estimated when the experiment is repeated a specified number of times with 10,000 samples in each repetition. The choice of the median as the estimator ensures that very large and very small values do not affect the results. For the present example, we chose to repeat the 250 trials three times. By counting the number of estimates that fell into the bins centered at 20, 28, and 36 for $P_F = 10^{-5}$; 40, 50, 70, and 90 for $P_F = 10^{-6}$; 100 and 150 for $P_F = 10^{-7}$, it was found that 88 percent of the estimates fell into these bins. Thus, even for this extremely large tailed example,

we believe that use of the GPD has allowed us to estimate useful values for the thresholds with sample sizes much smaller than $10/P_F$.

13.4. PERFORMANCE OF THE LOD FOR MULTIVARIATE STUDENT-*T* AND *K*-DISTRIBUTED DISTURBANCES

In radar problems involving weak signal applications, it is found that the large returns due to clutter can lead to a small signal to disturbance ratio. When the density function of the clutter exhibits an extended tail behavior relative to the Gaussian PDF, the PDF of the disturbance can no longer be modeled as Gaussian. The significance of a nonGaussian PDF with an extended tail is that many more large returns result than would be the case for a Gaussian PDF having the same variance. Hence, there is a need to be able to model nonGaussian random processes.

The multivariate Student-*T* distribution is a member of the class of joint PDFs arising from SIRPs. SIRPs are explained in Section 13.2. When an SIRP is sampled at N instants in time, the resulting vector is said to be a spherically invariant random vector (SIRV). The theory of SIRPs offers a way to model the joint density function on these N samples where the correlation between the individual random variables in the vector is accounted for. With this approach LOD structures can be derived for nonGaussian disturbances without the need to assume that the random variables are statistically independent. In this section we analyze the performance of the LOD for the known signal problem when the background disturbance consisting of clutter and noise can be modeled as having a multivariate Student-*T* distribution or a multivariate *K*-distribution.

13.4.1. The Multivariate Student-*T* Distribution

A convenient procedure for deriving the multivariate Student-*T* distribution from the representation theorem[28] is discussed in this section. Let the random vector $\underline{X}$ have a multivariate Gaussian distribution with zero mean and covariance matrix $\mathbf{M}$. The zero mean assumption will not affect the generality of the results that follow. The joint density function on the elements of $\underline{X}$ is given by

$$J_{\underline{X}}(\underline{x}) = \frac{1}{(2\pi)^N |\mathbf{M}|^{1/2}} e^{-(\underline{x}^T \mathbf{M}^{-1} \underline{x}/2)} \tag{13.148}$$

where the vector $\underline{X}$ has $2N$ elements from N inphase and N quadrature samples.

Consider the vector $\underline{W} = \underline{X}/\nu$, where ν is a nonnegative random variable statistically independent of X. Let $\underline{w}^T \mathbf{M}^{-1} \underline{w}$ be noted by the variable p. Then, the conditional density function of the vector $\underline{W}$ given ν can be written as

$$f_{\underline{W}}(\underline{w}|\nu) = \frac{1}{(2\pi)^N |\mathbf{M}|^{1/2}} \nu^{2N} e^{\frac{-\nu^2 p}{2}} \tag{13.149}$$

The unconditional density function of $\underline{W}$ is given by

$$f_{\underline{W}}(\underline{w}) = \int_0^\infty f_{\underline{w}}(\underline{w}|\nu) f_\nu(\nu) \mathrm{d}\nu \tag{13.150}$$

where $f_\nu(\nu)$ is the PDF of the random variables ν. Because $\underline{X}$ and ν are statistically independent, it follows that

$$E(\underline{W}) = E\left(\frac{\underline{X}}{\nu}\right) = E(\underline{X})E(\nu^{-1}) = \underline{0} \tag{13.151}$$

$$E(\underline{W}\,\underline{\dot{W}}^{\mathrm{T}}) = E(\underline{X}\,\underline{X}^{\mathrm{T}})E(\nu^{-2}) = E(\nu^{-1})\mathbf{M} \tag{13.152}$$

It can be seen from the above equation that the covariances of the elements of the vector $\underline{W}$ can be adjusted by appropriate choice of $E(\nu^{-2})$.

With respect to Equation 13.150, let $f_\nu(\nu)$ be the generalized chi PDF given by

$$f_\nu(\nu) = 2\frac{\nu^{2\beta-1}\mathrm{e}^{-\alpha\nu^2}\alpha^\beta}{\Gamma(\beta)} \qquad \nu \geq 0 \tag{13.153}$$

From Equation 13.153, $E(\nu^{-2})$ can be calculated. Specifically,

$$E(\nu^{-2}) = \int_0^\infty 2\nu^{-2}\frac{\nu^{2\beta-1}\mathrm{e}^{-\alpha\nu^2}\alpha^\beta}{\Gamma(\beta)}\mathrm{d}\nu = \int_0^\infty 2\frac{\nu^{2\beta-3}\mathrm{e}^{-\alpha\nu^2}\alpha^\beta}{\Gamma(\beta)}\mathrm{d}\nu \tag{13.154}$$

Letting $\alpha\nu^2 = x$ in the above equation we get

$$E(\nu^{-2}) = \alpha\int_0^\infty \frac{x^{\beta-2}\mathrm{e}^{-x}\,\mathrm{d}x}{\Gamma(\beta)} = \alpha\frac{\Gamma(\beta-1)}{\Gamma(\beta)} = \frac{\alpha}{\beta-1} \tag{13.155}$$

If we let $\alpha = \beta - 1$, then the generalized chi PDF in Equation 13.153 is such that $E(\nu^{-2}) = 1$ irrespective of the choice for the parameter β. Then the generalized chi PDF takes the form

$$f_\nu(\nu) = \frac{2\nu^{2\beta-1}\mathrm{e}^{-(\beta-1)\nu^2}(\beta-1)^\beta}{\Gamma(\beta)} \qquad \beta > 1 \tag{13.156}$$

In general, we can set the value of $E(\nu^{-2})$ to a desired constant c by choosing $\alpha = c(\beta - 1)$. Then the covariance matrix of $\underline{W}$ is guaranteed to equal $c\mathbf{M}$ independent of β.

Integrating the conditional density function $f_{\underline{W}}(\underline{w}|\nu)$, as given by Equation 13.149, over the PDF of the nonnegative random variable ν, we obtain the multivariate Student-T distribution. The details are given below. Choosing

$\alpha = \beta - 1$ in Equation 13.153 we can write

$$f_{\underline{W}}(\underline{w}) = \int_0^\infty \frac{1}{(2\pi)^N |\mathbf{M}|^{1/2}} \nu^{2N} e^{(-\nu^2 p/2)} \frac{2\nu^{2\beta-1} e^{-(\beta-1)\nu^2} (\beta-1)^\beta}{\Gamma(\beta)} d\nu$$

$$= \frac{(\beta-1)^\beta}{(2\pi)^N |\mathbf{M}|^{1/2} \Gamma(\beta)} \int_0^\infty 2\nu^{2N+2\beta-1} e^{-\nu^2(\beta-1+p/2)} d\nu \tag{13.157}$$

Letting $(\beta - 1 + p/2)\nu^2 = y$ we get

$$f_{\underline{W}}(\underline{w}) = \frac{(\beta-1)^\beta}{(2\pi)^N |\mathbf{M}|^{1/2} \Gamma(\beta)} \int_0^\infty \frac{y^{N+\beta-1} e^{-y}}{(\beta-1+p/2)^{N+\beta}} dy$$

$$= \frac{(\beta-1)^\beta \Gamma(N+\beta)}{(2\pi)^N |\mathbf{M}|^{1/2} \Gamma(\beta)(\beta-1+p/2)^{N+\beta}} \tag{13.158}$$

The above expression is defined to be the $2N$-dimensional multivariate Student-T distribution with parameters N and β. N represents the number of complex samples and β determines the tail behavior of the multivariate density function. The smaller the value of β, the larger will be the tail.

The density function in Equation 13.156 can be simulated as follows. The first step is to generate a standard Gamma variate having the density function $f_Y(y) = y^{\beta-1} e^{-y}/\Gamma(\beta)$. This was done using the subroutine DGAMDF from the IMSL package. The next step is to divide the generated random variable by the parameter $\beta - 1$. Let $Z = Y/(\beta - 1)$. The density function of Z is

$$f_Z(z) = \frac{(\beta-1)^\beta z^{\beta-1} e^{-z(\beta-1)}}{\Gamma(\beta)} \tag{13.159}$$

The positive square root of $Y/(\beta - 1)$ results in the desired density function. Let $\nu = Z^{1/2}$. Therefore $Z = \nu^2$. Introducing the Jacobian of the transformation, the density function of ν becomes

$$f_\nu(\nu) = \frac{2\nu^{2\beta-1} e^{-(\beta-1)\nu^2} (\beta-1)^\beta}{\Gamma(\beta)} \tag{13.160}$$

which is identical to that in Equation 13.156.

13.4.1.1. The Locally Optimum Detector

The LOD for the multivariate Student-T distribution can now be derived. From Equation 13.36, the LOD is given as

$$\frac{\left.\frac{\partial f_{\underline{D}}(\underline{r} - \theta\underline{s})}{\partial\theta}\right|_{\theta=0}}{f_{\underline{D}}(\underline{r})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{13.161}$$

Assuming the disturbance can be modeled by a multivariate Student-T distribution, $f_{\underline{D}}(\underline{r})$ is given by Equation 13.158, with the variable $\underline{R}$ replacing the variable $\underline{W}$, where $p = \underline{r}^{\mathrm{T}}\mathbf{M}^{-1}\underline{r}$. Since Equation 13.161 is a ratio test and all constants can be placed in the threshold determined by specifying a false alarm probability, all multiplicative constants are ignored for convenience. Hence, we will be concerned only with the terms containing the variable $\underline{R}$. Excluding the constant term the numerator in the ratio test is given by

$$\left.\frac{\partial f_{\underline{D}}(\underline{r}-\theta\underline{s})}{\partial\theta}\right|_{\theta=0} = \frac{\partial}{\partial\theta}\left[\frac{1}{(\beta-1+p/2)^{N+\beta}}\right]\Bigg|_{\theta=0} \tag{13.162}$$

Applying the chain rule, the derivative with respect to θ can be expressed as the derivative with respect to p times the derivative of p with respect to θ. The derivative of p with respect to θ at $\theta = 0$ can be derived as

$$\left.\frac{\partial_p}{\partial\theta}\right|_{\theta=0} = \left(\frac{\partial}{\partial\theta}(\underline{r}-\theta\underline{s})^{\mathrm{T}}\mathbf{M}^{-1}(\underline{r}-\theta\underline{s})\right)\Bigg|_{\theta=0} = -2\underline{s}^{\mathrm{T}}\mathbf{M}^{-1}\underline{r} \tag{13.163}$$

Therefore, the numerator in the ratio test, excluding the constant, is given by

$$\left.\frac{\partial f_{\underline{D}}(\underline{r}-\theta\underline{s})}{\partial\theta}\right|_{\theta=0} = (\beta-1+p/2)^{-(N+\beta+1)} \times \underline{s}^{\mathrm{T}}\mathbf{M}^{-1}\underline{r} \tag{13.164}$$

From the above equation, the sufficient statistic for the LOD for the multivariate Student-T distribution can be written as

$$T_{\mathrm{LOD}}(\underline{r}) = \frac{\underline{s}^{\mathrm{T}}\mathbf{M}^{-1}\underline{r}}{\beta-1+p/2} \tag{13.165}$$

The above result for the LOD statistic is very significant. The numerator in Equation 13.165 is recognized as the Gaussian linear detector. This detector is a matched filter that maximizes the signal-to-disturbance ratio whether or not the disturbance is Gaussian. In weak signal applications, by definition, the signal to disturbance ratio will still be low after matched filtering. The denominator of the LOD statistic is the nonlinear term in the statistic. The behavior of the nonlinearity is such that it scales down large values of p and enhances small values of p. The nonlinearity is plotted as a function of p in Figure 13.18. This is reasonable because, as shown in Section 13.1.3, large values of radar returns result in large p while small values of the returns yields small values of p. Because it is known *a priori* that we are dealing with the weak signal problem, large returns cannot be due to the signal. Consequently, the output of the matched filter is weighted by a small number. On the other hand, the matched filter output is weighted by a large number when the return is small and the contribution due to the signal, if present, can be detected. However, when the signal is present, the optimum nonlinearity alone is not sufficient to get

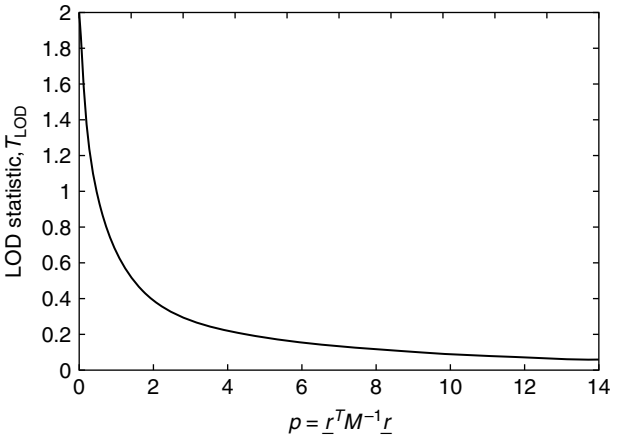

FIGURE 13.18 Nonlinearity of the LOD statistic for the Student-T distribution.

detections, though it brings the output close to the threshold value. The matched filter having a higher output value when the signal is present than when the signal is absent, contributes to raising the output value over the threshold when detections are obtained. The role of the matched filter is explained in greater detail in Section 13.4.1.3.

13.4.1.2. Computer Simulation of Performance

Because analytical closed form expressions for the detection and false alarm probabilities of the LOD in a multivariate Student-T distributed clutter are difficult to obtain, performance is evaluated through computer simulations for weak signal applications. For simulation purposes a multivariate Student-T distributed disturbance vector $\underline{D}$ and a transmitted signal vector $\underline{S}$ have to be generated. The covariance matrix of the clutter process is assumed known with unit elements along the diagonal. To get the covariance matrix of the disturbance we add a small number, determined by the clutter to noise ratio, to the diagonal elements of the clutter covariance matrix. This serves to limit the performance of the receiver even where the clutter power is negligible. In this simulation, the clutter to noise ratio is taken to be 80 dB. The simulation procedure for the disturbance vector is outlined as follows:

1. Generate a $2N$-dimensional white Gaussian random vector $\underline{X}'$. This was done by using the DRRNOA subroutine from the IMSL package.
2. Do a Cholesky decomposition on the matrix $\mathbf{M}$ to get $\mathbf{M} = \mathbf{K}\mathbf{K}^{\mathrm{T}}$ where $\mathbf{K}$ is a lower triangular matrix.
3. The vector $\underline{X} = \mathbf{K}\underline{X}'$ is the multivariate correlated Gaussian vector.
4. Generate a standard Gamma variate Y. This was done by using the subroutine DGAMDF from the IMSL package.

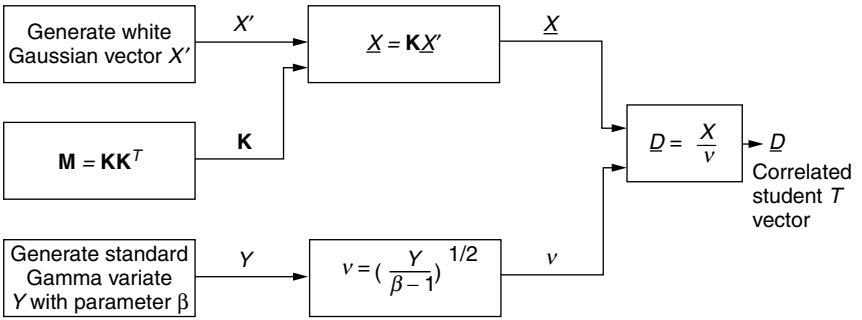

FIGURE 13.19 Generation scheme for the correlated multivariate Student-T distributed vector.

5. Obtain $\nu = (Y/(\beta - 1))^{1/2}$. The random variable ν has the generalized chi PDF.
6. Obtain the multivariate Student-T distributed disturbance vector $\underline{D}$ with the desired correlation properties from $\underline{D} = \underline{x}/\nu$.

The block diagram of the simulation procedure is shown in Figure 13.19. The autocorrelation of the clutter process is taken to be a geometric function in this problem. Assuming radar returns from clutter cells to be highly correlated, as is the case with ground clutter, the sample to sample time correlation is taken as 0.95 in this problem. Specifically, the sample autocorrelation function is chosen as

$$R_{CC}(n) = (0.95)^n \quad n = 0, 1, \ldots, N-1 \tag{13.166}$$

where $R_{CC}(n)$ is the discrete time autocorrelation function of the clutter process. Figure 13.20 and Figure 13.21 show the autocorrelation and the power spectral density of the clutter process. The power spectral density of clutter has a very small spread as the clutter is highly correlated. Using Equation 13.166 the elements of the covariance matrix of the disturbance can be filled appropriately. The elements of the signal vector are chosen such that the nth element $S_n = e^{j2\pi f_D(n-1)T}$, $n = 1,2,\ldots,N$. f_D represents the Doppler frequency shift of the received signal and T represents the time separation between sampling instants, as shown in Figure 13.22.

The detector in Equation 13.165 is now simulated. A value of $\beta = 1.5$ for the multivariate Student-T distribution is chosen because this value results in a relatively long tail for the corresponding marginal PDF of one element of the vector. By evaluating thresholds for specified false alarm probabilities, the Student-T distribution was seen to have heavier tails than the Gaussian distribution for false alarm probabilities less than 10^{-4} but smaller tails than the Gaussian otherwise.

The thresholds corresponding to false alarm probabilities 10^{-k}; $k = 1,2,3,4$ are obtained through the method of extreme value theory explained in Section

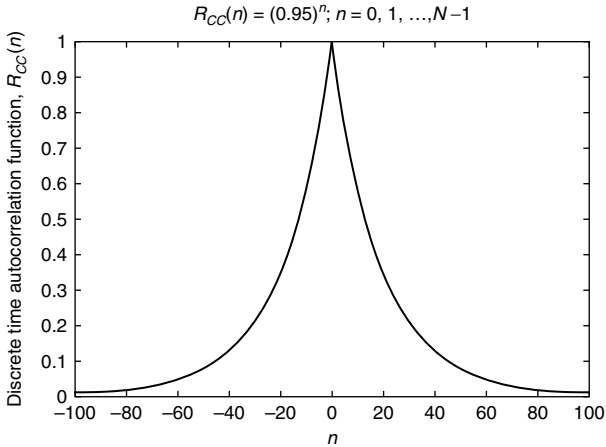

FIGURE 13.20 Autocorrelation of the clutter process for the Student-*T* distribution.

13.3. Once the threshold is set, the detection probabilities are obtained by simulating the LOD for received vectors consisting of the sum of the signal and disturbance vectors for various signal-to-disturbance ratios. The value of f_D is chosen to be zero in this simulation, resulting in a worst case situation. The number of trials in the Monte Carlo simulation for obtaining detection probabilities is equal to 10,000. The performance of the LOD is compared to that of the Gaussian detector for the same multivariate Student-*T* distributed clutter. The test statistic for the Gaussian detector is the same as the numerator of the LOD, i.e., $\underline{s}^T\mathbf{M}^{-1}r$.

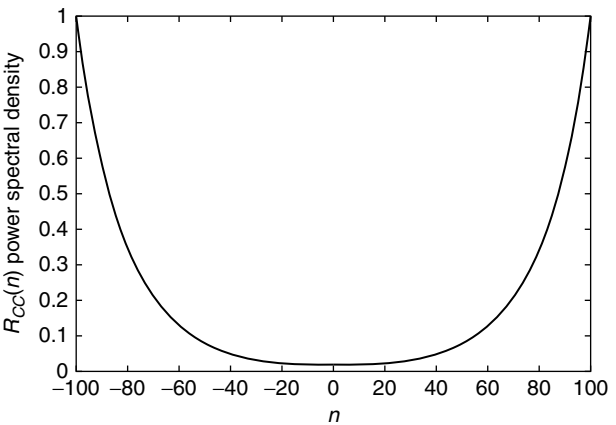

FIGURE 13.21 Power spectral density of the clutter process for the Student-*T* distribution.

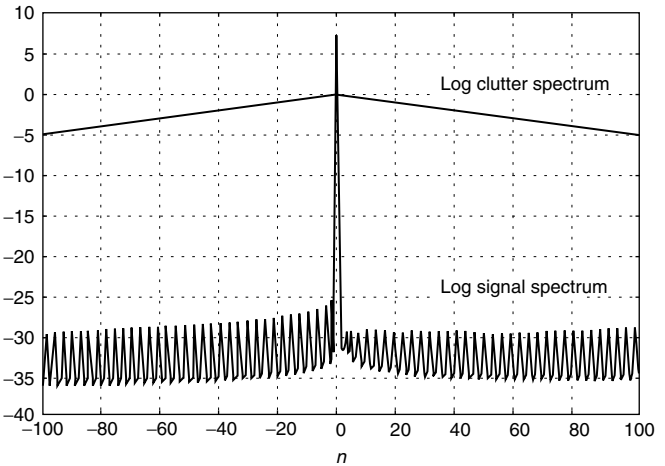

FIGURE 13.22 Log power spectral density of the clutter and signal processes for the Student-*T* distributions.

13.4.1.3. Results of the Computer Simulation

The results of the computer simulations are shown in Table 13.4. When SCR is less than 0 dB, it is seen from the tables that the LOD always outperforms the Gaussian receiver for all values of the tabulated false alarm probabilities. The difference is especially significant for false alarm probabilities equal to 10^{-3} and 10^{-4}.

The Student-*T* distribution, while being heavier tailed than the Gaussian, is not as heavy tailed as the *K*-distribution and Weibull distribution. In fact, the Student-*T* distribution may not be a likely candidate for modeling the radar disturbance. The Student-*T* distribution was chosen as the first distribution to be studied only because of the mathematical simplicity and well behaved nature of its multivariate PDF. Nevertheless, the analysis done with the Student-*T* distribution confirms that the LOD outperforms the Gaussian receiver for a nonGaussian weak signal application.

As the SCR is reduced, it can be observed from the tables that the Gaussian receiver performance degrades abruptly for false alarm probabilities less than or equal to 10^{-2}, whereas the LOD shows a gentler variation in performance. From the Tables it is seen that the performance of the LOD peaks around 0 dB and falls off for both higher and lower values of SCR. But as the SCR values fall below 0 dB, the degradation in the LOD's performance is not as drastic as that of the Gaussian receiver. Both the receivers show an improvement in performance as the number of samples is increased. However, the LOD shows a dramatic improvement in performance when the sample size is greater than 64. When the sample size is equal to 64 and the false alarm probability is as low as 10^{-4}, it is seen from Table 13.4(g) that the detection probability resulting from the LOD is

TABLE 13.4
Probability of Detection P_D for Student-T Distributed Disturbance, $\beta = 1.5$

	P_D	
SCR (dB)	**LOD**	**GR**
	(a) $N = 16$, $P_F = 10^{-2}$	
5	0.35	0.65
3	0.36	0.35
1	0.37	0.23
0	0.38	0.11
−3	0.32	0.052
−6	0.27	0.039
−8	0.23	0.03
−9	0.15	0.024
−10	0.10	0.019
	(b) $N = 16$, $P_F = 10^{-3}$	
5	0.08	0.007
3	0.11	0.005
1	0.17	0.004
0	0.16	0.003
−3	0.13	0.002
−6	0.10	0.001
−8	0.08	0.001
	(c) $N = 32$, $P_F = 10^{-2}$	
5	0.38	0.99
3	0.42	0.57
−1	0.45	0.32
0	0.46	0.18
−3	0.38	0.08
−6	0.31	0.06
−8	0.27	0.04
−9	0.20	0.033
−10	0.13	0.026
−11	0.10	0.019
	(d) $N = 32$, $P_F = 10^{-3}$	
8	0.09	0.15
7	0.11	0.08
5	0.14	0.014
2	0.22	0.008
0	0.26	0.004
−3	0.19	0.003
−6	0.14	0.002
−8	0.11	0.002
−9	0.08	0.001

Continued

TABLE 13.4 Continued

SCR (dB)	P_D LOD	P_D GR
	(e) $N = 64$, $P_F = 10^{-2}$	
5	0.44	0.96
3	0.47	0.68
1	0.53	0.42
0	0.55	0.30
−3	0.48	0.17
−6	0.40	0.10
−8	0.36	0.03
−10	0.14	0.023
−12	0.09	0.012
	(f) $N = 64$, $P_F = 10^{-3}$	
9	0.10	0.17
7	0.15	0.10
5	0.18	0.02
0	0.36	0.005
−3	0.28	0.002
−6	0.22	0.002
−8	0.18	0.001
−9	0.10	0.001
	(g) $N = 64$, $P_F = 10^{-4}$	
5	0.08	0.0007
0	0.25	0.0002
−3	0.19	0.0002
−6	0.13	0.0001
−8	0.11	0.0001
−9	0.06	0.0001
	(h) $N = 128$, $P_F = 10^{-2}$	
5	0.51	0.998
0	0.65	0.71
−1	0.63	0.57
−3	0.57	0.35
−6	0.51	0.16
−8	0.46	0.11
−10	0.22	0.032
−12	0.15	0.027
−13	0.11	0.023
−15	0.09	0.02
	(i) $N = 128$, $P_F = 10^{-3}$	
5	0.20	0.30
4	0.23	0.23

Continued

TABLE 13.4 Continued

	P_D	
SCR (dB)	**LOD**	**GR**
3	0.29	0.19
0	0.48	0.01
−3	0.36	0.006
−6	0.30	0.005
−8	0.26	0.004
−9	0.14	0.003
−10	0.09	0.002
	(j) $N = 128$, $P_F = 10^{-4}$	
7	0.08	0.001
5	0.13	0.0.0004
2	0.30	0.0004
0	0.37	0.0003
−3	0.27	0.0003
−6	0.21	0.0002
−8	0.17	0.0001
−9	0.09	0.0001

SCR = Signal to Clutter Ratio, LOD = Locally Optimum Detector, GR = Gaussian Receiver.

on the order of tenths for SCR as small as −8 dB. The corresponding detection probabilities resulting from the Gaussian receiver are negligibly small. In fact, there is an improvement factor in the vicinity of three orders of magnitude in favor of the LOD. When $-10\,\text{dB} \leq \text{SCR} < 0\,\text{dB}$, an interesting observation from the tables, for all values of N considered, is that the LOD outperforms the Gaussian receiver by close to one order of magnitude for $P_F = 10^{-2}$, by three orders of magnitude for $P_F = 10^{-3}$, and by three orders of magnitude for $P_F = 10^{-4}$. For SCR values lower than −10 dB, the LOD significantly outperforms the Gaussian receiver but with very small values of P_D. When both SCR ≥ 4 dB and $P_F > 10^{-3}$, the Gaussian receiver outperforms the LOD, as can be seen from the tables. For positive values of SCR in the range 0 to 5 dB and for false alarm probabilities equal to 10^{-3} and 10^{-4}, it is interesting to note that the LOD still outperforms the Gaussian receiver. The LOD shows a significant performance improvement over the Gaussian receiver over a dynamic range of about 14 dB. The end points of the range vary depending on the sample size and false alarm probability.

The test statistic is a product involving the outputs of matched filter and the optimum nonlinearity. In Section 13.4.1.1 it was explained how the nonlinearity present in the test statistic boosts small values of the received signal and attenuates large values. This not only serves to bring down the value of the threshold needed

to obtain the desired false alarm probability but also to bring the output for small received signals close to the threshold value whether or not the desired signal is present. The role of the matched filter is explained as follows. The matched filter has a larger output value when the signal is present as opposed to the signal absent case. This serves to increase the statistic in Equation 13.165. However, the quadratic form p, in general, also has a larger value when the signal is present than under the H_0 hypothesis. Thus, the factor, in the test statistic in Equation 13.165 due to the nonlinearity, decreases when the signal is present. This serves to lower the value of the test statistic. Therefore, detections are obtained only when the increase due to the matched filter dominates the decrease due to the nonlinearity. Simulations reveal that the matched filter's role is dominant only when the SCR is in the range $-10\ \text{dB} \leq \text{SCR} < 0\ \text{dB}$. This is expected since the linear receiver's performance is known to drop drastically for very low SCRs. The matched filter's effect is enhanced when a Doppler is present in the desired received signal since the clutter components become less correlated with the Doppler shifted reference signal. However, one must be careful when the Doppler frequency is so large that the signal spectrum appears in the tail of the clutter spectrum. Then a strong signal situation may exist and the nonlinearity that transforms large values into small values will cause performance to degrade. The LOD should not be used in strong signal situations.

The aim of using a LOD is to obtain detection in range–Doppler–azimuth cells where conventional space-time processing is unable to obtain acceptable performance. In present day radars these cells are ignored because the probability of detection is so small under a false alarm constraint. In general, when the SCR is relatively high (>0 dB) the LRT is the optimal test for target detection under a fixed false alarm constraint. In addition to not performing well when SCR is too large, the LOD does not perform well when SCR is too small. When the SCR drops below a certain value, depending upon N and P_{F}, the LOD receiver hardly shows any detections even though it still outperforms the Gaussian receiver. This is because the PDFs under H_0 and H_1 are so close to each other that it is impossible to discriminate between them without increasing the sample size by orders of magnitude.

The concepts of SIRP and LODs are particularly relevant in the context of modern radar applications. When the radar scans a volume searching for targets there might be certain regions in the volume where the clutter returns are significantly stronger than the desired target returns. It is in these regions that we can obtain detections with the LODs. There is a need to monitor the environment so that we are able to separate the clutter regions from volumes limited by weak background noise. When detection is limited by background noise alone, the LOD is not applicable. Using the concepts of artificial intelligence, clutter patches can be identified and the underlying multivariate PDF of the clutter returns can be approximated using the library of SIRPs that have been developed.[26] From the library of LODs the LOD corresponding to the approximated SIRP can be used in clutter regions to obtain detections if the target is present, where earlier it would not have been possible.

13.4.2. The Multivariate *K*-Distribution

In the previous section we analyzed the performance of the LOD for the multivariate Student-*T* distribution. The multivariate *K*-distribution is also a member of the class of joint PDFs arising from SIRPs. Jakeman[50] has shown that the *K*-distributed PDF has a physical interpretation in the sense that it arises from the random walk problem where the number of steps itself is random having a negative binomial distribution. Also, radar clutter has been empirically shown to have *K*-distributed PDF. In this chapter we analyze the performance of the LOD when the background disturbance consisting of the clutter and noise can be approximated as having a multivariate *K*-distribution.

Derivation of the multivariate *K*-distributed PDF from the representation theorem for SIRPs[28] is discussed next. Let the random vector $\underline{X}$ have a multivariate Gaussian distribution with zero mean and covariance matrix $\mathbf{M}$. The zero mean assumption will not affect the generality of the results that follow. The joint density function on the elements of $\underline{X}$ is given by Equation 13.148. Consider the vector $\underline{W} = \nu\underline{X}$, where ν is a nonnegative random variable statistically independent of $\underline{X}$. Let $\underline{w}^{\mathrm{T}}\mathbf{M}^{-1}\underline{w}$ be once again denoted by the variable p. Then, the conditional density function of the vector $\underline{W}$ given ν can be written as

$$f_{\underline{W}}(\underline{w}|\nu) = \frac{1}{(2\pi)^N|\mathbf{M}|^{1/2}}\,\nu^{-2N}\mathrm{e}^{-\frac{p}{2\nu^2}} \tag{13.167}$$

The unconditional density function on $\underline{W}$ is given by Equation 13.150, where $f_\nu(\nu)$ is the PDF of the random variable ν. Because $\underline{X}$ and ν are statistically independent, it follows that

$$E(\underline{W}) = E(\nu\underline{X}) = E(\underline{X})E(\nu) = \underline{0} \tag{13.168}$$

$$E(\underline{W}\,\underline{W}^{\mathrm{T}}) = E(\underline{X}\,\underline{X}^{\mathrm{T}})E(\nu^2) = E(\nu^2)\mathbf{M} \tag{13.169}$$

As is the case for the Student-*T* distribution, the variance of the elements of the vector $\underline{W}$ can be adjusted by appropriate choice of $E(\nu^2)$. Let $f_\nu(\nu)$ be the generalized chi PDF given by Equation 13.153. $E(\nu^2)$ is then given by

$$E(\nu^2) = \int_0^\infty 2\nu^2\,\frac{\nu^{2\beta-1}\mathrm{e}^{-\alpha\nu^2}\alpha^\beta}{\Gamma(\beta)}\,\mathrm{d}\nu = \int_0^\infty 2\,\frac{\nu^{2\beta+1}\mathrm{e}^{-\alpha\nu^2}\alpha^\beta}{\Gamma(\beta)}\,\mathrm{d}\nu \tag{13.170}$$

Letting $\alpha\nu^2 = x$ in the above equation we get

$$E(\nu^2) = \int_0^\infty \frac{x^\beta \mathrm{e}^{-x}\mathrm{d}x}{\alpha\Gamma(\beta)} = \frac{\Gamma(\beta+1)}{\alpha\Gamma(\beta)} = \frac{\beta}{\alpha} \tag{13.171}$$

If we let $\alpha = \beta$, then the generalized chi PDF in Equation 13.153 is such that $E(\nu^2) = 1$ irrespective of the choice for the parameter β. Then the

generalized chi PDF takes the form

$$f_\nu(\nu) = \frac{2\nu^{2\beta-1}\mathrm{e}^{-\beta\nu^2}\beta^\beta}{\Gamma(\beta)} \qquad \beta > 1 \tag{13.172}$$

In general, we can set the value of $E(\nu^2)$ to a desired constant C by choosing α appropriately. Integrating the conditional density function $f_{\underline{W}}(\underline{w}|\nu)$ given by Equation 13.167, over the PDF of the nonnegative random variable ν, we obtain the multivariate K-distribution. The details are given below. Choosing $\alpha = \beta$ in Equation 13.153 we can write

$$\begin{aligned} f_{\underline{W}}(\underline{w}) &= \int_0^\infty \frac{1}{(2\pi)^N|\mathbf{M}|^{1/2}}\nu^{-2N}\mathrm{e}^{-\frac{p}{2\nu^2}}\frac{2\nu^{2\beta-1}\mathrm{e}^{-\beta\nu^2}\beta^\beta}{\Gamma(\beta)}\mathrm{d}\nu \\ &= \frac{\beta^\beta}{(2\pi)^N|\mathbf{M}|^{1/2}\Gamma(\beta)}\int_0^\infty 2\nu^{-2N+2\beta-1}\mathrm{e}^{-\beta\nu^2-\frac{p}{2\nu^2}}\mathrm{d}\nu \end{aligned} \tag{13.173}$$

Letting $\beta\nu^2 = y$ and simplifying we get

$$f_{\underline{W}}(\underline{w}) = \frac{\beta^N}{(2\pi)^N|\mathbf{M}|^{1/2}\Gamma(\beta)}\int_0^\infty y^{-N+\beta-1}\mathrm{e}^{-\left(y+\frac{\beta p}{2y}\right)}\mathrm{d}y \tag{13.174}$$

From page 183 of Watson's book on Bessel functions,[51] we have the result

$$K_\beta(z) = \frac{1}{2}\left(\frac{z}{2}\right)^\beta\int_0^\infty y^{-\beta-1}\mathrm{e}^{-\left(y+\frac{z^2}{4y}\right)}\mathrm{d}y \tag{13.175}$$

provided that the real part of z^2 is Re $z^2 > 0$. $K_\beta(z)$ represents the modified Bessel function of the second kind of order β. Combining Equation 13.174 and Equation 13.175 results in

$$\begin{aligned} f_{\underline{W}}(\underline{w}) &= 2\frac{\beta^N}{(2\pi)^N|\mathbf{M}|^{1/2}\Gamma(\beta)}\left(\frac{2}{(2\beta p)^{1/2}}\right)^{N-\beta}K_{N-\beta}[(2\beta p)^{1/2}] \\ &= \frac{2^{-\frac{N+\beta}{2}+1}\beta^{\frac{N+\beta}{2}}K_{N-\beta}[(2\beta p)^{1/2}]}{\pi^N|\mathbf{M}|^{1/2}\Gamma(\beta)p^{\frac{N-\beta}{2}}} \end{aligned} \tag{13.176}$$

The above expression is defined to be the $2N$-dimensional multivariate K-distribution with parameters N and β. N represents the number of complex samples and β determines the tail behavior of the multivariate density function.

For simulation purposes, the density function in Equation 13.172 can be simulated as follows. The first step is to generate a standard Gamma variate having the density function $f_Y(y) = y^{\beta-1}\mathrm{e}^{-y}/\Gamma(\beta)$. The IMSL package is used to generate the standard Gamma variates. The next step is to divide the generated

random variable by the parameter β. Let $X = Y|\beta$. The density function of X is

$$f_X(x) = \frac{\beta^\beta x^{\beta-1} e^{-\beta x}}{\Gamma(\beta)} \tag{13.177}$$

The positive square root of Y/β results in a variate having the desired density function. Let $\nu = X^{1/2}$. Therefore, $X = \nu^2$. Introducing the Jacobian of the transformation, the density function of ν becomes

$$f_\nu(\nu) = \frac{2\nu^{2\beta-1} e^{-\beta\nu^2} \beta^\beta}{\Gamma(\beta)} \tag{13.178}$$

which is identical to that in Equation 13.172. This random variable is used to multiply the Gaussian vector $\underline{X}$ in order to generate the K-distributed vector $\underline{W}$.

13.4.2.1. The Locally Optimum Detector

The LOD for the multivariate K-distribution can now be derived. From Equation 13.36 the LOD is given as

$$\frac{\left.\dfrac{\partial f_{\underline{D}}(\underline{r} - \theta\underline{s})}{\partial\theta}\right|_{\theta=0}}{f_{\underline{D}}(\underline{r})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{13.179}$$

Assuming the disturbance can be modeled by a multivariate K-distribution, $f_D(\underline{r})$ is given by Equation 13.176 with the variable $\underline{R}$ replacing the variable $\underline{W}$, where $p = \underline{r}^T \mathbf{M}^{-1} \underline{r}$. Since Equation 13.179 is a ratio test and all constants can be placed in the threshold determined by specifying a false alarm probability, all multiplicative constants are ignored for convenience. Hence, we will be concerned only with the terms containing the vector $\underline{R}$. Excluding the constant term the numerator in the ratio test is given by

$$\left.\frac{\partial f_{\underline{D}}(\underline{r} - \theta\underline{s})}{\partial\theta}\right|_{\theta=0} = \left.\frac{\partial}{\partial\theta}\left[p^{\frac{\beta-N}{2}} K_{N-\beta}[(2\beta p)^{\frac{1}{2}}]\right]\right|_{\theta=0} \tag{13.180}$$

Applying the chain rule, the derivative with respect to θ can be expressed as the derivative with respect to p times the derivative of p with respect to θ. From Equation 13.163, the derivative of p with respect to θ at $\theta = 0$ is given by $-2s^T \mathbf{M}^{-1} r$. Therefore, the numerator in the ratio test, excluding the constant, is given by

$$\left.\frac{\partial f_{\underline{D}}(\underline{r} - \theta\underline{s})}{\partial\theta}\right|_{\theta=0} = \left(\frac{\beta - N}{2} p^{\frac{\beta-N}{2}-1} K_{N-\beta}[(2\beta p)^{\frac{1}{2}}] + p^{\frac{\beta-N}{2}} K'_{N-\beta}[(2\beta p)^{\frac{1}{2}}] \frac{\beta^{\frac{1}{2}}}{2p}\right)(-2\underline{s}^T \mathbf{M}^{-1} \underline{r}) \tag{13.181}$$

where $K'_{N-\beta}[x]$ denotes the derivative of $K_{N-\beta}[x]$ with respect to x. From the above equation, the sufficient statistic for the LOD for the multivariate K-distribution can be written as

$$T_{\mathrm{LOD}}(\underline{r}) = -\underline{s}^{\mathrm{T}}\mathbf{M}^{-1}\underline{r}\left[\frac{\beta - N}{p} + \left(\frac{2\beta}{p}\right)^{1/2}\frac{K'_{N-\beta}[(2\beta p)^{\frac{1}{2}}]}{K_{N-\beta}[(2\beta p)^{\frac{1}{2}}]}\right] \tag{13.182}$$

Equation 13.182 can be simplified further using the property of Bessel functions. From p. 79 of Watson's book on Bessel functions[51] we have the two results:

$$1.\ K'_{\nu}(z) = \frac{\nu}{z}K_{\nu}(z) - K_{\nu+1}(z) \tag{13.183}$$

$$2.\ K_{-\nu}(z) = K_{\nu}(z) \tag{13.184}$$

Use of Equation 13.183 and Equation 13.184 in Equation 13.182 and combining the constant term with the threshold results in

$$T_{\mathrm{LOD}}(\underline{r}) = -\underline{s}^{\mathrm{T}}\mathbf{M}^{-1}\underline{r}\left(\frac{K_{N-\beta+1}[(2\beta p)^{\frac{1}{2}}]}{p^{\frac{1}{2}}K_{N-\beta}[(2\beta p)^{\frac{1}{2}}]}\right) \tag{13.185}$$

The numerator in Equation 13.185 once again has as a factor a term which is the Gaussian linear detector or the matched filter that maximizes the signal-to-disturbance ratio whether or not the disturbance is Gaussian. The factor multiplying the linear detector is the optimum nonlinearity for weak signal detection when the disturbance is K-distributed. The nonlinearity is plotted as a function of p in Figure 13.23. The behavior of the nonlinearity is similar in form to the one obtained for the Student-T distributed disturbance. It scales down large

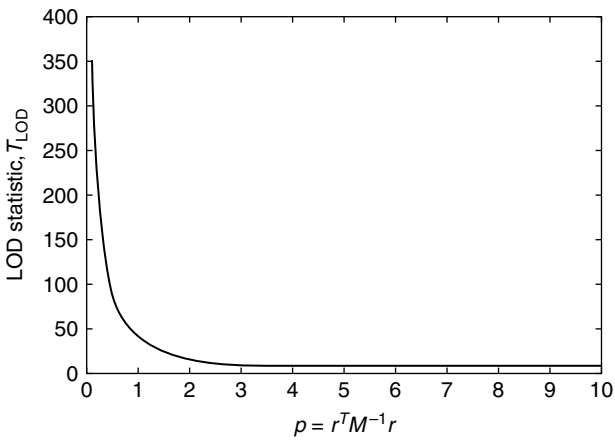

FIGURE 13.23 Nonlinearity of the LOD statistic for the K-distribution.

values of p and enhances small values of p. Since it is known that we are dealing with the weak signal problem, large returns cannot be due to the signal. Consequently, the output of the matched filter is weighted by a small number. On the other hand, when the return is small, there is a greater chance of the signal being detected, if present. Hence, the nonlinearity weights it by a large number. The role of the matched filter is also similar to the situation when the disturbance was modeled as having a Student-T distributed disturbance.

13.4.2.2. Computer Simulation of Performance

Because analytical closed form expressions for the detection and false alarm probabilities of the LOD in a multivariate K-distributed clutter are difficult to obtain, performance is evaluated through computer simulations for weak signal applications. For simulation purposes a multivariate K-distributed disturbance vector $\underline{D}$ and a transmitted signal vector $\underline{S}$ have to be generated. The covariance matrix of the clutter process is assumed known with unit elements along the diagonal. To get $\mathbf{M}$, the covariance matrix of the disturbance, we add a small number, determined by the clutter to noise ratio, to the diagonal elements of the clutter covariance matrix. This serves to limit the performance of the receiver even where the clutter power is negligible. In this simulation, the clutter to noise ratio is taken to be 80 dB. The simulation procedure for the disturbance vector is outlined as follows:

1. Generate a $2N$-dimensional white Gaussian random vector $\underline{X}'$. This was done by using the DRRNOA subroutine from the IMSL package.
2. Do a Cholesky decomposition on the matrix $\mathbf{M}$ to get $\mathbf{M} = KK^{\mathrm{T}}$ where K is a lower triangular matrix.
3. The vector $\underline{X} = K\underline{X}'$ is the multivariate correlated Gaussian vector.
4. Generate a standard Gamma variate Y. This was done by using the subroutine DGAMDF from the IMSL package.
5. Obtain $\nu = (Y/\beta)^{1/2}$. The random variable ν has the generalized chi PDF.
6. Obtain the multivariate K-distributed disturbance vector $\underline{D}$ with the desired correlation properties from $\underline{D} = \nu\underline{X}$.

The block diagram of the simulation procedure is shown in Figure 13.24. The autocorrelation of the clutter process is taken to be a Gaussian function in this problem. Assuming radar returns from clutter cells to be highly correlated, as is the case with ground clutter, the sample to sample time correlation is taken as 0.999 in this problem. Specifically, the sample autocorrelation function is chosen as

$$R_{\mathrm{CC}}(n) = \exp(-0.000801n^2) \qquad n = 0, 1, \ldots, N-1 \tag{13.186}$$

where $R_{\mathrm{CC}}(n)$ is the discrete time autocorrelation function of the clutter process. The autocorrelation and the power spectral density of the clutter process are

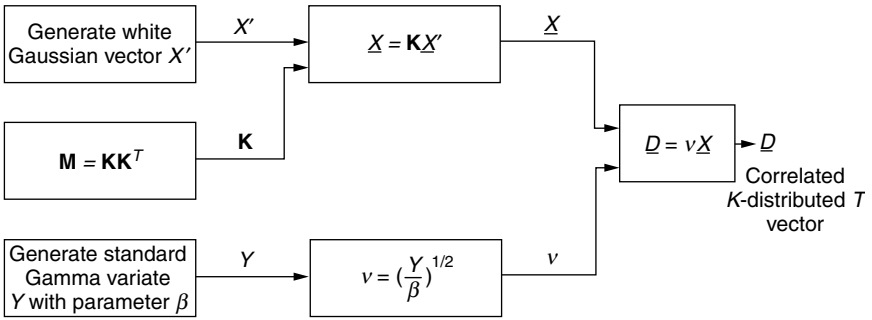

FIGURE 13.24 Generation scheme for the correlated multivariate *K*-distributed vector.

shown in Figure 13.25 and Figure 13.26, respectively. Once again, we see that the spectral spread of the clutter is very small. Using the above function the elements of the covariance matrix of the disturbance can be filled appropriately. The elements of the signal vector are chosen such that the nth element $S_n = e^{j2\pi f_D(n-1)T}$, $n = 1, 2, \ldots, N$. f_D represents the Doppler frequency shift of the received signal, and T represents the time separation between sampling instants, as shown in Figure 13.27.

The detector in Equation 13.165 is now simulated. As $\beta \to \infty$, the *K*-distribution tends to the Gaussian distribution. As $\beta \to 0.5$, the *K*-distribution deviates from that of the Gaussian in the sense of having very large tails. Four different values of $\beta = 0.5$, 1.0, 1.5, and 2.0 were used for performance evaluation.

The threshold corresponding to false alarm probabilities $10^{-k}; k = 1, 2, 3, 4$ are obtained through the method of extreme value theory explained in

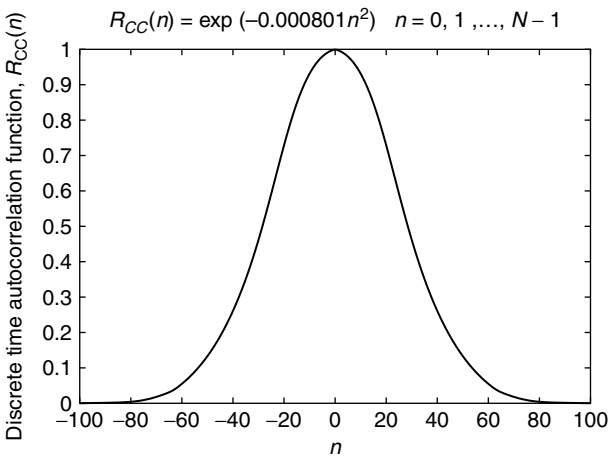

FIGURE 13.25 Autocorrelation function of the clutter process for the *K*-distribution.

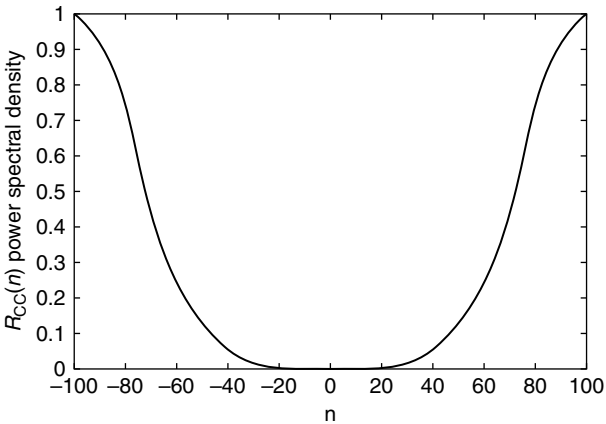

FIGURE 13.26 Power spectral density of the clutter process for the *K*-distribution.

Section 13.3. Once the threshold is set the detection probabilities are obtained by simulating the LOD for received vectors consisting of the sum of the signal and disturbance vectors for various signal to disturbance ratios. The value of f_D is chosen to be zero in this simulation, a worst case situation. The number of trials in the Monte Carlo simulation for obtaining detection probabilities is equal to 10,000. The performance of the LOD is compared to that of the Gaussian detector for the multivariate *K*-distributed clutter. The test statistic for the Gaussian detector is given by $\underline{s}^T\mathbf{M}^{-1}\underline{r}$. Simulations could not be carried out for sample sizes larger than 128 because of the behavior of the modified Bessel functions. The modified Bessel functions are highly nonlinear and numerical difficulties

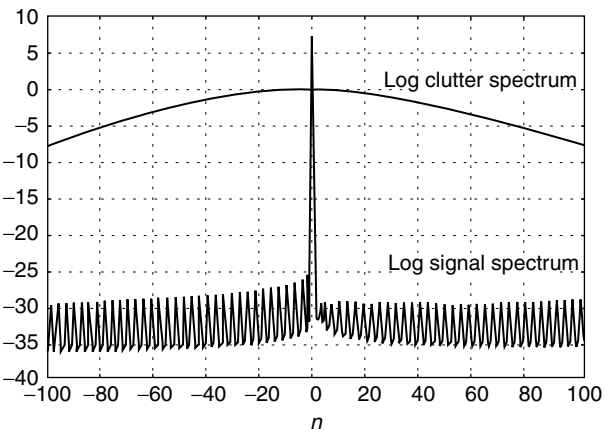

FIGURE 13.27 Log power spectral density of the clutter and signal processes for the *K*-distribution.

arise in the evaluation of the Bessel functions for either small arguments and large orders or large arguments and small orders. For the modified Bessel function of the second kind $K_\nu(x)$, ν must not be so large compared to x such that

$$K_\nu(x) \approx \frac{\Gamma(\nu)}{2(x/2)^\nu} \tag{13.187}$$

overflows. With reference to Equation 13.185 it is noticed that for $\nu = \beta - N > 128$, the value $x = 2(2\beta p)^{\frac{1}{2}}$ frequently is small enough to result in overflow. However, we notice that in Equation 13.185, the test statistic has a ratio of modified Bessel functions with the order differing by one. By using the small argument approximation given in Equation 13.187, the overflow problem for small arguments and large orders can be overcome. On the other hand, when the argument is large and the order is small underflow problems result. IMSL uses an iterative scheme to generate the modified Bessel functions of the second kind, starting from lower orders and building upto higher orders. The lower order Bessel functions, needed to generate the higher order Bessel functions, overflow for sufficiently large values of the argument. Because of this numerical difficulty analysis of performance is restricted to sample sizes smaller than 128.

13.4.2.3. Conclusions

When the sample size was less than 128, the probability of detection for the LOD was always less than 0.1 for false alarm probabilities less than or equal to 10^{-2}. Consequently, in this section results are tabulated for $N = 128$. As mentioned previously, numerical difficulties made it impossible to determine performance for sample sizes greater than 128. Results are presented in Table 13.5 for values of the shape parameter equal to 0.5, 1.0, 1.5, and 2.0 when $N = 128$ and $P_F = 10^{-2}$. Recall that $\beta = 0.5$ corresponds to a very large tail while $\beta = 2.0$ results in a distribution that begins to approximate the Gaussian tail. Note that P_D for the LOD is relatively constant when -10 dB $\leq$ SCR < 0 dB. The best performance is achieved for $\beta = 0.5$ where P_D peaks at a value of 0.23 and the LOD performs significantly better than the Gaussian receiver over a dynamic range of 20 dB extending from SCR equal to 0 dB to -20 dB. For $\beta = 1.0$ and 1.5, P_D peaks around 0.18 and 0.19 with a useful dynamic range of approximately 11 dB extending over the range -17 dB $\leq$ SCR ≤ -7 dB. Finally, for $\beta = 2.0$ the peak value of the LOD is 0.16 and the LOD performs better than the Gaussian receiver over a 7 dB dynamic range extending extending between -9 dB and -15 dB. The LOD for the K-distributed disturbance does not peak at values of P_D as large as those for the Student-T distribution. This result agrees with Spaulding[18] who shows that we cannot arbitrarily say, by inspection, that a noise process which is "tremendously" nonGaussian (i.e., the noise distribution has a very large tail) can result in "tremendous" improvement over the corresponding Gaussian or linear receiver situation. This behavior is explained by once again examining the role of the nonlinearity and the matched filter in determining the test statistic. As was the case with the Student-T distribution, the nonlinearity maps large values

TABLE 13.5
Probability of Detection P_D for *K*-Distributed Disturbance; $N = 128$, $P_F = 10^{-2}$

	P_D	
SCR (dB)	**LOD**	**GR**
	(a) $\beta = 0.5$	
5	0.13	0.99
1	0.21	0.25
0	0.21	0.16
−3	0.22	0.12
−6	0.22	0.09
−8	0.23	0.07
−10	0.23	0.05
−12	0.20	0.03
−14	0.17	0.02
−15	0.15	0.01
−18	0.12	0.01
−20	0.08	0.01
	(b) $\beta = 1.0$	
5	0.10	1.0
0	0.17	0.35
−3	0.17	0.33
−6	0.17	0.19
−7	0.18	0.15
−8	0.18	0.13
−10	0.18	0.04
−11	0.16	0.04
−13	0.14	0.04
−15	0.12	0.03
−17	0.10	0.03
−20	0.08	0.02
	(c) $\beta = 1.5$	
−5	0.11	1.0
0	0.18	0.40
−3	0.18	0.37
−6	0.19	0.24
−7	0.19	0.19
−8	0.19	0.17
−10	0.19	0.03
−12	0.16	0.03
−15	0.12	0.026
−17	0.09	0.02
	(d) $\beta = 2.0$	
5	0.06	0.99
0	0.14	0.44

Continued

TABLE 13.5 Continued

	P_D	
SCR (dB)	**LOD**	**GR**
−3	0.14	0.41
−6	0.15	0.25
−8	0.15	0.18
−9	0.16	0.14
−10	0.16	0.08
−12	0.14	0.05
−14	0.11	0.03
−15	0.09	0.02

SCR = Signal to Clutter Ratio, LOD = Locally Optimum Detector, GR = Gaussian Receiver.

into small values and vice versa. However, when $\beta = 0.5$, the tail of the K-distribution is much heavier than any of those encountered with the Student-T distribution. In addition to more large values of the received signal due to the clutter, there are also more small values. This prevents the threshold for the K-distribution from being lowered as much as it was for the Student-T distribution. As a result, if a detection is to occur, a larger boost must be generated by the matched filter when a signal is present. Unfortunately, with reference to Equation 13.185, the increase due to the matched filter under hypothesis H_1 does not dominate the decrease due to the nonlinearity. Consequently, not as many detections result as occurred with the Student-T distribution. Finally, it is pointed out that the detection probability was much less than 0.1 for all values of the shape parameter with $N = 128$ and $P_F = 10^{-3}$.

It should be noted that dramatically improved performance might occur when the sample size is greater than 128. Because the LOD is nonlinear, a threshold effect exists. It is not clear that $N = 128$ is a sufficiently large sample size to get over the threshold. On the other hand, larger sample sizes cause both numerical difficulties and may not be achievable in an actual application.

In the next chapter we come up with an alternative scheme and derive a detector with enhanced performance under weak signal conditions. This detector is termed the amplitude dependent locally optimum detector. This detector is not UMP. Thus, the thresholds for obtaining the desired false alarm probabilities and the detection probabilities are functions of the SCR and θ. It is shown that this detector offers a significant improvement in performance for smaller sample sizes compared to the LOD obtained on the basis of the UMP test.

13.4.3. Determining LOD Threshold with Real Data

Since the LODs corresponding to SIRP multivariate density functions are nonlinear, it is not possible to evaluate the thresholds corresponding to different

false alarm probabilities in closed form. However, it is shown in Section 13.3 that by using extreme value theory we can model the tail behavior of the test statistic using empirically available outputs of the LOD. It is also shown that this technique yields fairly accurate thresholds for sample sizes with orders of magnitude smaller than those required for Monte Carlo simulation. In practice, the cell false alarm probability is determined by specifying the number of false alarms allowed per scan of the surveillance volume. For example, if one false alarm per scan is specified and the surveillance volume consists of one million resolution cells, then the false alarm probability for each cell is set at 10^{-6}. The weak signal detector should be applied only to those cells for which a weak signal problem exists, as discussed in Section 13.1.3.

The number of such cells in typically a small fraction of the total number of cells in the surveillance volume. In addition, because of the difficulty in detecting weak signals, one should allow for a few more false alarms than is the case for strong signal detection. For example, assume a surveillance volume is comprised of one million resolution cells and that 1% of these cells can be classified as "weak signal". Allowing for one false alarm per scan due to the "weak signal" cells, the false alarm probability for each of these cells would be 10^{-4}. If reports from the tracker are used to sort out false alarms, even larger false alarm probabilities, such as 10^{-3} might be acceptable. The false alarm probabilities can be made even larger as the number of "weak signal" cells become smaller.

In practice, reference cells centered around the test cell are used to determine the threshold for the test cell. The problem is complicated by the fact that reference cells too far removed from the test cell may not be representative due to nonhomogeneities in the clutter. As a result, the number of representative reference cells is limited. In fact, there may not be available the number of representative reference cells required by the extreme value theory. This poses a practical problem in terms of implementation of the LOD.

Fortunately, the behavior of the LOD is such that this problem can be overcome. As seen from Figure 13.18 and Figure 13.23, the nonlinearity is such that, it transforms large values of the received observations (corresponding to large values of the quadratic form p) to small values and small values of the received observation (corresponding to small values of the quadratic for p) to large values. This implies that the tail behavior of the LOD test statistic is governed by the body of the multivariate SIRP disturbance PDF. This is due to the fact the tail region of the test statistic corresponds to large values of the test statistic. Due to the nonlinearity, these in turn, arise from small values of the quadratic form p, which correspond to the body of the disturbance PDF.

The LOD based on the K-distributed disturbance was simulated to see if, indeed, the tail of the test statistic corresponded to disturbance values that arise from the body of the K-distributed disturbance. Since the disturbance observation space in an N-dimensional space, corresponding to N points in the observation vector, the body and tail of the disturbance is defined through the corresponding quadratic form p. If the observations are uncorrelated, then p

corresponds to a hypersphere in N-dimensional space. The body of the disturbance density function, then can be defined as those points that lie within the sphere and the tail as those points outside the sphere. For a SIRP disturbance the univariate density function of p can be evaluated in closed form. This enables us to evaluate the point $p = p_0$ such that $\int_0^{p_0} f_P(p)\mathrm{d}p = \delta$, where $f_P(p)$ is the PDF of the quadratic form p and δ is a number between 0 and 1. By choosing δ say equal to 0.98, we define the point p_0. Then we classify the simulated disturbance vectors as follows: If $p > p_0$, the vector arises from the tail of the disturbance PDF. On the other hand, if $p < p_0$, the vector arises from the body of the disturbance PDF. If the vector is correlated the treatment is similar. In this case, p corresponds to $\underline{r}^{\mathrm{T}}\mathbf{M}^{-1}\underline{r}$ where $\mathbf{M}$ is the covariance matrix of the disturbance. A constant value of p then corresponds to an ellipse in a N-dimensional space. A value for p_0, can be defined in the same way as was done for the uncorrelated vector situation. When the vector is uncorrelated note that M is the identity matrix.

To verify this idea, simulations were carried out using the same covariance matrix as in Section 13.4.2.2. The value of $\delta = 0.98$ was chosen. From the analytical expression for the PDF of p corresponding to the K-distributed SIRP, the value of p_0 was found to be equal to 2.15. Hence, observation vectors that resulted in values of $p > 2.15$ were assigned to the tail of the disturbance PDF and those for which $p < 2.15$ were assigned to the body of the disturbance PDF. We now turn our attention to the test statistic. With respect to the PDF of the test statistic, the tail region is defined as the region corresponding to an area of 0.02 in the tail. From the simulations it was seen that the small values of p mapped on to either the left tail (i.e., negative) or the right tail (i.e., positive) of the test statistic PDF. Note that the output of the nonlinearity is nonnegative while the output of the matched filter can be either positive or negative. Consequently, the sign of the matched filter output determines whether small values of p map into the left or right tail. The PDF of the test statistic under the null hypothesis is symmetric with respect to the origin. Therefore, the right and the left tail were each assigned an equal area of 0.01. The threshold estimated using 10,000 computer simulations for a false alarm probability of 0.01 was 131.22. Thus, the right tail of the test statistic is defined to be those values exceeding 131.22. Ten thousand random vectors of dimension $N = 16$ were then simulated from the K-distributed disturbance. The value of the quadratic form and the test statistic were evaluated corresponding to simulated disturbance vectors. The test statistic was then sorted in ascending order. The test statistic values exceeding 131.22 and the corresponding quadratic form values were noted. It was found that the values of p corresponding to the test statistic values greater than 131.22 were all less than 1.72. From the data it is also seen that the smaller the value of the quadratic form, the higher is the value of the test statistic. This confirms that the tail values of the test statistic arise from the body of the disturbance PDF.

To utilize the Ozturk algorithm discussed in Section 13.3.1.1, it is required that approximately 100 of the neighboring cells be representative of the test cell.

Ozturk's algorithm can then be used to accurately approximate the body of the multivariate PDF. Once the body of the disturbance PDF is approximated accurately, it can be employed to generate the much larger sample size required to estimate the thresholds of the test statistic through extreme value theory. Good results are expected for the threshold estimation when the body of the disturbance PDF has been accurately approximated by the Ozturk algorithm because the tail of the test statistic is caused by the body of the disturbance PDF. Thus, we see that the number of reference cells required for threshold estimation is approximately 100, although we still have to generate from the approximated disturbance PDF much larger sample sizes off-line in order to make use of the extreme value theory.

When the disturbance random variables arise from the Student-T distributed disturbance, a similar nonlinear mapping is seen whereby the tail of the test statistic is caused by the body of the disturbance PDF and vice versa. A procedure similar to the one followed for the K-distributed disturbance case can also be used to reduce the number of reference cells. Such a reduction in the number of required reference cells makes the implementation of these LODs practical.

13.5. PERFORMANCE OF THE AMPLITUDE DEPENDENT LOD

In Section 13.4 we analyzed the performance of the LOD when the underlying disturbance was modeled by the Student-T and K-distributions, respectively. It was seen that the performance of the LOD was not as good as desired for the case where the disturbance is approximated as K-distributed. In this chapter we come up with an alternate form of the LOD, which takes into account the amplitude of the weak signal. Such a detector is no longer UMP in the sense that the thresholds for different false alarm probabilities and detection probabilities is a function of the SCR. We will then compare the performance of the ALOD with that of the LOD for the K-distributed and Student-T disturbance models.

The UMP test for a deterministic signal utilizes the ratio of the derivative with respect to the signal strength of the Nth order joint PDF under H_1 to the Nth order joint PDF under H_0. The limit of this ratio as the signal strength tends to zero is evaluated to obtain the test statistic for the decision rule. The ALOD also utilizes the ratio of the derivative with respect to the signal strength of the Nth order joint PDF under H_1 to the Nth order joint PDF under H_0. However, for the ALOD we do not evaluate this ratio as the signal strength tends to zero but leave it as a function of θ, the square root of the SCR. Note that θ also corresponds to the signal strength since, as explained in Section 13.2, the variance of the clutter is taken to be unity. Because θ is unknown a priori, a bank of receivers tuned to different values of θ must be implemented. Such an approach is analogous to that of a Doppler filter bank used in range–Doppler processing of radar signals. Instead of having a bank of Doppler filters, the ALOD employs a bank of

amplitude filters. The output of each filter will be maximized when the signal amplitude is matched to the amplitude for which the filter is designed.

13.5.1. The Amplitude Dependent LOD for the Multivariate *K*-Distributed Disturbance

The ALOD for a given value of the SCR, θ, is defined to be

$$\frac{\dfrac{\partial f_{\underline{D}}(\underline{r}-\theta\underline{s})}{\partial\theta}}{f_{\underline{D}}(\underline{r})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{13.188}$$

Assuming the disturbance can be modeled by a multivariate K-distribution, $f_{\mathrm{D}}(\underline{r})$ under H_0 is given by Equation 13.176 where $p = \underline{r}^{\mathrm{T}}\mathbf{M}^{-1}\underline{r}$. On the other hand, under H_1 the quadratic form in $f_{\mathrm{D}}(\underline{r}-\theta\underline{s})$ is $p_\theta = (\underline{r}-\theta\underline{s})^{\mathrm{T}}\mathbf{M}^{-1}(\underline{r}-\theta\underline{s})$ where the subscript θ is used to emphasize that we are not taking the limit as θ approaches zero. Since Equation 13.188 is a ratio test, all constants can be placed in the threshold that is determined by specifying a false alarm probability. Therefore, the multiplicative constants are ignored for convenience. Hence, we will be concerned only with the terms containing the vector $\underline{R}$. Excluding constants the numerator in the ratio test is given by

$$\frac{\partial f_{\underline{D}}(\underline{r}-\theta\underline{s})}{\partial\theta} = \frac{\partial}{\partial\theta}\left[(p_\theta)^{\frac{\beta-N}{2}}K_{N-\beta}[(2\beta p_\theta)^{\frac{1}{2}}].\right] \tag{13.189}$$

Applying the chain rule, the derivative with respect to θ can be expressed as the derivative with respect to $p\theta$ times the derivative of p_θ with respect to θ. The derivative of p_θ with respect to θ can be derived as

$$\frac{\partial p_\theta}{\partial\theta} = \left(\frac{\partial}{\partial\theta}(\underline{r}-\theta\underline{s})^{\mathrm{T}}\mathbf{M}^{-1}(\underline{r}-\theta\underline{s})\right) = -2\underline{s}^{\mathrm{T}}\mathbf{M}^{-1}(\underline{r}-\theta\underline{s}) \tag{13.190}$$

Therefore, differentiating the right hand side of Equation 13.189 with respect to θ and excluding the constant, we get

$$\begin{aligned}&\frac{\partial f_{\underline{D}}(\underline{r}-\theta\underline{s})}{\partial\theta}\\ &= -l_\theta\left(\frac{\beta-N}{2}(p_\theta)^{\frac{\beta-N}{2}-1}K_{N-\beta}\left[(2\beta p_\theta)^{\frac{1}{2}}\right] + (p_\theta)^{\frac{\beta-N}{2}}K'_{N-\beta}\left[(2\beta p_\theta)^{\frac{1}{2}}\right]\left(\frac{\beta}{2p_\theta}\right)^{\frac{1}{2}}\right)\end{aligned} \tag{13.191}$$

where $l_\theta = 2\underline{s}^{\mathrm{T}}\mathbf{M}^{-1}(\underline{r} - \theta\underline{s})$, the subscript θ on l is again used to emphasize that we are not taking the limit as θ approaches zero and the prime denotes the differentiation operation with respect to the argument of the function.

Equation 13.191 can be simplified further by using Equation 13.183 and Equation 13.184. The amplitude dependent locally optimum test can then be written as

$$T_{\mathrm{ALOD}}(\underline{r}) = \frac{l_\theta}{(p_\theta)^{\frac{1}{2}}}\left(\frac{p_\theta}{p}\right)^{\frac{\beta-N}{2}}\left(\frac{K_{N-\beta+1}[(2\beta p_\theta)^{\frac{1}{2}}]}{K_{N-\beta}[(2\beta p)^{\frac{1}{2}}]}\right) \tag{13.192}$$

Notice that the Bessel function in the numerator is one order higher than the Bessel function in the denominator. This is significant. In the LRT, the orders of the Bessel functions in the numerator and denominator are identical. This order difference between the Bessel function in the numerator and denominator of the ALOD makes the ALOD more sensitive to small perturbations than the classical form of the Neyman–Pearson test.

The computer simulation of performance was carried out in the same manner as described in Section 13.4.2.2. The values of the shape parameter β chosen in this simulation are 0.5, 1.0, 1.5, and 3.0. With reference to Equation 13.192, it is noticed for $\nu = N - \beta > 16$ that the value $x = (2\beta p)^{1/2}$ frequently assumes small enough values to result in overflow. Also, we notice in Equation 13.192 that the test statistic has a ratio of modified Bessel functions with the order differing by one. Since the Bessel functions in the numerator and the denominator of the test statistic differ in both the argument and the order, the small argument approximation given in Equation 13.187 cannot be used. Because of this numerical difficulty analysis of performance is restricted to sample size equal to 16.

13.5.1.1. Results of Computer Simulation

The results of the computer simulation are shown in Table 13.6 for 16 complex samples and $\beta = 0.5$, 1.0, 1.5, and 3.0, respectively. Performance is evaluated only for values of SCR less than or equal to 0 dB because we are interested in the problem of weak signal detection. From the tables we observe that the best performance of the ALOD is obtained for $\beta = 0.5$. For this value of β the K-distribution has a maximum deviation from the Gaussian in the sense of having a large tail. When $\beta = 3.0$, the tail of the K-distribution closely approximates that of the Gaussian.

It is seen from the tables that the ALOD for the K-distribution significantly outperforms both the Gaussian receiver as well as the LOD. The performance of the LOD for K-distributed disturbance is tabulated in Section 13.4, for $N = 128$. Some of those results are shown again in Table 13.6. Table 13.6(a) shows the detection probabilities obtained for different false alarm probabilities and SCRs

TABLE 13.6
Probability of Detection P_D for *K*-distributed Disturbance, $N = 16$

SCR (dB)	P_F	P_D ALOD	P_D GR	P_D LOD ($N = 128$)
		(a) $\beta = 0.5$		
0	10^{-2}	0.50	0.06	0.21
−5	10^{-2}	0.46	0.04	
−10	10^{-2}	0.40	0.02	0.23
−15	10^{-2}	0.32	0.01	0.15
−20	10^{-2}	0.22	0.01	0.08
−25	10^{-2}	0.17	0.01	
−30	10^{-2}	0.11	0.01	
0	10^{-3}	0.41	0.003	
−5	10^{-3}	0.39	0.001	
−10	10^{-3}	0.35	0.001	
−15	10^{-3}	0.27	0.001	
−20	10^{-3}	0.17	0.001	
−25	10^{-3}	0.13	0.001	
−30	10^{-3}	0.08	0.001	
0	10^{-4}	0.35	0.0003	
−5	10^{-4}	0.32	0.0001	
−10	10^{-4}	0.28	0.0001	
−15	10^{-4}	0.23	0.0001	
−20	10^{-4}	0.16	0.0001	
−25	10^{-4}	0.09	0.0001	
0	10^{-5}	0.30	10^{-5}	
−5	10^{-5}	0.28	10^{-5}	
−10	10^{-5}	0.25	10^{-5}	
−15	10^{-5}	0.18	10^{-5}	
−20	10^{-5}	0.10	10^{-5}	
		(b) $\beta = 1.0$		
0	10^{-2}	0.44	0.10	0.17
−5	10^{-2}	0.40	0.06	
−10	10^{-2}	0.33	0.03	0.18
−15	10^{-2}	0.25	0.02	0.12
−20	10^{-2}	0.17	0.01	0.08
−25	10^{-2}	0.12	0.01	
−30	10^{-2}	0.08	0.01	
0	10^{-3}	0.35	0.004	
−5	10^{-3}	0.30	0.001	
−10	10^{-3}	0.24	0.001	
−15	10^{-3}	0.18	0.001	
−20	10^{-3}	0.10	0.001	
0	10^{-4}	0.30	0.0005	

Continued

TABLE 13.6 Continued

SCR (dB)	P_F	P_D ALOD	P_D GR	P_D LOD ($N = 128$)
−5	10^{-4}	0.27	0.0001	
−10	10^{-4}	0.22	0.0001	
−15	10^{-4}	0.11	0.0001	
0	10^{-5}	0.19	10^{-5}	
−5	10^{-5}	0.16	10^{-5}	
−10	10^{-5}	0.11	10^{-5}	
		(c) $\beta = 1.5$		
0	10^{-2}	0.41	0.10	0.18
−5	10^{-2}	0.36	0.06	
−10	10^{-2}	0.27	0.02	0.19
−15	10^{-2}	0.21	0.01	0.12
−20	10^{-2}	0.12	0.01	
0	10^{-3}	0.29	0.009	
−5	10^{-3}	0.24	0.002	
−10	10^{-3}	0.19	0.001	
−15	10^{-3}	0.11	0.001	
0	10^{-4}	0.20	0.0005	
−5	10^{-4}	0.18	0.0004	
−10	10^{-4}	0.16	0.0003	
−15	10^{-4}	0.09	0.0002	
0	10^{-5}	0.13	10^{-5}	
−5	10^{-5}	0.10	10^{-5}	
−10	10^{-5}	0.07	10^{-5}	
		(d) $\beta = 3.0$		
0	10^{-2}	0.36	0.18	0.14
−5	10^{-2}	0.26	0.10	
−10	10^{-2}	0.17	0.03	0.16
−15	10^{-2}	0.11	0.01	0.09
0	10^{-3}	0.18	0.06	
−5	10^{-3}	0.13	0.009	
−10	10^{-3}	0.06	0.004	

SCR = Signal to Clutter Ratio, ALOD = Amplitude Dependent Locally Optimum Detector. GR = Gaussian Receiver, LOD = Locally Optimum Detector.

when $\beta = 0.5$. For $P_F = 10^{-2}$ the ALOD significantly outperforms the Gaussian receiver over a 30 dB dynamic range extending form 0 dB to −30 dB. The peak value of P_D equals 0.5. A similar result to that obtained for the performance of the LOD with a Student-T distributed disturbance is

observed here. The ALOD outperforms the Gaussian receiver by one order of magnitude for $P_F = 10^{-2}$, two orders of magnitude for $P_F = 10^{-3}$, three orders of magnitude for $P_F = 10^{-4}$ and four orders of magnitude for $P_F = 10^{-5}$. The useful dynamic range of the ALOD is smaller for decreasing values of P_F. The useful dynamic range is about 30 dB for false alarm probabilities equal to 10^{-2} and 10^{-3}, 25 dB for false alarm probability equal to 10^{-4}, and 20 dB for a false alarm probability equal to 10^{-5}. Recall that the detection probabilities resulting form the LOD for the *K*-distributed disturbance were negligibly small when P_F was lower than 10^{-2}. The ALOD, on the other hand, yields significant values of P_D for P_F as low as 10^{-5}.

As the value of β is increased it is noticed from the tables that the peak value of P_D as well as the dynamic range of the receiver decreases. This arises because the tail of the *K*-distribution becomes closer to that of the Gaussian. In fact, when $\beta = 3.0$, the detection probabilities obtained with the ALOD for $P_F = 10^{-4}$ and 10^{-5} are negligible. Also, the useful dynamic range of the ALOD is about 15 dB for $P_F = 10^{-2}$ and less than 10 dB for $P_F = 10^{-3}$. It is interesting to note that the Gaussian receiver shows an improvement in performance for increasing values of β. The improvement, however, is only marginal.

The behavior of the ALOD can be understood as follows. With reference to Equation 13.192, the ALOD can be factored into three terms. The first is l_θ. The second is

$$\frac{1}{(p_\theta)^{\frac{1}{2}}}\left(\frac{p_\theta}{p}\right)^{\frac{\beta-N}{2}}$$

and the third term is

$$\frac{K_{N-\beta+1}[(2\beta p_\theta)^{\frac{1}{2}}]}{K_{N-\beta}[(2\beta p)^{\frac{1}{2}}]}$$

The term l_θ behaves identically to that of the Gaussian receiver. It yields a higher value when the signal is present than when it is absent case. The random variable p_θ, in general, tends to assume a lower value whether or not the signal is present, compared to the term p. Since $\beta - N/2$ is negative with magnitude greater than one, the second term in the ALOD assumes a value greater than one whether or not the signal is present. However, from the simulations, it is seen in most cases that the second factor in the ALOD assumes a larger value when the desired signal is present than the value when it is not present. The modified Bessel function of the second kind, being a monotonically decreasing function assumes large values when the argument is small and small values when the argument is large. Since $p_\theta < p$ in general, the third term in the ALOD also has a value greater than one whether or not the signal is present. Once again, from the computer simulations, it is observed that the third term has a higher value in most cases when the desired signal is present. Thus, the ALOD for the *K*-distributed disturbance has three factors where each of the three factors, in general, assumes

larger values under hypothesis H_1 than under hypothesis H_0. This contributes to the increased sensitivity of the ALOD to weak signals for the K-distributed disturbance resulting in dramatically improved performance over that of the LOD with a much larger sample size. Recall that the LOD has two factors, one of which increases when the desired signal is present and the other decreases. The increase in the value of the test statistic of the ALOD when the signal is present is not restricted to weak signal situations. It will also hold true for strong signal situations. However, when there is a strong signal situation, the LRT is the optimal test and the ALOD should not be used.

13.5.2. The Amplitude Dependent LOD for the Student-T Distributed Disturbance

Assuming the disturbance can be modeled by a multivariate Student-T distribution, $f_{\underline{D}}(\underline{r})$ under H_0 is given by Equation 12.234 where $p = \underline{r}^{\mathrm{T}}\mathbf{M}^{-1}\underline{r}$. On the other hand, as before, under H_1 the quadratic form in $f_{\underline{D}}(\underline{r} - \theta\underline{s})$ is $p_\theta = (\underline{r} - \theta\underline{s})^{\mathrm{T}}\mathbf{M}^{-1}(\underline{r} - \theta\underline{s})$ where the subscript θ is used to emphasize that we are not taking the limit as θ approaches zero. Since Equation 13.188 is a ratio test, all constants can be placed in the threshold determined by specifying a false alarm probability. Therefore, the multiplicative constants are ignored for convenience. Hence, we will be concerned only with terms containing the vector $\underline{R}$. Excluding constants the numerator in the ratio test is given by

$$\frac{\partial f_{\underline{D}}(\underline{r} - \theta\underline{s})}{\partial\theta} = \frac{\partial}{\partial\theta}\left[\frac{1}{\Big(\beta - 1 + (p_\theta/2)\Big)^{N+\beta}}\right] \tag{13.193}$$

Applying the chain rule, as was done with the K-distributed disturbance differentiating with respect to θ and excluding the constant, the numerator in the ratio test is given by

$$\frac{\partial f_{\underline{D}}(\underline{r} - \theta\underline{s})}{\partial\theta} = \frac{l_\theta}{\Big(\beta - 1 + (p_\theta/2)\Big)^{(N+\beta+1)}} \tag{13.194}$$

where $l_\theta = 2\underline{s}^{\mathrm{T}}\mathbf{M}^{-1}(\underline{r} - \theta\underline{s})$. From the above equation the sufficient statistic for the ALOD for the multivariate Student-T distribution can be written as

$$T_{\mathrm{ALOD}}(\underline{r}) = \frac{l_\theta\Big(\beta - 1 + (p_\theta/2)\Big)^{N+\beta}}{\Big(\beta - 1 + (p_\theta/2)\Big)^{N+\beta+1}} \tag{13.195}$$

Notice that the exponent in the denominator of the test statistic is one more than that in the numerator. In the LRT the exponents in the numerator and the denominator of the test statistic are identical. But the factor l_θ appearing in the ALOD will not be present. The computer simulation of the ALOD for the

TABLE 13.7
Probability of Detection P_D for Student-T Distributed Disturbance, $\beta = 1.5$

SCR (dB)	P_F	P_D ALOD	P_D GR	P_D LOD
		(a) $N = 16$		
0	10^{-2}	0.48	0.11	0.38
−3	10^{-2}	0.36	0.052	0.32
−6	10^{-2}	0.30	0.039	0.27
−9	10^{-2}	0.19	0.024	0.15
−10	10^{-2}	0.12	0.019	0.10
0	10^{-3}	0.34	0.003	0.16
−3	10^{-3}	0.23	0.002	0.13
−6	10^{-3}	0.15	0.001	0.10
−8	10^{-3}	0.10	0.001	0.08
		(b) $N = 32$		
0	10^{-2}	0.48	0.18	0.46
−3	10^{-2}	0.39	0.08	0.38
−6	10^{-2}	0.34	0.06	0.31
−9	10^{-2}	0.24	0.033	0.20
−10	10^{-2}	0.15	0.026	0.13
−12	10^{-2}	0.11	0.013	
0	10^{-3}	0.37	0.004	0.26
−3	10^{-2}	0.25	0.003	0.19
−6	10^{-3}	0.16	0.002	0.14
−8	10^{-3}	0.11	0.002	0.11
		(c) $N = 64$		
0	10^{-2}	0.59	0.30	0.55
−3	10^{-2}	0.52	0.17	0.48
−6	10^{-2}	0.46	0.10	0.40
−9	10^{-2}	0.40	0.03	
−10	10^{-2}	0.16	0.023	0.14
−12	10^{-2}	0.12	0.012	0.09
0	10^{-3}	0.40	0.005	0.36
−3	10^{-2}	0.30	0.002	0.28
−6	10^{-3}	0.23	0.002	0.22
−9	10^{-3}	0.11	0.001	0.10
0	10^{-4}	0.36	0.0007	0.25
−3	10^{-4}	0.24	0.0002	0.19
−6	10^{-4}	0.15	0.0001	0.13
−8	10^{-4}	0.12	0.0001	0.11

SCR = Signal to Clutter Ratio, ALOD = Amplitude Dependent Locally Optimum Detector. GR = Gaussian Receiver, LOD = Locally Optimum Detector.

multivariate Student-T distributed disturbance was carried out as mentioned in Section 13.4.1.2. The results of the computer simulation are shown in Table 13.7.

13.5.2.1. Conclusions

As was the case for the K-distributed disturbance, performance is evaluated only for values of SCR less than or equal to 0 dB. Even though the ALOD may outperform the Gaussian receiver for values above 0 dB of SCR up to a certain level, the decision rule that should be used in such situations is the optimum one based on the LRT. The value of $\beta = 1.5$ is chosen because it represents a heavy tail situation. Three different sample sizes of $N = 16$, 32, and 64 were chosen for simulation purposes.

It is seen from the tables that the ALOD significantly outperforms the Gaussian receiver in the range of SCR values considered. The performance improvement of ALOD over the Gaussian receiver is similar to that of the LOD discussed in Section 13.4.1.3. Compared to the Gaussian receiver, the ALOD has about an order of magnitude improvement in performance for $P_F = 10^{-2}$, two orders of magnitude improvement in performance for $P_F = 10^{-3}$ and three orders of magnitude improvement in performance for $P_F = 10^{-4}$. The performance improvement of the ALOD for the Student-T distribution compared to that of the LOD is not as dramatic as was noticed for the K-distributed case. For the Student-T distribution the performance of the LOD and the ALOD are actually quite close. The ALOD for the K-distribution significantly outperformed both the LOD and the Gaussian receiver. The ALOD for the Student-T distribution, while significantly outperforming the Gaussian receiver in the range of SCR values considered, does not outperform the LOD significantly. In fact, comparing the detection probabilities resulting from the ALOD with the detection probabilities resulting form the LOD, it is noticed that the performance of the ALOD exceeds that of the LOD by less than a tenth for SCR values less than 0 dB.

With reference to Equation 13.195, the ALOD can be factored into two terms. The first term is l_θ. The behavior of l_θ was explained in Section 13.5.1.1. The output of l_θ and the test statistic increase when the signal is present as opposed to the signal absent case. The second term is a ratio of two polynomials. In general, as discussed in Section 13.5.1.1, $p_\theta < p$ whether or not the signal is present. Therefore, in general, the ratio of the two polynomials is a number greater than unity. It is seen in the simulations that this ratio has a higher value under the signal present hypothesis more often than it does under the signal absent hypothesis. This contributes to an increased sensitivity to the presence of weak signals. This explains why the ALOD outperforms the LOD. The increase in the value of the test statistic of the ALOD when the signal is present is not restricted to weak signal situations alone. The LOD also has two factors in the test statistic. However, when the signal is present, one factor increases and the other decreases. This is not a favorable situation for raising the output of the test statistic over the threshold for small signals. The LOD's nonlinearity for the Student-T distribution, which is the second factor in the test

statistic, is not as highly nonlinear as was observed for the *K*-distributed case. Thus, the performance improvement of the ALOD with respect to the LOD is not as significant.

13.6. CONCLUSIONS

13.6.1. Summary

Conclusions and suggestions for future work are presented in this chapter. We have addressed the problem of weak signal detection in correlated multivariate nonGaussian noise. Weak signal detectors were derived, based on the locally optimum decision rule, using the concept of SIRP for modeling the radar disturbance. LODs are useful only in the neighborhood of the point for which they are evaluated. For nonGaussian problems the test statistic derived for the LOD are nonlinear. Due to the nonGaussian and nonlinear nature of the problem, thresholds needed to set specified false alarm probabilities cannot be obtained in closed form. The GPD in conjunction with the method of Extreme Value Theory was used to obtain accurate approximations to thresholds for specified false alarm probabilities. This was achieved with orders of magnitude fewer samples compared to Monte Carlo simulation. Performance analyses of the LODs for the multivariate *K*-distribution and the Student-*T* distribution were carried out by means of computer simulations. Finally, the concept of the ALOD was introduced. For the *K*-distribution the ALOD was shown to greatly outperform the LOD.

The following significant contributions appear in this dissertation:

1. Under the assumption that the radar clutter can be modeled as a SIRP the canonical model for the LOD was shown to be the product of the Gaussian linear receiver and a zero memory nonlinearity.
2. The computational requirements needed to set thresholds for very small false alarm probabilities were reduced by orders of magnitude using the GPD in conjunction with the method of Extreme Value Theory. Accurate thresholds were determined by introducing Ordered Samples Least Squares technique to estimate the parameters of the GPD. For example, only 20,000 samples were required for different distributions to establish thresholds corresponding to a false alarm probability equal to 10^{-6}.
3. In contrast to the available literature, where LODs are evaluated based on the assumption of an infinite sample size, performance of the LODs were obtained for finite sample sizes.
4. The new concept of the ALOD was introduced.
5. Performance of the ALOD was evaluated for finite sample sizes. The ALOD had a significant performance improvement over the LOD when the disturbance was *K*-distributed.

As a result of the above contributions, practical decisions can be made with respect to use of LODs. These decisions will be based on the available sample size, the desired detection and false alarm probabilities, and the SCR.

13.6.2. Suggestion for Future Research

This research has led to many important results and has also raised a number of interesting questions. In particular, some of the issues arising an extensions to this work are to:

1. Compare performance of the LOD and the ALOD with that of the classical LRT for a broad range of SCRs and sample sizes. This will establish the conditions under which there is a need for weak signal detectors.
2. Analyze performance of the LOD and the ALOD for other multivariate PDFs in the SIRP class, such as, Weibull and Rician.
3. Establish confidence intervals for the thresholds estimated based on Extreme Value Theory.
4. Extend all of the above work to space-time processing.
5. Study the role of signal design in enhancing performance of the LOD.

14 A Generalization of Spherically Invariant Random Vectors (SIRVs) with an Application to Reverberation Reduction in a Correlation Sonar

T. J. Barnard

CONTENTS

14.1. Introduction 800
14.2. The SIRV Representation Theorem 803
14.2.1. The Traditional SIRV Model 803
14.2.2. The Generalized SIRV Model 805
14.2.3. A Comparison of the Traditional and Generalized Models 807
14.3. Generalized SIRV Properties 810
14.3.1. Linear Transformation 810
14.3.2. The Generalized SIRV "Bootstrap" Theorem 811
14.3.3. The Monotonicity of $h_{NM}(\alpha_{x1}, \ldots, \alpha_{xM})$ 812
14.3.4. Spherical Coordinates 812
14.3.5. The Generalized SIRV Bessel Function Representation 817
14.3.6. Minimum Mean Square Error Estimation 822
14.3.7. The Generalized SIRV Laplace Transform Representation 824
14.4. The Generalized SIRV Density Function 826
14.4.1. Direct Evaluation of $h_{NM}(a_x)$ 827
14.4.1.1. Case 1 828
14.4.1.2. Case 2 841

14.4.2. Evaluation of h_{NM} Using the Laplace Transform 844
14.4.2.1. Case 3 846
14.4.2.2. Case 4 854
14.5. Generalized SIRV Generation 856
14.5.1. Multivariate Rejection Theorem 857
14.5.2. Application of the Rejection Theorem 860
14.5.3. Examples of Random Variable Generation 861
14.5.3.1. Example 1 861
14.5.3.2. Example 2 863
14.6. Generalized SIRV Density Approximation 863
14.6.1. Univariate Density Approximation 865
14.6.2. 2-D Density Approximation 867
14.6.3. Multivariate Density Approximation 868
14.6.4. Real Data Analysis 869
14.7. Correlation Sonar Fundamentals 876
14.7.1. Correlation Sonar Basic Operation 876
14.7.2. Correlation Sonar Reverberation Model 880
14.7.2.1. Monostatic and Bistatic Reverberation 881
14.7.2.2. Reverberation as Heard on a Moving Correlation Sonar Platform 883
14.7.3. A Sub-Optimal Correlation Sonar Receiver 890
14.7.4. Performance in Previous Pulse Interference 895
14.8. *M*-ary Detection 896
14.8.1. Optimum *M*-ary Detection 897
14.8.2. Sub-Optimum *M*-ary Detection 901
14.8.3. Generalized SIRV *M*-ary Detection 904
14.9. Conclusion 909
14.9.1. Suggestions for Future Research 910

14.1. INTRODUCTION

An active sonar system operates by transmitting an acoustic pulse into the water and monitoring a target echo. Besides receiving an echo from a desired target, the system may also receive an echo from the ocean floor or the surface as shown in Figure 14.1. In sonar terminology, any nontarget echo arising explicitly from bottom or surface, is referred to as boundary reverberation. Typically, boundary reverberation power greatly exceeds that of a target. If it arrives at the same time as the target echo, this reverberation interferes with detection of the desired signal.

There currently exist several methods for finding a target hidden in loud reverberation. For example, if the target moves fast enough, the resulting Doppler shift lets a low-pass filter separate out the reverberation as shown in Figure 14.2.

In addition to such spectral processing, spatial processing techniques can also help remove reverberation, as explained in Refs. 1–3. For example, Figure 14.3

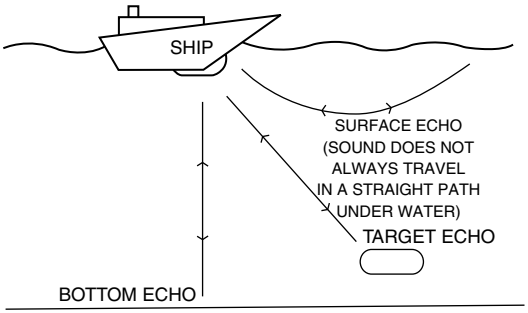

FIGURE 14.1 Sources of reverberation.

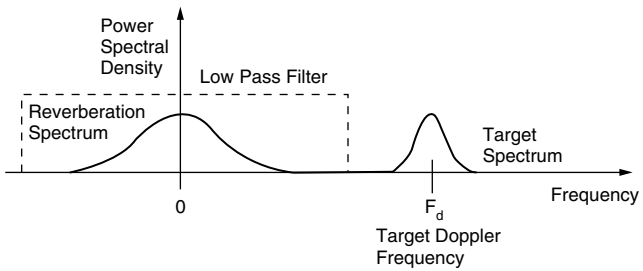

FIGURE 14.2 Spectral reverberation reduction.

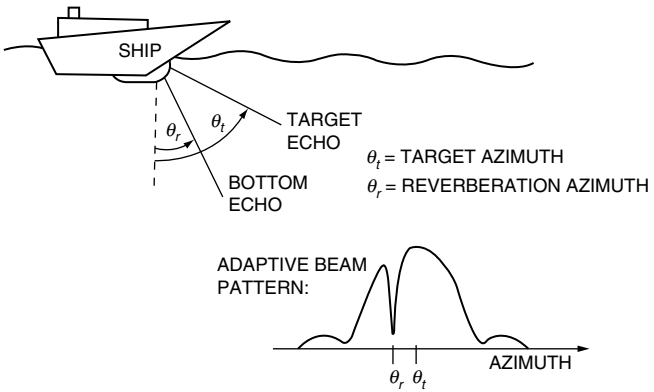

FIGURE 14.3 Spatial reverberation reduction.

shows how an adaptive beamformer steers a beam pattern null in the direction of the interference.

Consider the problem of finding a stationary target, lurking near the ocean floor. Bottom reverberation often masks such a target. Since the target does not move, its spectrum occupies the same band as the reverberation, meaning that

spectral processing, as shown in Figure 14.2, cannot separate the desired signal from the bottom echo interference. Furthermore, spatial processing cannot resolve the target since it lies at approximately the same azimuth as the reverberation. In other words, these traditional reverberation reduction methods cannot isolate this target.

The above example illustrates a shortcoming of existing reverberation reduction algorithms, i.e., they all require some sort of physical separation between the signal and the interference. This separation can be spectral or spatial, but it must exist. When the target and reverberation are not separated as such, these algorithms fail.

Given that the spectral and spatial reverberation reduction techniques fail, traditional sonar systems rely on a matched filter to uncover signals buried in interference. The matched filter is the optimum processor for finding a deterministic signal in Gaussian interference. If the interference is nonGaussian, the optimum processor utilizes a different nonlinear filter.

The first step in the design and evaluation of the optimum nonlinear filter involves determining the underlying distribution of the interference. For this dissertation, the interference is modeled as a generalized Spherically Invariant Random Vector (SIRV). As explained in Ref. 4, an SIRV is obtained by multiplying a Gaussian random vector by a nonnegative random scalar. In the case of radar clutter, the components of an SIRV serve to model the pulse train echoes from a single spatial range or bearing cell.[5–7] In the case of sonar reverberation, the components of an SIRV serve to model range cell returns from a single active pulse, or ping. Any two SIRV realizations result in independent vectors. Consequently, the radar returns, from different range cells or the sonar returns from different pings, must be modeled as statistically independent.

Section 14.2 presents a generalized SIRV model which allows for dependence between SIRV realizations. Following this analysis, Section 14.3 describes significant properties of the generalized SIRV. Although these properties are readily derived, specific examples of the generalized SIRV joint density functions are difficult to obtain because of the multi-dimensional integral involved. Nevertheless, Section 14.4 derives four nontrivial examples.

Given explicit expressions of density functions, Section 14.5 demonstrates how to generate generalized SIRV realizations. Section 14.6 then shows how to approximate the density of real data with a generalized SIRV. This section concludes by analyzing recorded sea data, and finding a close nonGaussian generalized SIRV approximation for density of surface reverberation.

An application particularly sensitive to surface reverberation interference is correlation sonar. As explained in Ref. 8–12, a correlation sonar measures ownship velocity by comparing the return from two active pulses which ensonify the same target volume at different times. Section 14.7 describes how interference arises when echoes from the first pulse intrude on data collected during the reception of the second pulse. When these previous pulse echoes ensonify an ocean boundary, the resulting interference power greatly exceeds that of the desired second pulse signal, leading to significant degradation in

performance. As shown in Section 14.6, these surface echoes may follow nonGaussian densities. Consequently, the cancellation of previous pulse boundary echoes, in correlation of sonar data, is a problem uniquely suited for the results derived in this dissertation.

Section 14.7 presents the optimum correlation sonar receiver. Since this optimum receiver is difficult to implement, a more practical sub-optimum receiver is proposed. Section 14.8 concludes the analysis by simulating correlation of sonar data and evaluating the performance of the sub-optimum receiver in an environment corrupted by nonGaussian interference.

14.2. THE SIRV REPRESENTATION THEOREM

This chapter first presents, the traditional SIRV representation theorem and describes how it applies to radar clutter or sonar reverberation. Following this work, the theorem is generalized to account for spatial dependence in the radar case, or temporal dependence in the sonar case.

14.2.1. The Traditional SIRV Model

Consider a collection of clutter range/bearing cells. Assume that a radar illuminates one of these cells with a pulse train of length N. Let the vector $\mathbf{X}$, of dimension $[N \times 1]$, model these pulse train samples. If $\mathbf{X}$ is given by

$$\mathbf{X} = \mathbf{Z}S \tag{14.1}$$

where $\mathbf{Z}$ is a zero mean Gaussian random vector and S is a random variable, then $\mathbf{X}$ is defined as a SIRV. Although s may be positive or negative, it is assumed nonnegative in this work without loss of generality, as explained in Ref. 13.

Equation 14.1 is referred to as the SIRV representation theorem. The vector $\mathbf{Z}$ models the rapid noisy fluctuations of the clutter while the scalar S models environmental effects associated with the illuminated spatial cell. The random process which generates S, is assumed to be slowly fluctuating so that the same value modulates the entire pulse train.

In case of active sonar, the transmitter emits a single ping, rather than a pulse train. Since the time interval between successive pulses is relatively large, the environment does not remain stable from ping to ping. As such, the sonar model does not treat the elements of an SIRV as successive pulse returns from a single reverberation cell, but rather as samples from different spatial cells ensonified by the same ping. The modulating scalar S then models temporal oceanic effects associated with the environment at the moment of ensonification, such as current turbulence. These effects are assumed to be "slowly fluctuating" across the region of space occupied by the SIRV. Therefore, the same random variable modulates each spatial cell.

The joint density function of the SIRV $\mathbf{X}$ is derived in Ref. 14. The following paragraphs summarize this work.

Assume that the vector $\mathbf{Z}$ contains N jointly Gaussian random variables with zero mean. Based on this assumption, the density function of $\mathbf{Z}$ is

$$f_{\mathbf{Z}}(\mathbf{z}) = (2\pi)^{-N/2}|\Sigma_{\mathbf{z}}|^{-1/2}\exp\left\{-\frac{1}{2}\mathbf{z}^{\mathrm{T}}\Sigma_z^{-1}\mathbf{z}\right\} \tag{14.2}$$

where $\Sigma_{\mathbf{z}}$ is the $[N \times N]$ covariance matrix of the random vector $\mathbf{Z}$. Based on the representation theorem, Equation 14.1, the conditional density of $\mathbf{X}$ given S is

$$f_{\mathbf{X}|\mathbf{S}}(\mathbf{x}|\mathbf{s}) = \frac{1}{s^N} f_{\mathbf{Z}}\left(\frac{\mathbf{x}}{\mathbf{s}}\right) \tag{14.3}$$

Since the joint density of $\mathbf{X}$ and s is given by

$$f_{\mathbf{X},S}(\mathbf{x}, s) = f_{\mathbf{X}|S}(\mathbf{x}|s) f_S(s) \tag{14.4}$$

the marginal density of $\mathbf{X}$ is

$$f_{\mathbf{X}}(\mathbf{x}) = \int_0^\infty f_{\mathbf{X},S}(\mathbf{x}, s)\mathrm{d}s = \int_0^\infty f_{\mathbf{X}|S}(\mathbf{x}|s) f_S(s)\mathrm{d}s \tag{14.5}$$

where the lower limit of integration is zero because S is nonnegative. Substituting Equation 14.2 and Equation 14.3 into Equation 14.5 gives the density of $\mathbf{X}$ as

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-N/2}|\Sigma_{\mathbf{z}}|^{-1/2}\int_0^\infty \frac{1}{s^N}\exp\left\{-\frac{1}{2}\frac{\mathbf{x}^{\mathrm{T}}\Sigma_{\mathbf{z}}^{-1}\mathbf{x}}{s^2}\right\} f_S(s)\mathrm{d}s \tag{14.6}$$

Define the scalar $\alpha_{\boldsymbol{x}}$ as

$$\alpha_{\boldsymbol{x}} = \mathbf{x}^{\mathrm{T}}\Sigma_{\mathbf{z}}^{-1}\mathbf{x} \tag{14.7}$$

Substituting Equation 14.7 into Equation 14.6 yields

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-N/2}|\Sigma_{\mathbf{z}}|^{-1/2}\int_0^\infty \frac{1}{s^N}\exp\left\{-\frac{1}{2}\frac{\alpha_{\mathbf{x}}}{s^2}\right\} f_S(s)\mathrm{d}s \tag{14.8}$$

In Equation 14.8, note that the density of $\mathbf{X}$ depends solely on the scalar quantity $\alpha_{\mathbf{x}}$. The vector is referred to as spherically invariant because when Σ_z

equals a constant K times the identity matrix, $\alpha_{\mathbf{x}}$ becomes

$$\alpha_{\mathbf{x}} = K \sum_{i=1}^{N} x_i^2 \tag{14.9}$$

the equation of a sphere in N-dimensional space.

In order to highlight the importance of the scalar $\alpha_{\mathbf{x}}$, Equation 14.8 is rewritten as

$$f_{\mathbf{X}}(\mathbf{x}) = Kh_N(\alpha_{\mathbf{x}}) \tag{14.10}$$

where K is a constant and $h_N(\cdot)$ is a one-dimensional nonlinear function. From Equation 14.8, the constant is

$$K = (2\pi)^{-N/2}|\Sigma_{\mathbf{Z}}|^{-1/2} \tag{14.11}$$

while the characteristic nonlinear function $h_N(\cdot)$ is

$$h_N(u) = \int_0^{\infty} \frac{1}{s^N} \exp\left\{-\frac{1}{2}\frac{u}{s^2}\right\} f_{\mathbf{S}}(s)\mathrm{d}s \tag{14.12}$$

In Equation 14.12, u is just a dummy variable.

Equation 14.10 to Equation 14.12 provide the basis for the traditional SIRV clutter model. Note that the density of $\mathbf{X}$ is completely specified by the characteristic univariate density $f_{\mathrm{S}}(s)$ and a covariance matrix, $\Sigma_{\mathbf{z}}$. Furthermore, as explained in Ref. 14, the components of the vector $\mathbf{X}$ are nonGaussian and those are similarly distributed except for location and scale variations.

14.2.2. The Generalized SIRV Model

As explained above, the elements of a given SIRV correspond to the pulse train samples from a given radar range or bearing cell. In the sonar case, these elements correspond to digitized temporal range samples received from a given ping. Now consider two vectors; i.e., the return from two radar cells or two sonar pings. Assume that the two underlying Gaussian vectors which generate these SIRV's are independent. Also assume that, a different random scalar modulates each of the Gaussian vector. Based on the representation theorem (see Equation 14.1), these two vectors are given by

$$\mathbf{X}_1 = \mathbf{Z}_1 S_1 \tag{14.13}$$

and

$$\mathbf{X}_2 = \mathbf{Z}_2 S_2 \tag{14.14}$$

Expand this model to account for M radar range cells or M sonar pings with each independent Gaussian vector modulated by a different random scalar.

The corresponding M vectors are

$$\begin{aligned} \mathbf{X}_1 &= \mathbf{Z}_1 S_1 \\ &\vdots \\ \mathbf{X}_M &= \mathbf{Z}_M S_M. \end{aligned} \tag{14.15}$$

From the previous section, we know that any vector $\mathbf{X}_i$ is an SIRV, for all i from one to M. From Equation 14.2 and Equation 14.3, the conditional density of any vector $\mathbf{X}_i$ given S_i is,

$$f_{\mathbf{X}_i|S_i}(\mathbf{x}_i|s_i) = \frac{1}{s^N}(2\pi)^{-N/2}|\Sigma_{\mathbf{z}i}|^{-1/2}\exp\left\{-\frac{1}{2}\frac{\mathbf{x}_i^{\mathrm{T}}\Sigma_{\mathbf{z}i}^{-1}\mathbf{x}_i}{s_i^2}\right\} \tag{14.16}$$

Define the scalar α_{xi} as

$$\alpha_{xi} = \mathbf{x}\boldsymbol{i}^{\mathrm{T}}\Sigma_{\mathbf{z}i}^{-1}\mathbf{x}_i \tag{14.17}$$

Based on this, the conditional density shown in Equation 14.16 becomes

$$f_{\mathbf{X}_i|\mathrm{S}_i}(\mathbf{x}_i|\mathrm{s}_i) = \frac{1}{\mathrm{s}_i^N}(2\pi)^{-N/2}|\Sigma_{\mathbf{z}i}|^{-1/2}\exp\left\{-\frac{1}{2}\frac{\alpha_{\mathbf{x}\mathbf{i}}}{s_i^2}\right\} \tag{14.18}$$

Since the Gaussian vectors $\{\mathbf{Z}_1 \ldots \mathbf{Z}_m\}$ are independent, the conditional vectors $\{\mathbf{X}_1/S_1 \ldots \mathbf{X}_M/S_M\}$ are also independent. This means that the joint conditional density of the M vectors is given by the product of the individual conditional densities, or,

$$f_{\mathbf{X}_1 \ldots \mathbf{X}_M|S_1 \ldots S_M}(\mathbf{x}_1 \ldots \mathbf{x}_M|S_1 \ldots S_M) = \prod_{i=1}^{M} f_{\mathbf{X}_i|\mathbf{S}_i}(\mathbf{x}_i|\mathbf{s}_i) \tag{14.19}$$

For the sake of notational simplicity, define the matrix $\bar{\mathbf{X}}$ as

$$\bar{\mathbf{X}} = \{\mathbf{X}_1, \mathbf{X}_2 \ldots \mathbf{X}_M\} \tag{14.20}$$

and the vector $\mathbf{S}$ as

$$\mathbf{S} = \{S_1, S_2, \ldots S_M\} \tag{14.21}$$

Based on this convention, Equation 14.19 becomes

$$f_{\bar{\mathbf{X}}|\mathbf{s}}(\bar{\mathbf{x}}|\mathbf{s}) = \prod_{i=1}^{M} f_{\mathbf{X}_i|\mathbf{S}_i}(\mathbf{x}_i|\mathbf{s}_i) \tag{14.22}$$

Since the joint density of $\bar{\mathbf{X}}$ and $\mathbf{S}$ is given by

$$f_{\bar{\mathbf{X}},\mathbf{S}}(\bar{\mathbf{x}}, \mathbf{s}) = f_{\bar{\mathbf{X}}|\mathbf{S}}(\bar{\mathbf{x}}|\mathbf{s}) f_{\mathbf{S}}(\mathbf{s}) \tag{14.23}$$

the density of $\bar{\mathbf{X}}$ alone is

$$f_{\bar{\mathbf{X}}}(\bar{\mathbf{x}}) = \int_0^\infty \cdots \int_0^\infty f_{\bar{\mathbf{X}}|\mathbf{S}}(\bar{\mathbf{x}}|\mathbf{s}) f_{\mathbf{S}}(\mathbf{s}) \mathrm{d}s_1 \ldots \mathrm{d}s_M \tag{14.24}$$

Substituting Equation 14.18 and Equation 14.19 into Equation 14.24 yields

$$f_{\bar{\mathbf{X}}}(\bar{\mathbf{x}}) = K h_{NM}(\alpha_{x1}, \alpha_{x2}, \ldots, \alpha_{xM}) \tag{14.25}$$

where

$$K = (2\pi)^{-NM/2} \left(\prod_{i=1}^{M} |\Sigma_{zi}|^{-1/2} \right) \tag{14.26}$$

$$h_{NM}(\alpha_{x1}, \alpha_{x2}, \ldots, \alpha_{xM}) = \int_{\mathbf{s}} \left(\prod_{i=1}^{M} \frac{1}{s_i^N} \right) \exp\left\{ -\frac{1}{2} \sum_{i=1}^{M} \frac{\alpha_{xi}}{s_i^2} \right\} f_{\mathbf{S}}(\mathbf{s}) \mathrm{d}\mathbf{s} \tag{14.27}$$

$$\int_{\mathbf{s}} = \int_0^\infty \ldots \int_0^\infty (M\text{-fold integration}) \tag{14.28}$$

and

$$\mathrm{d}\mathbf{s} = \mathrm{d}s_1 \mathrm{d}s_2 \ldots \mathrm{d}s_M \tag{14.29}$$

As seen from Equation 14.25, the joint density of the M vectors is a function of the M scalar products $\{\alpha_{x1} \ldots \alpha_{xM}\}$. This joint density is completely specified by the covariance matrices of the Gaussian random vectors $\{\mathbf{Z}_1 \ldots \mathbf{Z}_M\}$ and the multivariate characteristic density function $f_{\mathbf{S}}(\mathbf{s})$.

14.2.3. A Comparison of the Traditional and Generalized Models

This section compares the models presented above with respect to the correlation and independence among the elements of an SIRV realization. An examination of the correlation and independence between two separate SIRV realizations then follows.

The traditional and generalized SIRV's presented in the two previous sections, allow for correlation between the elements of a given vector. In fact, the correlation among elements of SIRV $\mathbf{X}$ (or $\mathbf{X}_i$ in the generalized case) is determined by the correlation matrix $\Sigma_{\mathbf{z}}$ (or $\Sigma_{\mathbf{z}i}$) of the generating Gaussian vector. Since the density shown in Equation 14.9 cannot be decomposed into a product of factors, the traditional SIRV radar clutter model cannot allow for independence between any of the pulse train samples. Similarly, in the sonar case, the model cannot allow for independence between any of the spatial cells. However, the independence among the elements of a generalized SIRV can be accounted for by multiplying each element by a different random scalar.

In Equation 14.25 to Equation 14.29, this is accomplished by letting M equal the desired length of the SIRV and N equal 1. The independence of the generalized SIRV elements is now tied to the independence of the vector $\mathbf{S}$. If $f_{\mathbf{S}}(\mathbf{s})$ is separable, so then is $f_{\bar{\mathbf{X}}}(\bar{\mathbf{x}})$.

Let us consider the case of correlation and independence between two radar clutter cells or two sonar pulses. Independence arises through different mechanisms for the two models. For the traditional SIRV, independence results when independent realizations of the vector $\mathbf{Z}$ and scalar S are used. On the other hand, the generalized SIRV models independent cells (pings), when the elements of $f_{\mathbf{S}}(\mathbf{s})$ are independent.

The traditional SIRV can account for dependence between two cells (pings) only if the same realization of S modulates each, generating Gaussian vector;

$$\mathbf{X}_1 = \mathbf{Z}_1 S \tag{14.30}$$

and

$$\mathbf{X}_2 = \mathbf{Z}_2 S \tag{14.31}$$

In effect, this is the same as treating the two vectors $\mathbf{X}_1$ and $\mathbf{X}_2$ as one large vector, or,

$$\begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_1 \\ \mathbf{Z}_2 \end{bmatrix} S \tag{14.32}$$

All of the elements of this "double" vector are dependent, and their correlation is determined by the correlation of the vector $[\mathbf{Z}_1^{\mathrm{T}}, \mathbf{Z}_2^{\mathrm{T}}]^{\mathrm{T}}$.

Consider the correlation among the elements of two generalized SIRV's. The covariance of any given generalized SIRV is given by

$$\mathbf{K}_{X_i X_i} = E\{\mathbf{X}_i \mathbf{X}_i^{\mathrm{T}}\} - E\{\mathbf{X}_i\} E\{\mathbf{X}_i\}^{\mathrm{T}} \tag{14.33}$$

where $E\{\cdot\}$ is the expected value operator. From the generalized SIRV representation theorem, the vector $\mathbf{X}_i$ is

$$\mathbf{X}_i = \mathbf{Z}_i \mathbf{S}_i \tag{14.34}$$

Since $\mathbf{Z}_i$ is zero mean and independent from $\mathbf{S}_i$, Equation 14.33 reduces to

$$\mathbf{K}_{X_i X_i} = E\{\mathbf{Z}_i \mathbf{Z}_i^{\mathrm{T}}\} E\{\mathbf{S}_i^2\} \tag{14.35}$$

If $E\{\mathbf{s}_i^2\}$ is arbitrarily set equal to unity, then the covariance among the elements of a generalized SIRV becomes the covariance of the generating Gaussian vector, or,

$$\mathbf{K}_{X_i X_i} = \Sigma_{\mathbf{z}i} \tag{14.36}$$

Let us consider the covariance between two different generalized SIRV's,

$$\mathbf{K}_{X_iX_j} = E\{\mathbf{X}_i\mathbf{X}_j^{\mathrm{T}}\} - E\{\mathbf{X}_i\}E\{\mathbf{X}_j\}^{\mathrm{T}} \tag{14.37}$$

Using the same logic as that applied above yields

$$\mathbf{K}_{X_iX_j} = E\{\mathbf{Z}_i\mathbf{Z}_j^{\mathrm{T}}\}E\{S_i\,S_j\} \tag{14.38}$$

However, recall that the generalized generating Gaussian vectors are independent and have zero mean. This fact reduces Equation 14.38 to

$$\mathbf{K}_{X_iX_j} = 0 \tag{14.39}$$

As seen, despite the fact that the modulating scalars S_i and S_j may be correlated, the final SIRV's $\mathbf{X}_i$ and $\mathbf{X}_j$ are uncorrelated. Herein lies a potential limitation of the generalized model; although it can model two dependent SIRV's, these vectors must be uncorrelated. If correlated SIRV's are required, note that the generalized model can "double-up" the vectors as well, just as shown in Equation 2.32 for the traditional case. As such, either of the models can account for correlation between radar spatial cells or sonar pings.

To illustrate potential limitations of the traditional model, consider the case of sonar reverberation. Recall that the scalar S modulates a spatial patch of reverberation cells. The traditional model either assumes this scalar to be rapidly fluctuating between successive pings, giving rise to independent reverberations with each pulse, or not fluctuating at all. However, the generalized model can account for dependent fluctuations of this scalar through the joint density $f_{\mathbf{S}}(\mathbf{s})$.

TABLE 14.1
A Comparison of the Traditional and Generalized Models

Model	Sonar	Radar
Traditional	Cannot model independent spatial cells	Cannot model independent pulse train samples
Generalized	Can model independent spatial cells	Cannot model independent pulse train samples
Traditional	Can model dependent pings only if the same random scalar modulates each	Can model dependent cells only if the same random scalar modulates each
Generalized	Can model dependent pings	Can model dependent cells
Traditional	Can model correlated pings only if the same random scalar modulates each	Can model correlated cells only if the same random scalar modulates each
Generalized	Can model correlated pings only if the same random scalar modulates each	Can model correlated cells only if the same random scalar modulates each

For convenience, Table 14.1 summarizes the two SIRV models. The next section derives significant properties of the generalized SIRV model.

14.3. GENERALIZED SIRV PROPERTIES

This chapter derives a series of significant properties associated with the generalized SIRV model.

14.3.1. Linear Transformation

Generalized SIRV's remains invariant under a linear transformation. In order to prove this assertion, first pass the M SIRV's through M different linear transformations. The ith transformation is specified by a matrix $\mathbf{A}_i$, and a vector $\mathbf{b}_i$, such that

$$\begin{aligned} \mathbf{Y}_1 &= \mathbf{A}_1\mathbf{X}_1 + \mathbf{b}_1 = \mathbf{A}_1\mathbf{Z}_1 S_1 + \mathbf{b}_1, \\ &\vdots \\ \mathbf{Y}_M &= \mathbf{A}_M\mathbf{X}_M + \mathbf{b}_M = \mathbf{A}_M\mathbf{Z}_M S_M + \mathbf{b}_M. \end{aligned} \tag{14.40}$$

In this case the conditional density of any $\mathbf{Y}_i$ given S_i is

$$f_{\mathbf{Y}_i|\mathbf{S}_i}(\mathbf{y}_i|s_i) = \frac{1}{s_i^N}(2\pi)^{-N/2}|(\mathbf{A}_i\Sigma_{\mathbf{z}i}\mathbf{A}_i)|^{-1/2}\exp\left\{-\frac{1}{2}\frac{(\mathbf{A}_i\mathbf{y}_i+\mathbf{b}_i)^{\mathrm{T}}\Sigma_{\mathbf{z}i}^{-1}(\mathbf{A}_i\mathbf{y}_i+\mathbf{b}_i)}{s_i^2}\right\} \tag{14.41}$$

Define the scalar α_{yi} as

$$\alpha_{yi} = (\mathbf{A}_i\mathbf{y}_i+\mathbf{b}_i)^{\mathrm{T}}\Sigma_{\mathbf{z}i}^{-1}(\mathbf{A}_i\mathbf{y}_i+\mathbf{b}_i) \tag{14.42}$$

With this definition, the derivation of the density of the matrix

$$\bar{\mathbf{Y}} = \{\mathbf{Y}_1, \mathbf{Y}_2, \ldots \mathbf{Y}_M\} \tag{14.43}$$

follows the derivation of $f_{\bar{\mathbf{X}}}(\bar{\mathbf{x}})$ given in Section 14.2.2. Based on this work, the density of $\bar{\mathbf{Y}}$ is given by

$$f_{\bar{\mathbf{Y}}}(\bar{\mathbf{y}}) = Kh_{NM}(\alpha_{\mathbf{Y}_1}, \alpha_{\mathbf{Y}_2}, \ldots \alpha_{\mathbf{Y}_M}) \tag{14.44}$$

where the constant K is

$$K = (2\pi)^{-NM/2}\left(\prod_{i=1}^{M}|(\mathbf{A}_i\Sigma_{\mathbf{z}i}\mathbf{A}_i)|^{-1/2}\right) \tag{14.45}$$

and $h_{NM}(\cdot)$ is the exact same function as that shown in Equation 14.27. As seen, passing generalized SIRV's through a set of linear transformations, yields another set of generalized SIRV's with the same characteristic nonlinear function $h_{NM}(u_1, \ldots, u_M)$. Since linear transformations can be used to transform each of the

covariance matrices $\Sigma_{\mathbf{z}i}$ into identity matrices, there is no loss in generality when assuming each of the quadratic forms α_{xi} to be spherically invariant, (i.e., $\alpha_{xi} = \mathbf{x}_i^{\mathrm{T}}\mathbf{x}_i$).

14.3.2. The Generalized SIRV "Bootstrap" Theorem

Higher order generalized SIRV's can be generated from lower order SIRV's. To help illustrate this point, define the vector $\mathbf{a}_x$ as

$$\mathbf{a}_x = \{\alpha_{x_1}, \alpha_{x_2}, \ldots, \alpha_{x_M}\} \tag{14.46}$$

where α_{xi} is defined in Equation 14.17. Based on this definition, the characteristic nonlinear function given in Equation 14.27 can be written as

$$h_{NM}(\mathbf{a}_x) = \int_{\mathbf{s}} \prod_{i=1}^{M} \left[s_i^{-N} \exp\left(\frac{-\alpha_{xi}}{2s_i^2} \right) \right] f_{\mathbf{S}}(\mathbf{s}) \mathrm{d}s \tag{14.47}$$

Now take a partial derivative of this function with respect to α_{x_1}. The result is

$$\frac{\partial h_{NM}(\mathbf{a}_x)}{\partial \alpha_{xi}} = \int_{\mathbf{s}} \left(\frac{-1}{2s_i^2} \right) \prod_{i=1}^{M} \left[s_i^{-N} \exp\left(\frac{-\alpha_{xi}}{2s_i^2} \right) \right] f_{\mathbf{S}}(\mathbf{s}) \mathrm{d}s \tag{14.48}$$

Taking more partial derivatives with respect to the other elements of $\mathbf{a}_x$ yields

$$\frac{\partial^M h_{NM}(\mathbf{a}_x)}{\partial \alpha_{xi} \ldots \partial \alpha_{xM}} = \int_{\mathbf{s}} \prod_{i=1}^{M} \left[\left(\frac{-1}{2s_i^2} \right) s_i^{-N} \exp\left(\frac{-\alpha_{xi}}{2s_i^2} \right) \right] f_{\mathbf{S}}(\mathbf{s}) \mathrm{d}s \tag{14.49}$$

Simplifying Equation 14.49 gives

$$\frac{\partial^M h_{NM}(\mathbf{a}_x)}{\partial \alpha_{xi} \ldots \partial \alpha_{xM}} = \left(-\frac{1}{2} \right)^M \int_{\mathbf{s}} \prod_{i=1}^{M} \left[s_i^{-(N+2)} \exp\left(\frac{-\alpha_{xi}}{2s_i^2} \right) \right] f_{\mathbf{S}}(\mathbf{s}) \mathrm{d}s \tag{14.50}$$

Compare the right side of Equation 14.50 with that of Equation 14.47. As seen, the M-fold integral in Equation 14.50 is of the same form as that in Equation 14.47 with N replaced by $(N + 2)$. It follows that

$$\frac{\partial^M h_{NM}(\mu_1, \ldots, \mu_{\boldsymbol{M}})}{\partial \mu_1 \ldots \partial \mu_{\boldsymbol{M}}} = \left(-\frac{1}{2} \right)^M h_{(N+2)M}(\mu_1, \ldots, \mu_{\boldsymbol{M}}) \tag{14.51}$$

Rearranging terms in Equation 14.51 yields

$$h_{(N+2)M}(\mu_1, \ldots, \mu_{\boldsymbol{M}}) = (-2)^M \frac{\partial^M h_{NM}(\mu_1, \ldots, \mu_{\boldsymbol{M}})}{\partial \mu_1 \ldots \partial \mu_{\boldsymbol{M}}} \tag{14.52}$$

Equation 14.52 is the generalized SIRV "bootstrap" theorem. As seen, the characteristic nonlinear function of a generalized SIRV of order $(N + 2)M$ can be generated from one of order NM. Consequently, the characteristic nonlinear function of all generalized SIRV's with an even number of elements can be generated from $h_{2M}(\cdot)$, and all odd order functions can be generated from $h_{1M}(\cdot)$.

Note that the bootstrap theorem is not valid if the density $f_{\mathbf{S}}(\mathbf{s})$ depends on the value of $\mathbf{N}$. This can be readily seen by substituting $f_{\mathbf{S}}(\mathbf{s}, N)$ for $f_{\mathbf{S}}(\mathbf{s})$ in the above analysis. With this substitution, the right side of Equation 14.50 is not equal to a constant times $h_{(N+2)M}(\mathbf{a}_x)$ because the density under the integral is still $f_{\mathbf{S}}(\mathbf{s}, N)$, not $f_{\mathbf{S}}(\mathbf{s}, N + 2)$.

14.3.3. The Monotonicity of $h_{NM}(\alpha_{x1},\ldots,\alpha_{xM})$

For the purposes of this dissertation, a multivariate function $f(x_1, \ldots, x_M)$ is defined as monotonically nonincreasing if

$$f(x_1 + \delta_1, \ldots x_M + \delta_M) \leq f(x_1, \ldots, x_M) \tag{14.53}$$

for all values $(x_1, \ldots, x_M)$ and for all nonnegative perturbations

$$(\delta_1, \ldots, \delta_M) \geq 0 \tag{14.54}$$

From Equation 14.27, the characteristic nonlinear function of a generalized SIRV is

$$h_{NM}(\alpha_{x1}, \alpha_{x2}, \ldots, \alpha_{xM}) = \int_{\mathbf{s}} \left(\prod_{i=1}^{M} \frac{1}{\mathbf{s}_i^N} \right) \exp\left\{ -\frac{1}{2} \sum_{i=1}^{M} \frac{\alpha_{xi}}{s_i^2} \right\} f_{\mathbf{S}}(\mathbf{s})\, \mathrm{ds} \tag{14.55}$$

Adding a nonnegative perturbation along each axis yields

$$h_{NM}(\alpha_{x1} + \delta_1, \ldots, \alpha_{xM} + \delta_M) = \int_{\mathbf{s}} \left(\prod_{i=1}^{M} \frac{1}{s_i^N} \right) \exp\left\{ -\frac{1}{2} \sum_{i=1}^{M} \frac{\alpha_{xi} + \delta_i}{s_i^2} \right\} f_{\mathbf{S}}(\mathbf{s})\, \mathrm{ds} \tag{14.56}$$

Since all values in the set $\{\delta_1, \ldots \delta_M\}$ are nonnegative, the exponential term in Equation 14.56 is always less than or equal to that in Equation 14.55. Also, since all elements of $\mathbf{S}$ are nonnegative, the integrand in Equation 14.56 is always less than or equal to that in Equation 14.55. This means that

$$h_{NM}(\alpha_{x1} + \delta_1, \ldots, \alpha_{xM} + \delta_M) \leq h_{NM}(\alpha_{x1}, \ldots, \alpha_{xM}) \tag{14.57}$$

and the characteristic nonlinear function of a generalized SIRV is monotonically nonincreasing.

14.3.4. Spherical Coordinates

The conversion of the traditional SIRV from a representation in Cartesian coordinates to one involving spherical coordinates is presented in Ref. 14. This

section follows a similar analysis to convert the generalized SIRV into spherical coordinates.

Consider a traditional SIRV $\mathbf{X}$ with a characteristic nonlinear function $h_N(\alpha_x)$. As explained in Ref. 14, the relations between cartesian and spherical coordinates are given by

$$\begin{aligned}
\langle \mathbf{X} \rangle_1 &= R \cdot \cos(\phi_1) \\
\langle \mathbf{X} \rangle_k &= R \cdot \cos(\phi_k) \prod_{j=1}^{k-1} \sin(\phi_j), \qquad \text{for all } k \in [2, \ldots, N-2] \\
\langle \mathbf{X} \rangle_{N-1} &= R \cdot \cos(\theta) \prod_{j=1}^{N-2} \sin(\phi_j)
\end{aligned}$$

and

$$\langle \mathbf{X} \rangle_N = R \cdot \sin(\theta) \prod_{j=1}^{N-1} \sin(\phi_j) \tag{14.58}$$

where $\langle \mathbf{X} \rangle_i$ is the ith element of $\mathbf{X}$.

If a random vector $\mathbf{Y}$ is a function of another random vector $\mathbf{X}$ such that

$$\mathbf{Y} = g(\mathbf{X}) \tag{14.59}$$

then the density of the vector $\mathbf{Y}$ is

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(g^{-1}(\mathbf{Y}))/J \tag{14.60}$$

where J is the Jacobian of the transformation. The Jacobian is defined by

$$J = \left\| \begin{bmatrix} \dfrac{\partial y_1}{\partial x_1} & \cdots & \dfrac{\partial y_1}{\partial x_N} \\ \vdots & & \vdots \\ \dfrac{\partial y_N}{\partial x_1} & \cdots & \dfrac{\partial y_N}{\partial x_N} \end{bmatrix} \right\| \tag{14.61}$$

From Ref. 8, the Jacobian of the transformation shown in Equation 14.58 is

$$J = \left[R^{N-1} \prod_{k=1}^{N-2} \sin^{N-1-k}(\phi_k) \right]^{-1} \tag{14.62}$$

Assume that the vector $\mathbf{X}$ has zero mean and an indentity covariance matrix. From Equation 14.49 and Equation 14.50, the density of $\mathbf{X}$ is

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{N/2}} h_N(\alpha_{\mathbf{x}}) \tag{14.63}$$

where

$$\alpha_{\mathbf{x}} = x_1^2 + \ldots + x_{N^2} = r^2 \tag{14.64}$$

Based on Equation 14.60 and Equation 14.62, the joint density of $\{R, \theta, \phi_1, \ldots \phi_{N-2}\}$ is

$$f_{(R,\Theta,\Phi_1,\ldots,\Phi_{N-2})}(r, \theta, \phi_1, \ldots, \phi_{N-2}) = \frac{r^{N-1}}{(2\pi)^{N/2}} h_N(r^2) \prod_{k=1}^{N-2} \sin^{N-1-k}(\phi_k) \tag{14.65}$$

where it is recognized that

$$\alpha_{\mathbf{x}} = r^2 \tag{14.66}$$

Equation 14.65 reveals that R, the envelope of the vector $\mathbf{X}$, is independent of the variables $\{\theta, \phi_1, \ldots, \phi_{N-2}\}$ and has a density of

$$f_R(r) = \frac{r^{N-1}}{2^{(N-2)/2}\Gamma(N/2)} h_N(r^2) u(r) \tag{14.67}$$

where $\Lambda(\cdot)$ is the Gamma function and $u(\cdot)$ is the unit step function. The densities of θ and $\{\phi_1, \ldots \phi_{N-2}\}$ are also derived from Equation 14.65. These densities are

$$f_\Theta(\theta) = (2\pi)^{-1} \cdot [u(\theta) - u(\theta - 2\pi)] \tag{14.68}$$

and

$$f_{\Phi_k}(\phi_k) = \frac{\Gamma\left(\frac{N-k+1}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{N-k}{2}\right)} \sin^{N-1-k}(\phi_k) \cdot [u(\phi_k) - u(\phi_k - \pi)] \tag{14.69}$$

As seen, all of the angles is the set $\{\theta, \phi_1, \ldots \phi_{N-2}\}$ are independent.

Now consider a set of generalized SIRVs $\bar{\mathbf{X}}$ with the characteristic nonlinear function $h_{NM}(\alpha_{x1}, \ldots, \alpha_{xM})$. The relations between cartesian and spherical coordinates are given by

$$\langle \mathbf{X_i} \rangle_1 = R_i \cdot \cos(\phi_{i,1})$$

$$\langle \mathbf{X_i} \rangle_k = R_i \cdot \cos(\phi_{i,k}) \cdot \prod_{j=1}^{k-1} \sin(\phi_{i,j}), \qquad \text{for all } k \in [2 \ldots N-2]$$

$$\langle \mathbf{X_i} \rangle_{N-1} = R_i \cdot \cos(\theta_i) \cdot \prod_{j=1}^{N-2} \sin(\phi_{i,j})$$

and

$$\langle \mathbf{X_i} \rangle_N = R_i \cdot \sin(\theta_i) \cdot \prod_{j=1}^{N-1} \sin(\phi_{i,j}) \tag{14.70}$$

where: $\langle \mathbf{X_i} \rangle_j$ is the jth element of the ith vector $\mathbf{X_i}$ and $(i = 1, \ldots, M)$.

Define J_i, the Jacobian of the transformation from the vector $\mathbf{X}_i$ to the set $\{R_i, \theta_i, \phi_{i,1}, \ldots, \phi_{i,N-2}\}$, as the determinant of a matrix $\mathbf{A}_i$, or

$$J_i = |\mathbf{A}_i| = \left\| \begin{bmatrix} \dfrac{\partial R_i}{\partial \langle \mathbf{X}_i \rangle_1} & \dfrac{\partial \theta_i}{\partial \langle \mathbf{X}_i \rangle_1} & \dfrac{\partial \phi_{i,1}}{\partial \langle \mathbf{X}_i \rangle_1} & \cdots & \dfrac{\partial \phi_{i,N-2}}{\partial \langle \mathbf{X}_i \rangle_1} \\ \vdots & \vdots & \vdots & & \vdots \\ \dfrac{\partial R_i}{\partial \langle \mathbf{X}_i \rangle_N} & \dfrac{\partial \theta_i}{\partial \langle \mathbf{X}_i \rangle_N} & \dfrac{\partial \theta_i}{\partial \langle \mathbf{X}_i \rangle_N} & \cdots & \dfrac{\partial \phi_{i,N-2}}{\partial \langle \mathbf{X}_i \rangle_N} \end{bmatrix} \right\| \tag{14.71}$$

Since the spherical coordinate transformation of any generalized SIRV $\mathbf{X}_i$ in the matrix $\bar{\mathbf{X}}$ does not depend on the spherical coordinates of any other vector in the matrix, and Jacobian of the entire matrix transformation is

$$J_{\bar{\mathbf{X}}} = \left\| \begin{bmatrix} \mathbf{A}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{A}_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \mathbf{A}_M \end{bmatrix} \right\| \tag{14.72}$$

Simplification of Equation 14.72 yields

$$J_{\bar{\mathbf{X}}} = |\mathbf{A}_1| \cdot |\mathbf{A}_2| \cdots |\mathbf{A}_M| \tag{14.73}$$

Equation 14.62 gives the determinant of any matrix $\mathbf{A}_i$ as

$$J_i = \left[R_i^{N-1} \cdot \prod_{k=1}^{N-2} \sin^{N-1-k}(\phi_{i,k}) \right]^{-1} \tag{14.74}$$

Based on Equation 14.73 and Equation 14.74, the Jacobian of the transformation of the matrix $\bar{\mathbf{X}}$ into spherical coordinate is

$$J_{\bar{\mathbf{X}}} = \prod_{i=1}^{M} \left[R_i^{N-1} \cdot \prod_{k=1}^{N-2} \sin^{N-1-k}(\phi_{i,k}) \right]^{-1} \tag{14.75}$$

For the sake of notational simplicity, define the vector $\mathbf{R}$ as

$$\mathbf{R} = [R_1, \ldots, R_M]^{\mathrm{T}} \tag{14.76}$$

the vector $\bar{\Theta}$ as,

$$\bar{\Theta} = [\theta_1, ..., \theta_M]^{\mathrm{T}} \tag{14.77}$$

and the matrix $\bar{\Phi}$ as

$$\bar{\Phi} = \begin{bmatrix} \phi_{1,1} & \cdots & \phi_{1,N-2} \\ & \vdots & \\ \phi_{M,1} & \cdots & \phi_{M,N-2} \end{bmatrix} \tag{14.78}$$

The transformation in Equation 14.70 can now be denoted as

$$\bar{X} = g^{-1}(\mathbf{r}, \bar{\theta}, \bar{\phi}) \tag{14.79}$$

The density of the transformed matrix is

$$f_{\mathbf{R},\bar{\Theta},\bar{\Phi}}(\mathbf{r}, \bar{\theta}, \bar{\phi}) = \left. \frac{f_{\bar{\mathbf{X}}}(\bar{\mathbf{x}})}{\mathbf{J}_{\bar{\mathbf{X}}}} \right|_{\bar{\mathbf{x}}=\mathbf{g}^{-1}(\mathbf{r},\bar{\theta},\bar{\phi})} \tag{14.80}$$

From Equation 14.25, the density of the matrix $\bar{\mathbf{X}}$ is

$$f_{\bar{\mathbf{X}}}(\bar{\mathbf{x}}) = Kh_{NM}(\alpha_{x1}, \alpha_{x2}, \ldots \alpha_{xM}) \tag{14.81}$$

Substituting Equation 14.75 and Equation 14.81 into Equation 14.80 produces the final density function of the spherically transformed matrix. Once again, all of the angles are independent, and their densities resemble those for a traditional SIRV given in Equation 14.68 and Equation 14.69,

$$f_{\Theta_i}(\theta_i) = (2\pi)^{-1}[u(\theta_i) - u(\theta_i - 2\pi)] \tag{14.82}$$

and

$$f_{\Phi_{i,k}}(\phi_{i,k}) = \frac{\Gamma\left(\frac{N-k+1}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{N-k}{2}\right)} \sin^{N-1-k}(\phi_{i,k})[u(\phi_{i,k}) - u(\phi_{i,k} - \pi)] \tag{14.83}$$

Not only are the angles independent from each other, but those are also independent of the envelope vector, **R**. However, the elements of **R** are not necessarily independent. From Equation 14.75, Equation 14.80, and Equation 14.81 the multivariate density of the envelope is

$$f_{\mathbf{R}}(r) = \frac{\left(\prod_{i=1}^{M} r_i\right)}{K_r} h_{NM}(r_1^2, \ldots, r_M^2)u(r_1)u(r_2)\ldots u(r_M) \tag{14.84}$$

where: K_r is a normalizing constant. Equation 14.84 reveals that the envelopes of the vectors in the generalized SIRV matrix $\bar{\mathbf{X}}$ are independent if and only if the vectors themselves are independent.

14.3.5. The Generalized SIRV Bessel Function Representation

The characteristic nonlinear function of a generalized SIRV, $h_{NM}(u_1, \ldots, u_M)$, can be expressed in terms of the vector's characteristic function $\phi(\cdot)$ and a product of Bessel functions. From Equation 14.47, the characteristic nonlinear function is

$$h_{NM}(a_x) = \int_{\mathbf{s}} \left(\prod_{i=1}^{M} \frac{1}{s_i^N} \right) \exp\left\{ \frac{1}{2} \sum_{i=1}^{M} \frac{\alpha_{xi}}{s_i^2} \right\} f_{\mathbf{S}}(\mathbf{s}) \mathrm{d}\mathbf{s} \tag{14.85}$$

where

$$\mathbf{a}_x = \{\alpha_{x_1}, \alpha_{x_2}, \ldots, \alpha_{x_M}\} \tag{14.86}$$

Let N, the length of each vector in the matrix $\bar{\mathbf{X}}$, equal unity. Based on this, the matrix $\bar{\mathbf{X}}$ is really a vector, or,

$$\bar{\mathbf{X}} = \{X_1, \ldots, X_M\} \tag{14.87}$$

Define the multivariate characteristic function of this vector $\bar{\mathbf{X}}$ as

$$\psi(\mathbf{w}) = \int_{\bar{\mathbf{X}}} \boldsymbol{f}_{\bar{\mathbf{X}}}(\bar{\mathbf{x}}) \exp\{j[\omega_1 x_1 + \ldots + \omega_M x_M]\} \mathrm{d}\bar{\mathbf{x}} \tag{14.88}$$

where:

$$\mathbf{w} = \{\omega_1, \omega_2, \ldots \omega_M\} \tag{14.89}$$

$$\mathrm{d}\bar{\mathbf{x}} = \mathrm{d}x_1 \mathrm{d}x_2 \ldots \mathrm{d}x_M \tag{14.90}$$

and

$$\int_{\bar{\mathbf{X}}} = \int_{X_1} \cdots \int_{X_M} (M\text{-fold integration}) \tag{14.91}$$

Taking the multidimensional inverse Fourier transform of Equation 14.88 yields

$$f_{\bar{\mathbf{X}}}(\bar{\mathbf{x}}) = (2\pi)^{-M} \int_{\mathbf{W}} \psi(\mathbf{w}) \exp\{ -j[\omega_1 x_1 + \ldots + \omega_M \mathrm{x}_M]\} \mathrm{d}\mathbf{w} \tag{14.92}$$

where:

$$\mathrm{d}\mathbf{w} = \mathrm{d}\omega_1 \mathrm{d}\omega_2 \ldots \mathrm{d}\omega_M \tag{14.93}$$

and

$$\int_{\mathbf{W}} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (M\text{-fold integration}) \tag{14.94}$$

Assume that the elements of $\bar{\mathbf{X}}$ have zero mean and unit variance. From Equation 14.25 and Equation 14.26, it follows that the density of $\bar{\mathbf{X}}$ becomes

$$f_{\bar{\mathbf{X}}}(\bar{\mathbf{x}}) = (2\pi)^{-M/2} \cdot h_{1M}(\mathbf{a}_x) \tag{14.95}$$

where $\cdot$ the ith element of $\mathbf{a}_x$ is

$$\alpha_{xi} = x_i^2 \tag{14.96}$$

Substituting Equation 14.95 into Equation 14.92 yields

$$h_{1M}(\mathbf{a}_x) = (2\pi)^{-M/2} \cdot \int_{\mathbf{W}} \psi(\mathbf{w}) \exp\{ -j[\omega_1 x_1 + \ldots + \omega_M x_M]\} \mathrm{d}\mathbf{w} \tag{14.97}$$

Substituting Equation 14.96 into Equation 14.97 produces

$$h_{1M}(\mathbf{a}_x) = (2\pi)^{-M/2} \int_{\mathbf{W}} \psi(\mathbf{w}) \prod_{i=1}^{M} \exp\{-j\omega_i \sqrt{\alpha_{xi}}\} \mathrm{d}\mathbf{w} \tag{14.98}$$

or

$$h_{1M}(\mathbf{a}_x) = (2\pi)^{-M/2} \cdot \int_{\mathbf{W}} \psi(\mathbf{w}) \cdot \prod_{i=1}^{M} \{\cos(\omega_i \sqrt{\alpha_{xi}}) - j \sin(\omega_i \sqrt{\alpha_{xi}})\} \mathrm{d}\mathbf{w} \tag{14.99}$$

Carrying out the product shown in Equation 14.99 yields

$$\begin{aligned} h_{1M}(\mathbf{a}_x) = {} & (2\pi)^{-M/2} \cdot \int_{\mathbf{W}} \psi(\mathbf{w}) \cdot \prod_{i=1}^{M} \cos(\omega_i \sqrt{\alpha_{xi}}) \, \mathrm{d}\mathbf{w} \\ & + (2\pi)^{-M/2} \cdot (-j)^M \cdot \int_{\mathbf{W}} \psi(\mathbf{w}) \cdot \prod_{i=1}^{M} \sin(\omega_i \sqrt{\alpha_{xi}}) \, \mathrm{d}\mathbf{w} \\ & + (2\pi)^{-M/2} \cdot \int_{\mathbf{W}} \psi(\mathbf{w}) \cdot \{\text{cosine/sine cross-terms}\} \mathrm{d}\mathbf{w} \end{aligned} \tag{14.100}$$

For the middle integral of Equation 14.100, separation of the M-fold product gives

$$\begin{aligned} \int_{\mathbf{W}} \psi(\mathbf{w}) \cdot \prod_{i=1}^{M} \sin(\omega_i \sqrt{\alpha_{xi}}) \mathrm{d}\mathbf{w} = {} & \left\{ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \psi(\mathbf{w}) \sin(\omega_i \sqrt{\alpha_{xi}}) \, \mathrm{d}\omega_1 \right\} \cdot \\ & \times \prod_{i=1}^{M} \sin(\omega_i \sqrt{\alpha_{xi}}) \mathrm{d}\omega_1 \ldots \mathrm{d}\omega_M \end{aligned} \tag{14.101}$$

Equation 14.88 gives the multivariate characteristic function as

$$\psi(\omega_1, \ldots, \omega_M) = \int_{\bar{\mathbf{X}}} f_{\bar{\mathbf{X}}}(\bar{\mathbf{x}}) \exp\{j[\omega_1 x_1 + \ldots + \omega_M x_M]\} \mathrm{d}\bar{\mathbf{x}} \tag{14.102}$$

Substituting $-\omega_1$ for ω_1 gives

$$\psi(-\omega_1,\ldots,\omega_M)=\int_{\bar{\mathbf{X}}}\boldsymbol{f}_{\bar{\mathbf{X}}}(x_1,\ldots,x_M)\exp\{j[-\omega_1x_1+\ldots+\omega_Mx_M]\}\mathrm{d}\bar{\mathbf{x}} \quad (14.103)$$

Letting the variable z equal $-x_1$ produces

$$\phi(-\omega_1,\ldots,\omega_M)=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\boldsymbol{f}_{\bar{\mathbf{X}}}(-z,x_2,\ldots,x_M)\exp\{j[\omega_1z+\omega_2x_2\ldots + \omega_Mx_M]\}\mathrm{d}z\,\mathrm{d}x_2\ldots\mathrm{d}x_M \quad (14.104)$$

However, note that Equation 14.95 to Equation 14.97 show $f_{\bar{\mathbf{X}}}(\bar{\mathbf{x}})$ as a function of x_i^2, for each i from 1 to M. This means that $f_{\bar{\mathbf{X}}}(\bar{\mathbf{x}})$ is an even function with respect to all axes, and

$$f_{\bar{\mathbf{X}}}(-z,x_2,\ldots,x_M)=f_{\bar{\mathbf{X}}}(z,x_2,\ldots,x_M) \quad (14.105)$$

Substituting this into Equation 14.104 and comparing the result to Equation 14.102 produces

$$\psi(-\omega_1,\ldots,\omega_M)=\psi(\omega_1,\ldots,\omega_M) \quad (14.106)$$

Equation 14.106 reveals that $\psi(\mathbf{w})$ is even with respect to the ω_1 axis. A similar analysis can be performed to show that $\psi(\mathbf{w})$ is even with respect to all other axes as well. Since $\psi(\mathbf{w})$ is even and $\sin(\cdot)$ is odd, the bracketed integral in Equation 3.62 is zero. This renders the entire integral zero, or

$$\int_{\mathrm{W}}\psi(\mathbf{w})\prod_{i=1}^{M}\sin\left(\omega_i\sqrt{\alpha_{xi}}\right)\mathrm{d}\mathbf{w}=0 \quad (14.107)$$

Equation 14.107 shows that middle integral of Equation 14.100 is zero. A similar analysis shows that the final integral is also zero, since a single odd sine term can be bracketed off with the even $\psi(\mathbf{w})$ here as well. This means that the characteristic nonlinear function is given by

$$h_{1M}(\mathbf{a}_x)=(2\pi)^{-M/2}\int_{\mathrm{W}}\psi(\mathbf{w})\prod_{i=1}^{M}\cos\left(\omega_i\sqrt{\alpha_{xi}}\right)\mathrm{d}\mathbf{w} \quad (14.108)$$

Recall that the Bessel Function $J_{-1/2}(\cdot)$ is related to the cosine function by

$$\cos(\theta)=\sqrt{\frac{\pi}{2}}\sqrt{\theta}\mathbf{J}_{-1/2}(\theta) \quad (14.109)$$

Substituting Equation 14.109 into Equation 14.108 yields

$$h_{1M}(\mathbf{a}_x)=2^{-M}\int_{\mathrm{W}}\psi(\mathbf{w})\prod_{i=1}^{M}\left\{\sqrt{\left(\omega_i\sqrt{\alpha_{xi}}\right)}\mathbf{J}_{-1/2}\left(\omega_i\sqrt{\alpha_{xi}}\right)\right\}\mathrm{d}\mathbf{w} \quad (14.110)$$

The limits on the M dimensional integral shown in Equation 14.110 are from minus infinity to infinity. Since the integrand is even, Equation 14.110 can be rewritten as

$$h_{1M}(\mathbf{a}_x) = \int_0^\infty \cdots \int_0^\infty \psi(\mathbf{w}) \cdot \prod_{i=1}^{M} \left\{ \sqrt{(\omega_i \sqrt{\alpha_{xi}})} \cdot \mathbf{J}_{-1/2}(\omega_i \sqrt{\alpha_{xi}}) \right\} \mathrm{d}\mathbf{w} \qquad (14.111)$$

Equation 14.111 expresses the generalized characteristic nonlinear function in terms of the characteristic function $\psi(\mathbf{w})$ and a product of Bessel functions, for the case where the length of each SIRV in the matrix $\bar{\mathbf{X}}$ is unity. Higher odd-order functions can now be generated by use of the bootstrap theorem. As shown below, these higher odd-order functions are given by

$$h_{(2N-1)M}(\mu_1, \ldots, \mu_M)$$

$$= \left\{ \prod_{i=1}^{M} \sqrt{\mu_i} \right\}^{\frac{3}{2}-N} \cdot \int_0^\infty \cdots \int_0^\infty \psi(\mathbf{w}) \cdot \prod_{i=1}^{M} \left\{ \omega_i^{N-\frac{1}{2}} \cdot \mathbf{J}_{N-\frac{3}{2}}(\omega_i \sqrt{\mu_i}) \right\} \mathrm{d}\mathbf{w} \qquad (14.112)$$

The following inductive proof verifies Equation 14.112. Equation 14.111 shows that the hypothesis holds for N equal to one. Next, assume that Equation 14.112 holds for some value of N. If Equation 14.112 can be shown to hold for $(N+1)$ as well, then the inductive proof is complete.

To show that Equation 14.112 holds for $(N+1)$, first take a series of partial derivatives with respect to $(\mu_1, \ldots, \mu_M)$. As provided by Ref. 14, an identity which helps in taking these derivatives is,

$$\frac{\partial}{\partial \alpha} \left[\sqrt{\alpha}^{\frac{3}{2}-N} \cdot \mathbf{J}_{N-\frac{3}{2}}(\omega \sqrt{\alpha}) \right] = \frac{-\omega}{2} \cdot \sqrt{\alpha}^{\frac{1}{2}-N} \cdot \mathbf{J}_{N-\frac{1}{2}}(\omega \sqrt{\alpha}) \qquad (14.113)$$

Based on Equation 14.112 and Equation 14.113, the derivative of $h_{(2N-1)M}(\mu_1, \ldots, \mu_M)$ with respect to μ_1 is

$$\frac{\partial h_{(2N-1)M}(\mu_1, \ldots, \mu_M)}{\partial \mu_1}$$

$$= \int_0^\infty \cdots \int_0^\infty \psi(\mathbf{w}) \left(\frac{-1}{2} \right) \omega_1^{N+\frac{1}{2}} \sqrt{\mu_1}^{\frac{1}{2}-N} \mathbf{J}_{N-\frac{1}{2}}(\omega_1 \sqrt{\mu_1})$$

$$\times \prod_{i=2}^{M} \left\{ \omega_i^{N-\frac{1}{2}} \sqrt{\mu_1}^{\frac{3}{2}-N} \cdot \mathbf{J}_{N-\frac{3}{2}}(\omega_i \sqrt{\mu_i}) \right\} \mathrm{d}\mathbf{w} \qquad (14.114)$$

Taking the remaining partial derivatives yields

$$\frac{\partial h_{(2N-1)M}(\mu_1,\ldots,\mu_M)}{\partial\mu_1\ldots\partial\mu_M} = \int_0^\infty\cdots\int_0^\infty \psi(\mathbf{w})\left(\frac{-1}{2}\right)\cdot\prod_{i=1}^{M}\left\{\omega_i^{N+\frac{1}{2}}\sqrt{\mu_1}^{\frac{1}{2}-N}\mathbf{J}_{N-\frac{1}{2}}\left(\omega_i\sqrt{\mu_i}\right)\right\}\mathrm{d}\mathbf{w} \quad (14.115)$$

This can be further simplified by use of the bootstrap theorem which states that

$$h_{(N+2)M}(\mu_1,\ldots,\mu_M) = (-2)^M\frac{\partial^M h_{NM}(\mu_1,\ldots,\mu_M)}{\partial\mu_1\ldots\partial\mu_M} \quad (14.116)$$

Substituting Equation 14.116 into Equation 14.115 gives

$$h_{(N+2)M}(\mu_1,\ldots,\mu_M) = \left\{\prod_{i=2}^{M}\sqrt{\mu_i}\right\}^{\frac{1}{2}-N}\int_0^\infty\cdots\int_0^\infty\psi(\mathbf{w})\cdot\prod_{i=2}^{M}\left\{\omega_i^{N+\frac{1}{2}}\mathbf{J}_{N-\frac{1}{2}}\left(\omega_i\sqrt{\mu_i}\right)\right\}\mathrm{d}\mathbf{w} \quad (14.117)$$

Equation 14.117 concludes the inductive proof because this is the same result as that which is obtained by letting (N) equal $(N+1)$ in Equation 14.112.

Note that Equation 14.112 derives odd order characteristic nonlinear functions. A similar result is provided for even order traditional SIRV's in Ref. 14. The extension of this work to higher even order generalized SIRV's follows in the same manner as that shown above for odd orders. The result is

$$h_{(N+2)M}(\mu_1,\ldots,\mu_M) = \left\{\prod_{i=1}^{M}\sqrt{\mu_i}\right\}^{-N}\int_0^\infty\cdots\int_0^\infty\psi(\mathbf{w})\prod_{i=1}^{M}\left\{\omega_i^{N+1}\mathbf{J}_N\left(\omega_i\sqrt{\mu_i}\right)\right\}\mathrm{d}\mathbf{w} \quad (14.118)$$

Equation 14.118 and Equation 14.112 can be combined to yield an expression for any value of N, odd or even. The result is

$$h_{(N+2)M}(\mu_1,\ldots,\mu_M) = \left\{\prod_{i=1}^{M}\sqrt{\mu_i}\right\}^{1-\frac{N}{2}}\int_0^\infty\cdots\int_0^\infty\psi(\mathbf{w})\cdot\prod_{i=1}^{M}\left\{\omega_i^{\frac{N}{2}}\mathbf{J}_{\frac{N}{2}-1}\left(\omega_i\sqrt{\mu_i}\right)\right\}\mathrm{d}\mathbf{w} \quad (14.119)$$

Equation 14.119 thus shows how to represent $h_{NM}(\mu_1,\ldots,\mu_M)$ in terms of the multivariate characteristic function $\psi(\mathbf{w})$ and a product of Bessel functions. Note that $\psi(\mathbf{w})$ is derived from the density $f_{\bar{\mathbf{X}}}(\bar{\mathbf{x}})$ which corresponds to the case where the length of each SIRV in the matrix $\bar{\mathbf{X}}$ equals unity.

14.3.6. Minimum Mean Square Error Estimation

As stated earlier, the generalized SIRV matrix $\bar{\mathbf{X}}$ contains a set of M column vectors, each of length N. The result of partitioning each of these column vectors into two sub-vectors is

$$\mathbf{X}_i = \begin{bmatrix} \mathbf{X}_{ai} \\ \mathbf{X}_{bi} \end{bmatrix} \qquad i \in [1, \ldots, M] \tag{14.120}$$

where $\mathbf{X}_{ai}$ is of dimension $[K \times 1]$ and $\mathbf{X}_{bi}$ is of dimension $[(N - K) \times 1]$. Based on this, the matrix $\bar{\mathbf{X}}$ can be partitioned into the two sub-matrices

$$\bar{\mathbf{X}}_a = \{\mathrm{X}_{a1}, \ldots, \mathrm{X}_{aM}\} \tag{14.121}$$

and

$$\bar{\mathbf{X}}_b = \{\mathrm{X}_{b1}, \ldots, \mathrm{X}_{bM}\} \tag{14.122}$$

where

$$\bar{\mathbf{X}} = \begin{bmatrix} \bar{\mathbf{X}}_a \\ \bar{\mathbf{X}}_b \end{bmatrix} \tag{14.123}$$

Note that $\bar{\mathbf{X}}_a$ is of dimension $[K \times M]$ and $\bar{\mathbf{X}}_b$ is of dimension $[(N - K) \times M]$. From Ref. 15, the minimum mean square error estimate of $\bar{\mathbf{X}}_b$ given $\bar{\mathbf{X}}_a$ is

$$\hat{\bar{\mathbf{X}}}_{\mathrm{b}} = E[\bar{\mathbf{X}}_b | \bar{\mathbf{X}}_a] \tag{14.124}$$

The following analysis reveals that when $\bar{\mathbf{X}}$ is a generalized SIRV, this estimate is a linear function of the data vectors in the matrix $\bar{\mathbf{X}}_a$.

Let Σ_i, the covariance matrix of vector $\mathbf{X}_i$, equal

$$\Sigma_{\mathrm{i}} = \begin{bmatrix} \mathbf{C}_{aa_i} & \mathbf{C}_{ab_i} \\ \mathbf{C}_{ba_i} & \mathbf{C}_{bb_i} \end{bmatrix} \tag{14.125}$$

where $\mathbf{C}_{ab_i}$ and $\mathbf{C}_{ba_i}$ are the cross covariance matrices of $\mathbf{X}_{ai}$ and $\mathbf{X}_{bi}$, $\mathbf{C}_{aa_i}$ is the covariance matrix of $\mathbf{X}_{ai}$, and $\mathbf{C}_{bb_i}$ is the covariance matrix of $\mathbf{X}_{bi}$. Recall that the representation theorem gives the ith vector of the generalized SIRV matrix as

$$\mathbf{X}_i = \mathbf{Z}_i \cdot S_i \tag{14.126}$$

As explained in Ref. 14, the expected value of $(\mathbf{X}_{bi} | \mathbf{X}_{ai})$ conditioned on S_i is

$$E[(\mathbf{X}_{bi} | \mathbf{X}_{ai}) | S_i] = \mathbf{C}_{ba_i} \cdot \mathbf{C}_{aa_i}^{-1} \cdot \mathbf{x}_{ai} \tag{14.127}$$

Since all of the conditional vectors $[(\mathbf{X}_{bi}|\mathbf{X}_{ai})|S_i]$ are statistically independent, the expected value of the conditional matrix $[(\bar{\mathbf{X}}_b|\bar{\mathbf{X}}_a)|S_1 \ldots S_M]$ is

$$E[(\bar{\mathbf{X}}_b|\bar{\mathbf{X}}_a)|\mathbf{S}] = E[(\mathbf{X}_{b1}|\mathbf{X}_{a1})]|S_1, \ldots, (\mathbf{X}_{bM}|\mathbf{X}_{aM})]|S_M]$$
$$= E[[(\mathbf{X}_{b1}|\mathbf{X}_{a1})]|S_1], \ldots, E(\mathbf{X}_{bM}|\mathbf{X}_{aM})]|S_M]] \tag{14.128}$$

where:

$$\mathbf{S} = \{S_1, \ldots, S_M\} \tag{14.129}$$

Substituting Equation 14.127 into Equation 14.128 gives

$$E[(\bar{\mathbf{X}}_b|\bar{\mathbf{X}}_a)|\mathbf{S}] = [(\mathbf{C}_{ba_1} \cdot \mathbf{C}_{aa_1}^{-1}) \cdot \mathbf{X}_{a1}, \ldots, (\mathbf{C}_{ba_M} \cdot \mathbf{C}_{aa_M}^{-1}) \cdot \mathbf{X}_{aM}] \tag{14.130}$$

or

$$E[(\bar{\mathbf{X}}_b|\bar{\mathbf{X}}_a)|\mathbf{S}] = \sum_{i=1}^{M} (\mathbf{C}_{ba_i} \cdot \mathbf{C}_{aa_i}^{-1}) \cdot \Delta\mathbf{X}_i \tag{14.131}$$

where:

$$\Delta\mathbf{X}_i = [\bar{0}, \bar{0}, \ldots, \mathbf{X}_{ai}, \ldots, \bar{0}] \tag{14.132}$$

The expected value of $(\bar{\mathbf{X}}_b|\bar{\mathbf{X}}_a)$ is related to the conditional expectation shown in Equation 14.131 by

$$E(\bar{\mathbf{X}}_b|\bar{\mathbf{X}}_a) = E_S[(\bar{\mathbf{X}}_b|\bar{\mathbf{X}}_a)|\mathbf{S}] \tag{14.133}$$

or

$$E[\bar{\mathbf{X}}_b|\bar{\mathbf{X}}_a] = \int_0^\infty \cdots \int_0^\infty E[(\bar{\mathbf{X}}_b|\bar{\mathbf{X}}_a)]|\mathbf{S}] f_{\mathbf{S}}(\mathbf{s}) \mathrm{d}s_1 \ldots \mathrm{d}s_M \tag{14.134}$$

Substituting Equation 14.131 into Equation 14.134 gives

$$E[\bar{\mathbf{X}}_b|\bar{\mathbf{X}}_a = \bar{\mathbf{x}}_a] = \int_0^\infty \cdots \int_0^\infty \left\{ \sum_{i=1}^{M} (\mathbf{C}_{aa_i} \cdot \mathbf{C}_{aa_i}^{-1}) \cdot \Delta\mathbf{X}_i \right\} \cdot f_{\mathbf{S}}(\mathbf{s}) \mathrm{d}s_1 \ldots \mathrm{d}s_M \tag{14.135}$$

or

$$E[\bar{\mathbf{X}}_b|\bar{\mathbf{X}}_a = \bar{\mathbf{x}}_a] = \left\{ \sum_{i=1}^{M} (\mathbf{C}_{ba_i} \cdot \mathbf{C}_{aa_i}^{-1}) \cdot \Delta\mathbf{X}_i \right\} \cdot \int_0^\infty \cdots \int_0^\infty f_{\mathbf{S}}(\mathbf{s}) \mathrm{d}s_1 \ldots \mathrm{d}s_M \tag{14.136}$$

Since $f_{\mathbf{S}}(\mathbf{s})$ is a density, Equation 14.136 reduces to

$$E[\bar{\mathbf{X}}_b|\bar{\mathbf{X}}_a = \bar{\mathbf{x}}_a] = \left\{ \sum_{i=1}^{M} (\mathbf{C}_{ba_i} \cdot \mathbf{C}_{aa_i}^{-1}) \cdot \Delta\mathbf{X}_i \right\} \tag{14.137}$$

Substituting Equation 14.137 into Equation 14.124 yields

$$\hat{\bar{\mathbf{X}}}_b = \left\{ \sum_{i=1}^{M} (\mathbf{C}_{ba_i} \cdot \mathbf{C}_{aa_i}^{-1}) \cdot \Delta \mathbf{X}_i \right\} \tag{14.138}$$

Equation 14.138 thus reveals that the minimum mean square error estimate of the matrix $\bar{\mathbf{X}}_b$ is a linear function of the data vectors in the matrix $\bar{\mathbf{X}}_a$.

14.3.7. The Generalized SIRV Laplace Transform Representation

Recall that the M dimensional Laplace transform of $g(\mathbf{z})$ is given by

$$G(\mathbf{a}_x) = \int_z \prod_{i=1}^{M} \exp[-\alpha_{xi} z_i] \cdot g(\mathbf{z}) \mathrm{d}\mathbf{z} \equiv L^M\{g(\mathbf{z})\} \tag{14.139}$$

where

$$\mathbf{a}_x = (\alpha_{x1}, \ldots, \alpha_{xM}) \tag{14.140}$$

$$\mathbf{z} = (z_1, \ldots, z_M) \tag{14.141}$$

$$\mathrm{d}\mathbf{z} = \mathrm{d}z_1 \ldots \mathrm{d}z_M \tag{14.142}$$

and

$$\int_{\mathbf{z}} = \int_0^\infty \cdots \int_0^\infty (M\text{-fold integration}) \tag{14.143}$$

The following analysis will show how the Laplace transformation can simplify the expression for the generalized SIRV characteristic nonlinear function $h_{NM}(\mathbf{a}_x)$. From Equation 14.27, the characteristic nonlinear function is

$$h_{NM}(\mathbf{a}_x) = \int_s \left[\prod_{i=1}^{M} s_i^{-N} \exp\left\{ -\frac{1}{2} \frac{\alpha_{xi}}{s_i^2} \right\} \right] f_{\mathbf{S}}(\mathbf{s}) \mathrm{d}\mathbf{s} \tag{14.144}$$

where

$$\mathbf{a}_x = (\alpha_{x1}, \alpha_{x2}, \ldots, \alpha_{xM}) \tag{14.145}$$

In Equation 14.144, let

$$z_i = \frac{1}{2s_i^2} \tag{14.146}$$

and

$$\mathrm{d}s_i = -(2z_i)^{-3/2} \mathrm{d}z_i \tag{14.147}$$

for all i in the set (1,…,M). Substituting Equation 14.147 and Equation 14.146 into Equation 14.144 gives

$$h_{NM}(\mathbf{a}_x) = \int_{\mathbf{s}} \prod_{i=1}^{M} [\exp\{ -\alpha_{xi} z_i \}] \cdot \prod_{i=1}^{M} \left[(2z_i)^{-\left(\frac{N-3}{2}\right)} \right] f_{\mathbf{S}}\left(\frac{1}{\sqrt{2z_1}}, \ldots, \frac{1}{\sqrt{2z_M}} \right) \mathrm{d}\mathbf{z} \tag{14.148}$$

Define the function $g(\mathbf{z})$ as

$$g(\mathbf{z}) = \frac{1}{K_{\mathrm{S}}} \left[\prod_{i=1}^{M} (2z_i)^{-\left(\frac{N-3}{2}\right)} u(z_i) \right] f_{\mathbf{S}}\left(\frac{1}{\sqrt{2z_1}}, \ldots, \frac{1}{\sqrt{2z_M}} \right) \tag{14.149}$$

where: K_{s} is a constant. Notice that $g(\mathbf{z})$ appears in the integrand of Equation 14.148. Comparison with Equation 14.139 thus results in

$$h_{NM}(\mathbf{s}) = K_{\mathrm{s}} \cdot L^{M}[g(\mathbf{z})] \tag{14.150}$$

From Equation 14.149 and Equation 14.146 it follows that

$$f_{\mathbf{S}}(\mathbf{s}) = K_{\mathrm{s}} (s_1 \ldots s_M)^{(N-3)} g\left(\frac{1}{2s_1^2}, \ldots, \frac{1}{2s_M^2} \right) u(s_1) \ldots u(s_M) \tag{14.151}$$

In this equation, the constant K_{s} is chosen to normalize the area under $f_{\mathbf{S}}(\mathbf{s})$ to unity. As seen, if $f_{\mathbf{S}}(\mathbf{s})$ can be expressed as shown in Equation 14.151, then the generalized SIRV characteristic nonlinear function is proportional to the Laplace transform of the function $g(\mathbf{z})$.

With this representation, it is possible to choose a Laplace transform pair where $f_{\mathbf{S}}(\mathbf{s})$ is not a density function, but the function $f_{\bar{\mathbf{X}}}(\bar{\mathbf{x}})$ generated from $h_{NM}(\mathbf{a}_{\mathbf{x}})$ is a density function. Also, note that the function $f_{\mathbf{S}}(\mathbf{s})$ depends on the value of N. As such, the bootstrap theorem does not hold for characteristic nonlinear functions generated in this fashion.

Since multi-dimensional Laplace transforms can be difficult to evaluate, the representation presented above will now be modified to include only one-dimensional transforms. Let the function $g(\mathbf{z})$ equal

$$g(\mathbf{z}) = \prod_{i=1}^{M} \left[z_i^N \exp(-\beta_i z_i) u(z_i) \right] \cdot \sum_{j=1}^{M} g_j(z_j) \tag{14.152}$$

Define the one-dimensional Laplace transform of each function $g_j(z_j)$ as

$$G_j(\alpha_{xi}) = L\{g_j(z_j)\} \tag{14.153}$$

From Ref. 16, two Laplace transform pairs are

$$L\{z_i^N \exp(-\beta_i z_i) u(z_i)\} = \frac{N!}{(\alpha_{xi} + \beta_i)^{N+1}} \tag{14.154}$$

and

$$L\{z_i^N \exp(-\beta_j z_j) g_j(z_j)\} = G_j^{\{N\}}(\alpha_{xj} + \beta_j) \tag{14.155}$$

where: N is an even number, and $G_j^{\{N\}}(\alpha_{xi})$ is the Nth derivative of $G_j(\alpha_{xi})$. Based on these equations, the M dimensional Laplace transform of the $g(\mathbf{z})$ in Equation 14.152 is

$$G(\mathbf{a}_x) = (N!)^{(M-1)} \cdot \sum_{j=1}^{M} \left\{ \frac{G_j^{\{N\}}(s_j + \beta_j)}{\prod_{\substack{i=1 \\ i \neq j}}^{M} (s_j + \beta_j)^{(N+1)}} \right\} \tag{14.156}$$

After substituting Equation 14.146 and Equation 14.152 into Equation 14.151, the function $f_{\mathbf{S}}(\mathbf{s})$ becomes

$$f_{\mathbf{S}}(\mathbf{s}) = \frac{K_s}{2^{NM}} \cdot \prod_{i=1}^{M} \left[s_i^{-(3+N)} \exp\left(\frac{-\beta_i}{2s_i^2} \right) u(s_i) \right] \cdot \sum_{j=1}^{M} g_j\left(\frac{1}{2s_j^2} \right) \tag{14.157}$$

Substituting Equation 14.156 into Equation 14.150 gives the characteristic nonlinear function $h_{NM}(\mathbf{a}_x)$ as

$$h_{NM}(\mathbf{a}_x) = K_s (N!)^{(M-1)} \cdot \sum_{j=1}^{M} \left\{ \frac{G_j^{\{N\}}(\alpha_{xj} + \beta_j)}{\prod_{\substack{i=1 \\ i \neq j}}^{M} (\alpha_{xj} + \beta_j)^{(N+1)}} \right\} \tag{14.158}$$

When $f_{\mathbf{S}}(\mathbf{s})$ can be expressed in terms of a set of functions $\{g_1(z_1), \ldots, g_M(z_M)\}$, as given in Equation 14.157, then Equation 14.158 shows how $h_{NM}(\mathbf{a}_x)$ can be expressed in terms of the Nth derivatives of the corresponding Laplace transforms.

This concludes the presentation of generalized SIRV fundamental properties. The next section provides specific closed-form examples of generalized SIRV density functions.

14.4. THE GENERALIZED SIRV DENSITY FUNCTION

Equation 14.25 gives the density function of the generalized SIRV matrix $\bar{\mathbf{X}}$ as

$$f_{\bar{\mathbf{X}}}(\bar{\mathbf{x}}) = K \cdot h_{NM}(\mathbf{a}_x) \tag{14.159}$$

where

$$\bar{\mathbf{X}} = \{\mathbf{X}_1, \mathbf{X}_2, \ldots \mathbf{X}_M\} \tag{14.160}$$

and

$$\mathbf{a}_x = \{\alpha_{x1}, \alpha_{x2}, \ldots, \alpha_{xM}\} \tag{14.161}$$

From Equation 14.26, the constant K is

$$K = (2\pi)^{-NM/2}\left(\prod_{i=1}^{M} |\Sigma_{\mathbf{z}i}|^{-1/2}\right) \tag{14.162}$$

and, from Equation 14.27, the characteristic nonlinear function is

$$h_{NM}(\mathbf{a}_x) = \int_{\mathbf{s}}\left(\prod_{i=1}^{M}\left[s_i^{-N}\cdot\exp\left\{\frac{-\alpha_{xi}}{2s_i^2}\right\}\right]\cdot f_{\mathbf{S}}(\mathbf{s})\mathrm{d}\mathbf{s}\right) \tag{14.163}$$

where

$$\mathbf{s} = \{s_1, s_2, \ldots, s_M\} \tag{14.164}$$

$$\int_{\mathbf{s}} = \int_0^\infty \cdots \int_0^\infty (M\text{-fold integration}) \tag{14.165}$$

and

$$\mathrm{d}\mathbf{s} = \mathrm{d}s_1\mathrm{d}s_2\ldots\mathrm{d}s_M \tag{14.166}$$

In order to use Equation 14.163 to find a closed-form expression for the generalized SIRV density, the multidimensional integral shown must be analytically evaluated. The properties developed in the previous section can help to simplify this task. For example, only $h_{1M}(\cdot)$ and $h_{2M}(\cdot)$, the first and second order characteristic nonlinear functions, need be derived. Once these are found, the bootstrap theorem can then be used to generate higher order functions. Employing the Bessel function representation of $h_{NM}(\cdot)$ may also help to evaluate the integral. In Ref. 14, this was done for traditional SIRV's, where M equals unity and the Bessel function representation shown in Equation 14.119 reduces to a Hankel transform. A table of these transforms is presented in Ref. 17. However, it is yet to be shown how to use this approach for the generalized SIRV case where M can exceed unity.

Two approaches are used in the following analysis to find closed-form expressions for $h_{NM}(\mathbf{a}_x)$. The first involves utilizing densities $f_{\mathbf{S}}(\mathbf{s})$ which lend themselves to direct evaluation of the multidimensional integral. The second approach involves utilizing the Laplace transform representation presented in Section 14.3.7.

14.4.1. Direct Evaluation of $h_{NM}(a_x)$

This section presents two closed form solutions for $h_{NM}(\boldsymbol{a}_x)$ as derived through direct evaluation of the integral shown in Equation 14.163.

14.4.1.1. Case 1

Consider the function

$$f_{\mathbf{S}}(\mathbf{s}) = K_{\mathrm{s}} \prod_{i=1}^{M} \left[s_i^{-Q_i} \exp\left(\frac{-\beta_i}{2s_i^2} \right) \right] \cdot \sum_{j=1}^{M} [s_j^{-R_j}] \cdot u(s_1) \ldots u(s_M) \tag{14.167}$$

where $\{Q_1,\ldots,Q_M\},\{R_1,\ldots,R_M\}$, and $\{\beta_1,\ldots,\beta_M\}$ are shape parameters and K_{s} is a normalizing constant. Because the variables $\{s_1,\ldots,s_M\}$ are nonnegative and the volume of $f_S(\mathbf{s})$ will be shown to be finite, the function can be interpreted as a probability density function. Since the volume under a density must be unity, the normalizing constant K_{S} satisfies

$$\frac{1}{K_{\mathrm{s}}} = \int_{\mathbf{S}} \prod_{i=1}^{M} \left[s_i^{-Q_i} \cdot \exp\left(\frac{-\beta_i}{2s_i^2} \right) \right] \cdot \sum_{j=1}^{M} [\mathrm{s}_j^{-R_j}] \mathrm{d}\mathbf{s} \tag{14.168}$$

Interchanging the order of summation and integration and taking out the jth term from the product produces

$$\frac{1}{K_s} = \sum_{j=1}^{M} \left\{ \int_0^{\infty} s_j^{-(Q_j+R_j)} \exp\left(\frac{-\beta_j}{2s_j^2} \right) \mathrm{d}s_j \cdot \prod_{\substack{i=1 \\ i \neq j}}^{M} \int_0^{\infty} s_j^{-Q_i} \exp\left(\frac{-\beta_i}{2s_i^2} \right) \mathrm{d}\mathbf{s}_i \right\} \tag{14.169}$$

In order to simplify Equation 14.169, consider the integral

$$\int_0^{\infty} s^{-x} \exp\left\{ \frac{-\beta}{2s^2} \right\} \mathrm{d}s \tag{14.170}$$

where β is greater than zero. In Equation 14.170, let

$$z = \frac{\beta}{2s^2} \tag{14.171}$$

and

$$\mathrm{d}s = -\frac{1}{2} \cdot 2^{-1/2} \cdot \beta^{1/2} \cdot z^{-3/2} \mathrm{d}z \tag{14.172}$$

Substituting Equation 14.172 and Equation 14.171 into Equation 14.170 gives

$$\int_0^{\infty} s^{-x} \exp\left\{ \frac{-\beta}{2s^2} \right\} \mathrm{d}s = 2^{\left(\frac{x-3}{2}\right)} \cdot \beta^{\left(\frac{1-x}{2}\right)} \cdot \int_0^{\infty} \mathrm{z}^{\left(\frac{x-3}{2}\right)} \exp\{-z\} \mathrm{d}z \tag{14.173}$$

Recall that the gamma function is given by the integral

$$\Gamma(x) = \int_0^{\infty} t^{x-1} \cdot \exp[-t] \mathrm{d}t \tag{14.174}$$

Substituting Equation 14.174 into Equation 14.173 produces

$$\int_0^\infty s^{-x}\exp\left\{\frac{-\beta}{2s^2}\right\}\mathrm{d}s = \frac{1}{2}\left(\frac{\beta}{2}\right)^{-\left(\frac{x-1}{2}\right)}\cdot\Gamma\left(\frac{1-x}{2}\right) \tag{14.175}$$

Equation 14.175 is now used to find the normalizing constant K_S. Substitution of Equation 14.175 into Equation 14.169 yields

$$\frac{1}{K_\mathrm{s}} = 2^{-M}\cdot\sum_{j=1}^{M}\left\{\left(\frac{\beta_j}{2}\right)^{-\left(\frac{Q_j+R_j-1}{2}\right)}\Gamma\left(\frac{Q_j+R_j-1}{2}\right)\cdot\sum_{\substack{i=1\\ i\neq j}}^{M}\left[\left(\frac{\beta_j}{2}\right)^{-\left(\frac{Q_i-1}{2}\right)}\Gamma\left(\frac{Q_i-1}{2}\right)\right]\right\} \tag{14.176}$$

The derivation of this constant completes the specification of the density $f_{\mathbf{S}}(\mathbf{s})$. Note that K_s is finite if and only if all of the parameters $\{\beta_1,\ldots,\beta_M\}$ are greater than zero.

The generalized SIRV characteristic nonlinear function is found by substituting $f_{\mathbf{S}}(\mathbf{s})$, as given in Equation 14.167, into Equation 14.346. The result of this operation is

$$h_{NM}(\mathbf{a}_x) = K_\mathrm{s}\int_{\mathbf{s}}\prod_{i=1}^{M}\left[s_i^{-(N+Q_i)}\cdot\exp\left\{\frac{-(\alpha_{xi}+\beta_i)}{2\mathrm{s}_i^2}\right\}\right]\cdot\sum_{j=1}^{M}[\mathrm{s}_j^{-R_j}]\mathrm{d}\mathbf{s} \tag{14.177}$$

Note the similarities between this and Equation 14.168. Following the same procedure as that used to derive Equation 14.176 gives

$$\begin{aligned} h_{NM}(\mathbf{a}_x) =& K_\mathrm{s}\cdot 2^{-M} \\ &\times\sum_{j=1}^{M}\left\{\left(\frac{\alpha_{xj}+\beta_j}{2}\right)^{-\left(\frac{Q_j+R_j+N-1}{2}\right)}\Gamma\left(\frac{Q_j+R_j+N-1}{2}\right)\right. \\ &\left.\times\prod_{\substack{i=1\\ i\neq j}}^{M}\left[\left(\frac{\alpha_{xi}+\beta_i}{2}\right)^{-\left(\frac{Q_i+N-1}{2}\right)}\Gamma\left(\frac{Q_i+N-1}{2}\right)\right]\right\} \end{aligned} \tag{14.178}$$

Equation 14.178 presents a closed-form solution for a generalized SIRV characteristic nonlinear function. The multivariate density which generated this function is shown in Equation 14.167.

The characteristic nonlinear function of an SIRV is monotonically nonincreasing, as described in Section 14.3.3. In order to form a valid density, this function must also be nonnegative. Note that the function shown in Equation 14.178 is nonnegative and monotonically nonincreasing, given that all values of $\{Q_1, \ldots, Q_M\}$ and $\{R_1, \ldots, R_M\}$ are positive.

Since $f_{\mathbf{S}}(\mathbf{s})$ is not a function of the vector length N, this characteristic nonlinear function must also satisfy the bootstrap theorem. From Equation 14.52, this theorem is,

$$h_{(N+2)M}(\mu_1, \ldots, \mu_M) = (-2)^M \cdot \frac{\partial^M h_{NM}(\mu_1, \ldots, \mu_M)}{\partial \mu_1 \ldots \partial \mu_M} \tag{14.179}$$

Taking M partial derivatives of Equation 14.178 yields

$$\begin{aligned} \frac{\partial^M h_{NM}(\mu_1, \ldots, \mu_M)}{(\partial \mu_1 \ldots \partial \mu_M)} = K_s 2^{-M} (-2)^{-M} \sum_{j=1}^{M} \left(\frac{Q_j + R_j + N - 1}{2} \right) \left(\frac{\mu_j + \beta_j}{2} \right)^{-\left(\frac{Q_j + R_j + N - 1}{2} \right)} \\ \times \Gamma\left(\frac{Q_j + R_j + N - 1}{2} \right) \\ \times \prod_{\substack{i=1 \\ i \neq j}}^{M} \left[\left(\frac{Q_i + N - 1}{2} \right) \left(\frac{\mu_i + \beta_i}{2} \right)^{-\left(\frac{Q_j + N - 1}{2} + 1 \right)} \Gamma\left(\frac{Q_j + N - 1}{2} \right) \right] \end{aligned} \tag{14.180}$$

With reference to Equation 14.174, a property of the gamma function is

$$\Gamma(x) = (x - 1)\Gamma(x - 1) \tag{14.181}$$

Substituting Equation 14.181 into Equation 14.180 produces

$$\begin{aligned} (-2)^M \frac{\partial^M h_{NM}(\mu_1 \ldots \partial \mu_M)}{\partial \mu_1 \ldots \partial \mu_M} = K_s 2^{-M} \sum_{j=1}^{M} \left(\frac{\mu_j + \beta_j}{2} \right)^{-\left(\frac{Q_j + R_j + (N+2) - 1}{2} \right)} \\ \times \Gamma\left(\frac{Q_j + R_j + (N+2) - 1}{2} \right) \\ \times \prod_{\substack{i=1 \\ i \neq j}}^{M} \left[\left(\frac{\mu_i + \beta_i}{2} \right)^{-\left(\frac{Q_j + (N+2) - 1}{2} \right)} \Gamma\left(\frac{Q_j + (N+2) - 1}{2} \right) \right] \end{aligned} \tag{14.182}$$

At this point, verification of the bootstrap theorem is complete since the right side of Equation 14.182 is $h_{(N+2)M}(\mu_1, \ldots, \mu_M)$.

Another interesting property of the generalized SIRV characteristic nonlinear function is its relation to the density of the vector envelopes. From Equation 14.84, this relationship is

$$f_{\mathbf{R}}(r) = \frac{\left(\prod_{i=1}^{M} r_i\right)^{N-1}}{K_R} h_{NM}(r_1^2, \ldots, r_M^2) u(r_1) u(r_2) \ldots u(r_M) \tag{14.183}$$

Without loss of generality, this equation assumes that the covariance matrices $\Sigma_{\mathbf{z}i}$ equal the identity matrix for all i in the set $(1, \ldots, M)$. Consider the case where the number of vectors equals unity, (i.e., $M = 1$). Substituting Equation 14.178 into Equation 14.183 gives

$$f_R(r) = K_{RS} \cdot r^{N-1} \left(\frac{r^2 + \beta}{2}\right)^{-\left(\frac{Q+R+N-1}{2}\right)} \cdot u(r) \tag{14.184}$$

where K_{rs} is a normalizing constant. Figure 14.4 to Figure 14.6 show plots of this density for various values of N, Q, R, and β. Note the heavy tails that can arise for suitable choices of the parameters.

For the case where the number of vectors is two, (i.e., $M = 2$), the joint density of the envelopes is

$$\begin{aligned} f_{R_1,R_2}(r_1, r_2) = K_{rs} \cdot (r_1 r_2)^{N-1} \\ \times \left\{ \left[\left(\frac{r_1^2 + \beta_1}{2}\right)^{-\left(\frac{Q_1+R_1+N-1}{2}\right)} \cdot \left(\frac{r_2^2 + \beta_2}{2}\right)^{-\left(\frac{Q_2+N-1}{2}\right)} \cdot \Gamma\left(\frac{Q_1 + R_1 + N - 1}{2}\right) \right.\right. \\ \left.\left. \times \Gamma\left(\frac{Q_2 + N - 1}{2}\right) \right] + \left(\frac{r_2^2 + \beta_2}{2}\right)^{-\left(\frac{Q_2+R_2+N-1}{2}\right)} \left(\frac{r_1^2 + \beta_1}{2}\right)^{-\left(\frac{Q_1+N-1}{2}\right)} \right. \\ \left. \times \Gamma\left(\frac{Q_2 + R_2 + N - 1}{2}\right) \Gamma\left(\frac{Q_1 + N - 1}{2}\right) \right\} u(r_1) u(r_2) \end{aligned} \tag{14.185}$$

Figure 14.7 and Figure 14.8 show plots of this two-dimensional density.

Consider, once again, the case where the number of vectors is unity, (i.e., $M = 1$). The marginal density of one element of this vector is derived from the expression

$$f_X(x) = (2\pi)^{-1/2} \cdot h_{11}(x^2) \tag{14.186}$$

for the case where the covariance equals unity. Substituting Equation 14.176 and

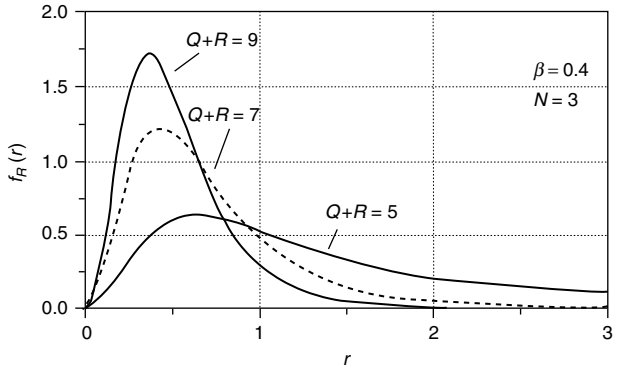

FIGURE 14.4 Case 1 envelope density with $(Q + R)$ as a parameter.

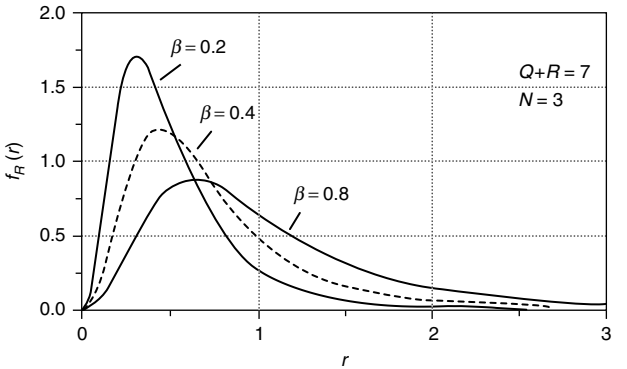

FIGURE 14.5 Case 1 envelope density with β as a parameter.

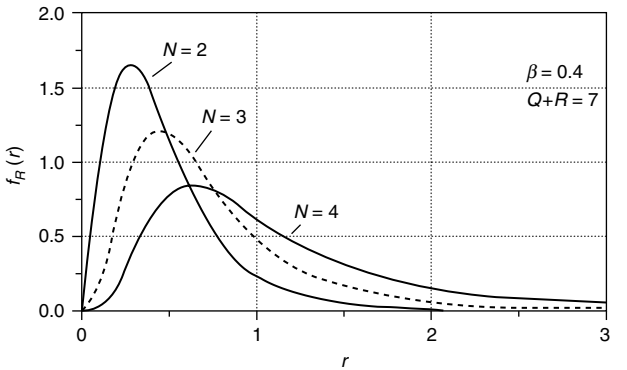

FIGURE 14.6 Case 1 envelope density with N as a parameter.

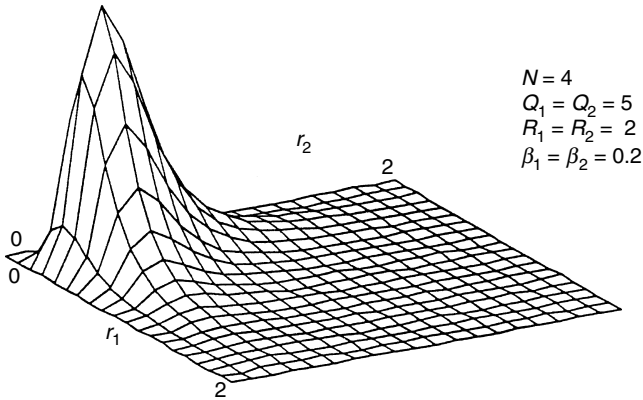

FIGURE 14.7 Case 1 envelope density ($M = 2$) symmetric in r_1 and r_2.

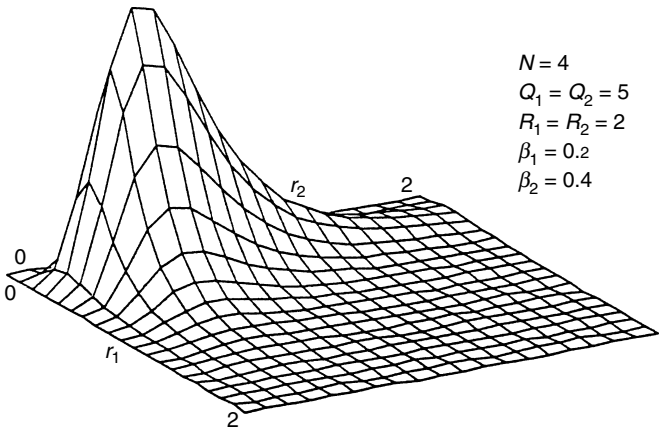

FIGURE 14.8 Case 1 envelope density ($M = 2$) asymmetric in r_1 and r_2.

Equation 14.178 into this expression produces

$$f_X(x) = \frac{(2\pi)^{-1/2} \cdot \left(\frac{x^2+\beta}{2}\right)^{-\left(\frac{Q+R}{2}\right)} \Gamma\left(\frac{Q+R}{2}\right)}{\left(\frac{\beta}{2}\right)^{-\left(\frac{Q+R-1}{2}\right)} \Gamma\left(\frac{Q+R-1}{2}\right)} \tag{14.187}$$

If $(Q+R)$ is an odd integer, then Equation 14.187 can be simplified to

$$f_X(x) = \frac{\left(\frac{x^2+\beta}{2}\right)^{-\left(\frac{Q+R}{2}\right)} \cdot (Q+R-2)!}{\left(\frac{\beta}{2}\right)^{-\left(\frac{Q+R-1}{2}\right)} \cdot 2^{\left(Q+R-\frac{3}{2}\right)} \cdot \left[\left(\frac{Q+R-3}{2}\right)!\right]^2} \tag{14.188}$$

Figure 14.9 and Figure 14.10 show plots of this function for various values of Q, R, and β.

In order to determine the covariance matrix of an SIRV, it is convenient to know the covariance matrix of the generating density $f_{\mathbf{S}}(\mathbf{s})$. The following analysis derives the expected values and covariance matrix of the generating density $f_{\mathbf{S}}(\mathbf{s})$ for the special case where β_i equals β, Q_i equals Q, and R_i equals R for all i in Equation 14.167. The resulting symmetric density is

$$f_{\mathbf{S}}(\mathbf{s}) = K_s \sum_{j=1}^{M} \left[s_j^{-(Q+R)} \prod_{\substack{n=1 \\ n\neq j}}^{M} s_n^{-Q} \exp\left\{ -\frac{\beta}{2} \sum_{i=1}^{M} s_i^{-2} \right\} \right] u(s_1)\ldots u(s_M) \tag{14.189}$$

where Equation 14.176 gives the normalizing constant as

$$\frac{1}{K_s} = M 2^{-M} \cdot \left(\frac{\beta}{2}\right)^{\left[-M\left(\frac{Q-1}{2}\right)-\frac{R}{2}\right]} \cdot \Gamma\left(\frac{Q-1}{2}\right)^{M-1} \Gamma\left(\frac{Q+R-1}{2}\right) \tag{14.190}$$

Based on these equations, the expected value of the kth component of the random vector $\mathbf{s}$ is

$$E\{s_k\} = \int_{\mathbf{S}} s_k \cdot f_{\mathbf{S}}(\mathbf{s}) d\mathbf{s} \tag{14.191}$$

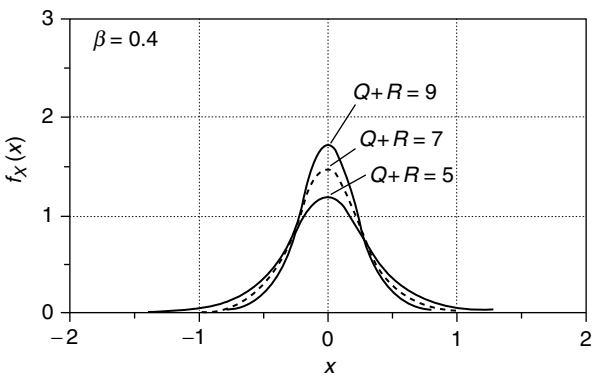

FIGURE 14.9 Case 1 marginal density with $(Q+R)$ as a parameter.

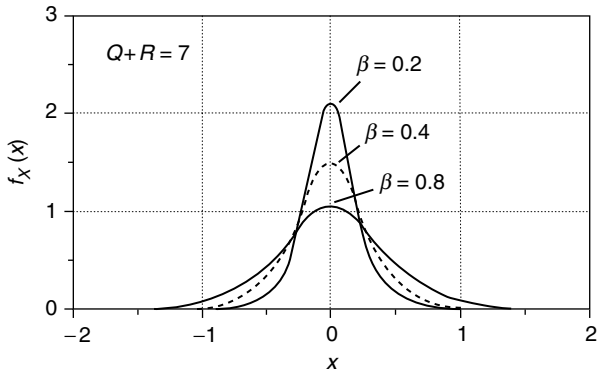

FIGURE 14.10 Case 1 marginal density with β as a parameter.

or

$$E\{s_k\} = \int_{\mathbf{s}} K_{\mathbf{s}} \mathbf{s}_k^{-(Q+R-1)} \prod_{\substack{n=1 \\ n\neq k}}^{M} \mathbf{s}_n^{-Q} \exp\left\{-\frac{\beta}{2}\sum_{i=1}^{M}\mathbf{s}_i^{-2}\right\} d\mathbf{s}$$
$$+ \int_{\mathbf{s}} K_{\mathbf{s}} \sum_{\substack{j=1 \\ j\neq k}}^{M} \left[\mathbf{s}_k^{-(Q-1)}\mathbf{s}_j^{-(Q+R)} \prod_{\substack{n=1 \\ n\neq j,k}}^{M} \mathbf{s}_n^{-Q} \exp\left\{-\frac{\beta}{2}\sum_{i=1}^{M}\mathbf{s}_i^{-2}\right\}\right] d\mathbf{s} \quad (14.192)$$

Equation 14.175 can be modified to produce

$$\int_{\mathbf{s}} \prod_{i=1}^{M} \mathbf{s}_i^{x_i} \exp\left\{-\frac{\beta}{2}\sum_{j=1}^{M}\mathbf{s}_j^{-2}\right\} d\mathbf{s} = 2^{-M}\left(\frac{\beta}{2}\right)^{-\sum_{i=1}^{M}\left(\frac{x_i-1}{2}\right)} \prod_{j=1}^{M} \Gamma\left(\frac{x_i-1}{2}\right) \quad (14.193)$$

Application of Equation 14.193 to Equation 14.192 yields

$$E\{s_k\} = K_{\mathbf{s}} 2^{-M}\left(\frac{\beta}{2}\right)^{-\left[(M-1)\left(\frac{Q-1}{2}\right)+\left(\frac{Q+R-2}{2}\right)\right]} \Gamma\left(\frac{Q-1}{2}\right)^{M-1} \Gamma\left(\frac{Q+R-2}{2}\right)$$
$$+ K_{\mathbf{s}}(M-1)2^{-M}\left(\frac{\beta}{2}\right)^{-\left[(M-2)\left(\frac{Q-1}{2}\right)+\left(\frac{Q-2}{2}\right)+\left(\frac{Q+R-1}{2}\right)\right]}$$
$$\times \Gamma\left(\frac{Q-1}{2}\right)^{M-2} \Gamma\left(\frac{Q-2}{2}\right)\Gamma\left(\frac{Q+R-1}{2}\right) \quad (14.194)$$

or, after simplification,

$$E\{s_k\} = K_{\mathbf{s}} 2^{-M} \left(\frac{\beta}{2}\right)^{-\left[M\left(\frac{Q-1}{2}\right)+\frac{R}{2}\right]} \Gamma\left(\frac{Q-1}{2}\right)^{M-1} \Gamma\left(\frac{Q+R-1}{2}\right) \times \left(\frac{\beta}{2}\right)^{1/2} \left\{ \frac{\Gamma\left(\frac{Q+R-2}{2}\right)}{\Gamma\left(\frac{Q+R-1}{2}\right)} + (M-1)\frac{\Gamma\left(\frac{Q-2}{2}\right)}{\Gamma\left(\frac{Q-1}{2}\right)} \right\} \tag{14.195}$$

Finally, substitution of Equation 14.190 into Equation 14.195 gives

$$E\{s_k\} = \frac{1}{M}\left(\frac{\beta}{2}\right)^{1/2} \left\{ \frac{\Gamma\left(\frac{Q+R-2}{2}\right)}{\Gamma\left(\frac{Q+R-1}{2}\right)} + (M-1)\frac{\Gamma\left(\frac{Q-2}{2}\right)}{\Gamma\left(\frac{Q-1}{2}\right)} \right\} \tag{14.196}$$

Because the shape parameters are identical for all s_k, and $f_{\mathbf{S}}(\mathbf{s})$ is symmetric with respect to all s_k, the mean is the same for all s_k.

The expected value of S_k^2 is

$$E\left\{S_k^2\right\} = \int_{\mathbf{S}} s_k^2 \cdot f_{\mathbf{S}}(\mathbf{s}) d\mathbf{s} \tag{14.197}$$

$$= \int_{\mathbf{S}} K_{\mathbf{s}} s_k^{-(Q+R-2)} \prod_{\substack{n=1 \\ n\neq k}}^{M} s_n^{-Q} \exp\left\{ -\frac{\beta}{2} \sum_{i=1}^{M} s_i^{-2} \right\} d\mathbf{s}$$

$$+ \int_{\mathbf{S}} K_{\mathbf{s}} \sum_{\substack{j=1 \\ j\neq k}}^{M} \left[s_k^{-(Q-2)} s_j^{-(Q+R)} \prod_{\substack{n=1 \\ n\neq j,k}}^{M} s_n^{-Q} \exp\left\{ -\frac{\beta}{2} \sum_{i=1}^{M} s_i^{-2} \right\} \right] d\mathbf{s} \tag{14.198}$$

Application of Equation 14.193 to Equation 14.198 yields

$$E\{S_k^2\} = K_{\mathbf{s}} 2^{-M} \left(\frac{\beta}{2}\right)^{-\left[(M-1)\left(\frac{Q-1}{2}\right)+\left(\frac{Q+R-3}{2}\right)\right]} \Gamma\left(\frac{Q-1}{2}\right)^{M-1} \times \Gamma\left(\frac{Q+R-3}{2}\right) + K_S (M-1) 2^{-M} \left(\frac{\beta}{2}\right)^{-\left[(M-2)\left(\frac{Q-1}{2}\right)+\left(\frac{Q-3}{2}\right)+\left(\frac{Q+R-1}{2}\right)\right]} \times \Gamma\left(\frac{Q-1}{2}\right)^{M-2} \Gamma\left(\frac{Q-3}{2}\right) \Gamma\left(\frac{Q+R-1}{2}\right) \tag{14.199}$$

or, after simplification,

$$E\{S_k^2\} = K_S 2^{-M}\left(\frac{\beta}{2}\right)^{-\left[M\left(\frac{Q-1}{2}\right)+\frac{R}{2}\right]}\Gamma\left(\frac{Q-1}{2}\right)^{M-1}\Gamma\left(\frac{Q+R-1}{2}\right) \times\left(\frac{\beta}{2}\right)\left\{\frac{\Gamma\left(\frac{Q+R-3}{2}\right)}{\Gamma\left(\frac{Q+R-1}{2}\right)}+(M-1)\frac{\Gamma\left(\frac{Q-3}{2}\right)}{\Gamma\left(\frac{Q-1}{2}\right)}\right\} \tag{14.200}$$

Finally, substitution of Equation 14.190 into Equation 14.200 gives

$$E\{S_k^2\} = \frac{1}{M}\left(\frac{\beta}{2}\right)\left\{\frac{\Gamma\left(\frac{Q+R-3}{2}\right)}{\Gamma\left(\frac{Q+R-1}{2}\right)}+(M-1)\frac{\Gamma\left(\frac{Q-3}{2}\right)}{\Gamma\left(\frac{Q-1}{2}\right)}\right\} \tag{14.201}$$

The variance of s_k is given by

$$\sigma_k^2 = E\{S_k^2\} - E\{S_k\}^2 \tag{14.202}$$

Substitution of Equation 14.201 and Equation 14.196 into Equation 14.202 produces

$$\sigma_k^2 = \frac{\beta}{2M}\left\{\frac{\Gamma\left(\frac{Q+R-3}{2}\right)}{\Gamma\left(\frac{Q+R-1}{2}\right)}+(M-1)\frac{\Gamma\left(\frac{Q-3}{2}\right)}{\Gamma\left(\frac{Q-1}{2}\right)}\right\} - \frac{\beta}{2M^2}\left\{\frac{\Gamma\left(\frac{Q+R-2}{2}\right)}{\Gamma\left(\frac{Q+R-1}{2}\right)}+(M-1)\frac{\Gamma\left(\frac{Q-2}{2}\right)}{\Gamma\left(\frac{Q-1}{2}\right)}\right\}^2 \tag{14.203}$$

Note that the variance is the same for all S_k.

The covariance between two different components of $\mathbf{S}$ is defined by

$$\mathbf{C}_{\mathbf{s}_{u,v}} = E\{[S_u - E(S_u)][S_v - E(S_v)]\} \tag{14.204}$$

Since the density is assumed symmetric, the expected value is the same for all components of $\mathbf{S}$. The covariance thus becomes

$$\mathbf{C}_{\mathbf{s}_{u,v}} = E(S_u S_v) - E(S_u)^2 \tag{14.205}$$

The expected value of the product between two different components of $\mathbf{S}$ is

$$E\{S_u S_v\} = \int_{\mathbf{s}} s_u s_v f_{\mathbf{S}}(\mathbf{s})\mathbf{ds} \tag{14.206}$$

or

$$E\{s_u s_v\} = \int_{\mathbf{S}} K_s s_u^{-(Q+R-1)} s_v^{-(Q-1)} \prod_{\substack{n=1 \\ i\neq u,v}}^{M} s_n^{-Q} \exp\left\{-\frac{\beta}{2}\sum_{i=1}^{M} s_i^{-2}\right\} \mathbf{ds}$$
$$+ \int_{\mathbf{S}} K_s s_v^{-(Q+R-1)} s_u^{-(Q-1)} \prod_{\substack{n=1 \\ i\neq u,v}}^{M} s_n^{-Q} \exp\left\{-\frac{\beta}{2}\sum_{i=1}^{M} s_i^{-2}\right\} \mathbf{ds}$$
$$+ \int_{\mathbf{S}} K_s \sum_{\substack{j=1 \\ j\neq u,v}}^{M} \left[s_u^{-(Q-1)} s_v^{-(Q-1)} s_j^{-(Q+R)} \prod_{\substack{n=1 \\ n\neq u,v,j}}^{M} s_n^{-Q} \exp\left\{-\frac{\beta}{2}\sum_{i=1}^{M} s_i^{-2}\right\} \right] \mathbf{ds} \tag{14.207}$$

Application of Equation 14.193 to Equation 14.207 yields

$$E\{s_u s_v\} = 2K_s 2^{-M} \left(\frac{\beta}{2}\right)^{-\left[(M-2)\left(\frac{Q-1}{2}\right)+\left(\frac{Q+R-2}{2}\right)+\left(\frac{Q-2}{2}\right)\right]}$$
$$\times \Gamma\left(\frac{Q-1}{2}\right)^{M-2} \Gamma\left(\frac{Q+R-2}{2}\right) \Gamma\left(\frac{Q-2}{2}\right)$$
$$+ K_s (M-2) 2^{-M} \left(\frac{\beta}{2}\right)^{-\left[(M-3)\left(\frac{Q-1}{2}\right)+2\left(\frac{Q-2}{2}\right)+\left(\frac{Q+R-1}{2}\right)\right]}$$
$$\times \Gamma\left(\frac{Q-1}{2}\right)^{M-3} \Gamma\left(\frac{Q-2}{2}\right)^{2} \Gamma\left(\frac{Q+R-1}{2}\right) \tag{14.208}$$

or, after simplification,

$$E\{s_u s_v\} = K_s 2^{-M} \left(\frac{\beta}{2}\right)^{-\left[M\left(\frac{Q-1}{2}\right)+\frac{R}{2}\right]} \Gamma\left(\frac{Q-1}{2}\right)^{M-1} \Gamma\left(\frac{Q+R-1}{2}\right)$$
$$\times \left(\frac{\beta}{2}\right) \left\{ 2\frac{\Gamma\left(\frac{Q-1}{2}\right)\Gamma\left(\frac{Q+R-2}{2}\right)}{\Gamma\left(\frac{Q-1}{2}\right)\Gamma\left(\frac{Q+R-1}{2}\right)} + (M-2)\frac{\Gamma\left(\frac{Q-2}{2}\right)^{2}}{\Gamma\left(\frac{Q-1}{2}\right)^{2}} \right\} \tag{14.209}$$

Substitution of Equation 14.190 into Equation 14.209 gives

$$E\{S_u S_v\} = \frac{\beta}{2M} \frac{\Gamma\left(\frac{Q-2}{2}\right)}{\Gamma\left(\frac{Q-1}{2}\right)} \times \left\{ 2 \frac{\Gamma\left(\frac{Q+R-2}{2}\right)}{\Gamma\left(\frac{Q+R-1}{2}\right)} + (M-2) \frac{\Gamma\left(\frac{Q-2}{2}\right)}{\Gamma\left(\frac{Q-1}{2}\right)} \right\} \tag{14.210}$$

Finally, Equation 14.210 and Equation 14.196 can be inserted into Equation 14.205 to produce the covariance,

$$C_{s_{u,v}} = \frac{\beta}{2M} \frac{\Gamma\left(\frac{Q-2}{2}\right)}{\Gamma\left(\frac{Q-1}{2}\right)} \left\{ 2 \frac{\Gamma\left(\frac{Q+R-2}{2}\right)}{\Gamma\left(\frac{Q+R-1}{2}\right)} + (M-2) \frac{\Gamma\left(\frac{Q-2}{2}\right)}{\Gamma\left(\frac{Q-1}{2}\right)} \right\} - \frac{\beta}{2M^2} \left\{ \frac{\Gamma\left(\frac{Q+R-2}{2}\right)}{\Gamma\left(\frac{Q+R-1}{2}\right)} + (M-1) \frac{\Gamma\left(\frac{Q-2}{2}\right)}{\Gamma\left(\frac{Q-1}{2}\right)} \right\}^2 \tag{14.211}$$

Once again, because the shape parameters are identical for all S_k, and $f_{\mathbf{S}}(\mathbf{s})$ is symmetric with respect to all S_k, the covariance is seen to be independent of u and v. Consequently, the full covariance matrix is simply evaluated by computing only two parameters. All of the diagonal elements are equal and are given by Equation 14.203. Similarly, all of the off-diagonal elements are equal and are given by Equation 14.211. This is referred to as a diagonal innovation matrix.

The covariance derivation for a nonsymmetric density $f_{\mathbf{S}}(\mathbf{s})$ is more complex. To simplify the analysis, let the number of SIRVs equal two, (i.e., $M = 2$). From Equation 14.167, the generating density $f_{\mathbf{S}}(\mathbf{s})$ is

$$f_{S_1,S_2}(s_1, s_2) = K_s s_1^{-Q_1} \exp\left\{ \frac{-\beta_1}{2s_1^2} \right\} s_2^{-Q_2} \exp\left\{ \frac{-\beta_2}{2s_2^2} \right\} [s_1^{-R_1} + s_2^{-R_2}] u(s_1) u(s_2) \tag{14.212}$$

where: the constant K_s satisfies

$$\frac{1}{K_s} = \frac{1}{4}\left(\frac{\beta_1}{2}\right)^{-\left(\frac{Q_1+R_1-1}{2}\right)}\left(\frac{\beta_2}{2}\right)^{-\left(\frac{Q_2-1}{2}\right)}\Gamma\left(\frac{Q_1+R_1-1}{2}\right)\Gamma\left(\frac{Q_2-1}{2}\right) + \frac{1}{4}\left(\frac{\beta_2}{2}\right)^{-\left(\frac{Q_2+R_2-1}{2}\right)}\left(\frac{\beta_1}{2}\right)^{-\left(\frac{Q_1-1}{2}\right)}\Gamma\left(\frac{Q_2+R_2-1}{2}\right)\Gamma\left(\frac{Q_1-1}{2}\right) \tag{14.213}$$

The expected value of S_1 is given by

$$E\{S_1\} = \int_0^\infty \int_0^\infty s_1 f_{S_1,S_2}(s_1, s_2)\mathrm{d}s_1\, \mathrm{d}s_2 \tag{14.214}$$

$$= \int_0^\infty \int_0^\infty K_s s_1^{-Q_1+1}\exp\left\{\frac{-\beta_1}{2s_1^2}\right\}s_2^{-Q_2}\exp\left\{\frac{-\beta_2}{2s_2^2}\right\}[s_1^{-R_1} + s_2^{-R_2}]\mathrm{d}s_1\, \mathrm{d}s_2 \tag{14.215}$$

Repeated use of Equation 14.175 produces

$$E\{S_1\} = \frac{K_s}{4}\left(\frac{\beta_1}{2}\right)^{-\left(\frac{Q_1+R_1-2}{2}\right)}\left(\frac{\beta_2}{2}\right)^{-\left(\frac{Q_2-1}{2}\right)}\Gamma\left(\frac{Q_1+R_1-2}{2}\right)\Gamma\left(\frac{Q_2-1}{2}\right) + \frac{K_s}{4}\left(\frac{\beta_2}{2}\right)^{-\left(\frac{Q_2+R_2-1}{2}\right)}\left(\frac{\beta_1}{2}\right)^{-\left(\frac{Q_1-2}{2}\right)}\Gamma\left(\frac{Q_2+R_2-1}{2}\right)\Gamma\left(\frac{Q_1-2}{2}\right) \tag{14.216}$$

In a similar fashion, the expected value of S_1^2 is found to be

$$E\{S_1^2\} = \frac{K_s}{4}\left(\frac{\beta_1}{2}\right)^{-\left(\frac{Q_1+R_1-3}{2}\right)}\left(\frac{\beta_2}{2}\right)^{-\left(\frac{Q_2-1}{2}\right)}\Gamma\left(\frac{Q_1+R_1-3}{2}\right)\Gamma\left(\frac{Q_2-1}{2}\right) + \frac{K_s}{4}\left(\frac{\beta_2}{2}\right)^{-\left(\frac{Q_2+R_2-1}{2}\right)}\left(\frac{\beta_1}{2}\right)^{-\left(\frac{Q_1-3}{2}\right)}\Gamma\left(\frac{Q_2+R_2-1}{2}\right)\Gamma\left(\frac{Q_1-3}{2}\right) \tag{14.217}$$

The variance of S_1 is found by substituting Equation 14.216 and Equation 14.217 into Equation 14.202. Note that the expected values of S_2 and S_2^2 are identical to those shown in Equation 14.216 and Equation 14.217, except for a one-to-one interchange of the one and two subscripts.

The expected value of the product (S_1S_2) is

$$E\{S_1S_2\} = \int_0^\infty \int_0^\infty s_1 s_2 f_{S_1,S_2}(s_1, s_2)\mathrm{d}s_1\, \mathrm{d}s_2 \tag{14.218}$$

$$= \frac{K_s}{4}\left(\frac{\beta_1}{2}\right)^{-\left(\frac{Q_1+R_1-2}{2}\right)}\left(\frac{\beta_2}{2}\right)^{-\left(\frac{Q_2-2}{2}\right)}\Gamma\left(\frac{Q_1+R_1-2}{2}\right)\Gamma\left(\frac{Q_2-2}{2}\right)$$
$$+ \frac{K_s}{4}\left(\frac{\beta_2}{2}\right)^{-\left(\frac{Q_2+R_2-2}{2}\right)}\left(\frac{\beta_1}{2}\right)^{-\left(\frac{Q_1-2}{2}\right)}\Gamma\left(\frac{Q_2+R_2-2}{2}\right)\Gamma\left(\frac{Q_1-2}{2}\right) \tag{14.219}$$

The covariance between s_1 and s_2 is given by $E\{s_1 s_2\} - E\{s_1\}E\{s_2\}$.

The density $f_{\mathbf{S}}(\mathbf{s})$ used in case one to find a closed form expression for $h_{NM}(\mathbf{a}_x)$ was chosen so as to factor the multidimensional integral into a product of single dimensional integrals. The following analysis presents another generalized SIRV derived through direct evaluation of $h_{NM}(\mathbf{a}_x)$.

14.4.1.2. Case 2

From Equation 14.27, the characteristic nonlinear function is

$$h_{NM}(\mathbf{a}_x) = \int_{\mathbf{s}} \prod_{i=1}^{M}\left[s_i^{-N}\exp\left\{\frac{-\alpha_{xi}}{2s_i^2}\right\}\right] f_{\mathbf{S}}(\mathbf{s})\mathrm{d}\mathbf{s} \tag{14.220}$$

In Equation 14.220 let

$$z_i = s_i^{-2} \tag{14.221}$$

Based on Equation 14.221,

$$\mathrm{d}s_i = -\frac{1}{2}z_i^{-3/2}\mathrm{d}z_i \tag{14.222}$$

for all i in the $\{1,\ldots,M\}$. Substituting Equation 14.222 and Equation 14.221 into Equation 14.220 produces

$$h_{NM}(\mathbf{a}_x) = 2^{-M}\int_{\mathbf{z}} \prod_{i=1}^{M}\left[z_i^{\left(\frac{N-3}{2}\right)}\exp\left\{-\frac{\alpha_{xi}}{2}z_i\right\}\right] f_{\mathbf{S}}(z_1^{-1/2},\ldots,z_M^{-1/2})\mathrm{d}\mathbf{z} \tag{14.223}$$

where:

$$\mathbf{z} = \{z_1, z_2, \ldots, z_M\} \tag{14.224}$$

$$\int_{\mathbf{z}} = \int_{-\infty}^{\infty}\ldots\int_{-\infty}^{\infty} \ (M\text{-fold integration}) \tag{14.225}$$

and

$$\mathrm{d}\mathbf{z} = \mathrm{d}z_1\mathrm{d}z_2\ldots\mathrm{d}z_M \tag{14.226}$$

In making this substitution, note that setting z_i equal to an inverse of s_i causes an exchange in the integral limits of zero and infinity. However, the negative sign in Equation 14.222 causes them to change back.

Consider the density

$$f_{\mathbf{S}} = \{z_1^{-1/2}, \ldots, z_M^{-1/2}\} = K_{s_{NM}}$$
$$\prod_{i=1}^{M} [z_i^{-Q_i} \exp\{-\beta_i z_i\}] \exp\left\{-\frac{\gamma}{z_1 \ldots z_M}\right\} u(z_1) \ldots u(z_M) \tag{14.227}$$

where $K_{s_{NM}}$ is a normalizing constant and $\{\beta_1, \ldots, \beta_M, \gamma\}$ are shape parameters. Define the parameters Q_i such that

$$\left(\frac{N-3}{2}\right) - Q_i = \left(\frac{i}{M+1}\right) - 1 \tag{14.228}$$

or

$$Q_i = \left(\frac{N-1}{2}\right) - \left(\frac{i}{M+1}\right) \tag{14.229}$$

for all i in the set $\{1, \ldots, M\}$. Substituting Equation 14.227 and Equation 14.229 into Equation 14.223 gives

$$h_{NM}(\mathbf{a}_x) = 2^{-M} K_{s_{NM}}$$
$$\int_{\mathbf{z}} \prod_{i=1}^{M} \left[z_i^{\left(\frac{1}{M+1} - 1\right)} \exp\left\{-\left(\frac{\alpha_{xi}}{2} + \beta_i\right) z_i\right\}\right] \exp\left\{-\frac{\gamma}{z_1 \ldots z_M}\right\} \mathrm{d}\mathbf{z} \tag{14.230}$$

In Equation 14.230 let

$$u_i = \left(\frac{\alpha_{xi}}{2} + \beta_i\right) z_i \tag{14.231}$$

and

$$\mathrm{d}z_i = \left(\frac{2}{\alpha_{xi} + 2\beta_i}\right) \mathrm{d}u_i \tag{14.232}$$

for all i in the set $\{1, \ldots, M\}$. Substituting Equation 14.231 and Equation 14.232 into Equation 14.230 produces

$$h_{NM}(\mathbf{a}_x) = 2^{-M} K s_{NM} \prod_{i=1}^{M} \left\{\frac{2}{\alpha_{xi} + 2\beta_i}\right\}^{\left(\frac{1}{M+1}\right)}$$
$$\times \int_u \prod_{i=1}^{M} \left[u_i^{\left(\frac{1}{M+1} - 1\right)} \exp\{-u_i\}\right] \exp\left\{-\frac{\lambda_M^{M+1}}{u_1 \ldots u_M}\right\} \mathrm{d}u \tag{14.233}$$

where

$$\mathbf{u} = \{u_1, u_2, \ldots u_M\} \tag{14.234}$$

$$\int_{\mathbf{u}} = \int_0^\infty \ldots \int_0^\infty \ (M\text{-fold integration}) \tag{14.235}$$

$$\mathrm{d}\mathbf{u} = \mathrm{d}u_1 \mathrm{d}u_2 \ldots \mathrm{d}u_M \tag{14.236}$$

and

$$\lambda_M = \left\{ \gamma \prod_{i=1}^{M} \left(\frac{\alpha_{xi}}{2} + \beta_i \right) \right\}^{\left(\frac{1}{M+1} \right)} \tag{14.237}$$

From Ref. 19, the solution to the multidimensional integral shown in Equation 14.233 is

$$h_{NM}(\mathbf{a}_x) = 2^{-M} K_{s_{NM}} \prod_{i=1}^{M} \left\{ \frac{2}{\alpha_{xi} + 2\beta_i} \right\}^{\left(\frac{1}{M+1} \right)} \left\{ \frac{(2\pi)^{M/2}}{\sqrt{M+1}} \right\} \times \exp\{ -(M+1)\lambda_M \} \tag{14.238}$$

The density function $f_{\mathbf{S}}(\mathbf{s})$ which generates this characteristic nonlinear function is found by substituting Equation 14.221 and Equation 14.229 into Equation 14.227. The result of these operations is

$$f_{\mathbf{S}}(\mathbf{s}) = K_{s_{NM}} \prod_{i=1}^{M} \left[s_i^{\left\{ (N-1) - \frac{2l}{M+1} \right\}} \exp\left\{ -\frac{\beta_i}{s_i^2} \right\} u(s_i) \right] \exp(-\gamma s_1^2 \ldots s_M^2) \tag{14.239}$$

The values of $K_{s_{NM}}$ can be found through numerical integration. For the special case where all of the shape parameters are unity, Table 14.2 shows values of this constant for several values of N and M.

With reference to Equation 14.239, note that the density function depends on the value of N. Because of this, the bootstrap theorem does not hold for the characteristic function $h_{NM}(\mathbf{a}_x)$ shown in Equation 14.223. Nevertheless, this function still generates a valid generalized SIRV density function.

TABLE 14.2
Case 2 Density Normalizing Constant

$K_{s_{NM}}$ for $(\beta_1,\ldots,\beta_M,\gamma) = (1.0,\ldots,1.0,1.0)$		
N	$M = 1$	$M = 2$
2	8.3377	19.644
4	5.5584	10.515
8	0.8663	1.0885
16	0.0012	0.0009

Consider the case where the number of vectors in the generalized SIRV matrix is one, (i.e., $M = 1$). From Equation 14.84, the density of the envelope of this vector is

$$f_{\mathrm{R}}(r) = K_{\mathrm{r1}} \cdot r^{(N-1)} \left(\frac{r^2}{2} + \beta \right)^{-1/2} \exp\left[-2\sqrt{\gamma} \left(\frac{r^2}{2} + \beta \right)^{1/2} \right] \tag{14.240}$$

where K_{r1} is a normalizing constant. Figure 14.11 to Figure 14.13 show plots of this density for various values of β, γ, and N.

Once again, consider the case where the number of vectors in the generalized SIRV matrix is one (i.e., $M = 1$). The marginal density of one component of this vector is found by setting N equal to unity in Equation 14.238. This marginal density is

$$f_X(x) = K_x \left(\frac{x^2}{2} + \beta \right)^{-1/2} \exp\left[-2\sqrt{\gamma} \left(\frac{x^2}{2} + \beta \right)^{1/2} \right] \tag{14.241}$$

where K_x is a normalizing constant. Figure 14.14 and Figure 14.15 show plots of this density of various values of β and γ.

This concludes discussion of the case 2 SIRV since it was not possible to obtain a closed form expression for the covariance matrix of $\mathbf{S}$.

14.4.2. Evaluation of h_{NM} Using the Laplace Transform

The previous section presented two closed form expressions for $h_{NM}(\mathbf{a}_x)$ as derived through direct evaluation of the integral shown in Equation 14.163. This section utilizes the Laplace transform representation derived in Section 14.3.7 to

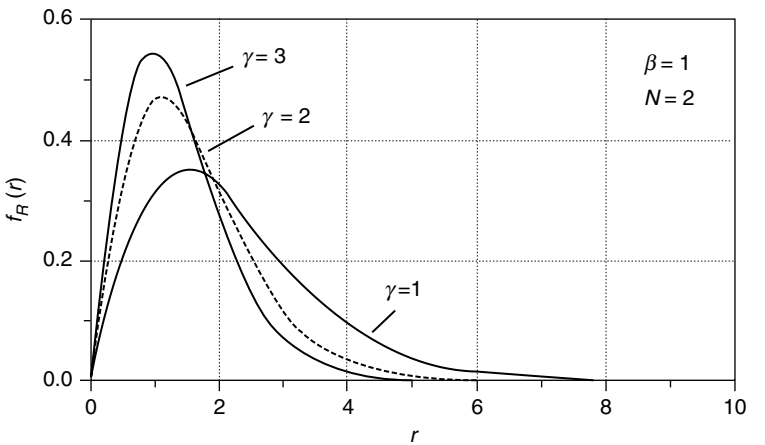

FIGURE 14.11 Case 2 envelope density with γ as a parameter.

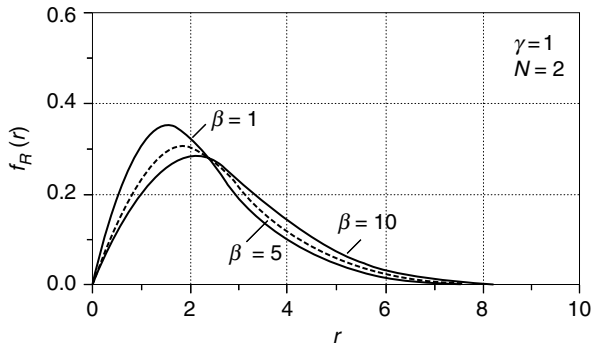

FIGURE 14.12 Case 2 envelope density with β as a parameter.

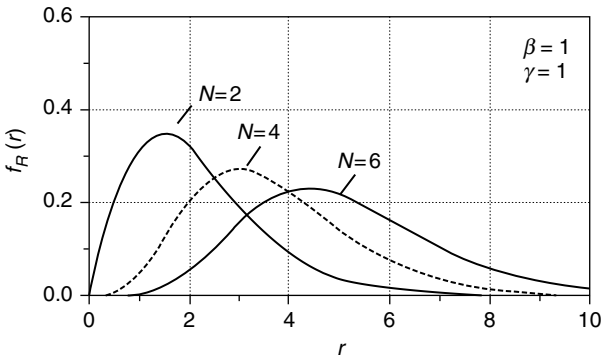

FIGURE 14.13 Case 2 envelope density with N as a parameter.

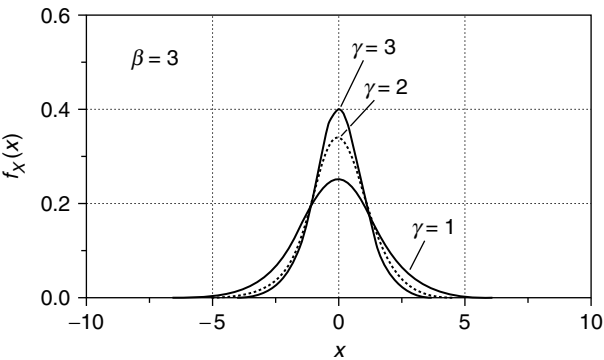

FIGURE 14.14 Case 2 marginal density with γ as a parameter.

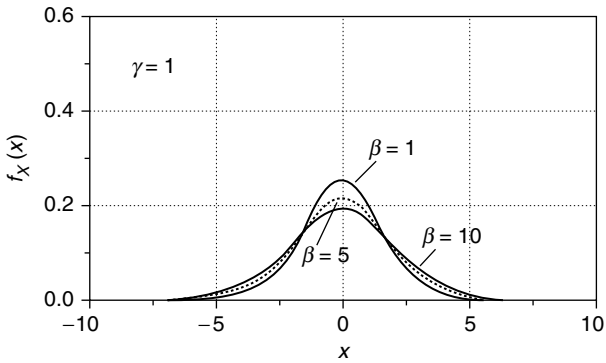

FIGURE 14.15 Case 2 marginal density with β as a parameter.

find closed form expressions for $h_{NM}(\mathbf{a}_x)$. From Equation 14.157 and Equation 14.158, the generating density,

$$f_{\mathbf{S}}(\mathbf{s}) = \frac{K_s}{2^{NM}} \prod_{i=1}^{M} \left[s_i^{-(3+N)} \exp\left\{ \frac{-\beta_i}{2s_i^2} \right\} u(s_i) \right] \sum_{j=1}^{M} g_j\left(\frac{1}{2s_j^2} \right) \tag{14.242}$$

results in

$$h_{NM}(\mathbf{a}_x) = K_s (N!)^{(M-1)} \sum_{j=1}^{M} \left\{ \frac{G_j^{\{N\}}(\alpha_{xj} + \beta_j)}{\prod_{\substack{i=1 \\ i \neq j}}^{M} (\alpha_{xj} + \beta_j)^{(N+1)}} \right\} \tag{14.243}$$

where the functions in the sets $\{g_1(z_1), \ldots, g_M(z_M)\}$ and $\{G_1(\alpha_{x1}), \ldots, G_M(\alpha_{xM})\}$ form one-to-one Laplace transform pairs. The following analysis presents two closed form expressions for $h_{NM}(\mathbf{a}_x)$ derived through use of these equations.

14.4.2.1. Case 3

Consider the case where the number of vectors in the generalized SIRV matrix is one (i.e., $M = 1$). Based on this, Equation 14.242 and Equation 14.243 reduce to

$$f_{\mathbf{S}}(\mathbf{s}) = \frac{K_s}{2^N} s^{-(3+N)} \exp\left\{ \frac{-\beta}{2s^2} \right\} g\left(\frac{1}{2s^2} \right) u(s) \tag{14.244}$$

and

$$h_N(\alpha_{\mathbf{x}}) = K_s G^{\{N\}}(\alpha_{\mathbf{x}} + \beta) \tag{14.245}$$

From Equation 14.84, the envelope density of this vector is

$$f_{\mathrm{R}}(r) = \frac{r^{(N-1)}}{K_r} G^{\{N\}}(r^2 + \beta)u(r) \tag{14.246}$$

where K_r is a normalizing constant. Consider the function

$$g(z) = \exp(-z) \tag{14.247}$$

From Ref. 16, the Laplace transform of this function is

$$G(\alpha_{\mathbf{x}}) = (\alpha_{\mathbf{x}} + 1)^{-1} \tag{14.248}$$

Taking N derivatives with respect to $\alpha_{\mathbf{x}}$ produces

$$G^{\{N\}}(\alpha_{\mathbf{x}}) = (-1)^N (N!)(\alpha_{\mathbf{x}} + 1)^{-(N+1)} \tag{14.249}$$

After assuming N is an even integer such that $G^{\{N\}}(\alpha_{\mathbf{x}})$ is nonnegative, substitution of Equation 14.249 into Equation 14.246 yields

$$f_{\mathrm{R}}(r) = \frac{N!}{K_r} r^{(N-1)}(r^2 + \beta + 1)^{-(N+1)}u(r) \tag{14.250}$$

Substituting Equation 14.247 into Equation 14.244 gives

$$f_{\mathbf{S}}(\mathbf{s}) = \frac{K_s}{2^N} s^{-(3+N)} \exp\left\{\frac{-(\beta+1)}{2s^2}\right\} u(s) \tag{14.251}$$

The function shown in Equation 14.251 is the generalized SIRV characteristic density which gives rise to the random vector whose envelope density is specified by Equation 14.250. This example SIRV is derived in Ref. 14 through direct evaluation of the integral shown in Equation 14.163, for the case where M equals unity. The components of this SIRV follow the student t distribution.[18]

The Laplace transformation representation readily allows for expansion to cases where M is greater than unity. Assume that all of the functions in the set $\{g_1(z_1), \ldots, g_M(z_M)\}$ are the same and that N is an even integer. Substituting Equation 14.249 into Equation 14.243 gives the characteristic nonlinear function as

$$h_{NM}(\mathbf{a}_x) = (N!)^M K_s \sum_{j=1}^{M} \left\{ \left[(\alpha_{xj} + \beta_j + 1) \prod_{\substack{i=1 \\ i \neq j}}^{M} (\alpha_{xj} + \beta_i) \right]^{-(N+1)} \right\} \tag{14.252}$$

Furthermore, from Equation 14.84 the joint density of the M vector envelopes is

$$f_{\mathrm{R}}(r) = \frac{K_{\mathrm{S}}(N!)^M}{K_r} (r_1 \ldots r_M)^{N-1}$$

$$\sum_{j=1}^{M}\left\{\left[\left(r_j^2+\beta_j+1\right)\prod_{\substack{i=1\\ i\neq j}}^{M}\left(r_i^2+\beta_i\right)\right]^{-(N+1)}\right\}u(r_1)\ldots u(r_M) \tag{14.253}$$

and from Equation 14.242 and Equation 14.247, the multivariate characteristic generating density is

$$f_{\mathbf{S}}(\mathbf{s})=\frac{K_s}{2^{NM}}\prod_{i=1}^{M}\left[s_i^{-(3+N)}\exp\left\{\frac{-\beta_i}{2s_i^2}\right\}u(s_i)\right]\sum_{j=1}^{M}\exp\left(\frac{-1}{2s_j^2}\right) \tag{14.254}$$

The following analysis concludes the investigation of the case three SIRV by deriving the expected values and covariance matrix of the generating density $f_{\mathbf{S}}(\mathbf{s})$, for the symmetric case where β_i equals β for all i in Equation 14.254. The resulting symmetric density is

$$f_{\mathbf{S}}(\mathbf{s})=\frac{K_s}{2^{NM}}\prod_{i=1}^{M}\left[s_i^{-(3+N)}\exp\left\{\frac{-\beta}{2s_i^2}\right\}u(s_i)\right]\sum_{j=1}^{M}\exp\left(\frac{-1}{2s_j^2}\right) \tag{14.255}$$

or

$$f_{\mathbf{S}}(\mathbf{s})=\frac{K_s}{2^{NM}}\sum_{j=1}^{M}\left\{s_j^{-(3+N)}\exp\left\{\frac{-(\beta+1)}{2s_j^2}\right\}u(s_j)\prod_{\substack{i=1\\ i\neq j}}^{M}\left[s_i^{-(3+N)}\exp\left\{\frac{-\beta}{2s_i^2}\right\}u(s_i)\right]\right\} \tag{14.256}$$

Since the volume under the density is unity, the normalizing constant K_{S} satisfies

$$\frac{1}{K_{\mathrm{S}}}=\frac{1}{2^{NM}}\sum_{j=1}^{M}\left\{\int_0^\infty s_j^{-(3+N)}\exp\left\{\frac{-(\beta+1)}{2s_j^2}\right\}\mathrm{d}s_j\prod_{\substack{i=1\\ i\neq j}}^{M}\left[\int_0^\infty s_i^{-(3+N)}\exp\left\{\frac{-\beta}{2s_i^2}\right\}\mathrm{d}s_i\right]\right\} \tag{14.257}$$

Assume that β is positive. From Equation 14.175, the solutions to the integrals shown in Equation 14.257 result in

$$\frac{1}{K_{\mathrm{S}}}=\frac{1}{2^{NM}}\sum_{j=1}^{M}\left\{\frac{1}{2}\left(\frac{\beta+1}{2}\right)^{-\left(\frac{N+2}{2}\right)}\Gamma\left(\frac{N+2}{2}\right)\prod_{\substack{i=1\\ i\neq j}}^{M}\left[\frac{1}{2}\left(\frac{\beta}{2}\right)^{-\left(\frac{N+2}{2}\right)}\Gamma\left(\frac{N+2}{2}\right)\right]\right\} \tag{14.258}$$

or

$$\frac{1}{K_S} = \frac{1}{2^{NM}} \frac{M}{2^M} \Gamma\left(\frac{N+2}{2}\right)^M \left(\frac{\beta+1}{2}\right)^{-\left(\frac{N+2}{2}\right)} \left(\frac{\beta}{2}\right)^{-(M-1)\left(\frac{N+2}{2}\right)} \tag{14.259}$$

Substitution of Equation 14.259 into Equation 14.255 gives

$$f_{\mathbf{S}}(\mathbf{s}) = \frac{2^M}{M} \frac{\left(\frac{\beta+1}{2}\right)^{\left(\frac{N+2}{2}\right)} \left(\frac{\beta}{2}\right)^{(M-1)\left(\frac{N+2}{2}\right)}}{\Gamma\left(\frac{N+2}{2}\right)^M}$$

$$\times \prod_{i=1}^{M} \left[s_i^{-(3+N)} \exp\left\{\frac{-\beta}{2s_i^2}\right\} u(s_i) \right] \sum_{j=1}^{M} \exp\left(\frac{-1}{2s_j^2}\right) \tag{14.260}$$

From Ref. 16, two properties of the gamma function are

$$\Gamma(m) = (m-1)! \tag{14.261}$$

and

$$\Gamma\left(m + \frac{1}{2}\right) = \sqrt{\pi} 2^{(1-2m)} \frac{(2m-1)!}{(m-1)!} \tag{14.262}$$

where m is an integer. Based on Equation 14.261, for the special case where N is an even integer, Equation 14.260 simplifies to

$$f_{\mathbf{S}}(\mathbf{s}) = \frac{(\beta+1)^{\left(\frac{N+2}{2}\right)} \beta^{(M-1)\left(\frac{N+2}{2}\right)}}{M 2^{MN/2} \left(\frac{N}{2}!\right)^M} \prod_{i=1}^{M} \left[s_i^{-(3+N)} \exp\left\{\frac{-\beta}{2s_i^2}\right\} u(s_i) \right]$$

$$\times \sum_{j=1}^{M} \exp\left(\frac{-1}{2s_j^2}\right) \tag{14.263}$$

In Equation 14.263, note that $f_{\mathbf{S}}(\mathbf{s})$ is always positive since β is positive. This fact, and the fact that the volume under the function is unity, verifies that $f_{\mathbf{S}}(\mathbf{s})$ is a probability density.

The expected value of the kth component of the random vector $\mathbf{s}$ is

$$E\{S_k\} = \int_S s_k f_{\mathbf{S}}(\mathbf{s}) d\mathbf{s} \tag{14.264}$$

or, for the case where N is even,

$$E\{S_k\} = \frac{(\beta+1)^{\left(\frac{N+2}{2}\right)}\beta^{(M-1)\left(\frac{N+2}{2}\right)}}{M2^{MN/2}\left(\frac{N}{2}!\right)^M}\int_{\mathbf{s}} s_k \prod_{i=1}^{M}\left[s_i^{-(3+N)}\exp\left(\frac{-\beta}{2s_i^2}\right)\right] \times \sum_{j=1}^{M}\exp\left(\frac{-1}{2s_j^2}\right)\mathrm{d}\mathbf{s} \tag{14.265}$$

Equation 14.265 can also be written as

$$E\{S_k\} = \frac{(\beta+1)^{\left(\frac{N+2}{2}\right)}\beta^{(M-1)\left(\frac{N+2}{2}\right)}}{M2^{MN/2}\left(\frac{N}{2}!\right)^M} \times \int_{\mathbf{s}} s_k \sum_{j=1}^{M}\left\{s_j^{-(3+N)}\exp\left(\frac{-(\beta+1)}{2s_j^2}\right)\prod_{\substack{i=1\\ i\neq j}}^{M}\left[s_i^{-(3+N)}\exp\left(\frac{-\beta}{2s_i^2}\right)\right]\right\}\mathrm{d}\mathbf{s} \tag{14.266}$$

Separating the integrals in Equation 14.266 produces

$$E\{S_k\} = \frac{(\beta+1)^{\left(\frac{N+2}{2}\right)}\beta^{(M-1)\left(\frac{N+2}{2}\right)}}{M2^{MN/2}\left(\frac{N}{2}!\right)^M} \times \left[\int_0^\infty s_k^{-(2+N)}\exp\left(\frac{-(\beta+1)}{2s_k^2}\right)\mathrm{d}s_k \prod_{\substack{i=1\\ i\neq j,k}}^{M}\left[\int_0^\infty s_i^{-(3+N)}\exp\left(\frac{-\beta}{2s_i^2}\right)\mathrm{d}s_i\right] + \sum_{\substack{j=1\\ j\neq 1}}^{M}\left\{\int_0^\infty s_k^{-(2+N)}\exp\left(\frac{-(\beta)}{2s_k^2}\right)\mathrm{d}s_k \int_0^\infty s_j^{-(3+N)}\exp\left(\frac{-(\beta+1)}{2s_j^2}\right)\mathrm{d}s_j \times \prod_{\substack{i=1\\ i\neq j,k}}^{M}\left[\int_0^\infty s_i^{-(3+N)}\exp\left(\frac{-\beta}{2s_i^2}\right)\mathrm{d}s_i\right]\right\}\right] \tag{14.267}$$

From Equation 14.175, the solutions to the integrals shown in Equation 14.267 result in

$$E\{S_k\} = \frac{(\beta+1)^{\left(\frac{N+2}{2}\right)}\beta^{(M-1)\left(\frac{N+2}{2}\right)}}{M2^{MN/2}\left(\frac{N}{2}!\right)^M}$$
$$\times\left[2^{-M}\left(\frac{\beta+1}{2}\right)^{-\left(\frac{N+1}{2}\right)}\Gamma\left(\frac{N+1}{2}\right)\left(\frac{\beta}{2}\right)^{-(M-1)\left(\frac{N+2}{2}\right)}\Gamma\left(\frac{N+2}{2}\right)^{(M-1)}\right.$$
$$+(M-1)2^{-M}\left(\frac{\beta}{2}\right)^{-\left(\frac{N+1}{2}\right)}\Gamma\left(\frac{N+1}{2}\right)\left(\frac{\beta+1}{2}\right)^{-\left(\frac{N+2}{2}\right)}$$
$$\left.\times\Gamma\left(\frac{N+2}{2}\right)\left(\frac{\beta}{2}\right)^{-(M-2)\left(\frac{N+2}{2}\right)}\Gamma\left(\frac{N+2}{2}\right)^{(M-2)}\right] \tag{14.268}$$

or

$$E\{S_k\} = \frac{(\beta+1)^{\left(\frac{N+2}{2}\right)}\beta^{(M-1)\left(\frac{N+2}{2}\right)}}{M2^{MN/2}\left(\frac{N}{2}!\right)^M}\frac{2^{MN/2}\left(\frac{N}{2}!\right)^M}{(\beta+1)^{\left(\frac{N+2}{2}\right)}\beta^{(M-1)\left(\frac{N+2}{2}\right)}}$$
$$\times\left[\sqrt{\frac{\beta+1}{2}}\frac{\Gamma\left(\frac{N+1}{2}\right)}{\Gamma\left(\frac{N+2}{2}\right)}+(M-1)\sqrt{\frac{\beta}{2}}\frac{\Gamma\left(\frac{N+1}{2}\right)}{\Gamma\left(\frac{N+2}{2}\right)}\right] \tag{14.269}$$

Since N is an even integer, and utilizing Equation 14.261 and Equation 14.262, Equation 14.269 simplifies to

$$E\{S_k\} = \frac{1}{M}\frac{\sqrt{\pi}2^{1/2-N}(N-1)!}{\left(\frac{N-2}{2}\right)!\left(\frac{N}{2}\right)!}[\sqrt{\beta+1}+(M-1)\sqrt{\beta}] \tag{14.270}$$

Because the shape parameters are identical for all S_k, and $f_{\mathbf{S}}(\mathbf{s})$ is symmetric with respect to all S_k, the mean is the same for all S_k.

Following the same approach as that used to derive Equation 14.268, the expected value of S_k^2 is

$$E\{S_k^2\} = \frac{(\beta+1)^{\left(\frac{N+2}{2}\right)}\beta^{(M-1)\left(\frac{N+2}{2}\right)}}{M2^{MN/2}\left(\frac{N}{2}!\right)^M}$$

$$\times\left[2^{-M}\left(\frac{\beta+1}{2}\right)^{-\left(\frac{N}{2}\right)}\Gamma\left(\frac{N}{2}\right)\left(\frac{\beta}{2}\right)^{-(M-1)\left(\frac{N+2}{2}\right)}\Gamma\left(\frac{N+2}{2}\right)^{(M-1)}\right.$$

$$+(M-1)2^{-M}\left(\frac{\beta}{2}\right)^{-\left(\frac{N}{2}\right)}\Gamma\left(\frac{N}{2}\right)\left(\frac{\beta+1}{2}\right)^{-\left(\frac{N+2}{2}\right)}$$

$$\left.\times\Gamma\left(\frac{N+2}{2}\right)\left(\frac{\beta}{2}\right)^{-(M-2)\left(\frac{N+2}{2}\right)}\Gamma\left(\frac{N+2}{2}\right)^{(M-2)}\right] \tag{14.271}$$

or

$$E\{S_k^2\} = \frac{2}{MN}\left[\frac{\beta+1}{2}+(M-1)\frac{\beta}{2}\right] \tag{14.272}$$

Simplification of Equation 14.272 produces

$$E\{s_k^2\} = \left[\frac{1}{NM}+\frac{\beta}{N}\right] \tag{14.273}$$

Substitution of Equation 14.273 and Equation 14.270 into Equation 14.202 gives the variance of s_k as

$$\sigma_k^2 = \left[\frac{1}{NM}+\frac{\beta}{N}\right]$$

$$-\left[\frac{1}{M}\frac{\sqrt{\pi}2^{1/2-N}(N-1)!}{\left(\frac{N+2}{2}\right)!\left(\frac{N}{2}\right)!}[\sqrt{\beta+1}+(M-1)\sqrt{\beta}]\right]^2 \tag{14.274}$$

Note that the variance is the same for all S_k.

The covariance between any two different components of $\mathbf{S}$ is

$$
\begin{aligned}
E\{S_u S_v\} = \int_{\mathbf{s}} s_u s_v f_{\mathbf{S}}(\mathbf{s})\mathrm{d}\mathbf{s} = {} & \frac{(\beta+1)^{\left(\frac{N+2}{2}\right)}\beta^{(M-1)\left(\frac{N+2}{2}\right)}}{M2^{MN/2}\left(\frac{N}{2}!\right)^M} \\
& \times \Bigg[2\int_0^\infty s_u^{-(2+N)}\exp\left(\frac{-(\beta+1)}{2s_u^2}\right)\mathrm{d}s_u \int_0^\infty s_v^{-(2+N)}\exp\left(\frac{-(\beta)}{2s_v^2}\right)\mathrm{d}s_v \\
& \times \prod_{\substack{i=1 \\ i\neq u,v}}^{M}\left[\int_0^\infty s_i^{-(3+N)}\exp\left(\frac{-\beta}{2s_i^2}\right)\mathrm{d}s_i\right] \\
& + \sum_{\substack{j=1 \\ j\neq u,v}}^{M}\Bigg\{\int_0^\infty s_u^{-(2+N)}\exp\left(\frac{-\beta}{2s_u^2}\right)\mathrm{d}s_u \int_0^\infty s_v^{-(2+N)}\exp\left(\frac{-(\beta)}{2s_v^2}\right)\mathrm{d}s_v \\
& \times \int_0^\infty s_j^{-(3+N)}\exp\left(\frac{-(\beta+1)}{2s_j^2}\right)\mathrm{d}s_j \\
& \times \prod_{\substack{i=1 \\ i\neq j,u,v}}^{M}\left[\int_0^\infty s_i^{-(3+N)}\exp\left(\frac{-\beta}{2s_i^2}\right)\mathrm{d}s_i\right]\Bigg\}\Bigg]
\end{aligned}
\tag{14.275}
$$

From Equation 14.175, evaluation of the integrals shown in Equation 14.275 results in

$$
\begin{aligned}
E\{S_u S_v\} = {} & \frac{(\beta+1)^{\left(\frac{N+2}{2}\right)}\beta^{(M-1)\left(\frac{N+2}{2}\right)}}{M2^{MN/2}\left(\frac{N}{2}!\right)^M}\Bigg[2^{-M}\left(\frac{\beta+1}{2}\right)^{-\left(\frac{N+1}{2}\right)}\Gamma\left(\frac{N+1}{2}\right) \\
& \times\left(\frac{\beta}{2}\right)^{-\left(\frac{N+1}{2}\right)}\Gamma\left(\frac{N+1}{2}\right)\left(\frac{\beta}{2}\right)^{-(M-2)\left(\frac{N+2}{2}\right)}\Gamma\left(\frac{N+2}{2}\right)^{(M-2)} \\
& + (M-2)2^{-M}\left(\frac{\beta}{2}\right)^{-2\left(\frac{N+1}{2}\right)}\Gamma\left(\frac{N+1}{2}\right)^2\left(\frac{\beta+1}{2}\right)^{-\left(\frac{N+2}{2}\right)} \\
& \times\Gamma\left(\frac{N+2}{2}\right)\left(\frac{\beta}{2}\right)^{-(M-3)\left(\frac{N+2}{2}\right)}\Gamma\left(\frac{N+2}{2}\right)^{(M-3)}\Bigg]
\end{aligned}
\tag{14.276}
$$

Based on Equation 14.261 and Equation 14.262, and recalling that N is an even integer, Equation 14.276 simplifies to

$$E\{S_u S_v\} = \left(\frac{\pi\sqrt{B}}{M2^{(2N-1)}}\right)\left[\frac{(N-1)!}{\left(\frac{N-2}{2}\right)!\left(\frac{N}{2}\right)!}\right]^2 (2\sqrt{\beta+1} + (M-2)\sqrt{\beta}) \tag{14.277}$$

The covariance between any two components of the vector $\mathbf{S}$ is

$$C_{s_{u,v}} = E\{S_u S_v\} - E\{S_u^2\} \tag{14.278}$$

Substitution of Equation 14.277 and Equation 14.270 into Equation 14.278 produces

$$C_{s_{u,v}} = \left(\frac{\pi}{M^2 2^{(2N-1)}}\right)\left[\frac{(N-1)!}{\left(\frac{N-2}{2}\right)!\left(\frac{N}{2}\right)!}\right]^2 (\sqrt{\beta+1} - \sqrt{\beta})^2 \tag{14.279}$$

Once again, because the shape parameters are identical for all S_k, and $f_{\mathbf{S}}(\mathbf{s})$ is symmetric with respect to all S_k, the covariance is independent of u and v. Consequently, the full covariance matrix is simply evaluated by computing only two parameters. All of the diagonal elements are equal and are given by Equation 14.274. Similarly, all of the off-diagonal elements are equal and are given by Equation 14.279.

This concludes the analysis of the Case 3 SIRV. The next case looks at another SIRV derived through use of the Laplace transform representation.

14.4.2.2. Case 4

Once again, let the number of vectors in the generalized SIRV matrix be unity, (i.e., $M = 1$). Consider the function

$$g(z) = \sin(z) \tag{14.280}$$

From Ref. 16, the Laplace transform of this function is

$$G(\alpha_{\mathbf{x}}) = (\alpha_{\mathbf{x}}^2 + 1)^{-1} \tag{14.281}$$

Taking N derivatives with respect to α_{x} produces

$$G^{\{N\}}(\alpha_{\mathbf{x}}) = (-1)^N (N!)(2\alpha)(\alpha_{\mathbf{x}}^2 + 1)^{-(N+1)} \tag{14.282}$$

After assuming that N is an even number, substitution of Equation 14.282 into Equation 14.246 yields

$$f_{\mathrm{R}}(r) = \left[\frac{2^N N!}{K_r}\right] r^{(N-1)} \frac{(r^2+\beta)^N}{[(r^2+\beta)^2+1]^{N+1}} u(r) \tag{14.283}$$

Figure 14.16 and Figure 14.17 show plots of this density for various values of β and N, where K_r has been numerically evaluated such that the area under $f_{\mathrm{R}}(r)$ is unity. Substituting Equation 14.280 into Equation 14.244 yields

$$f_{\mathrm{S}}(\mathbf{s}) = \frac{K_{\mathrm{s}}}{2^N} s^{-(3+N)} \exp\left(\frac{-\beta}{2s^2}\right) \sin\left(\frac{1}{2s^2}\right) u(s) \tag{14.284}$$

Figure 14.18 shows a plot of this function. As seen, this is not a valid density since it goes negative for certain values of s. However, it is known that it is not necessary for $f_{\mathbf{S}}(\mathbf{s})$ to be a density function, or even nonnegative.[13] As shown in Section 14.3.7, the Laplace transform representation used in this case guarantees that this function generates a valid SIRV density.

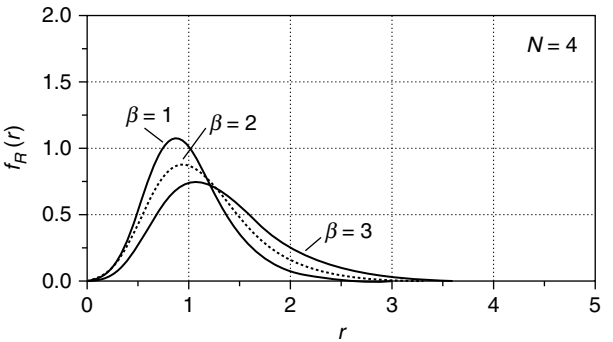

FIGURE 14.16 Case 4 envelope density with β as a parameter.

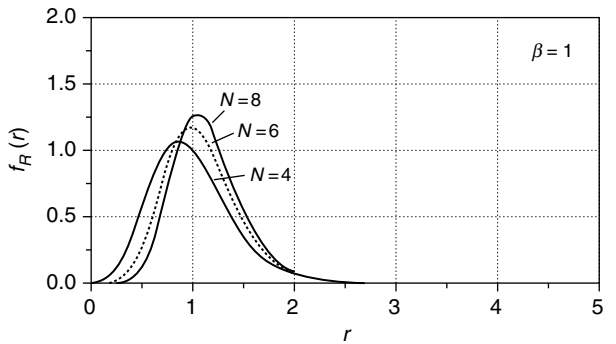

FIGURE 14.17 Case 4 envelope density with N as a parameter.

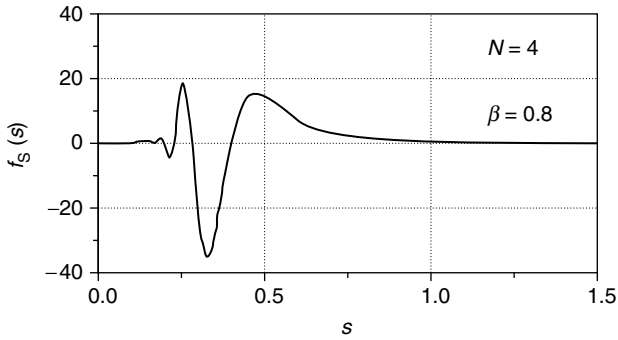

FIGURE 14.18 Case 4 generating function.

For the case where M is greater than unity, and all of the functions in the set $\{g_1(z_1), \ldots, g_M(z_M)\}$ are the same, $f_{\mathbf{S}}(\mathbf{s})$ and $h_{NM}(\alpha_{\mathbf{x}})$ are readily derived from Equation 14.242 and Equation 14.243 to be

$$f_{\mathrm{S}}(\mathbf{s}) = \frac{K_{\mathrm{S}}}{2^{NM}} \cdot \prod_{i=1}^{M} \left[s_i^{-(3+N)} \exp\left(\frac{-\beta_i}{2s_i^2} \right) u(s_i) \right] \sum_{j=1}^{M} \sin\left(\frac{1}{2s_j^2} \right) \tag{14.285}$$

and

$$h_{NM}(\mathbf{a}_{\mathrm{x}}) = (N!)^M 2^N K_{\mathrm{s}} \sum_{j=1}^{M} \left\{ \frac{(\alpha_{\mathrm{x}j} + \beta_j)^N (\alpha_{\mathrm{x}j} + \beta_j + 1)^{-(N+1)}}{\prod\limits_{\substack{i=1 \\ i \neq j}}^{M} (\alpha_{xi} + \beta_i)^{(N+1)}} \right\} \tag{14.286}$$

This concludes the analysis of specific generalized SIRVs. The next chapter describes how to generate random data which follows a particular SIRV density.

14.5. GENERALIZED SIRV GENERATION

This section describes a method for generating random data which follows a particular generalized SIRV density function. The method of generation is based on the generalized SIRV representation theorem, which is

$$\begin{aligned} \mathbf{X}_1 &= \mathbf{Z}_1 S_1 \\ &\vdots \\ \mathbf{X}_M &= \mathbf{Z}_M S_M, \end{aligned} \tag{14.287}$$

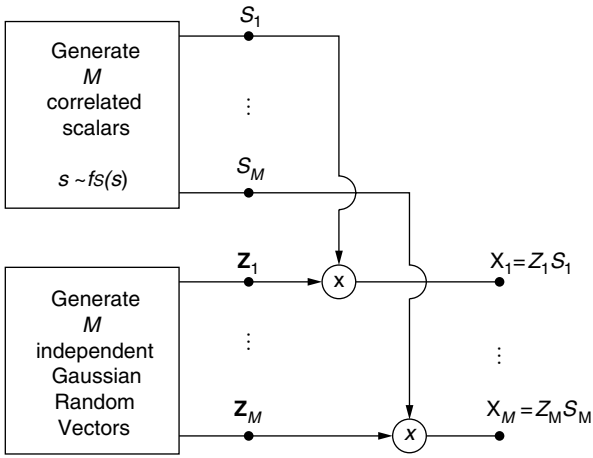

FIGURE 14.19 Generalized SIRV generation.

where $[\mathbf{X}_1, \ldots, \mathbf{X}_M]$ are M dependent SIRVs, $[\mathbf{Z}_1, \ldots, \mathbf{Z}_M]$ are M independent Gaussian random vectors, and $[S_1, \ldots, S_M]$ are M correlated random nonnegative scalars. Figure 14.19 shows the process used to generate generalized SIRVs based upon this theorem. As seen, the method involves first generating the M independent Gaussian random vectors. These vectors are then each multiplied by the appropriate random scalar to generate the M dependent SIRVs.

With reference to Figure 14.19, the generation of M independent Gaussian random vectors is a straightforward procedure. However, a scheme needs to be developed to generate the vector $\mathbf{S}$: $[S_1, \ldots, S_M]$ which follows a particular SIRV generating density $f_{\mathbf{S}}(\mathbf{s})$.

A scheme for generating random scalars is presented in Ref. 15, based on the rejection theorem. The following analysis presents an extension of this work to account for multivariate random variables.

14.5.1. Multivariate Rejection Theorem

The goal of this analysis is to create a method for generating a random vector $\mathbf{s}$, which follows the multivariate density $f_{\mathbf{S}}(\mathbf{s})$, based on generation of another random vector $\mathbf{Q}$, which follows the multivariate density $f_{\mathbf{Q}}(\mathbf{q})$.

Assume that both densities are nonzero over the same interval. In particular, if $f_{\mathbf{S}}(\mathbf{x})$ equals zero, then $f_{\mathbf{Q}}(\mathbf{x})$ also equal zero for all $\mathbf{x}$. If the densities are limited in this fashion, then there exists a finite positive scalar, a, such that

$$f_{\mathbf{Q}}(\mathbf{q})/f_{\mathbf{S}}(\mathbf{q}) \geq a \tag{14.288}$$

for all vectors $\mathbf{Q}$ in the region $\mathbf{R}$ where the densities are nonzero. Define the function $g(\mathbf{q})$ as

$$g(\mathbf{q}) = a{\cdot}f_{\mathbf{S}}(\mathbf{q})/f_{\mathbf{Q}}(\mathbf{q}) \tag{14.289}$$

where $g(\mathbf{q}) \leq 1$ for all vectors $\mathbf{Q}$ in $\mathbf{R}$.

Let V be a random scalar uniformly distributed over the interval [0, 1].

Assume that V is statistically independent of $\mathbf{Q}$. Define M as the event that v is less than or equal to $g(\mathbf{q})$, or

$$M\text{: } [v \leq g(\mathbf{q})] \tag{14.290}$$

The probability of event M is

$$P\{M\} = \int_{\mathbf{q}} \int_{0}^{g(\mathbf{q})} f_{\mathbf{Q},V}(\mathbf{q}, v)\mathrm{d}v\, \mathrm{d}\mathbf{q} \tag{14.291}$$

where

$$\int_{q} \mathrm{d}q = \int_{q_1} \ldots \int_{q_M} \mathrm{d}q_1 \ldots \mathrm{d}q_M \tag{14.292}$$

and $f_{\mathbf{Q},V}(\mathbf{q}, v)$ is the joint density of $\mathbf{q}$ and v.

Since $\mathbf{q}$ and v are independent, Equation 15.109 simplifies to become

$$P\{M\} = \int_{\mathbf{q}} \int_{0}^{g(\mathbf{q})} f_{\mathbf{Q}}(\mathbf{q})\mathrm{d}_v(v)\mathrm{d}v\, \mathrm{d}\mathbf{q} \tag{14.293}$$

Note that $f_V(v)$ is unity over the interval [0, 1]. Furthermore, since $g(\mathbf{q})$ is less than or equal to one, Equation 14.293 reduces to

$$P\{M\} = \int_{\mathbf{q}} g(\mathbf{q}) f_{\mathbf{Q}}(\mathbf{q})\mathrm{d}\mathbf{q} \tag{14.294}$$

Substituting Equation 14.289 into Equation 14.294 produces

$$P\{M\} = \int_{\mathbf{q}} a{\cdot}f_{\mathbf{S}}(\mathbf{q})\mathrm{d}\mathbf{q} = a{\cdot} \int_{\mathbf{q}} f_{\mathbf{S}}(\mathbf{q})\mathrm{d}\mathbf{q} \tag{14.295}$$

Since the densities $f_{\mathbf{Q}}(\mathbf{q})$ and $f_{\mathbf{S}}(\mathbf{q})$ are nonzero over the same region $\mathbf{R}$, the integrals shown in Equation 14.295 span the entire domain of $f_{\mathbf{S}}(\mathbf{q})$. Because the volume under a multivariate density is unity, Equation 14.295 simplifies to become

$$P\{M\} = a \tag{14.296}$$

Define A as the event that the random vector $\mathbf{Q}$ lies within an incremental distance away from a particular realization vector $\mathbf{q}$, or

$$A\text{: } \{q_1 < \mathbf{Q}_1 < q_1 < \Delta q_1, \ldots, q_M < \mathbf{Q}_M < q_M + \Delta q_M\} \tag{14.297}$$

The probability of event A is

$$P\{A\} = \int_{q_1}^{q_1+\Delta q_1} \cdots \int_{q_M}^{q_M+\Delta q_M} f_{\mathbf{Q}}(\mathbf{q}) \mathrm{d}q_1 \ldots \mathrm{d}q_M \tag{14.298}$$

Assuming $\Delta q_1, \ldots, \Delta q_M$ are sufficiently small such that the function $f_{\mathbf{Q}}(\mathbf{q})$ can be approximated by a constant over the region of integration, the probability of event A becomes

$$P\{A\} = f_{\mathbf{Q}}(\mathbf{q}) \Delta q \tag{14.299}$$

where

$$\Delta q = \{\Delta q_1 \cdot, \ldots, \cdot \Delta q_M\} \tag{14.300}$$

From Equation 14.298, the probability of event A given event M is

$$P\{A/M\} = \int_{q_1}^{q_1+\Delta q_1} \cdots \int_{q_M}^{q_M+\Delta q_M} f_{\mathbf{Q}|M}(\mathbf{q}) \mathrm{d}q_1 \ldots \mathrm{d}q_M = f_{\mathbf{Q}|M}(\mathbf{q}) \Delta q \tag{14.301}$$

Furthermore, with reference to Equation 14.293 and Equation 14.298, the joint probability of events A and M is

$$\begin{aligned} P\{A|M\} &= \int_{q_1}^{q_1+\Delta q_1} \cdots \int_{q_M}^{q_M+\Delta q_M} \int_0^{g(\mathbf{q})} f_{\mathbf{Q},v}(\mathbf{q}, v) \mathrm{d}v \, \mathrm{d}\mathbf{q} \\ &= \int_{q_1}^{q_1+\Delta q_1} \cdots \int_{q_M}^{q_M+\Delta q_M} \int_0^{g(\mathbf{q})} f_{\mathbf{Q}}(\mathbf{q}) f_v(v) \mathrm{d}v \, \mathrm{d}\mathbf{q} \\ &= \int_{q_1}^{q_1+\Delta q_1} \cdots \int_{q_M}^{q_M+\Delta q_M} g(\mathbf{q}) f_{\mathbf{Q}}(\mathbf{q}) \mathrm{d}\mathbf{q} = g(\mathbf{q}) f_{\mathbf{Q}}(\mathbf{q}) \Delta \mathbf{q} \end{aligned} \tag{14.302}$$

Bayes' rule gives the joint probability of two events as

$$P\{A, M\} = P\{A|M\} \cdot P\{M\} \tag{14.303}$$

Substituting Equation 14.302, Equation 14.301, and Equation 14.296 into Equation 14.303 produces

$$g(\mathbf{q}) f_{\mathbf{Q}}(\mathbf{q}) \Delta q = f_{\mathbf{Q}|M}(\mathbf{q}) \Delta \mathbf{q} \cdot a \tag{14.304}$$

Inserting the expression for $g(\mathbf{q})$ from Equation 14.289 into Equation 14.304 gives

$$a \cdot f_{\mathbf{S}}(\mathbf{q}) \Delta q = a \cdot f_{\mathbf{Q}|M}(\mathbf{q}) \Delta \mathbf{q} \tag{14.305}$$

Finally, canceling like terms from both sides of Equation 14.305 yields the multivariate rejection theorem,

$$f_{\mathbf{S}}(\mathbf{q}) = f_{\mathbf{Q}|M}(\mathbf{q}) \tag{14.306}$$

14.5.2. Application of the Rejection Theorem

With reference to Equation 14.306, the rejection theorem states that the density of **Q** conditioned on event M equals the density of **S**. To use the rejection theorem, first generate an ensemble of vectors **q** which follow the density $f_{\mathbf{Q}}(\mathbf{q})$. Next, remove all of those vectors **q** for which event **M** does not hold true. The remaining vectors are all conditioned on event M, and thus follow the density $f_{\mathbf{S}}(\mathbf{q})$.

As an example, let $f_{\mathbf{Q}}(\mathbf{q})$ equal a uniform density over the nonzero region **R** of $f_{\mathbf{S}}(\mathbf{q})$, as shown in Figure 14.20 for a univariate case. If $f_{\mathbf{S}}(\mathbf{q})$ does not have finite upper limits, then values must be chosen such that the volume under $f_{\mathbf{S}}(\mathbf{q})$ within these limits is very close to unity.

Assume that the lower boundaries of **R** are zero in all dimensions. Define the upper bounds of **R** to be $\{s_{h1}, \ldots, s_{hM}\}$. Based on these values, the multivariate uniform density $f_{\mathbf{Q}}(\mathbf{q})$ is

$$f_{\mathbf{Q}}(\mathbf{q}) = \begin{cases} 1.0/(s_{h1} \cdot s_{h2} \cdot \ldots \cdot s_{hM}) & \text{when } \{0 < q_1 < s_{h1}, \ldots, 0 < q_M < s_{hM}\} \\ 0.0 & \text{elsewhere} \end{cases} \tag{14.307}$$

From Equation 14.288, there exists a positive scalar, a, such that

$$f_{\mathbf{Q}}(\mathbf{q})/f_{\mathbf{s}}(\mathbf{q}) \geq a, \text{ for all vectors } \mathbf{q} \text{ in } \mathbf{R} \tag{14.308}$$

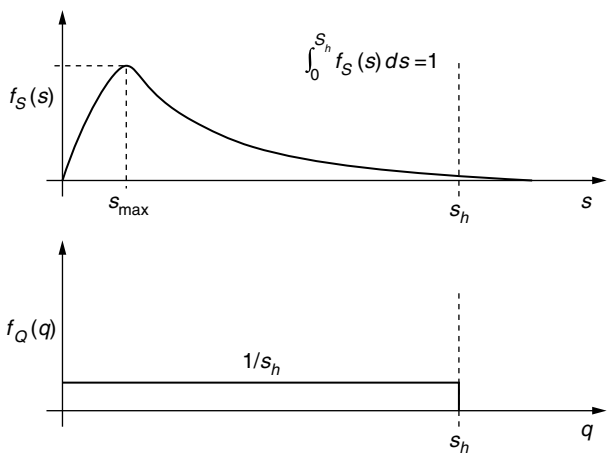

FIGURE 14.20 Rejection theorem densities.

From Equation 14.296, the scalar a equals the probability of event M, or the probability that we do not reject a given $\mathbf{q}$ realization. From a practical standpoint, it is desirable to pick this scalar as large as possible in order to minimize the number of rejections. In Equation 14.308, note that $f_{\mathbf{Q}}(\mathbf{q})$ is constant over $\mathbf{R}$. The maximum value of the scalar a is thus governed by the maximum value of $f_{\mathbf{S}}(\mathbf{q})$. Define $\mathbf{s}_{\max}$ as the vector at which the function $f_{\mathbf{S}}(\mathbf{s})$ is a maximum. Using this notation, the maximum value of the scalar a is

$$a = (s_{h1} \cdot s_{h1} \cdot \ldots \cdot s_{hM})^{-1} / f_{\mathbf{S}}(\mathbf{s}_{\max}) \tag{14.309}$$

Equation 14.289 gives the function $g(\mathbf{q})$ as

$$g(\mathbf{q}) = a \cdot f_{\mathbf{S}}(\mathbf{q}) / f_{\mathbf{Q}}(\mathbf{q}) \tag{14.310}$$

Substituting Equation 14.307 and Equation 14.309 into Equation 14.310 produces

$$g(\mathbf{q}) = f_{\mathbf{S}}(\mathbf{q}) / f_{\mathbf{S}}(\mathbf{s}_{\max}) \tag{14.311}$$

This is the function used to determine whether or not to reject a given realization of $\mathbf{q}$.

To summarize, the steps taken to generate random data which follows the density $f_{\mathbf{S}}(\mathbf{s})$ are listed below:

(1) Determine the upper bounds of $\mathbf{R}$, the nonzero region of $f_{\mathbf{S}}(\mathbf{s})$.
(2) Find the maximum value of $f_{\mathbf{S}}(\mathbf{s})$, $f_{\mathbf{S}}(\mathbf{s}_{\max})$.
(3) Generate an ensemble of vectors $\mathbf{q}$ which follow the uniform density $f_{\mathbf{Q}}(\mathbf{q})$ given in Equation 14.307.
(4) For each vector $\mathbf{q}$, generate the random scalar v uniformly distributed over [0, 1].
(5) Reject all those vectors $\mathbf{q}$ which do not satisfy $v < g(\mathbf{q})$, where $g(\mathbf{q})$ is given in Equation 14.311.
(6) The remaining vectors can be considered to be samples from $f_{\mathbf{S}}(\mathbf{q})$.

These vectors can now be used to generate the desired SIRV, as shown in Figure 14.19.

14.5.3. Examples of Random Variable Generation

This section presents two examples of random variable generation via the rejection theorem. The first example presents the generation of a univariate random scalar, while the second presents the generation of a multivariate random vector.

14.5.3.1. Example 1

Consider the Case 1 density derived in Section 14.4.1.1. Let the number of generalized SIRV's equal unity, (i.e., $M = 1$). Furthermore, let the shape

parameter β equal (1/2) and the sum $(Q + R)$ equal 5. From Equation 14.167 and Equation 14.166, the resulting SIRV generating density function $f_S(s)$ is

$$f_S(s) = \frac{1}{8s^5} \exp\left\{\frac{-1}{4s^2}\right\} \tag{14.312}$$

Since the random variable s is nonnegative, the lower bound on $\mathbf{R}$ is zero, and the upper bound is infinity. A more practical upper limit is

$$s_h = 3 \tag{14.313}$$

Since

$$\int_0^3 f_S(s)\mathrm{d}s > 0.9996 \tag{14.314}$$

The maximum value of $f_S(s)$ is found at $s_{\max} = 0.316$. This maximum is

$$f_S(s_{\max}) = 3.245 \tag{14.315}$$

Substituting Equation 14.313 and Equation 14.315 into Equation 14.309 gives the probability of acceptance as

$$a = 1/(3 \times 3.245) = 0.103 \tag{14.316}$$

The steps outlined above can now be used to generate random scalars which follows $f_S(s)$. Figure 14.21 shows the results. To generate this data, 1000 uniform random variable realizations q were generated over the interval [0, 3], of which 102 were accepted by the theorem. These 102 random variable realizations were used to create the histogram. As seen, the histogram closely follows the desired density.

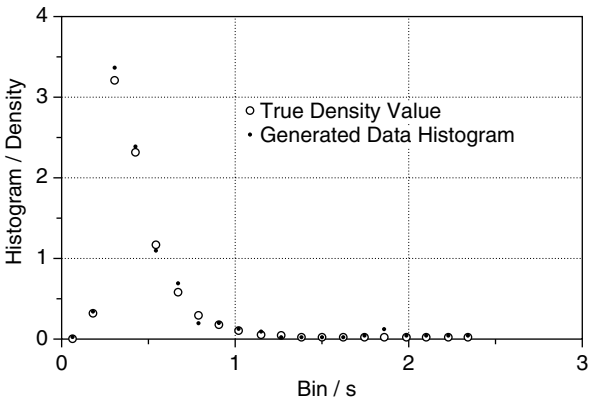

FIGURE 14.21 Rejection theorem Example 1.

14.5.3.2. Example 2

Consider once again the Case 1 density derived in Section 14.4.1.1. For this example, let the number of generalized SIRVs equal two, (i.e., $M = 2$). Furthermore, let the shape parameters β_1 and β_2 equal (1/2), Q_1 and Q_2 equal three, and R_1 and R_2 equal two. From Equation 14.167 and Equation 14.176, the resulting symmetric density is

$$f_{S_1,S_2}(s_1, s_2) = \frac{1}{32 s_1^3 s_2^3}\left(\frac{1}{s_1^2} + \frac{1}{s_2^2}\right)\exp\left\{\frac{-1}{4s_1^2} + \frac{1}{4s_2^2}\right\} \tag{14.317}$$

With maximum boundaries of eight, (i.e., $s_{h1}, s_{h2} = 8$), the volume under this density is greater than 0.9661. The maximum value of the density is 4.96. From Equation 14.309, the probability of acceptance is 0.003. Figure 14.22 shows a histogram of data which follows this density, as generated by the rejection theorem. To create these plots, 1,000,000 vectors (q_1, q_2) were generated from the multivariate density over the region given by $\{0 \leq (q_1, q_2) \leq 8\}$. Of these vectors, 3320 were accepted. These accepted vectors were used to generate the histogram. Once again, the histogram closely follows the desired density.

This example illustrates a potential problem associated with using a uniform density of $f_{\mathbf{Q}}(\mathbf{q})$; the probability of acceptance may be quite low (.003 in the above example). This problem becomes especially significant as the dimensionality of the density increases. In general, the probability of acceptance rises as $f_{\mathbf{Q}}(\mathbf{q})$ approaches $f_{\mathbf{S}}(\mathbf{s})$. As an example, the univariate density shown in Figure 14.23, which consists of three linear segments, yields a higher probability of acceptance than a uniform density. When $f_{\mathbf{Q}}(\mathbf{q})$ is made up of linear pieces in this fashion, note that the $\mathbf{q}$ samples can be generated from the inverse distribution function.

This section presented a method for generating random data which follows a particular SIRV density. The next section describes a method for approximating the SIRV density underlying a particular set of random data.

14.6. GENERALIZED SIRV DENSITY APPROXIMATION

This section describes a method for approximating the SIRV density underlying a particular set of random data, as based on the work of Ozturk[20] and Rangaswamy.[14] The Ozturk algorithm accurately approximates the underlying distribution of univariate random data, even when presented with a relatively small number of samples. Rangaswamy used this algorithm to approximate the underlying distribution of an SIRV by processing the quadratic form generated from the components of the vector. Assuming an identity covariance matrix, this quadratic form is the square of the envelope. As seen in Equation 14.84, the envelope density of an SIRV is unique. Approximating the density of the envelope thus reduces the multivariate random vector approximation problem to a univariate approximation problem.

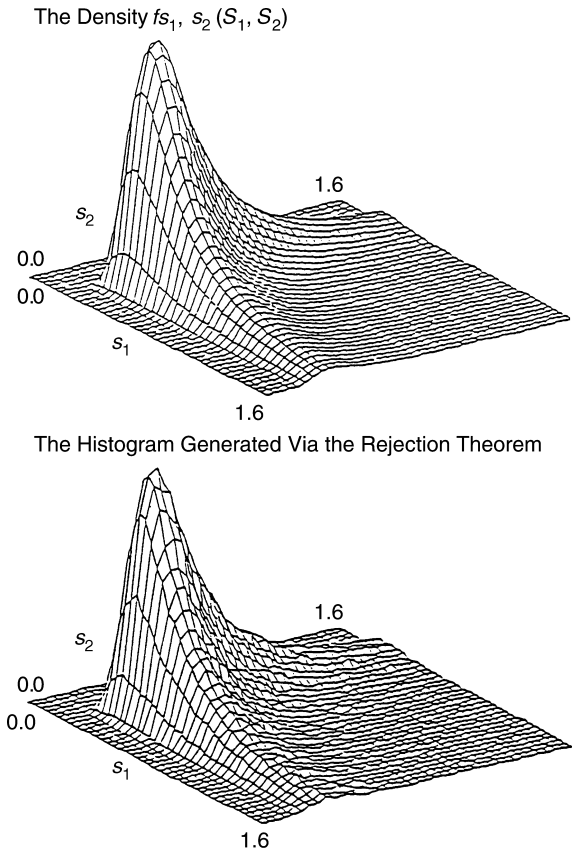

FIGURE 14.22 Rejection theorem Example 2.

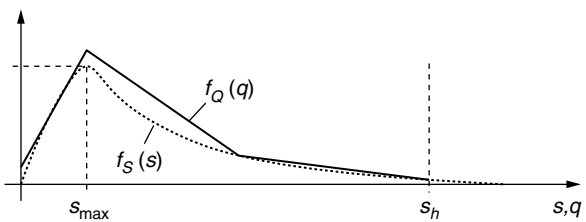

FIGURE 14.23 Alternate $f_Q(q)$ density.

However, generalized SIRV's have associated with them M different envelopes, hence the work described above must be extended. The analysis begins with a summary of Ozturk's algorithm in the following section. Subsequent sections describe how to extend the algorithm to account for

generalized SIRV approximation. The analysis concludes with an examination of recorded SONAR data.

14.6.1. Univariate Density Approximation

This chapter summarizes Ozturk's algorithm for approximating the underlying distribution of a random scalar. Consider a set of N_s samples, $\{x_1, \ldots, x_{N_s}\}$, generated from the density $f_X(x)$. Define $x_{(k)}$ as the kth ordered sample in the set. For each ordered sample, define the vector $\mathbf{u}_k$ as

$$\mathbf{u}_k = (|x_{(k)}|\cos\,\theta_k, |x_{(k)}|\sin\,\theta_k) \tag{14.318}$$

where

$$\theta_k = \pi F_{\text{ref}}(m_{k:N_s}) \tag{14.319}$$

$F_{\text{ref}}(\cdot)$ is the distribution function of a specified reference density, and $m_{k:N_s}$ is the expected value of the kth ordered statistic from an N_s sized sample of that reference density. The creation of the vector $\mathbf{u}_k$ involves mapping one dimensional data into a two dimensional plane, as shown in Figure 14.24. Note that the magnitude of the vector $\mathbf{u}_k$ equals the magnitude of the scalar $x_{(k)}$.

Define the vector $\mathbf{u}$ as the normalized sum of the individual vectors $\mathbf{u}_k$, or

$$\mathbf{u} = \frac{1}{N_s}\sum_{k=1}^{N_s}\mathbf{u}_k \tag{14.320}$$

As shown in Figure 14.25, the vector $\mathbf{u}$ is the end-point of a trajectory made up of the N_s individual vectors $\mathbf{u}_k$.

Ozturk's algorithm involves generating many samples of length N_s from a known density $f_X(x)$. A separate $\mathbf{u}$ vector is then created for each set of N_s samples. The end-points of these vectors form a cloud in (u_1, u_2) space, as shown in Figure 14.26. Define the average of all these $\mathbf{u}$ vectors as $\bar{\mathbf{u}}$.

The above procedure can be repeated for as many different densities as desired. This leads to formation of a "map" in (u_1, u_2) space, where different points correspond to the computed value of $\bar{\mathbf{u}}$ for different densities, as shown in Figure 14.27.

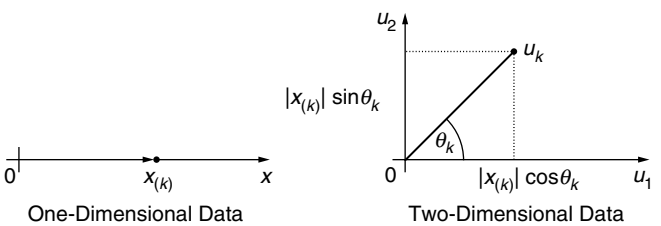

FIGURE 14.24 Ozturk algorithm univariate mapping.

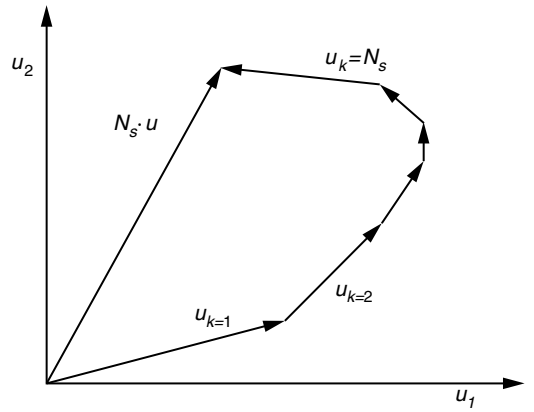

FIGURE 14.25 Ozturk algorithm trajectory.

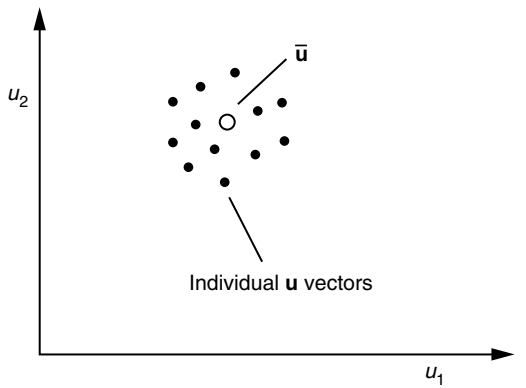

FIGURE 14.26 Ozturk algorithm **u** vector cloud.

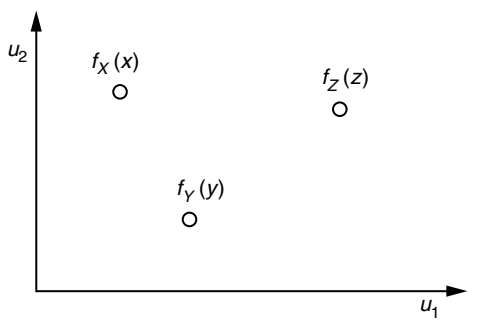

FIGURE 14.27 Ozturk algorithm map.

To approximate the density of an unknown sample, compute **u** as described above and plot the end-point location on the map. By definition, the chosen density is the one whose value of $\bar{\mathbf{u}}$ lies closest to the unknown data **u** vector. Obviously, the final accuracy of the approximation procedure improves as N_s and the number of $\bar{\boldsymbol{u}}$ points on the map increases.

14.6.2. 2-D Density Approximation

Now consider a sample set of N_s random vectors where each vector contains two components. The sample can thus be expressed as

$$\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N_s}\} \tag{14.321}$$

where $\mathbf{x}_i$ is the ith random vector with $\mathbf{x}_i^{\mathrm{T}} = (\mathbf{x}_{i,1}, \mathbf{x}_{i,2})$. Define $\mathbf{x}_{(k)}$ as the kth ordered sample of this set, as based on the magnitude of the vectors. In order to form the vector $\mathbf{u}_k$ associated with $\mathbf{x}_{(k)}$, rotate $\mathbf{x}_{(k)}$ into the third dimension as shown in Figure 14.28. In this figure, the angle of rotation equals the same angle as that shown in Figure 14.24.

Equation 14.319 gives the value of this angle.

Based on Figure 14.28, the components of vector $\mathbf{u}_k$ are

$$u_{k,1} = |x_{(k)}|\cos\,\theta_k \cos\,\phi = |x_{(k),1}|\cdot\cos\,\theta_k \tag{14.322}$$

$$u_{k,2} = |x_{(k)}|\cos\,\theta_k \sin\,\phi = |x_{(k),2}|\cdot\cos\,\theta_k \tag{14.323}$$

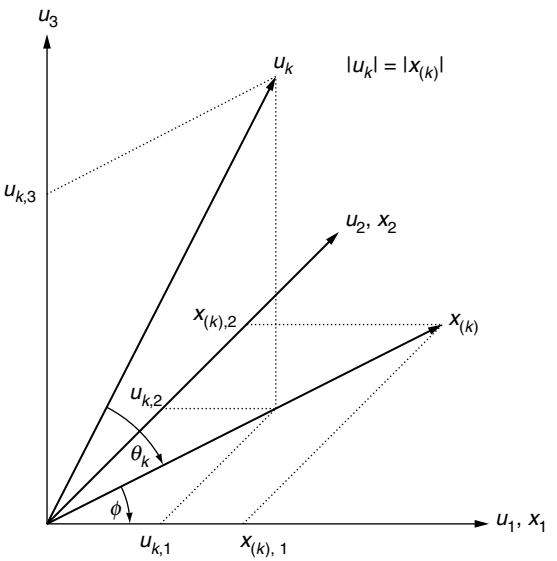

FIGURE 14.28 Ozturk algorithm 2-D mapping.

and

$$u_{k,3} = |x_{(k)}|\sin\,\theta_k = (x_{(k),1}^2 + x_{(k),2}^2)^{1/2}\cdot\sin\,\theta_k \tag{14.324}$$

In Equation 14.322 and Equation 14.323, the absolute value of $x_{(k),1}$ and $x_{(k),2}$ forces $\mathbf{u}_k$ into the positive sector of 3 D space. Note that the magnitude of $\mathbf{u}_k$ equals the magnitude of $\mathbf{x}_{(k)}$, just as in the one-dimensional case.

The set of vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_{N_s}\}$ are now summed and normalized to form the vector $\mathbf{u}$,

$$\mathbf{u} = \frac{1}{N_s}\sum_{k=1}^{N_s}\mathbf{u}_k \tag{14.325}$$

Once again, this vector is averaged over many trials to form one $\bar{\mathbf{u}}$ vector for each density on the map. This time, however, the map is three-dimensional.

14.6.3. Multivariate Density Approximation

Now consider a sample set of N_s random vectors, where each vector contains N_x components. The data sample can be expressed as

$$\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{N_s}\} \tag{14.326}$$

where $\mathbf{x}_i$ is the ith random vector with $\mathbf{x}_i^{\mathrm{T}} = (x_{i,1}, x_{i,2}, \ldots, x_{i,N_x})$. Define $\mathbf{x}_{(k)}$ as the kth ordered sample of this set, as based on the magnitude of the vectors. In order to form the vector $\mathbf{u}_k$ associated with $\mathbf{x}_{(k)}$, rotate $\mathbf{x}_{(k)}$ into the $(N_x + 1)$th dimension using the same angle of rotation as that defined by Equation 14.319.

The components of vector $\mathbf{u}_k$ are

$$u_{k,1} = |x_{(k),1}|\cdot\cos\,\theta_k \tag{14.327}$$

$$u_{k,2} = |x_{(k),2}|\cdot\cos\,\theta_k \tag{14.328}$$

$$\vdots$$

$$u_{k,N_x} = |x_{(k),N_x}|\cdot\cos\,\theta_k \tag{14.329}$$

and

$$u_{k,N_x+1} = (x_{(k),1}^2 + \cdots + x_{(k),N_x}^2)^{1/2}\cdot\sin\,\theta_k \tag{14.330}$$

Just as in the two-dimensional case, the vectors are added together and normalized to form the vector $\mathbf{u}$. Note that the final approximation map which arises by averaging the $\mathbf{u}$ vectors is of dimensionality (N_{x+1}).

The previous sections describe how to extend Ozturk's algorithm into multi-dimensions. The next section presents an example of real data processed in this fashion.

14.6.4. Real Data Analysis

As explained above, approximating the densities of real data requires creation of an approximation map. The first step in this process involves defining the angles $\{\theta_1, \ldots \theta_{N_s}\}$. From Equation 14.319, these angles are given by

$$\theta_k = \pi F_{\text{ref}}(m_{k:Ns}) \tag{14.331}$$

where $F_{\text{ref}}(m_{k:Ns})$ is the distribution function of a specified reference density, and $m_{k:Ns}$ is the expected value of the kth ordered statistic from that reference density.

For analytic simplicity, this example employs as a reference the Rayleigh density function,

$$f_{\text{ref}}(x) = x\, \mathrm{e}^{-x^2/2}, \qquad x > 0 \tag{14.332}$$

The Rayleigh distribution function is

$$F_{\text{ref}}(z) = \int_0^z f_{\text{ref}}(x)\mathrm{d}x = 1 - \mathrm{e}^{-z^2/2} \tag{14.333}$$

From Ref. 18, the density of the kth order statistic from a sample of size N_s is

$$f_{X_{(k)}}(x) = \left(\frac{N_s!}{(k-1)!(N_s-k)!}\right)[F_X(x)]^{(k-1)}[1 - F_X(x)]^{(N_s-k)} f_X(x) \tag{14.334}$$

Substituting Equation 14.333 and Equation 14.332 into Equation 14.334 produces

$$f_{\text{ref}_{(k)}}(x) = \left(\frac{N_s!}{(k-1)!(N_s-k)!}\right) x(1 - \mathrm{e}^{-x^2/2})^{(k-1)}\mathrm{e}^{-(N_s-k+1)x^2/2}, \quad x > 0 \tag{14.335}$$

The expected value of the kth ordered statistic is

$$\begin{aligned} m_{k:N_s} &= \int_0^\infty x \cdot f_{\text{ref}_{(k)}}(x)\mathrm{d}x \\ &= \left(\frac{N_s!}{(k-1)!(N_s-k)!}\right)\int_0^\infty x^2(1 - \mathrm{e}^{-x^2/2})^{(k-1)}\mathrm{e}^{-(N_s-k+1)x^2/2}\mathrm{d}x \end{aligned} \tag{14.336}$$

In Equation 14.336, let

$$u = x^2/2 \tag{14.337}$$

and

$$\mathrm{d}u = x\,\mathrm{d}x \tag{14.338}$$

Substituting Equation 14.337 and Equation 14.338 into Equation 14.336 gives

$$m_{k:N_s} = \left(\frac{\sqrt{2}N_s!}{(k-1)!(N_s-k)!} \right) \int_0^\infty \sqrt{u}(1-\mathrm{e}^{-u})^{(k-1)}\mathrm{e}^{-(N_s-k+1)u}\,\mathrm{d}u \quad (14.339)$$

From Ref. 19, the solution to the integral shown above results in

$$\begin{aligned} m_{k:N_s} = & \left(\frac{\sqrt{2}N_s!}{(k-1)!(N_s-k)!} \right) \frac{1}{2}\sqrt{\pi}(-1)^{(k-1)} \\ & \times \sum_{j=0}^{(k-1)} \left[(-1)^j \frac{(N_s-j)^{-3/2}}{j!(k-1-j)!} \right] \end{aligned} \quad (14.340)$$

Equation 14.340 and Equation 14.333 can now be combined with Equation 14.331 to find the desired angles θ_k. In practice, the specific values of the angles are not as important as the consistency; that is, the same set of angles must be used both to create every point on the map and to approximate the unknown data density.

The next step in the analysis involves creating the approximation map. This is accomplished by generating random vectors which follow a particular density. In this example, the length of each random vector is two. Define the number of vectors in a given sample set as N_s. With reference to Figure 14.25, N_s corresponds to the number of vectors used to form a given vector $\mathbf{u}$. Define the total number of sample sets as N_u, where N_u is the number of $\mathbf{u}$ vectors averaged together to form $\bar{\mathbf{u}}$, the final point on the approximation map.

The map created for this example utilizes the four SIRV envelope densities derived in Section 14.4. To begin with, consider Equation 14.185 which gives the envelope density for case 1 where the number of generalized SIRV's is two. This is a two-dimensional density function, where each dimension corresponds to the envelope of one SIRV. The shape parameters were chosen to be $Q_1 = Q_2 = 5$, $R_1 = R_2 = 2$, and $\beta_1 = \beta_2 = 0.2$, resulting in a symmetric density. A computer subroutine was developed to generate $N_\mathrm{u} = 2000$ sample sets of random vectors which follow this density. The generation scheme is based on the generalized rejection theorem, as outline in Section 14.5.

As described above, one $\mathbf{u}$ vector was computed for each generated sample set. Figure 14.29 shows plots of the estimated standard deviation and mean of the components of $\mathbf{u}$ as a function of the size of each sample set, N_s. As seen, the standard deviation is less than .02 for sizes above 200 samples. For this reason, a sample size of 200 was used to create the approximation map. Note that the statistics of u_1 equal those of u_2 because the density is symmetric.

The approximation map vector $\bar{\mathbf{u}}$ is the sample mean of the $N_\mathrm{u} = 2000$ $\mathbf{u}$ vectors and has components $\bar{u}_1, \bar{u}_2, \bar{u}_3$. Since the density is symmetric, the sample mean of $\bar{u}_1$ equals $\bar{u}_2$, and the map needs only consider two dimensions; $\bar{u}_3$ and either $\bar{u}_1$ or $\bar{u}_2$. Figure 14.30 shows the resulting approximation map for the case 1 density. Each point on this chart corresponds to a different set of shape

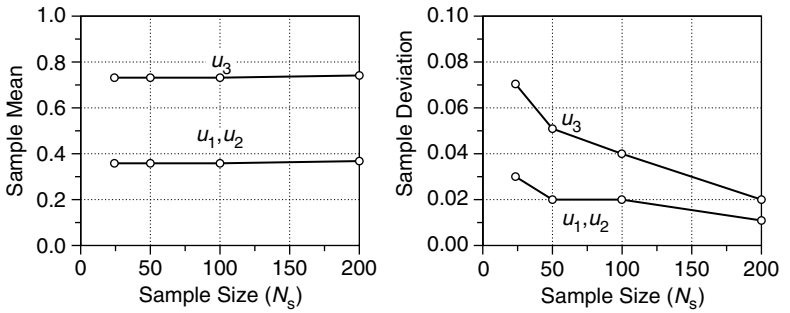

FIGURE 14.29 The effects of changing the sample size.

parameters. $N_u = 2000$ sets of length $N_s = 200$ were averaged to create each point.

The next density examined is the case 2 density from Section 14.4.1.2. The characteristic function of this generalized SIRV density, $h_{NM}(\mathbf{a}_x)$, is given in Equation 14.238. This function is inserted into Equation 14.84 to find the envelope density. Note that even if the shape parameters are identical in all dimensions, this density is not symmetric. Table 14.3 presents a list of the 3D approximation map values for various values of the shape parameters.

Sections. 14.4.2.1 and Section 14.4.2.2 derive the case 3 and case 4 densities. Equation 14.252 and Equation 14.286 give the characteristic functions for these cases. Figure 14.31 shows the approximation map formed after assuming symmetric densities, (i.e., the single shape parameter β is the same in all dimensions).

The final density examined is that which arises when the components of the random vector follow the standard normal Gaussian distribution. From Ref. 18,

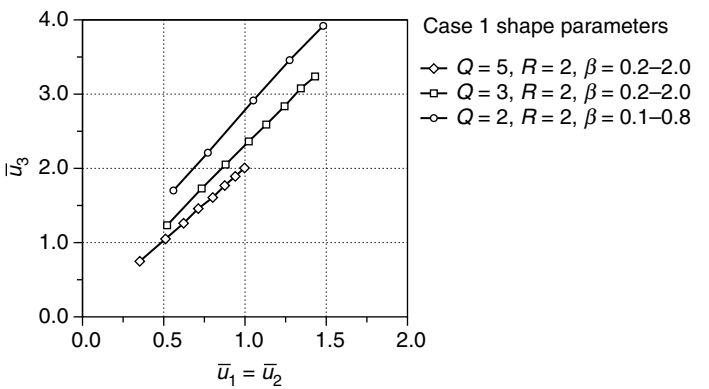

FIGURE 14.30 The case 1 approximation map.

TABLE 14.3
Case 2 Approximation Map

Shape Parameters		ID Chart Values		
γ	β	$\bar{u}_1$	$\bar{u}_2$	$\bar{u}_3$
4	1.0	1.76	1.62	3.05
3	1.0	1.90	1.73	3.31
2	1.0	2.08	1.88	3.60
1	1.0	2.51	2.21	4.36
4	0.5	1.79	1.57	3.16
3	0.5	1.92	1.67	3.36
2	0.5	2.14	1.83	3.77
1	0.5	2.60	2.16	4.59
4	0.1	2.05	1.50	3.77
3	0.1	2.18	1.60	3.97
2	0.1	2.45	1.76	4.43
1	0.1	2.97	2.08	5.36

the envelope of such a vector follows the density

$$f_R(r) = \frac{r^{(n-1)}e^{-r^2/2}}{2^{\left(\frac{n}{2}-1\right)}[(n/2)-1]!} \tag{14.341}$$

where n corresponds to the number of vector components (2 in this example).

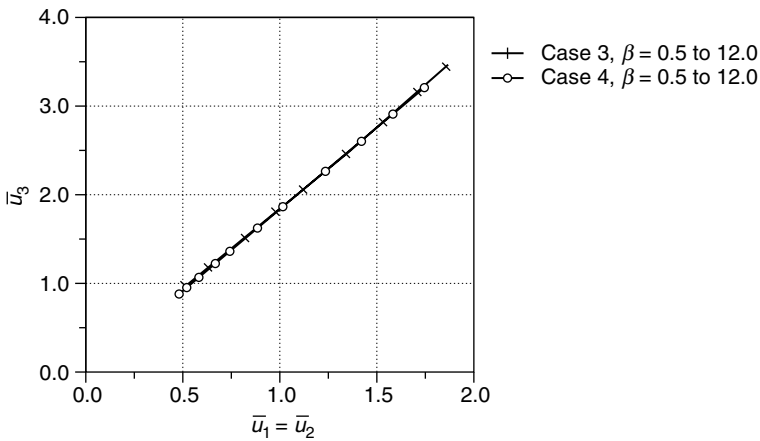

FIGURE 14.31 Cases 3 and 4 approximation maps.

$N_u = 2000$ sets of length $N_s = 200$ were processed to create $\bar{\mathbf{u}}$ on the approximation map. It was observed that $(\bar{u}_1, \bar{u}_2) = 1.62$, and $\bar{u}_3 = 2.70$.

The last step in the analysis involves processing the real data. In this example, two data tapes were examined. The first contains acoustic echoes off the surface of the ocean, called surface reverberation, and the second contains echoes from naturally occurring small organic particles suspended in the water, called volume reverberation. One sample set was processed from each tape. As explained in Section 14.2, the digitized spatially active returns from a given sonar ping make up the components of the SIRV. For this analysis, two aligned returns of 8 digital samples from the same range cells for successive pings make up two vectors in the generalized SIRV matrix.

A sample set of $N_s = 200$ independent vector pairs was read from each tape. Before being processed by the Ozturk algorithm, the sample mean was subtracted from each vector in the sample set, and the result divided by sample standard deviation. As explained in Section 14.3.1, the characteristic function of a generalized SIRV remains invariant to linear transforms of this nature. The envelope was then found for each vector to form the final generalized Ozturk algorithm sample set of $N_S = 200$ envelope pairs. Table 14.4 shows the result of processing this normalized envelope data with the generalized Ozturk algorithm, as a function of the sample size, N_S.

Note that the average $\bar{u}_1$ and $\bar{u}_2$ values lie close to one another, especially when compared to $\bar{u}_3$. Because of this, the data was plotted on the 2D approximate map after assuming a symmetric density and averaging the sample means of $\bar{u}_1$ and $\bar{u}_2$. Figure 14.32 shows the result for a sample size of $N_S = 200$ vectors.

From this chart it is possible to approximate the densities followed by the reverberation echoes. For surface reverberation, the closest point belongs to the case 1 density, with shape parameters of $Q = 2$, $R = 2$, and $\beta = 0.2$. For volume reverberation, the closest point belongs to either the Gaussian envelope or the case 4 density with shape parameter $\beta = 10.0$.

To help confirm the accuracy of these density approximations, Figure 14.33 shows histograms of the real data. Figure 14.34 shows the case 1 and case 4 approximation densities. Figure 14.35 shows the Gaussian envelope density.

TABLE 14.4
Real Data Results

Sample Size	Volume Reverberation			Surface Reverberation		
N_S	$\bar{u}_1$	$\bar{u}_2$	$\bar{u}_3$	$\bar{u}_1$	$\bar{u}_2$	$\bar{u}_3$
25	1.83	1.71	3.17	0.57	0.66	1.08
50	1.82	1.58	3.01	0.62	0.74	1.19
100	1.71	1.58	2.87	0.78	0.73	2.06
200	1.58	1.47	2.71	0.74	0.63	1.92

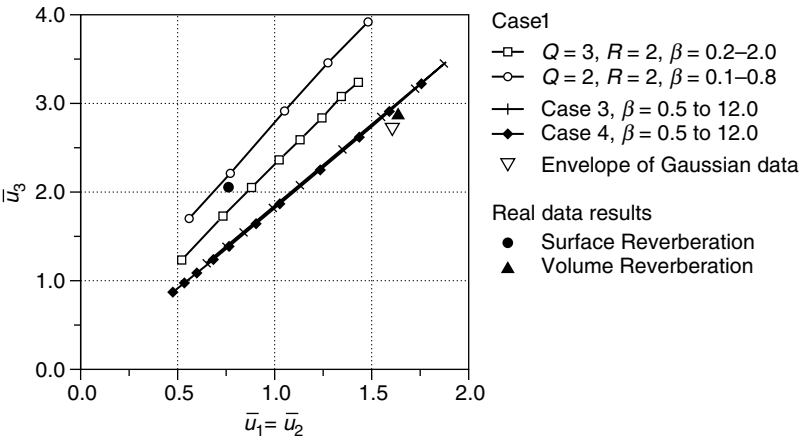

FIGURE 14.32 Ozturk algorithm: real data results.

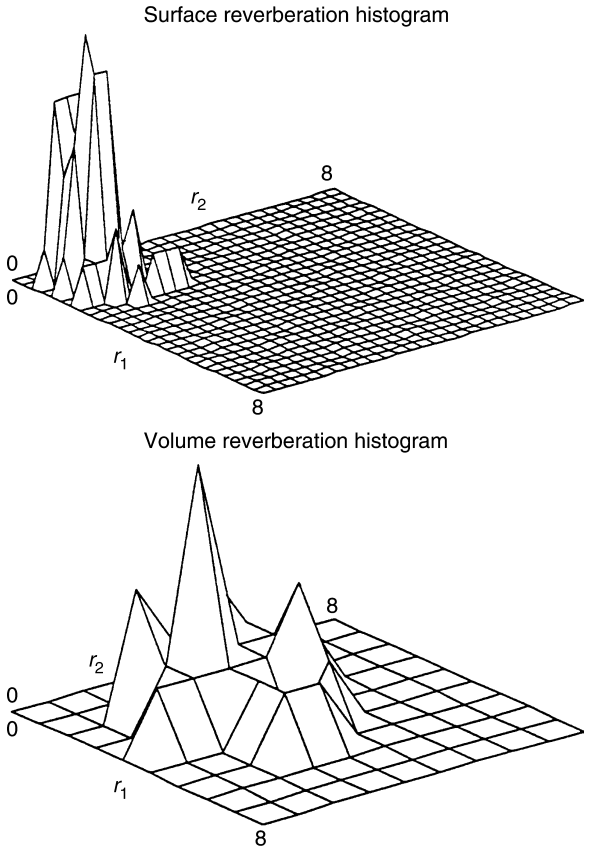

FIGURE 14.33 Real data histograms.

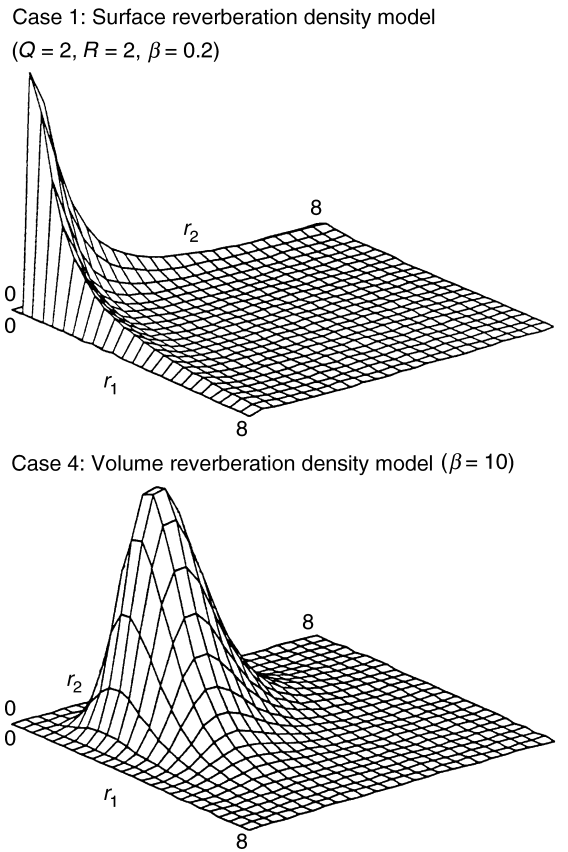

FIGURE 14.34 Approximation densities.

Since only $N_S = 200$ vector pairs make up the sample set, the resulting real data histograms are not an accurate indicator of the underlying density. Real data tends to be nonstationary and nonhomogeneous. Consequently, in practice, it is difficult to collect enough data to generate accurate histograms. It is for this reason that the generalized Ozturk algorithm is required in the first place.

In summary, this section presents a method for approximating the density function of multi-dimensional random data. This method was employed to find approximations for the densities of real sonar surface and volume reverberations. The volume reverberation is assumed to follow a Gaussian density. In contrast, the surface reverberation is nonGaussian. The density function of the examined surface echo data is approximated by case 1 which was derived in Section 14.4.1.

The processing of Gaussian data is a straightforward and well-studied problem. In contrast, this dissertation deals with the processing of nonGaussian data, as specifically modeled by generalized SIRVs. An application well suited

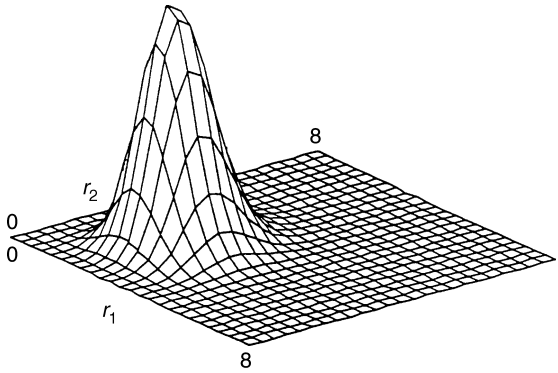

FIGURE 14.35 Envelope of Gaussian data density.

for this analysis is that of a correlation sonar. A correlation sonar measures own-ship velocity by comparing the return from two active pulses which ensonify the same target volume at different times. Interference arises when echoes from the first pulse intrude on the data collected during the second pulse window. When these previous pulse echoes ensonify an ocean boundary, the resulting interference power greatly exceeds that of the desired second pulse signal, leading to significant degradations in performance. As shown above, these surface echoes may follow nonGaussian densities. Consequently, the cancellation of previous pulse boundary echoes in correlation sonar data presents itself as a problem uniquely suited for this dissertation. The next section begins this analysis with an introduction of correlation sonar fundamentals.

14.7. CORRELATION SONAR FUNDAMENTALS

This section begins with a description of how a correlation sonar operates, as presented in Refs. 8–12. A derivation of a sub-optimal receiver then follows. A detailed analysis of the correlation sonar signal and interference concludes this section.

14.7.1. Correlation Sonar Basic Operation

A correlation sonar measures own-ship velocity by comparing the return from two active pulses which ensonify the same volume at different times. A typical system consists of two perpendicular receive arrays and a separate projector mounted to the hull of a ship, as shown in Figure 14.36. In this case, the projector generates two successive pulses which propagate down and echo off the bottom. By means of gating, the *X* array receives a bottom echo from the first pulse, while the *Y* array receives one from the second.

In the far field, any of the bistatic projector-receive element pairs shown in Figure 14.36 can be modeled as a monostatic sonar located halfway between the

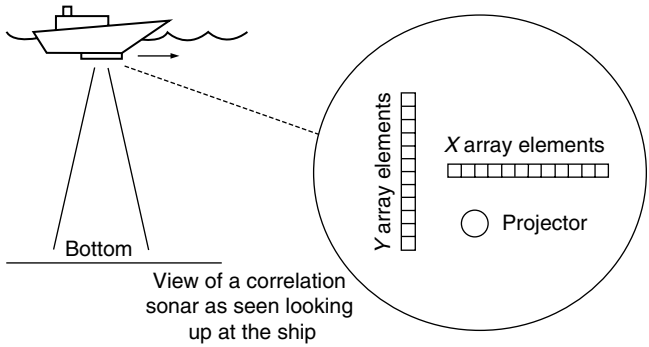

FIGURE 14.36 A hull mounted correlation sonar.

two. As such, in the far field, a correlation sonar system looks like a series of monostatic sonars aligned in a "T" pattern. To determine own-ship velocity, each X array sonar first samples a different region of the ocean floor, as shown in Figure 14.37. Then, after a small period of time, each Y array sonar samples the ocean floor. In both cases, the sonar beam pattern limits the size of the ensonified region.

Own-ship velocity allows a Y array element region to overlap with that of an X array element. After transmission of the first pulse, the platform moves, which brings the Y array over to the same general area as that previously occupied by the X array. It is now possible for one of the ensonified regions seen by a Y element to overlap with that seen by an X element, as shown in Figure 14.38.

The system uses samples from the bottom echo in order to estimate the correlation of the signal received on an X element with that received on a Y element. If the scattering characteristics of the bottom do not change during the time interval between pulses, then this correlation, in effect, measures how much overlap occurs between the region ensonified by the X pulse and the region

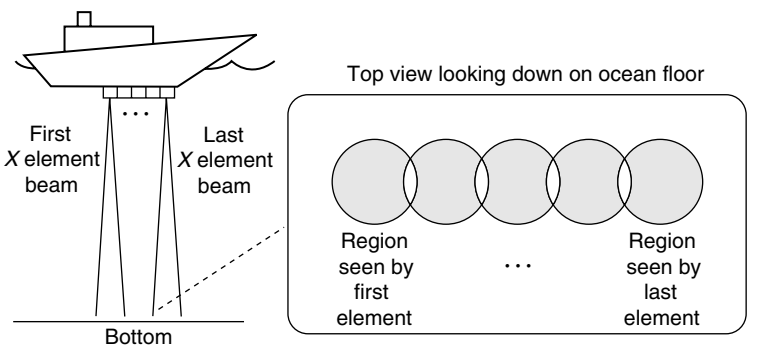

FIGURE 14.37 Regions ensonified by the first pulse.

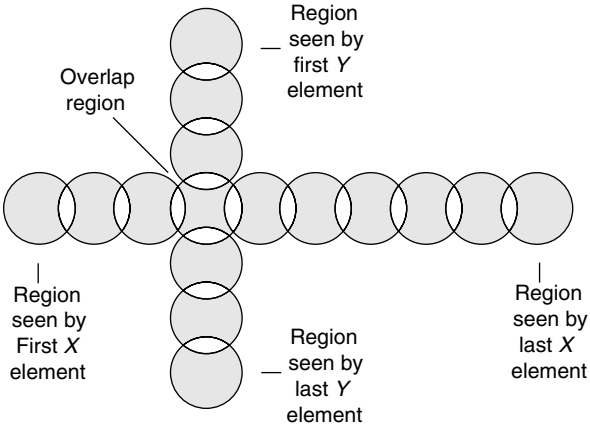

FIGURE 14.38 Overlap region.

ensonified by the Y pulse. If both elements receive echoes from the same region, then the correlation estimate returns a high value. To figure out own-ship velocity, therefore, the system must determine which $X-Y$ element pair produces the highest estimate of the correlation. Then, armed with precise knowledge of the array geometry, the system calculates the velocity which causes the ensonified region from these two elements to overlap.

Besides using the bottom echo, a correlation sonar can also estimate own-ship velocity by examining volume returns from small organic scatterers suspended throughout the water. The system typically operates in this volume echo mode when the bottom lies far away. This, of course, only works if the volume scattering characteristics do not significantly change during the time interval between pulses. When operating in this mode, the system treats depth as a discrete series of range bins, each as wide as one half of the pulse width. The system can use any one of these range bins to estimate own-ship velocity.

The above description does not account for the fact that the Y array receives volume echoes from both the first pulse and the second, as shown in Figure 14.39. With reference to the sketch drawn for time $t = t_i + T_d$, let the previous pulse echo be defined as the signal arising from scatterers in the volume ensonified by the first pulse. On the Y array, this previous pulse echo arrives at the same time as the second pulse, and thus interferes with the desired signal. Fortunately, the previous pulse volume usually lies far enough away that attenuation renders this interference negligible.

A problem arises, however, if the previous pulse ensonifies a boundary such as the ocean floor. Such boundary returns significantly exceed typical volume returns. In this case, the previous pulse boundary echo dominates the data received on the Y array. Since this boundary data comes from a completely different set of scatterers than those contributing to the X array data, it interferes

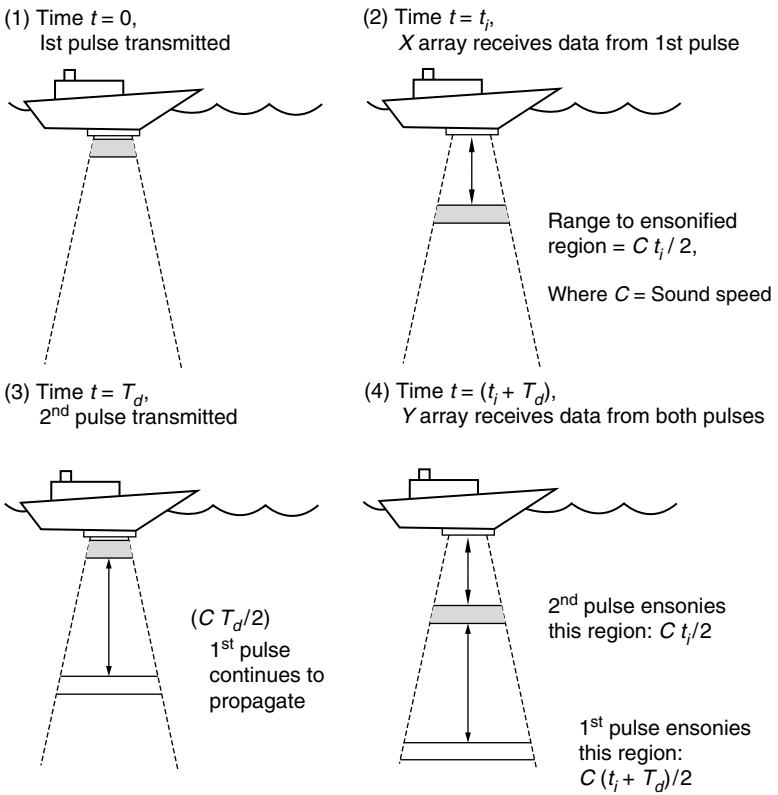

FIGURE 14.39 *Y* array previous pulse echo.

with the desired $X-Y$ correlation estimate. As the strength of the boundary interference increases, the correlation estimate becomes more and more corrupted. This means that the system cannot accurately determine own-ship velocity when a significant previous pulse echo contaminates the *Y* array data.

A possible solution to the problem of previous pulse echoes involves gating the data so as to exclude the interference. Figure 14.40 shows an example. In the top graph, the boundary lies for enough away such that its echo never intrudes on the *Y* array data. The second graph shows what can happen in shallower water. In this figure, a boundary echo from the first pulse intrudes on the data from the second. The bottom graph shows how decreasing the time interval between pulses removes the previous pulse boundary echo from the *Y* data.

This solution carries two undesirable features. First, for a given array configuration, longer time intervals are required in order to detect slower speeds. This means that decreasing the interpulse period increases the minimum detectable velocity of the correlation sonar. Furthermore, this solution also involves decreasing the size of the time window during which data is received,

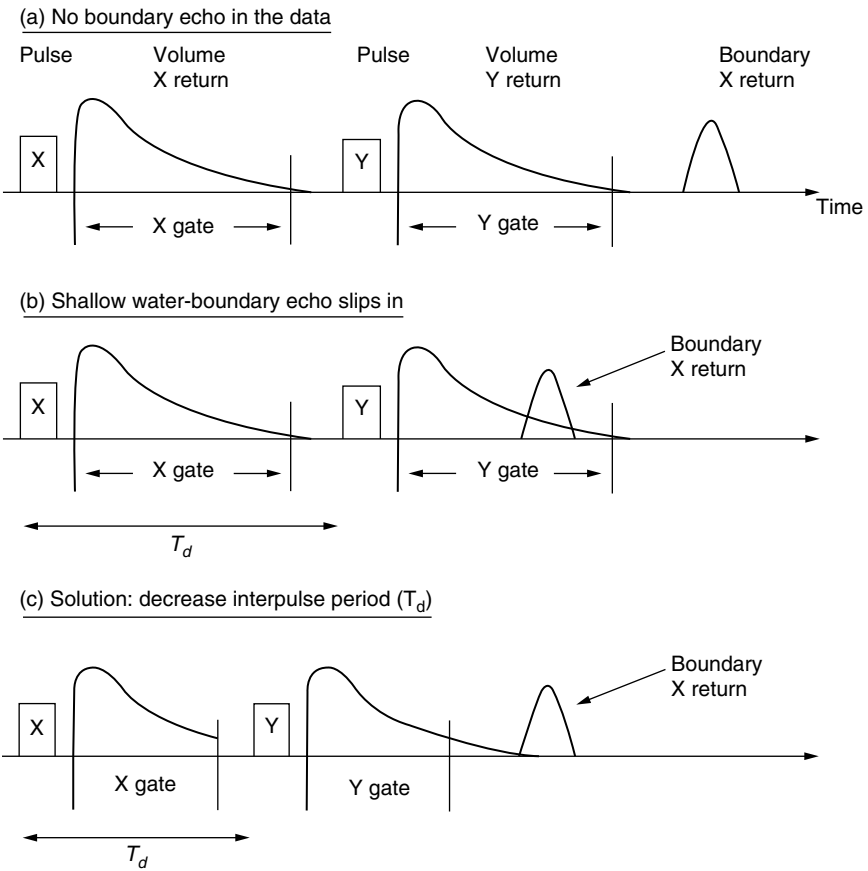

FIGURE 14.40 Boundary echoes in the data.

which limits the range extent over which the system can gather own-ship velocity estimates. To avoid these limitations, the system must have the ability to directly process data contaminated by a previous pulse boundary echo.

The following analysis begins by considering correlation sonar operation when the data is not corrupted by interference. First, an analytical model of volume reverberation signals is created, and applied to a correlation sonar system. A sub-optimal processor of uncorrupted data is then derived, based on the work of Van Trees.[21] The reverberation models are then applied to this sub-optimal processor, leading to an analytical expression for the output statistic.

14.7.2. Correlation Sonar Reverberation Model

This section derives explicit equations which model the desired correlation sonar volume reverberation signals. The analysis begins with a study of single pulse

monostatic volume reverberation, and then continues with an examination of bistatic reverberation. The analysis concludes by presenting a model for the two desired bistatic reverberation signals heard on a moving correlation sonar platform.

14.7.2.1. Monostatic and Bistatic Reverberation

To model monostatic reverberation, the following analysis borrows heavily from the work of Van Trees,[21] who derived a series of radar clutter models. The simplest model treats reverberation as a ring of slowly fluctuating point targets, all at the same range. "Slowly fluctuating" means that the scattering characteristics do not change during ensonification. An equation which models such reverberation is

$$\tilde{s}_r(\mathrm{t}) = \sqrt{E_\mathrm{t}} \sum_{i}^{N_i} \tilde{g}_i \tilde{f}(t - \lambda) \tag{14.342}$$

where:

$\tilde{s}_r(t) =$ Complex envelope of received signal (henceforth, the ($\sim$) superscript will denote a complex envelope)
$E_\mathrm{t} =$ Transmit energy,
$N_i =$ Total number of ensonified scatterers,
$\tilde{g}_i =$ Random scattering strength of scatterer i (this complex value also accounts for propagation effects and attenuation),
$\tilde{f}(\mathrm{t}) =$ Complex envelope of transmit waveform,
$\lambda =$ Two-way travel time to the scatterers.

Application of the central limit theorem reduces Equation 14.342 to

$$\tilde{s}_r(t) = \sqrt{E_\mathrm{t}} \tilde{b} \tilde{f}(t - \lambda) \tag{14.343}$$

where $\tilde{b}$ is a complex Gaussian random variable which models scattering strength, propagation loss, and attenuation.

A more sophisticated reverberation model allows the scattering characteristics to change with time during ensonification. Such reverberation is referred to as a Doppler spread target in Ref. 21. An equation which models this reverberation is

$$\tilde{s}_r(t) = \sqrt{E_\mathrm{t}} \tilde{b}(t - \lambda/2) \tilde{f}(t - \lambda) \tag{14.344}$$

In this model, $\tilde{b}(t)$ is a random process with an independent variable that corresponds to the ensonification time of the scatterers.

A third clutter model in Ref. 21 accounts for the fact that the ensonified volume may have a greater range extent than that of a point target. This range spread model treats reverberation as a collection of discrete point target volumes,

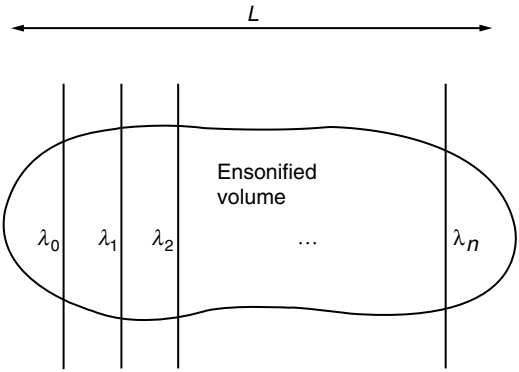

FIGURE 14.41 Range spread target.

as shown in Figure 14.41. Each volume (j) has its own two-way travel time, or range, λ_j.

The return from the jth discrete volume is modeled by

$$\tilde{s}_j(t) = \sqrt{E_t}\tilde{b}(\lambda_j)\tilde{f}(t - \lambda_j) \tag{14.345}$$

where $\tilde{b}(\lambda)$ is a random process with an independent variable that corresponds to range. For a given range λ_j, note that $\tilde{b}$ is a constant with respect to the time variable t. The echoes from these discrete volumes sum up to form the echo from the range spread target. In the limit, as the range extent of each discrete volume decreases, this sum approaches an integral. The final range spread target model thus becomes

$$\tilde{s}_r(t) = \sqrt{E_t}\int_L \tilde{b}(\lambda)\tilde{f}(t - \lambda)\mathrm{d}\lambda \tag{14.346}$$

where L indicates the range extent of the total volume, as shown in Figure 14.41.

A combination of Equation 14.344 and Equation 14.346 yields the model for a doubly spread target; that is, a range spread target whose scattering characteristics may change with time during ensonification. The doubly spread reverberation model is

$$\tilde{s}_r(t) = \sqrt{E_t}\int_L \tilde{b}(\lambda, t - \lambda/2)\tilde{f}(t - \lambda)\mathrm{d}\lambda \tag{14.347}$$

In this model, the random process $\tilde{b}(\lambda, t)$ has two independent variables. The first corresponds to range, or two-way travel time, and the second corresponds to ensonification time.

The limits of integration in Equation 14.347 depend on the transmit pulse length. Consider a monostatic omni-directional sonar transmitting a pulse of length T_p. At any time t, the sonar receives echoes from a locus of scatterers which lie within a spherical shell of width $(CT_p/2)$, as shown in Figure 14.42.

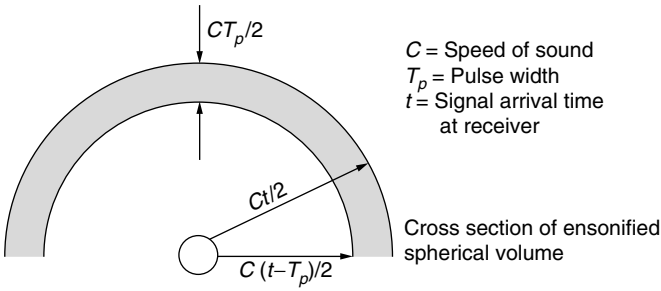

FIGURE 14.42 Monostatic reverberation volume.

Based on this figure, the model becomes

$$\tilde{s}_r(t) = \sqrt{E_t}\int_{t=T_p}^{t} \tilde{b}(\lambda, t-\lambda/2)\tilde{f}(t-\lambda)\mathrm{d}\lambda \tag{14.348}$$

Equation 14.348 serves as a doubly spread monostatic volume reverberation model. As will be shown, this model can be applied to a bistatic system, provided that the range to any scatterer greatly exceeds the projector–receiver separation. This is true even though the set of all bistatic point targets with the same two-way travel time forms an ellipsoidal locus, as shown in Figure 14.43, as opposed to a spherical locus, as shown in Figure 14.42.

The relationship between bistatic two-way travel time and range is a nonlinear function which varies with the target bearing. Consequently, Equation 14.348 does not model general bistatic reverberation, since it employs a linear monostatic relationship between range and travel-time. If the reverberation lies in the far-field, however, as shown in Figure 14.44, the distance to any ensonified scatterer greatly exceeds the projector–receiver separation, and the ellipsoid begins to approximate a sphere. At these ranges, the bistatic sonar can be modeled by an equivalent monostatic system. This equivalent sonar is located in such a fashion as to keep the ensonified volume symmetric about the original bistatic system. From Figure 14.43, this location corresponds to the point exactly half-way along a line drawn between the projector and receiver, henceforth referred to as the bistatic midpoint.

In summary, far field bistatic volume reverberation can be modeled under appropriate assumptions, as derived in this section. The next section applies this model to the correlation sonar.

14.7.2.2. Reverberation as Heard on a Moving Correlation Sonar Platform

A correlation sonar compares two far-field reverberation returns as heard on a moving platform. If the two pulses are examined independently, Equation 14.348 serves as an adequate model, but this equation does not account for correlation

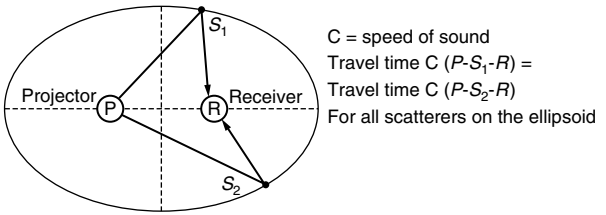

FIGURE 14.43 Bistatic constant travel-time ellipse.

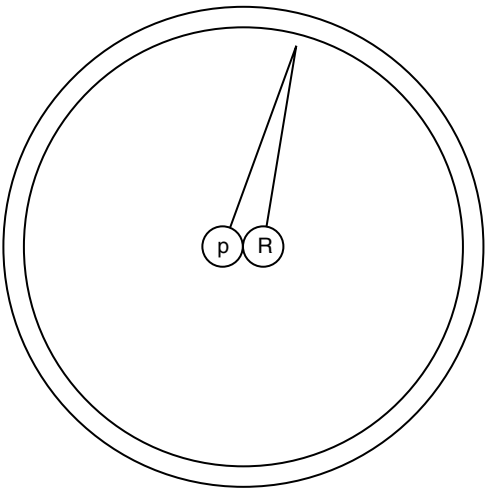

FIGURE 14.44 Far-field bistatic ellipsoid.

between these returns. As stated earlier, the overlap between the ensonified volumes causes the two pulse returns to be correlated. Figure 14.45 demonstrates how the amount of overlap, and hence the correlation, depends on the distance moved by the sonar. As seen, the moving sonar samples the reverberation process at different spatial locations with each pulse. In order to model this spatial sampling, a new dimension must be included in the far-field bistatic volume reverberation equation.

Define a coordinate system where the (x, y) plane lies parallel to the ocean surface, and the z dimension corresponds to depth. Assume that the platform depth does not significantly change during the interpulse period. Also assume that the platform moves with a constant speed V in the x direction. Under this convention, the spatial sampling model need only keep track of the sonar's x coordinate. Inserting this dimension into Equation 14.348 yields

$$\tilde{s}_r(x,t) = \sqrt{E_t}\int_{t-T_p}^{t} \tilde{b}(x,\lambda,t-\lambda/2)\tilde{f}(t-\lambda)\mathrm{d}\lambda \qquad (14.349)$$

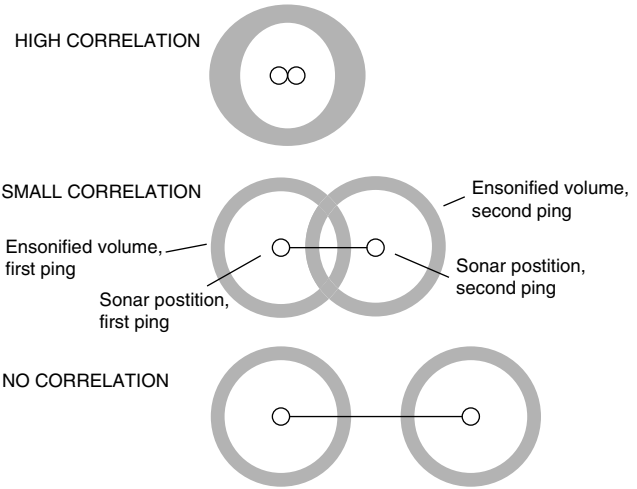

FIGURE 14.45 Overlapping volumes and correlation.

The random process $\tilde{b}(x, \lambda, t)$, which models scattering strength, propagation loss, and attenuation, now has three independent variables. The first variable corresponds to the location of the bistatic midpoint at transmission. The last two variables correspond to range and ensonification time, as before.

In order to help apply Equation 14.349 to the correlation sonar, consider the system geometry of the snap-shot shown in Figure 14.46. This system consists of a single Y element and N_x co-linear X array elements. The platform moves along the x axis with speed V. At the time of first pulse transmission, the coordinate set $\{x_1, x_2, \ldots, x_{N_x}\}$ defines the locations of the X array midpoints. Similarly, the bistatic midpoint of the Y element lies at coordinate (x_y). At the time of second pulse transmission, this midpoint lies at $(x_y + VT_d)$, where T_d equals the interpulse period.

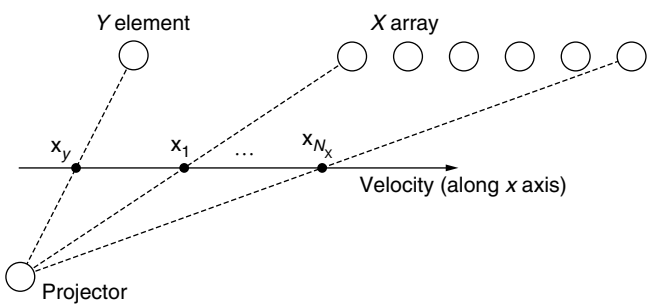

FIGURE 14.46 Correlation sonar geometry.

Based on Figure 14.46, Equation 14.349 gives the first pulse reverberation signal heard on the ith element of the X array as

$$\tilde{s}_{xi}(x_i, \mu) = \sqrt{E_t} \int_{\mu - T_p}^{\mu} \tilde{b}(x_i, \lambda, \mu - \lambda/2)\tilde{f}(\mu - \lambda)\mathrm{d}\lambda \tag{14.350}$$

where μ is a variable which measures receiver arrival time relative to transmission of the first pulse. Similarly, Equation 14.349 gives the second pulse Y element signal as

$$\begin{aligned}&\tilde{s}_y(x_y + VT_d, \nu + T_d)\\ &= \sqrt{E_t} \int_{\nu - T_p}^{\nu} \tilde{b}(x_y + VT_d, \lambda, \nu + T_d - \lambda/2)\tilde{f}(\nu - \lambda)\mathrm{d}\lambda\end{aligned} \tag{14.351}$$

This equation differs from Equation 14.350 in two key respects. First $(x_y + VT_d)$ corresponds to the coordinate location of the Y element bistatic midpoint at the time of the second pulse transmission. Second ν measures time relative to transmission of the second pulse. This is why the absolute time index shifts by T_d seconds, as seen in the second argument of $\tilde{s}_y$. Also, note that the ensonification time of the scatterers shifts with T_d, as reflected in the final argument of $\tilde{b}$. This argument corresponds to the explicit physical time that the waveform hits the scatterer.

To summarize the analysis thus far, Equation 14.350 and Equation 14.351 model the correlation sonar bistatic reverberation signals. Each signal arises from a different scattering volume. The correlation sonar attempts to derive platform velocity by determining which of the first pulse ensonified volumes completely overlaps with the second pulse volume. The following analysis proves that these volumes completely overlap if and only if their bistatic midpoints are colocated. The system thus needs only to determine the platform velocity which causes these midpoints to overlap.

The analysis begins by deriving a general expression for the round-trip distance traveled by the pulse in a bistatic system. This round-trip distance, $C\lambda$, equals the product of the two-way travel time to the scatterer, λ, and the speed of sound, C. Figure 14.47 shows the geometry of a general bistatic system, where bold-face indicates a vector. Based on this figure, the round-trip distance is

$$C\lambda = |\mathbf{D}_{ps}| + |\mathbf{D}_{rs}| \tag{14.352}$$

or

$$C\lambda = |\mathbf{D}_{ps}| + |\mathbf{D}_{ps} - \mathbf{D}_{pr}| \tag{14.353}$$

Expanding the second magnitude in Equation 14.353 produces

$$C\lambda = |\mathbf{D}_{ps}| + \{|\mathbf{D}_{ps}|^2 + |\mathbf{D}_{pr}|^2 - 2 \times \mathbf{D}_{pr}\mathbf{D}_{ps}\}^{1/2} \tag{14.354}$$

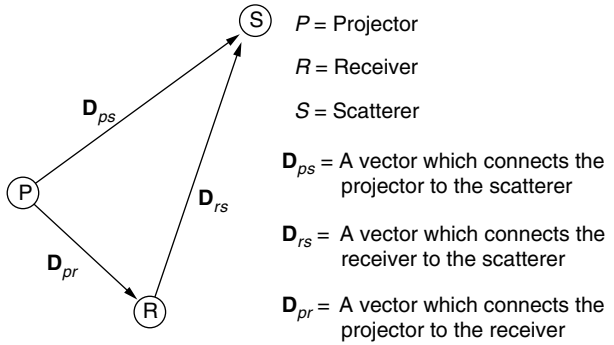

FIGURE 14.47 General bistatic geometry.

where $(\mathbf{D}_{pr}\cdot\mathbf{D}_{ps})$indicates a dot product between the vectors. Moving $|\mathbf{D}_{ps}|$ outside of the square root yields

$$C\lambda = |\mathbf{D}_{ps}| + |\mathbf{D}_{ps}| \times \left\{1 - 2 \times \frac{\mathbf{D}_{pr}\mathbf{D}_{ps}}{|\mathbf{D}_{ps}|^2} + \frac{|\mathbf{D}_{pr}|^2}{|\mathbf{D}_{ps}|^2}\right\}^{1/2} \tag{14.355}$$

Assume that the distance to the scatterer greatly exceeds the projector–receiver separation such that

$$|\mathbf{D}_{ps}| \gg |\mathbf{D}_{pr}| \tag{14.356}$$

This assumption places the scatterer in the far field of the bistatic system. Equation 14.356 renders the squared term in Equation 14.355 negligible, thus producing

$$C\lambda \approx |\mathbf{D}_{ps}| + |\mathbf{D}_{ps}| \times \left\{1 - 2 \times \frac{\mathbf{D}_{pr}\mathbf{D}_{ps}}{|\mathbf{D}_{ps}|^2}\right\}^{1/2} \tag{14.357}$$

Define a unit vector pointing in the direction of $\mathbf{D}_{ps}$ as

$$\mathbf{n}_{ps} = \frac{\mathbf{D}_{ps}}{|\mathbf{D}_{ps}|} \tag{14.358}$$

Substituting Equation 14.358 into Equation 14.357 gives

$$C\lambda \approx |\mathbf{D}_{ps}| + |\mathbf{D}_{ps}| \times \left\{1 - 2 \times \frac{\mathbf{D}_{pr}}{|\mathbf{D}_{ps}|}\mathbf{n}_{ps}\right\}^{1/2} \tag{14.359}$$

Recall the following quadratic approximation for small Δx:

$$\{1 - \Delta x\}^{1/2} \approx 1 - \Delta x/2 \tag{14.360}$$

The far-field assumption described by Equation 14.356 allows for use of this approximation. Applying Equation 14.360 to Equation 14.359 yields

$$C\lambda \approx |\mathbf{D}_{ps}| + |\mathbf{D}_{ps}| \times \left\{ 1 - \frac{\mathbf{D}_{pr}}{|\mathbf{D}_{ps}|} \cdot \mathbf{n}_{ps} \right\} \tag{14.361}$$

or

$$C\lambda = 2|\mathbf{D}_{ps}| - \mathbf{D}_{pr} \cdot \mathbf{n}_{ps} \tag{14.362}$$

From Equation 14.362, the set of all vectors $\mathbf{D}_{ps}$ which yield a constant time ellipsoid for a given $\mathbf{D}_{pr}$ must satisfy the equation

$$(2|\mathbf{D}_{ps}| - \mathbf{D}_{pr} \cdot \mathbf{n}_{ps})/C = \lambda \tag{14.363}$$

Now consider a second bistatic system, henceforth denoted by a prime superscript, situated relatively close to the one examined above. Figure 14.48 shows the geometry. The vectors $\mathbf{d}_p$ and $\mathbf{d}_r$ indicate the location of the second system (P' and R') as measured relative to the first (P and R). The round-trip distance to any given scatterer for this second system is

$$C\lambda' = |\mathbf{D}'_{ps}| + |\mathbf{D}'_{rs}| \tag{14.364}$$

or

$$C\lambda' = |\mathbf{D}_{ps} - \mathbf{D}_p| + \mathbf{D}_{ps} - (\mathbf{D}_{pr} + \mathbf{d}_r)| \tag{14.365}$$

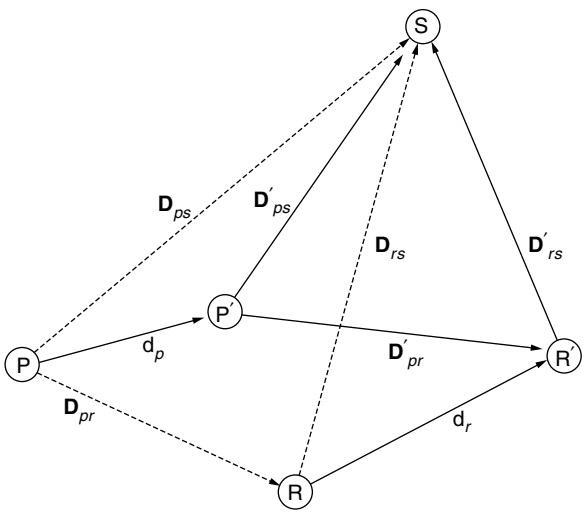

FIGURE 14.48 A second bistatic system.

Assume that the distance to the scatterer greatly exceeds the distance between any of the bistatic elements shown in Figure 14.48. This assumption allows for creation of the following two conditions:

$$|\mathbf{D}_{ps}| \gg |\mathbf{d}_p| \tag{14.366}$$

and

$$|\mathbf{D}_{ps}| \gg |\mathbf{D}_{pr} + \mathbf{d}_r| \tag{14.367}$$

A comparison of Equation (14.353) and Equation (14.361) shows that $|\mathbf{D}_{ps} - \mathbf{D}_{pr}|$ can be approximated as

$$|\mathbf{D}_{ps} - \mathbf{D}_{pr}| \approx |\mathbf{D}_{ps}| \times \left\{1 - \frac{\mathbf{D}_{pr}}{|\mathbf{D}_{ps}|} \cdot \mathbf{n}_{ps}\right\} = |\mathbf{D}_{ps}| - \mathbf{D}_{pr} \cdot \mathbf{n}_{ps} \tag{14.368}$$

when Equation 14.356 holds. In a similar fashion, the two magnitudes in Equation 14.365 can be approximated, due to the assumptions presented in Equation (14.366) and Equation (14.367). Applying these approximations yields

$$C\lambda' \approx (|\mathbf{D}_{ps}| - \mathbf{d}_p \cdot \mathbf{n}_{ps}) + (|\mathbf{D}_{ps}| - \mathbf{D}_{pr} \cdot \mathbf{n}_{ps} - \mathbf{d}_r \cdot \mathbf{n}_{ps}) \tag{14.369}$$

From Equation 14.369, the set of all vectors $\mathbf{D}_{ps}$ which yield a constant time ellipsoid for a given $\mathbf{D}_{pr}$, $\mathbf{d}_p$, and $\mathbf{d}_r$ must satisfy the equation

$$(2|\mathbf{D}_{ps}| - \mathbf{D}_{pr} \cdot \mathbf{n}_{ps})/C - (\mathbf{d}_p + \mathbf{d}_r) \cdot \mathbf{n}_{ps}/C = \lambda' \tag{14.370}$$

At this point in the analysis, Equation 14.363 defines the first system's constant-time ellipsoid, while Equation 14.370 defines that of the second bistatic system. Under the constraint that λ equals λ', a single set of vectors $\mathbf{D}_{ps}$ solves both equations if and only if the following condition holds true:

$$\mathbf{d}_p = -\mathbf{d}_r \tag{14.371}$$

Figure 14.49 shows the geometry associated with this result. The figure shows both bistatic systems, with a line connecting P to R, and another connecting P' to R'.

Because of Equation 14.371, the vectors $\mathbf{d}_p$ and $\mathbf{d}_r$ are parallel, and have the same magnitude. It follows that the interior angles measured from these vectors are all equal, or

$$\theta_1 = \theta_2 \tag{14.372}$$

and

$$\phi_1 = \phi_2 \tag{14.373}$$

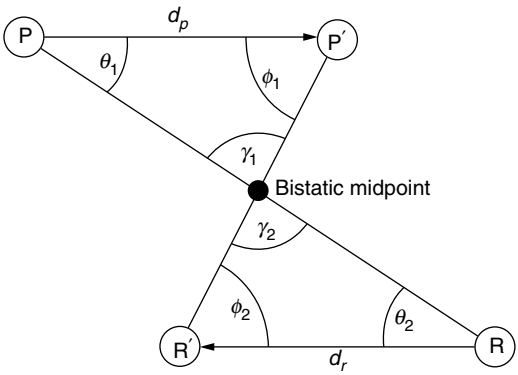

FIGURE 14.49 Bistatic midpoint.

In addition, because the interior angles of a triangle sum to 180°, the remaining two angles are equal, or

$$\gamma_1 = \gamma_2 \tag{14.374}$$

Because of Equation (14.372) to Equation (14.374), the two triangles shown in Figure 14.49 are similar, with sides having the same length opposite γ_1 and γ_2. Hence, the sides opposite θ_1 and θ_2 must have the same length, as must the sides opposite ϕ_1 and ϕ_2. Consequently, the point of intersection between the two lines is the bistatic midpoint for both systems. In other words, the far-field constant-time ellipsoids of two bistatic sonars completely overlap if and only if the systems have the same midpoint.

In summary, a correlation sonar determines platform velocity by first determining which **X** array ensonified volume completely overlaps with the *Y* element ensonified volume. It then calculates the speed required to cause the *Y* element bistatic midpoint to overlap with that **X** array midpoint. A technique still remains to be developed, however, that determines which of the ensonified volumes overlap. The following section derives a sub-optimal receiver which is used to accomplish this task.

14.7.3. A Sub-Optimal Correlation Sonar Receiver

The work of Van Trees[21] provides the background required to derive a sub-optimum correlation sonar processor. The analysis begins with a statement of the detection hypotheses.

Consider Figure 14.46, which shows a system that consists of N_x co-linear *X* elements and a single *Y* element. The platform moves with speed *V* along the *x* axis. As stated earlier, the system transmits two pulses which are separated by an interpulse period of length T_d. The **X** array elements receive echoes from the first pulse, and T_d seconds later the *Y* element receives echoes from the second pulse.

Let $\tilde{w}(t)$ represent a white Gaussian signal which models the ambient ocean noise received on these elements. With reference to Equation 14.350, let $\tilde{s}_{xi}(x_i, t)$ represent the volume echo received on the ith element of the X array, where x_i corresponds to the location of this element's bistatic midpoint at the time of first pulse transmission. Furthermore, with reference to Equation 14.351, let $\tilde{s}_y(x_y + VT_\mathrm{d}, t + T_\mathrm{d})$ represents the volume echo received T_d seconds later on the single Y element, where $(x_y + VT_\mathrm{d})$ corresponds to the location of the element's bistatic midpoint at the time of second pulse transmission. The time variable, t, ranges from T_i to T_f, two time instants which define the edges of the range bin under examination. Under these definitions, the total signal received on the ith element of the X array becomes

$$\tilde{x}_i(t) = \tilde{s}_{xi}(x_i, t) + \tilde{w}(t) \qquad T_i < t < T_f \tag{14.375}$$

and that received T_d seconds later on the single Y element becomes

$$\tilde{y}(t + T_\mathrm{d}) = \tilde{s}_y(x_y + VT_\mathrm{d}, t + T_\mathrm{d}) + \tilde{w}(t + T_\mathrm{d}) \qquad T_i < t < T_f \tag{14.376}$$

For practical purposes, assume during the interpulse period that the bistatic midpoint of the Y element moves into a location previously occupied by one of the X element midpoints. Given the identity of this X element, the system can determine the velocity which causes these midpoints to overlap. Based on this, the processor must decide between the following N_x possible hypotheses;

$$\begin{aligned} H_1&: \quad (x_y + VT_\mathrm{d}) = x_1 \\ H_2&: \quad (x_y + VT_\mathrm{d}) = x_2 \\ &\vdots \\ H_{N_x}&: \quad (x_y + VT_\mathrm{d}) = x_{N_x} \end{aligned} \tag{14.377}$$

Combining these hypotheses with Equation 14.376 produces an equivalent set of hypotheses based on the received waveform. These hypotheses are

$$\begin{aligned} H_1&: \quad \tilde{y}(t + T_\mathrm{d}) = \tilde{s}_y(x_1, t + T_\mathrm{d}) + \tilde{w}(t + T_\mathrm{d}), \\ H_2&: \quad \tilde{y}(t + T_\mathrm{d}) = \tilde{s}_y(x_2, t + T_\mathrm{d}) + \tilde{w}(t + T_\mathrm{d}), \\ &\vdots \\ H_{N_x}&: \quad \tilde{y}(t + T_\mathrm{d}) = \tilde{s}_y(x_{N_x}, t + T_\mathrm{d}) + \tilde{w}(t + T_\mathrm{d}), \quad T_i < t < T_f \end{aligned} \tag{14.378}$$

If T_d is short enough such that the scattering characteristics do not significantly change between the two pulses, then the volume echo waveforms are essentially invariant during the interpulse period. Under this assumption, the waveform $\tilde{s}_{xi}(x_i, t)$ approximately equals $\tilde{s}_y(x_i, t + T_\mathrm{d})$, for any given element. Consequently, the volume echo waveforms can be modeled as deterministic signals. Applying this assumption to the hypotheses listed in Equation 14.378

gives

$$\begin{aligned} H_1: &\quad \tilde{y}(t+T_\mathrm{d}) \approx \tilde{s}_{x1}(x_1,t) + \tilde{w}(t+T_\mathrm{d}), \\ H_2: &\quad \tilde{y}(t+T_\mathrm{d}) \approx \tilde{s}_{x2}(x_2,t) + \tilde{w}(t+T_\mathrm{d}), \\ &\quad\quad\vdots \\ H_{N_\mathrm{x}}: &\quad \tilde{y}(t+T_\mathrm{d}) \approx \tilde{s}_{xN_\mathrm{x}}(x_{N_\mathrm{x}},t) + \tilde{w}(t+T_\mathrm{d}), \quad T_i < t < T_f \end{aligned} \tag{14.379}$$

In this way, the correlation sonar problem is formulated as an M-ary deterministic signal detector.

The derivation of the optimal M-ary detector depends on the criterion being optimized. A criterion which lends itself to this purpose is probability of error, P_e, or equivalently, the probability of making a correct choice, P_c. Any receiver which minimizes P_e also maximizes P_c.

As outlined in Ref. 21, the procedure to optimize either criterion starts by approximating the received signal $\tilde{y}(t+T_\mathrm{d})$ by an equivalent data vector of finite length K. Define this vector as

$$\tilde{\mathbf{y}} = \{\tilde{y}_1, \tilde{y}_2, \ldots \tilde{y}_k\} \tag{14.380}$$

The ith element of this vector is given by

$$\tilde{y}_i = \int_{T_i}^{T_f} \tilde{y}(t+T_\mathrm{d})\phi_i(t)\mathrm{d}t \tag{14.381}$$

where $\phi_i(t)$ is the ith member in a complete ortho-normal function set.

Consider the set of all possible data vectors $\tilde{\mathbf{y}}$, defined as the decision space. This decision space is partitioned into a set of nonoverlapping regions Z_i, each corresponding to a different hypothesis, as shown in Figure 14.50. If the received vector $\tilde{\mathbf{y}}$ falls within region Z_i, the M-ary detector chooses hypothesis H_i. For a receiver operating in this fashion, the probability of making a correct choice is

$$P_\mathrm{c} = \sum_{i=1}^{M} P_i \cdot \int_Z f_{\tilde{\mathbf{Y}}|H_i}(\tilde{\mathbf{y}}|H_i)\mathrm{d}\tilde{\mathbf{y}} \tag{14.382}$$

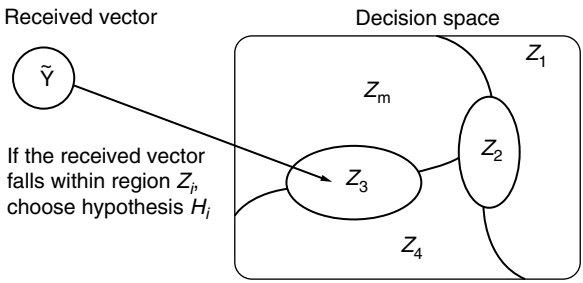

FIGURE 14.50 M-ary decision space.

where $f_{\tilde{\mathbf{Y}}|H_i}(\tilde{\mathbf{y}}|H_i)$ is the joint conditional density function of the received vector $\tilde{\mathbf{y}}$, given that hypothesis H_i is true.

Assume that all hypotheses are equally likely. Based on this assumption, Equation 14.382 reduces to

$$P_c = \frac{1}{M}\sum_{i=1}^{M}\int_{Z_i} f_{\tilde{\mathbf{Y}}|H_i}(\tilde{\mathbf{y}}|H_i)\mathrm{d}\tilde{\mathbf{y}} \tag{14.383}$$

As stated earlier, the optimum receiver is derived by maximizing this criterion. This is accomplished by defining the regional boundaries Z_i such that the M integrals in Equation 14.383 are maximized. To maximize a given integral, place into the region Z_i all of the data vectors $\tilde{\mathbf{y}}$ where $f_{\tilde{\mathbf{Y}}|H_i}(\tilde{\mathbf{y}}|H_i)$ exceeds every other $f_{\tilde{\mathbf{Y}}|H_j}(\tilde{\mathbf{y}}|H_j)$ (for $j \neq i$). Not only does defining the regions in this way maximize the probability of making a correct choice and minimize the probability of error, but it also results in a simple receiver structure. For a given received vector $\tilde{\mathbf{y}}$, the optimum M-ary detector calculates a set of statistics $\{f_{\tilde{\mathbf{Y}}|H_1}(\tilde{\mathbf{y}}|H_1),\ldots,f_{\tilde{\mathbf{Y}}|H_m}(\tilde{\mathbf{y}}|H_m)\}$, and then chooses the hypothesis which corresponds to the largest.

Define $\tilde{\mathbf{w}}$ as the vector arising from the k-term ortho-normal expansion of the white-noise shown in Equation 14.376. This data vector has a joint density function given by $f_{\tilde{\mathbf{w}}}(\tilde{\mathbf{w}})$. Define the likelihood ratio on hypothesis H_i as

$$\Lambda_i(\tilde{\mathbf{y}}) = f_{\tilde{\mathbf{Y}}|H_i}(\tilde{\mathbf{y}}|H_i)/f_{\tilde{\mathbf{w}}}(\tilde{\mathbf{y}}) \tag{14.384}$$

Since the denominator is the same for every hypothesis, the maximum $\Lambda_i(\tilde{\mathbf{y}})$ corresponds to the maximum $f_{\tilde{\mathbf{Y}}|H_i}(\tilde{\mathbf{y}}|H_i)$. As such, the optimal M-ary detector can make its decision based on choosing the largest likelihood ratio, $\Lambda_i(\tilde{\mathbf{y}})$.

The above "largest-of" detector is derived for the data vector $\tilde{\mathbf{y}}$, not the actual continuous waveform $\tilde{y}(t+T_d)$. Based on the ortho-normal function set, the original continuous received signal is

$$\tilde{y}(t+T_d) = \underset{K\to\infty}{\text{l.i.m}}\sum_{i=1}^{K}\tilde{y}_i\phi_i(t) \tag{14.385}$$

where l.i.m stands for "limit in the mean-squared sense". The optimum continuous likelihood ratio, therefore, corresponds to that shown in Equation 14.384 in the limit as K goes to infinity. This limit is calculated in Ref. 21 for a deterministic signal corrupted by white Gaussian noise. The resulting likelihood ratio reduces to the output of a matched filter correlator followed by an additive bias term. If all of the M signals have the same energy, the bias is the same for every likelihood ratio, and may be removed. The optimum M-ary detector, therefore, consists of a matched filter bank, as shown in Figure 14.51. The receiver picks the hypothesis which corresponds to the channel with the largest matched filter output.

A problem still exists, however. Although the signals $\tilde{s}_{x1}(x_1,t)$ through $\tilde{s}_{xnx}(x_{nx},t)$ are modeled as being deterministic, they are still unknown.

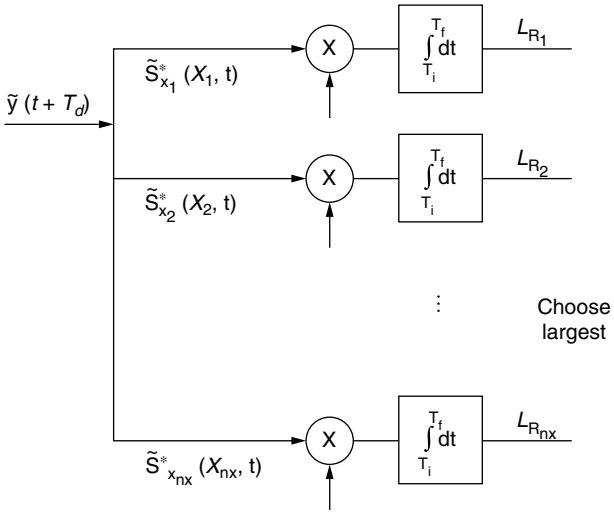

FIGURE 14.51 Optimum *M*-ary detector in Gaussian noise.

Fortunately, the received waveform $\tilde{x}_1(t)$ contains the signal $\tilde{s}_{x1}(x_1, t), \tilde{x}_2(t)$ contains $\tilde{s}_{x2}(x_2, t)$, and so on. Under the assumption that the additive noise is small relative to the reverberation, a possible sub-optimal receiver would correlate $\tilde{y}(t + T_d)$ with the received and stored waveforms $\tilde{x}_1(t)$ through $\tilde{x}_{nx}(t)$, in lieu of the desired unknown signals. Figure 14.52 shows a block diagram of this sub-optimal receiver. As the signal-to-noise ratios (SNR) in the signals $\tilde{x}_1(t)$ through $\tilde{x}_{nx}(t)$ increase, the sub-optimal processor approaches the optimum receiver in performance.

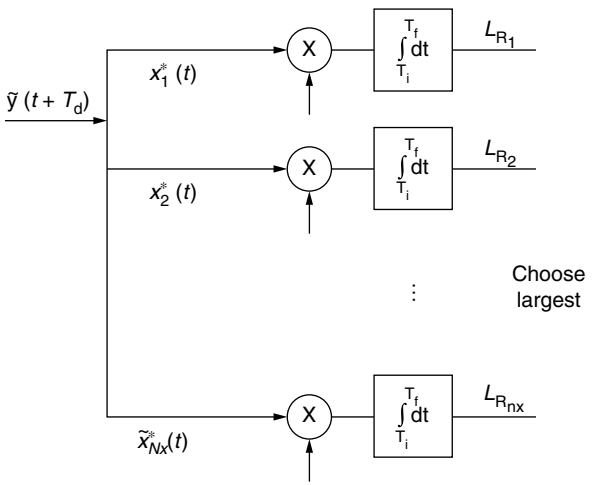

FIGURE 14.52 Sub-optimum *M*-ary detector in Gaussian noise.

To summarize the results thus far, the processor shown in Figure 14.52 approximates the optimum receiver, based on assumptions of equal energy signals and equally likely hypotheses. In addition, the ratio between the reverberation energy and the corrupting white Gaussian noise energy must greatly exceed unity.

This concludes the analysis of correlation sonar operation on data uncorrupted by previous pulse interference. The next section examines the effects of previous pulse interference interference on correlation sonar performance.

14.7.4. Performance in Previous Pulse Interference

The previous section derived an optimal correlation sonar designed to operate in the presence of additive Gaussian noise. With reference to Figure 14.46, the receiver determines which X array midpoint most closely overlaps with the Y element midpoint by first calculating a set of likelihood ratios. Each likelihood ratio corresponds to a different X element. The correlation sonar then makes a decision by choosing whichever X element has the largest likelihood ratio.

When given Gaussian noise, the likelihood ratio values equal the output of a matched filter bank, as shown in Figure 14.51. This optimal processor correlates the total signal received on the Y element with the noise free reverberation received on each X element. Since these reverberations are unknown, a suggested sub-optimal receiver instead correlates with the total signal received on each X element, as shown in Figure 14.52. This receiver is sub-optimal because additive noise corrupts the reverberations received on the X array. As the additive noise becomes less significant, this sub-optimal processor approaches the optimum receiver in performance.

Now consider what happens if interference corrupts the correlation sonar data. As explained in Section 14.7.1, one source of such interference arises when an echo from the first pulse intrudes on the second pulse data collected by the Y element. With reference to Figure 14.39, the previous pulse ensonifies a completely different range bin than the one under examination. Consequently, the previous pulse echo acts as a source of interference in the total signal received on the Y element. To account for this interference, the hypotheses given in Equation 14.379 should be rewritten as

$$H_i: \quad \tilde{y}(t + T_\mathrm{d}) = \tilde{s}_{x_i}(x_1, t + T_\mathrm{d}) + \tilde{d}(t) \qquad T_i < t < T_f \tag{14.386}$$

where $\tilde{d}(t)$ represents the combined effects of the previous pulse echo and the additive Gaussian noise. Note that the signal received on an X element is still given by

$$\tilde{x}_i(t) = \tilde{s}_{x_i}(x_i, t) + \tilde{w}(t) \qquad T_i < t < T_f \tag{14.387}$$

since no previous pulse exists before the first pulse.

The nature of the signal $\tilde{d}(t)$ governs the design of the optimum receiver. Typically, if the previous pulse ensonifies volume scatterers, the resulting echo is

assumed to be modeled by a Gaussian signal. An examination of actual volume echoes serves to bear out the validity of this assumption, as shown in Figure 14.32. If $\tilde{d}(t)$ is thus modeled as a Gaussian signal, the processor shown in Figure 14.51 still serves as the optimum receiver. In the presence of Gaussian noise and Gaussian interference, no other receiver has a lower probability of error.

If the previous pulse ensonifies an ocean boundary, however, a Gaussian signal may not serve as an adequate model for the resulting echo, as shown in Figure 14.32. Consequently, the optimum processor no longer consists of the traditional bank of matched filters. In addition, the previous pulse boundary echo power often greatly exceeds that of the desired volume echo. As such, this nonGaussian interference typically masks the desired signal, and significantly degrades the resulting correlation sonar performance.

To summarize, this section reveals that a correlation sonar can be treated as an M-ary detector. This section also demonstrates how the correlation sonar is sensitive to nonGaussian interference. The next section responds to these results by deriving and testing the optimum nonGaussian M-ary detector for the case in which the disturbance is much larger than the additive noise and can be modeled as a generalized SIRV.

14.8. *M*-ARY DETECTION

This section derives and evaluates several detectors for identifying one of M possible signals in a data set corrupted by nonGaussian interference. The problem is formulated as that of M-ary detection, where the mth hypothesis is

$$H_m: \mathbf{y} = \theta \mathbf{x}_m + \mathbf{d} \tag{14.388}$$

In this equation, $\mathbf{y}$ is a random vector which contains digital samples of the received data, $\mathbf{x}_m$ is a unit amplitude deterministic vector which contains samples of the mth signal, θ is a nonrandom scalar which accounts for the actual received amplitude of the mth signal, and $\mathbf{d}$ is a random vector which contains samples of the interference. If the interference is modeled as a traditional SIRV, its density function is given by

$$f_{\mathbf{D}}(\mathbf{d}) = K h_N(\alpha) \tag{14.389}$$

where K is a scalar, N is the number of samples in the vector $\mathbf{d}$, $h_N(\alpha)$ is the characteristic nonlinear function, and α is the inner product

$$\alpha = \mathbf{d}^{\mathrm{T}}\mathbf{d} \tag{14.390}$$

where, without loss of generality, the covariance matrix of the interference is assumed to be the identity matrix.

The analysis begins with an examination of the optimum detector, where the interference is modeled as a traditional SIRV. Subsequent subsections derive and evaluate detectors for the case where the interference is modeled with generalized SIRVs.

14.8.1. Optimum *M*-ary Detection

Section 14.7 reveals that the optimum M-ary detector is a "largest-of" receiver with statistic

$$\Lambda_m(\mathbf{y}) = f_{\mathbf{Y}|H_m}(\mathbf{y}|H_m)/f_{\mathbf{D}}(\mathbf{y}) \tag{14.391}$$

This statistic is computed for each m from one to M, and the receiver makes a decision by picking that hypothesis which returns the largest value. Since the denominator of Equation 14.391 is the same for all hypothesis, it may be removed without affecting the final outcome. The "largest-of" receiver thus need only compute

$$\Lambda_m(\mathbf{y}) = f_{\mathbf{Y}|H_m}(\mathbf{y}|H_m) \tag{14.392}$$

This statistic is derived based on an assumption of equi-probable hypotheses. From, Ref. 21, a more general statistic is

$$\Lambda_m(\mathbf{y}) = P_m f_{\mathbf{Y}|H_m}(\mathbf{y}|H_m) \tag{14.393}$$

where P_m is the probability of hypothesis m.

With reference to Equation 14.388, the interference under hypothesis m is given by

$$\mathbf{d} = \mathbf{y} - \theta\mathbf{x}_m \tag{14.394}$$

Based on Equation 14.394, the "largest-of" statistic can be expressed as

$$\Lambda_m(\mathbf{y}) = P_m f_{\mathbf{D}}(\mathbf{y} - \theta\mathbf{x}_m) \tag{14.395}$$

For an SIRV interference model, Equation 14.395 reduces to

$$\Lambda_m(\mathbf{y}) = P_m K h_N(\alpha) \tag{14.396}$$

where

$$\alpha = (\mathbf{y} - \theta\mathbf{x}_m)^{\mathrm{T}}(\mathbf{y} - \theta\mathbf{x}_m) \tag{14.397}$$

The SIRV constant K is the same on all hypotheses, and can thus be removed without affecting the final decision. Based on this, the "largest-of" statistic becomes

$$\Lambda_m(\mathbf{y}) = P_m h_N[(\mathbf{y} - \theta\mathbf{x}_m)^{\mathrm{T}}(\mathbf{y} - \theta\mathbf{x}_m)] \tag{14.398}$$

Assume that all hypotheses are equally likely. If this is the case, then the constant P_m can be removed from the previous equation, producing

$$\Lambda_m(\mathbf{y}) = h_N[(\mathbf{y} - \theta\mathbf{x}_m)^{\mathrm{T}}(\mathbf{y} - \theta\mathbf{x}_m)] \tag{14.399}$$

However, it is shown in Section 14.3.3 that the SIRV characteristic function $h_N(\alpha)$ is monotonically nonincreasing. The largest $\Lambda_m(\mathbf{y})$ thus corresponds to the smallest α. Note that α is nonnegative because it is the magnitude squared of a vector. To formulate the processor in terms of a "largest-of" receiver, $-\alpha$ can be

chosen as the statistic. Based on this, the "largest-of" statistic can be expressed as

$$\begin{aligned}\Lambda_m(\mathbf{y}) &= -\alpha = (-1)\cdot[(\mathbf{y}-\theta\mathbf{x}_m)^{\mathrm{T}}(\mathbf{y}-\theta\mathbf{x}_m)] \\ &= 2\theta\mathbf{x}_m^{\mathrm{T}}\mathbf{y} - \mathbf{y}^{\mathrm{T}}\mathbf{y} - \theta^2\mathbf{x}_m^{\mathrm{T}}\mathbf{x}_m \end{aligned} \tag{14.400}$$

Because $\mathbf{x}_m$ has unit amplitude for all m, and θ is assumed to be the same for each hypothesis, the final term ($\theta^2\mathbf{x}_m^T\mathbf{x}_m$) can be removed from the statistic. By the same token, the second term ($\mathbf{y}^T\mathbf{y}$) may also be removed. The resulting statistic is

$$\Lambda_m(\mathbf{y}) = 2\theta\mathbf{x}_m^{\mathrm{T}}\mathbf{y} \tag{14.401}$$

or, after removing like terms,

$$\Lambda_m(\mathbf{y}) = \mathbf{x}_m^{\mathrm{T}}\mathbf{y} \tag{14.402}$$

Equation 14.402 is simply a digital implementation of a matched filter bank. Just as for the Gaussian case, this analysis reveals, when the hypotheses are equally likely, that the optimum M-ary detector for interference modeled by *any* nonGaussian SIRV is a series of matched filters.

However, it is not reasonable to assume equi-probable hypotheses for a correlation sonar. Recall from Section 14.7.3 that each hypothesis corresponds to a different own-ship velocity estimate. Since own-ship velocity is not normally subject to rapid change, previous velocity estimates provide information regarding the probability distribution for the next velocity estimate. In practice, this probability distribution is approximated as Gaussian, where the mean corresponds to the previous velocity estimate (or the previously chosen hypothesis), and the standard deviation is one velocity resolution cell as determined by the spacing of the horizontal array elements. For five hypotheses centered about the previous choice, the hypothesis probabilities are determined from the standard normal density illustrated in Figure 14.53. The resulting probabilities are

$$\begin{aligned} P_1 &= .05856, \\ P_2 &= .24197, \\ P_3 &= .39894, \\ P_4 &= .24197, \\ P_5 &= .05856 \end{aligned} \tag{14.403}$$

With this in mind, Equation 14.398 gives the optimum M-ary "largest-of" statistic for the case where the interference can be modeled as a traditional SIRV and the hypotheses are not equally likely. The following paragraphs derive two explicit expressions for this statistic; one for the case where the interference is Gaussian, and another for the case where the interference is a traditional SIRV which follows from the case 1 Generalized SIRV density derived in Section 14.4.1 by letting M equal unity.

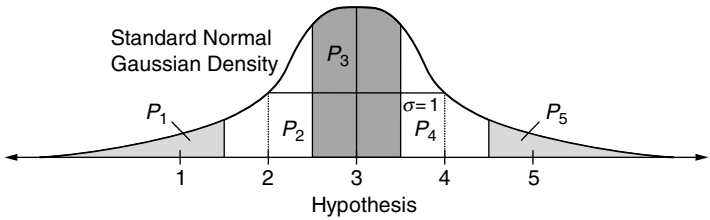

FIGURE 14.53 Hypothesis probability distribution.

Gaussian random vectors are SIRVs with a characteristic function given by

$$h_N(\alpha) = \exp(-\alpha/2) \tag{14.404}$$

Substituting Equation 14.404 into Equation 14.398 gives

$$\Lambda_m(\mathbf{y}) = P_m \cdot \exp[-(1/2)(\mathbf{y} - \theta\mathbf{x}_m)^{\mathrm{T}}(\mathbf{y} - \theta\mathbf{x}_m)] \tag{14.405}$$

Since the logarithm function is monotonically increasing, the logarithm of $\Lambda_m(\mathbf{y})$ can be used as a statistic without affecting the final decision. The "largest-of" statistic thus becomes

$$\begin{aligned} \ln\{\Lambda_m(\mathbf{y})\} &= \ln\{P_m\} - 0.5(\mathbf{y} - \theta\mathbf{x}_m)^{\mathrm{T}}(\mathbf{y} - \theta\mathbf{x}_m), \\ &= \ln\{P_m\} + \theta\mathbf{x}_m^{\mathrm{T}}\mathbf{y} - 0.5(\mathbf{y}^{\mathrm{T}}\mathbf{y} + \theta^2\mathbf{x}_m^{\mathrm{T}}\mathbf{x}_m) \end{aligned} \tag{14.406}$$

As with Equation 14.400 and Equation 14.401, the statistic in Equation 14.406 can be simplified to

$$\ln\{\Lambda_m(\mathbf{y})\} = \ln\{P_m\} + \theta\mathbf{x}_m^{\mathrm{T}}\mathbf{y} \tag{14.407}$$

Equation 14.407 is the optimum M-ary detector "largest-of" statistic for Gaussian interference. The inner product $(\mathbf{x}_m^{\mathrm{T}}\mathbf{y})$ is the output of a conventional matched filter. When the events are equi-probable, and θ is the same on every hypothesis, Equation 14.407 reduces to

$$\ln\{\Lambda_m(\mathbf{y})\} = \mathbf{x}_m^{\mathrm{T}}\mathbf{y} \tag{14.408}$$

As seen, the M-ary detector for Gaussian interference reduces to a matched filter bank, as shown in Figure 14.51 for analog data.

Now consider nonGaussian interference. Section 14.6.4 reveals that nonGaussian surface reverberation can be closely approximated by the case 1 density derived in Section 14.4.1, with shape parameters ($Q = R = 2$) and ($\beta = 0.2$). For the case of a tradition SIRV, (i.e., where the number of vectors in the generalized SIRV formulation is unity), Equation 14.178 gives the case 1 characteristic function as

$$h_N(\alpha) = (\alpha + 0.2)^{-[(3+N)/2]} \tag{14.409}$$

Substituting Equation 14.409 into Equation 14.396 and taking the logarithm gives

$$\ln\{\Lambda_m(\mathbf{y})\} = \ln\{P_m\} + \ln\{K\} - [(3+N)/2]\ln\{\alpha + 0.2\} \tag{14.410}$$

Removing the constant term $\ln\{K\}$ produces

$$\ln\{\Lambda_m(\mathbf{y})\} = \ln\{P_m\} - [(3+N)/2]\ln\{\alpha + 0.2\} \tag{14.411}$$

where

$$\alpha = (\mathbf{y} - \theta\mathbf{x}_m)^{\mathrm{T}}(\mathbf{y} - \theta\mathbf{x}_m) \tag{14.412}$$

Note that the optimum receiver for nonGaussian interference does not reduce to a familiar matched filter bank.

A computer simulation of correlation sonar data was used to evaluate the performance improvement of a nonGaussian receiver over a matched filter bank achieved in an environment corrupted with nonGaussian interference. Recall from Section 14.7.3 that the "deterministic" signal $\mathbf{x}_m$ in a correlation sonar is actually a sample from a slowly varying volume reverberation. Section 14.6.4 reveals that the density of volume reverberation can be closely approximated by the Gaussian density. Figure 14.54 thus shows the scheme used to simulate correlation sonar data.

As seen, the system generates 5 standard normal Gaussian reference vectors, $\{\mathbf{x}_1, \ldots, \mathbf{x}_5\}$. As shown in Figure 14.46, these signals correspond to the volume reverberation echoes received on 5 adjacent X-array elements. For each experimental trial, one of these signals is chosen according to the *a priori* probabilities $\{P_1, \ldots, P_5\}$ and multiplied by the scalar θ. The resulting vector represents the deterministic signal vector of the true hypothesis.

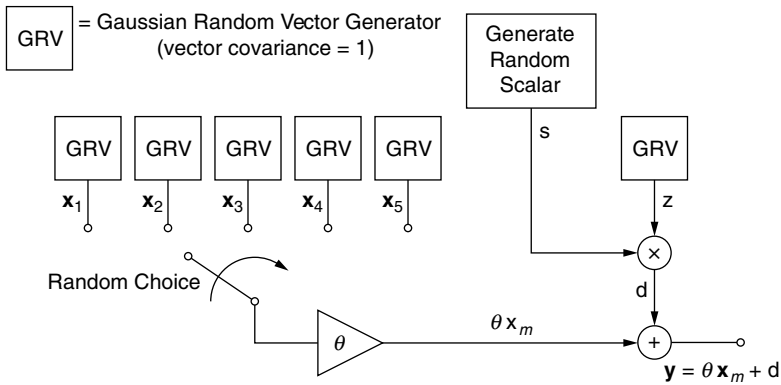

FIGURE 14.54 Correlation sonar data simulator.

The program also creates an interference signal $\mathbf{d}$ by multiplying another Gaussian random vector $\mathbf{z}$ by the random scalar s, whose probability density function is given by Equation 14.167. Section 14.5 explains in detail the procedure for generating SIRV interference data in this fashion. The scalar s is generated so as to make the interference follow the case 1 surface reverberation SIRV density. (When Gaussian interference is wanted, this scalar is set equal to a constant value of unity). Finally, the interference and signal are then combined to form the correlation sonar input vector $\mathbf{y}$.

The program creates 10,000 of these correlation sonar input vectors. It then computes the largest-of statistics as given in Equation (14.407) and Equation (14.411). A correct choice is made if the largest statistic Λ_m corresponds to the true hypothesis. The program tallies up the number of correct choices and then estimates the probability of a correct choice (P_c) by dividing this sum by the total number of trials, 10,000. This experiment is repeated for various values of the signal-to-background ratio (SBR). With reference to Figure 14.54, the SBR is

$$\text{SBR} = \frac{\theta^2}{E\{\mathbf{d}^T\mathbf{d}\}} = \frac{\theta^2}{E\{s^2\}E\{\mathbf{z}^T\mathbf{z}\}} = \frac{\theta^2}{NE\{s^2\}} \tag{14.413}$$

Figure 14.55 shows the performance metric P_c as a function of SBR for both the Gaussian and case 1 receivers. In the top graph the interference is Gaussian, while in the lower graph it is nonGaussian. As seen, when the disturbance $\mathbf{d}$ follows a Gaussian distribution, both receivers return approximately the same P_c. When processing nonGaussian data, however, the case 1 receiver yields a significantly higher value in the range of SBR from -30 to 0 dB.

With reference to Equation 14.407 and Equation 14.411, note that both receivers require explicit knowledge of θ. In practice, the amplitude of the volume reverberation is not known. The next section presents a sub-optimal receiver designed to address this issue.

14.8.2. Sub-Optimum *M*-ary Detection

From Equation 14.411, the optimum statistic for case 1 nonGaussian surface reverberation is

$$\ln\{\Lambda_m(\mathbf{y})\} = \ln\{P_m\} - [(3+N)/2]\ln\{\alpha + 0.2\} \tag{14.414}$$

where

$$\alpha = (\mathbf{y} - \theta\mathbf{x}_m)^T(\mathbf{y} - \theta\mathbf{x}_m) \tag{14.415}$$

Note that this receiver depends on a priori knowledge of the signal amplitude θ. When this value is unknown, the wrong optimum receiver may be employed, leading to poorer performance than that achieved with the correct receiver. Consequently, a robust sub-optimum receiver is designed.

The first class of sub-optimal receivers investigated uses a fixed value of θ, denoted as θ_0. Regardless of the true volume reverberation amplitude θ, the

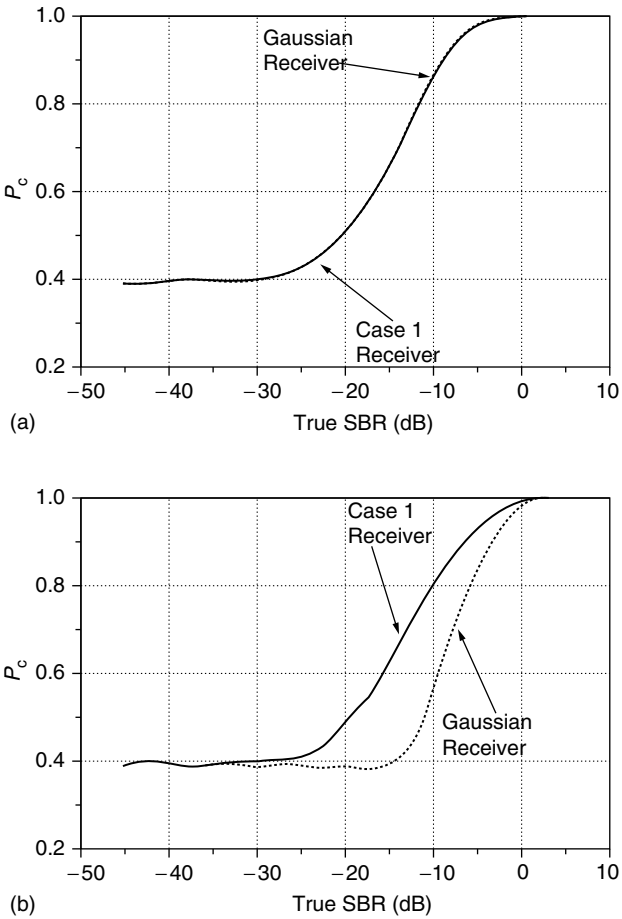

FIGURE 14.55 Correlation sonar optimum performance; (a) Gaussian clutter, (b) Case 1 clutter.

receiver uses the quadratic form

$$\alpha_0 = (\mathbf{y} - \theta_0 \mathbf{x}_m)^{\mathrm{T}} (\mathbf{y} - \theta_0 \mathbf{x}_m) \tag{14.416}$$

Correlation sonar data was simulated as shown in Figure 14.54, and the performance of this sub-optimal receiver was measured in terms of P_c. Figure 14.56 shows the results for three values of θ_0. The SBR is varied by changing the actual value of θ. As seen, the sub-optimal performance approaches the ideal receiver performance only when the true value of θ lies close to θ_0. This result suggests that a bank of filters, each with a different θ_0, may be employed to yield higher values of P_c.

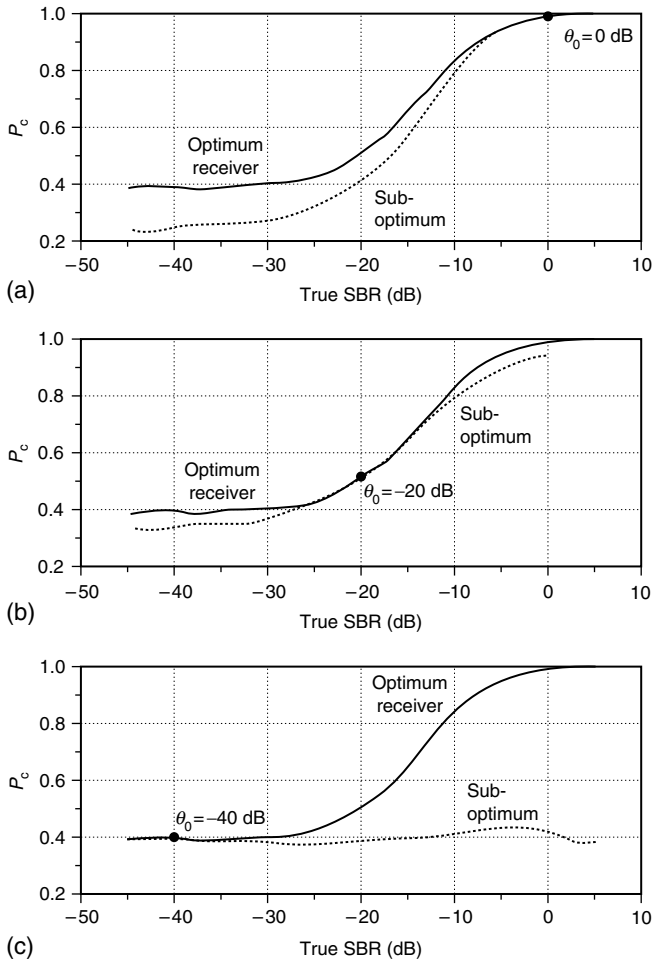

FIGURE 14.56 Correlation sonar sub-optimum performance; (a) $\theta_0 = 0$ dB, (b) $\theta_0 = -20$ dB, (c) $\theta_0 = -40$ dB.

The second class of sub-optimal receivers investigated uses a filter bank, as shown in Figure 14.57. Note that each filter outputs five statistics. For a receiver with 6 filters, the final decision must be made by examining all $5 \times 6 = 30$ output statistics.

This analysis considers two different methods for making a decision at the output of an M-ary filter bank. The first approach simply picks the channel which yields the largest output across all 30 statistics. Henceforth, the receiver which employs this method will be referred to as a maximum output filter bank. The second approach first finds the largest response within each filter. Think of this as each filter "voting" for one channel. The receiver then picks the channel which

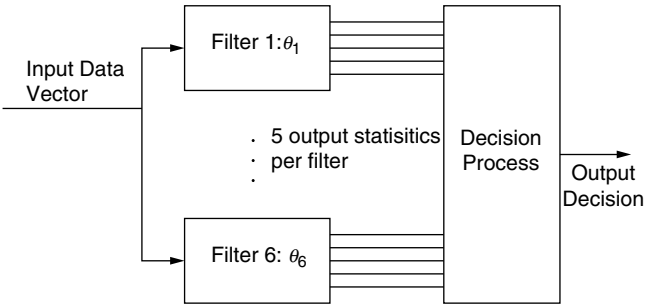

FIGURE 14.57 Sub-optimum filter bank.

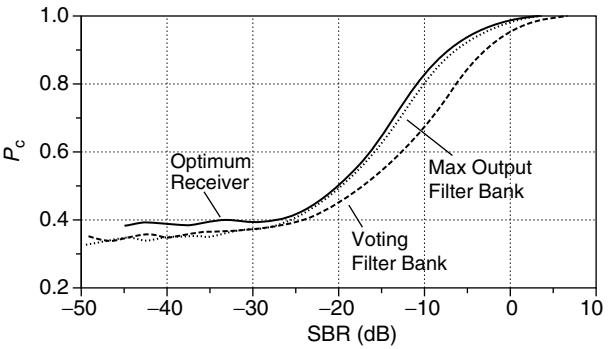

FIGURE 14.58 Correlation sonar filter-bank performance.

receives the most "votes." Henceforth, the receiver which employs this method will be referred to as a voting filter bank.

Correlation sonar data was simulated as shown in Figure 14.54, and the performance of these sub-optimal receivers was measured in terms of P_c. Figure 14.58 shows the results for both types of filter banks. As seen, the maximum output filter bank yields results closer to optimal than the voting filter bank.

All of the receivers designed so far have been based on a traditional SIRV interference model. The next section presents receivers derived to process generalized SIRV interference.

14.8.3. Generalized SIRV *M*-ary Detection

From Equation 14.393, the optimum largest-of statistic for *M*-ary detection is

$$\Lambda_m(\mathbf{y}) = P_m f_{\mathbf{Y}|H_m}(\mathbf{y}|H_m) \tag{14.417}$$

where:

$$\mathbf{y} = \theta \mathbf{x}_m + \mathbf{d} \tag{14.418}$$

In Equation 14.418, $\mathbf{y}$ is the received data vector, $\mathbf{x}_m$ is the mth deterministic signal vector with unit amplitude, and $\mathbf{d}$ is the interference vector.

In a correlation sonar, the interference $\mathbf{d}$ consists of ambient noise and the dominant previous pulse echo. As such, the dominant interference data in $\mathbf{d}$ arises from the first pulse, while the signal data in $\mathbf{x}_m$ arises from the second pulse. (Since the volume reverberation is assumed to be slowly varying, the signal information $\mathbf{x}_m$ received on the y element is identical to one of the signals received after first pulse transmission on the $\mathbf{x}$ array elements.) Therefore the received vector $\mathbf{y}$ contains information from both pulses, or one correlation sonar pulse pair. This data is processed to yield a single velocity estimate.

By transmitting successive pulse pairs, the system forms successive velocity estimates. Define the $\mathbf{y}$ vector which arises from the first pulse pair as

$$\mathbf{y}_1 = \theta \mathbf{x}_{m,1} + \mathbf{d}_1 \tag{14.419}$$

Similarly, define the $\mathbf{y}$ vector which arises from the second pulse pair as

$$\mathbf{y}_2 = \theta \mathbf{x}_{m,2} + \mathbf{d}_2 \tag{14.420}$$

and so on. In this model, note that $\mathbf{x}_{m,i}$ is the mth signal vector from the ith pulse pair. In a correlation sonar, the mth signal results from a slowly varying volume reverberation echo. Hence, it is reasonable to expect differences between pulse pairs.

The traditional SIRV model does not allow for dependence between successive interference vector $\mathbf{d}_1$ and $\mathbf{d}_2$. However, when the interference is modeled as a generalized SIRV, the density of the interference is given by

$$f_{\bar{\mathbf{D}}}(\bar{\mathbf{d}}) = Kh_{\mathrm{NP}}(\alpha_1, \alpha_2, \ldots, \alpha_p) \tag{14.421}$$

where:

$$\bar{\mathbf{d}} = [\mathbf{d}_1, \mathbf{d}_2, \ldots, \mathbf{d}_P] \tag{14.422}$$

In Equation 14.422, the set $\{\mathbf{d}_1, \ldots, \mathbf{d}_P\}$ represents P interference vectors received from P successive correlation sonar pulse pairs. The length of each vector is N. In Equation 14.421, the scalars $\{\alpha_1, \ldots \alpha_P\}$ are defined by the quadratic form

$$\alpha_i = \mathbf{d}_i^{\mathrm{T}} \mathbf{R}_i \mathbf{d}_i \tag{14.423}$$

for all i from 1 to P, where $\mathbf{R}_i$ is the covariance matrix of the ith interference vector. Henceforth, without loss of generality, this covariance matrix is assumed to be equal to the identity matrix for all i such that

$$\alpha_i = \mathbf{d}_i^{\mathrm{T}} \mathbf{d}_i \tag{14.424}$$

Define $\mathbf{y}_i$ as the data vector received after the ith pulse pair. From Equation 14.418,

$$\mathbf{y}_i = \theta \mathbf{x}_{m,i} + \mathbf{d}_i \tag{14.425}$$

From Equation 14.417, the generalized SIRV largest-of statistic is

$$\Lambda_m(\mathbf{y}_1, \ldots, \mathbf{y}_P) = P_m f_{(\mathbf{Y}_1,\ldots,\mathbf{Y}_P)|\mathbf{H}_m}(\mathbf{y}_1, \ldots, \mathbf{y}_P|\mathbf{H}_m) \tag{14.426}$$

or

$$\Lambda_m(\mathbf{y}_1, \ldots, \mathbf{y}_P) = P_m f_{(\mathbf{D}_1,\ldots,\mathbf{D}_P)}\{(\mathbf{y}_1 - \theta \mathbf{x}_{m,1}), \ldots, (\mathbf{y}_P - \theta \mathbf{x}_{m,P})\} \tag{14.427}$$

Substituting Equation 14.421 into Equation 14.427 gives

$$\Lambda_m(\mathbf{y}_1, \ldots, \mathbf{y}_P) = P_m K h_{NP}(\alpha_1, \ldots, \alpha_P) \tag{14.428}$$

where:

$$\alpha_i = (\mathbf{y}_i - \theta \mathbf{x}_{m,i})^{\mathrm{T}}(\mathbf{y}_i - \theta \mathbf{x}_{m,i}) \tag{14.429}$$

for all i from one to P.

Assume that the interference is independent from ping to ping and follows a Gaussian distribution. The generalized SIRV characteristic function for such interference is

$$h_{NP}(\alpha_1, \ldots, \alpha_P) = \exp\{-(1/2)(\alpha_1 + \alpha_2 + \cdots + \alpha_P)\} \tag{14.430}$$

Assume that the number of generalized SIRVs is two (i.e., $P = 2$). Substituting Equation 14.430 into Equation 14.428 gives

$$\Lambda_m(\mathbf{y}_1, \mathbf{y}_2) = P_m K \exp\{-\alpha_1/2\}\{-\alpha_2/2\} \tag{14.431}$$

Removing the constant K and taking the logarithm produces

$$\ln[\Lambda_m(\mathbf{y}_1, \mathbf{y}_2)] = \ln\{P_m\} - \alpha_1/2 - \alpha_2/2 \tag{14.432}$$

Substituting Equation 14.429 into Equation 14.432 yields

$$\begin{aligned}\ln[\Lambda_m(\mathbf{y}_1, \mathbf{y}_2)] = \ln\{P_m\} &- (1/2)\mathbf{y}_1^{\mathrm{T}}\mathbf{y}_1 - (1/2)\theta^2 \mathbf{x}_{m,1}^{\mathrm{T}}\mathbf{x}_{m,1} + \theta \mathbf{x}_{m,1}^{\mathrm{T}}\mathbf{y}_1 \\ &- (1/2)\mathbf{y}_2^{\mathrm{T}}\mathbf{y}_2 - (1/2)\theta^2 \mathbf{x}_{m,2}^{\mathrm{T}}\mathbf{x}_{m,2} + \theta \mathbf{x}_{m,2}^{\mathrm{T}}\mathbf{y}_2\end{aligned} \tag{14.433}$$

Removing terms which do not change from hypothesis to hypothesis gives

$$\ln[\Lambda_m(\mathbf{y}_1, \mathbf{y}_2)] = \ln\{P_m\} + \theta \mathbf{x}_{m,1}^{\mathrm{T}}\mathbf{y}_1 + \theta \mathbf{x}_{m,2}^{\mathrm{T}}\mathbf{y}_2 \tag{14.434}$$

As seen, for pulse pair independent Gaussian interference, the optimum receiver is the sum of the output from two matched filter banks.

Now assume that the interference follows the case 1 SIRV surface reverberation model identified in Section 14.6.4. For shape parameters ($Q_1 = Q_2 = R_1 = R_2 = 2$) and ($\beta = 0.2$), Equation 14.178 gives the generalized SIRV

characteristic function as

$$h_{N2}(\alpha_1, \alpha_2) = K\{(\alpha_1 + 0.2)(\alpha_2 + 0.2)\}^{-\left(\frac{1+N}{2}\right)} \times \left\{\frac{1}{\alpha_1 + 0.2} + \frac{1}{\alpha_2 + 0.2}\right\} \tag{14.435}$$

Substituting Equation 14.435 into Equation 14.428, taking the logarithm, and removing like terms produces

$$\Lambda_m(\mathbf{y}_1, \mathbf{y}_2) = \ln\{P_m\} - \left(\frac{1+N}{2}\right)\{\ln(\alpha_1 + 0.2) + \ln(\alpha_2 + 0.2)\} + \ln\left\{\frac{1}{\alpha_1 + 0.2} + \frac{1}{\alpha_2 + 0.2}\right\} \tag{14.436}$$

where:

$$\alpha_1 = (\mathbf{y}_1 - \theta\mathbf{x}_{m,1})^{\mathrm{T}}(\mathbf{y}_1 - \theta\mathbf{x}_{m,1}) \tag{14.437}$$

and

$$\alpha_2 = (\mathbf{y}_2 - \theta\mathbf{x}_{m,2})^{\mathrm{T}}(\mathbf{y}_2 - \theta\mathbf{x}_{m,2}) \tag{14.438}$$

Equation 14.436 is the optimum M-ary detector largest-of statistic for the case where the interference follows the generalized SIRV surface reverberation model.

A computer simulation was implemented to evaluate the effectiveness of the optimum receivers developed above. Figure 14.59 shows a block diagram of the scheme used to simulate correlation sonar data. The only difference between this and the previous scheme illustrated in Figure 14.54 is that the program generates a pair of Gaussian random vectors for each hypothesis. To create the interference,

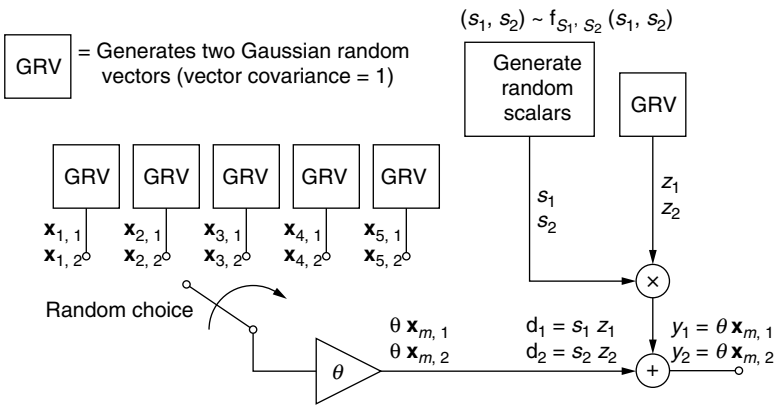

FIGURE 14.59 Correlation sonar 2-D data simulator.

the program simulates two dependent scalars s_1 and s_2. When multiplied by two independent Gaussian vectors, these scalars produce the desired pair of generalized SIRVs.

In order to evaluate the generalized SIRV receivers, a computer program creates 10,000 pairs of correlation sonar input vectors. It then computes the largest-of statistics as given in Equation (14.434) and Equation (14.436). Once again, a correct choice is made if the largest statistic Λ_m corresponds to the true hypothesis. This experiment is repeated for various values of the signal-to-background ratio (SBR). Figure 14.60 shows the resulting performance metric P_c, the probability of correct detection, as a function of the SBR for the Gaussian and case 1 receivers. Once again, when processing nonGaussian data, the case 1 receiver yields higher values of P_c for values of SBR from -30 to 0 dB.

With reference to Equation (14.434) and Equation (14.436), note that both receivers require explicit knowledge of θ. When this quantity is unknown, a robust sub-optimum processor is used, as previously described in Section 14.8.2. Figure 14.61 shows the result of processing the simulated correlation sonar data with two sub-optimum receivers; a maximum output filter bank and a voting filter

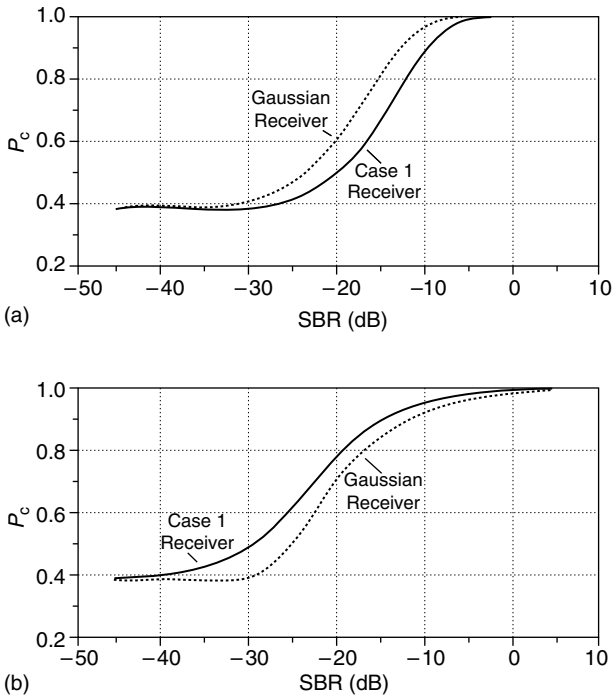

FIGURE 14.60 Generalized SIRV optimum M-ary performance; (a) 2D Gaussian clutter, (b) 2D Case 1 clutter.

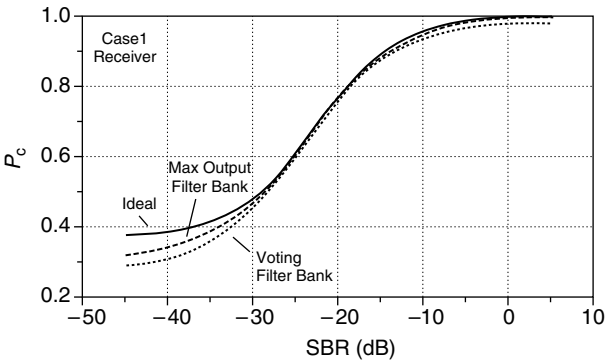

FIGURE 14.61 Generalized SIRV sub-optimum *M*-ary performance with 2D case 1 clutter.

bank. Once again, the maximum output filter bank yields results closer to those achieved with an optimum receiver using the correct value of θ.

In summary, this section demonstrates that correlation sonar performance against nonGaussian interference can be improved by utilizing a nonGaussian receiver. This holds true for both the traditional and generalized SIRV interference models. The improvement is especially significant for quiet targets with Signal-to-Background ratios in the range of -30 to 0 dB.

14.9. CONCLUSION

The original research presented within this dissertation includes several significant findings, as detailed in the paragraphs below.

Section 14.2 presents a generalized SIRV clutter model. The generalization represents an improvement over the traditional model in that the new version can account for dependence between SIRV realizations. In the radar case, the generalized model can account for dependent returns from neighboring spatial range/bearing cells. In the sonar case, the generalized model can account for dependent returns from successive pings.

Section 14.3 derives a significant properties associated with the generalized SIRV. One interesting property is that the generalized SIRV characteristic function $h_{NM}(\alpha_1, \ldots, \alpha_M)$ remains invariant when the random vector undergoes a linear transformation, just like the traditional SIRV. Another property reveals that the joint envelope density for a set of generalized SIRVs is unique. Because of this property, one can approximate data with a generalized SIRV by solely examining the vector envelopes. Section 14.3 also presents a method for deriving explicit expressions for the generalized SIRV density based on the Laplace transform.

Section 14.4 derives four generalized SIRV density functions. The first two are derived based on direct substitution into a multi-dimensional integral obtained from the generalized SIRV representation theorem. The remaining two density functions are derived with the Laplace transform representation.

Section 14.5 presents a method for generating data which follows a given generalized SIRV density function. This analysis includes an extension of the rejection theorem for generating a random variable[15] in order to account for multi-dimensional data.

Section 14.6 describes how to approximate with a generalized SIRV the density underlying random data. This work required extension of the Ozturk algorithm[20] in order to account for multi-dimensional data. Section 14.6 concludes with an analysis of real data. This analysis reveals that a suitable choice of shape parameters exists such that surface reverberation can be closely modeled by the first nonGaussian generalized SIRV density derived in Section 14.4.

Section 14.7 presents an analysis of a correlation sonar, an application particularly sensitive to surface reverberation. After deriving an optimum receiver that is difficult to implement, the analysis presents a practical sub-optimum receiver. When the ambient oceanic noise is small as compared to the received acoustic echoes, the performance of this sub-optimum receiver approaches that of the optimum receiver.

Section 14.8 concludes the analysis by simulating correlation sonar data and evaluating the receivers derived in Section 14.7. In this analysis, the interference is modeled by the case 1 nonGaussian generalized SIRV identified in Section 14.6. When processing nonGaussian data, the case 1 receiver significantly outperformed the Gaussian matched filter bank in the region of signal-to-background ratios from -30 to 0 dB.

14.9.1. Suggestions for Future Research

The following paragraphs outline four areas which serve as a logical extension of the work presented in this dissertation.

First, additional closed form generalized SIRV density functions need to be derived, in order to build a more complete library of interference models. The Laplace transform representation presented in Section 14.3 lends itself to this work. This presentation allows for the direct derivation of closed-form generalized SIRV density functions, without solving any multi-dimensional integrals.

Upon creation of the interference models described above, the Ozturk approximation map should be expanded based on these densities. A more complete map allows for a more accurate approximation of the underlying density followed by actual reverberation.

Following this work, the Ozturk algorithm should be employed to approximate the underlying distribution of additional recorded data. This

analysis should focus on surface and bottom reverberation, since the protection of shallow coastal water has become a priority in today's Navy.

Finally, nonGaussian filters should be derived to process the reverberation data which follows the densities approximated above. This work should be applied to more conventional sonar systems, as opposed to the lesser known correlation sonar.

15 Applications

CONTENTS

15.1. Statistical Normalization of Spherically Invariant NonGaussian Clutter 915
(*T. J. Barnard and F. Khan*)
15.1.1. Introduction 915
15.1.2. Background 916
15.1.3. SIRV Examples 919
15.1.4. Pareto SIRV GLRT 920
15.1.5. Statistical Normalization 925
15.1.6. Conclusion 927
15.2. NonGaussian Clutter Modeling and Application to Radar Target Detection 928
(*A. D. Keckler, D. L. Stadelman, and D. D. Weiner*)
15.2.1. Introduction 928
15.2.2. Summary of the SIRV Model 929
15.2.3. Distribution Approximation Using the Ozturk Algorithm 930
15.2.4. Approximation of SIRVs 933
15.2.5. NonGaussian Receiver Performance 936
15.2.6. Concluding Remarks 938
15.3. Adaptive Ozturk-Based Receivers for Small Signal Detection in Impulsive NonGaussian Clutter 938
(*D. L. Stadelman, A. D. Keckler, and D. D. Weiner*)
15.3.1. Introduction 938
15.3.2. Summary of the SIRV Model 940
15.3.3. The Ozturk Algorithm and SIRV PDF Approximation 941
15.3.4. NonGaussian SIRV Receivers 944
15.3.5. Graphical Representation of SIRV Receiver Behavior 945
15.3.6. Adaptive Ozturk-Based Receiver 951
15.3.7. Conclusions 953
15.4. Efficient Determination of Thresholds via Importance Sampling for Monte Carlo Evaluation of Radar Performance in NonGaussian Clutter 955
(*D. L. Stadelman, D. D. Weiner, and A. D. Keckler*)
15.4.1. Introduction 955
15.4.2. The Complex SIRV Clutter Model 956

15.4.3. NonGaussian SIRV Receivers 957
15.4.3.1. Known Covariance Matrix Case 959
15.4.3.2. Unknown Covariance Matrix Case 959
15.4.4. Importance Sampling 960
15.4.5. Estimation of SIRV Detector Thresholds with Importance Sampling 962
15.4.6. Extreme Value Theory Approximation 967
15.5. Rejection-Method Bounds for Monte Carlo Simulation of SIRVs 968
(*A. D. Keckler and D. D. Weiner*)
15.5.1. Introduction 968
15.5.2. Summary of the SIRV Model 969
15.5.3. Generation of SIRV Distributed Samples 970
15.5.4. Generation of PDF Bounds 975
15.5.5. Concluding Remarks 979
15.6. Optimal NonGaussian Processing in Spherically Invariant Interference 980
(*D. Stadelman and D. D. Weiner*)
15.6.1. Introduction 980
15.6.2. A Review of the SIRV Model 982
15.6.2.1. Definition of the SIRV Model 982
15.6.2.2. SIRV Properties 984
15.6.2.3. The Complex SIRV Model 987
15.6.2.4. Examples 988
15.6.3. Optimal Detection in NonGaussian SIRV Clutter 988
15.6.3.1. Introduction 988
15.6.3.2. Completely Known Signals 989
15.6.3.3. Signals with Random Parameters 990
15.6.3.4. Generalized Likelihood Ratio Test 1005
15.6.3.5. Maximum Likelihood Matched Filter 1008
15.6.4. Nonlinear Receiver Performance 1011
15.6.4.1. Introduction 1011
15.6.4.2. Indirect Simulation of SIRV Receiver Statistics 1012
15.6.4.3. Student t SIRV Results 1014
15.6.4.4. DGM Results 1018
15.6.4.5. NP vs. GLRT Receiver Comparison 1020
15.6.4.6. Additional Implementation Issues 1022
15.6.4.7. Summary 1023
15.7. Multichannel Detection for Correlated NonGaussian Random Processes Based on Innovations 1024
(*M. Rangaswamy, J. H. Michels, and D. D. Weiner*)
15.7.1. Introduction 1024
15.7.2. Preliminaries 1025
15.7.3. Minimum Mean-Square Estimation Involving SIRPs 1026
15.7.4. Innovations-Based Detection Algorithm for SIRPs Using Multichannel Data 1028

15.7.4.1. Block Form of the Multichannel Likelihood Ratio 1028
15.7.4.2. Sequential Form of the Multichannel Likelihood Ratio 1029
15.7.5. Detection Results Using Monte-Carlo Simulation 1032
15.7.6. Estimator Performance for SIRPs 1036
15.7.7. Conclusion 1037

15.1. STATISTICAL NORMALIZATION OF SPHERICALLY INVARIANT NONGAUSSIAN CLUTTER

(T. J. BARNARD AND F. KHAN)

15.1.1. Introduction

Three critical requirements of active sonar systems are as follows:

(1) constant false-alarm rate (CFAR) relative to undesired clutter;
(2) maximized probability of detection (PD) relative to desired contacts;
(3) uniform background on the display.

The third requirement, which necessitates a consistent background mean and variance, enables operator-assisted detection.

When searching for active sonar returns in background interference, the likelihood-ratio test (LRT) maximizes PD for a specified probability of false alarm (PFA). Given a Gaussian background, this LRT reduces to comparing the normalized matched filter output power to a threshold. Under a Gaussian assumption, this detector meets all three of the requirements above, as the output power consistently follows a unit-mean/unit-variance exponential density.

However, an active sonar operating in littoral waters faces interference from bottom and surface echoes, as well as particulates suspended in the water. Signal detection within such clutter or reverberation requires specialized processing[1] and interference modeling,[2–4] since strong interference raises the tail of the background distribution above that arising from Gaussian noise alone. Processing with a matched filter thus increases the PFA, as many of these undesired returns cross the Gaussian-based threshold. Such clutter also has an increased variance (even after unit-mean normalization), which leads to an inconsistent background density.

One solution involves modeling the non-Gaussian interference at the signal-processor output with spherically invariant random vectors (SIRVs). This model allows for application of an LRT and threshold which adapt to the dominant non-Gaussian background. The SIRV model also accounts for correlation between vector components and remains invariant to linear transforms such as demodulation and beamforming. This latter advantage means that the SIRV

model can be applied at the output of a conventional front end processing chain and still achieve optimal PD/PFA performance; it is unnecessary to implement a different non-Gaussian beamformer for each chosen SIRV. Instead, only the final detector stage is changed so that, the sonar system front end output is changed into a scalar for threshold comparison.

In radar applications, the SIRV components model successive wave-train returns from a single range/bearing cell of interest.[5–7] When present, a signal component exists on every element of the received vector. However, in many active sonar applications, propagation speed and environmental instability limit the effectiveness of a multipulse wave-train. Instead, the system employs a single long-pulse broadband waveform and the SIRV contains the echo from the resolution cell of interest and the surrounding neighborhood. For a spatial point target assumption, the signal now exists on only one element of the received vector. Depending on the clutter model, this assumption simplifies the LRT and enables closed-form derivation of the density at the detector output.

Given this density, the monotonic function, that converts the LRT output power to data which follows a unit-mean exponential distribution, can be found. Note that application of this monotonic function does not affect the optimum PD/PFA statistics of the detector. However, this transform statistically normalizes the data; despite a varying heavy-tailed input, the output always follows the same density and, thus, requires same threshold to achieve a desired CFAR output. Furthermore, the resulting display now has a consistent background distribution, which aids the sonar operator trying to make manual detection decisions.

15.1.2. Background

As shown in Ref. 8, multiplying a Gaussian random vector $\bar{Z}$ by a random scalar S gives rise to an SIRV

$$\bar{X} = [X_1 \cdots X_N]^{\mathrm{T}} = S \cdot \bar{Z} \tag{15.1}$$

assuming that a nonnegative random scalar allows for the derivation of several useful properties and does not lead to a loss of generality.[9] If the processor output undergoes unit-power normalization prior to detection, then $E(S^2) = 1$. As stated in Section 15.1, for active sonar systems, the components of the SIRV $\bar{X}$ contain the returns from the spatial cell of interest and the surrounding neighborhood. For example, the vector could contain the response from adjacent bearing cells (or beams) at a given range. Under this model, the covariance matrix $\mathbf{\Sigma}$ of the Gaussian vector $\bar{Z}$ models the underlying spatial correlation in this neighborhood. Also, the "slowly varying" random scalar S is assumed to remain constant across this region of space. Ref. 10 helps to validate the multiplicative model shown in Equation 15.1 by demonstrating how a finite number of scatterers in a resolution cell give rise to a K-distributed envelope density, which serves as a popular SIRV-based clutter model in the radar community.[11–18]

The underlying Gaussian nature of an SIRV gives rise to a canonical density or

$$f_{\bar{X}}(\bar{x}) = h_N(\bar{x}^H \boldsymbol{\Sigma}^{-1} \bar{x}) / \pi^N \det(\boldsymbol{\Sigma}) \tag{15.2}$$

where the characteristic nonlinear function is defined as

$$h_N(u) = \int_0^\infty s^{-2N} \mathrm{e}^{-u/s^2} f_S(s) \mathrm{d}s \tag{15.3}$$

As seen, this representation depends solely on the density of the random scalar $f_S(s)$ and the covariance matrix $\boldsymbol{\Sigma}$. Ref. 11 uses this representation to provide a tabulated library of SIRV density functions based on the desired distribution of the marginal envelope

$$r_1 = |x_1| = \sqrt{x_1 x_1^*} \tag{15.4}$$

It also is proven in Ref. 11 that the scalar quadratic form of the SIRV

$$\alpha = \bar{x}^H \cdot \boldsymbol{\Sigma}^{-1} \cdot \bar{x} \tag{15.5}$$

serves as a sufficient statistic for identifying the multivariate density function, given the known covariance. Ref. 12 comes to the same conclusion regarding the marginal envelope shown in Equation 15.4. In other words, when fitting data into a particular clutter model, the shape parameters in the univariate density $f_A(\alpha)$ need to be estimated rather than the multivariate SIRV density $f_{\bar{X}}(\bar{x})$. The univariate approach requires fewer independent samples and is more robust.

Based on the notation defined above, the densities of these sufficient statistics become

$$f_{R_1}(r) = 2r h_1(r^2) \tag{15.6}$$

and

$$f_A(\alpha) = \alpha^{N-1} h_N(\alpha) / (N-1)! \tag{15.7}$$

The density of the marginal power $\alpha_1 = |x_1|^2$ thus becomes

$$f_{A_1}(\alpha) = h_1(\alpha) \tag{15.8}$$

We use Equation 15.2, Equation 15.3 and Equation 15.8 to characterize the SIRV-based clutter model.

Next, consider the detection of random signals in non-Gaussian interference using an LRT matched to this SIRV-based model. We have two hypotheses regarding the received data vector $\bar{x}: H_0$ with no signal and a SIRV-based background

$$H_0 : \bar{x} = s \cdot \bar{z} \tag{15.9}$$

and H_1 with a signal present in the same interference

$$H_1 : \bar{x} = a \cdot \bar{g} + s \cdot \bar{z} \tag{15.10}$$

In these equations, $\bar{x}$ is an $[N \times 1]$ vector that contains the received data and $\bar{y}$ is an $[N \times 1]$ normalized replica of the expected signal structure ($\bar{g}^H \bar{g} = N$). Under a Swerling type-I assumption, the amplitude a is a slowly fluctuating random variable which remains constant for the duration of the return. The LRT

$$\Lambda(\bar{x}) = \frac{f_{\bar{X}|H_1}(\bar{x})}{f_{\bar{X}|H_0}(\bar{x})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{15.11}$$

which compares the function $\Lambda(\bar{x})$ to a threshold η, maximizes PD for a fixed PFA. The density in the denominator is simply that of the SIRV background or, from Equation 15.2.

$$f_{\bar{X}|H_0}(\bar{x}) = h_N(\bar{x}^H \boldsymbol{\Sigma}^{-1} \bar{x}) / \pi^N \det(\boldsymbol{\Sigma}) \tag{15.12}$$

The density in the numerator is difficult to derive in closed-form when facing a random amplitude signal. This means, the implementation of the optimum test requires computationally intensive numeric integration. In contrast, a generalized LRT (GLRT) assumes deterministic, yet still unknown, signal amplitude and inserts an Maximum Likelihood Estimate (MLE) thereof into the likelihood test.[13] This incorrect assumption enables a realizable closed-form, yet suboptimal, detector. Based on this deterministic model, the received density conditioned on H_1 becomes

$$f_{\bar{X}|H_1}(\bar{x}) = h_N((\bar{x} - a \cdot \bar{g})^H \boldsymbol{\Sigma}^{-1} (\bar{x} - a \cdot \bar{g})) / \pi^N \det(\boldsymbol{\Sigma}) \tag{15.13}$$

GLRT thus reduces to

$$\Lambda_{\mathrm{GLRT}}(\bar{x}) = \frac{h_N[(\bar{x} - \hat{a} \cdot \bar{g})^H \boldsymbol{\Sigma}^{-1} (\bar{x} - \hat{a} \cdot \bar{g})]}{h_N(\bar{x}^H \boldsymbol{\Sigma}^{-1} \bar{x})} \tag{15.14}$$

where $\hat{a}$ represents the MLE of the amplitude, given in Ref. 13 as

$$\hat{a} = \frac{1}{N} \cdot (\bar{g}^H \bar{x}) \tag{15.15}$$

Substituting this estimate into Equation 15.14 and assuming whitened data (i.e., $\boldsymbol{\Sigma} = I$, where I is the identity matrix) finally gives the GLRT as

$$\Lambda_{\mathrm{GLRT}}(\bar{x}) = \frac{h_N\left(\bar{x}^H \bar{x} - \frac{1}{N} \cdot |\bar{g}^H \bar{x}|^2\right)}{h_N(\bar{x}^H \bar{x})} \tag{15.16}$$

Equation 15.16 represents a canonical form of the SIRV-based GLRT. Specific examples are presented in Section 15.1.3. Note that, for a Gaussian background, where

$$h_N(u) = \mathrm{e}^{-u} \tag{15.17}$$

the GLRT reduces to

$$\Lambda(\bar{x}) = \frac{1}{N} \cdot |\bar{g}^H \bar{x}|^2 \tag{15.18}$$

which is a standard matched filter across the domain in question.

15.1.3. SIRV Examples

As an example of SIRV, consider the gamma density function with shape parameters λ and α or

$$f_V(v) = \frac{\lambda^\alpha}{\Gamma(\alpha)} v^{\alpha-1} \mathrm{e}^{-\lambda v}, \qquad s \geq 0 \text{ and } \lambda, \alpha \geq 0 \tag{15.19}$$

where the gamma function is defined by

$$\Gamma(u) = \int_0^\infty z^{u-1} \mathrm{e}^{-z} \, \mathrm{d}z \tag{15.20}$$

If we let $\lambda = \alpha =$ integer M and $s = v^{1/2}$, the generating density becomes

$$\begin{aligned} f_S(s) &= 2s f_V(s^2) = 2s \frac{M^M}{(M-1)!} s^{2(M-1)} \mathrm{e}^{-Ms^2} \\ &= \frac{2M^M}{(M-1)!} s^{2M-1} \mathrm{e}^{-Ms^2} \qquad s \geq 0, \text{ integer } M > 0 \end{aligned} \tag{15.21}$$

Note that the chosen shape parameters $\lambda = \alpha =$ integer M drive the expected value of s^2 to unity, i.e., $E(S^2) = 1$, which will happen after unit-power normalization in the signal-processor. Equation 15.3 then gives the resulting characteristic nonlinear function as

$$h_N(u) = \frac{2M^M}{(M-1)!} \left(\frac{u}{M} \right)^{(M-N)/2} K_{M-N}(2\sqrt{Mu}) \tag{15.22}$$

where $K_M(\cdot)$ is an Mth order modified Bessel function defined as

$$K_n(x) = \frac{\pi}{2} j^{n+1} [J_n(jx) + jY_n(jx)] \tag{15.23}$$

where $J_n(\cdot)$ is the Bessel function

$$J_n(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k! \Gamma(n+k+1)} \left(\frac{x}{2} \right)^{n+2k} \tag{15.24}$$

and $Y_n(\cdot)$ is the Weber function

$$Y_n(x) = \lim_{\nu \to n} \frac{J_\nu(x) \cos(\nu\pi) - J_{-\nu}(x)}{\sin(\nu\pi)} \tag{15.25}$$

Equation 15.6 gives the marginal envelope density as

$$f_R(r) = \frac{4M^{(M+1)/2}}{(M-1)!} r^M K_{M-1}(2\sqrt{M}r) \tag{15.26}$$

which is the K-distributed form common in radar clutter modeling.[11–13,18]

As another example, in the gamma density from Equation 15.19, let $\lambda = \beta$ and $\alpha = \beta + 1$. Furthermore, let $s = v^{-1/2}$, so that the generating density is

$$\begin{aligned} f_S(s) = 2s^{-3} f_V(s^{-2}) &= 2s^{-3} \frac{\beta^{\beta+1}}{\Gamma(\beta+1)} s^{-2\beta} \mathrm{e}^{-\beta s^{-2}} \\ &= \frac{2\beta^{\beta}}{\Gamma(\beta)} \frac{\mathrm{e}^{-\beta/s^2}}{s^{2\beta+3}}, \qquad s \geq 0, \quad \beta > 1. \end{aligned} \tag{15.27}$$

Once again, note that the chosen shape parameters $\lambda = \beta$ and $\alpha = \beta + 1$ drive the expected value of s^2 to unity, i.e., $E(s^2) = 1$. Equation 15.21 gives the resulting characteristic nonlinear function as

$$h_N(u) = \frac{\Gamma(\beta+N+1)}{\Gamma(\beta+1)} \frac{\beta^{(\beta+1)}}{(\beta-u)^{(N+\beta+1)}} \tag{15.28}$$

and Equation 15.26 gives the marginal power density as

$$f_A(\alpha) = \frac{(\beta+1)\beta^{(\beta+1)}}{(\beta+\alpha)^{(\beta+2)}} \tag{15.29}$$

This is normalized, i.e., $E(\alpha) = 1$, and location-shifted version of the generalized Pareto density (GPD)[19] (i.e., $\alpha = z - \beta$ where z follows the GPD). Figure 15.1 shows a plot of this function for various values of the shape parameter β. As this value increases, the Pareto tail approaches that of the complex Gaussian normalized power density.

15.1.4. Pareto SIRV GLRT

Substituting the characteristic nonlinear function from Equation 15.28 into Equation 15.16 gives the Pareto SIRV GLRT as

$$\Lambda(\bar{x}) = \frac{\frac{1}{N} \cdot |\bar{g}^H \bar{x}|^2}{\beta + \left(\bar{x}^H \bar{x} - \frac{1}{N} \cdot |\bar{g}^H \bar{x}|^2 \right)} + 1 \tag{15.30}$$

after applying the monotonic function $u^{1/(N+\beta+1)}$. Note that the application of such simplifying monotonic functions does not change the optimum PD and PFA detection statistics. Furthermore, removing the additive constant and multiplying

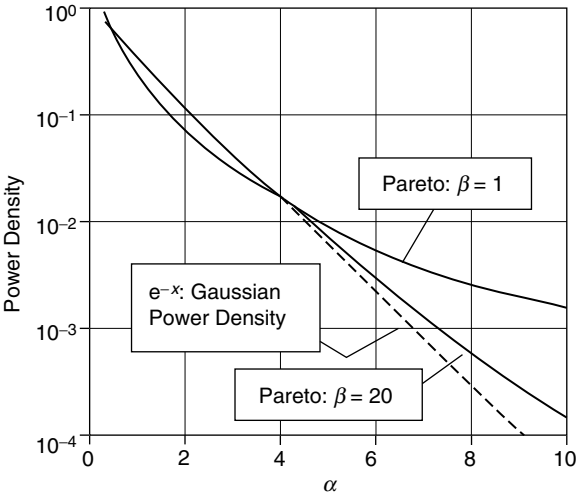

FIGURE 15.1 Pareto power density tail approach to the Gaussian as β grows.

by β produces another equivalent GLRT with the same PD and PFA performance

$$\Lambda(\bar{x}) = \frac{\frac{1}{N}\cdot|\bar{g}^H\bar{x}|^2}{1+\frac{1}{\beta}\cdot\left(\bar{x}^H\bar{x}-\frac{1}{N}\cdot|\bar{g}^H\bar{x}|^2\right)} \tag{15.31}$$

As stated previously, due to the SIRV invariance to linear transforms, we can apply this GLRT after conventional front end processing, such as beamforming and matched filtering. As such, the replica $\bar{g}$ contains the single echo from the resolution cell of interest, along with the echoes from neighboring cells. In this sonar context, a point target residing solely within a single resolution cell gives rise to the replica vector

$$\bar{g}^{\mathrm{T}} = [0,\ldots,\sqrt{N},\ldots,0] \tag{15.32}$$

Substituting this target model into the GLRT thus gives rise to the simplified test

$$\Lambda(\bar{x}) = \frac{|x_m|^2}{1+\frac{1}{\beta}\sum\limits_{\substack{n=1\\n\neq m}}^{N}|x_n|^2} \tag{15.33}$$

where m locates the resolution cell of interest within $\bar{g}$. This GLRT depends solely on the noncoherent power data from the matched filter and as such serves as a more robust detector than those that require a complex target model. Also note that, as β increases, this test approaches the point target GLRT for

a Gaussian noise background, derived from Equation 15.18 as

$$\Lambda_{\text{Gauss}}(\bar{x}) = |x_m|^2 \tag{15.34}$$

This result makes sense because, as β increases, the Pareto SIRV power density approaches the complex Gaussian power density (Figure 15.1). Thus, we can think of the GLRT from Equation 15.33 as a weighted normalizer: as the tail of the background density increases, the denominator becomes more significant.

The GLRT detector compares the output from Equation 15.33 to a threshold η to determine if a signal lies within the received data. When the data consists of background interference only (i.e., the null hypothesis H_0), all threshold crossings are false alarms. Thus, the PFA becomes

$$\mathrm{P}_{\text{fa}} = \int_0^\infty f_{\text{a}|H_0}(\lambda)\mathrm{d}\lambda \tag{15.35}$$

When processing *any* whitened SIRV with the Pareto GLRT shown in Equation 15.33, this PFA becomes

$$\mathrm{P}_{\text{fa}} = \left(\frac{\beta}{\eta+\beta}\right)^{N-1} \int_0^\infty \mathrm{e}^{-\eta/S^2} f_S(s)\mathrm{d}s = \left(\frac{\beta}{\eta+\beta}\right)^{N-1} h_0(\eta) \tag{15.36}$$

Derivation of this result requires use of the integral

$$\frac{\mathrm{e}^{-u}}{u^{(N-1)}} = \int_u^\infty \left[\frac{(N-1)}{z^N} + \frac{1}{z^{N-1}}\right]\mathrm{e}^{-z}\,\mathrm{d}z \tag{15.37}$$

In Equation 15.36, $h_0(\cdot)$ corresponds to the 0th-order characteristic nonlinear function of the SIRV in question, as defined by Equation 15.3.

When the data consists of unit power interference plus a random signal with variance σ_A^2 (the signal hypothesis H_1), all threshold crossings are desired detections. Thus, the PD becomes

$$\mathrm{P}_{\text{d}} = \int_\eta^\infty f_{A|H_1}(\lambda)\mathrm{d}\lambda \tag{15.38}$$

which, for data processed with the Pareto GLRT, reduces to

$$\mathrm{P}_{\text{d}} = \int_\eta^\infty \frac{\exp\left(-\dfrac{\eta}{N\sigma_A^2+s^2}\right) f_S(s)}{\left[\left(\dfrac{s^2}{N\sigma_A^2+s^2}\right)\dfrac{\eta}{\beta}+1\right]^{(N-1)}}\mathrm{d}s \tag{15.39}$$

Note that as the signal power σ_A^2 goes to zero, this PD approaches the PFA, as given by Equation 15.36. Furthermore, as σ_A^2 goes to infinity, P_{d} approach unity, as expected.

We refer to a plot of PD vs. PFA for a specific signal power σ_A^2 as a receiver operating characteristic (ROC) curve. To derive the ROC curve for *any* SIRV

processed with the Pareto GLRT, first invert Equation 15.36 to express the threshold η as a function of PFA and then insert the result into Equation 15.39. The solution to Equation 15.39 typically requires numeric integration. However, this does not present a significant computational burden, as the operating PD of interest typically lies high enough to obviate detailed sums down into the far tails of an integrand. In contrast, the *K*-distributed SIRV characteristic nonlinear function shown in Equation 15.22 does not lead to a simplified GLRT form which one can readily derive the densities $f_{a|H_0}(\lambda)$ and $f_{a|H_1}(\lambda)$. As such, computation of the threshold that gives rise to a low PFA value requires either multidimensional numeric integration into the far tails of known densities or simulating at least (10/PFA) random samples on the null hypothesis and processing with the GLRT from Equation 15.16.

Equation 15.36 and Equation 15.39 define the ROC curve that results when processing any SIRV with the Pareto GLRT Equation 15.33. When processing any SIRV with the point target Gaussian GLRT Equation 15.34, the resulting detection statistics become

$$\mathrm{P_{fa}} = h_0(\eta) \tag{15.40}$$

and

$$\mathrm{P_d} = \int_0^\infty \mathrm{e}^{-\eta/(N\sigma_a^2+s^2)} f_S(s)\mathrm{d}s \tag{15.41}$$

Note how the Pareto GLRT performance from Equation 15.36 and Equation 15.39 approaches that of the Gaussian test shown in Equation 15.40 and Equation 15.41 as β increases.

When processing real data or simulated data which follows some other density, implementation of the Pareto GLRT requires estimation of the shape parameter β. This generally is not an easy problem.[19] However, the MLE has been shown to return robust performance with a limited number of samples. To numerically implement the MLE, we chose a span of potential β values $(\beta_1, \beta_2, \ldots, \beta_M)$ and then pick the one that maximizes the test statistic

$$\Theta(\bar{\alpha}, \beta_m) = \prod_{n=1}^{N} f_A(\alpha_n, \beta_m) = \prod_{n=1}^{N} \frac{(\beta_m + 1)\beta_m^{(\beta_m+1)}}{(\beta_m + \alpha_n)^{(\beta_m+2)}} \tag{15.42}$$

for a given collection of N power samples $\bar{\alpha} = \{\alpha_1, \alpha_2, \ldots, \alpha_N\}$. In this equation, we are estimating β from the marginal power density shown in Equation 15.29.

To illustrate the impact of utilizing the MLE, Figure 15.2 shows the result at $\beta = 2$, of processing Pareto clutter (Equation 15.28 and Equation 15.29) with the Pareto GLRT (Equation 15.33). The thicker leftmost line shows the result of using Equation 15.36 and Equation 15.39 to form the ROC curve. This represents performance achieved with *a priori* knowledge of the true shape parameter. The thinner line to the immediate right of this curve shows the result of simulating

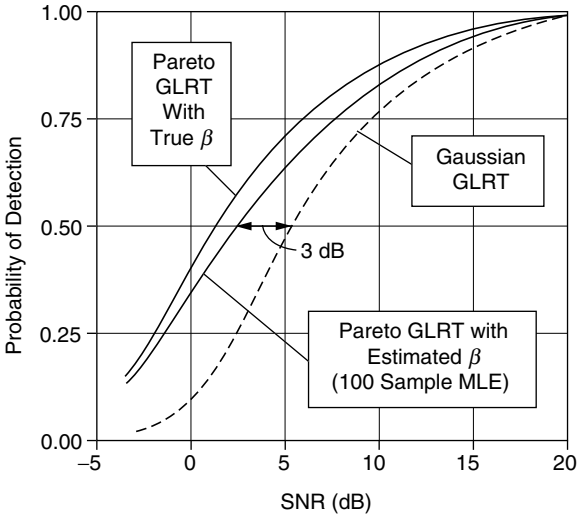

FIGURE 15.2 Processing Pareto clutter ($\beta = 2$) with the proposed GLRT gives rise to a 3 dB gain over the Gaussian GLRT at $P_d = 0.50$ and $P_{fa} = 0.001$.

Pareto clutter data and implementing an MLE and GLRT. In this case, the MLE uses 100 independent samples of the background to estimate the shape parameter. As seen, the estimator degrades performance by about 1 dB at PD = 0.50 and PFA = 0.001. However, this degraded curve still lies 3 dB above that found through application of the Gaussian GLRT (used in conventional sonar processing), shown as the dotted line on the far right and generated by Equation 15.40 and Equation 15.41.

Figure 15.2 shows the result of processing Pareto-SIRV data with the Pareto-SIRV GLRT Equation 15.33. In contrast, Figure 15.3 shows the result of processing heavy-tailed *K*-distributed clutter [($M = 1$) (Equation 15.21 and Equation 15.26)] with this detector. The thicker leftmost line shows the result of using simulated data and the GLRT from Equation 15.16 to form the ROC curve. This represents performance achieved with *a priori* knowledge of the true *K* density. The thinner line to the immediate right of this curve shows the result implementing a 100 sample MLE and Pareto GLRT (Equation 15.33). This time, the estimator degrades performance by less than 1 dB and, once again, the curve lies 3 dB above that found through application of the conventional sonar Gaussian GLRT.

Figure 15.3 demonstrates that the Pareto GLRT suffers a negligible loss compared to the optimal detection of *K-distributed* data. This result is not unexpected, as extreme value theory indicates that the Pareto density reasonably models the tail of *K*-distributed data.[19]

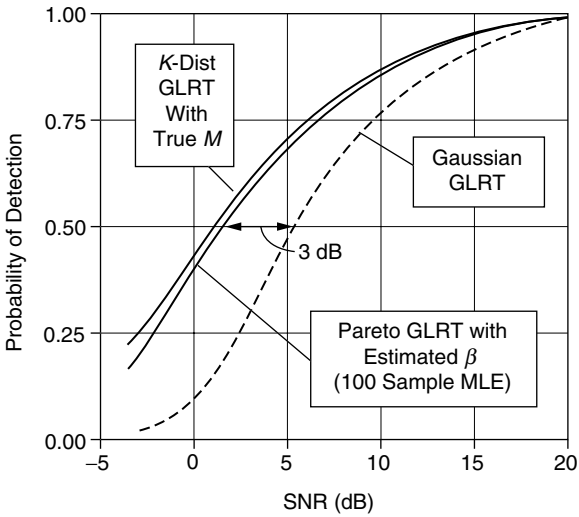

FIGURE 15.3 Processing K-distributed clutter ($M = 1$) with the Pareto GLRT at $P_{\text{fa}} = 0.001$ yields less than 1 dB loss when compared to processing with a matched GLRT.

15.1.5. Statistical Normalization

Use of the GLRT with a varying threshold ensures that the sonar system meets the first two requirements set forth in Section 15.1.1, that of CFAR and a maximized PD. However, an additional processing step is required to achieve the third requirement of a consistent background density. We refer to this process as statistical normalization (SN).[20]

When processing generic SIRV data with the Pareto GLRT, the output under the null hypothesis Equation 15.9 follows the density

$$f_{A|H_0}(\lambda) = \int_0^\infty f_{(A|H_0),s}(\lambda, s)\mathrm{d}s = \int_0^\infty f_{(A|H_0)|s}(\lambda, s) f_S(s)\mathrm{d}s$$

$$= \left(\frac{\hat{\beta}}{\lambda + \hat{\beta}}\right)^{N-1} \left\{\left(\frac{N-1}{\lambda + \hat{\beta}}\right) h_0(\lambda) + h_1(\lambda)\right\} \tag{15.43}$$

where $\hat{\beta}$ is the estimate of β used in the detector. For the Pareto SIRV Equation 15.28, given that the estimate $\hat{\beta}$ exactly equals the true value β, this reduces to

$$f_{A|H_0}(\lambda) = \frac{(N+\beta)\beta^{(N+\beta)}}{(\lambda+\beta)^{(N+\beta+1)}} \tag{15.44}$$

Consider the monotonic transform

$$\psi = (N + \hat{\beta})\ln\left(\frac{\lambda}{\hat{\beta}} + 1\right) \tag{15.45}$$

which yields the same PD and PFA as the original Pareto GLRT when applied to the output. However, under H_0 and assuming that $\hat{\beta}$ exactly equals β, this new test follows the exponential density or

$$f_{\psi|H_0}(\psi) = \frac{d\lambda}{d\psi} f_{A|H_0}(\lambda(\psi)) = e^{-\psi} \tag{15.46}$$

Note that, with an estimated value of β, the resulting density approximates the exponential density, with the quality of this approximation tied to the quality of the estimate.

The transform shown in Equation 15.45 statistically normalizes the heavy-tailed clutter; when the Pareto SIRV well represents the input, the output always approximates the same exponential distribution. This is the same power density as that followed by normalized complex Gaussian data. Since the statistically normalized Pareto GLRT output always approximates the exponential density, the detector achieves true CFAR processing with one threshold and, therefore, has a consistent background distribution.

Implementation of the proposed SN once again requires estimation of the shape parameter β, which can degrade performance. However, Figure 15.4 shows the result of applying Equation 15.45 with an MLE estimate of β to simulated K-distributed data sent through the Pareto GLT Equation 15.22. In fact, this is the same data and Pareto GLRT output used to generate the ROC curves from Figure 15.3, which shows the histogram of this data after application of the statistical normalizer. As seen, despite errors induced through estimation, the resulting normalized histogram lies close to the exponential density, as desired.

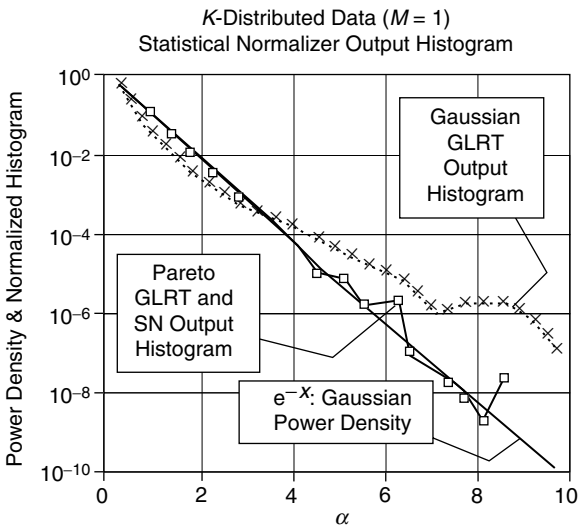

FIGURE 15.4 When presented with heavy-tailed K-distributed data, the statistical normalizer output resembles the exponential density.

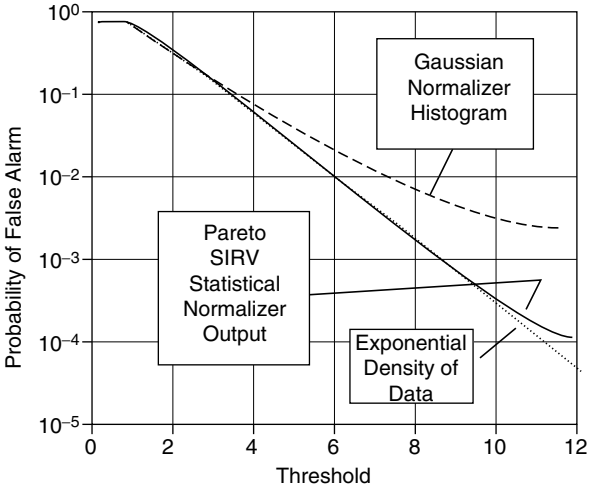

FIGURE 15.5 Applying the proposed processing string to recorded clutter returns drives the heavy tail down to the desired exponential density.

More formally, the exponential density serves as a "good fit" to this data, based on computation of the standard chi-squared testing statistic.[21] Finally, note that for the purpose of comparison, this figure also shows the heavy-tailed output density that arises when utilizing the conventional Gaussian processor Equation 15.34.

As a final validation, Figure 15.5 shows the result of applying an MLE of β, the Pareto GLRT, and statistical normalizer to real clutter-returns recorded in shallow water. This plot shows PFA vs. threshold for the proposed detector and the conventional Gaussian-based approach. Both are compared to the false alarms arising from data that follow the exponential density. As seen, the statistical normalizer drives the heavy tails of the recorded clutter envelope down toward the exponential density, which empirically demonstrates the validity of the proposed Pareto-SRIV model. However, note that, in the far tail of the density, below $P_{\mathrm{FA}} = 10^{-4}$, the model starts to break down due to estimation errors.

15.1.6. Conclusion

The most significant conclusion derived from this work is that the proposed Pareto GLRT allows for robust near-optimal processing and statistical normalization of heavy-tailed clutter. This was demonstrated against a limited set of recorded sea clutter and also against simulated clutter that follows the K-distribution that is popular in radar signal processing. In the latter case, the proposed approach gave rise to a 3-dB gain relative to conventional processing (at $P_{\mathrm{D}} = 0.50$ and $P_{\mathrm{FA}} = 0.001$).

The analysis also derives a closed-form expression for the likelihood output density conditioned on the null hypothesis, which enables direct computation of the threshold required to achieve a given PFA. In contrast, the likelihood test derived from a *K*-distributed clutter model does not reduce to a closed-form density, so threshold computation requires numeric integration. Additionally, the closed-form Pareto SIRV density allows for statistical normalization; without an expression for the density at the likelihood output, we cannot find a monotonic function that returns a consistent background distribution. Such statistical normalization approximates true CFAR by driving the Pareto GLRT output toward an exponential distribution, depending on the quality of the shape parameter estimate.

15.2. NONGAUSSIAN CLUTTER MODELING AND APPLICATION TO RADAR TARGET DETECTION

(A. D. KECKLER, D. L. STADELMAN, AND D. D. WEINER)

15.2.1. Introduction

Conventional radar receivers are based on the assumption of Gaussian distributed clutter. However, the Weibull and *K*-distribution are shown to approximate the envelope of some experimentally measured non-Gaussian clutter data.[1–5] The detection performance of the Gaussian receiver in this environment is significantly below that of the optimum non-Gaussian receiver, especially for weak target returns.

NonGaussian clutter is often observed to be "spiky," as illustrated in Figure 15.6. In such cases, the threshold of the conventional Gaussian receiver must be raised in order to maintain the desired false alarm rate. This results in a reduction of the probability of detection. In contrast, non-Gaussian receivers contain nonlinearities that limit large clutter spikes and allow a lower threshold to be used, which improves performance of targets with a low signal-to-clutter ratio

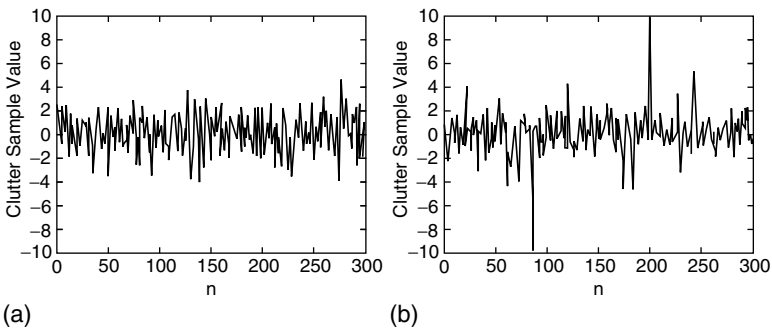

FIGURE 15.6 Comparison of Gaussian data with non-Gaussian data of equal variance: (a) Gaussian example, (b) non-Gaussian example.

(SCR). Determination of these non-Gaussian receivers requires specification of suitable PDFs for the clutter.

The nonhomogeneous and nonstationary clutter environment must be monitored to adapt detection algorithms over the surveillance volume. This is complicated by the need for an efficient technique to accurately approximate a joint clutter PDF that incorporates the pulse-to-pulse correlation. Spherically invariant random vectors (SIRVs), which are explained in this chapter, have been shown to be useful for modeling correlated non-Gaussian clutter.[6] The class includes many distributions of interest, such as the Gaussian, Weibull, Rician, and K-distributed, among others.[6–9]

This section extends the Ozturk algorithm for approximating univariate PDF's[6,10] to the case of multivariate SIRV clutter data. Several issues are addressed with regard to practical implementation of this approach, *viz*., computer simulation of correlated SIRVs, generation of Ozturk approximation charts for the corresponding multivariate PDFs, fit of approximations to the underlying distributions, and impact of using an estimated covariance matrix. The section concludes with a specific example illustrating how well an adaptive Ozturk-based receiver, which approximates the unknown clutter PDF, performs compared to a non-Gaussian receiver with *a priori* knowledge of the clutter PDF.

15.2.2. Summary of the SIRV Model

A random vector $\mathbf{Y}$ of dimension N is defined to be an SIRV if and only if its PDF has the form

$$f_{\mathbf{Y}}(\mathbf{y}) = (2\pi)^{-N/2}|\boldsymbol{\Sigma}|^{-1/2}h_N(q(\mathbf{y})) \tag{15.47}$$

where $\boldsymbol{\Sigma}$ is an $N \times N$ nonnegative definite matrix, $q(\mathbf{y})$ is the quadratic form defined by

$$q = q(\mathbf{y}) = (\mathbf{y} - \mathbf{b})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{y} - \mathbf{b}) \tag{15.48}$$

$\mathbf{b}$ is the $N \times 1$ mean vector, and $h_N(\cdot)$ is a positive, monotonic decreasing function for all N.[11] Equivalently, an SIRV $\mathbf{Y}$ can be represented by the linear transformation

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b} \tag{15.49}$$

where $\mathbf{X}$ is a zero-mean SIRV with uncorrelated components represented by

$$\mathbf{X} = S\mathbf{Z} \tag{15.50}$$

$\mathbf{Z}$ is a zero-mean Gaussian random vector with independent components, and S is a nonnegative random variable independent of $\mathbf{Z}$. The probability density function of S, $f_{\mathrm{S}}(s)$, uniquely determines the type of SIRV and is known as the characteristic PDF of $\mathbf{Y}$. Since the matrix $\mathbf{A}$ is specified independently of $f_{\mathrm{S}}(s)$, an

arbitrary covariance matrix, $\mathbf{\Sigma} = \mathbf{A}\mathbf{A}^{\mathrm{T}}$, can be introduced without altering the type of SIRV.

This representation is used to obtain

$$h_N(q) = \int_0^\infty s^{-N} \mathrm{e}^{-q/2s^2} f_S(s) \mathrm{d}s \tag{15.51}$$

and subsequently, the PDF of the quadratic form is

$$f_Q(q) = \frac{1}{2^{N/2}\Gamma(N/2)} q^{(N/2)-1} h_N(q) \tag{15.52}$$

Since $h_N(q)$ uniquely determines each type of SIRV, Equation 15.52 indicates that the multivariate approximation problem is reduced to an equivalent univariate problem.

It is not always possible to obtain the characteristic PDF $f_S(s)$ in closed-form. However, an N dimensional SIRV with uncorrelated elements can be expressed in generalized spherical coordinates R, θ, and ϕ_k for $k = 1, \ldots, N-2$, where the PDF of R is given by

$$f_R(r) = \frac{r^{N-1}}{2^{(N/2)-1}\Gamma(N/2)} h_N(r^2) u(r) \tag{15.53}$$

The angles θ and ϕ_k are statistically independent of the envelope R and do not vary with the type of SIRV. When $f_S(s)$ is unknown, Equation 15.53 can be used both to generate SIRVs and determine $h_N(q)$.[6]

It is desirable to develop a library of SIRV's for use in approximating unknown clutter-returns. Table 15.1 contains the characteristic PDF's and $h_N(q)'$s of some SIRVs for which analytical expressions are known. For simplicity, the results presented for the Weibull and Chi SIRV are valid only for even N. Additional SIRVs, such as the generalized Rayleigh, generalized Gamma, and Rician, have been developed in Ref. 6.

The discrete Gaussian mixture (DGM) is an SIRV of special interest. Its PDF is a simple finite weighted sum of Gaussian PDF's. It is useful for approximating many other SIRVs, as well as generating unique distributions.

15.2.3. Distribution Approximation Using the Ozturk Algorithm

It is important to suitably model the clutter PDF to obtain improved detection performance of weak signals in non-Gaussian clutter. Ozturk developed a general graphical method for testing whether random samples are statistically consistent with a specified univariate distribution.[10] The Ozturk algorithm is based upon sample order statistics and has two modes of operation. The first mode consists of the goodness-of-fit test. The second mode of the algorithm approximates the PDF

TABLE 15.1
Characteristic PDFs and $h_N(q)$ Functions for Known SIRVs

Marginal PDF	Characteristic PDF $f_S(s)$	$h_N(q)$
Gaussian	$\delta(s-1)$	$e^{-q/2}$
Student t	$\frac{2b}{\Gamma(\nu)2^{\nu}} b^{2\nu-1} s^{-(2\nu+1)} e^{-(b^2/2s^2)} u(s)$	$\frac{2^{N/2} b^{2\nu} \Gamma\left(\nu+\frac{N}{2}\right)}{\Gamma(\nu)(b^2+q)^{(N/2)+\nu}}$
Laplace	$b^2 s e^{-(b^2 s^2/2)} u(s)$	$b^N (b\sqrt{q})^{1-(N/2)} K_{(N/2)-1}(b\sqrt{q})$
K-distributed envelope	$\frac{2b}{2^{\alpha}\Gamma(\alpha)} (bs)^{2\alpha-1} e^{-(b^2 s^2/2)} u(s)$	$\frac{b^N}{\Gamma(\alpha)} \frac{(b\sqrt{q})^{\alpha-(N/2)}}{2^{\alpha-1}} K_{(N/2)-\alpha}(b\sqrt{q})$
Cauchy	$\sqrt{\frac{2}{\pi}} b s^{-2} e^{-(b^2/2s^2)} u(s)$	$\frac{2^{N/2} b \Gamma\left(\frac{N}{2}+\frac{1}{2}\right)}{\sqrt{\pi}(b^2+q)^{(N/2)+(1/2)}}$
Chi envelope $\nu \leq 1$	$\frac{2^{\nu+1} b^{2\nu}}{\Gamma(\nu)\Gamma(1-\nu)} \frac{s^{2\nu-1}}{(1-2b^2 s^2)^{\nu}} u(s) u\left(\frac{1}{b\sqrt{2}} - s\right)$	$\frac{(2)^{N/2} b^{2\nu}}{\Gamma(\upsilon)} \sum_{k=1}^{N/2} \binom{\frac{N}{2}-1}{k-1} q^{\nu-k} b^{N-2k} \frac{\Gamma(k-\nu)}{\Gamma(1-\nu)} e^{-b^2 q}$
Gaussian mixture $w_k > 0, \sum_{k=1}^{K} w_k = 1$	$\sum_{k=1}^{K} w_k \delta(s-s_k)$	$\sum_{k=1}^{K} w_k s_k^{-N} e^{-(q/2s_k^2)}$
Weibull envelope $0 < b < 2$	—	$(-2)^{N/2} e^{-a\sigma^b q^{b/2}} \sum_{k=1}^{N} B_k \frac{(a\sigma^b)^k}{k!} q^{(kb/2)-(N/2)}$, $B_k = \sum_{m=1}^{k} (-1)^m \binom{k}{m} \prod_{i=0}^{M-1} \left(\frac{mb}{2} - i\right)$

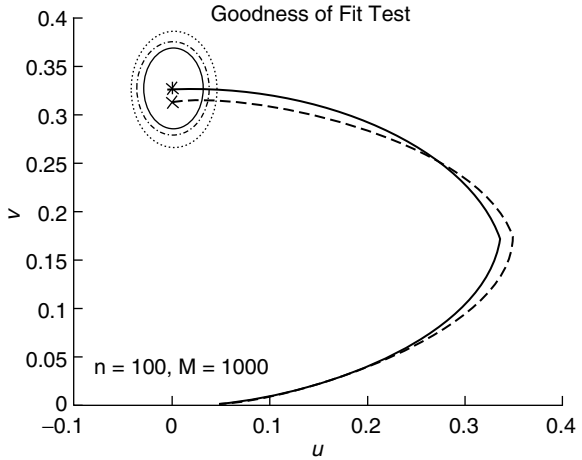

FIGURE 15.7 Linked vector and 90, 95, and 99% confidence intervals for the standard Gaussian distribution.

of the underlying data by using a test statistic generated from the goodness-of-fit test to select from a library of known PDFs.

The goodness-of-fit test is illustrated in Figure 15.7. The solid curve denotes the ideal locus of the null distribution, which is obtained by averaging 1000 Monte Carlo simulations of 100 data samples, where the Gaussian distribution is chosen as the null distribution. The 90, 95, and 99% confidence contours are shown. The dashed curve shows the locus of test data, which is accepted as being Gaussian distributed with significance 0.1. Figure 15.8 shows the scatter of locus

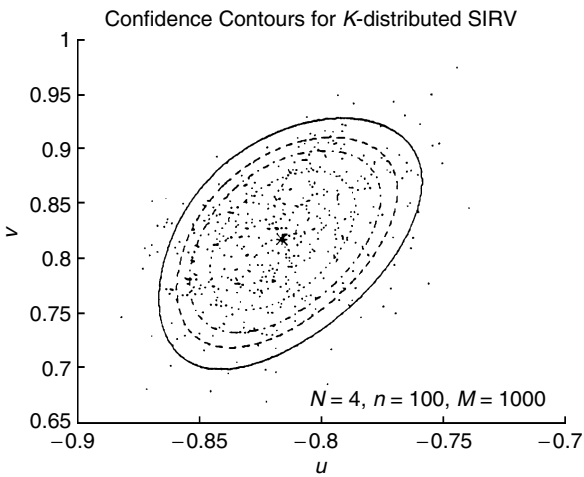

FIGURE 15.8 50, 70, 80, and 90% confidence contours for the *K*-distributed SIRV envelope.

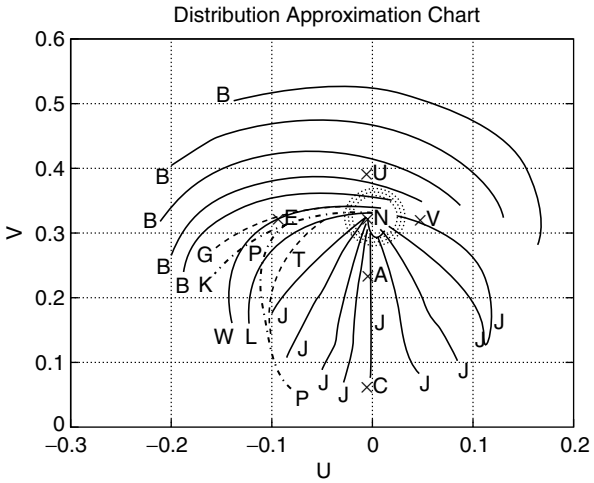

FIGURE 15.9 Ozturk approximation chart for univariate distributions (A = Laplace, B = Beta, C = Cauchy, E = Exponential, G = Gamma, J = SU Johnson, K = *K*-distribution, L = Lognormal, N = Normal, P = Pareto, S = Logistic, T = Gumbel, U = Uniform, V = Extreme Value, W = Weibull).

end points for 1000 simulations of *K*-distributed data. Each end point is obtained from 100 vectors of four components.

An approximation to an unknown distribution can be obtained by examining the location of the end point coordinate. An approximation chart is constructed for a library of PDFs by plotting the end point coordinates for each density in the library. A distribution that does not depend upon a shape parameter will appear as a single point on the approximation chart. Distributions that have a single shape parameter, such as the Weibull or *K*-distributions, will appear as trajectories. Distributions with more than one shape parameter are represented by a family of trajectories. A sample approximation chart for univariate distributions is shown in Figure 15.9 for 100 data samples and 1000 Monte Carlo simulations.

15.2.4. Approximation of SIRVs

The distribution approximation technique described above applies to univariate distributions. It is seen from Equation 15.47 and Equation 15.51 that the characteristic PDF of an SIRV is invariant with respect to the vector dimension N and uniquely determines the SIRV. If the data can be appropriately modeled as SIRV, then the marginal distribution can be used to uniquely distinguish it from all other SIRVs. Since the marginal distribution of an SIRV is univariate, the procedure discussed in Section 15.2.3 can be applied directly. However,

knowledge of the marginal distribution is insufficient to ensure that multivariate data can be modeled as an SIRV.

Multivariate sample data can be rejected as having a particular type of SIRV density if the envelope distribution is not supported by the Ozturk algorithm. In addition, the angle distributions must be checked for consistency. However, the angle distributions are independent of the type of SIRV considered and are useful only for verifying that sample data is not SIRV distributed.

The approximation problem in further complicated since the covariance matrix of the underlying SIRV distribution is usually unknown. The maximum likelihood (ML) estimate of the covariance matrix for a known zero-mean SIRV is given by

$$\hat{\boldsymbol{\Sigma}}_y = \frac{1}{K}\sum_{k=1}^{K}\frac{h_{N+1}(\mathbf{y}_k^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_y\mathbf{y}_k)}{h_N(\mathbf{y}_k^{\mathrm{T}}\hat{\boldsymbol{\Sigma}}_y\mathbf{y}_k)}\mathbf{y}_k\mathbf{y}_k^{\mathrm{T}} \tag{15.54}$$

Since Equation 15.54 depends upon $h_N(q)$, the ML estimate of the covariance matrix cannot be used in the approximation problem. Alternatively, a statistic formed using the well known sample covariance matrix is used in this chapter to select the appropriate approximation for the clutter distribution. This statistic is given by

$$\hat{R} = [(\mathbf{y} - \hat{\mathbf{b}}_y)^{\mathrm{T}} S_y^{-1}(\mathbf{y} - \hat{\mathbf{b}}_y)]^{1/2} \tag{15.55}$$

where $\hat{S}_y$ is the sample covariance matrix, given by

$$\hat{S}_y = \frac{1}{n-1}\sum_{k=1}^{n}(\mathbf{y}_k - \hat{\mathbf{b}}_y)(\mathbf{y}_k - \hat{\mathbf{b}}_y)^{\mathrm{T}} \tag{15.56}$$

and $\hat{\mathbf{b}}_y$ is the sample mean. Approximation charts using the envelope statistic $\hat{R}$ of Equation 15.55 are shown in Figure 15.10 and Figure 15.11 for vector dimensions $N = 2$ and $N = 4$, respectively. The 90% confidence contours for the K-distribution with shape parameter $\alpha_k = 0.4$ are shown on the charts. Surprisingly, the size of the confidence intervals does not significantly increase as the dimension of the SIRV increases. While the sample covariance matrix of Equation 15.56 may be a poor estimate of the actual covariance matrix, the statistic of Equation 15.55 appears to be insensitive to this estimation.

As seen in Figure 15.10 and Figure 15.11, the confidence contours overlap several trajectories on the approximation charts. Therefore, it is possible that any one of several different types of SIRV distributions may be selected by the Ozturk algorithm to approximate an SIRV distributed sample. Figure 15.12 compares the quadratic form PDF for two distributions that fall within the confidence contour shown in Figure 15.11. The locus end point of a K-distributed SIRV with shape parameter $\alpha_k = 0.4$ is marked by a 1 in Figure 15.11. The locus end point of a Weibull SIRV with shape parameter $\alpha_w = 0.8$ is labeled with a two. The close match between theses PDFs, even when their locus end points are separated

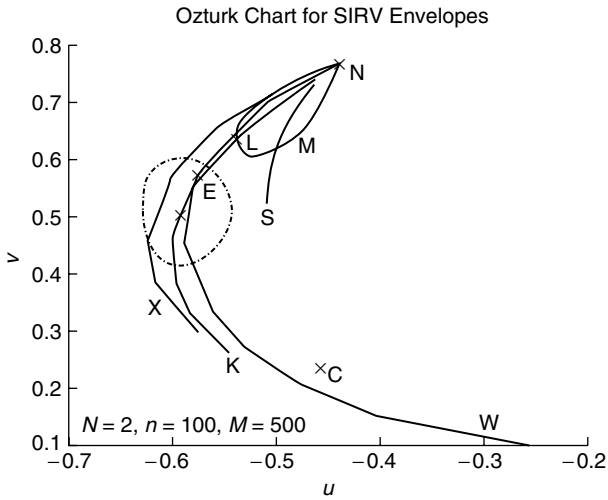

FIGURE 15.10 Approximation chart for SIRV envelope statistic, $N = 2$ (C = Cauchy, E = Exponential Envelope, K = K-distributed Envelope, L = Laplace, M = Discrete Gaussian Mixture, N = Normal, S = Student t, W = Weibull, X = Chi Envelope).

within the confidence contour, suggests similar distributions fall within a particular localized region of the Ozturk chart. Consequently, distributions whose locus end points are contained within a confidence contour are expected to be suitable approximations.

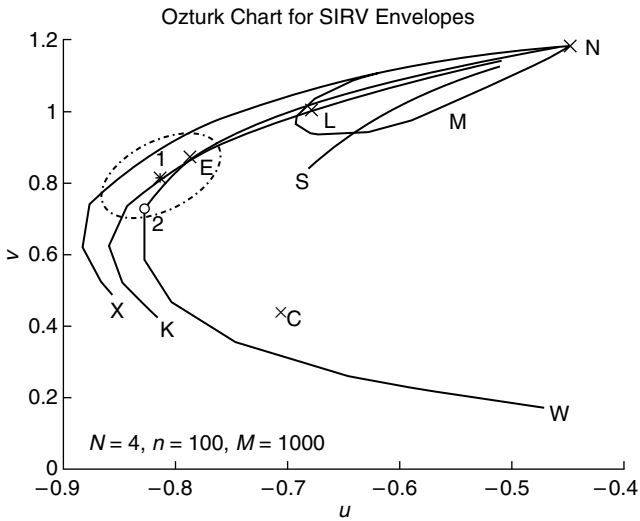

FIGURE 15.11 Approximation chart for SIRV envelope statistic, $N = 4$ (C = Cauchy, E = Exponential Envelope, K = K-distributed Envelope, L = Laplace, M = Discrete Gaussian Mixture, N = Normal, S = Student t, W = Weibull, X = Chi Envelope).

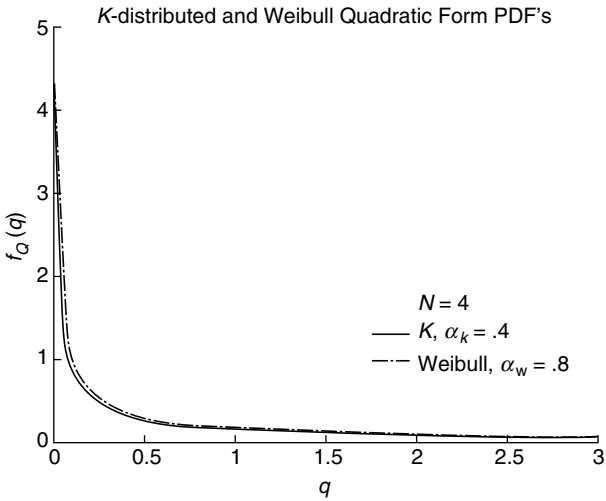

FIGURE 15.12 Comparison of Weibull and K-distributed quadratic form PDFs.

15.2.5. NonGaussian Receiver Performance

The performance of an adaptive detection scheme, which uses the Ozturk PDF approximation algorithm to regularly update the choice of receiver, is evaluated by simulating SIRV clutter. The clutter power is assumed to be much greater than the background noise power for the weak signal problem. Consequently, only the clutter PDF is used to model the total disturbance. The clutter is also assumed to have zero-mean and a known covariance matrix, $\boldsymbol{\Sigma}$. The amplitude of the desired signal is modeled as an unknown complex random variable, which is constant over each P pulse coherent processing interval. The phase of the complex amplitude is assumed to have a $U(0, 2\pi)$ PDF. Thus, the form of the ML estimate for the complex amplitude is the same for all SIRVs, and the generalized likelihood-ratio test (GLRT) is[8]

$$T_{\mathrm{GLRT}}(\tilde{\mathbf{r}}) = \frac{h_{2P}\left[2\left(\tilde{\mathbf{r}}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r} - \dfrac{|\tilde{\boldsymbol{s}}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\mathbf{r}}|^2}{\tilde{\boldsymbol{s}}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\boldsymbol{s}}}\right)\right]}{h_{2P}(2\tilde{\mathbf{r}}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \qquad (15.57)$$

where examples of $h_{2P}(\cdot)$ are given in Table 15.1. The GLRT of Equation 15.57 is formulated in terms of the complex low-pass envelopes of the receive data, $\tilde{\mathbf{r}}$, and known signal pattern, $\tilde{\mathbf{s}}$. Previous investigation has shown there is little or no degradation in performance of the GLRT for the known covariance problem, when compared with the Neyman–Pearson (NP) test.[5,12]

Figure 15.13 compares the two-pulse performance of the adaptive Ozturk-based receiver to several other receivers for a Swerling I target amplitude in K-distributed clutter. The shape parameter is chosen as $\alpha_k = 0.4$, which is within

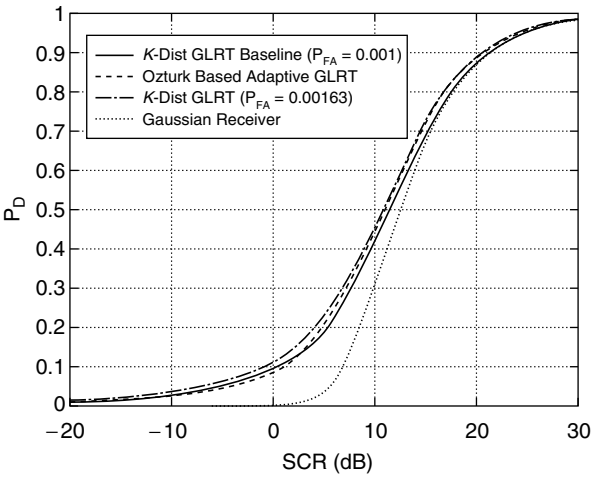

FIGURE 15.13 Adaptive Ozturk-based receiver comparison.

the range of values measured for real data.[1] The performance is evaluated for an identity covariance matrix, and may be interpreted as a function of the SCR at the output of a prewhitening filter. Detection results are obtained by processing 100,000 vector samples of K-distributed clutter. The solid curve shows the baseline detection performance of the K-distributed GLRT designed for 0.001 P_{FA}. The adaptive receiver performance, also indicated in Figure 15.13, is obtained by partitioning the data into 50 intervals of 2000 samples each. The first 100 samples of each interval are processed by the Ozturk algorithm to obtain the data end points shown in Figure 15.14. For each data end point, the corresponding

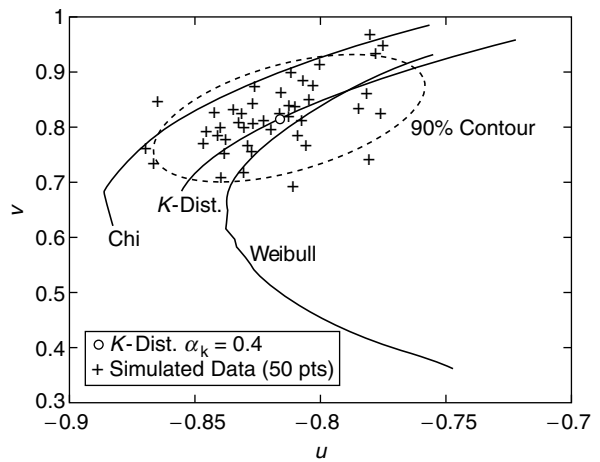

FIGURE 15.14 Ozturk algorithm end point scatter diagram for K-distributed data with $\alpha_k = 0.4$.

2000 sample interval is processed by a GLRT based upon the PDF associated with the closest library end point. While the known covariance matrix is used in the GLRT implementation, the sample covariance matrix for each 100 samples is used in the Ozturk algorithm, as described in Section 15.2.3 and Section 15.2.4.

Performance of the adaptive receiver closely matches the baseline performance and show significant improvement over the Gaussian receiver for SCR values below 10 dB. The measured P_{FA} for the adaptive receiver is 0.00163, which is slightly above the design value. This explains why the PD for the adaptive receiver exceeds that of the baseline receiver at large SCR values. Baseline receiver performance for the higher measured P_{FA} is also included in Figure 15.13 for comparison.

The adaptive receiver processed data associated with all the end points shown in the scatter diagram of Figure 15.14, including those that fell outside the 90% confidence contour. Nonetheless, the localized PD for each interval did not vary significantly form the average value reported in Figure 15.13 for a given SCR.

15.2.6. Concluding Remarks

This chapter provides significant contributions to the development of a novel adaptive non-Gaussian processing technique, which is based on the Ozturk PDF approximation algorithm. New results are presented, which allow the algorithm to adequately approximate multivariate SIRV PDFs from only 100 sample clutter vectors. Then, a simple example is presented for K-distributed clutter with known covariance matrix and 10^{-3} probability of false alarm. A receiver which adaptively processes the data based on the Ozturk PDF approximation has near optimum performance for this example, thus, demonstrating the successful application of the Ozturk algorithm to weak signal detection. Furthermore, the adaptive receiver has significantly better detection performance than the Gaussian receiver at low SCRs, with only a slight increase in the P_{FA}. The above results motivate investigation into application of the adaptive Ozturk algorithm to problems of more practical interest, such as unknown clutter covariance matrix and lower false alarm probabilities.

15.3. Adaptive Ozturk-Based Receivers for Small Signal Detection in Impulsive Nongaussian Clutter

(D. L. Stadelman, A. D. Keckler, and D. D. Weiner)

15.3.1. Introduction

Experimental measurement of radar clutter-returns shows the data may often be non-Gaussian and have a non-Rayleigh envelope distribution, such as the Weibull or K-distribution, particularly for data collected at low grazing angles or high resolution.[1–5] The detection performance of the Gaussian receiver in this environment is significantly less than the optimum non-Gaussian receiver

performance, especially for weak target returns. NonGaussian clutter is often observed to be impulsive or "spiky." Consequently, the threshold of the conventional Gaussian receiver must be raised in order to maintain the desired false alarm rate. This results in a reduction of the probability of detection. In contrast, non-Gaussian receivers contain nonlinearities that limit large clutter spikes and allow a lower threshold to be used, which improves performance for targets with a low SCR.

The nonhomogeneous and nonstationary clutter environment must be monitored to adapt detection algorithms over the surveillance volume. Determination of the appropriate non-Gaussian receiver to use for a region in the volume is based upon knowledge of the probability density function (PDF) for the received clutter data from that region. Thus, multivariate, non-Gaussian models which incorporate the pulse-to-pulse correlation of the clutter data are needed to describe the joint PDF of the received data. Furthermore, a means of choosing a particular one of these joint PDFs that sufficiently approximates the unknown, underlying PDF of the clutter data is required. The changing nature of the clutter environment limits the number of samples available to this PDF approximation method.

The application of SIRV models to many non-Gaussian clutter environments has both empirical and theoretical support. First, the SIRV model is equivalent to the compound clutter model, which is found to be an excellent fit to real sea clutter data in many instances.[2] Second, the SIRV PDF for radar clutter is derived from a generalization of the central limit theorem in which the number of scatterers in a range-azimuth cell is assumed to be a random variable.[6] The class of SIRVs includes many distributions of interest, such as the Gaussian, Weibull, Rician, and K-distributed, and has several properties which facilitate development and implementation of optimal receivers.[6–9]

The Ozturk algorithm[10] is a very efficient method for obtaining an approximation to an unknown PDF, requiring only about 100 samples from the unknown distribution. The efficiency of the Öztürk algorithm has prompted much investigation of its application to the non-Gaussian radar problem, particularly with regard to the SIRV clutter models.[7,11–13] Since the algorithm uses only about 100 points to approximate a PDF, few of these points are expected to fall within the tail region of the unknown PDF. Thus, while the Öztürk algorithm does a good job in selecting a PDF which closely approximates the body to the underlying PDF of the data, this approximating PDF is not expected to adequately match the tail behavior of the unknown PDF.

The low false alarm rates desired in radar applications result in receiver thresholds which fall in the tail region of the PDF of the receiver output statistic. Improper selection of this threshold yields an unacceptably high false alarm rate or a severe degradation of target detection capability. The optimal receiver for detection in Gaussian clutter is a matched filter, which is linear. Consequently, points in the tail region of the PDF for the output of the matched filter are generated by data in the tail region of the clutter PDF at the receiver input. The nonlinear behavior of the non-Gaussian receivers and the applicability

of the Öztürk algorithm to the radar detection problem in regard to threshold selection in an area of much interest.

An analysis is presented which shows the nonlinear nature of the non-Gaussian receivers causes most false alarms to arise from a small percentage of points within the body of the PDF for the received clutter data in some important cases of interest. An easily understood graphical representation is developed and illustrates the conditions that must be satisfied for the Öztürk-based receiver to control the false alarm rate in an acceptable manner.

15.3.2. Summary of the SIRV Model

A brief review of the significant properties of SIRVs is presented.[6,7,14] Any SIRV, $\mathbf{X}$, with zero-mean, uncorrelated components has the Gaussian mixture representation, $\mathbf{X} = S\mathbf{Z}$, where $\mathbf{Z}$ is a Gaussian random vector with N, zero-mean, independent components, and S is a real, nonnegative random variable which is independent of $\mathbf{Z}$. The PDF of S, denoted by $f_S(s)$, is called the characteristic PDF of the SIRV and is normalized to mean-square value, $E(s^2) = 1$, without loss of generality. This mixture model admits an interpretation which is consistent with observations often made on real clutter data. Clutter returns from a given range-azimuth cell are usually Gaussian ($\mathbf{Z}$), but the average clutter power level (S^2) varies among cells in the surveillance volume.

Correlation is introduced by the linear transformation, $\mathbf{Y} = \mathbf{AX} + \mathbf{b}$. The transformed vector, $\mathbf{Y}$, is always another SIRV and has mean vector, $\mathbf{b}$, covariance matrix, $\mathbf{\Sigma} = \mathbf{AA}^{\mathrm{T}}$, and the same characteristic PDF as $\mathbf{X}$. The PDF of $\mathbf{Y}$ is obtained form the Gaussian mixture representation as

$$f_{\mathbf{Y}}(\mathbf{y}) = (2\pi)^{-N/2}|\mathbf{\Sigma}|^{-1/2}h_N[(\mathbf{y}-\mathbf{b})^{\mathrm{T}}\mathbf{\Sigma}^{-1}(\mathbf{y}-\mathbf{b})] \tag{15.58}$$

where $h_N(\cdot)$ is a positive, real valued, monotonic decreasing function given by

$$h_N(q) = \int_0^\infty s^{-N}\exp\left(-\frac{q}{2s^2}\right)f_S(s)\mathrm{d}s \tag{15.59}$$

Thus, the PDF of any N-dimensional SIRV is uniquely specified by a covariance matrix, mean vector, and either the characteristic PDF or $h_N(q)$. Since a linear transformation is reversible, preprocessing to whiten the received data can be performed without loss of optimality.

The PDF of the quadratic form, $q = (\mathbf{y}-\mathbf{b})^{\mathrm{T}}\mathbf{\Sigma}^{-1}(\mathbf{y}-\mathbf{b})$, is

$$f_Q(q) = \frac{q^{(N/2)-1}}{2^{N/2}\Gamma(N/2)}h_N(q) \tag{15.60}$$

and has an important implication. Since $h_N(q)$ is unique for each type of SIRV, Equation 15.60 indicates the multivariate PDF for any SIRV is uniquely determined by the univariate PDF of the quadratic form. This significantly reduces the complexity of the PDF approximation required in the practical implementation of optimal, non-Gaussian processing.

The vector, $\mathbf{X}$, of the Gaussian mixture representation is an SIRV if and only if it can be expressed in generalized spherical coordinates in terms of N, statistically independent, random variables. The generalization of spherical coordinates to N-dimensions is not unique.[15,16] The possible generalizations differ only in their definition of the $N-1$ spherical angles. The vector magnitude, R, remains unchanged in all of these coordinate systems. Any of these generalizations may be used to specify an SIRV in spherical coordinates. A coherent radar processes N_p pulses of in-phase and quadrature data samples. Optimum processing requires the specification of a $2N_\mathrm{p}$ dimensional joint PDF. One convenient spherical coordinate transformation which is useful in the analysis of this case is, for $N = 2N_\mathrm{p}$ components,[17]

$$X_1 = R\cos\Theta_1\cos\Phi_1, \qquad X_2 = R\cos\Theta_1\cos\Phi_1 \tag{15.61a}$$

$$\begin{aligned} X_{2k-1} &= R\cos\Theta_k\cos\Phi_k\prod_{i=1}^{k-1}\sin\Phi_i, \\ X_{2k} &= R\sin\Theta_k\cos\Phi_k\prod_{i=1}^{k-1}\sin\Phi_i, \quad k=2,\ldots,N_\mathrm{p}-1 \end{aligned} \tag{15.61b}$$

$$X_{2N_\mathrm{p}-1} = R\cos\Theta_{N_\mathrm{p}}\prod_{i=1}^{N_\mathrm{p}-1}\sin\Phi_i, \quad X_{2N_\mathrm{p}} = R\sin\Theta_{N_\mathrm{p}}\prod_{i=1}^{N_\mathrm{p}-1}\sin\Phi_i \tag{15.61c}$$

with $R^2 = X_1^2 + X_2^2 + \cdots + X_{2N_\mathrm{p}}^2$ and Jacobian, $J = R^{2N_\mathrm{p}-1}\prod_{i=1}^{N_\mathrm{p}-1}(\sin\Phi_i)^{2N_\mathrm{p}-1-2i}\times\cos\Phi_i$. The spherical coordinate variables, R, $R, \Phi_1,\ldots,\Phi_{N_\mathrm{p}-1}$, $\Theta_1,\ldots,\Theta_{N_\mathrm{p}}$, are statistically independent with PDFs,

$$f_R(r) = \frac{r^{2N_\mathrm{p}-1}}{2^{N_\mathrm{p}-1}\Gamma(N_\mathrm{p})}h_{2N_\mathrm{p}}(r^2), \quad 0 \le r < \infty \tag{15.62}$$

$$\begin{aligned} f_{\Phi_k}(\phi_k) &= 2(N_\mathrm{p}-k)\cos\phi_k(\sin\phi_k)^{2(N_\mathrm{p}-k)-1} \\ &0 \le \phi_k \le \frac{\pi}{2}, \quad k=1,\ldots,N_\mathrm{p}-1 \end{aligned} \tag{15.63}$$

$$f_{\Theta_k}(\theta_k) = \frac{1}{2\pi}, \quad 0 \le \theta < 2\pi, \quad k=1,\ldots,N_\mathrm{p} \tag{15.64}$$

15.3.3. The Ozturk Algorithm and SIRV PDF Approximation

Öztürk developed a general graphical method for testing whether random samples are statistically consistent with a specified univariate distribution.[10] The Öztürk algorithm is based upon sample order statistics and has two modes of operation. The first mode performs a goodness-of-fit test. The second mode of the algorithm uses a test statistic generated from the goodness-of-fit test to select an approximation for the PDF of the underlying data from a library of known PDFs.

The goodness-of-fit test is illustrated in Figure 15.15a. The curve on the right denotes the ideal locus of the null distribution. This locus is obtained by

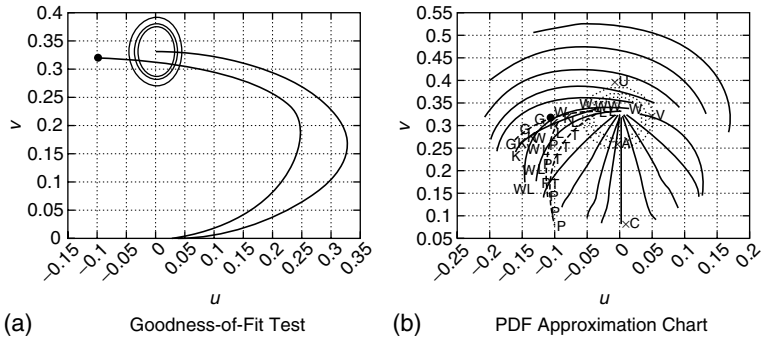

FIGURE 15.15 Ozturk algorithm charts for univariate densities (A = Laplace, B = Cauchy, E = Exponential, G = Gamma, J = SU Johnson, K = *K*-distribution, L = Lognormal, N = Normal, P = Pareto, S = Logisitic, T = Gumbel, U = Uniform, V = Extreme Value, W = Weibull).

averaging 1000 Monte Carlo simulations of 100 data samples. The Gaussian distribution is chosen as the null distribution in this example and the 90, 95, and 99% confidence contours are shown. The other curve is the locus for some non-Gaussian, sample data, which is rejected as Gaussian with significance 0.01.

An approximation chart is constructed for a library of PDFs by plotting the end point coordinates of the goodness-of-fit test locus for each density in the library, as shown in Figure 15.15b. A density which does not depend upon a shape parameter, such as the normal or uniform, appears as a single point on the approximation chart. Densities with a single shape parameter, such as the Weibull or *K*-distributions, appear as trajectories. Densities with more than one shape parameter are represented by a family of trajectories.

An approximation to the unknown PDF is obtained by examining the coordinates at the end point of the locus for the sample data. The endpoint of the goodness-of-fit locus in Figure 15.15a for the sample data is marked by an "o" at coordinates (−0.106, 0.319) on Figure 15.15b. The PDF in the library which is closest to this endpoint is chosen as the approximating PDF. Figure 15.16 compares the approximating PDF to the original PDF (which is a Weibull density). The Oztürk algorithm approximation is a good fit over most of the body of the PDF. However, as Figure 15.16 illustrates, the tail of the approximating PDF is below that of the original PDF.

The distribution approximation technique described above applies to univariate distributions. Since a multivariate SIRV is uniquely specified by its h_N function, Equation 15.17 indicates that the univariate PDF of the quadratic form is also unique for each type of SIRV. Thus, the Öztürk algorithm can be applied to the multivariate PDF approximation problem for SIRVs by using it to approximate either the quadratic form PDF or the envelope PDF in Equation 15.19. Multivariate sample data is rejected as coming from a particular type of SIRV density if the envelope PDF is not supported by the results of the Öztürk algorithm.

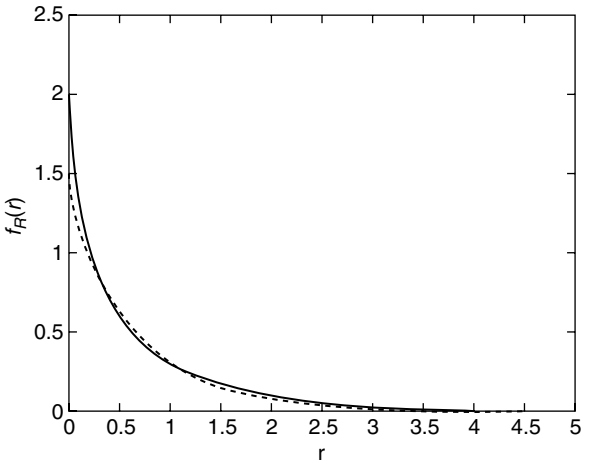

FIGURE 15.16 Ozturk PDF approximation comparison. Original PDF (—) and approximation (– –).

The covariance matrix of the underlying SIRV distribution is usually unknown, which complicates the multivariate PDF approximation problem. The ML estimate of the covariance matrix for a zero-mean SIRV depends upon h_N, which is unknown. Thus, the ML estimate of the covariance matrix cannot be used in the approximation problem. Alternatively, the statistic,

$$\hat{R} = [(\mathbf{y} - \hat{\mathbf{b}}_y)^{\mathrm{T}} \hat{S}_y^{-1} (\mathbf{y} - \hat{\mathbf{b}}_y)]^{1/2}$$

is formulated using the sample covariance matrix,

$$\hat{S}_y = \sum_{k=1}^{n} (\mathbf{y}_k - \hat{\mathbf{b}}_y)(\mathbf{y}_k - \hat{\mathbf{b}}_y)^{\mathrm{T}} / (n-1)$$

where $\hat{\mathbf{b}}_y = \sum_{k=1}^{n} \mathbf{b}_k / n$ is the sample mean. The secondary data vectors, $\mathbf{y}_k$, $k = 1, \ldots n$, are assumed to be homogeneous for the data of interest. Approximation charts are shown in Figure 15.17 for vector dimensions $N = 2$ and $N = 4$, respectively. The 90% confidence contour for the K-distribution with shape parameter, $\nu = 0.4$, is shown on each chart.

As seen in Figure 15.17, the confidence contour overlaps several trajectories on the approximation chart. Therefore, it is possible for any one of several different types of SIRV density functions to be selected by the Öztürk algorithm to approximate an SIRV distributed sample. Densities with locus end points inside the confidence contour are expected to be suitable approximations, since PDFs which are near each other on the approximation chart have similar looking shapes.

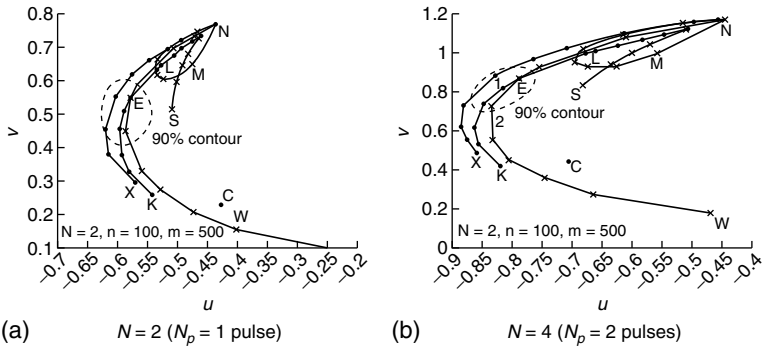

FIGURE 15.17 Ozturk approximation chart for SIRV envelope distributions [C = Cauchy, E = Exponential, K = *K*-distribution, L = Laplace, M = DGM (two components), N = Normal, S = Student *t*, W = Weibull, X = Chi].

15.3.4. NonGaussian SIRV Receivers

Target detection of a reflected radar waveform of N_{p} coherent pulses is described by the binary hypothesis problem, H_0: $\mathbf{r} = \mathbf{d}$ (target absent) vs. H_1: $\mathbf{r} = as(\phi) + \mathbf{d}$ (target present), where the elements of the $2N_{\mathrm{p}}$ dimensional vectors, $\mathbf{r}$, $\mathbf{d}$, and $\mathbf{s}$, are in-phase and quadrature samples of the received data, the disturbance (clutter plus background noise), and the desired signal, respectively. The clutter power is assumed to be much greater than the background noise power in the weak signal detection problem of interest. Consequently, only the clutter PDF is used to determine the statistics of the total disturbance. Signal attenuation and target reflection characteristics are modeled by the target amplitude parameter, a. The initial phase of the received waveform, which is pulse-to-pulse coherent, is represented by ϕ. Complete knowledge of a and ϕ is usually unavailable.

The Neyman–Pearson (NP) receiver, which maximizes PD, for a specified P_{FA}, is optimum for this binary detection problem when a and ϕ are assumed to be random. The NP test can be evaluated in closed-form for some SIRVs in this instance, but this is not typical. Consequently, analyzing the behavior of the optimum NP test is difficult.

However, if the target amplitude and phase remain constant over a single coherent processing interval (CPI) but vary between CPIs, then a generalized likelihood-ratio test (GLRT) receiver is more easily used. Closed-form solutions exist for the GLRT for many SIRVs. Furthermore, for several cases of interest, detection performance of the GLRT is essentially equivalent to that of the optimum NP receiver used in detection of targets with random amplitude and $U(0, 2\pi)$ random phase.[5,17] Thus, study of the GLRT receiver implementation is preferred here.

The linear transformation property for SIRVs allows preprocessing to whiten the received data and normalize the clutter power of each element without loss of optimality when the covariance matrix is known. Hence, the low-pass, complex envelope samples of the clutter are assumed to be uncorrelated with covariance

matrix, $\mathbf{\Sigma} = 2I$. Detection performance in correlated clutter is obtained by adding the processing gain of the whitening filter onto the input SCR before using detection performance curves for the uncorrelated clutter case.

The GLRT is obtained by using the maximum likelihood (ML) estimates for the target amplitude and phase in a likelihood ratio test. The GLRT in this case may be expressed as[6]

$$T_{\mathrm{GLRT}}(\tilde{\mathbf{r}}) = \frac{h_{2N_\mathrm{p}}(\tilde{\mathbf{r}}^H\tilde{\mathbf{r}} - |\tilde{\mathbf{s}}^H\tilde{\mathbf{r}}|^2)}{h_{2N_\mathrm{p}}[\tilde{\mathbf{r}}^H\tilde{\mathbf{r}}]} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{15.65}$$

Here $\tilde{\mathbf{r}}$ is the low-pass, complex envelope, $\tilde{\mathbf{r}} = \mathbf{r}_I + j\mathbf{r}_Q$. The joint PDF of $\tilde{\mathbf{r}}$ is the $2N_\mathrm{p}$ dimensional joint PDF of the elements of $\mathbf{r} = [\mathbf{r}_I\mathbf{r}_Q]^\mathrm{T}$. The low-pass complex envelope of the signal is defined similarly and $\tilde{s}^H\tilde{s} = \|\tilde{s}\|^2 = 1$ is assumed without loss of generality.

Representing the whitened data in the spherical coordinate system of Equation 15.61a–c allows an arbitrary rotation of the system such that the desired signal components are completely represented with only one of the coordinate planes, specifically the one in Equation 15.61a. Consequently, defining $R^2 = R_\mathrm{s}^2 + R_\mathrm{o}^2$, with $R_\mathrm{s} = |\tilde{s}^H\tilde{\mathbf{r}}| = R\cos\Phi_1$ and $R_\mathrm{o} = R\sin\mathbf{\Phi}_1$, simplifies the GLRT in Equation 15.65 to

$$\boxed{T_{\mathrm{GLRT}}(R_\mathrm{s}, R_\mathrm{o}) = \frac{h_{2N_\mathrm{p}}(R_\mathrm{o}^2)}{h_{2N_\mathrm{p}}(R_\mathrm{s}^2 + R_\mathrm{o}^2)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta} \tag{15.66}$$

Some examples of Equation 15.66 are given in Table 15.2 for Gaussian, Student t, and K-distributed SIRVs.

This GLRT depends only on, the envelope of the signal component of the received data (R_s), and on the envelope of the orthogonal component of the received data (R_o). Since R_s and R_o are nonnegative, D_0, the decision region for H_0, and D_1, the decision region for H_1, are located in the first quadrant of the $R_\mathrm{s} - R_\mathrm{o}$ plane. The boundary which separates these decision regions is a curve in this quadrant whose shape is determined by the type of SIRV. Each boundary curve denotes a contour level for the GLRT receiver function, T_{GLRT}, at the selected value of the threshold, η.

15.3.5. Graphical Representation of SIRV Receiver Behavior

The dependence of the GLRT on R_s and R_o in Equation 15.66 motivates the determination of the joint conditional densities, $f_{R_\mathrm{s},R_\mathrm{o}}(r_\mathrm{s}, r_\mathrm{o}|H_0)$ and $f_{R_\mathrm{s},R_\mathrm{o}}(r_\mathrm{s}, r_\mathrm{o}|H_1)$. The P_FA and P_D associated with a particular threshold equal the volumes under these two PDFs, respectively, within the decision region, D_1.

Under hypothesis, H_0, there is no signal component in the received data. Only SIRV interference is present. Consequently, the spherical coordinate random variables, R and $\mathbf{\Phi}_1$, are statistically independent with PDFs,

$$f_R(r) = \frac{r^{2N_\mathrm{p}-1}}{2^{N_\mathrm{p}-1}\Gamma(N_\mathrm{p})} h_{2N_\mathrm{p}}(r^2), \quad 0 \le r < \infty \tag{15.67}$$

TABLE 15.2
Summary of Results for Gaussian, Student *t*, and *K*-distributed SIRVs. (The Shape Parameter is ν, the Scale Parameter is b, and γ is the Threshold Corresponding to P_{FA}. The Dimension is $N = 2N_P$)

	Gaussian	Student t ($\nu > 0$)	*K*-Distribution
$h_{2N_p}(q)$	$\exp\left(-\frac{q}{2}\right)$	$\frac{2^{N_p} b^{2\nu} \Gamma(\nu + N_p)}{\Gamma(\nu)(b^2+q)^{\nu+N_p}}$	$\frac{b^{2N_p}}{\Gamma(\nu)} \frac{(b\sqrt{q})^{\nu-N_p}}{2^{\nu-1}} K_{N_p-\nu}(b\sqrt{q})$
$T_{GLRT}(r_s, r_o)$	r_s	$\frac{r_s^2}{b^2+r_o^2}$	$\left(\frac{r_o}{\sqrt{r_s^2+r_o^2}}\right)^{\nu-N_p} \frac{K_{N_p-\nu}(br_o)}{K_{N_p-\nu}(b\sqrt{r_s^2+r_o^2})}$
$f_{R_s,R_o\|H_0}(r_s, r_o\|H_0)$	$\frac{r_s r_o^{2N_p-3}}{2^{N_p-2}\Gamma(N_p-1)} \exp\left(-\frac{r_s^2+r_o^2}{2}\right)$	$\frac{4b^{2\nu}\Gamma(\nu+N_p)}{\Gamma(N_p-1)\Gamma(\nu)} \frac{r_s r_o^{2N_p-3}}{(b^2+r_s^2+r_o^2)^{\nu+N_p}}$	$\frac{8b^{2N_p} r_s r_o^{2N_p-3} K_{N_p-\nu}(b\sqrt{r_s^2+r_o^2})}{2^{N_p+\nu}\Gamma(N_p-1)\Gamma(\nu)(b\sqrt{r_s^2+r_o^2})^{N_p-\nu}}$
$f_R(r)$	$\frac{r^{2N_p-1}}{2^{N_p-1}\Gamma(N)} \exp\left(-\frac{r^2}{2}\right)$	$\frac{2b^{2\nu}\Gamma(\nu+N_p)}{\Gamma(\nu)\Gamma(N_p)} \frac{r^{2N_p-1}}{(b^2+r^2)^{\nu+N_p}}$	$\frac{4b^{\nu+N_p} r^{\nu+N_p-1}}{2^{\nu+N_p}\Gamma(N_p)\Gamma(\nu)} K_{N_p-\nu}(br)$
$f_{R\|FA}(r\|FA)$	$\frac{r(r^2-\gamma^2)^{N_p-1}}{P_{FA}2^{N_p-1}\Gamma(N_p)} \exp\left(-\frac{r^2}{2}\right)$ for $r > \gamma$ (0 elsewhere)	$\frac{2b^{2\nu}\Gamma(\nu+N_p)}{P_{FA}(\gamma+1)^{N_p-1}\Gamma(\nu)\Gamma(N_p)} \frac{r(r^2-\gamma b^2)}{(b^2+r^2)^{\nu+N_p}}$ for $r > b\sqrt{\gamma}$ (0 elsewhere)	(not available)

and

$$f_{\Phi_1}(\phi_1) = 2(N_p - 1)\cos(\phi_1)\sin^{(2N_p-3)}(\phi_1), \quad 0 \le \phi_1 \le \frac{\pi}{2} \tag{15.68}$$

for a $2N_p$ dimensional SIRV. The joint conditional PDF is easily obtained from

$$f_{R,\Phi}(r,\phi) = f_R(r)f_\Phi(\phi) = \frac{r^{2N_p-1}}{2^{N_p-2}\Gamma(N_p-1)}h_{2N_p}(r^2)\cos(\phi)\sin^{(2N_p-3)}(\phi) \tag{15.69}$$

via the variable transformations, $R_s = R\cos\Phi_1$ and $R_o = R\sin\Phi_1$, with Jacobian, $J(R,\Phi_1) = R^{-1}$, as

$$f_{R_s,R_o}(r_s,r_o|H_0) = f_{R,\Phi_1}(r,\phi_1)|J(r,\phi_1)|\Big|_{\substack{r=\sqrt{r_s^2+r_o^2}\\ \cos(\phi_1)=r_s/r\\ \sin(\phi_1)=r_o/r}} = \frac{r_s r_o^{2N_p-3}h_{2N_p}(r_s^2+r_o^2)}{2^{N_p-2}\Gamma(N_p-1)}. \tag{15.70}$$

Examples of this conditional probability density function are given in Table 15.2 for the Gaussian, Student t, and K-distributed SIRVs.

Under hypothesis, H_1, the signal coordinate envelope, R_s, contains both SIRV interference and a signal component. Consequently, R and Φ_1 are no longer statistically independent random variables and the conditional joint PDF, $f_{R_s,R_o}(r_s,r_o|H_1)$, is not as easily obtained. However, for a given value of the target amplitude, a, the conditional PDF is found to be[17]

$$f_{R_s,R_o}(r_s,r_o|a,H_1) = \frac{r_o^{2N_p-3}r_s}{2^{N_p-2}\Gamma(N_p-1)}G\left(\frac{r_s^2+r_o^2+a^2}{2}\right) \tag{15.71}$$

where $G(k)$ is the Laplace transform, $G(k) = \mathscr{L}\{(1/2)t^{N_p-(3/2)}f_S(t^{(-1/2)})I_0(ar_s t)\}$.

Combined contour plots of $f_{R_s,R_o}(r_s,r_o|H_0)$ and $T_{GLRT}(r_s,r_o)$ are shown in Figure 15.18 to Figure 15.20 for Gaussian, Student t, and K-distributed SIRV examples, respectively. The closed contours correspond to the conditional PDF and the open contours correspond to the GLRT receiver. Gaussian GLRT contours are shown in Figure 15.18 for thresholds associated with P_{FA} values of 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, and 10^{-5}. The contours are shown in Figure 15.19 for the Student t GLRT are associated with the same P_{FA} values, with a Student t threshold given by $\gamma = P_{FA}^{1/(\nu+N_p-1)} - 1$. The dependence of the Student t GLRT threshold on the dimension, $2N_p$, explains the change in the contours as N_P increases, even though the GLRT given in Table 15.2 does not depend on N_p. The contours shown in Figure 15.20 for the K-distributed GLRT correspond to arbitrary, unknown P_{FA} values.

These combined contour plots give a qualitative indication of how the received data, (r_s,r_o), maps into the decision regions for the test statistic of the receiver output. This is illustrated further by the scatter plot of 20,000 received

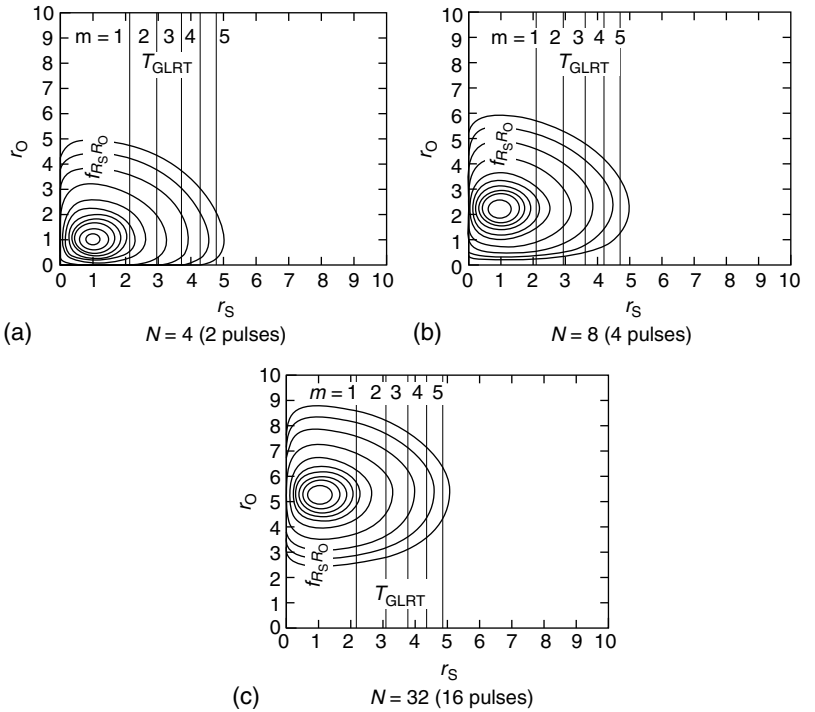

FIGURE 15.18 Contour plots of T_{GLRT} and $f_{R_s,R_o|H_0}(r_s, r_o|H_0)$ for a Gaussian SIRV. Numbered GLRT contours correspond to $P_{FA} = 10^{-m}$, where m is the number on the curve.

data points shown in Figure 15.21a for simulated, Student t, SIRV clutter with $\nu = 2$ and $N_p = 4$. The GLRT contours of Figure 15.19b are also indicated on this figure. The mapping of received data points into D_1, the decision region for H_1, when H_0 is true is of particular interest, since these are the points which cause false alarms. This is illustrated in the scatter plot of Figure 15.21b, where only received data points which cause a false alarm are plotted. The contour shown in Figure 15.21b, corresponds to $P_{FA} = 0.01$ for the same Student t SIRV clutter, so a sample of 100,000 received data points is simulated to obtain approximately 1000 points on this scatter plot. This mapping is determined by the Öztürk algorithm approximation of $f_R(r)$. However, as previously described, the Öztürk algorithm provides a good approximation to the body of the PDF, but this approximating PDF is not expected to adequately match the tail region. Consequently, the Öztürk approximation for $f_R(r)$ must adequately characterize the distribution of data points in Figure 15.21b if a receiver and threshold selection based on this approximation is to control the false alarm rate without a significant loss in detection performance.

When H_0 is true and the received data point maps into D_1, a false alarm occurs. The conditional distribution of the magnitude, R, of these data points is

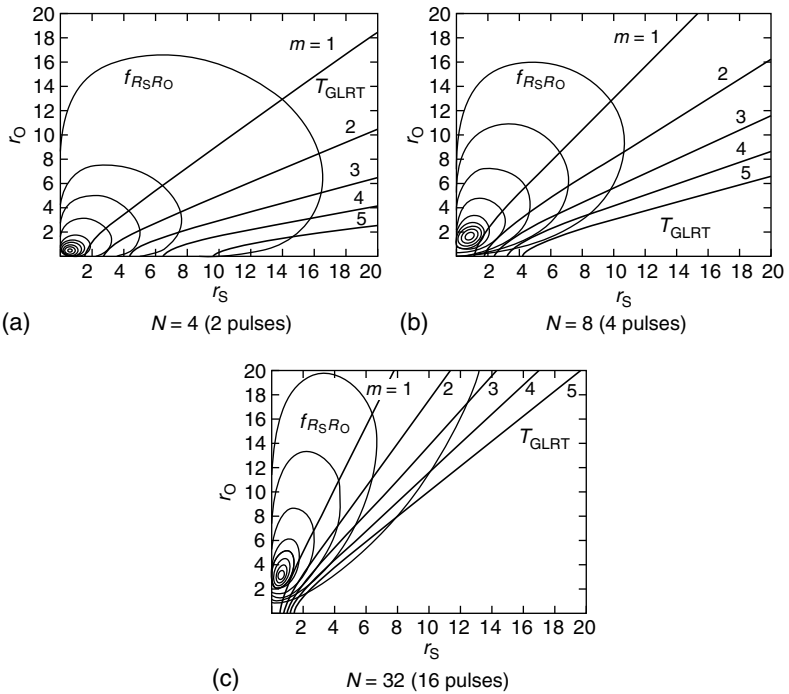

FIGURE 15.19 Contour plots of T_{GLRT} and $f_{R_s,R_o|H_0}(r_s, r_o|H_0)$ for a Student t SIRV with $\nu = 2$. Numbered GLRT contours correspond to $P_{\mathrm{FA}} = 10^{-m}$, where m is the number on the curve.

denoted as $F_{R|\mathrm{FA}}(r|a \text{ false alarm occurs})$ or $F_{R|\mathrm{FA}}(r|\mathrm{FA})$ and is given by

$$F_{R|\mathrm{FA}}(r|\mathrm{FA}) = \frac{\Pr\{R \le r \cap \mathrm{FA}\}}{\Pr\{\mathrm{FA}\}} \tag{15.72}$$

The denominator of this expression is the false alarm probability, P_{FA}. The conditional PDF is obtained from

$$f_{R|\mathrm{FA}}(r|\mathrm{FA}) = \frac{\mathrm{d}}{\mathrm{d}r} F_{R|\mathrm{FA}}(r|\mathrm{FA}) = \frac{\dfrac{\mathrm{d}}{\mathrm{d}r}\Pr\{R \le r \cap \mathrm{FA}\}}{P_{\mathrm{FA}}} \tag{15.73}$$

Determination of $f_{R|\mathrm{FA}}(r|\mathrm{FA})$ is closed-form is possible for the Student t and Gaussian SIRV cases.[17] These results are given in Table 15.2 and examples are plotted in Figure 15.22 and Figure 15.23. When a closed analytical form for $f_{R|\mathrm{FA}}(r|\mathrm{FA})$ cannot be found, such as for the K-distributed SIRV, a histogram of simulated data is sufficient. Figure 15.24 and Figure 15.25 compare histograms of $f_{R|\mathrm{FA}}(r|\mathrm{FA})$ with $f_R(r)$ for $P_{\mathrm{FA}} = 0.01$ and P_{FA} 0.001 in the two-pulse and 16 pulse cases, respectively.

A comparison of the conditional density, $f_{R|\mathrm{FA}}(r|\mathrm{FA})$, with the density function, $f_R(r)$, is very enlightening. It shows that for some important cases there

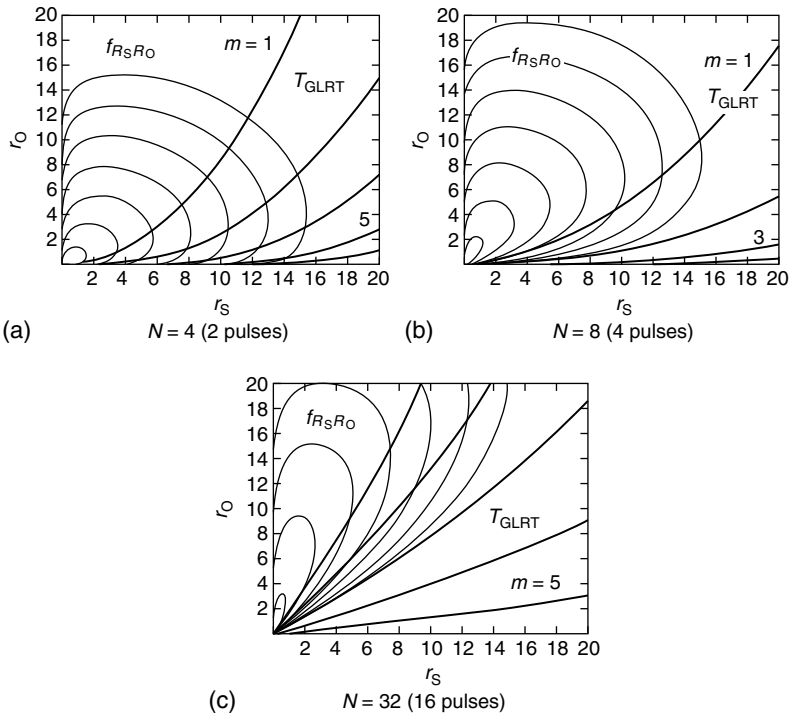

FIGURE 15.20 Contour plots of T_{GLRT} and $f_{R_s,R_o|H_0}(r_s, r_o|H_0)$ for a K-distributed SIRV with $\nu = 0.4$. Numbered GLRT contours correspond to $P_{\mathrm{FA}} = 10^{-m}$, where m is the number on the curve.

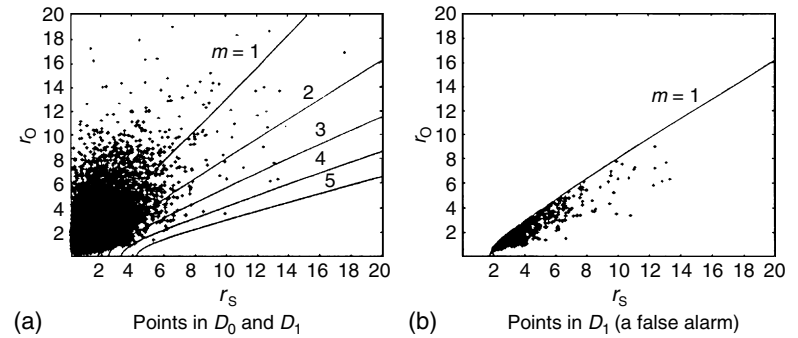

FIGURE 15.21 Scatter plots of (r_s, r_o) for Student t SIRV clutter with $\nu = 0.4$ and $N = 4$ (2 pulses). Numbered GLRT contours correspond to $P_{\mathrm{FA}} = 10^{-m}$, where m is the number on the curve.

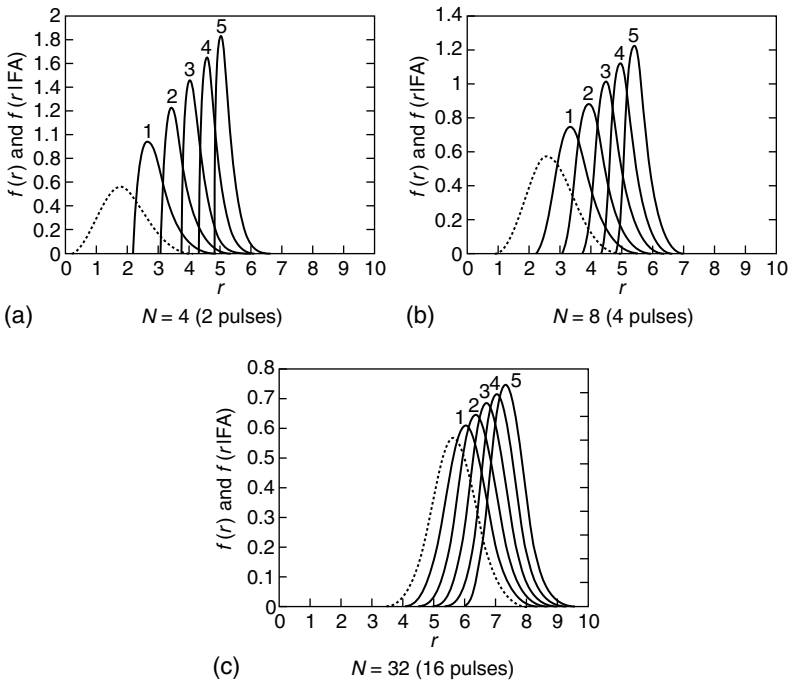

FIGURE 15.22 Comparison of the $f_R(r)$ envelope PDF (– –) and $f_{R|\mathrm{FA}}(r|\mathrm{FA})$ (—) for a Gaussian SIRV. $P_{\mathrm{FA}} = 10^{-m}$ for $m = 1, 2, 3, 4$, and 5, as indicated by the number on each curve.

is significant overlap in the body regions of these two PDFs. Furthermore, the comparisons indicate that the overlap increases as P_{FA} increases and as N_{D}, the number of pulses, increases. These trends are also indicated by the histogram comparisons shown in Figure 15.24 and Figure 15.25 for simulated, K-distributed data. A comparison between different SIRV types also shows that the overlap tends to be greater as the density is more non-Gaussian. There is more overlap for the K-distribution than in the Student t case for $N_{\mathrm{p}} = 2$ and $P_{\mathrm{FA}} = 0.001$. The Student t case also shows slightly more overlap than the Gaussian case in this same example. The performance of an adaptive Öztürk based receiver is evaluated for this K-distributed case.

15.3.6. Adaptive Ozturk-Based Receiver

The performance of an adaptive detection scheme, which uses the Öztürk PDF approximation algorithm to regularly update the choice of GLRT receiver, is evaluated by simulating SIRV clutter. The clutter power is assumed to be much greater than the background noise power for the weak signal problem. Consequently, only the clutter PDF is used to model the total disturbance.

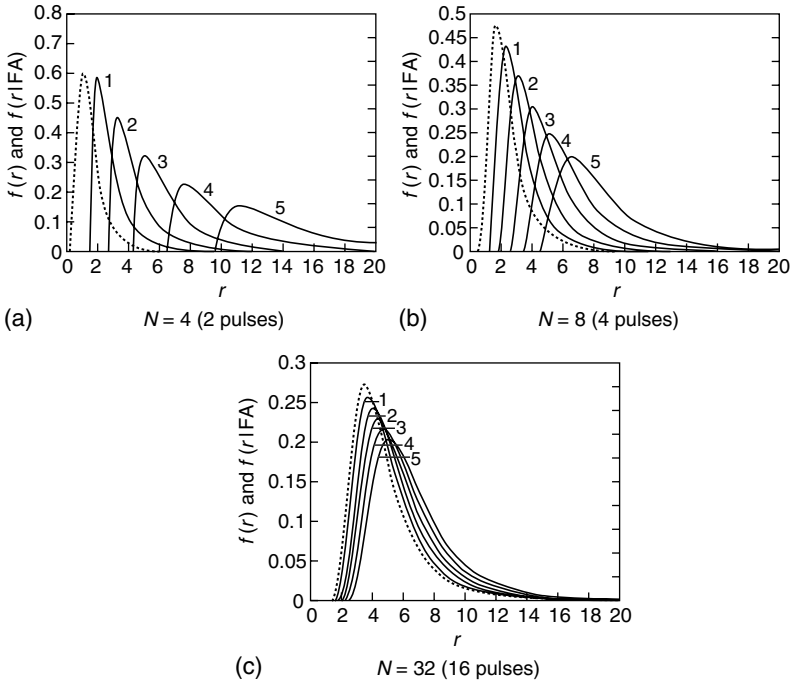

FIGURE 15.23 Comparison of the $f_R(r)$ envelope PDF (– –) and $f_{R|\mathrm{FA}}(r|\mathrm{FA})$ (—) for a Student t SIRV with $\nu = 2$. $P_{\mathrm{FA}} = 10^{-m}$ for $m = 1$, 2, 3, 4, and 5, as indicated by the number on each curve.

The clutter is also assumed to be zero-mean with identity covariance matrix. The target amplitude is modeled as an unknown random variable which is constant over each N_{p} pulse coherent processing interval. The initial target phase is assumed to be uniformly distributed on $(0, 2\pi)$.

Figure 15.26 compares the two-pulse performance of the adaptive Öztürk-based receiver to several other receivers for a Swerling I target amplitude in K-distributed clutter. The shape parameter is chosen as $\nu = 0.4$, which is within the range of values measured for real data.[1] The performance is evaluated for an identity covariance matrix, and may be interpreted as a function of the SCR at the output of a prewhitening filter. Detection results are obtained by processing 100,000 vector samples of K-distributed clutter. The solid curve shows the baseline detection performance of the K-distributed GLRT designed for $P_{\mathrm{FA}} = 0.001$. The adaptive receiver performance, also indicated in Figure 15.26, is obtained by partitioning the data into 50 intervals of 2000 samples each. The first 100 samples of each interval are processed by the Öztürk algorithm to obtain the data end points shown in Figure 15.27. For each data end point, the corresponding 2000 sample interval is processed by a GLRT based upon the PDF associated with the closest library end point. While the known

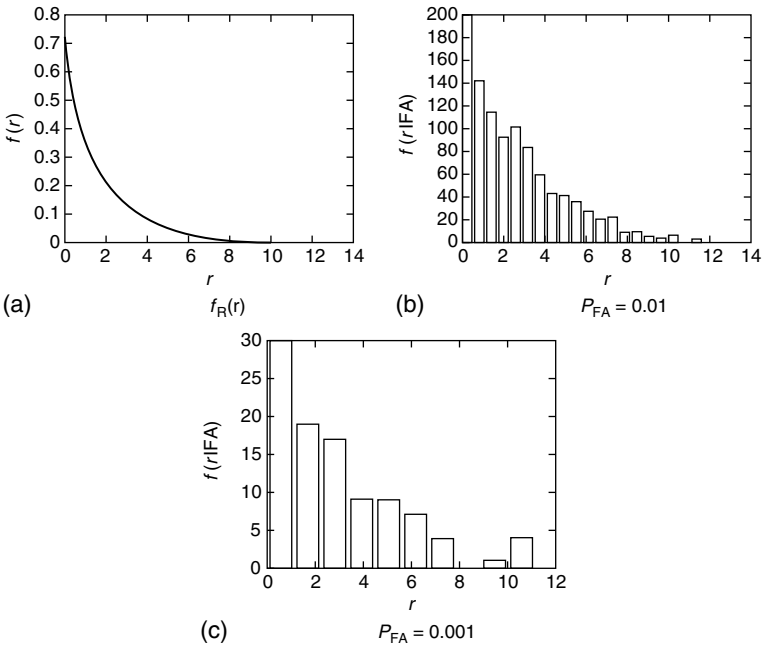

FIGURE 15.24 Conditional PDF, $f_{R|\mathrm{FA}}(r|\mathrm{FA})$, for a K-distributed SIRV with $N = 4$ (2 pulses) and $\nu = 0.4$.

covariance matrix is used in the GLRT implementation, the sample covariance matrix for each 100 samples is used in the Öztürk algorithm.

Performance of the adaptive receiver closely matches the baseline performance and shown significant improvement over the Gaussian receiver for SCR values below 10 dB. The measured P_{FA} for the adaptive receiver is 0.00163, which is slightly above the design value. This explains why the probability of detection (PD) for the adaptive receiver exceeds that of the baseline receiver at larger SCR values. Baseline receiver performance for the higher measured P_{FA} is also included in Figure 15.26 for comparison.

The adaptive receiver processed data associated with all the end points shown in the scatter diagram of Figure 15.27, including those that fell outside the 90% confidence contour. Nonetheless, the localized PD for each interval did not vary significantly from the average value reported in Figure 15.26 for a given SCR.

15.3.7. Conclusions

The performance potential of an adaptive Öztürk-based receiver for detection of weak targets in non-Gaussian SIRV clutter is demonstrated for a K-distributed example. The false alarm rate of this receiver is slightly higher than the design

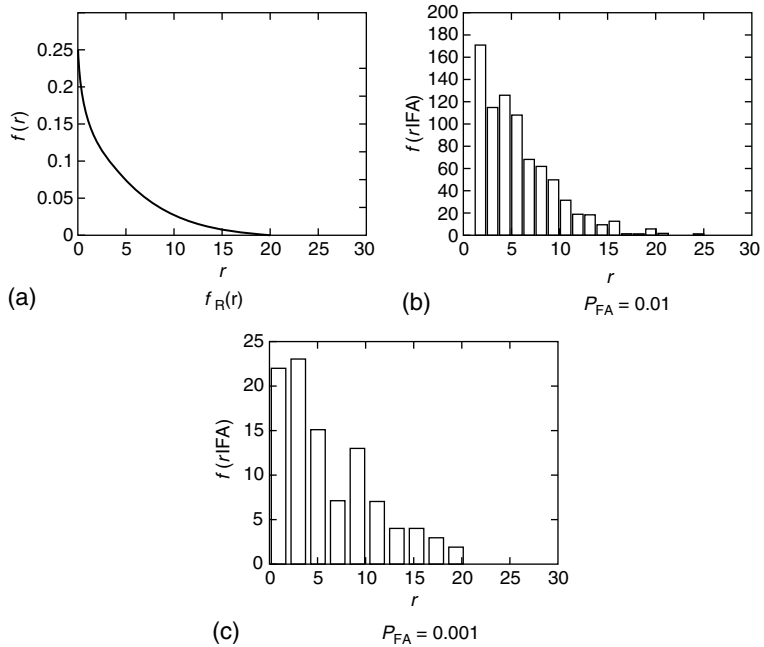

FIGURE 15.25 Conditional PDF, $f_{R|\mathrm{FA}}(r|\mathrm{FA})$, for a K-distributed SIRV with $N = 32$ (16 pulses) and $\nu = 0.4$.

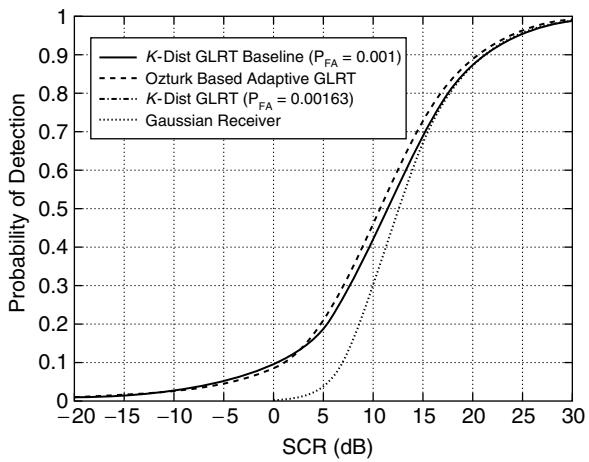

FIGURE 15.26 Adaptive Ozturk-based receiver comparison for K-distributed data with $\nu = 0.4$ and $N_{\mathrm{p}} = 2$.

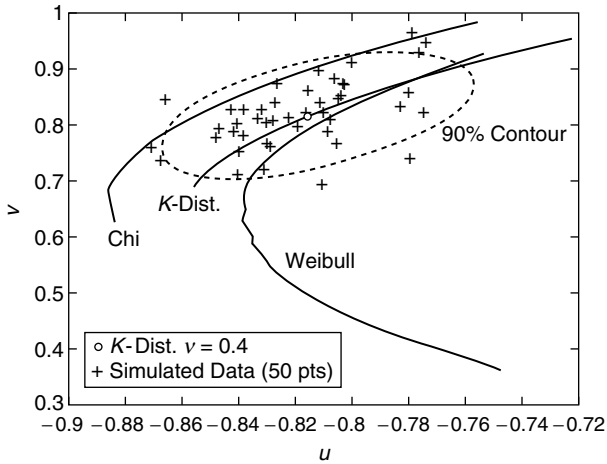

FIGURE 15.27 Ozturk algorithm end point scatter diagram for K-distributed data with $\nu = 0.4$ and $N_p = 2$.

value of 0.001. However, it is of the same order of magnitude and preserves the target detection capability. Furthermore, the graphical representation which is developed is useful for determining when the adaptive Öztürk-based receiver is expected to perform close to optimum. This representation provides guidelines based on P_{FA}, number of pulses, and non-Gaussianity of the data. It indicates that for a relatively small number of pulses, the Öztürk algorithm is able to adequately characterize the PDF of the received data, particularly the data which causes false alarms to occur. Thus, an appropriate receiver and threshold can be adaptively selected. The graphical contour representation of the GLRT is also useful in the analysis and specification of limiter approximations to the optimum non-Gaussian receivers.[17]

15.4. EFFICIENT DETERMINATION OF THRESHOLDS VIA IMPORTANCE SAMPLING FOR MONTE CARLO EVALUATION OF RADAR PERFORMANCE IN NONGAUSSIAN CLUTTER

(D. L. STADELMAN, D. D. WEINER, AND A. D. KECKLER)

15.4.1. Introduction

Interest in the detection of targets in correlated, non-Gaussian radar clutter environments has led to significant interest in the multivariate spherically random vector (SIRV) clutter model. The SIRV model[1,2] describes multivariate generalizations of many non-Gaussian distributions which are commonly used as statistical fits to real clutter data. It includes the K-distribution, which is commonly used to model sea clutter[3] and SAR clutter,[4] the Weibull distribution,

and the Student t distribution, which is also related to statistical models used for SAR data.[5] Several optimal and near-optimal detectors are available for the SIRV model. These include the generalized likelihood-ratio test (GLRT),[6] adaptive matched filter,[7,8] and the parametric adaptive matched filter.[9]

Due to the nonlinear nature of the non-Gaussian receivers and the additional mathematical complexity of the non-Gaussian distributions, it is frequently impossible to obtain closed form analytical results for the threshold, P_{FA}, and PD. This is especially the case for adaptive receivers which estimate the unknown covariance matrix of the clutter data. Consequently, Monte Carlo simulation is used to obtain estimates of performance.

Determination of a receiver threshold is the critical aspect of the radar detection performance evaluation. Conventional Monte Carlo simulation typically requires $100/P_{\mathrm{FA}}$ independent trials to obtain a suitable estimate of the threshold. For example, estimating a threshold for $P_{\mathrm{FA}} = 10^{-7}$ requires one billion samples of the receiver test statistic. The computational burden of calculating these samples for adaptive non-Gaussian receivers often prevents practical use of this technique. This problem is confirmed in the literature, where performance results from Monte Carlo simulation of adaptive non-Gaussian receivers are typically presented for false alarm probabilities on the order of 10^{-2} to 10^{-4}.[10,11] These values of P_{FA} exceed practical radar design requirements by several orders of magnitude. Analytical tractability requires assumptions which may not be characteristic of real clutter data.

A very efficient simulation method that uses importance sampling (IS) to estimate the threshold for very small P_{FA} values is developed. A significant reduction in the required number of trials is achieved.

15.4.2. The Complex SIRV Clutter Model

Let the clutter vector, $\mathbf{X} = \mathbf{X}_I + j\mathbf{X}_Q$, be modeled as a complex N-dimensional SIRV with zero-mean, where $\mathbf{X}_I$ and $\mathbf{X}_Q$ are real vectors of the in-phase (I) are quadrature (Q) components of the clutter samples, respectively. The marginal PDFs for the (I, Q) pair of each complex component are assumed to be identical and circularly symmetric.

The PDF of $\mathbf{X}$ is defined as the multivariate PDF for the concatenated, $2N$-dimensional real vector, $\mathbf{X}_r = [\mathbf{X}_I; \mathbf{X}_Q]$. If the covariance matrix of $\mathbf{X}_r$ has the structure,

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Gamma} & -\mathbf{\Phi} \\ \mathbf{\Phi} & \mathbf{\Gamma} \end{pmatrix} \tag{15.74}$$

where $\mathbf{\Gamma}$ is positive definite and $\mathbf{\Phi}$ is skew symmetric ($\mathbf{\Phi} = -\mathbf{\Phi}^{\mathrm{T}}$), then the PDF of the complex SIRV, $\mathbf{X}$, is[1,2]

$$f_{\mathbf{X}}(\mathbf{x}) = \pi^{-N} |\mathbf{\Sigma}|^{-1} h_{2N}(\mathbf{x}^H \mathbf{\Sigma}^{-1} \mathbf{x}) \tag{15.75}$$

where $h_{2N}(q)$ is a monotonic, decreasing function for all N.

The covariance matrix for $\mathbf{X}$ is $\mathbf{\Sigma} = 2(\mathbf{\Gamma} + j\mathbf{\Phi})$, which is Hermitian and positive definite. Therefore, it has an eigenvalue decomposition, $\mathbf{\Sigma} = \mathbf{U}\mathbf{D}\mathbf{U}^H$, where $\mathbf{D}$ is a diagonal matrix of positive eigenvalues and $\mathbf{U}$ is a unitary matrix of the orthonormal eigenvectors of $\mathbf{\Sigma}$. The representation theorem for SIRVs[2] is used to express $\mathbf{X}$ as the scale mixture,

$$\mathbf{X} = V\mathbf{U}\mathbf{D}^{1/2}\mathbf{Z} \tag{15.76}$$

where $\mathbf{Z}$ has the N dimensional, complex normal distribution, $N(\mathbf{0},\mathbf{I}_N)$, and V is a real, nonnegative random variable which is independent of $\mathbf{Z}$ and normalized to $E(V^2) = 1$. The PDF of $\mathbf{X}$ is then

$$f_{\mathbf{X}}(\mathbf{x}) = \int_0^\infty f_{\mathbf{X}|V}(\mathbf{x}|v) f_V(v)\mathrm{d}v \tag{15.77}$$

where $\mathbf{X}|v$ is conditionally Gaussian with PDF, $N(\mathbf{0}, v^2\mathbf{\Sigma})$, and

$$h_{2N}(q) = \int_0^\infty v^{-2N} \mathrm{e}^{-q/v^2} f_V(v)\mathrm{d}v \tag{15.78}$$

is easily obtained by equating Equation 15.75 and Equation 15.77.

A linear transformation, $\mathbf{\Sigma}^{-1/2} = \mathbf{D}^{-1/2}\mathbf{U}^H$, whitens $\mathbf{X}$ and gives

$$\mathbf{X}^H\mathbf{\Sigma}^{-1}\mathbf{X} = (\mathbf{\Sigma}^{-1/2}\mathbf{X})^H(\mathbf{\Sigma}^{-1/2}\mathbf{X}) = \mathbf{X}_{\mathrm{w}}^{\mathrm{T}}\mathbf{X}_{\mathrm{w}} = \|\mathbf{X}_{\mathrm{w}}\|^2 = R^2 \tag{15.79}$$

Hence, from Equation 15.75, the PDF of $\mathbf{X}$ is only a function of R, which is the norm of the whitened vector, $\mathbf{X}_{\mathrm{w}} = \mathbf{\Sigma}^{-1/2}\mathbf{X}$. The PDF of R obtained through a spherical coordinate transformation of $\mathbf{X}_{\mathrm{w}}$ is

$$f_R(r) = \frac{2r^{N-1}}{\Gamma(N)} h_{2N}(r^2), \quad 0 \le r < \infty \tag{15.80}$$

The PDF of V, denoted by $f_V(v)$, is the characteristic PDF of the SIRV and can be used to create SIRVs having specific marginal distributions. For example, a K-distributed SIRV is obtained when V has a gamma density function. The functions $h_{2N}(q)$ and $f_R(r)$ for the Gaussian, Student t, and K-distributed classes of SIRVs are given in Table 15.3, where b is a scale parameter and ν is a shape parameter that characterizes different PDFs within a class of SIRVs.

15.4.3. NonGaussian SIRV Receivers

The target detection problem in non-Gaussian SIRV clutter is formulated as the hypothesis test, $H_0 : \mathbf{r} = \mathbf{x}$ vs. $H_1 : \mathbf{r} = \tilde{a}\mathbf{s} + \mathbf{x}$, where $\mathbf{x}$ is a vector of N complex interference samples, $\mathbf{s}$ is a known signal vector, and $\tilde{a}$ is the unknown complex amplitude of the signal. The interference consists of clutter and additive white noise. However, the clutter-to-noise ratio is assumed large and the total interference is approximated by the SIRV clutter model only.

TABLE 15.3
Summary of Results for Gaussian, Student *t*, and *K*-Distributed SIRVs. (A Condensed Version of Table 15.2)

	Gaussian	Student t ($\nu > 0$)	K-Distribution
$h_{2N}(q)$	e^{-q}	$\dfrac{b^{2\nu}\Gamma(\nu+N)}{\Gamma(\nu)(b^2+q)^{\nu+N}}$	$\dfrac{b^{2N}}{\Gamma(\nu)}\dfrac{(b\sqrt{q})^{\nu-N}}{2^{N+\nu-1}}K_{N-\nu}(b\sqrt{q})$
$f_R(r)$	$\dfrac{2r^{2N-1}}{\Gamma(N)}\mathrm{e}^{-r^2}$	$\dfrac{2b^{2\nu}\Gamma(\nu+N)}{\Gamma(\nu)\Gamma(N)}\dfrac{r^{2N-1}}{(b^2+r^2)^{\nu+N}}$	$\dfrac{4b^{\nu+N}r^{\nu+N-1}}{2^{\nu+N}\Gamma(N)\Gamma(\nu)}K_{N-\nu}(br)$
$f_{R_\mathrm{s},R_\mathrm{o}\mid H_0}(r_\mathrm{s}, r_\mathrm{o}\mid H_0)$	$\dfrac{4r_\mathrm{s}r_\mathrm{o}^{2N-3}}{\Gamma(N-1)}\mathrm{e}^{-(r_\mathrm{s}^2+r_\mathrm{o}^2)}$	$\dfrac{4b^{2\nu}\Gamma(\nu+N)}{\Gamma(N-1)\Gamma(\nu)}\dfrac{r_\mathrm{s}r_\mathrm{o}^{2N-3}}{(b^2+r_\mathrm{s}^2+r_\mathrm{o}^2)^{\nu+N}}$	$\dfrac{8b^{2N}r_\mathrm{s}r_\mathrm{o}^{2N-3}K_{N-\nu}(b\sqrt{r_\mathrm{s}^2+r_\mathrm{o}^2})}{2^{N+\nu}\Gamma(N-1)\Gamma(\nu)(b\sqrt{r_\mathrm{s}^2+r_\mathrm{o}^2})^{N-\nu}}$
$T_{\mathrm{GLRT}}(r_\mathrm{s}, r_\mathrm{o})$	r_s	$\dfrac{r^2}{b^2+r_\mathrm{o}^2}$	$\left(\dfrac{r_\mathrm{o}}{\sqrt{r_\mathrm{s}^2+r_\mathrm{o}^2}}\right)^{\nu-N}\dfrac{K_{N-\nu}(br_\mathrm{o})}{K_{N-\nu}(b\sqrt{r_\mathrm{s}^2+r_\mathrm{o}^2})}$

15.4.3.1. Known Covariance Matrix Case

When the covariance matrix is known, nearly optimal detection performance is obtained with the generalized likelihood-ratio test (GLRT). The GLRT is obtained by using the maximum likelihood estimate (MLE) of the unknown complex signal amplitude, $\tilde{a}$, in the likelihood ratio test. The form of the test for SIRV clutter is[6]

$$T_{\mathrm{GLRT}}(\mathbf{r}) = \frac{h_{2N}(r_{\mathrm{o}}^2)}{h_{2N}(r_{\mathrm{o}}^2 + r_{\mathrm{s}}^2)} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma \tag{15.81}$$

with

$$r_{\mathrm{s}}^2 = \frac{|\mathbf{s}^H \boldsymbol{\Sigma}^{-1} \mathbf{r}|^2}{\mathbf{s}^H \boldsymbol{\Sigma}^{-1} \mathbf{s}} = \frac{|\mathbf{s}_{\mathrm{w}}^H \mathbf{r}_{\mathrm{w}}|^2}{\mathbf{s}_{\mathrm{w}}^H \mathbf{s}_{\mathrm{w}}} \tag{15.82}$$

and

$$r_{\mathrm{o}}^2 = \mathbf{r}^H \boldsymbol{\Sigma}^{-1} \mathbf{r} - r_{\mathrm{s}}^2 = \mathbf{r}_{\mathrm{w}}^H \mathbf{r}_{\mathrm{w}} - r_{\mathrm{s}}^2 = r^2 - r_{\mathrm{s}}^2 \tag{15.83}$$

where $\mathbf{r}_{\mathrm{w}} = \boldsymbol{\Sigma}^{-1/2}\mathbf{r}$ is now whitened. The transformed signal vector, $\mathbf{s}_{\mathrm{w}} = \boldsymbol{\Sigma}^{-1/2}\mathbf{s}$ may be arbitrarily chosen as coincident with a coordinate vector for the whitened data. Consequently, r_{s} is the magnitude of the signal component in the received data and r_{o} is the magnitude of the components orthogonal to the signal in the received data. GLRTs for the Student t and K-distributed SIRV cases are listed in Table 15.3.

Since $T_{\mathrm{GLRT}}(\mathbf{r})$ is dependent on $h_{2N}(q)$, the PDF of the interference must be known. In practice, a suitable approximation to the PDF of the SIRV must be obtained.

A suboptimum test statistic which does not require the PDF of the SIRV clutter is the normalized matched filter (NMF),[7]

$$T_{\mathrm{NMF}}(\mathbf{r}) = \frac{|\mathbf{s}^H \boldsymbol{\Sigma}^{-1} \mathbf{r}|^2}{(\mathbf{s}^H \boldsymbol{\Sigma}^{-1} \mathbf{s})(\mathbf{r}^H \boldsymbol{\Sigma}^{-1} \mathbf{r})} = \frac{|\mathbf{s}_{\mathrm{w}}^H \mathbf{r}_{\mathrm{w}}|^2}{(\mathbf{s}_{\mathrm{w}}^H \mathbf{s}_{\mathrm{w}})(\mathbf{r}_{\mathrm{w}}^H \mathbf{r}_{\mathrm{w}})} = \frac{r_{\mathrm{s}}^2}{r^2} \tag{15.84}$$

15.4.3.2. Unknown Covariance Matrix Case

The covariance matrix of the clutter is usually unknown and an estimate of the covariance matrix is used in the detectors of Equation 15.81 to Equation 15.84 instead. The MLE, $\hat{\boldsymbol{\Sigma}}$, of the SIRV covariance matrix is obtained from a set of independent secondary data vectors, $\{y_1, \ldots, y_K\}$. This estimate satisfies[12]

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{K} \sum_{k=1}^{K} - \frac{h'_{2N}(y_k^H \hat{\boldsymbol{\Sigma}} y_k)}{h_{2N}(y_k^H \hat{\boldsymbol{\Sigma}} y_k)} y_k y_k^H \tag{15.85}$$

which is solved using the expectation-maximization (EM) algorithm. The estimator reduces to the sample covariance matrix when the SIRV is Gaussian. Otherwise, it depends on suitable knowledge of the underlying clutter PDF.

15.4.4. Importance Sampling

The false alarm probability of the decision,

$$T(\mathbf{r}) \underset{H_0}{\overset{H_1}{\gtrless}} \gamma$$

is

$$P_{\mathrm{FA}} = \int_{\gamma}^{\infty} f_T(t)\mathrm{d}t \tag{15.86}$$

where $f_T(t)$ is the PDF of the test statistic, T. This is expressed in terms of the PDF of the input data as

$$P_{\mathrm{FA}} = \int_{\Omega_1(r)} I_\gamma(\mathbf{r}) f_{\mathbf{R}|H_0}(\mathbf{r}|H_0)\mathrm{d}\mathbf{r} \tag{15.87}$$

where $\Omega_1(\mathbf{r})$ denotes the decision region in $\mathbf{r}$ for H_1 and $I_\gamma(\mathbf{r})$ is the indicator function,

$$I_\gamma(\mathbf{r}) = \begin{cases} 1 & \text{if } T(\mathbf{r}) > \gamma \\ 0 & \text{if } T(\mathbf{r}) > \gamma \end{cases} \tag{15.88}$$

The unbiased estimate, $\hat{P}_{\mathrm{FA}}$, obtained from M conventional Monte Carlo trials is the sample mean,

$$\hat{P}_{\mathrm{FA}} = \frac{1}{M}\sum_{k=1}^{M} I_\gamma(\mathbf{r}_k) \tag{15.89}$$

which is approximately $N(0, P_{\mathrm{FA}}/M)$ for large M and $P_{\mathrm{FA}} \ll 1$. The quality of this estimate is characterized by the relative error,

$$e = \frac{P_{\mathrm{FA}} - \hat{P}_{\mathrm{FA}}}{P_{\mathrm{FA}}} \tag{15.90}$$

The relative RMS error is the standard deviation of e, which is $1/\sqrt{MP_{\mathrm{FA}}}$ in this case. Typically, $M = 100/P_{\mathrm{FA}}$ trials are required to obtain good threshold estimates in radar applications. The relative RMS error is then 0.1, which is used as a benchmark to measure the efficiency of importance sampling Monte Carlo simulations.

Importance sampling (IS) is a variance reduction technique that substantially reduces the number of Monte Carlo trials required to obtain accurate threshold estimates.[13,14] The basic idea behind importance sampling is to use random samples with a modified probability density function for the detector input. The modified PDF is selected to generate more threshold crossings at the detector output. Each detector output is then weighted in the computation of $\hat{P}_{\mathrm{FA}}$ to compensate for this modification.

The modified PDF, $f_{\mathbf{R}}^{(m)}(\mathbf{r})$, is introduced into Equation 15.87 as

$$P_{\mathrm{FA}} = \int_{\Omega_1(r)} w(\mathbf{r}) f_{\mathbf{R}}^{(m)}(\mathbf{r}) \mathrm{d}\mathbf{r} \tag{15.91}$$

where $w(\mathbf{r})$ is the weighted indicator function,

$$w(\mathbf{r}) = I_\gamma(\mathbf{r}) \frac{f_{\mathbf{R}|H_0}(\mathbf{r}|H_0)}{f_{\mathbf{R}}^{(m)}(\mathbf{r})} = \begin{cases} \dfrac{f_{\mathbf{R}|H_0}(\mathbf{r}|H_0)}{f_{\mathbf{R}}^{(m)}(\mathbf{r})} & \text{for } T(\mathbf{r}) > \gamma \\ 0 & \text{for } T(\mathbf{r}) > \gamma \end{cases} \tag{15.92}$$

The weighting depends only on the input data and is the ratio of the original PDF to the modified PDF.

The P_{FA}, now interpreted as the average of $w(\mathbf{r})$ when the detector inputs are generated from the modified PDF, is estimated by the sample mean,

$$\hat{P}_{\mathrm{FA,IS}} = \frac{1}{M_{\mathrm{is}}} \sum_{k=1}^{M_{\mathrm{is}}} w(\mathbf{r}_k) \tag{15.93}$$

of M_{is} importance sampling trials. This estimate is unbiased when the modified PDF spans the decision region. It is straightforward to show the variance of the relative error for this estimate is

$$\frac{1}{M_{\mathrm{is}} P_{\mathrm{FA}}^2} \int_{\Omega_1(\mathbf{r})} w(\mathbf{r}) f_{\mathbf{R}|H_0}(\mathbf{r}|H_0) \mathrm{d}\mathbf{r} \tag{15.94}$$

Then, for the same relative RMS error in both estimates,

$$\frac{M_{\mathrm{is}}}{M} = \frac{1}{P_{\mathrm{FA}}} \int_{\Omega_1(\mathbf{r})} w(\mathbf{r}) f_{\mathbf{R}|H_0}(\mathbf{r}|H_0) \mathrm{d}\mathbf{r} \tag{15.95}$$

provides a measure of the effectiveness of the IS technique. The objective is to make this ratio *very* small, such that $M_{\mathrm{is}} \nu M$.

Methods for selecting a modified PDF which realizes this effectiveness have been developed and very successfully implemented in many problems.[13–17] However, their application to a particular detection problem is not necessarily obvious and the optimum modification to minimize the variance may not be easy to find. Consequently, several modified PDFs may need to be tried and the question, of how many trials, M_{is}, should be chosen to achieve an accurate result, is raised.

The importance sampling technique based on a simple variance scaling to modify the input PDF is developed for the SIRV receivers. Confidence in the results is obtained by considering the sample variance of the relative error, known analytical results and PDF tail behavior related to extreme value theory.

15.4.5. Estimation of SIRV Detector Thresholds with Importance Sampling

Importance sampling is first applied to the estimation of thresholds when the covariance matrix of the received clutter data is known. This provides an upper bound on the detection performance obtained when an estimate of the covariance matrix is used. The behavior of the adaptive receivers should be similar to this case when good estimates of the covariance matrix are generated. A limited number of closed-form analytical expressions for the threshold are also available for the known covariance problem.[18] These provide excellent cases with which to evaluate and validate the importance sampling Monte Carlo methods.

The GLRT and NMF receivers described by Equation 15.81 to Equation 15.84 are functions of only the norms, R_s and R_o, of the signal and orthogonal components in the received data. The joint PDF of R_s and R_o for an SIRV with known covariance is[18]

$$f_{R_s,R_o}(r_s, r_o|H_0) = \frac{4r_s r_o^{2N_p-3}}{\Gamma(N_p-1)} h_{2N_p}(r_s^2 + r_o^2) \tag{15.96}$$

This joint density is used as the original input PDF for threshold estimation using importance sampling.

Representative contour plots of $f_{R_s,R_o}(r_s, r_o|H_0)$ and the GLRT are shown in Figure 15.28 and Figure 15.29 for $N = 4$. The decision region for H_1, $\Omega_1(\mathbf{r})$, tends towards larger values of R_s and smaller values of R_o as the threshold is increased. Consequently, the probability mass of the modified PDF for

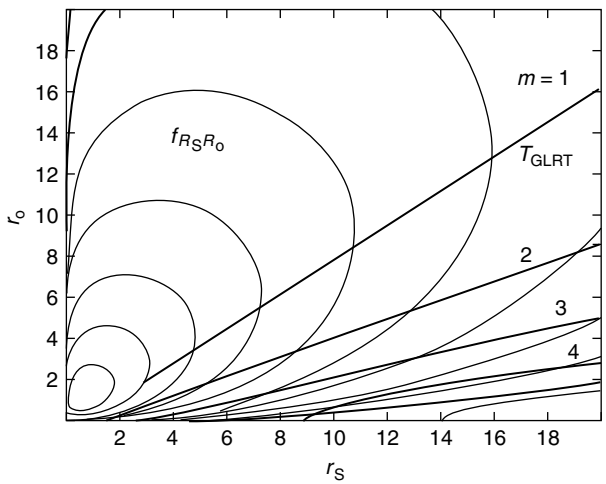

FIGURE 15.28 Student t SIRV contour plots of T_{GLRT} and $f_{R_s,R_o|H_0}(r_s, r_o|H_0)$ with $N = 4$, $\nu = 2$. Numbered GLRT contours correspond to $P_{FA} = 10^{-m}$, where m is the number on the curve.

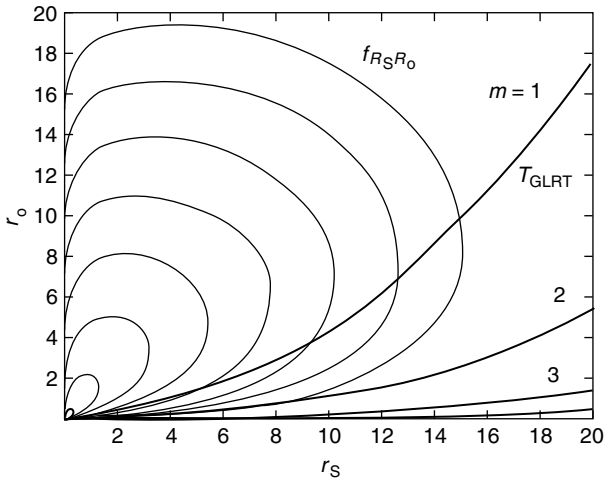

FIGURE 15.29 K-distributed SIRV contour plots of T_{GLRT} and $f_{R_s,R_o|H_0}(r_s, r_o|H_0)$ with $N = 4$, $\nu = 2$. Numbered GLRT contours correspond to $P_{\mathrm{FA}} = 10^{-m}$, where m is the number on the curve.

importance sampling should be shifted to larger values of R_s and smaller values of R_o. Such a shift is obtained from a simple variance scaling of these components. Since the peak of the original distribution occurs near the origin and that region is contributing very little to $\Omega_1(\mathbf{r})$, most of the IS improvement comes from the shift in R_s values. The modified PDF for IS has the form,

$$f^{(m)}_{R_s,R_o}(r_s, r_o|H_0) = \frac{4 r_s r_o^{2N-3} h_{2N}[(r_s^2/k_s^2) + (r_o^2/k_o^2)]}{\Gamma(N-1) k_o^2 k_o^{2(N-1)}} \tag{15.97}$$

where $k_s > 1$ scales the signal component and $k_o \leq 1$ scales the orthogonal term. This is illustrated in Figure 15.30 for the Student t SIRV with $k_s = 10$ and $k_o = 1$. The IS weighting function Equation 15.92 for this modification is

$$w(r_s, r_o) = k_s^2 k_o^{2(N-1)} = \frac{h_{2N}(r_s^2 + r_o^2)}{h_{2N}[(r_s^2/k_s^2) + (r_o^2/k_o^2)]} I_\gamma(\mathbf{r}) \tag{15.98}$$

Threshold estimates obtained from only 10,000 importance sampling trials for the Student t GLRT with $N = 4$ are shown in Figure 15.31 for the shape parameters, $\nu = 1.1$ and $\nu = 2$. The theoretical values of the threshold are given by $\gamma = P_{\mathrm{FA}}^{-1/(\nu-1)} - 1$ and marked with a "diamond" in the figure. Results for three values of k_s are overlaid and all show excellent agreement with the analytical results for a wide range of P_{FA} values. Figure 15.32 shows an approximation to the relative RMS error, $\hat{e} = \hat{\sigma}_w/(\hat{P}_{\mathrm{FA}}\sqrt{M_{\mathrm{is}}})$, where $\hat{\sigma}_w$ is the sample variance of the importance sampling weights. Similar results are shown in

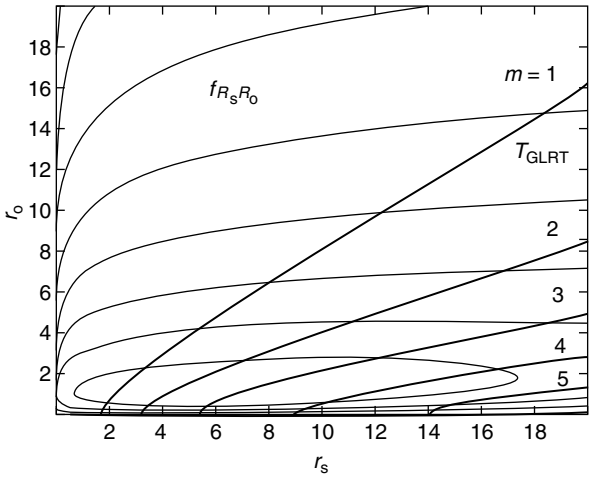

FIGURE 15.30 Modified Student t contour plots of T_{GLRT} and $f_{R_s,R_o|H_0}(r_s, r_o|H_0)$ with $N = 4$, $\nu = 2$, $k_s = 10$, $k_o = 1$. Numbered GLRT contours correspond to $P_{\mathrm{FA}} = 10^{-m}$, where m is the number on the curve.

Figure 15.33 and Figure 15.34 for a K-distributed SIRV with shape parameter, $\nu = 0.4$, although no theoretical results are available in this case.

Both examples exhibit a distinct minimum in the relative RMS error for a particular k_s. However, this minimum is very broad and excellent estimates of the threshold are obtained over many orders of magnitude in P_{FA} for a single choice of k_s. As the relative RMS error increases, the results become noticeably erratic. Where the relative RMS error is small, the results converge for several k_s values.

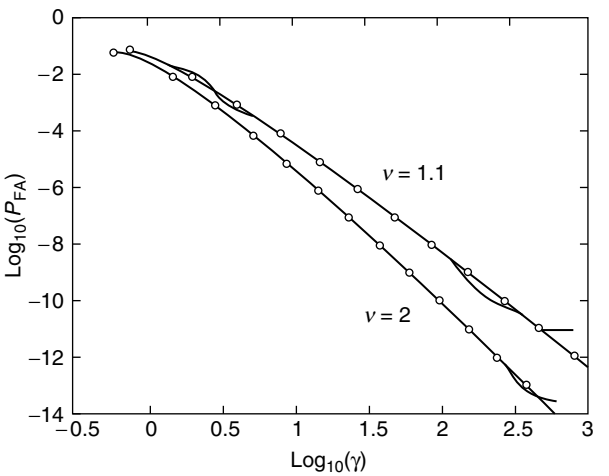

FIGURE 15.31 Importance sampling threshold estimates for a Student t GLRT with $N = 4$, $M_{\mathrm{is}} = 10000$, and $k_s = 10, 20, 50$.

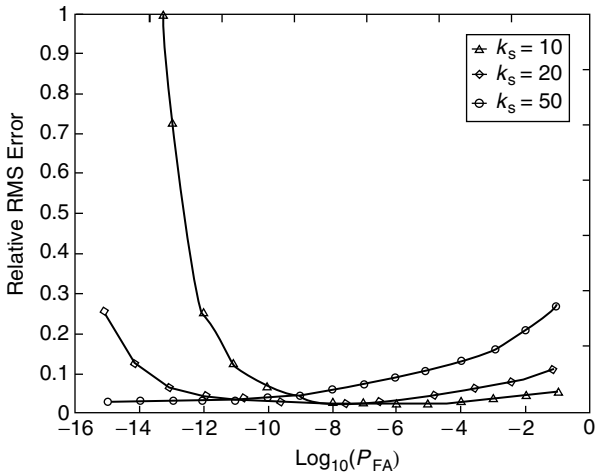

FIGURE 15.32 Relative RMS error of the importance sampling threshold estimates for a Student t GLRT with $N = 4$, $\nu = 2$, $M_{is} = 10000$.

The asymptotic linear nature of all these results is very significant. This is shown to be a consequence of extreme value theory in Section 15.4.6. It suggests that thresholds for very small P_{FA} values can be extrapolated from good estimates at higher P_{FA}. Figure 15.35 and Figure 15.36 show the threshold estimates the relative RMS error for the Student t SIRV when a MLE of the covariance matrix is obtained form K secondary data vectors. An identity covariance matrix is used for this simulation. The estimate of the covariance matrix is good when larger numbers of secondary data vectors ($10N$, $5N$, and $4N$) are used. Excellent

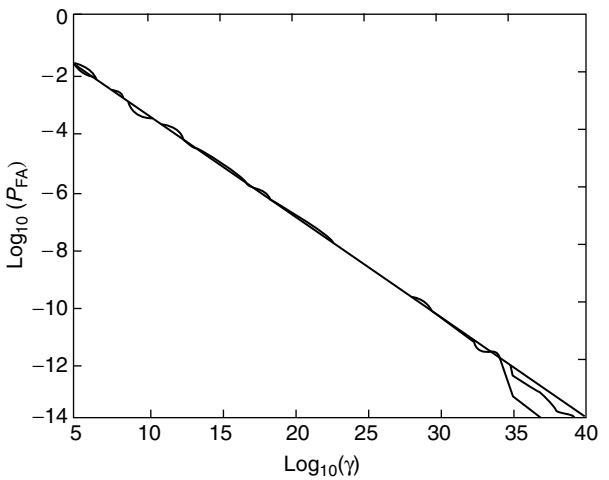

FIGURE 15.33 Importance sampling threshold estimates for K distributed GLRT with $N = 4$, $M_{is} = 10000$, and $k_s = 20, 40, 100, 200$.

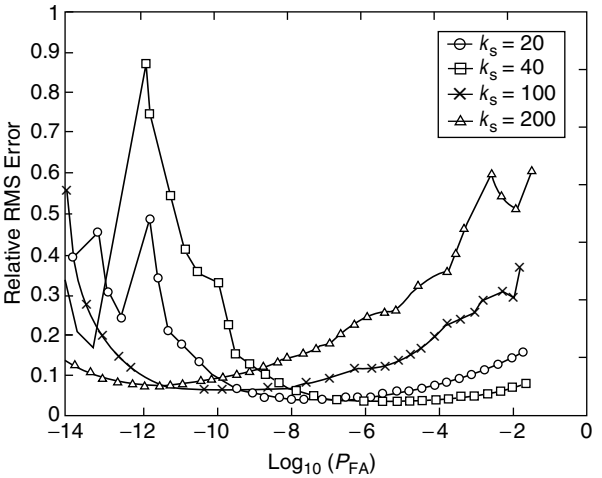

FIGURE 15.34 Relative RMS error of the importance sampling threshold estimates for a K-distributed GLRT with $N = 4$, $\nu = 0.4$, $M_{\text{is}} = 10000$.

threshold estimates down to $P_{\text{FA}} = 10^{-10}$ are obtained. The results degrade when the covariance matrix is estimated from fewer secondary data vectors ($2N$ and $3N$), but not severely. The asymptotic linear behavior is observed in this case and the slope of the asymptote is not very sensitive to K. Consequently, the slope of the known covariance matrix case might be applied to the estimated cases to extrapolate thresholds for the lower false alarm values.

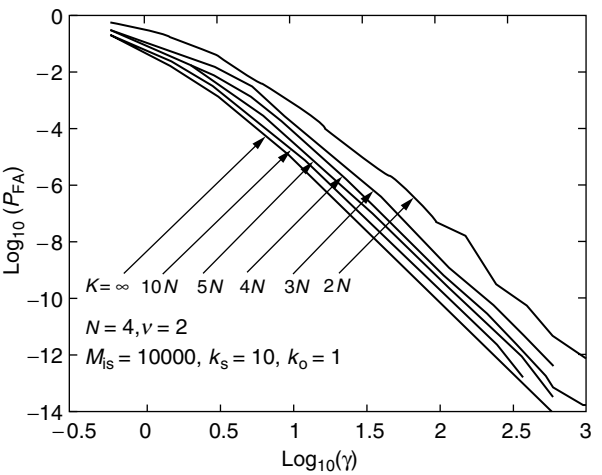

FIGURE 15.35 Importance sampling threshold estimates using estimated covariance matrices for a Student t GLRT with $N = 4$, $\nu = 2$, $M_{\text{is}} = 10000$.

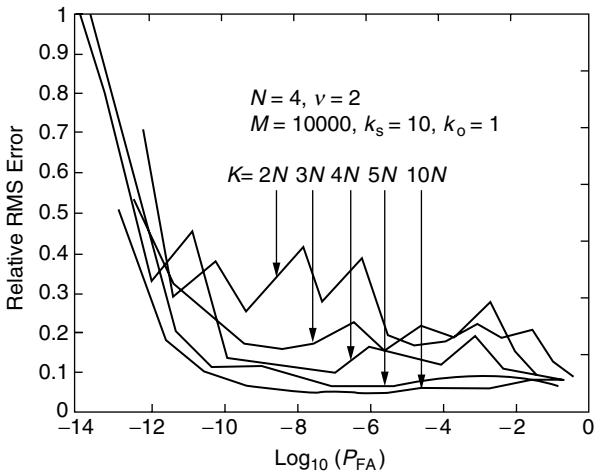

FIGURE 15.36 Relative RMS error of the importance sampling threshold estimates using estimated covariance matrices for a Student t GLRT with $N = 4$, $\nu = 2$, $M_{\text{is}} = 10000$.

15.4.6. Extreme Value Theory Approximation

The results exhibit a linear relationship between $\log(P_{\text{FA}})$ and $\log(\gamma)$ at very small P_{FA}. This is a consequence of extreme value theory,[19] which addresses the distribution of the maximum (and minimum) extreme values of independent, identically distributed (IID) random variables. The distribution of the extremes is related to the upper (and lower) tail of the underlying distribution of the random variables. The extreme value distributions are obtained as the limiting distributions for $N \rightarrow \infty$ of the largest (or smallest) value in the sample of N IID random variables.

Since the threshold is in the upper tail of the test statistic PDF for the very small false alarm probabilities of interest, extreme value theory as it relates to maximum extremes is applicable in the radar problem. The appropriate cumulative distribution function (CDF) for this case is the generalized extreme value (GEV) distribution,[19]

$$F_T(t) = \mathrm{e}^{-\left[1+\left(\xi\frac{t-\mu}{\sigma}\right)^{-1/\xi}\right]},\ \frac{\mu-\sigma}{\xi} \leq t < \infty,\ \xi > 0 \qquad (15.99)$$

A broad range of tail behavior is modeled by the shape parameter, ξ, and $\xi > 0$ corresponds to an infinitely long upper tail which is characteristic of the PDF for many radar test statistics.

Expanding Equation 15.99 in a Taylor series gives

$$F_T(t) = 1 - \left(1 + \xi\frac{t-\mu}{\sigma}\right)^{-1/\xi} + \frac{1}{2!}\left(1 + \xi\frac{t-\mu}{\sigma}\right)^{-2/\xi} - \cdots \qquad (15.100)$$

From $\xi > 0$ and sufficiently large threshold values, the higher order terms of the expansion are negligible and $F_T(t)$ is approximately

$$\hat{F}_T(t) = 1 - \left(1 + \xi\frac{t-\mu}{\sigma}\right)^{-1/\xi} \tag{15.101}$$

This is exactly the form of the generalized Pareto distribution, which has been successfully used to model tail behavior of locally optimum detector (LOD) test statistics.[20]

The P_{FA} obtained from the approximate CDF is

$$P_{\mathrm{FA}} \approx 1 - \hat{F}_T(t) = \left(1 + \xi\frac{t-\mu}{\sigma}\right)^{-1/\xi} \tag{15.102}$$

Applying the logarithm to both sides yields

$$\log(P_{\mathrm{FA}}) \approx -\frac{1}{\xi}\log\left(1 + \xi\frac{t-\mu}{\sigma}\right) \tag{15.103}$$

If t is sufficiently large, the approximation simplifies to

$$\log(P_{\mathrm{FA}}) \approx -\frac{1}{\xi}\log\left(\xi\frac{t}{\sigma}\right) = -\frac{1}{\xi}\log(t) - \frac{1}{\xi}\log\left(\frac{\xi}{\sigma}\right) \tag{15.104}$$

This is the equation of a straight line in the variables, $\log(P_{\mathrm{FA}})$ and $\log(t)$, and is valid for suitably large threshold values corresponding to very low probabilities of false alarm. The importance sampling simulation results presented in this chapter demonstrate this linear behavior at low false alarm probabilities.

15.5. REJECTION-METHOD BOUNDS FOR MONTE CARLO SIMULATION OF SIRVs

(A. D. KECKLER AND D. D. WEINER)

15.5.1. Introduction

In recent years, considerable interest has developed in the use of SIRV's as a model for non-Gaussian distributed radar clutter.[1–7,12–14] As a result, efficient techniques in the generation of SIRV distributed random samples are desirable for Monte Carlo simulation and system performance analysis. According to the representation theorem for SIRVs, an SIRV with zero-mean independent components can be generated as the product of a univariate random variable and a Gaussian vector with zero-mean independent components. In contrast with other multivariate non-Gaussian distributions, SIRVs exhibit closure under linear transformations, and the desired correlation can then be imposed by simple multiplication of the SIRV by the appropriate matrix. Additionally, the desired mean vector can be introduced by simply adding it to the SIRV.

SIRV's can be simulated by separately generating a Gaussian vector and an independent random scalar multiplier. The generation of independent zero-mean Gaussian random vectors is well understood. The univariate random multiplier is used to control the type of SIRV. However, generation of random data for the univariate multiplier is not always straightforward. Its probability density function (PDF), known as the characteristic PDF of the SIRV, may well have a mathematically complex form which does not lead to a convenient inverse for the corresponding cumulative distribution function (CDF), which is required to directly generate data for the scalar multiplier. If a function can be found that tightly bounds the characteristic PDF from above, and for which a convenient inverse CDF can be found, then the rejection theorem can be used to efficiently generate data from the characteristic PDF. A method for finding such a tight, computationally simple bound is presented in this section. This is significant, as the efficiency of the rejection method suffers dramatically when the bound is not tight. In the case where the characteristic PDF is not known, an alternative approach using the envelope of the SIRV can be employed.[8] Additionally, this technique can be applied directly to the PDFs of the envelope and/or the quadratic form for direct generation of these quantities, when only they are of interest.

15.5.2. Summary of the SIRV Model

A random vector $\mathbf{Y}$ of dimension N is defined to be an SIRV if and only if its PDF has the form[6,7]

$$f_Y(\mathbf{y}) = (2\pi)^{-N/2}|\boldsymbol{\Sigma}|^{-1/2}h_N(q(\mathbf{y})) \tag{15.105}$$

where $\boldsymbol{\Sigma}$ is an $N \times N$ nonnegative definite matrix, $q(\mathbf{y})$ is the quadratic form defined by

$$q = q(\mathbf{y}) = (\mathbf{y} - \mathbf{b})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{y} - \mathbf{b}) \tag{15.106}$$

$\mathbf{b}$ is the $N \times 1$ mean vector, and $h_N(\cdot)$ is a positive, monotonic decreasing function for all N.[8] Equivalently, an SIRV $\mathbf{Y}$ can be represented by the linear transformation

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b} \tag{15.107}$$

where $\mathbf{X}$ is a zero-mean SIRV with uncorrelated components represented by

$$\mathbf{X} = S\mathbf{Z} \tag{15.108}$$

$\mathbf{Z}$ is a zero-mean Gaussian random vector with independent components, and S is a nonnegative random variable independent of $\mathbf{Z}$. The probability density function of S, $f_S(s)$, uniquely determines the type of SIRV and is known as the characteristic PDF of $\mathbf{Y}$. Since the matrix $\mathbf{A}$ is specified independently of $f_S(s)$, an arbitrary covariance matrix, $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}^{\mathrm{T}}$, can be introduced without altering the type of SIRV.

The representation is used to obtain

$$h_N(q) = \int_0^\infty s^{-N} \mathrm{e}^{-q/2s^2} f_S(s) \mathrm{d}s \tag{15.109}$$

and subsequently, the PDF of the quadratic form is

$$f_Q(q) = \frac{1}{2^{N/2}\Gamma(N/2)} q^{(N/2)-1} h_N(q) \tag{15.110}$$

Since $h_N(q)$ uniquely determines each type of SIRV, Equation 15.110 indicates that the multivariate approximation problem is reduced to an equivalent univariate problem.

It is not always possible to obtain the characteristic PDF $f_S(s)$ in closed-form. However, an N dimensional SIRV with uncorrelated elements can be expressed in generalized spherical coordinates R, θ, and ϕ_k for $k = 1, \ldots, (N-2)$, where the PDF of R is given by

$$f_R(r) = \frac{r^{N-1}}{2^{(N/2)-1}\Gamma(N/2)} h_N(r^2) u(r) \tag{15.111}$$

The angles θ and ϕ_k are statistically independent of the envelope R and do not vary with the type of SIRV. When $f_S(s)$ is unknown, Equation 15.111 is used both to generate SIRVs and to determine $h_N(q)$.[6]

It is desirable to develop a library of SIRVs for use in approximating unknown clutter-returns. Table 15.4 contains the characteristic PDFs and $h_N(q)'$s of some SIRVs for which analytical expressions are known. For simplicity, the results presented for the SIRVs developed from the marginal envelope (Chi, Weibull, generalized Rayleigh, Rician, and generalized Gamma) are valid only for even N. Additional SIRVs, such as the generalized Pareto envelope, envelope SIRVs based on the Confluent Hypergeometric Function, and SIRVs based upon polynomial characteristic PDFs, have been developed.[9]

The discreete Gaussian mixture (DGM) is an SIRV of special interest. Its PDF is a simple finite weighted sum of Gaussian PDFs. It is useful for approximating many other SIRVs, as well as generating unique distributions.

15.5.3. Generation of SIRV Distributed Samples

If the characteristic PDF of the SIRV $f_S(s)$ is known, a zero-mean SIRV with uncorrelated components $\mathbf{X}$ can be generated from Equation 15.108, where $\mathbf{Z}$ is a zero-mean Gaussian vector with uncorrelated components. Without loss of generality, we assume that $E\{S^2\} = 1$, so that the covariance matrix of $\mathbf{X}$ is then identical to that of $\mathbf{Z}$. The desired covariance matrix and mean vector can be introduced using the linear transformation shown in Equation 15.107, where the covariance matrix of $\mathbf{Y}$ is $\mathbf{A}\mathbf{A}^{\mathrm{T}}$ and its mean vector is $\mathbf{b}$. This procedure is shown in Figure 15.37.

TABLE 15.4
Characteristic PDFs and $h_N(Q)$ Functions for Known SIRVs. (An Expanded Version of Table 15.1)

Marginal PDF	Characteristic PDF $f_S(s)$	$h_N(q)$
Gaussian	$\delta(s-1)$	$e^{-q/2}$
Student t	$\frac{2b}{\Gamma(\nu)2^{\nu}} b^{2\nu-1} s^{-(2\nu+1)} e^{-(b^2/2s^2)} u(s)$	$\frac{2^{N/2} b^{2\nu} \Gamma\left(\nu+\frac{N}{2}\right)}{\Gamma(\nu)(b^2+q)^{(N/2)+\nu}}$
Laplace	$b^2 s e^{(-b^2 s^2/2)} u(s)$	$b^N (b\sqrt{q})^{1-(N/2)} K_{(N/2)-1}(b\sqrt{q})$
K-distributed envelope	$\frac{2b}{2^{\alpha}\Gamma(\alpha)} (bs)^{2\alpha-1} e^{-(b^2 s^2/2)} u(s)$	$\frac{b^N}{\Gamma(\alpha)} \frac{(b\sqrt{q})^{\alpha-(N/2)}}{2^{\alpha-1}} K_{(N/2)-\alpha}(b\sqrt{q})$
Cauchy	$\sqrt{\frac{2}{\pi}} b s^{-2} e^{-(b^2/2s^2)} u(s)$	$\frac{2^{N/2} b \Gamma\left(\frac{N}{2}+\frac{1}{2}\right)}{\sqrt{\pi}(b^2+q)^{(N/2)+(1/2)}}$
Chi envelope $\nu < 1$	$\frac{2^{\nu+1} b^{2\nu}}{\Gamma(\nu)\Gamma(1-\nu)} \frac{s^{2\nu-1}}{(1-2b^2 s^2)^{\nu}} u(s) u\left(\frac{1}{b\sqrt{2}} - s\right)$	$\frac{(2)^{N/2} b^{2\nu}}{\Gamma(\upsilon)} \sum_{k=1}^{N/2} \binom{\frac{N}{2}-1}{k-1} q^{\nu-k} b^{N-2k} \frac{\Gamma(k-\nu)}{\Gamma(1-\nu)} e^{-b^2 q}$
Gaussian mixture $w_k > 0$, $\sum_{k=1}^{K} w_k = 1$	$\sum_{k=1}^{K} w_k \delta(s-s_k)$	$\sum_{k=1}^{K} w_k s_k^{-N} e^{-(q/2s_k^2)}$

Continued

TABLE 15.4 Continued

Marginal PDF	$h_N(Q)$	Components of $h_N(Q)$
Weibull envelope $0 < b < 2$	$(-2)^{N/2}\mathrm{e}^{-a\sigma^b q^{b/2}} \sum_{k=1}^{N} B_k \frac{(a\sigma^b)^k}{k!} q^{(kb/2)-(N/2)}$	$B_k = \sum_{m=1}^{k} (-1)^m \binom{k}{m} \prod_{i=0}^{M-1} \left(\frac{mb}{2} - i\right)$
Generalized Rayleigh $0 \le \alpha \le 2$	$\sum_{k=1}^{(N/2)-1} D_k q^{(k\alpha/2)-(N/2)+1} \mathrm{e}^{-\beta^{-\alpha}\sigma^{\alpha} q^{\alpha/2}}$	$D_k = \sum_{m=1}^{k} (-1)^{m+(N/2)-1} 2^{(N/2)-1} \frac{\beta^{-\alpha k}\sigma^{\alpha k}}{k!} \binom{k}{m} \frac{\Gamma\left(1 + \frac{m\alpha}{2}\right)}{\Gamma\left(2 + \frac{m\alpha}{2} - N\right)}$
Rician $0 < \rho \le 1$	$\frac{\sigma^N}{(1-\rho^2)^{\frac{N-1}{2}}} \sum_{k=0}^{(N/2)-1} \binom{\frac{N}{2} - 1}{k} (-1)^k \left(\frac{\rho}{2}\right)^k \xi_k \mathrm{e}^{\sigma^2 q/2(1-\rho^2)}$	$\xi_k = \sum_{m=0}^{k} \binom{k}{m} I_{k-2m}\left(\frac{\rho\sigma^2 q}{2(1-\rho^2)}\right)$
Generalized Gamma $c\alpha \le 2$ $a, c, \alpha > 0$	$\sum_{k=0}^{(N/2)-1} F_k q^{(c\alpha - N)/2} \mathrm{e}^{-a\sigma^c q^{c/2}}$	$F_k = (-2)^{(N/2)-1} \frac{c(a\sigma)^{c\alpha}}{\Gamma(\alpha)} \binom{\frac{N}{2} - 1}{k} \frac{\Gamma\left(\frac{c\alpha}{2}\right)}{\Gamma\left(\frac{c\alpha - N}{2} + k + 1\right)} \sum_{m=1}^{k} G_m q^{cm/2},$
		$G_m = \sum_{d=1}^{m} \frac{(-1)^{2md}}{m!} (a\sigma^c)^m \binom{m}{d} \frac{\Gamma\left(\frac{dc}{2} + 1\right)}{\Gamma\left(\frac{dc}{2} - k + 1\right)}$

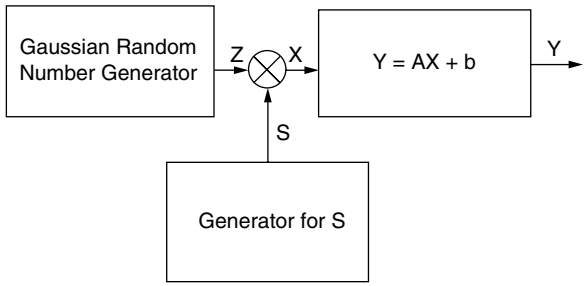

FIGURE 15.37 Generation of the SIRVs with known characteristic PDF $f_S(s)$.

When the characteristic PDF is unknown, an alternate approach is required. Recall that a SIRV with uncorrelated elements can be represented in spherical coordinates, elements can be represented in spherical coordinates, and that the PDFs of θ and ϕ_k remain unchanged, regardless of the type of SIRV under consideration. Only the PDF of R changes. Furthermore, since a Gaussian random vector is an SIRV, a white zero-mean Gaussian vector can be expressed in the spherical coordinates of Equation 15.111. It follows that[6]

$$\frac{X_k}{\|\mathbf{X}\|} = \frac{Z_k}{\|\mathbf{Z}\|}, \quad k = 1, 2..., N \tag{15.112}$$

where $\|\cdot\|$ denotes the vector norm. Consequently, the components of the desired white SIRV can be obtained from

$$X_k = Z_k \frac{\|\mathbf{X}\|}{\|\mathbf{Z}\|}, \quad k = 1, 2..., N \tag{15.113}$$

It should be noted that the norm of an N dimensional SIRV is the envelope R, which has the PDF given in Equation 15.111. In order to simulate the zero-mean white SIRV, it is only necessary to generate a sample from its envelope, and multiply it by a zero-mean white Gaussian distributed vector with norm of unity. Again, the desired covariance matrix and mean vector can be imposed upon the SIRV using Equation 15.107. This simulation procedure is illustrated in Figure 15.38.

A quick perusal of Table 15.4 shows that many of the characteristic PDFs and envelope PDFs are uncommon and mathematically complex. They do not lead to convenient closed-forms for the inverse CDFs, and thus are unsuitable for the direct generation of random samples. Some, such as the Student t characteristic PDF, can be generated by transforming samples from well known distributions (in this case, the gamma distribution), but involve a reciprocal. This can cause problems, due to the granularity of the original samples. The transformed samples do not uniformly sample the domain of the desired PDF, and is particularly relevant to the non-Gaussian problem. A preferred method for generating samples from the characteristic PDFs or the envelope PDFs is the rejection method.

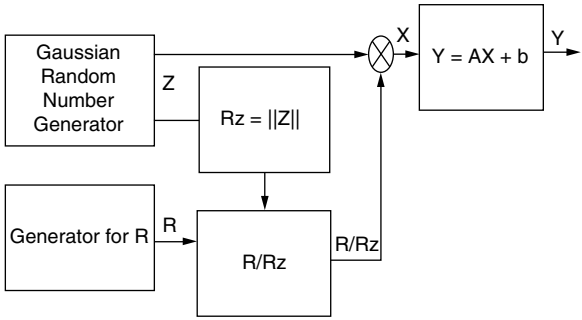

FIGURE 15.38 Generation of the SIRVs with unknown characteristic PDF $f_S(s)$.

The rejection theorem can be stated as[10]:

Theorem 1. *Let S be a random variable with density $f_S(s)$ and Φ be any random variable with density $f_\Phi(\phi)$ such that $f_s(s) = 0$ whenever $f_\Phi(\phi) = 0$. Then let U be uniformly distributed on the interval (0,1). If Φ and U are statistically independent and*

$$\eta = \{u \leq T(\phi)\} \tag{15.114}$$

where

$$T(\phi) = f_s(\phi)/[\alpha f_\Phi(\phi)] \leq 1 \tag{15.115}$$

then the rejection theorem states

$$f_{\Phi|\eta}(\phi|\eta) = f_S(\phi) \tag{15.116}$$

The density $f_\Phi(\phi)$ approximates $f_S(s)$ if the value

$$\alpha = \max_s[f_\Phi(s)/f_S(s)] \tag{15.117}$$

is a constant close to one. If α equals 1, then f_S is identical to f_ϕ. In Figure 15.39, the function αf_Φ is seen to bound f_S in the sense that $\alpha f_\Phi(s) \geq f_S(s)$ for all s in the support of S. It is desired to generate a random variable S with density $f_S(s)$ from variates generated from $f_\Phi(\phi)$. This can be accomplished using the following algorithm:

STEP 1. Generate ϕ from $f_\Phi(\phi)$ and compute $T(\phi) = f_S(\phi)/\alpha f_\Phi(\phi)$,
STEP 2. Generate u as a realization of a uniform random variable distributed over (0,1).
STEP 3. If $u > T(\phi)$, reject ϕ and return to STEP 1, else accept ϕ as a variate from $f_S(s)$.

If $f_S(s)$ is a time-consuming function to evaluate, and there exists a function $h(s)$ such that $h(s) \leq f_S(s)$ for all s in the support of S, then a fast, preliminary test can be made, as can be seen in Figure 15.39. The modified procedure becomes:

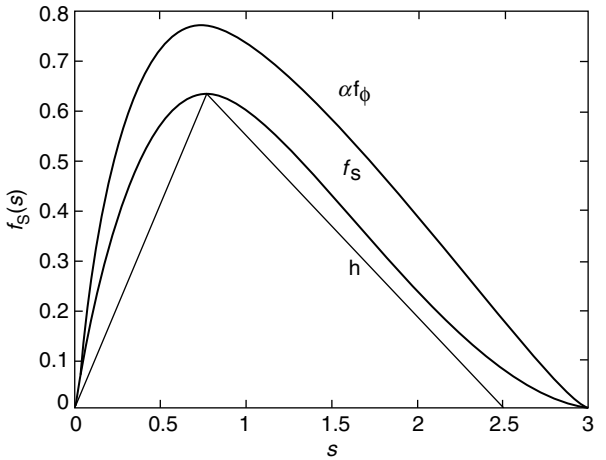

FIGURE 15.39 Illustration of the acceptance–rejection method.

STEP 1. Generate a realization ϕ from $f_\Phi(\phi)$ and compute $T_h(\phi) = h(\phi)/\alpha f_\Phi(\phi)$,
STEP 2. Generate u drawn from a uniform (0,1),
STEP 3. If $u \leq T_h(\phi)$, accept ϕ as a variate from $f_S(s)$,
STEP 4. Else, compute $T(\phi) = f_S(\phi)/\alpha f_\Phi(\phi)$. If $u \leq T(\phi)$, accept ϕ as a variate from $f_S(s)$, else reject ϕ and return to STEP 1.

The procedure has a geometric interpretation. A point (ϕ, y) is generated in the region bounded by $\alpha f_\Phi(\phi)$ and the ϕ-axis with probability $1/\alpha$. If the point falls within the region bounded by $h(\phi)$ and the ϕ axis, accept ϕ immediately. If not, then if the point falls within the region bounded by $f_S(\phi)$ and the ϕ-axis, accept ϕ. Otherwise reject ϕ. The parameter α equals the area under the bound function, and the average efficiency of the rejection algorithm is equal to $1/\alpha$.

15.5.4. Generation of PDF Bounds

For the rejection method to be viable, it is necessary to find a suitable bound for which random samples can easily be generated. As a practical matter, this bound should have an area as close to unity as possible, to avoid rejecting too many samples. Many simple bounds encountered may have extremely low acceptance rates. In the approach used here, the PDF is segmented into M equal intervals, such that horizontal line segments can be used to approximate the PDF, as shown in Figure 15.40. Samples can be generated from each segment with a simple uniform number generator, and each segment is chosen with a probability equal to its relative area.

Obviously, if the PDF has an infinite tail, the entire support cannot be segmented. The PDF is then divided into a body and a tail section at a point

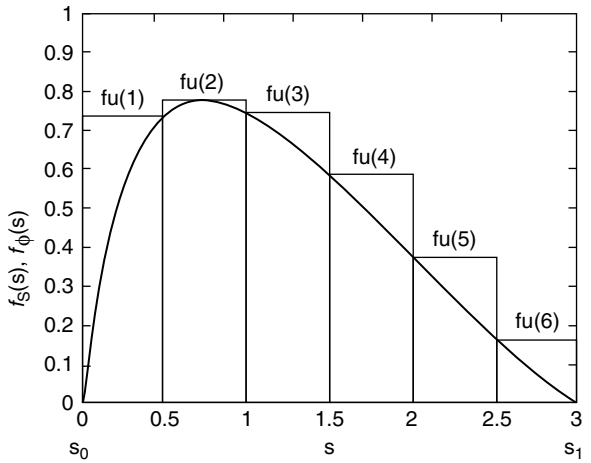

FIGURE 15.40 Piecewise constant bound of PDF.

sufficiently far into the tail. The generalized Pareto PDF, given by

$$f_S^{(s)}\Big|_{\text{Pareto}} = \frac{1}{\sigma}\left(1 + \frac{\gamma(s - s_1)}{\sigma}\right)^{-(1+\gamma)/\gamma} u(s - s_1) \qquad (15.118)$$

where σ is the scale parameter, γ is the shape parameter, and s_1 is the point where the tail begins, can be sued to bound the tail. This is illustrated in Figure 15.41. The parameters σ and γ in Equation 15.118 can be obtained by matching

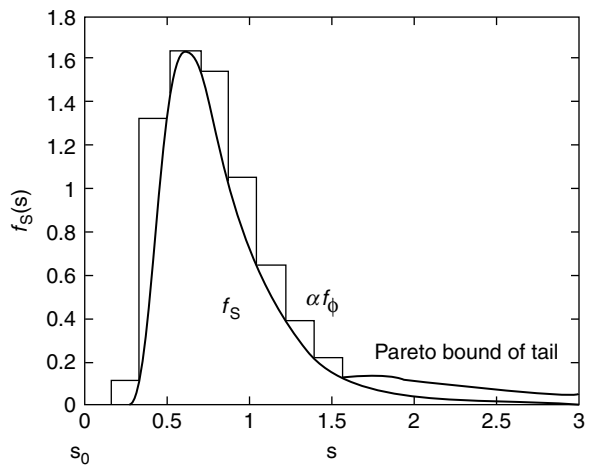

FIGURE 15.41 Pareto bound of PDF tail.

probability weighted moments. The estimates of the parameters are given by[11]

$$\hat{\sigma} = 2a_0a_1/(a_0 - 2a_1) \tag{15.119}$$

and

$$\hat{\gamma} = 2 - a_0/(a_0 - 2a_1) \tag{15.120}$$

where

$$a_0 = \sum_{i=1}^{m} z_i/m \tag{15.121}$$

and

$$a_1 = \sum_{i=1}^{m} (m - i)z_i/\{m(m - 1)\} \tag{15.122}$$

The samples z_i can be obtained using a bootstrapping approach, where the rejection method, using a piecewise constant bound, is applied to a truncated portion of the tail. While this is approximate in the sense that the samples generated are drawn only from a portion of the tail, it is sufficient for fitting the generalize Pareto bound to the tail. Similarly, if the PDF becomes infinite at its endpoints, an inverted generalized Pareto PDF, given by

$$f_S(s)\Big|_{\text{Inverted Pareto}} = \frac{\sigma}{1 + \gamma}(\sigma s)^{-\gamma/(1+\gamma)}u(s)u(\sigma - s) \tag{15.123}$$

can be used to bound the PDF near the singularity. This is illustrated in Figure 15.42. Random samples for the generalized Pareto PDF and the PDF of Equation 15.123 are readily generated, as the inverse of the CDF's for both are simple in form.

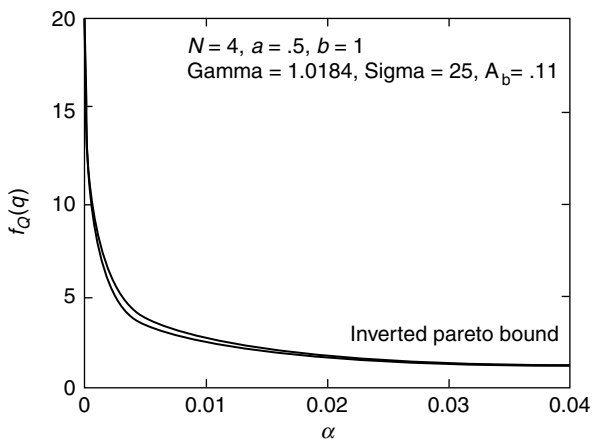

FIGURE 15.42 Inverted Pareto bound of singularity.

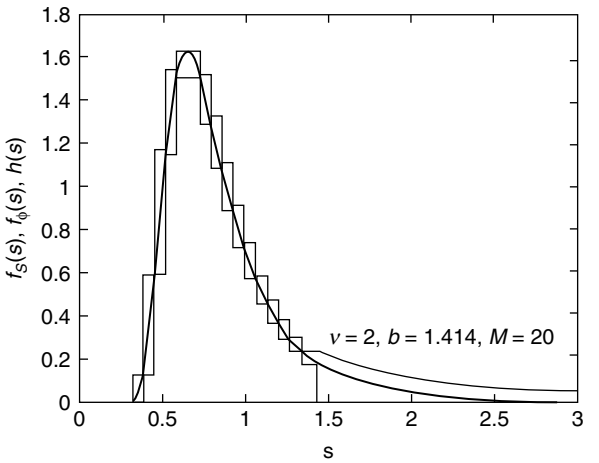

FIGURE 15.43 Bounds for the Student t SIRV characteristic PDF.

Random samples were obtained for the Student t SIRV characteristic PDF using this technique. Figure 15.43 shows the bounds generated, while Figure 15.44 shows a histogram for 10000 samples. As the total probability associated with the tail portion of the PDF is usually small, the lower bound $h(s)$ is not fitted to it, since relatively few evaluations of the PDF in the tail region will be required in any case. In this example, fully 89.6% of the samples generated were accepted using the bounds shown in Figure 15.43. Figure 15.45 shows a histogram of 10000 samples generated from the enveloped of the K-distributed

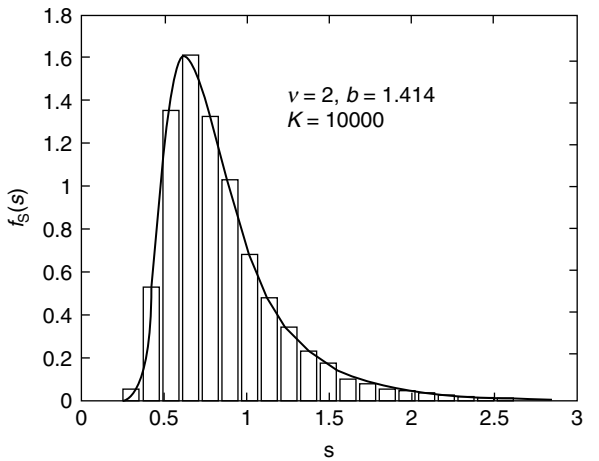

FIGURE 15.44 Histogram of characteristic PDF for the Student t SIRV.

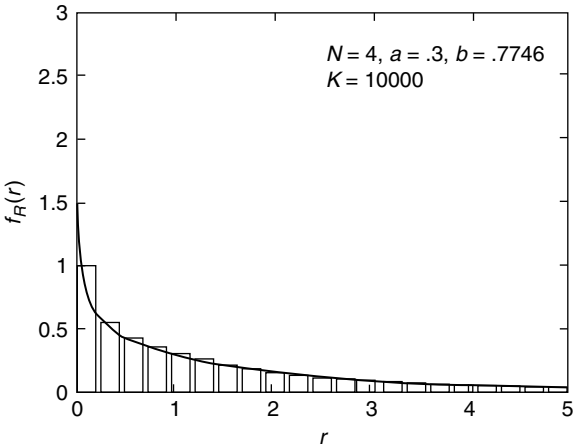

FIGURE 15.45 Histogram of envelope PDF for the K-distributed SIRV.

SIRV when $N = 4$. This PDF goes to infinity at the origin. Despite this, an efficiency of 74.2% was still obtained in this case.

15.5.5. Concluding Remarks

Efficient methods for simulating correlated multivariate non-Gaussian data are of interest. This chapter presents a technique for simulating non-Gaussian data based upon the SIRV model. The rejection method provides an elegant solution to the problem, provided a suitable bound can be found. A simple method is presented for producing tight bounds to fairly arbitrary PDFs, and for which the generation of random samples is easy and efficient. While the generation of the bound itself may require a fair amount of computation, it is straightforward and need be done only once for any particular PDF. Furthermore, this technique is not specific to any one SIRV, and can additionally be applied to a wide range of univariate densities. Of equal importance, this approach avoids the problems incurred when generating samples through the use of transformations. By using the generalized Pareto PDF to bound the tail, and its inverse to bound singularities, no truncation of the desired PDF is encountered, and samples can thus be drawn from the entire PDF.

The examples presented illustrate the efficiency of the bounds with respect to the acceptance of the samples generated. The acceptance rate for the bound can be readily adjusted by increasing or decreasing the number of segments used in the bound, at the cost of increasing the bound's computational complexity. This must be balanced against the cost of evaluating the original PDF. Further improvements to the bound can be readily achieved through the use of simple techniques, such as by using nonuniformly spaced segments, or by using piecewise linear instead of piecewise constant segments.

15.6. OPTIMAL NONGAUSSIAN PROCESSING IN SPHERICALLY INVARIANT INTERFERENCE

(D. STADELMAN AND D. D. WEINER)

15.6.1. Introduction

Conventional adaptive radar receiver designs are based on the assumption of Gaussian distributed clutter data, which corresponds to a Rayleigh distributed envelope. However, the statistics of a significant portion of the radar clutter samples from a surveillance volume are often non-Gaussian, particularly for data collected at low grazing angles or at high resolution. A Gaussian receiver is not optimum in this environment and may experience significant reduction in detection performance, especially for small targets, when compared to the optimal non-Gaussian processor. Other research[1] has shown that is some types of non-Gaussian clutter, near-optimal processing by locally optimum detectors gives detection probabilities of about .1 to .3, whereas, the Gaussian receiver applied to the same data has detection probabilities on the order of the false alarm rate, typically 10^{-3} to 10^{-5}.

This loss in detection performance of the Gaussian receiver may be understood by considering a probability density function (PDF), such as the Weibull or *K*-distribution, which is typically used to model real non-Gaussian radar clutter data. These PDFs have higher tails than the Gaussian PDF, which results in more frequent occurrences of very large clutter values. This effect is often described as "spiky" clutter and is illustrated in Figure 15.46.

The very low false alarm probabilities for which radar systems are designed cause the detection threshold to fall in the tail of the PDF for the clutter-only test statistic. The spiky clutter associated with the extended tail density of the non-Gaussian data generates significantly more false alarms in the Gaussian receiver. Consequently, the threshold in the Gaussian receiver must be raised to maintain the desired false alarm rate for the non-Gaussian clutter problem. This higher threshold causes a reduction in the probability of detection.

In contrast, the optimal non-Gaussian receiver is found to contain a nonlinearity which reduces the large clutter spikes. This allows the threshold to be maintained at a lower level, which provides increased target detection opportunities at the desired false alarm rate. This improvement in detection probability of the non-Gaussian receiver can be very significant, especially for low signal-to-clutter ratio (SCR) targets.

Optimal radar target detection requires the joint PDF of N pulse returns which are collected from a particular range-azimuth cell during a coherent processing interval (CPI). Pulse-to-pulse correlation may exist in the clutter-returns of this received data vector.

If the clutter samples in the received data vector are Gaussian, the form of the joint PDF is well known and optimal detection is accomplished by a matched filter. The matched filter is the linear filter which maximizes the output signal-to-interference ratio. Its design requires knowledge of only the second order moments

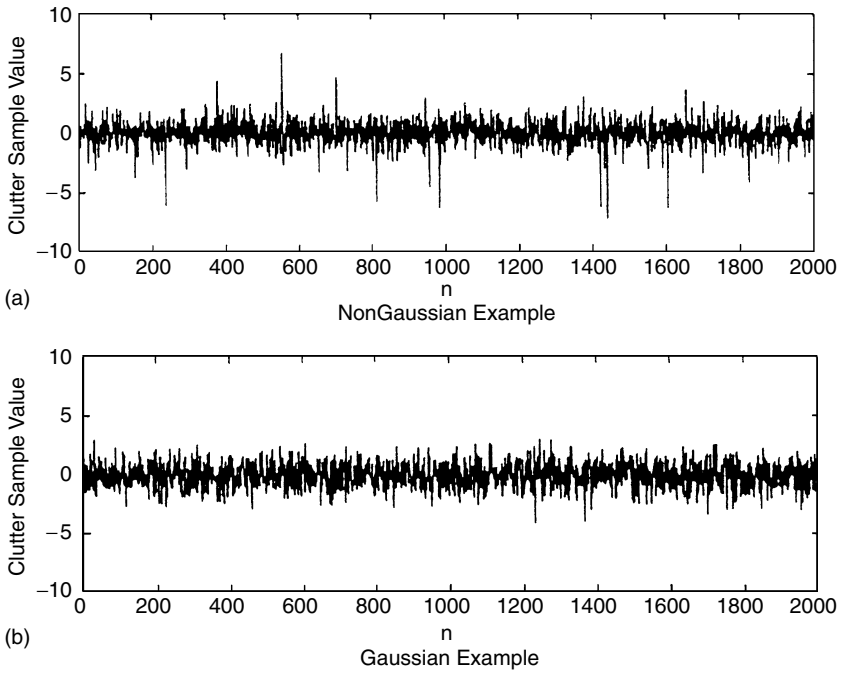

FIGURE 15.46 Time sequence of clutter data.

of the interference. Since the PDF of a zero-mean Gaussian random vector is completely determined by its covariance matrix, the matched filter is able to accomplish optimal detection in Gaussian clutter. A complete characterization of non-Gaussian clutter requires knowledge of the higher order moments of the clutter PDF. Thus, a matched filter alone is neither expected nor able to achieve optimum performance. Some form of nonlinear processing of the received data is necessary.

There are two major requirements of optimal non-Gaussian receiver design which must be addressed in any target detection application: (1) specification of an appropriate non-Gaussian PDF model, which must be approximated from the received clutter data, and (2) determination and implementation of the optimal (or near-optimal) non-Gaussian receiver for the PDF model which is selected.

Specification of the optimal receiver for targets in non-Gaussian clutter is typically limited by the lack of a sufficient mathematical model to completely describe the joint PDF of the non-Gaussian data. Furthermore, optimal detection is complicated by the lack of an efficient technique for accurately approximating this joint PDF from the received data. Frequently, the assumption of independent identically distributed (IID) interference samples is used for non-Gaussian data because it yields a closed-form solution for the joint PDF. This assumption also reduces the multivariate PDF approximation problem to a univariate one. However, the independent sample assumption implies a white clutter spectrum and can lead to inferior performance when the received data is correlated.

Spherically invariant random vectors (SIRVs) have recently been investigated for modeling correlated, non-Gaussian clutter data.[3,4] Closed-form expressions exist for the multivariate PDF of many types of SIRVs. Furthermore, these multivariate PDFs are uniquely identified by a single univariate PDF. The class of SIRV models includes many correlated distributions of interest in radar applications, such as Weibull, Student t, Gaussian, Chi, and K distributed.

The application of SIRV models to many non-Gaussian clutter environments has both empirical and theoretical support. First, the SIRV model is equivalent to the compound clutter model, which is found to be an excellent fit to real sea clutter data in many instances.[5,20] Second, the SIRV PDF for radar clutter has been derived from a generalization of the central limit theorem in which the number of scatterers in a range-azimuth cell from CPI to CPI is assumed to be random with a Poisson mixture distribution.[6]

Some general results related to the optimal detection of radar signals in SIRV clutter have been reported, with application to K-distributed clutter.[1,6–9] Design and performance of the LOD in Student t and K-distributed SIRV clutter[1] and the GLRT in K-distributed SIRV clutter[6,7] have also been presented. While the LOD and GLRT receivers have desirable characteristics, they are suboptimum for the detection of targets with random amplitude and phase parameters. A suboptimum channelized implementation of the optimal receiver for detecting targets that have an initial random phase which is uniformly distributed on $(0, 2\pi)$ has been considered.[8,9]

This chapter considers optimal detection of signals with $U(0, 2\pi)$ random phase in SIRV clutter. New results which illustrate the role of the matched filter and a whitening transformation in the optimum non-Gaussian SIRV receiver are presented. Furthermore, the first examples of closed-form expressions for the likelihood-ratio test of optimum receivers to detect a random amplitude and phase signal in non-Gaussian SIRV clutter are presented.

15.6.2. A Review of the SIRV Model

Recent research into the characteristics of the SIRV model has demonstrated its usefulness in approximating the joint probability density function of correlated, non-Gaussian radar clutter samples. The SIRV model has many nice properties, which can be attributed to its special relationship with Gaussian random vectors. A description of the SIRV model, examples, and some relevant properties are briefly presented in this chapter. Proofs and more detailed discussion of these and several other properties can be found in the references.[3,4,6]

15.6.2.1. Definition of the SIRV Model

The Representation Theorem. Every SIRV, $\mathbf{X} = [X_1 \ldots X_N]^{\mathrm{T}}$, with mean vector, $\mathbf{m}$, and covariance matrix, $\mathbf{\Sigma}$, can be represented in the form,

$$\mathbf{X} = S\mathbf{Z} + \mathbf{m} \tag{15.124}$$

where $\mathbf{Z} = [Z_1 \ldots Z_N]^{\mathrm{T}}$ is a zero-mean Gaussian random vector, also with covariance matrix, $\boldsymbol{\Sigma}$, and S is a nonnegative random variable which is independent of $\mathbf{Z}$ and $\mathbf{m}$. The mean-square value of S is intentionally assumed to be unity, so that the covariance matrix of the SIRV is equal to the covariance matrix of the underlying Gaussian random vector.

The representation theorem in Equation 15.124 provides a very useful description of the SIRV model because it relates non-Gaussian SIRVs to Gaussian random vectors. It clearly illustrates that $\mathbf{X}$, conditioned on a given $S = s$, is a Gaussian random vector with mean, $\mathbf{m}$, and covariance matrix, $s^2\boldsymbol{\Sigma}$. Hence, the conditional PDF of $\mathbf{X}|(S = s)$ is the multivariate Gaussian density function,

$$f_{\mathbf{X}|S}(\mathbf{x}|s) = (2\pi)^{-(N/2)}|s^2\boldsymbol{\Sigma}|^{-(1/2)}\exp\left(-\frac{(\mathbf{x}-\mathbf{m})^{\mathrm{T}}(s^2\boldsymbol{\Sigma})^{-1}(\mathbf{x}-\mathbf{m})}{2}\right) \quad (15.125)$$

The probability density function of $\mathbf{X}$ is obtained by substituting Equation 15.125 into

$$f_{\mathbf{X}}(\mathbf{x}) = \int_0^\infty f_{\mathbf{X}|S}(\mathbf{x}|s)f_S(s)\mathrm{d}s \quad (15.126)$$

which results in

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-(N/2)}|\boldsymbol{\Sigma}|^{-(1/2)}h_N(q) \quad (15.127)$$

where $h_N(\cdot)$ is defined by

$$h_N(q) = \int_0^\infty s^{-N}\exp\left(-\frac{q}{2s^2}\right)f_S(s)\mathrm{d}s \quad (15.128)$$

and q is the quadratic form,

$$q = (\mathbf{x}-\mathbf{m})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\mathbf{m}) \quad (15.129)$$

The probability density function, $f_S(s)$, of the random scale variable, S, is called the characteristic PDF of the SIRV. Different types of SIRVs can be modeled by changing the characteristic PDF.

The mixture model Equation 15.124 admits an interpretation which is consistent with observations often made on real clutter data. Clutter returns over a CPI from a given range-azimuth cell are usually Gaussian ($\mathbf{Z}$), but the average clutter power level (S^2) varies among cells in the surveillance volume. One consequence of Equation 15.128 is that $h_N(q)$ is a positive, monotonic decreasing function of q for all N. In addition, the PDF of any N-dimensional SIRV is uniquely and completely specified by a mean vector, covariance matrix, and either the characteristic PDF or $h_N(\cdot)$.

Finally, while the contours of constant density of $f_{\mathbf{X}}(\mathbf{x})$ are *ellipsoidal*, $\mathbf{X}$ is still frequently referred to as an *spherically* invariant random vector. This latter

terminology is chosen because $\mathbf{X}$ can always be transformed into a vector with spheroidal contours of constant density.

15.6.2.2. SIRV Properties

15.6.2.2.1. Linear Transformation

Let $\mathbf{X}$ be an SIRV with covariance matrix $\mathbf{\Sigma}_X$, mean vector $\mathbf{m}_X$, and characteristic PDF, $f_S(s)$. Then, if $\mathbf{A}\mathbf{A}^{\mathrm{T}}$ is nonsingular, the vector, $\mathbf{Y}$, defined by the linear transformation,

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{b} \tag{15.130}$$

is also an SIRV with the same characteristic PDF, a new mean value,

$$\mathbf{m}_Y = \mathbf{A}\mathbf{m}_X + \mathbf{b} \tag{15.131}$$

and a new covariance matrix,

$$\mathbf{\Sigma}_Y = \mathbf{A}\mathbf{\Sigma}_X\mathbf{A}^{\mathrm{T}} \tag{15.132}$$

Consequently, the class of SIRVs is closed under linear transformations.

This result is significant to non-Gaussian clutter modeling for two reasons. First, since a linear transformation is reversible, preprocessing to whiten the received data can be performed without loss of optimality. Second, the marginal PDF of any lower dimensional vector, which can be obtained by a linear transformation, is also an SIRV with the same characteristic PDF as the full vector.

15.6.2.2.2. Lack of Additive Closure for SIRVs

Let $\mathbf{X}_1$ and $\mathbf{X}_2$ be two statistically independent SIRVs with covariance matrices, $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$, respectively. The sum, $\mathbf{X}_1 + \mathbf{X}_2$ is not necessarily another SIRV, hence, the class of SIRVs is not closed under addition. Two circumstances under which this sum does yield another SIRV are:

1. $\mathbf{X}_1$ and $\mathbf{X}_2$ are both Gaussian random vectors.
2. The covariance matrices, $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$, are related by $\mathbf{\Sigma}_1 = k\mathbf{\Sigma}_2$, where k is any positive constant. The characteristic PDF can be different for each SIRV.

15.6.2.2.3. The Bootstrap Property

The probability density function of an N-dimensional SIRV may be obtained from the lower order density functions by using the recursive relation,

$$h_{2m+1}(\xi) = (-2)^m \frac{\mathrm{d}^m h_1(\xi)}{\mathrm{d}\xi^m}, \quad m = 0, 1, 2, \ldots \tag{15.133}$$

for odd orders, and

$$h_{2m+2}(\xi) = (-2)^m \frac{\mathrm{d}^m h_2(\xi)}{\mathrm{d}\xi^m}, \quad m = 0, 1, 2, \ldots \tag{15.134}$$

for even orders. Thus, only $h_1(\xi)$ and $h_2(\xi)$ are required to determine all the higher dimensional density functions for a particular type of SIRV.

It is sometimes desirable to define an SIRV probability density function which has a specific univariate marginal density function. The appropriate choice for the characteristic PDF is not always apparent. However, it is usually possible to determine $h_1(\xi)$ and $h_2(\xi)$ directly from the specified marginal PDF. Then, any higher order PDF may be obtained from the above recursions, *provided* that $h_N(\xi)$ derived in this way is a positive, monotonic decreasing function of any N. Otherwise, it is not possible to define a valid SIRV PDF which has the desired marginal density function.

15.6.2.2.4. Spherical Coordinate Representation

Any zero-mean random vector, $\mathbf{X} = [X_1 \ldots X_N]^{\mathrm{T}}$, with identity covariance matrix can be expressed as

$$X_1 = R\cos(\Phi_1) \tag{15.135}$$

$$X_k = R\cos(\Phi_k)\prod_{i=1}^{k-1}\sin(\Phi_i), \quad k = 2, \ldots, N-2 \tag{15.136}$$

$$X_{N-1} = R\cos(\Theta)\prod_{i=1}^{N-2}\sin(\Phi_i) \tag{15.137}$$

$$X_N = R\sin(\Theta)\prod_{i=1}^{N-2}\sin(\Phi_i) \tag{15.138}$$

where R, Θ, and $\Phi_1, \ldots, \Phi_{N-2}$ are random variables which uniquely specify the vector in a generalized, N dimensional spherical coordinate system. Such a vector, $\mathbf{X}$, is an SIRV if and only if the random variables, R, Θ, and $\Phi_1, \ldots, \Phi_{N-2}$ are all statistically independent and have the respective PDFs,

$$f_R(r) = \frac{r^{N-1}}{2^{(N/2)-1}\Gamma(N/2)} h_N(r^2), \quad 0 \le r < \infty \tag{15.139}$$

$$f_{\Phi_k}(\phi_k) = \frac{\Gamma\left(\dfrac{N-k+1}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\dfrac{N-k}{2}\right)} \sin^{N-k-1}(\phi_k) \tag{15.140}$$
$$0 \le \phi_k \le \pi, \quad k = 1, \ldots, N-2$$

$$f_\Theta(\theta) = \frac{1}{2\pi}, \quad 0 \le \theta < 2\pi \tag{15.141}$$

where $\Gamma(\cdot)$ denotes the gamma function.

The distance, $R = \sqrt{\mathbf{X}^T\mathbf{X}}$, is the only random spherical coordinate with a probability density function that varies for different SIRVs. The PDFs of the other $N - 1$ random spherical angles remain invariant to the type of SIRV which is considered. Thus, only the PDF of R is needed to characterize an SIRV with zero-mean and identity covariance matrix.

The generalization of spherical coordinates of N dimensions is not unique.[17,18,21,22] The possible generalizations differ only in how the set of $N - 1$ spherical angles and their PDFs are defined. The distance, R, remains unchanged in all of these coordinate systems. Therefore, any convenient generalization may be used in place of the one described above to specify an SIRV in spherical coordinates.

15.6.2.2.5. PDF of the Quadratic Form

For any SIRV, $\mathbf{Y}$, with mean vector, $\mathbf{m}$, and covariance matrix, $\mathbf{\Sigma}$, the PDF of the quadratic form,

$$Q = (\mathbf{Y} - \mathbf{m})^T\mathbf{\Sigma}^{-1}(\mathbf{Y} - \mathbf{m}) \tag{15.142}$$

is given by

$$f_Q(q) = \frac{q^{(N/2)-1}}{2^{N/2}\Gamma(N/2)} h_N(q) \tag{15.143}$$

This follows from Equation 15.139 because the linear transformation property implies

$$(\mathbf{Y} - \mathbf{m})^T\mathbf{\Sigma}^{-1}(\mathbf{Y} - \mathbf{m}) = \mathbf{X}^T\mathbf{X} = R^2 \tag{15.144}$$

where $\mathbf{X}$ is a zero-mean SIRV with identity covariance matrix.

Since $h_N(q)$ is unique for each type of SIRV, the multivariate density function for a particular type of SIRV can be uniquely determined based upon the univariate density function of its quadratic form. This property significantly reduces the complexity of the PDF approximation that must be performed for optimal non-Gaussian processing.

15.6.2.2.6. Unimodality

Since $h_N(q)$ is a positive, monotonic decreasing function for all N, and q describes contours of constant density for any SIRV, then clearly, the probability density function in Equation 15.127 is unimodal for all SIRVs. The peak value of the density function is $(2\pi)^{-(N/2)}|\mathbf{\Sigma}|^{-(1/2)}h_N(0)$, which occurs at the mean value of the SIRV.

15.6.2.2.7. Statistical Independence

If the components of an SIRV are statistically independent, then that SIRV must be Gaussian. It is not possible for any other types of SIRVs to have independent components.

15.6.2.3. The Complex SIRV Model

Sometimes it is more convenient to work with the N dimensional complex vector, $\tilde{\mathbf{Y}} = \mathbf{Y}_c + j\mathbf{Y}_s$, rather than the $2N$ dimensional real vector, $[\mathbf{Y}_c^T, \mathbf{Y}_s^T]^T$. Under certain conditions, the two approaches are equivalent and either one may be used. Otherwise, the real vector model should be used. A brief description of the complex SIRV Model in relation to complex Gaussian random vectors and the previous results for real SIRVs is given here.

Let $[\mathbf{Y}_c^T, \mathbf{Y}_s^T]^T$ be a $2N$ dimensional SIRV with mean vector, $[m_c^T, m_s^T]^T$, and covariance matrix,

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{\Gamma} & -\mathbf{\Phi} \\ \mathbf{\Phi} & \mathbf{\Gamma} \end{pmatrix} \tag{15.145}$$

where $\mathbf{\Gamma}$ is positive definite and $\mathbf{\Phi}$ is skew symmetric ($\mathbf{\Phi} = -\mathbf{\Phi}^T$). Then $\tilde{\mathbf{Y}} = \mathbf{Y}_c + j\mathbf{Y}_s$ is a complex SIRV with mean,

$$\tilde{\mathbf{m}} = \mathbf{m}_c + j\mathbf{m}_s \tag{15.146}$$

and the positive definite, Hermitian covariance matrix,

$$\tilde{\mathbf{\Sigma}} = 2(\mathbf{\Gamma} + j\mathbf{\Phi}) \tag{15.147}$$

Furthermore, the complex representation theorem, analogous to Equation 15.124, is

$$\tilde{\mathbf{Y}} = S\tilde{\mathbf{Z}} + \tilde{\mathbf{m}} \tag{15.148}$$

where $\tilde{\mathbf{Z}}$ is a complex, zero-mean, Gaussian random vector with covariance matrix, $\tilde{\mathbf{\Sigma}}$, and S is a real nonnegative random variable which is independent of $\tilde{\mathbf{Z}}$ and $\tilde{\mathbf{m}}$. Again, S is assumed to have unit-mean-square value so that the covariance matrices of $\tilde{\mathbf{Y}}$ and $\tilde{\mathbf{Z}}$ are equal.

The probability density functions of the N dimensional complex SIRV, $\tilde{\mathbf{Y}}$, is then found to be

$$f_{\tilde{Y}}(\tilde{y}) = \pi^{-N}|\tilde{\mathbf{\Sigma}}|^{-1}h_{2N}(2q) \tag{15.149}$$

where $h_{2N}(2q)$ is obtained from Equation 15.128 and given by

$$h_{2N}(2q) = \int_0^\infty s^{-2N}\exp\left(-\frac{q}{s^2}\right)f_S(s)\mathrm{d}s \tag{15.150}$$

with q denoting the quadratic form,

$$q = (\tilde{y} - \tilde{m})^H \tilde{\boldsymbol{\Sigma}}^{-1} (\tilde{y} - \tilde{m}) \tag{15.151}$$

The notation, $\tilde{x}^H$, denotes the Hermitian transpose of $\tilde{x}$.

15.6.2.4. Examples

Some univariate probability density functions which have multivariate generalizations corresponding to SIRVs include Weibull, Student t, chi, and K-distributed. Specific expressions for these and other SIRV density functions are given in Ref. 3.

15.6.3. Optimal Detection in NonGaussian SIRV Clutter

15.6.3.1. Introduction

Target detection in clutter for a transmitted radar waveform of N coherent pulses is described by the binary hypothesis problem,

$$\begin{aligned} H_0 &: \tilde{r} = \tilde{d} \text{ (target absent)} \\ H_1 &: \tilde{r} = a e^{j\phi} \tilde{s} + \tilde{d} \text{ (target present)} \end{aligned} \tag{15.152}$$

where the elements of the N dimensional vectors, $\tilde{r}$, $\tilde{d}$, and $\tilde{s}$, are low-pass complex enveloped samples of the received data, the disturbance (clutter plus background noise), and the desired signal, respectively. Signal attenuation and target reflection characteristics are modeled by the target amplitude parameter, a. The initial phase of the received waveform, which is pulse-to-pulse coherent, is represented by ϕ. This model has implicitly assumed that the time of arrival and doppler shift of the target return are known.

Complete knowledge and a and ϕ is usually unavailable. However, the optimal receiver for the completely known signal provides the basis for optimal and near-optimal detection when a and ϕ are modeled as either random variables or unknown constants. The optimal known signal receiver in non-Gaussian SIRV clutter is presented in Section 15.6.3.2, where it is also used to introduce a canonical structure for optimal SIRV receivers which also applies to the optimal and near-optimal receivers developed thereafter.

The primary emphasis of the research presented in this chapter is on optimal and near-optimal detection of target return signals with *random* amplitude and phase in non-Gaussian SIRV interference. Attention is focused on this problem for two major reasons. First, a signal with random phase and amplitude is usually the most realistic model for radar target returns. Second, this type of detection problem for SIRV clutter has not been adequately addressed in the literature.

General solutions for the optimum SIRV receiver are developed in Section 15.6.3.3 and illustrated by some examples for specific SIRV clutter models.

Suboptimum receivers which have near-optimal performance are investigated in Section 15.6.3.4 and Section 15.6.3.5.

Two significant assumptions are made about the statistics of the non-Gaussian SIRV clutter for the receiver development presented here. First, attention is focused only on situations where the disturbance is dominated by clutter, with the clutter power being much greater than the background Gaussian noise power. The PDF of the disturbance is then closely approximated by the PDF of the SIRV clutter. Second, the covariance matrix of the received clutter samples is assumed to be known. In practice, the clutter covariance matrix is usually unknown and changing. The impact of both assumptions is discussed latter.

15.6.3.2. Completely Known Signals

The Neyman–Pearson (NP) receiver is optimum for the detection problem of Equation 15.152. It maximizes probability of detection, P_D, for a specified probability of false alarm, P_{FA}. The NP receiver for a completely known signal is the likelihood-ratio test (LRT),[10]

$$T(\tilde{r}|a,\phi) = \frac{f_{\tilde{R}|H_1}(\tilde{r}|H_1)}{f_{\tilde{R}|H_0}(\tilde{r}|H_0)} = \frac{f_{\tilde{D}}(\tilde{r} - a\mathrm{e}^{\mathrm{j}\phi}\tilde{s})}{f_{\tilde{D}}(\tilde{r})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{15.153}$$

where $f_{\tilde{D}}(\tilde{d})$ is the PDF of the disturbance and the threshold, η, is determined from the design constraint on the probability of false alarm.

For the disturbance dominated by clutter which is assumed to be a zero-mean, complex SIRV with known covariance, matrix, $\tilde{\boldsymbol{\Sigma}}$, and PDF described by Equation 15.149, the optimum NP receiver of Equation 15.153 is

$$T(\tilde{r}|a,\phi) = \frac{h_{2N}[2(\tilde{r} - a\mathrm{e}^{\mathrm{j}\phi}\tilde{s})^H\tilde{\boldsymbol{\Sigma}}^{-1}(\tilde{r} - a\mathrm{e}^{\mathrm{j}\phi}\tilde{s})]}{h_{2N}[2\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}]} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{15.154}$$

The specific form of Equation 15.154 depends on the type of SIRV clutter which is present.

However, the general form of this optimum receiver for any SIRV has a canonical structure which incorporates the test statistic of the conventional Gaussian receiver. This canonical form, shown in Figure 15.47, is easily demonstrated by expanding the quadratic form in the numerator of Equation 15.154 to yield

$$T(\tilde{r}|a,\phi) = \frac{h_{2N}[2(\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r} - 2a\mathrm{Re}\{\mathrm{e}^{-\mathrm{j}\phi}\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}\} + a^2\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{s})]}{h_{2N}[2\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}]} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{15.155}$$

and recognizing that

$$T_g(\tilde{r}) = \mathrm{Re}\{\mathrm{e}^{-\mathrm{j}\phi}\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}\} \tag{15.156}$$

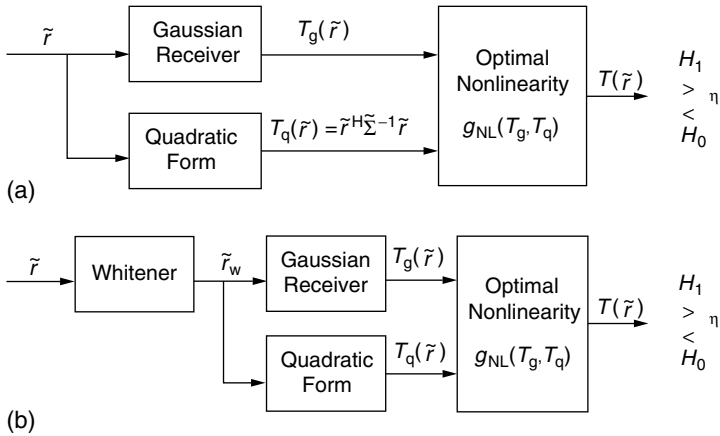

FIGURE 15.47 Optimal SIRV receiver: (a) canonical form, (b) equivalent canonical form with a single whitening filter.

is the optimum Gaussian test statistic. [The signal amplitude, a, is not included in this expression, because the optimum Gaussian statistic is uniformly most powerful (UMP) with respect to a and can be implemented without knowledge of its value.] The optimum SIRV receiver can now be written as

$$T(\tilde{r}|a,\phi) = g_{\mathrm{NL}}[\mathrm{Re}\{\mathrm{e}^{-\mathrm{j}\phi}\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}\}, \tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}] = g_{\mathrm{NL}}[T_{\mathrm{g}}(\tilde{r}), T_{\mathrm{q}}(\tilde{r})] \quad (15.157)$$

where $T_{\mathrm{q}}(\tilde{r})$ is defined as

$$T_{\mathrm{q}}(\tilde{r}) = \tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r} \quad (15.158)$$

and $g_{\mathrm{NL}}[\cdot,\cdot]$ is used to denote the nonlinear function of two arguments in Equation 15.155. This receiver is also a special case of the NP receivers considered in the derivation of Appendix Y.

The LRT given in Equation 15.153 is usually not sufficient for practical radar applications because the amplitudes and phase parameters are seldom known. However, the performance of this receiver, which assumes perfect *a priori* knowledge about the signal parameters, provides an upper bound on the detection performance when the signal parameters are not completely known. This performance bound is often referred to as the perfect measurement bound[10] or the envelope power function.[2]

15.6.3.3. Signals with Random Parameters

A more realistic target model assumes the target amplitude and phase are statistically independent random variables with PDFs, $f_A(a)$ and $f_\Phi(\phi)$, respectively. If a and ϕ are assumed to remain constant for a single CPI, then

the optimum NP receiver can be expressed as the average,[10]

$$T(\tilde{r}) = \int_{\Omega_\phi} \int_{\Omega_a} T(\tilde{r}|a,\phi) f_A(a) f_\Phi(\phi) \mathrm{d}a \, \mathrm{d}\phi \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{15.159}$$

which becomes

$$T(\tilde{r}) = \int_{\Omega_\phi} \int_{\Omega_a} \frac{h_{2N}[2(\tilde{r} - a\mathrm{e}^{\mathrm{j}\phi}\tilde{s})^H \tilde{\mathbf{\Sigma}}^{-1} (\tilde{r} - a\mathrm{e}^{\mathrm{j}\phi}\tilde{s})]}{h_{2N}[2\tilde{r}^H \tilde{\mathbf{\Sigma}}^{-1} \tilde{r}]} f_A(a) f_\Phi(\phi) \mathrm{d}a \, \mathrm{d}\phi \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{15.160}$$

for SIRV clutter when Equation 15.154 is substituted for $T(\tilde{r}|a,\phi)$.

It is shown in Appendix Y that this receiver also has the canonical form of Figure 15.47 for any SIRV, whenever the optimum Gaussian test statistic, $T_\mathrm{g}(\tilde{r})$, satisfies the sufficient condition,

$$T_\mathrm{g}(c\tilde{r}) = f(c) T_\mathrm{g}(\tilde{r}) \tag{15.161}$$

where c does not depend on $\tilde{r}$ and $f(\cdot)$ is any function. This condition is satisfied by the Gaussian receiver for many commonly used target models, including a $U(0, 2\pi)$ distribution for ϕ and any amplitude PDF, $f_A(a)$.

The canonical structure in Figure 15.47a is a significant result. It indicates that the optimal non-Gaussian SIRV receiver should incorporate the Gaussian receiver which is currently implemented in existing radar systems. Thus, techniques which have been developed to achieve optimal (or near-optimal) processing in Gaussian clutter are required to implement optimal non-Gaussian SIRV processing.

An equivalent representation of the canonical receiver structure is given in Figure 15.47b. The quadratic form, $\tilde{r}^H \tilde{\mathbf{\Sigma}}^{-1} \tilde{r}$, is equivalent to $\tilde{r}_\mathrm{w}^H \tilde{r}_\mathrm{w}$, where $\tilde{r}_\mathrm{w} = \tilde{\mathbf{\Sigma}}^{-1/2} \tilde{r}$ is obtained by passing $\tilde{r}$ through a whitener. The same whitening operation is also inherent to the Gaussian receiver. These identical whiteners are replaced by a single whitening filter at the input, as shown in Figure 15.47b. This equivalence is a consequence of the linear transformation property for SIRVs discussed in Section 15.6.2.2.

This alternative canonical form is significant because it indicates that detection performance depends upon only the signal energy and not the signal shape. Thus, detection performance in correlated clutter can be obtained by adding the processing gain of the whitening filter onto the input SCR, and then evaluating detection performance for the modified SCR in uncorrelated clutter.

The optimal Gaussian test statistic is uniformly most powerful (UMP) with respect to the target amplitude. Hence, the design of the optimal Gaussian receiver is independent of the probability density function of the random amplitude, a. Only the detection performance of the Gaussian receiver depends on the target amplitude model.

In contrast, the design of the optimal non-Gaussian SIRV receiver obtained from Equation 15.159 does depend on the target amplitude PDF. Thus, uncertainty in the target amplitude characteristics results in a detection performance loss.

The integrations necessary to determine the optimal SIRV receiver for a particular type of SIRV are usually very difficult, if not impossible, to evaluate in closed-form. The case of known amplitude, random phase signal is considered first, since the integration with respect to the phase has certain similarities to the Gaussian problem. Then the problem of both random amplitude and random phase is addressed for a channelized structure involving the discrete Gaussian mixture (DGM) SIRV.

15.6.3.3.1. Signals with Random Phase

The optimum NP receiver for the known amplitude, random phase target takes the form,

$$T(\tilde{r}|a) = \int_{\Omega_\phi} \frac{h_{2N}[2(\tilde{r} - a\mathrm{e}^{\mathrm{j}\phi}\tilde{s})^H \tilde{\mathbf{\Sigma}}^{-1}(\tilde{r} - a\mathrm{e}^{\mathrm{j}\phi}\tilde{s})]}{h_{2N}[2\tilde{r}^H \tilde{\mathbf{\Sigma}}^{-1}\tilde{r}]} f_\Phi(\phi)\mathrm{d}\phi \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{15.162}$$

The quadratic form in the numerator of the integrand is expanded to

$$\begin{aligned}&(\tilde{r} - a\mathrm{e}^{\mathrm{j}\phi}\tilde{s})^H \tilde{\mathbf{\Sigma}}^{-1}(\tilde{r} - a\mathrm{e}^{\mathrm{j}\phi}\tilde{s}) \\ &\quad = \tilde{r}^H \tilde{\mathbf{\Sigma}}^{-1}\tilde{r} - 2a\,\mathrm{Re}\{\mathrm{e}^{-\mathrm{j}\phi}\tilde{s}^H \tilde{\mathbf{\Sigma}}^{-1}\tilde{r}\} + a^2 \tilde{s}^H \tilde{\mathbf{\Sigma}}^{-1}\tilde{s}\end{aligned} \tag{15.163}$$

If the complex cross product, $\tilde{s}^H \tilde{\mathbf{\Sigma}}^{-1}\tilde{r}$, is defined to be

$$\tilde{s}^H \tilde{\mathbf{\Sigma}}^{-1}\tilde{r} = L_\mathrm{c} + jL_\mathrm{s} \tag{15.164}$$

then the quadratic form expansion in Equation 15.163 becomes

$$\begin{aligned}&(\tilde{r} - a\mathrm{e}^{\mathrm{j}\phi}\tilde{s})^H \tilde{\mathbf{\Sigma}}^{-1}(\tilde{r} - a\mathrm{e}^{\mathrm{j}\phi}\tilde{s}) \\ &\quad = \tilde{r}^H \tilde{\mathbf{\Sigma}}^{-1}\tilde{r} - 2a(L_\mathrm{c}\cos\phi + L_\mathrm{s}\sin\phi) + a^2 \tilde{s}^H \tilde{\mathbf{\Sigma}}^{-1}\tilde{s}\end{aligned} \tag{15.165}$$

Substituting this into Equation 15.162 yields

$$\begin{aligned}&T(\tilde{r}|a) \\ &\quad = \int_{\Omega_\phi} \frac{h_{2N}[2(\tilde{r}^H \tilde{\mathbf{\Sigma}}^{-1}\tilde{r} - 2a(L_\mathrm{c}\cos\phi + L_\mathrm{s}\sin\phi) + a^2 \tilde{s}^H \tilde{\mathbf{\Sigma}}^{-1}\tilde{s})]}{h_{2N}[2\tilde{r}^H \tilde{\mathbf{\Sigma}}^{-1}\tilde{r}]} f_\Phi(\phi)\mathrm{d}\phi\end{aligned} \tag{15.166}$$

for the optimum test statistic. This result is similar to one obtained for the Gaussian noise problem (in Ref. 10, p. 337).

A useful model for the random signal phase is the Viterbi phase density,[10,11]

$$f_\Phi(\phi) = \frac{\exp(\Lambda_m \cos\phi)}{2\pi I_0(\Lambda_m)}, \quad -\pi \le \phi \le \pi \tag{15.167}$$

where $I_0(\cdot)$ denotes the modified Bessel function of the first kind of order 0 and Λ_m is a shape parameter for the PDF. This phase density is uniformly distributed when Λ_m is zero, becomes more peaked as Λ_m increases, and tends to the known phase case when Λ_m approaches infinity. Substitution of this PDF into

Equation 15.166 gives

$$T(\tilde{r}|a) = \int_{-\pi}^{\pi} \frac{h_{2N}[2(\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r} - 2a(L_c\cos\phi + L_s\sin\phi) + a^2\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{s})]}{h_{2N}[2\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}]} \times \frac{\exp(\Lambda_m\cos\phi)}{2\pi I_0(\Lambda_m)}\mathrm{d}\phi \tag{15.168}$$

By definition, the monotonic decreasing function, $h_{2N}(\xi)$, is

$$h_{2N}(\xi) = \int_0^{\infty} s^{-2N}\exp\left(-\frac{\xi}{2s^2}\right)f_S(s)\mathrm{d}s \tag{15.169}$$

where s is the random scalar of the SIRV representation theorem with PDF, $f_S(s)$. Applying this definition for $h_{2N}(\xi)$ to the numerator of Equation 15.168 results in

$$\begin{aligned} T(\tilde{r}|a) = &\int_{-\pi}^{\pi} \frac{1}{h_{2N}[2\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}]}\frac{\exp(\Lambda_m\cos\phi)}{2\pi I_0(\Lambda_m)} \\ &\times\int_0^{\infty} s^{-2N}\exp\left(-\frac{\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r} - 2a(L_c\cos\phi + L_s\sin\phi) + a^2\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{s}}{s^2}\right) \\ &\times f_S(s)\mathrm{d}s\,\mathrm{d}\phi \end{aligned} \tag{15.170}$$

which, after interchanging the order of integration and combining the exponentials on ϕ, is

$$\begin{aligned} T(\tilde{r}|a) = &\int_0^{\infty} \frac{s^{-2N}f_S(s)}{h_{2N}[2\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}]}\exp\left(-\frac{\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r} + a^2\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{s}]}{s^2}\right) \\ &\times\int_{-\pi}^{\pi}\frac{1}{2\pi I_0(\Lambda_m)}\exp\left(\left(\Lambda_m + \frac{2aL_c}{s^2}\right)\cos\phi - \frac{2aL_s}{s^2}\sin\phi\right)\mathrm{d}\phi\,\mathrm{d}s \end{aligned} \tag{15.171}$$

The inner integral of Equation 15.171 is a standard form given by (in Ref. 25, p. 523)

$$\begin{aligned} &\int_{-\pi}^{\pi}\frac{1}{2\pi I_0(\Lambda_m)}\exp\left(\left(\Lambda_m + \frac{2aL_c}{s^2}\right)\cos\phi - \frac{2aL_s}{s^2}\sin\phi\right)\mathrm{d}\phi \\ &= \frac{1}{I_0(\Lambda_m)}I_0\left(\sqrt{\left(\Lambda_m + \frac{2aL_c}{s^2}\right)^2 + \left(\frac{2aL_s}{s^2}\right)^2}\right) \end{aligned} \tag{15.172}$$

The optimum receiver is now

$$T(\tilde{r}|a) = \int_0^\infty \frac{s^{-2N} f_S(s)}{I_0(\Lambda_m) h_{2N}[2\tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}]} I_0\left(\sqrt{\left(\Lambda_m + \frac{2aL_c}{s^2}\right)^2 + \left(\frac{2aL_s}{s^2}\right)^2}\right) \times \exp\left(-\frac{\tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r} + a^2 \tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{s}}{s^2}\right) \mathrm{d}s \tag{15.173}$$

which is of the general form,

$$T(\tilde{r}|a) = \int_0^\infty g(s) \exp\left(-\frac{k_{\tilde{r}}}{s^2}\right) \mathrm{d}s \tag{15.174}$$

where $k_{\tilde{r}}$ is defined to be

$$k_{\tilde{r}} = \tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r} + a^2 \tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{s} \tag{15.175}$$

and $g(s)$ is the remaining portion of the integrand in Equation 15.170,

$$g(s) = \frac{s^{-2N} f_S(s)}{I_0(\Lambda_m) h_{2N}[2\tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}]} I_0\left(\sqrt{\left(\Lambda_m + \frac{2aL_c}{s^2}\right)^2 + \left(\frac{2aL_s}{s^2}\right)^2}\right) \tag{15.176}$$

Applying the variable transformation, $s = t^{-1/2}$, to Equation 15.174 results in

$$T(\tilde{r}|a) = \int_0^\infty \frac{1}{2} t^{-3/2} g(t^{-1/2}) \exp(-k_{\tilde{r}} t) \mathrm{d}t \tag{15.177}$$

which is recognized as the Laplace transform,[24]

$$T(\tilde{r}|a) = \mathscr{L}\left\{\frac{1}{2} t^{-3/2} g(t^{-1/2})\right\} \tag{15.178}$$

The final form of the optimum SIRV receiver for the random Viterbi phase signal with known amplitude, a, is obtained by substituting Equation 15.176 into Equation 15.178. This yields

$$T(\tilde{r}|a) = \frac{\mathscr{L}\{t^{N-(3/2)} f_S(t^{-1/2}) I_0(\sqrt{(\Lambda_m + 2aL_c t)^2 + (2aL_s t)^2})\}}{2 I_0(\Lambda_m) h_{2N}[2\tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}]} \tag{15.179}$$

where $\mathscr{L}\{\cdot\}$ is the one-sided Laplace transform operator with frequency variable, $k_{\tilde{r}} = \tilde{r}^H \boldsymbol{\Sigma}^{-1} \tilde{r} + a^2 \tilde{s}^H \boldsymbol{\Sigma}^{-1} \tilde{s}$, and $f_S(s)$ is the characteristic PDF of the SIRV.

In general, the result in Equation 15.179 does not have a closed-form solution. However, for the case $\Lambda_m = 0$, which corresponds to a random phase variable which is uniformly distributed on $(-\pi, \pi)$, the general result reduces to

$$T(\tilde{r}|a) = \frac{\mathscr{L}\{t^{N-(3/2)} f_S(t^{-1/2}) I_0(2a\sqrt{L_c^2 + L_s^2} t)\}}{2 h_{2N}[2\tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}]} \tag{15.180}$$

Since L_c and L_s are defined by Equation 15.164, this optimum receiver can be expressed as

$$T(\tilde{r}|a) = \frac{\mathscr{L}\{t^{N-(3/2)} f_S(t^{-1/2}) I_0(2a|\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}|t)\}}{2h_{2N}[2\tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}]} \tag{15.181}$$

The Student t and DGM SIRVs are two examples for which a closed-form solution to Equation 15.180 does exist. The detailed solutions for these cases are now presented.

15.6.3.3.1.1. Student t Example

The characteristic PDF of the Student t SIRV is[3]

$$f_S(s) = \frac{2b^{2\nu}}{2^\nu \Gamma(\nu)} s^{-(2\nu+1)} \exp\left(-\frac{b^2}{2s^2}\right), \quad s \geq 0, \ \nu > 0 \tag{15.182}$$

where ν is a shape parameter and b is a scale parameter. Substituting this characteristic PDF into Equation 15.181, which is the optimum receiver for a $U(0, 2\pi)$ random phase signal, leads to

$$T(\tilde{r}|a) = \frac{b^{2\nu} \mathscr{L}\left\{t^{N+\nu-1} I_0(2a|\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}|t) \exp\left(-\frac{b^2 t}{2}\right)\right\}}{2^\nu \Gamma(\nu) h_{2N}[2\tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}]} \tag{15.183}$$

The $\exp(-b^2 t/2)$ factor in the argument of the Laplace transform is handled by the shifting property

$$\mathscr{L}\left\{\exp\left(-\frac{b^2 t}{2}\right) g(t)\right\} = G\left(k_{\tilde{r}} + \frac{b^2}{2}\right) \tag{15.184}$$

where $G(k_{\tilde{r}})$ is the Laplace transform of $g(t)$.

The solution to Equation 15.183 is now reduced to finding the transform,

$$G(k_{\tilde{r}}) = \mathscr{L}\{t^{N+\nu-1} I_0(2a|\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}|t)\} \tag{15.185}$$

This is similar to the standard Laplace transform (in Ref. 23, p. 149),

$$\begin{aligned} &\mathscr{L}\{t^{\mu_*} I_{\nu_*}(a_* t)\} \\ &\quad = \Gamma(\nu_* + \mu_* + 1)(k_{\tilde{r}}^2 - a_*^2)^{-(\mu_*+1)/2} P_{\mu_*}^{-\nu_*}[k_{\tilde{r}}(k_{\tilde{r}}^2 - a_*^2)^{-1/2}] \end{aligned} \tag{15.186}$$

where $P_{\mu_*}^{-\nu_*}[\cdot]$ is the associated Legendre function of the first kind, defined as

$$P_{\mu_*}^{-\nu_*}[z] = \frac{1}{\Gamma(1+\nu_*)}\left(\frac{z+1}{z-1}\right)^{-(\nu_*/2)} {}_2F_1\left(-\mu_*, \mu_*+1; \nu_*+1; \frac{1-z}{2}\right) \quad (15.187)$$

with ${}_2F_1(a, b; c; z)$ denoting the Gauss hypergeometric function. The two conditions which must be satisfied for Equation 15.186 to be a valid transform are[23]

$$\mathrm{Re}\{\nu_* + \mu_*\} > -1 \quad (15.188)$$

and

$$\mathrm{Re}\{k_{\tilde{r}}\} > |\mathrm{Re}\{a_*\}| \quad (15.189)$$

The standard transform of Equation 15.186 is related to the desired transform of Equation 15.185 by the parameter definitions,

$$\nu_* = 0 \quad (15.190)$$

$$\mu_* = N + \nu - 1 \quad (15.191)$$

and

$$a_* = 2a|\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}| \quad (15.192)$$

Substituting these parameter relations into the standard transform of Equation 15.186 results in

$$\mathcal{L}\left\{t^{N+\nu-1} I_0(2a|\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}|t)\right\} = \Gamma(N+\nu)[k_{\tilde{r}}^2 - (2a|\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}|)^2]^{-(N+\nu)/2} \times P_{N+\nu-1}\left(\frac{k_{\tilde{r}}}{\sqrt{k_{\tilde{r}}^2 - (2a|\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}|)^2}}\right) \quad (15.193)$$

where $P_\mu(\cdot)$ is the Legendre function of the first kind of degree μ. (Associated Legendre functions reduce to Legendre functions when the upper index is equal to zero, in which case the upper index is omitted.)

The monotonic decreasing function in the denominator of Equation 15.183 for the complex Student t SIRV is

$$h_{2N}(2\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}) = \frac{2^N b^{2\nu}\Gamma(\nu+N)}{\Gamma(\nu)(b^2+2\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r})^{\nu+N}} \tag{15.194}$$

Substituting this and Equation 15.193 into Equation 15.183 and applying the shifting property of Equation 15.184 leads to

$$\begin{aligned} T(\tilde{r}|a) = {} & \frac{b^{2\nu}}{2^\nu\Gamma(\nu)}\Gamma(N+\nu)\left[\left(k_{\tilde{r}}+\frac{b^2}{2}\right)^2-(2a|\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}|)^2\right]^{-(N+\nu)/2} \\ & \times P_{N+\nu-1}\left(\frac{k_{\tilde{r}}+\frac{b^2}{2}}{\sqrt{\left(k_{\tilde{r}}+\frac{b^2}{2}\right)^2-(2a|\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}|)^2}}\right)\left[\frac{\Gamma(\nu)(b^2+2\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r})^{\nu+N}}{2^N b^{2\nu}\Gamma(\nu+N)}\right] \end{aligned} \tag{15.195}$$

which simplifies to

$$\begin{aligned} T(\tilde{r}|a) = {} & \left(\frac{b^2}{2}+\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}\right)^{\nu+N}\left[\left(k_{\tilde{r}}+\frac{b^2}{2}\right)^2-(2a|\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}|)^2\right]^{-(N+\nu)/2} \\ & \times P_{N+\nu-1}\left(\frac{k_{\tilde{r}}+\frac{b^2}{2}}{\sqrt{\left(k_{\tilde{r}}+\frac{b^2}{2}\right)^2-(2a|\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}|)^2}}\right) \end{aligned} \tag{15.196}$$

The optimum receiver can be expressed completely in terms of the received data, r, by substituting Equation 15.175 for $k_{\tilde{r}}$ into the above equation. This results in

$$T(\tilde{r}|a) = \left(\frac{\left(\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}+\frac{b^2}{2}\right)\lambda_{\tilde{r}}}{\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}+\frac{b^2}{2}+a^2\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{s}}\right)^{\nu+N} P_{\nu+N-1}(\lambda_{\tilde{r}}) \tag{15.197}$$

where $\lambda_{\tilde{r}}$ is defined to be

$$\lambda_{\tilde{r}} = \left(\frac{\tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r} + \frac{b^2}{2} + a^2 \tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{s}}{\sqrt{\left(\tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r} + \frac{b^2}{2} + a^2 \tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{s} \right)^2 - (2a|\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}|)^2}} \right) \geq 1 \qquad (15.198)$$

b is the Student t scale parameter, ν is the Student t shape parameter, and a is the known target amplitude. Notice that Equation 15.197 is consistent with the canonical structure of Figure 15.47, where $|\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}|$ is identified as the Gaussian receiver output for the $U(0, 2\pi)$ random phase signal.

The Legendre function of the first kind of degree μ, $P_\mu(\cdot)$, reduces to a Legendre polynomial for integer μ. Thus, for integer values of the shape parameter, ν, $P_{\nu+N-1}(\lambda_{\tilde{r}})$ becomes a polynomial in $\lambda_{\tilde{r}}$.

It still remains to verify that conditions Equation 15.188 and Equation 15.189 are satisfied. If the parameter definitions of Equation 15.190 and Equation 15.191 are substituted into Equation 15.188, the first condition becomes

$$\mathrm{Re}\{N + \nu - 1\} > -1 \qquad (15.199)$$

This simplifies to

$$N + \nu > 0 \qquad (15.200)$$

which is obviously satisfied for all positive N and ν.

Since the shifting property of Equation 15.184 has been used to obtain the desired Laplace transform, the second condition to be satisfied is now

$$\mathrm{Re}\left\{ k_{\tilde{r}} + \frac{b^2}{2} \right\} > |\mathrm{Re}\{a_*\}| \qquad (15.201)$$

which is equivalent to

$$\mathrm{Re}\left\{ \tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r} + a^2 \tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{s} + \frac{b^2}{2} \right\} > |\mathrm{Re}\{2a|\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}|\}| \qquad (15.202)$$

The verification of this condition is more easily obtained by using the linearly transformed vectors, $\tilde{r}_\mathrm{w} = \tilde{\boldsymbol{\Sigma}}^{-1/2} \tilde{r}$ and $\tilde{s}_\mathrm{w} = \tilde{\boldsymbol{\Sigma}}^{-1/2} \tilde{s}$, in which case Equation 15.202 becomes

$$\mathrm{Re}\left\{ \|\tilde{r}_\mathrm{w}\|^2 + a^2 \|\tilde{s}_\mathrm{w}\|^2 + \frac{b^2}{2} \right\} > |\mathrm{Re}\{2a|\tilde{s}_\mathrm{w}^H \tilde{r}_\mathrm{w}|\}| \qquad (15.203)$$

This is further simplified to

$$\|\tilde{r}_w\|^2 + a^2\|\tilde{s}_w\|^2 + \frac{b^2}{2} > 2a|\tilde{s}_w^H \tilde{r}_w| \tag{15.204}$$

since the arguments of both $\mathrm{Re}\{\cdot\}$ operations in Equation 15.203 are real and positive.

The relation,

$$(\|\tilde{r}_w\| - a\|\tilde{s}_w\|)^2 = \|\tilde{r}_w\|^2 + a^2\|\tilde{s}_w\|^2 - 2a\|\tilde{s}_w\|\,\|\tilde{r}_w\| \geq 0 \tag{15.205}$$

implies

$$\|\tilde{r}_w\|^2 + a^2\|\tilde{s}_w\|^2 \geq 2a\|\tilde{s}_w\|\,\|\tilde{r}_w\| \tag{15.206}$$

Therefore, the condition in Equation 15.204 is satisfied whenever

$$2a\|\tilde{s}_w\|\,\|\tilde{r}_w\| + \frac{b^2}{2} > 2a|\tilde{s}_w^H \tilde{r}_w| \tag{15.207}$$

is true. The above relation is always satisfied, as seen by application of the Schwarz inequality,[29]

$$\|\tilde{s}_w\|\,\|\tilde{r}_w\| \geq |\tilde{s}_w^H \tilde{r}_w| \tag{15.208}$$

Thus, Equation 15.197 is a valid solution for the optimum receiver of a random phase signal in Student t SIRV clutter.

15.6.3.3.1.2. DGM Example

A DGM SIRV with K mixture components is generated by a discrete characteristic PDF of the form,

$$f_S(s) = \sum_{k=1}^{K} w_k \delta(s - s_k) \tag{15.209}$$

where $\delta(\cdot)$ is the impulse function and w_k, $k = 1, \ldots, K$, are probability weights which must be positive and sum to unity. The monotonic decreasing function, $h_{2N}(\xi)$, for this case is

$$h_{2N}(\xi) = \sum_{k=1}^{K} w_k s_k^{-2N} \exp\left(-\frac{\xi}{2s_k^2}\right) \tag{15.210}$$

The optimum receiver for detection of a signal with a random Viterbi phase density in DGM clutter can be obtained from Equation 15.179. However, because of the occurrence of impulse functions in $f_S(s)$, it is easier to use Equation 15.173 to find the optimum receiver.

Substituting Equation 15.209 into Equation 15.173 yields

$$T(\tilde{r}|a) = \int_0^\infty \frac{s^{-2N}}{I_0(\Lambda_m)h_{2N}[2\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}]} I_0\left(\sqrt{\left(\Lambda_m + \frac{2aL_c}{s^2}\right)^2 + \left(\frac{2aL_s}{s^2}\right)^2}\right) \times \exp\left(-\frac{\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r} + a^2\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{s}}{s^2}\right)\sum_{k=1}^K w_k\delta(s - s_k)\mathrm{d}s \tag{15.211}$$

Interchanging the order of the summation and integration results in

$$T(\tilde{r}|a) = \left[I_0(\Lambda_m)\sum_{k=1}^K w_k s_k^{-2N}\exp\left(-\frac{\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}}{s_k^2}\right)\right]^{-1} \times \sum_{k=1}^K w_k \int_0^\infty s^{-2N} I_0\left(\sqrt{\left(\Lambda_m + \frac{2aL_c}{s^2}\right)^2 + \left(\frac{2aL_s}{s^2}\right)^2}\right) \times \exp\left(-\frac{\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r} + a^2\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{s}}{s^2}\right)\delta(s - s_k)\mathrm{d}s \tag{15.212}$$

where Equation 15.210 has also been used. The optimum receiver now simplifies to

$$T(\tilde{r}|a) = \frac{\sum_{k=1}^K w_k s_k^{-2N}\exp\left(-\frac{\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r} + a^2\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{s}}{s_k^2}\right) I_0\left(\sqrt{\left(\Lambda_m + \frac{2aL_c}{s_k^2}\right)^2 + \left(\frac{2aL_s}{s_k^2}\right)^2}\right)}{I_0(\Lambda_m)\sum_{k=1}^K w_k s_k^{-2N}\exp\left(-\frac{\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}}{s_k^2}\right)} \tag{15.213}$$

where a is the known target amplitude, Λ_m is the Viterbi phase parameter and L_c and L_s are the real and imaginary parts of $\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}$, respectively.

A $U(0, 2\pi)$ random phase density is obtained for $\Lambda_m = 0$ and the optimum receiver in this case becomes

$$T(\tilde{r}|a) = \frac{\sum_{k=1}^K w_k s_k^{-2N}\exp\left(-\frac{a^2\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{s}}{s_k^2}\right)\exp\left(-\frac{\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}}{s_k^2}\right) I_0\left(\frac{2a|\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}|}{s_k^2}\right)}{\sum_{k=1}^K w_k s_k^{-2N}\exp\left(-\frac{\tilde{r}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}}{s_k^2}\right)} \tag{15.214}$$

The receiver in Equation 15.214 is also seen to have the canonical structure of Figure 15.47, where $|\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}|$ is identified as the Gaussian receiver output for the $U(0, 2\pi)$ random phase signal.

The DGM SIRV has the potential to closely approximate many SIRVs having continuous characteristic PDFs. This suggests the optimal receiver can be approximated by a receiver with the above structure, which is particularly useful when a closed-form solution for the optimal receiver cannot be obtained.

15.6.3.3.2. Signals with Random Amplitude and Phase

When both the phase, ϕ, and amplitude, a, of the target return are random, the optimum SIRV receiver given by Equation 15.160 can be expressed as

$$T(\tilde{r}) = \int_0^\infty T(\tilde{r}|a) f_A(a) \mathrm{d}a \tag{15.215}$$

A closed-form evaluation of this integral is usually not possible, even when $T(\tilde{r}|a)$ is known, and a near-optimal approximate receiver implementation is desired. A channelized approach which uses the DGM SIRV is discussed in this section and alternative approaches are discussed later.

The quantity, $v = a^2$, is directly related to the radar cross section and represents the amount of reflected energy received from the target. For convenience, the optimum receiver in Equation 15.215 is expressed in terms of the target energy as

$$T(\tilde{r}) = \int_0^\infty T(\tilde{r}|\sqrt{v}) f_V(v) \mathrm{d}v \tag{15.216}$$

The amplitude (or energy or radar cross section) of the target return is assumed to remain constant during a single CPI of N pulses. However, the amplitude is assumed to exhibit a random fluctuation across different CPIs. Specifically, the Swerling one and Swerling three-target models are considered.

The probability density function of v for a Swerling one model is

$$\text{Swerling 1:} \qquad f_V(v) = \frac{1}{\bar{v}} \exp\left(-\frac{v}{\bar{v}}\right), \quad v \geq 0 \tag{15.217}$$

which is the exponential density function, and the PDF for the Swerling three-target model is

$$\text{Swerling 3:} \qquad f_V(v) = \frac{4v}{\bar{v}^2} \exp\left(-\frac{2v}{\bar{v}}\right), \quad v \geq 0 \tag{15.218}$$

where $\bar{v}$ denotes the average value of v.

The Swerling one and Swerling three models are special cases of the chi-square distribution of degree $2m$, for $m = 1$ and $m = 2$, respectively.[30–32] The chi-square PDF is

$$f_V(v) = \frac{m}{(m-1)!\bar{v}} \left(\frac{mv}{\bar{v}}\right)^{m-1} \exp\left(-\frac{mv}{\bar{v}}\right), \quad v > 0 \tag{15.219}$$

The variance of the random target amplitude is reduced as m increases. In the limit, as m approaches infinity, a nonfluctuating target is obtained.

15.6.3.3.2.1. DGM Example

The optimum receiver for target detection in DGM SIRV interference is derived. The random phase of the target is assumed to have a $U(0, 2\pi)$ density and the target energy, $v = a^2$, is assumed to have a chi-square density. The optimum

receiver for this case is obtained by substituting Equation 15.214 and Equation 15.219 into Equation 15.216, which yields

$$T(\tilde{r}) = \int_0^\infty \frac{\sum_{k=1}^{K} w_k s_k^{-2N} \exp\left(-\frac{\tilde{s}^H \tilde{\Sigma}^{-1} \tilde{s} v}{s_k^2}\right) \exp\left(-\frac{\tilde{r}^H \tilde{\Sigma}^{-1} \tilde{r}}{s_k^2}\right) I_0\left(\frac{2|\tilde{s}^H \tilde{\Sigma}^{-1} \tilde{r}|\sqrt{v}}{s_k^2}\right)}{\sum_{k=1}^{K} w_k s_k^{-2N} \exp\left(-\frac{\tilde{r}^H \tilde{\Sigma}^{-1} \tilde{r}}{s_k^2}\right)} \times \frac{m}{(m-1)!\bar{v}} \left(\frac{mv}{\bar{v}}\right)^{m-1} \exp\left(-\frac{mv}{\bar{v}}\right) dv \tag{15.221}$$

This can be rearranged to the form,

$$T(\tilde{r}) = \frac{\sum_{k=1}^{K} w_k s_k^{-2N} \Delta_k \exp\left(-\frac{\tilde{r}^H \tilde{\Sigma}^{-1} \tilde{r}}{s_k^2}\right)}{\sum_{k=1}^{K} w_k s_k^{-2N} \exp\left(-\frac{\tilde{r}^H \tilde{\Sigma}^{-1} \tilde{r}}{s_k^2}\right)} \tag{15.221}$$

where Δ_k is given by

$$\Delta_k = \frac{(m/\bar{v})^m}{(m-1)!} \int_0^\infty \exp\left(-\frac{\tilde{s}^H \tilde{\Sigma}^{-1} \tilde{s} v}{s_k^2}\right) I_0\left(\frac{2|\tilde{s}^H \tilde{\Sigma}^{-1} \tilde{r}|\sqrt{v}}{s_k^2}\right) v^{m-1} \exp\left(-\frac{mv}{\bar{v}}\right) dv \tag{15.222}$$

which simplifies to

$$\Delta_k = \frac{(m/\bar{v})^m}{(m-1)!} \int_0^\infty v^{m-1} \exp\left[-\left(\frac{\tilde{s}^H \tilde{\Sigma}^{-1} \tilde{s}}{s_k^2} + \frac{m}{\bar{v}}\right) v\right] I_0\left(\frac{2|\tilde{s}^H \tilde{\Sigma}^{-1} \tilde{r}|\sqrt{v}}{s_k^2}\right) dv \tag{15.223}$$

This integral is related to the standard form (in Ref. 25, p. 741),

$$\int_0^\infty x^{\mu-(1/2)} \exp(-\alpha x) I_{2\nu}(2\beta\sqrt{x}) dx = \frac{\Gamma\left(\mu + \nu + \frac{1}{2}\right)}{\Gamma(2\nu+1)} \beta^{-1} \exp\left(\frac{\beta^2}{2\alpha}\right) \alpha^{-\mu} M_{-\mu,\nu}\left(\frac{\beta^2}{\alpha}\right) \tag{15.224}$$

where the function, $M_{-\mu,\nu}(z)$, is called the Whittaker function and is related to the confluent hypergeometric function, ${}_1F_1(a, b; z)$ by

$$M_{-\mu,\nu}(z) = z^{\nu+(1/2)} \exp\left(-\frac{z}{2}\right) {}_1F_1\left(\frac{1}{2} + \nu + \mu, 2\nu + 1; z\right) \tag{15.225}$$

Substituting

$$\mu = m - \frac{1}{2} \tag{15.226}$$

and

$$\nu = 0 \tag{15.227}$$

into Equation 15.224 yields

$$\int_0^\infty x^{m-1} \exp(-\alpha x) I_0(2\beta\sqrt{x}) \mathrm{d}x = \frac{\Gamma(m)}{\Gamma(1)} \beta^{-1} \exp\left(\frac{\beta^2}{2\alpha}\right) \alpha^{-m+(1/2)} \times M_{-m+(1/2),0}\left(\frac{\beta^2}{\alpha}\right) \tag{15.228}$$

which, from Equation 15.225, can be expressed as

$$\int_0^\infty x^{m-1} \exp(-\alpha x) I_0(2\beta\sqrt{x}) \mathrm{d}x = \frac{\Gamma(m)}{\Gamma(1)} \beta^{-1} \exp\left(\frac{\beta^2}{2\alpha}\right) \alpha^{-m+(1/2)} \left(\frac{\beta^2}{\alpha}\right)^{1/2} \times \exp\left(-\frac{\beta^2}{2\alpha}\right) {}_1F_1\left(m, 1; \frac{\beta^2}{\alpha}\right) \tag{15.229}$$

This simplifies to

$$\int_0^\infty x^{m-1} \exp(-\alpha x) I_0(2\beta\sqrt{x}) \mathrm{d}x = \Gamma(m) \alpha^{-m} {}_1F_1\left(m, 1; \frac{\beta^2}{\alpha}\right) \tag{15.230}$$

The evaluation of Δ_k in Equation 15.223 is obtained by substituting

$$\alpha = \frac{\tilde{s}^H \tilde{\mathbf{\Sigma}}^{-1} \tilde{s}}{s_k^2} + \frac{m}{\bar{\nu}} \tag{15.231}$$

and

$$\beta = \frac{|\tilde{s}^H \tilde{\mathbf{\Sigma}}^{-1} \tilde{r}|}{s_k^2} \tag{15.232}$$

into the above result and using the notational equivalence, $\Gamma(m) = (m-1)!$. This yields

$$\Delta_k = \left(\frac{m}{\bar{\nu}}\right)^m \left(\frac{\tilde{s}^H \tilde{\mathbf{\Sigma}}^{-1} \tilde{s}}{s_k^2} + \frac{m}{\bar{\nu}}\right)^{-m} {}_1F_1\left(m, 1; \frac{|\tilde{s}^H \tilde{\mathbf{\Sigma}}^{-1} \tilde{r}|^2}{s_k^4 \left(\frac{\tilde{s}^H \tilde{\mathbf{\Sigma}}^{-1} \tilde{s}}{s_k^2} + \frac{m}{\bar{\nu}}\right)}\right) \tag{15.233}$$

The final result for the optimum receiver is obtained by substituting Equation 15.233 into Equation 15.221, which results in

$$T(\tilde{r}) = \frac{\displaystyle\sum_{k=1}^{K} w_k s_k^{-2N} \left(\frac{\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{s}}{s_k^2} + \frac{m}{\bar{\nu}} \right)^{-m} \exp\left(- \frac{\tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}}{s_k^2} \right) {}_1F_1 \left(m, 1; \dfrac{|\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}|^2}{s_k^4 \left(\dfrac{\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{s}}{s_k^2} + \dfrac{m}{\bar{\nu}} \right)} \right)}{\left(\dfrac{\bar{\nu}}{m} \right)^m \displaystyle\sum_{k=1}^{K} w_k s_k^{-2N} \exp\left(- \frac{\tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}}{s_k^2} \right)} \tag{15.234}$$

where $2m$ is the number of degrees of freedom in the chi-square target amplitude model.

Swerling 1 Case. The Swerling one target amplitude model is obtained for $m = 1$. In this case, the optimum receiver for detection of targets with $U(0, 2\pi)$ random phase becomes

$$T(\tilde{r}) = \frac{\displaystyle\sum_{k=1}^{K} w_k s_k^{-2N} \left(\frac{\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{s}}{s_k^2} + \bar{\nu}^{-1} \right)^{-1} \exp\left(- \frac{\tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}}{s_k^2} \right) {}_1F_1 \left(1, 1; \dfrac{|\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}|^2}{s_k^4 \left(\dfrac{\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{s}}{s_k^2} + \bar{\nu}^{-1} \right)} \right)}{\bar{\nu} \displaystyle\sum_{k=1}^{K} w_k s_k^{-2N} \exp\left(- \frac{\tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}}{s_k^2} \right)} \tag{15.235}$$

However, the confluent hypergeometric function has the property,

$${}_1F_1(1, 1; z) = \exp(z) \tag{15.236}$$

and the optimum receiver in Equation 15.235 simplifies to

$$T(\tilde{r}) = \frac{\displaystyle\sum_{k=1}^{K} w_k s_k^{-2N} \left(\frac{\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{s}}{s_k^2} + \bar{\nu}^{-1} \right)^{-1} \exp\left(- \frac{\tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}}{s_k^2} \right) \exp \left(\dfrac{|\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}|^2}{s_k^4 \left(\dfrac{\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{s}}{s_k^2} + \bar{\nu}^{-1} \right)} \right)}{\bar{\nu} \displaystyle\sum_{k=1}^{K} w_k s_k^{-2N} \exp\left(- \frac{\tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}}{s_k^2} \right)} \tag{15.237}$$

Swerling 3 Case. The Swerling three model occurs for $m = 2$, and the corresponding optimum receiver obtained from Equation 15.234 is

$$T(\tilde{r}) = \frac{\displaystyle\sum_{k=1}^{K} w_k s_k^{-2N} \left(\frac{\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{s}}{s_k^2} + \frac{2}{\bar{\nu}} \right)^{-2} \exp\left(- \frac{\tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}}{s_k^2} \right) {}_1F_1 \left(2, 1; \dfrac{|\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}|^2}{s_k^4 \left(\dfrac{\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{s}}{s_k^2} + \dfrac{2}{\bar{\nu}} \right)} \right)}{\left(\dfrac{\bar{\nu}}{2} \right)^2 \displaystyle\sum_{k=1}^{K} w_k s_k^{-2N} \exp\left(- \frac{\tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}}{s_k^2} \right)} \tag{15.238}$$

The confluent hypergeometric function also has the reduction property

$${}_1F_1(a+1, \gamma; z) = \left(\frac{z + 2a - \gamma}{a} \right) {}_1F_1(a, \gamma; z) + \left(\frac{\gamma - a}{a} \right) {}_1F_1(a-1, \gamma; z) \tag{15.239}$$

which gives

$${}_1F_1(2, 1; z) = (1+z) {}_1F_1(1, 1; z) + 0 = (1+z) \exp(z) \tag{15.240}$$

The optimum Swerling three, $U(0, 2\pi)$ receiver becomes

$$T(\tilde{r}) = \frac{\displaystyle\sum_{k=1}^{K} w_k s_k^{-2N} \left[1 + \dfrac{|\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}|^2}{s_k^4 \left(\dfrac{\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{s}}{s_k^2} + \dfrac{2}{\bar{\nu}} \right)} \Big/ \left(\dfrac{\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{s}}{s_k^2} + \dfrac{2}{\bar{\nu}} \right)^2 \right] \exp\left(- \dfrac{\tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}}{s_k^2} \right) \exp\left(\dfrac{|\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}|^2}{s_k^4 \left(\dfrac{\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{s}}{s_k^2} + \dfrac{2}{\bar{\nu}} \right)} \right)}{\left(\dfrac{\bar{\nu}}{2} \right)^2 \displaystyle\sum_{k=1}^{K} w_k s_k^{-2N} \exp\left(- \dfrac{\tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}}{s_k^2} \right)} \tag{15.241}$$

As expected, these DGM SIRV receivers have the canonical structure of Figure 15.47. They can also be represented by the channelized canonical structure shown in Figure 15.48 for appropriate choices of $f(\tilde{r}; s_k, w_k)$ and $g(\tilde{r}; s_k, w_k)$.

The latter form is significant because it provides a suboptimal receiver structure which has the potential to closely approximate the performance of optimum receivers for the detection of random amplitude and phase targets in many types of SIRV clutter. This follows because any finite-valued SIRV PDF can be suitably approximated by some set of probability weights, w_k, and scale mixture values, s_k, for $k = 1, \ldots, K$.

15.6.3.4. Generalized Likelihood Ratio Test

When the signal amplitude the phase parameters are modeled as constants which randomly change from one observation interval (CPI) to the next in the decision problem of Equation 15.152, it is often impossible to obtain a closed-form solution for the integrals required by the optimum NP receiver in non-Gaussian

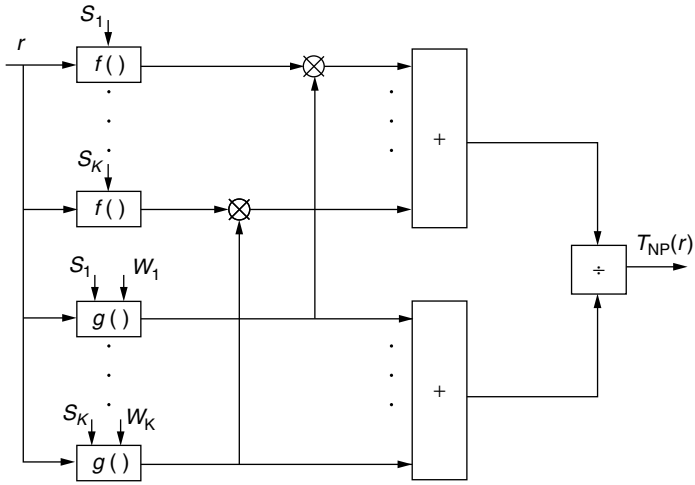

FIGURE 15.48 DGM SIRV optimal receiver.

interference. It is then useful to investigate suboptimal techniques which have potential for nearly optimum detection performance.

The GLRT, whereby ML estimates of the unknown parameters under each hypothesis are substituted into the LRT of Equation 15.154, is one such approach which is commonly used in radar applications. The performance of the GLRT in detection of signals with unknown amplitude and phase in K-distributed SIRV interference has been investigated.[6,7]

Consider the radar detection problem in the form,

$$\begin{aligned} H_0 &: \tilde{r} = \tilde{d} \\ H_1 &: \tilde{r} = \tilde{\alpha}\tilde{s} + \tilde{d}, \end{aligned} \tag{15.242}$$

where $\tilde{\alpha} = ae^{j\phi}$ is an unknown complex constant, $\tilde{s}$ is the low-pass complex envelope of the desired signal, $\tilde{r}$ is the low-pass complex envelope of the received data, and $\tilde{d}$ is the low-pass complex envelope of the SIRV disturbance. The likelihood ratio in SIRV interference is given by

$$T(\tilde{r}|\tilde{\alpha}) = \frac{h_{2N}[2(\tilde{r} - \tilde{\alpha}\tilde{s})^H \tilde{\boldsymbol{\Sigma}}^{-1}(\tilde{r} - \tilde{\alpha}\tilde{s})]}{h_{2N}[2\tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}]} \tag{15.243}$$

Hence, the GLRT for an SIRV is defined to be

$$T_{\text{GLRT}}(\tilde{r}) = T(\tilde{r}|\tilde{\alpha}) \tag{15.244}$$

which becomes

$$T_{\text{GLRT}}(\tilde{r}) = \frac{h_{2N}[2(\tilde{r} - \tilde{\alpha}\tilde{s})^H \tilde{\boldsymbol{\Sigma}}^{-1}(\tilde{r} - \tilde{\alpha}\tilde{s})]}{h_{2N}[2\tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}]} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{15.245}$$

where $\hat{\tilde{\alpha}}$ denotes the MLE of the unknown complex constant amplitude, $\tilde{\alpha}$, and is the value of $\tilde{\alpha}$ which maximizes the conditional probability density function,

$$f_{\tilde{r}|\tilde{\alpha},H_1}(\tilde{r}|\tilde{\alpha},H_1) = \pi^{-N}|\tilde{\mathbf{\Sigma}}|^{-1}h_{2N}[2(\tilde{r}-\tilde{\alpha}\tilde{s})^H\tilde{\mathbf{\Sigma}}^{-1}(\tilde{r}-\tilde{\alpha}\tilde{s})] \tag{15.246}$$

Since $h_{2N}(\cdot)$ is a monotonic decreasing function for every SIRV, this conditional PDF is maximized when the quadratic argument, $(\tilde{r}-\tilde{\alpha}\tilde{s})^H\tilde{\mathbf{\Sigma}}^{-1}(\tilde{r}-\tilde{\alpha}\tilde{s})$, is minimized. The minimization of this quadratic form is independent of the function, $h_{2N}(\cdot)$, which characterizes the type of SIRV interference. Consequently, the ML estimate of $\tilde{\alpha}$ is the same for all types of SIRV interference and must be identical to the estimate obtained in Gaussian interference. Therefore, $\hat{\tilde{\alpha}}$ is always

$$\hat{\tilde{\alpha}}(\tilde{r}) = \frac{\tilde{s}^H\tilde{\mathbf{\Sigma}}^{-1}\tilde{r}}{\tilde{s}^H\tilde{\mathbf{\Sigma}}^{-1}\tilde{s}} \tag{15.247}$$

The individual ML estimates of a and ϕ follow as

$$\hat{\alpha}(\tilde{r}) = \frac{|\tilde{s}^H\tilde{\mathbf{\Sigma}}^{-1}\tilde{r}|}{\tilde{s}^H\tilde{\mathbf{\Sigma}}^{-1}\tilde{s}} \tag{15.248}$$

and

$$\hat{\phi}(\tilde{r}) = \arg(\tilde{s}^H\tilde{\mathbf{\Sigma}}^{-1}\tilde{r}) \tag{15.249}$$

The general expression for the GLRT receiver in SIRV interference is obtained by substituting the ML estimate (MLE) given by Equation 15.247 into Equation 15.245. The resulting expression can be simplified to[6]

$$T_{\text{GLRT}}(\tilde{r}) = \frac{h_{2N}[2\tilde{r}^H\tilde{\mathbf{\Sigma}}^{-1}\tilde{r}(1-|\rho|^2)]}{h_{2N}[2\tilde{r}^H\tilde{\mathbf{\Sigma}}^{-1}\tilde{r}]} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{15.250}$$

where $|\rho|^2$ is determined by

$$|\rho|^2 = \frac{|\tilde{s}^H\tilde{\mathbf{\Sigma}}^{-1}\tilde{r}|^2}{(\tilde{r}^H\tilde{\mathbf{\Sigma}}^{-1}\tilde{r})(\tilde{s}^H\tilde{\mathbf{\Sigma}}^{-1}\tilde{s})} \tag{15.251}$$

It follows from the Schwarz inequality that ρ satisfies $0 \le |\rho| \le 1$.

In the case of a Gaussian SIRV with $h_{2N}(q) = \exp(-q/2)$, the generalized likelihood ratio can be reduced to the sufficient statistic, $|\tilde{s}^H\tilde{\mathbf{\Sigma}}^{-1}\tilde{r}|^2$. Hence, the GLRT for any SIRV is a nonlinear function of both the Gaussian receiver test statistic and the quadratic form, $\tilde{r}^H\tilde{\mathbf{\Sigma}}^{-1}\tilde{r}$, and can be implemented in the canonical receiver form of Figure 15.47.

A MLE is consistent, which means it converges in probability to the true values as the sample size approaches infinity. Consequently, the detection performance of the GLRT asymptotically approaches the perfect measurement bound. If acceptable (near-optimal) detection performance cannot be obtained by

the GLRT with the available number of samples, then other test procedures should be investigated. Simulation results for a few examples of GLRT detection in K-distributed SIRV clutter have been reported.[7] These results indicate the GLRT performance is near the perfect measurement bound for even relatively small sample sizes.

A performance comparison between the optimum NP receiver for the random parameter case and a GLRT used in the same random parameter case is of interest for several reasons. First, the optimum NP receiver for random signal parameters requires a knowledge of the joint probability density function of these parameters. If this density function is not known to sufficient accuracy, then it is possible that a distribution-free test, such as the GLRT, has better detection performance. Second, the GLRT may have a simpler functional form which leads to a less complicated and more accurate receiver implementation. In fact, if there is not a closed-form solution for the optimum receiver when the signal parameters are random, then a suboptimum implementation is necessary. That another suboptimum processor outperforms a GLRT without significantly more complexity is not guaranteed.

15.6.3.5. Maximum Likelihood Matched Filter

A suboptimal receiver which is independent of the type of SIRV clutter which is received is obtained by considering the nature of the SIRV model as described in the representation theorem. The clutter samples from the ith range-azimuth cell are assumed to be jointly Gaussian with zero-mean and covariance matrix, $s_i^2\tilde{\mathbf{\Sigma}}$, where $\tilde{\mathbf{\Sigma}}$ is the covariance matrix of the underlying Gaussian vector in the SIRV representation theorem. The commonality of the covariance structure for the received data from each range-azimuth cell is exploited in this development.

The likelihood-ratio test for Gaussian clutter in the ith range-azimuth cell is given by

$$T(\tilde{r}) = \frac{\exp\left[-(\tilde{r} - a\mathrm{e}^{\mathrm{j}\phi}\tilde{s})^H(s_i^2\tilde{\mathbf{\Sigma}})^{-1}(\tilde{r} - a\mathrm{e}^{\mathrm{j}\phi}\tilde{s})\right]}{\exp(-\tilde{r}^H(s_i^2\tilde{\mathbf{\Sigma}})^{-1}\tilde{r})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{15.252}$$

which is easily reduced to

$$T(\tilde{r}) = \exp\left(\frac{-a^2\tilde{s}^H\tilde{\mathbf{\Sigma}}^{-1}\tilde{s}}{s_i^2}\right)\exp\left(\frac{2a\,\mathrm{Re}\{\mathrm{e}^{\mathrm{j}\phi}\tilde{s}^H\tilde{\mathbf{\Sigma}}^{-1}\tilde{r}\}}{s_i^2}\right) \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{15.253}$$

The unknown signal phase is replaced by the MLE of Equation 15.249 to yield

$$T(\tilde{r}) = \exp\left(\frac{-a^2\tilde{s}^H\tilde{\mathbf{\Sigma}}^{-1}\tilde{s}}{s_i^2}\right)\exp\left(\frac{2a|\tilde{s}^H\tilde{\mathbf{\Sigma}}^{-1}\tilde{r}|}{s_i^2}\right) \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{15.254}$$

This result is independent of the phase density function. However, for a $U(0, 2\pi)$ random phase, it compares to the optimum NP receiver,

$$\exp\left(\frac{-a^2\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{s}}{s_i^2}\right)I_0\left(\frac{2a|\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}|}{s_i^2}\right)\underset{H_0}{\overset{H_1}{\gtrless}}\eta' \tag{15.255}$$

which can be obtained from Equation 15.214 for a single Gaussian mixture component. The Gaussian receivers in Equation 15.254 and Equation 15.255 are usually reduced to the envelope-detected matched filter,

$$|\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}|\underset{H_0}{\overset{H_1}{\gtrless}}\eta'' \tag{15.256}$$

by moving constants, including a and s_i^2, into the threshold and applying the inverses of monotonic functions. The threshold is then determined from the specified probability of false alarm, which can only be done if the clutter power, s_i^2, is known.

Since the unknown clutter power in each cell is the basis for the SIRV model, simplification to Equation 15.256 cannot be performed. Furthermore, a UMP test with respect to the signal amplitude, a, cannot be obtained without knowledge of s_i^2. This is consistent with the previous discussion concerning optimal SIRV receivers.

The ML matched filter (MLMF) is obtained by replacing a and s_i^2 in Equation 15.254 with the MLEs, $\hat{a}$ and $\hat{s}_i^2$. Taking the natural logarithm of Equation 15.254 yields

$$\frac{-a^2\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{s}}{s_i^2}+\frac{2a|\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}|}{s_i^2}\underset{H_0}{\overset{H_1}{\gtrless}}\ln\,\eta \tag{15.257}$$

Thus, the MLMF statistic is given by

$$T_{\mathrm{MLMF}}(\tilde{r})=\frac{-\hat{a}^2\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{s}}{\hat{s}_i^2}+\frac{2\hat{a}|\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}|}{\hat{s}_i^2} \tag{15.258}$$

which, for the signal amplitude estimate given by Equation 15.248, simplifies to

$$T_{\mathrm{MLMF}}(\tilde{r})=\frac{|\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{r}|^2}{(\tilde{s}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{s})\hat{s}_i^2} \tag{15.259}$$

Knowledge of the covariance matrix, $\tilde{\boldsymbol{\Sigma}}$, determines a whitening transformation which can be applied to the received data. The MLMF statistic is expressed in terms of the whitened vectors, $\tilde{r}_{\mathrm{w}}=\tilde{\boldsymbol{\Sigma}}^{-1/2}\tilde{r}$ and $\tilde{s}_{\mathrm{w}}=\tilde{\boldsymbol{\Sigma}}^{-1/2}\tilde{s}$, as

$$T_{\mathrm{MLMF}}(\tilde{r}_{\mathrm{w}})=\frac{|\tilde{s}_{\mathrm{w}}^H\tilde{r}_{\mathrm{w}}|^2}{(\tilde{s}_{\mathrm{w}}^H\tilde{s}_{\mathrm{w}})\hat{s}_i^2} \tag{15.260}$$

The whitening transformation does not affect s_i^2, which is seen by considering the SIRV representation theorem. In fact, s_i^2 is the variance of both the real and imaginary parts of every complex element in the whitened vector, $\tilde{r}_{\mathrm{w}}$. Furthermore, the spherical symmetry of the problem permits an arbitrary rotation of the coordinates, such that all of the unknown signal energy is contained in a single complex element and the remaining $N-1$ orthogonal complex elements contain only clutter components.

The $N-1$ signal-free components of the whitened vector are used to obtain the MLE of s_i^2, given by

$$\hat{s}_i^2 = \frac{1}{2(N-1)} \tilde{r}_{\mathrm{w}\perp}^H \tilde{r}_{\mathrm{w}\perp} \tag{15.261}$$

where the $\perp$ subscript indicates that the $N-1$ dimensional complex vector of these signal-free components is orthogonal to the complex signal component. This estimate is substituted into Equation 15.260 to yield

$$T_{\mathrm{MLMF}}(\tilde{r}_{\mathrm{w}}) = \frac{|\tilde{s}_{\mathrm{w}}^H \tilde{r}_{\mathrm{w}}|^2}{(\tilde{s}_{\mathrm{w}}^H \tilde{s}_{\mathrm{w}})(\tilde{r}_{\mathrm{w}\perp}^H \tilde{r}_{\mathrm{w}\perp})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_* \tag{15.262}$$

where the constant factor, $2(N-1)$, in incorporated into the threshold, η_*. The relation,

$$|\tilde{s}_{\mathrm{w}}^H \tilde{r}_{\mathrm{w}}|^2 = (\tilde{s}_{\mathrm{w}}^H \tilde{s}_{\mathrm{w}})(\tilde{r}_{\mathrm{ws}}^H \tilde{r}_{\mathrm{ws}}) \tag{15.263}$$

where $\tilde{r}_{\mathrm{ws}}$ denotes the projection of the whitened received vector, $\tilde{r}$, onto the whitened signal plane, is used to simplify Equation 15.262 to

$$T_{\mathrm{MLMF}}(\tilde{r}_{\mathrm{w}}) = \frac{\tilde{r}_{\mathrm{ws}}^H \tilde{r}_{\mathrm{ws}}}{\tilde{r}_{\mathrm{w}}^H \tilde{r}_{\mathrm{w}\perp}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_* \tag{15.264}$$

Substituting

$$\tilde{r}_{\mathrm{w}}^H \tilde{r}_{\mathrm{w}\perp} = \tilde{r}_{\mathrm{w}}^H \tilde{r}_{\mathrm{w}} - \tilde{r}_{\mathrm{ws}}^H \tilde{r}_{\mathrm{ws}} \tag{15.265}$$

into this expression gives

$$T_{\mathrm{MLMF}}(\tilde{r}_{\mathrm{w}}) = \frac{\tilde{r}_{\mathrm{ws}}^H \tilde{r}_{\mathrm{ws}}}{\tilde{r}_{\mathrm{w}}^H \tilde{r}_{\mathrm{w}} - \tilde{r}_{\mathrm{ws}}^H \tilde{r}_{\mathrm{ws}}} \tag{15.266}$$

which can also be written as

$$T_{\mathrm{MLMF}}(\tilde{r}_{\mathrm{w}}) = \frac{1}{\dfrac{\tilde{r}_{\mathrm{w}}^H \tilde{r}_{\mathrm{w}}}{\tilde{r}_{\mathrm{ws}}^H \tilde{r}_{\mathrm{ws}}} - 1} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_* \tag{15.267}$$

Further simplification of this statistic yields

$$T_{\mathrm{MLMF}}(\tilde{r}_{\mathrm{w}}) = \frac{\tilde{r}_{\mathrm{ws}}^H \tilde{r}_{\mathrm{ws}}}{\tilde{r}_{\mathrm{w}}^H \tilde{r}_{\mathrm{w}}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{**} \tag{15.268}$$

where the threshold, η_{**}, includes the appropriate modifications to η_*.

The MLMF statistic is written in terms of the original unwhitened received data by using Equation 15.263 to obtain

$$T_{\mathrm{MLMF}}(\tilde{r}_{\mathrm{w}}) = \frac{|\tilde{s}_{\mathrm{w}}^H \tilde{r}_{\mathrm{w}}|^2}{(\tilde{s}_{\mathrm{w}}^H \tilde{s}_{\mathrm{w}})(\tilde{r}_{\mathrm{w}}^H \tilde{r}_{\mathrm{w}})} \underset{H_0}{\overset{H_1}{\gtrless}} \eta_{**} \tag{15.269}$$

which leads to

$$T_{\mathrm{MLMF}}(\tilde{r}) = \frac{|\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r}|^2}{(\tilde{s}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{s})(\tilde{r}^H \tilde{\boldsymbol{\Sigma}}^{-1} \tilde{r})} \tag{15.270}$$

for the final form of the statistic.

During the course of this research, this MLMF receiver statistic also appeared in the literature.[26–28] However, the derivation presented here is different and provides a simpler explanation in terms of the Gaussian matched filter development.

The receiver statistic, $T_{\mathrm{MLMF}}(\tilde{r})$, in Equation 15.270 does not require any knowledge of the type of SIRV clutter which is present in the received data in order to be implemented. Furthermore, under hypothesis H_0, the probability density function of this test statistic is also independent of which SIRV is present. (Since the received data is only SIRV clutter under H_0, the SIRV representation theorem is used to replace the received data vector, $\tilde{r}$, in Equation 15.270 by the product of a Gaussian vector and a scalar. The scalar divides out of both the numerator and denominator, leaving the resulting expression independent of the scalar value and, hence, independent of the type of SIRV.) Consequently, the threshold required to maintain a given false alarm rate is constant for all SIRVs and can be determined from a Gaussian clutter assumption. Thus, this suboptimum receiver is the least complicated to implement for any non-Gaussian SIRV clutter environment. The performance of the MLMF statistic is compared to the NP and GLRT performance in Section 15.6.4.5.

15.6.4. Nonlinear Receiver Performance

15.6.4.1. Introduction

Since the optimum non-Gaussian receivers of Equation 15.154 and Equation 15.159 are nonlinear, it is usually impossible to obtain a closed-form expression for the PDF of the optimum test statistic. This makes analytical determination of receiver performance very difficult. Consequently, Monte Carlo simulation methods are considered for the determination of the receiver threshold and the evaluation of detection performance.

A direct approach to simulating the receiver test statistic, $T(\tilde{r})$, in Figure 15.47 requires the generation of independent realizations of the complex input vector, $\tilde{r}$, having the multivariate PDF given by Equation 15.127. Several techniques for generating SIRV samples with the appropriate PDF have been discussed.[12,13] This direct approach has two drawbacks, however. First, $2N$

independent, real data samples are required in order to obtain a single sample of the output statistic. This is very inefficient for large N. Second, undesirable correlation between vector samples may be created by grouping the outputs of a pseudo-random number generator. This problem is not well studied for even moderate N, but in some instances it becomes much worse as N increases. A significant correlation structure in the $2N$ dimensional vector sample space used for the Monte Carlo simulation could give erroneous performance results.[14–16]

A more efficient and reliable technique is based upon the canonical receiver in Figure 15.47. The inputs, $T_g(\tilde{r})$ and $T_q(\tilde{r})$, to the optimal nonlinearity can be shown to depend on only a small fixed number of independent random variables with known PDFs. Consequently, the potential correlation problems of high dimensional random vector generation are avoided for all N by computing samples of $T_g(\tilde{r})$ and $T_q(\tilde{r})$ from these random variables, instead of generating the elements of the received vector, $\tilde{r}$.

15.6.4.2. Indirect Simulation of SIRV Receiver Statistics

All the receivers in Section 15.6.3 for detection of signals in SIRV clutter are expressed in terms of low-pass complex envelope samples. The development of these receivers is often simplified by using the complex envelope, especially when considering signals with an unknown phase. The inputs to the optimal nonlinearity in Figure 15.47 are expressed in terms of the signal parameters and clutter disturbance, $\tilde{d}$, by substituting Equation 15.152 for the received vector under each hypothesis into the appropriate definitions for $T_g(\tilde{r})$ and $T_q(\tilde{r})$.

All the real and imaginary components of the complex clutter samples are assumed to be uncorrelated and to have unit-variance, without loss of generality. Thus, the covariance matrix of $\tilde{d}$ is $\tilde{\boldsymbol{\Sigma}} = 2I$, where I is the $N \times N$ identity matrix. The SCR is adjusted by changing the signal energy.

For a completely known signal in this uncorrelated clutter, the inputs to the optimal nonlinearity are given by Equation 15.156 and Equation 15.158. Since the signal is completely known, its phase is assumed to be $\phi = 0$ without loss of generality. Furthermore, the spherical symmetry of the clutter allows rotation of the signal vector, $\tilde{s}$, such that all of the signal energy is contained in the real part of one component. The inputs to the nonlinearity under hypotheses H_0 and H_1 are then

$$T_q(\tilde{\mathbf{r}}|H_0) = \|\tilde{\mathbf{d}}\|^2/2 \tag{15.271}$$

$$T_g(\tilde{\mathbf{r}}|H_0) = \|\tilde{\mathbf{s}}\|\,\|\tilde{\mathbf{d}}\|\cos(\phi_d)/2 \tag{15.272}$$

$$T_q(\tilde{\mathbf{r}}|H_1) = (\|\tilde{\mathbf{d}}\|^2 + 2a\|\tilde{\mathbf{s}}\|\,\|\tilde{\mathbf{d}}\|\cos(\phi_d) + a^2\|\tilde{\mathbf{s}}\|^2)/2 \tag{15.273}$$

and

$$T_g(\tilde{\mathbf{r}}|H_1) = (a\|\tilde{\mathbf{s}}\|^2 + \|\tilde{\mathbf{s}}\|\,\|\tilde{\mathbf{d}}\|\cos(\phi_d))/2 \tag{15.274}$$

These statistics depend on only two independent random variables, $q_d = \|\tilde{\mathbf{d}}\|^2 = \mathbf{d}^H\tilde{\mathbf{d}}$ and ϕ_d. The real, $2N$ dimensional vector, $\mathbf{d} = [\mathbf{d}_c^T, \mathbf{d}_s^T]^T$, where $\mathbf{d}_c$ and $\mathbf{d}_s$ are the in-phase and quadrature components of $\tilde{\mathbf{d}} = \mathbf{d}_c + j\mathbf{d}_s$, has the same norm as $\tilde{\mathbf{d}}$. Thus, it satisfied $\|\tilde{\mathbf{d}}\|^2 = \|\mathbf{d}\|^2$ and the PDF of q_d is given by Equation 15.171 with N replaced by $2N$. The second variable, ϕ_d, is the angle between the clutter disturbance vector and the signal vector, which is rotated as described above. This is precisely the spherical angle Φ_1 of Equation 15.135 for the spherical coordinate representation described in Section 15.6.2.2.4, but with dimension $2N$. Consequently, the PDF of f_d is

$$f_{\Phi_d}(f_d) = \frac{\Gamma(N)}{\sqrt{\pi}\Gamma\left(N - \frac{1}{2}\right)} \sin^{2(N-1)}(\phi_d), \qquad 0 \le f_d \le \pi \tag{15.275}$$

Only these two sequences of independent random numbers must be generated for computer simulation of the receiver performance.

Similarly, for the $U(0, 2\pi)$ random signal phase case, T_q and T_g under each hypothesis are derived as

$$T_q(\mathbf{r}|H_0) = \|\mathbf{d}\|^2/2 \tag{15.276}$$

$$T_g(\mathbf{r}|H_0) = \|\mathbf{s}\|\,\|\mathbf{d}\|\cos(\omega)/2 \tag{15.277}$$

$$T_q(\mathbf{r}|H_1) = (\|\mathbf{d}\|^2 + 2a\|\mathbf{s}\|\,\|\mathbf{d}\|\cos(\omega)\cos(\psi) + a^2\|\mathbf{s}\|^2)/2 \tag{15.278}$$

and

$$T_g(\mathbf{r}|H_1) = (\|\mathbf{s}\|\sqrt{a^2\|\mathbf{s}\|^2 + 2a\|\mathbf{s}\|\,\|\mathbf{d}\|\cos(\omega)\cos(\psi) + \|\mathbf{d}\|^2\cos^2(\omega)})/2 \tag{15.279}$$

Now, only three independent random variables, $q_d = \|d\|^2$, ω, and ψ, are needed to generate the above statistics. The random variable, q_d, is identical to the previous known signal problem and again has a PDF given by Equation 15.171.

Since the assumed signal has an unknown phase, it can only be rotated such that all the signal energy is contained in a single pair of real and imaginary components describing a plane. The variable, ω, is the angle between the clutter disturbance vector and this plane. This is precisely the spherical angle, Φ_1, from Equation Z-15 and Equation Z-16 of the alternative spherical coordinate representation described in Appendix Z. Its PDF, obtained from Equation Z-29 for $k = 1$, is

$$f_\Omega(\omega) = 2(N-1)\cos(\omega)\sin^{(2N-3)}(\omega), \quad 0 \le \omega \le \frac{\pi}{2} \tag{15.280}$$

Finally, ψ is a $U(0, 2\pi)$ distributed angle which occurs as a result of the random signal phase.

When a random signal amplitude case is considered, random samples of a must also be generated according to the amplitude PDF, $f_A(a)$.

The receiver performance is simulated by the following method. At least $100/P_{\mathrm{FA}}$ samples of either q_{d} and f_{d} for the known signal case or q_{d}, ω, and ψ for the random signal phase case are generated. Samples of the clutter-only output test statistic, $T(\tilde{\mathbf{r}}|H_0)$, are computed and used to determine the threshold which gives the desired P_{FA} value. Then samples of the output test statistic, $T(\tilde{\mathbf{r}}|H_1)$, are computed and the PD is determined from the number of threshold crossings.

15.6.4.3. Student *t* SIRV Results

Detection in uncorrelated Student t SIRV clutter is considered for both a fully known signal and a signal with known amplitude and $U(0, 2\pi)$ random phase. The marginal PDFs of the quadrature components, $d_{i,q}$, are the Student t density,[3]

$$f_{D_{i,q}}(d_{i,q}) = \frac{\Gamma\left(\nu + \frac{1}{2}\right)}{b\sqrt{\pi}\Gamma(\nu)}\left(1 + \frac{d_{i,q}^2}{b^2}\right)^{-\nu-(1/2)}, \quad \nu > 0 \tag{15.281}$$

where ν is a shape parameter, b is a scale parameter, and $\Gamma(\cdot)$ denotes the gamma function. The monotonic decreasing function, $h_{2N}(q)$, for the complex Student t SIRV is[3]

$$h_{2N}(q) = \frac{2^N b^{2\nu}\Gamma(\nu + N)}{\Gamma(\nu)(b^2 + q)^{\nu+N}}, \quad \nu > 0 \tag{15.282}$$

and the optimum known signal receiver is obtained by substituting this function into Equation 15.154. Since detection performance of this receiver depends only on the signal energy, $\phi = 0$ is used for the signal phase without loss of generality. The optimum receiver then reduces to

$$\frac{b^2 + \|\tilde{\mathbf{r}}\|^2}{b^2 + \|\tilde{\mathbf{r}} - a\tilde{\mathbf{s}}\|^2} = \frac{b^2 + \|\tilde{\mathbf{r}}\|^2}{b^2 + \|\tilde{\mathbf{r}}\|^2 + a^2\|\tilde{\mathbf{s}}\|^2 - 2a\,\mathrm{Re}\{\tilde{\mathbf{s}}^H\tilde{\mathbf{r}}\}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{15.283}$$

which is seen to have the canonical form of Figure 15.47a when $\mathrm{Re}\{\tilde{\mathbf{s}}^H\tilde{\mathbf{r}}\} = T_{\mathrm{G}}(\tilde{\mathbf{r}})$ is recognized as the optimal Gaussian statistic for this target model in uncorrelated clutter.

Optimum detection of the known amplitude, $U(0, 2\pi)$ random phase signal is given by Equation 15.159, where only an average over the phase density is required. The optimum receiver in Student t SIRV clutter for this case is derived in Section 15.6.3.3.1 and given by Equation 15.197 and Equation 15.198. For uncorrelated clutter samples, with $\tilde{\boldsymbol{\Sigma}} = 2\mathbf{I}$, the Student t receiver is easily simplified to

$$\left(\frac{(\|\tilde{\mathbf{r}}\|^2 + b^2)\lambda_{\tilde{\mathbf{r}}}}{\|\tilde{\mathbf{r}}\|^2 + b^2 + a^2\|\tilde{\mathbf{s}}\|^2}\right)^{\nu+N} P_{\nu+N-1}(\lambda_{\tilde{\mathbf{r}}}) \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \quad \nu > 0 \tag{15.284}$$

where $\lambda_{\tilde{\mathbf{r}}}$ is given by

$$\lambda_{\tilde{\mathbf{r}}} = \left(\frac{\|\tilde{\mathbf{r}}\|^2 + b^2 + a^2\|\tilde{\mathbf{s}}\|^2}{\sqrt{(\|\tilde{\mathbf{r}}\|^2 + b^2 + a^2\|\tilde{\mathbf{s}}\|^2)^2 - (2a|\tilde{\mathbf{s}}^H\tilde{\mathbf{r}}|)^2}} \right) \geq 1 \tag{15.285}$$

and $P_\mu(\cdot)$ is the Legendre function of the first kind of degree μ, which reduces to a Legendre polynomial for integer μ. Notice that Equation 15.165 is consistent with the canonical structure of Figure 15.47, where $|\tilde{\mathbf{s}}^H\tilde{\mathbf{r}}|$ is identified as the Gaussian receiver output for the $U(0, 2\pi)$ random phase signal.

Detection performance of the optimum receiver in uncorrelated Student t clutter for $P_{\text{FA}} = 0.01$, $\nu = 2$, and $N = 2$, 4, 8, and 16 pulses is shown in Figure 15.49 to Figure 15.51. The SCR is defined to be

$$\text{SCR} = \frac{a^2\|\tilde{\mathbf{s}}\|^2}{E(\|\tilde{\mathbf{d}}\|^2)} = \frac{a^2\|\tilde{\mathbf{s}}\|^2}{2N} \tag{15.286}$$

These figures indicate the optimal non-Gaussian receiver has significantly better detection performance than the Gaussian receiver for at least a 10 dB interval of SCR values.

The Gaussian receiver performance is essentially optimum for very strong signals. This is expected, since the signal appears more like a clutter spike as its strength increases. Therefore, nonlinear processing which reduces the large clutter spikes would also reduce the signal energy. The coherent integration gain

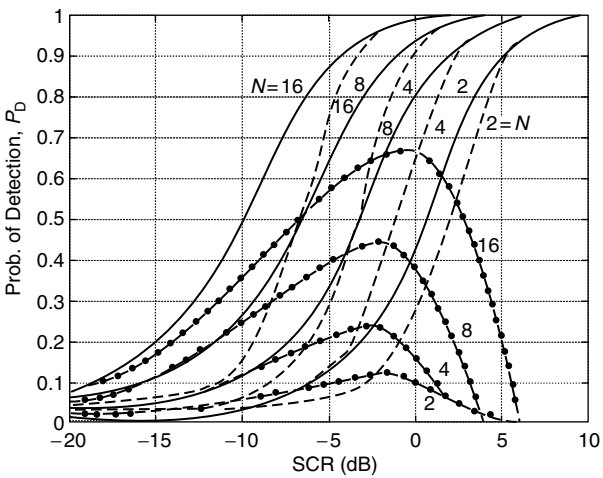

FIGURE 15.49 Detection probability for completely known signal in Student t SIRV clutter with $\nu = 2$, $P_{\text{FA}} = 0.01$, and $N = 2$, 4, 8, 16 pulses. (—) Optimal receiver, (– –) Gaussian receiver, (•—•) LOD.

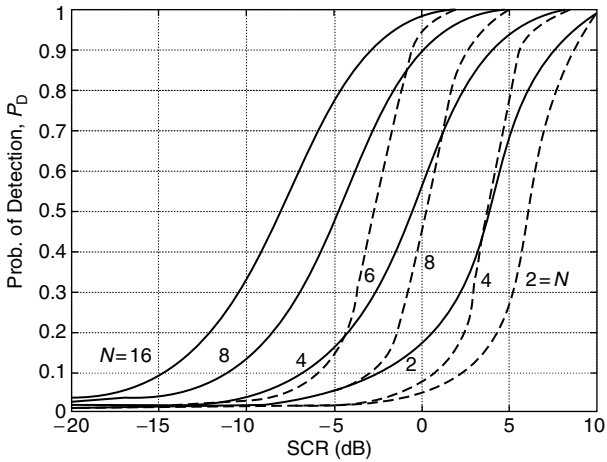

FIGURE 15.50 Detection probability for signal of known amplitude and $U(0, 2\pi)$ random phase in Student t SIRV clutter with $\nu = 2$, $P_{FA} = 0.01$, and $N = 2, 4, 8, 16$ pulses. (—) Optimal receiver, (– –) Gaussian receiver.

of the matched filter dominates the processing and provides enough margin between the signal and the clutter spikes to set a suitable threshold. The detection performance must approach the false alarm probability as the signal strength tends to zero. Detection of these very weak signals is accomplished only by using a very large number of integrated pulses. When the SCR lies between these two

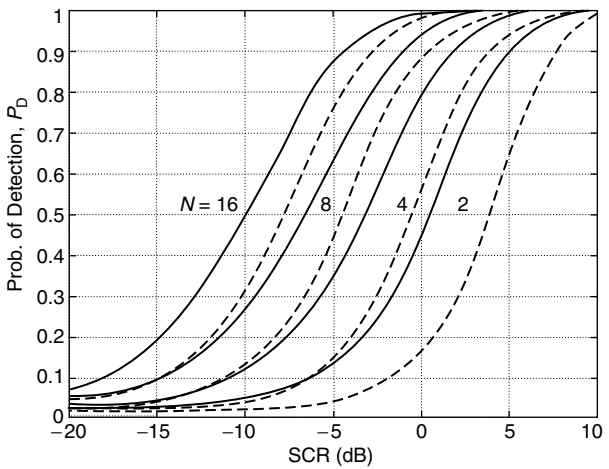

FIGURE 15.51 Detection probability with optimal receiver in Student t SIRV clutter with $\nu = 2$, $P_{FA} = 0.01$, and $N = 2, 4, 8, 16$ pulses. (–) Signal of known phase, (—) signal of $U(0, 2\pi)$ random phase.

extremes, the nonlinear processing of the optimal non-Gaussian receiver provides a significant detection capability for weak signals.

Figure 15.49 also compares performance of the optimal known signal receiver to the LOD,[1] given by

$$T_{\mathrm{LOD}}(\tilde{\mathbf{r}}) = \frac{T_{\mathrm{g}}(\tilde{\mathbf{r}})}{\nu - 1 + \dfrac{1}{2} T_{\mathrm{q}}(\tilde{\mathbf{r}})}, \quad \nu > 1 \tag{15.287}$$

for a Student t SIRV clutter case. The LOD is a suboptimum receiver which is often used for detection of unknown amplitude signals in small SCR situations. It is derived from a Taylor series expansion of the LRT about $a = 0$ and is also equivalent to maximizing the slope of the P_{d} vs. SCR curve at the origin.

One advantage of the LOD is that it does not require knowledge of the signal amplitude. The significant performance improvement of the LOD compared to the Gaussian receiver is verified by Figure 15.49. However, this figure also indicates that the LOD suffers a sizable detection loss for some SCR ranges when compared to the optimal known signal receiver performance. This detection loss is the motivation behind receiver designs which estimate the signal amplitude.[6,7]

Figure 15.51 compares optimal receiver performance between the known phase case and the $U(0, 2\pi)$ random phase case. These results indicate that an additional 2 to 4 dB increase in SCR is required for the unknown phase signal to obtain the same performance as a known phase signal.

Since the random phase of the received target return is assumed constant over the N pulse CPI, a commonly suggested receiver is the GLRT. The GLRT substitutes a ML estimate for the unknown phase into the likelihood-ratio test of Equation 15.153. The GLRT for the Student t SIRV has the form,

$$T_{\mathrm{GLRT}}(\tilde{\mathbf{r}}) = \frac{b^2 + \|\tilde{\mathbf{r}}\|^2}{b^2 + \|\tilde{\mathbf{r}}\|^2 - 2a|\tilde{\mathbf{s}}^H \tilde{\mathbf{r}}| + a^2 \|\tilde{\mathbf{s}}\|^2} \tag{15.288}$$

A comparison of the detection performance of the GLRT and the optimum NP receivers for $U(0, 2\pi)$ random phase is given in Figure 15.52. These results indicate that the two receivers have identical performance, even though the GLRT is not necessarily expected to be optimal for non-Gaussian SIRVs. This same result has been reported for a single example in K-distributed SIRV clutter.[9] It has also been observed for several examples involving the DGM SIRV, which is discussed in the next section.

The reason for this behavior is not thoroughly understood. However, investigation into whether or not it holds for all SIRVs is proceeding. Since the GLRT and NP receivers are known to be equivalent for the Gaussian clutter problem, the result is presumed to be related to the underlying conditionally Gaussian nature of SIRVs.

Performance has been shown for $P_{\mathrm{FA}} = 0.01$, which is a high value for the false alarm probability if the entire surveillance volume is considered. Certainly,

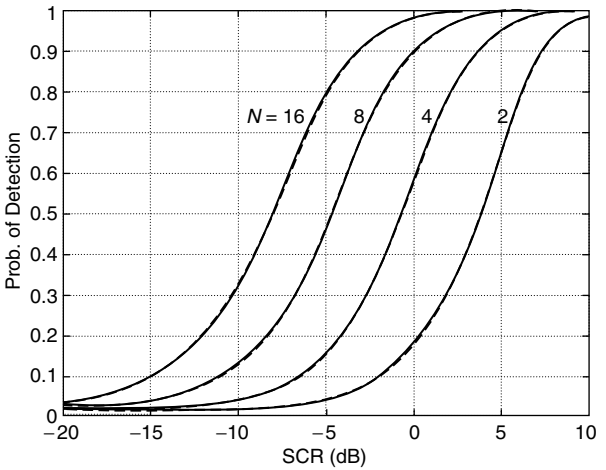

FIGURE 15.52 Detection probability for signal of $U(0, 2\pi)$ random phase in Student t SIRV clutter with $\nu = 2$, $P_{\text{FA}} = 0.01$, and $N = 2$, 4, 8, 16 pulses. (– –) NP receiver, (—) GLRT receiver.

performance for lower P_{FA} values is still of interest and being investigated. However, it is expected that optimum non-Gaussian processing is only necessary in a portion of this volume where there is very strong clutter. It seems reasonable to allow some increase in the false alarm rate within these regions to improve detection capability, with the understanding that the false alarm rate will be appropriately lowered in other regions of the volume.

15.6.4.4. DGM Results

DGM SIRVs are obtained when the characteristic PDF, $f_V(v)$, of the SIRV is a discrete probability density function described by

$$f_V(v) = \sum_{i=1}^{K} w_i \delta(v - v_i), \quad \text{for } w_i > 0 \text{ and } \sum_{i=1}^{K} w_i = 1 \tag{15.289}$$

and illustrated in Figure 15.53.

The monotonic decreasing function for the DGM SIRV is

$$h_{2N}(q) = \sum_{k=1}^{K} w_k v_K^{-2N} \exp\left(-\frac{q}{2v_k^2}\right) \tag{15.290}$$

which is a continuous function.

The DGM SIRV is especially useful because closed-form solutions exist for the optimum receiver for $U(0, 2\pi)$ random signal phase and chi-square random signal amplitude models, which includes the Swerling 1 and Swerling 3 models. The optimum receivers for these models are derived in Section 15.6.3.3.2.

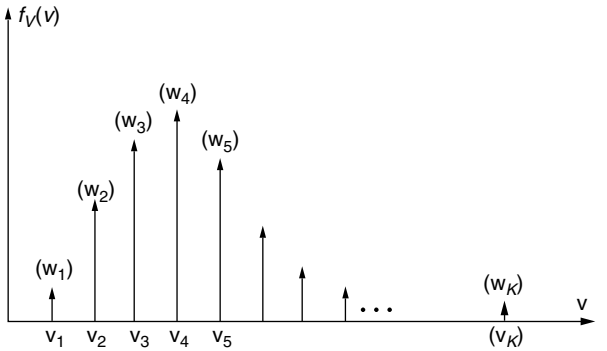

FIGURE 15.53 Characteristic PDF for a DGM SIRV.

Furthermore, any finite valued SIRV PDF can be approximated as closely as desired by some appropriate choice of K, w_i, and v_i for the characteristic PDF of Equation 15.289. Thus, the DGM SIRVs can be used to evaluate optimal performance for many other types of SIRVs. This is now illustrated in the following examples which use a ten-component DGM to approximate the $\nu = 2$ Student t SIRV density of the previous section.

Figure 15.54 compares the performance of the optimal Student t receiver and an approximating DGM receiver for detection of a known amplitude target with $U(0, 2\pi)$ random phase in Student t clutter with shape parameter, $\nu = 2$. It shows that the DGM receiver does an excellent job of realizing the full detection capability. This DGM receiver is then used to evaluate the PD for Swerling one- and Swerling three-targets in this same clutter. These results are shown in

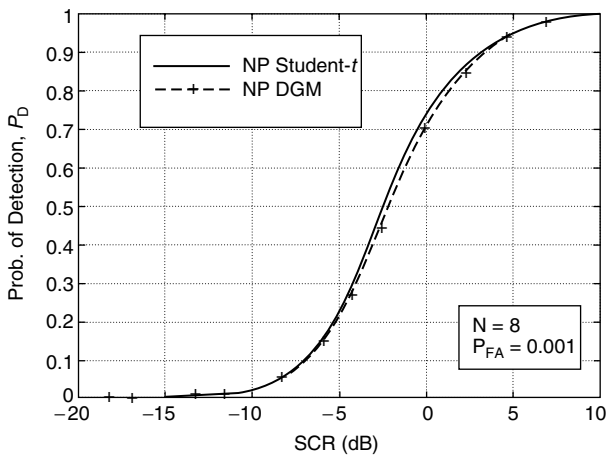

FIGURE 15.54 NP optimal receiver performance in DGM, 10 components and Student t SIRV clutter.

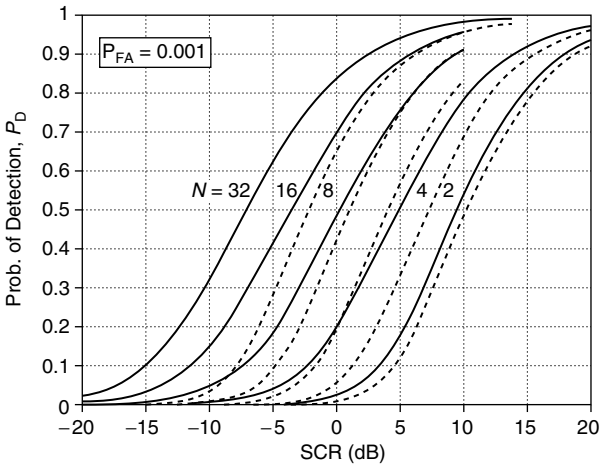

FIGURE 15.55 Optimal detection of a $U(0, 2\pi)$, Swerling one-target in DGM clutter.

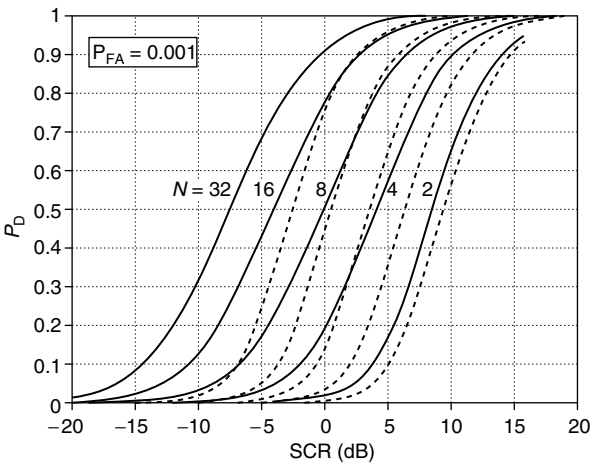

FIGURE 15.56 Optimal detection of a $U(0, 2\pi)$, Swerling three-target in DGM clutter.

Figure 15.55 and Figure 15.56 and again the optimal SIRV receiver significantly outperforms the Gaussian receiver for a significant range of SCR values.

15.6.4.5. NP vs. GLRT Receiver Comparison

A comparison of the detection probabilities of a Swerling one target in DGM clutter for the GLRT and the optimum NP receivers is shown in Figure 15.57 to Figure 15.59 at a .001 probability of false alarm. Once again, just as in the Student t SIRV clutter examples, the two receivers are seen to have essentially identical

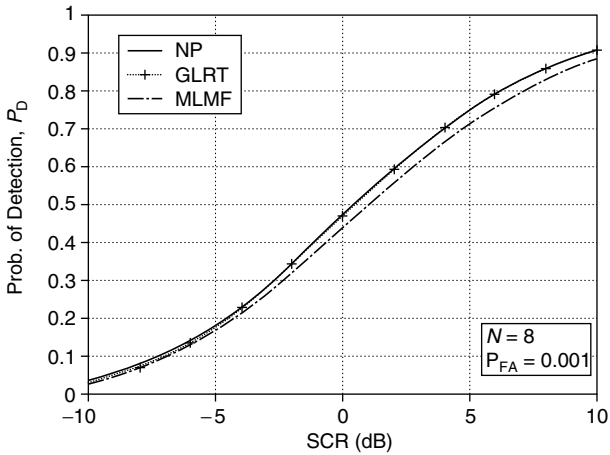

FIGURE 15.57 Detection probability of NP, GLRT, and maximum likelihood matched filter (MLMF) receivers in DGM clutter with $N = 8$ and $P_{FA} = 0.001$.

performance. Attempts to analytically verify the equivalent performance of these two receivers has been unsuccessful. However, several significant experimental results have been obtained from an investigation of this behavior.

1. As shown in Figure 15.57 to Figure 15.59, the number of samples used to estimate the signal amplitude and phase for the GLRT do not affect this phenomenon. Even when only two complex samples are used for

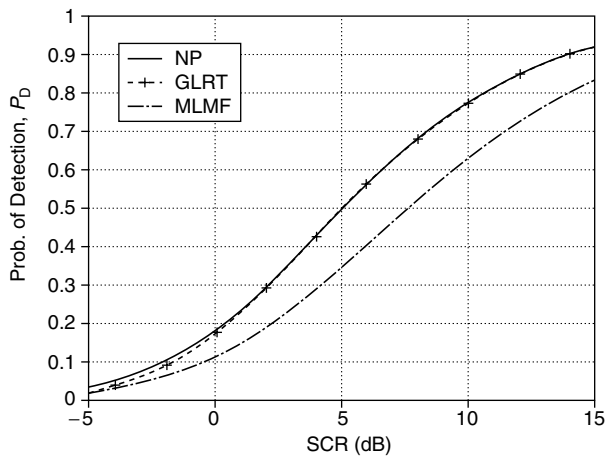

FIGURE 15.58 Detection probability of NP, GLRT, and maximum likelihood matched filter (MLMF) receivers in DGM clutter with $N = 4$ and $P_{FA} = 0.001$.

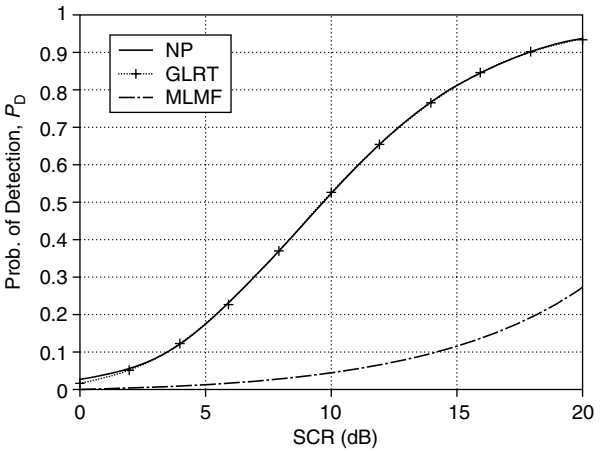

FIGURE 15.59 Detection probability of NP, GLRT, and maximum likelihood matched filter (MLMF) receivers in DGM clutter with $N = 2$ and $P_{FA} = 0.001$.

the estimates needed by the GLRT, it performs as well as the optimal NP receiver.

2. The same behavior is observed at both higher and lower false alarm probabilities.
3. The NP and GLRT are not equivalent test statistics, which does not necessarily rule out the possibility the two statistics have equivalent performance.
4. If these two receivers are used in SIRV clutter that is different from the clutter for which they are designed, the detection performances are still "equivalent", though no longer optimum.

These observations are important because they strengthen the case for using the GLRT in place of the NP statistic for optimal SIRV receivers. This is significant, because the GLRT is simpler than the NP receiver and has a closed-form solution for any SIRV for which the monotonic decreasing function, $h_N(q)$ is known. Figure 15.57 to Figure 15.59 also show the MLMF receiver performance. It is seen that the MLMF performance is significantly below optimum unless a sufficient number of samples are available for the required estimations. Nonetheless, if adequate samples are available for processing, this receiver performs very well and is by far the simplest of those considered.

15.6.4.6. Additional Implementation Issues

The receiver design and performance results presented above have been obtained by assuming exact knowledge of the clutter covariance matrix, target amplitude

PDF, and clutter PDF. In practical applications, these quantities are not known. Estimation of the covariance matrix for SIRV clutter includes all the problems associated with estimating the covariance matrix for Gaussian clutter, including selection of appropriate secondary data for the estimation. Problems raised by unknown target amplitude and clutter PDFs are not encountered in optimal Gaussian receiver design.

The optimal test statistic in Gaussian interference is UMP with respect to signal amplitude. Thus, the optimal receiver can be designed without knowledge of the signal amplitude or its PDF. No UMP test with respect to signal amplitude for optimal non-Gaussian receivers exists. Consequently, the receiver design depends on the value of a constant signal amplitude or the PDF of a random signal amplitude and it is expected that optimal receiver performance is degraded by the uncertain knowledge about the signal amplitude. A GLRT which uses a MLE of the amplitude may prove to be a suitable suboptimum receiver, depending on its detection loss for various target models.

Since the clutter PDF is not known *a priori* and only a relatively small number of clutter samples are available, an efficient PDF approximation algorithm such as the Öztürk algorithm[19] is necessary. Furthermore, since there are many possible types of non-Gaussian SIRV clutter, it is expected that the radar must select the best alternative from a library of receivers. The coarseness of this library depends on the sensitivity of detection performance to inaccuracy in the clutter PDF approximation. The optimal nonlinearities may be difficult to implement and in cases where no closed-form solution for the optimal SIRV receiver exists, a channelized receiver or other approximation becomes necessary. In addition, it is desirable for the library to cover a broad range of clutter densities. The DGM SIRV may prove particularly useful for this.

15.6.4.7. Summary

Several significant results in the area of optimal non-Gaussian receiver design in SIRV clutter have been presented. The significant improvement in detection performance of the optimal non-Gaussian receiver compared to conventional receiver design has been demonstrated. It is also shown that this optimal receiver has a canonical form which uses current Gaussian-based processing and has significant implications for space–time adaptive processing applications. The conventional matched filter is an integral component of this canonical architecture. It is shown that a whitening filter can be used to preprocess signals in correlated clutter without a loss in detection performance. The canonical form has also been used to develop a more efficient and reliable method of simulating the performance of the nonlinear receivers which arise in non-Gaussian processing. Finally, the first closed-form solution for an optimal non-Gaussian SIRV detector of a signal with random phase is presented.

15.7. MULTICHANNEL DETECTION FOR CORRELATED NONGAUSSIAN RANDOM PROCESSES BASED ON INNOVATIONS

(M. RANGASWAMY, J. H. MICHELS, AND D. D. WEINER)

15.7.1. Introduction

This work is motivated by a desire to detect signals in additive correlated non-Gaussian noise using multichannel data. The problem of signal detection in additive noise background is of interest in several areas such as radar, sonar, and digital communications. This problem has been addressed in great detail when the background noise is Gaussian.[1] However, the corresponding problem for the case of additive, correlated non-Gaussian noise has received limited attention.[2] Most of the previous work dealing with signal detection and estimation in non-Gaussian noise is based on the assumption that the noise samples are independent identically distributed (IID) random variables.[3,4] In many instances, the noise can be highly correlated. When the noise is a correlated non-Gaussian random process, there is no unique analytical expression for the joint PDF of N random variables obtained by sampling the noise process. The theory of spherically invariant random processes (SIRP) provides a powerful mechanism for modeling the joint PDF of the N correlated non-Gaussian random variables. Applications of SIRPs can be found in Refs. 5 and 6.

SIRPs are generalizations of the familiar Gaussian random process in that the PDF of every random vector obtained by sampling a SIRP is uniquely determined by the specification of a mean vector, a covariance matrix, and a characteristic first-order PDF. In addition, the PDF of a random vector obtained by sampling an SIRP is a monotonically decreasing function of a nonnegative quadratic form. However, the PDF does not necessarily involve an exponential dependence on the quadratic form, as in the Gaussian case. Many of the attractive properties of the Gaussian random process also apply to SIRPs. Every random vector obtained by sampling an SIRP is a SIRV.

Model-based parametric approaches for detection of time-correlated signals in nonwhite Gaussian noise for radar applications have received considerable attention.[1,7] An important feature of the model-based methods is their ability to utilize modern parameter estimators in the signal processing. In this scheme, random processes are whitened through a causal transformation of the observed data using prediction error filters. The resulting uncorrelated error processes are the innovations and contain, in a compact form, all the useful information about the processes. The innovations are useful for obtaining a sufficient statistic in hypothesis testing for the presence or absence of a desired signal.[8] Extension of these techniques for the multichannel problem using Gaussian noise has been considered in Refs. 7 and 9.

In this section, we present an innovations-based detection algorithm (IBDA) for multichannel signal detection in additive correlated non-Gaussian noise under

the assumption that the noise process can be modeled as an SIRP. Preliminary results of investigations for the single-channel case are available in Ref. 10. In particular, it is shown that the optimal estimator for obtaining the innovations process for SIRPs is linear and that the resulting detector has a canonical form. The detection architecture consists of a linear prediction filter followed by a memory less nonlinear transformation. Previous work[11] dealing with non-Gaussian processes has indicated that the innovations processes are obtained by using nonlinear prediction error filters. This approach, while having some interesting features, has several drawbacks and results in a suboptimal receiver. On the other hand, the work dealing with non-Gaussian random processes which can be modeled as SIRPs reveals that the optimal filter for obtaining the innovations process is linear. In addition, the IBDA developed in this section is optimal. Thus, the work of this chapter generalizes previous work in the area of signal detection in non-Gaussian noise.

This section is organized as follows:

In Section 15.7.2, we present a brief review of the theory of SIRPs. Section 15.7.3 describes the procedure for obtaining the innovations process. In Section 15.7.4, we present the IBDA for SIRPs. Section 15.7.5 presents detection results obtained from the IBDA for the case of the K-distributed SIRP. Estimator performance for SIRPs is discussed in Section 15.7.6. Finally, conclusions are presented in Section 15.7.7.

15.7.2. Preliminaries

The definitions and relevant mathematical preliminaries for complex SIRVs and complex SIRPs are presented in this section. A zero-mean random vector $\tilde{\mathbf{Y}} = \mathbf{Y}_c + j\mathbf{Y}_s$, where: $\mathbf{Y}_c = [Y_{c1}, Y_{c2}, \ldots, Y_{cN}]^T$ and $\mathbf{Y}_s = [Y_{s1}, Y_{s2}, \ldots, Y_{sN}]^T$ denote the vectors of the in-phase and out-of-phase quadrature components, is a complex SIRV if its PDF has the form

$$f_{\tilde{Y}}(\tilde{\mathbf{y}}) = (\pi)^{-N}|\tilde{\boldsymbol{\Sigma}}|^{-1}h_{2N}(p) \tag{15.291}$$

where: $p = \tilde{\mathbf{y}}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\mathbf{y}}$, $\tilde{\boldsymbol{\Sigma}}$ is a nonnegative definite Hermitian matrix, and $h_{2N}(\cdot)$ is a positive, real valued, monotonically decreasing function for all N. If every random vector obtained by sampling a complex random process $\tilde{y}(t)$ is a complex SIRV, regardless of the sampling instants or the number of samples, then the process $\tilde{y}(t)$ is defined to be a complex SIRP.

Yao, in Ref. 12, derived a representation theorem for real SIRV's. The representation theorem extends to complex SIRVs readily and is stated as follows.

The random vector $\tilde{Y}$ is a complex SIRV if and only if it is equivalent to the product of a complex Gaussian random vector $\tilde{Z}$ and an independent, nonnegative random variable V with PDF $f_V(v)$, which is defined to be the characteristic PDF of the complex SIRV.

Consequently,

$$\begin{aligned}\tilde{\mathbf{Y}} &= \tilde{\mathbf{Z}}V \\ f_{\tilde{\mathbf{Y}}}(\tilde{\mathbf{y}}) &= \int_0^\infty f_{\tilde{\mathbf{Y}}|V}(\tilde{\mathbf{y}}|v) f_V(v)\mathrm{d}v \\ f_{\tilde{\mathbf{Y}}|V}(\tilde{\mathbf{y}}|v) &= \pi^{-N}|\tilde{\boldsymbol{\Sigma}}|^{-1} v^{-2N} \exp\left(-\frac{p}{v^2}\right) \\ h_{2N}(p) &= \int_0^\infty v^{-2N} \exp\left(-\frac{p}{v^2}\right) f_V(v)\mathrm{d}v.\end{aligned} \tag{15.292}$$

It is assumed without loss of generality that $E(V^2) = 1$ so that the covariance matrix of the complex SIRV is equal to that of the complex Gaussian random vector. Due to the assumption $E(V^2) = 1$, the covariance matrix of the complex SIRV is $\tilde{\boldsymbol{\Sigma}}$.

The representation theorem and the assumption that $E(V^2) = 1$ give rise to the following necessary and sufficient conditions for representing $\tilde{\mathbf{Y}}$ as a complex SIRV

$$E\{\mathbf{Y}_\mathrm{c}\} = E\{\mathbf{Y}_\mathrm{s}\} = 0, \;\; \boldsymbol{\Sigma}_{\mathrm{cc}} = \boldsymbol{\Sigma}_{\mathrm{ss}}, \;\; \boldsymbol{\Sigma}_{\mathrm{cs}} = -\boldsymbol{\Sigma}_{\mathrm{sc}} \tag{15.293}$$

where:

$$\begin{aligned}\boldsymbol{\Sigma}_{\mathrm{cc}} &= E\{\mathbf{Y}_\mathrm{c}\mathbf{Y}_\mathrm{c}^\mathrm{T}\} \quad \boldsymbol{\Sigma}_{ss} = E\{\mathbf{Y}_\mathrm{s}\mathbf{Y}_\mathrm{s}^\mathrm{T}\} \\ \boldsymbol{\Sigma}_{\mathrm{cs}} &= E\{\mathbf{Y}_\mathrm{c}\mathbf{Y}_\mathrm{s}^\mathrm{T}\} \quad \boldsymbol{\Sigma}_{\mathrm{sc}} = E\{\mathbf{Y}_\mathrm{s}Y_\mathrm{c}^\mathrm{T}\}\end{aligned} \tag{15.294}$$

Under these conditions, it follows that

$$\tilde{\boldsymbol{\Sigma}} = 2[\boldsymbol{\Sigma}_{\mathrm{cc}} + j\boldsymbol{\Sigma}_{\mathrm{sc}}] \tag{15.295}$$

Several attractive properties of complex Gaussian random vectors generalize to complex SIRVs as a consequence of the representation theorem. Complex SIRVs satisfying the conditions of Equation 15.293 are also called SIRVs of the circular class.[13] For this chapter, the most important property of complex SIRVs of the circular class is the linearity of estimators in minimum mean-square error estimation (MMSE) problems.[13] This property is discussed in detail in Section 15.7.3.

15.7.3. Minimum Mean-Square Estimation Involving SIRPs

In MMSE problems, given a set of data, real SIRVs are found to result in linear estimators.[12,14,15] This property is readily extended to complex SIRV's in this section. Let $\tilde{\mathbf{Y}} = [\tilde{\mathbf{Y}}_1^\mathrm{T}\tilde{\mathbf{Y}}_2^\mathrm{T}]^\mathrm{T}$ where $\tilde{\mathbf{Y}}_1 = [\tilde{Y}_1, \tilde{Y}_2, \ldots, \tilde{Y}_m]^\mathrm{T}$ and $\tilde{\mathbf{Y}}_2 = [\tilde{Y}_{m+1}, \tilde{Y}_{m+2}, \ldots, \tilde{Y}_N]^\mathrm{T}$ denote the partitions of $\tilde{\mathbf{Y}}$. Note that $\tilde{Y}_i = Y_{ci} + jY_{si}$, $i = 1, 2, \ldots, N$. It can be readily shown that the MMSE estimate of $\tilde{\mathbf{Y}}_2$, given the

observations from $\tilde{\mathbf{Y}}_1$, is

$$\hat{\tilde{\mathbf{Y}}}_2 = E[\tilde{\mathbf{Y}}_2|\tilde{\mathbf{Y}}_1] = [C_{21}C_{11}^{-1}\tilde{\mathbf{y}}_1] \tag{15.296}$$

where:

C_{11}:	covariance matrix of $\tilde{\mathbf{Y}}_1$
C_{21}:	cross-covariance matrix of the vectors $\tilde{\mathbf{Y}}_2$ and $\tilde{\mathbf{Y}}_1$
$E[\tilde{\mathbf{Y}}_2\|\tilde{\mathbf{Y}}_1]$:	conditional mean or the expected value of $\tilde{\mathbf{Y}}_2$, given $\tilde{\mathbf{Y}}_1$.

Details of the derivation of Equation 15.296 are available in Ref. 16.

We assume that the complex SIRV $\tilde{\mathbf{Y}}$ is obtained by sampling a complex SIRP $\tilde{y}(k)$ at different time instants. Thus, for a given k, $\tilde{y}(k)$ is a complex random variable. The complex innovations sequence is defined as

$$\tilde{\epsilon}(k) = \tilde{y}(k) - \hat{\tilde{y}}(k) \tag{15.297}$$

where $\hat{\tilde{y}}(k)$ is the MMSE estimate of $\tilde{y}(k)$, given the observations $\tilde{y}(m)$, $m = 1, 2, \ldots k-1$. Since $\tilde{y}(k)$ is a complex SIRP, it follows that $\hat{\tilde{y}}(k)$ can be obtained by using Equation 15.296. This is achieved by the use of a linear prediction error filter whose coefficients are chosen to be equivalent to Equation 15.296. In particular, we use a complex autoregressive (AR) process of order two (AR[2] process) for approximating the complex SIRP. Details for specifying the linear prediction filter matrix coefficients are available in Refs. 9 and 17. The complex innovations process of Equation 15.297 has zero-mean, is uncorrelated, and is a complex SIRP having the same characteristic PDF as $\tilde{y}(k)$.

The problem of obtaining the single-channel innovations sequence for correlated non-Gaussian processes has also been considered in Refs. 11 and 18. The approach of Farina et al.[11,18] involved zero memory nonlinear (ZMNL) transformations, which transformed the processes from non-Gaussian to Gaussian. This was followed by a linear prediction filter and another ZMNL transformation, which gave rise to the innovations process. This approach has the following drawbacks.

(1) The correlation function at the output of the ZMNL transformation is related in a rather complicated manner to the correlation function at the input.
(2) The correlation function at the output of the ZMNL transformation is not guaranteed to be nonnegative definite.[19]
(3) If the process at the input of the ZMNL transformation is bandlimited, then the process at the output is also bandlimited if and only if the nonlinearity is a polynomial.[20]

Therefore, the approach using nonlinear transformations to obtain the innovations process is suboptimal. However, for non-Gaussian complex SIRPs, we have shown that the linear prediction filter is optimal for obtaining the

innovations sequence. The complex innovations sequence obtained in this section is used for developing the detection procedure of Section 15.7.4.

15.7.4. Innovations-Based Detection Algorithm for SIRPs Using Multichannel Data

We concern ourselves with the multichannel innovations-based detection algorithm (IBDA) for SIRPs in this section. In particular, we consider the binary multichannel detection problem for a known (nonrandom) signal in an additive SIRP interference. The interference is allowed to be correlated within a given channel as well as across channels. We present a model-based approach for this problem and show that the resulting receiver has a canonical form. The underlying interference process is assumed to be characterized by a multichannel auto regressive (AR) process. The detection procedure implements a likelihood ratio that is sensitive to the differences between the estimated model parameters under each hypothesis. Thus, the model-based approach is based on the contention that the coefficients of the received process are distinct for each of the two hypotheses. The innovations process arises naturally in this procedure.

We consider the following multichannel detection problem and derive the relevant likelihood ratio by two different methods. Consider the two hypotheses

$$\begin{aligned} H_0&: \tilde{\mathbf{x}}(\mathbf{k}) = \tilde{y}(\mathbf{k}) \quad k = 1, 2, \ldots, N \\ H_1&: \tilde{\mathbf{x}}(\mathbf{k}) = \tilde{s}(\mathbf{k}) + \tilde{y}(\mathbf{k}) \quad k = 1, 2, \ldots, N \end{aligned} \tag{15.298}$$

where:

H_0 and H_1:	absence and presence of the signal, respectively
$\tilde{\mathbf{x}}(\mathbf{k})$:	$J \times 1$ complex observation data vector
$\tilde{y}(\mathbf{k})$:	zero-mean complex SIRP
$\tilde{s}(\mathbf{k})$:	known $J \times 1$ constant complex signal
J:	number of channels.

15.7.4.1. Block Form of the Multichannel Likelihood Ratio

We first express the likelihood ratio for the multivariate SIRV PDFs as

$$\Lambda\{\tilde{\mathbf{x}}\} = \frac{f_{\tilde{X}|H_1}(\tilde{\mathbf{x}}|H_1)}{f_{\tilde{X}|H_0}(\tilde{\mathbf{x}}|H_0)} \tag{15.299}$$

where $\tilde{\mathbf{x}} = [\tilde{\mathbf{x}}(1)^T, \tilde{\mathbf{x}}(2)^T, \ldots, \tilde{\mathbf{x}}(N)^T]^T$. From the complex SIRV PDF of Equation 15.291, it follows that

$$f_{\tilde{\mathbf{X}}|H_i}(\tilde{\mathbf{x}}|H_i) = \pi^{-JN} |\mathbf{\Sigma}_{JN,H_i}|^{-1} h_{2JN}(q_{\tilde{\mathbf{x}}}|H_i) \;\; i = 0, 1 \tag{15.300}$$

where $q_{\tilde{\mathbf{x}}} = \tilde{\mathbf{x}}^H (\mathbf{\Sigma}_{JN,H_i})^{-1} \tilde{\mathbf{x}}$, $\mathbf{\Sigma}_{JN,H_i}$ is the $JN \times JN$ covariance matrix of the observed process $\tilde{\mathbf{x}}$ under the hypothesis H_i, $h_{2JN}(\cdot)$ is a positive, real valued,

monotonically decreasing function for all N and J, and $h_{2JN}(\cdot)$ is obtained by using Equation 15.292. Considering an $\mathbf{LDL}^H$ decomposition of $\mathbf{\Sigma}_{JN,H_i}$, where L is a lower triangular unit diagonal matrix, it can be readily shown that

$$\begin{aligned}\mathbf{\Sigma}_{JN,H_i}^{-1} &= [\mathbf{L}_{\tilde{\beta},H_i}^H]^{-1}\mathbf{D}_{\tilde{\beta},H_i}^{-1}(\mathbf{L}_{\tilde{\beta},H_i})^{-1} \\ |\mathbf{\Sigma}_{JN,H_i}| &= |\mathbf{L}_{\tilde{\beta},H_i}||\mathbf{D}_{\tilde{\beta},H_i}||\mathbf{L}_{\tilde{\beta},H_i}^H| = |\mathbf{D}_{\tilde{\beta},H_i}|\end{aligned} \tag{15.301}$$

where $\mathbf{D}_{\tilde{\beta},H_i}$ is a diagonal matrix. Since $\mathbf{\Sigma}_{JN,H_i} = \mathbf{L}_{\tilde{\beta},H_i}\mathbf{D}_{\tilde{\beta},H_i}\mathbf{L}_{\tilde{\beta},H_i}^H$, it follows that

$$\mathbf{D}_{\tilde{\beta},H_i} = E[\tilde{\beta}\tilde{\beta}^H] \tag{15.302}$$

where $\tilde{\beta} = \mathbf{L}_{\tilde{\beta},H_i}^{-1}\tilde{\mathbf{x}}$. Consequently

$$\begin{aligned}q_{\tilde{\mathbf{x}}}|H_i &= \tilde{\mathbf{x}}^H(\mathbf{\Sigma}_{JN,H_i})^{-1}\tilde{\mathbf{x}} = [\mathbf{L}_{\tilde{\beta},H_i}\tilde{\beta}]^H[\mathbf{L}_{\tilde{\beta},H_i}^H]^{-1}(\mathbf{D}_{\tilde{\beta},H_i})^{-1}\mathbf{L}_{\tilde{\beta},H_i}^{-1}[\mathbf{L}_{\tilde{\beta},H_i}\tilde{\beta}] \\ &= [\tilde{\beta}^H(\mathbf{D}_{\tilde{\beta},H_i})^{-1}\tilde{\beta}] = \sum_{j=1}^{J}\sum_{n=1}^{N}\frac{|\tilde{\beta}|^2}{\sigma_{jn}^2}(H_i) = q_{\tilde{\beta}}|H_i\end{aligned} \tag{15.303}$$

where σ_{jn}^2 is the jnth diagonal component of $\mathbf{D}_{\tilde{\beta},H_i}$. The vector $\tilde{\beta}$ is the multichannel complex innovations process.[7,9] Thus, a block form of a statistically equivalent innovations-based likelihood ratio can be written as

$$\Lambda\{\tilde{\beta}\} = \frac{f_{\tilde{\beta}|H_1}(\tilde{\beta}|H_1)}{f_{\tilde{\beta}|H_0}(\tilde{\beta}|H_0)} = \frac{|\mathbf{D}_{\tilde{\beta},H_1}|^{-1}h_{2JN}(q_{\tilde{\beta}}|H_1)}{|\mathbf{D}_{\tilde{\beta},H_0}|^{-1}h_{2JN}(q_{\tilde{\beta}}|H_0)} \tag{15.304}$$

Taking the natural logarithm of Equation 15.304, we obtain

$$\begin{aligned}\ln[\Lambda\{\tilde{\beta}\}] = &\left\{\sum_{j=1}^{J}\sum_{n=1}^{N}\ln[\sigma_{jn}^2(H_0)] - \ln[\sigma_{jn}^2(H_1)]\right\} + \ln[h_{2JN}(q_{\tilde{\beta}|H_1})] \\ &- \ln[h_{2JN}(q_{\tilde{\beta}|H_0})]\end{aligned} \tag{15.305}$$

Although Equation 15.296 for the MMSE estimate of $\tilde{\mathbf{Y}}_2$ is independent of the type of SIRP, observe that the test statistic of Equation 15.305 involves a nonlinearity $h_{2JN}(\cdot)$, which does depend on the type of SIRP.

15.7.4.2. Sequential Form of the Multichannel Likelihood Ratio

In order to obtain the sequential form of the multichannel likelihood ratio, we make use of the representation theorem. We start with the PDF of Equation 15.300. Due to the representation theorem for complex SIRVs, we can express the PDF of Equation 15.300 as

$$f_{\tilde{\mathbf{X}}|H_i}(\tilde{x}|H_i) = \int_0^\infty f_{\tilde{\mathbf{X}}|H_i,V}(\tilde{\mathbf{x}}|H_i, v)f_V(v)\mathrm{d}v \tag{15.306}$$

where

$$f_{\tilde{\mathbf{X}}|H_i,V}(\tilde{\mathbf{x}}|H_i,v) = \pi^{-JN}|\mathbf{\Sigma}_{JN,H_i}|^{-1}v^{-2JN}\exp\left[-\frac{q_{\tilde{\mathbf{x}}}|H_i}{v^2}\right] \tag{15.307}$$

The PDF of Equation 15.307 can be expressed as

$$\begin{aligned} f_{\tilde{\mathbf{X}}|H_i,V}(\tilde{\mathbf{x}}|H_i,v) = & f_{\tilde{\mathbf{X}}(1)|H_i,V}(\tilde{\mathbf{x}}(1)|H_i,v)\prod_{k=1}^{N} f_{\tilde{\mathbf{X}}(k)|H_i,\tilde{\mathbf{X}}(k-1),V} \\ & \times[\tilde{\mathbf{X}}(k)|H_i,\tilde{\mathbf{x}}(k-1),v]. \end{aligned} \tag{15.308}$$

Since the conditional PDF of Equation 15.307 is a complex Gaussian PDF, it follows that the conditional PDF's

$$f_{\tilde{\mathbf{X}}(k)|H_i,\tilde{\mathbf{X}}(k-1)}[\tilde{\mathbf{X}}(k)|H_i,\tilde{\mathbf{x}}(k-1),\tilde{\mathbf{x}}(k-2),\ldots,\tilde{\mathbf{x}}(1),v],\quad k=2,3,\ldots,N$$

are also complex Gaussian. Therefore, all that is needed to determine the PDF's

$$f_{\tilde{\mathbf{X}}(k)|H_i,\tilde{\mathbf{X}}(k-1)}[\tilde{\mathbf{X}}(k-2),\ldots,\tilde{\mathbf{X}}(1),V\cdot[\tilde{\mathbf{x}}(k)|H_i,\tilde{\mathbf{x}}(k-1),\tilde{\mathbf{x}}(k-2),\ldots,\tilde{\mathbf{x}}(1),v]],\ k=2,3,\ldots,N$$

is the specification of their conditional means and conditional covariance matrices. In particular, we need to specify

$$\begin{aligned} \hat{\tilde{\mathbf{X}}}(k) &= E[\tilde{\mathbf{X}}(k)|\tilde{\mathbf{X}}(k-1),\tilde{\mathbf{X}}(k-2),\ldots,\tilde{\mathbf{X}}(1),V] \\ \mathbf{\Sigma}_{\mathbf{k}|\mathbf{k}-1},V &= E[\{\tilde{\mathbf{X}}(\mathbf{k})-\hat{\tilde{\mathbf{X}}}(\mathbf{k})\}\{\tilde{\mathbf{X}}(\mathbf{k})-\hat{\tilde{\mathbf{X}}}(\mathbf{k})\}^H] \end{aligned} \tag{15.309}$$

Let

$$\begin{aligned} \tilde{\phi}(k) &= [\{\tilde{\mathbf{X}}(\mathbf{k})-\hat{\tilde{\mathbf{X}}}(\mathbf{k})\}]|H_i,V \\ M_{k|k-1,H_i,V} &= E[\tilde{\phi}(k)\tilde{\phi}^H(k)]. \end{aligned} \tag{15.310}$$

A further LDL^H decomposition of $\mathbf{M}_{k|k-1,H_i,V}$ results in the quadratic form $q_{\tilde{\mathbf{x}}}|H_i$ being expressed as

$$q_{\tilde{\mathbf{x}}}|H_i = \sum_{j=1}^{J}\sum_{n=1}^{N}\frac{|\tilde{\mathbf{\Gamma}}|^2}{\sigma_{jn}^2}(H_i) = q_{\tilde{\mathbf{\Gamma}}}|H_i \tag{15.311}$$

where $\tilde{\mathbf{\Gamma}}(k)$ is the complex zero-mean, uncorrelated innovations vector having a diagonal covariance matrix $\mathbf{D}_{\tilde{\Gamma},H_i}$, and σ_{jn} is the jnth diagonal component of $\mathbf{D}_{\tilde{\Gamma},H_i}$. Therefore, it follows that

$$f_{\tilde{\mathbf{X}}|H_i,V}(\tilde{\mathbf{x}}|H_i,v) = \pi^{-JN}|\mathbf{D}_{\tilde{\Gamma},H_i}|^{-1}v^{-2JN}\exp\left[-\frac{q_{\tilde{\mathbf{\Gamma}}}|H_i}{v^2}\right] \tag{15.312}$$

where $q_{\tilde{\mathbf{\Gamma}}}|H_i$ is defined in Equation 15.311. Using Equation 15.312 in Equation 15.306, and recognizing that the resulting integral is of the form of Equation

15.292, it follows that the PDF of Equation 15.306 can be expressed in terms of the corresponding multichannel innovations process. The likelihood ratio of the form of Equation 15.305 follows, using the approach Section 15.7.4.1. Thus, we have shown that the block form and sequential form of the innovations-based likelihood ratio are equivalent. We use the sequential approach for ease of implementation. The implementation of the linear prediction filter for obtaining the multichannel innovations process is discussed in detail in Ref. 9.

The detection procedure consists of comparing the innovations-based likelihood ratio of Equation 15.305 with a suitable threshold η. The threshold η is determined based on a desired false alarm probability. Therefore, the IBDA takes the form

$$\ln\{\Lambda\{\tilde{\mathbf{\Gamma}}\}] \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{15.313}$$

where

$$\Lambda\{\tilde{\mathbf{\Gamma}}\} = \frac{f_{\tilde{\mathbf{\Gamma}}|H_1}(\tilde{\mathbf{\Gamma}}|H_1)}{f_{\tilde{\mathbf{\Gamma}}|H_0}(\tilde{\mathbf{\Gamma}}|H_0)} = \frac{|D_{\tilde{\Gamma},H_1}|^{-1} h_{2JN}(q_{\tilde{\mathbf{\Gamma}}}|H_1)}{|D_{\tilde{\Gamma},H_0}|^{-1} h_{2JN}(q_{\tilde{\mathbf{\Gamma}}}|H_0)} \tag{15.314}$$

The resulting receiver consists of a whitening filter followed by a zero memory nonlinear (ZMNL) transformation under H_0 and H_1, where the difference is compared with a threshold. The detection architecture is shown in Figure 15.60. The detection architecture of Figure 15.60 is canonical for all complex SIRPs.

The innovations-based likelihood ratio enables understanding of the role of the nonlinear transformation in the detection architecture of Figure 15.60 and is presented here. Since $h_{2JN}(\cdot)$ is a monotonically decreasing function of the argument for all JN, $h_{2JN}(\cdot)$ decreases as $q\mathbf{\Gamma}|H_i$ increases. Furthermore, the error term $|\tilde{\Gamma}(k|H_i)|^2$ increases under the incorrect hypothesis in a statistical

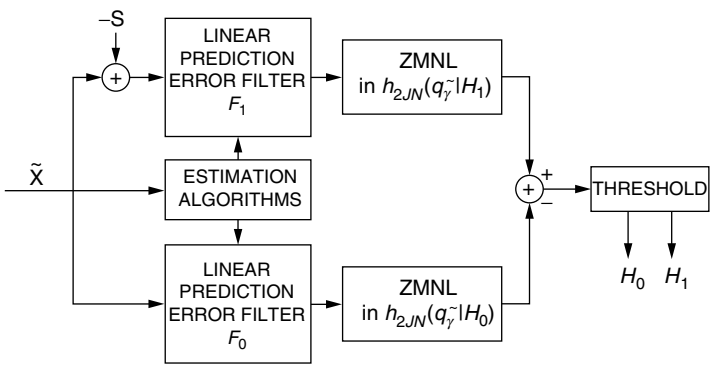

FIGURE 15.60 Multichannel detection architecture of the IBDA for the known constant signal case.

sense. Thus, $q_{\tilde{\Gamma}}|H_i$ increases proportionally via Equation 15.313 under the mismatched hypothesis condition. Therefore, the basis of the detection strategy is that under hypothesis H_0, the log-likelihood ratio of Equation 15.313 decreases, thus remaining below a fixed threshold. This is due to the fact that $h_{2JN}(q_{\tilde{\Gamma}}|H_1)$ decreases with increasing error signals under the mismatched hypothesis. Conversely, $h_{2JN}(q_{\tilde{\Gamma}}|H_0)$ increases, yielding a larger negative contribution in Equation 15.313. When H_1 is true, the respective roles of $h_{2JN}(q_{\tilde{\Gamma}}|H_0)$ and $h_{2JN}(q_{\tilde{\Gamma}}|H_1)$ are reversed, thus causing an increase in the log-likelihood ratio.

15.7.5. Detection Results Using Monte-Carlo Simulation

We present detection results using the detection architecture of Figure 15.60 in this section. Performance of the IBDA developed in Section 15.7.4 is determined by using computer simulation for the case of a known (nonrandom) signal in an additive *K*-distributed SIRP interference. We consider both the case of an uncorrelated SIRP and an AR SIRP of order two.

The *K*-distributed envelope PDF, whose quadrature components have a generalized-Laplace distribution, is useful for modeling non-Gaussian interference in radar[2] and communications[21] and is given by

$$f_R(r) = \frac{2b}{\Gamma(\alpha)}\left(\frac{br}{2}\right)^{\alpha} K_{\alpha-1}(br) \qquad 0 \le r \le \infty \tag{15.315}$$

where:

α: shape parameter of the distribution
b: scale parameter of the distribution
$K_N(t)$: Nth order modified Bessel function of the second kind.

It has been pointed out in Ref. 22 that the *K*-distribution is a member of the family of SIRV's. The characteristic PDF for the *K*-distributed SIRV and the corresponding $h_{2JN}(q)$ are

$$\begin{aligned} f_V(v) &= \frac{2b}{\Gamma(\alpha)}(bv)^{2\alpha-1}\exp(-b^2v^2), \quad 0 \le v \le \infty \\ h_{2JN}(q) &= \frac{2b^{2JN}}{\Gamma(\alpha)}(b\sqrt{q})^{\alpha-JN}K_{\alpha-NJ}(2b\sqrt{q}) \end{aligned} \tag{15.316}$$

Details of the Monte Carlo procedure are described in Ref. 9. The input observation data processes are generated using the process synthesis procedure described in Ref. 17, which has been modified to include the SIRP generation procedure developed in Ref. 22. Detection performance results are determined as a function of the channel signal-to-noise $(\text{SNR})_j$ ratios and the temporal and cross-channel correlation of the interference processes.

In several practical applications, *a priori* information may be available for determining the filter order and coefficients. Specifically, a "secondary data set"

is used to assess the hypothesis H_0 condition while a "primary data set" (statistically independent of the secondary set) is used for that of H_1. Under these conditions, the two data sets are used for estimating the filter coefficients F_0 and F_1, respectively.

In this section, the filter orders are assumed to be known *a priori*. For the case of a deterministic signal and SIRP interference modeled as an AR process of order two, the F_1 and F_0 filters are also of order two. However, the prediction error filter coefficients need to be determined. For the case where the interference spectrum is unknown, the filter coefficients are estimated with the Strand–Nuttall algorithm[23,24] using *a priori* data consisting of N_{TC} time samples. With the estimated prediction error filter coefficients, the detection results are then computed for $N_T = 2$ time samples per channel using $N_R = 10^5$ Monte-Carlo realizations and a false alarm probability (P_{FA}) of 10^{-4}.

Figure 15.61 shows the plot of PD vs. signal-to-noise ratio (SNR) for a known deterministic signal in an additive white *K*-distributed SIRP. Since for this case, the SIRP is known to be white *a priori* and uncorrelated across channels, the F_1 and F_0 filters are omitted.

Figure 15.61(a) shows the detection results for the single channel ($J = 1$) case. Curve one describes the case where the white *K*-distributed SIRP has shape parameter $\alpha = 0.1$. Curve two corresponds to the case of $\alpha = \infty$, i.e., when the interference is a white Gaussian SIRP. The $h_{2JN}(\cdot)$ for this case is simply $\exp(-q)$. Curve three shows the mismatch case, where the receiver designed for the *K*-distributed SIRP ($\alpha = 0.1$) is used to cancel white Gaussian interference. Finally, curve 4 shows the case where a receiver matched to the white Gaussian SIRP is used when the interference is a white *K*- distributed SIRP with $\alpha = 0.1$.

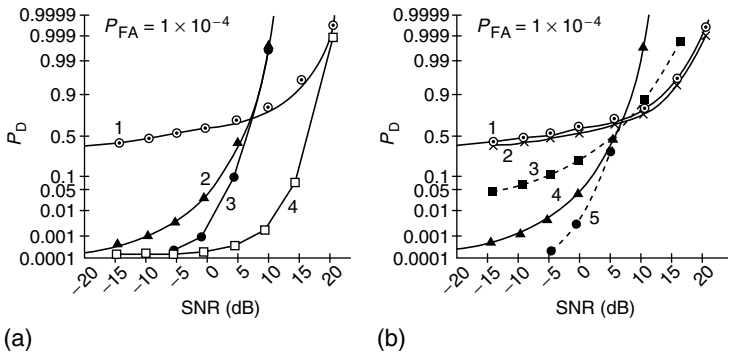

FIGURE 15.61 Probability of detection vs. SNR, known constant signal in white noise ($J = 1$, $N_T = 2$, shape parameter α is for a *K*-distributed PDF). (a) Curves one and two: receivers matched to *K* ($\alpha = 0.1$) and Gaussian noise, respectively. Curve three: *K* ($\alpha = 0.1$) receiver with Gaussian noise. Curve 4: Gaussian receiver with *K* ($\alpha = 0.1$) noise. (b) Curves one, three, and four: receivers matched to *K* ($\alpha = 0.1$), *K* ($\alpha = 0.5$), and Gaussian noise, respectively. Curves two and five: *K* ($\alpha = 0.5$) receiver with *K* ($\alpha = 0.1$) and Gaussian noise, respectively.

These curves reveal several important features. First, curves one and two show the potential for significant performance improvement at low SNRs when dealing with non-Gaussian processes. For example, at SNR ratio of -5 dB, the results for the K-distributed SIRP for $\alpha = 0.1$ show performance improvement of more than two orders of magnitude compared with the case of the Gaussian SIRP. Comparison of curves one and four reveals that there is a significant degradation of performance due to receiver mismatch. Comparison of curves two and three also shows a degradation in performance for the case where the non-Gaussian receiver is used, and the interference is Gaussian. However, the degradation for the latter case is less severe than in the former. Therefore, there is an improvement in performance when the interference is modeled correctly and the receiver matched appropriately to the interference is used, i.e., performance is degraded severely when the Gaussian receiver is used to cancel the non-Gaussian interference. However, we notice that the non-Gaussian receiver shows considerable robustness for shape parameter values in the range $\alpha = 0.1 - \infty$.

The robustness of the non-Gaussian receiver is further explored in Figure 15.61(b). Curves one and four of this figure are identical to curves one and two, respectively, of Figure 15.61(a). Curve three shows the case where the K-distributed SIRP has a shape parameter $\alpha = 0.5$, and the receiver is matched to this condition. Curve two corresponds to the case where the interference is a K-distributed SIRP with $\alpha = 0.1$ while the receiver is designed for $\alpha = 0.5$. Curve five pertains to the case where the interference is Gaussian and the receiver is designed for a K-distributed SIRP with $\alpha = 0.5$. Comparison of curves four and five shows the loss in performance for this mismatched receiver case. Comparison of curves one and two indicates that performance penalty is insignificant for the mismatch. However, we observe that the receiver designed for $\alpha = 0.5$ performs better than the receiver designed for $\alpha = 0.1$ while dealing with Gaussian interference. In summary, we point out that the receiver designed for $\alpha = 0.5$ is robust over a considerable range of shape parameter values.

The role of the ZMNL transformation is apparent in curves one, three, and four. For the Gaussian receiver, $h_{2JN}(q) = \exp(-q)$. It follows that the ZMNL transformation, i.e., $\ln[h_{2JN}(q)]$ is simply $-q$. However, for non-Gaussian SIRPs, $\ln[h_{2JN}(q)]$ is a monotonically decreasing function. Therefore, large values of q are mapped to smaller values of the test statistic and vice versa. Consequently, the threshold can be lowered in a weak signal situation, resulting in improved detection performance.

Figure 15.62 contains the plots of P_d vs. SNR for a known constant signal in known interference modeled as a K-distributed SIRP. The K-distributed SIRP was modeled as an AR(2) process with a normalized Doppler shift of 0.5. The results were obtained using $J = 2$, $N_T = 2$, and equal SNR's $(\text{SNR})_j$, $j = 1, 2$. The one lag temporal correlation parameter[9,17] on each channel was $\lambda_j = 0.7, j = 1, 2$. The K-distributed SIRP with shape parameter values $\alpha = 0.1$, 0.5, and ∞ was used. For each shape parameter value, the K-distributed SIRP was modeled with cross-channel correlation parameters $\rho_{ij} = 0$ and 0.99.

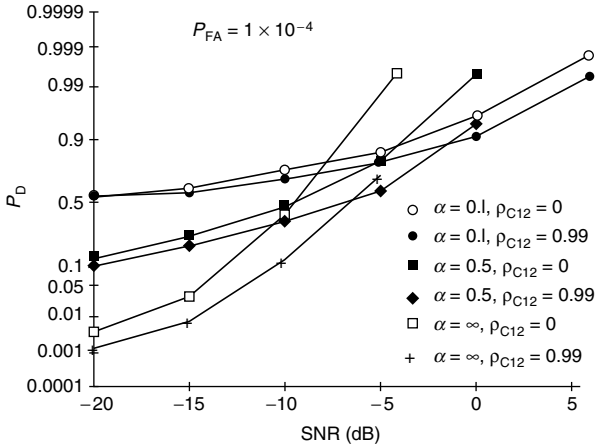

FIGURE 15.62 Probability of detection vs. SNR, known constant signal, AR(2) *K*-distributed SIRP noise, normalized clutter Doppler = 0.5, $J = 2$, $N_T = 2$, $\lambda_c = 0.7$.

Figure 15.63 considers the case where the *K*-distributed SIRP is approximated by an AR(2) process with unknown coefficients. These coefficients are estimated from the data for a signal to white noise ratio (SWNR) $= -10$ dB. In this case, N_{TC} data samples were used to estimate the AR(2) parameters. The plots show the convergence of the detection results to the corresponding known interference

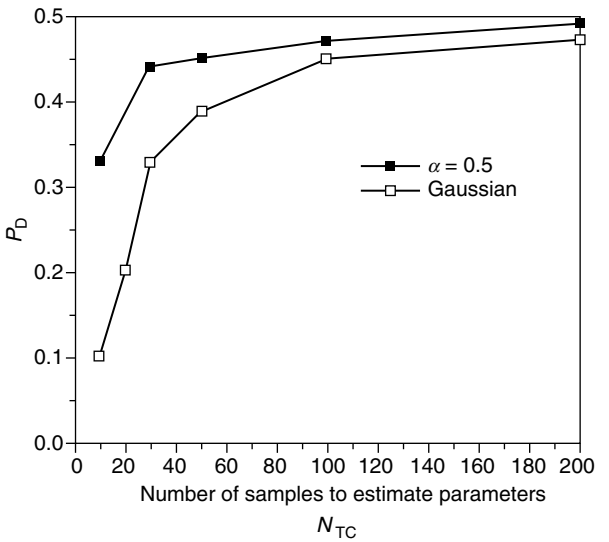

FIGURE 15.63 Probability of detection vs. number of time samples N_{TC} used to estimate AR(2) clutter parameters. $J = 2$, $N_T = 2$, $\lambda_{11} = \lambda_{22} = 0.7$, $\rho_c = 0$, SNR $= -10$ dB, $P_{FA} = 10^{-4}$.

case presented in Figure 15.62 as the estimator performance improves for large values of N_{TC}.

15.7.6. Estimator Performance for SIRPs

We discuss the effect of the estimated AR(2) coefficients on the detection performance in this section. In particular, we consider the $J \times 1$ multichannel AR vector as

$$\tilde{\mathbf{Y}}(k) = -\sum_{l=1}^{p} \mathbf{A}(l)\tilde{\mathbf{Y}}(k-l) + \tilde{\mathbf{U}}(k) \tag{15.317}$$

where:

p: model order
$\tilde{\mathbf{A}}(l)$: $J \times J$ matrix
$\tilde{\mathbf{U}}(k)$: $J \times 1$ white noise driving vector with covariance matrix $\boldsymbol{\Sigma}_u$.

We concern ourselves with the estimates of the 2×2 matrices $\tilde{\mathbf{A}}(l)$ and $\boldsymbol{\Sigma}_u$.

The PD shown in Figure 15.63 is obtained by averaging the results of nine independent detection runs. It is important to note that the detection probabilities converge to the values obtained for the known interference case of Figure 15.62 for increasing values of N_{TC}. The results are shown for a K-distributed SIRP with $\alpha = 0.5$ as well as the Gaussian case ($\alpha = \infty$). It is important to note that every realization of the SIRP corresponds to a fixed value of the random variable V. However, V is random from one realization to another according to Equation 15.292. We emphasize that the SIRP is nonergodic (except for the Gaussian case).

We have observed that both the magnitude of the bias as well as the error variance of the elements of $\tilde{\mathbf{A}}(k)$ fall off rapidly as N_{TC} increases. This result applies for both the K-distributed SIRP as well as the Gaussian process. A similar result was noted for the magnitude of the bias of the estimator for $\boldsymbol{\Sigma}_u$. However, Figure 15.64 shows an interesting result for the error variance of $\hat{\boldsymbol{\sigma}}_{u11}$, which is the one-one element of $\boldsymbol{\Sigma}_u$. Specifically, the error variance decreases for increasing N_{TC} for the Gaussian process. However, for the K-distributed SIRP, the error variance reaches a saturation level with increasing N_{TC}. Hence, the estimate $\hat{\boldsymbol{\sigma}}_{u11}$ is inconsistent for SIRPs. The inconsistency of the estimator is due to nonergodicity of the SIRP and the large tail behavior of the SIRP arising from small values of α. Consequently, the error variance of the estimator reaches a saturation level and is reported here for SIRPs for the first time. A similar result was reported in Ref. 4 while dealing with a different class of non-Gaussian phenomena.

We can now argue that the reduction in the level of the bias magnitude results in improvement of detection performance as N_{TC} increases. Our results also show that the receiver for the K-distributed SIRP is less sensitive to the estimation error as opposed to the Gaussian receiver. This is explained by the argument that $h_{2JN}(\cdot)$ appearing in the likelihood ratio of Equation 15.314 is a monotonically

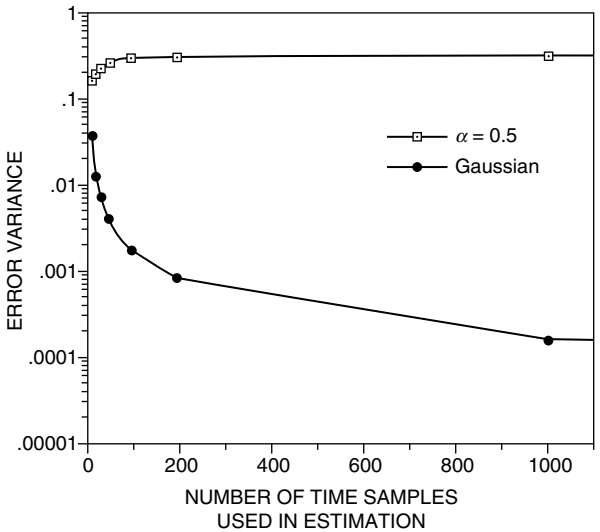

FIGURE 15.64 Error variance vs. number of time samples N_{TC} used in the time-averaged estimate of $\boldsymbol{\Sigma}_{11}$ for K-distributed and Gaussian processes. One-lag temporal correlation parameter $\lambda_{ii} = 0.7$.

decreasing function. Thus, large values of the argument of $h_{2JN}(\cdot)$ are mapped into small values. Large values in the argument of $h_{2JN}(\cdot)$ arise on account of the errors in the AR parameter estimation resulting from the nonergodicity of the SIRP (different realizations of the SIRP corresponding to different values of the random variable V). However, since $h_{2JN}(\cdot)$ arises from a representation of the form of Equation 15.292, an averaging over the random variable V results. Hence, the detection results are insensitive to the errors in the parameter estimation problem.

15.7.7. Conclusion

This section has made several significant original contributions. First, the complex SIRP model for correlated non-Gaussian random processes was developed. Next, it was shown that the innovations process for complex SIRPs can be obtained by using linear prediction filters. The IBDA was then developed for known signals in additive complex SIRP interference using multichannel observations and showed that the resulting receiver was canonical. Finally, a performance analysis of the IBDA was carried out showing excellent agreement with that of Ref. 2. Performance analysis of the IBDA also indicates that the receiver is robust for a variety of shape parameters. Furthermore, it was found that the receiver designed for the K-distributed SIRP is less sensitive to estimation errors arising from the estimates of the parameters of the AR process than the Gaussian receiver.

Appendices

CONTENTS

Appendix A. Stochastic Representation for the Normalized Generalized Inner Product (Section 3.1) 1040
Appendix B. Expectation-Maximization Algorithm for Covariance Matrix Estimation (Section 3.2) 1041
Appendix C. Algebraic Derivations for Johnson Distributions (Section 4.2) 1044
C.1. Johnson S_U Distribution 1044
C.2. Johnson S_B Distribution 1049
C.3. Johnson S_L Distribution 1056
Appendix D. Connections Between g_a, k_a, P_a (Section 4.2) 1058
Appendix E. Cancellation for an Analog Hybrid Canceler (Section 10.1) 1059
Appendix F. Cancellation for a Digital Hybrid Canceler (Section 10.1) 1060
Appendix G. Matrix Elements in Equation 10.10 (Section 10.1) 1062
Appendix H. Asymptotic Cancellation Curves (Section 10.1) 1063
Appendix I. Optimum Values of N and M (Section 10.1) 1065
Appendix J. Effect of Near-Field Nulling Constraint (Section 10.2) 1067
Appendix K. Equivalence of Element Space and Beam Space Results (Section 10.4) 1069
Appendix L. Evaluation of the Integrals in Equation 10.128 and Equation 10.129 (Section 10.4) 1070
Appendix M. Calculation of the Adaptive Weights (Section 10.5) 1072
Appendix N. Elimination of False Targets (Section 10.5) 1075
Appendix O. Approximate Derivation of Equation 10.165 (Section 10.5) 1076
Appendix P. Interference Covariance Matrix (Section 10.10) 1079
Appendix Q. Number of Time Taps Required (Section 10.10) 1082
Appendix R Inclusion of Polarization (Section 10.10) 1084
Appendix S. Signal Cancellation in First Stage Beamformer (Section 10.11) 1085
Appendix T. Interferer-Free Limit of Equation 10.298 (Section 10.11) 1088

Appendix U. Properties of SIRVs (Section 12.2) 1089
U.1. Statistical Independence 1089
U.2. Spherically Symmetric Characteristic Function 1090
U.3. Relationship between Higher Order and Lower Order SIRV PDFs 1091
Appendix V. Computer Generation of SIRVs Using Rejection Method (Section 12.4) 1092
V.1. Rejection Method 1092
V.2. Rejection Theorem 1093
Appendix W. Maximum Likelihood Estimation Involving SIRVs (Section 12.5) 1095
Appendix X. Issues Related to Extreme Value Theory (Section 13.3) 1102
X.1. Limiting Forms for the Largest Order Statistic 1102
X.1.1. Case 1 1103
X.1.2. Cases 2 and 3 1105
X.2. Tails of Probability Density Functions 1106
X.2.1. Case 1 1107
X.2.2. Case 2 1107
X.2.3. Case 3 1107
X.3. PDF of the rth Order Statistic 1108
Appendix Y. Canonical Form Derivation (Section 15.6) 1109
Appendix Z. Alternative Spherical Coordinate SIRV Representations (Section 15.6) 1111

APPENDIX A. STOCHASTIC REPRESENTATION FOR THE NORMALIZED GENERALIZED INNER PRODUCT (SECTION 3.1)

Let $\mathbf{Z}$ denote a data matrix whose columns are the previously defined $\mathbf{z}_i$, $i = 1, 2, \ldots, K$. The maximum likelihood estimate of the covariance matrix is given by $\hat{\mathbf{R}} = (1/k)\mathbf{Z}\mathbf{Z}^H$. The data matrix $\mathbf{Z}$ and the test data vector $\mathbf{x}$ admit a representation of the form

$$\mathbf{Z} = \mathbf{R}^{1/2}\mathbf{Y} \qquad \mathbf{x} = \mathbf{R}^{1/2}\mathbf{y} \tag{A.1}$$

where $\mathbf{Y}$ is a data matrix whose columns $\mathbf{y}_i$, $i = 1, 2, \ldots, K$ are IID $CN(0, \mathbf{I})$ random vectors and $\mathbf{y}$ is a $CN(0, \mathbf{I})$ random vector which is statistically independent of $\mathbf{Y}$. Hence, the normalized GIP is expressed as

$$P' = \mathbf{y}^H \mathbf{S}_y^{-1} \mathbf{y} \tag{A.2}$$

where $\mathbf{S}_y = (1/K)\mathbf{Y}\mathbf{Y}^H$. Next, we use a Householder transformation defined by $\mathbf{A} = \left(\mathbf{I} - 2(\mathbf{u}\mathbf{u}^H/\mathbf{u}^H\mathbf{u})\right)$, where $\mathbf{u} = \mathbf{y} - \|\mathbf{y}\|\mathbf{e}$ and $\mathbf{e} = [100\ldots0]^{\mathrm{T}}$ so that $\tilde{\mathbf{y}} = \mathbf{A}\mathbf{y}$. Also let $\tilde{\mathbf{Y}} = \mathbf{A}\mathbf{Y}$. Since $\mathbf{A}\mathbf{A}^H = \mathbf{A}^H\mathbf{A} = \mathbf{I}$, it follows that the statistics of $\tilde{\mathbf{Y}}$ are

identical to that of $\mathbf{Y}$. Consequently, the normalized GIP is expressed as

$$P' = \tilde{\mathbf{y}}^H S_y^{-1} \tilde{\mathbf{y}}. \tag{A.3}$$

Furthermore, we partition $\tilde{\mathbf{Y}} = \begin{bmatrix} \mathbf{y}_1^H \\ \mathbf{Y}_{11}^H \end{bmatrix}$ where $\mathbf{y}_1^H$ is the first row of $\tilde{\mathbf{Y}}$ and $\mathbf{Y}_{11}^H$ denotes the $(M-1)\times K$ matrix formed from the remaining rows of $\tilde{\mathbf{Y}}$. Consequently,

$$\mathbf{S}_{\tilde{y}} = \begin{bmatrix} \mathbf{y}_1^H \mathbf{y}_1 & \mathbf{y}_1^H \mathbf{Y}_{11} \\ \mathbf{Y}_{11}^H \mathbf{y}_1 & \mathbf{Y}_{11}^H \mathbf{Y}_{11} \end{bmatrix}. \tag{A.4}$$

Also, $\mathbf{S}_{\tilde{y}}^{-1}$ admits a representation of the form

$$\mathbf{S}_{\tilde{y}}^{-1} = \begin{bmatrix} \mathbf{S}_{11} & \mathbf{S}_{12} \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix}. \tag{A.5}$$

Finally, the normalized GIP is expressed as

$$P' = \|\mathbf{y}\|^2 \mathbf{S}_{11}.$$

However, from the matrix inversion Lemma it follows that $\mathbf{S}_{11} = (\mathbf{y}_1^H \mathbf{P}_\perp \mathbf{y}_1)^{-1}$ where $\mathbf{P}_\perp = \mathbf{I} - \mathbf{Y}_{11}(\mathbf{Y}_{11}^H \mathbf{Y}_{11})^{-1} \mathbf{Y}_{11}^H$. Since $\mathbf{Y}_{11}(\mathbf{Y}_{11}^H \mathbf{Y}_{11})^{-1} \mathbf{Y}_{11}^H$ is a projection matrix of rank $M-1$, it follows that $\mathbf{P}_\perp$ is a projection matrix of rank $K - M + 1$. Consequently,

$$\mathbf{y}_1^H \mathbf{P}_\perp \mathbf{y}_1 = \sum_{i=1}^{K-M+1} |y(i)|^2 \tag{A.6}$$

where $y(i) \sim CN(0,1)$. Hence $\mathbf{S}_{11}$ is simply the reciprocal of a Chi-Squared distributed random variable with $(K - M + 1)$ complex degrees-of-freedom. Also, since $\mathbf{y}$ is a $CN(0,\mathbf{I})$ random vector, $\|\mathbf{y}\|^2$ is a Chi-Squared distributed random variable with M complex degrees of freedom. Consequently, the representation of Equation 3.5 follows.

APPENDIX B. EXPECTATION-MAXIMIZATION ALGORITHM FOR COVARIANCE MATRIX ESTIMATION (SECTION 3.2)

We discuss the maximum likelihood estimation of the SIRV covariance matrix in this appendix. Let $\mathbf{X}$ denote a data matrix, whose columns $x_i, i = 1,2,\ldots,K$ are independent identically distributed training data vectors, which are distributed as SIRV$[0,\mathbf{R}_x,f_V(v)]$. The likelihood function for estimating $\mathbf{R}$ if given by

$$g[\mathbf{X}|\mathbf{R}] = \prod_{i=1}^{K} \pi^{-M} |\mathbf{R}|^{-1} h_{2M}(q_i). \tag{B.1}$$

Direct maximization of the likelihood function of Equation B.1 over $\mathbf{R}$ is rendered difficult because of the missing information. Consequently, it is helpful to treat the problem in the context of a complete–incomplete data problem.[28] Recall from the representation theorem for SIRVs[22] that $x_i = z_i V_i$, where z_i, $i = 1, 2, \ldots, K$ are statistically independent $CN(0, \mathbf{R})$ random vectors, and V_i, $i = 1, 2, \ldots, K$ are statistically independent random variables with PDF $f_V(v)$. For this problem, the complete data is either z_i, V_i, $i = 1, 2, \ldots, K$ or x_i, V_i, $i = 1, 2, \ldots, K$. However, the observed data x_i, $i = 1, 2, \ldots, K$, contains no explicit information about V_i, $i = 1, 2, \ldots, K$ and thus constitutes the incomplete data. The complete data likelihood function is given by the joint PDF of x_i, V_i, $i = 1, 2, \ldots, K$, which is expressed as

$$g_c[\mathbf{X}, V_i|\mathbf{R}] = \prod_{i=1}^{K} f(\mathbf{x}_i|V_i) \prod_{i=1}^{K} f(v_i) \tag{B.2}$$

Taking the natural logarithm of Equation B.2 yields the complete-data log-likelihood function of the form

$$L[\mathbf{X}, V_i|\mathbf{R}] = -KM \log(\pi) - K \log(|\mathbf{R}|) - \sum_{i=1}^{K} q_i v_i^{-2} + \sum_{i=1}^{K} \log[v_i^{-2M} f(v_i)]. \tag{B.3}$$

Note that given an initial estimate of $\mathbf{R}$ denoted by $\hat{\mathbf{R}}$, the quantity,

$$E\{\log[v_i^{-2M} f(v_i)]|\hat{\mathbf{R}}\} \tag{B.4}$$

depends only on $\hat{\mathbf{R}}$ and not on $\mathbf{R}$. Consequently, the relevant terms for the maximization over $\mathbf{R}$ are given by

$$L_1[\mathbf{X}, V_i|\mathbf{R}] = -K \log(|\mathbf{R}|) - \sum_{i=1}^{K} q_i v_i^{-2} \tag{B.5}$$

The missing data, v_i, $i = 1, 2, \ldots, K$, are assumed to be missing at random (MAR).[28] Consequently, given an initial estimate of $\mathbf{R}$ denoted by $\hat{\mathbf{R}}$, the complete data sufficient statistic[28] is given by

$$c_i = E[V_i^{-2}|\hat{\mathbf{R}}, \mathbf{x}_i]. \tag{B.6}$$

Thus, c_i is simply, the minimum mean squared error (MMSE) estimate of V_i^{-2} given $\mathbf{x}_i$. Note that $f(v_1) = f(v_2) = \ldots = f(v_K) = f_V(v)$ (since v_i, $i = 1, 2, \ldots, K$ are independent identically distributed random variables). Therefore,

$$f_{V_i|\mathbf{X}_i, \hat{\mathbf{R}}}(v_i|\mathbf{x}_i, \hat{\mathbf{R}}) = \frac{f(\mathbf{x}_i|v_i, \hat{\mathbf{R}}) f_V(v_i)}{f_{\mathbf{X}_i|\hat{\mathbf{R}}}(\mathbf{x}_i|\hat{\mathbf{R}})}. \tag{B.7}$$

However,

$$\frac{f(\mathbf{x}_i|v_i, \hat{\mathbf{R}}) f_V(v_i)}{f_{\mathbf{X}_i|\hat{\mathbf{R}}}(\mathbf{x}_i|\hat{\mathbf{R}})} = \frac{v_i^{-2M} \exp(q_i v_i^{-2}) f_V(v_i)}{h_{2M}(q_i)}. \tag{B.8}$$

Consequently,

$$c_i = \frac{\int_0^\infty v_i^{-2M-2}\exp(q_i v_i^{-2}) f_V(v_i)\mathrm{d}v_i}{h_{2M}(q_i)} = -\frac{h'_{2M}(q_i)}{h_{2M}(q_i)} = \frac{h_{2M+2}(q_i)}{h_{2M}(q_i)} \qquad \text{(B.9)}$$

Having specified the complete data sufficient statistic, we seek the maximization of Equation B.5. For this purpose, we reproduce the following matrix differentiation identities from Ref. 42.

$$\delta[\mathbf{R}^{-1}] = -\mathbf{R}^{-1}\delta[\mathbf{R}]\mathbf{R}^{-1} \qquad \delta[\log|\mathbf{R}^{-1}|] = -tr\{\mathbf{R}^{-1}\delta[\mathbf{R}]\} \qquad \text{(B.10)}$$

Further, we recognize that $q_i = \sum_{i=1}^K tr[\mathbf{R}^{-1}\mathbf{x}_i\mathbf{x}_i^H]$. Consequently,

$$\delta L_1[\mathbf{X}, V_i|\mathbf{R}] = Ktr\{\mathbf{R}^{-1}\delta[\mathbf{R}]\} - tr\left[\mathbf{R}^{-1}\delta[\mathbf{R}]\mathbf{R}^{-1}\sum_{i=1}^K c_i\mathbf{x}_i\mathbf{x}_i^H\right]. \qquad \text{(B.11)}$$

Maximization of Equation B.5 results from setting Equation B.11 equal to zero. Therefore, the maximum likelihood estimate of $\mathbf{R}$ is given by

$$\hat{\mathbf{R}} = \frac{1}{K}\sum_{i=1}^K c_i\mathbf{x}_i\mathbf{x}_i^H. \qquad \text{(B.12)}$$

Since $\hat{\mathbf{R}}$ appears on both sides of Equation B.12 (implicitly on the right hand side through the calculation of c_i), it is not possible to obtain a closed form solution for the resulting estimate. Consequently, an iterative method is needed for calculating $\hat{\mathbf{R}}$. The EM algorithm, which provides an iterative solution to this problem, is summarized below.

1. *E Step*: Given an initial estimate of $\mathbf{R}$ denoted by $\hat{\mathbf{R}}$, calculate c_i for $i = 1, 2, \ldots, K$.
2. *M Step*: Calculate $\hat{\mathbf{R}} = \frac{1}{K}\sum_{i=1}^K c_i\mathbf{x}_i\mathbf{x}_i^H$.
3. Use $\hat{\mathbf{R}}$ from step two to calculate a new value of c_i.
4. Iterate until convergence. Convergence is determined through a suitable error criterion.

In Section 3.2, the convergence criterion is an error of 10^{-6} defined to be the Frobenius norm of the difference between the values of $\hat{\mathbf{R}}$ resulting from two successive iterations. At convergence, the resultant $\hat{\mathbf{R}}$, is to within a multiplicative constant of the sample covariance matrix. This is due to the fact that the outer product of each training data realization with itself is scaled by the MMSE estimate of V_i^{-2}. This fact has been verified for all the simulated data examples presented in the paper. In particular, we examined the diagonal matrix of eigen values of the estimated covariance matrix. We found that they were to within a multiplicative constant of the eigenvalues of the sample covariance matrix formed by averaging the outer products of the realizations $\mathbf{z}_i$, $i = 1, 2, \ldots, K$, of the Gaussian component of the SIRV $\mathbf{x}_i$. Convergence of the

algorithm is dictated by the choice of the initial estimate of $\mathbf{R}$. Any positive definite Hermitian matrix is suitable for the initial estimate of $\mathbf{R}$. Two choices, which readily arise are the $M \times M$ identity matrix, $\mathbf{I}_M$, and the sample covariance matrix given by $\mathbf{S} = \frac{1}{K} \sum_{i=1}^{K} \mathbf{x}_i \mathbf{x}_i^H$. We employ the latter choice in this paper due to the fact that it yields faster convergence.

The simulated data examples considered in this paper involve the calculation of the modified Bessel function of the second kind for specifying $h_{2M}(\cdot)$ and its derivative. Numerical errors in their calculation for $\alpha = 0.1$ tend to be rather large. Consequently, convergence of the algorithm is extremely slow for $\alpha = 0.1$.

APPENDIX C. ALGEBRAIC DERIVATIONS FOR JOHNSON DISTRIBUTIONS (SECTION 4.2)

In this appendix, the criteria given in Equation 4.52 of Section 4.2 are established and the parameter estimates given in Equation 4.57, Equation 4.59, and Equation 4.61 are developed.

C.1. Johnson S_U Distribution

The transformation for the *Johnson S_U Distribution* is of the form

$$R = \gamma + \eta \sinh^{-1}\left(\frac{G - \epsilon}{\lambda}\right) \tag{C.1}$$

where R is a standard normal variable ϵ and is a location parameter, λ is a scale parameter, and γ and η are shape parameters for the PDF of the *S_U Distribution*. Solving Equation C.1 for G in terms of R, there results

$$G = \epsilon + \lambda \sinh\left(\frac{R - \gamma}{\eta}\right) = \epsilon - \lambda \sinh\left(\frac{\gamma - R}{\eta}\right) \tag{C.2}$$

where we have made use of the fact that $\sinh(A)$ is an odd function of A. Define

$$m = g_{3r} - g_r \qquad l = g_{-r} - g_r \qquad p = g_r - g_{-r} \tag{C.3}$$

where $g_{\pm r}$, $g_{\pm 3r}$ are obtained from Equation C.2 for $\pm r$ and $\pm 3r$, respectively.

For a fixed positive value of r, Equation C.2 and Equation C.3 give

$$\begin{aligned} m &= \lambda\left[\sinh\left(\frac{\gamma - r}{\eta}\right) - \sinh\left(\frac{\gamma - 3r}{\eta}\right)\right] \\ &= \lambda\left[\sinh\left(\frac{\gamma - 2r + r}{\eta}\right) - \sinh\left(\frac{\gamma - 2r - r}{\eta}\right)\right], \\ l &= \lambda\left[\sinh\left(\frac{\gamma + 3r}{\eta}\right) - \sinh\left(\frac{\gamma + r}{\eta}\right)\right] \\ &= \lambda\left[\sinh\left(\frac{\gamma + 2r + r}{\eta}\right) - \sinh\left(\frac{\gamma + 2r - r}{\eta}\right)\right], \\ p &= \lambda\left[\sinh\left(\frac{\gamma + r}{\eta}\right) - \sinh\left(\frac{\gamma - r}{\eta}\right)\right]. \end{aligned} \tag{C.4}$$

Using the standard formula

$$\sinh(A + B) - \sinh(A - B) = 2\cosh A \sinh B. \tag{C.5}$$

we obtain the values of m, l and p from Equation C.4 as

$$\begin{aligned} m &= 2\lambda \cosh\left(\frac{\gamma - 2r}{\eta}\right)\sinh\left(\frac{r}{\eta}\right) \\ l &= 2\lambda \cosh\left(\frac{\gamma + 2r}{\eta}\right)\sinh\left(\frac{r}{\eta}\right) \\ p &= 2\lambda \cosh\left(\frac{\gamma}{\eta}\right)\sinh\left(\frac{r}{\eta}\right). \end{aligned} \tag{C.6}$$

Thus,

$$\frac{m}{p} = \frac{\cosh\left(\frac{\gamma - 2r}{\eta}\right)}{\cosh\left(\frac{\gamma}{\eta}\right)} \qquad \frac{l}{p} = \frac{\cosh\left(\frac{\gamma + 2r}{\eta}\right)}{\cosh\left(\frac{\gamma}{\eta}\right)}, \tag{C.7}$$

which gives

$$\frac{ml}{p^2} = \frac{\cosh\left(\frac{\gamma - 2r}{\eta}\right)\cosh\left(\frac{\gamma + 2r}{\eta}\right)}{\cosh^2\left(\frac{\gamma}{\eta}\right)}. \tag{C.8}$$

Using the identity

$$\cosh(A + B)\cosh(A - B) = \cosh^2 A + \cosh^2 B - 1, \tag{C.9}$$

we obtain from Equation C.8

$$\frac{ml}{p^2} = \frac{\cosh^2\left(\frac{\gamma}{\eta}\right) + \cosh^2\left(\frac{2r}{\eta}\right) - 1}{\cosh^2\left(\frac{\gamma}{\eta}\right)} = 1 + \frac{\cosh^2\left(\frac{2r}{\eta}\right) - 1}{\cosh^2\left(\frac{\gamma}{\eta}\right)}. \tag{C.10}$$

Since $\cosh^2(2r/\eta) > 1$, it is clear that $ml/p^2 > 1$ for the S_U distribution.

Applying the identity.

$$\cosh(A + B) + \cosh(A - B) = 2\cosh A \cosh B. \tag{C.11}$$

to the sum of m/p and l/p in Equation C.7, we get

$$\frac{m}{p} + \frac{l}{p} = \frac{2\cosh\left(\frac{\gamma}{\eta}\right)\cosh\left(\frac{2r}{\eta}\right)}{\cosh\left(\frac{\gamma}{\eta}\right)} = 2\cosh\left(\frac{2r}{\eta}\right). \tag{C.12}$$

Solving Equation C.12 for η, we get

$$\eta = \frac{2r}{\cosh^{-1}\left[\frac{1}{2}\left(\frac{m}{p} + \frac{l}{p}\right)\right]} \qquad \eta > 0. \tag{C.13}$$

Since $\eta > 0$ is assumed, the positive value of the inverse of $\cosh(\cdot)$ must be chosen.

The expression for $\cosh(2r/\eta)$ in Equation C.12 can be substituted into Equation C.10 to give

$$\frac{ml}{p^2} = \frac{\cosh^2\left(\frac{\gamma}{\eta}\right) + \left(\frac{m+l}{2p}\right)^2 - 1}{\cosh^2\left(\frac{\gamma}{\eta}\right)}. \tag{C.14}$$

Solving for $\cosh^2(\gamma/\eta)$,

$$\cosh^2\left(\frac{\gamma}{\eta}\right) = \frac{(m+l)^2 - 4p^2}{4(ml - p^2)}. \tag{C.15}$$

Since $\cosh^2(A) - \sinh^2(A) = 1$, Equation C.15 leads to

$$\sinh^2\left(\frac{\gamma}{\eta}\right) = \frac{(m-l)^2}{4(ml - p^2)} \tag{C.16}$$

Thus,

$$\sinh\left(\frac{\gamma}{\eta}\right) = \frac{[(m-l)^2]^{1/2}}{2(ml - p^2)^{1/2}}. \tag{C.17}$$

Unlike η, the parameter γ may be either positive or negative. Thus, a determination of the sign of the numerator in Equation C.17 must be made. Taking the difference of *l*/*p* and *m*/*p* in Equation C.7 and applying the hyperbolic identity,

$$\cosh(A + B) - \cosh(A - B) = 2 \sinh A \sinh B, \tag{C.18}$$

yields

$$\frac{l - m}{p} = \frac{2 \sinh\left(\frac{\gamma}{\eta}\right)\sinh\left(\frac{2r}{\eta}\right)}{\cosh\left(\frac{\gamma}{\eta}\right)}. \tag{C.19}$$

Both $\cosh(\gamma/\eta) > 0$ and $\sinh(2r/\eta) > 0$ (since $r > 0$) and hence the sign of $\sinh(\gamma/\eta)$ is the same as that of $l - m$. It follows that

$$\sinh\left(\frac{\gamma}{\eta}\right) = \frac{l - m}{2(ml - p^2)^{1/2}} \tag{C.20}$$

and γ is given by

$$\gamma = \eta \sinh^{-1}\left[\frac{\left(\frac{l}{p} - \frac{m}{p}\right)}{2\left(\frac{m}{p}\frac{l}{p} - 1\right)^{1/2}}\right]. \tag{C.21}$$

As indicated earlier, the expression for p is given by Equation C.6 as

$$p = 2\lambda\gamma \cosh\left(\frac{\gamma}{\eta}\right)\sinh\left(\frac{r}{\eta}\right). \tag{C.22}$$

$\text{Sinh}(r/\eta)$ can be obtained by using the relationship $\sinh^2 A = \dfrac{\cosh 2A - 1}{2}$ in conjunction with Equation C.12. This gives

$$\sinh^2\left(\frac{r}{\eta}\right) = \left(\frac{\frac{m + l}{2p} - 1}{2}\right) = \frac{m + l - 2p}{4p}. \tag{C.23}$$

$\cosh(\gamma/\eta)$ is known from Equation C.15. Using Equation C.22, Equation C.15, and Equation C.23, we get

$$p = 2\lambda \frac{(m + l - 2p)(m + l + 2p)^{1/2}}{4[p(ml - p^2)]^{1/2}}. \tag{C.24}$$

Consequently,

$$\lambda = \frac{2p\left(\frac{ml}{p^2} - 1\right)^{1/2}}{\left(\frac{m}{p} + \frac{l}{p} - 2\right)\left(\frac{m}{p} + \frac{l}{p} + 2\right)^{1/2}}. \tag{C.25}$$

Finally, consider

$$g_r + g_{-r} = 2\epsilon - \lambda\left[\sinh\left(\frac{\gamma - r}{\eta}\right) + \sinh\left(\frac{\gamma + r}{\eta}\right)\right]. \tag{C.26}$$

Using the standard formula,

$$\sinh(A + B) + \sinh(A - B) = 2 \sinh A \cosh B. \tag{C.27}$$

Equation C.26 can be rewritten as

$$g_r + g_{-r} = 2\epsilon - 2\lambda \sinh\left(\frac{\gamma}{\eta}\right)\cosh\left(\frac{r}{\eta}\right). \tag{C.28}$$

$\text{Cosh}(r/\eta)$ in Equation C.28 is obtained using the relation, $\cosh^2 A = (1 + \cosh(2A))/2$. It follows from Equation C.12 that

$$\cosh^2\left(\frac{r}{\eta}\right) = \frac{1 + \frac{(m + l)}{2p}}{2} = \frac{2p + m + l}{4p}. \tag{C.29}$$

Thus, using the values of $\sinh(\gamma/\eta)$ from Equation C.20, $\cosh(r/\eta)$ from Equation C.29 and λ from Equation C.25, it is seen that

$$\begin{aligned} g_r + g_{-r} &= 2\epsilon - 2\left[\frac{2p[p(ml - p^2)]^{1/2}}{(m + l - 2p)(m + l + 2p)^{1/2}}\right]\left[\frac{l - m}{2(ml - p^2)^{1/2}}\right] \\ &\quad \times \left[\frac{(2p + m + l)^{1/2}}{2p^{1/2}}\right] \\ &= 2\epsilon - \frac{p(l - m)}{m + l - 2p}. \end{aligned} \tag{C.30}$$

Consequently, we get

$$\epsilon = \frac{1}{2}\left[g_r + g_{-r} + \frac{p\left(\frac{l}{p} - \frac{m}{p}\right)}{\left(\frac{m}{p} + \frac{l}{p} - 2\right)} \right]. \tag{C.31}$$

C.2. Johnson S_B Distribution

The transformation of the *Johnson S_B Distribution* is of the form

$$R = \gamma + \eta \ln\left(\frac{G - \epsilon}{\lambda + \epsilon - G}\right) \tag{C.32}$$

where R is the standard normal random variable and ϵ is a location parameter, λ is a scale parameter, and γ and η are the shape parameters of the *S_B Distribution*. Solving Equation C.32 for G, we obtain

$$G = \epsilon + \frac{\lambda}{1 + \exp\dfrac{\gamma - R}{\eta}}. \tag{C.33}$$

As was done in the previous section, we now proceed to find m, l and p in terms of the parameters γ, λ, η, and ϵ. Thus,

$$m = g_{3r} - g_r = \frac{\lambda}{1 + \exp\left(\frac{\gamma - 3r}{\eta}\right)} - \frac{\lambda}{1 + \exp\left(\frac{\gamma - r}{\eta}\right)}. \tag{C.34}$$

Therefore,

$$m = \lambda\left[\frac{\exp\left(\frac{\gamma - r}{\eta}\right) - \exp\left(\frac{\gamma - 3r}{\eta}\right)}{1 + \exp\left(\frac{\gamma - 3r}{\eta}\right) + \exp\left(\frac{\gamma - r}{\eta}\right) + \exp\left(\frac{2\gamma - 4r}{\eta}\right)} \right] \tag{C.35}$$

and consequently,

$$m = \lambda \frac{\exp\left(\frac{\gamma}{\eta}\right)\exp\left(\frac{-2r}{\eta}\right)\left[\exp\left(\frac{r}{\eta}\right) - \exp\left(\frac{-r}{\eta}\right)\right]}{1 + \exp\left(\frac{\gamma - 3r}{\eta}\right) + \exp\left(\frac{\gamma - r}{\eta}\right) + \exp\left(\frac{2\gamma - 4r}{\eta}\right)}. \tag{C.36}$$

This implies

$$m=\lambda\frac{2\exp\left(\frac{\gamma-2r}{\eta}\right)\sinh\left(\frac{r}{\eta}\right)}{\exp\left(\frac{\gamma-2r}{\eta}\right)\left[\exp\left(-\frac{\gamma-2r}{\eta}\right)+\exp\left(\frac{-r}{\eta}\right)+\exp\left(\frac{r}{\eta}\right)+\exp\left(\frac{\gamma-2r}{\eta}\right)\right]}. \quad \text{(C.37)}$$

Finally, we get

$$m=\lambda\frac{\sinh\left(\frac{r}{\eta}\right)}{\cosh\left(\frac{\gamma-2r}{\eta}\right)+\cosh\left(\frac{r}{\eta}\right)}. \quad \text{(C.38)}$$

Proceeding in a similar fashion, we find

$$l=g_{-r}-g_{-3r}=\lambda\frac{\sinh\left(\frac{r}{\eta}\right)}{\cosh\left(\frac{r}{\eta}\right)+\cosh\left(\frac{\gamma+2r}{\eta}\right)} \quad \text{(C.39)}$$

and

$$p=g_{r}-g_{-r}=\lambda\frac{\sinh\left(\frac{r}{\eta}\right)}{\cosh\left(\frac{r}{\eta}\right)+\cosh\left(\frac{\gamma}{\eta}\right)}. \quad \text{(C.40)}$$

In summary, the values of m, l and p are

$$\begin{aligned} m&=\frac{\lambda\sinh\left(\frac{r}{\eta}\right)}{\cosh\left(\frac{r}{\eta}\right)+\cosh\left(\frac{\gamma-2r}{\eta}\right)}\\ l&=\frac{\lambda\sinh\left(\frac{r}{\eta}\right)}{\cosh\left(\frac{r}{\eta}\right)+\cosh\left(\frac{\gamma+2r}{\eta}\right)}\\ p&=\frac{\lambda\sinh\left(\frac{r}{\eta}\right)}{\cosh\left(\frac{r}{\eta}\right)+\cosh\left(\frac{\gamma}{\eta}\right)}. \end{aligned} \quad \text{(C.41)}$$

Hence,

$$\frac{p}{m} = \frac{\cosh\left(\frac{r}{\eta}\right) + \cosh\left(\frac{\gamma - 2r}{\eta}\right)}{\cosh\left(\frac{r}{\eta}\right) + \cosh\left(\frac{\gamma}{\eta}\right)}$$
$$\frac{p}{l} = \frac{\cosh\left(\frac{r}{\eta}\right) + \cosh\left(\frac{\gamma + 2r}{\eta}\right)}{\cosh\left(\frac{r}{\eta}\right) + \cosh\left(\frac{\gamma}{\eta}\right)}, \tag{C.42}$$

which gives

$$\frac{p}{m}\frac{p}{l} = \frac{A}{B} \tag{C.43}$$

where

$$A = \cosh^2\left(\frac{r}{\eta}\right) + \cosh\left(\frac{\gamma - 2r}{\eta}\right)\cosh\left(\frac{\gamma + 2r}{\eta}\right) + \cosh\left(\frac{r}{\eta}\right)\cosh\left(\frac{\gamma - 2r}{\eta}\right) + \cosh\left(\frac{r}{\eta}\right)\cosh\left(\frac{\gamma + 2r}{\eta}\right) \tag{C.44}$$

and

$$B = \cosh^2\left(\frac{r}{\eta}\right) + \cosh^2\left(\frac{\gamma}{\eta}\right) + 2\cosh\left(\frac{\gamma}{\eta}\right)\cosh\left(\frac{r}{\eta}\right). \tag{C.45}$$

Using the identities

$$\begin{aligned} \cosh(A + B)\cosh(A - B) &= \cosh^2(A) + \cosh^2(B) - 1 \\ \cosh(A + B) + \cosh(A - B) &= 2\cosh(A)\cosh(B) \\ \cosh^2(A) - \sinh^2(A) &= 1 \end{aligned} \tag{C.46}$$

and applying them to Equation C.44, we get

$$A = \cosh^2\left(\frac{r}{\eta}\right) + \cosh^2\left(\frac{\gamma}{\eta}\right) + 2\cosh\left(\frac{r}{\eta}\right)\cosh\left(\frac{\gamma}{\eta}\right)\cosh\left(\frac{2r}{\eta}\right) + \sinh^2\left(\frac{2r}{\eta}\right). \tag{C.47}$$

B is smaller than A because the cosh function assumes only positive values and since $\cosh(2r/\eta) > 1$ and $\sinh^2(2r/\eta) > 0$. This gives $(p^2/ml) > 1$ or equivalently, $(ml/p^2) < 1$.

From Equation C.41

$$1+\frac{p}{m}=\frac{2\cosh\left(\frac{r}{\eta}\right)+\cosh\left(\frac{\gamma}{\eta}\right)+\cosh\left(\frac{\gamma-2r}{\eta}\right)}{\cosh\left(\frac{r}{\eta}\right)+\cosh\left(\frac{\gamma}{\eta}\right)}$$
$$1+\frac{p}{l}=\frac{2\cosh\left(\frac{r}{\eta}\right)+\cosh\left(\frac{\gamma}{\eta}\right)+\cosh\left(\frac{\gamma+2r}{\eta}\right)}{\cosh\left(\frac{r}{\eta}\right)+\cosh\left(\frac{\gamma}{\eta}\right)}. \tag{C.48}$$

Multiplying the expressions in Equation C.48 and using the identities in Equation C.46 we get

$$\left(1+\frac{p}{m}\right)\left(1+\frac{p}{l}\right)=\frac{C}{D} \tag{C.49}$$

where

$$C=4\cosh^2\left(\frac{r}{\eta}\right)+2\cosh^2\left(\frac{\gamma}{\eta}\right)\left[1+\cosh\left(\frac{2r}{\eta}\right)\right]$$
$$+4\cosh\left(\frac{\gamma}{\eta}\right)\cosh\left(\frac{r}{\eta}\right)\left[\cosh\left(\frac{2r}{\eta}\right)+1\right]+\cosh^2\left(\frac{2r}{\eta}\right)-1 \tag{C.50}$$

and

$$D=\cosh^2\left(\frac{r}{\eta}\right)+\cosh^2\left(\frac{\gamma}{\eta}\right)+2\cosh\left(\frac{r}{\eta}\right)\cosh\left(\frac{\gamma}{\eta}\right). \tag{C.51}$$

Using the property, $\cosh(2A)+1=2\cosh^2 A$ in Equation C.50, we get

$$C=4\cosh^2\left(\frac{r}{\eta}\right)+4\cosh^2\left(\frac{\gamma}{\eta}\right)\cosh^2\left(\frac{r}{\eta}\right)+8\cosh\left(\frac{\gamma}{\eta}\right)\cosh^3\left(\frac{r}{\eta}\right)$$
$$+4\cosh^4\left(\frac{r}{\eta}\right)-4\cosh^2\left(\frac{r}{\eta}\right). \tag{C.52}$$

thus

$$C=4\cosh^2\left(\frac{r}{\eta}\right)D. \tag{C.53}$$

Finally, from Equation C.48 and Equation C.52 we get

$$\left(1+\frac{p}{m}\right)\left(1+\frac{p}{l}\right)=4\cosh^2\left(\frac{r}{\eta}\right). \tag{C.54}$$

Solution for η results in

$$\eta = \frac{r}{\cosh^{-1}\left[\frac{1}{2}\left\{\left(1+\frac{p}{m}\right)\left(1+\frac{p}{l}\right)\right\}^{1/2}\right]}. \tag{C.55}$$

Now consider the sum of the terms in Equation C.41. Using the identities in Equation C.45 on this sum we get

$$\begin{aligned}\frac{p}{m}+\frac{p}{l} &= \frac{2\cosh\left(\frac{r}{\eta}\right)+\cosh\left(\frac{\gamma+2r}{\eta}\right)+\cosh\left(\frac{\gamma-2r}{\eta}\right)}{\cosh\left(\frac{r}{\eta}\right)+\cosh\left(\frac{\gamma}{\eta}\right)} \\ &= \frac{2\cosh\left(\frac{r}{\eta}\right)+2\cosh\left(\frac{\gamma}{\eta}\right)\cosh\left(\frac{2r}{\eta}\right)}{\cosh\left(\frac{r}{\eta}\right)+\cosh\left(\frac{\gamma}{\eta}\right)}\end{aligned} \tag{C.56}$$

Solving the above equation for $\cosh(\gamma/\eta)$, we get

$$\cosh\left(\frac{\gamma}{\eta}\right) = \frac{\cosh\left(\frac{r}{\eta}\right)\left(\frac{p}{m}+\frac{p}{l}-2\right)}{\left(\frac{p}{m}+\frac{p}{l}+2\right)-4\cosh^2\left(\frac{r}{\eta}\right)} \tag{C.57}$$

where again we have made use of the identities in Equation C.45. From Equation C.53 expressions of $\cosh(r/\eta)$ and $\sinh^2(r/\eta)$ are obtained and substituted in Equation C.56. The result is

$$\begin{aligned}\cosh\left(\frac{\gamma}{\eta}\right) &= \frac{\left(\frac{p}{m}+\frac{p}{l}-2\right)\left[\left(1+\frac{p}{m}\right)\left(1+\frac{p}{l}\right)\right]^{1/2}}{2\left(1+\frac{p}{m}+\frac{p}{l}+\frac{p}{m}\frac{p}{l}-\frac{p}{m}-\frac{p}{l}-2\right)} \\ &= \frac{\left(\frac{p}{m}+\frac{p}{l}-2\right)\left[\left(1+\frac{p}{m}\right)\left(1+\frac{p}{l}\right)\right]^{1/2}}{2\left(\frac{p^2}{ml}-1\right)}\end{aligned} \tag{C.58}$$

Rather than solve this for γ, it is preferable to derive $\sinh(\gamma/\eta)$ because $\sinh^{-1}(\cdot)$ yields the correct sign of γ due to it being single valued. Again we use the identity

in Equation C.45 to get sinh(·) from cosh(·). Thus

$$\sinh\left(\frac{\gamma}{\eta}\right) = \frac{\left[\left(\frac{p}{m} - \frac{p}{l}\right)^2\right]^{1/2}\left[\left(1+\frac{p}{m}\right)\left(1+\frac{p}{l}\right) - 4\right]^{1/2}}{2\left(\frac{p}{m}\frac{p}{l} - 1\right)}. \tag{C.59}$$

To see which root should be taken in the first factor in the numerator, observe from Equation C.41 that

$$\frac{p}{l} - \frac{p}{m} = \frac{2\sinh\left(\frac{\gamma}{\eta}\right)\sinh\left(\frac{2r}{\eta}\right)}{\cosh\left(\frac{r}{\eta}\right) + \cosh\left(\frac{\gamma}{\eta}\right)}, \tag{C.60}$$

where the hyperbolic identity, $\cosh(A+B) - \cosh(A-B) = 2\sinh(A)\sinh(B)$, is used. The denominator on the right of Equation C.60 is always positive. Also, $\sinh(2r/\eta) > 0$ since $r > 0$. Hence, it follows that the sign of $\sinh(\gamma/\eta)$ is the same as that of $(p/l) - (p/m)$. Equation C.58 then becomes

$$\sinh\left(\frac{\gamma}{\eta}\right) = \frac{\left(\frac{p}{l} - \frac{p}{m}\right)\left[\left(1+\frac{p}{m}\right)\left(1+\frac{p}{l}\right) - 4\right]^{1/2}}{2\left(\frac{p}{m}\frac{p}{l} - 1\right)} \tag{C.61}$$

The value of γ is consequently given by

$$\gamma = \eta\sinh^{-1}\left[\frac{\left(\frac{p}{l} - \frac{p}{m}\right)\left\{\left(1+\frac{p}{m}\right)\left(1+\frac{p}{l}\right) - 4\right\}^{1/2}}{2\left(\frac{p}{m}\frac{p}{l} - 1\right)}\right]. \tag{C.62}$$

The parameter λ can be evaluated using the expression for p in Equation C.40,

$$\lambda = \frac{p\left[\cosh\left(\frac{r}{\eta}\right) + \cosh\left(\frac{\gamma}{\eta}\right)\right]}{\sinh\left(\frac{r}{\eta}\right)}. \tag{C.63}$$

Cosh(r/η) and sinh(r/η) are known in terms of p, l and m from Equation C.53 and cosh(γ/η) is obtained from Equation C.58. Thus

$$\lambda = p\frac{\left\{\left[\left(1+\frac{p}{m}\right)\left(1+\frac{p}{l}\right)\right]^{1/2}+\frac{\left[\left(1+\frac{p}{m}\right)\left(1+\frac{p}{l}\right)\right]^{1/2}\left(\frac{p}{m}+\frac{p}{l}-2\right)}{\frac{p^2}{ml}-1}\right\}}{\left[\left(1+\frac{p}{m}\right)\left(1+\frac{p}{l}\right)-4\right]^{1/2}}. \quad \text{(C.64)}$$

Going through the algebra we get

$$\lambda = p\frac{\left[\left(1+\frac{p}{m}\right)\left(1+\frac{p}{l}\right)\right]^{1/2}\left[1+\frac{\frac{p}{m}+\frac{p}{l}-2}{\left(\frac{p^2}{ml}-1\right)}\right]}{\left[\left(1+\frac{p}{m}\right)\left(1+\frac{p}{l}\right)-4\right]^{1/2}} \quad \text{(C.65)}$$

which gives

$$\lambda = p\frac{\left[\left(1+\frac{p}{m}\right)\left(1+\frac{p}{l}\right)\right]^{1/2}\left[\left(1+\frac{p}{m}\right)\left(1+\frac{p}{l}\right)-4\right]^{1/2}}{\left(\frac{p^2}{ml}-1\right)}. \quad \text{(C.66)}$$

Consequently,

$$\lambda = p\frac{\left\{\left[\left(1+\frac{p}{m}\right)\left(1+\frac{p}{l}\right)\right]^2-4\left(1+\frac{p}{m}\right)\left(1+\frac{p}{l}\right)\right\}^{1/2}}{\left(\frac{p^2}{ml}-1\right)}. \quad \text{(C.67)}$$

Finally the formula for λ is obtained as

$$\lambda = p\frac{\left\{\left[\left(1+\frac{p}{m}\right)\left(1+\frac{p}{l}\right)-2\right]^2-4\right\}^{1/2}}{\left(\frac{p^2}{ml}-1\right)}. \quad \text{(C.68)}$$

For ϵ, Equation C.33 is used to determine the sum

$$g_r + g_{-r} = 2\epsilon + \lambda\left[\frac{1}{1+\exp\left(\frac{\gamma - r}{\eta}\right)} + \frac{1}{1+\exp\left(\frac{\gamma + r}{\eta}\right)}\right]. \tag{C.69}$$

This is easily shown to be equivalent to

$$g_r + g_{-r} = 2\epsilon + \lambda\left[1 - \frac{\sinh\left(\frac{\gamma}{\eta}\right)}{\cosh\left(\frac{\gamma}{\eta}\right) + \cosh\left(\frac{r}{\eta}\right)}\right]. \tag{C.70}$$

From Equation C.61, this reduces to

$$g_r + g_{-r} = 2\epsilon + \lambda - \frac{p\sinh\left(\frac{\gamma}{\eta}\right)}{\sinh\left(\frac{r}{\eta}\right)}. \tag{C.71}$$

Substituting the previously determined values of $\sinh(\gamma/\eta)$ and $\sinh(r/\eta)$ in terms of p, l and m from Equation C.60 and Equation C.53 respectively, quickly yields the desired result for ϵ, as

$$\epsilon = \frac{g_r + g_{-r}}{2} - \frac{\lambda}{2} + \frac{p\left(\frac{p}{l} - \frac{p}{m}\right)}{2\left(\frac{p}{m}\frac{p}{l} - 1\right)}. \tag{C.72}$$

C.3. JOHNSON S_L DISTRIBUTION

The *Johnson S_L Distribution* is given by the transformation of the form

$$R = \gamma^* + \eta \ln(G - \epsilon) \tag{C.73}$$

where again R is a standard normal variable, ϵ is a location parameter, and γ and η are the shape parameters. Solving Equation C.73 for G we get

$$G = \epsilon + \exp\left(\frac{R - \gamma^*}{\eta}\right). \tag{C.74}$$

Then

$$m = g_{3r} - g_r = \exp\left(\frac{3r - \gamma^*}{\eta}\right) - \exp\left(\frac{r - \gamma^*}{\eta}\right)$$
$$l = g_{-r} - g_{-3r} = \exp\left(\frac{-r - \gamma^*}{\eta}\right) - \exp\left(\frac{-3r - \gamma^*}{\eta}\right) \quad \text{(C.75)}$$
$$p = g_r - g_{-r} = \exp\left(\frac{r - \gamma^*}{\eta}\right) - \exp\left(\frac{-r - \gamma^*}{\eta}\right).$$

It follows that

$$\frac{m}{p} = \exp\left(\frac{2r}{\eta}\right) \qquad \frac{l}{p} = \exp\left(\frac{-2r}{\eta}\right) \quad \text{(C.76)}$$

and therefore $(ml/p^2) = 1$.

Moreover, the value of η is obtained from Equation C.76 as

$$\eta = \frac{2r}{\ln\left(\frac{m}{p}\right)}. \quad \text{(C.77)}$$

From Equation C.75 and Equation C.76,

$$p = \exp\left(-\frac{\gamma^*}{\eta}\right)\left[\left(\frac{m}{p}\right)^{1/2} - \left(\frac{l}{p}\right)^{1/2}\right] \quad \text{(C.78)}$$

which yields γ^* as

$$\gamma^* = \eta \ln\left[\frac{\frac{m}{p} - 1}{p\left(\frac{m}{p}\right)^{1/2}}\right]. \quad \text{(C.79)}$$

Finally,

$$g_r + g_{-r} = 2\epsilon + \exp\left(-\frac{\gamma^*}{\eta}\right)\left[\exp\left(\frac{r}{\eta}\right) + \exp\left(-\frac{r}{\eta}\right)\right]. \quad \text{(C.80)}$$

Substituting the known expressions for $\exp(-\gamma^*/\eta)$ and $\exp(r/\eta)$ from Equation C.75 and Equation C.76 respectively, we get the desired result for ϵ, as

$$\epsilon = \frac{g_r + g_{-r}}{2} - \frac{p\frac{m}{p} + 1}{2\frac{m}{p} - 1}. \quad \text{(C.81)}$$

APPENDIX D. CONNECTIONS BETWEEN g_a, k_a, P_a (SECTION 4.2)

According to Equation 4.75 of Section 4.2, the Gaussian random variable R is related to the nonGaussian random variable G by the transformation

$$R = \gamma + \eta f_i(G; \lambda, \epsilon) \tag{D.1}$$

where $f_i(g; \lambda,\epsilon)$ are single valued monotonically increasing functions, $i = 1, 2, 3$. Let r_0 and g_0 satisfy Equation D.1. Because of the single-valued monotonically increasing nature of the transformation

$$\Pr(G \leq g_0) = \Pr(R \leq r_0). \tag{D.2}$$

From a relative frequency point of view

$$\Pr(R \leq r_0) \approx \frac{\text{Number of observations less than equal to } r_0}{n} \tag{D.3}$$

where n is the total number of observations.

Now assume that n observations of the random variable G are obtained. Ordering the observations of G from the smallest to the largest, denote the kth ordered observation by g^k. Then k equals the number of observations less than or equal to g^k. Corresponding to the ordered observations of the random variable G are ordered realizations of the random variable R [See Equation 4.45]. Denote the kth ordered realization of R by r^k. Because the transformation in Equation D.1 is single valued and monotonically increasing, it follows that

$$\Pr(R \leq r^k) = \Pr(G \leq g^k) \approx \frac{k}{n} \tag{D.4}$$

Introducing the "continuity correction", as was done with the q–q and p–p plots in Chapter 1, $\Pr(R \leq r^k)$ is approximated by

$$\Pr(R \leq r^k) \approx \frac{k - \frac{1}{2}}{n}. \tag{D.5}$$

By definition,

$$P_a = \Pr(R \leq a) \tag{D.6}$$

where $a = 3r, r, -r, -3r$. We define the integer k_a such that

$$P_a \approx \frac{k_a - \frac{1}{2}}{n} \tag{D.7}$$

where

$$k_a = \left[nP_a + \frac{1}{2} \right], \tag{D.8}$$

$[\cdot]$ denotes the closest integer and $a = 3r, r, -r, -3r$.

Equation D.6 and Equation D.7 imply that a approximately equals the k_ath ordered sample of R. Note that the k_ath ordered sample of G is g^{k_a}, where r^{k_a} and g^{k_a} satisfy Equation D.1. From Equation D.4 and Equation D.5

$$\Pr(G \leq g^{k_a}) = \Pr(R \leq r^{k_a}) \approx \Pr(R \leq a) \approx \frac{k_a - \frac{1}{2}}{n}. \tag{D.9}$$

It follows, given P_a, that one can determine k_a, and given k_a, one can determine g_a by the simple relation

$$g_a = g^{k_a}. \tag{D.10}$$

APPENDIX E. CANCELLATION FOR AN ANALOG HYBRID CANCELER (SECTION 10.1)

Here we derive the cancellation for an analog canceler which employs bandwidth partitioning and time delays as shown in Figure 10.6. Let the system function $H_k(f)$ for the kth band-pass filter have the characteristic

$$\begin{aligned} H_k(f) &= 1, \quad f_0 + \left(k - \frac{1}{2}\right)B < f < f_0 + \left(k + \frac{1}{2}\right)B \\ &= 0, \quad \text{otherwise} \end{aligned} \tag{E.1}$$

where f_0 is the center frequency and B is the filter bandwidth. Then, in the main channel the signal $y_k(t)$ at the filter output in the kth frequency bin is

$$y_k(t) = \int_{-\infty}^{\infty} H_k(f)Y(f)\mathrm{e}^{\mathrm{j}2\pi ft}\,\mathrm{d}f \tag{E.2}$$

where $Y(f)$ is the Fourier transform of $y(t)$. Also, if $G_k(f)$ is the system function of the section containing the time delays, the output in the kth frequency bin of the auxiliary channel is

$$x_k(t) = \int_{-\infty}^{\infty} \mathrm{d}f\ H_k(f)G_k(f)X(f)\mathrm{e}^{\mathrm{j}2\pi ft}. \tag{E.3}$$

After adding the outputs of all the frequency bins, the output signal $r(t)$ is

$$r(t) = \sum_{k=0}^{N-1} r_k(t) = \sum_{k=0}^{N-1} [y_k(t) - x_k(t)]. \tag{E.4}$$

The average output power is

$$\int_{-\infty}^{\infty} |r(t)|^2 \mathrm{d}t = \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \int_{-\infty}^{\infty} r_k(t) r_l^*(t) \mathrm{d}t. \tag{E.5}$$

The terms with $k \neq l$ can be shown to vanish because

$$\int_{-\infty}^{\infty} \mathrm{d}t\, r_k(t) r_l^*(t) = \int_{-\infty}^{\infty} \mathrm{d}f \int_{-\infty}^{\infty} \mathrm{d}f' B_k(f) B_l^*(f') H_k(f) H_l^*(f') \times \int_{-\infty}^{\infty} \mathrm{d}t \exp[\mathrm{j}2\pi(f - f')t] \tag{E.6}$$

where

$$B_k(f) = Y(f) - G_k(f)X(f). \tag{E.7}$$

If we use the result that

$$\int_{-\infty}^{\infty} \mathrm{d}t \exp[\mathrm{j}2\pi(f - f')t] = \delta(f - f') \tag{E.8}$$

we can rewrite Equation E.6 as

$$\int_{-\infty}^{\infty} \mathrm{d}t\ r_k(t) r_l^*(t) = \int_{-\infty}^{\infty} \mathrm{d}f\ B_k(f) B_l^*(f) H_k(f) H_l^*(f). \tag{E.9}$$

However, because $H_k(f)$ and $H_l(f)$ do not overlap, it is clear that the integral on the right-hand side of Equation E.9 vanishes unless $k = l$. Consequently, Equation E.5 reduces to

$$\int_{-\infty}^{\infty} |r(t)|^2 \mathrm{d}t = \sum_{k=0}^{N-1} \int_{-\infty}^{\infty} |r_k(t)|^2 \mathrm{d}t. \tag{E.10}$$

A similar result applies to the case when the auxiliary channel is absent. If the output in that case is denoted by $r_0(t)$ then the average output is

$$\int_{-\infty}^{\infty} |r_0(t)|^2 \mathrm{d}t = \sum_{k=0}^{N-1} \int_{-\infty}^{\infty} |y_k(t)|^2 \mathrm{d}t. \tag{E.11}$$

Consequently, if we define the cancellation as the average output power with the auxiliary channel to that without the auxiliary we find

$$C = \frac{\sum_{k=0}^{N=1} \langle |r_k|^2 \rangle}{\sum_{k=0}^{N=1} \langle |y_k|^2 \rangle} \tag{E-12}$$

where we have replaced the time average with an ensemble average.

APPENDIX F. CANCELLATION FOR A DIGITAL HYBRID CANCELER (SECTION 10.1)

Let us discuss here how the cancellation is defined for a hybrid canceler that employs both bandwidth partitioning and time delays, as shown in Figure 10.7

when the delay Δ is chosen to be equal to pT_s, where T_s is the inter-sample period. If we define $Y(k)$ as the DFT of the main channel signal y_n and $X(k)$ as the DFT of the auxiliary channel signal x_n then, after weighting and summing, the signal $R(k)$ in the kth frequency bin is

$$R(k) = Y(k) - H(k)X(k) \tag{F.1}$$

where, for a three-tap system,

$$H(k) = w_{k0} + w_{k1}\exp\left(-\mathrm{j}\frac{2\pi kp}{N}\right) + w_{k2}\exp\left(-\mathrm{j}\frac{4\pi kp}{N}\right).$$

The time sequence r_n at the output of the IDFT is

$$r_n = \frac{1}{N}\sum_{k=0}^{N-1} R(k)\exp\left(\mathrm{j}\frac{2\pi nk}{N}\right). \tag{F.2}$$

The average power in the nth output sample is

$$\langle |r_n|^2\rangle = \frac{1}{N^2}\sum_{k=0}^{N-1}\sum_{m=0}^{N-1}\langle R(k)R^*(m)\rangle\exp\left[\mathrm{j}\frac{2\pi n(k-m)}{N}\right]. \tag{F.3}$$

We now consider the average power over the block of N consecutive output samples. This is

$$\langle P\rangle = \frac{1}{N}\sum_{n=0}^{N-1}\langle |r_n|^2\rangle. \tag{F.4}$$

If Equation F.3 is substituted into Equation F.4 and the summation over n is performed, we obtain

$$\langle P\rangle = \frac{1}{N^2}\sum_{k=0}^{N-1}\langle |R(k)|^2\rangle \tag{F.5}$$

where we have used the result that

$$\frac{1}{N}\sum_{n=0}^{N-1}\exp\left[\mathrm{j}\frac{2\pi n(k-m)}{N}\right] = \delta_{km} \tag{F.6}$$

and δ_{km} is the Kronecker delta. If we substitute Equation F.1 into Equation F.5 we then have

$$\langle P\rangle = \frac{1}{N^2}\sum_{k=0}^{N-1}\langle |Y(k) - H(k)X(k)|^2\rangle. \tag{F.7}$$

Next, consider the block-averaged power $\langle P_0 \rangle$ when there is no adaptive processing (auxiliary channel absent). In this case

$$\langle P_0 \rangle = \frac{1}{N^2} \sum_{k=0}^{N-1} \langle |Y(k)|^2 \rangle. \tag{F.8}$$

Consequently, we may define the cancellation C as

$$C = \frac{\langle P \rangle}{\langle P_0 \rangle} = \frac{\sum_{k=0}^{N-1} \langle |Y(k) - H(k)X(k)|^2 \rangle}{\sum_{k=0}^{N-1} \langle |Y(k)|^2 \rangle}. \tag{F.9}$$

APPENDIX G. MATRIX ELEMENTS IN EQUATION 10.10 (SECTION 10.1)

For the signals in Equation 10.15 and Equation 10.16 it is readily shown that for the kth subband the elements in the matrices $[\mathbf{Z}_k]$ and $[\mathbf{R}_k]$ are, for $n = 1, 2, \ldots, N$, $m = 1, 2, \ldots, M$

$$Z_n = F(\alpha_n) + \sum_{q=1}^{Q} \rho_q^* F(\alpha_n + \tau_q) \exp(\mathrm{i}\phi_{kq}) + \sum_{q=1}^{Q} \rho_q F(\alpha_n - t_q) \exp(-\mathrm{i}\psi_{kq})$$
$$+ \sum_{q=1}^{Q} \sum_{l=1}^{Q} \rho_q \rho_l^* F(\alpha_n - t_q + \tau_l) \times \exp[\mathrm{i}(\phi_{kl} - \psi_{kq})] \tag{G.1}$$

$$R_{nm} = F(\alpha_n - \alpha_m) + \sum_{q=1}^{Q} \rho_q^* F(\alpha_n - \alpha_m + \tau_q) \exp(\mathrm{i}\phi_{kq})$$
$$+ \sum_{q=1}^{Q} \rho_q F(\alpha_n - \alpha_m - \tau_q) \exp(-\mathrm{i}\phi_{kq})$$
$$+ \sum_{q=1}^{Q} \sum_{l=1}^{Q} \rho_q^* \rho_l F(\alpha_n - \alpha_m + \tau_q - \tau_l) \exp[\mathrm{i}(\phi_{kq} - \phi_{kl}) \tag{G.2}$$

where $Y_n = (n-1)\Delta + T_d - D$, $D = (M-1)\Delta/2$, T_d =delay of direct jammer ray between main and auxiliary elements, $\tau_q = T_{mq} + t_q - T_d$, T_{mq} =delay of qth multipath ray between main and auxiliary elements, $\phi_{kq} = \omega_k \tau_q$, $\psi_{kq} = \omega_k t_q$, t_q = delay of qth multipath ray relative to the direct jammer ray at the main element, ω_k is the radian frequency at the center of the kth subband and

$F(x) = \text{sinc}(\pi Bx)$. Also,

$$\langle |y_k|^2 \rangle = F(0) + \sum_{q=1}^{Q} \rho_q F(t_q)\exp(-i\psi_{kq}) + \sum_{q=1}^{Q} \rho_q^* F(t_q)\exp(i\psi_{kq}) + \sum_{q=1}^{Q}\sum_{l=1}^{Q} \rho_q \rho_l^* F(t_q - t_l)\exp[i(\psi_{kl} - \psi_{kq})]. \quad \text{(G.3)}$$

APPENDIX H. ASYMPTOTIC CANCELLATION CURVES (SECTION 10.1)

In this Appendix it has been chosen to present asymptotic cancellation curves for $M = 2$, $M = 3$ in the limit when the multipath is specular. In Figure H.1, the cancellation (averaged over 2000 subbands) for $M = 2$, $B\Delta = 0.25$, and $T_m/t_{\text{spec}} = 0.05$ will be shown. The averages were also performed over 10,000 subbands, and the results were indistinguishable from those averaged over 2000 subbands; therefore a 2000 subband average is sufficient to ensure asymptotic conditions. Figure H.2 shows the relative insensitivity of the average cancellation level to the value chosen for the intertap delay Δ, although it is evident that the results are best if $\Delta = t_{\text{spec}}$. The sensitivity of the cancellation to the value chosen for T_m/t_{spec} was also studied. The results were quite insensitive to the value chosen, as long as $T_m/t_{\text{spec}} \ll 1$.

In Figure H.3 the cancellation, averaged over 2,000 subbands, for $M = 3$, $B\Delta = 0.25$, and $T_m/t_{\text{spec}} = 0.05$ is shown. The sensitivity of these results to the

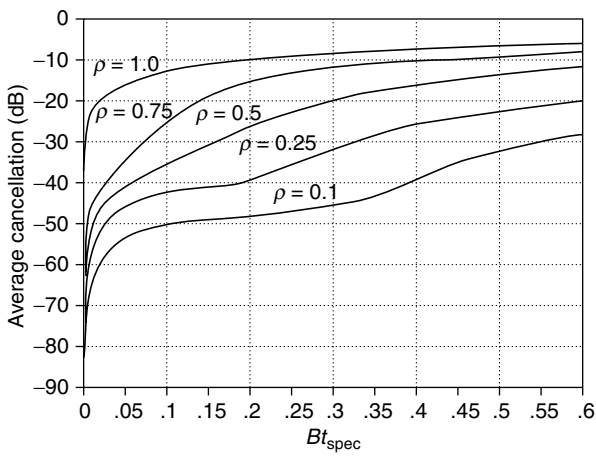

FIGURE H.1 Asymptotic cancellation for specular multipath if $M = 2$ and $B\Delta = 0.25$.

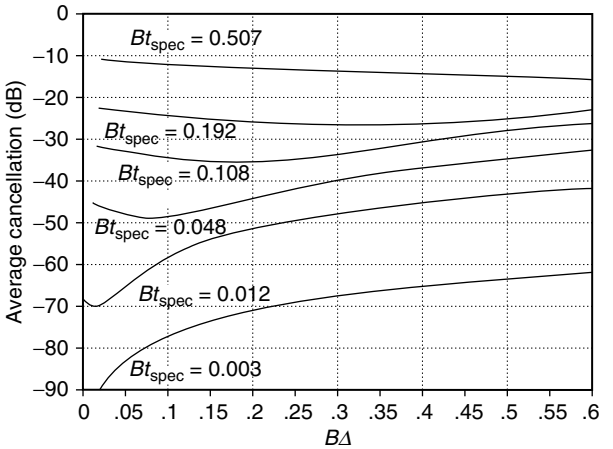

FIGURE H.2 Effect of varying $B\Delta$ on the asymptotic cancellation for specular multipath when $\rho = 0.5$ and $M = 2$.

value chosen for the intertap delay Δ is investigated in Figure H.4. It may be noted that the cancellation is again best when the intertap delay Δ is of the order of the multipath delay t_{spec}.

It is important to note that the conclusions which was reached in Section 10.1.4 are not altered if these asymptotic curves are used. For example, suppose we require 40 dB of cancellation, over a band $B_r = 5$ MHz, of a jammer plus specular multipath with a time delay $t_{spec} = 10^{-6}$ sec. Then for $M = 1$ we find from either Figure 10.4 or Equation 10.8a that $Bt_{spec} = 0.0075$ is required, so that

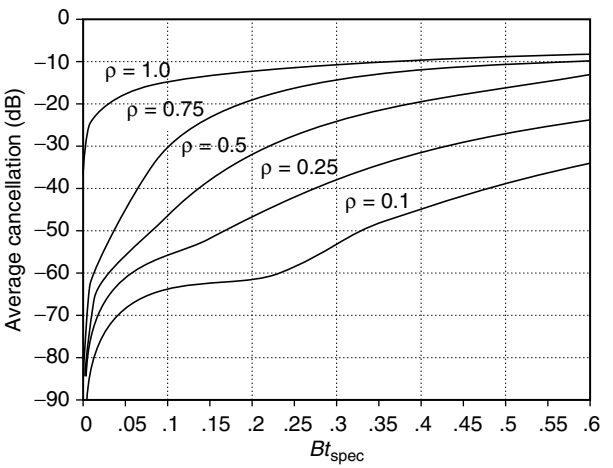

FIGURE H.3 Asymptotic cancellation for specular multipath if $M = 3$ and $B\Delta = 0.25$.

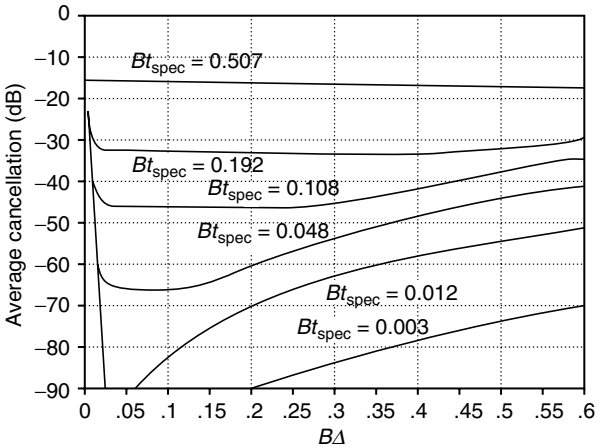

FIGURE H.4 Effect of varying $B\Delta$ on the asymptotic cancellation for specular multipath when $\rho = 0.5$ and $M = 3$.

the subband width is $B = 7.5 \times 10^3$ Hz. For $M = 2$ we find from Figure H.2 that if $\Delta \approx t_{spec}$ (optimum design) then $Bt_{spec} \simeq 0.075$ so that $B \simeq 7.5 \times 10^4$ Hz. Finally, for $M = 3$ it is evident from Figure H.4 that for the optimum choice $\Delta \sim t_{spec}$ we obtain $Bt_{spec} \simeq 0.130$, so that $B = 1.3 \times 10^5$ Hz. Because $B_r = 5$ MHz this implies $N = B_r/B = 667$ subbands are needed for $M = 1$, $N = 67$ subbands are required for $M = 2$ and we need $N = 39$ subbands for $M = 3$.

APPENDIX I. OPTIMUM VALUES OF *N* AND *M* (SECTION 10.1)

One important question that arises is whether there is an optimum combination of M and $B = B_r/N$ (remember B_r is the total bandwidth and N is the number of subbands) that results in the smallest number of computations. In order to discuss this point we need to decide on a processing scheme. One possible approach is to group the samples $X_m(k)$ in each subband (k) in Figure 10.7 into blocks of length S. The S samples in the pth data block of the kth subband would then be used to estimate the covariance matrix

$$[R_k] = \frac{1}{S} \sum_{m=1}^{S} [X_m^*(k)][X_m(k)]^{\mathrm{T}} \tag{I.1}$$

and compute the weight vector $[\mathbf{w}_k] = [R_k]^{-1}[Z_k]$. This weight would then be applied to all samples in the $(p+1)$th block of S samples. Likewise, the S samples in the $(p+1)$th data block would be used to compute the weight applied to all S samples in the $(p+2)$th block. The residue r_k for each sample in the $(p+1)$th data block of the kth subband is calculated from $R_m(k) = Y_m(k) - [\mathbf{w}_k]^{\mathrm{T}}[X_m(k)]$.

We can readily estimate the number of operations necessary to obtain each sample of the residue $R_m(k)$. To form S samples of $[X_m(k)]$ in all the

subbands of the auxiliary and S samples of $Y_m(k)$ in the main channel requires $S(M+1)N\log_2 N$, assuming a radix two FFT is used to perform the subbanding operation. Once the block of S snapshots is obtained, it is shown in [in Ref. 11, Section 5.3.3] that if the "normal-equation method" is used, $M^2(S+M/3)$ operations are required to form the weight vector $[\mathbf{w}_k]$ in each subband. Therefore, the total number of operations necessary to form weight vectors in all N subbands is $NM^2(S+M/3)$. Finally, once $[\mathbf{w}_k]$ is known, we require $2MS$ operations to form the residue r_k in each subband, assuming, as noted earlier, that the same weight is applied to all S snapshots (samples) in the data block. This result must be multiplied by the total number of subbands N.

There is one other set of operations that could be considered, the IDFT shown in Figure 10.7. We do not include this in our calculations, because as mentioned earlier, it may not be necessary. Furthermore, including it in our calculations does not alter the conclusions we reach.

If we add up all the aforementioned operations, we find that the total number of operations needed to form a block of S samples in each of the N subbands is $S(M+1)N\log_2 N + M^2(S+M/3) + 2MNS$. The number of operations per sample is obtained by dividing this result by S, and is

$$\text{Operations/Sample} = N\left[(M+1)\log_2 N + M^2\left(1+\frac{M}{3S}\right)+2M\right]. \qquad \text{(I.2)}$$

Brennan et al.[12] have shown computationally that a good approximation for $[R_k]$ is obtained whenever $S \geq 2M$. Therefore, if we choose the minimum allowable value, $S = 2M$, we obtain from Equation I.2

$$\text{Operations/Sample} = N\left[(M+1)\log_2 N + \frac{7}{6}M^2 + 2M\right]. \qquad \text{(I.3)}$$

Equation I.3 gives the number of operations per output sample, but because bandwidth partitioning has reduced the number of independent samples by a factor of N, all of the output samples are no longer independent. That is, if one samples at the Nyquist rate each input sample is independent, but because bandwidth partitioning reduces the bandwidth by a factor of N, and hence increases the decorrelation time by the same factor, one must choose output samples separated by at least N for independence. Thus, the number of operations per independent output sample is given by multiplying Equation I.3 by N, (this multiplication by N is unnecessary if we use a block transform rather than a sliding window transform) giving

$$\text{Operations/Independent Sample} = N^2\left[(M+1)\log_2 N + \frac{7}{6}M^2 + 2M\right]. \qquad \text{(I.4)}$$

The result in Equation I.4 has been used in conjunction with those in Figure 10.15 to obtain the plots in Figure I.1, which shows the number of operations per

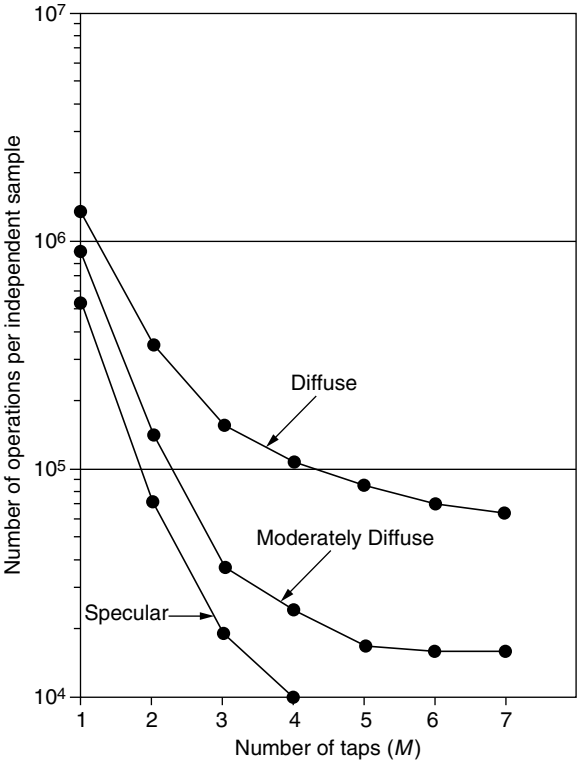

FIGURE I.1 Number of operations per independent sample required to achieve 30 dB of jammer suppression over 2 MHz band when $\rho = 0.5$.

independent output sample when 30 dB cancellation is required. The results clearly indicate that it is preferable to use multiple temporal taps. What particular value of M should be chosen is a compromise, because choosing $M \gg 1$ implies holding the weight constant for $S = 2M$ snapshots, whereas ideally one would like to vary the weight from sample to sample.

APPENDIX J. EFFECT OF NEAR-FIELD NULLING CONSTRAINT (SECTION 10.2)

In this Appendix we demonstrate that enforcing a near-field nulling constraint degrades the adapted signal-to-clutter ratio. This, of course, is to be expected because, in absence of a constraint, the processor automatically places nulls on near-field obstacles in an optimum manner.

In order to demonstrate this point, suppose we forced the processor to place a null on the near-field obstacle. Mathematically, this constraint can be written as

$$[\mathbf{w}]^{\mathrm{T}}[\mathbf{Q}] = 0 \tag{J.1}$$

where $\mathbf{w}$ is the weight vector and $[\mathbf{Q}]$ is a vector given by

$$[\mathbf{Q}]^{\mathrm{T}} = [\alpha_1 \alpha_2 \cdots \alpha_N \; \alpha_1 \cdots \alpha_N \cdots \alpha_1 \cdots \alpha_N], \tag{J.2}$$

$$\alpha_n = \frac{\exp(ikR_n)}{R_n} \tag{J.3}$$

and R_n is the distance from the nth antenna to the near-field obstacle. If there are L pulses the sequence $\alpha_1 \cdots \alpha_N$ in Equation J.2 is repeated L times.

We now minimize the clutter power subject to the constraint in Equation J.1, along with the constraint that the signal be fixed; this latter constraint is specified by $\mathbf{w}^{\mathrm{T}}\mathbf{s} = C$. This requires that we minimize

$$H = [\mathbf{w}]^{+}[\mathbf{M}][\mathbf{w}] - \lambda([\mathbf{w}]^{\mathrm{T}}[\mathbf{s}] - C) - \beta[\mathbf{w}]^{\mathrm{T}}[\mathbf{Q}] \tag{J.4}$$

where $[\mathbf{w}]^{+}$ is the conjugate-transpose of $[\mathbf{w}]$. The optimum weight vector is obtained by differentiating Equation J.4, setting which results equal to zero and then using the resulting solution in the constraint equations to determine the Lagrange multipliers λ and β. The result is

$$[\mathbf{w}] = \lambda^{*}[\mathbf{M}]^{-1}[\mathbf{s}^{*}] + \beta^{*}[\mathbf{M}]^{-1}[\mathbf{Q}^{*}] \tag{J.5}$$

where

$$\begin{aligned}
\lambda &= \frac{\alpha_{22}C^{*}}{\Delta} \\
\beta &= \frac{\alpha_{21}C^{*}}{\Delta} \\
\Delta &= \lambda_{11}\lambda_{22} - \lambda_{12}\lambda_{21} \\
\alpha_{11} &= [\mathbf{s}]^{\mathrm{T}}[\mathbf{M}]^{-1}[\mathbf{s}^{*}] \\
\alpha_{12} &= [\mathbf{Q}]^{\mathrm{T}}[\mathbf{M}]^{-1}[\mathbf{s}^{*}] \\
\alpha_{21} &= [\mathbf{s}]^{\mathrm{T}}[\mathbf{M}]^{-1}[\mathbf{Q}^{*}] \\
\alpha_{22} &= [\mathbf{Q}]^{\mathrm{T}}[\mathbf{M}]^{-1}[\mathbf{Q}^{*}].
\end{aligned}$$

If the weight vector in Equation J.5 is used in the expression

$$\frac{S}{C+N} = \frac{|[\mathbf{w}]^{\mathrm{T}}[\mathbf{s}]|^2}{[\mathbf{w}]^{+}[\mathbf{M}][\mathbf{w}]} \tag{J.6}$$

for the signal-to-noise-plus-clutter ratio, we get

$$\frac{S}{C+N} = \left(\frac{S}{C+N}\right)_0 - \frac{|[\mathbf{Q}]^{\mathrm{T}}[\mathbf{M}]^{-1}[\mathbf{s}^*]|^2}{[\mathbf{Q}]^{\mathrm{T}}[\mathbf{M}]^{-1}[\mathbf{Q}^*]} \tag{J.7}$$

where

$$\left(\frac{S}{C+N}\right)_0 = [\mathbf{s}]^{\mathrm{T}}[\mathbf{M}]^{-1}[\mathbf{s}^*] \tag{J.8}$$

is the signal-to-clutter-plus-noise ratio in the absence of a near-field constraint. Because the second term in Equation J.7 is a positive-definite quantity, it is clear from Equation J.7 that the near-field constraint always produces a loss in signal-to-clutter-plus-noise relative to that which would be obtained if no near-field constraint were employed. Thus, it is best to allow the processor to use whatever degrees of freedom are required to deal with near-field obstacle, rather than attempting to apply a specific constraint.

APPENDIX K. EQUIVALENCE OF ELEMENT SPACE AND BEAM SPACE RESULTS (SECTION 10.4)

In this Appendix we show that, for a fully adaptive array, the SIR after adaptation is the same for both element space and beam space.

In Equation 10.103, we noted that for the element space the SIR after adaptation is

$$\left(\frac{S}{I}\right)_a = s^{\mathrm{T}}\mathbf{\Phi}^{-1}s^*. \tag{K.1}$$

Now suppose a beamforming network is inserted between the antenna elements and the tapped-delay-lines of the processor in Figure 10.52. Then instead of having the element space vector $\mathbf{U}$ as input to the space-time processor, the new beam space input is

$$\mathbf{U}' = \mathbf{B}\mathbf{U} \tag{K.2}$$

where $\mathbf{B}$ is a matrix representing the beamforming operation. This should not be confused with the radar bandwidth. If we assume that the beams are orthogonal and the number of beam is equal to the number of elements, then $\mathbf{B}$ will be a square invertible matrix.

Now let us compute the new covariance matrix $\mathbf{\Phi}'$ in beam space. We have

$$\mathbf{\Phi}' = \langle \mathbf{U}'^*\mathbf{U}'^{\mathrm{T}} \rangle = \mathbf{B}^*\langle \mathbf{U}^*\mathbf{U}^{\mathrm{T}} \rangle \mathbf{B}^{\mathrm{T}} = \mathbf{B}^*\mathbf{\Phi}\mathbf{B}^{\mathrm{T}}. \tag{K.3}$$

Likewise, the steering vector $\mathbf{s}'$ in beam space is

$$\mathbf{s}' = \mathbf{B}\mathbf{s}. \tag{K.4}$$

Consequently, the adapted SIR is

$$\left(\frac{S}{I}\right)_a' = (\mathbf{s}')^{\mathrm{T}}(\mathbf{\Phi}')^{-1}\mathbf{s}'^{*} = \mathbf{s}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}(\mathbf{B}^{*}\mathbf{\Phi}\mathbf{B}^{\mathrm{T}})^{-1}\mathbf{B}^{*}\mathbf{s}^{*}$$

$$= \mathbf{s}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}(\mathbf{B}^{\mathrm{T}})^{-1}\mathbf{\Phi}^{-1}(\mathbf{B}^{*})^{-1}\mathbf{B}^{*}\mathbf{s}^{*} = \mathbf{s}^{\mathrm{T}}\mathbf{\Phi}^{-1}\mathbf{s}^{*} = \left(\frac{S}{I}\right)_a. \qquad \text{(K.5)}$$

Consequently, the SIR after adaptation is the same whether we work in element space or beam space, as long as the number of beam is equal to the number of antenna elements. When there are fewer beams than elements, we may expect some degradation.

APPENDIX L. EVALUATION OF THE INTEGRALS IN EQUATION 10.128 AND EQUATION 10.129 (SECTION 10.4)

Let us consider the first integral on the right-hand side of Equation 10.127. If we make the transformation $\zeta = \tau + t_0 - t - t_1$, the integral in Equation 10.128 becomes

$$I_{kl}(n, m, t_1, t_2, t) = \int_{-\infty}^{\infty} \mathrm{d}\zeta z^{*}(\zeta) z(\alpha_{kl}\zeta + Q) \qquad \text{(L.1)}$$

where $Q = (\alpha_{kl} - 1)(t - t_0) + \gamma + \alpha_{kl}t_1 - t_2$. The integral l' in Equation 10.129 is of the same form, except with $\alpha_{kl} = 1$ and $\gamma = 0$. Now because $z(\zeta) \neq 0$ for $0 \leq \zeta \leq t_s$, and $z(\zeta)$ is equal to zero outside this interval, the integrand is nonzero over only a limited region. Now, we set $t_1 = p_1 T_{\mathrm{p}} + l_1 T$, and $t_2 = p_2 T_{\mathrm{p}} + l_2 T$, where T_{p} is the interpulse period and T is the intertap spacing. It is evident that the right-hand side of Equation L.1 vanishes unless $p_1 = p_2$, because there is no pulse overlap otherwise. Thus, in all that follows, we implicitly assume $p_1 = p_2$.

By considering the various possible parameters we can rewrite Equation L.1 as

$$I_{kl} = \int_{at_s}^{bt_s} \mathrm{d}\zeta z^{*}(\zeta) z(\alpha_{kl}\zeta + Q) \qquad \text{(L.2)}$$

where a and b are given in Table L.1, and the arguments n, m, t_1, t_2, t have been suppressed. If we now substitute Equation 10.117 for $z(\zeta)$ we find

$$I_{kl} = \int_{at_s}^{bt_s} \mathrm{d}\zeta \exp\left\{\mathrm{i}2\pi f_c[(\alpha_{kl} - 1)\zeta + Q] + \mathrm{i}\frac{\pi B}{t_s}[(\alpha_{kl}^2 - 1)\zeta^2 + 2\alpha_{kl}\zeta Q + Q^2]\right\}. \qquad \text{(L.3)}$$

Because the maximum value of b is roughly unity, the term in Equation L.3 that is quadratic in ζ is at most equal to $\pi B(\alpha_{kl}^2 - 1)t_s$. By recalling the definition of α_{kl},

TABLE L.1
Value of *a* and *b*

Condition	*a*	*b*
$Q > 0, 0 < \frac{1}{\alpha_{k\ell}}\left(1 - \frac{Q}{t_s}\right) < 1$	0	$\frac{1}{\alpha_{k\ell}}\left(1 - \frac{Q}{t_s}\right)$
$Q > 0, \frac{1}{\alpha_{k\ell}}\left(1 - \frac{Q}{t_s}\right) > 1$	0	1
$Q < 0, \frac{1}{\alpha_{k\ell}}\left(1 + \left\|\frac{Q}{t_s}\right\|\right) < 1$	$\left\|\frac{Q}{\alpha_{k\ell} t_s}\right\|$	$\frac{1}{\alpha_{k\ell}}\left(1 + \left\|\frac{Q}{t_s}\right\|\right)$
$Q < 0, \frac{1}{\alpha_{k\ell}}\left(1 + \left\|\frac{Q}{t_s}\right\|\right) > 1$	$\left\|\frac{Q}{\alpha_{k\ell} t_s}\right\|$	1

it is evident that $(\alpha_{kl}^2 - 1)$ is at most of order $2v/c$ where v is the magnitude of the net velocity between the platform and the jammer and c is the speed of light. For typical velocities the quantity $v/c \approx 10^{-6}$. Also, for typical high performance radars $Bt_s \approx 10^3$, so that $\pi Bt_s(\alpha_{kl}^2 - 1) \approx 2\pi(10^{-3})(10^{-6}) \sim 2\pi \times 10^{-3}$, and the term in Equation L.3 that is quadratic in ζ can consequently be neglected, so that the resulting integration is readily performed. The result is

$$\begin{aligned} I_{kl} &= (b-a)t_s \exp\{\mathrm{i}[\varphi + \pi(b+a)\Omega t_s]\} \cdot \mathrm{sinc}[\pi(b-a)\Omega t_s], \quad a \le b; \\ &= 0, \quad a > b \end{aligned} \tag{L.4}$$

where

$$\Omega = f_c(\alpha_{kl} - 1) + \frac{\alpha_{kl} BQ}{t_s} \tag{L.5}$$

$$\varphi = 2\pi f_c Q + \frac{\pi B Q^2}{t_s}. \tag{L.6}$$

It is important to note that the integral I_{kl} is a function of time, because Q is a function of time.

Cancellation effectiveness depends on the degree of correlation between the interference from scatterers located in the radar main beam and those located in the side lobes. Therefore, it is now important to substitute for a, b, etc. to determine the conditions under which $|I_{kl}/t_s|$ becomes smaller than unity, because as is evident from Equation 10.130, this will indicate when the interference from the kth and lth scatterers is becoming decorrelated. Decorrelated interference usually results in a reduction in cancellation effectiveness. Let us therefore evaluate Equation L.4 for the case when $n = m$ (same element) and $t_1 = t_2 = 0$. If we also neglect higher order terms in v_A/c and

v_J/c, assume $(v/c)(T_p/t_s) \ll 1, \pi(Bt_s)(T_p/t_s)(v/c) \ll 1$, and ignore an unimportant phase term, we then obtain from Equation L.4

$$\frac{1}{t_s} I_{kl}(n,n,0,0,t) \approx \left(1 - \frac{|\Delta\tau_{lk}|}{t_s}\right)\exp[i2\pi\Delta f_{kl}(t-t_0)]\cdot\mathrm{sinc}\left[\pi\left(1 - \frac{|\Delta\tau_{lk}|}{t_s}\right)(\Delta f_{kl}t_s + B|\Delta\tau_{lk}|)\right], \tag{L.7}$$

where $\Delta\tau_{lk} = \tau_{ln} - \tau_{kn}$ is the differential delay, as measured on the nth antenna element, between the jammer interference scattered by the lth scatterer and that scattered by the kth scatterer. The quantity Δf_{kl} is the Dopper difference between the jammer signal scattered by the kth and lth scatterers, and is defined as

$$\Delta f_{kl} = \frac{f_c}{C}[\mathbf{v}_A\cdot(\hat{\mathbf{R}}''_k - \hat{\mathbf{R}}''_l) - \mathbf{v}_J\cdot(\hat{\mathbf{R}}'_k - \hat{\mathbf{R}}'_l)] \tag{L.8}$$

where $\hat{\mathbf{R}}''_k, \hat{\mathbf{R}}'_k$, etc. are shown in Figure 10.53. From Equation L.7 we note that the conditions necessary for $|I_{kl}/t_s| \approx 1$, and hence good cancellation of the jammer multipath, are:

$$|\Delta\tau_{lk}| \ll t_s \tag{L.9}$$

$$|\Delta f_{kl}|t_s \ll 1 \tag{L.10}$$

$$|B\Delta\tau_{lk}| \ll 1. \tag{L.11}$$

Consequently, the necessary conditions for effective cancellation of the jammer multipath are that (1) the differential delay between the main beam and side lobe scatterers be small in comparison with the uncompressed pulse length, (2) the product of the differential Doppler frequency and the uncompressed pulse length must be small in comparison with unity, and (3) the product of the radar bandwidth and the differential delay between the scatterers must be small in comparison with unity. The term $\exp[i2\pi\Delta f_{kl}(t-t_0)]$ illustrates the nonstationary behavior of the covariance, and its dependence on the Doppler frequency difference.

APPENDIX M. CALCULATION OF THE ADAPTIVE WEIGHTS (SECTION 10.5)

Let us refer to the one-dimensional array, shown in Figure 10.63, and make the following assumptions: (1) all the auxiliary arrays are identical, (2) there are N_J jammers located in the Fraunhofer (or far) zone of the array at azimuths $(\theta_1, \theta_2, \cdots, \theta_{NJ})$, (3) the voltage that would be received from the qth jammer by an isotropic antenna at the location of the main array is $j_q(t)$, and (4) the array is scanned to an angle θ_0. We also define $g_M(\theta,f)$ as the voltage gain of the main array at azimuth θ and frequency f, and assume the temporal Fourier transform

of this quantity if $\tilde{g}_{\mathrm{M}}(\theta,t)$. Likewise, the voltage gain of each auxiliary array is $g_{\mathrm{A}}(\theta,f)$ with a corresponding Fourier transform $\tilde{g}_{\mathrm{A}}(\theta,t)$. Then, referring to Figure 10.63 and using all of the above definitions, we can write the output voltage $U(t)$ in the main array as

$$U(t) = \sum_{q=1}^{N_J} \int_{-\infty}^{\infty} \mathrm{d}\tau \tilde{g}_{\mathrm{M}}(\theta_q,\tau) j_q(t-\tau-T_0) + z(t) \tag{M.1}$$

where $z(t)$ is the main channel noise. Likewise, the voltage received on the kth tap of the nth auxiliary is

$$V_k(n,t) = \sum_{q=1}^{N_J} \int_{-\infty}^{\infty} \mathrm{d}\tau \tilde{g}_{\mathrm{A}}(\theta_q,\tau) j_q(t-\tau-\alpha_{n,k}) + z_n(t-kT) \tag{M.2}$$

where

$$\alpha_{nk} = \frac{L_n}{C}(\sin\,\theta_q - \sin\,\theta_0) - kT \tag{M.3}$$

L_n is the separation between the centers of the main and nth auxiliary arrays, and $z_n(t)$ is the noise in the nth auxiliary. If we refer to Figure 10.63, it is evident that the output voltage y is

$$y = U - \frac{1}{\sqrt{N(K+1)}} \sum_{n=1}^{N} \sum_{k=0}^{K} w_k(n) V_k(n) \tag{M.4}$$

Let us define

$$\tilde{w}_k(n) = \frac{1}{\sqrt{N(K+1)}} w_k(n)$$

along with the vectors

$$\tilde{\mathbf{w}}^{\mathrm{T}} = [\tilde{w}_0(1)\cdots\tilde{w}_K(1)\tilde{w}_0(2)\cdots\tilde{w}_K(2)\cdots]$$
$$\mathbf{v}^{\mathrm{T}} = [V_0(1)\cdots V_K(1)V_0(2)\cdots V_K(2)\cdots].$$

Equation M.4 can be rewritten as

$$y = U - \tilde{\mathbf{w}}^{\mathrm{T}} v. \tag{M.5}$$

It can we shown that the jammer residue y is minimized if the weight vector $\tilde{\mathbf{w}}$ is chosen as

$$\tilde{\mathbf{w}} = \mathbf{\Phi}^{-1}\mathbf{g} \tag{M.6}$$

where $\mathbf{\Phi}$ is an $N(K+1)\times N(K+1)$ matrix given by

$$\mathbf{\Phi} = \langle \mathbf{v}^{*}\mathbf{v}^{\mathrm{T}} \rangle \tag{M.7}$$

$\mathbf{g}$ is an $N(K+1)\times 1$ vector defined as

$$\mathbf{g} = \langle \mathbf{v}^* U \rangle \tag{M.8}$$

and $\langle \cdots \rangle$ denotes an expectation.

We can now use Equation M.1 and Equation M.2 to compute the components of $\mathbf{\Phi}$ and $\mathbf{g}$. Let us write

$$j_q(t) = J_q(t)\exp(\mathrm{i}2\pi f_0 t) \tag{M.9}$$

and assume the voltages from different jammers are statistically independent, zero mean random variables so that

$$\langle J_q(t) J_p^*(t+\zeta)\rangle = R_q(\zeta)\delta_{pq} \tag{M.10}$$

where δ_{pq} is the Kronecker delta and $R_q(\zeta)$ is the autocorrelation function for the qth jammer. Then we obtain

$$\begin{aligned}\langle V_k(n) V_l^*(m)\rangle = \sum_{q=1}^{N_J} \int_{-\infty}^{\infty} \mathrm{d}\tau \int_{-\infty}^{\infty} \mathrm{d}\tau' \ \tilde{\mathbf{g}}_{\mathrm{A}}(\theta_q,\tau)\tilde{\mathbf{g}}_{\mathrm{A}}^*(\theta_q,\tau')\cdot R_q(\tau-\tau'-\beta) \\ \times \exp[-\mathrm{i}2\pi f_0(\tau-\tau'-\beta)] + R_{zz}[(k-\ell)T]\delta_{nm} \\ \times \exp[-\mathrm{i}2\pi f_0(k-\ell)T]\end{aligned} \tag{M.11}$$

where $\beta(n,k,m,l,q) \equiv \alpha_{nk} - \alpha_{ml}$ and $R_{zz}(\tau)$ is the autocorrelation function of the noise. A similar expression is obtained for the components

$$\langle U(t) V_k^*(n,t)\rangle$$

of the vector $\mathbf{g}$, but for brevity, these will be omitted. We can simplify the result in Equation M.11 by recalling that $\mathbf{g}_{\mathrm{A}}(\theta,f)$ by

$$\tilde{\mathbf{g}}_{\mathrm{A}}(\theta,\tau) = \int_{-\infty}^{\infty} \mathrm{d}f \ \mathbf{g}_{\mathrm{A}}(\theta,f) e^{\mathrm{i}2\pi f\tau}. \tag{M.12}$$

Next, assume that all jammers are identical with an auto-correlation function

$$R_q(\tau) = JB\ \mathrm{sinc}(\pi B\tau) \tag{M.13}$$

where J is the power density of the jammer and B is its bandwidth. Then, if Equation M.12 and Equation M.13 are used in Equation M.11 and the integrations performed we obtain

$$\langle V_k(n) V_l^*(m)\rangle = J\sum_{q=1}^{N_J} \int_{f_0-B/2}^{f_0+B/2} \mathrm{d}f\ |\mathbf{g}_{\mathrm{A}}(\theta_q,f)|^2 \cdot \exp(\mathrm{i}2\pi\beta f) + \text{noise}. \tag{M.14}$$

Equation M.14 and a similar expression for $\langle U V_k^*(n)\rangle$ are the wideband expressions required in Equation M.7 and Equation M.8. Note that β is a function of n, m, k, l, and q, but the argument has been suppressed for conciseness.

APPENDIX N. ELIMINATION OF FALSE TARGETS (SECTION 10.5)

As mentioned earlier, false targets may be produced when the received signals due to targets are subjected to the delays in the auxiliary antennas, if those delays are larger than $c/2B$ (i.e., one resolution cell). These spurious targets can be eliminated by using a second bank of matched filters that place the false targets in different range bins than they are placed in by the conventional matched filter in Equation 10.162. Then using AND gates, such as shown in Figure N.1, one can eliminate the false targets while retaining the true targets.

A filter that achieves the aforementioned goal is $H''_{\text{sub}}(f)$ where $H_{\text{sub}}(f)$ has the same denominator as $H_{\text{sub}}(f)$ in Equation 10.162, but instead of $Y_s^*(f)$ in the numerator, we have

$$Y_p^*(f) = CS^*(f)\Gamma''(\theta_0, f) \tag{N.1}$$

where $\Gamma''(\theta_0, f)$ is the same as $\Gamma(\theta_0, f)$ in Equation 10.156, except for the minus sign in front of $\mathbf{g}_{\text{A}}$ replaced by a plus sign and C is a constant.

To understand how this approach works, suppose there is only one auxiliary antenna and only one jammer. Suppose further that the main and auxiliary arrays are identical ($\mathbf{g}_{\text{M}} = \mathbf{g}_{\text{A}}$) and that only one weight w_L in the tapped delay line in Figure 10.150 is nonzero. Then, by taking the inverse transform of $Y_s(f)H_{\text{sub}}(f)$ we find that the temporal response (for a target on boresight) after conventional

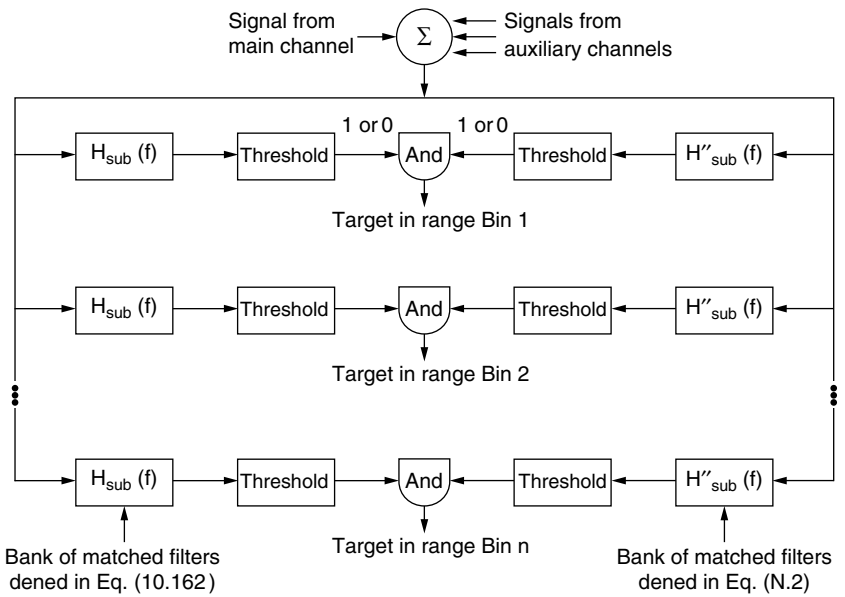

FIGURE N.1 Approach to discriminate true targets from false targets.

matched filtering is

$$y_s(t) = (1 + |w'_L|^2)r(t) - w'_L r(t - \tau_L) - w'^{*}_L r(t + \tau_L) \tag{N.2}$$

where $r(t)$ is the inverse Fourier transform of $|S(f)g_{\mathrm{M}}(\theta_0,f)|^2$, $X(f)$ is the quantity in the denominator of Equation 10.162, and $w'_L = w_L/(k+1)^{1/2}$. Thus, the conventional matched filter produces the true target response $r(t)$ pulse two false targets—one delayed by τ_L and one advanced by τ_L.

Next, consider the output of $H''_{\mathrm{sub}}(f)$. We now find that the response to a target is

$$y''_s(t) = r(t) - (w'_L)^2 r(t - 2\tau_L). \tag{N.3}$$

Therefore, $H''_{\mathrm{sub}}(t)$ again produces the true target response $\tau(t)$, but now there is only a single false target delayed by $2\tau_L$.

If AND logic is used, only the true target survives, and the false targets are all rejected. This approach can be generalized to multiple targets. For example, if there were two targets present then, unless the differential delay happens to be an odd multiple of τ_L, the approach again rejects all the spurious targets while retaining the two true targets. It is also noteworthy that waveforms having sufficient bandwidth to produce spurious targets would normally only be used in track modes. Thus, the number of matched filters in the hardware (software) shown in Figure N.1 should be small.

APPENDIX O. APPROXIMATE DERIVATION OF EQUATION 10.165 (SECTION 10.5)

Assume that the jammer bandwidth is sufficiently small that the temporal taps shown in Figure 10.150 are unnecessary. We also recognize that when N auxiliaries are used, we have the capability of forming up to N cancellation beams. Therefore, we can insert an imaginary beamformer behind the N auxiliary arrays such that the ports b_1 through b_N are now beam ports, and the weights $w'_1, w'_2, \cdots, w'_n$ are now applied in beamspace rather than element space. Because each of the N beams is formed using all N auxiliaries, each beam has a gain $G_b = NG_{\mathrm{A}}$, where G_{A} is the gain per auxiliary. Suppose further that the jammers are spaced in azimuth so that there is only one jammer located within the 3 dB beamwidth of each new auxiliary beam. If there are N_{J} mainbeam jammers present and we ignore any target signals, then the main channel voltage (ignoring constants of proportionality) is

$$v_M = G_{\mathrm{M}}^{1/2} \sum_{p=1}^{N_J} j_p(t) + n_{\mathrm{M}}(t) \tag{O.1}$$

where G_M is main-array gain, $j_p(t)$ is the voltage of the pth jammer, and $n_M(t)$ is the main-channel noise. The auxiliary voltage v_A is

$$v_A = G_b^{1/2} \sum_{p=1}^{N} w_p' j_p(t) + \sum_{p=1}^{N} w_p' n_p(t) \tag{O.2}$$

where $G_b = NG_A, n_p(t)$ is the noise voltage at the pth beam port, we have ignored the delay in $j_p(t)$ between the main and auxiliary channels because the delay bandwidth product has been assumed to be much less than unity, and we have assumed that $N_J \leq N$. The residue $y = v_M - v_A$ is given by

$$y = G_M^{1/2} \sum_{p=1}^{N_J} j_p - G_b^{1/2} + \sum_{p=1}^{N} w_p' j_p + n_M - \sum_{p=1}^{N} w_p' n_p. \tag{O.3}$$

Because dispersive effects are compensated by the tapped delay line, we expect that the jammer residue will be nearly completely cancelled and the remaining residue is just the noise. For $N > N_J$, this has been verified by computation of the two terms in Equation 10.158. Thus, jammer portion of the residue in Equation O.3 vanishes if the weights are chosen so that $w_p' = (G_M/G_b)^{1/2}$ for $p = 1, 2, \ldots, N_J$, and $w_p' = 0$ for $p = N_J + 1, \ldots, N$. Then, if we define the $N \times 1$ weight and noise vectors

$$w' = \left(\frac{G_M}{G_b}\right)^{1/2} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} \left.\vphantom{\begin{matrix}1\\1\\\vdots\\1\end{matrix}}\right\} N_J \\ \left.\vphantom{\begin{matrix}0\\\vdots\\0\end{matrix}}\right\} N - N_J \end{matrix}$$

$$n = \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_N \end{bmatrix} \tag{O.4}$$

the residue after adaptation in Equation O.3 can be written as

$$y = n_M - w'^T n. \tag{O.5}$$

The noise vector n at the beam ports is related to the noise vector n' at the input to the beamformer by the transformation

$$n = Qn' \tag{O.6}$$

where: $\mathbf{Q}$ is the $N \times N$ beamforming matrix. One way to form this transformation matrix is to use the eigenvectors of the covariance matrix $\mathbf{\Phi}$; the latter approach will place a cancellation beam on each well-resolved jammer.[2] Because $\mathbf{\Phi}$ is Hermitian, its eigenvectors are orthonormal so that

$$\mathbf{QQ}^{+} = I_N. \tag{O.7}$$

If Equation O.6 is used in Equation O.5, and we take the expected value of the magnitude squared of the residue voltage, we find

$$\langle |y|^2 \rangle = \langle |n_{\mathrm{M}}|^2 \rangle + w^T Q \langle n' n'^{+} \rangle Q^{+} w^{*} \tag{O.8}$$

where $\mathbf{Q}^{+}$ is the conjugate transpose of $\mathbf{Q}$, $\langle\rangle$ denotes an expectation and we have used the result that the main channel and auxiliary noises are independent. Because the noises in each of the auxiliary arrays are independent, we may write

$$\langle n' n'^{+} \rangle = \sigma^2 I_N \tag{O.9}$$

where I_N is the $N \times N$ identity matrix. Also, $\langle |n_{\mathrm{M}}|^2 \rangle = \sigma^2$ so that Equation O.8 becomes

$$\langle |y|^2 \rangle = \sigma^2 [1 + \mathbf{w}^{\mathrm{T}} \mathbf{QQ}^{+} \mathbf{w}^{*}]. \tag{O.10}$$

If Equation O.7 is used we find

$$\langle |y|^2 \rangle = \sigma^2 [1 + \mathbf{w}^{\mathrm{T}} \mathbf{w}^{*}]. \tag{O.11}$$

Finally, if Equation O.4 is used in Equation O.11, we obtain

$$\langle |y|^2 \rangle = \sigma^2 \left[1 + \frac{N_J G_{\mathrm{M}}}{G_b} \right] = \sigma^2 \left[1 + \frac{N_J G_{\mathrm{M}}}{N G_A} \right] \tag{O.12}$$

where G_{A} is the gain of each auxiliary.

In reality, Equation O.12 is valid only for the artificial geometry where the jammers are well-resolved in angle so that the weight of each cancellation beam can be independently estimated. To extend our result to the general case when the jammer separations are arbitrary, we have postulated that the residue after cancellation will still depend on $N_J G_{\mathrm{M}}/N G_{\mathrm{A}}$, but that this quantity will be multiplied by a factor β that depends only on N and N_J, but not on G_{M}, G_{A}. We have obtained this factor β by using 25 Monte Carlos over jammer locations for each of the following cases:

$N = 1$ to 8 and $N_J = 1$ to N. For all cases, except $N = 2$, which is a pathological case because of grating lobes, we found that the coefficient β, given in Equation 10.166, gave a residue after cancellation that was within 1 dB of the simulated results for $0.3 \leq N_J G_{\mathrm{M}}/N G_{\mathrm{A}} \leq 50$.

APPENDIX P. INTERFERENCE COVARIANCE MATRIX (SECTION 10.10)

In this Appendix we calculate the interference covariance matrix. Consider an array of K antennas mounted on an arbitrary platform and suppose the position of antenna k is (x_k, y_k, z_k), and that multipath is represented by L point scatterers with scatterer p having a bistatic scattering function[†] given by $S_p(\theta_s, \phi_s; \theta_i, \phi_i)$, where (θ_i, ϕ_i) is the angular direction of the incident radiation and (θ_s, ϕ_s) is the angular direction of the scattered radiation. This array is illuminated by N far-field interferers such that interferer n produces a voltage $j_n(t)\exp(\mathrm{i}2\pi f_0 t)$ on the reference element of the antenna array in the absence of multipath, where f_0 is the carrier frequency. Then, if both the direct and multipath voltages are included (and polarization effects are temporarily ignored) the voltage produced by the N interferers on antenna element k, after down-conversion (by multiplying by $\exp(-\mathrm{i}2\pi f_0 t)$) is[‡]

$$y_k(t) = \sum_{n=1}^{N} j_n(t - \tau_{nk})\exp(-\mathrm{i}2\pi f_0 \tau_{nk}) + \sum_{n=1}^{N}\sum_{p=1}^{L} b_{kp} S_p(\theta_k, \phi_k; \theta_n, \phi_n) \times j_n(t - \tilde{\tau}_{np} - T_{pk})\exp[-\mathrm{i}2\pi f_0(\tilde{\tau}_{np} + T_{pk})] + u_k(t) \tag{P.1}$$

where τ_{kn} is the direct-path delay at antenna k from interferer n, $\tilde{\tau}_{np}$ is the delay from interferer n to scatterer p, T_{pk} is the delay from scatterer p to antenna k, b_{kp} is a term that contains the range from scatterer p to antenna k and $u_k(t)$ is the receiver noise on antenna k. In the frequency domain we can write Equation P.1 as

$$Y_k(f) = \sum_{n=1}^{N} J_n(f) G_{nk}(f) + U_k(f) \tag{P.2}$$

where Y_k, J_n, and U_k, are the Fourier transforms of y_k, j_n, and u_k, respectively, and

$$G_{nk}(f) = \exp[\mathrm{i}2\pi(f - f_0)\tau_{nk}] + \sum_{p=1}^{L} b_{kp} S_p \exp[-\mathrm{i}2\pi(f - f_0)(\tilde{\tau}_{np} + T_{pk})]. \tag{P.3}$$

[†] The bistatic radar cross section is $|S_p|^2$.

[‡] The result in Equation P.1 ignores mutual coupling between antennas. This can be approximately included by writing

$$y'_k(t) = y_k(t) + \sum_{l \neq k} \rho_{kl} y_l(t - \xi_{kl})$$

where ρ_{kl} is a complex coupling coefficient and ξ_{kl} is the propagation delay from antenna k to antenna l. We have included mutual coupling in our covariance matrix, but our calculations indicate that it has a negligible effect on the results unless $|\rho| > -15$ dB, which represents a highly coupled antenna array. For a detailed treatment of the effect of mutual coupling on adaptive arrays the reader can consult Ref. 21.

Because the channels in the antenna array are not perfectly equalized, we must account for the different frequency response of each channel across the band of operation. We include this effect by defining $H_k(f)$ as the frequency response of receiver k, so that the interference voltage at the output of receiver k is

$$V_k(f) = H_k(f)Y_k(f). \tag{P.4}$$

Then, the time domain interference on time tap q of antenna k (see Figure 10.94) is

$$\begin{aligned} v_k(t - qT) = \sum_{n=1}^{N} \int_{-\infty}^{\infty} \mathrm{d}f \; H_k(f)J_n(f)G_{nk}(f) \times \exp[\mathrm{i}2\pi f(t - qT)] \\ + \tilde{u}_k(t - qT) \end{aligned} \tag{P.5}$$

where $\tilde{u}_k(t - qT)$ is the noise on time tap q of antenna k.

Let us now derive the interference covariance matrix for the case when the interference consists of N_1 monochromatic (i.e., very narrowband) interferers and N_2 statistically independent, stationary, broadband noise interferers. For a monochromatic noise interferer radiating a frequency f_n we have

$$J_n(f) = J_n\delta(f - f_n) \tag{P.6}$$

and for the independent, wideband noise interferers we have

$$\langle J_n(f)J_m^*(f')\rangle = P_n(f)\delta_{nm}\delta(f - f') \tag{P.7}$$

where $P_n(f)$ is the power spectrum for interferer n, and δ_{nm} is the Kronecker delta. Thus, by using Equation P.6 and Equation P.7 we can show that the general term in the covariance matrix is

$$\begin{aligned} R_{kl}(p - q) &\equiv \langle v_k(t - pT)v_l^*(t - qT)\rangle \\ &= \sum_{n=1}^{N_1} \bar{P}_n H_k(f_n)H_l^*(f_n)G_{nk}(f_n)G_{nl}^*(f_n) \times \exp[\mathrm{i}2\pi f_n(q - p)T] \\ &\quad + \sum_{m=1}^{N_2} \int_{-\infty}^{\infty} df \; P_m(f)H_k(f)H_l^*(f)G_{mk}(f)G_{ml}^*(f) \\ &\quad \times \exp[\mathrm{i}2\pi f(q - p)T] + \delta_{kl}R_N[(p - q)T] \end{aligned} \tag{P.8}$$

where $\bar{P}_n = |J_n|^2$ and

$$R_N[(p - q)T] = \langle \tilde{u}_k(t - pT)\tilde{u}_k^*(t - qT)\rangle \tag{P.9}$$

is the noise covariance.

The response $H_k(f)$ of channel k can be written as

$$H_k(f) = [1 + \alpha_k(f)]\exp[\mathrm{i}\beta_k(f)]. \tag{P.10}$$

The phase fluctuation β_k from channel-to-channel (the mean group delay has been removed from Equation P.10 is usually small in comparison with unity, so that the exponential can be expanded in a Taylor series giving

$$H_k(f) = 1 + \alpha_k(f) + \cdots \tag{P.11}$$

where $a_k = \alpha_k + i\beta_k$. Typically, $a_k(f)$ can be represented as

$$a_k(f) = b_k(f) + c_k \sin(2\pi\gamma_k f + \phi_k) \tag{P.12}$$

where the first term on the right-hand side of Equation P.12 is the differential variation that is linear in frequency and the second term represents a ripple across the passband that also varies from channel-to-channel. In general, the coefficients b_k and c_k may be complex, but γ_k and ϕ_k are real. If Equation P.12 is used in Equation P.8 the integration over frequency is readily performed, but the results are omitted here.

The cancellation ratio is a measure of the mismatch from antenna to antenna. For a white-noise jammer illuminating antennas k and P the cancellation ratio in decibels is defined as $-10\log_{10}(\mathrm{CR})$ where

$$\mathrm{CR} = \frac{\langle |\nu_k(t) - \nu_\ell(t)|^2 \rangle}{P_J B}. \tag{P.13}$$

If Equation P.12 is used in Equation P.5 and the resulting expression substituted into Equation P.13 we find

$$\mathrm{CR} = \mathrm{CR}' + \mathrm{CR}'' \tag{P.14}$$

where CR' is the contribution to the cancellation ratio by the linear fluctuations from channel to channel, and CR'' is the contribution due to the differential ripples across the passband. If the real and imaginary parts of the coefficient b_k are random and uniformly distributed between $\pm b_{\max}$ we find that

$$\mathrm{CR}' = \frac{B^2}{9} b_{\max}^2. \tag{P.15}$$

Also, if ϕ_k is random and uniformly distributed between zero and 2π and the real and imaginary parts of the coefficient c_k and each random and uniformly distributed between $\pm c_{\max}$ we find

$$\mathrm{CR}'' = \frac{2}{3} c_{\max}^2. \tag{P.16}$$

Typically, it is very easy to equalize the variations that are linear in frequency, but it is much more difficult to equalize multiple ripples across the passband. Thus, the value of CR'' is much more important than the value of CR'.

It should be noted that the Doppler effects due to either moving interferers or a moving platform are not included in this analysis. We have shown that Doppler

effects can be ignored if

$$2\pi f_{\mathrm{d}}(P-1)T \ll 1 \tag{P.17}$$

where f_{d} is the Doppler shift (differential speed divided by wavelength) and P is the total number of taps in each delay line. For all of the cases considered in this analysis the condition in Equation P.17 is easily satisfied.

APPENDIX Q. NUMBER OF TIME TAPS REQUIRED (SECTION 10.10)

In order to better understand the condition expressed by Equation 10.275 suppose we have the two antenna configuration shown in Figure Q.1, and this configuration is illuminated by an interferer at an angle θ, so that the delay between the two antennas is $\tau = d \sin \theta / c$. We now desire to examine how well the tapped delay line can track the delay across the frequency band. If the interferer voltage has a Fourier transform $J(f)$ then the output $O(f)$ in Figure Q.1 is

$$O(f) = J(f)[\exp(\mathrm{i}2\pi f\tau) - H(f)] \tag{Q.1}$$

where:

$$H(f) = \sum_{q=0}^{P-1} w_q \exp(\mathrm{i}2\pi q f T). \tag{Q.2}$$

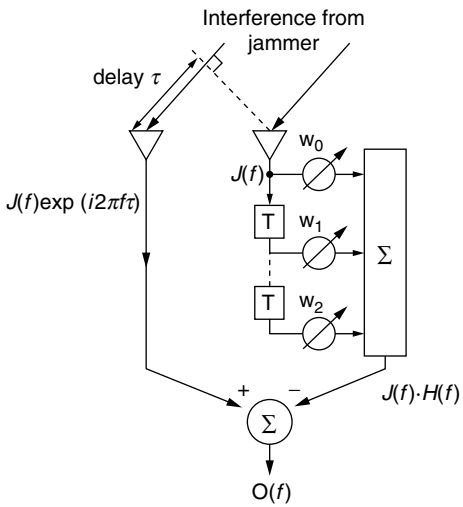

FIGURE Q.1 Two-element processor.

If we use Parseval's Theorem, we can express the mean output power as

$$\varepsilon^2 = \int_{-\infty}^{\infty} \mathrm{d}(f) P_J(f) |\exp(\mathrm{i}2\pi f\tau) - H(f)|^2 \tag{Q.3}$$

where $P_J(f)$ is the interferer power-spectral-density. If we assume the interferer radiates white noise, so that $P_J(f) = P_J$ for $|f| < B/2$ and $P_J = 0$, otherwise, we can write Equation Q.3 as

$$\frac{\varepsilon^2}{P_J B} = 1 - \sum_{q=0}^{P-1} (w_q + w_q^*) \operatorname{sinc} \pi B(\tau - qT) + \sum_{q=0}^{P-1} \sum_{r=0}^{P-1} w_q w_r^* \operatorname{sinc} \pi B(q - r)T. \tag{Q.4}$$

If we minimize ε^2 with respect to w_q we find that the normalized residual power after adaptation is

$$\text{residue} = 1 - h^T \Gamma^{-1} h \tag{Q.5}$$

where h is $P \times 1$ vector with components

$$h_q = \operatorname{sinc} \pi B(\tau - (q-1)T) \tag{Q.6}$$

and Γ is $P \times P$ matrix with components

$$\Gamma_{qr} = \operatorname{sinc} \pi BT(q - r). \tag{Q.7}$$

A plot of the residue power versus the normalized delay $B\tau$ that must be matched is shown in Figure Q.2, for the case when the normalized tap spacing $BT = 0.8$. For

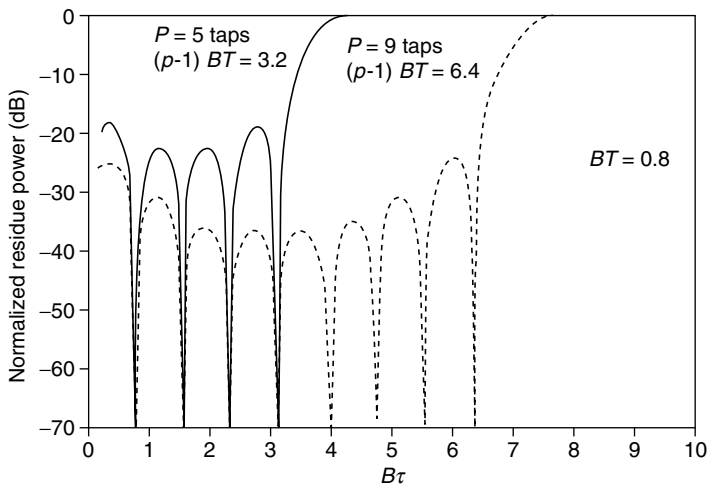

FIGURE Q.2 Normalized residue power for $BT = 0.8$.

the tapped delay line with five taps, the total normalized delay encompassed is $(P-1)BT = 3.2$. Note that as long as $B\tau \leq 3.2$ the normalized interferer residue is small, but for $B\tau > 3.2$ it rises rapidly. When nine taps are used the total delay encompassed is $(P-1)BT = 6.4$, and as long as $B\tau \leq 6.4$ the interferer residue is small, but for $B\tau > 6.4$ the residue rises rapidly, because the delay line is not long enough to match the delay.

APPENDIX R. INCLUSION OF POLARIZATION (SECTION 10.10)

The effects of polarization are readily included in our analysis. In order to do so, let us define a two by one vector J_n as

$$J_n^T = [j_{n\mathrm{H}}(t) \qquad j_{n\mathrm{V}}(t)] \tag{R.1}$$

where $j_{n\mathrm{H}}$ and $j_{n\mathrm{V}}$ are, respectively, the voltages emitted by interferer n on horizontal and vertical polarization. Also define the two by one vector α_k as

$$\alpha_k^T = [\alpha_{\mathrm{H}}(k) \qquad \alpha_{\mathrm{V}}(k)] \tag{R.2}$$

where $\alpha_{\mathrm{H}}(k)$, $\alpha_{\mathrm{V}}(k)$ are the voltages gains of antenna k on horizontal and vertical polarization, respectively. Finally, define the two by one scattering matrix

$$S_P = \begin{bmatrix} S_{\mathrm{HH}}(p) & S_{\mathrm{HV}}(p) \\ S_{\mathrm{VH}}(p) & S_{\mathrm{VV}}(p) \end{bmatrix} \tag{R.3}$$

where s_{HH}, s_{HV}, s_{VH}, s_{VV} are the scattering amplitudes of multipath scatterer p. Then, in place of Equation P.1 we can write the general expression for the interference voltage received on antenna k as

$$y_k(t) = \sum_{n=1}^{N} \alpha_k^T j_n(t-\tau_{nk}) \exp(-\mathrm{i}2\pi f_0 \tau_{nk}) + \sum_{n=1}^{N} \sum_{p=1}^{L} b_{kp} \alpha_k^T S_p J_n(t-\tilde{\tau}_{np} - T_{pk}) \times \exp[-\mathrm{i}2\pi f_0(\tilde{\tau}_{np} + T_{pk})] + u_k(t). \tag{R.4}$$

All of the remainder of the analysis then proceeds exactly as in Appendix P.

It is important that polarization be included in the analysis, because interferers may radiate independent voltages on each of two orthogonal polarizations. In this case, depending on the differences in the cross-polarization response from antenna to antenna, it may like up to twice the number of degrees of freedom to cancel the interferers than it does for interferers that do not use independent signals.

APPENDIX S. SIGNAL CANCELLATION IN FIRST STAGE BEAMFORMER (SECTION 10.11)

In this appendix, we demonstrate that if beam formation is done before pulse compression, there will generally be no signal cancellation as long as SNR in the beam before nulling is less than unity (target return less than receiver noise). In order to demonstrate the point, we have modeled a linear array of M elements separated by one-half wavelength (i.e., each subarray consists of only one element). Suppose we adaptively weight the M elements so as to form a beam steered to θ_0 in the presence of a nonfluctuating signal incident as $\theta_s = \theta_0 + \alpha\theta_3$, where θ_3 is the 3 dB beamwidth. In the absence of interference, the voltage on element n at discrete time i is

$$v_n(i) = s_n + x_n(i) \tag{S.1}$$

where: the signal voltage on the nth element is given by

$$s_n = A \exp[\, \mathrm{j}\pi n(\sin\theta_s + \sin\theta_0)] \tag{S.2}$$

and x_n is the complex, zero mean, unit variance noise voltage on element n. The covariance matrix of the M elements is then given by

$$R_I = I^{-1} \sum_{i=1}^{I} V_i^* V_i^{\mathrm{T}} \tag{S.3}$$

where

$$V_i^{\mathrm{T}} = [v_i(i) v_2(i) \ldots v_M(i)].$$

In order to form a beam steered to θ_0, we apply the weight vector

$$W = \frac{M^{1/2} R_I^{-1} P}{P^H R_I^{-1} P} \tag{S.4}$$

to the received voltage V_i, where P is an $M \times 1$ steering vector of ones. The numerator of Equation S.4 is the whitening filter and the denominator is a scale factor that constrains the gain at the peak of the beam. The signal power in the beam (in the absence of any interference) before nulling is

$$\xi_0 = \frac{|S^{\mathrm{T}} P|^2}{M} \tag{S.5}$$

where

$$S^{\mathrm{T}} = [s_1 s_2 \ldots s_M] \tag{S.6}$$

The signal cancellation ratio is then given by

$$CR_s = \frac{|W^{\mathrm{T}}S|^2}{\xi_0}. \tag{S.7}$$

A MATLAB program was written for a 16 element linear array to study the signal cancellation problem. The beam was steered 20° off normal and the angle of arrival and SNR of the incident signal were varied. The left insert in Figure S.1 plots signal cancellation ratio versus output SNR before nulling, for a case where $4M$ samples ($I = 64$) were used to generate the sample covariance matrix and the target return was present in each sample. The right insert shows the same cases, but the target was present in only two of the 64 training samples.

The left insert shows that if the signal is present in all samples used to form the covariance matrix, no signal cancellation occurs until SNR in the beam before nulling (adaptive weights set to unity) is within 10 dB of the receiver noise level, and even at the signal level, nulling only begins if the signal is near the 3 dB point on the beam. As the signal approaches the peak of the beam (the constraint point), SNR before nulling must get larger before any signal cancellation occurs. The right insert demonstrate that if the signal is not present in all samples, the SNR at which cancellation begins increases, as the number of samples containing the signal decreases.

The issue at hand is whether to perform pulse compression at the subarray level or the beam level. Let us first assume that compression is done at

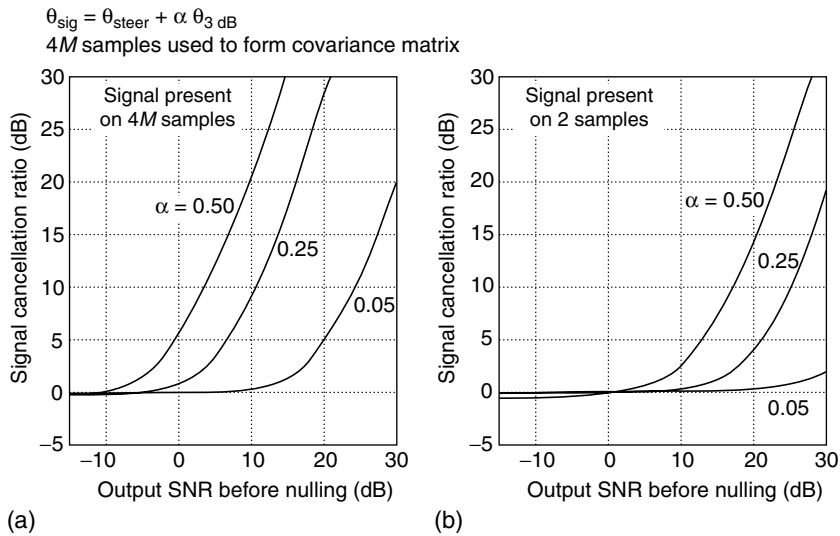

FIGURE S.1 Signal cancellation prior to pulse compression with no diagonal loading (M = number of array elements). (a) Signal present on $4M$ samples. (b) Signal present on two samples.

the subarray level prior to beam formation and adaptive weighting. After compression, the signal will typically straddle two adjacent range cells. The right insert in Figure S.1 shows that significant signal cancellation will occur a large percent of the time since SNR will be the 10 to 20 dB region, in two cells, at the beam level, after compression but prior to nulling. One solution is to recalculate both the first and second stage covariance matrices and their inverses for every range cell in which we attempt a detection; and, omit the detection cell and its nearest neighbors from the training region. The latter requirement would substantially increase the computational complexity. A more reasonable approach appears to be to do compression at the beam level — at least in radars having high pulse compression gain.

If compression is done at the beam level, then in radars having high pulse compression gain, SNR will usually be much less than unity in each of the first-stage beams prior to compression and adaptive nulling, and no signal cancellation will occur. If the signal is so large that it is above noise at the beam level prior to compression and the compression gain is large, the radar could insert attenuation to lower the signal or could just tolerate a little cancellation. If, however, the wave form has low compression gain, even if we compress the signal at the beam level, it may still be above noise prior to compression and we could be in trouble. In the latter case, we have two options: 1) recalculate the covariance matrix for each range cell in which a detection is attempted, omit the detection cell and its nearest neighbors from the training region, and live with the increased computation, or 2) apply diagonal loading to the covariance matrix of subarray voltages.[13] Figure S.2 shows that adding 8 dB of diagonal loading solves

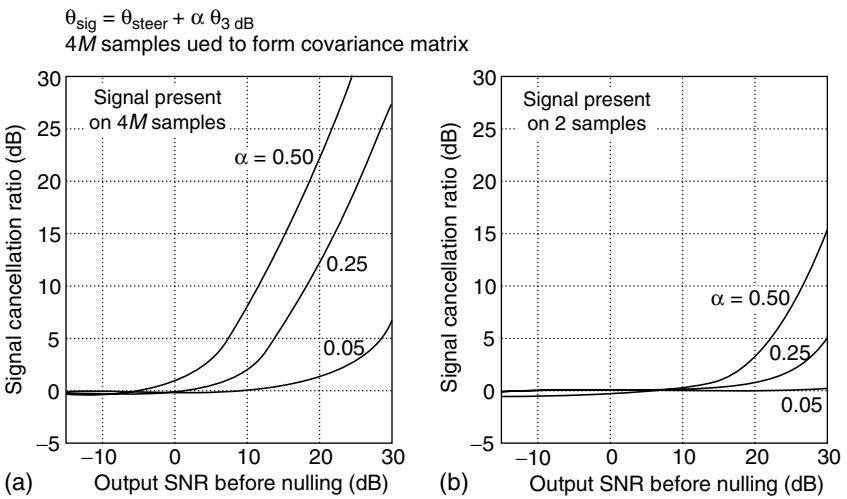

FIGURE S.2 Signal cancellation prior to pulse compression with 8 dB diagonal loading (M = number of array elements). (a) Signal present on $4M$ samples. (b) Signal present on two samples.

the problem. The diagonal loading was accomplished by adding a term $d|z_i|^2$ to the ith diagonal component of the covariance matrix, where d is a scalar and z_i is a complex, zero-mean, unit variance Gaussian variate. Computer simulation predicted that in the presence of sidelobe interference, d could be as large as one-tenth the trace of the covariance matrix prior to loading before seriously affecting the ability of the canceller to null the interference.

APPENDIX T. INTERFERER-FREE LIMIT OF EQUATION 10.298 (SECTION 10.11)

Suppose that there are no interferers present and that each subarray consists of only a single isotropic element, so that $g_m(\theta, \Phi) = 1$. Then, if the noise in each subarray is independent and a target of amplitude A is located at the center of a beam, it is well known that the output signal to noise ratio is $M|A|^2/\sigma^2$ where M is the number of elements and σ^2 is the noise power on each element. Let us show that Equation 10.298 correctly reduces to this limit.

In the absence of interferers, $R = \sigma^2 I$, where I is the $M \times M$ identity matrix. Therefore, Equation 10.291 becomes

$$C = \sigma^2 BB^H. \tag{T.1}$$

If Equation T.1 and Equation 10.299 are used, we can calculate

$$T^H \xi = A + \frac{G^H C^{-1} \rho}{G^H C^{-1} G}. \tag{T.2}$$

Because ρ includes only noise, then if the SNR after pulse compression is large, we can ignore the second term on the right-hand side of Equation T.2. Consequently, if we use Equation T.1 and Equation T.2, along with the defined quantity $G = Bh$ in Equation 10.298, we get

$$\text{SIR} = \frac{|A|^2}{\sigma^2} h^H B^H (BB^H) Bh. \tag{T.3}$$

Next, suppose there is only a single constraint (beam) at (θ_1, ϕ_1) and the target is located at the peak of the beam. Then, because $S_1^H S_1 = M$, we see from Equation 10.282 and Equation 10.283 that $w_1 = S_1/M$ and $B = S_1^H/M$. Also, upon comparing Equation 10.285 and Equation 10.287, it is evident that when $\theta = \theta_1, \phi = \phi_1$, we have $h = S_1$. Therefore, if these results are used in Equation T.3, we find

$$\text{SIR} = \frac{|A|^2}{\sigma^2} \left(\frac{S_1^H S_1}{M} \right) \left(\frac{S_1^H S_1}{M^2} \right)^{-1} \left(\frac{S_1^H S_1}{M} \right) = M \frac{|A|^2}{\sigma^2} \tag{T.4}$$

which is the correct limit.

APPENDIX U. PROPERTIES OF SIRVs (SECTION 12.2)

In this appendix we present some original proofs for properties of SIRPs stated in the literature.

U.1. Statistical Independence

An SSRV $\mathbf{X} = [X_1, X_2, \ldots, X_N]^{\mathrm{T}}$ has statistically independent components X_i $i = 1, 2, \ldots, N$ if and only if the SSRV is Gaussian.

Proof. Recall that the PDF of $\mathbf{X}$ can be expressed as

$$f_{\mathbf{X}}(\mathbf{x}) = k h_N[(x_1^2 + x_2^2 + \cdots + x_N^2)^{1/2}] = (2\pi)^{-N/2} h_N(\sqrt{\mathbf{x}^{\mathrm{T}}\mathbf{x}}). \tag{U.1}$$

If the components of $\mathbf{X}$ are statistically independent, then the PDF given by Equation U.1 must factor into the product of the marginal PDFs of the components of $\mathbf{X}$. It then follows that

$$h_N[(x_1^2 + x_2^2 + \cdots + x_N^2)^{1/2}] = \prod_{i=1}^{N} g(x_i). \tag{U.2}$$

Letting $r = (x_1^2 + x_2^2 + \cdots + x_N^2)^{1/2}$ and differentiating both sides of Equation U.2 with respect to x_i, results in

$$\frac{x_i}{r} h_N'(r) = \prod_{\substack{j=1 \\ j \neq i}}^{N} g(x_j) g'(x_i). \tag{U.3}$$

Dividing both sides of Equation U.3 by $x_i h_N(r)$ results in

$$\frac{h_N'(r)}{r h_N(r)} = \frac{g'(x_i)}{x_i g(x_i)}. \tag{U.4}$$

Equality holds in Equation U.4 if and only if the left and right sides of Equation U.4 are equal to the same constant. Denoting this constant by $-\lambda$, we have

$$\frac{h_N'(r)}{r h_N(r)} = -\lambda. \tag{U.5}$$

Integrating both sides of Equation U.15 with respect to r gives

$$h_N(r) = a \exp\left(-\frac{\lambda r^2}{2}\right) \tag{U.6}$$

where a is the constant of integration. Hence,

$$h_N[(x_1^2 + x_2^2 + \cdots + x_N^2)^{1/2}] = a \exp\left[-\frac{\lambda}{2}(x_1^2 + x_2^2 + \cdots + x_N^2)\right] \tag{U.7}$$

Substitution of Equation U.7 into Equation U.1 clearly results in the Gaussian PDF. The constraint of unity volume under the PDF results in $a = \lambda^{N/2}$.
In order to prove the sufficient part of the property, we start with the PDF of a Gaussian SSRV $\mathbf{X}$ given by

$$f_{\mathbf{X}}(\mathbf{x}) = \left(\frac{2\pi}{\lambda}\right)^{-N/2} \exp\left(-\frac{\lambda}{2}\sum_{i=1}^{N} x_i^2\right). \tag{U.8}$$

Clearly the PDF given by Equation U.8 can be expressed as

$$f_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^{N} f_{X_i}(x_i). \tag{U.9}$$

where

$$f_{X_i}(x_i) = \left(\frac{2\pi}{\lambda}\right)^{-1/2} \exp\left(-\frac{\lambda x_i^2}{2}\right). \tag{U.10}$$

Hence, the sufficient part of the property follows.
An alternate proof of this property can be obtained by using the representation theorem. The representation theorem allows us to express the SSRV $\mathbf{X}$ as a product of a Gaussian random vector $\mathbf{Z}$ having zero mean and identity covariance matrix and a non-negative random variable S. More precisely, we can write

$$\mathbf{X} = \mathbf{Z}S. \tag{U.11}$$

The components of $\mathbf{X}$ can be statistically independent if and only if S is a constant. When S is a constant, $\mathbf{X}$ is a Gaussian SSRV. As is often the case, the representation theorem provides a simplified approach for determining properties of SIRVs.

U.2. Spherically Symmetric Characteristic Function

In this section, we prove that the characteristic function of an SSRV is spherically symmetric.

Proof. We consider the SSRV $\mathbf{X} = [X_1, X_2, \ldots, X_N]^{\mathrm{T}}$. From the representation theorem, we can write $\mathbf{X} = \mathbf{Z}S$ where $\mathbf{Z}$ is a Gaussian random vector having zero mean and identity covariance matrix of S is a non-negative random variable with PDF $f_S(s)$. The characteristic function of $\mathbf{X}$ given by

$$\mathbf{\Phi}_{\mathbf{X}}(\omega) = E[\exp(\mathrm{j}\omega^{\mathrm{T}}\mathbf{X})] \tag{U.12}$$

where $\omega = [\omega_1, \omega_2, \ldots, \omega_N]^{\mathrm{T}}$, can be expressed as

$$\mathbf{\Phi}_{\mathbf{X}}(\omega) = E_S[\mathbf{\Phi}_{\mathbf{X}|S=s}(\omega)] \tag{U.13}$$

where $\Phi_{\mathbf{X}|S=s}(\omega) = E[\exp(\mathrm{j}\omega^{\mathrm{T}}\mathbf{Z}s)]$. However,

$$E[\exp(\mathrm{j}\omega^{\mathrm{T}}\mathbf{Z}s)] = \exp\left(-\frac{s^2}{2}\sum_{i=1}^{N}\omega_i^2\right). \tag{U.14}$$

Using Equation U.14 in Equation U.13 results in

$$\mathbf{\Phi}_{\mathbf{X}}(\omega) = \int_0^\infty \exp\left(-\frac{s^2}{2}\sum_{i=1}^{N}\omega_i^2\right) f_S(s)\mathrm{d}s. \tag{U.15}$$

The characteristic function given by Equation U.15 can be expressed as a function of $\sqrt{\omega^{\mathrm{T}}\omega}$. Hence it is spherically symmetric.

U.3. Relationship between Higher Order and Lower Order SIRV PDFs

In this section we examine the relationship between the higher order and lower order SIRV PDFs. More precisely we consider an SIRV $\mathbf{Y} = [Y_1, Y_2, \ldots, Y_N]^{\mathrm{T}}$ having mean vector μ, covariance matrix $\mathbf{\Sigma}$ and characteristic PDF $f_S(s)$. The PDF of $\mathbf{Y}$ is given by

$$f_{\mathbf{Y}}(y) = (2\pi)^{-N/2}|\mathbf{\Sigma}|^{-1/2}h_N(p) \tag{U.16}$$

where $p = (y-\mu)^{\mathrm{T}}\mathbf{\Sigma}^{-1}(y-\mu)$ and

$$h_N(p) = \int_0^\infty s^{-N}\exp\left(-\frac{p}{2s^2}\right) f_S(s)\,\mathrm{d}s. \tag{U.17}$$

The vector $\mathbf{Y}$ can be partitioned as $\mathbf{Y} = [\mathbf{Y}_1^{\mathrm{T}}\ \mathbf{Y}_2^{\mathrm{T}}]^{\mathrm{T}}$ where $\mathbf{Y}_1 = [Y_1, Y_2, \ldots, Y_m]^{\mathrm{T}}$ and $\mathbf{Y}_2 = [Y_{m+1}, Y_{m+2}, \ldots, Y_N]^{\mathrm{T}}$. Let μ_1 and μ_2 denote the mean vectors of $\mathbf{Y}_1$ and $\mathbf{Y}_2$, respectively, and $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ denote the corresponding covariance matrices. We wish to obtain the PDF of $\mathbf{Y}_1$ from the PDF of $\mathbf{Y}$ by integrating out over the $N-m$ random variables (i.e., the components of $\mathbf{Y}_2$). Let $p_1 = (y_1-\mu_1)^{\mathrm{T}}\mathbf{\Sigma}_1^{-1}(y_1-\mu_1)$ and $p_2 = (y_2-\mu_2)^{\mathrm{T}}\mathbf{\Sigma}_2^{-1}(y_2-\mu_2)$. The PDF of $\mathbf{Y}_1$ is given by

$$f_{\mathbf{Y}_1}(y_1) = (2\pi)^{-N/2}|\mathbf{\Sigma}|^{-1/2}\int_{-\infty}^{\infty}\int_0^\infty s^{-N}\exp\left(-\frac{p}{2s^2}\right) f_S(s)\mathrm{d}s\,\mathrm{d}y_2. \tag{U.18}$$

From Ref. 26 (p. 17 Equation 8, p. 18 Equation 11) we have

$$(2\pi)^{-N/2}|\mathbf{\Sigma}|^{-1/2}\int_0^\infty \exp\left(-\frac{p}{2s^2}\right)\mathrm{d}y_2 = (2\pi)^{-m/2}|\mathbf{\Sigma}_1|^{-1/2}s^{N-m}\exp\left(-\frac{p_1}{2s^2}\right). \tag{U.19}$$

Using Equation U.19 in Equation U.18 gives

$$f_{\mathbf{Y}_1}(y_1) = (2\pi)^{-m/2}|\mathbf{\Sigma}_1|^{-1/2}\int_0^\infty s^{-m}\exp\left(-\frac{p_1}{2s^2}\right) f_S(s)\mathrm{d}s. \tag{U.20}$$

The PDF of $\mathbf{Y}_1$ can be expressed as

$$f_{\mathbf{Y}_1}(y_1)=(2\pi)^{-m/2}|\mathbf{\Sigma}_1|^{-1/2}h_m(p_1) \tag{U.21}$$

where

$$h_m(p_1)=\int_0^\infty s^{-m}\exp\left(-\frac{p_1}{2s^2}\right)f_S(s)\mathrm{d}s. \tag{U.22}$$

Clearly, $h_m(p_1)$ given by Equation U.22 can be obtained from Equation U.17 by simply replacing N by m and p by p_1. To determine the PDF of $\mathbf{Y}_1$, all that is needed is the specification of its mean vector and covariance matrix. As a special case, when $m=1$, Equation U.20 gives us the first order SIRV PDF. Therefore, to obtain the first order SIRV PDF of the ith component of $\mathbf{Y}$ starting from the Nth order SIRV PDF, we simply use Equation U.20 with $m=1$, $\mathbf{\Sigma}_1=\sigma_i$ and $p_1=(y_i-\mu_i)^2/\sigma_i^2$.

APPENDIX V. COMPUTER GENERATION OF SIRVs USING REJECTION METHOD (SECTION 12.4)

V.1. Rejection Method

We present a proof of the rejection procedure[30] used for generating the norm R of the white SIRV $\mathbf{X}$ in Section 12.4. In many instances, it is likely that the PDF of a random variable is known explicitly, but its cumulative distribution function is either unknown or has a complicated functional form. Consequently, the cumulative distribution function cannot be inverted easily. Therefore, the use of the inverse distribution function for generating the random variable does not offer a practical solution for this problem. Hence, it is necessary to use a different scheme for generating the random variable. We consider the problem of generating a sequence of random numbers with PDF $f_R(r)$ of a random variable R, in terms of a random number sequence with PDF $f_{U_1}(u_1)$ of a random variable U_1. The underlying assumption is that the random number sequence from the PDF of U_1 can be readily generated.

The rejection method used in Section 12.4 is based on the relative frequency interpretation of the conditional PDF

$$f_{U_1}(u_1|\mathscr{M})\mathrm{d}u_1=\frac{P\{u_1<U_1\le u_1+\mathrm{d}u_1,\mathscr{M}\}}{P(\mathscr{M})} \tag{V.1}$$

of a random variable U_1 given the event $\mathscr{M}$. $\mathscr{M}$ is expressed in terms of the random variable U_1 and another random variable U_2 and is chosen so that the resulting conditional PDF $f_{U_1}(u_1|\mathscr{M})$ equal $f_R(r)$. The desired sequence is generated by setting $R=U_1$ given that the event $\mathscr{M}$ has occurred and rejecting U_1 otherwise. The problem has a solution only if the domains of r and u_1 are such

that $f_R(r) = 0$ in every interval for which $f_{U_1}(u_1) = 0$. Therefore, we can assume that the ratio $f_{U_1}(u_1)/f_{R_1}(u_1)$ is bounded from below by some positive constant a:

$$\frac{f_{U_1}(u_1)}{f_{R_1}(u_1)} \geq a > 0 \qquad \text{for every } u_1 \tag{V.2}$$

V.2. Rejection Theorem

It is desired to generate a random variable R with PDF $f_R(r)$. Let U_1 be any random variable with PDF $f_{U_1}(u_1)$ such that $f_{U_1}(u_1) = 0$ whenever $f_R(r) = 0$. Let U_2 be a uniformly distributed random variable on the interval (0,1). If the random variables U_1 and U_2 are statistically independent and

$$\mathscr{M} = \{U_2 \leq g(U_1)\} \tag{V.3}$$

where

$$g(u_1) = a\frac{f_{R_1}(u_1)}{f_{U_1}(u_1)} \leq 1, \tag{V.4}$$

then

$$f_{U_1}(u_1|\mathscr{M}) = f_R(u_1). \tag{V.5}$$

Proof. The joint PDF of the random variables U_1 and U_2 can be written as $f_{U_1,U_2}(u_1,u_2) = f_{U_1}(u_1)f_{U_2}(u_2)$, since U_1 and U_2 are statistically independent. Hence, we have

$$P(\mathscr{M}) = \int_{-\infty}^{\infty}\int_{0}^{g(u_1)} f_{U_1}(u_1)f_{U_2}(u_2)\mathrm{d}u_1\,\mathrm{d}u_2. \tag{V.6}$$

However, since U_2 is uniformly distributed in the interval (0,1) and $g(u_1) \leq 1$,

$$\int_{0}^{g(u_1)} f_{U_2}(u_2)\mathrm{d}u_2 = g(u_1). \tag{V.7}$$

Using Equation V.7 in Equation V.6 gives

$$P(\mathscr{M}) = \int_{-\infty}^{\infty} g(u_1)f_{U_1}(u_1)\mathrm{d}u_1. \tag{V.8}$$

However, $g(u_1) = a[f_R(u_1)/f_{U_1}(u_1)]$. Therefore, we have

$$P(\mathscr{M}) = a\int_{-\infty}^{\infty} f_R(u_1)\mathrm{d}u_1 = a. \tag{V.9}$$

We can express the numerator of Equation V.1 as

$$P\{u_1 < U_1 \le u_1 + \mathrm{d}u_1, \mathscr{M}\} = \int_0^{g(u_1)} f_{U_1}(u_1) f_{U_2}(u_2)\mathrm{d}u_1\, \mathrm{d}u_2$$
$$= g(u_1) f_{U_1}(u_1)\mathrm{d}u_1 = a f_R(u_1)\mathrm{d}u_1. \qquad \text{(V.10)}$$

Using Equation V.9 and Equation V.10 in Equation V.1 results in Equation V.5. Thus, we have the following algorithm for generating the sequence of random numbers from the PDF of R.

1. Generate U_1 and U_2.
2. If $U_2 \le a[f_R(u_1)/f_{U_1}(u_1)]$, then $U_1 = R$
3. Otherwise reject U_1.

Referring now to the generation of samples of the norm R in Section 12.4, note that U_1 and U_2 were uniformly distributed random variables. Let c denote the maximum value of the PDF of R and b denote a finite range for the PDF of R such that the area under the curve of the PDF is close to unity. U_1 is assumed to be uniformly distributed in the interval $(0,b)$. Clearly,

$$\frac{f_{U_1}(u_1)}{f_R(u_1)} \ge \frac{1}{bc}.$$

Hence,

$$\frac{f_R(u_1)}{bcf_{U_1}(u_1)} \le 1.$$

Therefore, $a = 1/bc$. Step 2 above becomes: If

$$U_2 \le \frac{f_R(u_1)}{bcf_{U_1}(u_1)} = \frac{f_R(u_1)}{c},$$

then $U_1 = R$. This can be rewritten as: *If* $cU_2 \le f_R(u_1)$, then $U_1 = R$. For ease of implementation, this latter form is used in conjunction with a uniform random variable U_2' that is uniformly distributed over the interval $(0,c)$. This is the procedure followed in Section 12.4.

The method used in Section 12.4 becomes inefficient if U_1 is rejected frequently in step 3, resulting in the necessity to generate the two uniformly distributed random variables of step 1 an inordinate number of times. This problem can be overcome by using for U_1 a PDF which bounds the PDF of R and satisfies the conditions stated in Section V.1 and in the rejection theorem. Then a random variable from this PDF is used in step 1 instead of the uniform random variable U_1.

A second drawback of using a uniformly distributed random variable U_1 is that it may not be possible to efficiently generate SIRVs of length greater than eight. This is due to the fact that the PDF of R depends on N. Consequently, the uniform

distribution for U_1 may not satisfactorily bound the PDF of the norm R for all N. This drawback can be overcome by choosing a different PDF for U_1 for each choice of N, such that the conditions stated in Section V.1 and in the rejection theorem are satisfied. This method would require the use of an exhaustive table which tabulates the appropriate PDF of U_1 for each desired value of N.
Finally, it is pointed out that by using a composite function for the PDF of U_1, it is possible to improve the simulation procedure by making it possible to generate random numbers from the body and the tail of the PDF of R. These issues are suitable topics for future investigation as an extension of this work.

APPENDIX W. MAXIMUM LIKELIHOOD ESTIMATION INVOLVING SIRVs (SECTION 12.5)

The objective of this appendix is to determine maximum likelihood estimates for the mean vector and covariance matrix of an N-dimensional SIRV obtained by sampling a wide sense stationary (WSS) SIRP. Ideally, n independent data vectors $\mathbf{Y}_i,\ i = 1,2,\ldots,n$ should be processed corresponding to n independent trials of the basic experiment. This corresponds to the sample space given by the product

$$\mathscr{S} = \mathscr{S}_1 \times \mathscr{S}_2 \times \cdots \times \mathscr{S}_3 \tag{W.1}$$

where $\mathscr{S}_i,\ i = 1,2,\ldots,n$ denotes the ith ensemble of the SIRP and $\mathbf{Y}_i$ is obtained from $\mathscr{S}_i$. This is shown in Figure W.1.

However, the approach becomes unwieldy from a practical point of view because n ensembles are required. An alternate approach[49] makes use of a single ensemble as shown in Figure W.2, where $\mathbf{Y} = [\mathbf{Y}_1^{\mathrm{T}}, \mathbf{Y}_2^{\mathrm{T}}, \ldots, \mathbf{Y}_n^{\mathrm{T}}]^{\mathrm{T}}$ is obtained by

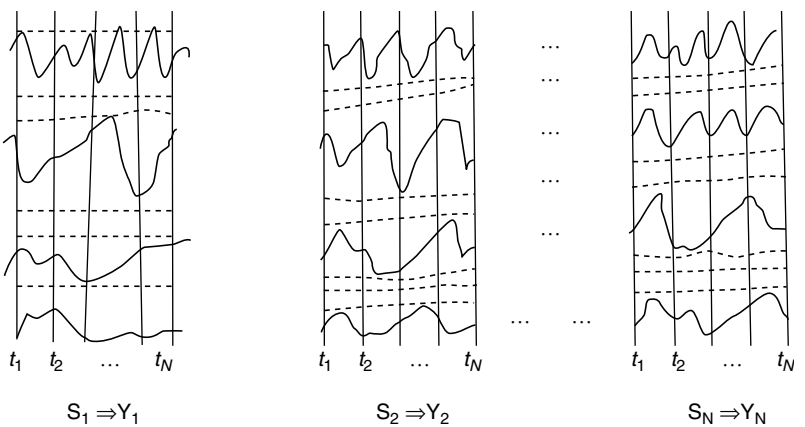

FIGURE W.1 Independent sampling.

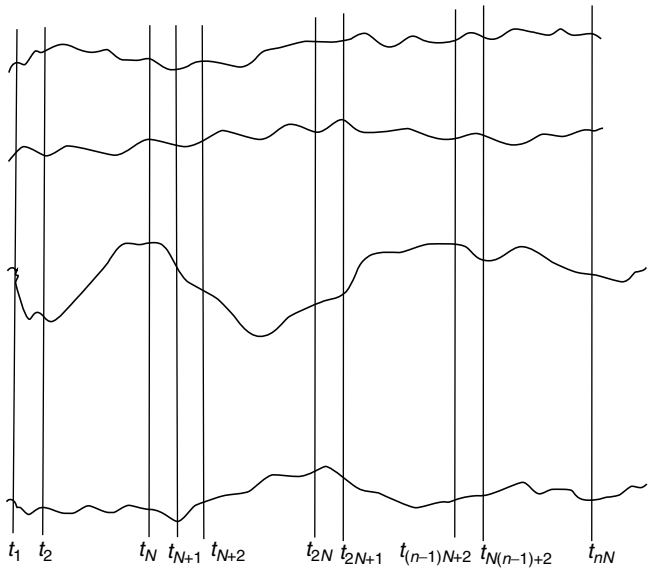

FIGURE W.2 Sampling from a single ensemble.

sampling a WSS SIRP at nN time instants, such that

$$\begin{aligned} &t_{mN+j} - t_{mN+k} = t_j - t_k\, j, \\ &m = 1,2,\ldots,n-1;\ \ j,k = 1,2,\ldots,N \end{aligned} \tag{W.2}$$

and the $\mathbf{Y}_i,\ i = 1,2,\ldots,n,$ are obtained from n different sample functions of the same ensemble. Due to this, the mean of $\mathbf{Y}$ is

$$\alpha = [\mathbf{b}^{\mathrm{T}}, \mathbf{b}^{\mathrm{T}}, \ldots, \mathbf{b}^{\mathrm{T}}]^{\mathrm{T}} \tag{W.3}$$

and

$$E[(\mathbf{Y}_i - \mathbf{b})^{\mathrm{T}}(\mathbf{Y}_i - \mathbf{b})^{\mathrm{T}}] = \Sigma\delta_{jk} \tag{W.4}$$

where δ_{jk} is the Kronecker delta function, so that the covariance matrix of $\mathbf{Y}$ is

$$C = \begin{bmatrix} \Sigma & 0 & \cdots & 0 \\ 0 & \Sigma & \cdots & 0 \\ 0 & 0 & \Sigma & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \Sigma \end{bmatrix}. \tag{W.5}$$

In the context of the radar problem, we consider a surveillance volume that represents a single ensemble for the SIRP. Each cell within the volume generates a sample function of the SIRP. The ith data vector $\mathbf{Y}_i$ is obtained from the ith cell of the volume by sampling at the time instants $t_{(i-1)N+k}$ as shown in Figure W.2.

In terms of the representation theorem, each cell corresponds to a different value of the random variables S whose density function is the characteristic PDF $f_S(s)$.

The PDF of $\mathbf{Y}$ given $\mathbf{b}$ and $\boldsymbol{\Sigma}$ is

$$f_{\mathbf{Y}|\mathbf{b},\boldsymbol{\Sigma}}(\mathbf{y}|\mathbf{b},\boldsymbol{\Sigma}) = (2\pi)^{-nN/2}|\boldsymbol{\Sigma}|^{-n/2}h_{nN}(p) \tag{W.6}$$

where $p = (\mathbf{Y}_i - \alpha)^{\mathrm{T}}C^{-1}(\mathbf{Y}_i - \alpha) = \sum_{j=1}^{n}(y_j - \mathbf{b})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(y_j - \mathbf{b})$ and

$$h_{nN}(p) = \int_0^\infty s^{-nN}\exp\left(-\frac{p}{2s^2}\right)f_S(s)\mathrm{d}s. \tag{W.7}$$

Note that $h_{nN}(\cdot)$ is a monotonically decreasing function for all n and N.

Since p is a scalar, we have

$$p = \sum_{j=1}^{n}(\mathbf{y}_j - b)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{y}_j - b) = \mathrm{tr}\sum_{j=1}^{n}(\mathbf{y}_j - b)^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\mathbf{y}_j - b) \tag{W.8}$$

where $\mathrm{tr}(\cdot)$ denotes the trace of the matrix $(\cdot)$. However, for any two square matrices $\mathbf{A}$ and $\mathbf{B}$,

$$\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA}). \tag{W.9}$$

Consequently,

$$p = \mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{G}) \tag{W.10}$$

where

$$\mathbf{G} = \sum_{j=1}^{n}(\mathbf{y}_j - b)(\mathbf{y}_j - b)^{\mathrm{T}}. \tag{W.11}$$

Thus, we have

$$f_{\mathbf{Y}|\mathbf{b},\boldsymbol{\Sigma}}(\mathbf{y}|\mathbf{b},\boldsymbol{\Sigma}) = (2\pi)^{-nN/2}|\boldsymbol{\Sigma}|^{-n/2}h_{nN}[\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{G})]. \tag{W.12}$$

Let

$$\mathbf{W} = \sum_{j=1}^{n}(\mathbf{y}_j - \bar{\mathbf{y}})(\mathbf{y}_j - \bar{\mathbf{y}})^{\mathrm{T}}. \tag{W.13}$$

where

$$\bar{\mathbf{y}} = \frac{1}{n}\sum_{j=1}^{n}\mathbf{y}_j. \tag{W.14}$$

Then, we can express p as

$$\begin{aligned} p &= \sum_{j=1}^{n} (\mathbf{y}_j - \mathbf{b})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{y}_j - \mathbf{b}) \\ &= \sum_{j=1}^{n} (\mathbf{y}_j - \bar{\mathbf{y}} + \bar{\mathbf{y}} - \mathbf{b})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{y}_j - \bar{\mathbf{y}} + \bar{\mathbf{y}} - \mathbf{b}) \\ &= \sum_{j=1}^{n} (\mathbf{y}_j - \bar{\mathbf{y}})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{Y}_j - \bar{\mathbf{y}}) + \sum_{j=1}^{n} (\mathbf{y}_j - \bar{\mathbf{y}})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{y}}_j - \mathbf{b}) \\ &\quad + \sum_{j=1}^{n} (\bar{\mathbf{y}} - \mathbf{b})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{y}_j - \bar{\mathbf{y}}) + \sum_{j=1}^{n} (\bar{\mathbf{y}} - \mathbf{b})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{y}} - \mathbf{b}) \end{aligned} \tag{W.15}$$

However, from Equation W.14,

$$\begin{aligned} \sum_{j=1}^{n} (\mathbf{y}_j - \bar{\mathbf{y}})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{y}} - \mathbf{b}) = 0 \\ \sum_{j=1}^{n} (\bar{\mathbf{y}} - \mathbf{b})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{y}_j - \bar{\mathbf{y}}) = 0. \end{aligned} \tag{W.16}$$

Therefore,

$$p = \sum_{j=1}^{n} (\mathbf{y}_j - \bar{\mathbf{y}})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\mathbf{y}_j - \bar{\mathbf{y}}) + n(\bar{\mathbf{y}} - \mathbf{b})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{y}} - \mathbf{b}). \tag{W.17}$$

Using Equation W.9, we have

$$p = \mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{W}) + n(\bar{\mathbf{y}} - \mathbf{b})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{y}} - \mathbf{b}). \tag{W.18}$$

Thus, the likelihood function for $\mathbf{b}$ and $\boldsymbol{\Sigma}$ can be expressed as

$$L(\mathbf{b}, \boldsymbol{\Sigma}) = f_{\mathbf{Y}|\mathbf{b},\boldsymbol{\Sigma}}(\mathbf{y}|\mathbf{b}, \boldsymbol{\Sigma}) = (2\pi)^{-nN/2} |\boldsymbol{\Sigma}|^{-n/2} h_{nN}(p) \tag{W.19}$$

where p is given by Equation W.18. We first prove that $\mathbf{W}$ is positive definite with probability one. We can express $\mathbf{W}$ as

$$\mathbf{W} = \sum_{i=1}^{n} \mathbf{W}_i \tag{W.20}$$

where $\mathbf{W}_i = (\mathbf{y}_i - \bar{\mathbf{y}}_i)(\mathbf{y}_i - \bar{\mathbf{y}}_i)^{\mathrm{T}}$. $\mathbf{W}$ is positive definite if $\mathbf{W}_i$, $i = 1, 2, \ldots, n$ are positive definite. We consider a vector $\mathbf{a} = [a_1, a_2, \cdots a_N]^{\mathrm{T}}$ such that $a_i \neq 0$, $i = 1, 2, \ldots, n$. Then $\mathbf{W}_i$ is positive definite if and only if

$$\mathbf{a}^{\mathrm{T}} \mathbf{W}_i \mathbf{a} > 0. \tag{W.21}$$

We have

$$q_i = \mathbf{a}^{\mathrm{T}}\mathbf{W}_i\mathbf{a} = \sum_{j=1}^{N}\sum_{k=1}^{N} a_j a_k (y_{ij} - \overline{y_{ij}})(y_{ik} - \overline{y_{ik}})$$
$$= \left[\sum_{j=1}^{N} a_j (y_{ij} - \overline{y_{ij}})\right]^2 \geq 0. \tag{W.22}$$

However, the probability of $Q_i = 0$ is zero. Therefore, $\mathbf{W}_i$ is positive definite with probability one. It follows that $\mathbf{W}$ is positive definite with probability one.

Before proceeding with the maximum likelihood estimate of $\mathbf{b}$ and $\boldsymbol{\Sigma}$, we present an important lemma[49] which is useful for the maximization problem.

Lemma. *Let $g(\cdot)$ be a monotonically decreasing differentiable function such that $cg(x_1^2 + x_2^2 + \cdots + x_K^2)$ is a PDF of $\mathbf{X} = [X_1, X_2, \ldots, X_K]^{\mathrm{T}}$,* where c is a nonzero constant. Then the function $h(x) = x^{K/2}g(x)$ for $x > 0$ has a maximum at some finite x_0 and is a solution of

$$g'(x) + \frac{K}{2x}g(x) = 0. \tag{W.23}$$

Proof. Since $cg(\cdot)$ is a PDF,

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g\left(\sum_{i=1}^{K} x_i^2\right)\mathrm{d}x_1\cdots\mathrm{d}x_K = \frac{1}{c} < \infty. \tag{W.24}$$

Also, using the transformation to generalized spherical coordinates of Equation 12.22 and integrating over θ and Φ_k, $k = 1, 2, \ldots K-2$, it follows that

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g\left(\sum_{i=1}^{K} x_i^2\right)\mathrm{d}x_1\cdots\mathrm{d}x_K = \frac{\pi^{K/2}}{\Gamma\left(\frac{K}{2}\right)}\int_0^{\infty} 2r^{K-1}g(r^2)\mathrm{d}r. \tag{W.25}$$

Making the change of variable $r^2 = \alpha$, we have

$$\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} g\left(\sum_{i=1}^{K} x_i^2\right)\mathrm{d}x_1\cdots\mathrm{d}x_K = \frac{\pi^{K/2}}{\Gamma\left(\frac{K}{2}\right)}\int_0^{\infty} \alpha^{(K/2)-1}g(\alpha)\mathrm{d}\alpha. \tag{W.26}$$

Since $g(\cdot)$ is a monotonically decreasing function,

$$g(2x)[2x - x] = xg(2x) \leq \int_x^{2x} g(t)\mathrm{d}t. \tag{W.27}$$

Hence,

$$2^{-K/2}(2x)^{K/2}g(2x) = x^{K/2}g(2x) \leq x^{(K/2)-1}\int_x^{2x} g(t)\mathrm{d}t. \tag{W.28}$$

Since $g(\cdot)$ is a monotonically decreasing function,

$$x^{K/2}g(2x) \leq x^{(K/2)-1}\int_x^{2x} g(t)\mathrm{d}t \leq \int_x^{2x} t^{(K/2)-1}g(t)\mathrm{d}t. \tag{W.29}$$

Since $t^{(K/2)-1}g(t)$ is the PDF of $t = \sum_{i=1}^{K} x_i^2$, to within a multiplicative constant, and since the PDF of $t \rightarrow 0$ as $t \rightarrow \infty$ (see Section 12.3 for details), it follows that

$$\int_x^{2x} t^{(K/2)-1}g(t)\mathrm{d}t \rightarrow 0 \qquad \text{as } x \rightarrow \infty. \tag{W.30}$$

Also, $h(0) = 0$ and $h(x) \geq 0, x > 0$. Therefore, $h(x)$ has a maximum at some finite $x_0 > 0$. The first assertion of the lemma follows. Differentiating $h(x)$ with respect to x, we have

$$h'(x) = \frac{K}{2}x^{(K/2)-1}g(x) + x^{K/2}g'(x). \tag{W.31}$$

Since $h(x)$ has a maximum at some finite $x_0 > 0$, it follows that x_0 is a solution to the equation

$$\frac{K}{2x_0}g(x_0) + g'(x_0) = 0. \tag{W.32}$$

Letting $K = nN$ and $x = N/\lambda$, we have $h(N/\lambda) = (N/\lambda)^{nN/2}g(N/\lambda)$. It follows from the above lemma that the function

$$f(\lambda) \equiv \lambda^{-nN/2}g\left(\frac{N}{\lambda}\right) \tag{W.33}$$

arrives at its maximum at some finite positive λ_0 and arises as a solution of

$$\frac{n\lambda_0}{2}g\left(\frac{N}{\lambda}\right) + g'\left(\frac{N}{\lambda}\right) = 0 \tag{W.34}$$

We now return to the problem of maximization of $L(\mathbf{b}, \boldsymbol{\Sigma})$. Since $h_{nN}(\cdot)$ is a monotonically decreasing function and $\boldsymbol{\Sigma}$ is positive definite with probability one, $L(\mathbf{b}, \boldsymbol{\Sigma})$ arrives at its maximum when $\hat{\mathbf{b}} = \bar{\mathbf{y}}$. We then focus on the concentrated likelihood function

$$L(\mathbf{b}, \boldsymbol{\Sigma}) = (2\pi)^{-nN/2}|\boldsymbol{\Sigma}|^{-n/2}h_{nN}[\mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{W})]. \tag{W.35}$$

Since $\mathbf{W}$ is a positive definite matrix, it can be represented as $\mathbf{W} = \mathbf{C}\mathbf{C}^{\mathrm{T}}$ where $\mathbf{C}$ is a nonsingular matrix. We define the matrix $\tilde{\boldsymbol{\Sigma}} = \mathbf{C}^{-1}\boldsymbol{\Sigma}(\mathbf{C}^{\mathrm{T}})^{-1}$, so that $\boldsymbol{\Sigma} = \mathbf{C}\tilde{\boldsymbol{\Sigma}}\mathbf{C}^{\mathrm{T}}$.

Also, it follows that

$$\begin{aligned} \mathrm{tr}(\boldsymbol{\Sigma}^{-1}\mathbf{W}) &= \mathrm{tr}[(\mathbf{C}\tilde{\boldsymbol{\Sigma}}\mathbf{C}^{\mathrm{T}})^{-1}\mathbf{W}] = \mathrm{tr}(\tilde{\boldsymbol{\Sigma}}^{-1}[\mathbf{C}\mathbf{C}^{\mathrm{T}}]^{-1}\mathbf{W}) = \mathrm{tr}(\tilde{\boldsymbol{\Sigma}}^{-1}) \\ |\boldsymbol{\Sigma}| &= |\mathbf{C}\tilde{\boldsymbol{\Sigma}}\mathbf{C}^{\mathrm{T}}| = |\tilde{\boldsymbol{\Sigma}}||\mathbf{W}| \end{aligned} \tag{W.36}$$

Hence, the likelihood function can be rewritten as

$$L(\mathbf{b},\boldsymbol{\Sigma}) = (2\pi)^{-nN/2}|\mathbf{W}|^{-n/2}|\tilde{\boldsymbol{\Sigma}}|^{-n/2}h_{nN}[\mathrm{tr}(\tilde{\boldsymbol{\Sigma}}^{-1})]. \tag{W.37}$$

Let $\lambda_i > 0,\ i = 1,2,\ldots N$ denote the eigenvalues of $\tilde{\boldsymbol{\Sigma}}$. Then,

$$\begin{aligned} L(\mathbf{b},\boldsymbol{\Sigma}) &= (2\pi)^{-nN/2}|\mathbf{W}|^{-n/2}\left(\prod_{i=1}^{N}\lambda_i^{-1}\right)^{-n/2} h_{nN}\left(\textstyle\sum_{i=1}^{N}\lambda_i^{-1}\right) \\ &= (2\pi)^{-nN/2}|\mathbf{W}|^{-n/2}\left(\prod_{i=1}^{N}\lambda_i^{-1/N}\right)^{-Nn/2} h_{nN}\left(\textstyle\sum_{i=1}^{N}\lambda_i^{-1}\right) \end{aligned} \tag{W.38}$$

Since the arithmetic mean is always greater than or equal to the geometric mean, it follows that

$$L(\mathbf{b},\boldsymbol{\Sigma}) \le (2\pi)^{-nN/2}|\mathbf{W}|^{-n/2}(\bar{\lambda})^{nN/2}h_{nN}(N\bar{\lambda}) \tag{W.39}$$

where $\bar{\lambda} = (1/N)\sum_{i=1}^{N}\lambda_i^{-1}$. Equality between the arithmetic and geometric mean holds only if $\lambda_1 = \lambda_2 = \cdots = \lambda_N = \lambda$. Therefore, $L(\mathbf{b},\boldsymbol{\Sigma})$ arrives at its maximum when the eigenvalues of $\tilde{\boldsymbol{\Sigma}}$ are equal. Consequently,

$$\begin{aligned} \max L(\mathbf{b},\boldsymbol{\Sigma}) &= \max(2\pi)^{-nN/2}|\mathbf{W}|^{-n/2}\lambda^{-nN/2}h_{nN}\left(\frac{N}{\lambda}\right) \\ &= \max(2\pi)^{-nN/2}|\lambda\mathbf{W}|^{-n/2}h_{nN}\left(\frac{N}{\lambda}\right) \\ &= \max(2\pi)^{-nN/2}|\mathbf{W}|^{-n/2}f(\lambda). \end{aligned} \tag{W.40}$$

where $f(\lambda)$ is given by Equation W.33 with $g(\cdot)$ replaced by $h_{nN}(\cdot)$. Note that $h_{nN}(\cdot)$ satisfies all the conditions of the lemma dealing with the maximization of $h(\cdot)$. Let the value of λ resulting in the maximum be denoted by λ_0. Comparing Equation W.40 with Equation W.33, it follows that the maximum likelihood estimate of $\boldsymbol{\Sigma}$ is

$$\hat{\boldsymbol{\Sigma}} = \lambda_0\mathbf{W}. \tag{W.41}$$

In summary, the maximum likelihood estimates of $\mathbf{b}$ and $\boldsymbol{\Sigma}$ are:

$$\begin{aligned} \hat{\mathbf{b}} &= \bar{y} \\ \hat{\boldsymbol{\Sigma}} &= \lambda_0\mathbf{W} \end{aligned} \tag{W.42}$$

In order to guarantee the nonnegative definite property of $\hat{\boldsymbol{\Sigma}}$, it is required that $n > N$. It has been pointed out in Refs. 52 and 53 that a rule of thumb for obtaining a reasonably good estimate of $\boldsymbol{\Sigma}$ is that $n \ge 2N - 3$.

APPENDIX X. ISSUES RELATED TO EXTREME VALUE THEORY (SECTION 13.3)

X.1. Limiting Forms for the Largest Order Statistic

Let $X_1 \leq X_2 \leq \cdots \leq X_n$ be the ordered statistics of n random variables having a common distribution function $F(x)$. Assuming that the trials of drawing the random variables from the distribution function $F(x)$ are independent, the distribution function of the largest order statistic X_n is given by

$$\begin{aligned} P(X_n \leq x) &= P(X_1 \leq x, X_2 \leq x, \ldots, X_n \leq x) \\ &= F^n(x). \end{aligned} \tag{X.1}$$

When F is continuous but unknown, an asymptotic theory is developed for F in the range 0_+ to 1_-.[45] It is shown that positive sequences $\{a_n\}$ and $\{b_n\}$ exist such that

$$\lim_{n\to\infty} P\left(\frac{X_n - b_n}{a_n} \leq x\right) = \lim_{n\to\infty} P(X_n \leq a_n x + b_n) \to \Lambda(x) \tag{X.2}$$

or equivalently, by means of Equation X.1, that

$$\lim_{n\to\infty} F^n(a_n x + b_n) \to \Lambda(x). \tag{X.3}$$

Let $n = md$ in Equation X.3. d is a fixed positive constant so that as $n \to \infty$, $m \to \infty$. Using the fact that $n = md$, we can write

$$\lim_{m\to\infty} F^{md}(a_{md} x + b_{md}) = \lim_{n\to\infty} F^n(a_n x + b_n) \to \Lambda(x). \tag{X.4}$$

It is also true that

$$\lim_{m\to\infty} [F^m(a_m x + b_m)]^d = \lim_{m\to\infty} F^{md}(a_m x + b_m) \to \Lambda^d(x). \tag{X.5}$$

If Equation X.4 and Equation X.5 hold, then from a theorem of Hintchin,[52] there exist numbers $A_d > 0$ and $B_d > 0$ such that

$$\Lambda^d(A_d x + B_d) = \Lambda(x) \tag{X.6}$$

for all integer values of d.

Solution of the above functional equation yields all the possible limiting forms for the distribution function $F^n(x)$. The constant A_d may or may not be unity. If it is unity, then the functional equation to be solved is given by

$$\Lambda^d(x + B_d) = \Lambda(x). \tag{X.7}$$

On the other hand, if A_d is not unity, the form of Equation X.6 stands and there exists a value $x_{0d} = B_d/(1 - A_d)$ such that

$$\Lambda^d(x_{0d}) = \Lambda(x_{0d}). \tag{X.8}$$

Constraining the solution to the above equation to be real and nonnegative, the solution is either $\Lambda = 0$ or 1. However, because $\Lambda(x)$ is a distribution function the value of Λ can be zero only if x_{0d} is the lower endpoint at which $\Lambda(x_{0d}) = 0_+$ and Λ can be one only if x_{0d} is the upper end at which $\Lambda(x_{0d}) = 1_-$. Since A_d and B_d are assumed to be finite, x_{0d} must also be finite. Consequently, there is no loss in generality by assuming that the endpoint of interest is located at the origin (i.e., $x_{0d} = 0$). When $A_d \neq 1$, note that $x_{0d} = 0$ implies $B_d = 0$. As a result, the solutions for Equation X.6 fall into three cases which are given below.

$$(1)\ \ \Lambda^d(x + B_d) = \Lambda(x) \quad A_d = 1 \tag{X.9}$$

$$(2)\ \ \Lambda^d(A_d x) = \Lambda(x) \quad A_d \neq 1\ F = 0\ \text{ when } x = 0 \tag{X.10}$$

$$(3)\ \ \Lambda^d(A_d x) = \Lambda(x) \quad A_d \neq 1\ F = 1\ \text{ when } x = 0 \tag{X.11}$$

X.1.1. Case 1

Case (1) of Equation X.9 is solved as follows. Taking the logarithm, we have

$$\log \Lambda(x) = d \log \Lambda(x + B_d). \tag{X.12}$$

Multiplying through by a minus sign and taking the logarithm of both sides, we obtain

$$\log[-\log \Lambda(x)] = \log d + \log[-\log \Lambda(x + B_d)]. \tag{X.13}$$

For simplicity, let

$$g(x) = \log[-\log \Lambda(x)]. \tag{X.14}$$

Then Equation X.13 becomes

$$g(x) = \log d + g(x + B_d). \tag{X.15}$$

Equivalently,

$$g(x - B_d) = \log d + g(x) \tag{X.16}$$

or

$$g(x) = g(x - B_d) - \log d. \tag{X.17}$$

Adding Equation X.15 and Equation X.17, we obtain

$$g(x + B_d) + g(x - B_d) = 2g(x). \tag{X.18}$$

The above equation is valid for all x if and only if $g(x)$ is linear in x. Specifically, let

$$g(x) = kx + j \tag{X.19}$$

where j and k are constants. Then

$$g(x + B_d) = k(x + B_d) + j = g(x) - \log d = kx + j - \log d. \tag{X.20}$$

It follows that

$$kB_d = -\log d \text{ or } k = -\frac{\log d}{B_d}. \tag{X.21}$$

Substituting Equation X.21 in Equation X.19, we see that

$$g(x) + \frac{x \log d}{B_d} = j. \tag{X.22}$$

Using Equation X.14, this result becomes

$$\log[-\log \Lambda(x)] + \frac{x \log d}{B_d} = j. \tag{X.23}$$

Thus, we have

$$\log[-\log \Lambda(x)] = -\frac{x \log d}{B_d} + j. \tag{X.24}$$

Hence, for case (1) of Equation X.9 to hold, $\log[-\log \Lambda(x)]$ must be linear in x.

We now solve for the sequence $\{B_d\}$. For this purpose, let $d = pq$ where p and q are both integers. Note that

$$\Lambda^{pq}(x + B_{pq}) = \Lambda(x). \tag{X.25}$$

From the above equation we get

$$\begin{aligned} \Lambda(x + B_{pq}) &= \Lambda^{1/pq}(x) = [\Lambda^{1/p}(x)]^{1/q} \\ &= [\Lambda(x + B_p]^{1/q} = \Lambda^{1/q}(x + B_p) \\ &= \Lambda((x + B_p) + B_q) = \Lambda(x + B_p + B_q). \end{aligned} \tag{X.26}$$

Equation X.26 implies that

$$B_{pq} = B_p + B_q. \tag{X.27}$$

We now determine the functional dependence of the sequence (B_d) on the subscript d. To emphasize this functional dependence, we rewrite Equation X.27 as

$$B(pq) = B(p) + B(q). \tag{X.28}$$

From the above equation, it is clear that the functional dependence is logarithmic. Thus, the solution for B_d is given by

$$B(d) = B_d = \log d \tag{X.29}$$

Substituting Equation X.29 into Equation X.24 yields

$$\log[-\log \Lambda(x)] = -x + j \tag{X.30}$$

where j plays the role of a location parameter. Hence, without loss of generality, j is chosen to be zero. The above equation then simplifies to

$$\log[-\log \Lambda(x)] = -x. \tag{X.31}$$

Solution for $\Lambda(x)$ results in

$$\Lambda(x) = \exp(-e^{-x}). \tag{X.32}$$

Equation X.32 is the solution of Equation X.9 for case one.

X.1.2. Cases 2 and 3

The solutions to Cases (2) and (3) of Equation X.10 and Equation X.11 are now derived. In both cases we have

$$\Lambda^d(A_d x) = \Lambda(x). \tag{X.33}$$

From Equation X.33 we get

$$\log \Lambda(x) = d \log \Lambda(A_d x). \tag{X.34}$$

Multiplying through by a minus sign and taking the logarithm of both sides, we obtain

$$\log[-\log \Lambda(x)] = \log d + \log[-\log \Lambda(A_d x)]. \tag{X.35}$$

As in case 1, let

$$g(x) = \log[-\log \Lambda(x)]. \tag{X.36}$$

Then Equation X.35 becomes

$$g(x) = \log d + g(A_d x). \tag{X.37}$$

Alternatively,

$$g\left(\frac{x}{A_d}\right) = \log d + g(x) \tag{X.38}$$

or equivalently,

$$g(x) = -\log d + g\left(\frac{x}{A_d}\right). \tag{X.39}$$

Adding Equation X.37 and Equation X.39 results in

$$g(A_d x) + g\left(\frac{x}{A_d}\right) = 2g(x). \tag{X.40}$$

The solution to the above equation is

$$g(x) = \pm k \log x \quad \text{for } x > 0 \tag{X.41}$$

and

$$g(x) = \pm k \log(-x) \quad \text{for } x < 0 \tag{X.42}$$

where k is a positive constant. Use of Equation X.36 in Equation X.41 and Equation X.42 yields

$$\log[-\log \Lambda(x)] = \pm k \log x \quad \text{for } x > 0 \tag{X.43}$$

$$\log[-\log \Lambda(x)] = \pm k \log(-x) \quad \text{for } x < 0. \tag{X.44}$$

For Case 2, $\Lambda = 0$ when $x = 0$. This implies $x = 0$ is the lower end point of $\Lambda(x)$. Hence, $\Lambda(x)$ is nonzero for $x \geq 0$. Therefore, our solution is given by Equation X.43 where we must choose the sign in front of k to be negative. Then

$$\log[-\log \Lambda(x)] = -k \log x \quad x \geq 0 \tag{X.45}$$

which results in

$$\Lambda(x) = \exp(-x^{-k}) \quad x \geq 0. \tag{X.46}$$

For Case 3, $\Lambda = 1$ when $x = 0$. This implies that $x = 0$ is the upper endpoint of $\Lambda(x)$. Hence, $\Lambda(x)$ is nonzero for $x \leq 0$. Consequently, the solution is given by Equation X.44 where we choose the sign in front of k to be positive. Then

$$\log[-\log \Lambda(x)] = k \log(-x) \quad x \leq 0 \tag{X.47}$$

resulting in

$$\Lambda(x) = \exp(-(-x)^k) \quad x \leq 0. \tag{X.48}$$

Thus, the three possible forms for the limiting distribution $\Lambda(x)$ that arise as solutions to Equation X.6 are given as follows:

$$(1) \quad \Lambda(x) = \exp(-e^{-x}) \tag{X.49}$$

$$(2) \quad \Lambda(x) = \exp(-x^{-k}) \quad x \geq 0,\ k > 0 \tag{X.50}$$

$$(3) \quad \Lambda(x) = \exp(-(-x)^k) \quad x \leq 0,\ k > 0. \tag{X.51}$$

X.2. Tails of Probability Density Functions

Equation X.49 to Equation X.51 represent the three possible limiting forms of the distribution function for almost all smooth and continuous probability density functions. By differentiating the three functions, we obtain analytical expressions

for the limiting forms of the probability density functions. However, because of the differentiation, it should be recognized that these expressions may not be good approximations to the density functions. In practice, extreme value theory should always be applied to a distribution function, or equivalently, the area under the density function.

X.2.1. Case 1

The derivative of $\Lambda(x)$ is given by

$$H(x) = \frac{\mathrm{d}}{\mathrm{d}x}\Lambda(x) = \exp(-\mathrm{e}^{-x})(-\mathrm{e}^{-x})(-1) = \mathrm{e}^{-x}\exp(-\mathrm{e}^{-x})$$
$$= \exp(-x - e^{-x}). \tag{X.52}$$

In our application we are interested in the right tail of the probability density function. Since we have to set thresholds corresponding to small false alarm probabilities, the thresholds will be in the right tail of the probability density function. When x is very large, $x \gg \mathrm{e}^{-x}$. Therefore, Equation X.52 can be simplified to obtain the PDF of the tail as

$$H(x) = \mathrm{e}^{-x} \quad x \text{ large}. \tag{X.53}$$

X.2.2. Case 2

The derivative of $\Lambda(x)$ is given by

$$H(x) = \frac{\mathrm{d}}{\mathrm{d}x}\Lambda(x) = \exp(-x^{-k})(kx^{-k-1}) = k\exp(-x^{-k})\mathrm{e}^{(-k-1)\log x}$$
$$= k\exp[-x^{-k} - (k+1)\log x]. \tag{X.54}$$

When x is very large $\log x \gg x^{-k}$. Therefore, Equation X.54 can be simplified to obtain the PDF of the tail as

$$H(x) = k\mathrm{e}^{-(k+1)\log x} = kx^{-(k+1)} \qquad x > 0, \;\; x \text{ large } k > 0. \tag{X.55}$$

X.2.3. Case 3

The derivative of $\Lambda(x)$ for this case is given by

$$H(x) = \frac{\mathrm{d}}{\mathrm{d}x}\Lambda(x) = \exp(-(-x)^k)(k(-x)^{k-1}) = k\exp(-(-x)^k)\mathrm{e}^{(k-1)\log(-x)}$$
$$= k\exp[-(-x)^k + (k-1)\log x]. \tag{X.56}$$

When $-x$ is very large, $(-x)^k \gg \log x$. Therefore, Equation X.56 can be simplified to obtain the PDF of the tail as

$$H(x) = k\, \mathrm{e}^{-(-x)^k} \qquad x < 0, \ -x \text{ large } k > 0. \tag{X.57}$$

A basic assumption in the above development is that successive trials are independent. This led to Equation X.1. In practice, as n becomes large, it may be difficult to ensure the independence of successive trials. To the extent that the assumption holds, the results in Equation X.49 to Equation X.51 are valid.

X.3. PDF OF THE rTH ORDER STATISTIC

Suppose that the ordered samples $X_1 \le X_2 \le \cdots \le X_n$ are drawn from the distribution function $F(x)$. Let us further assume that the trials used to draw the samples from the distribution are independent. Consider the rth order statistic X_r. Recall that $P(X_r \le x)$ is the distribution function of X_r. This, in turn, is the probability that at least r of the X_i's are less than or equal to x. Treating this as a Binomial problem, the distribution function is

$$Fx_r(x) = P(X_r \le x) = \sum_{i=r}^{n} \frac{n!}{i!(n-i)!} F^i(x)[1 - F(x)]^{n-i} \tag{X.58}$$

where the ith term in the summation is the binomial probability that exactly i of $X_1, X_2, \ldots, X_n$ are less than or equal to x. Equation X.58 can also be represented in the form of an integral,

$$F_{X_r}(x) = \frac{n!}{(r-1)!(n-r)!} \int_0^{F(x)} t^{r-1}(1-t)^{n-r}\, \mathrm{d}t \tag{X.59}$$

which can be verified by using integration by parts in Equation X.59. The probability density function of the rth order statistic is the derivative of $F_{X_r}(x)$ and is given by

$$\begin{aligned} f_{X_r}(x) &= \frac{\mathrm{d}}{\mathrm{d}x} F_{X_r}(x) = \frac{n!}{(r-1)!(n-r)!} \frac{\mathrm{d}}{\mathrm{d}x} \int_0^{F(x)} t^{r-1}(1-t)^{n-r}\mathrm{d}t \\ &= \frac{n!}{(r-1)!(n-r)!} F^{r-1}(x)[1-F(x)]^{n-r} f(x) \end{aligned} \tag{X.60}$$

where $f(x) = (\mathrm{d}/\mathrm{d}x)F(x)$. Equation X.60 represents the general form of the PDF of the rth order statistic. If $F(x)$ is known, then the mean and the variance

of the rth order statistic can be calculated. The expected value of X_r is given by

$$E(X_r) = \frac{n!}{(r-1)!(n-r)!} \int_{-\infty}^{\infty} xF^{r-1}(x)[1-F(x)]^{n-r} f(x)\mathrm{d}x. \tag{X.61}$$

An alternate form for the expected value of X_r can be obtained by letting $u = F(x)$. Therefore, $x = F^{-1}(u)$. The infinite limits of the integral in the above equation then become finite after the transformation. The transformed integral is

$$E(X_r) = \frac{n!}{(r-1)!(n-r)!} \int_0^1 F^{-1}(u)u^{r-1}(1-u)^{n-r}\,\mathrm{d}u. \tag{X.62}$$

The variance of the rth order statistic is expressed as

$$\mathrm{Var}(X_r) = E\{[X_r - E(X_r)]^2\} = E(X_r^2) - E^2(X_r). \tag{X.63}$$

Making use of Equation X.60, $E(X_r^2)$ can be written as follows.

$$E(X_r^2) = \frac{n!}{(r-1)!(n-r)!} \int_{-\infty}^{\infty} x^2F^{r-1}(x)[1-F(x)]^{n-r} f(x)\mathrm{d}x. \tag{X.64}$$

An alternate form for the expected value of X_r can be obtained by again letting $u = F(x)$. We then get

$$E(X_r^2) = \frac{n!}{(r-1)!(n-r)!} \int_0^1 [F^{-1}(u)]^2u^{r-1}(1-u)^{n-r}\,\mathrm{d}u. \tag{X.65}$$

The variance of X_r can be calculated from Equation X.64 and Equation X.65 when $F^{-1}(u)$ is known.

APPENDIX Y. CANONICAL FORM DERIVATION (SECTION 15.6)

A derivation of the canonical form given in Figure 15.47 for the optimal SIRV receiver specified by Equation 15.159 is presented. The target return is assumed to depend on an unknown parameter vector, $\mathbf{v}$, which has a probability density function, $f_{\mathbf{V}}(v)$, defined on a parameter space, Ω_v. The optimum NP test is given by

$$T(\tilde{\mathbf{r}}) = \frac{\displaystyle\int_{\Omega_v} \pi^{-N}|\tilde{\boldsymbol{\Sigma}}|^{-1}h_{2N}[2(\tilde{\mathbf{r}} - \tilde{\mathbf{s}}(\mathbf{v}))^H\tilde{\boldsymbol{\Sigma}}^{-1}(\tilde{\mathbf{r}} - \tilde{\mathbf{s}}(\mathbf{v}))]f_{\mathbf{V}}(\mathbf{v})d\mathbf{v}}{\pi^{-N}|\tilde{\boldsymbol{\Sigma}}|^{-1}h_{2N}[2\tilde{\mathbf{r}}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\mathbf{r}}]} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \tag{Y.1}$$

After the integral form for $h_{2N}(\cdot)$ is substituted into the numerator, the test becomes

$$T(\tilde{\mathbf{r}}) = \frac{\int_{\Omega_v}\int_0^\infty s^{-2N} f_S(s)\exp\left[-\dfrac{(\tilde{\mathbf{r}}-\tilde{\mathbf{s}}(\mathbf{v}))^H\tilde{\boldsymbol{\Sigma}}^{-1}(\tilde{\mathbf{r}}-\tilde{\mathbf{s}}(\mathbf{v}))}{s^2}\right] f_{\mathbf{V}}(\mathbf{v})\mathrm{d}s\,\mathrm{d}\mathbf{v}}{h_{2N}[2\tilde{\mathbf{r}}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\mathbf{r}}]} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \quad \text{(Y.2)}$$

The denominator does not depend on $\mathbf{v}$ or s and can be moved inside the integrand to yield

$$T(\tilde{\mathbf{r}}) = \int_{\Omega_v}\int_0^\infty \frac{s^{-2N} f_S(s)\exp\left[\dfrac{(\tilde{\mathbf{r}}-\tilde{\mathbf{s}}(\mathbf{v}))^H\tilde{\boldsymbol{\Sigma}}^{-1}(\tilde{\mathbf{r}}-\tilde{\mathbf{s}}(\mathbf{v}))}{s^2}\right] f_{\mathbf{V}}(\mathbf{v})\mathrm{d}s\,\mathrm{d}\mathbf{v}}{h_{2N}[2\tilde{\mathbf{r}}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\mathbf{r}}]} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \quad \text{(Y.3)}$$

Interchanging the order of integration and expanding the quadratic form in the numerator of the integrand leads to

$$T(\tilde{\mathbf{r}}) = \int_0^\infty \frac{\exp\left(-\dfrac{\tilde{\mathbf{r}}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\mathbf{r}}}{s^2}\right)}{h_{2N}[2\tilde{\mathbf{r}}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\mathbf{r}}]} s^{-2N} f_S(s)\Gamma(\tilde{\mathbf{r}}|s)\mathrm{d}s, \quad \text{(Y.4)}$$

where $\Gamma(\tilde{\mathbf{r}}|s)$ is defined to be

$$\Gamma(\tilde{\mathbf{r}}|s) = \int_{\Omega_v}\exp\left[-\frac{-2\mathrm{Re}\{\tilde{\mathbf{s}}(\mathbf{v})^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\mathbf{r}}\}+\tilde{\mathbf{s}}(\mathbf{v})^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\mathbf{s}}(\mathbf{v})}{s^2}\right] f_V(\mathbf{v})\mathrm{d}\mathbf{v}. \quad \text{(Y.5)}$$

$\Gamma(\tilde{\mathbf{r}}|s)$ is recognized as the Gaussian receiver, for a given value of s, which would usually be simplified to a sufficient statistic, $T_g(\tilde{\mathbf{r}})$ (by incorporating constant terms into a threshold and applying the inverses of monotonic functions). This simplification cannot be performed in Equation Y.5 because s is the variable of integration in Equation Y.4. Nonetheless, $\Gamma(\tilde{\mathbf{r}}|s)$ can be written in the general functional form,

$$\Gamma(\tilde{\mathbf{r}}|s) = g\left[T_g\left(\frac{\tilde{\mathbf{r}}}{s^2}\right), s\right]. \quad \text{(Y.6)}$$

Substitution of this expression for $\Gamma(\tilde{\mathbf{r}}|s)$ into Equation Y.4 yields

$$T(\tilde{\mathbf{r}}) = \int_0^\infty \frac{\exp\left(-\dfrac{\tilde{\mathbf{r}}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\mathbf{r}}}{s^2}\right)}{h_{2N}[2\tilde{\mathbf{r}}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\mathbf{r}}]} s^{-2N} f_S(s) g\left[T_g\left(\frac{\tilde{\mathbf{r}}}{s^2}\right), s\right]\mathrm{d}s \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \quad \text{(Y.7)}$$

In general, the integral in the above result is not an explicit function of the sufficient statistic, $T_g(\tilde{\mathbf{r}})$, of the Gaussian problem. However, if $T_g(\tilde{\mathbf{r}})$ satisfies the separability condition,

$$T_g(c\tilde{\mathbf{r}}) = f(c)T_g(\tilde{\mathbf{r}}), \tag{Y.8}$$

where c does not depend on $\tilde{\mathbf{r}}$ and $f(\cdot)$ is any function, then Equation Y.7 simplifies to

$$T(\tilde{\mathbf{r}}) = \int_0^\infty \frac{\exp\left(-\dfrac{\tilde{\mathbf{r}}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\mathbf{r}}}{s^2}\right)}{h_{2N}[2\tilde{\mathbf{r}}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\mathbf{r}}]} s^{-2N} f_S(s) g[f(s^{-2})T_g(\tilde{\mathbf{r}}), s]\mathrm{d}s \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \tag{Y.9}$$

This result is an explicit function of the Gaussian sufficient statistic and can be written in general as

$$T(\tilde{\mathbf{r}}) = g_{\mathrm{NL}}[T_g(\tilde{\mathbf{r}}), \tilde{\mathbf{r}}^H\tilde{\boldsymbol{\Sigma}}^{-1}\tilde{\mathbf{r}}] \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \tag{Y.10}$$

where $g_{\mathrm{NL}}[\cdot,\cdot]$ is the optimal nonlinearity which must be determined for each SIRV receiver. This general form is implemented by the canonical receiver structure given in Figure 15.47.

APPENDIX Z. ALTERNATIVE SPHERICAL COORDINATE SIRV REPRESENTATIONS (SECTION 15.6)

Any zero mean random vector, $\mathbf{X} = [X_1 \cdots X_N]$, with identity covariance matrix can be expressed in terms of a commonly used set of N generalized spherical coordinate variables, R, Θ, and $\Phi_1, \ldots, \Phi_{N-2}$, by the transformation equations,

$$X_1 = R\cos\Phi_1 \tag{Z.1}$$

$$X_k = R\cos\Phi_k \prod_{i=1}^{k-1} \sin\Phi_i, \quad k = 2, \ldots, N-2 \tag{Z.2}$$

$$X_{N-1} = R\cos\Theta \prod_{i=1}^{N-2} \sin\Phi_i \tag{Z.3}$$

$$X_N = R\sin\Theta \prod_{i=1}^{N-2} \sin\Phi_i, \tag{Z.4}$$

with

$$R^2 = X_1^2 + X_2^2 + \cdots + X_N^2 = \mathbf{X}^{\mathrm{T}}\mathbf{X} \tag{Z.5}$$

and

$$dX_1 dX_2 \cdots dx_N = |J| d\Phi_{N-2} \cdots d\Phi_2 d\Phi_1 d\Theta dR, \tag{Z.6}$$

where J is the Jacobian of the transformation, given by

$$J = R^{N-1} \prod_{i=1}^{N-2} \sin^{N-1-i}(\Phi_i). \tag{Z.7}$$

The transformation is one-to-one over the entire N-dimensional space for angular coordinate variables restricted to the intervals $0 \leq \Theta < 2\pi$ and $0 \leq \Phi_k \leq \pi$ for $k = 1, \ldots, N-2$.

This generalized spherical coordinate representation is not unique for $N > 3$. Two distinct representations are possible for $N = 4$ and more distinct spherical coordinate system choices are available as N increases.[17,18] The transformation equations above may be interpreted as the decomposition of the generalized N-dimensional sphere into a diameter (one-dimensional spherical coordinate system) and an $(N-1)$-dimensional spherical coordinate system. The $(N-1)$-dimensional sphere is then decomposed into another diameter and an $(N-2)$-dimensional spherical coordinate system, which is further decomposed in this way.

An alternative spherical coordinate representation is derived here. The particular alternative spherical coordinate system of interest may be interpreted as a decomposition of a generalized N-dimensional spherical coordinate system into a circle (two-dimensional spherical coordinate system) and an $(N-2)$-dimensional spherical coordinate system. Further decomposition of the $(N-2)$-dimensional sphere is arbitrary for applications considered in this research. However, N is assumed to be an even integer and the decomposition into another circle at each stage is continued for this example.

A technique for developing alternative spherical coordinate system representations is presented in Ref. 18. It is based on

$$k(u) = \int_0^\infty x^u \exp\left(-\frac{x^2}{2}\right) dx, \quad \mathrm{Re}\{u\} > -1, \tag{Z.8}$$

which is a variant of the gamma function, and the resulting form of the Euler beta integral,

$$\int_0^{\pi/2} \cos^a\theta \sin^b\theta \, d\theta = \frac{k(a)k(b)}{k(a+b+1)}, \quad \mathrm{Re}\{a\}, \mathrm{Re}\{b\} > -1. \tag{Z.9}$$

These two relations are substituted for the appropriate factors in the identity,

$$
\begin{aligned}
&k(a_1)k(a_2)\cdots k(a_N)\\
&\quad = k(a_1+a_2+\cdots+a_N+N-1)\\
&\qquad \times \frac{k(a_1+a_2+1)k(a_3+a_4+\cdots+a_N+N-3)}{k(a_1+a_2+\cdots+a_N+N-1)}\,\frac{k(a_1)k(a_2)}{k(a_1+a_2+1)}\\
&\qquad \times \frac{k(a_3+a_4+1)k(a_5+a_6+\cdots+a_N+N-5)}{k(a_3+a_4+\cdots+a_N+N-3)}\,\frac{k(a_3)k(a_4)}{k(a_3+a_4+1)}\\
&\qquad \times \frac{k(a_5+a_6+1)k(a_7+a_8+\cdots+a_N+N-7)}{k(a_5+a_6+\cdots+a_N+N-5)}\,\frac{k(a_5)k(a_6)}{k(a_5+a_6+1)}\\
&\qquad \vdots\\
&\qquad \times \frac{k(a_{n-3}+a_{N-2}+1)k(a_{N-1}+a_N+1)}{k(a_{N-3}+a_{N-2}+a_{N-1}+a_N+3)}\,\frac{k(a_{N-3})k(a_{N-2})}{k(a_{N-3}+a_{N-2}+1)}\\
&\qquad \times \frac{k(a_{N-1})k(a_N)}{k(a_{N-1}+a_N+1)}.
\end{aligned}
\tag{Z.10}
$$

There are many ways to construct the right hand side of this identity. This choice results in the desired decomposition which is discussed above.

The integral in Equation Z.8 is used to express the product on the left hand side of Equation Z.10 as

$$
\begin{aligned}
k(a_1)k(a_2)\cdots k(a_N) = \int_0^\infty \cdots \int_0^\infty x_1^{a_1}x_2^{a_2}\cdots x_N^{a_N}\\
\times \exp\left(-\frac{x_1^2+x_2^2+\cdots+x_N^2}{2}\right)\mathrm{d}x_1\cdots\mathrm{d}x_N
\end{aligned}
\tag{Z.11}
$$

and to express the first factor on the right hand side of Equation Z.10 as

$$
\begin{aligned}
k(a_1+a_2+\cdots+a_N+N-1) = \int_0^\infty R^{a_1+a_2+\cdots+a_N+N-1}\\
\times \exp\left(-\frac{R^2}{2}\right)\mathrm{d}R,
\end{aligned}
\tag{Z.12}
$$

where R is used as the variable of integration in anticipation of the spherical coordinate determinations. The Euler beta integral in Equation Z.9 is used to express the remaining factors on the right hand side of Equation Z.10 as integrals over a set of angular variables. If $\Phi_1, \Phi_2, \ldots, \Phi_{(N/2)-1}$ and $\Theta_1, \Theta_2, \ldots, \Theta_{N/2}$

are used to denote these integration variables, then Equation Z.10 becomes

$$\begin{aligned}
&\int_0^\infty \cdots \int_0^\infty x_1^{a_1} x_2^{a_2} \cdots x_N^{a_N} \exp\left(-\frac{x_1^2 + x_2^2 + \cdots + x_N^2}{2}\right) \mathrm{d}x_1 \cdots \mathrm{d}x_N \\
&\quad = \int_0^\infty R^{a_1 + a_2 + \cdots + a_N + N - 1} \exp\left(-\frac{R^2}{2}\right) \mathrm{d}R \\
&\qquad \times \int_0^{\pi/2} (\cos \Phi_1)^{a_1 + a_2 + 1} (\sin \Phi_1)^{a_3 + a_4 + \cdots + a_N + N - 3} \mathrm{d}\Phi_1 \\
&\qquad \times \int_0^{\pi/2} (\cos \Theta_1)^{a_1} (\sin \Theta_1)^{a_2} \mathrm{d}\Theta_1 \\
&\qquad \times \int_0^{\pi/2} (\cos \Phi_2)^{a_3 + a_4 + 1} (\sin \Phi_2)^{a_5 + a_6 + \cdots + a_N + N - 5} \mathrm{d}\Phi_2 \\
&\qquad \times \int_0^{\pi/2} (\cos \Theta_2)^{a_3} (\sin \Theta_2)^{a_4} \mathrm{d}\Theta_2 \\
&\qquad \times \int_0^{\pi/2} (\cos \Phi_3)^{a_5 + a_6 + 1} (\sin \Phi_3)^{a_7 + a_8 + \cdots + a_N + N - 7} \mathrm{d}\Phi_3 \\
&\qquad \times \int_0^{\pi/2} (\cos \Theta_3)^{a_5} (\sin \Theta_3)^{a_6} \mathrm{d}\Theta_3 \\
&\qquad \vdots \\
&\qquad \times \int_0^{\pi/2} \left(\cos \Phi_{(N/2)-1}\right)^{a_{N-3} + a_{N-2} + 1} \left(\sin \Phi_{(N/2)-1}\right)^{a_{N-1} + a_N + 1} \mathrm{d}\Phi_{(N/2)-1} \\
&\qquad \times \int_0^{\pi/2} \left(\cos \Phi_{(N/2)-1}\right)^{a_{N-3}} \left(\sin \Phi_{(N/2)-1}\right)^{a_{N-2}} \mathrm{d}\Phi_{(N/2)-1} \\
&\qquad \times \int_0^{\pi/2} \left(\cos \Theta_{N/2}\right)^{a_{N-1}} \left(\sin \Theta_{N/2}\right)^{a_N} \mathrm{d}\Theta_{N/2} \qquad (Z.13)
\end{aligned}$$

Grouping terms with common exponents in the above expression leads to

$$\begin{aligned}
&\int_0^\infty \cdots \int_0^\infty x_1^{a_1} x_2^{a_2} \cdots x_N^{a_N} \exp\left(-\frac{x_1^2 + x_2^2 + \cdots + x_N^2}{2}\right) \mathrm{d}x_1 \cdots \mathrm{d}x_N \\
&= \int_0^\infty \int_0^{\pi/2} \cdots \int_0^{\pi/2} (R \cos \Phi_1 \cos \Theta_1)^{a_1} (R \cos \Phi_1 \sin \Theta_1)^{a_2} \\
&\quad \times \left(R \sin \Phi_1 \cos \Phi_2 \cos \Theta_2\right)^{a_3} (R \sin \Phi_1 \cos \Phi_2 \sin \Theta_2)^{a_4} \\
&\quad \vdots \\
&\quad \times \left(R \sin \Phi_1 \sin \Phi_2 \cdots \sin \Phi_{(N/2)-2} \cos \Phi_{(N/2)-1} \cos \Theta_{(N/2)-1}\right)^{a_{N-3}} \\
&\quad \times \left(R \sin \Phi_1 \sin \Phi_2 \cdots \sin \Phi_{(N/2)-2} \cos \Phi_{(N/2)-1} \sin \Theta_{(N/2)-1}\right)^{a_{N-2}} \\
&\quad \times \left(R \sin \Phi_1 \sin \Phi_2 \cdots \sin \Phi_{(N/2)-1} \cos \Theta_{N/2}\right)^{a_{N-1}} \\
&\quad \times \left(R \sin \Phi_1 \sin \Phi_2 \cdots \sin \Phi_{(N/2)-1} \sin \Theta_{N/2}\right)^{a_N} \exp\left(-\frac{R^2}{2}\right) \\
&\quad \times R^{N-1} \left(\cos \Phi_1 \sin^{N-3} \Phi_1\right) \left(\cos \Phi_2 \sin^{N-5} \Phi_2\right) \cdots \left(\cos \Phi_{(N/2)-1} \sin \Phi_{(N/2)-1}\right) \\
&\quad \times \mathrm{d}\Phi_1 \mathrm{d}\Phi_2 \cdots \mathrm{d}\Phi_{(N/2)-1} \mathrm{d}\Theta_1 \mathrm{d}\Theta_2 \cdots \mathrm{d}\Theta_{N/2} \mathrm{d}R. \qquad (Z.14)
\end{aligned}$$

Equating the integrands on both sides of this expression results in the transformation equations,

$$X_1 = R \cos \Theta_1 \cos \Phi_1 \tag{Z.15}$$

$$X_2 = R \sin \Theta_1 \cos \Phi_1 \tag{Z.16}$$

$$X_{2k-1} = R \cos \Theta_k \cos \Phi_k \prod_{i=1}^{k-1} \sin \Phi_i, \;\; k = 2, \ldots, \frac{N}{2} - 1 \tag{Z.17}$$

$$X_{2k} = R \sin \Theta_k \cos \Phi_k \prod_{i=1}^{k-1} \sin \Phi_i, \;\; k = 2, \ldots, \frac{N}{2} - 1 \tag{Z.18}$$

$$X_{N-1} = R \cos \Theta_{N/2} \prod_{i=1}^{(N/2)-1} \sin \Phi_i \tag{Z.19}$$

$$X_N = R \sin \Theta_{N/2} \prod_{i=1}^{(N/2)-1} \sin \Phi_i \tag{Z.20}$$

with

$$R^2 = X_1^2 + X_2^2 + \cdots + X_N^2 \tag{Z.21}$$

and the Jacobian,

$$J = R^{N-1} \prod_{i=1}^{(N/2)-1} (\sin \Phi_i)^{N-1-2i} \cos \Phi_i, \tag{Z.22}$$

for this alternative spherical coordinate representation.

The integration in Equation Z.14 covers only the positive orthant of the N-dimensional space. The transformation, Equation Z.15 to Equation Z.20, is one-to-one over all 2^N orthants for

$$0 \leq \Phi_i \leq \frac{\pi}{2}, \qquad i = 1, \ldots \frac{N}{2} - 1, \tag{Z.23}$$

$$0 \leq \Theta_i < 2\pi, \qquad i = 1, \ldots \frac{N}{2}, \tag{Z.24}$$

and $0 \leq R < \infty$. The extension of the transformation to cover the entire N-dimensional space is accomplished by expanding the range of the Θ_i coordinate variables in Equation Z.24.

The multivariate probability density function of the N spherical coordinate variables is given by

$$F_{R, \Phi_1, \ldots \Phi_{(N/2)-1}, \Theta_1, \ldots \Theta_{N/2}}(r, \phi_1, \ldots \phi_{(N/2)-1}, \theta_1, \ldots \theta_{N/2}) = |J| f_{\mathbf{X}}(\mathbf{x}) \tag{Z.25}$$

The PDF of the SIRV, $\mathbf{X}$, is

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-N/2} h_N(\mathbf{x}^{\mathrm{T}} \mathbf{x}). \tag{Z.26}$$

Substitution of Equation Z.21, Equation Z.22, and Equation Z.26 into Equation Z.25 for the joint PDF of the spherical coordinate variables leads to

$$f_{R,\Phi_1,\ldots\Phi_{(N/2)-1},\Theta_1,\ldots\Theta_{N/2}}(r,\phi_1,\ldots\phi_{(N/2)-1},\theta_1,\ldots\theta_{N/2})$$

$$= (2\pi)^{-N/2} r^{N-1} h_N(r^2) \prod_{i=1}^{(N/2)-1} (\sin\phi_i)^{N-1-2i} \cos\phi_i. \tag{Z.27}$$

Since the right hand side of this equation is factored into separate functions of each of the spherical coordinate variables, $R, \Phi_1,\ldots,\Phi_{(N/2)-1},\Theta_1,\ldots,\Theta_{N/2}$, then these coordinate variables are all statistically independent. Their respective PDFs are determined to be

$$f_R(r) = \frac{r^{N-1}}{2^{(N/2)-1}\Gamma\left(\frac{N}{2}\right)} h_N(r^2), \quad 0 \le r < \infty \tag{Z.28}$$

$$f_{\Phi_k}(\Phi_k) = (N-2k)\cos\phi_k(\sin\phi_k)^{N-2k-1},$$
$$0 \le \phi_k \le \frac{\pi}{2},\ k = 1,\ldots,\frac{N}{2}-1 \tag{Z.29}$$

$$f_{\Theta_k}(\theta_k) = \frac{1}{2\pi}, \quad 0 \le \theta < 2\pi,\ k = 1,\ldots,\frac{N}{2}. \tag{Z.30}$$

The integral,

$$\int_0^{\pi/2} \cos\omega(\sin\omega)^{N-2k-1} \mathrm{d}\omega = \frac{1}{N-2k},\ k = 1,\ldots,\frac{N}{2}-1, \tag{Z.31}$$

is used in determining the $f_{\Phi_k}(\phi_k)$.

An interpretation of the angular coordinate variables in this alternative generalized spherical coordinate representation is obtained by comparing it to the standard representation of Equation Z.2 to Equation Z.4. The Φ_i variables of Equation Z.23 are seen to be analogous to the angular coordinate variables of a standard $N/2$-dimensional generalized spherical coordinate representation, except in this case each of the $N/2$ orthogonal dimensions is a plane. Points on each of these planes are represented in polar coordinates local to each plane and the Θ_i variables of Equation Z.24 are the localized polar angles. This is particularly useful for representing a vector of complex envelope samples.

Acronyms

PREFACE

CFAR	Constant False-Alarm Rate
LOD	Locally Optimum Detector
LR	Likelihood Ratio
PDF	Probability Density Function
SCV	SubClutter Visibility

CHAPTER 1

ASCAPE	Automatic Statistical Characterization and Partitioning of Environments
CFAR	Constant False-Alarm Rate
IPUS	Integrated Processing and Understanding of Signals
NHD	Nonhomogeneous Detector
PDF	Probability Density Function

CHAPTER 2

SECTION 2.1

(H_1, H_0)	Binary Hypothesis
LOD	Locally Optimum Detector
LR	Likelihood Ratio
LRT	Likelihood Ratio Test
PD	Probability of Detection
PDF	Probability Density Function
PFA	Probability of False Alarm
SIRP	Spherically Invariant Random Process
SIRV	Spherically Invariant Random Vector

SECTION 2.2

BN	Background Noise
CL	Clutter

GLRT	Generalized Likelihood Ratio Test
ISC	Intermediate Signal Case
LOD	Locally Optimum Detector
LRT	Likelihood Ratio Test
PDF	Probability Density Function
SIRP	Spherically Invariant Random Process
SSC	Strong Signal Case
WSC	Weak Signal Case
SUD/FFES	Signal Understanding and Detection using a Feedforward Expert System
WSUD/FFES	Weak Signal Understanding and Detection using a Feedforward Expert System

SECTION 2.3

GLRT	Generalized Likelihood Ratio Test
IPUS	Integrated Processing and Understanding of Signals
KS	Knowledge Source
LOD	Locally Optimum Detector
PDF	Probability Density Function
RESUN	Resolving Sources of Uncertainty
SOU	Sources of Uncertainty
SPA	Signal Processing Algorithm
WSUD/FFES	Weak Signal Understanding and Detection using a Feedforward Expert System

SECTION 2.4

BN	Background Noise
CL	Clutter
FFT	Fast Fourier Transform
IPUS	Integrated Processing and Understanding of Signals
PRI	Pulse Repetition Interval
R/A	Range-Azimuth
R/A/D	Range-Azimuth-Doppler
RF	Radio-Frequency

SECTION 2.5

BN	Background Noise
$(\text{BN} \rightarrow \text{CL})_i$	Percentage of the total number of cells in the surveillance volume that were below the threshold but are reclassified as Clutter cells after the *i*th correction stage
BNCCP	Percentage of Background-Noise cells in the corrected-corrected volume

BNCQP	Percentage of Background Noise cells in the corrected-Quantized Volume
BNQP	Percentage of Background Noise cells in the Quantized Volume
CCV	Corrected-Corrected Volume (second corrected volume)
CL	Clutter
$(CL \rightarrow BN)_i$	Percentage of the total number of cells in the surveillance volume that were above the threshold but are reclassified as backgound noise cells after the *i*th correction stage
CLE	Clutter Edge
CNR	Clutter-to-Noise Ratio
CQV	Corrected-Quantized Volume (first corrected volume)
CQV → CCV	Mapping procedure during the second-correction stage
FFT	Fast-Fourier Transform
IPUS	Integrated Processing and Understanding of Signals
NCC	Number of neighboring Cells in the first Corrected volume
NCQ	Number of neighboring Cells in the Quantized volume
NS	Neighboring cells in the Smoothed Volume
PDF	Probability Density Function
PRI	Pulse Repetition Interval
QV	Quantized Volume
QV → CCV	Mapping procedure at the end of the two correction stages
QV → CQV	Mapping procedure during the first-correction stage
R/A	Range-Azimuth
SNR	Signal-to-Noise Ratio

SECTION 2.6

BN	Background Noise
CL	Clutter
CNR	Clutter-to-Noise Ratio
ISC	Intermediate Signal Case
NCC	Number of neighboring Cells in the first Corrected volume
NCQ	Number of neighboring Cells in the Quantized Volume
PDF	Probability Density Function
PLCCP	Percentage number of cells of the subpatch with the Lowest power among the possible subpatches of a Clutter patch in the second Corrected stage
PLCQP	Percentage number of cells of the subpatch with the lowest power among the possible subpatches of a clutter patch in the first-corrected stage
PLQP	Percentage number of cells of the subpatch with the lowest power among the possible subpatches of a Clutter patch in the quantized stage
R/A	Range/Azimuth

R/A/D	Range/Azimuth/Doppler
SSC	Strong Signal Case
SUD/FBES	Signal Understanding and Detection using a Feedback Expert System
SUD/FFES	Signal Understanding and Detection using a Feed Forward Expert System
WSC	Weak Signal Case

SECTION 2.7

BN	Background Noise
BNCCP	Percentage of Background-Noise cells in the Corrected-Corrected volume
BNCQP	Percentage of Background Noise cells in the Corrected-Quantized Volume
BNQP	Percentage of Background Noise cells in the Quantized Volume
$BNQP_t$	True value for the fraction of Background Noise cells in the generated scene
CL	Clutter
CNR	Clutter-to-Noise Ratio
FFES	Feed-Forward Expert System
GLRT	Generalized Likelihood Ratio Test
IPUS	Integrated Processing and Understanding of Signals
ISC	Intermediate Signal Case
LOD	Locally Optimum Detector
LRT	Likelihood Ratio Test
N/A	Not Applicable
NCC	Number of the neighboring Cells in the first Corrected volume
NCQ	Number of neighboring Cells in the Quantized volume
PDF	Probability Density Function
PLQP	Percentage number of cells of the subpatch with the lowest power among the possible subpatches of a clutter patch in the quantized stage
$PLQP_t$	True value of PLQP
SPA	Signal Processing Algorithm
SSC	Strong Signal Case
WSC	Weak Signal Case

CHAPTER 3

SECTION 3.1

AMF	Adaptive Matched Filter
CFAR	Constant False-Alarm Rate

GIP	Generalized Inner Product
IID	Independent Identically Distributed
MCARM	Multi-Channel Airborne Radar Measurement
MF	Matched Filter
NAMF	Normalized Adaptive Matched Filter
NHD	Nonhomogeneity Detector
N PAMF	Normalized Parametric Adaptive Matched Filter
PAMF	Parametric Adaptive Matched Filter
PDF	Probability Density Function
SINR	Signal-to-Interference Noise Ratio
SNR	Signal-to-Noise Ratio
STAP	Space–Time Adaptive Processing

SECTION 3.2

CFAR	Constant False-Alarm Rate
DFT	Discrete Fourier Transform
EM	Expectation-Maximization
GIP	Generalized Inner Product
IID	Independent Identically Distributed
MAR	Missing at Random
MCARM	Multichannel Airborne Radar Measurement
ML	Maximum Likelihood
MMSE	Minimum Mean Squared Error
NAMF	Normalized Adaptive Matched Filter
NHD	Nonhomogeneity Detector
PDF	Probability Density Function
SIRP	Spherically Invariant Random Process
SIRV	Spherically Invariant Random Vector
STAP	Space–Time Adaptive Processing

CHAPTER 4

SECTION 4.1

CDF	Cumulative Distribution Function
PDF	Probability Density Function
P–P	Probability–Probability
Q–Q	Quantile–Quantile

SECTION 4.2

SIRV	Spherically Invariant Random Vector
VLSI	Very Large–Scale Integrated

SECTION 4.4

SIRP	Spherically Invariant Random Process

CHAPTER 5

SECTION 5.1

IMSL	International Mathematical and Statistical Libraries
MLE	Maximum Likelihood Estimate
Q–Q	Quantile–Quantile

SECTION 5.2

FORTRAN	Formula + Translation
IMSL	International Mathematical and Statistical Libraries

CHAPTER 6

SECTION 6.1

PDF	Probability Density Function
PDFs	Probability Density Functions

SECTION 6.2

BN	Background Noise
BNCCP	Percentage of BN cells in the "Corrected-Corrected" volume
BNCQP	Percentage of BN cells in the "Corrected-Quantized" volume
BNQP	Percentage of BN cells in the Quantized volume
CCV	"Corrected-Corrected" Volume
CL	Clutter
CLE	Clutter edge
CNR	Clutter-to-Noise Ratio
CQV	"Corrected-Quantized" Volume
LOD	Locally Optimum Detector
LRT	Likelihood Ratio Test
NCC	Number of neighboring Cells in the first Corrected volume
NCQ	Number of neighboring Cells in the Quantized Volume
PDF	Probability Density Function
PRI	Pulse Repetition Interval
R/A	Range-Azimuth
SIRP	Spherically Invariant Random Process
SV	Smoothed Volume

SECTION 6.3

BN	Background Noise
BNCCP	Percentage of BN cells in the "Corrected-Corrected" volume
$(BN \rightarrow CL)_i$	Percentage of the total number of cells in the surveillance volume that were below the threshold but are reclassified as CL cells after the ith correction stage
BNCQP	Percentage of BN cells in the "Corrected-Quantized" Volume
BNQP	Percentage of BN cells in the Quantized Volume
CCV	"Corrected-Corrected" volume
CL	Clutter
$(CL \rightarrow BN)_i$	Percentage of the total number of cells in the surveillance volume that were above the threshold but are reclassified as BN cells during the ith correction stage
CQV	"Corrected-Quantized" Volume
NCC	Number of neighboring Cells in the first Corrected volume
NCQ	Number of neighboring Cells in the Quantized Volume
PDF	Probability Density Function
QV	Quantized Volume

SECTION 6.4

ASCAPE	Automatic Statistical Characterization and Partitioning of Environments
IPUS	Integrated Processing and Understanding of Signals
LP	Lowest Patch (patch with lowest average power level
LPE	LP Edge cell (cell with power level below latest threshold)
NR	Sample size
PDF	Probability Density Function
RP	Remaining Patch (patch excluding LP)
RPE	RP Edge cell (cell with power levels above latest threshold)
SPA	Signal Processing Algorithm

SECTION 6.5

ASCAPE	Automated Statistical Characterization and Partitioning of Environments
GLRT	Generalized Likelihood Ratio Test
INR	Interference-to-Noise Ratio
IPUS	Integrated Processing and Understanding of Signals
LOD	Locally Optimum Detector
LRT	Likelihood Ratio Test
PD	Probability of Detection
PDF	Probability Distribution Function

SCR	Signal-to-Clutter Ratio
SDR	Signal-to-Disturbance Ratio

SECTION 6.6

ASCAPE	Automated Statistical Characterization and Partitioning of Environments
GLRT	Generalized Likelihood Ratio Test
ML	Maximum Likelihood
MMW	MilliMeter Wave
PDF	Probability Distribution Function
SCR	Signal-to-Clutter Ratio
SIRV	Spherically Invariant Random Vector

SECTION 6.7

CFAR	Constant False Alarm Rate
KBMapSTAP	Knowledge Based Map Space Time Adaptive Processing
LULC	Land Use and Land Cover
MCARM	Multi-Channel Airborne Radar Measurements
MSMI	Modified Sample Matrix Inversion
PRF	Pulse Repetition Frequency
STAP	Space–Time Adaptive Processing
USGS	United States Geological Survey

SECTION 6.8

AFRL	Air Force Research Laboratory
CPI	Coherent Processing Interval
DOF	Degrees Of Freedom
ECEF	Earth-Centered Earth-Fixed
GMTI	Ground Moving Target Indicator
MCARM	Multi-Channel Airborne Radar Measurements
NLCD	National Land Cover Data
MSMI	Modified Sample Matrix Inversion
MTS	Moving Target Simulator
NGA	National Geospatial-Intelligence Agency
NIMA	National Imagery and Mapping Agency
PPM	Preferred Performance Measure
PRF	Pulse Repetition Frequency
SDGC	Secondary Data Guard Cells
SPEAR	Signal Processing Evaluation Analysis and Research

STAP	Space–Time Adaptive Processing
USGS	United States Geological Survey

CHAPTER 7

CFAR	Constant False-Alarm Rate
JDL-GLR	Joint-Domain Localized General Likelihood Ratio
PDF	Probability Density Function

CHAPTER 8

SECTION 8.1

CPI	Coherent Processing Interval
GLR	Generalized Likelihood Ratio
IID	Independent and Identically Distributed
SMI	Sample-Matrix Inversion
S-T	Space–Time
T-S	Time–Space

SECTION 8.2

PRF	Pulse Repetition Frequency

SECTION 8.3

CFAR	Constant False Alarm Rate
INR	Clutter/Interference-to-Noise-Ratio
SINR	Signal-to-Clutter/Interference-plus-Noise-Ratio
S-T	Space–Time
T-S	Time–Space

SECTION 8.4

CFAR	Constant False Alarm Rate
DFT	Discrete Fourier Transform
GLR	Generalized Likelihood Ratio
IID	Independent and Identically Distributed
INR	Clutter/Interference-to-Noise Ratio
JDL-GLR	Joint-Domain Localized GLR
MTD	Moving Target Detector
MTI	Moving Target Indicator
PSD	Pulse Spectrum Density
SMI	Sample-Matrix Inversion
SNR	Signal-to-Noise Ratio

S-T	Space–Time
T-S	Time–Space

SECTION 8.5

CFAR	Constant False Alarm Rate
GLR	Generalized Likelihood Ratio
IID	Independent and Identically Distributed
JDL-GLR	Joint-Domain Localized GLR

CHAPTER 10

SECTION 10.1

DFT	Discrete Fourier Transform
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
IDFT	Inverse Discrete Fourier Transform

SECTION 10.2

DPCA	Displaced Phase Center Antenna
MTI	Moving Target Indicator
PRF	Pulse Repetition Frequency

SECTION 10.3

PRF	Pulse Repetition Frequency

SECTION 10.4

JNR	Jammer-plus-Noise-to-Ratio
PRF	Pulse-Repetition Frequency
PRI	Pulse-Repetition Interval
SIR	Signal-to-Interference Ratio
UHF	Ultra High Frequency

SECTION 10.5

FIR	Finite Impulse Response

SECTION 10.6

FFT	Fast Fourier Transform
MCARM	Multichannel Airborne Radar Measurements

MTI	Moving Target Indicator
PRI	Pulse Repetition Interval

SECTION 10.7

STAP	Space–Time Adaptive Processing

SECTION 10.8

DPCA	Displaced Phase Center Antenna
FIR	Finite Impulse Response

SECTION 10.9

SAR	Synthetic Aperture Radar

SECTION 10.10

A/D	Analog-to-Digital
FIR	Finite Impulse Response
GPS	Global Positioning System
FFT	Fast Fourier Transform
IFFT	Inverse Fast Fourier Transform
STAP	Space–Time Adaptive Processing

SECTION 10.11

ADC	Analog-to-Digital convertor
AOA	Angle of Arrival
CFAR	Constant False Alarm Rate
DBF	Digital Beam Forming
INR	Interference-to-Noise Ratio
ML	Maximum Likelihood
MLBP	Maximum Likelihood Beamspace Processor
PC	Pulse Compression
PRF	Pulse Repetition Frequency
SIR	Signal-to-Interference Ratio
SNR	Signal-to-Noise Ratio

CHAPTER 11

LOD	Locally Optimum Detector
PDF	Probability Density Function
SIRP	Spherically Invariant Random Process
SIRV	Spherically Invariant Random Vector

CHAPTER 12

SECTION 12.1

PDF	Probability Density Function

SECTION 12.2

CLT	Central Limit Theorem
MMSE	Minimum Mean Square Error
PDF	Probability Density Function
SIRP	Spherically Invariant Random Process
SIRV	Spherically Invariant Random Vector
SSRV	Spherically Symmetric Random Vector

SECTION 12.3

PDF	Probability Density Function
SIRP	Spherically Invariant Random Process
SIRV	Spherically Invariant Random Vector
SSRV	Spherically Symmetric Random Vector

SECTION 12.4

IMSL	International Mathematical and Statistical Libraries
PDF	Probability Density Function
SIRP	Spherically Invariant Random Process
SIRV	Spherically Invariant Random Vector
ZMNL	Zero Memory NonLinear

SECTION 12.5

PDF	Probability Density Function
SIRP	Spherically Invariant Random Process
SIRV	Spherically Invariant Random Vector

SECTION 12.6

SIRP	Spherically Invariant Process
SIRV	Spherically Invariant Random Vector

CHAPTER 13

SECTION 13.1

LOD	Locally Optimum Detector
LRT	Likelihood Ratio Test

PDF	Probability Density Function
SCNR	Signal to Clutter plus Noise Ratio
SIRP	Spherically Invariant Random Process
UMP	Uniformly Most Powerful

SECTION 13.2

ARE	Asymptotically Relative Efficiency
CPFSK	Continuous Phase Frequency Shift Keying
LOBD	Locally Optimum Binary Detector
LOD	Locally Optimum Detector
LRT	Likelihood Ratio Test
PDF	Probability Density Function
SIRP	Spherically Invariant Random Process
SIRV	Spherically Invariant Random Vector
UMP	Uniformly Most Powerful

SECTION 13.3

GPD	Generalized Pareto Distributionr
LOD	Locally Optimum Detector
OSLS	Ordered Sample Least Squares
PDF	Probability Density Function
UMP	Uniformly Most Powerful

SECTION 13.4

IMSL	International Mathematical and Statistical Libraries
LOD	Locally Optimum Detector
PDF	Probability Density Function
SCR	Signal to Clutter Ratio
SIRP	Spherically Invariant Random Process
SIRV	Spherically Invariant Random Vector
UMP	Uniformly Most Powerful

SECTION 13.5

ALOD	Amplitude Dependent Locally Optimum Detector
LOD	Locally Optimum Detector
SCR	Signal to Clutter Ratio
UMP	Uniformly Most Powerful

SECTION 13.6

ALOD	Amplitude Dependent Locally Optimum Detector
LOD	Locally Optimum Detector

GPD	Generalized Pareto Distributionr
SIRP	Spherically Invariant Random Process

CHAPTER 14

SECTION 14.1 TO SECTION 14.9

SIRV	Spherically Invariant Random Vector

SECTION 14.7

SNR	Signal-to-Noise Ratio

SECTION 14.8

SBR	Signal-to-Background Ratio

CHAPTER 15

SECTION 15.1

CFAR	Constant False-Alarm Rate
GLRT	Generalized Likelihood Ratio Test
LRT	Likelihood Ratio Test
MLE	Maximum Likelihood Estimate
PD	Probability of Detection
PFA	Probability of False Alarm
ROC	Receiver Operating Characteristic
SIRV	Spherically Invariant Random Vector
SN	Statistical Normalizer (or Statistical Normalization)

SECTION 15.2

GLRT	Generalized Likelihood Ratio Test
ML	Maximum Likelihood
NP	Neyman–Pearson
PDF	Probability Density Function
SCR	Signal-to-Clutter Ratio
SIRV	Spherically Invariant Random Vector

SECTION 15.3

CPI	Coherent Processing Interval
GLRT	Generalized Likelihood Ratio Test

ML	Maximum Likelihood
NP	Neyman–Pearson
PD	Probability of Detection
PDF	Probability Density Function
SCR	Signal-to-Clutter Ratio
SIRV	Spherically Invariant Random Vector

SECTION 15.4

EM	Expectation-Maximization
GEV	Generalized Extreme Value
GLRT	Generalized Likelihood Ratio Test
IID	Independent, Identically Distributed
LOD	Locally Optimum Detector
IS	Importance Sampling
ML	Maximum Likelihood
PD	Probability of Detection
PDF	Probability Density Function
RMS	Root-Mean-Squared
SAR	Synthetic Aperture Radar
SIRV	Spherically Invariant Random Vector

SECTION 15.5

CDF	Cumulative Distribution Function
PDF	Probability Distribution Function
SIRV	Spherically Invariant Random Vector

SECTION 15.6

CPI	Coherent Processing Interval
DGM	Discrete Gaussian Mixture
GLRT	Generalized Likelihood Ratio Test
IID	Independent Identically Distributed
LOD	Locally Optimum Detector
LRT	Likelihood Ratio Test
ML	Maximum-Likelihood
MLE	Maximum Likelihood Estimate
MLMF	Maximum-Likelihood Matched Filter
NP	Neyman-Pearson
PDF	Probability Density Function
SCR	Signal-to-Clutter Ratio
SIRV	Spherically Invariant Random Vector
UMP	Uniformly Most Powerful

SECTION 15.7

AR	Auto Regressive
AR(2)	Auto Regressive of order Two
IBDA	Innovations-Based Detection Algorithm
IID	Independent Identically Distributed
MMSE	Minimum Mean Square Error
PDF	Probability Density Function
SIRP	Spherically Invariant Random Process
SIRV	Spherically Invariant Random Vector
SNR	Signal-to-Noise ratio
SWNR	Signal-to-White-Noise Ratio
ZMNL	Zero Memory NonLinear

References

PREFACE

1. Weiner, M.M., *Monopole Antennas*, Marcel Dekker, New York, pp. 310–318, 2003.
2. Howell, P.W., Intermediate Frequency Sidelobe Canceller. U.S. Patent 3,202,990, August 24, 1965.
3. Appelbaum, S.P., Adaptive Arrays. Syracuse University Research Corporation, Report SU-SEL-66-12, Technical Report 6764-6, December 1966.
4. Widrow, B., Martey, P.E., Griffiths, L.J., and Goode, B.S., Adaptive Antenna Systems, *Proc. IEEE*, 55, 2143–2159, 1967.
5. Middleton, D., *Introduction to Statistical Communication Theory*, N.Y.: McGraw-Hill, Section 19.4, 1960.
6. Ozturk, A. A New Method for Univariate and Multivariate Distribution Identification. *J. Am. Stat. Assoc.*, submitted but not accepted for publication, 1990.
7. Ozturk, A., A General Algorithm for Univariate and Multivariate Goodness of Fit Test Based on Graphical Representation, *Commun. in Statistics, Part A–Theory and*, Methods-20(10), 3111–3137, 1991.
8. Rangaswamy, M., Weiner, D.D., and Ozturk, A., Simulation of correlated non-Gaussian interference for radar signal detection, *Proceedings of the 25th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA., November, 1991.
9. Ozturk, A., Rangaswamy, M., and Weiner, D.D., Distribution identification of spherically invariant vectors, *Proceedings of the 26th Conference on Information Sciences and System*, Princeton, NJ, March, 1992.
10. Rangaswamy, M., Chakravarthi P., Weiner, D.D., Cai, L., Wang, H. Ozturk, A., Signal Detection in Correlated Gaussian and Non-Gaussian Radar. Kaman Sciences Corporation, Final Technical Report, Rome Laboratory, U.S. Air Force Materiel Command, RL-TR-93-79, May, 1993.
11. Slamani, M.A. and Weiner, D.D., Use of image processing to partition a radar surveillance volume into background noise and clutter patches, *Proc. Conference on Information Sciences and Systems*, Baltimore, Md., March, 24-26, 1993.
12. Rangaswamy, M., Modeling of Non-Gaussian Radar Clutter, *Invited talk, Sensor Center, The MITRE Corporation*, Sept., 10, 1993.
13. Weiner, D.D., Weak Signal Detection in a Non-Gaussian Environment Invited talk, *Sensor Center*, The MITRE Corporation, Oct., 12, 1993.

CHAPTER 2

1. Van Trees, H.L., *Detection, Estimation, and Modulation Theory, Part I*, Wiley, New York, 1968.
2. Rangaswamy, M., Weiner, D., and Ozturk, A., Spherically invariant random processes for modeling and distribution identification of non-Gaussian random vectors, *IEEE Trans. AES*, 3(July), 111–124, 1992.
3. Chakravarthi, P., Subclutter Visibility: The Problem of Weak Signal Detection, Ph.D. dissertation, Department of Electrical and Computer Engineering, Syracuse University, 1993.
4. Sekine, M., Musha, T., Tomita, Y., Hagisawa, T., Irabu, T., and Kiuchi, E., On Weibull distributed weather clutter, *IEEE Trans. AES*, 15, 824–828, 1979.
5. Jakeman, E., On the statistics of *K*-distributed noise, *J. Phys. A*, 13, 31–48, 1980.
6. Watts, S. and Ward, K.D., Spatial correlation in *K*-distributed sea clutter, *IEE Proc. F*, 134(6), 526–532, 1987.
7. Hawkes, C.W. and Haykin, S.S., Modeling of clutter for coherent pulsed radar, *IEEE Trans. IT*, 21, 703–707, 1975.
8. Farina, A., Russo, A., and Studer, F.A., Coherent radar detection in lognormal clutter, *IEE Proc. F*, 133(1), 39–54, 1986.
9. Conte, E. and Longo, M., On a coherent model for log-normal clutter, *IEE Proc. F*, 134(2), 198–201, 1987.
10. Watts, S., Radar detection prediction in *K*-distributed sea clutter and thermal noise, *IEEE Trans. AES*, 23, 40–45, 1987.
11. Rangaswamy, M., Spherically Invariant Random Processes for Modeling, Simulation and Distribution Identification of Correlated Non-Gaussian Radar Clutter, Ph.D. dissertation, Department of Electrical and Computer Engineering, Syracuse University, 1992.
12. Kingman, J.F.C., Random walks with spherical symmetry, *Acta Math.*, 109, 11–53, Stokholm, 1963.
13. Yao, K., A representation theorem and its applications to spherically invariant random processes, *IEEE Trans. IT*, 19, 600–608, 1973.
14. Goldman, J., Detection in the presence of spherically symmetric random processes, *IEEE Trans. IT*, 22, 52–58, 1976.
15. Brehm, H., Description of spherically invariant random processes by means of *G*-functions, *Springer Lect. Notes*, 969, 39–73, 1982.
16. Conte, E. and Longo, M., Characterization of radar clutter as a spherically invariant random process, *IEE Proc. F*, 134(2), 191–197, 1987.
17. Kassam, S.A., *Signal Detection in Non-Gaussian Noise*, Springer-Verlag, New York, 1988.
18. Chakravarthi, P., Weiner, D., and Ozturk, A., On determining the radar thresholds for non-Gaussian pulses from experimental data, *Proceedings of the 25th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA*, pp. 594–598, November 1990.
19. Wang, H. and Cai, L., On adaptive multiband signal detection with GLR algorithm, *IEEE Trans. AES*, 27(2) 1991.
20. Lesser, V., Nawab, H., Bhandaru, M., Carver, N., Cvetanovic, Z., Gallastegi, I., and Klassner, F., Integrated Signal Processing and Signal Understanding, University of Massachusetts Technical Report 91-34, December 1991.

21. Nawab, H. and Lesser, V., *Integrated Processing and Understanding of Signals in Symbolic and Knowledge-Based Signal Processing*, Oppenheim, A.V. and Nawab, H., eds., Prentice Hall, Englewood Cliffs, NJ, 1991.
22. Nawab, H. and Lesser, V., High-Level Adaptive Signal Processing, NAIC Final Report, Vol. 17, October 1989.
23. Chakravarthi, P. and Weiner, D., Performance of the Locally Optimum Detector in a Correlated Non-Gaussian Disturbance, *35th Midwest Symposium on Circuits & Systems*, Washington, DC, August 9–12, 1992.
24. Gonzalez, R. and Wintz, P., *Digital Image Processing*, 2nd ed., Addison Wisley Publishing, New York, 1987.
25. Hall, E., *Computer Image Processing and Recognition*, Academic Press, London, 1979.
26. Shah, R.R., A New Technique for Distribution Approximation of Random Data. Master's thesis, Department of Electrical and Computer Engineering, Syracuse University, December 1993.
27. Ozturk, A., A new method for distribution identification, *J. Am. Stat. Assoc.*, 1990, (submitted but not accepted).
28. Ozturk, A., A general algorithm for univariate and multivariate goodness of fit tests based on graphical representation, *Commun. Stat. A. Theor. Methods*, 20(10), 3111–3137, 1991.
29. Ozturk, A. and Dudewicz, E.J., A new statistical goodness of fit test based on graphical representation, *Biomedical J.*, 34, 403–427, 1992.
30. Slaski, L.K., An Introduction to Dr. Ozturk's Algorithm for PDF Approximation, Rome Laboratory Internal Report, USAF March 1993.
31. Chair, Z. and Varshney, P.K., Optimal data fusion in multiple sensor detection systems, *IEEE Trans. AES*, 22(January), 98–101, 1998.

CHAPTER 3

SECTION 3.1

1. Chen, P., On Testing the Equality of Covariance Matrices Under Singularity, Rome Laboratory Technical Report, AFOSR Summer Faculty Research Program, August 1994.
2. Chen, P., Partitioning Procedure in Radar Signal Processing Problems, Rome Laboratory Technical Report, AFOSR Summer Faculty Research Program, August 1995.
3. Melvin, W., Wicks, M., and Brown, R., Assessment of multichannel airborne radar measurements for analysis and design of space–time adaptive processing architectures and algorithms, *Proceedings of the IEEE National Radar Conference,* Ann Arbor, MI, 1996.
4. Melvin, W. and Wicks, M., Improving practical space–time adaptive radar, *Proceedings of the IEEE National Radar Conference,* Syracuse, NY, 1997.
5. Himed, B., Salama, Y., and Michels, J.H., Improved detection of close proximity targets using two-step NHD, *Proceedings of the Inter-National Radar Conference,* Alexandria, VA, 2000.
6. Nitzberg, R., An effect of range-heterogeneous clutter on adaptive Doppler filters, *IEEE Trans. AES*, 26(3), 475–480, 1990.

7. Melvin, W.L., Guerci, J.R., Callahan, M.J., and Wicks, M.C., Design of Adaptive Detection Algorithms for Surveillance Radar, *Proceedings of the International Radar Conference*, Alexandria, VA, 2000.
8. Melvin, W.L., Space-time radar performance in heterogeneous clutter, *IEEE Trans. AES*, 36(3), 621–633, 2000.
9. Rangaswamy, M., Weiner, D., and Ozturk, A., Non-Gaussian random vector identification using spherically invariant random processes, *IEEE Trans. AES*, 29, 111–124, 1993.
10. Chen, P., Melvin, W., and Wicks, M., Screening among multivariate normal data, *J. Multivar. Anal.*, 69, 10–29, 1999.
11. Rangaswamy, M., Himed, B., and Michels, J., Statistical analysis of the nonhomogeneity detector, *Proceedings of the 34th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, 2000.
12. Cai, L. and Wang, H., On adaptive filtering with CFAR feature and its performance sensitivity of non-Gaussian interference, *Proceedings of the 24th Annual Conference on Information Sciences and Systems*, Princeton, NJ, pp.558–563, 1990.
13. Robey, F., Fuhrmann, D., Kelly, E., and Nitzberg, R., A CFAR adaptive matched filter detector, *IEEE Trans. AES*, 28, 208–216, 1992.
14. Chen, W.S. and Reed, I.S., A new CFAR detection test for radar, *Digit. Signal Process.*, October, 198–214, 1991.
15. MCARMDATA, View www@http://128.132.2.229, Data from the Multichannel Airborne Radar Measurement Program of the U.S. Air Force Research Laboratory, Rome, NY.
16. Kraut, S., Scharf, L.L., and McWhorter, L.T., Adaptive subspace detectors, *IEEE Trans. SP*, 49, 1–16, 2001.
17. Roman, J.R., Rangaswamy, M., Davis, D., Zhang, Q., Himed, B., and Michels, J.H., Parametric adaptive matched filter for airborne radar applications, *IEEE Trans. AES*, 36, 677–692, 2000.
18. Michels, J.H., Himed, B., and Rangaswamy, M., Performance of STAP tests in Gaussian and compound-Gaussian clutter, *Digit. Signal Process.*, 10(4), 309–324, 2000.
19. Steiner, M. and Gerlach, K., Fast converging adaptive processor for a structured covariance matrix, *IEEE Trans. AES*, 36, 1115–1126, 2000.
20. Anderson, T.W., *An Introduction to Multivariate Statistical Analysis*, Wiley, New York, 1958.
21. Siotani, M., Hayakawa, T., and Fujikoshi, Y., *Modern Multivariate Statistical Analysis*, American Science, Ohio, 1985.
22. Abramovich, Y.I., Analysis of a direct adaptive tuning method for interference compensation systems with auxiliary linear constraints, *Sov. J. Commun. Technol. Electron.*, 35(1), 30–37, 1990, English translation of Radioteknika Elektronika.
23. Abramovich, Y.I., Mikhaylyukov, V.N., and Malyavin, I.P., Test of interference stationarity in adaptive filtering systems, *Sov. J. Commun. Technol. Electron.*, 37(3), 1–10, 1992, English translation of Radioteknika Electronika.
24. Abramovich, Y.I., Convergence analysis of linearly constrained {SMI}, and {LSMI} adaptive algorithms, *Proceedings of ASSPCC*-2000, Lake Louise, Canada, pp. 255–259, 2000.

SECTION 3.2

1. Chen, P., On Testing the Equality of Covariance Matrices Under Singularity, Rome Laboratory Technical Report, AFOSR Summer Faculty Research Program, August 1994.
2. Chen, P., Partitioning Procedure in Radar Signal Processing Problems, Rome Laboratory Technical Report, AFOSR Summer Faculty Research Program, August 1995.
3. Melvin, W., Wicks, M., and Brown, R., Assessment of multichannel airborne radar measurements for analysis and design of space–time adaptive processing architectures and algorithms. *Proceedings of the IEEE National Radar Conference*, Ann Arbor, MI, 1996.
4. Melvin, W. and Wicks, M., Improving practical space-time adaptive radar, *Proceedings of the IEEE National Radar Conference*, Syracuse, NY, 1997.
5. Melvin, W.L., Space-time radar performance in heterogeneous clutter, *IEEE Trans. AES*, 36(3), 621–633, 2000.
6. Nitzberg, R., An effect of range-heterogeneous clutter on adaptive Doppler filters, *IEEE Trans. AES*, 26(3), 475–480, 1990.
7. Melvin, W.L., Guerci, J.R., Callahan, M.J., and Wicks, M.C., Design of adaptive detection algorithms for surveillance radar, *Proceedings of the International Radar Conference*, Alexandria, VA, 2000.
8. Himed, B., Salama, Y., and Michels, J.H., Improved Detection of Close Proximity Targets Using Two-Step NHD, *Proceedings of the Inter-National Radar Conference*, Alexandria, VA, 2000.
9. Chen, P., Melvin, W., and Wicks, M., Screening among multivariate normal data, *J. Multivar. Anal.*, 69, 10–29, 1999.
10. Rangaswamy, M., Himed, B., and Michels, J., Statistical analysis of the nonhomogeneity detector, *Proceedings of the 34th Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, 2000.
11. Rangaswamy, M., Himed, B., and Michels, J., Performance analysis of the nonhomogeneity detector for STAP applications, *Proceedings of the 2001 IEEE Radar Conference*, Atlanta, GA, May 2001.
12. Rangaswamy, M., Weiner, D., and Ozturk, A., Non-Gaussian random vector identification using spherically invariant random processes, *IEEE Trans. AES*, 29, 111–124, 1993.
13. Rangaswamy, M., Weiner, D.D., and Ozturk, A., Computer generation of correlated non-Gaussian radar clutter, *IEEE Trans. AES*, 31, 106–116, 1995.
14. Sangston, K. and Gerlach, K., Coherent detection of targets in a non-Gaussian background, *IEEE Trans. AES*, 30, 330–340, 1994.
15. Rangaswamy, M., Lin, F., and Gerlach, K., Robust adaptive signal processing methods for heterogeneous radar clutter scenarios, *Proceedings of IEEE Radar Conference*, Huntsville, AL, May 2003.
16. Wang, Y.L., Chen, J.W., Bao, Z., and Peng, Y.N., Robust space-time adaptive processing for airborne radar in nonhomogeneous clutter environments, *IEEE Trans. AES*, 39(1), 70–81, 2003.
17. Guerci, J. and Bergin, J., Principal components, covariance matrix tapers, and the subspace leakage problem, *IEEE Trans. AES*, 38(1), 152–162, 2002.
18. Gerlach, K.R., Outlier resistant adaptive matched filtering, *IEEE Trans. AES*, 38(3), 885–901, 2002.

19. Vorobyov, S., Gershman, A., and Luo, Z., Robust adaptive beamforming using worst-case performance optimization: a solution to the signal mismatch problem, *IEEE Trans. SP*, 51, 313–324, 2003.
20. Sangston, K. and Gerlach, K., Non-Gaussian Noise Models and Coherent Detection of Radar Targets, Technical Report, NRL, 1992.
21. Rangaswamy, M., Michels, J., and Weiner, D., Multichannel detection for correlated non-Gaussian random processes based on innovations, *IEEE Trans. SP*, 43, 1915–1922, 1995.
22. Yao, K., A representation theorem and its applications to spherically invariant random processes, *IEEE Trans. IT*, 19, 600–608, 1973.
23. Pulsone, N., Adaptive Signal Detection In Non-Gaussian Interference, Ph.D. thesis, Northeastern University, 1997.
24. Rangaswamy, M., Michels, J.H., and Himed, B., Statistical analysis of the nonhomogeneity detector for STAP applications, *Digit. Signal Process.*, 14(1), 2004.
25. Kraut, S., Scharf, L.L., and McWhorter, L., Adaptive subspace detectors, *IEEE Trans. SP*, 49, 1–16, 2001.
26. Michels, J., Himed, B., and Rangaswamy, M., Performance of STAP tests in Gaussian and compound-Gaussian clutter, *Digit. Signal Process.*, 10(4), 309–324, 2000.
27. Michels, J., Rangaswamy, M., and Himed, B., Performance of parametric and covariance based STAP tests in compound-Gaussian clutter, *Digit. Signal Process.*, 12(2,3), 307–328, 2002.
28. Little, R.J. and Rubin, D.B., *Statistical Analysis with Missing Data*, Wiley, New York, 1987.
29. Rangaswamy, M. and Michels, J.H., Adaptive processing in non-Gaussian noise backgrounds, *Proceedings of the Ninth IEEE Workshop on Statistical Signal and Array Processing*, Portland, OR, 1998.
30. Conte, E., DeMaio, A., and Ricci, G., Recursive estimation of the covariance matrix of a compound-Gaussian process and its application to adaptive CFAR, *IEEE Trans. SP*, 50, 1908–1916, 2002.
31. Gini, F. and Greco, M., Covariance matrix estimation for CFAR detection in correlated heavy-tailed clutter, *Signal Process.*, 82(12), 1847–1859, 2002.
32. Gini, F. and Greco, M., Suboptimum approach to adaptive coherent radar detection in compound-Gaussian clutter, *IEEE Trans. AES*, 35(3), 1095–1103, 1999.
33. Gini, F. and Michels, J.H., Performance analysis of two covariance matrix estimators in compound Gaussian clutter, *IEE Proc. F*, 146(3), 133–140, 1999.
34. Michels, J., Rangaswamy, M., and Himed, B., Performance of STAP tests in compound gaussian clutter, *Proceedings of the First IEEE Workshop on Sensor Array and Multichannel Processing (SAM-2000)*, Cambridge, MA, 2000.
35. Jakeman, E. and Pusey, P., A model for non-Rayleigh sea echo, *IEEE Trans. AP*, 24, 806–814, 1976.
36. Jao, J., Amplitude distribution of composite terrain radar clutter and the *K*-distribution, *IEEE Trans. AP*, 32, 1049–1062, 1984.
37. Michels, J., Tsao, T., Himed, B., and Rangaswamy, M., Space-time adaptive processing (STAP) in airborne radar applications, *Proceedings of the IASTED Conference on Signal Processing and Communications*, Canary Islands, Spain, 1998.

38. MCARMDATA, View www@http://128.132.2.229, Data from the Multichannel Airborne Radar Measurement Program of the U.S. Air Force Research Laboratory, Rome, NY.
39. Himed, B. and Melvin, W., Analyzing Space-time adaptive processors using measured data, *Proceedings of the 31st Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, 1997.
40. Lin, X. and Blum, R., Robust STAP algorithms using prior knowledge for airborne radar application, *Signal Process.*, 79, 273–287, 1999.
41. Michels, J., Rangaswamy, M., and Himel, B., Robust multichannel detection in heterogeneous airborne radar disturbance, *Proceedings IEEE Radar Conference*, Long Beach, CA, April 2002.
42. Burg, J.P., Luenberger, D.G., and Wenger, D.L., Estimation of structured covariance matrices, *Proc. IEEE*, 70(9), 963–974, 1982.

CHAPTER 4

1. Ozturk, A., A new method for univariate and multivariate distribution identification, *J. Am. Stat. Assoc.*, 1990, (submitted but not accepted).
2. Ozturk, A., A general algorithm for univariate and multivariate goodness of fit tests based on graphical representation, *Commn. Stat. A-Theor. Method.*, 20(10), 3111–3137, 1991.
3. Ozturk, A. and Dudewicz, E.J., A new statistical goodness of fit test based on graphical representation, *Biom. J.*, 34, 403–427, 1992.
4. Slifker, J. and Shapiro, S., Johnson system: selection and parameter estimation, *Technometrics*, 22(May), 239–246, 1980.
5. Aly, E.E.A.A. and Ozturk, A., Hodges Lehman quantile–quantile plots, *Comput. Stat. Data Anal.*, 6, 99–108, 1988.
6. Wilk, M.B. and Gnanadesikan, R., Probability plotting methods for the analysis of data, *Biometrika*, 55, 1–8, 1968.
7. Small, N.J.H., Plotting squared radii, *Biometrika*, 65, 657–658, 1978.
8. Filliben, J.J., The probability plot correlation coefficient test for normality, *Technometrics*, 17(February), 111–112, 1975.
9. Massey, F.J., The Kolmogorov–Smironov test for goodness of fit, *J. Am. Stat. Assoc.*, 46, 68–78, 1951.
10. Slaski, L.K., An Introduction to Dr. Ozturk's Algorithm for PDF Approximation, Rome Laboratory Internal Report, USAF, March 1993.
11. Rangaswamy, M., Spherically Invariant Random Processes for Radar Clutter Modeling, Simulation, and Distribution Identification, Ph.D. dissertation, Department of Electrical and Computer Engineering, Syracuse University, December 1992.
12. Lindgren, B.W. and McElrath, G.W., *Introduction to Probability and Statistics*, 4th ed., Macmillan Publishing Co., New York, 1978.
13. Lindgren, B.W. and McElrath, G.W., *Introduction to Probability and Statistics*, 3rd ed., Macmillan Publishing Co., New York, 1969.
14. Johnson, R.A. and Wichern, D.W., *Applied Multivariate Statistical Analysis*, 3rd ed., Prentice Hall, New Jersey, 1992.
15. Johnson, N. and Kotz, S., *Distributions in Statistics: Continuous Univariate Distributions*, Wiley, New York, 1970.

16. Johnson, N. and Kotz, S., *Distributions in Statistics: Continuous Multivariate Distributions*, Wiley, New York, 1976.
17. Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, McGraw Hill, New York, 1984.
18. Bratley, P., Fox, B., and Schrage, L., *A Guide to Simulation*, Springer Verlag, New York, 1987.
19. Johnson, M.E., *Multivariate Statistical Simulation*, Wiley, New York, 1987.

CHAPTER 5

SECTION 5.1

1. Wilk, M.B. and Gnanadesikan, R., Probability plotting methods for the analysis of Data, *Biometrika*, 55, 1–17, 1968.
2. Easton, G.S. and McCulloch, R.H., A multivariate generalization of quantile–quantile plots, *J. Am. Stat. Assoc.*, 85(410), 376–386, 1990.
3. Johnson, N.L. and Kotz, S., *Distributions in Statistics: Continuous Univariate Distributions*, Wiley, New York, 1970.
4. Ozturk, A., A general algorithm for univariate and multivariate goodness-of-fit test based on graphical representation, *Commun. Stat. A. Theor. Methods*, 20(10), 3111–3137, 1991.
5. Ozturk, A. and Dudewicz, E.J., A new statistical goodness-of-fit test based on graphical representation, *Biom. J.*, 34, 403–427, 1992.
6. Gnanadesikan, R., *Methods for Statistical Data Analysis of Multivariate Observations*, Wiley, New York, 1977.
7. Hawkins, D.M., A new test for multivariate normality and homoscedasticity, *Techonometrics*, 23, 105–110, 1981.
8. Koziol, J., A class of invariant procedures for assessing multivariate normality, *Biometrika*, 69, 423–427, 1982.
9. Gnanadesikan, R. and Kettenring, J.R., Robust estimates, residuals, and outlier detection with multiresponse data, *Biometrika*, 59, 81–124, 1972.
10. Johnson, M.E., *Multivariate Statistical Simulation*, Wiley, New York, 1987.
11. Preston, E.J., A graphical method for the analysis of statistical distributions into two normal components, *Biometrika*, 40, 460–464, 1953.
12. Bowman, K.O. and Shenton, H.L., Simulation and estimation problems associated with the 3-parameter gamma distribution, *Commun. Stat. Simul.*, 16(4), 1147–1188, 1987.
13. Fisher, R.A., The use of multiple measurements in taxonomic problems, *Ann. Eugen.*, 8, 179–188, 1936.
14. Royston, J.P., Some techniques for assessing multivariate normality based on Shapiro-Wilk W, *Appl. Stat.*, 32, 121–133, 1983.
15. Small, N.J.H., Marginal skewness and kurtosis in testing multivariate normality, *Appl. Stat.*, 29, 85–87, 1980.

SECTION 5.2

1. Ozturk, A. and Dudewicz, E.J., A New Statistical Goodness-of-Fit Test Based on Graphical Representation, Technical Report, No. 52, Department of Mathematics, Syracuse University, NY, 1990.

2. Shapiro, S.S. and Gross, A.J., *Statistical Modeling Techniques*, Marcel Dekker, New York, p. 227, 1981.
3. Lilliefors, H.F., On the Kolmogorov–Smirnov tests for normality with mean and variance unknown, *J. Am. Stat. Assoc.*, 62, 399–402, 1967.
4. Massey, F.J., The Kolmogorov–Smirnov test for goodness-of-fit, *J. Am. Stat. Assoc.*, 46, 68–78, 1951.
5. Kac, M., Kiefer, J., and Wolfovitz, J., On tests of normality and other tests of fit based on distance method, *Am. Math. Soc.*, 25, 189–198, 1955.
6. Wilk, M.B. and Gnanadesikan, R., Probability plotting methods for the analysis of data, *Biometrika*, 55, 1968.
7. Shapiro, S.S. and Wilk, M.B., An analysis of variance test for the exponential distribution, *Technometrics*, 14, 239–246, 1980.
8. D'Agostino, R.B. and Pearson, E.S., Tests for departures from normality: empirical results for the distributions of b_2 and $(b_1)^{1/2}$, *Biometrika*, 60, 613–622, 1973.
9. Shapiro, S.S. and Wilk, M.B., An analysis of variance test for normality, *Biometrika*, 52, 591–611, 1965.
10. Ozturk, A., On the *W* test for the extreme value distribution, *Biometrika*, 73, 738–740, 1986.
11. Ozturk, A. and Korukoglu, S., A new test for the extreme value distribution, *Commun. Stat.*, 17, 1375–1393, 1988.
12. Oturk, A., Korukoglu, S., and Okur, M.C., A goodness-of-fit test for the exponential distribution, *J. Comput. Stat.*, 2, 21–37, 1990.
13. Stephens, M.A., Tests based on EDF statistics, In *Goodness-of-Fit Techniques*, D'Agostino, R.B. and Stephens, M.A. eds., Marcel Decker, New York, 1986a, pp. 97–103.
14. Stephens, M.A., Tests based on regression and correlation, In *Goodness-of-Fit Techniques*, D'Agostino, R.B. and Stephens, M.A. eds., Marcel Decker, New York, 1986b, pp. 195–233.
15. Shapiro, S.S. and Francia, R.S., Approximate analysis of variance test for normality, *J. Am. Stat. Assoc.*, 67, 215–216, 1972.
16. Weisberg, S. and Bingham, C., An approximate analysis of variance test for non-normality suitable for machine calculation, *Technometrics*, 17, 133–134, 1975.
17. Filliben, J.J., The probability plot correlation coefficient test for normality, *Technometrics*, 17, 111–117, 1975.
18. Ozturk, A., On a test statistic for testing normality, *Comput. Stat. Data Anal.*, 6, 1–8, 1987.
19. Aly, E.E.A.A. and Ozturk, A., Hodges-Lehman quantile-quantile plots, *Comput. Stat. Data Anal.*, 6, 99–108, 1988.
20. Aly, E.E.A.A. and Ozturk, A., A modified one-sample Q–Q plot and a test for normality, *J. Stat. Comput. Simul.*, 29, 1–5, 1988.
21. D'Agostino, R.B., An omnibus test of normality for moderate and large size samples, *Biometrika*, 58, 341–348, 1971.
22. Slifker, J.F. and Shapiro, S.S., The Johnson system: selection and parameter estimation, *Technometrics*, 22, 355–370, 1980.
23. Cox, D.R. and Small, N.J.H., Testing multivariate normality, *Biometrika*, 65, 263–272, 1978.
24. Gnanadesikan, R., *Methods for Statistical Data Analysis*, Wiley, New York, pp. 888–999, 1977.

25. Gnanadesikan, R., Robust estimates, residuals, and outlier detection with multiresponse data, *Biometrics*, 28, 81–124, 1972.
26. Small, N.J.H., Plotting squared radii, *Biometrika*, 65, 657–658, 1978.
27. Andrews, D.F., Gnanadesikan, R., and Warner, J.L., *Methods for Assessing Multivariate Normality in Multivariate Analysis III*, Krisnaiah, P.R., ed., Academic Press, New York, pp. 95–116, 1973.
28. Fisher, R.A., The use of multiple measurements in taxonomic problems, *Ann. Eugen.*, 7, 179–188, 1936.
29. Kendall, M.G., *Multivariate Analysis*, Griffin, London, 1975.
30. Johnson, N.L. and Kotz, S., *Distributions in Statistics*, *Continuous Univariate Distributions*, Vol. 1, Wiley, New York, pp. 187–190, 1970.
31. Johnson, N.L. and Kotz, S., *Distributions in Statistics*, *Continuous Univariate Distributions*, Vol. 2, Wiley, New York, p. 87, 1970.
32. Koziol, J., A class of invariant procedures for assessing multivariate normality, *Biometrika*, 69, 423–427, 1982.
33. Small, N.J.H., Marginal skewness and kurtosis in testing multivariate normality, *Appl. Stat.*, 29, 85–87, 1980.
34. Royston, J.P., Some techniques for assessing multivariate normality based on the Shapiro-Wilk W, *Appl. Stat.*, 32, 121–133, 1983.
35. Fowkes, E.B. and Folio, *A Folio of Distributions and Collection of Theoretical Quantile-Quantile Plots*, Marcel Dekker, New York, 1990.
36. Gan, F.F. and Koether, K.J., Goodness-of-fit tests based on P–P probability plots, *Technometrics*, 32(3), 289–303, 1990.

CHAPTER 6

SECTION 6.1

1. Ozturk, A., A general algorithm for univariate and multivariate goodness of fit tests based on graphical representation, *Commun. Stat. A. Theo. Methods*, 20(10), 3111–3137, 1991.
2. Shah, R.R., A New Technique for Distribution Approximation of Random Data, Master of Science thesis, Department of Electrical and Computer Engineering, Syracuse University, Syracuse, NY.
3. Hastings, N.A. and Peacock, J.B., *Statistical Distributions*, Wiley, New York, 1975.
4. Ozturk, A. and Dudewicz, E.J., A new statistical goodness of fit test based on graphical representation, *Biom. J.*, 34, 403–427, 1992.
5. Liberman, C.M., Auditory-nerve response tom cats raised in a low-noise chamber, *J. Acoust. Soc. Am.*, 63, 442–455, 1978.
6. Gaumond, R.P., Molnar, R.P., and Kim, D.O., Stimulus and recovery dependence of cat cochlear nerve fiber spike discharge probability, *J. Neurophysiol.*, 48, 856–873, 1982.
7. Gaumond, R.P., Kim, D.O., and Molnar, C.E., Response of cochlear nerve fibers to brief acoustic stimuli: role of discharge-history effects, *J. Acoust. Soc. Am.*, 74, 1392–1398, 1983.
8. Relkin, E.M. and Pelli, D.G., Probe tone thresholds in the auditory nerve measured by two-interval forced-choice procedures, *J. Acoust. Soc. Am.*, 82, 1679–1691, 1987.

9. Barlow, R.B. Jr., Chamberlain, S.C., and Levinson, J.Z., *Limulus* brain modulates structure and function of the lateral eyes, *Science*, 210, 1037–1039, 1980.
10. Kass, L. and Barlow, R.B. Jr., Efferent neurotransmission of circadian rhythms in the *Limulus* lateral eye: I. octopamine-induced changes in retinal sensitivity, *J. Neurosci.*, 4, 908–917, 1992.
11. Chamberlain, S.C. and Barlow, R.B. Jr., Transient membrane shedding in *Limulus* photoreceptors: control mechanisms under natural lighting, *J. Neurosci.*, 4, 2792–2810, 1984.
12. Chamberlain, S.C. and Barlow, R.B. Jr., Control of structural rhythms in the lateral eye of *Limulus*: interactions of natural lighting and circadian efferent activity, *J. Neurosci.*, 7, 2135–2144, 1987.
13. Barlow, R.B. Jr., Chamberlain, S.C., and Lehman, H.K., *Circadian Rhythms in the Intervertebrate Retina: Facets of Vision*, Stavenga, D. and Hardie, R., eds., Springer-Verlag, Berlin, 1989.
14. Kier, C.K. and Chamberlian, S.C., Dual controls for screening pigment movement in photoreceptors of the *Limulus* lateral eye: circadian efferent input and light, *Vis. Neurosci.*, 4, 237–255, 1990.
15. Barlow, R.B. Jr., Circadian rhythms in the *Limulus* visual system, *J. Neurosci.*, 3, 856–870, 1983.
16. Kass, L. and Barlow, R.B., A circadian clock in the *Limulus* brain transmits synchronous efferent signals to all eyes, *Vis. Neurosci.*, 9, 493–504, 1992.
17. Passaglia, C.L., Weiner, W.W., Raghavan, M., and Barlow, R.B. Jr., Circadian regulation of efferent optic nerve activity in *Limulus*, *Soc. Neurosci. Abstr.*, 18, 5118, 1992.
18. Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, 1984.
19. Weiner, W.W. and Chamberlain, S.C., The visual fields of American horseshoe crabs: two different eye shapes in *Limulus polyphemus*, *Vis. Neurosci.*, 11, 333–346, 1994.
20. Beyer, W.H., ed., *CRC Handbook of Tables for Provbability and Statistics*, Chemical Rubber Co., Cleveland, 1966.
21. Bowman, K.O. and Shenton, L.R., *Properties of Estimators for the Gamma Distribution*, Marcel Dekker, New York, 1988.
22. Fowlkes, E.B., *A Folio of Distributions*, Marcel Dekker, New York, 1987.
23. *Handbook of Mathematical, Scientific, and Engineering Formulas, Tables, Functions, Graphs, Transforms*, Staff of Research and Education Association, 1980.
24. Wells, H.G., *The Time Machine*, Bantam Classic Reissue, 1991 of his 1895 book, New York, 1991.

SECTION 6.2

1. Rangaswamy, M., Weiner, D., and Ozturk, A., Non-Gaussian random vector identification using spherically invariant random processes, *IEEE Trans. AES*, 29, 111–124, 1993.
2. Chakravathi, P., Subclutter Visibility: The Problem of Weak Signal Detection, Ph.D. dissertation, Department of Electrical and Computer Engineering, Syracuse University, 1992.
3. Conte, E. and Longo, M., Characterization of radar clutter as a spherically invariant random process, *IEE Proc. F*, 134(2), 91–197, 1987.

4. Gonzalez, R. and Wintz, P., *Digital Image Processing*, 2nd ed., Addison-Wesley Publishing Co., Reading, MA, 1987.
5. Hall, E., *Computer Image Processing and Recognition*, Academic Press, New York, 1979.

SECTION 6.3

1. Chakravathi, P., Subclutter Visibility: the Problem of Weak Signal Detection, Ph.D. dissertation, Department of Electrical and Computer Engineering, Syracuse University, 1992.
2. Slamani, M.A. and Weiner, D.D., Use of image processing to partition a radar surveillance volume into background noise and clutter patches, *Proceedings of the 1993 Conference on Information Sciences and Systems, Johns Hopkins University*, Baltimore, MD, March 24–26, 1993.

SECTION 6.4

1. Hawkes, C.W. and Haykin, S.S., Modeling of clutter for coherent pulsed radar, *IEEE Trans. IT*, 21, 703–707, 1975.
2. Slamani, M.A. and Weiner, D.D., Rationale behind the use of image processing to partition a radar surveillance volume into background noise and clutter patches, *Proc. 36th Midwest Symposium on Circuits and Systems*, Detroit, MI, Aug. 10–18, 1993.
3. Slamani, M.A. and Weiner, D.D., Use of image processing to partition a radar surveillance volume into background noise and clutter patches, *Proceedings of the 1993 Conference on Information Sciences and Systems, Johns Hopkins University*, Baltimore, MD, March 24–26, 1993.
4. Ozturk, A. and Dudewicz, E.J., A new statistical goodness of fit test based on graphical representation, *Biomedical J.*, 34, 403–427, 1992.
5. Nawab, H. and Lesser, V. Integrated processing and understanding of signals, In *Symbolic and Knowledge-Based Signal Processing*, Oppenheim, A.V. and Nawab, H., eds., Prentice Hall, Englewood Cliffs, NJ, 1992.
6. Slamani, M.A., A New Approach to Radar Detection Based on the Partitioning and Statistical Characterization of the Surveillance Volume, Ph.D. dissertation, Department of Electrical and Computer Engineering, Syracuse University, December 1994.
7. Sahoo, P.K., Soltami, S., Wong, A.K.C., and Chen, Y.C., A survey of thresholding techniques, *Comput. Vis. Graph. Image Process.*, 41, 233–260, 1988.
8. Gonzalez, R. and Wintz, P., *Digital Image Processing*, 2nd ed., Addison-Wesley Publishing Co., Reading, MA, 1987.
9. Hall, E., *Computer Image Processing and Recognition*, Academic Press, New York, 1979.

SECTION 6.5

1. Wang, H., Rangaswamy, M., Chakravarthi, P., Weiner, D., Cai, L., and Ozturk, A., Signal Detection in Correlated Gaussian and Non-Gaussian Radar Clutter, Final Technical Report RL-TR-93-79, Rome Lab, Griffiss AFB, New York, May 1993.

2. Chakravathi, P., Slamani, M.A., and Weiner, D., Performance of the locally optimum detector in a correlated *K*-distributed disturbance, *Proceedings of the 1993 National Radar Conference on Revolutionary Developments in Radar*, Wakefield, MA, April 20–23, 1993.
3. Slamani, M.A. and Weiner, D.D., Use of image processing to partition a radar surveillance volume into background noise and clutter patches, *Proceedings of the 1993 Conference on Information Sciences and Systems, Johns Hopkins University*, Baltimore, MD, March 24–26, 1993.
4. Slamani, M.A., A New Approach to Radar Detection Based on the Partitioning and Statistical Characterization of the Surveillance Volume, Ph.D. dissertation, Department of Electrical and Computer Engineering, Syracuse University, December 1994.
5. Slamani, M.A., Weiner, D.D., and Ozturk, A., New approach to the analysis of IR images, *Proceedings of the 1995 SPIE Conference on Signal and Data Processing of Small Targets*, San Diego, CA, July 11–13, 1995.
6. Ozturk, A. and Dudewicz, E.J., A new statistical goodness of fit test based on graphical representation, *Biomedical J.*, 34, 403–427, 1992.
7. Nawab, H. and Lesser, V. Integrated processing and understanding of signals, In *Symbolic and Knowledge-Based Signal Processing*, Oppenheim, A.V. and Nawab, H., eds., Prentice Hall, Englewood Cliffs, NJ, 1992.

SECTION 6.6

1. Slamani, M.A. and Weiner, D.D., Use of image processing to partition a radar surveillance volume into background noise and clutter patches, *Proceedings of the 1993 Conference on Information Sciences and Systems, Johns Hopkins University*, Baltimore, MD, March 24–26, 1993.
2. Slamani, M.A. and Weiner, D.D., ASCAPE: An automated approach to the partitioning and statistical characterization of a surveillance volume, *Proceedings of the Sixth International Conference on Signal Processing Applications and Technology*, Boston, MA, October 24–26, 1955.
3. Ozturk, A. and Dudewicz, E.J., A new statistical goodness-of-fit test based on graphical representation, *Biomedical J.*, 34, 403–427, 1992.
4. Ward, K.D., Baker, C.J., and Watts, S., Maritime surveillance radar, part I: radar scattering from the ocean surface, *IEE Proc. F*, 137(April), 51–62, 1990.
5. Baker, C.J., *K*-distributed coherent sea clutter, *IEE Proc. F*, 138(2), 89–92, 1991.
6. Jao, J.K., Amplitude distribution of composite terrain radar clutter and the *K*-distribution, *IEEE Trans. AP*, 32(10) 1049–1062, 1984.
7. Sekine, M., Ohtani, S., Musha, T., Irabu, T., Kiuchi, E., Hagisawa, T., and Tomita, Y., Weibull distributed ground clutter, *IEEE Trans. AES*, 17(4), 596–598, 1981.
8. Conte, E., Lops, M., and Ricci, G., Radar detection in *K*-distributed clutter, *IEEE Proc. Radar, Sonar Navig.*, 141(2) 1994.
9. Rangaswamy, M., Weiner, D., and Ozturk, A., Non-Gaussian random vector identification using spherically invariant random processes, *IEEE Trans. AES*, 29, 111–124, 1993.
10. Rangaswamy, M., Weiner, D., and Ozturk, A., Computer generation of correlated non-Gaussian radar clutter, *IEEE Trans. AES*, 31(January), 106–116, 1995.

11. Sangston, K.J. and Gerlach, K.R., Coherent detection of radar targets in a non-Gaussian background, *IEEE Trans. AES*, 30(April), 330–340, 1984.
12. Keckler, A.D., Non-Gaussian Clutter Modeling by Spherically Invariant Random Vectors, AFOSR Final Report, Rome Laboratory, Griffith AFB, Rome, NY, September, 1996.
13. Yao, K., A representation theorem and its applications to spherically invariant random processes, *IEEE Trans. IT*, 19, 600–608, 1973.
14. Stadelman, D.L., Ph.D. Part 2 Examination, Department of Electrical Engineering & Computer Science, Syracuse University, Syracuse NY, 1995.

SECTION 6.7

1. Antonik, P., Shuman, H., Li, P., Melvin, W., and Wicks, M., Knowledge-based space-time adaptive processing, *Proceedings of the IEEE 1997 National Radar Conference*, Syracuse, NY, May 1997.
2. Baldygo, W., Wicks, M., Brown, R., Antonik, P., Capraro, G., and Hennington, L., Artificial intelligence applications to constant false alarm rate (CFAR) Processing, *Proceedings of the IEEE 1993 National Radar Conference*, Boston, MA, April 1993.
3. Ward, J., Space-Time Adaptive Processing for Airborne Radar, MIT Lincoln Laboratory, Technical Report 1015, December 13, 1994.
4. Westinghouse Electronic Systems, Multi-Channel Airborne Radar Measurement (MCARM) Flight Report, Vol. 1 of 4, MCARM Flight Test, Contract F30602-92-C-0161 for Rome Laboratory, USAF.
5. Chen, W.S. and Reed, I.S., *A new CFAR detection test for radar*, *Digital Signal Processing*, Vol. 1, Academic Press, New York, pp. 198–214, 1991.

SECTION 6.8

1. Reed, I.S., Mallet, J.D., and Brennan, L.E., Rapid convergence rate in adaptive arrays, *IEEE Trans. AES*, 10(6), 853–863, 1974.
2. Melvin, W. and Wicks, M., Improving practical space-time adaptive radar, *Proceedings of the 1997 IEEE National Radar Conference*, Syracuse, NY, pp. 48–53, May 1997.
3. Rangaswamy, M., Himed, B., and Michels, J., Performance Analysis of the Nonhomogeneity Detector for STAP Applications, *Proceedings of the 2001 IEEE Radar Conference*, Atlanta, GA, pp. 193–197, May 2001.
4. Rangaswamy, M., Michels, J., and Himed, B., Statistical analysis of the nonhomogeneity detector for non-Gaussian interference backgrounds, *Proceedings of the 2002 IEEE Radar Conference*, Long Beach, CA, pp. 304–310, April 2002.
5. Baldygo, W., Wicks, M., Brown, R., Antonik, P., Capraro, G., and Hennington, L., Artificial intelligence applications to constant false alarm rate (CFAR) processing, *Proceedings of the 1993 IEEE National Radar Conference*, Boston, MA, April 1993.
6. Antonik, P., Shuman, H., Li, P., Melvin, W., and Wicks, M., Knowledge-based space-time adaptive processing, *Proceedings of the 1997 IEEE National Radar Conference*, Syracuse, NY, pp. 372–377, May 1997.

7. Capraro, C., Capraro, G., Weiner, D., and Wicks, M., Knowledge based space time adaptive processing (KBMapSTAP), *Proceedings of the International Conference on Imaging Science, Vol. 2*, Las Vegas, NV, pp.533–538, June 2001.
8. Guerci, J.R., Knowledge-aided sensor signal processing and expert reasoning, *Proceedings of the 2002 Workshop on Knowledge-Aided Sensor Signal Processing and Expert Reasoning (KASSPER)*, Washington, DC, April 2002 (CD-ROM).
9. Westinghouse Electronic Systems, Multi-Channel Airborne Radar Measurement (MCARM) Flight Test, Vol. 1, Contract F2062-92-C-0161, Rome Laboratory, USAF, September 1995.
10. National Land Cover Data (NLCD), http://landcover.usgs.gov/prodescription.asp.
11. Ward, J., Space-Time Adaptive Processing for Airborne Radar, Technical Report 1015, MIT Lincoln Laboratory, Lexington, MA, December 1994.
12. Guerci, J.R., *Space-Time Adaptive Processing for Radar*, Artech House, Norwood, MA, 2003.
13. Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P., *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, New York, 1992.
14. Borsari, G.K., Mitigating Effects on STAP Processing Caused by an Inclined Array, *Proceedings of the 1998 IEEE National Radar Conference*, Dallas, TX, pp. 135–140, May 1998.
15. Chen, W.S., and Reed, I.S., *A new CFAR detection test for radar*, *Digital Signal Processing*, Vol. 1, Academic Press, New York, pp. 198–214, 1991.

CHAPTER 8

1. Brennan, L.E. and Reed, I.S., Theory of adaptive radar, *IEEE Trans. AES*, 9, 237–252, 1973.
2. Reed, I.S., Mallet, J.D., and Brennan, L.E., Rapid convergence rate in adaptive arrays, *IEEE Trans. AES*, 10, 853–863, 1974.
3. Kelly, E.J., An adaptive detection algorithm, *IEEE Trans. AES*, 22, 115–127, 1986.
4. Khatri, C.G. and Rao, C.R., Effects of estimated noise covariance matrix in optimum signal detection, *IEEE Trans. ASSP*, 35, 671–679, 1987.
5. Cai, L. and Wang, H., On adaptive filtering with the CFAR feature and its performance sensitivity to non-Gaussian interference, *Proceedings of the 24th Annual Conference on Information Sciences and Systems*, pp. 558–563, 1990.
6. Cai, L. and Wang, H., Performance comparisons of the modified SMI and GLR algorithms, *IEEE Trans. AES*, 27 1991.
7. Wang, H. and Cai, L., On adaptive implementation of optimum MTI in severely nonhomogeneous environments, *Proceedings of the 1990 IEEE International Radar Conference*, pp. 351–355, 1990.
8. Wang, H. and Cai, L., A localized adaptive MTD processor, *IEEE Trans. AES*, 27 1991.
9. Adams, R.N., Horowitz, L.L., and Senne, K.D., Adaptive main-beam nulling for narrow beam antenna arrays, *IEEE Trans. AES*, 16, 509–516, 1980.

10. Monzingo, R.A. and Miller, T.W., *Introduction to Adaptive Arrays*, Wiley, New York, 1980.
11. Hudson, J.E., *Adaptive Array Principles*, Peter Peregrinus, New York, 1981.
12. Wang, H. and Cai, L., On adaptive multiband signal detection with the GLR algorithm, *IEEE Trans. AES*, 27 1991.
13. Wang, H. and Cai, L., CFAR Performance of Adaptive Spatial-Temporal Processors, Technical Report, Rome Laboratory, 1991.

CHAPTER 9

1. Ploussios, G., Crossed Bent Monopole Doublets, U.S. Patent No. 6,356,242 B1, March 12, 2002.
2. Weiner, M.M., *Monopole Antennas*, Marcel Dekker, New York, pp. 52–58, 2003.

CHAPTER 10

SECTION 10.1

1. Gabriel, W.E., Adaptive arrays-an introduction, *Proc. IEEE*, 64(February), 239, 1976.
2. Compton, R.T., *Adaptive Antennas*, Prentice Hall, Englewood Cliffs, NJ, 1988.
3. Orfanides, S.J., *Optimum Signal Processing*, 2nd ed., Macmillan, New York, 1988.
4. Applebaum, S.P., Adaptive arrays, *IEEE Trans. AP*, 24(September), 585–594, 1976.
5. Monzingo, R.A. and Miller, T.W., *Introduction to Adaptive Arrays*, Wiley, New York, 1980.
6. Beckmann, P. and Spizzichino, A., *The Scattering of Electromagnetic Waves from Rough Surfaces*, Artech House, Dedham, MA, 1987.
7. Mayhan, J.T., Simmons, A.J., and Cummings, W.C., Wideband adaptive antenna nulling using tapped delay lines, *IEEE Trans. AP*, 29(November), 923, 1981.
8. White, W.D., Wideband interference cancellation in adaptive sidelobe cancellers, *IEEE Trans. AES*, 19(November), 915, 1983.
9. Compton, R.T., The bandwidth performance of a two-element adaptive array with tapped delay-line processing, *IEEE Trans. AP*, 36(January), 5, 1988.
10. Compton, R.T., The relationship between tapped delay line and FFT processing in adaptive arrays, *IEEE Trans. AP*, 36(January), 15–26, 1988.
11. Golub, G.H. and van Loan, C.F., *Matrix Calculations*, 2nd ed., The Johns Hopkins University Press, Baltimore, MD, 1989.
12. Brennan, L.E., Mallet, J.D., and Reed, I.S., Adaptive arrays in airborne MTI radar, *IEEE Trans. AP*, 24(September), 607, 1976.

SECTION 10.2

1. Brennan, L. and Reed, I.S., Theory of adaptive radar, *IEEE Trans. AES*, 9, 237–252, 1973.
2. Brennan, L.E., Mallet, J.D., and Reed, I.S., Adaptive arrays in airborne MTI radar, *IEEE Trans. AP*, 24(September), 607–615, 1976.
3. Monzingo, R. and Miller, T., *Introduction to Adaptive Arrays*, Wiley, New York, 1980.

4. Klemm, R., Adaptive clutter suppression for airborne phased array radars, *IEEE Proc. F and H*, 130, 125–132, 1983.
5. Compton, R.T., *Adaptive Antennas*, Prentice Hall, Englewood Cliffs, NJ, 1988.
6. Andrews, G., Radar pattern design for platform motion compensation, *IEEE Trans. AP*, 26, 566–571, 1978.
7. Skolnik, M., *Radar Handbook*, McGraw-Hill, New York, pp. 16.8–16.14, 1990.
8. Lightstone, L., Faubert, D., and Rempel, G., Multiple phase center DPCA for airborne radar, *Proceedings of the 1991 IEEE National Radar Conference*, pp. 36–40, 1991.
9. Tsandoulas, G., Tolerance control in an array antenna, *Microw. J.*, Oct, 24–35, 1977.

SECTION 10.3

1. Brennan, L. and Reed, I.S., Theory of adaptive radar, *IEEE Trans. AES*, 9, 237–252, 1973.
2. Barile, E.C., Fante, R.L., and Torres, J.A., Some limitations on the effectiveness of airborne adaptive radar, *IEEE Trans. AES*, 28, 1015–1032, 1992.
3. Reed, I.S., Mallett, J.D., and Brennan, L.E., Rapid convergence rate in adaptive arrays, *IEEE Trans. AES*, 10, 853–861, 1974.
4. Takeo, K. and Kikuma, N., An adaptive array utilizing an adaptive spatial averaging technique for multipath environments, *IEEE Trans. AP*, 35, 1389–1396, 1987.
5. Ali, M.E. and Schreib, F., Adaptive single snapshot beamforming: a new concept for the rejection of nonstationary and coherent interferers, *IEEE Trans. SP*, 40, 3055–3058, 1992.
6. Shan, T.J. and Kailath, T., Adaptive beamforming for coherent signals and interference, *IEEE Trans. ASSP*, 33, 527–536, 1985.

SECTION 10.4

1. Monzingo, R. and Miller, T., *Introduction to Adaptive Arrays*, Wiley, New York, 1980.
2. Fante, R.L., Cancellation of specular and diffuse jammer multipath using a hybrid adaptive array, *IEEE Trans. AES*, 27, 823–837, 1991.
3. Morgan, D.R. and Aridgides, A., Adaptive sidelobe cancellation of wideband multipath interference, *IEEE Trans. AES*, 33, 908–917, 1985.
4. Barile, E.C., Fante, R.L., and Torres, J.A., Some limitations on the effectiveness of airborne adaptive radar, *IEEE Trans. AES*, 28, 1015–1032, 1992.
5. Brennan, L. and Reed, I.S., Theory of adaptive radar, *IEEE Trans. AES*, 9, 237–252, 1973.
6. Compton, R.T., *Adaptive Antennas*, Prentice Hall, Englewood Cliffs, NJ, 1988.
7. Farina, A., *Antenna-Based Signal Processing Techniques for Radar Systems*, Artech House, Norwood, MA, 1992.

SECTION 10.5

1. Ganz, M., Ward, J., and Carlson, B.D., Mainbeam nulling with adaptive array interferometry, *Proceedings of the 25th Asilomar Conference on Signals, Systems, and Computers*, Vol. 2, 1991.

2. Monzingo, R. and Miller, T., *Introduction to Adaptive Arrays*, Wiley, New York, p. 72, 1980.
3. Fante, R.L., Cancellation of specular and diffuse jammer multipath using a hybrid adaptive array, *IEEE Trans. AES*, 27, 823–837, 1991.
4. Morgan, D.R. and Aridgides, A., Adaptive sidelobe cancellation of wideband multipath interference, *IEEE Trans. AES*, 33, 908–917, 1985.
5. Mayhan, J.T., Simmons, A.J., and Cummings, W., Wideband adaptive nulling using tapped delay lines, *IEEE Trans. AP*, 29(November), 923–936, 1981.
6. White, W., Wideband interference cancellation in adaptive sidelobe cancellers, *IEEE Trans. AES*, 19(November), 915–925, 1983.
7. Compton, R.T., *Adaptive Antennas*, Prentice Hall, Englewood Cliffs, NJ, 1988.
8. Gabriel, W.E., Adaptive arrays-an introduction, *Proc. IEEE*, 64(February), 239, 1976.
9. Hudson, J.E., *Adaptive Array Principles*, Peregrinus, London, 1981.
10. Brennan, L. and Reed, I.S., Theory of adaptive radar, *IEEE Trans. AES*, 9, 237–252, 1973.
11. Farina, A., *Antenna-Based Signal Processing Techniques for Radar Systems*, Artech House, Norwood, MA, 1992.
12. Applebaum, S.P., Adaptive arrays, *IEEE Trans. AP*, 24(September), 585–594, 1976.
13. Widrow, B. and Stearns, S.D., *Adaptive Signal Processing*, Prentice Hall, Englewood Cliffs, NJ, 1985.

Section 10.6

1. Brennan, L. and Reed, I.S., Theory of adaptive radar, *IEEE Trans. AES*, 9, 237–252, 1973.
2. Friedman, B., *Principles and Techniques of Applied Mathematics*, Wiley, New York, 1962.
3. Ward, J., Space-Time Adaptive Processing for Airborne Radar, Technical Report 1015, MIT Lincoln Laboratory Report, 1994.
4. Brennan, L. and Staudaher, F., Subclutter Visibility Demonstration, Technical Report RL-TR-92-21, 1992.
5. DiPietro, R., Extended factored space-time processing for airborne radar systems, *Proceedings of the 26th Asilomer Conference on Signals, Systems, and Computing*, pp. 425–430, 1992.
6. Haimovich, A., Puch, M., and Baren, M., Training and signal cancellation in adaptive radar, *Digest of 1996 IEEE Radar Conference*, Ann Arbor, MI, pp. 124–128, 1996.
7. Suresh-Babu, B., Torres, J., and Melvin, W., *Processing and Evaluation of Multichannel Airborne Radar Measurements (MCARM)*, *MTR96B0000053*, The MITRE Corporation, Bedford, MA, 1996.
8. Fante, R. and Torres, J., Cancellation of diffuse jammer multipath by an airborne adaptive radar, *IEEE Trans. AES*, 31, 805–820, 1995.

Section 10.7

1. Nickel, U., Monopulse estimation with adaptive arrays, *IEE Proc.-F*, 40(October), 303–308, 1993.

2. Frost, O., An algorithm for linearly constrained adaptive array processing, *Proc. IEEE*, 60(August), 926–935, 1972.
3. Monzingo, R. and Miller, T., *Introduction to Adaptive Arrays*, Wiley, New York, 1980.
4. Brennan, L. and Reed, I.S., Theory of adaptive radar, *IEEE Trans. AES*, 9, 237–252, 1973.
5. VanVeen, B., Maximum variance beamforming, In *Adaptive Radar Detection & Estimation*, Haykin, S. and Steinhardt, A., eds., Wiley, New York, 1992.

SECTION 10.8

1. VanVeen, B., Maximum variance beamforming, In *Adaptive Radar Detection & Estimation*, Haykin, S. and Steinhardt, A., eds., Wiley, New York, 1992.
2. Compton, R.T., *Adaptive Antennas*, Prentice Hall, Englewood Cliffs, NJ, 1988.

SECTION 10.9

1. Mensa, D., *High Resolution Radar Cross-Section Imaging*, Artech House, Norwood, MA, 1991.
2. Carrara, W., Goodman, R., and Majewski, R., *Spotlight Synthetic Aperture Radar*, Artech House, Norwood, MA, 1995.
3. Curlander, J. and McDonough, R., *Synthetic Aperture Radar*, Wiley, New York, 1991.
4. Jakowitz, C., Wahl, D., Eichel, P., and Ghiglia, D., *Spotlight-Mode Synthetic Aperture Radar*, Kluwer Academic Publishing, Boston, 1996.
5. Walker, J., Range-Doppler imaging of rotating objects, *IEEE Trans. AES*, 16, 23–52, 1980.
6. DeGraaf, S., SAR imaging via modern 2-D spectral estimation methods, *IEEE Trans. IP*, 7, 729–761, 1998.
7. DeGraaf, S., Sidelobe reduction via adaptive FIR filtering in SAR imagery, *IEEE Trans. IP*, 3, 292–301, 1994.
8. DiPietro, R., Fante, R., and Perry, R., Synthetic aperture imaging of moving targets, *IEEE Trans. AES*, 35, 188–220, 1999.
9. Capon, J., High resolution frequency–wavenumber spectral estimation, *Proc. IEEE*, 57, 1408–1418, 1969.
10. Compton, R.T., *Adaptive Antennas*, Prentice Hall, Englewood Cliffs, NJ, 1988.
11. Brennan, L. and Reed, I.S., Theory of adaptive radar, *IEEE Trans. AES*, 9, 237–252, 1973.
12. Monzingo, R. and Miller, T., *Introduction to Adaptive Arrays*, Wiley, New York, 1980.
13. Hudson, J., *Adaptive Array Principles*, Peter Peregrinus, London, 1989.
14. VanVeen, B., Maximum variance beamforming, In *Adaptive Radar Detection & Estimation*, Haykin, S. and Steinhardt, A., eds., Wiley, New York, 1992.
15. Lacoss, R., Data adaptive spectral analysis methods, *Geophysics*, 36, 661–675, 1976.
16. Brennan, L.E., Mallet, J.D., and Reed, I.S., Adaptive arrays in airborne MTI radar, *IEEE Trans. AP*, 24(September), 607–615, 1976.

17. Baird, C. and Zahm, C., Performance criteria for narrowband array processing, *IEEE Conference on Decision and Control*, Miami, FL, December 1971.

SECTION 10.10

1. Paulraj, A. and Papadias, C.B., Space-time processing for wireless communications, *IEEE Signal Process. Mag.*, November, 49–83, 1997.
2. Monzingo, R. and Miller, T., *Introduction to Adaptive Arrays*, Wiley, New York, 1980.
3. Fante, R. and Torres, J., Cancellation of diffuse jammer multipath by an airborne adaptive radar, *IEEE Trans. AES*, 31, 805–820, 1995.
4. Techau, P., Guerci, J., Slocumb, T., and Griffiths, L., Performance bounds for interference mitigation in radar systems, *Proceedings of the 1999 IEEE Radar Conference*, pp. 12–17, 1999.
5. Techau, P., Effects of receiver filtering on hot clutter mitigation, *Proceedings of the 1999 Radar Conference*, pp. 84–89, 1999.
6. Marshall, D. and Gabel, R., Simultaneous mitigation of multipath jamming and ground clutter, *Proceedings of the Adaptive Sensor Array Processing (ASAP) Workshop*, Report ASAP-4, MIT Lincoln Laboratory, March 1996.
7. Ward, J., Space-Time Adaptive Processing for Airborne Radar, Technical Report 1015, MIT Lincoln Laboratory Report, 1994.
8. Brennan, L.E., Mallet, J.D., and Reed, I.S., Adaptive arrays in airborne MTI radar, *IEEE Trans. AP*, 24(September), 607–615, 1976.
9. Compton, R.T., *Adaptive Antennas*, Prentice Hall, Englewood Cliffs, NJ, 1988.
10. Hudson, J., *Adaptive Array Principles*, Peter Peregrinus, London, 1989.
11. VanVeen, B., Maximum variance beamforming, In *Adaptive Radar Detection & Estimation*, Haykin, S. and Steinhardt, A., eds., Wiley, New York, 1992.
12. Brennan, L. and Reed, I.S., Theory of adaptive radar, *IEEE Trans. AES*, 9, 237–252, 1973.
13. Farina, A., *Antenna-Based Signal Processing Techniques for Radar Systems*, Artech House, Norwood, MA, 1992.
14. Hildebrand, F., *Methods of Applied Mathematics*, 2nd ed., Prentice Hall, Englewood Cliffs, NJ, 1965.
15. Kaplan, E., ed., *Understanding GPS: Principles and Applications*, Artech House, Norwood, MA, 1996.
16. Parkinson, B. and Spilker, J., eds., *Global Positioning System: Theory and Applications*, American Institute of Aeronautics and Astronautics, Washington, DC, 1996.
17. Weibull, L.B., Conquering multipath; the GPS accuracy battle, *GPS World*, April 59–66, 1997.
18. DiPietro, R., Extended factored space-time processing for airborne radar systems, *Proceedings of the 26th Asilomer Conference on Signals, Systems, and Computing*, pp. 425–430, 1992.
19. Compton, R.T., The relationship between tapped delay line and FFT processing in adaptive arrays, *IEEE Trans. AP*, 36(January), 15–26, 1988.
20. Fante, R. and Vaccaro, J., Cancellation of Jammers and Jammer Multipath in a GPS Receive Array, Technical report MTR98B0000084, The MITRE Corporation, November 1998.

21. Gupta, I. and Ksienski, A., Effect of mutual coupling on the performance of adaptive arrays, *IEEE Trans. AP*, 31, 785–791, 1983.
22. Fante, R. and Vaccaro, J., Ensuring GPS availability in an interference environment, *Proc. IEEE PLANS* 2000, San Diego, CA, March 2000.

SECTION 10.11

1. Baronoski, E. and Ward, J., Source allocation using adaptive subspace beamformer outputs, *Proceedings of the 1997 IEEE International Conference on Speech Signal Processing (ICAASP 97)*, Vol. 5, pp. 3993–3776, 1997.
2. Applebaum, S.P., Adaptive arrays, *IEEE Trans. AP*, 24(September), 585–594, 1976.
3. Kelly, E., Reed, I., and Root, W., The detection of radar echoes in noise II, *J. Soc. Ind. Appl. Math.*, 8(September), 481–510, 1960.
4. Urkowitz, H., The accuracy of maximum likelihood angle estimates in radar and sonar, *IEEE Trans. ME*, 8(January), 39–45, 1964.
5. Davis, R., Brennnan, L., and Reed, I., Angle estimation with adaptive arrays in external noise fields, *IEEE Trans. AES*, 12(March), 179–186, 1976.
6. Whalen, A., *Detection of Signals in Noise*, Academic Press, New York, 1971.
7. Brennan, L. and Reed, I.S., Theory of adaptive radar, *IEEE Trans. AES*, 9, 237–252, 1973.
8. Reed, I.S., Mallett, J.D., and Brennan, L.E., Rapid convergence rate in adaptive arrays, *IEEE Trans. AES*, 10, 853–861, 1974.
9. Picinbono, B., *Random Signals and Systems*, Prentice Hall, Englewood Cliffs, NJ, 1993.
10. Hatke, G.F., Superresolution source location with planar arrays, *Lincoln Lab J.*, 10(2) 1997.
11. Barton, D.K. and Ward, H.R., *Handbook of Radar Measurements*, Prentice Hall, Englewood Cliffs, NJ, p. 24, 1969.
12. Davis, R.M., Fante, R.L., Crosby, W.J., and Balla, R.J., A Maximum Likelihood Beamspace Processor, Record of IEEE International Radar Conference, May 2000.
13. Carlson, B.D., Covariance matrix estimation errors and diagonal loading in adaptive arrays, *IEEE Trans. AES*, 24(July), 397–401, 1988.

CHAPTER 12

1. Sekine, M., Musham, T., Tomitam, Y., Hagisawa, T., Irabu, T., and Kiuchi, E., On Weibull distributed weather clutter, *IEEE Trans. AES*, 15, 824–828, 1979.
2. Jakeman, E. and Pusey, P., A model for non-Rayleigh sea echo, *IEEE Trans. AP*, 24, 806–814, 1976.
3. Watts, S. and Ward, K., Spatial correlation in *K*-distributed sea clutter, *IEE Proc. F*, 134(6), 526–532, 1987.
4. Hawkes, C. and Haykin, S., Modeling of clutter for coherent pulsed radar, *IEEE Trans. IT*, 21, 703–707, 1975.
5. Farina, A., Russo, A., and Studer, F., Coherent radar detection in log-normal clutter, *IEE Proc. F*, 133(1), 39–54, 1986.
6. Conte, E. and Longo, M., On a coherent model for log-normal clutter, *IEE Proc. F*, 134(2), 198–201, 1987.

7. Watts, S., Radar detection prediction in *K*-distributed sea clutter and thermal noise, *IEEE Trans. AES*, 23, 40–45, 1987.
8. Li, G. and Yu, K.B., Modeling and simulation of coherent Weibull clutter, *IEE Proc. F*, 136(1), 2–12, 1989.
9. Jakeman, E., On the statistics of *K*-distributed noise, *J. Phys. A*, 13, 31–48, 1980.
10. Szajnowski, W., The generation of correlated Weibull clutter for signal detection problems, *IEEE Trans. AES*, 13, 536–540, 1977.
11. Peebles, P., The generation of correlated lognormal clutter for radar simulation, *IEEE Trans. AES*, 7, 1215–1217, 1971.
12. Farina, A., Russo, A., Scannapieco, F., and Barbarossa, S., Theory of radar detection in coherent Weibull clutter, *IEE Proc. F*, 134(2), 174–190, 1987.
13. Conte, E., Galati, G., and Longo, M., Exogenous modeling of non-Gaussian clutter, *J. IERE*, 57(4), 151–155, 1987.
14. Kingman, J., Random walks with spherical symmetry, *Acta Math.*, 109, 11–53, 1963.
15. Yao, K., A representation theorem and its applications to spherically invariant random processes, *IEEE Trans. IT*, 19, 600–608, 1973.
16. Goldman, J., Detection in the presence of spherically symmetric random vectors, *IEEE Trans. IT*, 22, 52–58, 1976.
17. Brehm, H., Description of spherically invariant random processes by means of *G*-functions, *Springer Lect. Notes*, 969, 39–73, 1982.
18. Brehm, H. and Stammler, W., Description and generation of spherically invariant speech-model signals, *Signal Process.*, 12(2), 119–141, 1987.
19. Conte, E. and Longo, M., Characterization of radar clutter as a spherically invariant random process, *IEE Proc. F*, 134(2), 191–197, 1987.
20. Conte, E., Longo, M., and Lops, M., Modeling and simulation of non-Rayleigh radar clutter, *IEE Proc. F*, 138(2), 121–130, 1991.
21. Rangaswamy, M., Weiner, D., and Ozturk, A., Non-Gaussian random vector identification using spherically invariant random processes, *IEEE Trans. AES*, 29, 111–124, 1993.
22. Rangaswamy, M., Weiner, D., and Ozturk, A., Computer generation of correlated non-Gaussian clutter, *IEEE Trans. AES*, 31(January), 106–116, 1955.
23. Kingman, J., On random sequences with spherical symmetry, *Biometrika*, 59, 492–494, 1972.
24. Wise, G. and Gallagher, N., On spherically invariant random processes, *IEEE Trans. IT*, 24, 118–120, 1978.
25. Johnson, M., *Multivariate Statistical Simulation*, Wiley, New York, 1987.
26. Miller, K., *Multidimensional Gaussian Distributions*, Wiley, New York, 1964.
27. Blake, I. and Thomas, J., On a class of processes arising in linear estimation theory, *IEEE Trans. IT*, 14, 12–16, 1968.
28. Abramowitz, M. and Stegun, I., *Handbook of Mathematical Functions*, Dover Publications, New York, 1972.
29. Chu, K., Estimation and decision for linear systems with elliptical random processes, *IEEE Trans. AC*, 18, 499–505, 1973.
30. Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, 1984.
31. Frame, J., Matrix functions and applications, part 1-matrix operations and generalized inverses, *IEEE Spectrum*, 3, 212, 1964.

32. Picinbono, B., Spherically invariant and compound Gaussian stochastic processes, *IEEE Trans. IT*, 16, 77–79, 1970.
33. Wise, G., Traganitis, A., and Thomas, J., The effect of a memoryless nonlinearity on the spectrum of a random process, *IEEE Trans. IT*, 23, 84–89, 1977.
34. Liu, B. and Munson, D., Generation of a random sequence having a jointly specified marginal distribution and autocovariance function, *IEEE Trans. ASSP*, 6, 973–983, 1982.
35. Gradshteyn, I. and Ryzhik, I., *Table of Integrals, Series, and Products*, Academic Press, New York, 1980.
36. Erdelyi, A., Magnus, W., and Oberhettinger, F., *Tables of Integral Transforms*, McGraw-Hill, New York, 1954.
37. Ritcey, J. and Tran, P., Tail behavior of compound clutter models, *Proceedings of the 26th Annual Conference on Information Sciences and Systems*, Princeton, NJ, 1992.
38. Modestino, J. and Ningo, A., Detection of weak signals in narrowband non-Gaussian nose, *IEEE Trans. IT*, 25, 592–600, 1979.
39. Kassam, S., *Signal Detection in Non-Gaussian Noise*, Springer Verlag, New York, 1987.
40. Bratley, P., Fox, B., and Schrage, L., *A Guide to Simulation*, Springer Verlag, New York, 1987.
41. Ozturk, A., A new method for univariate and multivariate distribution identification, *J. Am. Stat. Assoc.*, 1990, (submitted but not accepted).
42. Koziol, J., A class of invariant procedures for assessing multivariate normality, *Biometrika*, 69(423), 427, 1982.
43. Martinez, A., Swaszek, P., and Thomas, J., Locally optimal detection in multivariate non-Gaussian noise, *IEEE Trans. IT*, 30, 815–822, 1984.
44. Ozturk, A. and Romeu, J., A new method for assessing multivariate normality with graphical applications, *Commun. Stat. B. Simul. Comput.*, 21, 15–34, 1992.
45. Johnson, N. and Kotz, S., *Distributions in Statistics–Continuous Multivariate Distributions*, Wiley, New York, 1977.
46. Gnanadesikan, R., *Methods of Statistical Data Analysis of Multivariate Observations*, Wiley, New York, 1977.
47. Mardia, K., Test of univariate and multivariate normality, *Handbook of Statistics*, 1, 279–320, 1980.
48. Goodman, N., Statistical analysis based on a certain multivariate complex distribution, *Ann. Math. Stat.*, 34, 152–177, 1963.
49. Fang, K. and Anderson, T., *Statistical Inference in Elliptically Contoured and Related Distributions*, Allerton Press, New York, 1990.
50. Rangaswamy, M., Weiner, D., and Ozturk, A., Simulation of correlated non-Gaussian interference for radar signal detection, *Proceedings of the 25th Asilomar Conference on Signals, Systems, and Computer*, Pacific Grove, CA, 1990.
51. Fisher, R., The use of multiple measurements in taxonomic problems, *Ann. Eugen.*, 7, 179–188, 1936.
52. Reed, I., Mallett, J., and Brennan, L., Rapid convergence rate in adaptive arrays, *IEEE Trans. AES*, 10, 853–863, 1974.
53. Khatri, C. and Rao, C., Effects of estimated noise covariance matrix in optimum signal detection, *IEEE Trans. ASSP*, 35, 671–679, 1987.

CHAPTER 13

1. Neyman, J. and Pearson, E.S., Contributions to the theory of testing statistical hypothesis, *Stat. Res. Mem.*, 1, 1–37, 1936.
2. Neyman, J. and Pearson, E.S., Sufficient statistics and uniformly most powerful tests of statistical hypothesis, *Stat. Res. Mem.*, 1, 113–137, 1936.
3. Middleton, D., Canonically Optimum Threshold Detection, Memo RM-4687-PR, RAND Corporation, 1965.
4. Middleton, D., *Statistical Communication Theory*, McGraw-Hill, New York, 1960.
5. Middleton, D., Threshold detection and estimation in correlated interference, *Proceedings of the ninth Zurich Symposium on EMC*, Switzerland, EMC-91, pp. 7–12, 1991.
6. Middleton, D., Threshold vector field detectors, *IEEE Trans. EC*, 30(4), 538–552, 1988.
7. Miller, J.H. and Thomas, J., Detectors for discrete time signals in non-Gaussian noise, *IEEE Trans. IT*, 18(2), 241–250, 1972.
8. Kassam, S.A., *Signal Detection in Non-Gaussian Noise*, Springer Verlag, New York, 1989.
9. Song, I. and Kassam, S.A., Locally optimum detection of signals in a generalized observation model: the known signal case, *IEEE Trans. IT*, 36(3), 502–515, 1990.
10. Song, I. and Kassam, S.A., Locally optimum detection of signals in a generalized observation model: the random signal case, *IEEE Trans. IT*, 36(3), 516–530, 1990.
11. Capon, J., On the asymptotic efficiency of locally optimum detectors, *IRE Trans. IT*, 7(2), 67–71, 1961.
12. Conte, E., Izzo, L., and Pauro, L., Asymptotically optimum radar detectors in non-Rayleigh clutter, *IEE Proc. F*, 134(7), 667–672, 1987.
13. Valeyev, V.G. and Aspisov, I.Y., Asymptotically effective nonparametric algorithms for detecting signals in non-Gaussian interference, *Radiotekhnika i Elektronika*, 6, 1–7, 1989.
14. Raveendra, K.R. and Srinivasan, R., Threshold detection of CPFSK in non-Gaussian noise, *IEE Proc. F*, 1984.
15. Ingram, R.F. and Houle, R., Performance of the Optimum and Several Suboptimum Receivers for Threshold Detection of Known Signals in Additive, White, non-Gaussian Noise, NUSC Technical Report-6339, pp. 1–26, 1980.
16. Gardner, W., Structural characterization of locally optimum detectors in terms of locally optimum estimators and correlators, *IEEE Trans. IT*, 28(6), 924–932, 1982.
17. Johnson, D.H. and Orsack, G.C., Relation of Signal Set Choice to the Performance of Optimal Non-Gaussian Detectors, IEEE International Symposium on Information Theory, 1990.
18. Spaulding, A.D., Locally optimum and suboptimum detector performance in a non-Gaussian interference environment, *IEEE Trans. COM*, 33(6), 509–517, 1985.
19. Bouvet, M., Expansions of the likelihood ratio and applications, *IEEE Trans. ASSP*, 34(4), 653–660, 1986.
20. Shishkov, B.B. and Penev, S.I., Asymptotically optimum algorithms for detection of signals in a background of correlated interference and white noise, *Radiotekhnika i Electronika*, 7, 1419–1424, 1988.

21. Modestino, J.W. and Ningo, A., Detection of weak signals in non-Gaussian noise, *IEEE Trans. IT*, 25(5), 592–600, 1979.
22. Martinez, A.B., Swaszek, P.F., and Thomas, J.B., Locally optimum detection in multivariate non-Gaussian noise, *IEEE Trans. IT*, 30(6), 815–822, 1984.
23. Sangston, K.J. and Gerlach, K., Results on the Detection of Signals in Spherically Invariant Random Noise, NRL Report 9202, pp. 1–12, 1989.
24. Conte, E. and Longo, M., Characterization of radar clutter as a spherically invariant random process, *IEE Proc. F*, 134, 191–197, 1987.
25. Conte, E., Longo, M., and Lops, M., Modeling and simulation of non-Rayleigh radar clutter, *IEE Proc. F*, 138, 121–130, 1991.
26. Rangaswamy, M., Weiner, D.D., and Ozturk, A., Spherically invariant random processes for modeling and distribution identification of non-Gaussian random vectors, *IEEE Trans. AES*, 3(July), 111–124, 1992.
27. Rangaswamy, M., Weiner, D.D., and Ozturk, A., Computer generation of correlated non-Gaussian radar clutter, *IEEE Trans. AES*, 31(January), 106–116, 1995.
28. Yao, K., A representation theorem and its application to spherically invariant random processes, *IEEE Trans. IT*, 19, 600–608, 1973.
29. Pentini, F.A., Farina, A., and Zirilli, F., Radar detection of targets located in a coherent *K*-distributed clutter background, *IEE Proc. F, Commun. Radar Signal Process.*, 139(3), 239–245, 1992.
30. Van Trees, H., *Detection, Estimation, and Modulation Theory, Part I*, Wiley, New York, 1968.
31. Hoaking, J.R.M. and Wallis, J.R., Parameter and quantile estimation of the generalized Pareto distribution, *Technometrics*, 29(3), 339–349, 1987.
32. Harrel, F.E. and Davis, C.E., A new distribution-free quantile estimator, *Biometrika*, 69, 635–640, 1982.
33. Ozturk, A., A new method for univariate and multivariate distribution identification, *J. Am. Stat. Assoc.*, 1990, (submitted but not accepted).
34. Hill, B.M., A simple general approach to inference about the tail of a distribution, *Ann. Stat.*, 3, 1163–1174, 1975.
35. Pickands, J., Statistical inference using extreme order statistics, *Ann. Stat.*, 3, 119–131, 1975.
36. Weissman, I., Estimation of parameters and large quantiles based on the *k* largest observations, *J. Am. Stat. Assoc.*, 73, 812–815, 1978.
37. Boos, D.D., Using extreme value theory to estimate large samples, *Technometrics*, 20, 33–39, 1984.
38. Mandelbrot, B., The Pareto levy law and the distribution of income, *Int. Econ. Rev.*, 1, 79–106, 1960.
39. Patel, G.P., *Statistical Distributions in Scientific Work*, Reidal Dordrecht, Holland, 1975.
40. McCullough, J.H., Continuous time processes with stable increments, *J. Business (*The University of Chicago), 15, 601–619, 1978.
41. Paulson, A.S., Rutherford, D.C., and Ulseth, G.W., Some Application of Stable Laws to Collective Risk Theory, Technical Report, School of Management, Rensselaer Polytechnic Institute, 1981.
42. Kahn, P.M., *Credibility Theory and Applications*, Academic Press, New York, 1975.
43. Guida, M., Iovino, D., and Longo, M., Comparative performance analysis of some extrapolative estimators of probability tails, *IEEE J. Sel. Area. Commun.*, 6, 76–84, 1988.

44. SAS User's Guide, Cary NC: SAS Institute Inc., Version 5 Edition, 1990.
45. de Oliviera, T., *Statistical Extremes and Applications*, D. Reidel Publishing Co., Boston, 1984.
46. Davison, A.C. and Smith, R.L., Models for exceedances over high thresholds, *J. R. Stat. Soc.*, 52, 393–442, 1990.
47. DuMouchel, W.H., Estimating the stable index α in order to measure tail thickness: a critique, *Ann. Stat.*, 11, 1019–1031, 1983.
48. Nelder, J.A. and Mead, R., A simplex method for function minimization, *Comput. J.*, 7, 308–313, 1965.
49. Gradshteyn, I.S. and Ryzhik, I.M., *Tables of Integrals, Series, and Products*, Academic Press, New York, 1980.
50. Jakeman, E., On the statistics of *K*-distributed noise, *J. Phys. A, Math. Gen.*, 13, 31–48, 1980.
51. Watson, G.N., *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, London, 1966.
52. Galambos, J., *Asymptotic Theory of Extreme Order Statistics*, Wiley, New York, 1978.

CHAPTER 14

1. Kailath, T. and Shan, T., Adaptive beamforming for coherent signals and interference, *IEEE Trans. ASSP*, 33(3), 527–536, 1985.
2. Gray, D.A., Formulation of the maximum signal-to-noise ratio array processor in beam space, *J. Acoust. Soc. Am.*, 1195–1201, 1982.
3. Vural, A.E., A comparative performance study of adaptive array processors, *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 695–700, 1977.
4. Yao, K., A representation theorem and its application to spherically invariant random processes, *IEEE Trans. IT*, 19(July), 600–608, 1973.
5. Conte, E., Longo, M., and Lops, M., Modeling and simulation of non-Rayleigh radar clutter, *IEE Proc. F*, 138, 121–130, 1991.
6. Sangston, K.J. and Gerlach, K., Results on the detection of signals in spherically invariant random noise, NRL Report 9209, pp. 1–12, 1989.
7. Conte, E. and Longo, M., Characterization of radar clutter as a spherically invariant random process, *IEE Proc. F*, 134, 191–197, 1987.
8. Gross, B.L., The correlation sonar: an absolute velocity sensor for autonomous underwater vehicle navigation, EDO Report 11189, pp. 15–16, 1991.
9. Bookheimer, W.C. and Edward, J.A., Navigation sonar for deep water operation, *Proceedings of the IEEE Position, Location, and Navigation Symposium (PLANS) 1984*, San Diego, CA, pp. 224–228, November 26–29, 1984.
10. Bookheimer, W.C., Dickey, F.R., and Rhoades, K.W., Implementation and testing of a deep water correlation sonar, *Proceedings of the 1983 Offshore Technology Conference*, Houston, TX, pp. 437–446, May 2–5, 1983.
11. Dickey, F.R., Velocity Measuring Correlation Sonar, U.S. Patent No. 4,244,026, January 6, 1981.
12. Dickey, F.R. and Edward, J.A., Velocity measurement using correlation sonar, *Proceedings, IEEE Position, Location, and Navigation Symposium (PLANS) 1978*, San Diego, CA, pp. 255–264, November 6–9, 1978.

13. Chu, K., Estimation and detection for linear systems with elliptical random processes, *IEEE Trans. AC*, 18(August), 499–505, 1973.
14. Rangaswamy, M., Spherically Invariant Random Processes for Radar Clutter Modeling, Simulation, and Distribution Identification, Ph.D. dissertation, Department of Electrical and Computer Engineering, Syracuse University NY, December 1992.
15. Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, 1984.
16. Abramovitz, M. and Stegun, I.A., *Handbook of Mathematical Functions*, Dover Publications, New York, 1965.
17. Erdelyi, A., Magnus, W., and Oberhettinger, F., *Tables of Integral Transforms*, McGraw-Hill, New York, 1954.
18. Bickel, P.J. and Doksum, K.A., *Mathematical Statistics: Basic Ideas and Selected Topics*, Holden-Day, Oakland, CA, 1977.
19. Gradshteyn, I.S. and Ryzhik, I.M., *Tables of Integrals, Series, and Products*, Academic Press, Orlando, CA, 1980.
20. Ozturk, A., Rangaswamy, M., and Weiner, D.D., Non-Gaussian random vector identification using spherically invariant random processes, *IEEE Trans. AES*, 29, 111–124, 1993.
21. Van Trees, H.L., *Detection, Estimation, and Modulation Theory, Parts I, II, and III*, Wiley, New York, 1968.

CHAPTER 15

SECTION 15.1

1. Ginolhac, G. and Jourdain, G., 'Principal component inverse' algorithm for detection in the presence of reverberation, *IEEE J. Ocean. Eng.*, 27(April), 310–321, 2002.
2. Kuo, E.Y.T., Joint perturbation scattering characterization of a littoral ocean bottom reverberation: theory, scattering strength predictions, and data comparisons, *IEEE J. Ocean. Eng.*, 20(July), 198–210, 1995.
3. Miklovic, D.W., Huster, M.E., and Moser-Breaker, M., The spatial variability of surface reverberation under high and low wind speeds, *IEEE J. Ocean. Eng.*, 20(October), 337–339, 1995.
4. Novarini, J.C. and Caruther, J.W., A simplified approach to backscattering from a rough sea-floor with sediment in homogeneities, *IEEE J. Ocean. Eng.*, 23(July), 157–166, 1998.
5. Conte, E. and Longo, M., Characterization of radar clutter as a spherically invariant random process, *IEE Proc. F*, 134, 191–197, 1987.
6. Sangston, K.J. and Gerlach, K., Results on the Detection of Signals in Spherically Invariant Random Noise, NRL Report 9209, pp. 1–12, 1989.
7. Conte, E., Longo, M., and Lops, M., Modeling and simulation of non-Rayleigh radar clutter, *IEE Proc. F*, 138, 121–130, 1991.
8. Yao, K., A representation theorem and its application to spherically invariant random processes, *IEEE Trans. IT*, 19(July), 600–608, 1973.
9. Yao, K., A representation theorem and its application to spherically invariant random processes, *IEEE Trans. IT*, 19(July), 600–608, 1973.

10. Abraham, D.A., Signal excess in *K*-distributed reverberation, *IEEE J. Ocean. Eng.*, 28(July), 526–536, 2003.
11. Ozturk, A., Rangaswamy, M., and Weiner, D.D., Non-Gaussian random vector identification using spherically invariant random processes, *IEEE Trans. AES*, 29, 111–124, 1993.
12. Gerlach, K.R. and Sangston, K.J., Coherent detection of radar targets in a non-Gaussian Background, *IEEE Trans. AES*, 3(April), 599–604, 1994.
13. Michels, J.H. and Rangaswamy, M., Adaptive signal processing in non-Gaussian backgrounds, *Proceedings of the ninth IEEE Workshop Statistical Signal Array Processing*, Portland, OR, 1998.
14. Oztgurk, A., Rangaswamy, M., and Weiner, D.D., Computer generation of correlated non-Gaussian radar clutter, *IEEE Trans. AES*, 31(January), 106–116, 1995.
15. DiBisceglie, M., Galdi, C., and Griffiths, H.D., Statistical scattering model for high resolution sonar images: characterization and parameter estimation, *IEE Proc. F*, 146, 1999.
16. Conte, E., DeMaio, A., and Galdi, C., Signal detection in compound Gaussian noise: Neyman–Pearson and CFAR detectors, *IEEE Trans. SP*, 48(February), 419–428, 2000.
17. Conte, E., DeMaio, A., and Ricci, G., Adaptive CFAR detection in compound Gaussian clutter with circulant covariance matrix, *IEEE Signal Process. Lett.*, 7(March), 63–65, 2000.
18. Gini, F., Montanari, M., and Verrazzani, L., Estimation of chirp radar signals in compound Gaussian clutter: a cyclostationary approach, *IEEE Trans. SP*, 48(April), 1029–1039, 2000.
19. Castillo, E. and Hadi, A.S., Fitting the generalized Pareto distribution to data, *J. Am. Stat. Assoc.*, 92(440) 1997.
20. Abraham, D.A., Statistical normalization of non-Rayleigh reverberation, *Proceedings of the MTS/IEEE OCEANS'97 Conference*, 1997.
21. Dudewicz, E.J. and Mishra, S.N., *Modern Mathematical Statistics*, Wiley, New York, 1988.

Section 15.2

1. Ward, K.D., Baker, C.J., and Watts, S., Maritime surveillance radar, part 1: radar scattering from the ocean surface, *IEE Proc. F*, 137(April), 51–62, 1990.
2. Baker, C.J., *K*-distributed coherent sea clutter, *IEE Proc. F*, 138(April), 89–92, 1991.
3. Jao, J.K., Amplitude distribution of composite terrain radar clutter and the *k*-distribution, *IEEE Trans. AP*, 32(October), 1049–1062, 1984.
4. Sekine, M., Ohtani, S., Musha, T., Irabu, T., Kiuchi, E., Higisawa, T., and Tomita, Y., Weibull-distributed ground clutter, *IEEE Trans. AES*, 17(July), 596–598, 1981.
5. Conte, E., Lops, M., and Ricci, G., Radar detection in *K*-distributed clutter, *IEE Proc. F*, 141(April), 116–118, 1994.
6. Rangaswamy, M., Weiner, D.D., and Ozturk, A., Non-Gaussian random vector identification using spherically random processes, *IEEE Trans. AES*, 29(January), 111–124, 1993.
7. Rangaswamy, M., Weiner, D.D., and Ozturk, A., Computer generation of correlated non-Gaussian radar clutter, *IEEE Trans. AES*, 31(January), 106–116, 1995.

8. Sangston, K.J. and Gerlach, K.R., Coherent detection of radar targets in a non-Gaussian background, *IEEE Trans. AES*, 30(April), 330–340, 1994.
9. Keckler, A.D., Non-Gaussian Clutter Modeling by Spherically Invariant Random Vectors, Technical Report, AFOSR, Rome Laboratory, Rome, NY, September 1996.
10. Ozturk, A. and Dudewicz, E.J., A new statistical goodness-of-fit test based on graphical representation, *Biom. J.*, 34(4), 403–427, 1992.
11. Yao, K., A representation theorem and its applications to spherically invariant random processes, *IEEE Trans. IT*, 19, 600–608, 1973.
12. Stademan, D.L., Ph.D. Part 2 Examination, Department of Electrical and Computer Engineering, Syracuse University, 1995.

SECTION 15.3

1. Ward, K.D., Baker, C.J., and Watts, S., Maritime surveillance radar, part 1: radar scattering from the ocean surface, *IEE Proc. F*, 137(April), 51–62, 1990.
2. Baker, C.J., *K*-distributed coherent sea clutter, *IEE Proc. F*, 138(April), 89–92, 1991.
3. Jao, J.K., Amplitude distribution of composite terrain radar clutter and the *k*-distribution, *IEEE Trans. AP*, 32(October), 1049–1062, 1984.
4. Sekine, M., Ohtani, S., Musha, T., Irabu, T., Kiuchi, E., Hagisawa, T., and Tomita, Y., Weibull-distributed ground clutter, *IEEE Trans. AES*, 17(July), 596–598, 1981.
5. Conte, E., Lops, M., and Ricci, G., Radar detection in *K*-distributed clutter, *IEE Proc. F*, 141(April), 116–118, 1994.
6. Sangston, K.J., and Gerlach, K.R., Coherent detection of radar targets in a non-Gaussian background, *IEEE Trans. AES*, 30(April), 330–340, 1994.
7. Rangaswamy, M., Weiner, D.D., and Ozturk, A., Non-Gaussian random vector identification using spherically random processes, *IEEE Trans. AES*, 29(January), 111–124, 1993.
8. Rangaswamy, M., Weiner, D.D., and Ozturk, A., Computer generation of correlated non-Gaussian radar clutter, *IEEE Trans. AES*, 31(January), 106–116, 1995.
9. Keckler, A.D., Non-Gaussian Clutter Modeling by Spherically Invariant Random Vectors, Technical Report, AFOSR, Rome Laboratory, Rome NY, September 1996.
10. Ozturk, A. and Dudewicz, E.J., A new statistical goodness-of-fit test based on graphical representation, *Biom. J.*, 34(4), 403–427, 1992.
11. Keckler, A.D., NonGaussian Clutter Simulation and Distrubute Approximation using Spherically Invariant Random Vectors, Ph.D. dissertation, Dept. of Electrical Engineering and Computer Science, Syracuse University, Nov. 22, 2004.
12. Slamani, M.A., A New Approach to Radar Detection Based on the Partitioning and Statistical Characterization of the Surveillance Volume, Ph.D. thesis, Department of Electrical and Computer Engineering, Syracuse University, 1994.
13. Slamani, M.A. and Weiner, D.D., ASCAPE: an automated approach to the partitioning and statistical characterization of a surveillance volume, *Proceedings of the sixth International Conference on Signal Processing Applications and Technology*, 1995.

14. Conte, E. and Longo, M., Characterization of radar clutter as a spherically invariant random process, *IEE Proc. F*, 134(April), 191–197, 1987.
15. Miller, K.S., *Multidimensional Gaussian Distributions*, Wiley, New York, 1964.
16. Shelupsky, D., An introduction of spherical coordinates, *Am. Math. Mon.*, 69(September), 644–646, 1962.
16. Stademan, D.L., Ph.D. Part 2 Examination, Department of Electrical and Computer Engineering, Syracuse University, 1995.

SECTION 15.4

1. Conte, E. and Longo, M., Characterization of radar clutter as a spherically invariant random process, *IEE Proc. F*, 132(2), 191–197, 1987.
2. Rangaswamy, M., Weiner, D.D., and Ozturk, A., Non-Gaussian random vector identification using spherically invariant random processes, *IEEE Trans. AES*, 29(1), 111–124, 1993.
3. Ward, K.D., Baker, C.J., and Watts, S., Maritime surveillance radar, Part 1: radar scattering from the ocean surface, *IEE Proc. F*, 137(2), 51–62, 1990.
4. Blacknell, D., Target detection in correlated SAR clutter, *IEE Proc. F*, 147(1), 9–16, 2000.
5. Frery, A.C., Muller, H.J., Yanasse, C.F., and Sant'Anna, S.J.S., A model for extremely heterogeneous clutter, *IEEE Trans. GRS*, 35(3), 648–659, 1997.
6. Sangston, K.J. and Gerlach, K.R., Coherent detection of radar targets in a non-Gaussian background, *IEEE Trans. AES*, 30(2), 330–340, 1994.
7. Conte, E., Lops, M., and Ricci, G., Adaptive matched filter detection in spherically invariant noise, *IEEE Signal Process. Lett.*, 3(8), 248–250, 1996.
8. Kraut, S. and Scharf, L., The CFAR adaptive subspace detector is a scale-invariant GLRT, *IEEE Trans. SP*, 47(9), 2538–2546, 1999.
9. Roman, J.R., Rangaswamy, M., Davis, D.W., Zhang, Q., Himed, B., and Michels, J.H., Parametric adaptive matched filter for airborne radar applications, *IEEE Trans. AES*, 36(2), 677–692, 2000.
10. Michels, J.H., Himed, B., and Rangaswamy, M., Performance of STAP tests in Gaussian and compound-Gaussian clutter, *Digit. Signal Process.*, 10, 309–324, 2000.
11. Pulsone, N.B. and Raghavan, R.S., Analysis of an adaptive CFAR detector in non-Gaussian interference, *IEEE Trans. AES*, 35(3), 903–916, 1999.
12. Raghaven, R.S. and Pulsone, N.B., Space-Time Adaptive Processing for Radars, Technical Report AFRL-SN-RS-TR-1998-25, Air Force Research Lab, Rome, NY, March 1998.
13. Mitchell, R.L., Importance sampling applied to simulation of false alarm statistics, *IEEE Trans. AES*, 17(1), 15–24, 1981.
14. Robert, C.P. and Casella, G., *Monte Carlo Statistical Methods*, Springer-Verlag, New York, 1999.
15. Gerlach, K., New results in importance sampling, *IEEE Trans. AES*, 35(3), 917–925, 1999.
16. Grajal, J. and Asensio, A., Multiparametric importance sampling for simulation of radar systems, *IEEE Trans. AES*, 35(1), 123–137, 1999.
17. Lank, G.W., Theoretical aspects of importance sampling applied to false alarms, *IEEE Trans. IT*, 29(1), 1983.

18. Stadelman, D.L., Keckler, A.D., and Weiner, D.D., Adaptive Ozturk-based receivers for small signal detection in impulsive non-Gaussian clutter, 1999 Conference on Signal and Data Processing of Small Targets, *Proceedings of SPIE-3809*, pp. 17–28, 1999.
19. Kotz, S. and Nadarajah, S., *Extreme Value Distributions — Theory and Applications*, Imperial College Press, London, 2000.
20. Chakravarthi, P., Subclutter Visibility: the Problem of Weak Signal Detection, Ph.D. thesis, Department of Electrical and Computer Engineering, Syracuse University, 1993.

SECTION 15.5

1. Ward, K.D., Baker, C.J., and Watts, S., Maritime surveillance radar, part 1: radar scattering form the ocean surface, *IEE Proc. F*, 137(2), 51–62, 1990.
2. Baker, C.J., *K*-distributed coherent sea clutter, *IEE Proc. F*, 138(2), 89–92, 1991.
3. Jao, J.K., Amplitude distribution of composite terrain radar clutter and the *k*-distribution, *IEEE Trans. AP*, 32(October), 1049–1062, 1984.
4. Sekine, M., Ohtani, S., Musha, T., Irabu, T., Kiuchi, E., Hagisawa, T., and Tomita, Y., Weibull-distributed ground clutter, *IEEE Trans. AES*, 17(July), 596–598, 1981.
5. Conte, E., Lops, M., and Ricci, G., Radar detection in *K*-distributed clutter, *IEE Proc. F*, 141(April), 116–118, 1994.
6. Rangaswamy, M., Weiner, D.D., and Ozturk, A., Non-Gaussian random vector identification using spherically invariant random processes, *IEEE Trans. AES*, 29(1), 111–124, 1993.
7. Rangaswamy, M., Weiner, D.D., and Ozturk, A., Computer generation of correlated non-Gaussian radar clutter, *IEEE Trans. AES*, 31(January), 106–116, 1995.
8. Yao, K., A representation theorem and its applications to spherically invariant random processes, *IEEE Trans. IT*, 19, 600–608, 1973.
9. Keckler, A.D., NonGaussian Clutter Simulation and Distrubute Approximation using Spherically Invariant Random Vectors, Ph.D. dissertation, Dept. of Electrical Engineering and Computer Science, Syracuse University, Nov. 22, 2004.
10. Brately, P., Fox, B.L., and Schrage, L.E., *A Guide to Simulation*, Springer-Verlag, New York, 1987.
11. Ozturk, A., Chakravarthi, P.R., and Weiner, D.D., On determining the radar threshold for non-Gaussian processes from experimental data, *Proceedings of the 25th Asilomar Conference on Signals, Systems, and Computers:* Pacific Grove, CA, pp. 594–598, November 1990.
12. Sangston, K.J. and Gerlach, K.R., Coherent detection of radar targets in a non-Gaussian background, *IEEE Trans. AES*, 30(2), 330–340, 1994.
13. Keckler, A.D., Non-Gaussian Clutter Modeling by Spherically Invariant Random Vectors, Technical Report, AFOSR, Rome Laboratory, Rome NY, September 1996.
14. Stademan, D.L., Ph.D. Part 2 Examination, Department of Electrical and Computer Engineering, Syracuse University, 1995.

SECTION 15.6

1. Chakravarthi, P.R., Subclutter Visibility: The Problem of Weak Signal Detection, Ph.D. dissertation, Department of Electrical and Computer Engineering, Syracuse University, 1993.

2. Kassam, S.A., *Signal Detection in Non-Gaussian Noise*, Springer-Verlag, New York, 1988.
3. Rangaswamy, M., Weiner, D.D., and Ozturk, A., Non-Gaussian random vector identification using spherically invariant random processes, *IEEE Trans. AES*, 29(1), 111–124, 1993.
4. Conte, E. and Longo, M., Characterization of radar clutter as a spherically invariant random process, *IEE Proc. F*, 132(2), 191–197, 1987.
5. Ward, K.D., Baker, C.J., and Watts, S., Maritime surveillance radar, Part 1: radar scattering from the ocean surface, *IEE Proc. F*, 137(2), 51–62, 1990.
6. Sangston, K.J. and Gerlach, K.R., Coherent detection of radar targets in a non-Gaussian background, *IEEE Trans. AES*, 30(2), 330–340, 1994.
7. Conte, E., Longo, M., Lops, M., and Ullo, S.L., Radar detection of signals with unknown parameters in *K*-distributed clutter, *IEE Proc. F*, 138(April), 131–138, 1991.
8. Pentini, F.A., Farina, A., and Zirilli, F., Radar detection of targets located in a coherent *K*-distributed clutter background, *IEE Proc. F*, 139(June), 239–245, 1992.
9. Conte, E., Lops, M., and Ricci, G., Radar detection in *K*-distributed clutter, *IEE Proc. F*, 141(April), 116–118, 1994.
10. Van Trees, H.L., *Detection, Estimation, and Modulation Theory, Part I*, Wiley, New York, 1968.
11. Stadelman, D., Coherent detection of signals in nonGaussian spherically invariant interference, Ph.D. dissertation, Dept. of Electrical and Computer Science, Syracuse University, Nov. 23, 2004.
12. Rangaswamy, M., Weiner, D.D., and Ozturk, A., Computer generation of correlated non-Gaussian radar clutter, *IEEE Trans. AES*, 31(January), 106–116, 1995.
13. Conte, E., Longo, M., and Lops, M., Modeling and simulation of non-Rayleigh radar clutter, *IEE Proc. F*, 138(April), 121–130, 1991.
14. L'Ecuyer, P., Recent advances in uniform random generation, *Proceedings of the 1994 Winter Simulation Conference*, pp. 176–183, 1994.
15. Nuth, D.E., *The Art of Computer Programming, Vol. 2: Seminumerical Algorithms*, Addison–Wesley, Reading, MA, 1981.
16. Coddington, P.D., *Analysis of Random Number Generators Using Monte Carlo Simulation*, Northeast Parallel Architecture Center, Syracuse University, 1993.
17. Miller, K.S., *Multidimensional Gaussian Distributions*, Wiley, New York, 1964.
18. Shelupsky, D., An introduction to spherical coordinates, *Am. Math. Monthly*, 69(September), 644–646, 1962.
19. Ozturk, A. and Dudewicz, E.J., A new statistical goodness-of-fit test based on graphical representation, *Biom. J.*, 34, 403–427, 1992.
20. Watts, S., Baker, C.J., and Ward, K.D., Maritime surveillance radar, part 2: detection performance prediction in sea clutter, *IEE Proc. F*, 137(April), 63–72, 1990.
21. Yao, K., A representation theorem and its applications to spherically invariant random processes, *IEEE Trans. IT*, 19, 600–608, 1973.
22. Shah, R.R., A New Technique for Distribution Approximation of Random Data, M.S. thesis, Department of Electrical and Computer Engineering, Syracuse University, 1993.

23. Oberhettinger, F. and Badii, L., *Tables of Laplace Transforms*, Springer-Verlag, New York, 1973.
24. Abramovitz, M. and Stegun, I.A., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical TablesQ*, Dover Publications, New York, 1972.
25. Gradshteyn, I.S. and Ryzhik, I.M., *Tables of Integrals, Series, and Products*, Academic Press, Boston, 1994.
26. Conte, E., Lops, M., and Ricci, G., Asymptotically optimum radar detection in compound-Gaussian clutter, *IEEE Trans. AES*, 31(April), 617–625, 1995.
27. Kelly, E.J., An adaptive detection algorithm, *IEEE Trans. AES*, 22(March), 115–127, 1986.
28. Khatri, C.G. and Rao, C.R., Test for a specified signal when the noise covariance matrix is unknown, *J. Multivar. Anal.*, 22(August), 177–188, 1987.
29. Noble, B. and Daniel, J.W., *Applied Linear Algebra*, Prentice Hall, Englewood Cliffs, NJ, 1988.
30. Nathanson, F.E., *Radar Design Principles*, McGraw-Hill, New York, 1969.
31. Anderson, T.W., *An Introduction to Multivariate Statistical Analysis*, Wiley, New York, 1984.
32. Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P., *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, New York, 1992.

Section 15.7

1. Metford, P., Haykin, S., and Taylor, D., An innovations approach to discrete-time detection theory, *IEEE Trans. IT*, 28, 376–380, 1982.
2. Conte, E., Longo, M., Lops, M., and Ullo, S., Radar detection of signals with unknown parameters in *K*-distributed clutter, *IEEE Proc. F*, 138(2), 131–138, 1991.
3. Kassam, S., *Signal Detection in Non-Gaussian Noise*, Springer-Verlag, New York, 1988.
4. Sengupta, D. and Kay, S., Efficient estimation of parameters for non-Gaussian autoregressive processes, *IEEE Trans. ASSP*, 37(6), 785–794, 1989.
5. Rangaswamy, M., Weiner, D.D., and Ozturk, A., Non-Gaussian random vector identification using spherically invariant random processes, *IEEE Trans. AES*, 29(1), 111–124, 1993.
6. Rangaswamy, M., Weiner, D.D., and Ozturk, A., Computer generation of correlated non-Gaussian radar clutter, *IEEE Trans. AES*, 31(January), 106–116, 1995.
7. Michels, J.H., Detection of partially correlated signals in clutter using a multichannel model-based approach, *Proceedings of the Natural Telesystems Conference*, Ashburn, VA, 1992.
8. Schweppe, F., Evaluation of likelihood functions for Gaussian signals, *IEEE Trans. IT*, 11(1), 61–70, 1965.
9. Michels, J.H., Multichannel Detection Using the Discrete-Time Model-Based Innovations Approach, Ph.D. thesis, Department of Electrical and Computer Engineering, Syracuse University, 1991.

10. Rangaswamy, M., Weiner, D.D., and Michels, J.H., Innovations Based Detection Algorithm for Correlated Non-Gaussian Random Processes, *Proceedings of the Sixth IEEE ASSP Workshop*, Victoria, Canada, 1992.
11. Farina, A., Russo, A., Scannapieco, F., and Barbarossa, S., Theory of radar detection in coherent Weibull clutter, *IEE Proc. F*, 134(2), 174–190, 1987.
12. Yao, K., A representation theorem and its applications to spherically invariant random processes, *IEE Trans. IT*, 19, 600–608, 1973.
13. Picinbono, B., On circularity, *IEEE Trans. SP*, 42, 3473–3482, 1994.
14. Blake, I. and Thomas, J., On a class of processes arising in linear estimation theory, *IEEE Trans. IT*, 14, 12–16, 1968.
15. Chu, K., Estimation and decision for linear systems with elliptical random processes, *IEEE Trans. AC*, 18, 499–505, 1973.
16. Rangaswamy, M., Weiner, D.D., and Michels, J.H., Multichannel detection for correlated non-Gaussian random processes based on innovations, *Proceedings of the SPIE International Symposium on Optical Engineering and Photonics in Aerospace and Remote Sensing*, Orlando, FL, 1993.
17. Michels, J.H., Synthesis of Multichannel Autoregressive Random Processes and Ergodicity Considerations, Technical Report RADC-TR0-90-211, Rome Laboratory, July 1990.
18. Farina, A., Russo, A., and Studer, F., Coherent radar detection in lognormal clutter, *IEE Proc. F*, 133(1), 39–54, 1986.
19. Liu, B. and Munson, D., Generation of a random sequence having a jointly specified marginal distribution and autocovariance function, *IEEE Trans. ASSP*, 6, 973–983, 1982.
20. Wise, G., Traganitis, A., and Thomas, J., The effect of a memoryless nonlinearity on the spectrum of a random process, *IEEE Trans. IT*, 23, 84–89, 1977.
21. Goldman, J., Detection in the presence of a spherically symmetric random vectors, *IEEE Trans. IT*, 22, 52–58, 1976.
22. Rangaswamy, M., Spherically Invariant Random Processes for Radar Clutter Modeling, Simulation, and Distribution Identification, Ph.D. thesis, Department of Electrical and Computer Engineering, Syracuse University, 1992.
23. Strand, O., Multichannel complex maximum entropy (auto-regressive) spectral analysis, *IEEE Trans. AC*, 22, 634–640, 1977.
24. Nuttall, A., Multivariate Linear Predictive Spectral Analysis Employing Weighted Forward and Backward Averaging, Technical Report TR-5501, Naval Underwater Systems Center, New London CT, October 1976.

APPENDIX B

SECTION 3.2

22. Yao, K., A representation theorem and its applications to spherically invariant random processes, *IEEE Trans.*, IT-19, 600–608, 1973.
28. Little, R.J. and Rubin, D.B., *Statistical Analysis with Missing Data*, Wiley, New York, 1987.
42. Burg, J.P., Luenberger, D.G., and Wenger, D.L., Estimation of structured covariance matrices, *Proceedings*, IEEE-70(9), 963–974, 1982.

APPENDIX I

SECTION 10.1

12. Brennan, L.E., Mallet, J.D., and Reed, I.S., Adaptive arrays in airborne MTI radar, *IEEE Trans.*, AP-24, 607, 1976.

APPENDIX O

SECTION 10.5

2. Monzingo, R. and Miller, T., Introduction to Adaptive Arrays, New York: John Wiley & Sons, p. 72 1980.

APPENDIX P

SECTION 10.10

21. Gupta, I. and Ksienski, A., Effect of mutual coupling on the performance of adaptive arrays, *IEEE Trans.*, AP-31, 785–791, 1983.

APPENDIX S

SECTION 10.11

13. Carlson, B.D., Covariance matrix estimation errors and diagonal loading in adaptive arrays, *IEEE Trans.*, AES-24, 397–401, 1988.

APPENDIX U

SECTION 12.2

26. Miller, K., *Multidimensional Gaussian Distributions*, Wiley, New York, 1964.

APPENDIX V

SECTION 12.4

30. Papoulis, A., *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York, 1984.

APPENDIX W

SECTION 12.5

49. Fang, K. and Anderson, T., *Statistical Inference in Elliptically Contoured and Related Distributions*, Allerton Press, New York, 1990.

52. Reed, I., Mallett, J., and Brennan, L., Rapid convergence rate in adaptive arrays, *IEEE Trans.*, AES-10, 853–863, 1974.
53. Khatri, C. and Rao, C., Effects of estimated noise covariance matrix in optimum signal detection, *IEEE Trans.*, ASSP-35, 671–679, 1987.

APPENDIX X

SECTION 13.3

45. de Oliviera, T., *Statistical Extremes and Applications*, Reidel, Boston, 1984.
52. Galambos, J., *Asymptotic Theory of Extreme Order Statistics*, Wiley, New York, 1978.

APPENDIX Z

SECTION 15.6

17. Miller, K.S., *Multidimensional Gaussian Distributions*, Wiley, New York, 1964.
18. Shelupsky, D., An introduction to spherical coordinates, *American Mathematical Monthly-69*, 644–646, September, 1962.

Computer Programs Available at CRC Press Website

CONTENTS

A. GENREJ — Generalized Acceptance–Rejection Method Random Number Generator 1169
(*A. D. Keckler*)
B. OZTURK — Univariate Probability Distribution Approximation Algorithm 1171
(*A. D. Keckler*)
C. OZSIRV — Multivariate Probability Distribution Algorithm for Spherically Invariant Random Vectors (SIRVs) 1174
(*A. D. Keckler*)
D. GMIXEM — Approximation of SIRVs With Gaussian Mixtures Using the Expectation–Maximization (EM) Algorithm 1177
(*A. D. Keckler*)
E. SIRVCOV — Maximum Likelihood Estimation of the Covariance Matrix for an SIRV 1179
(*A. D. Keckler*)
F. THRESHOLD — Generation of Receive Thresholds for Various False Alarm Probabilities and Sampled Unknown Noise Distributions 1181
(*P. Chakravarthi*)

These programs may be downloaded free of charge at the following Universal Resouce Locator (URL) address: http://www.crcpress.com/e_products/downloads/download.asp?cat_no=DK6045

A. GENREJ — GENERALIZED ACCEPTANCE–REJECTION METHOD RANDOM NUMBER GENERATOR

(A. D. KECKLER)

Program GENREJ uses the generalized acceptance–rejection method to generate random samples from a large range of probability distributions, as discussed in

Chapter 15. When using the acceptance–rejection algorithm, a random number, Φ, is generated from a probability density function (PDF), $f_\Phi(\phi)$, associated with a bound of the desired PDF, $f_S(s)$. The bound function is given by $\alpha f_\Phi(\phi)$, where α is the area under the bound. A random variate, Y, which is uniformly distributed on the interval $(0,\alpha f_\Phi(\phi = \Phi))$, is generated and used to form a uniformly distributed random point, (Φ, Y), in the region bounded by $\alpha f_\Phi(\phi)$ and the Φ-axis. If the point falls within the region bounded by $f_S(\Phi)$ and the Φ-axis, then the random variate, Φ, is accepted; otherwise it is rejected.

For the acceptance–rejection method to be viable, it is necessary to find a suitable easily obtained bound for which random samples can be readily generated. As a practical matter, this bound should have an area as close to unity as possible to avoid rejecting too many samples, as the efficiency of the generalized rejection algorithm is equal to $1/\alpha$. In the approach developed here a piecewise constant bound is used, as shown in Figure A.1. Samples can be generated from each segment with a simple uniform number generator. Each segment is chosen with a probability equal to its relative area. Obviously, if the PDF has a tail of infinite extent, the entire support cannot be segmented. The PDF is then divided into a body and a tail section at a point sufficiently far into the tail. The generalized Pareto PDF, given by

$$p_S(s) = \frac{1}{\sigma}\left(1 + \frac{\gamma(s - s_1)}{\sigma}\right)^{\frac{1+\gamma}{\gamma}} u(s - s_1) \tag{A.1}$$

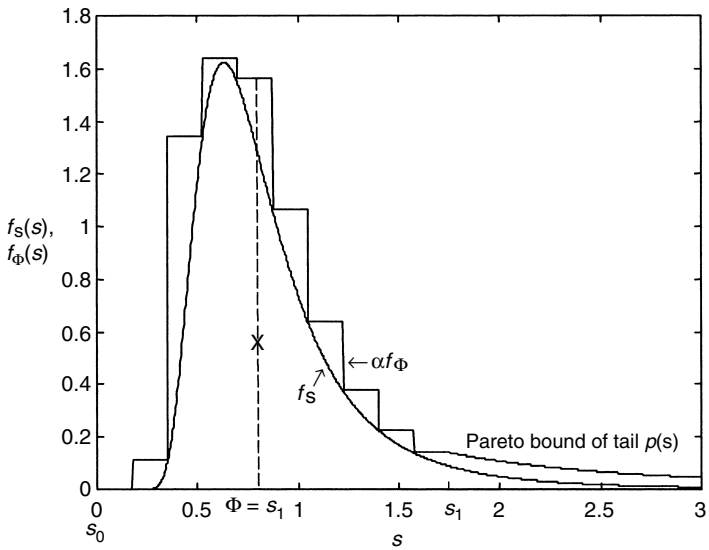

FIGURE A.1 Illustration of the acceptance–rejection method.

where σ is the sale parameter, γ is the shape parameter, and s_1 is the point where the tail is determined to begin, is used to bound the tail. The parameters, σ and γ, are obtained by matching probability weighted moments.

Program GENBND is used to generate the bound function, $\alpha f_{\Phi}(\phi)$. It is invoked in MATLAB® using the command

$$[\text{BND}] = \text{GENBND}(\text{PDF}, [A], [B], \text{NS}, \text{MARGIN}, P1, P2, \ldots)$$

The parameter PDF is a string containing the name of the program used to calculate the PDF, $f_S(s)$, which is to be bounded. The parameters A and B represent the left and right endpoints of the PDF. If A or B are passed as arrays of the form [a, inf] and [b, inf], respectively, then GENBND fits a generalized Pareto distribution to the corresponding tail and the parameters a and b represent the transition point from the body to the tail of the PDF. Typically values for a and b are selected such that the tail region has an area of 0.1. The parameter NS represents the number of piecewise constant segments used to bound the body of the PDF. Increasing the number of segments tightens the fit of the bound function, which increases the efficiency of the generalized acceptance–rejection algorithm, but also increases the cost of evaluating the bound. Typically, the number of segments, NS, is set to 10. The parameter MARGIN represents a small factor used to raise the bound to account for numerical inaccuracies and guarantee that the generated function is actually greater than the PDF across its entire support. The optional parameters $P1$, $P2$, etc., represent parameters of the PDF. Program GENBND returns a structure, BND, containing parameters of the fitted bound.

Once the bound function has been created, random variates for the PDF, $f_S(s)$, are generated by invoking the program GENREJ using the command,

$$[\text{S}, \text{EFF}] = \text{GENREJ}(\text{PDF}, N, \text{M}, \text{BND}, U, P1, P2, \ldots)$$

Again, the parameter PDF is a string containing the name of the PDF program, and BND is the structure generated by GENBND for this PDF. The parameters N and M are the dimensions of the random matrix to be generated. The optional parameter U specifies the source of the univariate random numbers used by the acceptance–rejection algorithm, and again the optional parameters $P1$, $P2$, etc., represent parameters of the PDF.

B. OZTURK — UNIVARIATE PROBABILITY DISTRIBUTION APPROXIMATION ALGORITHM

(A. D. KECKLER)

Program OZTURK approximates the underlying probability distribution of a collection of random samples using hypothesis testing to assess whether

a distribution is statistically consistent with the data sample via a goodness-of-fit test. The goodness-of-fit test for the Öztürk algorithm is based upon the sample order statistics, which act as estimators for the $[k/(K+1)] \times 100$ percentile of the sample data. The sample order statistics are compared against the expected order statistics produced by a reference distribution. The relative order statistics are plotted in a two-dimensional space to produce a unique curve, which is formed as a set of linked vectors such that any deviations from the reference are cumulative. This dramatically increases the power of the test. The endpoint of the linked vector is an accumulation of all the deviations observed along the length of the curve. The endpoint is used as the test statistic and forms a unique point in the associated two-dimensional space for each distribution family tested. The expected values of the endpoints generated for the hypotheses are used to identify candidate distributions. This is illustrated in Figure B.1.

The expected value of the test statistic for the Öztürk goodness-of-fit test produces a unique point in the associated space. This can be exploited when the null hypothesis is rejected. The expected value of the test statistic for a hypothesis with the nearest endpoint to the test statistic of the data is considered a likely candidate. If the test statistic from the data falls within the specified confidence contour of the closest hypothesis, then the hypothesis is accepted. If not, the next few closest neighbors are checked since the confidence contours may not be the same for neighboring distributions. The Öztürk approximation chart is constructed by calculating the expected value of the test statistic for a library of normalized (with respect to scale and location parameters) distributions. Distributions that do not have shape parameters appear as single points on the Öztürk chart. For distributions that have a shape parameter, the

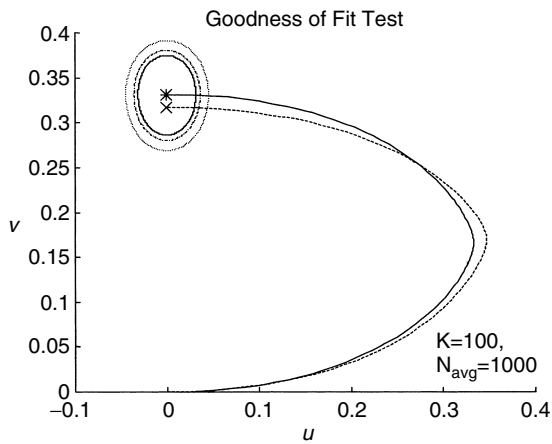

FIGURE B.1 Öztürk goodness-of-fit linked vector with 90, 95, and 99% confidence intervals for the standard Gaussian distribution.

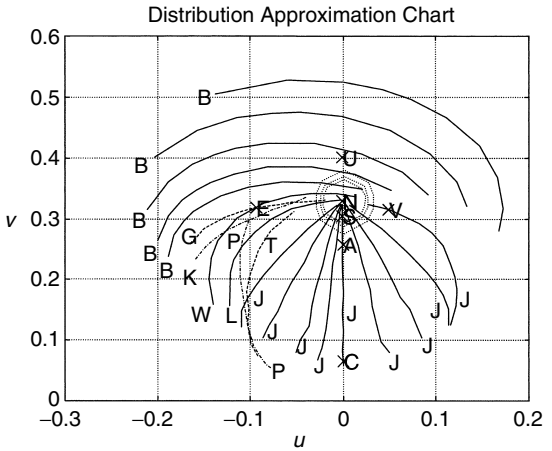

FIGURE B.2 Öztürk Distribution approximation chart for univariate distributions: B = Beta, J = SU Johnson, G = Gamma, K = K-distribution, P = Pareto, L = Log-normal T = Gumbel, E = Exponential, V = Extreme Value, A = Laplace, S = Logistic, U = Uniform, N = Normal, W = Weibull, C = Cauchy.

shape parameter is varied to produce points on the chart which are connected with straight line segments to form trajectories. Each point on a trajectory is considered a separate hypothesis even though they are drawn from the same distribution. A trajectory can be considered as a family of distributions having the same form but different shape parameters. For distributions with two shape parameters, one shape parameter is held constant while the other is varied to form a trajectory, resulting in multiple trajectories. The Öztürk approximation chart is illustrated in Figure B.2.

When program OZTURK is invoked, it brings up a graphical user interface (GUI) for generating and displaying the Öztürk approximation chart. A collection of random samples stored in a MATLAB® vector can be loaded using the LOAD tab, whereupon program OZTURK generates the appropriate Öztürk chart and plots the test statistic for the random data sample on it, identifying the best approximations for the underlying distribution of the random data sample. The library of probability distributions available to program OZTURK is listed in Table B.1. The Öztürk algorithm can provide meaningful results with as few as 40 random samples, but it has been observed that the best results are obtained for sample sizes of 100 or more. It should be noted that the Öztürk chart is unique for each sample size, and must be regenerated if the sample size changes, which can be a lengthy calculation. Therefore, it is advisable to process all data sets of the same size together.

TABLE B.1
Univariate Öztürk Chart Library of Probability Distributions

Distribution	Öztürk Chart Symbol	Standardized PDF, $f_X(x)$
Beta	B	$\frac{\mathbf{x}^{\nu-1}(1-\mathbf{x})^{\lambda-1}}{\beta(\nu,\lambda)}u(\mathbf{x})u(\mathbf{x}-1)$
Cauchy	C	$\frac{1}{\pi(1+\mathbf{x}^2)}$
Exponential	E	$e^{-\mathbf{x}}u(\mathbf{x})$
Extreme value I	V	$e^{-\mathbf{x}}\exp(-e^{-\mathbf{x}})$
Extreme value I (Gumbel, Frechet)	T	$\nu\mathbf{x}^{-\nu-1}\exp(-\mathbf{x}^{-\nu})u(\mathbf{x})$
Gamma	G	$\frac{1}{\Gamma(\nu)}\mathbf{x}^{\nu-1}e^{-\mathbf{x}}u(\mathbf{x})$
Johnson S.U.	J	$\lambda\, e^{\frac{\lambda^2}{2}(\sinh^{-1}(\mathbf{x})-\nu)^2}(2\pi(1+\mathbf{x}^2))^{-\frac{1}{2}}$
K-distribution	K	$\frac{2}{\Gamma(\nu)}\left(\frac{\mathbf{x}}{2}\right)^{\nu}K_{\nu-1}(\mathbf{x})u(\mathbf{x})$
Laplace	A	$\frac{1}{2}e^{-\lvert\mathbf{x}\rvert}$
Logistic	S	$e^{-\mathbf{x}}(1+e^{-\mathbf{x}})^{-2}$
Log-normal	L	$\frac{1}{\nu\sqrt{2\pi}}\mathbf{x}^{-1}e^{-\frac{(\ln\mathbf{x})^2}{2\nu^2}}u(\mathbf{x})$
Normal	N	$\frac{1}{\sqrt{2\pi}}e^{-\frac{\mathbf{x}^2}{2}}$
Pareto	P	$\frac{\nu}{\mathbf{x}^{\nu+1}}u(\mathbf{x}-1)$
Uniform	U	$u(\mathbf{x})-u(\mathbf{x}-1)$
Weibull	W	$\nu\mathbf{x}^{\nu-1}e^{-x^{\nu}}u(\mathbf{x})$

C. OZSIRV — MULTIVARIATE PROBABILITY DISTRIBUTION ALGORITHM FOR SPHERICALLY INVARIANT RANDOM VECTORS (SIRVs)

(A. D. KECKLER)

The PDF of a Spherically Invariant Random Vector (SIRV), $\mathbf{Y}$, of dimension N is given by

$$f_{\mathbf{Y}}(\mathbf{y}) = (2\pi)^{-\frac{N}{2}}|\Sigma_{\mathbf{Y}}|^{-\frac{1}{2}}h_N(q) \tag{C.1}$$

where $\mathbf{b}$ is the mean vector $\sum_{\mathbf{Y}}$ is the covariance matrix of $\mathbf{Y}$, respectively. The quadratic form, q, is given by $q = q(\mathbf{y}) = (\mathbf{y} - \mathbf{b})^{\mathrm{T}} \Sigma_{\mathbf{Y}}^{-1} (\mathbf{y} - \mathbf{b})$, and the function $h_N(q)$ is given by

$$h_N(q) = \int_0^{\infty} s^{-N} \, \mathrm{e}^{-\frac{q}{2s^2}} f_{\mathrm{S}}(s) \mathrm{d}s \tag{C.2}$$

where $f_{\mathrm{S}}(s)$ is the characteristic PDF of the SIRV and is unique for each type of SIRV. Implicit in Equation C.1 is the assumption that $E\{S^2\}$ is unity. The PDF of the corresponding Spherically Symmetric Random Vector (SSRV), $\mathbf{X}$, where $\mathbf{b} = \mathbf{0}$ and $\Sigma_{\mathbf{X}} = \mathbf{I}$, is given by

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{\frac{N}{2}} h_N(q) \tag{C.3}$$

where $q = \mathbf{x}^{\mathrm{T}}\mathbf{x}$ is characterized completely by $h_N(q)$. The SIRV, $\mathbf{Y}$, can be related to the SSRV, $\mathbf{X}$, by the whitening transformation

$$\mathbf{X} = \Sigma_{\mathbf{Y}}^{-\frac{1}{2}} (\mathbf{Y} - \mathbf{b}) \tag{C.4}$$

Thus, the PDF of $\mathbf{X}$ can be considered the standard form of the PDF, which defines the type of SIRV.

Since the characteristic PDF of an SIRV, $f_{\mathrm{S}}(s)$, is unique to each type of SIRV, $h_N(q)$ is also unique and Equation C.1 shows that the PDF of the SIRV, $\mathbf{Y}$, is completely characterized by specification of the mean vector, $\mathbf{b}$, the covariance matrix, $\Sigma_{\mathbf{Y}}$, and the function, $h_N(q)$. The quadratic form is also invariant to the mean vector, $\mathbf{b}$, and the covariance matrix, $\Sigma_{\mathbf{Y}}$. The PDF of the quadratic form is given by

$$f_{\mathrm{Q}}(q) = \frac{q^{\frac{N}{2}-1}}{2^{\frac{N}{2}} \Gamma\left(\frac{N}{2}\right)} h_N(q) u(q) \tag{C.5}$$

The PDF of the quadratic form given in Equation C.5 is uniquely determined by $h_N(q)$, and contains all the information required to uniquely identify the PDF of the SSRV.

In general the mean vector, $\mathbf{b}$, and the covariance matrix, $\Sigma_{\mathbf{Y}}$, which are required to form the quadratic form, $q = (\mathbf{y} - \mathbf{b})^{\mathrm{T}} \Sigma_{\mathbf{Y}}^{-1} (\mathbf{y} - \mathbf{b})$, are unknown, Thus, the quadratic form, q, is not directly observable. Instead, an estimate of the quadratic form, $\hat{q}$ is substituted, where the sample mean and covariance matrix have been substituted, such that $\hat{q} = (\mathbf{y} - \bar{\mathbf{y}})^{\mathrm{T}} \bar{\mathrm{S}}_{\mathbf{y}}^{-1} (\mathbf{y} - \bar{\mathbf{y}})$, where the sample mean is given by $\bar{\mathbf{y}} = \frac{1}{K} \sum_{i=1}^{K} \mathbf{y}_i$ and the sample covariance matrix is given by $\bar{\mathrm{S}}_{\mathrm{y}} = \left[\frac{1}{K-1} \sum_{i=1}^{K} (\mathbf{y}_i - \bar{\mathbf{y}})(\mathbf{y}i - \bar{\mathbf{y}})^{\mathrm{T}}\right]$. Samples of SIRVs are first converted to equivalent univariate distributions, using estimates of the quadratic form, and the procedure described in Computer programs, Section B is used to approximate the underlying SIRV distribution. Essentially, the distributions, for which the test statistic are plotted on the Öztürk chart, are the distributions of the estimate, $\hat{q}$, and not those of the actual quadratic form. The results are then related back to

the parent SIRV distributions. Therefore, the linked vector endpoint of each hypothesis is constructed using the estimates of the quadratic form, $\hat{q}$. This models the error introduced to the quadratic form when the sample mean and sample covariance matrix have been substituted. In this way the errors are accounted for by the Öztürk chart.

The PDFs of the SSRV, $f_{\mathbf{X}}(\mathbf{x})$ and the quadratic form, $f_Q(q)$, are expressed in terms of the function, $h_N(q)$. Thus, the quadratic form is the straightforward choice for the statistic for determining an approximate PDF for the SIRV. However, empirical results suggest the generalized envelope of the SSRV, R, may provide better discrimination on the Öztürk chart. The generalized envelope of the SSRV, $\mathbf{X}$, can be expressed in terms of the SIRV, $\mathbf{Y}$, by $r_{\mathrm{w}} = \sqrt{q} = [(\mathbf{y} - \mathbf{b})^{\mathrm{T}} \Sigma_{\mathbf{Y}}^{-1} (\mathbf{y} - \mathbf{b})]^{\frac{1}{2}}$, where r_{w} is termed the whitened envelope. An example of the Öztürk approximation chart for SIRVs is illustrated in Figure C.1.

Similar to program OZTURK described in Computer Programs, Section, the program OZSIRV brings up a graphical user interface (GUI) when invoked. A collection of random vectors stored in row–column format as a MATLAB® matrix can be loaded using the LOAD tab, whereupon program OZSIRV generates the appropriate Öztürk SIRV approximation chart and plots the test statistic for the random data samples on it, identifying the best approximations for the underlying SIRV distribution. The library, of SIRV probability distributions available to program OZSIRV and the shape parameters used to generate the Öztürk SIRV approximation charts, is listed in Table C.1. The Öztürk SIRV approximation chart is specific to both the number of vectors in the random data sample and the dimension of the random vectors, and must be regenerated if the sample size or vector size changes from test to test.

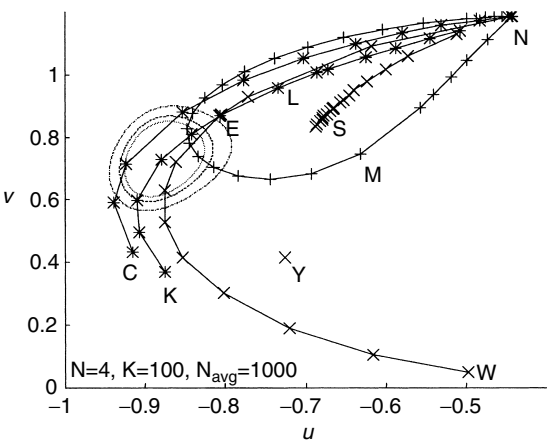

FIGURE C.1 Example Öztürk approximation chart using whitened envelopes for $N = 4$.

TABLE C.1
Öztürk SIRV Approximation Chart Distributions and Shape Parameters

SIRV Distribution	Symbol on Öztürk Chart	Shape Parameters
K-distributed	K	0.1, 0.15, 0.2, 0.3, 0.4, 0.5, 0.75, 1, 1.1, 1.5, 2, 3, 5
Chi	C	0.1, 0.15, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.99
Contaminated Gaussian ($s_c = 6$)	M	0, 0.01, 0.02, 0.03, 0.04, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 0.96, 0.97, 0.98, 0.99, 1
Student-t	S	1.1, 1.12, 1.14, 1.16, 1.18, 1.2, 1.22, 1.24, 1.26, 1.28, 1.3, 1.35, 1.4, 1.45, 1.5, 1.7, 2.0, 2.5, 5.0
Gaussian	N	N/A
Exponential	E	N/A
Laplace	L	N/A
Weibull	W	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 1, 1.1, 1.5, 2.0
Cauchy	Y	N/A

D. GMIXEM — APPROXIMATION OF SIRVs WITH GAUSSIAN MIXTURES USING THE EXPECTATION–MAXIMIZATION (EM) ALGORITHM

(A. D. KECKLER)

Often, the probability density function (PDF) of an SIRV is mathematically complex, and it is desirable to approximate it with a simpler distribution, and a discrete Gaussian mixture SIRVG can often be used. The PDF of the normalized discrete Gaussian mixture SIRV is given by

$$f_{\mathbf{X}}(\mathbf{x}) = (2\pi)^{-\frac{N}{2}} \sum_{k=1}^{K} \frac{w_k}{s_k^N} \mathrm{e}^{\frac{q}{2S_k^2}} \tag{D.1}$$

For $w_k > 0$, $\sum_{k=1}^{K} w_k = 1$, and with covariance matrix $\Sigma_{\mathbf{X}} = \mathbf{I}$ and mean vector $\mathbf{b} = \mathbf{0}$. Once and adequate approximation has been obtained, analytical results for receiver performance and other problems of interest can often be obtained for the discrete Gaussian mixture SIRV, due to its simple form. If analytical results cannot be obtained, computer simulation of the discrete Gaussian mixture SIRV is readily performed. The approximation by a discrete Gaussian mixture SIRV can be made arbitrarily accurate but may require many terms in the mixture, especially if accuracy is required in the tails of the PDF. The main difficulty encountered with using the discrete Gaussian mixture SIRV as an approximation is in performing the fit. In one approach, the Expectation Maximization (EM)

algorithm is applied to the problem of approximation a SIRV. Program GMIXEM fits the quadratic form PDF of the discrete Gaussian mixture to the corresponding PDF of the SIRV being approximated.

In this approach, the error between the PDFs of the quadratic form is minimized. If a set of data is obtained from the SIRV distribution being approximated and each sample is considered to have been generated by a specific component of the discrete Gaussian mixture SIRV, then the approximation problem can be viewed as an estimation involving incomplete data. Each element of the complete data consists of the random sample generated and the component, k, from which it was drawn. The observed data, however, consists only of the random sample and, therefore, is incomplete. An elegant solution to this problem is provided by the Expectation–Maximization (EM) algorithm, which is an iterative method that produces maximum likelihood (ML) estimates of the parameters. An example of an approximation of an SIRV using this approach is shown in Figure D.1.

Program GMIXEM is called from MATLAB® using

$$[\mathrm{W}, \mathrm{S}, \mathrm{ES2}] = \mathrm{GMIXEM}(Q, K, W0, V0, \mathrm{TOL}, \mathrm{MAXITER})$$

The parameter Q is a vector of random samples of the quadratic form drawn from the SIRV distribution to be approximated. The parameter K is the number of components in the Gaussian mixture SIRV. The parameters $W0$ and $V0$ are initial guesses for the component weights and variances, respectively. TOL specifies the convergence criteria, with regards to the relatives change from iteration to iteration, and MAXITER is the maximum number of iterations for the EM algorithm.

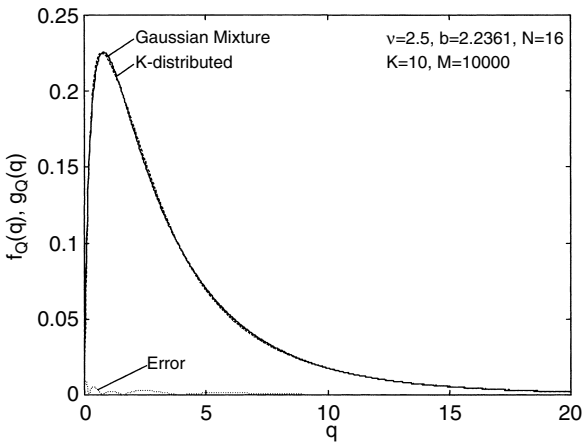

FIGURE D.1 Normalized K-distributed SIRV quadratic form PDF and EM Gaussian mixture approximation.

While the approximation using the EM algorithm is determined based upon a particular order of the quadratic form PDF, the resulting approximation is not restricted to that value of N because the approximation is a discreate Gaussian mixture SIRV that is defined for any $N > 0$. It is recommended, however, that the approximation be obtained using PDFs of the order of interest, as the error in the approximation will not necessarily be invariant with respect to N.

E. SIRVCOV — MAXIMUM LIKELIHOOD ESTIMATION OF THE COVARIANCE MATRIX FOR AN SIRV

(A. D. KECKLER)

The maximum likelihood (ML) estimate of the covariance matrix is obtained using K independent and identically distributed (IID) realizations of the SIRV, designated as $\mathbf{y}_k$ for $k = 1, 2, \ldots K$, and assuming the form of the PDF of the SIRV, $\mathbf{Y}$, is known and that the mean vector $\mathbf{b}$ is zero. With the mean vector $\mathbf{b} = \mathbf{0}$, the SIRV, $\mathbf{Y}$, can be expressed as

$$\mathbf{Y} = S\mathbf{W} \tag{E.1}$$

where S is a positive, independent random multiplier, and $\mathbf{W}$ is a zero-mean Gaussian random vector with covariance matrix, $\Sigma_{\mathbf{W}}$. The covariance matrix of the SIRV, $\mathbf{Y}$, is given by

$$\Sigma_{\mathbf{Y}} = E\{S^2\}\Sigma_{\mathbf{W}} \tag{E.2}$$

and, thus, $\Sigma_{\mathbf{Y}} = \Sigma_{\mathbf{W}}$ when the SIRV is normalized such that $E\{S^2\}$ is unity.

It is assumed that each realization of the SIRV, $\mathbf{Y}$, has been obtained from an independent realization of the random multiplier, S, as well as independent realizations of the Gaussian random vector, $\mathbf{W}$. If each realization, s_k, of the random multiplier is known, then the realizations $\mathbf{w}_k$ of the Gaussian random vector, $\mathbf{W}_k$ can be obtained from the observed realizations of the SIRV such that

$$\mathbf{w}_k = \frac{\mathbf{y}_k}{S_k} \tag{E.3}$$

It is well known that the ML estimate of the covariance matrix, $\Sigma_{\mathbf{W}}$, for the zero-mean Gaussian random vector $\mathbf{W}$ is the sample covariance matrix, such that

$$\hat{\Sigma}_{\mathbf{W}} = \left[\frac{1}{K}\sum_{k=1}^{K}\mathbf{W}_k\mathbf{W}_k^{\mathrm{T}}\right] \tag{E.4}$$

Since the covariance matrix of the SIRV, $\mathbf{Y}$, is related to the covariance matrix of the underlying Gaussian random vector, $\mathbf{W}$, by Equation E.2, the ML estimate of

$\Sigma_{\mathbf{Y}}$ is given by

$$\hat{\Sigma}_{\mathbf{Y}} = E\{S^2\}\hat{\Sigma}_{\mathbf{W}} \tag{E.5}$$

which, if the SIRV is normalized such that $E\{S^2\}$ is unity, reduces to

$$\hat{\Sigma}_{\mathbf{Y}} = \hat{\Sigma}_{\mathbf{W}} \tag{E.6}$$

In practice, however, the realizations of the random multiplier, S_k, are not known, and estimation of the convariance matrix, $\Sigma_{\mathbf{W}}$, becomes an estimation problem with incomplete data. This is similar to the problem of estimating the parameters of the discrete Gaussian mixture SIRV approximation discussed in Computer Programs, Section D.

The ML estimate of $\Sigma_{\mathbf{W}}$ can be obtained from the incomplete data iteratively by applying the EM algorithm. The complete data consists of pairs of the observed SIRVs, $\mathbf{y}_k$, and the corresponding realizations of the independent multiplier, s_k, for $k = 1, 2, \ldots, K$. The incomplete data consists of just the realizations of the SIRV, $\mathbf{y}_k$, for $k = 1, 2, \ldots, K$. If the complete data were available, it would be possible to determine the Gaussian vectors, $\mathbf{w}_k$, directly and then the sample covariance matrix could be used to estimate $\Sigma_{\mathbf{W}}$. However, since they are not, the expectation of the log-likelihood function for the complete data is required, conditioned on the incomplete data. The resulting ML estimate for the covariance matrix is obtained iteratively, and is given by

$$\begin{aligned}\hat{\Sigma}_{\mathbf{W}}(i+1) &= \frac{1}{\mathbf{K}}\sum_{k=1}^{K}\frac{h_{N+2}(\mathbf{y}_k^{\mathrm{T}}[\hat{\Sigma}_{\mathbf{W}}(i)]^{-1}\mathbf{y}_k)}{h_N(\mathbf{y}_k^{\mathrm{T}}[\hat{\Sigma}_{\mathbf{W}}(i)]^{-1}\mathbf{y}_k)}\mathbf{y}_k\mathbf{y}_k^{\mathrm{T}} \\ &= -\frac{2}{\mathbf{K}}\sum_{k=1}^{K}\frac{h'_N(\mathbf{y}_k^{\mathrm{T}}[\hat{\Sigma}_{\mathbf{W}}(i)]^{-1}\mathbf{y}_k)}{h_N(\mathbf{y}_k^{\mathrm{T}}[\hat{\Sigma}_{\mathbf{W}}(i)]^{-1}\mathbf{y}_k)}\mathbf{y}_k\mathbf{y}_k^{\mathrm{T}}\end{aligned} \tag{E.7}$$

Equation E.7 is iterated until convergence is reached, such that $\hat{\Sigma}_{\mathbf{W}}(i+1) = \hat{\Sigma}_{\mathbf{W}}(i) = \hat{\Sigma}_{\mathbf{W}}$. The convergence criterion can be specified using the Frobenius norm of the difference between the estimates of the covariance matrix from successive iterations. The Frobenius norm is a matrix norm defined to be equal to the square root of the sum of the absolute squares of its elements, and can be calculated using

$$\|\mathbf{A}\|_{\mathrm{F}} = \sqrt{\mathrm{tr}(\mathbf{A}\mathbf{A}^{\mathrm{H}})} \tag{E.8}$$

Where $\mathrm{tr}(\cdot)$ is the matrix trace, and $\mathbf{A}^{\mathrm{H}}$ is the conjugate transpose of the matrix, $\mathbf{A}$. Therefore, the convergence criterion using Equation E.8 is given by

$$\|\hat{\Sigma}_{\mathbf{W}}(i+1) - \hat{\Sigma}_{\mathbf{W}}(i)\|_{\mathrm{F}} < \eta \tag{E.9}$$

where η is a threshold specifying the allowable error.

Once the ML estimate for the covariance matrix, $\Sigma_{\mathbf{W}}$, of the underlying Gaussian vector, $\mathbf{W}$, has been determined, the estimate of the covariance matrix for the SIRV, $\mathbf{Y}$, is obtained using

$$\hat{\Sigma}_{\mathbf{Y}} = E\{S^2\}\hat{\Sigma}_{\mathbf{W}} \tag{E.10}$$

which, if the SIRV is normalized such that $E\{S^2\} = 1$, reduces to

$$\hat{\Sigma}_{\mathbf{Y}} = \hat{\Sigma}_{\mathbf{W}} \tag{E.11}$$

It should be noted, since the PDF of the SIRV, $\mathbf{Y}$, is assumed known by the EM algorithm, that $E\{S^2\}$ does not need to be estimated. Therefore, the ML estimate, $\hat{\Sigma}_{\mathbf{Y}}$, of the covariance matrix for the SIRV, $\mathbf{Y}$, is obtained directly from the ML estimate of the covariance matrix, $\hat{\Sigma}_{\mathbf{W}}$, of the underlying Gaussian random vector, $\mathbf{W}$.

The EM algorithm converges to a local minimum, and is dependent upon the initial guess for the estimate, $\hat{\Sigma}_{\mathbf{W}}(0)$. Any symmetric positive definite matrix can be used for the initial guess, with the $N \times N$ identity matrix, $\mathbf{I}$, or the sample covariance matrix, $\tilde{\mathbf{S}}_{\mathbf{y}} = \left[\frac{1}{K}\sum_{k=1}^{K} \mathbf{y}_k \mathbf{y}_k^{\mathrm{T}}\right]$, as two readily available choices.

The ML estimates for the covariance matrix of a set of SIRV samples is obtained using the function call

$$[C, S, E1, E2] = \mathrm{EMCOV}(X, \mathrm{HN}, \mathrm{C0}, \mathrm{TOL}, \mathrm{MAXITER}, P1, P2, \ldots)$$

In this call, the parameter X is a MATLAB® matrix whose K rows contain samples of SIRVs, each of dimension N. The parameter HN is a string containing the name of the program used to evaluate the characteristic function, $h_N(q)$, of the SIRV, C0 is the initial guess for the covariance matrix, which must be a symmetric positive definite matrix of dimension $N \times N$. The parameter TOL is the convergence criteria indicated in Equation E.9, and MAXITER is the maximum number of iterations allowed before termination of the algorithm. Finally, the parameters $P1$, $P2$, etc., are parameters for the characteristic function specified by HN.

F. THRESHOLD — GENERATION OF RECEIVE THRESHOLDS FOR VARIOUS FALSE ALARM PROBABILITIES AND SAMPLED UNKNOWN NOISE DISTRIBUTIONS

(P. CHAKRAVARTHI)

Program THRESHOLD computes the thresholds corresponding to various false alarm rates when there is no information on the underlying noise distribution. For signal detection in non-Gaussian noise it is important to accurately estimate the threshold corresponding to a given false alarm probability. The estimation of these thresholds becomes especially hard when the underlying noise distribution is unknown and the false alarm probabilities are very small, i.e., $< 10^{-5}$.

Extreme Value Theory is an approach that enables estimating the tail of probability density functions with far samples than conventionally required by the Monte Carlo method. Generating typically between 1000 and 20,000 samples of the underlying noise, the tail of the non-Gaussian clutter is fitted by the Generalized Pareto Distribution (GPD) described in Section 13.3.5. The user supplies as input a column vector of observation samples from the underlying noise distribution and the false alarm rates, as well as specifies the method to use in estimating the parameters of the GPD. In particular, the Maximum Likelihood (ML) method and the Ordered Sample Least Squares (OSLS) method are used for estimating the GPD parameters.

Let $\mathbf{x} \in \mathbf{R}$ denote a column vector composed of the observation samples of the noise and α denote the desired false rate. Program thresholdByGPD ($\mathbf{x}$, α, 'ML') will return the estimated threshold at false alarm rate α through estimating the GPD parameters by the ML method. Similarly, program thresholdByGPD ($\mathbf{x}$, α, 'OSLS') gives the result through the OSLS method.

The program is written in MATLAB and can be executed on a PC.