Adaptive Array Signal Processing and Performance Analysis in Non-Gaussian Environments

by Christ D. Richmond

S.B. Mathematics Bowie State University, 1990
S.B. Electrical Engineering University of Maryland College Park, 1990
S.M./E.E. Electrical Engineering Massachusetts Institute of Technology, 1993/1995

Submitted to the Department of Electrical Engineering and Computer Science in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

at the Massachusetts Institute of Technology

May 1996

Signature of Author ______________________
Department of Electrical Engineering and Computer Science
May 21, 1996

Certified by ______________________
Arthur B. Baggeroer, *Ford Professor* of Electrical Engineering and Ocean Engineering
Department of Electrical Engineering and Computer Science
Thesis Advisor

Certified by ______________________
Prof/Dr. Allan O. Steinhardt, Technical Research Staff, MIT Lincoln Laboratory
Co-thesis Advisor

Accepted by ______________________
Professor F. R. Morgenthaler
Chairman of the Department Graduate Committee

Adaptive Array Signal Processing and Performance Analysis in Non-Gaussian Environments

by Christ D. Richmond

Submitted to the Department of Electrical Engineering and Computer Science Massachusetts Institute of Technology in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Ph.D.)

Abstract

Adaptive array signal processing has been based historically on the premise of data Gaussianity. This is primarily due to (1) the analytic tractability of the Gaussian, (2) the wealth of Gaussian based statistical sampling theory available and (3) the central limit theorem which often makes for good approximations even in non-Gaussian environments. The widely known results on adaptive array (i) signal detection (ii) signal estimation/beamforming (iii) power spectral estimation, and (iv) structured covariance estimation are all Gaussian based. It has been known for some time, however, that the Gaussian distribution does not always adequately model the data obtained in several radar, sonar, and other array application areas. In spite of this statistical mismatch, Gaussian based processors have demonstrated robust performance. Consequently, there exists a need for consideration of non-Gaussian models which will lend insight into optimal structures and performance issues surrounding adaptive arrays. Non-Gaussian models have been considered by a number of authors. None, however, have considered the problem of *adaptive* array detection/estimation as originally posed by Reed, Mallett, and Brennan and later by Kelly. Specifically, previous authors consider non-adaptive situations and do not employ nor propose the use of a theoretical framework which allows for general estimation of the data covariance, typically an unknown parameter in adaptive array scenarios, via a *secondary (or training)* data set. In this thesis we ergo reformulate several problems of adaptive array processing under a weaker assumption of complex multivariate elliptically contoured (MEC) distributed data, and subsequently (i) derive optimal processing structures, (ii) consider in detail their performance analysis in these non-Gaussian environments.

Thesis Supervisor: Arthur B. Baggeroer
Title: Ford Professor of Electrical Engineering and Ocean Engineering

Co-Advisor: Dr. Allan O. Steinhardt
Title: Former Professor Cornell University and present MIT LL Technical Staff

Acknowledgements

Many persons have contributed to my success and well being while pursuing this degree. First I'd like to thank my Dad, Dr. Clinton T. Richmond, for being a great example and support to me ever since I can remember. I recall him teaching my brother and me algebra, while we were still in elementary school...fourth grade in fact! Not long after, he'd teach us aspects of what I now know to be vector calculus. As an eigth-grader, I remember the many nights I sat up at the kitchen table with my Dad trying desparately to understand the initially elusive methodology for solving word problems. It has all paid off Dad. Math and science have never intimidated me, because I saw it come home everyday with the wonderbread, pepsi colas, hotdogs, and pork n' beans my Dad brought with him. Thanks Dad, your enthusiasm with physics and science in general played a major role in the directions I decided to take in life, including the pursuit of the present doctorate degree.

Second I thank my Mom, Gloria H. Snoddy, for her unstinted love and support. A voice of sound reason, compassion, love, and sensitivity. Whether I fail or succeed, whether I'm happy or sad, whether the sun shines or rain falls, her song is always the same...an unconditional love invariant to any circumstance. Thank you for believing in me Mom! Thank you for a love I can only compare to God's.

My thanks to my immediate family: Clinton T. Richmond Jr. (brother), Vivie, Katonya, and Kiwan Richmond (sisters), and Dejuene J. Jones (niece). Thank you for being my brother and sisters. Special thanks to my brother Clinton for the many personal pep talks, and big brother advice given over the years. Thanks for believing in me Clinton, as I also believe in you! That's at least one aspect of life we don't have to figure out...the indispensible need to love, support, and care for one another. To neglect this need is to risk the possibility of our own peril.

Special thanks to Prof Arthur B. Baggeroer for providing guidiance, and support in the midst of personal difficulties. Thanks Art, for pressing your way to my thesis Defense against odds. My prayer's are with you and for you. Hang in there, and try to believe and to remember that it will pass.

My true sincere thanks to Prof Allan O. Steinhardt for his constant, insightful, fueling, and enthusiastic support throughout this entire thesis. He is truly a friend that has given new life to my graduate school experience, and set an example in many ways I can only hope to imitate. Thanks Allan, I'm so glad our paths did intersect with timing I can only describe as miraculous!

Thanks to Pauline Bennett for being such a good friend and confidant. Always enthusiastic, supportive, edifying, and strengthening. My thanks likewise for participating in and contributing to the several meetings on adaptive beamforming. You're next my friend!!!

My thanks to Prof George C. Verghese and Prof Gregory W. Wornell for serving on my committee.

My thanks to Dr. M. Rangaswamy for initiating my interest in non-Gaussian models last summer (1995). I had no idea, and certainly nor did he, that I would catch so many fish by simply casting my net on the other side of the boat!

My thanks to Prof Barry D. Van Veen at the University of Wisconsin for his instructive feedback given on the adaptive beamforming results.

My thanks to the folks at Lincoln Labs for support, without which this work would have taken much longer to complete. Thanks to Dr. Steve Krich for agreeing to support me financially as a Research Assistant. Thanks also to my friend "hyper-drive" Dr. Steven Smith for lending an ear at my sporadic requests. Thanks to Doctors Edward Baranoski, Jim Ward, Michael Zatman, Dan Rabideau, Tim Barton, and others for participation in the signal processing seminars, from which I'm still learning.

Thanks to Pastor Linda Pitt for her many prayers and support. Thank you Pastor Pitt for believing in me, and helping me to reinforce when my citadel of confidence was challenged. Also thanks to Evangelist Gloria Lindsay and Interdenominational Church for their many prayers.

My thanks to Dr. Steven Isabelle for being a mentor, good friend, and big brother to me during my entire time at MIT. Thanks for lending an ear, sharing your wisdom, and giving the right advice at key points in my MIT experience. It made a difference!

My thanks to those who have served as refueling tanks socially, mentally, and/or spiritually: Pastors Gilbert and Ivonne Thompson and the NCCC family, George and Lynnette Russell and the NCCC musicians, Youth Pastors Chris and Joy Turner-Hill, Reggie Brothers, J. Ofori-Tenkorang, Deron Jackson, Robert Selders, and Cedric Logan.

My thanks and regards to the faculty of Bowie State University's (BSU) Math and Science Department, who encouraged me to stand up and pilot my ship because I was one who could, when I was planning, after a wearisome high school experience, to simply be a passenger along for the ride. My thanks to a brother who was in BSU's Adminstrations department at the time, but whose name escapes me. For he was the first to tell of MIT's existence. My thanks to the many esteemed African Americans out there who always strive for excellence, but may never be known as those who have appeared on television, or have been written about in the newspapers. In the words of Dr. Maya Angelou, "Black Stallions," indeed.

My thanks to Prof Blankenship of the University of Maryland College Park, whom I've met only once. He told me of the existence of the Office of Naval Research (ONR) fellowship program a couple weeks before the application deadline. Needless to say, I was one of 55 awarded this prestigious fellowship competing with 863 applicants on a national level. It was of great benefit to me. Indeed, my thanks to ONR and the American Society of Engineering Education for their financial support I can only describe as gracious.

And last, but giga-light-years from being least, my hands go up in praise and my spirit in worship to the one and only, Lord God Almighty. Lord Jesus, I thank You for things of which only You and I can possibly know the depths and heights.

May God's grace and peace be with you all, and fill you with the hope of the promise that is in Christ Jesus.

Biographical Note

Christ D. Richmond was born in Washington, D.C. on December 16, 1967. He received both his S.B. in Electrical Engineering from the University of Maryland College Park (UMCP) and his S.B. in Mathematics from Bowie State University (BSU) in 1990 graduating with honors. He received his S.M., E.E., and Ph.D. degrees in Electrical Engineering from the Massachusetts Institute of Technology (MIT) in February 1993, February 1995, and May 1996 respectively.

He is the recipient of the Second Annual White House Initiative Award for Outstanding Achievement in Science and Technology 1988 (Carnegie Corporation, NY. Honored by President Ronald Reagan), and the John Hope Franklin Award for Academic Excellence 1990 (UMCP). He is the recipient of several annual Math and Science departmental awards given by BSU. He is the recipient of the Alan Berman Research Publications Award (ARPAD) March 1994 (NRL) and is a member of the Tau Beta Pi Engineering Honor Society. He is also an Office of Naval Research Graduate Fellow 1990–1994.

During the summers he conducted research in the Radar Division and in the Acoustics Division of the Naval Research Laboratory (NRL) in Washington, D.C. and the MITRE Corporation in McLean, VA. He served as a Research Assistant with Group 108 of MIT Lincoln Laboratory under Prof Allan Steinhardt's supervision. He was nominated for the Baker Foundation Award for Excellence in Undergraduate Teaching 1995 at MIT for his efforts as a Teaching Assistant of the Circuits, Signals and Systems class (an award intended for professors, not teaching assistants). He has authored four journal publications and five conference proceeding. His research interests include statistical signal and array processing, detection, estimation, radar, sonar, multivariate statistical analysis, adaptive systems, and controls.

He is actively involved in the music ministries of New Covenant Christian Center where he faithfully serves as an organist, pianist, keyboard player. He is a composer of contemporary Gospel music and has toured with the nationally acclaimed Bowie State Gospel choir.

C. D. Richmond has recently accepted a technical research staff position at MIT Lincoln Laboratories. He will begin working this summer in Division 10 in Group 108 under Dr. Steve Krich. After a few years of work experience he will pursue an academic position involving both teaching and research.

Contents

List of Figures

Chapter 1

Introduction

1.1 History and Motivation

Adaptive array signal processing has been based historically on the premise of data Gaussianity. This is primarily due to (1) the analytic tractability of the Gaussian, (2) the wealth of Gaussian based statistical sampling theory available, and (3) the central limit theorem which often makes for good approximations even in non-Gaussian environments. The widely known results on adaptive array (i) signal detection [16]–[26], (ii) signal estimation and beamforming [22], [28]–[31] (iii) power spectral estimation [32]–[34], and (iv) structured covariance estimation [35]–[36] are all Gaussian based. It has been known for some time, however, that the Gaussian distribution does not always adequately model the data obtained in several radar, sonar, and other array application areas. Non-Gaussian models, therefore, have been considered by a number of authors [1]–[7].

The class of non-Gaussian distributions explored in this thesis is known as *elliptically contoured* distributions (ECs) of which the Gaussian is simply a special case [53]–[60]. Historically speaking, special cases of EC distributions include *spherically invariant random vectors or processes (SIRVs or SIRPs)*, *Gaussian mixtures or compounds*, *Rayleigh mixtures*, *symmetric distributions*, and *sub-Gaussian alpha-stable distributions*. Specific examples of these types of distributions which have found use in practice include the *Weibull*, t-distribution, K-distribution, etc. These types of distributions have been shown to success-fully model radar sea clutter, for example [2, 3, 4]. Although no physical basis for applying such models to arrays existed originally, Sangston and Gerlach [5] recently gave a compelling theoretical argument for the use of such models in array processing; namely, based on a phenomenological scattering model, they derived a central limit theorem of which the Gaussian mixture model representation is the end product. Thus, EC distributions have had utility historically, but under the guise of different aliases, and show promise for future applications.

The common characteristic shared by *all* previously cited references on non-Gaussian models is that none considered the problem of adaptive array detection and estimation as originally posed by Reed, Mallett, and Brennan [16] and later by Kelly [19]. Specifically, they all consider non-adaptive situations and do not employ nor propose the use of a theoretic framework which allows for general estimation of the data covariance, typically an unknown parameter in adaptive array scenarios, via a *secondary (or training)* data set. This avoidance of adaptive non-Gaussian scenarios is primarily due to the potential analytic complexities that can arise, which are only exacerbated by the presence of an estimated data covariance. In addition there exist a plurality of possible non-Gaussian distributions.

Khatri and Rao [17, 18] appear to be the first and only ones to have considered adaptive array signal detection (but not estimation) in non-Gaussian environments via data models known as *multivariate elliptically contoured* distributions (MECs). Specifically, their work on detection included estimation of the data covariance. MECs provide an alternative to classical Gaussian based sampling theory allowing for the consideration of data samples which are nonnormal and dependent. It appears that no other literature has been published on adaptive array processing which further pursues the potential possibilities and insights offered by these data models. MECs in fact represent a very attractive set of data models for adaptive arrays for several reasons: (1) similar models have proven successful historically and contemporarily, under the guise of SIRVs/SIRPs and Gaussian mixtures/compounds, (2) MECs along with classical decision theory provide a theoretic framework which (i) allows one to consider a plurality of possible non-Gaussian distributions simultaneously, and (ii) provides for potential optimal estimation of the typically unknown data covariance, and (3) MECs often allow for tractable performance analysis even for processors based on the data covariance estimate known as the sample covariance matrix (SCM). The first reason provides confidence for the physical validity of such models. The significant implications of the second reason are that (i) the resulting analysis is not exclusive to a single non-Gaussian distribution, and (ii) there is room for potential nuisance parameters. The third reason is important for adaptive arrays because most if not all adaptive array processors are functions of the SCM. [1] Although a formidable task, performance analyses of SCM based processors are indispensable, insightful, and practically useful. This analytic tractability will allow (1) with regard to array detection structures, exact computation of probabilities of false alarm (PFA) and detection (PD), and (2) with regard to adaptive beamforming, exact assessment of processor optimality. As we shall show, the SCM, whose presence is commonly attributed to a Gaussian data assumption, is in fact optimal and pervasive within the entire class of MECs considered. *Indeed, analysis suggests that many of the array signal processing structures we typically attribute to a Gaussian data assumption, arise not because of the Gaussian per se, but rather are by products of the elliptical symmetry which the Gaussian happens to possess.*

Raghavan/Pulsone [8] and Tsihrintzis/Nikias [9] have recently considered adaptive array detection with independent identically distributed (i.i.d.) EC distributed samples. They each employ data models differing from the MEC model explored in this thesis with respect to snapshot dependence.

1.2 Goal of Thesis and Results

The goals of this thesis are twofold. The first objective is to relax the traditional constraint of data Gaussianity and reformulate the problems of adaptive array signal detection and estimation under a weaker distributional assumption. Specifically, we will not assume that

[1] Most of the literature on arrays appears to communicate that the prefix "adaptive" is synonymous with dependence on the SCM, or more formally the need to estimate the unknown data covariance. This definition will therefore be adopted in this thesis as well.

the data is Gaussian, but rather that its distribution is a member of the MEC class. Optimal detection and estimation structures/processors will be sought via classical approaches and heuristic ones. Concerning adaptive array detection, of primary interest are the *generalized likelihood ratio test* (GLRT) approach to adaptive detection, and the *adaptive matched filter* (AMF) approach. There exists numerous approaches to adaptive array signal detection, due its natural formulation as a composite binary hypothesis testing problem. From a practical point of view, however, the GLRT and AMF approaches have proven to work well in real world systems, and consequently have grown in popularity. In addition, Bose and Steinhardt [25, 26] have recently proven that when one casts the problem of adaptive array detection into an invariant hypothesis testing framework, both the GLRT and AMF are maximal invariants under the general linear group, and the GLRT is asymptotically uniformly most powerful invariant (UMPI). Hence, we restrict attention to these two approaches, although not to discourage future pursuit of better array detection structures. One major advantage to giving consideration to adaptive array signal detection first is that results on signal estimation, adaptive beamforming, and power spectral estimation serendipitously follow.

Understanding exactly how these processors perform in non-Gaussian environments is likewise vital to the adaptive array community, because these processors may ultimately be engaged in real physical systems, some of which cannot afford the risk associated with processors whose properties and weaknesses are not well understood. In addition Gaussian based processors have demonstrated robust performance in environments observed to be non-Gaussian. Thus, the second goal of this thesis is to provide an exact statistical performance analysis of the adaptive detection, estimation, beamforming, etc. structures when employed in these MEC environments. For the detection structures, we derive exact general expressions for the PFA and PD over this MEC class. Concerning adaptive signal estimation and beamforming, exact densities, confidence regions and relative statistics are derived for the SCM based: (1) Maximum-Likelihood (ML) signal vector estimator, (2) Linearly Constrained Minimum Variance beamformer (LCMV), (3) Minimum Variance Distortionless Response beamformer (MVDR), and (4) Generalized Sidelobe Cancellor (GSC) implementation of the LCMV beamformer. In addition some attention is devoted to the effects of jammers, and limited attention is given to the effects of steering vector mismatch. As further contributions, the Capon/Goodman result [33] is generalized over the MEC class, and we initiate the consideration/discussion of structured covariance estimation in the MEC class.

One of the major contributions of this thesis is the development of a large body of complex MEC based statistical measures which are of value in many adaptive array settings. In the process of analyzing processor performances, several new contributions to the theory of EC distributions are likewise made. These contributions are labeled as Theorems in Chapter 3, whereas results that are well known are labeled as Propositions. Although some contributions are rather straightforward from existing theoretical methods, they seemingly do not appear to be published in the open literature and are given for the first time in this thesis. Among them is a complex MEC generalization of the complex Wishart partition Theorem, which provides exact joint and marginal distributions for the Schur complements and regression coefficients of the MEC based SCM. Likewise we consider for seemingly the first time the density of singular complex EC distributions. Miller [62] appears to be the first to introduce the concept of a singular real normal distribution. Following his lead and that of Khatri [70], we develop similar results for complex ECs in general.

1.3 Organization of Thesis

This thesis is organized as follows. In Chapter 2 a brief review of several classical Gaussian based results on adaptive array signal processing is provided. Notational conventions are likewise established, and precise statements are given for what is historically meant by the problems of adaptive array detection and estimation, and adaptive beamforming. An excellent contemporary synopsis of these results and ideas, can be found in Haykin and Steinhardt [39]. Readers may also find the conference proceedings [10, 11, 15] helpful as well. Since most readers will be unfamiliar with the concept of an MEC distribution, a synoptic, yet fairly extensive list of properties and theorems is provided and developed in

Chapter 3. This chapter will in fact provide the theoretical foundation for most of the analysis encountered in this thesis. It is recommended that it be read one time through (or even browsed depending on the reader), and subsequently referenced throughout when needed. In Chapter 4 we consider both (i) the reformulation of the problem of adaptive array detection under a complex MEC data assumption, and (ii) the exact performance analyses of the resulting detection structures. Specifically, the GLRT and AMF approaches are considered, and general expressions for the PFA and PD are found. In Chapter 5 a detailed performance analysis of classical beamformers based on the SCM is given when the data is MEC distributed. In Chapter 6 several numerical simulations and calculations are made, illustrating the potential utility of the results of this thesis. In Chapter 7 further results on power spectral estimation and structured covariance estimation appear. Lastly, in Chapter 8 we attempt to summarize the key contributions made in this thesis and to outline directions for future work. There are many symbols used in this thesis. Appendix A, therefore, contains a list of the symbols, notation, abbreviations, and acronyms frequently used.

Chapter 2

Classical Gaussian Based Adaptive Array Signal Processing

In this chapter we review some of the more well known results on adaptive arrays, all of which are based on the premise of complex Gaussian data. In Chapter 4 it is shown that several of these processors are solely by products of the elliptical symmetry of the complex Gaussian, and in fact are obtained under a weaker assumption of complex MEC data.

2.1 Adaptive Array Signal Detection/Estimation

Given an array of spatially configured sensors, it is typically of interest to detect the presence of signals impinging upon the array in a background of noise. Arrays offer the added benefit of spatial discrimination of signals as well as frequency discrimination, via spatial filtering techniques. This ability to differentiate spatially diverse signals allows one, for example, to restrict attention to prespecified spatial look directions, or angular sectors. The process of spatial filtering is often referred to as beamforming, and typically makes up the front-end processor of array systems. Although beamforming constitutes most front-end processing, the internal processor is faced with the task of determining signal presence, which fundamentally speaking is one of classical binary hypothesis testing. In practice, the internal system processes data which sometimes comprises the outputs of array sensors (assume the effects of anti-aliasing filters, sampling, channel equalization, etc. to be implicit), or, for example, the outputs of multiple beams (and/or Doppler bins) steered to prespecified directions (or Doppler) [44]. In either case, the array data that the internal processor receives can be modeled as the following vector observation

$$\mathbf{x}_{(N\times 1)} = \begin{cases} \mathbf{n}_{(N\times 1)} & H_0 \\ \mathbf{G}_{(N\times E)}\,\mathbf{s}_{(E\times 1)} + \mathbf{n}_{(N\times 1)} & H_1 \end{cases} \tag{2.1}$$

where the dimensions have been indicated in subscript. (The notational convention will be boldfaced capitals to indicate complex matrices and boldfaced lowercase to indicate complex vectors, unless stated otherwise. Italic indicates complex scalar quantities in general, although several will be real and some integers. All will be apparent from the context in which they are used.). The vector **x** is the measured array data, with a covariance denoted by $\kappa \cdot \mathbf{R}$, where κ is a real positive scalar. It is often referred to as the *primary data vector* since it is the array observation or *snapshot* under interrogation. The vector of signal parameters we denote by **s** which, for example, in a radar and sonar context would contain target signal range, Doppler (velocity), and azimuth information; **n** is additive noise which models the uncertainty in the data observation; **G** is the matrix of array response vectors, also known as system transfer functions, and steering vectors. We choose to not be more specific concerning **G**, as the present analysis is primarily theoretical and applicable to many diverse systems which involve the processing of data obtained from an array of sensors. It is noted, however, that in a radar and sonar context **G** summarizes both the physical constraints of electromagnetic or acoustic wave propagation, and any front-end preprocessing and/or beamspace conversions the data undergoes prior to internal processing. In addition it is assumed that **G** is full column rank and known exactly (unknown **G** is considered in Section 5.6).

2.1.1 The LRT: R and s Known

In array signal detection the objective is simply to declare **x** as either "signal free," which we denote by hypothesis H_0, or "signal bearing," which we denote by hypothesis H_1. Traditionally this problem has been cast in a decision theoretic framework as a binary hypothesis testing problem, as we've already hinted. This turns out to be an attractive and rather prolific choice of framework for arrays since it (i) naturally allows one to deal with uncertainty in data measurements, and (ii) provides for potential consideration of a host of unknown parameters. When both the data covariance parameter $\mathbf{R}$ and the signal parameter vector $\mathbf{s}$ are known, then the optimal decision rule under a Neyman/Pearson criterion is the likelihood ratio test (LRT):

$$\frac{g_{H_1}(\cdot)}{g_{H_0}(\cdot)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{2.2}$$

where $g_{H_i}(\cdot)$ represent the likelihood functions of $\mathbf{x}$ given H_i is true, $i = 0, 1$. When $\mathbf{x}$ is assumed complex Gaussian [45, 61, 63], *i.e.*

$$\begin{array}{ll} g_{H_i} = \pi^{-N}|\mathbf{R}|^{-1}\exp\left[(\mathbf{x}-\mathbf{m}_i)^H\mathbf{R}^{-1}(\mathbf{x}-\mathbf{m}_i)\right] & i = 0, 1 \\ \mathbf{m}_0 = \mathbf{0} \qquad\qquad \mathbf{m}_1 = \mathbf{G}\mathbf{s}, & \end{array} \tag{2.3}$$

where $|\cdot|$ denotes the matrix determinant, a superscript T denotes matrix transposition, and a superscript H denotes conjugate transposition, the resulting test statistic is known to be the matched filter output $\mathrm{Re}(\mathbf{m}_1^H\mathbf{R}^{-1}\mathbf{x})$ where $\mathrm{Re}(\cdot)$ denotes the real part. This filter sets the upper bound on performance of all complex Gaussian based detectors in the Neyman/Pearson sense.

In adaptive array signal processing it is typically assumed that both $\mathbf{R}$ and $\mathbf{s}$ are unknowns, thus forming the LRT statistic is not possible. Indeed, these unknowns must be practically dealt with. In the next section we consider first the possibility of unknown $\mathbf{s}$. Subsequently, we consider the case in which both $\mathbf{R}$ and $\mathbf{s}$ are unknown. This latter situation sets the stage for the need of what is commonly referred to as classical *adaptive array signal proccssing.* As will become clear, $\mathbf{s}$ is assumed unknown but deterministic, and the need for adaptivity comes from the necessity to estimate $\mathbf{R}$, which is typically accomplished via multiple signal free data observations referred to collectively as a *secondary data set.*

2.1.2 The Clairvoyant MF Detector: R Known but s Unknown

An unknown $\mathbf{s}$ can be anything. Still assuming complex Gaussian data, it follows that the array signal detection problem is in fact a *composite* binary hypothesis testing problem;

namely, we are testing the composite hypotheses

$$H_0 : \mathbf{s} = \mathbf{0} \qquad \text{against} \qquad H_1 : \mathbf{s} \neq \mathbf{0}, \tag{2.4}$$

i.e. testing H_0 against an infinite number of possible H_1. In such a case a UMP test is desired, *i.e.* one would seek a test decision statistic t_{UMP} such that for any fixed PFA, would yield a PD greater than or equal to that of any other test with the same or smaller PFA. Unfortunately, no such test exists for this problem, and therefore other test procedures must be sought. One possible alternative test procedure is the GLRT approach.

The GLRT approach is heuristic and not necessarily optimal in any finite sample sense. Asymptotically, however, it does possess desirable properties [25, 52]. It attempts to mimic the optimal LRT. Indeed, the procedure is to form the LRT first, and subsequently replace all unknown parameters of the distributions by their ML estimates. Specifically, for the problem at hand the GLRT approach suggests the following test:

$$\frac{\max_{\mathbf{s}} g_{H_1}(\cdot)}{g_{H_0}(\cdot)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \tag{2.5}$$

Under the complex Gaussian assumption this approach is known to lead to what is commonly referred to as the following *clairvoyant* (known $\mathbf{R}$) matched filter (MF):

$$t_{MF} = \mathbf{x}^H\mathbf{R}^{-1}\mathbf{G}(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}\mathbf{G}^H\mathbf{R}^{-1}\mathbf{x} = \left\|(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1/2}\mathbf{G}^H\mathbf{R}^{-1}\mathbf{x}\right\|^2. \tag{2.6}$$

The ML estimate for $\mathbf{s}$ obtained in the process is well known to be given by

$$\mathbf{s}_{ML} = \mathbf{W}_{ML}^H\mathbf{x} = (\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}\mathbf{G}^H\mathbf{R}^{-1}\mathbf{x}. \tag{2.7}$$

The clairvoyant MF is commonly accepted as the detector yielding the upper bound on performance of any Gaussian based adaptive array detector. An adaptive detector can only fall below the performance bounds set by the MF, since the adaptive detector will suffer some loss for not knowing $\mathbf{R}$ exactly.

2.1.3 The GLRT and AMF: Both R and s Unknown

This section begins with a detailed statement of the adaptive array detection problem as originally posed by Reed *et al* [16] and later by Kelly [19], which primarily represents the

the class of problems to be addressed in this thesis. The GLRT approach taken by Kelly and the resulting test statistic is summarized. Likewise, the AMF approach and test statistic proposed by Robey *et al* [23], is also summarized.

Adaptive Array Detection Problem Statement

In the problem of adaptive array signal detection, signal presence is sought in a single $N \times 1$ array observation (or array *snapshot*) $\mathbf{x}$ often called the *primary* data vector. This primary data vector has an unknown covariance which we denote by $\kappa \cdot \mathbf{R}$ where κ is a known constant. Essentially, it is desired to classify this snapshot as one of two things:

$$\begin{array}{lcl} H_0 & : & \mathbf{x} = \mathbf{n} \\ H_1 & : & \mathbf{x} = \mathbf{G}\mathbf{s} + \mathbf{n}; \end{array} \tag{2.8}$$

either the primary data vector consists of noise only (denoted by the null hypothesis H_0), or it consists of signal plus noise (denoted by hypothesis H_1). When there is an actual target signal present in the primary data vector, the signal model consists of a known $N \times E$ matrix of system transfer function models $\mathbf{G}$, multiplied by an $E \times 1$ vector of unknown signal parameters $\mathbf{s}$. The matrix $\mathbf{G}$ is assumed full column rank and may represent, *e.g.* steering vectors for multiple pulses of radar echo returns, in which case the vector $\mathbf{s}$ will contain target range, Doppler, and azimuth information. It is worth noting that determination of parameter $\mathbf{s}$ is typically the goal many adaptive beamforming applications [39, 40]. Since it is an assumed unknown in this detection problem, the resulting detection structures must ultimately estimate it in some form. Consequently, the resulting detectors typically imply an optimal beamforming structure as well.

This detection problem has two unknowns $\mathbf{R}$ and $\mathbf{s}$. To compensate for the assumed ignorance of these nuisance parameters, we make the additional assumption that a *secondary* data set (or training set) $\mathbf{X} = [\mathbf{x}_1|\cdots|\mathbf{x}_L]$ is available. It is assumed that each snapshot $\mathbf{x}_i$ is noise only, zero mean, and shares the same covariance as the primary data vector, *i.e.* $\text{cov}(\mathbf{x}_i) = \kappa \cdot \mathbf{R}$ for $i = 1, \ldots, L$. The decision on signal presence, therefore, will not be based on observing $\mathbf{x}$ alone, but rather will be based on the totality of the data summarized by the $N \times (L+1)$ data matrix $\mathbf{X}_0 = [\mathbf{X}|\mathbf{x}]$. Under both hypotheses, it is assumed that $L \geq N$. We define the symbol

$$\mathbf{D} \triangleq \mathbf{X}\mathbf{X}^H, \tag{2.9}$$

which is the noise only SCM (unnormalized), as this quantity will predominate in many subsequent adaptive processors.

Gaussian Based GLRT and AMF

Kelly's approach to adaptive array detection was to use the GLRT approach. Specifically, he assumed that the data matrix $\mathbf{X}_0$ is complex Gaussian distributed with densities given by

$$g_{H_i} = |\mathbf{R}|^{-(L+1)}\pi^{-N(L+1)}\exp[\,-\mathrm{tr}\mathbf{R}^{-1}(\mathbf{X}_0 - \mathbf{M}_i)(\mathbf{X}_0 - \mathbf{M}_i)^H\,], \quad i = 0, 1. \tag{2.10}$$

Under the corresponding hypotheses, differences exist only in the means such that

$$\begin{array}{lcl} H_0 & : & \mathbf{M}_0 = \mathbf{0}_{N\times(L+1)} \\ H_1 & : & \mathbf{M}_1 = [\mathbf{0}_{N\times L}|\,\mathbf{G}\mathbf{s}\,]. \end{array} \tag{2.11}$$

In the GLRT method one first forms the optimal detector under the Neyman-Pearson criterion, *i.e.* the likelihood ratio test [46]; namely,

$$\frac{g_{H_1}(\cdot)}{g_{H_0}(\cdot)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \longrightarrow \frac{\max\limits_{\mathbf{s},\mathbf{R}} g_{H_1}(\cdot)}{\max\limits_{\mathbf{R}} g_{H_0}(\cdot)}. \tag{2.12}$$

Subsequently, each density is maximized over all its unknown parameters, *i.e.* $\mathbf{R}$ and $\mathbf{s}$ are replaced by their maximum likelihood (ML) estimates. Kelly showed that when both $\mathbf{R}$ and $\mathbf{s}$ are unknown the ML estimate of $\mathbf{R}$ is given by the SCM and the ML estimate of $\mathbf{s}$ is given by

$$\hat{\mathbf{s}}_{ML} = (\mathbf{G}^H\mathbf{D}^{-1}\mathbf{G})^{-1}\mathbf{G}^H\mathbf{D}^{-1}\mathbf{x}, \tag{2.13}$$

i.e. the SCM based LCMV beamformer. The GLRT technique results in the following decision statistic

$$t_{GLRT}(\mathbf{X}_0) = \left[\frac{\max\limits_{\mathbf{s},\mathbf{R}} g_{H_1}(\cdot)}{\max\limits_{\mathbf{R}} g_{H_0}(\cdot)}\right]^{\frac{1}{(L+1)}} = \frac{1+\mathbf{x}^H\mathbf{D}^{-1}\mathbf{x}}{1+\mathbf{x}^H\mathfrak{P}(\mathbf{G}|\mathbf{D}^{-1})\mathbf{x}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{2.14}$$

where the weighted projection operator $\mathfrak{P}$ is defined:

$$\mathfrak{P}(\mathbf{V}|\mathbf{H}) = \mathbf{H} - \mathbf{H}\mathbf{V}(\mathbf{V}^H\mathbf{H}\mathbf{V})^{-1}\mathbf{V}^H\mathbf{H}. \tag{2.15}$$

An alternate approach to this detection problem which originates from a more or less heuristic viewpoint was also suggested and called the AMF approach. Robey *et al* [23] likewise adopts the GLRT school of thought, but only to a point. Specifically, the GLRT method is applied first assuming $\mathbf{R}$ is known and only $\mathbf{s}$ is unknown:

$$\frac{\max\limits_{\mathbf{s}} g_{H_1}(\cdot)}{g_{H_0}(\cdot)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta. \tag{2.16}$$

Under the complex Gaussian assumption this approach is known to lead to the clairvoyant MF encountered earlier in Section 2.1.2. Robey, then says to estimate $\mathbf{R}$ from the secondary data via the SCM and subsequently replace $\mathbf{R}$ with the SCM in the expression for the MF. This leads to the SCM based MF

$$t_{AMF}(\mathbf{X}_0) = \mathbf{x}^H\mathbf{D}^{-1}\mathbf{G}^H(\mathbf{G}^H\mathbf{D}^{-1}\mathbf{G})^{-1}\mathbf{G}^H\mathbf{D}^{-1}\mathbf{x} \tag{2.17}$$

which is commonly called the adaptive matched filter (AMF) in the literature.

Consider an alternate interpretation of the AMF statistic. Note again from Section 2.1.2 that t_{MF} is just the magnitude squared of a normalized version of the clairvoyant ML signal estimate, *i.e.* $t_{MF} = ||(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{1/2}\mathbf{s}_{ML}||^2$. Hence, it is reasonable to suggest that if an optimal estimate of the signal parameter $\mathbf{s}$ can be found, then its magnitude along with a subsequent normalization can be used for detection. Specifically, for a complex Gaussian distributed $\mathbf{x}$ and known $\mathbf{R}$, the ML estimate of $\mathbf{s}$ is known to be

$$\mathbf{s}_{ML} = (\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}\mathbf{G}^H\mathbf{R}^{-1}\mathbf{x}. \tag{2.18}$$

This ML estimate of s is unbiased with a covariance given by the Cramer-Rao bound $(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}$. The transformation $(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{1/2}\mathbf{s}_{ML}$ whitens this estimate and removes its dependence on the true value of $\mathbf{R}$, under the null hypothesis. This leads to the constant false alarm rate (CFAR) property of the AMF detector. Lastly, as Robey suggests, we use the noise only SCM $\mathbf{D}$ in place of the unknown $\mathbf{R}$. We shall consider this interpretation of the AMF approach to adaptive detection, seeking detection structures for several non-Gaussian distributions. The key ingredient determining detection structure then, is the ML estimate of s based on $\mathbf{x}$ when $\mathbf{R}$ is known.

The properties and performance of both detectors t_{GLRT} and t_{AMF} under the complex Gaussian assumption are well known and documented [19]–[26]. For example, the AMF is known to be a CFAR detector as is the GLRT. Likewise, the AMF performance is comparable to the GLRT. In Chapter 4 we consider similar performance analyses under a weaker complex MEC data assumption.

First, however, we make a remark about the weighted projection operator $\mathfrak{P}(\mathbf{V}|\mathbf{H})$ defined earlier, as it will play a prominent role in the present analysis. Note that when $\mathbf{H} = \mathbf{I}$ we obtain the unweighted and true projection matrix onto the space orthogonal to the columns of $\mathbf{V}$. The weighted projector can in fact be written in terms of a true projector

$$\mathfrak{P}(\mathbf{V}|\mathbf{H}) = \mathbf{H}^{1/2}\, \mathfrak{P}(\mathbf{H}^{1/2}\mathbf{V}|\mathbf{I})\, \mathbf{H}^{1/2}. \tag{2.19}$$

One can interpret the operation of this projector, therefore, as one which first transforms all the data via the linear transformation $\mathbf{H}^{1/2}$, and subsequently truly projects, but within the transformed space.

Remarks

Signal detection can often be one step in a series of operations toward a specific goal. Once signal presence is determined, it is often of interest to estimate the signal parameters s. For example, space-time adaptive processing (STAP) methods [44], several of which are directly or indirectly related to the processors discussed in the next section, are employed to estimate radar target ranges, Doppler, and azimuth. Likewise, such techniques are employed in radar to suppress clutter and/or hostile directional interferences. In the next section we review a class of classical minimum variance beamforming and signal estimation structures widely present in the literature and of interest in this thesis.

2.2 Adaptive Array Signal Estimation/Beamforming

The specific array processors which we wish to consider are presented in this section. The brevity of this section's presentation is due to the vast amount of literature available on these classical processors (see [39]–[45] and their bibliographies).

2.2.1 Complex Gaussian Assumption

Recall that in array processing the multisensor array data is often modeled by the following vector observation

$$\mathbf{x} = \mathbf{G}\mathbf{s} + \mathbf{n} = \sum_{i=1}^{E} \mathbf{g}_i S_i + \mathbf{n}. \tag{2.20}$$

The vector $\mathbf{x}$ is again the measured or received array data containing the desired signal parameter vector $\mathbf{s} = [S_1, S_2, \ldots, S_E]^T$. The additive noise $\mathbf{n}$ models the uncertainty present in the data measurement. The columns of the matrix $\mathbf{G} = [\mathbf{g}_1 | \mathbf{g}_2 | \cdots | \mathbf{g}_E]$ model the system transfer functions, also known as the *steering vectors* or *look directions.* It is assumed that $\mathbf{G}$ has full column rank and is known exactly.

Given such data the objective is often to determine the signal $\mathbf{s}$ from the array data $\mathbf{x}$, or to spatially filter (beamform) the data, passing only select look directions $\mathbf{g}_i$, $i = 1, \ldots, E$, while minimizing output power (variance). Under the assumption of complex Gaussian data the optimal processor in the Bayesian mean square error sense always consists of performing a linear transformation on the data; namely, the signal estimate/beamformer output will respectively be of the form

$$\tilde{\mathbf{s}} = \mathbf{W}^H \mathbf{x} \qquad \text{and} \qquad \tilde{y} = \mathbf{w}^H \mathbf{x}. \tag{2.21}$$

The *clairvoyant* (when $\mathbf{R}$ is known exactly) array processors performing these linear transformations which we wish to consider in this thesis are (1) the Maximum-Likelihood (ML) signal vector estimator, (2) the Linearly Constrained Minimum Variance beamformer (LCMV), (3) the Minimum Variance Distortionless Response beamformer (MVDR), and (4) the Generalized Sidelobe Cancellor (GSC) implementation of the LCMV beamformer. The specific clairvoyant weightings for these array processors are summarized in eq(2.22):

$$\begin{array}{rl} & \underline{\textit{Clairvoyant Weightings}} \\ \text{ML:} & \mathbf{W}_{ML} = \mathbf{R}^{-1}\mathbf{G}(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1} \\ \text{LCMV/MVDR:} & \mathbf{w}_{LCMV} = \mathbf{W}_{ML}\mathbf{f} \\ \text{GSC:} & \mathbf{w}_{gsc} = \mathbf{W}_{GSC}\mathbf{f} \\ & \mathbf{W}_{GSC} = [\mathbf{I}_N - \mathbf{G}_\perp \mathbf{\Omega}_{GSC}]\, \mathbf{G}(\mathbf{G}^H\mathbf{G})^{-1} \\ & \mathbf{\Omega}_{GSC} = (\mathbf{G}_\perp^H\, \mathbf{R}\, \mathbf{G}_\perp)^{-1}\mathbf{G}_\perp^H\, \mathbf{R}. \end{array} \tag{2.22}$$

A brief discussion of each clairvoyant processor is given in the following subsections.

2.2.2 Maximum-Likelihood Signal Estimator/LCMV Beamformer

When the objective is to estimate the received signal $\mathbf{s}$, then assuming that the noise $\mathbf{n}$ is complex Gaussian distributed with zero mean and covariance $\mathbf{R}$, which we indicate by the notation $\mathbf{n} \sim \mathcal{CN}_N(\mathbf{0}, \mathbf{R})$, it is known that the ML signal estimate is given by the following linear transformation on the data

$$\mathbf{s}_{ML} = \mathbf{W}_{ML}^H \mathbf{x}, \tag{2.23}$$

where the weight matrix performing the linear transformation is given by

$$\mathbf{W}_{ML} = \mathbf{R}^{-1}\mathbf{G}(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}. \tag{2.24}$$

The mathematical form of this signal estimator appears under several aliases representative of differing methods of derivation. For example, it is also the mathematical form of (1) the *weighted or generalized* least-squares estimator, (2) the *Gauss-Markov* minimum variance estimator, and (3) the Best Linear Unbiased Estimator (BLUE) [45]. Under *appropriate* assumptions they all coincide, but in the most general context they can differ. A derivation and discussion of the optimality of the ML signal estimator is found in [45] pp. 524–530. In particular this Gaussian based ML estimator is both unbiased and efficient.

In beamforming applications the goal is typically to spatially/temporally filter the signal bearing data as a front-end processor [40, 44]. To this end the processor's *output* and *scalar beam response* (which is look direction, and frequency dependent) are defined to be

$$y = \mathbf{w}^H \mathbf{x} \quad \text{and} \quad b(\theta,\omega) = \mathbf{w}^H \mathbf{d}_\theta(\omega) \tag{2.25}$$

respectively. The vector $\mathbf{d}_\theta(\omega)$ is a vector of array element transfer functions called the *array response vector* or *steering vector*. It embeds in a rather compact form the constraints imposed on signal waveforms by the physics of wave propagation as they traverse the array. The magnitude square $|b(\theta,\omega)|^2$ as a function of look direction θ from array broadside and/or temporal frequency ω is referred to as the array *beampattern*.

The Linearly Constrained Minimum Variance beamformer (LCMV) chooses the data weighting vector $\mathbf{w}$ according to the following constrained optimization problem

$$\min_{\mathbf{w}} \mathbf{w}^H \mathbf{R} \mathbf{w} \quad \text{subject to} \quad \mathbf{G}^H \mathbf{w} = \mathbf{f}. \tag{2.26}$$

The LCMV beamformer constrains the beam response to have fixed gains $\mathbf{f}$ along specific look directions/frequencies conveyed by the columns of $\mathbf{G}$, [1] while minimizing the output power $E\{|y|^2\} - |E\{y\}|^2 = \mathbf{w}^H \mathbf{R} \mathbf{w}$. This problem is typically solved via Lagrange multiplier methods. The solution, however, follows from elementary linear algebra. First transform (whiten) the data such that

$$\mathbf{R}^{-1/2}\mathbf{x} = \mathbf{R}^{-1/2}\mathbf{G}\mathbf{s} + \mathbf{R}^{-1/2}\mathbf{n} = \mathbf{G}_0\mathbf{s} + \mathbf{n}_0 = \mathbf{x}_0. \tag{2.27}$$

Now the equivalent problem becomes

$$\min_{\mathbf{w}_0} ||\mathbf{w}_0||^2 \quad \text{subject to} \quad \mathbf{G}_0^H \mathbf{w}_0 = \mathbf{f}, \tag{2.28}$$

i.e. find the minimum norm solution to a set of *underdetermined* linear equations, which have an infinite number of solutions. This is a classical problem whose solution can be found

[1] In general the constraints can be of the form $\mathbf{C}^H\mathbf{w} = \mathbf{f}$ where, *e.g.* $\mathbf{C} = [\mathbf{G}|\mathbf{C}_0]$. The matrix $\mathbf{C}_0$ can include directions for desired nulls, derivative constraints, etc. [40]. The extension of our analysis to this case is fairly straightforward and rather trivial. Therefore, we restrict ourselves to the form implied by eq(2.26), where the constraints are chosen to guarantee an unbiased signal estimate. This choice likewise facilitates comparisons between different structures.

in almost any text on linear algebra [48, 49]; namely, $\mathbf{w}_0 = \mathbf{G}_0(\mathbf{G}_0^H\mathbf{G}_0)^{-1}\mathbf{f}$. Equating the outputs

$$\begin{aligned} y = \mathbf{w}_0^H\mathbf{x}_0 &= \mathbf{f}^H(\mathbf{G}_0^H\mathbf{G}_0)^{-1}\mathbf{G}_0^H\mathbf{x}_0 \\ &= \mathbf{f}^H(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}\mathbf{G}^H\mathbf{R}^{-1}\mathbf{x} \\ &= \mathbf{w}^H\mathbf{x}, \end{aligned} \tag{2.29}$$

the solution to this classical problem is shown to be given by the weighting

$$\mathbf{w}_{LCMV} = \mathbf{W}_{ML}\mathbf{f} = \mathbf{R}^{-1}\mathbf{G}(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}\mathbf{f}. \tag{2.30}$$

Hence, the beamformer output

$$y_{LCMV} = \mathbf{w}_{LCMV}^H\mathbf{x} = \mathbf{f}^H\mathbf{s}_{ML}. \tag{2.31}$$

Note that the LCMV beamformer output is a linear combination of the ML signal estimates $[S_{ML,1}, S_{ML,2}, \ldots, S_{ML,E}]^T = \mathbf{s}_{ML}$ where each true signal S_i originates from a different direction $\mathbf{g}_i$. The i-th column of $\mathbf{W}_{ML}$ is a beamformer for direction $\mathbf{g}_i$. Correspondingly, we define the following *vector beam response*

$$\mathbf{b}_{LCMV}(\theta,\omega) = \mathbf{W}_{ML}^H\mathbf{d}_\theta(\omega). \tag{2.32}$$

The LCMV scalar beam response is given by

$$b_{LCMV}(\theta,\omega) = \mathbf{w}_{LCMV}^H\mathbf{d}_\theta(\omega) = \mathbf{f}^H\mathbf{b}_{LCMV}(\theta,\omega), \tag{2.33}$$

a linear combination of the rows of $\mathbf{b}_{LCMV}(\theta,\omega)$.

In the special case of $E = 1$ and f = 1 note that the LCMV beamformer reduces to the Minimum Variance Distortionless Response (MVDR) beamformer with weighting given by

$$\mathbf{w}_{MVDR} = \mathbf{R}^{-1}\mathbf{g}/(\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}). \tag{2.34}$$

2.2.3 Generalized Sidelobe Cancellor

The generalized sidelobe cancellor (GSC) represents an alternate yet equivalent implementation of the LCMV beamformer. In the GSC, the LCMV beamformer weights $\mathbf{w}_{LCMV}$ are decomposed into a constrained component and an unconstrained component. The fixed constrained portion satisfies the constraints of the optimization problem given in eq(2.26), leaving the unconstrained portion free for adaption. The GSC beamformer vector weighting, say $\mathbf{w}_{gsc}$, is given by

$$\mathbf{w}_{gsc} = \mathbf{\Phi f} - \mathbf{G}_\perp \mathbf{w}_\perp \tag{2.35}$$

where $\mathbf{\Phi} = \mathbf{G}(\mathbf{G}^H\mathbf{G})^{-1}$ and $\mathbf{G}^H\mathbf{G}_\perp = \mathbf{0}_{E\times(N-E)}$. The *quiescent* (when $\mathbf{R} \propto \mathbf{I}_N$) beamformer $\mathbf{\Phi f}$ has been chosen to satisfy the constraint on $\mathbf{w}_{gsc}$. Note that

$$\mathbf{G}^H\mathbf{w}_{gsc} = \mathbf{G}^H\mathbf{\Phi f} - \mathbf{G}^H\mathbf{G}_\perp\mathbf{w}_\perp = \mathbf{f} - \mathbf{0}. \tag{2.36}$$

The orthogonality of the manifold $\mathbf{G}_\perp$ to the array manifold $\mathbf{G}$ suggests that the constraints are satisfied independent of how $\mathbf{w}_\perp$ is chosen. The LCMV equivalent optimization problem for the GSC implementation is hence given by

$$\min_{\mathbf{w}_\perp} (\mathbf{\Phi f} - \mathbf{G}_\perp\mathbf{w}_\perp)^H \mathbf{R} (\mathbf{\Phi f} - \mathbf{G}_\perp\mathbf{w}_\perp). \tag{2.37}$$

The solution is shown to be given by

$$\mathbf{w}_\perp = \mathbf{\Omega}_{GSC}\mathbf{\Phi f} \quad \text{where} \quad \mathbf{\Omega}_{GSC} = (\mathbf{G}_\perp^H \mathbf{R}\, \mathbf{G}_\perp)^{-1}\mathbf{G}_\perp^H \mathbf{R}, \tag{2.38}$$

which follows by simply completing the square

$$\begin{gathered}(\mathbf{\Phi f} - \mathbf{G}_\perp\mathbf{w}_\perp)^H \mathbf{R} (\mathbf{\Phi f} - \mathbf{G}_\perp\mathbf{w}_\perp) = \\ (\mathbf{w}_\perp - \mathbf{\Omega}_{GSC}\mathbf{\Phi f})^H (\mathbf{G}_\perp^H\mathbf{R}\mathbf{G}_\perp)(\mathbf{w}_\perp - \mathbf{\Omega}_{GSC}\mathbf{\Phi f}) + \mathbf{f}^H\mathbf{\Phi}^H\mathfrak{P}(\mathbf{G}_\perp|\mathbf{R})\mathbf{\Phi f}.\end{gathered} \tag{2.39}$$

Minor algebra shows that the GSC weight vector can be written as

$$\mathbf{w}_{gsc} = \mathbf{W}_{GSC}\mathbf{f} = [\mathbf{I}_N - \mathbf{G}_\perp\mathbf{\Omega}_{GSC}]\mathbf{\Phi f}. \tag{2.40}$$

This form will be useful in later analysis.

Analogous to eq(2.31) we define respectively the following GSC signal estimate, vector beam response, and scalar beam response

$$\mathbf{s}_{GSC} = \mathbf{W}_{GSC}^H\mathbf{x} \qquad \mathbf{b}_{GSC}(\theta,\omega) = \mathbf{W}_{GSC}^H\mathbf{d}_\theta(\omega) \qquad b_{GSC}(\theta,\omega) = \mathbf{w}_{gsc}^H\mathbf{d}_\theta(\omega). \tag{2.41}$$

The GSC beamformer output is given by $y_{GSC} = \mathbf{w}_{gsc}^H\mathbf{x}$. The overall GSC structure is illustrated in fig(2.1). An alternative structure is given in fig(2.2) which allows comparison with other structures. Note in fig(2.2) that the input to the $\mathbf{f}$ block is the GSC signal estimate $\mathbf{s}_{GSC}$. In principle linear combinations of the columns of $\mathbf{G}_\perp$ are subtracted from the sidelobes of the quiescient beam response in an attempt to cancel noise leaking through its sidelobes. Hence, the name generalized sidelobe cancellor. The normality of $\mathbf{G}_\perp$ to the

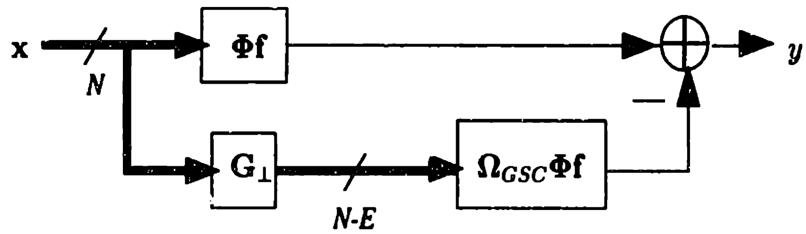


Figure 2.1: Generalized Sidelobe Cancellor

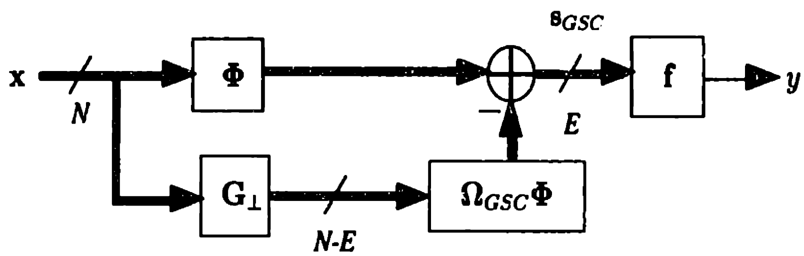


Figure 2.2: GSC Alternate Block Representation

array manifold $\mathbf{G}$ guarantees that the processor is careful not to subtract desired signal from the mainlobe. For this reason $\mathbf{G}_\perp$ is often referred to as the *signal blocking matrix*. Although conceptually intuitive and appealing, the concepts of a mainlobe and sidelobes are actually a bit blurred in light of the problem formulation eq(2.26). It is more accurate to say that linear combinations of the columns of $\mathbf{G}_\perp$ are subtracted from the unconstrained portions of the quiescent beam response in an attempt to cancel noise, and the normality of $\mathbf{G}_\perp$ to $\mathbf{G}$ guarantees that the constraints are not affected by this noise cancellation process. We will at times, however, choose the former more appealing dialog in terms of sidelobes and mainlobes.

Although different in mathematical form, the weight vector $\mathbf{w}_{gsc}$ is an algebraic equivalent implementation of the LCMV beamformer weight vector $\mathbf{w}_{LCMV}$ given in eq(2.30) [37]. This equivalence can and will be exploited at times to infer results about one processor based on those of the other.

2.2.4 Unknown Data Covariance

The first generation optimal array processors presented in the previous section were designed/derived under the rather idealized assumption of known data covariance $\kappa \cdot \mathbf{R}$. In practice, however, such an assumption rarely holds. Consequently, one is lead to (1) attempt derivation of second generation array processors which drop the ideal assumption of known data covariance, or (2) yield to the common heuristic method of using the SCM estimate of this nuisance parameter, namely $\mathbf{D}$, in place of its true unknown value. An example of the first method is ML estimation performed over both the unknown data covariance parameter $\mathbf{R}$ and unknown signal parameters $\mathbf{s}$. The second method is known as *sample matrix inversion (SMI)* which, because of its fast convergence rate, grew in popularity over slowly converging gradient based methods such as least mean squares (LMS) and recursive least squares (RLS) [16]. In the case of zero mean Gaussian data it is known that the two approaches lead to the same processor [39, 22]. In particular both these approaches lead to the following SCM based (SCB) processors:

<u>*SCB Weightings*</u>

$$\begin{aligned}
\text{ML:}\quad & \widehat{\mathbf{W}}_{ML} = \mathbf{D}^{-1}\mathbf{G}(\mathbf{G}^H\mathbf{D}^{-1}\mathbf{G})^{-1} \\
\text{LCMV/MVDR:}\quad & \hat{\mathbf{w}}_{LCMV} = \widehat{\mathbf{W}}_{ML}\mathbf{f} \\
\text{GSC:}\quad & \hat{\mathbf{w}}_{gsc} = \widehat{\mathbf{W}}_{GSC}\mathbf{f} \\
& \widehat{\mathbf{W}}_{GSC} = \left[\mathbf{I}_N - \mathbf{G}_\perp\widehat{\Omega}_{GSC}\right]\mathbf{G}(\mathbf{G}^H\mathbf{G})^{-1} \\
& \widehat{\Omega}_{GSC} = (\mathbf{G}_\perp^H\,\mathbf{D}\,\mathbf{G}_\perp)^{-1}\mathbf{G}_\perp^H\,\mathbf{D}.
\end{aligned} \tag{2.42}$$

The hat " $\widehat{\ }$ " accent will be used to denote the SCM as well as the dependence of the weightings and other quantities on the SCM. What is not transparent is whether both approaches lead to the SCB processors under a weaker assumption of MEC data. This question will be considered in Chapter 4.

Regardless of the approach taken, use of estimated data covariances pervades much of adaptive array processing. Either directly, as in both the SMI and ML methods, or indirectly as in the LMS, RLS, and Kalman filter based approaches. Estimated covariances inevitably result in a loss in overall system performance. Use of the SCM traditionally has been shown to yield analytically tractable performance analyses under the Gaussian data assumption. These performance analyses have proven to be of great value to the practicing engineer/scientist in several respects. These analyses quantify in a statistical sense the loss experienced due to the processors reliance upon the covariance estimate $\mathbf{D}$ as opposed to the true covariance. For example, the adaptivity loss typically appears as a multiplicative factor which is solely a function of the sample size of the SCM and the degrees of freedom available for adaptation [16]–[34]. The advantage in knowing these losses a priori allows engineers to adjust parameters of the system in order to make compensations such that a desired level of performance is ultimately achieved. A goal in this thesis is to extend such performance analyses for the SCB processors when the data is MEC distributed. In particular we seek exact densities, confidence regions, means and covariances of each SCB weighting listed in eq(2.42), and their corresponding beam responses, signal estimates and beamformer outputs.

Remarks

This concludes review of longstanding Gaussian based adaptive processing methods. Additional Gaussian based results on power spectral estimation and structured covariance estimation can be found in Chapter 7. We now turn attention to analysis under the class of non-Gaussian models known as complex MECs.

Chapter 3

The Theory of Complex Multivariate Elliptically Contoured Distributions

3.1 Introduction

In this chapter we provide a synoptic introduction to the theory of complex elliptically contoured (EC) distributions. No attempt has been made here, however, to provide a full exposition of this elegant theory, for there are several excellent texts on the subject [54]–[57]. In addition a systematic theoretic treatment of the concept of complex EC theory is found in [60]. Our goal, rather, will be two fold: (i) to define the concept of a complex multivariate elliptically contoured (MEC) distributed matrix, and (ii) to list (and for most prove) properties of these types of distributions. These properties will be used extensively in subsequent analysis. Most properties have interpretations with respect to both the density (sample space domain) and the characteristic function (c.f.) (frequency domain). We will work primarily with densities since their existence is presumed in this thesis and is a common assumption in array signal processing. Familiarity with the properties presented and developed in this chapter will be essential in later analysis. Indeed, we will draw heavily upon these theorems, some of which seemingly appear for the first time in this thesis.

Since the normal distribution is a member of the EC class, and familiar to most, we start with a brief summary of normal theory. It will be evident in later sections that many of the properties making analysis under a Gaussian assumption attractive, are shared by all members of the EC class.

The notational convention adopted in this chapter differs from the rest of the thesis in that we find it necessary and useful to distinguish real quantities from complex ones, and statistics related to each. Therefore, a tilde "~" is used to indicate that a vector/matrix quantity is complex, and to indicate that a statistic was obtained from a complex random quantity. For example, the symbol $\mathbf{x}$ indicates a real vector, and $\tilde{\mathbf{x}}$ indicates a complex one. Likewise, the statistics $\rho = ||\mathbf{x}||^2$ and $\tilde{\rho} = ||\tilde{\mathbf{x}}||^2$, although both real quantities, follow the same convention. Prior and subsequent to this chapter, all vector and matrix quantities are assumed complex unless stated otherwise. Thus, superseding the need for a tilde to differentiate.

3.2 Multivariate Normal Theory

Recall from basic probability that a random variable y is univariate Gaussian if it has pdf and c.f. given by

$$P_y = \frac{1}{\sigma\sqrt{2\pi}} e^{-(y-m)^2/2\sigma^2} \overset{\mathcal{F}}{\longleftrightarrow} e^{jtm} \cdot e^{-t^2\sigma^2/2} = \psi_y(t) \tag{3.1}$$

where $E\{y\} = m$ and $\text{var}(y) = \sigma^2$.

Proposition 3.2.0.1 Cramer-Wold: *The distribution for a random vector (r.vc.) $\mathbf{x} \in \Re^N$ is completely determined by that of $y = \mathbf{a}^T\mathbf{x}$ if the pdf for y is known for arbitrary $\mathbf{a} \in \Re^N$.*

The proof follows by noting that from the pdf of y we can obtain the c.f.

$$\psi_y(t_y) = E\{e^{jt_y y}\} = E\{e^{jt_y \mathbf{a}^T\mathbf{x}}\}\Big|_{t_y=1} = \psi_{\mathbf{x}}(\mathbf{t}_x = \mathbf{a}). \tag{3.2}$$

That we can uniquely determine the c.f. for $\mathbf{x}$ completes the proof.

Definition 3.2.1 *A random vector $\mathbf{x} \in \Re^N$ is said to be normal if and only if for every vector $\mathbf{a} \in \Re^N$ the random variable $y = \mathbf{a}^T\mathbf{x}$ is univariate normal with mean $\mathbf{a}^T\mathbf{m}$ and variance $\mathbf{a}^T\mathbf{R}\mathbf{a}$ where we define $\mathbf{m} \triangleq E\{\mathbf{x}\}$ and $\mathbf{R} \triangleq \text{cov}(\mathbf{x})$.*

Note from the Cramer-Wold Proposition 3.2.0.1 and the normal definition with $\mathbf{a} = \mathbf{t}_x$ the following proposition is obtained:

Proposition 3.2.0.2 *A random vector is normally distributed if its c.f. is given by*

$$E\{e^{jt_y \mathbf{t}_x^T \mathbf{x}}\}\Big|_{t_y=1} = E\{e^{j\mathbf{t}_x^T \mathbf{x}}\} = e^{j\mathbf{t}_x^T \mathbf{m}} \cdot \exp\left(-\mathbf{t}_x^T \mathbf{R} \mathbf{t}_x/2\right) \triangleq \psi_{\mathbf{x}}(\mathbf{t}_x). \tag{3.3}$$

We indicate that a real random vector has this distribution by the notation $\mathbf{x} \sim \mathcal{N}_N(\mathbf{m}, \mathbf{R})$.

Proposition 3.2.0.3 *If* $\mathbf{z} \sim \mathcal{N}_N(\mathbf{0}, \mathbf{I}_N)$ *then it is said to be white and has the spherically symmetric density*

$$P_{\mathbf{z}} = (2\pi)^{-N/2} \exp\left(-||\mathbf{z}||^2/2\right). \tag{3.4}$$

This follows by noting that the c.f. is given by

$$\psi_{\mathbf{z}}(\mathbf{t}_z) = \exp\left(-\mathbf{t}_z^T \mathbf{t}_z/2\right) = \prod_{i=1}^{N} e^{-t_i^2/2} \stackrel{\mathcal{F}}{\longleftrightarrow} \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}} e^{-z_i^2/2} = (2\pi)^{-N/2} \exp\left(-||\mathbf{z}||^2/2\right). \tag{3.5}$$

Proposition 3.2.0.4 *The Gaussian regenerates under linear transformations. Specifically, if* $\mathbf{y} = \mathbf{A}^T \mathbf{x}$ *where* $\mathbf{A}$ *is* $M \times N$ *(*$M \leq N$*) and* $\mathbf{x} \sim \mathcal{N}_N(\mathbf{m}, \mathbf{R})$, *then* $\mathbf{y} \sim \mathcal{N}_N(\mathbf{A}^T \mathbf{m}, \mathbf{A}^T \mathbf{R} \mathbf{A})$.

This follows by noting that

$$E\{e^{j\mathbf{t}_y^T \mathbf{y}}\} = E\{e^{j\mathbf{t}_x^T \mathbf{x}}\}\Big|_{\mathbf{t}_x = \mathbf{A}^T \mathbf{t}_y} = e^{j\mathbf{t}_y^T (\mathbf{A}^T \mathbf{m})} \cdot \exp\left[-\mathbf{t}_y^T (\mathbf{A}^T \mathbf{R} \mathbf{A}) \mathbf{t}_y/2\right], \tag{3.6}$$

i.e. $\mathbf{y}$ is Gaussian distributed with a mean of $\mathbf{A}^T \mathbf{m}$ and a covariance $\mathbf{A}^T \mathbf{R} \mathbf{A}$.

Proposition 3.2.0.5 *If* $\mathbf{z} \sim \mathcal{N}_N(\mathbf{0}, \mathbf{I}_N \sigma^2)$, *and* $\mathbf{y} = \mathbf{Q}\mathbf{z}$ *where* $\mathbf{Q}^T \mathbf{Q} = \mathbf{Q}\mathbf{Q}^T = \mathbf{I}_N$, *then* $\mathbf{y}$ *and* $\mathbf{z}$ *are identically distributed.*

This spherical invariance property easily follows from Proposition 3.2.0.4.

Proposition 3.2.0.6 *If* $\mathbf{z} \sim \mathcal{N}_N(\mathbf{0}, \mathbf{I})$, *then* $\mathbf{x} \sim \mathcal{N}_N(\mathbf{m}, \mathbf{R})$ *can be written identically distributed to*

$$\mathbf{x} \stackrel{d}{=} \mathbf{m} + \mathbf{R}^{1/2} \mathbf{z}. \tag{3.7}$$

Above we have introduced the symbol "$\stackrel{d}{=}$", which shall be used to indicate that two quantities share the same pdf. This is commonly referred to as an equivalent *stochastic representation* for $\mathbf{x}$. The stochastic representation is an extremely powerful and insightful concept which will be revisited in Section 3.4.2. To prove the equality in distribution note that

$$E\{e^{j\mathbf{t}_x^T(\mathbf{m}+\mathbf{R}^{1/2}\mathbf{z})}\} = e^{j\mathbf{t}_x^T\mathbf{m}} \cdot \exp\left(-\mathbf{t}_x^T\mathbf{R}\mathbf{t}_x/2\right) = \psi_{\mathbf{x}}(\mathbf{t}_x) \tag{3.8}$$

where the first equality follows from the Gaussian linear regeneration property of Proposition 3.2.0.4.

Proposition 3.2.0.7 *In general if* $\mathbf{x} \sim \mathcal{N}_N(\mathbf{m}, \mathbf{R})$ *and* $\mathbf{R}$ *is nonsingular, then it has the density*

$$P_{\mathbf{x}} = (2\pi)^{-N/2}|\mathbf{R}|^{-1/2}\exp\left[-(\mathbf{x}-\mathbf{m})^T\mathbf{R}^{-1}(\mathbf{x}-\mathbf{m})/2\right]. \tag{3.9}$$

This result easily follows by noting that (i) the pdf for $\mathbf{z} \sim \mathcal{N}_N(\mathbf{0}, \mathbf{I}_N)$ is given in Proposition 3.2.0.3, (ii) $\mathbf{x}$ can be written as a linear combination of $\mathbf{z}$ by Proposition 3.2.0.6, recalling that (iii) the Jacobian for the linear transformation in Proposition 3.2.0.6 is $J(\mathbf{z} \to \mathbf{x}) = |\mathbf{R}|^{-1/2}$.

Let $\mathbf{x} \sim \mathcal{N}_N(\mathbf{m}, \mathbf{R})$ and make the partition

$$\mathbf{x} = \begin{bmatrix} (\mathbf{x}_1)_{(M\times 1)} \\ (\mathbf{x}_2)_{(N-M)\times 1} \end{bmatrix} \qquad \mathbf{m} = \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \end{bmatrix} \qquad \mathbf{R} = \begin{bmatrix} \mathbf{R}_{11} & \mathbf{R}_{12} \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{bmatrix}. \tag{3.10}$$

Proposition 3.2.0.8 *The marginal pdfs are given by* $\mathbf{x}_1 \sim \mathcal{N}_M(\mathbf{m}_1, \mathbf{R}_{11})$ *and* $\mathbf{x}_2 \sim \mathcal{N}_{(N-M)}(\mathbf{m}_2, \mathbf{R}_{22})$.

This follows from Proposition 3.2.0.4 by choosing for example $\mathbf{A}^T = [\mathbf{I}_M, \mathbf{0}_{M\times(N-M)}]$.

Proposition 3.2.0.9 *The conditional distribution of* $\mathbf{x}_1$ *given* $\mathbf{x}_2$ *is* $\mathcal{N}_M(\mathbf{m}_{1.2}, \mathbf{R}_{11.2})$ *where*

$$\begin{aligned} \mathbf{m}_{1.2} &= \mathbf{m}_1 + \mathbf{R}_{12}\mathbf{R}_{22}^{-1}(\mathbf{x}_2 - \mathbf{m}_2) \\ \mathbf{R}_{11.2} &= \mathbf{R}_{11} - \mathbf{R}_{12}\mathbf{R}_{22}^{-1}\mathbf{R}_{21}. \end{aligned} \tag{3.11}$$

This proposition is proved by considering the linear transformation

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{I}_M & -\mathbf{R}_{12}\mathbf{R}_{22}^{-1} \\ \mathbf{0}_{(N-M)\times M} & \mathbf{I}_{(N-M)} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \mathbf{A}\mathbf{x}. \tag{3.12}$$

By Proposition 3.2.0.4 it follows that

$$\mathbf{y} \sim \mathcal{N}_N\left(\begin{bmatrix} \mathbf{m}_1 - \mathbf{R}_{12}\mathbf{R}_{22}^{-1}\mathbf{m}_2 \\ \mathbf{m}_2 \end{bmatrix}, \begin{bmatrix} \mathbf{R}_{11.2} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{22} \end{bmatrix}\right). \tag{3.13}$$

Noting that (i) the Jacobian $J(\mathbf{y} \to \mathbf{x}) = |\mathbf{A}| = 1$, (ii) $|\mathbf{R}| = |\mathbf{R}_{11.2}||\mathbf{R}_{22}|$, and (iii)

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{I}_M & \mathbf{R}_{12}\mathbf{R}_{22}^{-1} \\ \mathbf{0}_{(N-M)\times M} & \mathbf{I}_{(N-M)} \end{bmatrix}, \tag{3.14}$$

it follows that the density for $\mathbf{x}$ can be written in the form

$$\begin{aligned} P_{\mathbf{x}_1,\mathbf{x}_2} = (2\pi)^{-M/2}|\mathbf{R}_{11.2}|^{-1/2}\exp\left[-(\mathbf{x}_1 - \mathbf{m}_{1.2})^T\mathbf{R}_{11.2}^{-1}(\mathbf{x}_1 - \mathbf{m}_{1.2})\right] \times \\ (2\pi)^{-(N-M)/2}|\mathbf{R}_{22}|^{-1/2}\exp\left[-(\mathbf{x}_2 - \mathbf{m}_2)^T\mathbf{R}_{22}^{-1}(\mathbf{x}_2 - \mathbf{m}_2)\right]. \end{aligned} \tag{3.15}$$

Thus, the proposition is proven by recalling Bayes' formula for conditional probability. The following proposition easily follows:

Proposition 3.2.0.10 *The vectors* $\mathbf{x}_1$ *and* $\mathbf{x}_2$ *are independent if and only if they are uncorrelated, i.e.* $\mathbf{R}_{12} = \mathbf{R}_{21}^T = \mathbf{0}$.

Although traditionally attributed exclusively to Gaussianity, most of these properties, as we shall see, are shared by all members of the EC class. Indeed, most of these properties are a consequence of the elliptical symmetry which the Gaussian possesses.

3.2.1 Sampling from a Real Gaussian

Consider a collection of data samples $\mathbf{x}_i$ from the distribution $\mathcal{N}_N(\mathbf{m}, \mathbf{R})$, and defined the totality of this data by the the *data matrix*

$$\mathbf{X} = [\mathbf{x}_1|\mathbf{x}_2|\cdots|\mathbf{x}_L]. \tag{3.16}$$

If these samples are mutually uncorrelated (and therefore independent) the pdf is given by

$$\begin{aligned} &\prod_{i=1}^{L}(2\pi)^{-N/2}|\mathbf{R}|^{-1/2}\exp\left[-(\mathbf{x}_i - \mathbf{m})^T\mathbf{R}^{-1}(\mathbf{x}_i - \mathbf{m})/2\right] \\ &= (2\pi)^{-NL/2}|\mathbf{R}|^{-L/2}\exp\left[-\frac{1}{2}\sum_{i=1}^{L}(\mathbf{x}_i - \mathbf{m})^T\mathbf{R}^{-1}(\mathbf{x}_i - \mathbf{m})\right]. \end{aligned} \tag{3.17}$$

Recalling that $\mathbf{a}^T\mathbf{B}\mathbf{a} = \mathrm{tr}(\mathbf{B}\mathbf{a}\mathbf{a}^T)$ and that $\mathrm{tr}(\mathbf{A}+\mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$ where $\mathrm{tr}(\cdot)$ denotes the matrix trace, the pdf for the matrix $\mathbf{X}$ can be written in the form

$$\mathbf{X} \sim (2\pi)^{-NL/2}|\mathbf{R}|^{-L/2}\exp\left[-\frac{1}{2}\mathrm{tr}\mathbf{R}^{-1}(\mathbf{X}-\mathbf{M})(\mathbf{X}-\mathbf{M})^T\right] \tag{3.18}$$

where $\mathbf{M} = [\mathbf{m}|\cdots|\mathbf{m}]$ is $N \times L$. This is known as a *multivariate* real Gaussian. [1] The form of the pdf in eq(3.18) is analytically more convenient. Its associated c.f. can be written

$$E\left\{\exp\left[j\mathrm{tr}(\mathbf{T}^T\mathbf{X})\right]\right\} = \exp\left[j\mathrm{tr}(\mathbf{T}^T\mathbf{M})\right]\ \exp\left[-\mathrm{tr}(\mathbf{T}^T\mathbf{R}\mathbf{T})/2\right] \tag{3.19}$$

where $\mathbf{T}$ is $N \times L$.

Define the operator vec$(\cdot)$ by

$$\mathrm{vec}(\mathbf{X}) \triangleq \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \vdots \\ \mathbf{x}_L \end{bmatrix} \triangleq [\mathbf{x}_1; \mathbf{x}_2; \cdots; \mathbf{x}_L]. \tag{3.20}$$

The covariance of a random matrix is then defined to be

$$\mathrm{cov}(\mathbf{X}) \triangleq \mathrm{cov}\left[\mathrm{vec}(\mathbf{X})\right]. \tag{3.21}$$

For the particular case eq(3.18) note that $\mathrm{cov}(\mathbf{X}) = \mathbf{I}_L \otimes \mathbf{R}$ where $\otimes$ represents the Kronecker matrix product:

$$\mathbf{A}_{C\times D} \otimes \mathbf{B}_{F\times G} \triangleq \begin{bmatrix} A_{11}\mathbf{B} & A_{12}\mathbf{B} & \cdots & A_{1D}\mathbf{B} \\ A_{21}\mathbf{B} & A_{22}\mathbf{B} & \cdots & A_{2D}\mathbf{B} \\ \vdots & & \ddots & \vdots \\ A_{C1}\mathbf{B} & A_{C2}\mathbf{B} & \cdots & A_{CD}\mathbf{B} \end{bmatrix}_{CF\times DG}. \tag{3.22}$$

This matrix product is known to have the properties

$$\mathbf{A}\otimes\mathbf{B} + \mathbf{A}\otimes\mathbf{C} = \mathbf{A}\otimes(\mathbf{B}+\mathbf{C}) \ \text{ and } \ (\mathbf{A}\otimes\mathbf{B})\times(\mathbf{C}\otimes\mathbf{D}) = (\mathbf{A}\mathbf{C})\otimes(\mathbf{B}\mathbf{D}) \tag{3.23}$$

We have already proven the following:

Proposition 3.2.1.1

$$\mathbf{X} \ \textit{has pdf eq(3.18)} \overset{\textit{iff}}{\Longleftrightarrow} \mathrm{vec}(\mathbf{X}) \sim \mathcal{N}_{NL}(\mathbf{1}_L \otimes \mathbf{m}, \mathbf{I}_L \otimes \mathbf{R}). \tag{3.24}$$

where the vector of ones $\mathbf{1}_L = [1, 1, \ldots, 1]^T$ *is* $L \times 1$.

---

[1]It is common practice in the statistical analysis literature to indicate that a random quantity is a data matrix by the prefix "multivariate."

Multivariate distributions for data matrices are useful for modeling a collection of population samples as one random object. Classical multivariate statistical analysis is based on multivariate normal theory. Within the last twenty years or so, extensive progress has been made extending classical analysis to the non-Gaussian distribution class known as EC. Before we introduce the EC distribution, several integral theorems are listed which we shall find repeatedly useful.

3.3 Useful Integral Theorems

The following integral theorems are well known and listed here without proof. The real versions can be found in [54]. Most of the complex versions follow directly from the real, but with twice as many degrees of freedom. Proposition 3.3.2.2 is proved in the appendix of [22].

3.3.1 Real Variables

Proposition 3.3.1.1 *Let* $d\mathbf{a} \triangleq da_1 da_2 \cdots da_N$. *Integrating a well-behaved function* $q(\mathbf{a}^T\mathbf{a})$ *over all* $\mathbf{a} \in \Re^N$ *has the equivalent integral representation*

$$\int_{\mathbf{a}\in\Re^N} q(\mathbf{a}^T\mathbf{a})d\mathbf{a} = \frac{\pi^{N/2}}{\Gamma(N/2)} \int_0^\infty r^{\frac{N}{2}-1} q(r)dr \tag{3.25}$$

where $\Gamma(x)$ *is the well known univariate Gamma function [72], which has the property that* $\Gamma(n+1) = n!$ *when* n *is a non-negative integer.*

Proposition 3.3.1.2 *Let* $d\mathbf{A} \triangleq d\mathbf{a}_1 d\mathbf{a}_2 \cdots d\mathbf{a}_L$ *where* $\mathbf{A}^T = [\mathbf{a}_1|\cdots|\mathbf{a}_L]$, *each* $\mathbf{a}_i$ *is* $N \times 1$, *and* $L \geq N$. *Integrating the function* $q(\mathbf{A}^T\mathbf{A})$ *over all* $\mathbf{A} \in \Re^{NL}$ *has the equivalent integral representation*

$$\int_{\mathbf{A}\in\Re^{NL}} q(\mathbf{A}^T\mathbf{A})d\mathbf{A} = \frac{\pi^{NL/2}}{\Gamma_N(L/2)} \int_{\mathbf{D}=\mathbf{D}^T>0} |\mathbf{D}|^{\frac{1}{2}(L-N-1)} q(\mathbf{D})d\mathbf{D}. \tag{3.26}$$

where the multivariate Gamma function is given by $\Gamma_N(x) = \pi^{N(N-1)/4} \prod_{i=1}^N \Gamma[x-(i-1)/2]$, *and the latter integral is over the space of all symmetric positive definite matrices* $\mathbf{D}$.

Proposition 3.3.1.3 *Let $B_N = 2\pi^{N/2}/\Gamma(N/2)$, i.e. the surface area of a unit sphere in $\Re^N$. Then*

$$\int_{\{\mathbf{a}\in\Re^N \mid \mathbf{a}^T\mathbf{a}=c\}} q(\mathbf{a}^T\mathbf{a})d\mathbf{a} = B_{N-1}\int_{-c}^{c} q(||\mathbf{a}||r)(c^2-r^2)^{\frac{1}{2}(N-3)}dr. \tag{3.27}$$

Proposition 3.3.1.4 *For $q(\cdot)$ a well-behaved function we have the following equivalence*

$$\int_{\mathbf{a}\in\Re^N} q(\mathbf{b}^T\mathbf{a},\mathbf{a}^T\mathbf{a})d\mathbf{a} = \frac{\pi^{\frac{1}{2}(N-1)}}{\Gamma[(N-1)/2]}\int_{-\infty}^{\infty} du \int_0^{\infty} r^{\frac{1}{2}(N-1)-1} q(u||\mathbf{b}||, u^2+r)dr. \tag{3.28}$$

3.3.2 Complex Variables

Proposition 3.3.2.1 *Let $\tilde{\mathbf{a}} = \mathbf{a}_R + j\mathbf{a}_I$ and $d\tilde{\mathbf{a}} \triangleq da_{R1}da_{R2}\cdots da_{RN}da_{I1}da_{I2}\cdots da_{IN}$. Integrating a well-behaved function $q(\tilde{\mathbf{a}}^H\tilde{\mathbf{a}})$ over all $\tilde{\mathbf{a}} \in \mathfrak{C}^N$ has the equivalent integral representation*

$$\int_{\tilde{\mathbf{a}}\in\mathfrak{C}^N} q(\tilde{\mathbf{a}}^H\tilde{\mathbf{a}})d\tilde{\mathbf{a}} = \frac{\pi^N}{(N-1)!}\int_0^{\infty} r^{N-1}q(r)dr. \tag{3.29}$$

Proposition 3.3.2.2 *Let $d\tilde{\mathbf{A}} \triangleq d\tilde{\mathbf{a}}_1 d\tilde{\mathbf{a}}_2\cdots d\tilde{\mathbf{a}}_L$ where $\tilde{\mathbf{A}}^H = [\tilde{\mathbf{a}}_1|\cdots|\tilde{\mathbf{a}}_L]$, each $\tilde{\mathbf{a}}_i$ is $N\times 1$, and $L \geq N$. Integrating the function $q(\tilde{\mathbf{A}}^H\tilde{\mathbf{A}})$ over all $\tilde{\mathbf{A}} \in \mathfrak{C}^{NL}$ has the equivalent integral representation*

$$\int_{\tilde{\mathbf{A}}\in\mathfrak{C}^{NL}} q(\tilde{\mathbf{A}}^H\tilde{\mathbf{A}})d\tilde{\mathbf{A}} = \frac{\pi^{NL}}{\tilde{\Gamma}_N(L)}\int_{\tilde{\mathbf{D}}=\tilde{\mathbf{D}}^H>0} |\tilde{\mathbf{D}}|^{L-N}q(\tilde{\mathbf{D}})d\tilde{\mathbf{D}}, \tag{3.30}$$

where the latter integral is over the space of all hermitian positive definite (hpd) matrices $\tilde{\mathbf{D}}$, and the multivariate complex Gamma function is defined to be the following product of univariate Gamma functions:

$$\tilde{\Gamma}_N(L) \triangleq \pi^{N(N-1)/2}\prod_{i=1}^{N}\Gamma(L-i+1). \tag{3.31}$$

Proposition 3.3.2.3 *Let $\tilde{B}_N = 2\pi^N/(N-1)!$, i.e. the surface area of a unit sphere in $\mathbb{C}^N$. Then*

$$\int_{\{\tilde{\mathbf{a}}\in\mathbb{C}^N \ | \ \tilde{\mathbf{a}}^H\tilde{\mathbf{a}}=c\}} q(\tilde{\mathbf{a}}^H\tilde{\mathbf{a}})d\tilde{\mathbf{a}} = \tilde{B}_{N-1}\int_{-c}^{c} q(||\tilde{\mathbf{a}}||r)(c^2-r^2)^{N-\frac{3}{2}}dr. \tag{3.32}$$

Proposition 3.3.2.4 *For $q(\cdot)$ a well-behaved function we have the following equivalence*

$$\int_{\tilde{\mathbf{a}}\in\mathbb{C}^N} q(\tilde{\mathbf{b}}^H\tilde{\mathbf{a}}, \tilde{\mathbf{a}}^H\tilde{\mathbf{a}})d\tilde{\mathbf{a}} = \frac{\pi^{(N-1)}}{(N-2)!}\int_{u\in\mathbb{C}^1} d\tilde{u}\int_0^\infty r^{N-2}q(\tilde{u}||\tilde{\mathbf{b}}||, |\tilde{u}|^2+r)dr. \tag{3.33}$$

3.4 Real Elliptically Contoured Distributions

It is of interest to consider non-Gaussian densities which have the same elliptical symmetry as the normal, but whose modal and tail behavior may differ significantly relative to the normal.

Definition 3.4.1 *A random vector $\mathbf{x}\in\Re^N$ is said to be EC distributed if its c.f. is of the form*

$$\psi_{\mathbf{x}}(\mathbf{t}) = E\{e^{j\mathbf{t}^T\mathbf{x}}\} = e^{j\mathbf{t}^T\mathbf{m}}\cdot\phi(\mathbf{t}^T\mathbf{R}\mathbf{t}) \tag{3.34}$$

where $\phi(\cdot)$ is some valid functional form and $\mathbf{R}>0$ (positive definite). [2] *We indicate this distribution and its three parameters by the notation*

$$\mathbf{x}\sim\mathcal{EC}_N(\mathbf{m},\mathbf{R},\phi). \tag{3.35}$$

The functional form of $\phi(\cdot)$ distinguishes one type of EC distribution from another. For example, we've seen that if $\phi(u)=e^{-u/2}$, then $\mathbf{x}$ is real Gaussian.

[2]In this thesis the parameter $\mathbf{R}$ is always assumed to be positive definite, unless stated otherwise. In general $\mathbf{R}$ can be positive semidefinite, and hence singular. This leads to singular EC distributions. See Section 3.7 for a discussion of densities of singular ECs.

3.4.1 Spherically Symmetric Distributions

A spherically symmetric (SS) distribution is the special case of the EC distribution with parameters $\mathbf{m} = \mathbf{0}$ and $\mathbf{R} \propto \mathbf{I}$. From the definition of EC distribution it follows that the c.f. of a SS r.vc. is

$$\psi_{\mathbf{z}}(\mathbf{t}_z) = E\{e^{j\mathbf{t}_z^T\mathbf{z}}\} = \phi(||\mathbf{t}_z||^2). \tag{3.36}$$

Proposition 3.4.1.1 *If* $\mathbf{z}$ *is SS then its distribution is invariant to any orthogonal transformation, i.e. for* $\mathbf{Q}$ *orthogonal* $\mathbf{y} = \mathbf{Q}\mathbf{z}$ *and* $\mathbf{z}$ *are identically distributed.*

This follows by noting that

$$E\{e^{j\mathbf{t}_y^T\mathbf{y}}\} = E\{e^{j\mathbf{t}_z^T\mathbf{z}}\}\Big|_{\mathbf{t}_z=\mathbf{Q}^T\mathbf{t}_y} = \phi(\mathbf{t}_y^T\mathbf{Q}\mathbf{Q}^T\mathbf{t}_y) = \phi(\mathbf{t}_y^T\mathbf{t}_y). \tag{3.37}$$

This proposition is also a necessary and sufficient definition of a SS random vector. The density of a SS random vector $\mathbf{z}$ is necessarily a function of its magnitude $||\mathbf{z}||$. This follows from (i) the Fourier integral definition of the density in terms of the c.f. $\phi(||\mathbf{t}||^2)$ and (ii) Proposition 3.3.1.4. Indeed, the magnitude is known to be the maximal invariant on the orthogonal linear group. Hence, the following proposition:

Proposition 3.4.1.2 *If a random vector* $\mathbf{z} \in \Re^N$ *is SS then its density is necessarily of the form*

$$2^{-N/2}g\left(||\mathbf{z}||^2/2\right). \tag{3.38}$$

The multiple 2 factors in eq(3.38) are included to make notation of the complex SS and EC distributions, our primary focus in this thesis, convenient. [3]

Proposition 3.4.1.3 *If a random vector* $\mathbf{z} \in \Re^N$ *is SS with density* $P_{\mathbf{z}} = 2^{-N/2}g(||\mathbf{z}||^2/2)$, *then the density of the random variable* $\rho = ||\mathbf{z}||^2$ *is*

$$P_\rho = \frac{2^{-N/2}\pi^{N/2}}{\Gamma(N/2)}\rho^{\frac{N}{2}-1}g(\rho/2) \qquad 0 \le \rho < \infty. \tag{3.39}$$

[3] This is to be contrasted with the notation of [60]. If the factors are not included here, then they would continuously appear in subsequent analysis.

To prove consider the expectation

$$E\left\{q(||\mathbf{z}||^2)\right\} = \int_{\mathbf{z}\in\Re^N} q(\mathbf{z}^T\mathbf{z})2^{-N/2}g(\mathbf{z}^T\mathbf{z}/2)d\mathbf{z} = \int_0^\infty q(\rho)\frac{2^{-N/2}\pi^{N/2}}{\Gamma(N/2)}\rho^{\frac{N}{2}-1}g(\rho/2)d\rho \quad (3.40)$$
$$= \int_0^\infty q(\rho)\, P_\rho\, d\rho \triangleq E\{q(\rho)\}.$$

The second equality follows from Proposition 3.3.1.1. By comparison with the definition of expectation the result follows.

A very important example of a SS random vector is $\mathbf{u}^{(N)} \in \Re^N$ uniformly distributed on the unit spheroid in $\Re^N$. Clearly the density of $\mathbf{u}^{(N)}$ is *singular* (degenerate) since it is known with certainty that $||\mathbf{u}^{(N)}|| = 1$ independent of variation in the elements of $\mathbf{u}^{(N)}$. The c.f. nevertheless does exist and is given by the following integral

$$\psi_{\mathbf{u}^{(N)}}(\mathbf{t}) = \frac{1}{B_N}\int_{\{\mathbf{a}\in\Re^N|\mathbf{a}^T\mathbf{a}=1\}} e^{j\mathbf{t}^T\mathbf{a}}d\mathbf{a} = \frac{\Gamma(N/2)}{\sqrt{\pi}\Gamma[(N-1)/2]}\int_{-1}^{1} e^{j||\mathbf{t}||b}(1-b^2)^{\frac{1}{2}(N-3)}db \quad (3.41)$$

where the second equality follows from Proposition 3.3.1.3. Making the change of variables $b = \cos\theta$, and recognizing the resulting integral with respect to θ as a Bessel function it follows (see [71] p. 482 no. 5) that

$$\psi_{\mathbf{u}^{(N)}}(\mathbf{t}) = \Gamma(N/2)(2/||\mathbf{t}||^2)^{\frac{1}{2}(N-2)}J_{\frac{1}{2}(N-2)}(||\mathbf{t}||^2) \triangleq \Omega_N(||\mathbf{t}||^2). \quad (3.42)$$

The following significant result is due to Schoenberg [66]. It essentially set the stage for many subsequent results on EC distributions which appeared in the last twenty years:

Proposition 3.4.1.4 *The function $\psi_{\mathbf{z}}(\mathbf{t})$ is the c.f. of a SS random vector $\mathbf{z} \in \Re^N$ if and only if*

$$\psi_{\mathbf{z}}(\mathbf{t}) = \int_0^\infty \Omega_N(||\mathbf{t}||^2\rho)\, P_\rho\, d\rho = \phi(||\mathbf{t}||^2) \quad (3.43)$$

where $\Omega_N(\cdot)$ is defined in eq(3.42) and P_ρ is the pdf of some r.v. ρ defined on $0 \le \rho < \infty$.

To prove let $\mathbf{t} = ||\mathbf{t}|| \cdot [a_1, a_2, \ldots, a_N]^T$ where $||\mathbf{a}|| = 1$, and hence

$$\phi(||\mathbf{t}||^2) = \psi_{\mathbf{z}}(||\mathbf{t}||a_1, ||\mathbf{t}||a_2, \ldots, ||\mathbf{t}||a_N). \quad (3.44)$$

Since $\psi_{\mathbf{z}}$ is only a function of $||\mathbf{t}||$ by definition of SS, then for any fixed value of $\mathbf{t}$ the function $\psi_{\mathbf{z}}$ is unchanged by averaging over the surface of a spheroid of radius $||\mathbf{t}||$, *i.e.*

$$\begin{aligned} \phi(||\mathbf{t}||^2) &= \frac{1}{B_N}\int_{||\mathbf{a}||=1} \psi_{\mathbf{z}}(||\mathbf{t}||a_1, ||\mathbf{t}||a_2, \ldots, ||\mathbf{t}||a_N)d\mathbf{a} \\ &= \frac{1}{B_N}\int_{||\mathbf{a}||=1} \left[\int_{\mathbf{z}\in\Re^N} e^{i||\mathbf{t}||\mathbf{a}^T\mathbf{z}}\, P_{\mathbf{z}}\, d\mathbf{z}\right] d\mathbf{a} \end{aligned} \tag{3.45}$$

where $P_{\mathbf{z}}$ is the pdf for the SS random vector $\mathbf{z}$. Switching the order of integration we obtain

$$\begin{aligned} \phi(||\mathbf{t}||^2) &= \int_{\mathbf{z}\in\Re^N} \left[\frac{1}{B_N}\int_{||\mathbf{a}||=1} e^{i||\mathbf{t}||\mathbf{z}^T\mathbf{a}} d\mathbf{a}\right] P_{\mathbf{z}}\, d\mathbf{z} \\ &= \int_{\mathbf{z}\in\Re^N} \Omega_N(||\mathbf{t}||^2||\mathbf{z}||^2)\, P_{\mathbf{z}}\, d\mathbf{z} \\ &= E\{\, \Omega_N(||\mathbf{t}||^2||\mathbf{z}||^2)\,\}. \end{aligned} \tag{3.46}$$

Clearly, if we let $\rho \overset{d}{=} ||\mathbf{z}||$ with pdf P_ρ it follows that

$$\begin{aligned} E\{\, \Omega_N(||\mathbf{t}||^2||\mathbf{z}||^2)\,\} &= E\{\, \Omega_N(||\mathbf{t}||^2\rho)\,\} \\ &= \int_0^\infty \Omega_N(||\mathbf{t}||^2\rho)\, P_\rho\, d\rho \end{aligned} \tag{3.47}$$

which completes proof. If the SS $\mathbf{z}$ has the density eq(3.38) then by Proposition 3.4.1.3

$$P_\rho = \frac{2^{-N/2}\pi^{N/2}}{\Gamma(N/2)}\rho^{\frac{N}{2}-1}g(\rho/2), \tag{3.48}$$

and therefore the functional forms $\phi(\cdot)$ and $g(\cdot)$ are related by the integral

$$\boxed{\phi(a) = \frac{2^{-N/2}\pi^{N/2}}{\Gamma(N/2)}\int_0^\infty \Omega_N(a\,\rho)\rho^{\frac{N}{2}-1}g(\rho/2)d\rho.} \tag{3.49}$$

Recalling eq(3.42) this relation is seen to be the well-known *Hankel Transform*. Therefore, knowing $\phi(\cdot)$ uniquely determines $g(\cdot)$, and vice versa, *i.e.* their relationship is one-to-one. Shortly, it will be apparent that the actual functional form of $g(\cdot)$ in general distinguishes one EC distribution from another.

3.4.2 The Stochastic Representation and Its Uses

In our review of multivariate normal theory recall from Proposition 3.2.0.6 that any normal vector can be written as a linearly transformed white random vector. Such a generalization of this concept exists across the class of EC distributions providing a very elegant theory which unifies the normal distribution with its host of EC kindred.

Proposition 3.4.2.1 *If* $\mathbf{z} \in \Re^N$ *is SS with c.f.* $\phi(||\mathbf{t}||^2)$ *and density* $2^{-N/2}g(||\mathbf{z}||^2/2)$ *necessarily related as in eq(3.49), then it can be written equal in distribution to*

$$\mathbf{z} \stackrel{d}{=} \mathbf{u}^{(N)} \cdot \sqrt{\rho} \tag{3.50}$$

where $\rho \sim P_\rho$ *given by eq(3.48) with* ρ *and* $\mathbf{u}^{(N)}$ *independent of each other.*

This follows immediately from the expectation

$$\begin{aligned} E\{e^{j\mathbf{t}^T\mathbf{u}^{(N)}\cdot\sqrt{\rho}}\} &= \int_0^\infty P_\rho \, d\rho \, \frac{1}{B_N} \int_{||\mathbf{a}||=1} e^{j\mathbf{t}^T\mathbf{a}\cdot\sqrt{\rho}} d\mathbf{a} \\ &= \int_0^\infty P_\rho \, d\rho \, \Omega_N(||\mathbf{t}||^2 \rho) \\ &= \phi(||\mathbf{t}||^2) = \psi_{\mathbf{z}}(\mathbf{t}). \end{aligned} \tag{3.51}$$

The next proposition easily follows from this fact:

Proposition 3.4.2.2 *If* $\mathbf{z} \in \Re^N$ *is SS with c.f.* $\phi(||\mathbf{t}||^2)$ *and density* $2^{-N/2}g(||\mathbf{z}||^2/2)$ *necessarily related as in eq(3.49), then* $||\mathbf{z}||^2 \stackrel{d}{=} \rho \sim P_\rho$ *and* $\mathbf{z}/||\mathbf{z}|| \stackrel{d}{=} \mathbf{u}^{(N)}$ *and they are independent.*

Note that the r.v. ρ, the functional form of $g(\cdot)$, and that of $\phi(\cdot)$ are uniquely coupled and determine which distribution of the SS (and EC) class we're dealing with. The mental picture should look something like

$$\rho \leftrightarrow g(\cdot) \leftrightarrow \phi(\cdot), \tag{3.52}$$

i.e. the pdf of ρ is determined by $g(\cdot)$, which in turn uniquely determines $\phi(\cdot)$ via the Hankel transform. Note from Proposition 3.4.2.2 that any SS r.vc. $\mathbf{z}$, regardless of the functional form of its density or c.f., generates a uniformly distributed r.vc. by forming the statistic $\mathbf{z}/||\mathbf{z}|| \stackrel{d}{=} \mathbf{u}^{(N)}\sqrt{\rho}/||\mathbf{u}^{(N)}||\sqrt{\rho} \stackrel{d}{=} \mathbf{u}^{(N)}$. This follows because the resulting statistic washes out ρ, and therefore washes out dependence on $g(\cdot)$ and $\phi(\cdot)$.

Proposition 3.4.2.3 *Concerning* $\mathbf{u}^{(N)}$

$$E\{\mathbf{u}^{(N)}\} = \mathbf{0} \qquad \text{and} \qquad \text{cov}(\mathbf{u}^{(N)}) = \frac{1}{N}\mathbf{I}_N. \tag{3.53}$$

The simplest SS vector is $\mathbf{z} \sim \mathcal{N}_N(\mathbf{0}, \mathbf{I}_N)$. In this case it is known that ρ is a real chi-squared of N degrees of freedom. Hence, $E\{\mathbf{z}\} = \mathbf{0} = E\{\mathbf{u}^{(N)}\} \cdot E\{\sqrt{\rho}\}$ and $\text{cov}(\mathbf{z}) = \mathbf{I}_N = \text{cov}(\mathbf{u}^{(N)}) \cdot N$. That $E\{\sqrt{\rho}\} > 0$ completes proof.

3.4.3 The Density of a Real EC

Proposition 3.4.3.1 *If* $\mathbf{x} \sim \mathcal{EC}_N(\mathbf{m}, \mathbf{R}, \phi)$ *then it has the equivalent stochastic representation*

$$\mathbf{x} \stackrel{d}{=} \mathbf{m} + \mathbf{R}^{1/2}\mathbf{u}^{(N)}\sqrt{\rho} \tag{3.54}$$

where P_ρ *is related to* ϕ *via eq(3.43).*

That

$$\begin{aligned} E\{e^{j\mathbf{t}^T(\mathbf{m}+\mathbf{R}^{1/2}\mathbf{u}^{(N)}\sqrt{\rho})}\} &= e^{j\mathbf{t}^T\mathbf{m}} \cdot E\{\ E\{e^{j(\sqrt{\rho}\mathbf{R}^{1/2}\mathbf{t})^T\mathbf{u}^{(N)}} \big| \ \rho\}\ \} \\ &= e^{j\mathbf{t}^T\mathbf{m}} \cdot E\{\ \Omega_N(\mathbf{t}^T\mathbf{R}\mathbf{t}\rho)\ \} \\ &= e^{j\mathbf{t}^T\mathbf{m}} \cdot \int_0^\infty \Omega_N(\mathbf{t}^T\mathbf{R}\mathbf{t}\rho)\, P_\rho\, d\rho = e^{j\mathbf{t}^T\mathbf{m}} \cdot \phi(\mathbf{t}^T\mathbf{R}\mathbf{t}) \end{aligned} \tag{3.55}$$

where the last equality follows from eq(3.43), completes proof.

Proposition 3.4.3.2 *If* $\mathbf{x} \sim \mathcal{EC}_N(\mathbf{m}, \mathbf{R}, \phi)$ *then it has mean and covariance*

$$E\{\mathbf{x}\} = \mathbf{m} \qquad\qquad \text{cov}(\mathbf{x}) = \frac{E\{\rho\}}{N} \cdot \mathbf{R}. \tag{3.56}$$

Proposition 3.4.3.3 *If* $\mathbf{x} \sim \mathcal{EC}_N(\mathbf{m}, \mathbf{R}, \phi)$ *has a density, then it is necessarily of the form*

$$2^{-N/2}|\mathbf{R}|^{-1/2} g\left[(\mathbf{x}-\mathbf{m})^T\mathbf{R}^{-1}(\mathbf{x}-\mathbf{m})/2\right] \tag{3.57}$$

where $\phi(\cdot)$ *and* $g(\cdot)$ *are related by eq(3.49). In this case we may choose to denote this by* $\mathbf{x} \sim \mathcal{EC}_N(\mathbf{m}, \mathbf{R}, g(\cdot))$.

Recall Proposition 3.4.3.1 and let $\mathbf{z} \stackrel{d}{=} \mathbf{u}^{(N)}\sqrt{\rho}$. By Proposition 3.4.1.2 the density of $\mathbf{z}$ is of the form $2^{-N/2}g\left(||\mathbf{z}||^2/2\right)$. Finally, note that the Jacobian for the linear transformation is $J(\mathbf{z} \to \mathbf{x}) = |\mathbf{R}|^{-1/2}$. This completes proof.

3.4.4 Sampling from Real EC Distributions: The MEC

As in our discussion of the real Gaussian, consider a collection of data samples $\mathbf{x}_i$ from the distribution $\mathcal{EC}_N(\mathbf{m}, \mathbf{R}, g(\cdot))$, and define the totality of this data by the the data matrix

$$\mathbf{X} = [\mathbf{x}_1|\mathbf{x}_2|\cdots|\mathbf{x}_L]. \tag{3.58}$$

If these samples are independent then the pdf of $\mathbf{X}$ is given by

$$\prod_{i=1}^{L} 2^{-N/2}|\mathbf{R}|^{-1/2} g\left[(\mathbf{x}_i - \mathbf{m})^T \mathbf{R}^{-1}(\mathbf{x}_i - \mathbf{m})/2\right]. \tag{3.59}$$

An alternate model for the data matrix $\mathbf{X}$ is one in which its columns still share the same mean and covariance, but are only uncorrelated. The density of $\mathbf{X}$ is then given by

$$2^{-NL/2}|\mathbf{R}|^{-L/2} g\left[\frac{1}{2}\mathrm{tr}\mathbf{R}^{-1}(\mathbf{X} - \mathbf{M})(\mathbf{X} - \mathbf{M})^T\right] \tag{3.60}$$

where $\mathbf{M} = [\mathbf{m}|\cdots|\mathbf{m}]$ is $N \times L$. This is a special case of a real *multivariate elliptically contoured (MEC)* distribution often referred to as a *Louiville elliptically contoured (LEC)* distribution. The models in eq(3.59) and eq(3.60) coincide if and only if $\mathbf{X}$ is a normal data matrix. As with the Gaussian we have the following property:

Proposition 3.4.4.1

$$\mathbf{X} \ \textit{has pdf eq(3.60)} \overset{\textit{iff}}{\Longleftrightarrow} \mathrm{vec}(\mathbf{X}) \sim \mathcal{EC}_{NL}\left[\mathbf{1}_L \otimes \mathbf{m}, \mathbf{I}_L \otimes \mathbf{R},\ g(\cdot)\right]. \tag{3.61}$$

Hence, we indicate that $\mathbf{X}$ is real MEC by the notation

$$\mathbf{X} \sim \mathcal{MEC}_{NL}\left[\mathbf{1}_L \otimes \mathbf{m}, \mathbf{I}_L \otimes \mathbf{R},\ g(\cdot)\right]. \tag{3.62}$$

At this point we have developed enough theory to proceed to complex EC distributions, and will therefore do so. Further properties of real ECs can be inferred from those we developed for the complex EC.

3.5 Complex Elliptically Contoured Distributions

In array processing the data encountered is typically complex or modeled as complex. Complex data can arise from in phase quadrature demodulation of band pass signals. This is a common occurrence in radar/sonar, for example. In addition, the Fourier domain (yielding complex data) is sometimes preferred over time domain processing. It was for reasons like these that (i) the complex Gaussian distribution was first introduced [65], and has since gained popularity, and (ii) here we focus primarily on complex distributions (although the analogous results for the real case are directly obtainable from the complex). As the complex exponential $e^{j\theta}$ has historically proven itself to be an extremely convenient and quite prolific mathematical artifice, the concept of a complex density likewise proves to be very useful.

3.5.1 Definition

An $N \times 1$ complex vector $\mathbf{x}_R + j\mathbf{x}_I \in \mathfrak{C}^N$ is said to be complex EC distributed if and only if the $2N \times 1$ real vector $[\mathbf{x}_R; \mathbf{x}_I] \in \Re^{2N}$ is jointly real EC distributed. As with the complex Gaussian, we restrict attention to ECs with certain covariance structures imposed. If we let

$$\mathbf{y} = \begin{bmatrix} \mathbf{x}_R \\ \mathbf{x}_I \end{bmatrix} \qquad \mathbf{m}_y = \begin{bmatrix} \mathbf{m}_R \\ \mathbf{m}_I \end{bmatrix} \qquad \mathbf{t} = \begin{bmatrix} \mathbf{t}_R \\ \mathbf{t}_I \end{bmatrix} \qquad \tilde{\mathbf{t}} = \mathbf{t}_R + j\mathbf{t}_I$$
$$\tilde{\mathbf{x}} = \mathbf{x}_R + j\mathbf{x}_I \quad \tilde{\mathbf{m}}_x = \mathbf{m}_R + j\mathbf{m}_I \quad \tilde{\mathbf{R}}_x = \mathbf{C}_1 + j\mathbf{C}_2 \quad \mathrm{cov}(\tilde{\mathbf{x}}) \triangleq E\{(\tilde{\mathbf{x}} - \tilde{\mathbf{m}}_x)(\tilde{\mathbf{x}} - \tilde{\mathbf{m}}_x)^H\} \tag{3.63}$$

then we require that

$$\mathbf{y} \sim \mathcal{EC}_{2N}\left(\mathbf{m}_y = \begin{bmatrix} \mathbf{m}_R \\ \mathbf{m}_I \end{bmatrix}, \mathbf{R}_y = \frac{1}{2}\begin{bmatrix} \mathbf{C}_1 & -\mathbf{C}_2 \\ \mathbf{C}_2 & \mathbf{C}_1 \end{bmatrix}, \phi(\cdot); g(\cdot)\right) \tag{3.64}$$

where $\mathbf{C}_1$ is symmetric and $\mathbf{C}_2$ is skew symmetric. This covariance structure is typically referred to as circularity [67], and leads to the equivalences (see Chapter 15 of [45])

$$\begin{array}{llcl} (i) & \mathbf{t}^T\mathbf{y} & = & \mathrm{Re}(\tilde{\mathbf{t}}^H\tilde{\mathbf{x}}) \\ (ii) & \mathbf{t}^T\mathbf{R}_y\mathbf{t} & = & \tilde{\mathbf{t}}^H\tilde{\mathbf{R}}_x\tilde{\mathbf{t}}/2 \\ (iii) & (\mathbf{y}-\mathbf{m}_y)^T\mathbf{R}_y^{-1}(\mathbf{y}-\mathbf{m}_y)/2 & = & (\tilde{\mathbf{x}}-\tilde{\mathbf{m}}_x)^H\tilde{\mathbf{R}}_x^{-1}(\tilde{\mathbf{x}}-\tilde{\mathbf{m}}_x) \\ (iv) & 2^{-(2N)/2}|\mathbf{R}_y|^{-1/2} & = & |\mathbf{R}_x|^{-1} \end{array} \tag{3.65}$$

which ultimately allow the c.f. and density of $\mathbf{y}$ to be rewritten in terms of the complex variable $\tilde{\mathbf{x}}$:

$$\begin{array}{c} E\{e^{j\mathbf{t}^T\mathbf{y}}\} = e^{j\mathbf{t}^T\mathbf{m}_y} \cdot \phi(\mathbf{t}^T\mathbf{R}_y\mathbf{t}) = e^{j\mathrm{Re}(\tilde{\mathbf{t}}^H\tilde{\mathbf{m}}_x)} \cdot \phi(\tilde{\mathbf{t}}^H\tilde{\mathbf{R}}_x\tilde{\mathbf{t}}/2) = E\{e^{j\mathrm{Re}(\tilde{\mathbf{t}}^H\tilde{\mathbf{x}})}\} \\ 2^{-(2N)/2}|\mathbf{R}_y|^{-1/2}g\left[(\mathbf{y}-\mathbf{m}_y)^T\mathbf{R}_y^{-1}(\mathbf{y}-\mathbf{m}_y)/2\right] = |\tilde{\mathbf{R}}_x|^{-1}g\left[(\tilde{\mathbf{x}}-\tilde{\mathbf{m}}_x)^H\tilde{\mathbf{R}}_x^{-1}(\tilde{\mathbf{x}}-\tilde{\mathbf{m}}_x)\right]. \end{array} \tag{3.66}$$

We denote the complex EC by

$$\tilde{\mathbf{x}} \sim \mathcal{CEC}_N\left[\tilde{\mathbf{m}}_x, \tilde{\mathbf{R}}_x, \phi(\cdot) \text{ and/or } g(\cdot)\right] \tag{3.67}$$

If $\mathbf{R}_y$ does not possess the indicated covariance structure, then the mapping representation is not possible in terms of $\tilde{\mathbf{x}}$ and one must work exclusively with the real $\mathbf{y}$.

Examples of complex EC distributions include, but are not limited to:

complex Gaussian	$\longrightarrow$	$g(a) = \pi^{-N} e^{-a}$
complex Gaussian mixture	$\longrightarrow$	$g(a) = \pi^{-N} \int_0^\infty \frac{1}{\tau^N} e^{-a/\tau} P_\tau \, d\tau$
complex multivariate t w/ L d.o.f.	$\longrightarrow$	$g(a) = \pi^{-N} L! \, (1+a)^{-(L+1)}/(L-N)!$
complex K w/ parameters ν, ϖ	$\longrightarrow$	$g(a) = 2\pi^{-N} \varpi^\nu (a/\varpi)^{\frac{(\nu-N)}{2}} K_{\nu-N}(2\sqrt{a\varpi})/\Gamma(\nu)$.

The K and t distributions are in fact special cases of the Gaussian mixture.

It should be clear that one can obtain a complex EC distribution from a real EC by simply making the replacements indicated by eq(3.65). Because there does exist this direct mapping of a real EC (with some structure) to a corresponding complex EC and vice versa, all of the theory developed for the real EC applies to the complex. Note also that the mapping preserves the functional forms of $g(\cdot)$ and $\phi(\cdot)$ when going from real to complex, or vice versa.

We close this section with the following proposition which is useful and well known [22, 65]:

Proposition 3.5.1.1 *If $\tilde{\mathbf{x}}$ has pdf $P_{\tilde{\mathbf{x}}}(\tilde{\mathbf{x}}_0)$ and $\tilde{\mathbf{y}} = \tilde{\mathbf{A}}\tilde{\mathbf{x}} + \tilde{\mathbf{b}}$ with $\tilde{\mathbf{A}}$ nonsingular, then $J(\tilde{\mathbf{x}} \to \tilde{\mathbf{y}}) = |\tilde{\mathbf{A}}|^2$ and $P_{\tilde{\mathbf{y}}}(\tilde{\mathbf{y}}_0) = |\tilde{\mathbf{A}}|^2 P_{\tilde{\mathbf{x}}}[\tilde{\mathbf{A}}^{-1}(\tilde{\mathbf{y}}_0 - \tilde{\mathbf{b}})]$.*

3.5.2 Complex Gaussian Related Distributions

The complex Gaussian distribution is a special case of the complex EC distribution, and has been studied in great detail. As with the real Gaussian, there are several distributions which arise from the complex Gaussian distribution. As we'll later demonstrate, some of these complex Gaussian related distributions likewise often arise from complex EC distributions in general. In this section we simply list some common distributions related to the complex Gaussian. The following results are well-known for the real case and found in most text on statistical analysis [51, 52, 54]. These complex analogs are developed and appear in Appendix A of [22].

Let $\tilde{\mathbf{z}} \sim \mathcal{CN}_N(\mathbf{0}, \mathbf{I}_N)$ and $\tilde{\mathbf{z}}_\delta \sim \mathcal{CN}_N(\tilde{\mathbf{m}}, \mathbf{I}_N)$.

Proposition 3.5.2.1 *The pdf of the random variable $\tilde{\rho} = ||\tilde{\mathbf{z}}||^2$ is given by a complex central chi-square of N degrees of freedom:*

$$\tilde{\rho}^{N-1} e^{-\tilde{\rho}}/(N-1)! \qquad \tilde{\rho} \geq 0. \tag{3.68}$$

We denote this distribution by $\tilde{\rho} \sim \chi^2_N$.

Proposition 3.5.2.2 *The pdf of the random variable $\tilde{\rho}_\delta = ||\tilde{\mathbf{z}}_\delta||^2$ is given by a complex non-central chi-square of N degrees of freedom:*

$$\tilde{\rho}_\delta^{N-1} e^{-\tilde{\rho}_\delta} e^{-\delta^2} {}_0F_1(N; \delta^2\, \tilde{\rho}_\delta)/(N-1)! \qquad \tilde{\rho}_\delta \geq 0 \tag{3.69}$$

where ${}_0F_1(\cdot)$ is a hypergeometric function [73, 74], and the non-centrality parameter is given by $\delta^2 = ||\tilde{\mathbf{m}}||^2$. We denote this distribution by $\tilde{\rho}_\delta \sim \chi^2_N(\delta)$.

Proposition 3.5.2.3 *The pdf of the random variable $\tilde{F}_{N,M} = \chi^2_N/\chi^2_M$ is given by*

$$\frac{(N+M-1)!}{(N-1)!(M-1)!} \frac{\tilde{F}^{N-1}}{(1+\tilde{F})^{(N+M)}} \qquad \tilde{F} \geq 0, \tag{3.70}$$

a complex central $\tilde{F}$-distribution with N, M degrees of freedom.

Proposition 3.5.2.4 *The pdf of the random variable $\tilde{F}_{N,M}(\delta) = \chi^2_N(\delta)/\chi^2_M$ is given by*

$$\frac{(N+M-1)!}{(N-1)!(M-1)!} \frac{\tilde{F}^{N-1}}{(1+\tilde{F})^{(N+M)}} e^{-\delta^2} {}_1F_1[N+M; N; \delta^2\, \tilde{F}/(1+\tilde{F})] \qquad \tilde{F} \geq 0 \tag{3.71}$$

a complex non-central $\tilde{F}$-distribution with N, M degrees of freedom. In addition its cumulative distribution function (cdf) is given by

$$Pr(\tilde{F} \leq x) = \frac{x^N}{(1+x)^{N+M-1}} \sum_{k=0}^{M-1} \begin{pmatrix} N+M-1 \\ k+N \end{pmatrix} \cdot x^k \cdot IG_{k+1}\left(\frac{\delta^2}{1+x}\right) \tag{3.72}$$

where $IG_m(a) = e^{-a} \sum_{k=0}^{m-1} a^k/k!$ is the incomplete Gamma function. The cdf for the central complex F is the special case of $\delta = 0$.

Proposition 3.5.2.5 *The pdf of the random variable* $\tilde{\beta}_{N,M} = 1/(1+\tilde{F}_{M,N})$ *is given by*

$$\frac{(N+M-1)!}{(N-1)!(M-1)!}\tilde{\beta}^{N-1}(1-\tilde{\beta})^{M-1} \qquad 0 \leq \tilde{\beta} \leq 1, \tag{3.73}$$

a complex central beta distribution. In addition its cdf is given by the convenient finite sum

$$Pr(\tilde{\beta} \leq x) = x^{N+M-1}\sum_{k=0}^{M-1}\binom{N+M-1}{k}\left(\frac{1-x}{x}\right)^k. \tag{3.74}$$

Noting that $Pr(\tilde{F}_{N,M} \leq x) = Pr(-1+1/\tilde{\beta}_{M,N} \leq x)$, a second representation for the cdf of the central $\tilde{F}$ is possible.

Proposition 3.5.2.6 *The pdf of the random variable* $\tilde{\beta}_{N,M}(\delta) = 1/(1+\tilde{F}_{M,N}(\delta))$ *is given by*

$$\frac{(N+M-1)!}{(N-1)!(M-1)!}\tilde{\beta}^{N-1}(1-\tilde{\beta})^{M-1}e^{-\delta^2}{}_1F_1[N+M; M; \delta^2\,(1-\tilde{\beta})] \tag{3.75}$$
$$0 \leq \tilde{\beta} \leq 1$$

a complex non-central beta distribution.

3.5.3 Complex Spherically Symmetric Distributions

A complex SS (CSS) distribution is the special case of the complex EC distribution with parameters $\tilde{\mathbf{m}} = \mathbf{0}$ and $\tilde{\mathbf{R}} \propto \mathbf{I}$. From the definition of a complex EC distribution it follows that the c.f. of a CSS r.vc. is

$$\psi_{\tilde{\mathbf{z}}}(\tilde{\mathbf{t}}_z) = E\{e^{j\mathrm{Re}(\tilde{\mathbf{t}}_z^H\tilde{\mathbf{z}})}\} = \phi(||\tilde{\mathbf{t}}_z||^2/2). \tag{3.76}$$

The proofs for most of the following propositions are analogous to the real case, and therefore omitted.

Proposition 3.5.3.1 *If* $\tilde{\mathbf{z}}$ *is CSS then its distribution is invariant to any unitary transformation, i.e. for* $\tilde{\mathbf{Q}}$ *unitary* $\tilde{\mathbf{y}} = \tilde{\mathbf{Q}}\tilde{\mathbf{z}}$ *and* $\tilde{\mathbf{z}}$ *are identically distributed.*

Clearly, from eq(3.66) the density of a CSS distributed $\tilde{\mathbf{z}} \in \mathfrak{C}^N$ is of the form $g(||\tilde{\mathbf{z}}||^2)$, and that of its squared magnitude $\tilde{\rho} = ||\tilde{\mathbf{z}}||^2$ is given by

$$P_{\tilde{\rho}} = \frac{\pi^N}{(N-1)!}\tilde{\rho}^{N-1}g(\tilde{\rho}). \tag{3.77}$$

A very important example of a CSS is the uniformly distributed complex random vector.

Definition 3.5.1 *A complex random vector* $\tilde{\mathbf{u}}^{(N)} = \mathbf{u}_R + j\mathbf{u}_I \in \mathfrak{C}^N$ *with* $||\tilde{\mathbf{u}}^{(N)}||^2 = 1$ *is said to be uniformly distributed if*

$$\begin{bmatrix} \mathbf{u}_R \\ \mathbf{u}_I \end{bmatrix} \in \Re^{2N} \text{ has c.f. } \begin{array}{rcl} \Omega_{2N}(||\mathbf{t}||^2) & = & (N-1)!(2/||\tilde{\mathbf{t}}||^2)^{(N-1)}J_{(N-1)}(||\tilde{\mathbf{t}}||^2) \\ & \triangleq & \tilde{\Omega}_N(||\tilde{\mathbf{t}}||^2) = E\{e^{j\mathrm{Re}(\tilde{\mathbf{t}}^H\tilde{\mathbf{u}}^{(N)})}\}. \end{array} \tag{3.78}$$

Proposition 3.5.3.2 *Concerning* $\tilde{\mathbf{u}}^{(N)}$

$$E\{\tilde{\mathbf{u}}^{(N)}\} = \mathbf{0} \qquad \text{and} \qquad \mathrm{cov}(\tilde{\mathbf{u}}^{(N)}) = \frac{1}{N}\mathbf{I}_N. \tag{3.79}$$

Proposition 3.5.3.3 *The function* $\psi_{\tilde{\mathbf{z}}}(\tilde{\mathbf{t}})$ *is the c.f. of a CSS random vector* $\tilde{\mathbf{z}} \in \mathfrak{C}^N$ *if and only if*

$$\psi_{\tilde{\mathbf{z}}}(\tilde{\mathbf{t}}) = \int_0^\infty \tilde{\Omega}_N(||\tilde{\mathbf{t}}||^2\tilde{\rho})\; P_{\tilde{\rho}}\; d\tilde{\rho} = \phi(||\tilde{\mathbf{t}}||^2/2) \tag{3.80}$$

where $\tilde{\Omega}_N(\cdot)$ *is defined in eq(3.78) and* $P_{\tilde{\rho}}$ *is the pdf of some r.v.* $\tilde{\rho}$ *defined on* $0 \le \tilde{\rho} < \infty$.

Note from this proposition that the functions $\phi(\cdot)$ and $g(\cdot)$ for CSS and complex ECs in general, are likewise biuniquely related via the Hankel transform

$$\boxed{\phi(a/2) = \frac{\pi^N}{(N-1)!}\int_0^\infty \tilde{\Omega}_N(a\,\tilde{\rho})\;\tilde{\rho}^{N-1}g(\tilde{\rho})\; d\tilde{\rho}.} \tag{3.81}$$

Proposition 3.5.3.4 *If* $\tilde{\mathbf{z}} \in \mathfrak{C}^N$ *is CSS with c.f.* $\phi(||\tilde{\mathbf{t}}||^2/2)$ *and density* $g(||\tilde{\mathbf{z}}||^2)$ *necessarily related as in eq(3.81), then it can be written equal in distribution to*

$$\tilde{\mathbf{z}} \stackrel{d}{=} \tilde{\mathbf{u}}^{(N)} \cdot \sqrt{\tilde{\rho}} \tag{3.82}$$

where $\tilde{\rho} \sim P_{\tilde{\rho}}$ *is given by eq(3.77) with* $\tilde{\rho}$ *and* $\tilde{\mathbf{u}}^{(N)}$ *independent of each other.*

This follows immediately from the expectation

$$\begin{aligned} E\{e^{j\mathrm{Re}(\tilde{\mathbf{t}}^H\tilde{\mathbf{u}}^{(N)}\cdot\sqrt{\tilde{\rho}})}\} &= \int_0^\infty P_{\tilde{\rho}}\, d\tilde{\rho}\, \frac{1}{\tilde{B}_N}\int_{||\tilde{\mathbf{a}}||=1} e^{j\mathrm{Re}(\tilde{\mathbf{t}}^H\tilde{\mathbf{a}}\cdot\sqrt{\tilde{\rho}})}d\tilde{\mathbf{a}} \\ &= \int_0^\infty P_{\tilde{\rho}}\, d\tilde{\rho}\, \tilde{\Omega}_N(||\tilde{\mathbf{t}}||^2\,\tilde{\rho}) \\ &= \phi(||\tilde{\mathbf{t}}||^2/2) = \psi_{\tilde{\mathbf{z}}}(\tilde{\mathbf{t}}). \end{aligned} \tag{3.83}$$

The next proposition easily follows from this fact:

Proposition 3.5.3.5 *If* $\tilde{\mathbf{z}} \in \mathbb{C}^N$ *is CSS with c.f.* $\phi(||\tilde{\mathbf{t}}||^2/2)$ *and density* $g(||\tilde{\mathbf{z}}||^2)$ *necessarily related as in eq(3.81), then* $||\tilde{\mathbf{z}}||^2 \overset{d}{=} \tilde{\rho} \sim P_{\tilde{\rho}}$ *and* $\tilde{\mathbf{z}}/||\tilde{\mathbf{z}}|| \overset{d}{=} \tilde{\mathbf{u}}^{(N)}$ *and they are independent.*

Proposition 3.5.3.6 *With* $\tilde{\mathbf{u}}^{(N)}$ *complex uniformly distributed, its partition satisfies*

$$\tilde{\mathbf{u}}^{(N)} = \begin{bmatrix} (\tilde{\mathbf{u}}_1)_{M\times 1} \\ (\tilde{\mathbf{u}}_2)_{(N-M)\times 1} \end{bmatrix} \overset{d}{=} \begin{bmatrix} \tilde{\mathbf{u}}^{(M)}\cdot\sqrt{\tilde{\beta}_{M,(N-M)}} \\ \tilde{\mathbf{u}}^{(N-M)}\cdot\sqrt{1-\tilde{\beta}_{M,(N-M)}} \end{bmatrix} \tag{3.84}$$

where $\tilde{\mathbf{u}}^{(M)}$ *and* $\tilde{\mathbf{u}}^{(N-M)}$ *are independent.*

To prove note that with $\tilde{\mathbf{z}} \sim \mathcal{CN}_N(\mathbf{0}, \mathbf{I}_N)$ it follows that

$$\tilde{\mathbf{u}}^{(N)} \overset{d}{=} \frac{\tilde{\mathbf{z}}}{||\tilde{\mathbf{z}}||} = \begin{bmatrix} \tilde{\mathbf{z}}_1/||\tilde{\mathbf{z}}|| \\ \tilde{\mathbf{z}}_2/||\tilde{\mathbf{z}}|| \end{bmatrix} \overset{d}{=} \begin{bmatrix} \tilde{\mathbf{u}}^{(M)}\dfrac{||\tilde{\mathbf{z}}_1||}{||\tilde{\mathbf{z}}||} \\ \tilde{\mathbf{u}}^{(N-M)}\dfrac{||\tilde{\mathbf{z}}_2||}{||\tilde{\mathbf{z}}||} \end{bmatrix}. \tag{3.85}$$

That $||\tilde{\mathbf{z}}_1||^2/||\tilde{\mathbf{z}}||^2 = ||\tilde{\mathbf{z}}_1||^2/(||\tilde{\mathbf{z}}_1||^2 + ||\tilde{\mathbf{z}}_2||^2) = 1/(1 + ||\tilde{\mathbf{z}}_2||^2/||\tilde{\mathbf{z}}_1||^2) \overset{d}{=} 1/(1+\chi^2_{(N-M)}/\chi^2_M)$ with the aid of propositions in Section 3.5.2 completes proof.

Proposition 3.5.3.7 *If* $\tilde{\mathbf{z}} = [(\tilde{\mathbf{z}}_1)_{M\times 1}; (\tilde{\mathbf{z}}_2)_{(N-M)\times 1}]$ *is CSS, then the conditional distribution of* $\tilde{\mathbf{z}}_1$ *given that* $\tilde{\mathbf{z}}_2 = \tilde{\mathbf{a}}$ *is such that*

$$(\tilde{\mathbf{z}}_1|\,\tilde{\mathbf{z}}_2 = \tilde{\mathbf{a}}) \overset{d}{=} \tilde{\mathbf{u}}^{(M)}\sqrt{\tilde{\rho}_{\tilde{\mathbf{a}}}} \qquad \text{where} \qquad \tilde{\rho}_{\tilde{\mathbf{a}}} \overset{d}{=} \tilde{\rho} - ||\tilde{\mathbf{a}}||^2. \tag{3.86}$$

From Propositions 3.5.3.4 and 3.5.3.6 we know that the directions of $\tilde{\mathbf{z}}_1$ and $\tilde{\mathbf{z}}_2$ are independent and uniformly distributed. Thus, *the conditioning can only effect the magnitude.* By Proposition 3.5.3.5 it follows that $\tilde{\rho} \overset{d}{=} ||\tilde{\mathbf{z}}||^2 = ||\tilde{\mathbf{z}}_1||^2 + ||\tilde{\mathbf{z}}_2||^2$. With $||\tilde{\mathbf{z}}_1||^2 \overset{d}{=} \tilde{\rho} - ||\mathbf{z}_2||^2$ the result is proved.

Theorem 3.5.1 *If* $\tilde{\mathbf{z}} = [\tilde{\mathbf{z}}_1; \tilde{\mathbf{z}}_2]$ *is CSS then* $\tilde{\mathbf{z}}^* \stackrel{d}{=} \tilde{\mathbf{z}}$, *and* $[\tilde{\mathbf{z}}_1; \tilde{\mathbf{z}}_2^*] \stackrel{d}{=} \tilde{\mathbf{z}}$, *where superscript* $*$ *represents elementwise complex conjugation.*

To prove consider the real counterpart of $\tilde{\mathbf{z}} = \mathbf{z}_R + j\mathbf{z}_I$. Let $\mathbf{y} = [\mathbf{z}_R; \mathbf{z}_I]$ and note that $\mathbf{y}$ is SS, and thus $\mathbf{a} = \mathbf{Q}\mathbf{y} \stackrel{d}{=} \mathbf{y}$ for all orthogonal $\mathbf{Q}$. Choosing

$$\mathbf{Q} = \begin{bmatrix} \mathbf{I}_N & \mathbf{0} \\ \mathbf{0} & -\mathbf{I}_N \end{bmatrix} \tag{3.87}$$

shows that $\mathbf{a} = [\mathbf{z}_R; -\mathbf{z}_I] \stackrel{d}{=} \mathbf{y}$, *i.e.* $\mathbf{z}_R - j\mathbf{z}_I = \tilde{\mathbf{z}}^* \stackrel{d}{=} \tilde{\mathbf{z}}$, and completes proof.

Partition a CSS random vector such that $\tilde{\mathbf{z}} = [(\tilde{\mathbf{z}}_1)_{M\times 1}; (\tilde{\mathbf{z}}_2)_{(N-M)\times 1}]$. If $\tilde{\mathbf{z}}$ has pdf $g(||\tilde{\mathbf{z}}||)$, then the marginal density for $\tilde{\mathbf{z}}_1$ is clearly given by

$$P_{\tilde{\mathbf{z}}_1} = \int_{\tilde{\mathbf{z}}_2 \in \mathbb{C}^{(N-M)}} g(||\tilde{\mathbf{z}}_1||^2 + ||\tilde{\mathbf{z}}_2||^2) d\tilde{\mathbf{z}}_2 = \frac{\pi^{(N-M)}}{(N-M-1)!} \int_0^\infty u^{(N-M)-1} g(||\tilde{\mathbf{z}}_1||^2 + u) du \tag{3.88}$$

where the last equality follows from Proposition 3.3.2.1. This integral operation will be prominent in subsequent analysis. It is worthwhile, therefore, to make the following definition:

Definition 3.5.2 *When* $0 \le M < N$, *let*

$$\boxed{g_{M:N}(a) \stackrel{\Delta}{=} \frac{\pi^{(N-M)}}{(N-M-1)!} \int_0^\infty u^{(N-M)-1} g(a+u) du,} \tag{3.89}$$

when $M = N$, *then we say that* $g_{N:N}(a) \stackrel{\Delta}{=} g(a)$, *otherwise its undefined.*

The subscript $M:N$ can be read "M-dimensional density, obtained from an N-dimensional density." Clearly $P_{\tilde{\mathbf{z}}_1} = g_{M:N}(||\tilde{\mathbf{z}}_1||^2)$. This operation actually has several interesting properties. Several of which appear in [54, 59]. We close this section by listing a few of these properties.

Proposition 3.5.3.8 *The exponential is an eigenfunction under the operator* $g_{M:N}$; *namely, if* $g(a) = \pi^{-N} e^{-a}$ *then* $g_{M:N}(a) = \pi^{-M} e^{-a}$.

Proposition 3.5.3.9 *The functional forms of Gaussian mixtures/compounds regenerate* *under the operator $g_{M:N}$, i.e.*

$$g(a) = \pi^{-N} \int_0^\infty \frac{1}{\tau^N} e^{-a/\tau} P_\tau \, d\tau \Longrightarrow g_{M:N}(a) = \pi^{-M} \int_0^\infty \frac{1}{\tau^M} e^{-a/\tau} P_\tau \, d\tau. \tag{3.90}$$

Theorem 3.5.2 *If $g(||\tilde{\mathbf{z}}||^2)$ is an N-dimensional pdf, then $g_{0:N}(a = 0) = 1$.*

This follows from the fact that eq(3.77) is a density.

Theorem 3.5.3 *If $g(||\tilde{\mathbf{z}}||^2)$ is an N-dimensional pdf and $0 \le M \le L \le N$ and $q(a) = g_{L:N}(a)$ then*

$$q_{M:L}(a) = [\ g_{L:N}(\cdot)\ ]_{M:L}(a) = g_{M:N}(a). \tag{3.91}$$

The proof follows from known properties of densities. Let

$$\tilde{\mathbf{z}} = [(\tilde{\mathbf{z}}_1)_{M\times 1}; (\tilde{\mathbf{z}}_2)_{(N-M)\times 1}] = [(\tilde{\mathbf{z}}_1)_{M\times 1}; (\tilde{\mathbf{z}}_A)_{(L-M)\times 1}; (\tilde{\mathbf{z}}_B)_{(N-L)\times 1}]. \tag{3.92}$$

That

$$P_{\tilde{\mathbf{z}}_1} = \int_{\tilde{\mathbf{z}}_2 \in \mathbb{C}^{(N-M)}} d\tilde{\mathbf{z}}_2 \, P_{\tilde{\mathbf{z}}_1,\tilde{\mathbf{z}}_2} = \int_{\tilde{\mathbf{z}}_A \in \mathbb{C}^{(L-M)}} d\tilde{\mathbf{z}}_A \int_{\tilde{\mathbf{z}}_B \in \mathbb{C}^{(N-L)}} d\tilde{\mathbf{z}}_B \, P_{\tilde{\mathbf{z}}_1,\tilde{\mathbf{z}}_A,\tilde{\mathbf{z}}_B} \tag{3.93}$$

completes proof.

Proposition 3.5.3.10 *$g_{M:N}(r)$ for $M \le N-1$ is guaranteed to be a non-increasing function on $0 < r < \infty$ for well-behaved $g(r)$.*

3.5.4 Properties of Complex ECs

Proposition 3.5.4.1 *If $\tilde{\mathbf{x}} \sim \mathcal{CEC}_N(\tilde{\mathbf{m}}, \tilde{\mathbf{R}}, \phi)$ then it has the equivalent stochastic representation*

$$\tilde{\mathbf{x}} \stackrel{d}{=} \tilde{\mathbf{m}} + \tilde{\mathbf{R}}^{1/2} \tilde{\mathbf{u}}^{(N)} \sqrt{\tilde{\rho}} \tag{3.94}$$

where $P_{\tilde{\rho}}$ is related to ϕ via eq(3.80).

Noting that

$$\begin{aligned} E\{e^{j\mathrm{Re}[\tilde{\mathbf{t}}^H(\tilde{\mathbf{m}}+\tilde{\mathbf{R}}^{1/2}\tilde{\mathbf{u}}^{(N)}\sqrt{\tilde{\rho}})]}\} &= e^{j\mathrm{Re}(\tilde{\mathbf{t}}^H\tilde{\mathbf{m}})} E\{e^{j\mathrm{Re}[(\sqrt{\tilde{\rho}}\,\tilde{\mathbf{R}}^{1/2}\tilde{\mathbf{t}})^H\tilde{\mathbf{u}}^{(N)}]}\} \\ &= e^{j\mathrm{Re}(\tilde{\mathbf{t}}^H\tilde{\mathbf{m}})} \int_0^\infty \tilde{\Omega}_N(\tilde{\mathbf{t}}^H\tilde{\mathbf{R}}\tilde{\mathbf{t}}\tilde{\rho}) \, P_{\tilde{\rho}} \, d\tilde{\rho} = e^{j\mathrm{Re}(\tilde{\mathbf{t}}^H\tilde{\mathbf{m}})} \phi(\tilde{\mathbf{t}}^H\tilde{\mathbf{R}}\tilde{\mathbf{t}}/2) \end{aligned} \tag{3.95}$$

completes proof.

Proposition 3.5.4.2 *If* $\tilde{\mathbf{x}} \sim \mathcal{CEC}_N[\tilde{\mathbf{m}}, \tilde{\mathbf{R}}, \phi(\cdot);\ g(\cdot)]$, *then*

$$E\{\tilde{\mathbf{x}}\} = \tilde{\mathbf{m}} \qquad \textit{and} \qquad \operatorname{cov}(\tilde{\mathbf{x}}) = \frac{E\{\tilde{\rho}\}}{N}\tilde{\mathbf{R}} \tag{3.96}$$

where the pdf for $\tilde{\rho}$ *is* $P_{\tilde{\rho}}$ *as in eq(3.77).*

Proposition 3.5.4.3 *If* $\tilde{\mathbf{x}} \sim \mathcal{CEC}_N[\tilde{\mathbf{m}}, \tilde{\mathbf{R}}, \phi; (\cdot)\ g(\cdot)]$ *and* $\tilde{\mathbf{A}}^H$ $(M \times N)$ *is full rank with* $M \leq N$, *then*

$$\tilde{\mathbf{y}} = \tilde{\mathbf{A}}^H\tilde{\mathbf{x}} \sim \mathcal{CEC}_M\left[\tilde{\mathbf{A}}^H\tilde{\mathbf{m}}, \tilde{\mathbf{A}}^H\tilde{\mathbf{R}}\tilde{\mathbf{A}}, \phi(\cdot);\ g_{M:N}(\cdot)\right]. \tag{3.97}$$

The c.f. of $\tilde{\mathbf{y}}$ follows directly from multiplication of eq(3.94) by $\tilde{\mathbf{A}}^H$. Concerning the pdf, let $\tilde{\mathbf{z}} \overset{d}{=} \tilde{\mathbf{u}}^{(N)}\sqrt{\tilde{\rho}}$ and note that for unitary $\tilde{\mathbf{Q}}$

$$\tilde{\mathbf{y}} \overset{d}{=} \tilde{\mathbf{A}}^H\tilde{\mathbf{m}} + \tilde{\mathbf{A}}^H\tilde{\mathbf{R}}^{1/2}\tilde{\mathbf{z}} \overset{d}{=} \tilde{\mathbf{A}}^H\tilde{\mathbf{m}} + \tilde{\mathbf{A}}^H\tilde{\mathbf{R}}^{1/2}\tilde{\mathbf{Q}}\tilde{\mathbf{z}}. \tag{3.98}$$

Choosing $\tilde{\mathbf{Q}}$ such that $\tilde{\mathbf{A}}^H\tilde{\mathbf{R}}^{1/2}\tilde{\mathbf{Q}} = [(\Delta_A^H)_{M\times M}, \mathbf{0}_{M\times(N-M)}]$ where $\Delta_A^H\Delta_A = \tilde{\mathbf{A}}^H\tilde{\mathbf{R}}\tilde{\mathbf{A}}$, and partitioning $\tilde{\mathbf{z}}$ accordingly implies that

$$\tilde{\mathbf{y}} \overset{d}{=} \tilde{\mathbf{A}}^H\tilde{\mathbf{m}} + (\tilde{\mathbf{A}}^H\tilde{\mathbf{R}}\tilde{\mathbf{A}})^{1/2}\tilde{\mathbf{z}}_1. \tag{3.99}$$

That $(\tilde{\mathbf{A}}^H\tilde{\mathbf{R}}\tilde{\mathbf{A}})^{1/2}$ is nonsingular and $\tilde{\mathbf{z}}_1$ has marginals given by eq(3.88) completes proof.

Proposition 3.5.4.4 *Let* $\tilde{\mathbf{x}} \sim \mathcal{CEC}_N[\tilde{\mathbf{m}}, \tilde{\mathbf{R}}, g(\cdot)]$ *and make the partitions*

$$\tilde{\mathbf{x}} = \begin{bmatrix} (\tilde{\mathbf{x}}_1)_{M\times 1} \\ (\tilde{\mathbf{x}}_2)_{(N-M)\times 1} \end{bmatrix} \qquad \tilde{\mathbf{m}} = \begin{bmatrix} \tilde{\mathbf{m}}_1 \\ \tilde{\mathbf{m}}_2 \end{bmatrix} \qquad \tilde{\mathbf{R}} = \begin{bmatrix} \tilde{\mathbf{R}}_{11} & \tilde{\mathbf{R}}_{12} \\ \tilde{\mathbf{R}}_{21} & \tilde{\mathbf{R}}_{22} \end{bmatrix}. \tag{3.100}$$

The conditional distribution of $\tilde{\mathbf{x}}_1$ *given* $\tilde{\mathbf{x}}_2$ *is such that*

$$(\tilde{\mathbf{x}}_1 | \tilde{\mathbf{x}}_2 = \tilde{\mathbf{a}}) \overset{d}{=} \tilde{\mathbf{m}}_1 + \tilde{\mathbf{R}}_{12}\tilde{\mathbf{R}}_{22}^{-1}(\tilde{\mathbf{a}} - \tilde{\mathbf{m}}_2) + \tilde{\mathbf{R}}_{11.2}^{1/2}\tilde{\mathbf{u}}^{(M)} \cdot \sqrt{\tilde{\rho}_{\tilde{\mathbf{a}}}} \tag{3.101}$$

where $\tilde{\rho}_{\tilde{\mathbf{a}}} \overset{d}{=} \tilde{\rho} - (\tilde{\mathbf{a}} - \tilde{\mathbf{m}}_2)^H\tilde{\mathbf{R}}_{22}^{-1}(\tilde{\mathbf{a}} - \tilde{\mathbf{m}}_2)$ *is independent of* $\tilde{\mathbf{u}}^{(M)}$, *and* $\tilde{\mathbf{R}}_{11.2} \overset{\Delta}{=} \tilde{\mathbf{R}}_{11} - \tilde{\mathbf{R}}_{12}\tilde{\mathbf{R}}_{22}^{-1}\tilde{\mathbf{R}}_{21}$.

To prove note that

$$\tilde{\mathbf{x}} = \begin{bmatrix} \tilde{\mathbf{x}}_1 \\ \tilde{\mathbf{x}}_2 \end{bmatrix} \overset{d}{=} \begin{bmatrix} \tilde{\mathbf{m}}_1 \\ \tilde{\mathbf{m}}_2 \end{bmatrix} + \begin{bmatrix} \tilde{\mathbf{R}}_{11.2}^{1/2} & \tilde{\mathbf{R}}_{12}\tilde{\mathbf{R}}_{22}^{-1/2} \\ \mathbf{0} & \tilde{\mathbf{R}}_{22}^{1/2} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{z}}_1 \\ \tilde{\mathbf{z}}_2 \end{bmatrix} \tag{3.102}$$

where $\tilde{\mathbf{z}} \sim \mathcal{CEC}_N[\mathbf{0}, \mathbf{I}_N, \phi(\cdot); g(\cdot)]$. Consequently, it follows that

$$\begin{aligned}(\tilde{\mathbf{x}}_1 | \tilde{\mathbf{x}}_2 = \tilde{\mathbf{a}}) &\stackrel{d}{=} \left(\tilde{\mathbf{m}}_1 + \tilde{\mathbf{R}}_{11.2}^{1/2}\tilde{\mathbf{z}}_1 + \tilde{\mathbf{R}}_{12}\tilde{\mathbf{R}}_{22}^{-1/2}\tilde{\mathbf{z}}_2\right)\Big| \tilde{\mathbf{x}}_2 \stackrel{d}{=} \tilde{\mathbf{m}}_2 + \tilde{\mathbf{R}}_{22}^{1/2}\tilde{\mathbf{z}}_2 = \tilde{\mathbf{a}} \\ &\stackrel{d}{=} \left[\left(\tilde{\mathbf{m}}_1 + \tilde{\mathbf{R}}_{12}\tilde{\mathbf{R}}_{22}^{-1/2}\tilde{\mathbf{z}}_2\right)\Big| \tilde{\mathbf{m}}_2 + \tilde{\mathbf{R}}_{22}^{1/2}\tilde{\mathbf{z}}_2 = \tilde{\mathbf{a}}\right] + \left[\tilde{\mathbf{R}}_{11.2}^{1/2}\tilde{\mathbf{z}}_1 \Big| \tilde{\mathbf{m}}_2 + \tilde{\mathbf{R}}_{22}^{1/2}\tilde{\mathbf{z}}_2 = \tilde{\mathbf{a}}\right].\end{aligned}$$

Solving we get $\tilde{\mathbf{z}}_2 = \tilde{\mathbf{R}}_{22}^{-1/2}(\tilde{\mathbf{a}} - \tilde{\mathbf{m}}_2)$, which implies that

$$(\tilde{\mathbf{x}}_1 | \tilde{\mathbf{x}}_2 = \tilde{\mathbf{a}}) \stackrel{d}{=} \tilde{\mathbf{m}}_1 + \tilde{\mathbf{R}}_{12}\tilde{\mathbf{R}}_{22}^{-1}(\tilde{\mathbf{a}} - \tilde{\mathbf{m}}_2) + \tilde{\mathbf{R}}_{11.2}^{1/2}\left[\tilde{\mathbf{z}}_1 | \tilde{\mathbf{z}}_2 = \tilde{\mathbf{R}}_{22}^{-1/2}(\tilde{\mathbf{a}} - \tilde{\mathbf{m}}_2)\right]. \qquad (3.103)$$

The result then follows from Proposition 3.5.3.7. Note especially that when $\tilde{\mathbf{x}}_1$ and $\tilde{\mathbf{x}}_2$ are uncorrelated, *i.e.* $\tilde{\mathbf{R}}_{12} = \tilde{\mathbf{R}}_{21}^H = \mathbf{0}$, the effect of the conditioning is completely summarized in the r.v. $\tilde{\rho}_{\tilde{\mathbf{a}}}$.

Proposition 3.5.4.5 *Let* $\tilde{\mathbf{x}} \sim \mathcal{CEC}_N[\mathbf{0}, \tilde{\mathbf{R}}, \phi(\cdot); g(\cdot)]$ *and* $\tilde{\mathbf{F}}(\tilde{\mathbf{x}})$ *be any matrix of statistics formed from* $\tilde{\mathbf{x}}$. *If the matrix of statistics is scale invariant such that* $\tilde{\mathbf{F}}(\tilde{c}\tilde{\mathbf{x}}) = \tilde{\mathbf{F}}(\tilde{\mathbf{x}})$ *at least for all real* $\tilde{c} > 0$, *then the distribution of* $\tilde{\mathbf{F}}(\tilde{\mathbf{x}})$ *is completely independent of the the functional form of* $g(\cdot)$ *(and* $\phi(\cdot)$*).*

To prove, simply note that $\tilde{\mathbf{F}}(\tilde{\mathbf{x}}) \stackrel{d}{=} \tilde{\mathbf{F}}\left(\tilde{\mathbf{R}}^{1/2}\tilde{\mathbf{u}}^{(N)} \cdot \sqrt{\tilde{\rho}}\right)$. Consider the conditional distribution of $\tilde{\mathbf{F}}(\cdot)$ given $\tilde{\rho}$, and note that by the scale invariance

$$\tilde{\mathbf{F}}(\tilde{\mathbf{x}}) \stackrel{d}{=} \tilde{\mathbf{F}}\left(\tilde{\mathbf{R}}^{1/2}\tilde{\mathbf{u}}^{(N)} \cdot \sqrt{\tilde{\rho}}\right) = \tilde{\mathbf{F}}\left(\tilde{\mathbf{R}}^{1/2}\tilde{\mathbf{u}}^{(N)}\right). \qquad (3.104)$$

That $\tilde{\mathbf{F}}(\cdot)$ is independent of $\tilde{\rho}$ when conditioned on $\tilde{\rho}$ implies that it is marginally independent as well. Since the presence of r.v. $\tilde{\rho}$ is the only way the functional form of $g(\cdot)$ can influence the distribution of $\tilde{\mathbf{F}}(\cdot)$, the proof is complete.

Theorem 3.5.4 *Let* $\tilde{\mathbf{x}} \sim \mathcal{CEC}_N[\tilde{\mathbf{m}}, \tilde{\mathbf{R}}, g(\cdot)]$ *and consider the partitioning in eq(3.100) with* $\tilde{\mathbf{m}}_2 = \mathbf{0}$ *and with* $\tilde{\mathbf{R}}_{12} = \tilde{\mathbf{R}}_{21}^H = \mathbf{0}$, *i.e with* $\tilde{\mathbf{x}}_1$ *and* $\tilde{\mathbf{x}}_2$ *uncorrelated. Consider two matrices of statistics* $\tilde{\mathbf{F}}_1(\tilde{\mathbf{x}}_1)$ *and* $\tilde{\mathbf{F}}_2(\tilde{\mathbf{x}}_2)$ *where the second is scale invariant such that* $\tilde{\mathbf{F}}_2(\tilde{c}\tilde{\mathbf{x}}_2) = \tilde{\mathbf{F}}_2(\tilde{\mathbf{x}}_2)$ *at least for all real* $\tilde{c} > 0$. *It follows that (i) the distribution of* $\tilde{\mathbf{F}}_2(\cdot)$ *does not depend on* $g(\cdot)$, *and (ii)* $\tilde{\mathbf{F}}_1(\cdot)$ *and* $\tilde{\mathbf{F}}_2(\cdot)$ *are completely independent.*

Part (i) follows from Proposition 3.5.4.5. Concerning part (ii), note from Proposition 3.5.4.4 that the conditional distribution of $\tilde{\mathbf{F}}_2(\cdot)$ given $\tilde{\mathbf{x}}_1 = \tilde{\mathbf{a}}$ is such that $\tilde{\mathbf{F}}_2(\tilde{\mathbf{x}}_2) \stackrel{d}{=}$

$\tilde{\mathbf{F}}_2(\tilde{\mathbf{R}}_{22}^{1/2}\tilde{\mathbf{u}}^{(N-M)}\sqrt{\tilde{\rho}_{\tilde{\mathbf{a}}}})$ where $\tilde{\rho}_{\tilde{\mathbf{a}}} \stackrel{d}{=} \tilde{\rho} - (\tilde{\mathbf{a}} - \tilde{\mathbf{m}}_1)^H \tilde{\mathbf{R}}_{11}^{-1}(\tilde{\mathbf{a}} - \tilde{\mathbf{m}}_1)$. Now condition on $\tilde{\rho}_{\tilde{\mathbf{a}}}$ and note by the scale invariance that $\tilde{\mathbf{F}}_2(\tilde{\mathbf{R}}_{22}^{1/2}\tilde{\mathbf{u}}^{(N-M)}\sqrt{\tilde{\rho}_{\tilde{\mathbf{a}}}}) = \tilde{\mathbf{F}}_2(\tilde{\mathbf{R}}_{22}^{1/2}\tilde{\mathbf{u}}^{(N-M)})$. Thus, $\tilde{\mathbf{F}}_2(\cdot)$ is independent of $\tilde{\rho}_{\tilde{\mathbf{a}}}$, and therefore independent of $\tilde{\mathbf{x}}_1$ which completes proof.

Proposition 3.5.4.6 *With $\tilde{\mathbf{x}} \sim \mathcal{CEC}_N[\tilde{\mathbf{m}}, \tilde{\mathbf{R}}, g(\cdot)]$ the k-th moment of $\tilde{\mathbf{x}}$ exists if and only if $E\{[\sqrt{\tilde{\rho}}]^k\} < \infty$.*

This follows from Proposition 3.5.4.1.

3.5.5 Sampling from Complex EC Distributions: The Complex MEC

When processing sensor array data one often considers a collection of complex data samples $\tilde{\mathbf{x}}_i \in \mathbb{C}^N$. Assume each sample is taken from the distribution $\mathcal{CEC}_N(\tilde{\mathbf{m}}, \tilde{\mathbf{R}}, g(\cdot))$, and define the totality of this data by the the complex data matrix

$$\tilde{\mathbf{X}} = [\tilde{\mathbf{x}}_1 | \tilde{\mathbf{x}}_2 | \cdots | \tilde{\mathbf{x}}_L]. \tag{3.105}$$

If these samples are independent, then the pdf of $\tilde{\mathbf{X}}$ is given by

$$\prod_{i=1}^{L} |\tilde{\mathbf{R}}|^{-1} g\left[(\tilde{\mathbf{x}}_i - \tilde{\mathbf{m}})^H \tilde{\mathbf{R}}^{-1}(\tilde{\mathbf{x}}_i - \tilde{\mathbf{m}})\right]. \tag{3.106}$$

This type of model has recently been considered for adaptive array detection by [8, 9].

An alternate model for the data matrix $\tilde{\mathbf{X}}$ is one in which its columns still share the same mean and covariance, but are only uncorrelated. The density for $\tilde{\mathbf{X}}$ is then given by

$$|\tilde{\mathbf{R}}|^{-L} g\left[\mathrm{tr}\tilde{\mathbf{R}}^{-1}(\tilde{\mathbf{X}} - \tilde{\mathbf{M}})(\tilde{\mathbf{X}} - \tilde{\mathbf{M}})^H\right] \tag{3.107}$$

where $\tilde{\mathbf{M}} = [\tilde{\mathbf{m}} | \cdots | \tilde{\mathbf{m}}]$ is $N \times L$. This is a special case of a complex MEC distribution often referred to as a complex LEC distribution. The models in eq(3.106) and eq(3.107) coincide if and only if $\tilde{\mathbf{X}}$ is a complex normal data matrix. As with the Gaussian we have the following property:

Proposition 3.5.5.1

$$\tilde{\mathbf{X}} \text{ has pdf eq(3.107)} \stackrel{iff}{\Longleftrightarrow} \mathrm{vec}(\tilde{\mathbf{X}}) \sim \mathcal{CEC}_{NL}\left[\mathbf{1}_L \otimes \tilde{\mathbf{m}}, \mathbf{I}_L \otimes \tilde{\mathbf{R}},\ g(\cdot)\right]. \tag{3.108}$$

More generally speaking, we have

Proposition 3.5.5.2

$$\tilde{\mathbf{X}} \text{ has pdf } |\tilde{\mathbf{R}}|^{-L}|\tilde{\boldsymbol{\Sigma}}|^{-N} g\left[\mathrm{tr}\tilde{\mathbf{R}}^{-1}(\tilde{\mathbf{X}}-\tilde{\mathbf{M}})\tilde{\boldsymbol{\Sigma}}^{-1}(\tilde{\mathbf{X}}-\tilde{\mathbf{M}})^H\right] \overset{iff}{\Longleftrightarrow} \tag{3.109}$$

$$\mathrm{vec}(\tilde{\mathbf{X}}) \sim \mathcal{CEC}_{NL}\left([\tilde{\mathbf{m}}_1;\tilde{\mathbf{m}}_2;\cdots;\tilde{\mathbf{m}}_L], \tilde{\boldsymbol{\Sigma}}^* \otimes \tilde{\mathbf{R}},\ g(\cdot)\right), \tag{3.110}$$

where $\tilde{\mathbf{M}} = [\tilde{\mathbf{m}}_1|\tilde{\mathbf{m}}_2|\cdots|\tilde{\mathbf{m}}_L]$ *and matrices* $\tilde{\mathbf{R}}$ *and* $\tilde{\boldsymbol{\Sigma}}$ *are assumed hpd.*

To prove let $\mathrm{vec}(\tilde{\mathbf{X}}_{N\times L}) = \tilde{\mathbf{x}}_v$. Given square matrices $\tilde{\mathbf{A}}_{N\times N}$, $\tilde{\mathbf{B}}_{L\times L}$, $\tilde{\mathbf{C}}_{N\times N}$, and $\tilde{\mathbf{D}}_{L\times L}$ the following properties hold

- $\mathrm{tr}(\tilde{\mathbf{A}}\tilde{\mathbf{X}}\tilde{\mathbf{B}}\tilde{\mathbf{X}}^H) = \tilde{\mathbf{x}}_v^H(\tilde{\mathbf{B}}^T \otimes \tilde{\mathbf{A}})\tilde{\mathbf{x}}_v$
- $|\tilde{\mathbf{B}} \otimes \tilde{\mathbf{A}}| = |\tilde{\mathbf{A}}|^L|\tilde{\mathbf{B}}|^N$
- $\mathrm{vec}(\tilde{\mathbf{C}}\tilde{\mathbf{X}}\tilde{\mathbf{D}}^*) = (\tilde{\mathbf{D}}^H \otimes \tilde{\mathbf{C}})\tilde{\mathbf{x}}_v$.

The proof of Proposition 3.5.5.2 is an immediate consequence of the first two properties. From Propositions 3.5.5.2 and 3.5.4.1 we obtain the rather convenient and useful stochastic representation

$$\tilde{\mathbf{x}}_v \overset{d}{=} \tilde{\mathbf{m}}_v + (\tilde{\boldsymbol{\Sigma}}^* \otimes \tilde{\mathbf{R}})^{1/2}\tilde{\mathbf{z}}_v \Longrightarrow \tilde{\mathbf{X}} \overset{d}{=} \tilde{\mathbf{M}} + \tilde{\mathbf{R}}^{1/2}\,\tilde{\mathbf{Z}}\,\tilde{\boldsymbol{\Sigma}}^{1/2} \tag{3.111}$$

where $\mathrm{vec}(\tilde{\mathbf{M}}) = \tilde{\mathbf{m}}_v$, $\mathrm{vec}(\tilde{\mathbf{Z}}) = \tilde{\mathbf{z}}_v \overset{d}{=} \tilde{\mathbf{u}}^{(NL)}\sqrt{\tilde{\rho}}$ is CSS with pdf $g(||\tilde{\mathbf{z}}_v||^2) = g(\mathrm{tr}\tilde{\mathbf{Z}}\tilde{\mathbf{Z}}^H)$. Hence, we sometimes indicate that $\tilde{\mathbf{X}}$ is complex MEC by the notation

$$\tilde{\mathbf{X}} \sim \mathcal{CMEC}_{NL}\left[\tilde{\mathbf{M}}, \tilde{\boldsymbol{\Sigma}}^* \otimes \tilde{\mathbf{R}},\ g(\cdot)\right]. \tag{3.112}$$

This model allows for the possibility of correlation among snapshots when $\tilde{\boldsymbol{\Sigma}}$ is not proportional to a diagonal matrix. Several properties of complex MECs follow from this stochastic relation. A very useful one is given in the following proposition:

Proposition 3.5.5.3 *If* $\tilde{\mathbf{X}} \sim \mathcal{CMEC}_{NL}\left[\tilde{\mathbf{M}}, \tilde{\boldsymbol{\Sigma}}^* \otimes \tilde{\mathbf{R}}, g(\cdot)\right]$ *and* $\tilde{\mathbf{A}}_{M\times N}$ *and* $\tilde{\mathbf{B}}_{L\times K}$ *are full rank matrices, then the matrix* $\tilde{\mathbf{Y}} = \tilde{\mathbf{A}}\tilde{\mathbf{X}}\tilde{\mathbf{B}}^* + \tilde{\mathbf{C}}$ *is distributed as*

$$\mathcal{CMEC}_{MK}\left[\tilde{\mathbf{A}}\tilde{\mathbf{M}}\tilde{\mathbf{B}}^* + \tilde{\mathbf{C}}, \tilde{\mathbf{B}}^H\tilde{\boldsymbol{\Sigma}}^*\tilde{\mathbf{B}} \otimes \tilde{\mathbf{A}}\tilde{\mathbf{R}}\tilde{\mathbf{A}}^H, g_{MK:NL}(\cdot)\right]. \tag{3.113}$$

This proposition follows from (i) last of three property following Proposition 3.5.5.2, (ii) eq(3.111), and (iii) Proposition 3.5.4.3. Many other properties of complex MECs can be inferred from the stochastic representation eq(3.111) based on the properties we know for vector complex ECs. Therefore, we omit stating them explicitly, but will use them in subsequent analysis.

The convenience of considering complex MEC models for the data matrix is that all the theory previously developed for the complex vector ECs applies to these matrices by Proposition 3.5.5.2. For example, note by Proposition 3.5.5.2 that

$$E\{\tilde{\mathbf{X}}\} = \tilde{\mathbf{M}} \qquad \text{and} \qquad \text{cov}(\tilde{\mathbf{X}}) = \tilde{\mathbf{\Sigma}}^* \otimes \kappa \cdot \tilde{\mathbf{R}} \tag{3.114}$$

where $\kappa = E\{\tilde{\rho}\}/(NL)$ and the pdf for $\tilde{\rho}$ is $P_{\tilde{\rho}} = \pi^{NL}\ \tilde{\rho}^{(NL-1)}\ g(\tilde{\rho})/(NL\ -1)!$. In addition we find that these models provide a natural framework for the consideration of unknown parameters like the data covariance.

The remaining two sections of this chapter consider (i) a generalization of the complex Wishart partition Theorem (WPT) to complex MECs, and (ii) the density representation of a singular complex EC.

3.6 New Generalized Complex Wishart Theorems

Estimation of unknown data covariances play a central role in much of adaptive array processing. Since the convergence rate of an adaptive algorithm is important, the introduction of the rapidly converging SMI technique by Reed *et al* [16] was well received . Indeed, several adaptive algorithms are functions of the SCM. Exact statistical distributions of quantities related to the SCM are presented in this section. Specifically, we consider statistics formed from the zero mean complex MEC data matrix $\tilde{\mathbf{X}}$ partitioned and distributed as

$$\left[\tilde{\mathbf{X}}\right]_{N\times L} = \left[\begin{array}{c} (\tilde{\mathbf{X}}_1)_{M\times L} \\ (\tilde{\mathbf{X}}_2)_{(N-M)\times L} \end{array}\right] \sim \mathcal{CMEC}_{NL}\left[\tilde{\mathbf{M}} = \mathbf{0}, \mathbf{I}_L \otimes \tilde{\mathbf{R}},\ g(\cdot)\ \right] = P_{\tilde{\mathbf{X}}_1,\tilde{\mathbf{X}}_2} \tag{3.115}$$

where it is assumed that $N \geq 2M$. The main and new result appears in Theorem 3.6.1. The lengthy proof is given in Appendix B.

Proposition 3.6.0.4 Wishart Expectation Theorem (WET): *The density of the sample* *covariance matrix (SCM) formed from this data matrix is given by* $(L \geq N)$

$$\tilde{\mathbf{D}} \triangleq \tilde{\mathbf{X}}\tilde{\mathbf{X}}^H \sim \frac{\pi^{NL}}{\tilde{\Gamma}_N(L)} |\tilde{\mathbf{R}}|^{-L} |\tilde{\mathbf{D}}|^{L-N} \, g[\mathrm{tr}\tilde{\mathbf{R}}^{-1}\tilde{\mathbf{D}}]. \tag{3.116}$$

We indicate this generalized complex Wishart distribution by $\tilde{\mathbf{D}} \sim \mathcal{GCW}_N\left[L, \tilde{\mathbf{R}}, g(\cdot)\right]$.

Considering the expectation

$$E\{q(\tilde{\mathbf{X}}\tilde{\mathbf{X}}^H)\} = \int_{\tilde{\mathbf{X}} \in \mathbb{C}^{NL}} q(\tilde{\mathbf{X}}\tilde{\mathbf{X}}^H)|\tilde{\mathbf{R}}|^{-L} g[\mathrm{tr}\tilde{\mathbf{R}}^{-1}\tilde{\mathbf{X}}\tilde{\mathbf{X}}^H] d\tilde{\mathbf{X}} \tag{3.117}$$

$$= \frac{\pi^{NL}}{\tilde{\Gamma}_N(L)} \cdot \int_{\tilde{\mathbf{D}}=\tilde{\mathbf{D}}^H>0} q(\tilde{\mathbf{D}}) \, |\tilde{\mathbf{D}}|^{L-N} |\tilde{\mathbf{R}}|^{-L} \, g[\mathrm{tr}\tilde{\mathbf{R}}^{-1}\tilde{\mathbf{D}}] \, d\tilde{\mathbf{D}} = E\{q(\tilde{\mathbf{D}})\}$$

completes proof. When the data matrix is complex Gaussian, *i.e.* when $g(a) = \pi^{-NL} e^{-a}$ then the density for $\tilde{\mathbf{D}}$ is the well known complex Wishart.

Theorem 3.6.1 Generalized Wishart Partition Theorem (GWPT): *Let the SCM be partitioned such that*

$$\tilde{\mathbf{D}} = \begin{bmatrix} \tilde{\mathbf{D}}_{11} & \tilde{\mathbf{D}}_{12} \\ \tilde{\mathbf{D}}_{21} & \tilde{\mathbf{D}}_{22} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{X}}_1\tilde{\mathbf{X}}_1^H & \tilde{\mathbf{X}}_1\tilde{\mathbf{X}}_2^H \\ \tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_1^H & \tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H \end{bmatrix} = \tilde{\mathbf{X}}\tilde{\mathbf{X}}^H \tag{3.118}$$

and define the Schur complements of dimension $M \times M$

$$\begin{aligned} \tilde{\mathbf{D}}_{11.2} &\triangleq \tilde{\mathbf{D}}_{11} - \tilde{\mathbf{D}}_{12}\tilde{\mathbf{D}}_{22}^{-1}\tilde{\mathbf{D}}_{21} \\ &= \tilde{\mathbf{X}}_1 \mathfrak{P}(\tilde{\mathbf{X}}_2^H|\tilde{\mathbf{I}}_L)\tilde{\mathbf{X}}_1^H \end{aligned} \qquad \tilde{\mathbf{R}}_{11.2} \triangleq \tilde{\mathbf{R}}_{11} - \tilde{\mathbf{R}}_{12}\tilde{\mathbf{R}}_{22}^{-1}\tilde{\mathbf{R}}_{21} \tag{3.119}$$

and the regression coefficients of dimension $(N-M) \times M$

$$\tilde{\mathcal{T}} \triangleq \tilde{\mathbf{D}}_{22}^{-1}\tilde{\mathbf{D}}_{21} \qquad \tilde{\mathbf{T}} \triangleq \tilde{\mathbf{R}}_{22}^{-1}\tilde{\mathbf{R}}_{21}. \tag{3.120}$$

(i) The partition quantities have the following marginal densities

$$\begin{aligned} \tilde{\mathbf{X}}_1 &\sim |\tilde{\mathbf{R}}_{11}|^{-L} g_{ML:NL}\left[\mathrm{tr}\tilde{\mathbf{R}}_{11}^{-1}\tilde{\mathbf{X}}_1\tilde{\mathbf{X}}_1^H\right] & \tilde{\mathbf{D}}_{11} &\sim \mathcal{GCW}_M\left[L, \tilde{\mathbf{R}}_{11}, g_{ML:NL}(\cdot)\right] \\ \tilde{\mathbf{X}}_2 &\sim |\tilde{\mathbf{R}}_{22}|^{-L} g_{(N-M)L:NL}\left[\mathrm{tr}\tilde{\mathbf{R}}_{22}^{-1}\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H\right] & \tilde{\mathbf{D}}_{22} &\sim \mathcal{GCW}_{(N-M)}\left[L, \tilde{\mathbf{R}}_{22}, g_{(N-M)L:NL}(\cdot)\right]. \end{aligned} \tag{3.121}$$

(ii) The matrices $\tilde{\mathbf{D}}_{11.2}$ *and* $\tilde{\mathcal{T}}$ *are independent and have the following marginals*

$$\frac{\pi^{M(L-N+M)}}{\tilde{\Gamma}_M(L-N+M)} |\tilde{\mathbf{R}}_{11.2}|^{-(L-N+M)} |\tilde{\mathbf{D}}_{11.2}|^{L-N} g_{M(L-N+M):NL}\left[\mathrm{tr}\tilde{\mathbf{D}}_{11.2}\tilde{\mathbf{R}}_{11.2}^{-1}\right] \tag{3.122}$$

$$\left[\frac{\tilde{\Gamma}_{(N-M)}(L+M)}{\tilde{\Gamma}_{(N-M)}(L)\pi^{M(N-M)}}\right] |\tilde{\mathbf{R}}_{22}|^{-L} |\tilde{\mathbf{R}}_{11.2}|^{-(N-M)} \left|(\tilde{\mathcal{T}} - \tilde{\mathbf{T}})\tilde{\mathbf{R}}_{11.2}^{-1}(\tilde{\mathcal{T}} - \tilde{\mathbf{T}})^H + \tilde{\mathbf{R}}_{22}^{-1}\right|^{-(L+M)} \tag{3.123}$$

(iii) The joint density of $(\tilde{\mathbf{D}}_{11.2}, \tilde{\mathcal{T}}, \tilde{\mathbf{D}}_{22})$ *is given by*

$$\left[\frac{\pi^{M(L-M+N)+(M-N)L}}{\tilde{\Gamma}_M(L-M+N)\tilde{\Gamma}_{(N-M)}(L)}\right] |\tilde{\mathbf{D}}_{11.2}|^{(L-N)} |\tilde{\mathbf{D}}_{22}|^{(L-N+2M)} |\tilde{\mathbf{R}}_{11.2}|^{-L} |\tilde{\mathbf{R}}_{22}|^{-L} \tag{3.124}$$

$$\times g\left[\mathrm{tr}\tilde{\mathbf{R}}_{11.2}^{-1}(\tilde{\mathcal{T}}-\tilde{\mathbf{T}})^H\tilde{\mathbf{D}}_{22}(\tilde{\mathcal{T}}-\tilde{\mathbf{T}})+\mathrm{tr}\tilde{\mathbf{R}}_{11.2}^{-1}\tilde{\mathbf{D}}_{11.2}+\mathrm{tr}\tilde{\mathbf{R}}_{22}^{-1}\tilde{\mathbf{D}}_{22}\right]$$

(iv) The joint density of $(\tilde{\mathbf{D}}_{11.2}, \tilde{\mathbf{D}}_{22})$ *is given by*

$$\left[\frac{\pi^{L(N-M)+M(L-N+M)}}{\tilde{\Gamma}_{(N-M)}(L)\tilde{\Gamma}_M(L-N+M)}\right]|\tilde{\mathbf{R}}_{22}|^{-L}|\tilde{\mathbf{R}}_{11.2}|^{-(L-N+M)}|\tilde{\mathbf{D}}_{22}|^{(L-N+M)}|\tilde{\mathbf{D}}_{11.2}|^{L-N} \quad (3.125)$$
$$\times\, g_{NL-M(N-M):NL}\left[\mathrm{tr}\tilde{\mathbf{R}}_{11.2}^{-1}\tilde{\mathbf{D}}_{11.2}+\mathrm{tr}\tilde{\mathbf{R}}_{22}^{-1}\tilde{\mathbf{D}}_{22}\right].$$

(v) The joint density of $(\tilde{\mathcal{T}}, \tilde{\mathbf{D}}_{22})$ *is given by*

$$\left[\frac{\pi^{L(N-M)}}{\tilde{\Gamma}_{(N-M)}(L)}\right]|\tilde{\mathbf{R}}_{22}|^{-L}|\tilde{\mathbf{R}}_{11.2}|^{-(N-M)}|\tilde{\mathbf{D}}_{22}|^{(L-N+2M)} \quad (3.126)$$
$$\times\, g_{NL-M(L-N+M):NL}\left[\mathrm{tr}\tilde{\mathbf{R}}_{11.2}^{-1}(\tilde{\mathcal{T}}-\tilde{\mathbf{T}})^H\tilde{\mathbf{D}}_{22}(\tilde{\mathcal{T}}-\tilde{\mathbf{T}})+\mathrm{tr}\tilde{\mathbf{R}}_{22}^{-1}\tilde{\mathbf{D}}_{22}\right].$$

This appears to be the first generalization of the WPT.

3.7 Singular Complex EC Distributions

The covariances of most complex random quantities considered in sensor array processing are hermitian positive definite and therefore nonsingular. In subsequent chapters, however, we consider the distributions of random spatial filters. As discussed in Chapter 2, spatial filter designs typically incorporate constraints in an attempt to guarantee desired responses in certain look directions. These constraints ultimately result in random spatial filter weights which have singular covariances, and necessarily have singular distributions. Conceptually, singular distributions arise when there is something partially deterministic about a random quantity. For example, a normally distributed random vector with a singular covariance matrix yields a singular distribution (see [63] or [51] pp. 41–43). There is a known subspace to which all these normal random vectors are orthogonal; namely, the nullspace of the singular covariance. In this section we consider for seemingly the first time the concept of *singular complex EC distributions*. In particular, it is of interest to specify the form of the density of such quantities.

Consider a random vector $\tilde{\mathbf{x}} \in \mathbb{C}^N$ distributed as $\mathcal{CEC}_N[\tilde{\mathbf{m}}, \tilde{\mathbf{R}}, \phi(\cdot)]$ where $\tilde{\mathbf{R}} \geq 0$ with rank $M \leq N$. By definition the c.f. of $\tilde{\mathbf{x}}$ is given by

$$\psi_{\tilde{\mathbf{x}}}(\tilde{\mathbf{t}}) = e^{j\text{Re}(\tilde{\mathbf{t}}^H\tilde{\mathbf{m}})} \cdot \phi(\tilde{\mathbf{t}}^H\tilde{\mathbf{R}}\tilde{\mathbf{t}}/2), \tag{3.127}$$

and when $M < N$ its density does not exist, *i.e.* it is degenerate. Note that $\tilde{\mathbf{R}}$ has an eigen decomposition such that

$$\tilde{\mathbf{R}} = \tilde{\mathbf{Q}}^H \begin{bmatrix} \tilde{\mathbf{\Lambda}}_{M\times M} & \mathbf{0} \\ \mathbf{0} & \mathbf{0}_{(N-M)\times(N-M)} \end{bmatrix} \tilde{\mathbf{Q}}. \tag{3.128}$$

Let $\tilde{\mathbf{v}} = \tilde{\mathbf{Q}}\tilde{\mathbf{t}}$, $\tilde{\boldsymbol{\mu}} = \tilde{\mathbf{Q}}\tilde{\mathbf{m}}$ and make the partitions $\tilde{\mathbf{v}} = [\tilde{\mathbf{v}}_1; \tilde{\mathbf{v}}_2]$ and $\tilde{\boldsymbol{\mu}} = [\tilde{\boldsymbol{\mu}}_1; \tilde{\boldsymbol{\mu}}_2]$. Note that

$$\tilde{\mathbf{t}}^H\tilde{\mathbf{m}} = \tilde{\mathbf{t}}^H\tilde{\mathbf{Q}}^H\tilde{\mathbf{Q}}\tilde{\mathbf{m}} = \tilde{\mathbf{v}}_1^H\tilde{\boldsymbol{\mu}}_1 + \tilde{\mathbf{v}}_2^H\tilde{\boldsymbol{\mu}}_2 \qquad \text{and} \qquad \tilde{\mathbf{t}}^H\tilde{\mathbf{R}}\tilde{\mathbf{t}} = \tilde{\mathbf{v}}_1^H\tilde{\mathbf{\Lambda}}\tilde{\mathbf{v}}_1. \tag{3.129}$$

Consequently, if we let $\tilde{\mathbf{y}} = \tilde{\mathbf{Q}}\tilde{\mathbf{x}}$ then

$$E\{e^{j\text{Re}(\tilde{\mathbf{t}}^H\tilde{\mathbf{x}})}\} = E\{e^{j\text{Re}(\tilde{\mathbf{v}}^H\tilde{\mathbf{y}})}\} = \psi_{\tilde{\mathbf{y}}}(\tilde{\mathbf{v}}) = e^{j\text{Re}(\tilde{\mathbf{v}}_2^H\tilde{\boldsymbol{\mu}}_2)} \cdot e^{j\text{Re}(\tilde{\mathbf{v}}_1^H\tilde{\boldsymbol{\mu}}_1)} \cdot \phi(\tilde{\mathbf{v}}_1^H\tilde{\mathbf{\Lambda}}\tilde{\mathbf{v}}_1/2) \tag{3.130}$$

$$= \psi_{\tilde{\mathbf{y}}_2}(\tilde{\mathbf{v}}_2) \cdot \psi_{\tilde{\mathbf{y}}_1}(\tilde{\mathbf{v}}_1). \tag{3.131}$$

Clearly, $\tilde{\mathbf{y}}_1$ and $\tilde{\mathbf{y}}_2$ are independent. The density for $\tilde{\mathbf{y}}_2$ is

$$\delta\left([\mathbf{y}_{R2}; \mathbf{y}_{I2}] - [\boldsymbol{\mu}_{R2}; \boldsymbol{\mu}_{I2}]\right) \triangleq \tilde{\delta}(\tilde{\mathbf{y}}_2 - \tilde{\boldsymbol{\mu}}_2) \qquad \text{that is} \qquad Pr(\tilde{\mathbf{y}}_2 = \tilde{\boldsymbol{\mu}}_2) = 1 \tag{3.132}$$

where $\delta(\mathbf{a}_{K\times 1}) \triangleq \delta(a_1) \cdot \delta(a_2) \cdots \delta(a_K)$ and $\delta(\cdot)$ is the Dirac impulse function. The density for $\tilde{\mathbf{y}}_1$ is clearly complex EC such that when integrated over $\mathfrak{C}^M$ one obtains unity. We can denote this by saying $\tilde{\mathbf{y}}_1 \sim \mathcal{CEC}_M[\tilde{\boldsymbol{\mu}}_1, \tilde{\mathbf{\Lambda}}, g_{M:M}(\cdot)]$. The subscript $M : M$ on $g(\cdot)$ emphasizes that the density is well defined as a function over the domain $\mathfrak{C}^M$.

Following Miller/Khatri's lead [62, 50, 70], we can obtain a closed formed expression for the density of $\tilde{\mathbf{x}}$ via the use of the *pseudo-inverse*. Let $\tilde{\mathbf{Q}}^H = [\tilde{\mathbf{Q}}_1|\tilde{\mathbf{Q}}_2]$. The pseudo-inverse of $\tilde{\mathbf{R}}$ is defined and denoted by

$$\tilde{\mathbf{R}}^+ = \tilde{\mathbf{Q}}^H \begin{bmatrix} \tilde{\mathbf{\Lambda}}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tilde{\mathbf{Q}} = \tilde{\mathbf{Q}}_1\tilde{\mathbf{\Lambda}}^{-1}\tilde{\mathbf{Q}}_1^H. \tag{3.133}$$

Note that

$$\begin{aligned}
(\tilde{\mathbf{y}}_1 - \tilde{\boldsymbol{\mu}}_1)^H \tilde{\mathbf{\Lambda}}^{-1} (\tilde{\mathbf{y}}_1 - \tilde{\boldsymbol{\mu}}_1) &= (\tilde{\mathbf{y}} - \tilde{\boldsymbol{\mu}})^H \begin{bmatrix} \mathbf{I}_M \\ \mathbf{0}_{(N-M)\times M} \end{bmatrix} \tilde{\mathbf{\Lambda}}^{-1} [\mathbf{I}_M | \mathbf{0}_{M\times(N-M)}] (\tilde{\mathbf{y}} - \tilde{\boldsymbol{\mu}}) \\
&= (\tilde{\mathbf{y}} - \tilde{\boldsymbol{\mu}})^H \begin{bmatrix} \tilde{\mathbf{Q}}_1^H \\ \tilde{\mathbf{Q}}_2^H \end{bmatrix} \tilde{\mathbf{Q}}_1 \tilde{\mathbf{\Lambda}}^{-1} \tilde{\mathbf{Q}}_1^H [\tilde{\mathbf{Q}}_1 | \tilde{\mathbf{Q}}_2] (\tilde{\mathbf{y}} - \tilde{\boldsymbol{\mu}}) \\
&= (\tilde{\mathbf{Q}}^H \tilde{\mathbf{y}} - \tilde{\mathbf{Q}}^H \tilde{\boldsymbol{\mu}})^H \tilde{\mathbf{Q}}^H \begin{bmatrix} \tilde{\mathbf{\Lambda}}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \tilde{\mathbf{Q}} (\tilde{\mathbf{Q}}^H \tilde{\mathbf{y}} - \tilde{\mathbf{Q}}^H \tilde{\boldsymbol{\mu}}) \\
&= (\tilde{\mathbf{x}} - \tilde{\mathbf{m}})^H \tilde{\mathbf{R}}^+ (\tilde{\mathbf{x}} - \tilde{\mathbf{m}}).
\end{aligned} \tag{3.134}$$

Thus, we have the following representation for the pdf of a singular complex EC random vector $\tilde{\mathbf{x}} \in \mathcal{C}^N$:

$$|\tilde{\mathbf{\Lambda}}|^{-1} g_{M:M} \left[(\tilde{\mathbf{y}}_1 - \tilde{\boldsymbol{\mu}}_1)^H \tilde{\mathbf{\Lambda}}^{-1} (\tilde{\mathbf{y}}_1 - \tilde{\boldsymbol{\mu}}_1) \right] \times \tilde{\delta}(\tilde{\mathbf{y}}_2 - \tilde{\boldsymbol{\mu}}_2) = \tag{3.135}$$

$$\left[\prod_{i=1}^{M} \lambda_i(\tilde{\mathbf{R}}) \right]^{-1} g_{M:M} \left[(\tilde{\mathbf{x}} - \tilde{\mathbf{m}})^H \tilde{\mathbf{R}}^+ (\tilde{\mathbf{x}} - \tilde{\mathbf{m}}) \right] \times \tilde{\delta}[\tilde{\mathbf{Q}}_2^H (\tilde{\mathbf{x}} - \tilde{\mathbf{m}})]$$

which we denote by $\quad \tilde{\mathbf{x}} \sim \mathcal{SCEC}_N \left[\tilde{\mathbf{m}}, \tilde{\mathbf{R}}, g_{M:M}(\cdot) \right].$

We can use this representation to write the pdf of singular complex EC distributed quantities. It is good to remember that one can always introduce a linear transformation producing a new variate whose density does exist, *e.g* the transformation $\tilde{\mathbf{Q}}_1^H \tilde{\mathbf{x}}$ yields the r.vc. $\tilde{\mathbf{y}}_1$.

As an example of how these singular distributions can arise, let $\tilde{\mathbf{x}} \sim \mathcal{CEC}_N[\tilde{\mathbf{m}}, \tilde{\mathbf{R}}, \phi(\cdot); g(\cdot)]$, thus $\tilde{\mathbf{R}} > 0$. Let $\tilde{\mathbf{A}}$ be $J \times N$ with $J > N$. Consider the linear transformation

$$\tilde{\mathbf{y}} = \tilde{\mathbf{A}} \tilde{\mathbf{x}} \stackrel{d}{=} \tilde{\mathbf{A}} \tilde{\mathbf{m}} + \tilde{\mathbf{A}} \tilde{\mathbf{R}}^{1/2} \tilde{\mathbf{u}}^{(N)} \sqrt{\tilde{\rho}}. \tag{3.136}$$

Clearly, $\psi_{\tilde{\mathbf{y}}}(\tilde{\mathbf{t}}_y) = \psi_{\tilde{\mathbf{x}}}(\tilde{\mathbf{t}}_x = \tilde{\mathbf{A}}^H \tilde{\mathbf{t}}_y) = e^{j\mathrm{Re}(\tilde{\mathbf{t}}^H \tilde{\mathbf{m}})} \cdot \phi(\tilde{\mathbf{t}}_y^H \tilde{\mathbf{A}} \tilde{\mathbf{R}} \tilde{\mathbf{A}}^H \tilde{\mathbf{t}}_y)$ with matrix $\tilde{\mathbf{A}} \tilde{\mathbf{R}} \tilde{\mathbf{A}}^H$ singular by construction. Hence,

$$\tilde{\mathbf{y}} \sim \mathcal{SCEC}_J \left[\tilde{\mathbf{A}} \tilde{\mathbf{m}}, \tilde{\mathbf{A}} \tilde{\mathbf{R}} \tilde{\mathbf{A}}^H, g_{N:N}(\cdot) \right]. \tag{3.137}$$

This completes the overview of complex MECs, and their properties. Attention will be given next to adaptive array signal detection in complex EC environments. As a reminder, the use of tilde to differentiate complex quantities from real ones is exclusive to this chapter. Quantities in all other chapters are complex unless implied or stated otherwise.

3.8 Assumed Data Model for Thesis

Concerning the problems of adaptive array detection and adaptive beamforming as discussed in Chapter 2, the totality of the data is given by the $N \times (L+1)$ data matrix $\mathbf{X}_0 = [\mathbf{X}|\mathbf{x}]$. The first L columns contain the secondary data set, and the last contains the primary data vector. Recall that we assume a sample support $L \geq N$. The data model type we shall assume throughout the remainder of this thesis shall be complex MEC with a density of the form

$$\mathbf{X}_0 \sim |\mathbf{R}|^{-(L+1)} g[\,\mathrm{tr}\mathbf{R}^{-1}(\mathbf{X}_0 - \mathbf{M})(\mathbf{X}_0 - \mathbf{M})^H\,] \tag{3.138}$$

where the mean is of the form $\mathbf{M} = [\mathbf{0}_{N\times L}|\mathbf{m}]$. At times we may consider alternate models, but only when stated otherwise. The model eq(3.138) leads to the following complex MEC marginals

$$\mathbf{x} \sim |\mathbf{R}|^{-1} g_{N:N(L+1)}[(\mathbf{x}-\mathbf{m})^H\mathbf{R}^{-1}(\mathbf{x}-\mathbf{m})] \quad \mathbf{X} \sim |\mathbf{R}|^{-L} g_{NL:N(L+1)}[\,\mathrm{tr}\mathbf{R}^{-1}\mathbf{X}\mathbf{X}^H]. \tag{3.139}$$

By the nature of the models the snapshots are uncorrelated, yet dependent except for the the complex Gaussian case in which uncorrelation implies independence. Note that (i) the primary data vector is marginally complex EC, and (ii) the secondary data matrix is zero mean complex LEC.

Chapter 4

Complex MEC Based Adaptive Array Detection and Performance Analysis

4.1 Introduction

In this chapter we consider in detail the complex MEC based GLRT and AMF approaches to the problem of adaptive array detection (AAD). Recall that a precise statement of the AAD problem appears in Section 2.1.3, where a summary of Gaussian based results likewise appears. The present objective is to relax the traditional constraint of data Gaussianity somewhat by assuming that the data is complex MEC distributed. Subsequently, we (i) reformulate the problem of AAD under this weaker distributional assumption and seek the detection structures implied by the GLRT and AMF approaches, and (ii) consider performance issues such as CFAR properties, the PFA, and the PD.

4.2 The GLRT Approach

Under both hypotheses we assume that $L \geq N$ and that the data matrix $\mathbf{X}_0$ is MEC distributed with densities given by

$$g_{H_i} = |\mathbf{R}|^{-(L+1)} g^{(i)}[\ \mathrm{tr}\mathbf{R}^{-1}(\mathbf{X}_0 - \mathbf{M}_i)(\mathbf{X}_0 - \mathbf{M}_i)^H\ ], \quad i = 0, 1 \tag{4.1}$$

where under the corresponding hypotheses we have

$$H_0 : \mathbf{M}_0 = \mathbf{0}_{N\times(L+1)}, \quad H_1 : \mathbf{M}_1 = [\mathbf{0}_{N\times L}|\ \mathbf{Gs}\ ]. \tag{4.2}$$

The following assumptions about the functions $g^{(i)}(r)$, $i = 0, 1$ are made:

- Assume $g^{(i)}(r)$, $i = 0, 1$ on $r \geq 0$ (1) are continuous and (2) at least have tails which decrease monotonically.

- These assumptions guarantee that there exists $r_{MAX}^{(i)} > 0$ such that $h^{(i)}(r) = r^{N(L+1)} g^{(i)}(N \cdot r)$ attains a maximum at $r_{MAX}^{(i)}$, $i = 0, 1$ [55, 57]; call them $h_{MAX}^{(i)} = h^{(i)}(r_{MAX})$.

(Recall that this data model includes the complex Gaussian considered by Kelly [22, 19] when we choose $g^{(i)}(r) = e^{-r}\pi^{-N(L+1)}$, $i = 0, 1$. In addition this model allows for multimodal behavior since it is only required that the tails decrease monotonically. Thus, it is more general than the complex Gaussian mixture representation commonly used in statistical signal and array processing [1]–[7], which necessarily requires strict monotonicity [68].) Following the GLRT approach we maximize the likelihood functions over the unknown variables $\mathbf{R}$ and s to obtain the resulting decision statistic. These unknown parameters are hence replaced by their maximum-likelihood (ML) estimates:

$$\frac{\max_{\mathbf{s},\mathbf{R}} g_{H_1}(\cdot)}{\max_{\mathbf{R}} g_{H_0}(\cdot)} \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \tilde{\eta}. \tag{4.3}$$

Let $\mathbf{D}_i = (\mathbf{X}_0 - \mathbf{M}_i)(\mathbf{X}_0 - \mathbf{M}_i)^H$, $\mathbf{A}_i = \mathbf{R}^{-1/2}\mathbf{D}_i\mathbf{R}^{-1/2}$, and consider the following maximization problem

$$\begin{aligned} \max_{\mathbf{R}} \mathcal{L}^{(i)} &= \max_{\mathbf{R}} |\mathbf{R}|^{-(L+1)} g^{(i)}[\, \mathrm{tr}\mathbf{R}^{-1}(\mathbf{X}_0 - \mathbf{M}_i)(\mathbf{X}_0 - \mathbf{M}_i)^H \,] \\ &= \max_{\mathbf{A}_i} |\mathbf{D}_i|^{-(L+1)} |\mathbf{A}_i|^{(L+1)} g^{(i)}[\, \mathrm{tr}\mathbf{A}_i \,]. \end{aligned} \tag{4.4}$$

First we maximize over parameter $\mathbf{R}$. The second equality in eq(4.4) follows from the modulo property of the trace operator, and the fact that there exists a one-to-one mapping between $\mathbf{R}$ and $\mathbf{A}_i$. Thus, maximizing over either will yield the same result. Note the following upper bound

$$\begin{aligned} \mathcal{L}^{(i)} &= |\mathbf{D}_i|^{-(L+1)} \left[\prod_{k=1}^{N} \lambda_k^{\frac{1}{N}}(\mathbf{A}_i) \right]^{N(L+1)} g^{(i)} \left[\sum_{k=1}^{N} \lambda_k(\mathbf{A}_i) \right] \\ &\leq |\mathbf{D}_i|^{-(L+1)} \left[\frac{1}{N} \sum_{k=1}^{N} \lambda_k(\mathbf{A}_i) \right]^{N(L+1)} g^{(i)} \left[\sum_{k=1}^{N} \lambda_k(\mathbf{A}_i) \right]. \end{aligned} \tag{4.5}$$

Equality between the geometric and arithmetic mean is obtained if and only if $\lambda_k(\mathbf{A}_i) = \lambda_0(i)$ for all k. This implies that the upper bound on the likelihood function $\mathcal{L}^{(i)}$ is obtained only by the choice $\mathbf{A}_i = \lambda_0(i) \cdot \mathbf{I} \Longrightarrow \mathbf{R} = \mathbf{D}_i/\lambda_0(i)$, which leads to the equivalence

$$\max_{\mathbf{R}} \mathcal{L}^{(i)} = |\mathbf{D}_i|^{-(L+1)} \max_{\lambda_0(i)} [\lambda_0(i)]^{N(L+1)} g^{(i)}[N\lambda_0(i)] = |\mathbf{D}_i|^{-(L+1)} \cdot h^{(i)}_{MAX} \ . \tag{4.6}$$

Note that because $\mathbf{A}_i$ is hermitian positive definite (hpd) it follows that $\lambda_0(i)$ is real and positive. The second equality follows from the previously stated assumptions about the functions $g^{(i)}(\cdot)$, $i = 0, 1$. The maximum over $\mathbf{R}$ is therefore obtained with $\widehat{\mathbf{R}} = \mathbf{D}_i/r^{(i)}_{MAX}$. If $\mathbf{X}_0$ is complex normal, *i.e.* if $g^{(i)}(r) \propto e^{-r}$, then it is easy to show that $r^{(i)}_{MAX} = (L+1)$, $i = 0, 1$ yielding the well-known SCM as the ML estimate of $\mathbf{R}$. It is interesting to note that subsequent to this point in the analysis we now overlaps with the work of Kelly [22, 19]. In particular when Kelly maximized over $\mathbf{R}$, the resulting likelihood functions were likewise constants times $|\mathbf{D}_i|^{-(L+1)}$.

Under the null hypothesis, $\mathbf{R}$ is the only parameter over which g_{H_0} need be maximized. Under hypothesis H_1, further maximization of g_{H_1} over the unknown signal parameters s is necessary.

To maximize $\mathcal{L}^{(1)}[\mathbf{s}, \mathbf{D}_1/r^{(1)}_{MAX}]$ over s note that $\mathbf{D}_1 = \mathbf{X}\mathbf{X}^H + (\mathbf{x} - \mathbf{G}\mathbf{s})(\mathbf{x} - \mathbf{G}\mathbf{s})^H$ and recall that $\mathbf{D} = \mathbf{X}\mathbf{X}^H$. Maximization of the likelihood function is obtained by minimizing over s the determinant

$$\begin{aligned} |\mathbf{D}_1| &= |\mathbf{D}^{1/2}[\mathbf{I} + \mathbf{D}^{-1/2}(\mathbf{x} - \mathbf{G}\mathbf{s})(\mathbf{x} - \mathbf{G}\mathbf{s})^H\mathbf{D}^{-1/2}]\mathbf{D}^{1/2}| \\ &= |\mathbf{D}| \cdot [1 + (\mathbf{x} - \mathbf{G}\mathbf{s})^H\mathbf{D}^{-1}(\mathbf{x} - \mathbf{G}\mathbf{s})]. \end{aligned} \tag{4.7}$$

Let $\widehat{\mathbf{W}}_{ML} = \mathbf{D}^{-1}\mathbf{G}(\mathbf{G}^H\mathbf{D}^{-1}\mathbf{G})^{-1}$. Completing the square of the quadratic form appearing in the determinant $|\mathbf{D}_1|$ yields

$$(\mathbf{x} - \mathbf{G}\mathbf{s})^H\mathbf{D}^{-1}(\mathbf{x} - \mathbf{G}\mathbf{s}) = (\mathbf{s} - \widehat{\mathbf{W}}^H_{ML}\mathbf{x})^H(\mathbf{G}^H\mathbf{D}^{-1}\mathbf{G})(\mathbf{s} - \widehat{\mathbf{W}}^H_{ML}\mathbf{x}) + \mathbf{x}^H\mathfrak{P}(\mathbf{G}|\mathbf{D}^{-1})\mathbf{x}. \tag{4.8}$$

Since $\mathbf{G}^H\mathbf{D}^{-1}\mathbf{G}$ is hpd, the ML signal estimate for this class of distributions is shown to be

$$\arg\max_{\mathbf{s}} \mathcal{L}^{(1)}[\mathbf{s}, \mathbf{D}_1/r_{MAX}] = \arg\min_{\mathbf{s}} |\mathbf{D}_1| = (\mathbf{G}^H\mathbf{D}^{-1}\mathbf{G})^{-1}\mathbf{G}^H\mathbf{D}^{-1}\mathbf{x} = \widehat{\mathbf{s}}_{ML}. \tag{4.9}$$

Taking the (L+1) root of resulting decision statistic leads to the following form

$$
\boxed{t_{GLRT}(\mathbf{X}_0) = \left[\frac{\max_{\mathbf{s},\mathbf{R}} g_{H_1}(\cdot)}{\max_{\mathbf{R}} g_{H_0}(\cdot)}\right]^{\frac{1}{(L+1)}} = \frac{1+\mathbf{x}^H\mathbf{D}^{-1}\mathbf{x}}{1+\mathbf{x}^H\mathfrak{P}(\mathbf{G}|\mathbf{D}^{-1})\mathbf{x}} \underset{H_0}{\overset{H_1}{\gtrless}} \left[\frac{h_{MAX}^{(0)}}{h_{MAX}^{(1)}}\cdot\tilde{\eta}\right]^{\frac{1}{(L+1)}} \triangleq \eta} \quad (4.10)
$$

which is exactly the mathematical form of Kelly's GLRT statistic [22, 19]. Thus, we obtain the unanticipated, but quite remarkable result that the GLRT approach leads to the same test statistic for a large number of non-Gaussian distributions. The key properties contributing to this result include the joint elliptical symmetry of the data matrix, and the mutual snapshot dependence implicated by such symmetry. It is also quite remarkable that the functional forms under each hypothesis, *i.e.* $g^{(i)}(r)$, $i = 0, 1$, can differ, yet the GLRT approach results in the same test.

In the process of deriving the test t_{GLRT}, it was shown that the SCM based LCMV beamformer is likewise optimal in the ML sense for many complex MECs as an estimator of s when both $\mathbf{R}$ but s are unknown. This latter fact has rather significant implications concerning robustness in adaptive beamforming and nulling problems. The optimality of $\hat{\mathbf{s}}_{ML}$ across a class of distributions suggests a plausible non-asymptotic theory for the observed robustness of this adaptive beamformer in non-Gaussian environments. In Chapter 5 more attention is devoted to the signal estimator $\hat{\mathbf{s}}_{ML}$ and adaptive beamforming. Next we consider the AMF approach, and subsequently the PFA and PD under the complex MEC data assumption.

4.3 The AMF Approach

4.3.1 Partial GLRT Approach

Robey *et al* [23] suggest applying the GLRT approach assuming $\mathbf{R}$ is known and s is unknown:

$$
\frac{\max_{\mathbf{s}} g_{H_1}(\cdot)}{g_{H_0}(\cdot)} \underset{H_0}{\overset{H_1}{\gtrless}} \tilde{\eta}. \quad (4.11)
$$

The likelihood functions for the primary data vector are given by

$$g_{H_i} = g^{(i)}_{N:N(L+1)}[(\mathbf{x}-\mathbf{m}_i)^H\mathbf{R}^{-1}(\mathbf{x}-\mathbf{m}_i)] \qquad i=0,1 \qquad \mathbf{m}_0 = \mathbf{0} \qquad \mathbf{m}_1 = \mathbf{G}\mathbf{s}. \tag{4.12}$$

By Proposition 3.5.3.10 note that $g^{(i)}_{N:N(L+1)}(r)$ is guaranteed to be non-increasing, and often strictly decreasing on $r > 0$. Hence, completing the square

$$(\mathbf{x}-\mathbf{G}\mathbf{s})^H\mathbf{R}^{-1}(\mathbf{x}-\mathbf{G}\mathbf{s}) = (\mathbf{s}-\mathbf{W}^H_{ML}\mathbf{x})^H(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})(\mathbf{s}-\mathbf{W}^H_{ML}\mathbf{x}) + \mathbf{x}^H\mathfrak{P}(\mathbf{G}|\mathbf{R}^{-1})\mathbf{x} \tag{4.13}$$

we conclude that the clairvoyant ML estimate of s remains as in eq(2.7) under the complex Gaussian assumption. Replacing **R** with **D**, the AMF approach leads to the test

$$\frac{g^{(1)}_{N:N(L+1)}[\mathbf{x}^H\mathfrak{P}(\mathbf{G}|\mathbf{D}^{-1})\mathbf{x}]}{g^{(0)}_{N:N(L+1)}[\mathbf{x}^H\mathbf{D}^{-1}\mathbf{x}]} \underset{H_0}{\overset{H_1}{\gtrless}} \tilde{\eta}. \tag{4.14}$$

Note also that when $g^{(0)}(r) = g^{(1)}(r) = g(r)$ then the decision test can be written

$$g_{N:N(L+1)}[\mathbf{x}^H\mathfrak{P}(\mathbf{G}|\mathbf{D}^{-1})\mathbf{x}]/\sqrt{\tilde{\eta}} \underset{H_0}{\overset{H_1}{\gtrless}} g_{N:N(L+1)}[\mathbf{x}^H\mathbf{D}^{-1}\mathbf{x}]/\sqrt{\tilde{\eta}}. \tag{4.15}$$

The monotonicity of $g_{N:N(L+1)}(r)$ guarantees that there exists a *threshold dependent* inverse function $h_{\tilde{\eta}}(\cdot)$ such that $h_{\tilde{\eta}}[g_{N:N(L+1)}(r)/\sqrt{\tilde{\eta}}] = r$. When applied to the test, it yields the usual Gaussian based AMF as the decision rule.

Due to the complexity of these AMF detection structures we will not further pursue their analysis. Rather we consider the alternate interpretation of the Gaussian based AMF.

4.3.2 Alternate Interpretation of AMF

Recall from Section 2.1.3 that the complex Gaussian based AMF statistic is likewise obtainable by normalizing the clairvoyant ML estimate of s, substituting **D** for **R**, and subsequently taking its magnitude. In the previous section it was shown that the complex MEC based clairvoyant ML estimate of s is the same as the Gaussian based one. Since the determining factor is the clairvoyant ML signal estimate, it follows that this alternate approach to the

AMF likewise leads to the same AMF structure as the Gaussian:

$$\boxed{t_{AMF}(\mathbf{X}_0) = \mathbf{x}^H\mathbf{D}^{-1}\mathbf{G}(\mathbf{G}^H\mathbf{D}^{-1}\mathbf{G})^{-1}\mathbf{G}^H\mathbf{D}^{-1}\mathbf{x}.} \tag{4.16}$$

4.4 Probabilities of False Alarm and Detection

4.4.1 Complex MEC Based

It is of interest to determine the PFA and PD of both the GLRT and AMF detectors. The PFA can in fact be deduced immediately from Kelly's work and Robey's. Indeed, under the null hypothesis both decision statistics $t_{GLRT}(\mathbf{X}_0)$ and $t_{AMF}(\mathbf{X}_0)$ are scale invariant functions of the data matrix $\mathbf{X}_0$, *e.g.* $t_{GLRT}(c\mathbf{X}_0) = t_{GLRT}(\mathbf{X}_0)$ for all nonzero complex scalars c. By Proposition 3.5.4.5 the distributions of these statistics do not depend on $g^{(0)}(\cdot)$. Hence, if we known their distributions, for example, when $g^{(0)}(\cdot) = q(\cdot)$, then by invariance we likewise known them for all valid $g^{(0)}(\cdot) \neq q(\cdot)$. Choosing $g^{(0)}(r) = \pi^{-N(L+1)}e^{-r}$ leads to the complex Gaussian assumption made by both Kelly and Robey. Thus, the PFA's under the general complex MEC assumption are those given in [19, 23], where both the GLRT and AMF were likewise shown to be CFAR detectors. The CFAR property is likewise maintained under the complex MEC data assumption. These assertions will also be validated from the resulting expressions for the PD, *i.e.* when we set the signal to zero $\mathbf{s} = \mathbf{0}$.

To find the PD we will reduce each test statistic to an equivalent statistic involving random variables of much lower dimension. To this end note that

$$\begin{aligned}\tilde{t}_{GLRT}(\mathbf{X}_0) &= t_{GLRT}(\mathbf{X}_0) - 1 \\ &= \frac{t_{AMF}(\mathbf{X}_0)}{1 + \mathbf{x}^H\mathbf{D}^{-1}\mathbf{x} - t_{AMF}(\mathbf{X}_0)}.\end{aligned} \tag{4.17}$$

Recall from eq(3.111) that we can write $\mathbf{X} \stackrel{d}{=} \mathbf{R}^{1/2}\mathbf{Z}$. Since $\mathbf{Z}$ is CSS MEC and invariant to unitary transformations we note that $\mathbf{X} \stackrel{d}{=} \mathbf{R}^{1/2}\mathbf{Q}^H\mathbf{Z}$, which leads to the following stochastic representation for the inverse noise only SCM

$$\mathbf{D}^{-1} \stackrel{d}{=} \mathbf{R}^{-1/2}\mathbf{Q}^H(\mathbf{Z}\mathbf{Z}^H)^{-1}\mathbf{Q}\mathbf{R}^{-1/2}. \tag{4.18}$$

The unitary $\mathbf{Q}$ can be chosen such that

$$\mathbf{Q}\mathbf{R}^{-1/2}\mathbf{G} = \begin{bmatrix} \Delta_{E\times E} \\ \mathbf{0}_{(N-E)\times E} \end{bmatrix}, \tag{4.19}$$

i.e. to roll all the energy of $\mathbf{R}^{-1/2}\mathbf{G}$ into the first E rows. Next consider and define the transformed primary vector such that

$$\begin{aligned} \mathbf{x}_0 &= \mathbf{Q}\mathbf{R}^{-1/2}\mathbf{x} \\ &= \mathbf{Q}\mathbf{R}^{-1/2}\mathbf{G}\mathbf{s} + \mathbf{Q}\mathbf{R}^{-1/2}\mathbf{n} \\ &= \begin{bmatrix} \Delta\mathbf{s} \\ \mathbf{0}_{(N-E)\times E} \end{bmatrix} + \mathbf{z}_0 \triangleq \begin{bmatrix} (\mathbf{x}_1)_{E\times 1} \\ (\mathbf{x}_2)_{(N-E)\times 1} \end{bmatrix} \end{aligned} \tag{4.20}$$

where $\mathbf{z}_0$ is zero mean CSS. Likewise partition the white secondary data matrix $\mathbf{Z}$ such that

$$\mathbf{Z} = \begin{bmatrix} [\mathbf{Z}_1^H]_{E\times L} \\ [\mathbf{Z}_2^H]_{(N-E)\times L} \end{bmatrix} \tag{4.21}$$

and recall the relations for the inverse of a partitioned matrix to obtain

$$(\mathbf{Z}\mathbf{Z}^H)^{-1} = \begin{bmatrix} \Theta_{E\times E}^{-1} & -\Theta^{-1}\mathcal{T}^H \\ -\mathcal{T}\Theta^{-1} & \Psi_{(N-E)\times(N-E)} \end{bmatrix} \tag{4.22}$$

where

$$\begin{aligned} \mathcal{T} &= (\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{Z}_2^H\mathbf{Z}_1 & \Psi &= \mathcal{T}\Theta^{-1}\mathcal{T}^H + (\mathbf{Z}_2^H\mathbf{Z}_2)^{-1} \\ \Theta &= \mathbf{Z}_1^H\mathfrak{P}(\mathbf{Z}_2|\mathbf{I}_L)\mathbf{Z}_1. \end{aligned} \tag{4.23}$$

Employing eq(4.18) the stochastic representation for $\mathbf{D}^{-1}$, the AMF detector can be written

$$\begin{aligned} t_{AMF}(\mathbf{X}_0) \overset{d}{=} \mathbf{x}^H\mathbf{R}^{-1/2}\mathbf{Q}^H(\mathbf{Z}\mathbf{Z}^H)^{-1}\mathbf{Q}\mathbf{R}^{-1/2}\mathbf{G}\left[\mathbf{G}^H\mathbf{R}^{-1/2}\mathbf{Q}^H(\mathbf{Z}\mathbf{Z}^H)^{-1}\mathbf{Q}\mathbf{R}^{-1/2}\mathbf{G}\right]^{-1} \\ \times\mathbf{G}^H\mathbf{R}^{-1/2}\mathbf{Q}^H(\mathbf{Z}\mathbf{Z}^H)^{-1}\mathbf{Q}\mathbf{R}^{-1/2}\mathbf{x}. \end{aligned} \tag{4.24}$$

Recalling the choice made for $\mathbf{Q}$ and using the partitioned matrix inverse relations, the AMF statistic can be written

$$\begin{aligned} t_{AMF}(\mathbf{X}_0) &\overset{d}{=} \mathbf{x}_0^H \begin{bmatrix} \mathbf{I}_E \\ -\mathcal{T} \end{bmatrix} \Theta^{-1}\Delta(\Delta^H\Theta^{-1}\Delta)^{-1}\Delta^H\Theta^{-1}[\mathbf{I}_E \quad -\mathcal{T}^H]\mathbf{x}_0 \\ &= \mathbf{x}_0^H \begin{bmatrix} \mathbf{I}_E \\ -\mathcal{T} \end{bmatrix} \Theta^{-1}[\mathbf{I}_E \quad -\mathcal{T}^H]\mathbf{x}_0. \end{aligned} \tag{4.25}$$

Note that the inverse white noise SCM can be written in the form

$$(\mathbf{Z}\mathbf{Z}^H)^{-1} = \begin{bmatrix} \mathbf{I}_E \\ -\mathcal{T} \end{bmatrix} \Theta^{-1}[\mathbf{I}_E \quad -\mathcal{T}^H] + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & (\mathbf{Z}_2^H\mathbf{Z}_2)^{-1} \end{bmatrix} \tag{4.26}$$

from which it follows that

$$\begin{aligned} \mathbf{x}^H\mathbf{D}^{-1}\mathbf{x} &\stackrel{d}{=} \mathbf{x}^H\mathbf{R}^{-1/2}\mathbf{Q}^H(\mathbf{Z}\mathbf{Z}^H)^{-1}\mathbf{Q}\mathbf{R}^{-1/2}\mathbf{x} = \mathbf{x}_0^H(\mathbf{Z}\mathbf{Z}^H)^{-1}\mathbf{x}_0 \\ &= t_{AMF}(\mathbf{X}_0) + \mathbf{x}_2^H(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{x}_2. \end{aligned} \tag{4.27}$$

This relation along with eq(4.17) yields the very interesting GLRT/AMF stochastic relation

$$\boxed{\tilde{t}_{GLRT} \stackrel{d}{=} t_{AMF} \cdot \beta \qquad \text{where} \quad \beta \triangleq \frac{1}{1 + \mathbf{x}_2^H(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{x}_2}.} \tag{4.28}$$

Concerning the PD, we consider the GLRT statistic first. Subsequently, we provide parallel results for the AMF which follow based on this relation.

Recalling eq(4.25) and the partition for $\mathbf{x}_0$, note that

$$\tilde{t}_{GLRT} \stackrel{d}{=} \left\| \Theta^{-1/2}[\mathbf{x}_1 - \mathbf{Z}_1^H\mathbf{Z}_2(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{x}_2] \cdot \sqrt{\beta} \right\|^2. \tag{4.29}$$

Define the following vectors

$$\mathbf{b}^H = -\mathbf{x}_2^H(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{Z}_2^H \qquad\qquad \mathbf{b}_\beta^H = \mathbf{b}^H \cdot \sqrt{\beta}. \tag{4.30}$$

That the eigenvalues of projection matrices are ones and zeros says there exists a matrix $\mathbf{\Lambda}$ $L \times (L-N+E)$ which satisfies $\mathbf{\Lambda}\mathbf{\Lambda}^H = \mathfrak{P}(\mathbf{Z}_2|\mathbf{I}_L)$, $\mathbf{\Lambda}^H\mathbf{Z}_2 = \mathbf{0}$, and $\mathbf{\Lambda}^H\mathbf{\Lambda} = \mathbf{I}_{L-N+E}$. Next let vec($\mathbf{Z}_1$)= $\mathbf{z1}_v$ $EL \times 1$ and consider the following linear transformation

$$\boldsymbol{\eta} \triangleq \begin{bmatrix} (\boldsymbol{\vartheta})_{E\times 1} \\ \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_E \end{bmatrix}_{E(L-N+E)\times 1} \end{bmatrix} \triangleq \begin{bmatrix} \mathbf{I}_E \cdot \sqrt{\beta} & \begin{matrix} \mathbf{b}_\beta^H & & \\ & \mathbf{b}_\beta^H \;\; E \text{ terms} & \\ & \ddots & \\ & & \mathbf{b}_\beta^H \end{matrix} \\ \mathbf{0} & \begin{matrix} \mathbf{\Lambda}^H & & \\ & \mathbf{\Lambda}^H \;\; E \text{ terms} & \\ & \ddots & \\ & & \mathbf{\Lambda}^H \end{matrix} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{z1}_v \end{bmatrix} \tag{4.31}$$

$$\triangleq \mathbf{J}_\beta \cdot \boldsymbol{\xi}$$

where all matrix terms not shown are zero. Note that the quantities $\mathfrak{P}(\mathbf{Z}_2|\mathbf{I}_L)$, $\mathbf{b}$, and β

are scale invariant functions of their matrix arguments, *e.g.* $\mathfrak{P}(c\mathbf{Z}_2|\mathbf{I}_L) = \mathfrak{P}(\mathbf{Z}_2|\mathbf{I}_L)$ for all nonzero complex scalars c. Thus, calling $\boldsymbol{\xi}$ some statistic $\mathbf{F}_1(\mathbf{x}_1, \mathbf{Z}_1)$ and $\mathbf{J}_\beta$ some other statistic $\mathbf{F}_2(\mathbf{x}_2, \mathbf{Z}_2)$, it follows by Theorem 3.5.4 that the matrix $\mathbf{J}_\beta$ is independent of $\boldsymbol{\xi}$, which is distributed as

$$\boldsymbol{\xi} \sim \mathcal{CEC}_{E(L+1)}\left\{[\Delta \mathrm{s}; \mathbf{0}_{EL\times 1}],\ \mathbf{I}_{E(L+1)},\ g^{(1)}_{E(L+1):N(L+1)}(\cdot)\ \right\}. \tag{4.32}$$

Thus, treating $\mathbf{J}_\beta$ as deterministic, it follows by Proposition 3.5.4.3 that

$$\boldsymbol{\eta} \sim \mathcal{CEC}_{E(L-N+E+1)}\left\{[\Delta \mathrm{s}\sqrt{\beta}; \mathbf{0}_{E(L-N+E)\times 1}], \mathbf{I}_{E(L-N+E+1)},\ g^{(1)}_{E(L-N+E+1):N(L+1)}(\cdot)\ \right\}. \tag{4.33}$$

Observe that the variable β is the only quantity from the original matrix $\mathbf{J}_\beta$ which survives the linear transformation. We will therefore condition on β until the last step of the PD derivation.

For convenience let

$$K \triangleq L - N + E + 1 \qquad \text{and} \qquad \mathbf{Z}_\gamma^H \triangleq [\boldsymbol{\gamma}_1|\boldsymbol{\gamma}_2|\cdots|\boldsymbol{\gamma}_E]_{(L-N+E)\times E} \tag{4.34}$$

and note that the conditional pdf for $\boldsymbol{\eta}$ can be written in terms of $\mathbf{Z}_\gamma$

$$\boldsymbol{\eta} \sim g^{(1)}_{EK:N(L+1)}\left[||\boldsymbol{\vartheta} - \Delta \mathrm{s}\sqrt{\beta}||^2 + \mathrm{tr}(\mathbf{Z}_\gamma \mathbf{Z}_\gamma^H)\right] = P_{\boldsymbol{\vartheta},\, \mathbf{Z}_\gamma|\beta}. \tag{4.35}$$

Observing that $\Theta = \mathbf{Z}_1^H \Lambda \Lambda^H \mathbf{Z}_1 = \mathbf{Z}_\gamma \mathbf{Z}_\gamma^H$, it follows from the WET Proposition 3.6.0.4 that the joint density of $\boldsymbol{\vartheta}$ and Θ is given by

$$(\boldsymbol{\vartheta}, \Theta) \sim \frac{\pi^{E(K-1)}}{\tilde{\Gamma}_E(K-1)}|\Theta|^{L-N} g^{(1)}_{EK:N(L+1)}\left[||\boldsymbol{\vartheta} - \Delta \mathrm{s}\sqrt{\beta}||^2 + \mathrm{tr}(\Theta)\right] = P_{\boldsymbol{\vartheta},\, \Theta|\beta}. \tag{4.36}$$

The final observation needed to tie things back to the desired distribution of $\tilde{t}_{GLRT}$ is the following

$$\begin{aligned} \mathbf{x}_1 - \mathbf{Z}_1^H \mathbf{Z}_2 (\mathbf{Z}_2^H \mathbf{Z}_2)^{-1} \mathbf{x}_2 &\triangleq \mathbf{x}_1 - \mathbf{z}_A \\ &\stackrel{d}{=} \mathbf{x}_1 - \mathbf{z}_A^* = \boldsymbol{\vartheta}/\sqrt{\beta} \end{aligned} \tag{4.37}$$

where the superscript * indicates elementwise complex conjugation. The equality in distribution follows from Theorem 3.5.1. Hence, by eq(4.29) it follows that

$$\tilde{t}_{GLRT} \stackrel{d}{=} \boldsymbol{\vartheta}^H \Theta^{-1} \boldsymbol{\vartheta} = \boldsymbol{\vartheta}^H (\mathbf{Z}_\gamma \mathbf{Z}_\gamma^H)^{-1} \boldsymbol{\vartheta}. \tag{4.38}$$

Our next objective will be to reduce the effects of $\mathbf{Z}_\gamma$ to an equivalent statistic, and subsequently reduce $\boldsymbol{\vartheta}$.

Note from eq(4.35) that if we condition on $\boldsymbol{\vartheta}$, then $\mathbf{Z}_\gamma$ has a CSS MEC distribution

$$P_{\mathbf{Z}_\gamma|\boldsymbol{\vartheta},\beta} = \frac{P_{\boldsymbol{\vartheta},\mathbf{Z}_\gamma|\beta}}{P_{\boldsymbol{\vartheta}|\beta}} = \text{CSS MEC}. \tag{4.39}$$

With this conditioning maintained it follows that

$$\tilde{t}_{GLRT} \stackrel{d}{=} \boldsymbol{\vartheta}^H(\mathbf{Z}_\gamma\mathbf{Z}_\gamma^H)^{-1}\boldsymbol{\vartheta} \stackrel{d}{=} \boldsymbol{\vartheta}^H\mathbf{Q}_{\vartheta}^H(\mathbf{Z}_\gamma\mathbf{Z}_\gamma^H)^{-1}\mathbf{Q}_{\vartheta}\boldsymbol{\vartheta} \stackrel{d}{=} ||\boldsymbol{\vartheta}||^2/\alpha_g \tag{4.40}$$

with unitary $\mathbf{Q}_\vartheta$ chosen such that $\mathbf{Q}_\vartheta\boldsymbol{\vartheta} = [||\boldsymbol{\vartheta}||, 0, 0, \ldots, 0]^T$, leading to the definition $1/\alpha_g = [(\mathbf{Z}_\gamma\mathbf{Z}_\gamma^H)^{-1}]_{1,1}$. The relations for the inverse of a partitioned matrix show that α_g is the Schur complement of the matrix $\mathbf{Z}_\gamma\mathbf{Z}_\gamma^H$ of dimension 1 by 1. Hence, by the GWPT Theorem 3.6.1 part (ii), it follows that the conditional density for α_g is given by

$$P_{\alpha_g|\boldsymbol{\vartheta},\beta} = \frac{\pi^{(L-N+1)}}{(L-N)!}\,\alpha_g^{L-N}\,g^{(1)}_{K:N(L+1)}\left[||\boldsymbol{\vartheta} - \Delta\mathbf{s}\sqrt{\beta}||^2 + \alpha_g\right]\Big/P_{\boldsymbol{\vartheta}|\beta}. \tag{4.41}$$

The joint density of α_g and $\boldsymbol{\vartheta}$ is easily obtained by multiplication by $P_{\boldsymbol{\vartheta}|\beta}$.

Concerning $\boldsymbol{\vartheta}$, note that the decision statistic is a function of its magnitude only. Observe that $||\boldsymbol{\vartheta}||^2 = \boldsymbol{\vartheta}^H\boldsymbol{\vartheta} \stackrel{d}{=} \boldsymbol{\vartheta}^H\mathbf{Q}_a^H\mathbf{Q}_a\boldsymbol{\vartheta}$ for all unitary $\mathbf{Q}_a$, and consider a choice such that

$$\mathbf{Q}_a\boldsymbol{\vartheta} \stackrel{d}{=} \sqrt{\beta}\mathbf{Q}_a\Delta\mathbf{s} + \mathbf{Q}_a\mathbf{z}_\vartheta \stackrel{d}{=} \mathbf{e}_1\sqrt{\beta\mathbf{s}^H\Delta^H\Delta\mathbf{s}} + \mathbf{z}_\vartheta \tag{4.42}$$

where $\mathbf{e}_1$ is the standard basis vector with unity in the first position and zeros elsewhere. Note that

$$\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G} = \mathbf{G}^H\mathbf{R}^{-1/2}\mathbf{Q}^H\mathbf{Q}\mathbf{R}^{-1/2}\mathbf{G} = \begin{bmatrix}\Delta^H & \mathbf{0}^H\end{bmatrix}\begin{bmatrix}\Delta \\ \mathbf{0}\end{bmatrix} = \Delta^H\Delta. \tag{4.43}$$

Defining

$$\delta \triangleq \sqrt{\mathbf{s}^H\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G}\mathbf{s}} \qquad\qquad \delta_\beta \triangleq \delta\cdot\sqrt{\beta}, \tag{4.44}$$

it is therefore sufficient to restrict attention to the joint density

$$P_{\boldsymbol{\vartheta},\alpha_g|\beta} = \frac{\pi^{(L-N+1)}}{(L-N)!}\,\alpha_g^{L-N}\,g^{(1)}_{K:N(L+1)}\left[||\boldsymbol{\vartheta} - \delta_\beta\mathbf{e}_1||^2 + \alpha_g\right] \tag{4.45}$$

in the effort to specify the PD via the pdf of $\tilde{t}_{GLRT} \stackrel{d}{=} ||\ \boldsymbol{\vartheta}\ ||^2/\alpha_g$.

The argument of this density can be written

$$\begin{aligned} ||\boldsymbol{\vartheta} - \delta_\beta\mathbf{e}_1||^2 &= (\vartheta_{1R} - \delta_\beta)^2 + \vartheta_{1I}^2 + \vartheta_{2R}^2 + \vartheta_{2I}^2 + \cdots + \vartheta_{ER}^2 + \vartheta_{EI}^2 \\ &= (\vartheta_{1R} - \delta_\beta)^2 + \rho_1. \end{aligned} \tag{4.46}$$

Using the integral theorem of Proposition 3.3.1.1 (the real version of the WET) to integrate over all $\vartheta_{(\cdot)}$'s, save ϑ_{1R}, we obtain the joint density

$$P_{\vartheta_{1R},\rho_1,\alpha_g|\beta} = \frac{\pi^{(L-N+E+\frac{1}{2})}}{\Gamma(E-\frac{1}{2})(L-N)!}\,\alpha_g^{L-N}\,\rho_1^{E-\frac{3}{2}}\,g^{(1)}_{K:N(L+1)}\left[(\vartheta_{1R}-\delta_\beta)^2+\rho_1+\alpha_g\right]. \quad (4.47)$$

Introducing the change of variables

$$\begin{aligned} \vartheta_{1R} &= \sqrt{T}\cos\theta \\ \rho_I &= T\sin^2\theta \\ 0\le\theta\le\pi, & \qquad\qquad 0\le\rho_I<\infty \end{aligned} \quad (4.48)$$

note that $\|\,\boldsymbol{\vartheta}\,\|^2=\vartheta_{1R}^2+\rho_1=T$, and thus $\tilde{t}_{GLRT}\stackrel{d}{=}T/\alpha_g$. A little calculus shows that the Jacobian for this transformation is given by $J(\vartheta_{1R},\rho_I\to T,\theta)=\sqrt{T}\,\sin\theta$, and hence the density

$$P_{T,\theta,\alpha_g|\beta} = \frac{\pi^{(L-N+E+\frac{1}{2})}}{\Gamma(E-\frac{1}{2})(L-N)!}\,\alpha_g^{L-N}\,(T\sin^2\theta)^{E-1}\,g^{(1)}_{K:N(L+1)}\left(T-2\delta_\beta\sqrt{T}\cos\theta+\delta_\beta^2+\alpha_g\right).$$

Integrating over θ to obtain the joint density for T and α_g, and subsequent application of the Ratio Density Theorem F.0.0.1 of Appendix F to $\tilde{t}_{GLRT}\stackrel{d}{=}T/\alpha_g$ yields the desired density

$$\begin{aligned} P_{\tilde{t}_{GLRT}|\beta} &= k_1\cdot\int_0^\infty d\alpha_g\cdot\alpha_g^{(L-N+1)}\int_0^\pi d\theta\;(\sin^2\theta)^{E-1}(\tilde{t}_{GLRT}\,\alpha_g)^{E-1} \\ &\times g^{(1)}_{K:N(L+1)}\left[(\tilde{t}_{GLRT}\,\alpha_g)-2\sqrt{\tilde{t}_{GLRT}\,\alpha_g}\,\delta_\beta\cos\theta+\delta_\beta^2+\alpha_g\right] \end{aligned} \quad (4.49)$$

where $k_1=\pi^{(L-N+E+\frac{1}{2})}/\Gamma(E-\frac{1}{2})(L-N)!$. This integral expression can be manipulated into a more insightful form by considering the change of variables

$$r=\alpha_g(1+\tilde{t}_{GLRT}) \qquad\qquad d\alpha_g=dr/(1+\tilde{t}_{GLRT}). \quad (4.50)$$

A little algebra shows that

$$\boxed{\begin{aligned} P_{\tilde{t}_{GLRT}|\beta} &= \frac{\tilde{t}_{GLRT}^{E-1}}{(1+\tilde{t}_{GLRT})^{L-N+E+1}}\cdot k_1\cdot\int_0^\pi d\theta\;(\sin^2\theta)^{E-1} \\ &\times\int_0^\infty dr\;r^{K-1}g^{(1)}_{K:N(L+1)}\left[r-2\sqrt{\frac{r\,\tilde{t}_{GLRT}}{(1+\tilde{t}_{GLRT})}}\,\delta_\beta\cos\theta+\delta_\beta^2\right]. \end{aligned}} \quad (4.51)$$

Notice that the first factor is proportional to a complex central F-distribution (see Section 3.5.2). The pdf in eq(4.51) is in fact the complex MEC generalization of the complex non-central F-distribution.

Under the null hypothesis $\mathbf{s} = \mathbf{0}$ and therefore $\delta_\beta = 0$. Note the integral identities

$$\begin{aligned}\int_0^\infty r^{K-1} g^{(i)}_{K:N(L+1)}(r)dr &= \frac{(K-1)!}{\pi^K}\left[g^{(i)}_{K:N(L+1)}\right]_{0:K}(a=0) \\ &= \frac{(K-1)!}{\pi^K} g^{(i)}_{0:N(L+1)}(a=0) = \frac{(K-1)!}{\pi^K}\end{aligned} \quad (4.52)$$

(which is true for both $i = 0, 1$)

$$\int_0^\pi (\sin^2\theta)^{E-1} d\theta = \sqrt{\pi}\Gamma[E-(1/2)]/(E-1)!,$$

where the first identity follows from Theorems 3.5.2 and 3.5.3. Applying these under H_0, it is shown that $P_{\tilde{t}_{GLRT}}$ is complex central F-distributed such that

$$P_{\tilde{t}_{GLRT}|\beta} = \left[\frac{(L-N+E)!}{(L-N)!(E-1)!}\right]\frac{\tilde{t}^{E-1}_{GLRT}}{(1+\tilde{t}_{GLRT})^{L-N+E+1}} \stackrel{d}{=} F_{E,(K-E)}. \quad (4.53)$$

This distribution is independent of β, $\mathbf{R}$, and the actual functional form of the data matrix density, *i.e.* $g^{(0)}(\cdot)$. This validates the detector as CFAR, and coincides with known complex Gaussian based results [19] under the null hypothesis, as anticipated by scale invariance. Indeed, note from Section 3.5.2 that the PFA is obtainable via the cdf relations for the complex central F and β:

$$\begin{aligned}\mathrm{PFA}_{GLRT} = Pr(t_{GLRT} \geq \eta) &= 1 - Pr(\tilde{t}_{GLRT} \leq \eta - 1) \\ &= 1 - Pr(F_{E,K-E} \leq \eta - 1) \\ &= 1 - [1 - Pr(\beta_{K-E,E} \leq 1/\eta)].\end{aligned} \quad (4.54)$$

Hence, we obtain the following simple invariant expression for the PFA of a complex MEC based GLRT detector:

$$\boxed{\mathrm{PFA}_{GLRT} = \left(\frac{1}{\eta}\right)^{(L-N+E)} \sum_{k=0}^{E-1} \binom{L-N+E}{k} (\eta - 1)^k.} \quad (4.55)$$

When $E = 1$ we obtain the rather simple and well known formula $\mathrm{PFA}_{GLRT} = (1/\eta)^{L-N+1}$.

For the AMF it follows from eq(4.28) and eq(4.29) that

$$t_{AMF} \stackrel{d}{=} \tilde{t}_{GLRT}/\beta \stackrel{d}{=} \left\|\Theta^{-1/2}[\mathbf{x}_1 - \mathbf{Z}_1^H\mathbf{Z}_2(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{x}_2]\right\|^2. \quad (4.56)$$

By simply replacing the matrix $\mathbf{J}_\beta$ with the matrix $\mathbf{J}$

$$\mathbf{J} \triangleq \begin{bmatrix} \mathbf{I}_E & \begin{matrix} \mathbf{b}^H & & & \\ & \mathbf{b}^H & E \text{ terms} & \\ & & \ddots & \\ & & & \mathbf{b}^H \end{matrix} \\ \mathbf{0} & \begin{matrix} \mathbf{\Lambda}^H & & & \\ & \mathbf{\Lambda}^H & E \text{ terms} & \\ & & \ddots & \\ & & & \mathbf{\Lambda}^H \end{matrix} \end{bmatrix} \tag{4.57}$$

in eq(4.31), analogous reasoning shows that

$$\boldsymbol{\eta} \sim \beta^E \cdot g^{(1)}_{EK:N(L+1)} \left[||\boldsymbol{\vartheta} - \Delta \mathbf{s}||^2 \beta + \mathrm{tr}(\mathbf{Z}_\gamma \mathbf{Z}_\gamma^H) \right] = P_{\boldsymbol{\vartheta}, \mathbf{Z}_\gamma | \beta}. \tag{4.58}$$

Following exactly the same procedures as in the analysis of the $\tilde{t}_{GLRT}$, it can be shown that

$$\boxed{\begin{aligned} P_{t_{AMF}|\beta} &= \frac{\beta^E \ t_{AMF}^{E-1}}{(1+\beta \ t_{AMF})^{L-N+E+1}} \cdot k_1 \cdot \int_0^\pi d\theta \ (\sin^2\theta)^{E-1} \\ &\times \int_0^\infty dr \ r^{K-1} g^{(1)}_{K:N(L+1)} \left[r - 2\sqrt{\frac{r \ \beta \ t_{AMF}}{(1+\beta t_{AMF})}} \ \delta_\beta \cos\theta + \delta_\beta^2 \right]. \end{aligned}} \tag{4.59}$$

Under the null hypothesis, we reason as in the GLRT case that

$$P_{t_{AMF}|\beta} = \left[\frac{(L-N+E)!}{(L-N)!(E-1)!} \right] \frac{\beta^E \ t_{AMF}^{E-1}}{(1+\beta \ t_{AMF})^{L-N+E+1}} \overset{d}{=} \frac{1}{\beta} \cdot F_{E,K-E} \tag{4.60}$$

independent of $\mathbf{R}$ and the actual functional form of the data matrix density, *i.e.* $g^{(0)}(\cdot)$. The AMF is still CFAR, as anticipated earlier. In the AMF case, however, note that as in Robey [23] the PFA depends on β.

Concerning the density of β, note that by scale invariance and Proposition 3.5.4.5, its pdf is independent of $g^{(i)}(\cdot)$, $i = 0, 1$. Thus, without loss of generality assume that $[\mathbf{x}_2, \mathbf{Z}_2^H]$ is a zero mean complex white Gaussian data matrix, from which it follows that

$$\mathbf{x}_2^H (\mathbf{Z}_2^H \mathbf{Z}_2)^{-1} \mathbf{x}_2 \overset{d}{=} ||\mathbf{x}_2||^2 / \chi^2_{L-N+E+1} \overset{d}{=} \frac{\chi^2_{N-E}}{\chi^2_K}. \tag{4.61}$$

The first equality from Capon/Goodman [33] or part (ii) of GWPT 3.6.1, and the second is well known. Thus, by eq(4.28) and the propositions of Section 3.5.2, it follows that

$$P_\beta = \frac{L!}{(K-1)!(N-E-1)!} \beta^{K-1} (1-\beta)^{(N-E)-1} \qquad 0 \leq \beta \leq 1. \tag{4.62}$$

Clearly, the a priori densities for $\tilde{t}_{GLRT}$ and t_{AMF} are given by

$$\boxed{\begin{aligned} P_{\tilde{t}_{GLRT}} &= \int_0^1 P_\beta \; P_{\tilde{t}_{GLRT}|\beta} \, d\beta \\ P_{t_{AMF}} &= \int_0^1 P_\beta \; P_{t_{AMF}|\beta} \, d\beta \end{aligned}} \tag{4.63}$$

from which we can in theory compute the PD's.

4.4.2 Special Case of Complex Gaussian

Recall from Proposition 3.5.3.8 that for Gaussian data $g_{K:N(L+1)}(r) = \pi^{-K} e^{-r}$. Employing the integral identities

$$\int_0^\pi e^{\pm x \cos y} (\sin^2 y)^M \, dy = \sqrt{\pi}\,\Gamma[M + (1/2)]\, {}_0F_1[M+1; x^2/4]/M! \tag{4.64}$$

$$\int_0^\infty x^{J-1} e^{-x} {}_0F_1[M+1; ax] \, dx = (J-1)!\, {}_1F_1[J; M+1; a] \tag{4.65}$$

eq(4.51) reduces to a complex non-central F, *i.e.* given β the distribution is $\tilde{t}_{GLRT} \sim F_{E,K-E}(\delta_\beta)$ in the notation of Section 3.5.2. The PD for the Gaussian based GLRT test is shown to be obtainable from

$$\begin{aligned} \mathrm{PD}_{GLRT\,|\,\beta} = Pr(t_{GLRT} \geq \eta \mid \beta) &= 1 - Pr(\tilde{t}_{GLRT} \leq \eta - 1 \mid \beta) \\ &= 1 - Pr(F_{E,K-E}(\delta_\beta) \leq \eta - 1 \mid \beta). \end{aligned} \tag{4.66}$$

Recalling eq(3.72) the conditional PD is given by the finite sum

$$\boxed{\mathrm{PD}_{GLRT\,|\,\beta} = 1 - \frac{(\eta-1)^{(E-1)}}{\eta^{(L-N+E)}} \sum_{k=0}^{L-N+1} \binom{L-N+E}{k+E-1} (\eta-1)^k \cdot IG_k(\delta_\beta^2/\eta).} \tag{4.67}$$

Note that when $E = 1$ this reduces to the well known PD expression for a scalar CFAR detector (see p. 123 of [19]). The a priori PD is given by

$$\mathrm{PD}_{GLRT} = \int_0^1 P_\beta \cdot \mathrm{PD}_{GLRT\,|\,\beta} \, d\beta. \tag{4.68}$$

The same integral identities used to reduce eq(4.51) to Kelly's result when the data is complex Gaussian, can be used to show that eq(4.59) reduces to the distribution $t_{AMF} \sim F_{E,K-E}(\delta_\beta)/\beta$, when the data is Gaussian. This special case is Robey's [23] result.

4.4.3 Special Case of Complex Gaussian Mixtures/Compounds

As implicated in Chapter 3, the pdf functional forms for any complex Gaussian mixture or compound distributed data matrix can be written

$$g^{(i)}(r) = \int_0^\infty \frac{1}{\tau_i} e^{-r/\tau_i} P_{\tau_i}, \qquad i = 0, 1. \tag{4.69}$$

Gaussian mixtures can be written identically distributed to a normal data matrix times an independent scalar r.v.:

$$\mathbf{X}_0 \stackrel{d}{=} \mathbf{M}_i + \mathbf{N}\sqrt{\tau_i} \tag{4.70}$$

where $\mathbf{N} \sim \mathcal{CN}_{N(L+1)}(\mathbf{0}, \mathbf{I}_{L+1} \otimes \mathbf{R})$ is independent of the real positive r.v.'s τ_i, $i = 0, 1$. Using the H_1 stochastic representation $\mathbf{X}_0 = [\mathbf{X}|\mathbf{x}] \stackrel{d}{=} [\mathbf{0}_{N\times L}|\mathbf{Gs}] + \mathbf{N}\sqrt{\tau}_1$, it can be shown that given β and τ_1 the GLRT statistic is distributed as $\tilde{t}_{GLRT} \sim F_{E,K-E}(\delta_\beta/\sqrt{\tau_1})$. The unconditional PD can be found via

$$\boxed{P_{\tilde{t}_{GLRT}} = \int_0^\infty P_{\tau_1}\, d\tau_1 \int_0^1 P_\beta \;\; P_{\tilde{t}_{GLRT}|\beta,\tau_1}\, d\beta.} \tag{4.71}$$

Thus, simply dividing the non-centrality parameter δ_β^2 *by* τ_1*, and subsequently averaging Kelly's or Robey's finite sum expressions over* τ_1 *gives the desired a priori PD's for the Gaussian mixture models.* An approach which is useful for simulations.

In the next section we develop statistics for adaptive beamformers. Results from this chapter and the next will be demonstrated via simulations in Chapter 6.

Chapter 5

Adaptive Array Signal Estimation and Beamforming with Complex MEC Data

5.1 Introduction

As discussed in Chapter 2, first generation optimal array processors are designed/derived under rather idealized assumptions, which often includes a priori knowledge of the statistics of the data observations; namely, the data covariance parameter $\mathbf{R}$ is assumed known. In practice, however, such an assumption rarely holds. Consequently, one is lead to (1) attempt derivation of second generation array processors which drop the ideal assumption of known data covariance, *e.g.* concerning AAD the GLRT approach is one such second generation array processor (second to the first generation MF), or (2) yield to the common heuristic method of using an estimate of this nuisance parameter in place of its true unknown value (SMI), *e.g.* with regard to AAD the AMF approach is more in the flavor of SMI. Concerning adaptive signal estimation and beamforming, it was proven in Section 4.2 that both approaches lead to the same processors for a large class of complex MECs. Indeed, the SCB LCMV beamformer was shown to be optimal in the ML sense for complex MEC data. This in effect serves as a motivation for the use of SMI in non-Gaussian environments.

The effects of estimated data covariances on the performance of adaptive signal estimators and beamformers in Gaussian environments has been studied in great detail [28]–[33]. Given that the use of estimated data covariances inevitably results in a loss in overall system performance, the intended goal of such analyses is typically to quantify in a statistical sense the performance loss experienced as a consequence of using an SCM as opposed to the true covariance. In this chapter we add to the many results on SCB array processors by (i) weakening the traditional assumption of data Gaussianity, and (ii) subsequently providing for a class of array processors additional performance measures we believe to be of value in practice. The data matrix of snapshots is assumed complex MEC distributed. The performance measures include the exact pdfs, confidence regions, and first and second moments of the weight vector/matrix, beam response, and beamformer output of certain SCB array processors. The array processors considered include the SCB (1) ML signal vector estimator, (2) LCMV and (3) MVDR beamformers, and (4) GSC implementation of the LCMV. We relate the results obtained in this chapter to the works of references [14, 16, 18, 22, 28].

5.2 Derivation of PDFs for SCB Weightings

The computation and use of the exact pdfs for the SCB weightings is the intended goal of this chapter. The resulting statistical decompositions of the adaptive weight vectors/matrices from which the pdfs are deduced, interestingly lend themselves to a structurally consistent Generalized Sidelobe Cancellor interpretation which we likewise illustrate.

5.2.1 Invariance in Adaptive Beamforming

As in Chapter 4, it is assumed that the data matrix $\mathbf{X}_0 = [\mathbf{X}|\mathbf{x}]$ is distributed as

$$\mathbf{X}_0 \sim \mathcal{CMEC}_{N(L+1)}\left\{[\mathbf{0}_{N\times L}|\mathbf{Gs}], \mathbf{I}_{L+1}\otimes\mathbf{R}, g(\cdot)\right\}. \tag{5.1}$$

Recall from Section 3.8 that

$$\mathbf{x} \sim \mathcal{CEC}_N\left[\mathbf{Gs}, \mathbf{R}, g_{N:N(L+1)}(\cdot)\right] \qquad \mathbf{X} \sim \mathcal{CMEC}_{NL}\left[\mathbf{0}, \mathbf{I}_L\otimes\mathbf{R}, g_{NL:N(L+1)}(\cdot)\right]. \tag{5.2}$$

Recalling that the SCM is formed from the secondary data via $\mathbf{D} = \mathbf{X}\mathbf{X}^H$, the SCB weightings considered are:

$$\begin{array}{rl} & \underline{SCB\ Weightings} \\ \text{ML:} & \widehat{\mathbf{W}}_{ML} = \mathbf{D}^{-1}\mathbf{G}(\mathbf{G}^H\mathbf{D}^{-1}\mathbf{G})^{-1} \\ \text{LCMV/MVDR:} & \widehat{\mathbf{w}}_{LCMV} = \widehat{\mathbf{W}}_{ML}\mathbf{f} \\ \text{GSC:} & \widehat{\mathbf{w}}_{gsc} = \widehat{\mathbf{W}}_{GSC}\mathbf{f} \\ & \widehat{\mathbf{W}}_{GSC} = \left[\mathbf{I}_N - \mathbf{G}_\perp\widehat{\boldsymbol{\Omega}}_{GSC}\right]\mathbf{G}(\mathbf{G}^H\mathbf{G})^{-1} \\ & \widehat{\boldsymbol{\Omega}}_{GSC} = (\mathbf{G}_\perp^H\,\mathbf{D}\,\mathbf{G}_\perp)^{-1}\mathbf{G}_\perp^H\,\mathbf{D}. \end{array} \tag{5.3}$$

Note that each SCB weighting can be considered a matrix of statistics $\mathbf{F}(\mathbf{X})$ which satisfy the scale invariant property $\mathbf{F}(c\mathbf{X}) = \mathbf{F}(\mathbf{X})$ for all nonzero complex scalars c. Hence, by Proposition 3.5.4.5 the pdfs for these weightings do not depend on $g(\cdot)$ the functional form of the original density. We will, ergo, assume with out loss of generality over complex MECs that $g_{NL:N(L+1)}(r) = \pi^{-NL}e^{-r}$, *i.e.* $\mathbf{X} \sim \mathcal{CN}_{NL}[\mathbf{0}, \mathbf{I}_L \otimes \mathbf{R}]$.

5.2.2 SCB ML/LCMV Weightings

The clairvoyant ML and LCMV beamformer weightings and their corresponding SCB counterparts, recall are given respectively in eq(2.22) and eq(5.3). Note that specifying the pdf of $\widehat{\mathbf{W}}_{ML}$ is clearly the common goal. As in Chapter 4 we start by representing the data matrix as a linearly transformed white data matrix, *i.e.* $\mathbf{X} \stackrel{d}{=} \mathbf{R}^{1/2}\mathbf{Q}^H\mathbf{Z}$ where $\mathbf{Q}$ is unitary. The matrix $\mathbf{Z}$ represents a white data matrix whose L columns are i.i.d. as $\mathcal{CN}_N(\mathbf{0}, \mathbf{I}_N)$. Recalling eq(4.18), the following stochastic representation for the SCB weight matrix is obtained

$$\widehat{\mathbf{W}}_{ML} \stackrel{d}{=} \mathbf{R}^{-1/2}\mathbf{Q}^H(\mathbf{Z}\mathbf{Z}^H)^{-1}\mathbf{Q}\mathbf{R}^{-1/2}\mathbf{G}\left[\mathbf{G}^H\mathbf{R}^{-1/2}\mathbf{Q}^H(\mathbf{Z}\mathbf{Z}^H)^{-1}\mathbf{Q}\mathbf{R}^{-1/2}\mathbf{G}\right]^{-1}. \tag{5.4}$$

Choosing $\mathbf{Q}$ as in eq(4.19), yields the relation $\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G} = \Delta^H\Delta$ (as shown in eq(4.43)), and implies that

$$\mathbf{R}^{-1}\mathbf{G} = \mathbf{R}^{-1/2}\mathbf{Q}^H\mathbf{Q}\mathbf{R}^{-1/2}\mathbf{G} = \mathbf{R}^{-1/2}\mathbf{Q}^H\begin{bmatrix}\Delta \\ \mathbf{0}\end{bmatrix}, \tag{5.5}$$

allowing for the alternate representation of the clairvoyant weighting

$$\begin{aligned} \mathbf{W}_{ML} &= \mathbf{R}^{-1}\mathbf{G}(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1} &&= \mathbf{R}^{-1/2}\mathbf{Q}^H \begin{bmatrix} \Delta \\ \mathbf{0} \end{bmatrix} (\Delta^H\Delta)^{-1} \\ &= \mathbf{R}^{-1/2}\mathbf{Q}^H \begin{bmatrix} \Delta\Delta^{-1} \\ \mathbf{0} \end{bmatrix} \Delta^{-H} &&= \mathbf{R}^{-1/2}\mathbf{Q}^H \begin{bmatrix} \mathbf{I}_E \\ \mathbf{0} \end{bmatrix} \Delta^{-H}. \end{aligned} \tag{5.6}$$

Partition the white data matrix, and its SCM inverse as in eq(4.21)–eq(4.24). Because $\mathbf{Z}$ is a normal data matrix with columns i.i.d. as $\mathcal{CN}_N(\mathbf{0}, \mathbf{I}_N)$, it follows that $\mathbf{Z}_1$ and $\mathbf{Z}_2$ are independent. The partitioning allows the weight matrix to be expressed in a more compact form. Note that

$$\widehat{\mathbf{W}}_{ML} \stackrel{d}{=} \mathbf{R}^{-1/2}\mathbf{Q}^H(\mathbf{Z}\mathbf{Z}^H)^{-1}\begin{bmatrix} \Delta \\ \mathbf{0} \end{bmatrix} \cdot \left\{ \begin{bmatrix} \Delta^H & \mathbf{0}^H \end{bmatrix} (\mathbf{Z}\mathbf{Z}^H)^{-1} \begin{bmatrix} \Delta \\ \mathbf{0} \end{bmatrix} \right\}^{-1}, \tag{5.7}$$

which can be simplified to obtain

$$\widehat{\mathbf{W}}_{ML} \stackrel{d}{=} \mathbf{R}^{-1/2}\mathbf{Q}^H \begin{bmatrix} \Theta^{-1}\Delta \\ -\mathcal{T}\Theta^{-1}\Delta \end{bmatrix} (\Delta^H\Theta^{-1}\Delta)^{-1} = \mathbf{R}^{-1/2}\mathbf{Q}^H \begin{bmatrix} \mathbf{I}_E \\ -\mathcal{T} \end{bmatrix} \Delta^{-H}. \tag{5.8}$$

Observing that

$$\begin{aligned} \widehat{\mathbf{W}}_{ML} &\stackrel{d}{=} \mathbf{R}^{-1/2}\mathbf{Q}^H \begin{bmatrix} \mathbf{I}_E \\ -\mathcal{T} \end{bmatrix} \Delta^{-H} \\ &= \mathbf{R}^{-1/2}\mathbf{Q}^H \begin{bmatrix} \mathbf{I}_E \\ \mathbf{0}_{(N-E)\times E} \end{bmatrix} \Delta^{-H} + \mathbf{R}^{-1/2}\mathbf{Q}^H \begin{bmatrix} \mathbf{0}_{E\times E} \\ -\mathcal{T} \end{bmatrix} \Delta^{-H} \end{aligned} \tag{5.9}$$

and recalling eq(5.6) and that

$$\mathcal{T}_{(N-E)\times E} = (\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{Z}_2^H\mathbf{Z}_1, \tag{5.10}$$

we write $\widehat{\mathbf{W}}_{ML}$ as

$$\widehat{\mathbf{W}}_{ML} \stackrel{d}{=} \mathbf{W}_{ML} + \mathbf{R}^{-1/2}\mathbf{Q}^H \begin{bmatrix} \mathbf{0}_{E\times E} \\ -(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{Z}_2^H\mathbf{Z}_1 \end{bmatrix} \Delta^{-H}. \tag{5.11}$$

Eq(5.11) is the matrix analog of eq(17) of reference [28]. From this equation one can reason that the SCB ML/LCMV weight matrix remains an unbiased estimate of the true weight matrix. Note that the conditional expectation given $\mathbf{Z}_2$ is $E\{\widehat{\mathbf{W}}_{ML}|\mathbf{Z}_2\} = \mathbf{W}_{ML}$, which is completely independent of $\mathbf{Z}_2$, and is therefore also the unconditional expectation.

Partition $\mathbf{R}^{-1/2}\mathbf{Q}^H$ such that

$$\mathbf{R}^{-1/2}\mathbf{Q}^H = \left[[\mathcal{W}_{\mathcal{ML}}]_{N\times E} \;\middle|\; [\mathbf{G}_\perp]_{N\times(N-E)} \right]. \tag{5.12}$$

Substituting this partitioned form into eq(4.19), note that $\mathbf{G}^H\mathcal{W}_{\mathcal{ML}} = \Delta^H$ and $\mathbf{G}^H\mathbf{G}_\perp = \mathbf{0}$.

The span($\mathbf{G}_\perp$) consists of the space orthogonal to the array manifold $\mathbf{G}$. Considering that $\mathbf{R}^{-1/2}\mathbf{Q}^H$ is nonsingular and therefore full rank, and that $\mathcal{W}_{\mathcal{ML}} = \mathbf{W}_{ML}\Delta^H$ by substitution of eq(5.12) into eq(5.6), it is shown that rank($\mathbf{G}$) = rank($\mathcal{W}_{\mathcal{ML}}$) and that span($\mathcal{W}_{\mathcal{ML}}$) = span($\mathbf{W}_{ML}$). The above partition allows us to express the SCB weight matrix as

$$\boxed{\widehat{\mathbf{W}}_{ML} \stackrel{d}{=} \mathbf{W}_{ML} - \mathbf{G}_\perp \mathcal{T} \Delta^{-H}.} \tag{5.13}$$

From this stochastic decomposition of the weight matrix note that it consists of two additive parts. One is a fixed and deterministic component, while the other is random. The fixed part lies in the span($\mathcal{W}_{\mathcal{ML}}$) and is given by the clairvoyant weight matrix. The random part lies in the span($\mathbf{G}_\perp$). The uncertainty present in the SCB weight matrix is summarized by the quantity $\mathcal{T}$. Consequently, finding the pdf for this quantity is the desired goal. Consider first, however, an interesting interpretation of this stochastic representation. Observe the analogous structure and form of eq(5.13) to eq(2.35). This analogy essentially provides a GSC viewpoint of the stochastic representation of $\widehat{\mathbf{W}}_{ML}$, which is illustrated in fig(5.1).

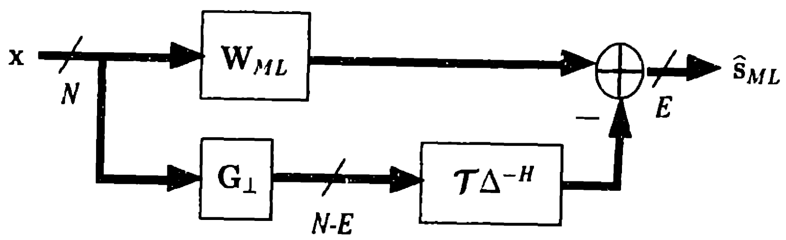


Figure 5.1: GSC Interpretation of Stochastic Representation of $\widehat{\mathbf{W}}_{ML}$

The role of $\mathcal{T}$ can be interpreted as a source of undesired random and sporadic "sidelobe" variation in view of fig(5.1), where the "sidelobes" are those of the ideal clairvoyant beamformer $\mathbf{W}_{ML}$. The signal blocking matrix $\mathbf{G}_\perp$ assures that the imposed beamformer constraints are always satisfied. Although the intended goal here is to derive the distribution of $\widehat{\mathbf{W}}_{ML}$, it is quite fascinating and encouraging that the intuitive structure motivating the

GSC implementation of the LCMV beamformer falls out of the mathematics in the process.

Concerning the pdf of $\mathcal{T}$, recall from the partitioned matrix inverse relations that $\mathcal{T}$ is simply the regression coefficient of the GWPT Theorem 3.6.1. Hence, we readily conclude from part (ii) of the GWPT that the density of $\mathcal{T}$ is given by the following *standardized* complex multivariate *t*-distribution

$$\boxed{P_{\mathcal{T}}(\mathcal{T}_0) = \left[\frac{\tilde{\Gamma}_{N-E}(L+E)}{\tilde{\Gamma}_{N-E}(L)\pi^{(N-E)E}}\right] \times \left|\mathbf{I}_{N-E} + \mathcal{T}_0\mathcal{T}_0^H\right|^{-(L+E)}.} \tag{5.14}$$

A *generalized* complex multivariate t–distributed $M \times J$ ($M \geq J$) matrix, say $\widetilde{\mathcal{T}}$, with L' degrees of freedom has pdf

$$\left[\frac{\tilde{\Gamma}_J(L'+M)}{\tilde{\Gamma}_J(L')\pi^{MJ}}\right] |\mathbf{V}|^{-J}|\mathbf{H}|^{-L'} \left|\mathbf{H}^{-1} + (\widetilde{\mathcal{T}}_0 - \mathbf{T})^H\mathbf{V}^{-1}(\widetilde{\mathcal{T}}_0 - \mathbf{T})\right|^{-(L'+M)} \tag{5.15}$$

which we shall denote by the shorthand notation

$$\widetilde{\mathcal{T}}_{M\times J} \sim \mathcal{CT}_{M\times J}\left[\mathbf{T}_{M\times J}, L', \mathbf{V}_{M\times M}, \mathbf{H}_{J\times J}\right]. \tag{5.16}$$

Recall that we encountered this distribution type in the GWPT Theorem 3.6.1. A little linear algebra shows that the shorthand notation for eq(3.123) is $\mathcal{CT}_{(N-M)\times M}[\tilde{\mathbf{T}}, L - N + 2M, \tilde{\mathbf{R}}_{22}^{-1}, \tilde{\mathbf{R}}_{11.2}^{-1}]$ (also see Lemma 2 of [18]). Noting that $|\mathbf{I}_{N-E} + \mathcal{T}_0\mathcal{T}_0^H| = |\mathbf{I}_E + \mathcal{T}_0^H\mathcal{T}_0|$, and that $\tilde{\Gamma}_{N-E}(L+E)/\tilde{\Gamma}_{N-E}(L) = \tilde{\Gamma}_E(L+E)/\tilde{\Gamma}_E(L-N+2E)$, the shorthand notation for the pdf in eq(5.14) is similarly seen to be $\mathcal{CT}_{(N-E)\times E}\left[\mathbf{0}_{(N-E)\times E},\, L-N+2E, \mathbf{I}_{N-E}, \mathbf{I}_E\right]$.

This completes the derivation for the pdf of $\widehat{\mathbf{W}}_{ML}$. Eq(5.14) together with eq(5.13) completely specifies the model of uncertainty for the SCB weight matrix. The explicit pdf for $\widehat{\mathbf{W}}_{ML}$, *i.e.* an equation of the form "$P_{\widehat{\mathbf{W}}_{ML}}$ equals," is in fact singular. As discussed in Section 3.7, singular distributions arise when there is something partially deterministic about a random quantity. In the case of the SCB weight matrix it is known with certainty that $\mathbf{G}^H\widehat{\mathbf{W}}_{ML} = \mathbf{I}_E$ independent of variation in $\widehat{\mathbf{R}}$. This is the result of the design constraint imposed by eq(2.26).

Singular distributions are expressible in two forms. The first form has been illustrated via eq(5.13) and eq(5.14). Recall from Section 3.7 that an alternative form which allows for an explicit pdf of the form "$P_{\widehat{\mathbf{W}}_{ML}}$ equals" is possible by introducing the matrix pseudo-inverse [62, 71]. The latter form will result in the following generalized singular complex multivariate t-distribution:

$$P_{\widehat{\mathbf{W}}_{ML}} = \left[\frac{\tilde{\Gamma}_E(L+E)}{\tilde{\Gamma}_E(L-N+2E)\pi^{(N-E)E}}\right] |\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G}|^{-(L-N+2E)} \left\{\prod_{i=1}^{N-E} \lambda_i \left[\mathfrak{P}(\mathbf{G}|\mathbf{R}^{-1})\right]\right\}^{-E} \quad (5.17)$$
$$\times \left|(\widehat{\mathbf{W}}_{ML} - \mathbf{W}_{ML})^H \mathfrak{P}(\mathbf{G}|\mathbf{R}^{-1})^+ (\widehat{\mathbf{W}}_{ML} - \mathbf{W}_{ML}) + (\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}\right|^{-(L+E)}$$
$$\times \tilde{\delta}\left[\mathrm{vec}(\mathbf{G}^H\widehat{\mathbf{W}}_{ML}) - \mathrm{vec}(\mathbf{G}^H\mathbf{W}_{ML})\right].$$

See Appendix D for proof.

The pdf for the LCMV beamformer $\widehat{\mathbf{w}}_{LCMV}$ is simply

$$\boxed{\begin{aligned} \widehat{\mathbf{w}}_{LCMV} &\stackrel{d}{=} \mathbf{w}_{LCMV} - \mathbf{G}_\perp \mathcal{T} \Delta^{-H} \mathbf{f} \\ &\stackrel{d}{=} \mathbf{w}_{LCMV} - \mathbf{G}_\perp \tilde{\mathbf{t}}_{LCMV} \sqrt{\mathbf{f}^H(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}\mathbf{f}} \end{aligned}} \quad (5.18)$$

with $\mathcal{T}$ as in eq(5.14) or $\tilde{\mathbf{t}}_{LCMV}$ as $\mathcal{CT}_{(N-E)\times 1}[\mathbf{0}, K, \mathbf{I}_{N-E}, 1]$, where $K = L-N+E+1$ as in eq(4.34). The slightly simpler form is possible by noting that for any vector $\mathbf{q}_{E\times 1}$ it is true that $\mathbf{Z}_1\mathbf{q} \stackrel{d}{=} \mathbf{z}_0 \times ||\mathbf{q}||$ where $\mathbf{z}_0 \sim \mathcal{CN}_L(\mathbf{0}, \mathbf{I}_L)$. Thus, $\mathcal{T}\Delta^{-H}\mathbf{f} \stackrel{d}{=} \tilde{\mathbf{t}}_{LCMV}\sqrt{\mathbf{f}^H(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}\mathbf{f}}$. The resulting GSC interpretation of the stochastic representation of $\widehat{\mathbf{w}}_{LCMV}$ is shown in fig(5.2). It is known with certainty that $\mathbf{G}^H\widehat{\mathbf{w}}_{LCMV} = \mathbf{f}$. Thus, the explicit pdf for $\widehat{\mathbf{w}}_{LCMV}$ is likewise a generalized singular complex t; namely,

$$P_{\widehat{\mathbf{w}}_{LCMV}} = \left[\frac{L!}{(K-1)!\pi^{(N-E)}}\right] \left[\mathbf{f}^H(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}\mathbf{f}\right]^K \left\{\prod_{i=1}^{N-E} \lambda_i \left[\mathfrak{P}(\mathbf{G}|\mathbf{R}^{-1})\right]\right\}^{-1} \quad (5.19)$$
$$\times \left[(\widehat{\mathbf{w}}_{LCMV} - \mathbf{w}_{LCMV})^H \mathfrak{P}(\mathbf{G}|\mathbf{R}^{-1})^+ (\widehat{\mathbf{w}}_{LCMV} - \mathbf{w}_{LCMV}) + \mathbf{f}^H(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}\mathbf{f}\right]^{-(L+1)}$$
$$\times \tilde{\delta}\left[\mathbf{G}^H\widehat{\mathbf{w}}_{LCMV} - \mathbf{f}\right].$$

One utility of these stochastic representations for the SCB weightings is that they afford explicit expressions for their covariances. One can show that the covariances for the

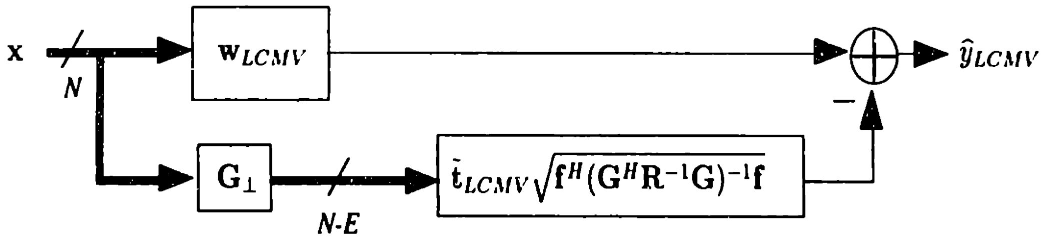


Figure 5.2: GSC Interpretation of Stochastic Representation of $\hat{\mathbf{w}}_{LCMV}$

ML/LCMV beamformer weightings are given by

$$\begin{aligned} \mathrm{cov}(\widehat{\mathbf{W}}_{ML}) &= \frac{(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-T}}{(L-N+E)} \otimes \mathfrak{P}(\mathbf{G}|\mathbf{R}^{-1}) \\ \mathrm{cov}(\hat{\mathbf{w}}_{LCMV}) &= \frac{\mathbf{f}^H(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}\mathbf{f}}{(L-N+E)} \times \mathfrak{P}(\mathbf{G}|\mathbf{R}^{-1}). \end{aligned} \tag{5.20}$$

See Appendix C for the derivation of these covariances. As noted earlier the random parts of these SCB weightings lie completely in the space orthogonal to $\mathbf{G}$; namely, span($\mathbf{G}_\perp$). This is also evidenced by the fact that the null spaces of their covariances belong to span($\mathbf{G}$).

For the particular case of the MVDR beamformer, *i.e.* $E = 1$ and $\mathrm{f} = 1$, note that eq(5.18) and eq(5.14) simplify to

$$\begin{aligned} \hat{\mathbf{w}}_{MVDR} &\stackrel{d}{=} \mathbf{w}_{MVDR} - \mathbf{G}_\perp \mathbf{t}_{MVDR}/\sqrt{(\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g})} \\ P_{\mathbf{t}_{MVDR}}(\mathbf{t}_0) &= \left[\frac{L!}{(L-N+1)!\pi^{(N-1)}}\right] \times \left(1+||\mathbf{t}_0||^2\right)^{-(L+1)} \end{aligned} \tag{5.21}$$

respectively, and its covariance is given by

$$\mathrm{cov}(\hat{\mathbf{w}}_{MVDR}) = \frac{1/(\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g})}{(L-N+1)} \times \left[\mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{g}\mathbf{g}^H\mathbf{R}^{-1}/(\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g})\right]. \tag{5.22}$$

5.2.3 SCB GSC Weightings

Recall the associated weightings of the clairvoyant GSC beamformer eq(2.22) and their SCB counterparts eq(5.3). The goal is clearly to specify the pdf of the SCB adaptive unconstrained portion of these weightings given by $\hat{\mathbf{\Omega}}_{GSC}$, from which the pdfs of $\widehat{\mathbf{W}}_{GSC}$ and $\hat{\mathbf{w}}_{gsc}$ follow. Recall from the discussion in Chapter 4 that the SCM can be written equal in distribution such that $\mathbf{D} \overset{d}{=} \mathbf{R}^{1/2}\mathbf{Q}_0^H\mathbf{Z}\mathbf{Z}^H\mathbf{Q}_0\mathbf{R}^{1/2}$. The unitary matrix $\mathbf{Q}_0$ has the property $\mathbf{Q}_0^H\mathbf{Q}_0 = \mathbf{Q}_0\mathbf{Q}_0^H = \mathbf{I}_N$, and is independent of the white data matrix $\mathbf{Z}$. Consider a partition of the white data matrix $\mathbf{Z}$ similar in fashion to eq(4.21)

$$\mathbf{Z} = \begin{bmatrix} \left[\mathbf{Z}_3^H\right]_{(N-E)\times L} \\ \left[\mathbf{Z}_4^H\right]_{E\times L} \end{bmatrix}. \tag{5.23}$$

Because $\mathbf{Z}$ is a normal matrix with columns i.i.d. as $\mathcal{CN}_N(\mathbf{0}, \mathbf{I}_N)$, $\mathbf{Z}_3$ and $\mathbf{Z}_4$ are statistically independent. Using this partition and choosing the unitary matrix $\mathbf{Q}_0$ such that

$$\mathbf{Q}_0\mathbf{R}^{1/2}\mathbf{G}_\perp = \begin{bmatrix} \nabla_{(N-E)\times(N-E)} \\ \mathbf{0}_{E\times(N-E)} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{N-E} \\ \mathbf{0}_{E\times(N-E)} \end{bmatrix} \nabla, \tag{5.24}$$

note that the adaptive weight matrix can be written equal in distribution to the following

$$\begin{aligned} \hat{\mathbf{\Omega}}_{GSC} &\overset{d}{=} (\mathbf{G}_\perp^H\mathbf{R}^{1/2}\mathbf{Q}_0^H\mathbf{Z}\mathbf{Z}^H\mathbf{Q}_0\mathbf{R}^{1/2}\mathbf{G}_\perp)^{-1}\mathbf{G}_\perp^H\mathbf{R}^{1/2}\mathbf{Q}_0^H\mathbf{Z}\mathbf{Z}^H\mathbf{Q}_0\mathbf{R}^{1/2} \\ &= (\nabla^H\mathbf{Z}_3^H\mathbf{Z}_3\nabla)^{-1}\nabla^H\left[\,\mathbf{Z}_3^H\mathbf{Z}_3 \,\middle|\, \mathbf{Z}_3^H\mathbf{Z}_4\right]\mathbf{Q}_0\mathbf{R}^{1/2} \\ &= \nabla^{-1}\left[\,\mathbf{I}_{N-E} \,\middle|\, (\mathbf{Z}_3^H\mathbf{Z}_3)^{-1}\mathbf{Z}_3^H\mathbf{Z}_4\right]\mathbf{Q}_0\mathbf{R}^{1/2}. \end{aligned} \tag{5.25}$$

Let

$$\mathcal{T}_{GSC} = (\mathbf{Z}_3^H\mathbf{Z}_3)^{-1}\mathbf{Z}_3^H\mathbf{Z}_4 \tag{5.26}$$

and partition $\mathbf{R}^{1/2}\mathbf{Q}_0^H$ such that

$$\mathbf{R}^{1/2}\mathbf{Q}_0^H = \left[\,\left[\tilde{\mathbf{\Omega}}\right]_{N\times(N-E)} \,\middle|\, \left[\widetilde{\mathbf{G}}_{GSC}\right]_{N\times E}\,\right]. \tag{5.27}$$

Note that the choice of $\mathbf{Q}_0$ implied by eq(5.24) suggests that

$$\begin{aligned} \left[\,\mathbf{I}_{N-E} \,\middle|\, \mathbf{0}_{(N-E)\times E}\right] &= \nabla^{-H}\mathbf{G}_\perp^H\mathbf{R}^{1/2}\mathbf{Q}_0^H \qquad \text{and that} \\ \mathbf{G}_\perp^H\,\mathbf{R}\,\mathbf{G}_\perp &= \mathbf{G}_\perp^H\,\mathbf{R}^{1/2}\mathbf{Q}_0^H\mathbf{Q}_0\mathbf{R}^{1/2}\,\mathbf{G}_\perp = \nabla^H\nabla. \end{aligned} \tag{5.28}$$

Using eq(5.28) the adaptive weight matrix can be written as

$$\begin{aligned} \hat{\mathbf{\Omega}}_{GSC} &\overset{d}{=} \nabla^{-1}\left[\,\mathbf{I}_{N-E} \,\middle|\, \mathbf{0}_{(N-E)\times E}\right]\mathbf{Q}_0\mathbf{R}^{1/2} + \nabla^{-1}\left[\,\mathbf{0}_{N-E} \,|\, \mathcal{T}_{GSC}\right]\mathbf{Q}_0\mathbf{R}^{1/2} \\ &= \mathbf{\Omega}_{GSC} + \nabla^{-1}\left[\,\mathbf{0}_{N-E} \,|\, \mathcal{T}_{GSC}\right]\mathbf{Q}_0\mathbf{R}^{1/2} \end{aligned} \tag{5.29}$$

Recalling eq(5.27) we conclude with the representation

$$\boxed{\widehat{\boldsymbol{\Omega}}_{GSC} \stackrel{d}{=} \boldsymbol{\Omega}_{GSC} + \nabla^{-1}\, \boldsymbol{\mathcal{T}}_{GSC}\, \widetilde{\mathbf{G}}_{GSC}^{H}} \tag{5.30}$$

where the pdf for $\mathcal{T}_{GSC}$ is also given by eq(5.14). To see this, simply note that the mathematical form and probabilistic assumptions made for $\mathcal{T}_{GSC}$ are exactly the same as those made for $\boldsymbol{\mathcal{T}}$. Thus, $\boldsymbol{\mathcal{T}}_{GSC} \stackrel{d}{=} \boldsymbol{\mathcal{T}}$. Similarly, it can be deduced that $\widehat{\boldsymbol{\Omega}}_{GSC}$ is an unbiased estimate of $\boldsymbol{\Omega}_{GSC}$ since $E\{\mathcal{T}_{GSC}\} = E\{\mathcal{T}\} = \mathbf{0}$. The latter expectation was proven in the previous section. Consequently, $\widehat{\mathbf{W}}_{GSC}$ and $\hat{\mathbf{w}}_{gsc}$ are likewise unbiased estimates.

Substituting eq(5.27) into eq(5.24) implies that $\tilde{\boldsymbol{\Omega}}^{H}\mathbf{G}_{\perp} = \nabla$ and $\widetilde{\mathbf{G}}_{GSC}^{H}\mathbf{G}_{\perp} = \mathbf{0}$. Since $\mathbf{Q}_0\mathbf{R}^{1/2}$ has full rank, it follows that $\text{span}(\tilde{\boldsymbol{\Omega}}) = \text{span}(\boldsymbol{\Omega}_{GSC}^{H})$ and that $\text{span}(\widetilde{\mathbf{G}}_{GSC}) = \text{span}(\mathbf{G})$.

Eq(5.30) together with eq(5.3) suggest the following stochastic representations for the GSC weightings

$$\boxed{\begin{aligned} \widehat{\mathbf{W}}_{GSC} &\stackrel{d}{=} \mathbf{W}_{GSC} - \mathbf{G}_{\perp}\nabla^{-1}\boldsymbol{\mathcal{T}}_{GSC}\, \widetilde{\mathbf{G}}_{GSC}^{H}\boldsymbol{\Phi} \\ \hat{\mathbf{w}}_{gsc} &\stackrel{d}{=} \mathbf{w}_{gsc} - \mathbf{G}_{\perp}\nabla^{-1}\boldsymbol{\mathcal{T}}_{GSC}\, \widetilde{\mathbf{G}}_{GSC}^{H}\boldsymbol{\Phi}\mathbf{f} \\ &\stackrel{d}{=} \mathbf{w}_{gsc} - \mathbf{G}_{\perp}\nabla^{-1}\tilde{\mathbf{t}}_{GSC}\sqrt{\mathbf{f}^{H}\boldsymbol{\Phi}^{H}\mathfrak{P}(\mathbf{G}_{\perp}|\mathbf{R})\boldsymbol{\Phi}\mathbf{f}}. \end{aligned}} \tag{5.31}$$

These representations with $\mathcal{T}_{GSC}$ multivariate t-distributed as in eq(5.14) or $\tilde{\mathbf{t}}_{GSC} \stackrel{d}{=} \tilde{\mathbf{t}}_{LCMV}$ completely specify the pdfs for the SCB GSC weightings. Since it is known with certainty that

$$\mathbf{G}_{\perp}^{H}\widehat{\boldsymbol{\Omega}}_{GSC}^{H} = \mathbf{I}_{N-E}, \quad \mathbf{G}^{H}\widehat{\mathbf{W}}_{GSC} = \mathbf{I}_{E}, \quad \text{and that } \mathbf{G}^{H}\hat{\mathbf{w}}_{gsc} = \mathbf{f} \tag{5.32}$$

independent of variation in $\widehat{\mathbf{R}}$, the explicit pdfs for these weightings are singular. The analogous GSC interpretation of the GSC SCB weightings' stochastic representations are shown in fig(5.3)–fig(5.5). One minor difference in fig(5.3) is the swapping of the roles played by $\mathbf{G}_{\perp}$ and $\widetilde{\mathbf{G}}_{GSC}$ in the dashed line block representing $\widehat{\boldsymbol{\Omega}}_{GSC}$. The variable $\mathcal{T}_{GSC}$ again represents a source of undesired random sidelobe variation.

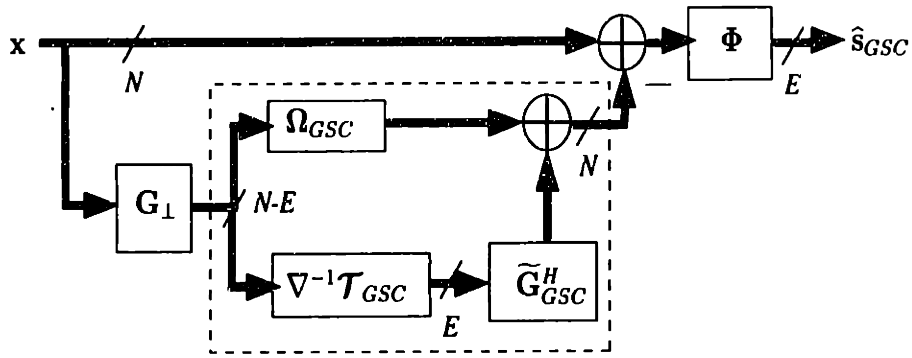


Figure 5.3: GSC Interpretation of Stochastic Representation of $\widehat{\mathbf{W}}_{GSC}$

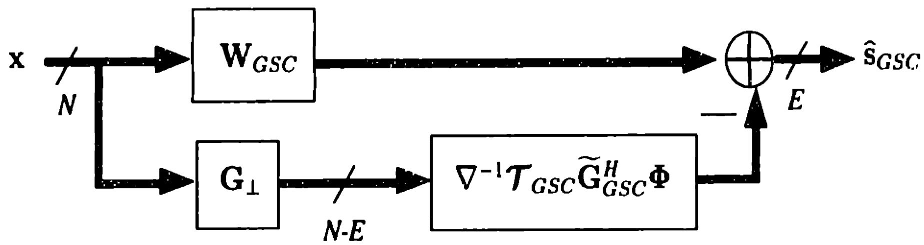


Figure 5.4: Alternate GSC Interpretation of $\widehat{\mathbf{W}}_{GSC}$

The derived stochastic representations for the adaptive weight matrices can be used to show the following covariances

$$\begin{aligned} \text{cov}(\widehat{\boldsymbol{\Omega}}_{GSC}) &= \mathfrak{P}(\mathbf{G}_\perp|\mathbf{R})^T \otimes \frac{(\mathbf{G}_\perp^H\mathbf{R}\mathbf{G}_\perp)^{-1}}{(L-N+E)} \\ \text{cov}(\widehat{\mathbf{W}}_{GSC}) &= \left[\boldsymbol{\Phi}^H\mathfrak{P}(\mathbf{G}_\perp|\mathbf{R})\boldsymbol{\Phi}\right]^T \otimes \frac{\mathbf{G}_\perp(\mathbf{G}_\perp^H\mathbf{R}\mathbf{G}_\perp)^{-1}\mathbf{G}_\perp^H}{(L-N+E)} \\ \text{cov}(\widehat{\mathbf{w}}_{gsc}) &= \mathbf{f}^H\boldsymbol{\Phi}^H\mathfrak{P}(\mathbf{G}_\perp|\mathbf{R})\boldsymbol{\Phi}\mathbf{f} \times \frac{\mathbf{G}_\perp(\mathbf{G}_\perp^H\mathbf{R}\mathbf{G}_\perp)^{-1}\mathbf{G}_\perp^H}{(L-N+E)}. \end{aligned} \tag{5.33}$$

5.2.4 LCMV and GSC Equivalence

As mentioned earlier the LCMV and GSC beamformers implement exactly the same overall weightings [37, 17]. This equivalence is completely algebraic and therefore includes their SCB counterparts, *i.e.* $\widehat{\mathbf{W}}_{ML} = \widehat{\mathbf{W}}_{GSC}$ and $\widehat{\mathbf{w}}_{LCMV} = \widehat{\mathbf{w}}_{gsc}$. Comparing eq(5.20) to eq(5.33)

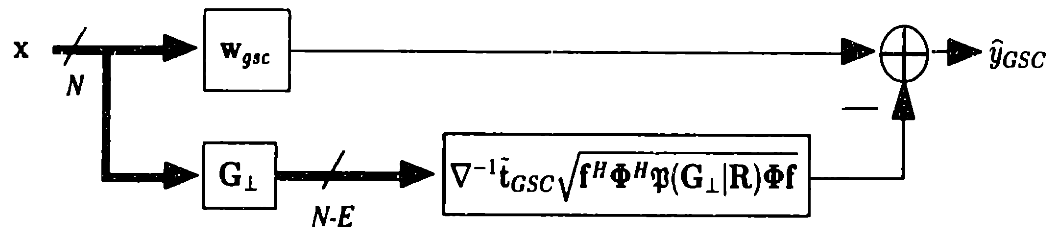


Figure 5.5: GSC Interpretation of Stochastic Representation of $\hat{\mathbf{w}}_{gsc}$

note that

$$\begin{aligned} \mathfrak{P}(\mathbf{G}|\mathbf{R}^{-1}) &= \mathbf{G}_\perp(\mathbf{G}_\perp^H\mathbf{R}\mathbf{G}_\perp)^{-1}\mathbf{G}_\perp^H \\ (\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1} &= \mathbf{\Phi}^H\mathfrak{P}(\mathbf{G}_\perp|\mathbf{R})\mathbf{\Phi}. \end{aligned} \tag{5.34}$$

Both of these equalities are versions of Khatri's Lemma 1 [17], and one can use them to immediately conclude/infer parallel results for the GSC from those derived for the LCMV beamformer.

5.3 PDFs of Adaptive SCB Beam Responses

In this section we derive the pdfs for the beam responses which result from the SCB weightings. In Section 5.4 the pdf for the beamformer output will be derived, which also accounts for the random variations in the primary data observation $\mathbf{x}$ from which the signal is estimated.

5.3.1 ML/LCMV Beamformer

The clairvoyant LCMV vector beam response is $\mathbf{b}_{LCMV}(\theta,\omega) = \mathbf{W}_{ML}^H\mathbf{d}_\theta(\omega)$. Recall from eq(5.13) that the SCB vector beam response can be written as

$$\hat{\mathbf{b}}_{LCMV}(\theta,\omega) = \widehat{\mathbf{W}}_{ML}^H\mathbf{d}_\theta(\omega) \stackrel{d}{=} \mathbf{W}_{ML}^H\mathbf{d}_\theta(\omega) - \Delta^{-1}\mathbf{Z}_1^H\mathbf{Z}_2(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{G}_\perp^H\mathbf{d}_\theta(\omega). \tag{5.35}$$

Let $\mathbf{p}_\theta(\omega) = \mathbf{G}_\perp^H\mathbf{d}_\theta(\omega)$ and $\mathbf{q} = \mathbf{Z}_2(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{p}_\theta(\omega)$. Note first that $\mathbf{Z}_1^H\mathbf{q} \stackrel{d}{=} \mathbf{z}_0 \times ||\mathbf{q}||$ with $\mathbf{z}_0 \sim \mathcal{CN}_E(\mathbf{0},\mathbf{I}_E)$ and $||\mathbf{q}||^2 = \mathbf{p}_\theta^H(\omega)(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{p}_\theta(\omega)$, and second that if we condition on $\mathbf{Z}_2$ then the a posteriori distribution of the adaptive beam response is complex normal with mean $\mathbf{W}_{ML}^{H}\mathbf{d}_\theta(\omega)$ and covariance

$$\mathrm{cov}[\widehat{\mathbf{b}}_{LCMV}(\theta,\omega)|\mathbf{Z}_2] = (\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1} \times \mathbf{p}_\theta^H(\omega)(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{p}_\theta(\omega). \tag{5.36}$$

Note by the GWPT Theorem 3.6.1, or the well known Capon and Goodman result [33] that

$$\mathbf{p}_\theta^H(\omega)(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{p}_\theta(\omega) \stackrel{d}{=} ||\mathbf{p}_\theta(\omega)||^2/\chi^2_{L-N+E+1}, \tag{5.37}$$

and from Appendix C it is shown that $||\mathbf{p}_\theta(\omega)||^2 = \mathbf{d}_\theta^H(\omega)\ \mathfrak{P}(\mathbf{G}|\mathbf{R}^{-1})\mathbf{d}_\theta(\omega)$. These can be used to remove the conditioning on $\mathbf{Z}_2$ to obtain the exact covariance for the SCB LCMV vector beam response

$$\boxed{\mathrm{cov}[\widehat{\mathbf{b}}_{LCMV}(\theta,\omega)] = \frac{(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}}{(L-N+E)} \times ||\mathbf{p}_\theta(\omega)||^2.} \tag{5.38}$$

In addition we now have the following conditional distribution

$$\widehat{\mathbf{b}}_{LCMV}(\theta,\omega) \sim \mathcal{CN}_E\left[\mathbf{W}_{ML}^H\mathbf{d}_\theta(\omega), (\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1} \times \frac{||\mathbf{p}_\theta(\omega)||^2}{\chi^2_{L-N+E+1}}\right] = P_{\widehat{\mathbf{b}}_{LCMV}(\theta,\omega)|\chi^2_{L-N+E+1}} \tag{5.39}$$

Consequently, to obtain the marginal distribution for the SCB LCMV vector beam response, one simply integrates over the complex chi-squared variable $\chi^2_{L-N+E+1}$, that is

$$P_{\widehat{\mathbf{b}}_{LCMV}(\theta,\omega)}(\widehat{\mathbf{b}}_0) = \int_0^\infty P_{\chi^2_{L-N+E+1}}(a) P_{\widehat{\mathbf{b}}_{LCMV}(\theta,\omega)|\chi^2_{L-N+E+1}}\left(\widehat{\mathbf{b}}_0|a\right) da. \tag{5.40}$$

Using the integral identity obtained from the fact that eq(3.68) must integrate to unity it is shown that the SCB beam response has the following generalized complex multivariate t-distribution

$$\boxed{\widehat{\mathbf{b}}_{LCMV}(\theta,\omega) \sim \mathcal{CT}_{E\times 1}\left[\mathbf{W}_{ML}^H\mathbf{d}_\theta(\omega),\ L-N+E+1,\ (\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1},\ ||\mathbf{p}_\theta(\omega)||^{-2}\right].} \tag{5.41}$$

The pdf for the SCB LCMV scalar beam response $\widehat{b}_{LCMV}(\theta,\omega) = \widehat{\mathbf{w}}_{LCMV}^H\mathbf{d}(\theta,\omega)$ is

$$\boxed{\widehat{b}_{LCMV}(\theta,\omega) \sim \mathcal{CT}_{1\times 1}\left[\mathbf{w}_{LCMV}^H\mathbf{d}_\theta(\omega),\ L-N+E+1,\ \mathbf{f}^H(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}\mathbf{f},\ ||\mathbf{p}_\theta(\omega)||^{-2}\right].}$$

(5.42)

Note that for the special case of the SCB MVDR beamformer where $E = 1$ and $\mathrm{f} = 1$ eq(5.42) simplifies to show that the pdf of the MVDR beam response $\hat{b}_{MVDR}(\theta,\omega) = \hat{\mathbf{w}}_{MVDR}^H \mathbf{d}_\theta(\omega)$ is given by

$$P_{\hat{b}_{MVDR}(\theta,\omega)}\left(\hat{b}_0\right) = \left[\frac{L-N+2}{\pi}\right] \times ||\mathbf{p}_\theta(\omega)||^{2(L-N+3)-2} \times (\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}) \times \left[|\hat{b}_0 - \mathbf{w}_{MVDR}^H \mathbf{d}_\theta(\omega)|^2(\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}) + ||\mathbf{p}_\theta(\omega)||^2\right]^{-(L-N+3)}. \tag{5.43}$$

If a standard basis look direction $\mathbf{e}_i = \mathbf{d}_\theta(\omega)$ is chosen, then this equation reduces to eq(25) of Steinhardt [28] yielding the marginal pdfs for the MVDR weight vector as a special case of a more general statement.

5.3.2 Generalized Sidelobe Cancellor

Exploiting the equivalence of the LCMV and GSC beamformers we note that letting $\tilde{\mathbf{p}}_\theta(\omega) = \nabla^{-H}\mathbf{G}_\perp^H \mathbf{d}_\theta(\omega)$, the SCB vector beam response is conditionally complex normal such that

$$\hat{\mathbf{b}}_{GSC}(\theta,\omega) \sim \mathcal{CN}_E\left[\mathbf{W}_{GSC}^H \mathbf{d}_\theta(\omega), \mathbf{\Phi}^H \mathfrak{P}(\mathbf{G}_\perp|\mathbf{R})\mathbf{\Phi} \times \frac{||\tilde{\mathbf{p}}_\theta(\omega)||^2}{\chi^2_{L-N+E+1}}\right] = P_{\hat{\mathbf{b}}_{GSC}(\theta,\omega)|\chi^2_{L-N+E+1}} \tag{5.44}$$

where $||\tilde{\mathbf{p}}_\theta(\omega)||^2 = \mathbf{d}_\theta^H(\omega)\mathbf{G}_\perp(\mathbf{G}_\perp^H\mathbf{R}\mathbf{G}_\perp)^{-1}\mathbf{G}_\perp^H\mathbf{d}_\theta(\omega) = ||\mathbf{p}_\theta(\omega)||^2$ (second equality follows from eq(5.34)). The exact covariance is given by

$$\mathrm{cov}[\hat{\mathbf{b}}_{GSC}(\theta,\omega)] = \frac{||\tilde{\mathbf{p}}_\theta(\omega)||^2}{(L-N+E)} \times \mathbf{\Phi}^H\mathfrak{P}(\mathbf{G}_\perp|\mathbf{R})\mathbf{\Phi}, \tag{5.45}$$

and the exact distribution by

$$\hat{\mathbf{b}}_{GSC}(\theta,\omega) \sim \mathcal{CT}_{E\times 1}\left[\mathbf{W}_{GSC}^H\mathbf{d}_\theta(\omega),\ L-N+E+1,\ \mathbf{\Phi}^H\mathfrak{P}(\mathbf{G}_\perp|\mathbf{R})\mathbf{\Phi},\ ||\tilde{\mathbf{p}}_\theta(\omega)||^{-2}\right].$$

The pdf for the SCB GSC scalar beam response $\widehat{b}_{GSC}(\theta,\omega) = \widehat{\mathbf{w}}_{gsc}^H \mathbf{d}_\theta(\omega)$ is

$$\widehat{b}_{GSC}(\theta,\omega) \sim \mathcal{CT}_{1\times 1}\left[\mathbf{w}_{gsc}^H \mathbf{d}_\theta(\omega),\ L-N+E+1,\ \mathbf{f}^H \mathbf{\Phi}^H \mathfrak{P}(\mathbf{G}_\perp|\mathbf{R})\mathbf{\Phi}\mathbf{f},\ ||\tilde{\mathbf{p}}_\theta(\omega)||^{-2}\right].$$

Remarks

It is interesting to note that each SCB beam response is likewise a scale invariant function of the secondary data matrix $\mathbf{X}$. This follows from the invariance of each $\widehat{\mathbf{W}}$. Hence, the generalized t-distributions for these beam responses do not depend on $g(\cdot)$. In Section 5.5 we shall obtain confidence regions for these random spatial filter responses. That their statistics are MEC invariant implies, for example, that from a spatial filtering point of view the "ability" of the SCB spatial filter to adaptively suppress hostile interferences is invariant to $g(\cdot)$. By the word "ability" we mean the likelihood that the adaptive response is within some standard deviation of the optimal clairvoyant response. This type reasoning suggests a plausible theory for the observed robustness of these beamformers to non-Gaussian environments [27].

5.4 PDFs of Beamformer Outputs

In this section we specify the pdfs for the resulting beamformer outputs, via the use of the derived distributions for the SCB weightings. The pdf for the output of the SCB ML/LCMV beamformer is known to be given by the confluent hypergeometric function known as *Kummer's* function [14, 22] when $\mathbf{x}$ is Gaussian. To close the loop between the results of this chapter and those of [14, 18, 16, 22] we use the derived distributions for the SCB ML/LCMV/GSC weightings to arrive at the same result, and in fact to generalize over MECs.

5.4.1 ML Signal Estimate/LCMV Beamformer Output

First note that if we call the primary data vector $\mathbf{x}$ some statistic $\mathbf{F}_1(\mathbf{x})$, and call $\widehat{\mathbf{W}}_{ML}$ some other statistic $\mathbf{F}_2(\mathbf{X})$, then by Theorem 3.5.4 it follows that they are completely independent. Thus, without loss of generality over complex MECs we assume in this section that $\mathbf{x}$ and $\mathbf{X}$ are independent and distributed as

$$\mathbf{x} \sim \mathcal{CEC}_N\left[\mathbf{Gs}, \mathbf{R}, \phi(\cdot); g_{N:N(L+1)}(\cdot)\right] \qquad \mathbf{X} \sim \mathcal{CN}_{NL}\left[\mathbf{0}, \mathbf{I}_L \otimes \mathbf{R}\right]. \tag{5.46}$$

Consider that the SCB ML signal estimate is given by

$$\hat{\mathbf{s}}_{ML} = \widehat{\mathbf{W}}_{ML}^H \mathbf{x} = \mathbf{s} + \widehat{\mathbf{W}}_{ML}^H \mathbf{n}. \tag{5.47}$$

Conditioning on $\widehat{\mathbf{W}}_{ML}$, note that by Proposition 3.5.4.3

$$\hat{\mathbf{s}}_{ML} \sim \mathcal{CEC}_E\left[\mathbf{s}, \widehat{\mathbf{W}}_{ML}^H \mathbf{R} \widehat{\mathbf{W}}_{ML}, \phi(\cdot); g_{E:N(L+1)}(\cdot)\right] = P_{\hat{\mathbf{s}}_{ML}|\widehat{\mathbf{W}}_{ML}}. \tag{5.48}$$

This a posteriori distribution is completely parameterized by $g(\cdot)$, the constant $\mathbf{s}$, and the random quantity $\widehat{\mathbf{W}}_{ML}^H \mathbf{R} \widehat{\mathbf{W}}_{ML}$. Hence, we seek the distribution of the latter. Note from eq(5.9) and minor algebra that

$$\widehat{\mathbf{W}}_{ML}^H \mathbf{R} \widehat{\mathbf{W}}_{ML} \;\stackrel{d}{=}\; \Delta^{-1}\left(\mathbf{I}_E + \mathcal{T}^H \mathcal{T}\right)\Delta^{-H}. \tag{5.49}$$

The exact a priori covariance is found by the following expectation

$$\operatorname{cov}(\hat{\mathbf{s}}_{ML}) = \Delta^{-1}\left(\mathbf{I}_E + E\left\{\mathcal{T}^H \mathcal{T}\right\}\right)\Delta^{-H} \cdot \kappa \tag{5.50}$$

where we recall that $\kappa = E\{\rho\}/N(L+1)$ (see eq(3.114)). It will be proven that $E\left\{\mathcal{T}^H \mathcal{T}\right\} = \mathbf{I}_E \cdot (N-E)/(L-N+E)$, and thus

$$\boxed{\operatorname{cov}(\hat{\mathbf{s}}_{ML}) = \kappa \cdot \left[\frac{L}{L-N+E}\right] \times (\mathbf{G}^H \mathbf{R}^{-1} \mathbf{G})^{-1}.} \tag{5.51}$$

Defining the variable $\mathbf{\Lambda}$ by the relation

$$\left(\widehat{\mathbf{W}}_{ML}^H \mathbf{R} \widehat{\mathbf{W}}_{ML}\right)^{-1} \stackrel{d}{=} \Delta^H \left(\mathbf{I}_E + \mathcal{T}^H \mathcal{T}\right)^{-1} \Delta = \Delta^H \mathbf{\Lambda} \Delta, \tag{5.52}$$

and invoking Proposition F.0.0.2 of Appendix F, the pdf of $\mathbf{\Lambda}$ is shown to be given by the following E-variate complex beta distribution

$$P_{\Lambda}(\Lambda_0) = \left[\tilde{\beta}_E(L-N+2E, N-E)\right]^{-1} |\Lambda_0|^{L-N+E}|\mathbf{I}-\Lambda_0|^{N-2E}. \tag{5.53}$$

The beta function is defined by the ratio of complex multivariate gamma functions $\tilde{\beta}_E(a,b) = \tilde{\Gamma}_E(a)\tilde{\Gamma}_E(b)\ /\tilde{\Gamma}_E(a+b)$; in addition the following statement of positive definiteness is true $\mathbf{0} < \Lambda_0 < \mathbf{I}$, *i.e.* $0 < \mathbf{q}^H\Lambda_0\mathbf{q} < \mathbf{q}^H\mathbf{q}$ for all complex $E \times 1$ vectors $\mathbf{q} \neq \mathbf{0}$. The variable Λ appears as statistic (4) in Lemma 2 of Khatri and Rao [18]. There it was shown to be the multiple signal matrix generalization of the normalized SNR beta statistic found in Reed *et al* [16]. As in [14] note that the marginal pdf for the ML signal estimate is given implicitly by the following integral

$$P_{\widehat{\mathbf{s}}_{ML}}(\widehat{\mathbf{s}}_0) = \int_0^I P_{\Lambda}(\Lambda_0)P_{\widehat{\mathbf{s}}_{ML}|\Lambda}(\widehat{\mathbf{s}}_0|\Lambda_0)(d\Lambda_0). \tag{5.54}$$

In general this integral can be difficult to evaluate in any sense for arbitrary complex MECs. We can, however, simplify things rather significantly.

Note from Proposition 3.5.4.3 that the conditional c.f. of $\widehat{\mathbf{s}}_{ML}$ given $\widehat{\mathbf{W}}_{ML}$ is

$$e^{j\mathrm{Re}(\mathbf{t}^H\mathbf{s})} \cdot \phi\left(\mathbf{t}^H\widehat{\mathbf{W}}_{ML}^H\mathbf{R}\widehat{\mathbf{W}}_{ML}\mathbf{t}/2\right). \tag{5.55}$$

Let $\mathbf{t}_0 = \Delta^{-H}\mathbf{t}$ and note that

$$\mathbf{t}^H\widehat{\mathbf{W}}_{ML}^H\mathbf{R}\widehat{\mathbf{W}}_{ML}\mathbf{t} \stackrel{d}{=} \mathbf{t}_0^H\mathbf{t}_0 + \mathbf{t}_0^H(\mathcal{T}^H\mathcal{T})\mathbf{t}_0. \tag{5.56}$$

That $\mathbf{Z}_1\mathbf{t}_0 \stackrel{d}{=} \mathbf{z}_0 \cdot ||\mathbf{t}_0||$, where $\mathbf{z}_0 \sim \mathcal{CN}_L(\mathbf{0}, \mathbf{I}_L)$ implies that $\mathcal{T}\mathbf{t}_0 = (\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{Z}_2^H\mathbf{Z}_1\mathbf{t}_0 \stackrel{d}{=} ||\mathbf{t}_0||(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1/2}\mathbf{z}_B$ where $\mathbf{z}_B \sim \mathcal{CN}_{N-E}(\mathbf{0}, \mathbf{I}_{N-E})$. Hence, $\mathbf{t}_0^H(\mathcal{T}^H\mathcal{T})\mathbf{t}_0 \stackrel{d}{=} ||\mathbf{t}_0||^2\mathbf{z}_B^H(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1} \times\mathbf{z}_B \stackrel{d}{=} ||\mathbf{t}_0||^2\chi^2_{N-E}/\chi^2_K \stackrel{d}{=} ||\mathbf{t}_0||^2 \cdot F_{N-E,K}$, where the last two equalities in distribution follow from the discussion succeeding eq(4.61) and the propositions of Section 3.5.2. That $\mathbf{t}_0^H(\mathcal{T}^H\mathcal{T})\mathbf{t}_0 \stackrel{d}{=} ||\mathbf{t}_0||^2\chi^2_{N-E}/\chi^2_K$ for all $\mathbf{t}_0$ proves that $E\{\mathcal{T}^H\mathcal{T}\} = \mathbf{I}_E \cdot E\{\chi^2_{N-E}\} \cdot E\{1/\chi^2_K\} = \mathbf{I}_E(N-E)/(K-1)$. Recall also from Section 3.5.2 that $(1+F_{N-E,K})^{-1} \stackrel{d}{=} \beta_{K,N-E}$. Ergo, the conditional c.f. and concomitant pdf can be written

$$\begin{gathered} e^{j\mathrm{Re}(\mathbf{t}^H\mathbf{s})} \cdot \phi\left\{\mathbf{t}^H[(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})/(1+F_{N-E,K})]^{-1}\mathbf{t}/2\right\} \stackrel{\mathcal{F}}{\longleftrightarrow} \\ |\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G}| \cdot \beta_{K,N-E}^E \cdot g_{E:N(L+1)}\left[\, z(\widehat{\mathbf{s}}_{ML}) \cdot \beta_{K,N-E}\,\right] = P_{\widehat{\mathbf{s}}_{ML}|\beta} \end{gathered} \tag{5.57}$$

where the real scalar argument is the quadratic

$$z(\widehat{\mathbf{s}}_{ML}) = ||(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{1/2}(\widehat{\mathbf{s}}_{ML} - \mathbf{s})||^2 = (\widehat{\mathbf{s}}_{ML} - \mathbf{s})^H\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G}(\widehat{\mathbf{s}}_{ML} - \mathbf{s}). \tag{5.58}$$

Thus, the a priori density is given by the much simpler integral

$$\boxed{P_{\widehat{\mathbf{s}}_{ML}} = |\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G}| \cdot \int_0^1 P_\beta \cdot \beta^E \cdot g_{E:N(L+1)}\left[\, z(\widehat{\mathbf{s}}_{ML}) \cdot \beta \,\right] d\beta} \tag{5.59}$$

where P_β is still given by eq(4.62).

The pdf for the output of the SCB LCMV beamformer $\widehat{y}_{LCMV} = \widehat{\mathbf{w}}^H_{LCMV}\mathbf{x} = \mathbf{f}^H\widehat{\mathbf{s}}_{ML}$ follows directly from Proposition 3.5.4.3 and is similarly given by

$$\boxed{\begin{array}{c} P_{\widehat{y}_{LCMV}} = [\mathbf{f}^H(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}\mathbf{f}]^{-1} \\ \times \int_0^1 P_\beta \cdot \beta^E \cdot g_{1:N(L+1)}\left[|\widehat{y}_{LCMV} - \mathbf{f}^H\mathbf{s}|^2\beta/\mathbf{f}^H(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}\mathbf{f}\,\right] d\beta. \end{array}} \tag{5.60}$$

The SCB ML signal estimate $\widehat{\mathbf{s}}_{ML} = \widehat{\mathbf{W}}^H_{ML}\mathbf{x}$ and SCB LCMV beamformer output $\widehat{y}_{LCMV} = \widehat{\mathbf{w}}^H_{LCMV}\mathbf{x}$ respectively can be written equal in distribution to:

$$\boxed{\begin{array}{lcl} \widehat{\mathbf{s}}_{ML} & \stackrel{d}{=} & \mathbf{s} + (\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1/2}\,\mathbf{z}_g/\sqrt{\beta_{K,N-E}} \\ \widehat{y}_{LCMV} & \stackrel{d}{=} & \mathbf{f}^H\mathbf{s} + Z_g\sqrt{\mathbf{f}^H(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}\mathbf{f}/\beta_{K,N-E}} \end{array}} \tag{5.61}$$

where $\mathbf{z}_g \sim \mathcal{CEC}_E[\mathbf{0}, \mathbf{I}_E, g_{E:N(L+1)}(\cdot)]$ and $Z_g \sim \mathcal{CEC}_1[0, 1, g_{1:N(L+1)}(\cdot)]$.

For the complex Gaussian case recall that $g_{E:N(L+1)}(r) = \pi^{-E}e^{-r}$. Eq(5.59) leads to the exact pdf given in [14]; namely the confluent hypergeometric function [73]

$$P_{\widehat{\mathbf{s}}_{ML}} = \frac{|\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G}|}{\pi^E} \times \left[\frac{\tilde{\beta}_E(L-N+2E+1, N-E)}{\tilde{\beta}_E(L-N+2E, N-E)}\right] \times {}_1F_1\left[L-N+2E+1; L+E+1; -z(\widehat{\mathbf{s}}_{ML})\right]. \tag{5.62}$$

5.4.2 Generalized Sidelobe Cancellor Output

The signal estimate given by the SCB GSC can be written as

$$\widehat{\mathbf{s}}_{GSC} = \widehat{\mathbf{W}}^H_{GSC}\mathbf{x} = \mathbf{s} + \widehat{\mathbf{W}}^H_{GSC}\mathbf{n}. \tag{5.63}$$

Exploiting the LCMV GSC equivalence, the GSC analogous results follow immediately. The covariance of the GSC signal estimate is given by

$$\boxed{\operatorname{cov}(\widehat{\mathbf{s}}_{GSC}) = \kappa \cdot \left[\frac{L}{L-N+E}\right] \times \mathbf{\Phi}^H \mathfrak{P}(\mathbf{G}_\perp|\mathbf{R})\mathbf{\Phi}.} \tag{5.64}$$

Let a matrix variate $\mathbf{\Lambda}_{GSC}$ be distributed according to the $E-$variate complex beta distribution given in eq(5.53). It is true that

$$\widehat{\mathbf{s}}_{GSC} \sim \mathcal{CEC}_E\left[\mathbf{s}, \mathbf{\Phi}^H \widetilde{\mathbf{G}}_{GSC}\ \ \mathbf{\Lambda}_{GSC}^{-1}\ \ \widetilde{\mathbf{G}}_{GSC}^H \mathbf{\Phi}, g_{E:N(L+1)}(\cdot)\right] = P_{\widehat{\mathbf{s}}_{GSC}|\mathbf{\Lambda}_{GSC}} \tag{5.65}$$

from the LCMV GSC equivalence [1]. The pdfs for the SCB GSC signal estimate $\widehat{\mathbf{s}}_{GSC}$ and beamformer output $\widehat{y}_{GSC} = \widehat{\mathbf{w}}_{GSC}^H \mathbf{x}$ are implicit from the previous section.

5.5 Confidence Regions

5.5.1 Beam Responses

It is useful to know the likelihood of the event that the SCB vector beam response is within, say R_{CR} times a matrix standard deviation of the the optimal clairvoyant response. Quantifying the probability of this event essentially consists of specifying confidence regions. To specify confidence regions for the SCB vector beam responses we first define the variables

$$\begin{aligned} \mathbf{t}_{LCMV} &= \frac{(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{1/2}}{||\mathbf{p}_\theta(\omega)||}\left[\widehat{\mathbf{b}}_{LCMV}(\theta,\omega) - \mathbf{W}_{ML}^H \mathbf{d}_\theta(\omega)\right] \\ \mathbf{t}_{GSC} &= \frac{(\mathbf{\Phi}^H \mathfrak{P}(\mathbf{G}_\perp|\mathbf{R})\mathbf{\Phi})^{-1/2}}{||\widetilde{\mathbf{p}}_\theta(\omega)||}[\widehat{\mathbf{b}}_{GSC}(\theta,\omega) - \mathbf{W}_{GSC}^H \mathbf{d}_\theta(\omega)]. \end{aligned} \tag{5.66}$$

The pdf for these new variables is found to be a standardized complex multivariate $t-$distribution by noting that the Jacobians for these linear transformations are $J(\widehat{\mathbf{b}}_{LCMV}(\theta,\omega) \to \mathbf{t}_{LCMV}) = |\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G}|^{-1} \cdot ||\mathbf{p}_\theta(\omega)||^{2E}$ and $J(\widehat{\mathbf{b}}_{GSC}(\theta,\omega) \to \mathbf{t}_{GSC}) = |\mathbf{\Phi}^H\mathfrak{P}(\mathbf{G}_\perp|\mathbf{R})\mathbf{\Phi}| \cdot ||\widetilde{\mathbf{p}}_\theta(\omega)||^{2E}$. Hence, by Proposition 3.5.1.1 it follows that $\mathbf{t}_{LCMV} \overset{d}{=} \mathbf{t}_{GSC} \overset{d}{=} \mathbf{t}_{BR}$ where

$$\mathbf{t}_{BR} \sim \mathcal{CT}_{E\times 1}\left[\mathbf{0}_{E\times 1},\ L-N+2E,\ \mathbf{I}_E, 1\right] = P_{\mathbf{t}_{BR}}. \tag{5.67}$$

[1] This result likewise can be derived using eq(5.25)–eq(5.28). See Appendix E.

In terms of this new variable it is desired to specify the probability of the event $||\mathbf{t}_{BR}|| < R_{CR}$. The pdf for $\rho = ||\mathbf{t}_{BR}||^2$ can be found by employing Proposition 3.3.2.1. Note that

$$\begin{aligned} E\{h(||\mathbf{t}_{BR}||^2)\} &= \int h(\mathbf{t}_0^H \mathbf{t}_0) P_{\mathbf{t}_{BR}}(\mathbf{t}_0) d\mathbf{t}_0 = \frac{\pi^E}{\tilde{\Gamma}_1(E)} \times \left[\frac{(L-N+2E)!}{(L-N+E)! \pi^E} \right] \\ &\times \int_{\rho_0 > 0} h(\rho_0) \ (1+\rho_0)^{-(L-N+2E+1)} (\rho_0)^{E-1} d\rho_0 = E\{h(\rho)\}. \end{aligned} \tag{5.68}$$

Comparing the integrand to the definition of expectation it is shown that the pdf for ρ is a central complex F-distribution, *i.e.* $\rho \sim F_{E,K}$. Thus, the likelihood of the desired event is given by the integral

$$Pr\left(||\mathbf{t}_{BR}|| < R_{CR}\right) = \frac{(L-N+2E)!}{(L-N+E)!(E-1)!} \int_0^{R_{CR}^2} \frac{\rho_0^{E-1} d\rho_0}{(1+\rho_0)^{(L-N+2E+1)}}. \tag{5.69}$$

This integral can be written in closed form as the following finite sum (see Section 3.5.2) to yield, *e.g.* the desired LCMV confidence regions

$$\boxed{\begin{aligned} &Pr\left(\left|\left| (\mathbf{G}^H \mathbf{R}^{-1} \mathbf{G})^{1/2} \left[\hat{\mathbf{b}}_{LCMV}(\theta,\omega) - \mathbf{W}_{ML}^H \mathbf{d}_\theta(\omega) \right] \right|\right| < R_{CR} \times ||\mathbf{p}_\theta(\omega)|| \right) = \\ &\frac{R_{CR}^{2E}}{(1+R_{CR}^2)^{(L-N+2E)}} \times \sum_{k=0}^{(L-N+E)} \binom{L-N+2E}{k+E} R_{CR}^{2k}. \end{aligned}}$$

The actual confidence region consists of an ellipsoid in complex Euclidean E–space $\mathfrak{C}^E$ centered on the SCB vector beam response, where the principle axes are given by the eigenvectors of the matrix $\mathbf{G}^H \mathbf{R}^{-1} \mathbf{G}$. The boundaries on these principle axes are determined by the eigenvalues in the form $R_{CR} \times ||\mathbf{p}_\theta(\omega)|| / \sqrt{\lambda_k(\mathbf{G}^H \mathbf{R}^{-1} \mathbf{G})}$, $k = 1, 2, \ldots, E$.

The confidence region for the SCB LCMV scalar beam response is given by

$$\boxed{\begin{aligned} &Pr\left(|\hat{b}_{LCMV}(\theta,\omega) - \mathbf{w}_{LCMV}^H \mathbf{d}_\theta(\omega)| < R_{CR} \, ||\mathbf{p}_\theta(\omega)|| \sqrt{\mathbf{f}^H (\mathbf{G}^H \mathbf{R}^{-1} \mathbf{G})^{-1} \mathbf{f}} \right) = \\ &\frac{R_{CR}^2}{(1+R_{CR}^2)^{(L-N+E+1)}} \times \sum_{k=0}^{(L-N+E)} \binom{L-N+E+1}{k+1} R_{CR}^{2k}. \end{aligned}} \tag{5.70}$$

Note that the likelihood of $\hat{b}_{LCMV}(\theta,\omega)$ falling within R_{CR} times a standard deviation of the clairvoyant beam response $b_{LCMV}(\theta,\omega)$ yields some fairly intuitive insights. First note that the standard deviation is given by the product of a projection term $||\mathbf{p}_\theta(\omega)||$ and the square root of the clairvoyant LCMV beamformer output power. The projection term simply indicates that the SCB beam response variation is only present where there are no constraints; namely, in the space orthogonal to $\mathbf{G}$. The closer $\mathbf{d}_\theta(\omega)$ is to a constrained direction the smaller the variation. The second term of the product indicates that the smaller the clairvoyant beamformer output power the tighter the confidence region for its SCB counterpart, as one might suspect. When $E = 1$ and f $= 1$ the confidence regions are given for the special case of the MVDR beamformer by

$$\begin{array}{|c|}\hline Pr\left(|\hat{b}_{MVDR}(\theta,\omega) - \mathbf{w}_{MVDR}^H \mathbf{d}_\theta(\omega)| < R_{CR}||\mathbf{p}_\theta(\omega)||/(\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g})^{1/2}\right) = \\ \dfrac{R_{CR}^2}{(1+R_{CR}^2)^{(L-N+2)}} \times \sum_{k=0}^{(L-N+1)} \begin{pmatrix} L-N+2 \\ k+1 \end{pmatrix} R_{CR}^{2k}. \\ \hline \end{array} \tag{5.71}$$

The confidence regions for the SCB GSC vector/scalar beam response are implicit.

5.5.2 Signal Estimates/Beamformer Outputs

To specify confidence regions for the SCB signal estimates we first define the variables

$$\begin{aligned} \mathbf{v}_{ML} &= (\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{1/2}(\hat{\mathbf{s}}_{ML} - \mathbf{s}) \\ \mathbf{v}_{GSC} &= (\mathbf{\Phi}^H \mathfrak{P}(\mathbf{G}_\perp|\mathbf{R})\mathbf{\Phi})^{-1/2}\,(\hat{\mathbf{s}}_{GSC} - \mathbf{s})\,. \end{aligned} \tag{5.72}$$

Noting that the Jacobian for these transformations are given by $J(\hat{\mathbf{s}}_{ML} \to \mathbf{v}_{ML}) = |\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G}|^{-1}$ and $J(\hat{\mathbf{s}}_{GSC} \to \mathbf{v}_{GSC}) = |\mathbf{\Phi}^H\mathfrak{P}(\mathbf{G}_\perp|\mathbf{R})\mathbf{\Phi}|$, it can be shown by Proposition 3.5.1.1 that $\mathbf{v}_{ML} \overset{d}{=} \mathbf{v}_{GSC} \overset{d}{=} \mathbf{v}$, where

$$P_{\mathbf{v}} = \int_0^1 P_\beta \cdot \beta^E \cdot g_{E:N(L+1)}\left[\,||\mathbf{v}||^2 \cdot \beta\,\right] d\beta. \tag{5.73}$$

We seek the likelihood of the event $||\mathbf{v}|| < R_{CR}$. The pdf for $\rho = ||\mathbf{v}||^2$ is found by considering the expectation $E\left\{h(||\mathbf{v}||^2)\right\}$ and using the WET of Proposition 3.6.0.4 to reduce the resulting integral. The pdf for ρ is shown to be

$$P_\rho = \frac{\pi^E}{(E-1)!}\,\rho^{E-1} \int_0^1 P_\beta \cdot \beta^E \cdot g_{E:N(L+1)}\left[\,\rho \cdot \beta\,\right] d\beta. \tag{5.74}$$

The probability of the desired event is given by the definite integral

$$Pr\left(||(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{1/2}(\widehat{\mathbf{s}}_{ML}-\mathbf{s})||<R_{CR}\right)=$$
$$\frac{\pi^E}{(E-1)!}\int_0^{R_{CR}^2}\rho^{E-1}\,d\rho\int_0^1 P_\beta\cdot\beta^E\cdot g_{E:N(L+1)}\left[\,\rho\cdot\beta\,\right]d\beta.$$

A detailed statistical analysis of the SCB LCMV/MVDR beamformer output under the complex Gaussian assumption, which includes its confidence regions is found in [14]. The analogous results for the SCB GSC signal estimate and beamformer output are implicit.

5.6 Statistics for Array Manifold Mismatch

It has thus far been assumed that the array manifold $\mathbf{G}$ is known exactly. Reality, however, often precludes exact knowledge of $\mathbf{G}$. Although in many applications good estimates can be obtained, an understanding of how uncertainties in its assumed structure can effect performance is of practical concern. We will therefore consider the following general scenario. The form of the SCB weightings will remain the same, *i.e.* with $\mathbf{G}$ as the assumed manifold. The signal bearing data vector will be changed such that $\mathbf{x}=\mathbf{V}\mathbf{s}+\mathbf{n}\sqrt{\tau}$, *i.e.* the actual true array manifold will be denoted by the $N\times E$ matrix $\mathbf{V}$. For this discussion we assume that $\mathbf{n}\sim\mathcal{CN}_N(\mathbf{0},\mathbf{R})$ and $\tau\sim P_\tau$. Exact pdfs for the beamformer outputs/signal estimates will be sought, assuming this Gaussian mixture subclass of complex MECs.

Note from eq(5.9) that the SCB ML signal estimator is distributed such that

$$\widehat{\mathbf{s}}_{ML}=\widehat{\mathbf{W}}_{ML}^H\mathbf{x}\stackrel{d}{=}\Delta^{-1}\left[\mathbf{I}_E,-\mathbf{Z}_1^H\mathbf{Z}_2(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\right]\mathbf{Q}\mathbf{R}^{-1/2}\mathbf{x}. \tag{5.75}$$

Define the following

$$\mathbf{Q}\mathbf{R}^{-1/2}\mathbf{x}=\mathbf{Q}\mathbf{R}^{-1/2}\mathbf{V}\mathbf{s}+\mathbf{Q}\mathbf{R}^{-1/2}\mathbf{n}\sqrt{\tau}\stackrel{d}{=}\mathbf{V}_0\mathbf{s}+\mathbf{z}_0\sqrt{\tau}\stackrel{\Delta}{=}\mathbf{x}_0 \tag{5.76}$$

where $\mathbf{z}_0\sim\mathcal{CN}_N(\mathbf{0},\mathbf{I}_N)$ independent of the real and positive r.v. τ. Partition the white primary data vector such that

$$\mathbf{x}_0=\begin{bmatrix}(\mathbf{x}_1)_{E\times 1}\\(\mathbf{x}_2)_{(N-E)\times 1}\end{bmatrix}. \tag{5.77}$$

Given τ note that $\mathbf{x}_0 \sim \mathcal{CN}_N(\mathbf{V}_0\mathbf{s}, \mathbf{I}_{N\tau})$. Thus, it follows that $\mathbf{x}_1$ and $\mathbf{x}_2$ are independent given τ. In terms of the partition variables the signal estimate can be written

$$\widehat{\mathbf{s}}_{ML} \stackrel{d}{=} \Delta^{-1}\mathbf{x}_1 - \Delta^{-1}\mathbf{Z}_1^H\mathbf{Z}_2(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{x}_2. \tag{5.78}$$

Unitary $\mathbf{Q}$ chosen as in eq(4.19) implies the following expectation

$$E\{\mathbf{x}_1\} = \left[\mathbf{I}_E, \mathbf{0}_{E\times(N-E)}\right]\mathbf{V}_0\mathbf{s} = \Delta^{-H}\mathbf{G}^H\mathbf{R}^{-1/2}\mathbf{Q}^H\mathbf{Q}\mathbf{R}^{-1/2}\mathbf{V}\mathbf{s} = \Delta^{-H}\mathbf{G}^H\mathbf{R}^{-1}\mathbf{V}\mathbf{s}. \tag{5.79}$$

Thus, conditioning on τ, $\mathbf{Z}_2$ and $\mathbf{x}_2$ the signal estimate is complex normal such that

$$\widehat{\mathbf{s}}_{ML} \sim \mathcal{CN}_E\left[\mathbf{W}_{ML}^H\mathbf{V}\mathbf{s},\ \tau\cdot(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}\times(1+\mathbf{x}_2^H(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{x}_2/\tau)\right]. \tag{5.80}$$

First conditioning on $\mathbf{x}_2$, the GWPT Theorem 3.6.1, or the Capon/Goodman result shows that

$$\mathbf{x}_2^H(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{x}_2/\tau \stackrel{d}{=} ||\mathbf{x}_2/\sqrt{\tau}||^2/\chi^2_{L-N+E+1} \stackrel{d}{=} \frac{\chi^2_{N-E}(\delta_\perp/\sqrt{\tau})}{\chi^2_{L-N+E+1}} \stackrel{d}{=} F_{N-E,K}(\delta_\perp/\sqrt{\tau}). \tag{5.81}$$

From Section 3.5.2 it follows that $||\mathbf{x}_2/\sqrt{\tau}||^2$ is non-central chi-squared $\chi^2_{N-E}(\delta_\perp/\sqrt{\tau})$ with non-centrality parameter $\delta_\perp/\sqrt{\tau} = ||E\{\mathbf{x}_2\}||/\sqrt{\tau}$. Note that

$$\begin{aligned} ||E\{\mathbf{x}_2\}||^2 &= ||E\{\mathbf{x}_0\}||^2 - ||E\{\mathbf{x}_1\}||^2 \\ &= \mathbf{s}^H\mathbf{V}^H\mathbf{R}^{-1}\mathbf{V}\mathbf{s} - \mathbf{s}^H\mathbf{V}^H\mathbf{R}^{-1}\mathbf{G}(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}\mathbf{G}^H\mathbf{R}^{-1}\mathbf{V}\mathbf{s} \\ &= \mathbf{s}^H\mathbf{V}^H\mathfrak{P}(\mathbf{G}|\mathbf{R}^{-1})\mathbf{V}\mathbf{s} = \delta_\perp^2 \end{aligned} \tag{5.82}$$

Ergo, the a priori density is obtained by

$$\begin{aligned} P_{\widehat{\mathbf{s}}_{ML}} &= \pi^{-E}|\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G}|\cdot\int_0^\infty P_\tau\, d\tau \int_0^1 P_{\beta_0|\tau} d\beta_0 \\ &\times (\beta_0/\tau)^E\cdot\exp\left\{-\left[\widehat{\mathbf{s}}_{ML} - \mathbf{W}_{ML}^H\mathbf{V}\mathbf{s}\right]^H(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})\left[\widehat{\mathbf{s}}_{ML} - \mathbf{W}_{ML}^H\mathbf{V}\mathbf{s}\right]\beta_0/\tau\right\} \end{aligned} \tag{5.83}$$

where $P_{\beta_0|\tau}$ is a non-central beta, *i.e* $\beta_{K,N-E}(\delta_\perp/\sqrt{\tau})$.

Chapter 6

Selected Simulations

The results of this thesis have wide utility and the potential to provide insight into the effects of non-Gaussianity on processor performance within an adaptive array setting. In this section we illustrate via a few examples some of these performance assessment possibilities. Note first that both the AMF and GLRT detectors possess the systematic structure illustrated in fig(6.1). Both detectors are essentially beamformers followed by an energy detect scheme. In the following simulations we will at times exploit this structure to gain physical insights.

6.1 The GLRT, LCMV, and the K-Distribution

Attention is restricted to the GLRT detector, the LCMV beamformer, and the K-distribution as the example of non-Gaussianity. The K-distribution is in fact a Gaussian mixture, which has received significant attention in the radar literature [2]–[7]. The mixture variable τ has the Gamma density:

$$P_\tau = \frac{\varpi}{\Gamma(\nu)}\tau^{\nu-1}e^{-\varpi\tau}, \qquad \tau \geq 0. \tag{6.1}$$

It has mean ν/ϖ and variance ν/ϖ^2. If we choose $\varpi = \nu/\varepsilon$ then the density becomes

$$P_\tau = \frac{(\nu/\varepsilon)}{\Gamma(\nu)}\tau^{\nu-1}e^{-(\nu/\varepsilon)\tau} \qquad \tau \geq 0, \tag{6.2}$$

and $E\{\tau\} = \varepsilon$, $\text{var}(\tau) = \varepsilon^2/\nu$. This choice coincides with that made in [5], and will be

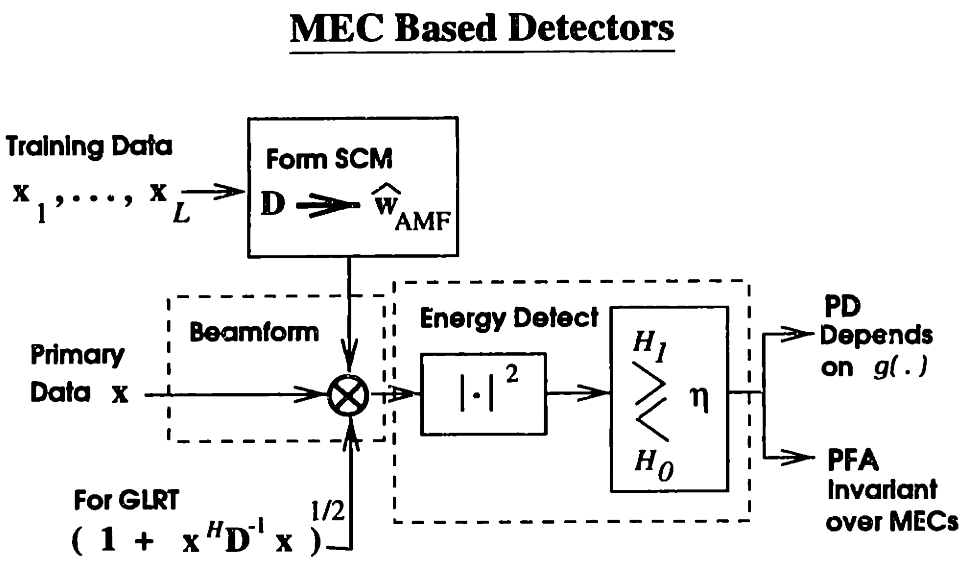


Figure 6.1: Basic Detection Structure

used in the present simulations. In fig(6.2)–fig(6.3) plots of P_τ are given for $\varepsilon = 1, 0.5$ as ν varies from 1 to 30. The density P_τ becomes more concentrated about unity for $\varepsilon = 1$ and about 0.5 for $\varepsilon = 0.5$, as anticipated from the mean and variance of τ. As $\nu \to \infty$ then $P_\tau \to \delta(\tau - \varepsilon)$, and thus the K-distribution approaches the Gaussian distribution $\mathcal{CN}_{N(L+1)}(\mathbf{M}_i, \mathbf{I}_L \otimes \mathbf{R})$ for $\varepsilon = 1.0$ and $\mathcal{CN}_{N(L+1)}(\mathbf{M}_i, \mathbf{I}_L \otimes 0.5\,\mathbf{R})$ for $\varepsilon = 0.5$.

6.2 PD vs. SNR

Consider the two plots of PD versus SNR $\triangleq \mathbf{s}^H\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G}\mathbf{s}$ in fig(6.4)–fig(6.5). The parameters L, N, E were chosen to match those in fig(3) of [19]. In fig(6.4) $\varepsilon = 1$, and ν varies from 1 to 30 as in fig(6.2). The PD curve of the K-distribution approaches that of the Gaussian as anticipated. Choosing $\varepsilon = 0.5$, fig(6.5) demonstrates that the PD curve of the K-distribution approaches the performance of a Gaussian data matrix with half the covariance of the Gaussian in the plot. Hence, the 3dB shift.

6.3 ROC Curves and Convergence Issues

In fig(6.6) we plot receiver operation characteristic (ROC) curves as the SNR varies from 5dB to 15dB. For the K-distribution we arbitrarily keep $\varepsilon = 1$ and $\nu = 9.2857$. This figure illustrates that the K-distribution can in fact have a different adaptive convergence rate than the Gaussian. Recall from fig(6.4) that for $L = 40$, SNR = 15dB, the PD for the Gaussian is higher than that of the K-distribution. In fig(6.6), however, for $L = 25$, the K-distribution has a higher PD than the Gaussian. To illustrate the varied convergence, we set SNR = 5dB, PFA = 0.4, and plot the PD as a function of L in fig(6.7). Although the K-distribution starts out with a higher PD for smaller L, the Gaussian eventually over takes the K-distribution approximately at $L = 29$. This difference in convergence rates can be attributed to the differing distributional head and tail behavior as exhibited in their functional forms, *i.e.* the $g(\cdot)$ for each.

6.4 Trade-Off Between SNR and Noise Adaptation

6.4.1 White Noise

In [14] and [15] it was demonstrated for the SCB LCMV beamformer that there exists a trade off in performance between the SNR (and spatial resolution) and noise adaptation. As mentioned earlier, the GLRT and AMF detectors are in fact front-end adaptive beamformers directly related to the SCB LCMV beamformer. In fig(6.8) we demonstrate that this very same trade off exists for these adaptive detectors. As in [14] we set $\mathbf{R} = \mathbf{I}$. Subsequently we fix $L = 40$, $E = 1$, PFA = 1e-6, and plot the resulting PD as N varies from 2 to L. The SNR is inherently a function of N. The SNR's indicated in fig(6.8) represent peak SNR's obtained at $N = L$. Analogous to the beamformer's performance, these plots suggests the existence of an optimal number of adaptive degrees of freedom N for which the best performance (highest PD) is obtained. This optimal N, however, depends on the SNR, L, E, PFA, and the functional form of the MEC distribution $g(\cdot)$.

6.4.2 Jammer/Interferer Effects

As in the previous section we give attention to the SNR/noise adaptation trade off. We assume, however, a more complex clairvoyant covariance structure; namely, $\mathbf{R} = \sigma^2\mathbf{I}_N + \sigma_v^2\mathbf{v}(\theta_J)\mathbf{v}^H(\theta_J)$, where $\sigma^2 = \sigma_v^2$. This structure represents a jamming interferer of direction θ_J degrees from array broadside in a background of white noise, as illustrated in fig(6.9). We assume planewave propagation in homogeneous media with a uniform linear array steered to a desired look direction of 30° from broadside, *i.e.* $\mathbf{G} = \mathbf{g}(30°)$. In fig(6.10)–fig(6.12) we plot the GLRT PD as a function of N and θ_J with fixed parameters $L = 40$, $E = 1$, and PFA = 1e-6. The peak SNR is 20.5dB occurring at $N = 40$. The jammer's direction is varied from 20° to 30° where it sits exactly on top of the desired signal. The plots indicate that the SNR/noise adaptation trade off is still present. Likewise, there exists a value of N which will yield the highest PD. When $\theta_J = 30°$, the PD ramps up to unity as the jammer and true signal become indistinguishable to the GLRT processor.

In fig(6.13)–fig(6.17) we plot the adaptive spatial filter patterns including both (i) the clairvoyant LCMV pattern and (ii) a typical SCB LCMV pattern, *i.e.* $|\mathbf{w}_{LCMV}^H\mathbf{d}_\theta(\omega)|^2$ and $|\hat{\mathbf{w}}_{LCMV}^H\mathbf{d}_\theta(\omega)|^2$. These patterns are plotted in an attempt to gain a more physical interpretation of the previous PD plots. In the left columns of each figure we plot the spatial patterns obtained from a fully adaptive array, which corresponds to $N = 40$ on the PD plots. The right columns show an improved adaptive pattern for a partially adaptive array. In particular, the values of N shown in the right column plots were chosen to correspond the peak PD obtained by the Gaussian in fig(6.10)–fig(6.12). The GLRT processor gives up some spatial resolution and SNR for a smaller variation in the adaptive spatial patterns about the clairvoyant ones in an attempt to maximize PD. As θ_J gets very close to 30°, the effectiveness of the array processor to accomplish this maximization diminishes. Indeed, as fig(6.12) and fig(6.17) indicate, when $\theta_J = 30°$, the battle is lost. The optimal value of $N = 2$ indicates that the benefit of having an array completely vanishes.

6.5 Figures

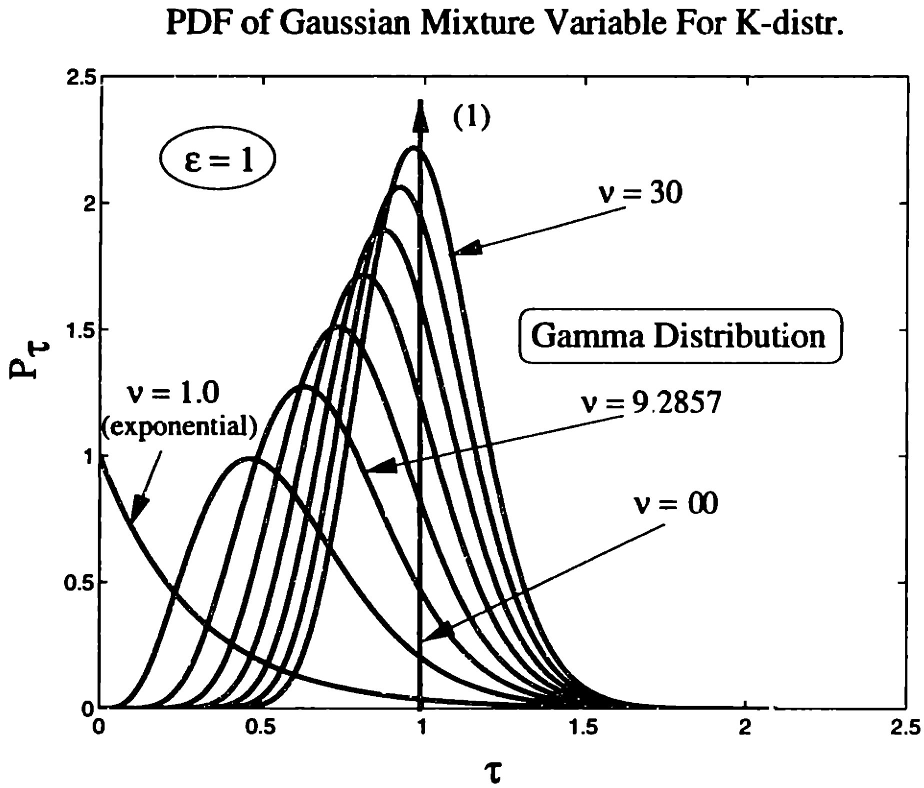


Figure 6.2: Gamma Distribution, $\varepsilon = 1$

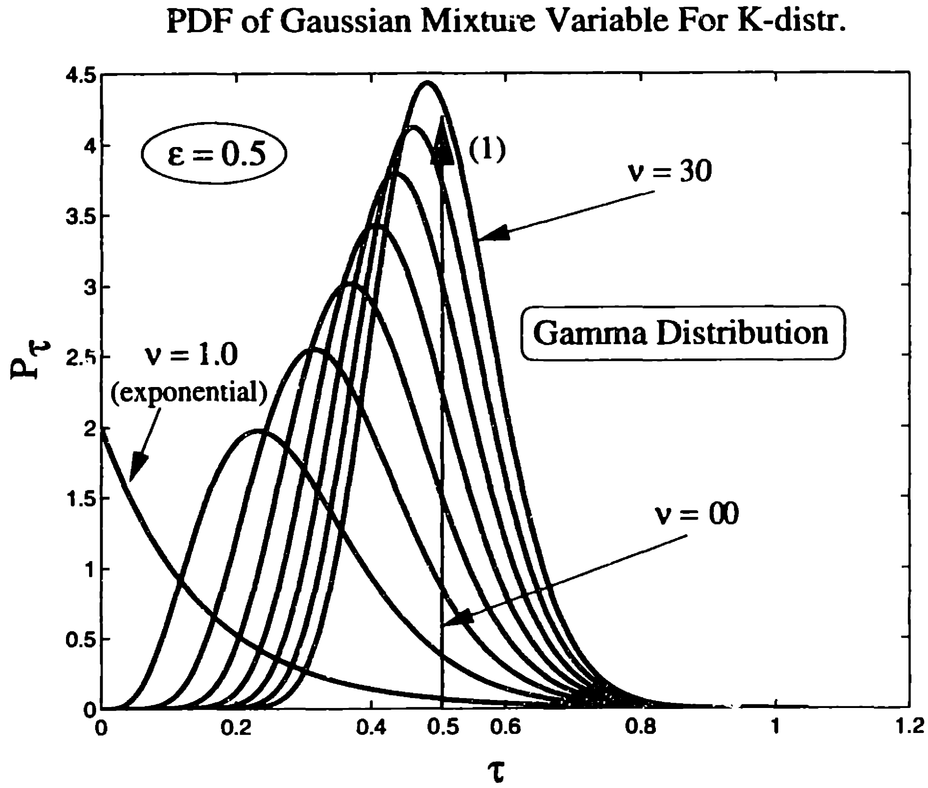


Figure 6.3: Gamma Distribution, $\varepsilon = 0.5$

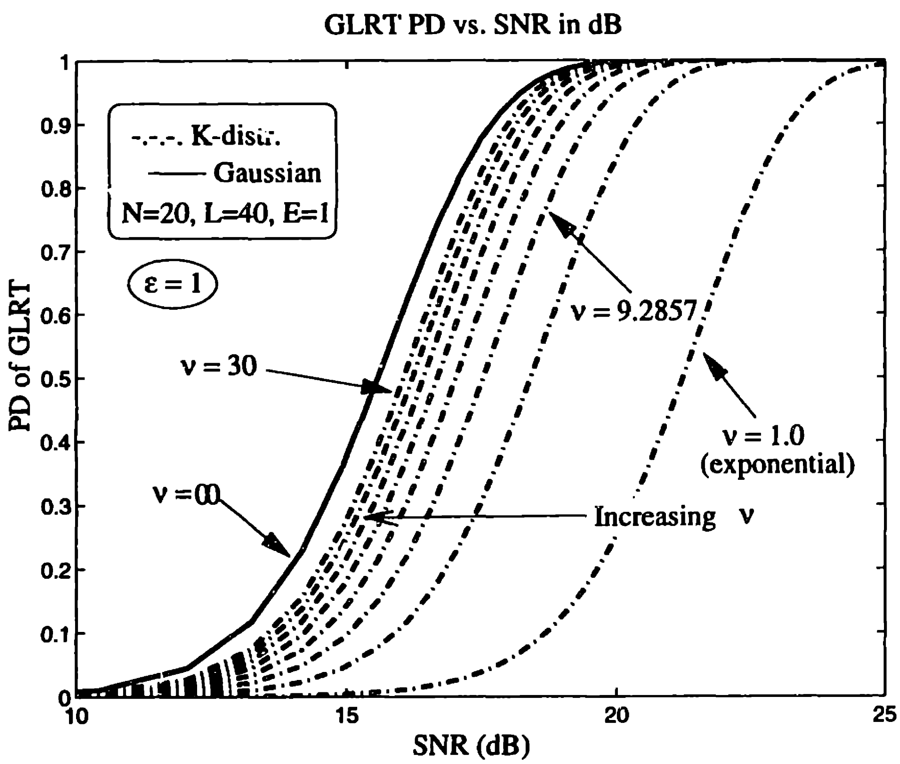


Figure 6.4: PD vs. SNR, $\varepsilon = 1$

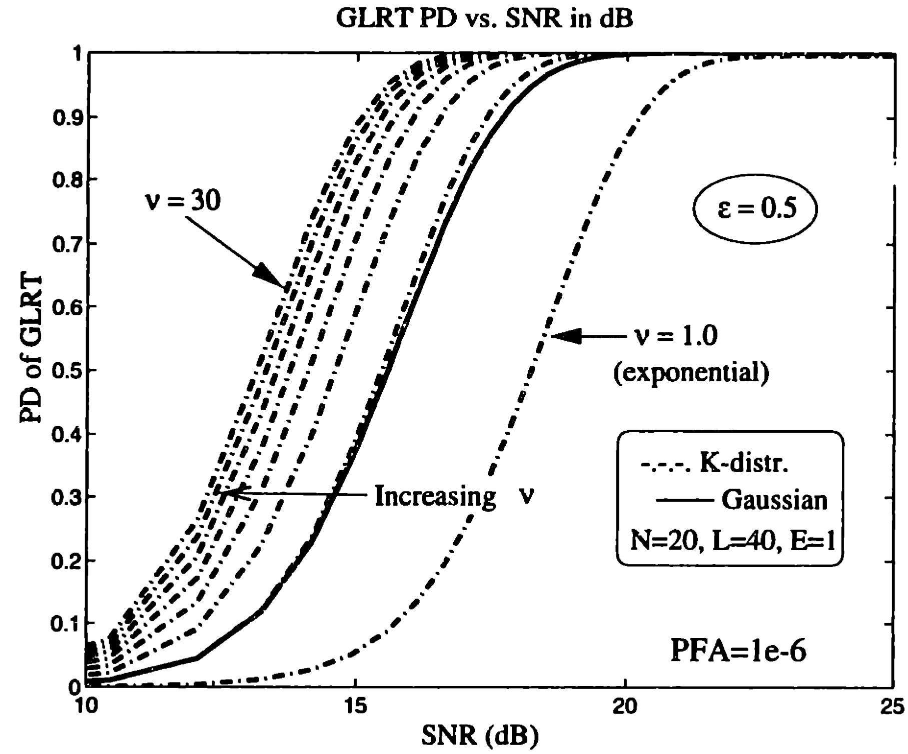


Figure 6.5: PD vs. SNR, $\varepsilon = 0.5$

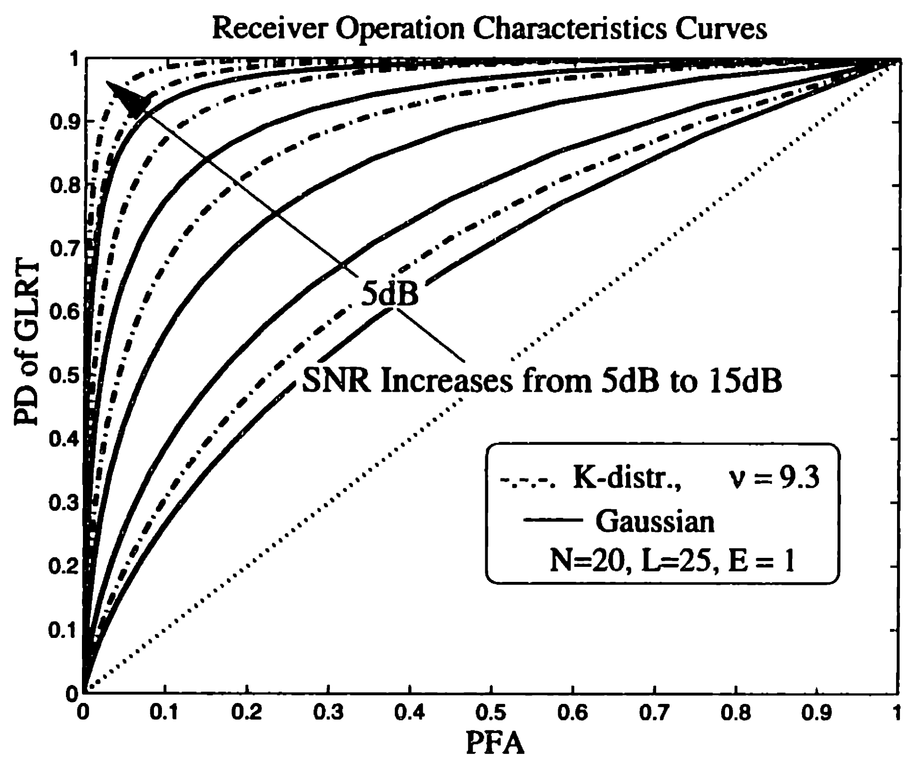


Figure 6.6: ROC Curves for various SNR

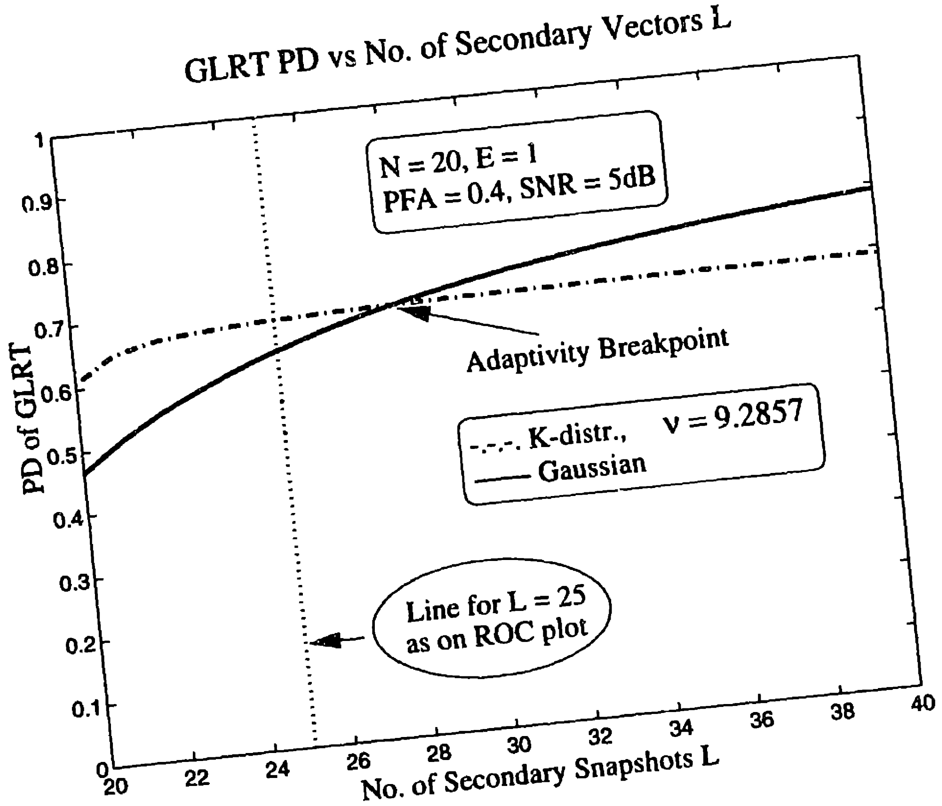


Figure 6.7: Convergence Rates as L Varies

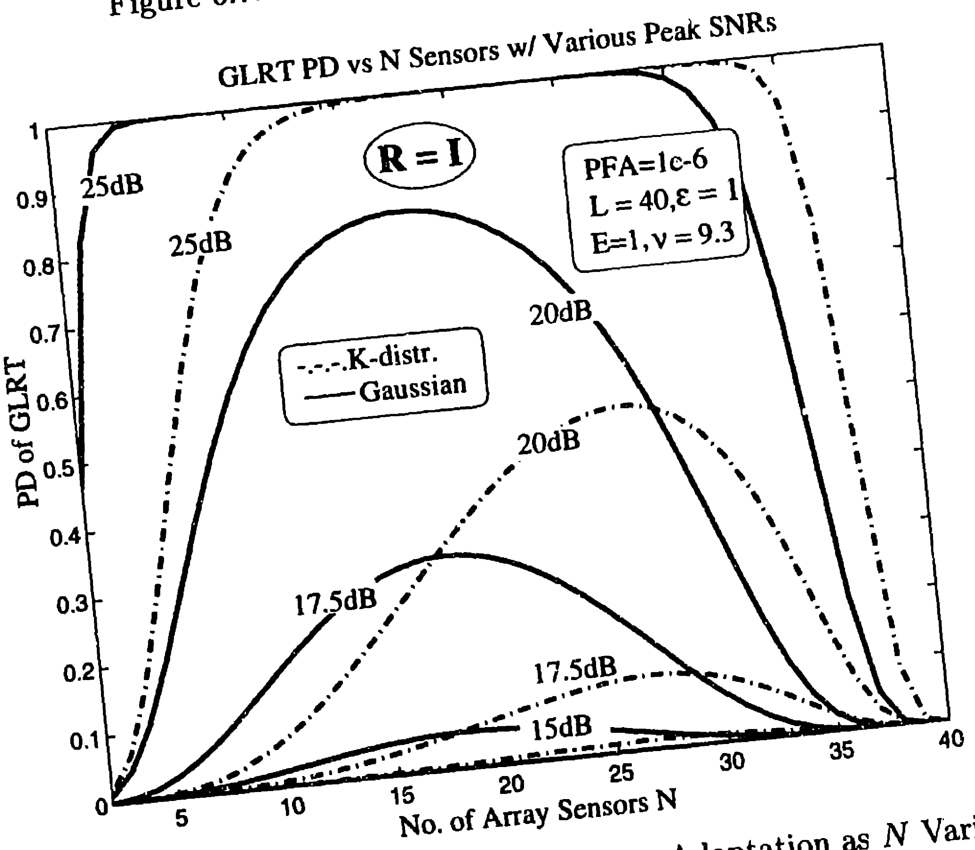


Figure 6.8: Trade Off between SNR and Noise Adaptation as N Varies

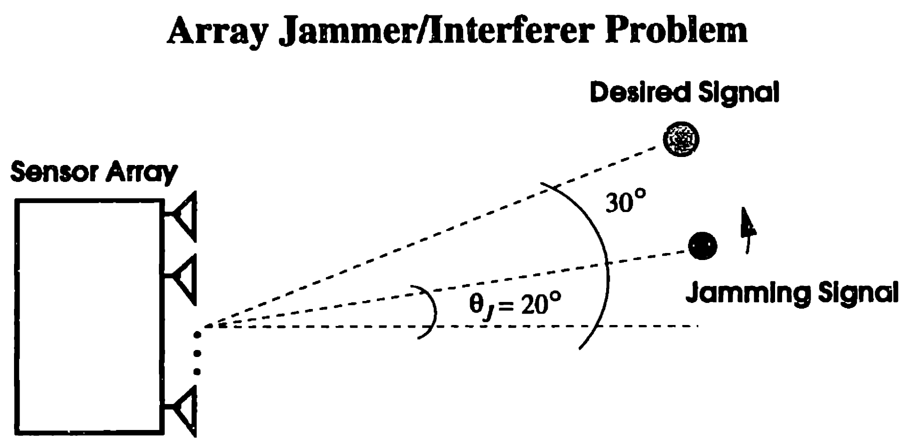
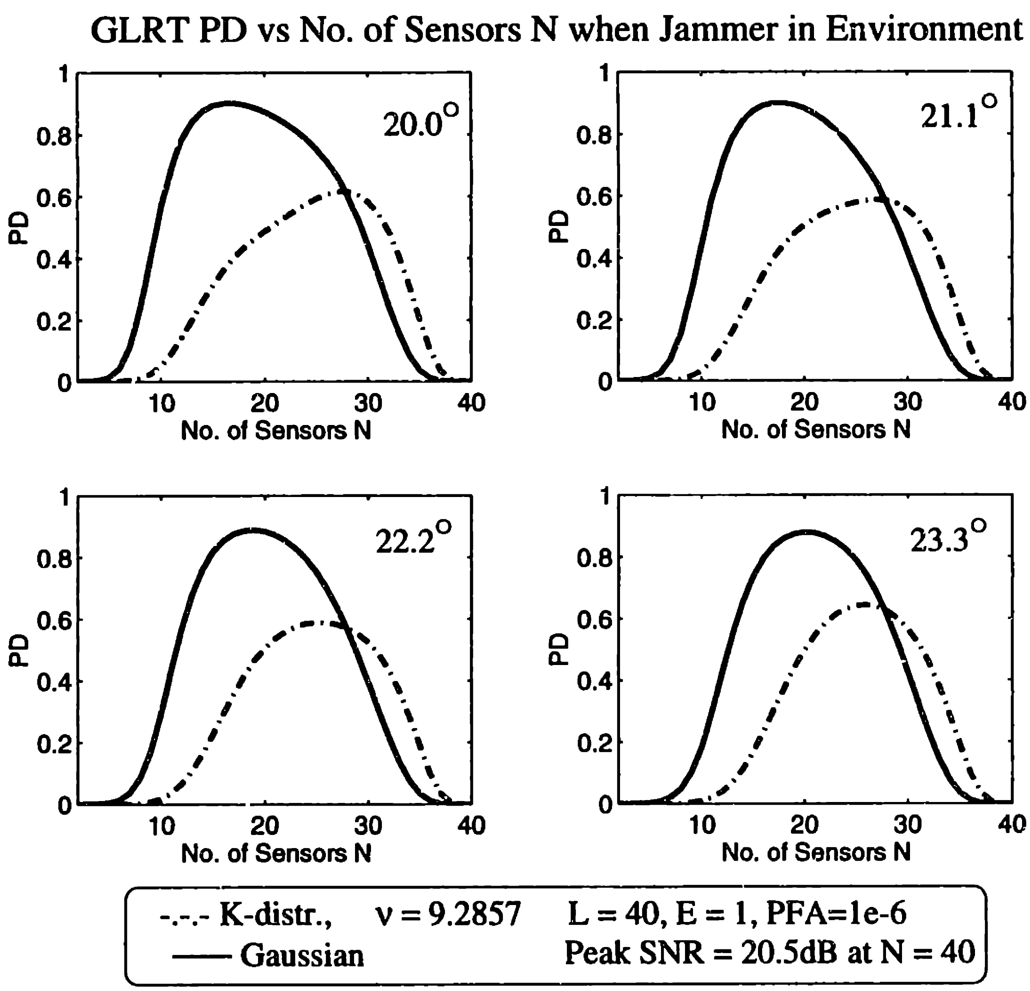


Figure 6.9: Jammer/Interferer Noise Environment



Array Steered to 30°, Jammer Direction in Corner

Figure 6.10: PD vs. N when Jammer Present (i)

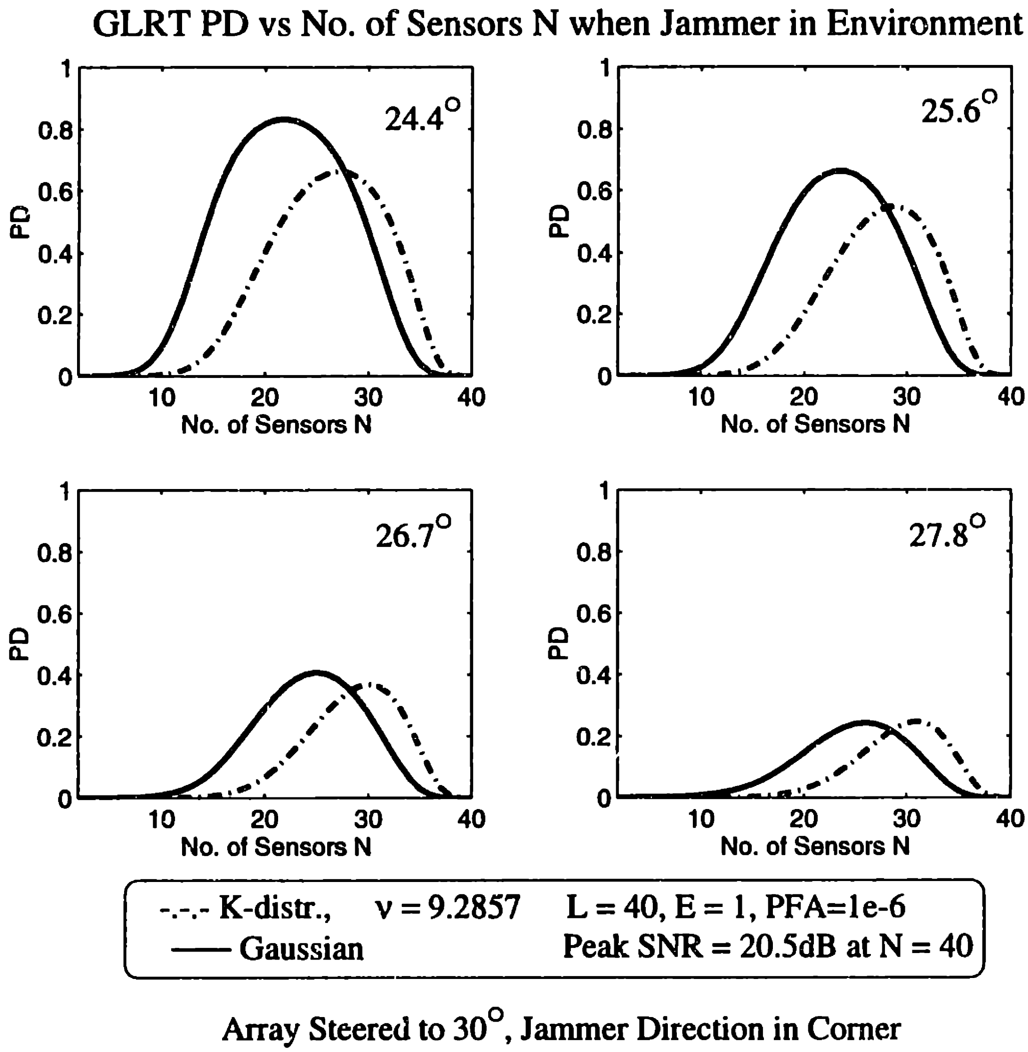


Figure 6.11: PD vs. N when Jammer Present (ii)

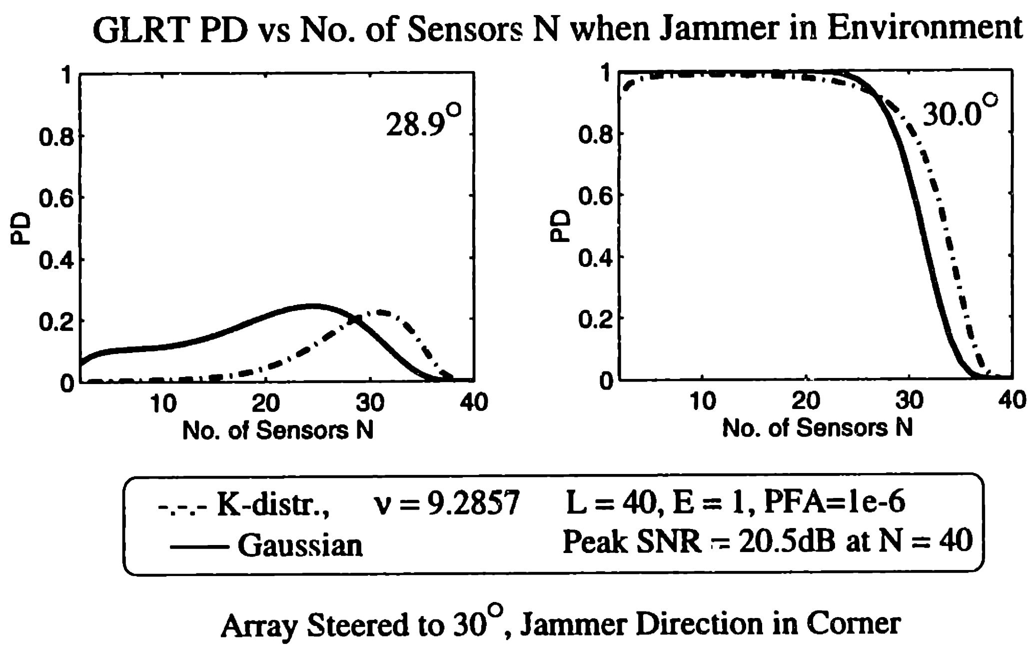


Figure 6.12: PD vs. N when Jammer Present (iii)

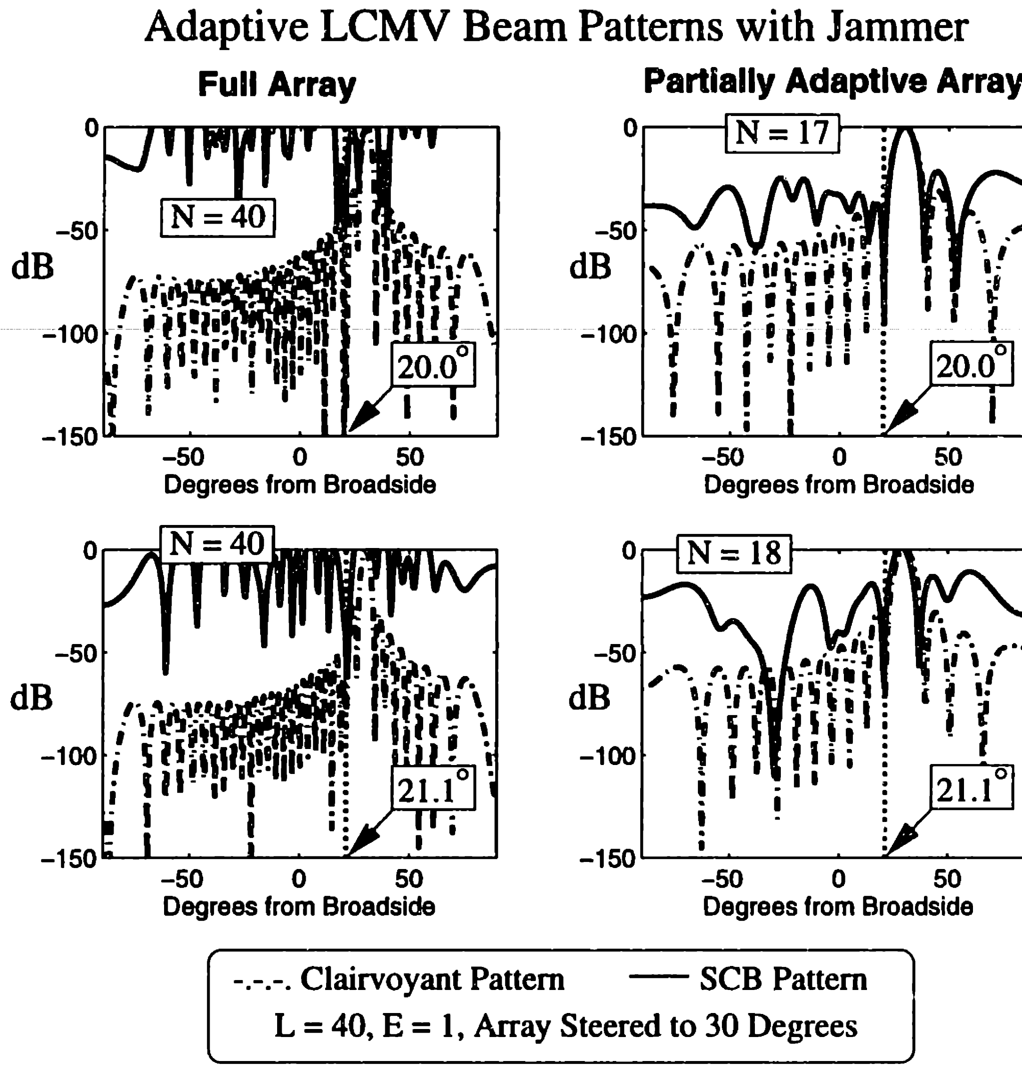


Figure 6.13: Adaptive LCMV Beampatterns (i)

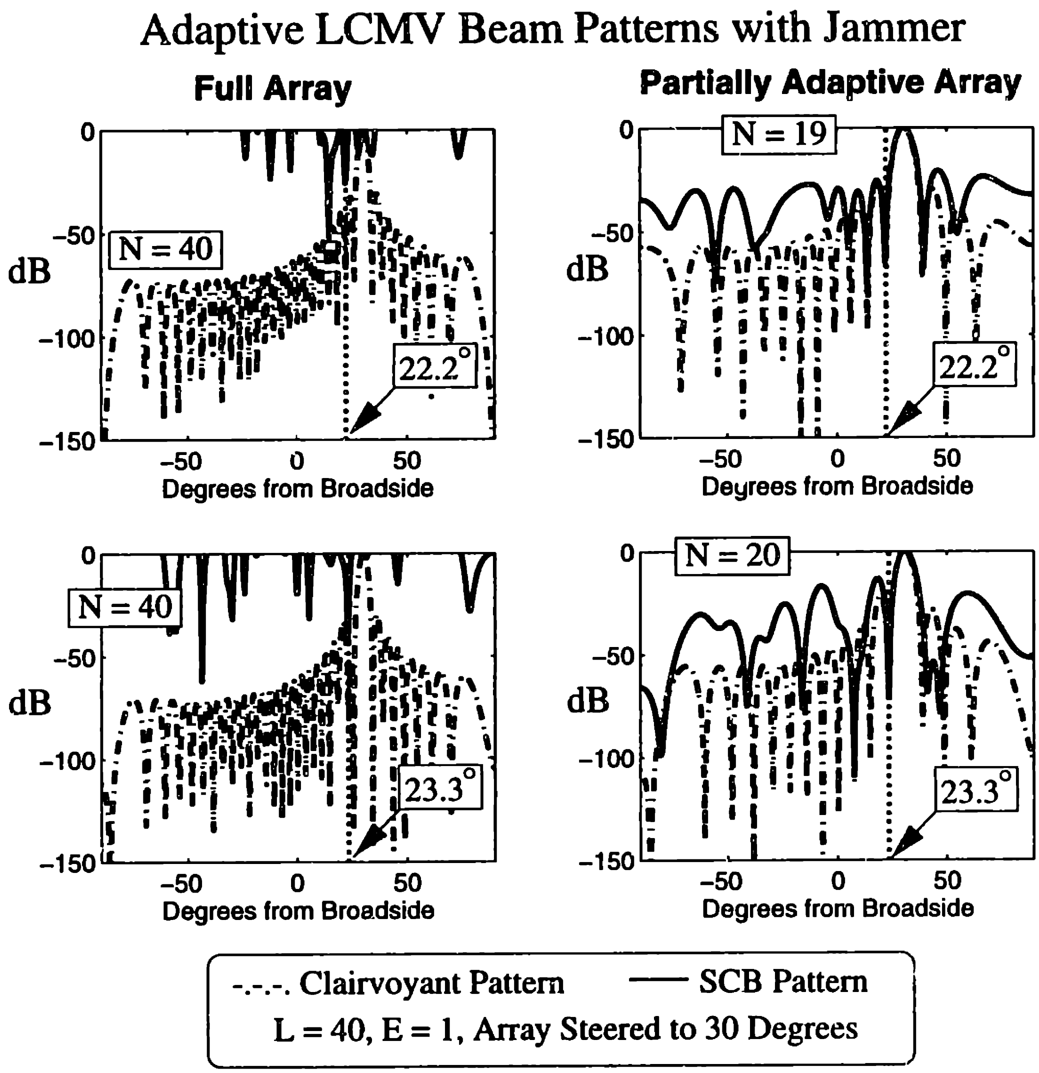


Figure 6.14: Adaptive LCMV Beampatterns (ii)

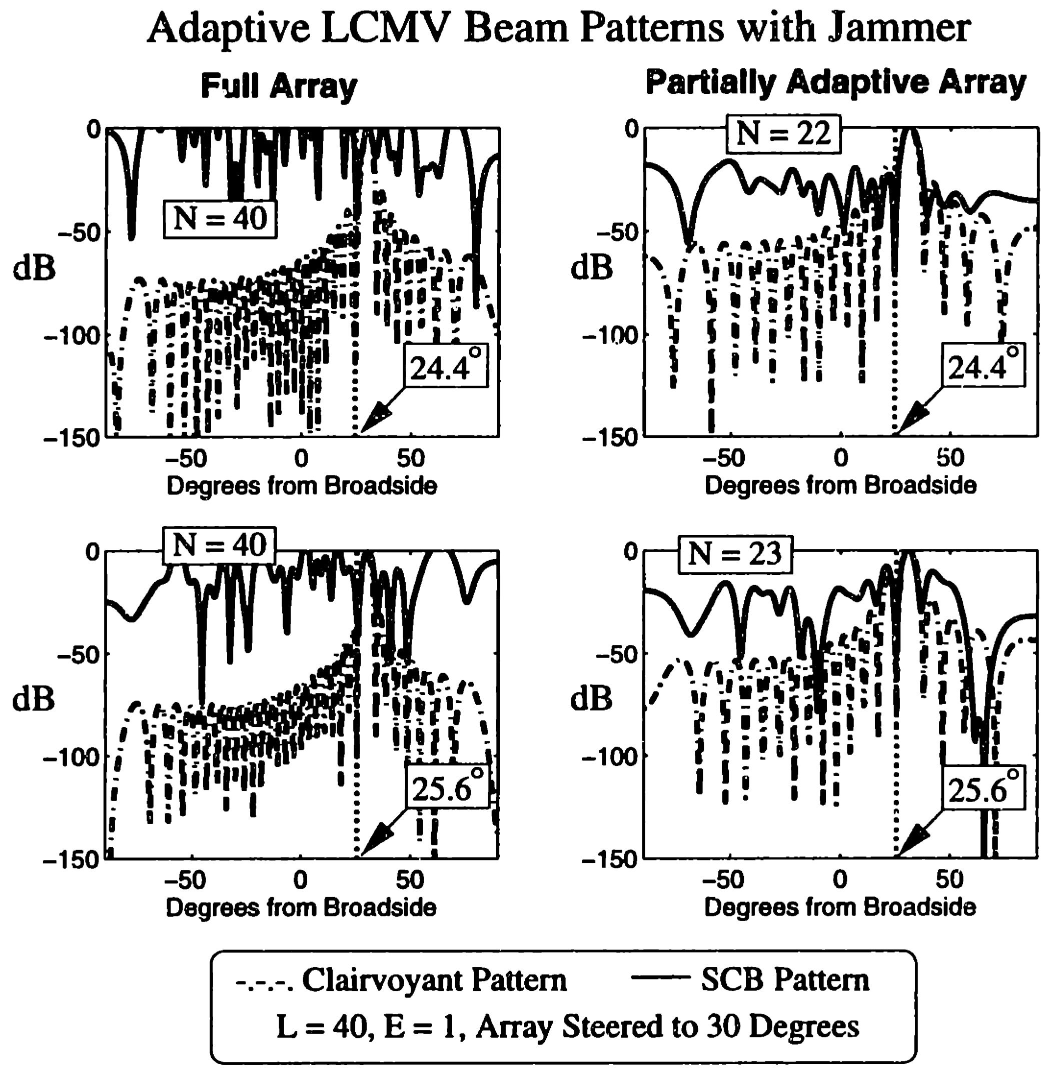


Figure 6.15: Adaptive LCMV Beampatterns (iii)

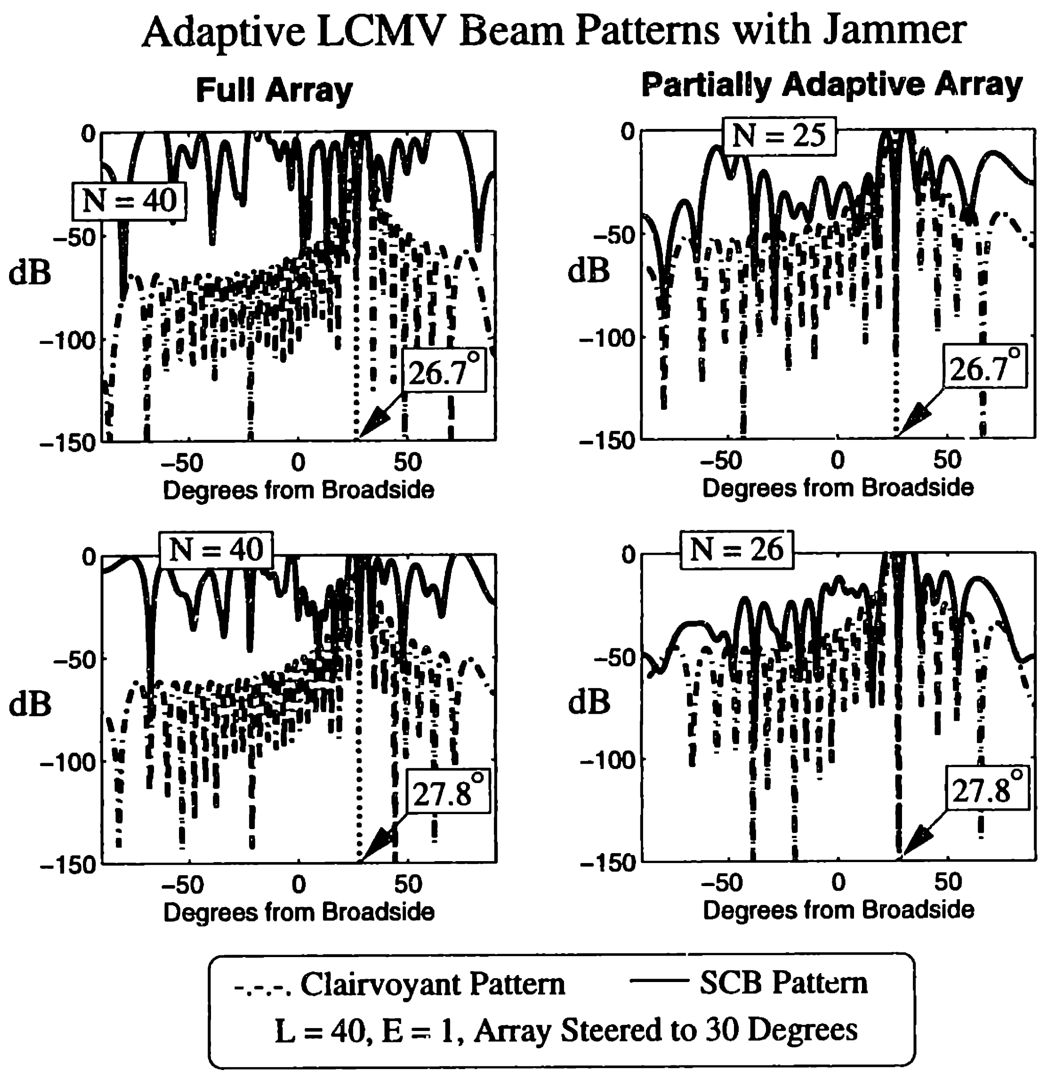


Figure 6.16: Adaptive LCMV Beampatterns (iv)

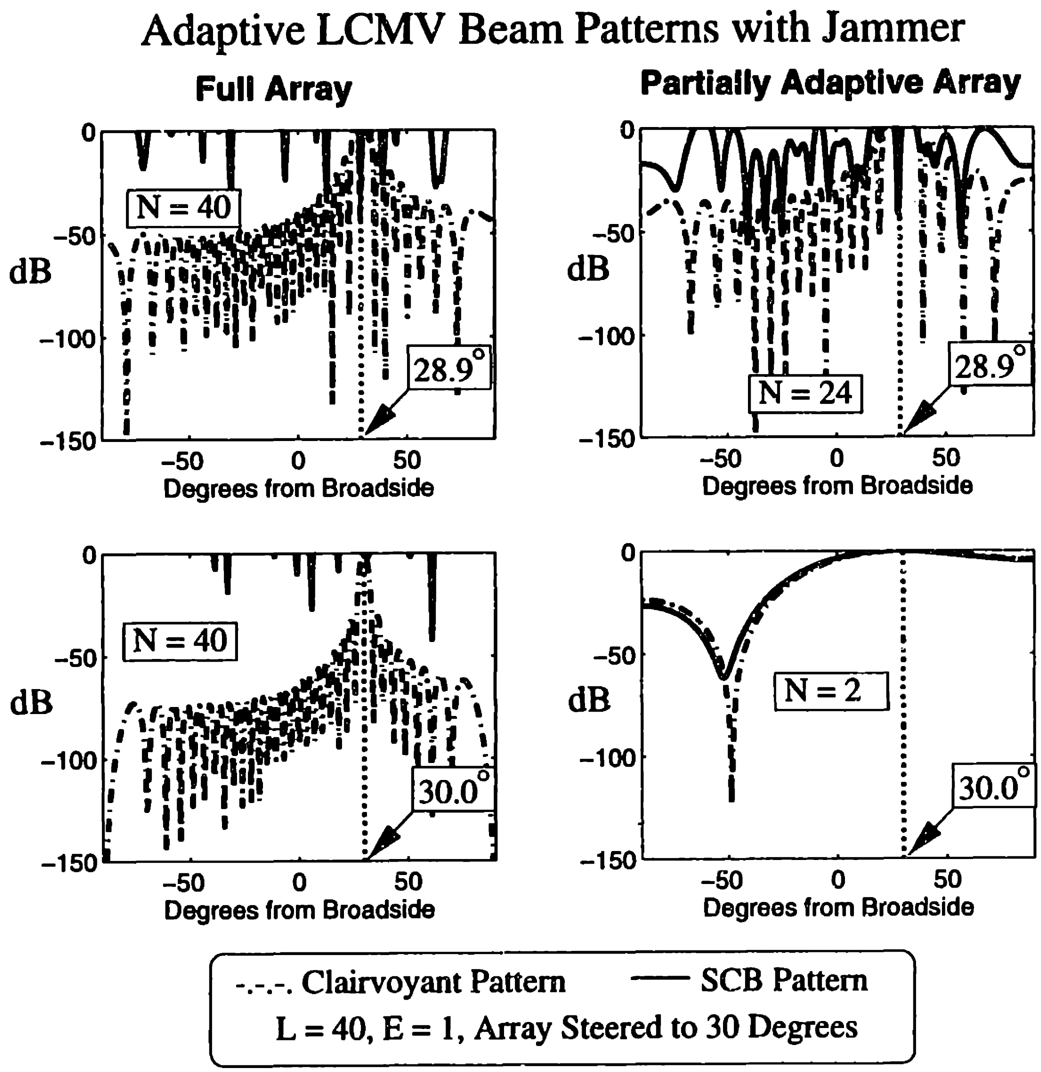


Figure 6.17: Adaptive LCMV Beampatterns (v)

Chapter 7

Additional Results on Adaptive Arrays

In this chapter we present results on power spectral estimation (PSE) and structured covariance estimation, which essentially extend the classical Gaussian based results of Capon and Goodman [32, 33] and those of Burg *et al* [35].

7.1 Adaptive Power Spectral Estimation

Capon's approach to PSE is a rather heuristic one. Essentially, he reasons that the variance (or covariance) of the clairvoyant complex Gaussian ML signal parameter estimator

$$\mathrm{cov}(\mathbf{s}_{ML}) = (\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1} = P_{CAPON}(\boldsymbol{\theta}, \boldsymbol{\omega}) \tag{7.1}$$

can be used as an estimator of signal power(s) where, *e.g.* the matrix

$$\mathbf{G}_{\boldsymbol{\theta}}(\boldsymbol{\omega}) = [\mathbf{g}_{\theta_1}(\omega_1)|\mathbf{g}_{\theta_2}(\omega_2)|\cdots|\mathbf{g}_{\theta_E}(\omega_E)] \tag{7.2}$$

is one of steering vectors $\mathbf{g}_i = \mathbf{g}_{\theta_i}(\omega_i)$ which are direction $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_E]^T$, and frequency $\boldsymbol{\omega} = [\omega_1, \ldots, \omega_E]^T$ dependent. Again, there's a need for estimation of the typically unknown data covariance parameter $\mathbf{R}$. Hence, Capon [32, 33] takes the SMI approach, although not referred to as SMI at the time of publication, and suggests the SCB estimator:

$$\widehat{P}_{CAPON}(\boldsymbol{\theta}, \boldsymbol{\omega}) = (\mathbf{G}_{\boldsymbol{\theta}}^H(\boldsymbol{\omega})\widehat{\mathbf{R}}^{-1}\mathbf{G}_{\boldsymbol{\theta}}(\boldsymbol{\omega}))^{-1}. \tag{7.3}$$

Capon/Goodman showed that estimator has a complex Wishart distribution when $\widehat{\mathbf{R}}$ is the usual SCM based on complex zero mean Gaussian data.

In Section 4.3 it was shown that the clairvoyant ML estimate for the signal parameters remains the same for many complex MECs. Hence, we use the same power spectral estimator as that suggested by Capon. Now assume that the secondary data is complex MEC distributed according to

$$\mathbf{X} \sim \mathcal{CMEC}_{NL}\left[\mathbf{0}_{N\times L}, \mathbf{I}_L \otimes \mathbf{R}, g(\cdot)\right]. \tag{7.4}$$

Let $\widehat{\mathbf{R}} = \mathbf{D}/r_{MAX}$ where r_{MAX} is chosen in a fashion analogous to that described in Section 4.2 so that it corresponds to the ML estimate of $\mathbf{R}$. Using eq(4.18) and choosing $\mathbf{Q}$ as in eq(4.19) we obtain the stochastic representation

$$\begin{aligned} \widehat{P}_{CAPON}(\boldsymbol{\theta},\boldsymbol{\omega}) &= \kappa\cdot(\mathbf{G}_{\boldsymbol{\theta}}^H(\boldsymbol{\omega})\widehat{\mathbf{R}}^{-1}\mathbf{G}_{\boldsymbol{\theta}}(\boldsymbol{\omega}))^{-1} \stackrel{d}{=} \kappa\cdot(\Delta^H\Theta^{-1}\Delta)^{-1}/r_{MAX} \\ &\stackrel{d}{=} \kappa\cdot(\mathbf{G}_{\boldsymbol{\theta}}^H(\boldsymbol{\omega})\mathbf{R}^{-1}\mathbf{G}_{\boldsymbol{\theta}}(\boldsymbol{\omega}))^{-1/2}\cdot\Theta\cdot(\mathbf{G}_{\boldsymbol{\theta}}^H(\boldsymbol{\omega})\mathbf{R}^{-1}\mathbf{G}_{\boldsymbol{\theta}}(\boldsymbol{\omega}))^{-1/2}/r_{MAX}. \end{aligned} \tag{7.5}$$

Recalling the partition matrix inversion lemmas of eq(4.24), it follows that $\Theta = \mathbf{Z}_1^H\, \mathfrak{P}(\mathbf{Z}_2|\mathbf{I}_L)\, \mathbf{Z}_1$. Thus, by the GWPT Theorem 3.6.1 part (ii) it follows that

$$\Theta \sim \mathcal{GCW}_E\left[L-N+E, \mathbf{I}_E, g_{E(L-N+E):NL}(\cdot)\right] \tag{7.6}$$

which generalizes the Capon/Goodman result over complex MECs. Indeed, for the complex Gaussian it follows that $r_{MAX} = L$, $\kappa = 1$, and $g_{E(L-N+E):NL}(r) = \pi^{E(L-N+E)}e^{-r}$ which yield the complex Wishart distribution.

7.2 Necessary Gradient Condition for Real MEC Based Structured Covariance Estimation

In this section we assume all data is real and assume the notation of Burg *et al* [35]. Burg's analysis assumed the data to be zero mean Gaussian distributed. In this section we weaken that assumption somewhat by assuming the objective function (pdf) of the data set is of the general real MEC form

$$g(S,R) = |R|^{-M/2}h\left[\frac{M}{2}\mathrm{tr}(R^{-1}S)\right]. \tag{7.7}$$

Note that for the Gaussian data assumption made in [35] one would choose the functional dependence of the density to be $h(u) = (2\pi)^{-MN/2}e^{-u}$. It is assumed that $h(\cdot)$ is differentiable and satisfies certain conditions which will guarantee that its maximum exist. The idea of this section is simply to initiate the discussion of extending Burg's theoretic paradigm to other distributions. This analysis, if considered to reasonable extent, would make a nice dissertation.

Following [35] we consider the variation in g w.r.t. variation in R

$$\begin{aligned}\delta g(S,R) = -(M/2)|R|^{-\frac{M}{2}-1}\delta|R|h\left[\frac{M}{2}\text{tr}(R^{-1}S)\right] + \\ |R|^{-\frac{M}{2}}\delta_u h\left[\frac{M}{2}\text{tr}(R^{-1}S)\right]\delta[\text{tr}(R^{-1}S)]M/2.\end{aligned} \tag{7.8}$$

The subscripted differential δ_u is simply a reminder to indicate that the chain rule operation differentiating the argument of $h(\cdot)$ has already been carried out. Analogous to analysis under the Gaussian assumption it follows that

$$\delta g(S,R) = -\frac{M}{2}|R|^{-M/2}\left\{\text{tr}(R^{-1}\delta R)h\left[\frac{M}{2}\text{tr}(R^{-1}S)\right] + \text{tr}(R^{-1}\delta R R^{-1}S)\delta_u h\left[\frac{M}{2}\text{tr}(R^{-1}S)\right]\right\}.$$

Setting this equal to zero the *generalized trace condition* becomes

$$\boxed{\text{tr}\left[\left(R^{-1}h\left[\tfrac{M}{2}\text{tr}(R^{-1}S)\right] + R^{-1}SR^{-1}\delta_u h\left[\tfrac{M}{2}\text{tr}(R^{-1}S)\right]\right)\delta R\right] \triangleq \langle\nabla_R\, g, \delta R\rangle = 0} \tag{7.9}$$

Note that in Burg's defined inner product space the variation δR must be perpendicular to direction of the generalized gradient

$$\nabla_R\, g \propto R^{-1}h\left[\frac{M}{2}\text{tr}(R^{-1}S)\right] + R^{-1}SR^{-1}\delta_u h\left[\frac{M}{2}\text{tr}(R^{-1}S)\right]. \tag{7.10}$$

For Gaussian data note that

$$\begin{aligned}\delta_u h\left[\frac{M}{2}\text{tr}(R^{-1}S)\right] &= -(2\pi)^{-MN/2}\exp\left[-\tfrac{M}{2}\text{tr}(R^{-1}S)\right] \\ &= -h\left[\frac{M}{2}\text{tr}(R^{-1}S)\right]\end{aligned} \tag{7.11}$$

and trace condition reduces to the original form in [35].

As the only example to be considered here, note that when no structure is assumed for R, the gradient of g must be zero. This implies that

$$R = -aS \qquad \text{where} \tag{7.12}$$

$$a = \delta_u h\left[\frac{M}{2}\text{tr}(R^{-1}S)\right] \Big/ h\left[\frac{M}{2}\text{tr}(R^{-1}S)\right]. \tag{7.13}$$

This tells us that the answer is some $R \propto S$. If we let $R = (MN/2\alpha)S$ and substitute this

R back into the zero gradient equation, then one can show that α must be chosen to satisfy

$$h(\alpha) + \left(\frac{2\alpha}{MN}\right) \delta_u h(\alpha) = 0. \tag{7.14}$$

or written another way

$$\boxed{\frac{dh(\alpha)}{d\alpha} + \left(\frac{MN}{2\alpha}\right) h(\alpha) = 0.} \tag{7.15}$$

This is consistent with the literature on unconstrained covariance estimates [55]. If $h(u) \propto e^{-u}$ (Gaussian) then it is easy to show that $\alpha = MN/2$.

Chapter 8

Concluding Remarks

8.1 Summary

In this dissertation we considered for the first time adaptive array signal processing in a non-Gaussian setting. Adaptive array analyses is non-trivial in a Gaussian setting, as the work of Kelly [19, 22] and other bear witness. Weakening the Gaussian assumption ergo potentially increases difficulty. Using the concept of a complex MEC distribution, however, we were able to make headway. Consideration was given to both (i) problem reformation, and (ii) performance analysis issues.

In Chapter 4 the problem of adaptive array detection, as posed by Kelly, was considered in detail under a weaker complex MEC data assumption. It was proven that both the GLRT and AMF approaches lead to the same test statistic as that obtained in the Gaussian case. Indeed, for the GLRT approach we found that even the functional forms of the likelihood functions can differ, *i.e.* $g^{(0)}(r) \neq g^{(1)}(r)$, yet the test remains the same. Both the GLRT and AMF tests likewise remain CFAR detectors and have distributions under the null hypothesis which are invariant over the entire complex MEC class. Thus, the MEC based PFA's coincide with Kelly's [19] and Robey's [23]. The PD for both detectors, however, depend on $g^{(1)}(r)$ in addition to the usual parameters $(L, N, E, \mathbf{s}^H\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G}\mathbf{s})$, as the closed form expressions derived in Chapter 4 indicate.

Concerning adaptive signal estimation and beamforming we proved that the SCB ML, LCMV, MVDR, and GSC beamformers, although based on the heuristic SMI method, are in fact optimal in the ML sense under many complex MEC distributions. Regarding their performance in these non-Gaussian environments, it was shown that there is a unified stochastic representation across the MEC class; namely, the complex multivariate t–distribution is the underlying *joint* pdf for all the SCB weightings, generalizing Steinhardt's [28] Gaussian based result which only showed that the *marginals* of the SCB MVDR weightings were proportional to a univariate t. The exact means and covariances were computed for the SCB weightings, demonstrating them all to be (1) unbiased estimates of their clairvoyant counterparts, and (2) convergent in probability due to a covariance which approaches the zero matrix in the limit as $L \to \infty$. In addition means, covariances, pdfs, and confidence regions were computed for the resulting vector beam responses, scalar beam responses, signal estimates, and beamformer outputs. The beam responses were likewise shown to be MEC invariant with a generalized complex multivariate t–distribution. The signal estimates and beamformer outputs, however, were shown to depend on the functional form of the MEC distribution. Exact closed form expressions for their pdfs, confidence regions, and relevant statistics appear in Chapter 5.

In retrospect, recall that the MEC based adaptive array detectors can be block diagrammed as in fig(6.1). Both the GLRT and AMF detectors are fundamentally front-end adaptive beamformers followed by an energy (magnitude squared) detect scheme. Our analysis of both detectors concentrated on finding the PD and PFA shown in the diagram. The analysis of the adaptive beamformers focused on the statistical fidelity of the beamformer box in the diagram. Under the null hypothesis all statistics are invariant over the MEC class. Under H_1 the fidelity of the spatial filtering characteristics are unchanged, but the outputs of the beamform and energy detect boxes depend on the functional form of the MEC distribution.

Further, in Chapter 7 the distribution for the SCB MVDR or Capon power spectral estimator was generalized over MECs. In addition a generalization of Burg's necessary gradient condition for structured covariance estimation was given.

New contributions to the theory of EC distributions were also made. These contributions were labeled as Theorems in Chapter 3, whereas results that are well known were labeled as Propositions. A complex MEC generalization of the complex Wishart partition theorem, which gives exact joint and marginal distributions for the Schur complements and regression coefficients appears in Chapter 3. Likewise we considered the density of singular complex EC distributions.

In several applications (*e.g.* radar, sonar, etc.) the physics suggests and empirical analysis supports the notion that the snapshots of the data matrix $\mathbf{X}_0$ are not Gaussian nor independent; however, empirical analysis of certain statistics, *e.g.* the normalized SNR statistic [18, 16], appear rather robust in that they match well those under the Gaussian assumption [27]. The analysis in this dissertation under the assumption of complex MEC data provides an alternate (to central limiting theorems) non-asymptotic theory in harmony with [18] for the robustness of the array processors and their corresponding statistics in non-Gaussian data.

8.2 Future Work

One of the major contributions of this thesis is the provision of a large body of complex MEC based statistical measures for adaptive arrays. A full comprehensive exploitation of all these measures, of course, cannot be presented in one thesis. It is hoped that the few simulations and examples we chose to demonstrate the utility of the results will serve to stimulate further research in this area. Indeed, this dissertation sets the stage for the consideration of a wide variety of analyses surrounding the joint effects of non-Gaussianity and loss due to the processor's need to adapt.

In this thesis we implicitly assumed that each snapshot was dependent, *i.e.* except when $g(r) \propto e^{-r}$. The work of Raghavan [8] on adaptive detection, however, assumes independent complex EC snapshots. Reality, however, may lie somewhere in between. The consideration of adaptive detection, estimation and/or beamforming with a model representing a hybrid of the one chosen in this thesis and that assumed in [8] is an open problem. This hybrid model would be a mix between eq(3.106) and eq(3.107).

In Chapter 4 Robey's AMF procedure involving the partial GLRT approach lead to an MEC based detector having the same form as the Gaussian based AMF. The analysis of this detector at first glance appears rather elusive due to the inverse functions dependence threshold. A comprehensive analysis of this detector seeking to understand the net effects and surrounding issues of this inversion should prove useful.

The problem of structured covariance estimation from non-Gaussian data remains a open problem. We initiated the analysis in Chapter 7 by deriving a generalization of the gradient trace condition. Common covariance structures, such as a diagonal matrix, or the identity plus a low rank noise subspace, would be good starting points.

Several of the statistics derived in this thesis involve the functional operator $g_{M:N}(\cdot)$ defined in eq(3.89). Numerical approximations involving of this operator, which for example are necessary for computing detector PD's, can become numerically unstable with finite precision software packages like MATLAB. Investing time in an efficient numerical approximation of this functional operator for arbitrary continuous functions $g(\cdot)$, which would run on MATLAB or in C or FORTRAN, would be worthwhile. Infinite precision packages like MAPLE, for example, in theory should have no problems. [1]

[1] The thesis author is presently writing such a routine for MAPLE.

Appendix A

Symbols and Notation

In this thesis the notational convention is boldfaced capitals indicating complex matrices and boldfaced lowercase indicative of complex vectors, unless stated otherwise. Italic indicates complex scalar quantities in general, although several are real and some integers. All are apparent from the context in which they are used. (The notational convention of Chapter 3 differs from the rest of the thesis in that a tilde " ~ " is used to indicate complex quantities and statistic related to complex quantities. See introduction to Chapter 3.)

Symbol	Definition
$\mathbf{x}_{N\times 1}$	Array snapshot under interrogation
$\mathbf{X}_{N\times L}$	Training data matrix $[\mathbf{x}_1\vert\cdots\vert\mathbf{x}_L]$
$\mathbf{X}_0 = [\mathbf{X}\vert\mathbf{x}]$	Totality of data
$\mathbf{R}_{N\times N}$	True but unknown data covariance cov($\mathbf{x}$)
$\widehat{\mathbf{R}}_{N\times N} = \mathbf{X}\mathbf{X}^H/L$	Sample covariance matrix
$\mathbf{D} = \mathbf{X}\mathbf{X}^H$	Unnormalized sample covariance
$\mathbf{G}_{N\times E}$	Matrix of E steering vectors (look directions)
$[\mathbf{G}_\perp]_{N\times(N-E)}$	$\mathbf{G}_\perp^H\mathbf{G} = \mathbf{0}$ such that $\mathbf{G}_\perp\mathbf{G}_\perp^H = \mathfrak{P}(\mathbf{G}\vert\mathbf{R}^{-1})$
$\mathbf{s}_{E\times 1}$	True but unknown signal parameters
$\hat{\mathbf{s}}_{(\cdot)} = \widehat{\mathbf{W}}_{(\cdot)}^H\mathbf{x}$	Signal parameter estimate, $(\cdot)$ = ML or GSC
$\mathbf{W}_{ML}$ and $\widehat{\mathbf{W}}_{ML}$	$\mathbf{R}^{-1}\mathbf{G}(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}$ and $\mathbf{D}^{-1}\mathbf{G}(\mathbf{G}^H\mathbf{D}^{-1}\mathbf{G})^{-1}$
$\mathbf{W}_{GSC}$ and $\widehat{\mathbf{W}}_{GSC}$	$[\mathbf{I}_N - \mathbf{G}_\perp\mathbf{\Omega}_{GSC}]\,\mathbf{G}(\mathbf{G}^H\mathbf{G})^{-1}$ and $\left[\mathbf{I}_N - \mathbf{G}_\perp\hat{\mathbf{\Omega}}_{GSC}\right]\mathbf{G}(\mathbf{G}^H\mathbf{G})^{-1}$
$\mathbf{\Omega}_{GSC}$ and $\hat{\mathbf{\Omega}}_{GSC}$	$(\mathbf{G}_\perp^H\,\mathbf{R}\,\mathbf{G}_\perp)^{-1}\mathbf{G}_\perp^H\,\mathbf{R}$ and $(\mathbf{G}_\perp^H\,\mathbf{D}\,\mathbf{G}_\perp)^{-1}\mathbf{G}_\perp^H\,\mathbf{D}$
$\mathbf{w}_{LCMV} = \mathbf{W}_{ML}\mathbf{f}$	LCMV beamformer with constraints $\mathbf{G}^H\mathbf{w}_{LCMV} = \mathbf{f}$
$\mathbf{w}_{gsc} = \mathbf{W}_{GSC}\mathbf{f}$	GSC beamformer with constraints $\mathbf{G}^H\mathbf{w}_{gsc} = \mathbf{f}$
$\hat{\mathbf{w}}_{LCMV}$ and $\hat{\mathbf{w}}_{gsc}$	$\widehat{\mathbf{W}}_{ML}\mathbf{f}$ and $\widehat{\mathbf{W}}_{GSC}\mathbf{f}$
$y_{(\cdot)}$ and $\hat{y}_{(\cdot)}$	$\mathbf{w}_{(\cdot)}^H\mathbf{x}$ and $\hat{\mathbf{w}}_{(\cdot)}^H\mathbf{x}$ beamformer outputs, $(\cdot)$ = LCMV/GSC (gsc)
$\mathbf{d}_\theta(\omega)$	Array response (or steering) vector
$\mathbf{b}_{(\cdot)}(\theta,\omega)$ and $\hat{\mathbf{b}}_{(\cdot)}(\theta,\omega)$	$\mathbf{W}_{(\cdot)}^H\mathbf{d}_\theta(\omega)$ and $\widehat{\mathbf{W}}_{(\cdot)}^H\mathbf{d}_\theta(\omega)$, $(\cdot)$ = LCMV/GSC
$b_{(\cdot)}(\theta,\omega)$ and $\hat{b}_{(\cdot)}(\theta,\omega)$	$\mathbf{w}_{(\cdot)}^H\mathbf{d}_\theta(\omega)$ and $\hat{\mathbf{w}}_{(\cdot)}^H\mathbf{d}_\theta(\omega)$, $(\cdot)$ = LCMV/GSC (gsc)
$\mathbf{\Phi} = \mathbf{G}(\mathbf{G}^H\mathbf{G})^{-1}$	Where the quiescent beamformer is $\mathbf{\Phi}\mathbf{f}$
$\Delta_{E\times E}$	Non-unique square root of $\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G}$
$\nabla_{(N-E)\times(N-E)}$	Non-unique square root of $\mathbf{G}_\perp^H\mathbf{R}\mathbf{G}_\perp$
N	Array Dimension
E	Number of signals
L	Sample support of secondary data
K	$L - N + E + 1$
$\stackrel{\mathcal{F}}{\longleftrightarrow}$	Fourier transform pair
$\Gamma(\cdot)$	Univariate Real Gamma function
$\tilde{\Gamma}_N(\cdot)$	Multivariate Complex Gamma function (see eq(3.31))
$\mathfrak{P}(\mathbf{V}\vert\mathbf{H})$	Weighted projector $\mathbf{H} - \mathbf{H}\mathbf{V}(\mathbf{V}^H\mathbf{H}\mathbf{V})^{-1}\mathbf{V}^H\mathbf{H}$
$\vert\vert\mathbf{p}_\theta(\omega)\vert\vert^2$	LCMV variation projector $\mathbf{d}_\theta^H(\omega)\ \mathfrak{P}(\mathbf{G}\vert\mathbf{R}^{-1})\mathbf{d}_\theta(\omega)$
$\vert\vert\tilde{\mathbf{p}}_\theta(\omega)\vert\vert^2$	GSC variation projector $\mathbf{d}_\theta^H(\omega)\ \mathbf{G}_\perp(\mathbf{G}_\perp^H\ \mathbf{R}\ \mathbf{G}_\perp)^{-1}\mathbf{G}_\perp^H\mathbf{d}_\theta(\omega)$
t_{GLRT}	GLRT test decision statistic
$\tilde{t}_{GLRT} = t_{GLRT} - 1$	Related GLRT statistic
t_{AMF}	AMF test decision statistic
η	Threshold for GLRT and AMF detectors

Symbol	Definition
$IG_k(\cdot)$	Incomplete Gamma function (see Section 3.5.2)
δ^2	SNR $\mathbf{s}^H\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G}\mathbf{s}$
δ_β	$\delta \cdot \sqrt{\beta}$
$\delta_\perp^2$	$\mathbf{s}^H\mathbf{V}^H\mathfrak{P}(\mathbf{G}\vert\mathbf{R}^{-1})\mathbf{V}\mathbf{s}$
$\delta(\cdot)$	Real Dirac Delta Impulse function
$\tilde{\delta}(\cdot)$	Complex Dirac Delta Impulse function (see Section 3.7)
$\sim$	"Distributed as" and/or "has density"
$\chi_N^2, F_{N,M}, \beta_{N,M}$	See Section 3.5.2
$P_{CAPON}(\boldsymbol{\theta},\boldsymbol{\omega})$	$\kappa \cdot (\mathbf{G}_{\boldsymbol{\theta}}^H(\boldsymbol{\omega})\mathbf{R}^{-1}\mathbf{G}_{\boldsymbol{\theta}}(\boldsymbol{\omega}))^{-1}$, also MVDR PSE
$\widehat{P}_{CAPON}(\boldsymbol{\theta},\boldsymbol{\omega})$	$\kappa \cdot (\mathbf{G}_{\boldsymbol{\theta}}^H(\boldsymbol{\omega})\widehat{\mathbf{R}}^{-1}\mathbf{G}_{\boldsymbol{\theta}}(\boldsymbol{\omega}))^{-1}$, also MVDR PSE
$\mathcal{CN}(\cdot)$	Complex Gaussian distributed
$\mathcal{CEC}(\cdot)$	Complex elliptically contoured (EC) distributed
$\mathcal{CMEC}(\cdot)$	Complex multivariate EC distributed
$\mathcal{SCEC}(\cdot)$	Singular Complex EC distributed
$\mathcal{CT}(\cdot)$	Complex multivariate t distributed
$\mathcal{T}_{(N-E)\times E}$	Complex t-distributed matrix
$\phi(\cdot)$, $g(\cdot)$	Functional form of data matrix c.f., pdf
$g_{M:N}(\cdot)$	Function operator defined in eq(3.89)
κ	See sentence following eq(3.114)
$\triangleq$	"Defined as"
$\text{vec}(\cdot)$	Stacks all columns of matrix into vector
${}_pF_q(\cdot)$	Generalized hypergeometric function
$J(\cdot \rightarrow \cdot)$	Jacobian
$\lambda_i(\mathbf{A})$	i-th eigenvalue of matrix $\mathbf{A}$
$E\{\cdot\}$	Statistical a priori expectation
$E\{\cdot\vert\cdot\}$	Statistical conditional expectation
$Pr(A)$	Probability of event A
$Pr(A\vert B)$	Probability of event A conditioned on B
$P_{\mathbf{A}}$	PDF of random matrix $\mathbf{A}$
$P_{\mathbf{A}}(\mathbf{A}_0)$	Dummy variable $\mathbf{A}_0$ introduced for integrals involving $P_{\mathbf{A}}$
$\Re^N$	Euclidean N-space
$\mathfrak{C}^N$	Complex Euclidean N-space
j	$\sqrt{-1}$
$(\cdot)^*$	Superscript for elementwise conjugation
$(\cdot)^T$	Superscript for matrix transposition
$(\cdot)^H$	Superscript for matrix conjugate transposition
$(\cdot)^+$	Superscript for matrix pseudo-inverse
$\text{Re}(\cdot)$ and $\text{Im}(\cdot)$	Real and Imaginary parts
$\text{tr}(\cdot)$	Matrix trace
$\vert\cdot\vert$	Matrix determinant or absolute value
$\vert\vert\cdot\vert\vert^2$	Vector norm
$\otimes$	Kronecker matrix product

Acronym	Meaning
AAD	Adaptive Array Detection
AMF	Adaptive Matched Filter
cdf	cumulative distribution function
c.f.	characteristic function
CFAR	Constant False Alarm Rate
CSS	Complex Spherically Symmetric
d.o.f.	degrees of freedom
EC	Elliptically Contoured
GLRT	Generalized Likelihood Ratio Test
GSC	Generalized Sidelobe Cancellor
GWPT	Generalized Wishart Partition Theorem 3.6.1
hpd	hermitian positive definite
iid	independent identically distributed
iff	if and only if
LCMV	Linearly Constrained Minimum Variance
LEC	Louiville EC
LMS	Least Mean Squares
LRT	Likelihood Ratio Test
MEC	Multivariate EC
MF	Matched Filter
ML	Maximum Likelihood
MVDR	Minimum Variance Distortionless Response
PD	Probability of Detection
pdf	probability density function
PFA	Probability of False Alarm
PSE	Power Spectral Estimation
RLS	Recursive Least Squares
r.v.	random variable
r. vc.	random vector
SCB	Sample Covariance Based
SCM	Sample Covariance Matrix
SMI	Sample Matrix Inversion
SNR	Signal-to-Noise Ratio
SS	Spherically Symmetric
UMP	Uniformly Most Powerful
UMPI	UMP Invariant
WET	Wishart Expectation Theorem 3.6.0.4
WPT	Wishart Partition Theorem

Appendix B

Proof of the Complex GWPT

In this appendix we provide a detailed proof of the GWPT Theorem 3.6.1.

Part (i): The marginals for both $\tilde{\mathbf{X}}_1$ and $\tilde{\mathbf{X}}_2$ follow directly from Proposition 3.5.5.3 with appropriate choices of matrices $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ of the proposition. The resulting pdfs for the matrices $\tilde{\mathbf{D}}_{11}$ and $\tilde{\mathbf{D}}_{22}$ then follow directly from the WET Theorem 3.6.0.4.

Parts (ii)-(v): Let $\tilde{\mathbf{R}}$ be partitioned as in eq(3.100) and let

$$\begin{aligned} \tilde{\mathbf{R}}_{11.2} &\triangleq \tilde{\mathbf{R}}_{11} - \tilde{\mathbf{R}}_{12}\tilde{\mathbf{R}}_{22}^{-1}\tilde{\mathbf{R}}_{21} \\ \tilde{\mathbf{M}}_{1.2} &\triangleq \tilde{\mathbf{M}}_1 + \tilde{\mathbf{R}}_{12}\tilde{\mathbf{R}}_{22}^{-1}(\tilde{\mathbf{X}}_2 - \tilde{\mathbf{M}}_2). \end{aligned} \tag{B.1}$$

Consider the linearly transformed data matrix given by

$$\tilde{\mathbf{A}}_0\tilde{\mathbf{X}} = \begin{bmatrix} \mathbf{I}_M & -\tilde{\mathbf{R}}_{12}\tilde{\mathbf{R}}_{22}^{-1} \\ \mathbf{0} & \mathbf{I}_{(N-M)} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{X}}_1 \\ \tilde{\mathbf{X}}_2 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{Y}}_1 \\ \tilde{\mathbf{Y}}_2 \end{bmatrix} = \tilde{\mathbf{Y}} \tag{B.2}$$

where the distribution of $\tilde{\mathbf{Y}}$ easily follows from Proposition 3.5.5.3; namely,

$$\tilde{\mathbf{Y}} \sim \mathcal{CMEC}_{NL}\left\{ \begin{bmatrix} \tilde{\mathbf{M}}_1 - \tilde{\mathbf{R}}_{12}\tilde{\mathbf{R}}_{22}^{-1}\tilde{\mathbf{M}}_2 \\ \tilde{\mathbf{M}}_2 \end{bmatrix}, \mathbf{I}_L \otimes \begin{bmatrix} \tilde{\mathbf{R}}_{11.2} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{R}}_{22} \end{bmatrix}, g(\cdot) \right\}. \tag{B.3}$$

Since the transformation is nonsingular, the matrix $\tilde{\mathbf{X}}$ can be written in terms of the matrix $\tilde{\mathbf{Y}}$:

$$\begin{bmatrix} \tilde{\mathbf{X}}_1 \\ \tilde{\mathbf{X}}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{I}_M & \tilde{\mathbf{R}}_{12}\tilde{\mathbf{R}}_{22}^{-1} \\ \mathbf{0} & \mathbf{I}_{(N-M)} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{Y}}_1 \\ \tilde{\mathbf{Y}}_2 \end{bmatrix}. \tag{B.4}$$

Noting that the Jacobian for this transformation is given by $J(\tilde{\mathbf{Y}} \to \tilde{\mathbf{X}}) = |\tilde{\mathbf{A}}_0|^{2L} = 1$ and that $|\tilde{\mathbf{R}}| = |\tilde{\mathbf{R}}_{11.2}||\tilde{\mathbf{R}}_{22}|$, eq(B.3) together with Proposition 3.5.1.1 allow the density of $\tilde{\mathbf{X}}$ to written in the convenient form

$$|\tilde{\mathbf{R}}_{11.2}|^{-L}|\tilde{\mathbf{R}}_{22}|^{-L}g\left[\mathrm{tr}\tilde{\mathbf{R}}_{11.2}^{-1}(\tilde{\mathbf{X}}_1-\tilde{\mathbf{M}}_{1.2})(\tilde{\mathbf{X}}_1-\tilde{\mathbf{M}}_{1.2})^H+\mathrm{tr}\tilde{\mathbf{R}}_{22}^{-1}(\tilde{\mathbf{X}}_2-\tilde{\mathbf{M}}_2)(\tilde{\mathbf{X}}_2-\tilde{\mathbf{M}}_2)^H\right] = P_{\tilde{\mathbf{X}}_1,\tilde{\mathbf{X}}_2}. \quad \text{(B.5)}$$

Next we consider the conditional density of $\tilde{\mathbf{X}}_1$ given $\tilde{\mathbf{X}}_2$. The density of $\tilde{\mathbf{X}}_1$ conditioned on $\tilde{\mathbf{X}}_2$ is likewise complex MEC as seen from Bayes' theorem $P_{\tilde{\mathbf{X}}_1|\tilde{\mathbf{X}}_2}=P_{\tilde{\mathbf{X}}_1,\tilde{\mathbf{X}}_2}/P_{\tilde{\mathbf{X}}_2}$. With $\tilde{\mathbf{M}}_1=\mathbf{0}$ and $\tilde{\mathbf{M}}_2=\mathbf{0}$, it follows that $\tilde{\mathbf{M}}_{1.2}=\tilde{\mathbf{T}}^H\tilde{\mathbf{X}}_2$ where $\tilde{\mathbf{T}}=\tilde{\mathbf{R}}_{22}^{-1}\tilde{\mathbf{R}}_{21}$. From part (i) of this theorem recall that $P_{\tilde{\mathbf{X}}_2}=|\tilde{\mathbf{R}}_{22}|^{-L}g_{(N-M)L:NL}\left[\mathrm{tr}\tilde{\mathbf{R}}_{22}^{-1}\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H\right]$. Conditioning on $\tilde{\mathbf{X}}_2=\tilde{\mathbf{A}}$, pick $\tilde{\mathbf{A}}_\perp$ $(L-N+M)\times L$ such that

$$\tilde{\mathbf{X}}_2\tilde{\mathbf{A}}_\perp^H=\mathbf{0}_{(N-M)\times(L-N+M)} \qquad \text{and} \qquad \tilde{\mathbf{A}}_\perp\tilde{\mathbf{A}}_\perp^H=\mathbf{I}_{(L-N+M)}. \quad \text{(B.6)}$$

Clearly the matrix

$$\left[\tilde{\mathbf{X}}_2^H|\tilde{\mathbf{A}}_\perp^H\right]=\tilde{\mathbf{C}}_{L\times L}^H \quad \text{(B.7)}$$

is nonsingular. Note that

$$\begin{aligned}\mathbf{I}_L=\tilde{\mathbf{C}}^H\tilde{\mathbf{C}}^{-H}\tilde{\mathbf{C}}^{-1}\tilde{\mathbf{C}}=\tilde{\mathbf{C}}^H(\tilde{\mathbf{C}}\tilde{\mathbf{C}}^H)^{-1}\tilde{\mathbf{C}}&=\left[\tilde{\mathbf{X}}_2^H|\tilde{\mathbf{A}}_\perp^H\right]\begin{bmatrix}\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H & \tilde{\mathbf{X}}_2\tilde{\mathbf{A}}_\perp^H\\ \tilde{\mathbf{A}}_\perp\tilde{\mathbf{X}}_2^H & \tilde{\mathbf{A}}_\perp\tilde{\mathbf{A}}_\perp^H\end{bmatrix}^{-1}\begin{bmatrix}\tilde{\mathbf{X}}_2\\ \tilde{\mathbf{A}}_\perp\end{bmatrix} \\ &=\tilde{\mathbf{X}}_2^H(\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H)^{-1}\tilde{\mathbf{X}}_2+\tilde{\mathbf{A}}_\perp^H\tilde{\mathbf{A}}_\perp\end{aligned} \quad \text{(B.8)}$$

where the last equality follows from eq(B.6). Recall the following relations

$$\tilde{\mathbf{X}}\tilde{\mathbf{X}}^H=\begin{bmatrix}\tilde{\mathbf{X}}_1\tilde{\mathbf{X}}_1^H & \tilde{\mathbf{X}}_1\tilde{\mathbf{X}}_2^H\\ \tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_1^H & \tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H\end{bmatrix}=\begin{bmatrix}\tilde{\mathbf{D}}_{11} & \tilde{\mathbf{D}}_{12}\\ \tilde{\mathbf{D}}_{21} & \tilde{\mathbf{D}}_{22}\end{bmatrix}=\tilde{\mathbf{D}} \quad \text{(B.9)}$$

from which it follows that

$$\begin{aligned}\tilde{\mathbf{D}}_{11.2}&=\tilde{\mathbf{D}}_{11}-\tilde{\mathbf{D}}_{12}\tilde{\mathbf{D}}_{22}^{-1}\tilde{\mathbf{D}}_{21}\\ &=\tilde{\mathbf{X}}_1\tilde{\mathbf{X}}_1^H-\tilde{\mathbf{X}}_1\tilde{\mathbf{X}}_2^H(\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H)^{-1}\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_1^H\\ &=\tilde{\mathbf{X}}_1\left[\mathbf{I}_L-\tilde{\mathbf{X}}_2^H(\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H)^{-1}\tilde{\mathbf{X}}_2\right]\tilde{\mathbf{X}}_1^H\\ &=\tilde{\mathbf{X}}_1\tilde{\mathbf{A}}_\perp^H\tilde{\mathbf{A}}_\perp\tilde{\mathbf{X}}_1^H.\end{aligned} \quad \text{(B.10)}$$

The last equality follows from eq(B.8). With the conditioning maintained, consider the nonsingular linear transformation

$$\tilde{\mathbf{X}}_1\tilde{\mathbf{C}}^H=\tilde{\mathbf{X}}_1\left[\tilde{\mathbf{X}}_2^H|\tilde{\mathbf{A}}_\perp^H\right]=\left[(\tilde{\mathbf{D}}_{12})_{M\times(N-M)}|(\tilde{\mathbf{V}})_{M\times(L-N+M)}\right]\triangleq\tilde{\mathbf{U}}_{M\times L} \quad \text{(B.11)}$$

and note that $\tilde{\mathbf{V}}\tilde{\mathbf{V}}^H=\tilde{\mathbf{D}}_{11.2}$. By eq(B.6) and Proposition 3.5.5.3 it follows that

$$P_{\tilde{\mathbf{V}},\tilde{\mathbf{D}}_{12}|\tilde{\mathbf{X}}_2}=|\tilde{\mathbf{R}}_{11.2}|^{-L}|\tilde{\mathbf{R}}_{22}|^{-L}|\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H|^{-M} \quad \text{(B.12)}$$

$$\times g\left[\mathrm{tr}\tilde{\mathbf{R}}_{11.2}^{-1}(\tilde{\mathbf{D}}_{12}-\tilde{\mathbf{T}}^H\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H)(\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H)^{-1}(\tilde{\mathbf{D}}_{12}-\tilde{\mathbf{T}}^H\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H)^H+\mathrm{tr}\tilde{\mathbf{R}}_{11.2}^{-1}\tilde{\mathbf{V}}\tilde{\mathbf{V}}^H+\mathrm{tr}\tilde{\mathbf{R}}_{22}^{-1}\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H\right]$$
$$\div \qquad P_{\tilde{\mathbf{X}}_2}.$$

Note that the density $P_{\tilde{\mathbf{V}},\tilde{\mathbf{D}}_{12}|\tilde{\mathbf{X}}_2}$ is complex MEC in the variables $(\tilde{\mathbf{V}},\tilde{\mathbf{D}}_{12})$. Thus, considering the linear transformation

$$\tilde{\mathbf{R}}_{11.2}^{-1/2}\cdot\left(\tilde{\mathbf{U}}-[\tilde{\mathbf{T}}^H(\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H)\ \mid\ \mathbf{0}\ ]\right)\cdot\begin{bmatrix}(\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H)^{-1/2} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}_{L-N+M}\end{bmatrix}=[\tilde{\mathbf{D}}_0|\tilde{\mathbf{V}}] \qquad \text{(B.13)}$$

which simply subtracts the mean off of $\tilde{\mathbf{D}}_{12}$ and subsequently whitens $\tilde{\mathbf{D}}_{12}$, the density for $(\tilde{\mathbf{V}},\tilde{\mathbf{D}}_0)$ given $\tilde{\mathbf{X}}_2$ likewise follow from Proposition 3.5.5.3:

$$P_{\tilde{\mathbf{V}},\tilde{\mathbf{D}}_0|\tilde{\mathbf{X}}_2}=|\tilde{\mathbf{R}}_{11.2}|^{-(L-N+M)}|\tilde{\mathbf{R}}_{22}|^{-L}g\left[\mathrm{tr}\tilde{\mathbf{D}}_0\tilde{\mathbf{D}}_0^H+\mathrm{tr}\tilde{\mathbf{R}}_{11.2}^{-1}\tilde{\mathbf{V}}\tilde{\mathbf{V}}^H+\mathrm{tr}\tilde{\mathbf{R}}_{22}^{-1}\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H\right]/P_{\tilde{\mathbf{X}}_2} \text{(B.14)}$$

Multiplication by $P_{\tilde{\mathbf{X}}_2}$ yields the joint density for $(\tilde{\mathbf{D}}_0,\tilde{\mathbf{V}},\tilde{\mathbf{X}}_2)$, which apparently is likewise complex MEC. Thus, integrating out $\tilde{\mathbf{D}}_0$ to obtain the joint pdf for $(\tilde{\mathbf{V}},\tilde{\mathbf{X}}_2)$ we lose $M(N-M)$ degrees of freedom (d.o.f.) from the total NL d.o.f. The result is ergo given by

$$P_{\tilde{\mathbf{V}},\tilde{\mathbf{X}}_2}=|\tilde{\mathbf{R}}_{11.2}|^{-(L-N+M)}|\tilde{\mathbf{R}}_{22}|^{-L}g_{NL-N(M-N):NL}\left[\mathrm{tr}\tilde{\mathbf{R}}_{11.2}^{-1}\tilde{\mathbf{V}}\tilde{\mathbf{V}}^H+\mathrm{tr}\tilde{\mathbf{R}}_{22}^{-1}\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H\right]. \qquad \text{(B.15)}$$

Recalling that $\tilde{\mathbf{D}}_{11.2}=\tilde{\mathbf{V}}\tilde{\mathbf{V}}^H$ and $\tilde{\mathbf{D}}_{22}=\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H$, subsequent application of the WET 3.6.0.4 twice gives the joint pdf for $(\tilde{\mathbf{D}}_{11.2},\tilde{\mathbf{D}}_{22})$ found in part (iv) of the GWPT.

Similarly, considering the joint pdf for $(\tilde{\mathbf{D}}_0,\tilde{\mathbf{V}},\tilde{\mathbf{X}}_2)$, we can whiten $\tilde{\mathbf{X}}_2$ and subsequently integrate out both $\tilde{\mathbf{D}}_0$ and the whitened $\tilde{\mathbf{X}}_2$. We then lose a total of $M(N-M)+(N-M)L=(N-M)(L+M)$ d.o.f. Hence, noting that $NL-(N-M)(L+M)=M(L-N+M)$ the marginals for $\tilde{\mathbf{V}}$ are shown to be given by

$$P_{\tilde{\mathbf{V}}}=|\tilde{\mathbf{R}}_{11.2}|^{-(L-N+M)}g_{M(L-N+M):NL}\left[\mathrm{tr}\tilde{\mathbf{R}}_{11.2}^{-1}\tilde{\mathbf{V}}\tilde{\mathbf{V}}^H\right]. \qquad \text{(B.16)}$$

Recalling the relation $\tilde{\mathbf{D}}_{11.2}=\tilde{\mathbf{V}}\tilde{\mathbf{V}}^H$, and applying the WET 3.6.0.4 yields the marginals for $\tilde{\mathbf{D}}_{11.2}$ given in part (ii) of the GWPT.

Conditioning on $\tilde{\mathbf{X}}_2$, we can likewise entertain the nonsingular transformation

$$\tilde{\mathbf{U}}\cdot\begin{bmatrix}(\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H)^{-1} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}_{L-N+M}\end{bmatrix}=[\tilde{\mathcal{T}}^H|\tilde{\mathbf{V}}] \qquad \text{(B.17)}$$

which by eq(B.13) and Proposition 3.5.5.3 has density

$$P_{\tilde{\mathbf{V}},\tilde{\mathcal{T}}|\tilde{\mathbf{X}}_2}=|\tilde{\mathbf{R}}_{11.2}|^{-L}|\tilde{\mathbf{R}}_{22}|^{-L}|\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H|^M \qquad \text{(B.18)}$$

$$\times\ g\left[\mathrm{tr}\tilde{\mathbf{R}}_{11.2}^{-1}(\tilde{\mathcal{T}}-\tilde{\mathbf{T}})^H(\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H)(\tilde{\mathcal{T}}-\tilde{\mathbf{T}})+\mathrm{tr}\tilde{\mathbf{R}}_{11.2}^{-1}\tilde{\mathbf{V}}\tilde{\mathbf{V}}^H+\mathrm{tr}\tilde{\mathbf{R}}_{22}^{-1}\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H\right]/P_{\tilde{\mathbf{X}}_2}.$$

Multiplication by $P_{\tilde{\mathbf{X}}_2}$ yields the joint density for $(\tilde{\mathbf{V}},\tilde{\mathcal{T}},\tilde{\mathbf{X}}_2)$. Recalling that $\tilde{\mathbf{D}}_{11.2}=\tilde{\mathbf{V}}\tilde{\mathbf{V}}^H$ and $\tilde{\mathbf{D}}_{22}=\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H$, subsequent application of the WET 3.6.0.4 twice gives the joint pdf for $(\tilde{\mathbf{D}}_{11.2},\tilde{\mathcal{T}},\tilde{\mathbf{D}}_{22})$ found in part (iii) of the GWPT.

Note likewise from eq(B.18) that multiplication by $P_{\tilde{\mathbf{X}}_2}$ and subsequently integrating over $\tilde{\mathbf{V}}$ we lose $M(L-N+M)$ d.o.f. and are left with the joint pdf for $(\tilde{\mathcal{T}},\tilde{\mathbf{X}}_2)$:

$$P_{\tilde{\mathcal{T}},\tilde{\mathbf{X}}_2}=|\tilde{\mathbf{R}}_{11.2}|^{-(N-M)}|\tilde{\mathbf{R}}_{22}|^{-L}|\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H|^M \times g_{NL-M(L-N+M):NL}\left[\mathrm{tr}\tilde{\mathbf{R}}_{11.2}^{-1}(\tilde{\mathcal{T}}-\tilde{\mathbf{T}})^H(\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H)(\tilde{\mathcal{T}}-\tilde{\mathbf{T}})+\mathrm{tr}\tilde{\mathbf{R}}_{22}^{-1}\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H\right]. \tag{B.19}$$

Again, applying the WET 3.6.0.4 to $\tilde{\mathbf{D}}_{22}=\tilde{\mathbf{X}}_2\tilde{\mathbf{X}}_2^H$, yields the joint density for $(\tilde{\mathcal{T}},\tilde{\mathbf{D}}_{22})$ found in part (v) of the GWPT:

$$(\tilde{\mathcal{T}},\tilde{\mathbf{D}}_{22})\sim\left[\frac{\pi^{L(N-M)}}{\tilde{\Gamma}_{(N-M)}(L)}\right]|\tilde{\mathbf{R}}_{22}|^{-L}|\tilde{\mathbf{R}}_{11.2}|^{-(N-M)}|\tilde{\mathbf{D}}_{22}|^{(L-N+2M)} \times g_{NL-M(L-N+M):NL}\left[\mathrm{tr}\tilde{\mathbf{R}}_{11.2}^{-1}(\tilde{\mathcal{T}}-\tilde{\mathbf{T}})^H\tilde{\mathbf{D}}_{22}(\tilde{\mathcal{T}}-\tilde{\mathbf{T}})+\mathrm{tr}\tilde{\mathbf{R}}_{22}^{-1}\tilde{\mathbf{D}}_{22}\right]. \tag{B.20}$$

Rearranging this density function's argument slightly, the marginals for $\tilde{\mathcal{T}}$ are in theory given by the multivariate integral

$$\tilde{\mathcal{T}}\sim k_3\cdot\int_{\tilde{\mathbf{D}}_{22}=\tilde{\mathbf{D}}_{22}^H>0}|\tilde{\mathbf{D}}_{22}|^{(L+M)-(N-M)} \times g_{(L+M)(N-M):NL}\left\{\mathrm{tr}\tilde{\mathbf{D}}_{22}\left[(\tilde{\mathcal{T}}-\tilde{\mathbf{T}})\tilde{\mathbf{R}}_{11.2}^{-1}(\tilde{\mathcal{T}}-\tilde{\mathbf{T}})^H+\tilde{\mathbf{R}}_{22}^{-1}\right]\right\}d\tilde{\mathbf{D}}_{22} \tag{B.21}$$

where $k_3=|\tilde{\mathbf{R}}_{22}|^{-L}|\tilde{\mathbf{R}}_{11.2}|^{-(N-M)}\pi^{L(N-M)}/\tilde{\Gamma}_{(N-M)}(L)$. To evaluate this integral note the rather useful integration lemma for an $N'\times N'$ hpd matrix $\tilde{\mathbf{B}}$

$$\int_{\tilde{\mathbf{B}}=\tilde{\mathbf{B}}^H>0}|\tilde{\mathbf{B}}|^{L'-N'}\ \tilde{g}_{N'L':N'L'}[\ \mathrm{tr}\tilde{\mathbf{B}}\tilde{\Sigma}^{-1}]\ d(\tilde{\mathbf{B}})=\frac{\tilde{\Gamma}_{N'}(L')}{\pi^{N'L'}}|\tilde{\Sigma}|^{L'} \tag{B.22}$$

which follows from the fact that the density in eq(3.116) of the WET must have unit area. Note that $\tilde{g}(r)$ is the functional form for a density of $N'L'$ complex d.o.f., *e.g.* a pdf for some data matrix of dimensions $N'\times L'$. Thus, the integral identity holds whenever integrating an $N'L'$ dimensional pdf over an hpd matrix of $(N')^2$ elements. Equating $N'=(N-M)$ and $L'=(L+M)$ it follows therefore that the marginals for $\tilde{\mathcal{T}}$ are as given in part (ii) of the GWPT.

Indeed, it is possible to employ this same reasoning to deduce the marginals for $\tilde{\mathcal{T}}$ starting from the joint pdf for $(\tilde{\mathbf{D}}_{11.2}, \tilde{\mathcal{T}}, \tilde{\mathbf{D}}_{22})$. Note that one can divide eq(3.124) by $P_{\tilde{\mathbf{D}}_{11.2}}$ to obtain the conditional distribution $P_{\tilde{\mathcal{T}},\tilde{\mathbf{D}}_{22}|\tilde{\mathbf{D}}_{11.2}}$. Since $P_{\tilde{\mathcal{T}},\tilde{\mathbf{D}}_{22}|\tilde{\mathbf{D}}_{11.2}}$ is a density for $(L+M)(N-M)$ d.o.f., the integration lemma in eq(B.22) still applies and it yields precisely the marginals for $\tilde{\mathcal{T}}$ found in part (ii) of the GWPT. This likewise proves that $\tilde{\mathcal{T}}$ and $\tilde{\mathbf{D}}_{11.2}$ are independent. Note that the pdf for $\tilde{\mathcal{T}}$ is independent of $g(\cdot)$ the functional form of the density of $\tilde{\mathbf{X}}$. Considering that $\tilde{\mathcal{T}}$ is a scale invariant function of $\tilde{\mathbf{X}}$, the invariance of its pdf to $g(\cdot)$ follows from Proposition 3.5.4.5. Q.E.D.

Appendix C

Covariance of ML/LCMV Weightings

Let $\mathbf{Z}_1 = [\boldsymbol{\alpha}_1|\boldsymbol{\alpha}_2|\cdots|\boldsymbol{\alpha}_E]$, $\widehat{\mathbf{W}}_{ML} - \mathbf{W}_{ML} = [\hat{\mathbf{w}}_1|\hat{\mathbf{w}}_2|\cdots|\hat{\mathbf{w}}_E]$, $\Delta^{-H} = [\boldsymbol{\delta}_1|\boldsymbol{\delta}_2|\cdots|\boldsymbol{\delta}_E]$, and $\boldsymbol{\delta}_k = (b_{1,k}, b_{2,k}, \ldots, b_{E,k})^T$. The k–th column of the centered weight matrix is $\hat{\mathbf{w}}_k = -\mathbf{G}_\perp \mathcal{T} \boldsymbol{\delta}_k$. Hence, $\hat{\mathbf{w}}_k \hat{\mathbf{w}}_l^H = \mathbf{G}_\perp \mathcal{T} \boldsymbol{\delta}_k \boldsymbol{\delta}_l^H \mathcal{T}^H \mathbf{G}_\perp^H$. Noting that $\mathbf{Z}_1 \boldsymbol{\delta}_k = b_{1,k}\boldsymbol{\alpha}_1 + b_{2,k}\boldsymbol{\alpha}_2 + \cdots + b_{E,k}\boldsymbol{\alpha}_E$, we obtain the equivalence

$$\mathcal{T}\boldsymbol{\delta}_k\boldsymbol{\delta}_l^H\mathcal{T}^H = (\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{Z}_2^H\left(\sum_{n=1}^{E}\sum_{m=1}^{E} b_{n,k}\boldsymbol{\alpha}_n\boldsymbol{\alpha}_m^H b_{m,l}^*\right)\mathbf{Z}_2(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}. \tag{C.1}$$

Conditioning on $\mathbf{Z}_2$, recalling that $\mathbf{Z}_1$ is independent of $\mathbf{Z}_2$ and has independent columns, it follows that

$$E\{\boldsymbol{\alpha}_n\boldsymbol{\alpha}_m^H\} = \begin{cases} \mathbf{I}_L & , n = m \\ \mathbf{0}_{L\times L} & , \text{otherwise.} \end{cases} \tag{C.2}$$

Thus, the following conditional expectation holds

$$E\left\{\mathcal{T}\boldsymbol{\delta}_k\boldsymbol{\delta}_l^H\mathcal{T}^H|\mathbf{Z}_2\right\} = (\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{Z}_2^H\left(\sum_{n=1}^{E} b_{n,k}\mathbf{I}_L b_{n,l}^*\right)\mathbf{Z}_2(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1} = (\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\cdot\boldsymbol{\delta}_l^H\boldsymbol{\delta}_k. \tag{C.3}$$

Removing the conditioning on $\mathbf{Z}_2$ yields $E\{\mathcal{T}\boldsymbol{\delta}_k\boldsymbol{\delta}_l^H\mathcal{T}^H\} = \mathbf{I}_{N-E}\cdot\boldsymbol{\delta}_l^H\boldsymbol{\delta}_k/(L-N+E)$. From eq(4.43) note that $\boldsymbol{\delta}_l^H\boldsymbol{\delta}_k = [\Delta^{-1}\Delta^{-H}]_{l,k} = [(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}]_{l,k}$ from which we conclude that $E\{\hat{\mathbf{w}}_k\hat{\mathbf{w}}_l^H\} = [(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}]_{l,k}\,\mathbf{G}_\perp\mathbf{G}_\perp^H\,/(L-N+E)$. The relation

$$\mathbf{R}^{-1} = \mathbf{R}^{-1/2}\mathbf{Q}^H\mathbf{Q}\mathbf{R}^{-1/2} = \mathcal{W}_{\mathcal{MC}}\mathcal{W}_{\mathcal{MC}}{}^H + \mathbf{G}_\perp\mathbf{G}_\perp^H \tag{C.4}$$

implies that $\mathbf{G}_\perp\mathbf{G}_\perp^H = \mathbf{R}^{-1} - \mathcal{W}_{\mathcal{MC}}\mathcal{W}_{\mathcal{MC}}{}^H$. Recall that $\mathcal{W}_{\mathcal{MC}} = \mathbf{W}_{ML}\Delta^H$. Consequently, it follows that $\mathcal{W}_{\mathcal{MC}}\mathcal{W}_{\mathcal{MC}}{}^H = \mathbf{R}^{-1}\mathbf{G}(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}\mathbf{G}^H\mathbf{R}^{-1}$ and that $\mathbf{G}_\perp\mathbf{G}_\perp^H = \mathfrak{P}(\mathbf{G}|\mathbf{R}^{-1})$.

Hence, the desired expectation $E\{\hat{\mathbf{w}}_k\hat{\mathbf{w}}_l^H\} = [(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}]_{l,k}\mathfrak{P}(\mathbf{G}|\mathbf{R}^{-1})/(L-N+E)$ and the covariance $\text{cov}(\widehat{\mathbf{W}}_{ML}) = (\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-T} \otimes \mathfrak{P}(\mathbf{G}|\mathbf{R}^{-1})/(L-N+E)$. Q.E.D.

Appendix D

Proof for Singular PDF of ML/LCMV Weightings

Recall the stochastic relation

$$\widehat{\mathbf{W}}_{ML} \overset{d}{=} \mathbf{W}_{ML} - \mathbf{G}_{\perp}(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{Z}_2^H\mathbf{Z}_1\Delta^{-H}. \tag{D.1}$$

Note that the matrix $\mathbf{Z}_1$ is an $L \times E$ complex normal data matrix distributed as $\mathcal{CN}_{LE}(\mathbf{0}, \mathbf{I}_E \otimes \mathbf{I}_L)$. Conditioning on $\mathbf{Z}_2$ it follows from Proposition 3.5.5.3 that

$$\mathbf{G}_{\perp}(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{Z}_2^H\mathbf{Z}_1\Delta^{-H} \sim \mathcal{CN}_{LE}\left[\mathbf{0}, (\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-T} \otimes \mathbf{G}_{\perp}(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{G}_{\perp}^H\right]. \tag{D.2}$$

Hence, the conditional c.f. of $\widehat{\mathbf{W}}_{ML}$ is given by

$$\exp\left\{j\mathrm{Re}[\mathrm{tr}(\mathbf{T}^H\mathbf{W}_{ML})]\right\} \cdot \exp\left\{-\mathrm{tr}[\mathbf{T}^H\mathbf{G}_{\perp}(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{G}_{\perp}^H\mathbf{T}(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1}]/4\right\}. \tag{D.3}$$

Let

$$[\mathbf{T}_0]_{(N-E)\times E} = \mathbf{G}_{\perp}^H\mathbf{T}(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1/2}. \tag{D.4}$$

Since the columns of $\mathbf{Z}_2^H$ are CSS it follows that $\mathbf{Q}_b\mathbf{Z}_2^H$ is identically distributed to $\mathbf{Z}_2$ and thus the argument of the c.f. exponent can be written

$$\mathbf{T}_0^H(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{T}_0 \overset{d}{=} \mathbf{T}_0^H\mathbf{Q}_b(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{Q}_b^H\mathbf{T}_0 \overset{d}{=} [\Delta_0^H\ \mathbf{0}](\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\begin{bmatrix}\Delta_0 \\ \mathbf{0}\end{bmatrix} \overset{d}{=} \Delta_0^H\Theta_2^{-1}\Delta_0 \tag{D.5}$$

where the unitary matrix was chosen in a fashion similar to eq(4.19). This choice implies that $\mathbf{T}_0^H\mathbf{T}_0 = \Delta_0^H\Delta_0$, and thus $\Delta_0 = (\mathbf{T}_0^H\mathbf{T}_0)^{1/2}$. The relations for the inverse of a partitioned matrix shown in eq(4.24) indicate that Θ_2 is the Schur complement of the white SCM $\mathbf{Z}_2^H\mathbf{Z}_2$. Hence, by part (ii) of the GWPT Theorem 3.6.1 it follows that $\Theta_2 \sim \mathcal{CW}_{N-E}(L - N +$

$2E, \mathbf{I}_{N-E})$. The conditional c.f. can then be written

$$\exp\left\{j\mathrm{Re}[\mathrm{tr}(\mathbf{T}^H\mathbf{W}_{ML})]\right\}\cdot\exp\left\{-\mathrm{tr}[\Theta_2^{-1}(\mathbf{T}_0^H\mathbf{T}_0)]/4\right\} =$$
$$\exp\left\{j\mathrm{Re}[\mathrm{tr}(\mathbf{T}^H\mathbf{W}_{ML})]\right\}\cdot\exp\left\{-\mathrm{tr}[\mathbf{T}^H\mathbf{G}_\perp\mathbf{G}_\perp^H\mathbf{T}(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1/2}\Theta_2^{-1}(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{-1/2}]/4\right\}.$$

Since the matrix $\mathbf{G}_\perp\mathbf{G}_\perp^H = \mathfrak{P}(\mathbf{G}|\mathbf{R}^{-1})$ (see Appendix C for this equality) is singular the conditional distribution is complex singular normal [63, 51]. As discussed in Section 3.7, this density can be written in the form

$$k_w\cdot|\Theta_2|^{(N-E)}\exp\left[-\mathrm{tr}\Theta_2\Sigma\right]\times\tilde{\delta}\left[\mathrm{vec}(\mathbf{G}^H\widehat{\mathbf{W}}_{ML}) - \mathrm{vec}(\mathbf{G}^H\mathbf{W}_{ML})\right] = P_{\widehat{\mathbf{W}}_{ML}|\Theta_2} \tag{D.6}$$

where

$$k_w = \pi^{-(N-E)E}|\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G}|^{(N-E)}\left\{\prod_{i=1}^{N-E}\lambda_i\left[\mathfrak{P}(\mathbf{G}|\mathbf{R}^{-1})\right]\right\}^{-E} \tag{D.7}$$

$$\Sigma = (\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{1/2}(\widehat{\mathbf{W}}_{ML} - \mathbf{W}_{ML})^H\mathfrak{P}(\mathbf{G}|\mathbf{R}^{-1})^+(\widehat{\mathbf{W}}_{ML} - \mathbf{W}_{ML})(\mathbf{G}^H\mathbf{R}^{-1}\mathbf{G})^{1/2}. \tag{D.8}$$

That the complex Wishart distribution integrated over the space of all hpd matrices is unity yields the following integral identity

$$\int_{\Theta_0=\Theta_0^H>0}|\Theta_0|^{K'-M'}\exp\left[-\mathrm{tr}(\Theta_0\Sigma_0)\right](d\Theta_0) = |\Sigma_0|^{-K'}\tilde{\Gamma}_{M'}(K') \tag{D.9}$$

from which eq(5.17) follows. A similar proof can be used to derive eq(5.19).

Appendix E

GSC Beamformer Output

The signal estimate given by the SCM based GSC can be written as

$$\hat{\mathbf{s}}_{GSC} = \widehat{\mathbf{W}}_{GSC}^H \mathbf{x} = \mathbf{s} + \widehat{\mathbf{W}}_{GSC}^H \mathbf{n}. \tag{E.1}$$

The data observation $\mathbf{x}$ from which the signal is estimated is independent of the GSC weight matrix by Theorem 3.5.4. The signal estimate is conditionally distributed as

$$\hat{\mathbf{s}}_{GSC} \sim \mathcal{CEC}_E \left[\mathbf{s}, \widehat{\mathbf{W}}_{GSC}^H \mathbf{R} \widehat{\mathbf{W}}_{GSC}, g_{E:N(L+1)}(\cdot)\right] = P_{\hat{\mathbf{s}}_{GSC}|\widehat{\mathbf{W}}_{GSC}}. \tag{E.2}$$

Note that

$$\widehat{\mathbf{W}}_{GSC}^H \mathbf{R} \widehat{\mathbf{W}}_{GSC} = \mathbf{\Phi}^H \left[\mathbf{R} - \hat{\mathbf{\Omega}}_{GSC}^H \mathbf{G}_\perp \mathbf{R} - \mathbf{R}\mathbf{G}_\perp \hat{\mathbf{\Omega}}_{GSC} + \hat{\mathbf{\Omega}}_{GSC}^H \mathbf{G}_\perp^H \mathbf{R} \mathbf{G}_\perp \hat{\mathbf{\Omega}}_{GSC}\right] \mathbf{\Phi}. \tag{E.3}$$

Recalling eq(5.25)–eq(5.28) for the adaptive weight matrix $\hat{\mathbf{\Omega}}_{GSC}$ the conditional covariance parameter can be written

$$\begin{aligned} &\widehat{\mathbf{W}}_{GSC}^H \mathbf{R} \widehat{\mathbf{W}}_{GSC} \stackrel{d}{=} \\ &\mathbf{\Phi}^H \Bigg\{ \mathbf{R}^{1/2}\mathbf{Q}_0^H \mathbf{I}_N \mathbf{Q}_0^H \mathbf{R}^{1/2} - \mathbf{R}^{1/2}\mathbf{Q}_0^H \begin{bmatrix} \mathbf{I}_{N-E} \\ \mathcal{T}_{GSC}^H \end{bmatrix} \begin{bmatrix} \mathbf{I}_{N-E} & | \mathbf{0}_{(N-E)\times E} \end{bmatrix} \mathbf{Q}_0 \mathbf{R}^{1/2} \\ &\quad - \mathbf{R}^{1/2}\mathbf{Q}_0^H \begin{bmatrix} \mathbf{I}_{N-E} \\ \mathbf{0}_{E\times(N-E)} \end{bmatrix} \begin{bmatrix} \mathbf{I}_{N-E} & | \mathcal{T}_{GSC} \end{bmatrix} \mathbf{Q}_0 \mathbf{R}^{1/2} \\ &\quad + \mathbf{R}^{1/2}\mathbf{Q}_0^H \begin{bmatrix} \mathbf{I}_{N-E} \\ \mathcal{T}_{GSC}^H \end{bmatrix} \begin{bmatrix} \mathbf{I}_{N-E} & | \mathcal{T}_{GSC} \end{bmatrix} \mathbf{Q}_0 \mathbf{R}^{1/2} \Bigg\} \mathbf{\Phi}. \end{aligned} \tag{E.4}$$

Simplifying yields the form

$$\widehat{\mathbf{W}}_{GSC}^H \mathbf{R} \widehat{\mathbf{W}}_{GSC} \stackrel{d}{=} \mathbf{\Phi}^H \left\{ \mathbf{R}^{1/2}\mathbf{Q}_0^H \begin{bmatrix} \mathbf{0}_{N-E} & \mathbf{0}_{(N-E)\times E} \\ \mathbf{0}_{E\times(N-E)} & (\mathbf{I}_E + \mathcal{T}_{GSC}^H \mathcal{T}_{GSC}) \end{bmatrix} \mathbf{Q}_0 \mathbf{R}^{1/2} \right\} \mathbf{\Phi}. \tag{E.5}$$

Recalling the matrix partition given in eq(5.27) we obtain

$$\begin{aligned} \widehat{\mathbf{W}}_{GSC}^{H}\mathbf{R}\widehat{\mathbf{W}}_{GSC} &\stackrel{d}{=} \mathbf{\Phi}^{H}\widetilde{\mathbf{G}}_{GSC}\left(\mathbf{I}_{E}+\mathcal{T}_{GSC}^{H}\mathcal{T}_{GSC}\right)\widetilde{\mathbf{G}}_{GSC}^{H}\mathbf{\Phi} \\ &= \mathbf{\Phi}^{H}\widetilde{\mathbf{G}}_{GSC}\ \mathbf{\Lambda}_{GSC}^{-1}\ \widetilde{\mathbf{G}}_{GSC}^{H}\mathbf{\Phi}. \end{aligned} \tag{E.6}$$

Appendix F

Selected Theorems

Proposition F.0.0.1 Ratio Density Theorem: *Given two real random variables x and y with joint density $P_{x,y}$, the pdf for the ratio $a = y/x$ is given by*

$$P_a = \int x_0 P_{x,y}(x_0, ax_0) dx_0. \tag{F.1}$$

To prove first condition on x, and note by Bayes' theorem that $P_{y|x} = P_{x,y}/P_x$. Thus, treating x as a deterministic constant it follows that the conditional cdf for a is

$$Pr(a \leq a_0|x) = Pr(y/x \leq a_0|x) = Pr(y \leq a_0 x|x). \tag{F.2}$$

Differentiate to obtain the conditional pdf $P_{a|x}(a_0) = x \cdot P_{y|x}(a_0 x) = x \cdot P_{xy}(x, a_0 x)/P_x(x)$. Finally, multiplying by the marginal for x to obtain the joint for a and x, and subsequent integration over x proves the intended.

Proposition F.0.0.2 *If the random matrix $\mathcal{T}$ is distributed according to the complex multivariate t in eq(5.14),* i.e. $\mathcal{T} \sim C\mathcal{T}_{(N-E)\times E}[\mathbf{0}_{(N-E)\times E},\ L - N + 2E,\ \mathbf{I}_{N-E},\ \mathbf{I}_E]$ *where $N \geq 2E$, then the matrix $\mathbf{\Lambda} = (\mathbf{I}_E + \mathcal{T}^H\mathcal{T})^{-1}$ is distributed according to the complex multivariate Beta pdf shown in eq(5.53).*

Proof: To prove this result we first derive the pdf for $\mathbf{D} = \mathcal{T}^H\mathcal{T}$, and subsequently perform a change of variables. Let $k = \tilde{\Gamma}_{N-E}(L+E)/[\tilde{\Gamma}_{N-E}(L)\tilde{\Gamma}_E(N-E)]$ be a normalizing constant.

Consider the expectation

$$
\begin{aligned}
E\{\ h(\mathcal{T}^H\mathcal{T})\ \} &\triangleq \int h(\mathcal{T}_0^H\mathcal{T}_0)\ P_{\mathcal{T}}(\mathcal{T}_0)\ d(\mathcal{T}_0) \\
&= k\cdot\int_{\mathbf{D}_0=\mathbf{D}_0^H>0} h(\mathbf{D}_0)\cdot|\mathbf{I}_E+\mathbf{D}_0|^{-(L+E)}\,|\mathbf{D}_0|^{N-2E}\,d(\mathbf{D}_0) \qquad \text{(F.3)} \\
&\triangleq E\{\ h(\mathbf{D})\ \}.
\end{aligned}
$$

The second equality follows from eq(5.14), the equivalence $|\mathbf{I}_{N-E}+\mathcal{T}_0\mathcal{T}_0^H| = |\mathbf{I}_E+\mathcal{T}_0^H\mathcal{T}_0|$, and Proposition 3.3.2.2. Comparing the integrand of the second equality to the definition of expectation it is shown that the pdf for $\mathbf{D}$ is given by $P_{\mathbf{D}}(\mathbf{D}_0) = k\cdot|\mathbf{I}_E+\mathbf{D}_0|^{-(L+E)}\,|\mathbf{D}_0|^{N-2E}$, a matrix generalization of the central complex F-distribution. Making the change of variables $\mathbf{\Lambda} = (\mathbf{I}_E+\mathbf{D})^{-1}$ note that the Jacobian is given by $J(\mathbf{D}\to\mathbf{\Lambda}) = |\mathbf{\Lambda}|^{-2E}$ [22, 64]. Considering the equivalences $(\mathbf{I}_E+\mathbf{D}) = \mathbf{\Lambda}^{-1}$ and $\mathbf{D} = \mathbf{\Lambda}^{-1}(\mathbf{I}_E-\mathbf{\Lambda})$, elementary probability proves the intended via $P_{\mathbf{\Lambda}}(\mathbf{\Lambda}_0) = J(\mathbf{D}\to\mathbf{\Lambda})\cdot P_{\mathbf{D}}[\mathbf{\Lambda}_0^{-1}(\mathbf{I}_E-\mathbf{\Lambda}_0)]$. Noting that $\tilde{\Gamma}_{N-E}(L+E)/\tilde{\Gamma}_{N-E}(L) = \tilde{\Gamma}_E(L+E)/\tilde{\Gamma}_E(L+2E-N)$ it is clear that $k = 1/\tilde{\beta}_E(L+2E-N,N-E)$. Q.E.D.

Bibliography

[1] E. J. Wegman, S. C. Schwartz, J. B. Thomas (Editors), *Topics in Non-Gaussian Signal Processing*, Springer-Verlag, New York (1989).

[2] E. Conte, M. Longo, "Characterization of Radar Clutter as a Spherically Invariant Random Process," *IEE Proceedings*, Part F, **134**(2), 191–197 (1987).

[3] E. Jakeman, P. N. Pusey, "Significance of the K-Distribution in Scattering Experiments," *Phys. Rev. Letters*, **4**(9), 546–550 (1978).

[4] K. D. Ward, "A Radar Sea Clutter Model and Its Application to Performance Assessment," *IEE Conference Publication 216*, Radar '82, London, (Oct. 1982).

[5] K. J. Sangston, K. R. Gerlach, "Coherent Detection of Radar Targets in a Non-Gaussian Background," *IEEE Trans. Aeros. Electr. Sys.*, Vol. 30, No. 2, 330–340 (1994).

[6] M. Rangaswamy, J. H. Michels, D. D. Weiner, "Multichannel Detection for Correlated Non-Gaussian Random Processes Based on Innovations," Vol.43, No. 8, 1915–1922 (1995).

[7] M. Rangaswamy, D. Weiner, A. Öztürk, "Non-Gaussian Random Vector Identification Using Spherically Invariant Random Processes," *IEEE Trans. Aeros. Electr. Sys.*, Vol. 29, No. 1, 111–124 (1993).

[8] R. Raghavan, N. Pulsone, "A Generalization of the Adaptive Matched Filter Receiver for Array Detection in a Class of Non-Gaussian Interference," *Proceedings of the Adaptive Sensor Array Processing Workshop March 1996*, MIT Lincoln Labs, Lexington, MA.

[9] G. A. Tsihrintzis, C. L. Nikias, "STAP Detection with Sub-Gaussian Distributions and Fractional Lower-Order Statistics for Airborne Radar," *Proceedings of the Adaptive Sensor Array Processing Workshop March 1996*, MIT Lincoln Labs, Lexington, MA.

[10] C. D. Richmond, "Adaptive Array Processing in Non-Gaussian Environments," *Proceedings of the Adaptive Sensor Array Processing Workshop March 1996*, MIT Lincoln Labs, Lexington, MA.

[11] C. D. Richmond, "A Note on Non-Gaussian Adaptive Array Detection and Signal Parameter Estimation," to appear *IEEE Signal Processing Letters* August 1996. Some

discussion found in the *Proceedings of Post STAP Detection Technical Interchange Meeting November 1995*, MIT Lincoln Labs, Lexington, MA.

[12] C. D. Richmond, "PDFs, Confidence Regions, and Relevant Statistics for a Class of Sample Covariance Based Array Processors," to appear *IEEE Trans. Signal Processing* July 1996.

[13] C. D. Richmond, "Exact PDFs for Sample Covariance Based Array Processors with Elliptically Contoured Data," *Proceedings of the ICASSP 1996.*

[14] C. D. Richmond, "Derived PDF of Maximum-Likelihood Signal Estimator which Employs an Estimated Noise Covariance," *IEEE Trans. Signal Processing* Vol. 44, No. 2, pp. 305–315, February 1996.

[15] C. D. Richmond, "Derived PDF of Maximum-Likelihood Signal Estimator which Employs an Estimated Noise Covariance," *Proceedings of the Adaptive Sensor and Array Processing Workshop March 1995*, Vol. 2, pp. 395–410, MIT Lincoln Laboratory, Lexington, MA.

[16] I. S. Reed, J. D. Mallett, L. E. Brennan, "Rapid Convergence Rate in Adaptive Arrays," *IEEE Trans. Aerospace and Electronic Systems*, Vol. AES-10, No. 6, 853–863 (1974).

[17] C. G. Khatri, C. Radhakrishna Rao, "Test for a Specified Signal when the Noise Covariance Matrix is Unknown," *Technical Report 85-47*, Center for Multivariate Analysis, University of Pittsburgh, PA (Nov. 1985).

[18] C. G. Khatri, C. Radhakrishna Rao, "Effects of Estimated Noise Covariance Matrix in Optimal Signal Detection," *IEEE Trans. Acoust. Speech Sig. Proc.*, Vol. ASSP-35, No. 5, 671–679 (1987).

[19] E. J. Kelly, "An Adaptive Detection Algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, Vol. AES-22, 115 - 127 (1986).

[20] E. J. Kelly, "Adaptive Detection in Non–Stationary Interference, Part I and Part II," *Technical Report 724*, Lincoln Laboratory, MIT (1985).

[21] E. J. Kelly, "Adaptive Detection in Non–Stationary Interference, Part III," *Technical Report 761*, Lincoln Laboratory, MIT (1987).

[22] E. J. Kelly, K. Forsythe, "Adaptive Detection and Parameter Estimation for Multidimensional Signal Models," *Technical Report 848*, Lincoln Laboratory, MIT (1989).

[23] F. C. Robey, D. R. Furhmann, E. J. Kelly, R. A. Nitzberg, "A CFAR Adaptive Matched Filter Detector," *IEEE Trans. Aerosp. Electron. Syst.* Vol. 28, No. 1, 208–216 (1992).

[24] S. T. Smith, "Statistical Expressions for Adaptive Detectors," Technical Memorandum No. 102L-0002, Lincoln Laboratory, MIT (1995).

[25] S. Bose, *Invariant Hypothesis Testing with Sensor Arrays*, Ph.D. Thesis, Cornell University, January 1995.

[26] S. Bose, A. O. Steinhardt, "A Maximal Invariant Framework for Adaptive Detection with Structured and Unstructured Covariance Matrices," *IEEE Trans. Signal Proc.*, Vol. 43, No. 9, 2164–2175 (1995).

[27] A. O. Steinhardt, Private Communication, MIT Lincoln Laboratory (1995).

[28] A. O. Steinhardt, "The PDF of Adaptive Beamforming Weights," *IEEE Trans. Signal Processing*, Vol. 39, No. 5, 1232–1235 (1991).

[29] B. D. Van Veen, "Adaptive Convergence of Linearly Constrained Beamformers Based on the Sample Covariance Matrix," *IEEE Trans. on Signal Processing*, Vol. 39, No. 6, 1470–1473 (1991).

[30] J. L. Krolik, D. N. Swingler, "On the Mean–Square Error Performance of Adaptive Minimum Variance Beamformers Based on the Sample Covariance Matrix," *IEEE Trans. Signal Processing*, Vol. 42, No. 2, 445–448 (1994).

[31] M. Wax, Y. Anu, "Performance Analysis of the Minimum Variance Beamformer," *IEEE Trans Signal Processing*, Vol. 44, No. 4, 928–937 (1996).

[32] J. Capon, "High–Resolution Frequency–Wavenumber Spectrum Analysis," *Proc. of IEEE*, Vol. 57, 1408–1418 (1969).

[33] J. Capon, N.R. Goodman,"Probability Distributions for Estimators of Frequency Wavenumber Spectrum," *Proc. IEEE*, Vol. 58, No. 10, 1785–1786 (1970).

[34] A. B. Baggeroer, "Confidence Intervals for Regression (MEM) Spectral Estimates," *IEEE Trans. on Information Theory*, Vol. IT–22, 534–545 (1976).

[35] J. P. Burg, D. G. Luenberger, D. L. Wegner, "Estimation of Structured Covariance Matrices," *Proc. IEEE*, Vol. 70, No. 9, 963–974 (Sept 1982).

[36] T. Barton, *Structured Covariance Estimation*, Sc.D. Thesis, Washington University, December 1993.

[37] C. W. Jim, "A Comparison of two LMS Constrained Optimal Array Structures," *Proceedings of the IEEE*, Vol. 65, 1730–1731 (1977).

[38] A. B. Baggeroer, H. Van Trees, *Optimal Array Signal Processing*, Lecture Notes for 6.456J Signal Processing II, Fall 1995.

[39] A. O. Steinhardt, "Adaptive Multisensor Detection and Estimation," Chapter 3 of *Adaptive Radar Detection and Estimation*, Edited by S. Haykin and A. O. Steinhardt, John Wiley & Sons, Inc. (1992).

[40] B. Van Veen, "Minimum Variance Beamforming," Chapter 4 of *Adaptive Radar Detection and Estimation*, Edited by S. Haykin and A. O. Steinhardt, John Wiley & Sons, Inc. (1992).

[41] A. O. Steinhardt, B. D. Van Veen, "Adaptive Beamforming," *International Jour. Adaptive Control and Sig. Proc.*, Vol. 3, 253–281 (1989).

[42] B. D. Van Veen, K. M. Buckley, "Beamforming: A Versatile Approach to Spatial Filtering," *IEEE ASSP Magazine*, Vol. 5, 4–24 (1988).

[43] R. A. Monzingo, T. W. Miller, *Introduction to Adaptive Arrays*,John Wiley & Sons, Inc. (1980).

[44] J. Ward, "Space-Time Adaptive Processing for Airborne Radar," *Technical Report 1015*, Lincoln Labs, MIT (1994).

[45] S. M. Kay, *Fundamentals of Statistical Signal Processing*, PTR Prentice-Hall, Englewood Cliffs, NJ (1993).

[46] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, Addison–Wesley, Reading, MA (1991).

[47] H. Van Trees, *Detection, Estimation, and Modulation Theory Part I*, Wiley, New York (1968).

[48] S. Haykin, J. Litva, T. J. Shepherd, *Radar Array Processing*, Springer-Verlag, New York (1993).

[49] G. Strang, *Linear Algebra and Its Applications, 3rd ed.*, Harcourt Brace Jovanovich, San Diego, CA (1988).

[50] B. Noble and J. W. Daniel, *Applied Linear Algebra, 2nd ed.*, Prentice Hall, Englewood Cliffs, NJ (1977).

[51] M. S. Srivastava, C. G. Khatri, *An Introduction to Multivariate Statistics*, Elsevier North Holland, Inc., New York (1979).

[52] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis 2nd Edition*, John Wiley & Sons, Inc. (1984).

[53] E. L. Lehmann, *Testing Statistical Hypotheses 2nd Edition*, John Wiley & Sons, Inc. (1986).

[54] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, John Wiley & Sons, Inc. (1982).

[55] Fang Kai-Tai, Zhang Yao-Ting, *Generalized Multivariate Analysis*, Springer–Verlag, New York (1990).

[56] Fang Kai-Tai and T. W. Anderson (Editors), *Statistical Inference in Elliptically Contoured and Related Distributions*, Allerton Press Inc., New York (1990).

[57] Fang Kai-Tai, S. Kotz, Ng Kai-Wang, *Symmetric Multivariate and Related Distributions*, Chapman and Hall, New York (1980).

[58] A. K. Gupta, T. Varga, *Elliptically Contoured Models in Statistics*, Kluwer Academic Publishers, Boston, MA (1993).

[59] S. Cambanis, S. Huang, G. Simons, "On the Theory of Elliptically Contoured Distributions," *Journal of Multivariate Analysis*, Vol. 11, 368-385 (1981).

[60] D. Kelker, "Distribution Theory of Spherical Distributions and a Location- Scale Parameter Generalization," *Sankhyā: The Indian Journal of Statistics: Series A*, Vol. 32, 419–430 (1970).

[61] P. R. Krishnaiah, J. G. Lin, "Complex Elliptically Symmetric Distributions," *Communcations in Statistics A: Theory & Methods*, Vol. 15, No. 12, 3693–3718 (1986).

[62] K. S. Miller, *Complex Stochastic Processes*, Addison–Wesley, Reading, MA (1974).

[63] K. S. Miller, *Multidimensional Gaussian Distributions*, John Wiley and Sons, Inc. New York, pp 88–92 (1964).

[64] N. R. Goodman,"Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution," *Ann. Math. Stat.*, Vol. 34, 152–177 (1963).

[65] C. G. Khatri, "Classical Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution," *Ann. Math. Stats.*, Vol. 36, pp. 98–114 (1965).

[66] R. A. Wooding, "The Multivariate Distribution of Complex Normal Variables," *Biometrika*, Vol. 43, pp. 212–215 (1956).

[67] I. J. Schoenberg, "Metric Spaces and Completely Monotone Functions," *Ann. Mathematics*, Vol. 39, pp. 811–841 (1938).

[68] B. Picinbono, "On Circularity," *IEEE Trans. Sig. Proc.* Vol. 42, No. 12, pp. 3473–3482 (1994).

[69] K. Yao, "A Representation Theorem and Its Applications to Spherically Invariant Random Processes," *IEEE Trans. Inform. Theory*, **IT-19**, 600–608, (1973).

[70] S. Finette, P. Mignerey, J. Smith, C. D. Richmond, "Broadband Source Signature Extraction Using Vertical Array," *J. Acoust. Soc. of Amer.*, **94**(1), 309 - 318 (1993).

[71] C. G. Khatri, "Some Results for the Singular Normal Multivariate Regression Models," *Sankhyā: The Indian Journal of Statistics: Series A* Vol. 30, 267–280 (1968).

[72] I. S. Gradshteyn, I. M. Ryzhik, *Table of Integrals, Series,and Products*, Academic Press, Inc., San Diego, CA (1980).

[73] M. Abramowitz, I. A. Stegun, *Handbook of Mathematical Functions*, Dover Publications, Inc., New York, NY (1970).

[74] L. J. Slater, *Confluent Hypergeometric Functions*, Cambridge University Press, (1960).