Statistical Analysis of Adaptive Maximum-Likelihood Signal Estimator

by Christ D. Richmond

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

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

Electrical Engineer

at the Massachusetts Institute of Technology

February 1995

Signature of Author ....
Department of Electrical Engineering and Computer Science
January 24, 1995

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

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

Statistical Analysis of Adaptive Maximum-Likelihood Signal Estimator

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 Electrical Engineer (E.E.)

Abstract

A classical problem in many radar and sonar applications is the adaptive detection/estimation of a given signal in the presence of zero mean Gaussian noise. Reed, Mallett, and Brennan (RMB) derived and analyzed an adaptive detection scheme where the noise adaptation and non-trivial nature of their analysis resulted from the use of a noise sample covariance matrix (SCM). The case now considered is that of adaptive signal estimation. Specifically, the exact probability density function (pdf) for the ML signal estimator, also referred to as the *Minimum Variance Distortionless Response (MVDR)* and as the *Linearly Constrained Minimum Variance (LCMV) Beamformer*, is derived when the estimator relies on a SCM for evaluation. The observation from which the signal ML estimate is made is assumed *linear* in the signal and corrupted by additive complex Gaussian noise. The SCM assumes a Complex Wishart (CW) distribution when each of the noise samples is *i.i.d.* Thus, by using the CW probabilistic model for the distribution of the estimated noise covariance it is shown that the pdf of the *Adaptive* ML (AML) signal estimator, *i.e.* the ML signal estimator which employs a SCM for evaluation, is in general the confluent hypergeometric function known as *Kummer's Function.* The AML signal estimator remains unbiased, but asymptotically efficient; moreover, the AML signal estimator converges in distribution to the Gaussian non-adaptive beamformer output (known noise covariance). When the sample size of the estimated noise covariance matrix is fixed, there exist a dynamic tradeoff between Signal-to-Noise Ratio (SNR) and *noise adaptivity* as the dimensionality of array data is varied suggesting the existence of an optimal array data dimension which will yield the best performance.

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

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. and E.E. degrees in Electrical Engineering from the Massachusetts Institute of Technology (MIT) in February 1993 and February 1995 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 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 is presently a Teaching Assistant and a candidate for the PhD at MIT. His research interest include statistical signal and array processing, detection/estimation, radar, controls, multivariate analysis and adaptive systems. His journal publications include

- "Derived PDF of Maximum Likelihood Signal Estimator which Employs an Estimated Noise Covariance," to appear in *IEEE Transactions on Signal Processing* 1995.
- "Broadband Source Signature Extraction Using a Vertical Array," *Journal Acoustical Society of America*, **94**(1), 309–318, July 1993, with S. Finette, P. Mignerey, and J. Smith.
- "Exact Joint PDF of Weight Matrix of Sample Covariance Based Minimum Variance Distortionless Processor (MVDP)," to be submitted to *IEEE Transactions on Signal Processing* 1995.

Acknowledgements

At all times my *premier* thanks is to my Lord and God Jesus the Christ. " For by *him* were all things created, that are in heaven, and that are in earth, visible and invisible, whether they be thrones, or dominions, or principalities, or powers: all things were created by *him*, and for *him*: And *he* is before all things, and by *him* all things consist" (Colossians 1:16–17). I always try to bear in mind that these "all things" include the forces/powers behind the principles of Maximum–Likelihood estimation/detection, optimal array processing, Maxwell's equations, etc. *He* is the source and light of all true wisdom including that reported in this thesis. Indeed, hid in Christ are all the treasures of wisdom and knowledge (Colossians 2:3).

I'd liked to thank the Office of Naval Research (ONR) for gracious funding of my graduate studies. I'd also like to thank my advisor Prof Arthur B. Baggeroer for making me aware of references which lead to this work, and his support throughout this endeavor.

Special thanks to Dr. Kevin Sangston at the Naval Research Laboratory (NRL) for several encouraging discussions. I'd like to thank my father Dr. Clinton T. Richmond for his constant encouragement and his profitable experiential advice. Also, thanks to my mother Gloria H. Snoody for her perpetual and unbounded love.

Thanks to my mentor/colleague and good friend Dr. Steven H. Isabelle for his support, advice, and just lending a good ear.

Thanks to the many who keep me in their prayers for you are also in mine. *God's grace and peace through Christ Jesus be with you all.*

Contents

List of Figures

Symbols and Notation

Symbol	Meaning
$\mathbf{R}$	$B \times B$ true noise covariance matrix
$\hat{\mathbf{R}}$	$B \times B$ noise sample covariance matrix
Q	Sample size of noise sample covariance matrix
$\mathbf{x}$	$B \times 1$ received array data vector
$\mathbf{g}$	$B \times 1$ steering vector or look direction vector
$\mathbf{n}$	$B \times 1$ additive complex Gaussian noise
$\mathbf{w}_{OPT}$	True optimum weight vector $\mathbf{R}^{-1}\mathbf{g}/(\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g})$
$\hat{\mathbf{w}}_{OPT}$	Estimate of optimum weight vector $\hat{\mathbf{R}}^{-1}\mathbf{g}/(\mathbf{g}^H\hat{\mathbf{R}}^{-1}\mathbf{g})$
B	Dimension of array data vector
S	True complex signal value $S_R + jS_I$ (non-random)
S_{ML}	Maximum-Likelihood signal estimator $\mathbf{w}_{OPT}^H\mathbf{x}$
S_{AML}	Adaptive Maximum-Likelihood signal estimator $\hat{\mathbf{w}}_{OPT}^H\mathbf{x}$
ϑ	Normalized SNR statistic from [4]
$Pr(A)$	Probability of event A
$p_x(\cdot)$	Probability density function for x
$CN_D(\cdot,\cdot)$	D–variate Complex Gaussian distributed
$CW_D(\cdot,\cdot)$	$D \times D$–variate complex Wishart distributed
$\|\|\cdot\|\|$	Vector norm
$\|\cdot\|$	Matrix determinant or absolute value of complex scalar
$\text{tr}(\cdot)$	Matrix trace
T	Matrix transpose
H	Conjugate transposition
$E\{\cdot\}$	Expectation
$\text{Var}(\cdot)$	Variance
σ_x^2	Variance of x
$\Phi(\cdot)$	Moment generating function
${}_pF_q(\cdot)$	Generalized hypergeometric function
${}_1F_1(\cdot)$	Kummer's Function
$I_0(\cdot)$	Zeroth order Bessel function
$Re(\cdot)$	Real part
$Im(\cdot)$	Imaginary part
SCM	"Sample Covariance Matrix"
AML	"Adaptive Maximum-Likelihood" signal estimator
MVDR	"Minimum Variance Distortionless Response"
LCMV	"Linearly Constrained Minimum Variance" beamformer
SNR	"Signal-to-Noise Ratio" $\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}$

Table 0.1: Symbols/Notation

Chapter 1

Introduction

1.1 Motivation

A classical problem in many radar and sonar applications is the adaptive detection/estimation of a given signal in the presence of zero mean Gaussian noise. A search radar, for example, scans a broad surveillance area seeking to detect the possible presence of a target. After target presence is determined, it may be desired to estimate the signal parameters. In pulsed Doppler and narrowband active radar systems, for example, it is desired to estimate the complex signal amplitude parameters [1] of the echo returns of each pulse burst [2]. These signal parameters carry target range, and Doppler (velocity) information.

The success or performance of any signal detector or estimator is highly dependent

[1] Amplitude of signal is complex due to demodulated *in-phase* and *quadrature* components.

upon one's knowledge of the probabilistic character of the noise interference; namely, the *noise covariance matrix*, denote by $\mathbf{R}$. Indeed, estimating $\mathbf{R}$ lies at the core of every adaptive detection/estimation procedure. When $\mathbf{R}$ is known *exactly* then much can be said about how well the signal detectors/estimators which rely on $\mathbf{R}$ perform. In practice the true noise covariance is typically unknown, and must be approximated from *noise only* observations (data samples). The introduction of an estimate of $\mathbf{R}$, say a sample covariance matrix (SCM) denoted by $\hat{\mathbf{R}}$, complicates the performance analysis of the signal detectors/estimators which employ $\hat{\mathbf{R}}$ instead of $\mathbf{R}$. The performance analysis becomes complex because $\hat{\mathbf{R}}$ is a random matrix whose probability distribution must be incorporated in the analysis. Although tedious and difficult to obtain, such performance analyses are indispensable for the design of truly optimal array processors.

1.2 Objectives

Reed, Mallett, and Brennan (RMB) derived and analyzed an adaptive detection scheme where the noise adaptation and non-trivial nature of their analysis resulted from the use of a noise SCM [4]. The case considered in this thesis is that of adaptive signal estimation; namely, it is desired to statistically characterize the behavior/performance of the Adaptive Maximum-Likelihood (AML) signal estimator, alias the sample covariance based *Minimum Variance Distortionless Response (MVDR)* [2, 3] or the *Linearly Constrained Minimum Variance (LCMV) Beamformer* (see Ch. 4 of [2]). When the noise covariance is known then the non-adaptive (fixed/unchanging covariance) ML signal estimator is Gaussian distributed. When the noise covariance is unknown the distribution of the adaptive signal estimator relying on $\hat{\mathbf{R}}$ is not Gaussian. Primary attention is devoted to finding a statistical description that accounts for the random variations experienced in both (1) the additive noise corrupting the data observation from which the signal is estimated, and (2) the estimated noise covariance matrix. An SCM will adaptively provide an estimate of the true noise covariance matrix. For this reason when the ML signal estimator employs a sample covariance matrix in place of the true noise covariance it is referred to as the *Adaptive* ML signal estimator. The noise will be assumed complex normally distributed [2, 10, 11, 15] and consequently allow for an analytically tractable development. Although the most natural application of the results obtained in this thesis is to radar and to active sonar where coherent processing can be done due to the known form of the signals, the mathematical model, however, is applicable to many diverse systems involving the processing of data obtained from an array of sensors. The reader is referred to [2, 10, 11] for such applications.

1.3 Organization of Thesis

In Chapter 2 the assumed system model for the observed array data is given, and a subsequent derivation of the ML signal estimator (array data processor) to be investigated follows. The probabilistic model for the estimated noise covariance is then discussed in preparation for the AML signal estimator pdf derivation given in Chapter 3. In Chapter 4 a detailed statistical analysis of the AML estimator is provided, shedding light on the statistical asymptotic behavior of the AML estimator, its moments, and its confidence regions; moreover, a somewhat recondite phenomena is observed in the estimator's performance as we consider the dimensionality of the array data as a parameter over which to optimize performance. Specifically, when the sample size of the estimated noise covariance matrix is fixed, there exist a dynamic tradeoff between Signal-to-Noise Ratio (SNR) and *noise adaptivity* as the dimensionality of array data is varied suggesting the existence of an optimal array data dimension which will yield the best performance. This optimal array dimension is computed explicitly for the very common scenario of a uniform linear array in a spatially white noise environment, and validated via numerical simulation. In Chapter 5 a summary of the results is given along with suggestions for further research in this area of array processing.

Because the resulting pdf for the AML estimator is a *special function* and unfamiliar to most readers, extensive appendices are provided which give detailed information on Kummer's Function, and derivations of the main results.

Chapter 2

Theoretical Background

2.1 System Model and Signal Estimator

2.1.1 Model of Array Data Output

The array output is modeled as the following complex vector observation

$$\mathbf{x}_{(B\times 1)} = \mathbf{g}_{(B\times 1)} S_{(1\times 1)} + \mathbf{n}_{(B\times 1)} \tag{2.1}$$

where the dimensions of the corresponding matrices in the general linear system model are indicated in subscript [1]. $\mathbf{x}$ is the measured or received data [2] containing the desired signal scalar S corrupted by the additive noise $\mathbf{n}$. The vector $\mathbf{g}$ models the system transfer

[1]The notational convention will be boldfaced capitals indicating matrices and boldface lowercase indicating vectors.

[2]Data is complex in general due to in-phase/quadrature components of demodulated data.

function, also known as the *steering vector* or *look direction*, and is assumed to be known exactly. In pulsed Doppler radar, for example, the look direction one can assume to be known, since a known fixed direction is being illuminated during each pulse burst [2]. In underwater acoustic applications good estimates of the transfer functions (Green's Functions) can be made which allow them to be assumed known for all practical purposes [17, 18]. [3]

2.1.2 Maximum-Likelihood Signal Estimator

If one assumes the complex Gaussian distribution $CN_B(\mathbf{0}, \mathbf{R})$ for the noise $\mathbf{n}$, *i.e.* with zero mean and $B \times B$ covariance matrix $\mathbf{R}$, then the ML estimate for the signal S can be shown to be given by

$$S_{ML} = \frac{\mathbf{g}^H \mathbf{R}^{-1} \mathbf{x}}{\mathbf{g}^H \mathbf{R}^{-1} \mathbf{g}} \tag{2.2}$$

[7, 10] where superscript H denotes conjugate transposition.

This estimator can be derived in numerous ways, and hence appears under many aliases. A method of derivation with intuitive geometric appeal generates the signal estimate by choosing the "optimal" linear combination of the array outputs; namely, optimally choosing the weight vector or filter $\mathbf{w}$ such that $\mathbf{w}^H \mathbf{x} = \hat{S}$. The optimality

[3]When the steering vector $\mathbf{g}$ is not known exactly the resulting ML signal estimate introduces a multiplicative bias. The AML estimator is likewise biased by an unknown $\mathbf{g}$, but the multiplicative biasing is stochastic in this case. The statistical analysis of such a situation remains an open problem. Empirical studies indicate, however, that accurate knowledge of $\mathbf{g}$ is crucial to the success of the signal estimators [17].

criterion is conveyed by the following minimization problem:

$$\mathbf{w}_{OPT} = \begin{cases} \min_{\mathbf{w}} \mathrm{Var}(\hat{S}) = \min_{\mathbf{w}} \mathbf{w}^H \mathbf{R} \mathbf{w} \\ \text{Subject to the Constraint } \mathbf{w}^H \mathbf{g} = 1. \end{cases} \tag{2.3}$$

Note that the solution will yield a minimum variance estimate. The given constraint forces the estimate to be unbiased since $\mathbf{w}^H\mathbf{x} = \mathbf{w}^H\mathbf{g} \cdot S + \mathbf{w}^H\mathbf{n} = S+$ *zero mean noise.* If a solution exists then it *must* correspond to the ML solution [16]. To solve we first whiten [4] the data yielding the new variables

$$\begin{aligned} \mathbf{x}_0 = \mathbf{R}^{-1/2}\mathbf{x} &= \mathbf{R}^{-1/2}\mathbf{g}S + \mathbf{R}^{-1/2}\mathbf{n} \\ &= \mathbf{g}_0 S + \mathbf{n}_0. \end{aligned} \tag{2.4}$$

This data whitening can be thought of as a change of variables, or a defining weighting metric to the complex euclidean B-space $\mathcal{C}^B$. The optimization problem then becomes

$$\mathbf{w}_{0_{OPT}} = \begin{cases} \min_{\mathbf{w}_0} \mathrm{Var}(\hat{S}) = \min_{\mathbf{w}_0} \mathbf{w}_0^H \mathbf{w}_0 = \min_{\mathbf{w}_0} ||\mathbf{w}_0||^2 \\ \text{Subject to the Constraint } \mathbf{w}_0^H \mathbf{g}_0 = 1 \end{cases} \tag{2.5}$$

which says to find the vector $\mathbf{w}_0$ with minimum norm that satisfies the constraint $\mathbf{w}_0^H\mathbf{g}_0 =$

[4]Although $\mathbf{R}^{-1/2}$ is not unique the choice of $\mathbf{R}^{-1/2}$ made to perform the data whitening does not influence the resulting solution in terms of the original variables since the matrix square roots will recombine to produce $\mathbf{R}^{-1}$.

$||\mathbf{w}_0||||\mathbf{g}_0||\cos\theta(\mathbf{w}_0, \mathbf{g}_0) = 1$, or

$$||\mathbf{w}_0|| = \frac{1}{||\mathbf{g}_0||\cos\theta(\mathbf{w}_0, \mathbf{g}_0)}. \tag{2.6}$$

To minimize the norm $||\mathbf{w}_0||$ subject to the given constraint we must choose $\mathbf{w}_0$ such that $|\cos\theta(\mathbf{w}_0, \mathbf{g}_0)|$, which is always ≤ 1, is as large as possible; namely, we must choose $\mathbf{w}_0$ to be in exactly the same direction as $\mathbf{g}_0$ so that $\cos\theta(\mathbf{w}_0, \mathbf{g}_0) = 1$. Hence, $\mathbf{w}_{0_{OPT}} = \beta \cdot \mathbf{g}_0$. The constant of proportionality β can be found from the constraint to be $\beta = 1/||\mathbf{g}_0||^2$. The optimal choice for the weight vector is therefore $\mathbf{w}_{0_{OPT}} = \mathbf{g}_0/||\mathbf{g}_0||^2$ or in terms of the original variables $\mathbf{w}_{OPT} = \mathbf{R}^{-1}\mathbf{g}/(\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g})$ yielding the ML signal estimate given by eq(2.2). Note that the optimal weight vector can be thought of as a *matched filter.* Indeed, a filter matched to the signal direction $\mathbf{g}$.

2.1.3 Remarks

It is of interest to characterize the statistical behavior of the ML signal estimator when one has to rely on an estimate $\hat{\mathbf{R}}$ of the true noise covariance $\mathbf{R}$ in order to compute S_{ML}. The chosen noise covariance estimator and its associated probabilistic model is discussed in the next section.

2.2 Noise Covariance Estimator and Distribution

2.2.1 Sample Covariance Matrix

From eq(2.2) note that exact knowledge of the noise covariance $\mathbf{R}$ is required to compute the signal estimate. In practice, however, $\mathbf{R}$ must be approximated. When there is foreknowledge about the *structure* of the noise covariance matrix (e.g. a Toeplitz configuration), a *true* maximum likelihood estimate of this matrix can be made which incorporates the constraints implied by this known structure [13]. If this knowledge is not available, *i.e.* no constraints are made on the family of feasible noise covariances, then the following *sample covariance matrix (SCM)* is the unbiased ML estimate of the zero mean complex Gaussian noise covariance $\mathbf{R}$

$$\hat{\mathbf{R}} = \frac{1}{Q}\sum_{i=1}^{Q}\mathbf{n}_i\mathbf{n}_i^H. \tag{2.7}$$

Therefore, this sample covariance matrix will serve as an *adaptive* estimate of $\mathbf{R}$ in the absence of its exact knowledge. If statistical variations in the noise field arise then the sample covariance matrix adapts to these changes via the inclusion of the appropriate noise samples. Q is the total number of independent identically distributed *(i.i.d)* sample noise vectors $\mathbf{n}_i$ included in the average [5] and is often termed the *degrees of freedom* because of $\hat{\mathbf{R}}$'s inherent probabilistic relationship to the χ^2-distribution [12]. Note that if the complex vector quantity $\mathbf{n}_i$ were a one dimensional complex scalar (*i.e.* $B = 1$), then $\hat{\mathbf{R}}$ is the sum of the squares of *i.i.d* real normal random variables. This sum is known to

[5] Note that the signal data vector $\mathbf{x}$ is not used to estimate $\hat{\mathbf{R}}$.

produce a random variable proportional to a χ^2; hence the perception of their kindredness. Indeed, $\hat{\mathbf{R}}$ is the multivariate extension of the χ^2 random variable. Note that Q must at least equal or exceed B in order for $\hat{\mathbf{R}}$ to be nonsingular (with probability 1, see [2]), and hence of use here. [6] [7]

2.2.2 Complex Wishart Model

The joint distribution of the distinct elements of the hermitian positive definite matrix $\mathbf{A} = Q \times \hat{\mathbf{R}}$ is known to have the following Complex Wishart (CW) distribution

$$p_{CW}(\mathbf{A}) = \left[|\mathbf{A}|^{Q-B}|\mathbf{R}|^{-Q}/\tilde{\Gamma}_B(Q)\right] \exp\left[-\mathrm{tr}(\mathbf{R}^{-1}\mathbf{A})\right] \tag{2.8}$$

where the differential volume element is given by $(d\mathbf{A}) = dA_{11}dA_{22}\cdots dA_{BB} \cdot dRe(A_{12}) \cdot dIm(A_{12}) \cdot dRe(A_{13})dIm(A_{13})\cdots dRe(A_{B-1,B})dIm(A_{B-1,B})$, and where $Re(\cdot)$ and $Im(\cdot)$ denote the real and imaginary parts respectively. This distribution is sometimes denoted as $CW_B(Q, \mathbf{R})$. The complex multivariate gamma function is given by the following

[6] Augmenting $\hat{\mathbf{R}}$ by ***diagonal loading*** for situations in which $Q < B$ or $Q \approx B$ is sometimes done in practice due to the scarce availability of noise samples. A rigorous mathematical probabilistic/stochastic treatment of the performance analysis of such methods still remains an open problem. Empirical investigations have shown, however, that such techniques lead to improved sidelobe levels (noise rejection), but reduces the beamformer's nulling capability against weak interferences [8, 9]. Which makes intuitive sense, since such loading tells the beamformer that there are no weak signals (small eigenvalues) present in the data. In addition such techniques require an estimate of the small eigenvalues of $\hat{\mathbf{R}}$ to set the load level correctly. Excessive loading could result in unacceptable performance.

[7] In this presentation we simply replaced the true noise covariance with a sample noise covariance in order to produce an estimate of the signal parameters from quantities we can measure. If, however, one assumed (1) the totality of the unknown parameters includes both S and $\mathbf{R}$, (2) totality of the observation from which we can estimate the unknowns is given by the data matrix $\mathbf{X} = [\mathbf{x}|\mathbf{n}_1|\mathbf{n}_2|\cdots|\mathbf{n}_Q]$ and (3) one proceeds to estimate S and $\mathbf{R}$ from $\mathbf{X}$ via the ML procedure, then the resulting estimate is in fact given by simply replacing the true noise covariance with a sample noise covariance in order to produce an estimate of the signal parameters, see [2], and [6, 7].

product of univariate gamma functions

$$\tilde{\Gamma}_B(Q) = \pi^{B(B-1)/2} \prod_{i=1}^{B} \Gamma(Q - i + 1). \tag{2.9}$$

The symbol $|\cdot|$ denotes determinant (and sometimes the absolute value of a complex scalar), and $\text{tr}(\cdot)$ denotes the matrix trace. Note that the CW pdf does not exist when $Q < B$ because the rank of $\mathbf{A}$ is $< B$, hence $|\mathbf{A}| = 0$. Its cumulative distribution, nevertheless, will exist [12]. We avoid such issues by requiring $Q \geq B$.

2.2.3 Summary

When no a priori information about the structure of the true noise covariance is assumed then the SCM is the unbiased ML estimator of $\mathbf{R}$, and hence the noise covariance estimator of choice. The CW distribution model will be assumed for the SCM $\hat{\mathbf{R}}$ and thus allows an associated pdf to be obtained for the AML signal estimator which relies on $\hat{\mathbf{R}}$ for evaluation. The derivation of the AML pdf is given in the next chapter.

Chapter 3

Derivation of PDF

3.1 A Posteriori Distribution of AML Estimator

The distribution of the signal ML estimator S_{ML} is known to be Gaussian when the noise covariance $\mathbf{R}$ is known exactly, *i.e.* because $\mathbf{R}$ is a deterministic quantity, and therefore computation can be made via eq(2.2); namely, it is distributed according to

$$S_{ML} \sim CN_1\left(S, \frac{1}{\mathbf{g}^H \mathbf{R}^{-1} \mathbf{g}}\right) = p_{S_{ML}}. \tag{3.1}$$

Gaussian random variables are known to regenerate under linear transformations by definition. Note from the above pdf that S_{ML} is an *unbiased* and *efficient* estimator of S (see [10] pp. 524–530). It will be shown that these desirable estimator properties are in fact preserved (the latter asymptotically) when the SCM $\hat{\mathbf{R}}$ replaces $\mathbf{R}$ in the signal estimation procedure.

The distribution of the AML signal estimator given by

$$S_{AML} = \frac{\mathbf{g}^H \hat{\mathbf{R}}^{-1} \mathbf{x}}{\mathbf{g}^H \hat{\mathbf{R}}^{-1} \mathbf{g}} \tag{3.2}$$

is not on the other hand such a trivial matter since the estimated noise covariance $\hat{\mathbf{R}}$ is in general a matrix variate with an associated distribution. A distribution can be found, nevertheless, for the signal AML estimator rather directly by making two important observations:

1. S_{AML} is *conditionally* complex Gaussian distributed *given* $\hat{\mathbf{R}}$.
2. The variance of S_{AML} as a function of $\hat{\mathbf{R}}$ has a distribution proportional to that of a beta distributed random variable when $Q \times \hat{\mathbf{R}}$ assumes the distribution $CW_B(Q, \mathbf{R})$.

These two observations constitute sufficient information to determine uniquely the pdf of S_{AML} because Gaussian distributions are completely characterized by the first and second moments. Consider the derivation in the following section.

3.2 Derivation of AML Marginal Distribution

Simplifying the expression given in eq(3.2) for the AML estimator via eq(2.1) yields the following form for the AML estimator

$$S_{AML} = S + \frac{\mathbf{g}^H \hat{\mathbf{R}}^{-1} \mathbf{n}}{\mathbf{g}^H \hat{\mathbf{R}}^{-1} \mathbf{g}} = S + \mathcal{N}. \tag{3.3}$$

The estimator is the sum of the non-random true value of the source S and a noise term. The noise term $\mathcal{N}$ is a complex random variable whose distribution is jointly dependent upon the normality of $\mathbf{n}$ and the associated distribution of the SCM.

Recall that the weight vector which produces the AML signal estimate $\hat{\mathbf{w}}_{OPT} = \hat{\mathbf{R}}^{-1}\mathbf{g}/(\mathbf{g}^H\hat{\mathbf{R}}^{-1}\mathbf{g})$ is in fact a matched filter. Matched filters pervade the front ends of most detection algorithms. In 1974 Reed, Mallett, and Brennan [4] derived an adaptive detection scheme with a matched filter as the front end processor. The adaptive matched filter of [4] is in fact proportional to $\hat{\mathbf{w}}_{OPT}$; however, the proportionality constant is a random variable which arises from the constraint imposed by eq(2.3). A mathematical linkage between the work of [4] and the AML estimator studied here can be made which aids in the AML estimator pdf derivation. The authors of [4] were the first to proved that the following random variable

$$\vartheta = \frac{1}{\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}}\left[\frac{(\mathbf{g}^H\hat{\mathbf{R}}^{-1}\mathbf{g})^2}{\mathbf{g}^H\hat{\mathbf{R}}^{-1}\mathbf{R}\hat{\mathbf{R}}^{-1}\mathbf{g}}\right] \tag{3.4}$$

is beta distributed according to the pdf

$$p_{\vartheta}(\theta) = \begin{cases} \dfrac{Q!}{(B-2)!(Q+1-B)!}(1-\theta)^{B-2}\theta^{Q+1-B} & ,0 \le \theta \le 1 \\ 0 & ,\text{otherwise.} \end{cases} \tag{3.5}$$

Since then many other authors have derived this same result in a variety of ways [2, 5, 7]. This random variable is referred to as a normalized SNR statistic and represents the SNR of the random output of the matched filter used to detect signal presence against the zero mean Gaussian noise background.

A sagacious observation which affords us the exact pdf of the AML estimator rather directly relates this random variable ϑ to an *a posteriori* distribution of the AML estimator. Note that the AML signal estimator is conditionally complex Gaussian distributed given $\hat{\mathbf{R}}$ such that its conditional variance is given by

$$\sigma^2_{S_{AML}|\hat{\mathbf{R}}} = E\{|\mathcal{N}|^2|\hat{\mathbf{R}}\} = \frac{\mathbf{g}^H\hat{\mathbf{R}}^{-1}\mathbf{R}\hat{\mathbf{R}}^{-1}\mathbf{g}}{(\mathbf{g}^H\hat{\mathbf{R}}^{-1}\mathbf{g})^2} = \left(\frac{1}{\vartheta}\right)\frac{1}{\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}}, \tag{3.6}$$

and hence

$$S_{AML} \sim CN_1\left(S, \frac{1/\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}}{\vartheta}\right) = p_{S_{AML}|\vartheta}. \tag{3.7}$$

Elementary probability theory suggests therefore that the distribution of the AML signal estimator is given by the following integral

$$p_{S_{AML}}(\hat{S}_0) = \int_0^1 p_\vartheta(\theta)p_{S_{AML}|\vartheta}(\hat{S}_0|\theta)d\theta \tag{3.8}$$

which, as one can easily verify (see Gradshteyn and Ryzhik [19], p. 318 no. 3.383), yields

$$\boxed{p_{S_{AML}}(\hat{S}_0) = \frac{\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}}{\pi}\left[\frac{Q-B+2}{Q+1}\right]{}_1F_1\left(Q-B+3; Q+2; -|\hat{S}_0 - S|^2\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}\right).} \tag{3.9}$$

${}_1F_1(a, b; z)$ is the *confluent hypergeometric function* also known as *Kummer's Function* [20, 21]. More is said about this special function, its properties, and its corresponding power series in Appendix A. The attraction of eq(3.9) is the statistical summary it provides. It provides an exact/complete probabilistic characterization of the AML signal estimator that accounts for the uncertainty present in both (1) the noisy data observation from which the signal is estimated, and (2) the estimated noise covariance. Note that this pdf is explicitly parameterized by the dimensionality of the array data B, the sample size Q of the SCM, the steering vector $\mathbf{g}$, and the true but unknown noise covariance matrix $\mathbf{R}$, allowing investigation of variations in all these parameters. In the next chapter we use the AML pdf eq(3.9) to infer aspects of the processor's performance; in particular to determine the effects of estimated noise covariances on the AML processor's performance and to discover how all the parameters of the pdf play a role. Before moving to the next chapter, more is said about the random variable ϑ from reference [4].

3.3 Normalized SNR Statistic

If $\eta \in [0, 1]$ represents the average normalized SNR one wishes to maintain in the detection process, a trivial algebraic exercise will show that the associated lower bound on Q necessary to accomplish this is

$$\boxed{Q \geq Q_{TH}(\eta) = \frac{B - 2 + \eta}{1 - \eta}.} \tag{3.10}$$

For example, reference [4] suggested the "rule-of-thumb" $Q \geq Q_{TH} = 2B - 3$, obtained by choosing $\eta = 1/2$. Clearly, the closer η is to unity the greater the lower bound on Q. The mean of the beta distributed variable ϑ is

$$E\{\vartheta\} = \frac{Q+2-B}{Q+1} \tag{3.11}$$

and its variance is given by

$$\sigma_\vartheta^2 = \frac{(Q+2-B)(B-1)}{(Q+1)^2(Q+2)}. \tag{3.12}$$

The mean of ϑ approaches one and its variance approaches zero in the limit of large Q; hence, ϑ converges "with probability one" to unity as Q becomes arbitrarily large. Noting that ϑ will be unity if and only if $\hat{\mathbf{R}} = \mathbf{R}$, reference [4] argued that the sample covariance matrix $\hat{\mathbf{R}}$ must also converge in probability to the true noise covariance matrix $\mathbf{R}$ in the limit of large Q. It is therefore useful in practice to think of η as a representative measure of the reliability or confidence level of the resulting estimated noise covariance which dictates a corresponding lower bound $Q_{TH}(\eta)$ on the degrees of freedom via eq(3.10) necessary to accomplish this level of confidence.

Note that because ϑ converges in probability to unity as Q becomes arbitrarily large, its pdf converges to $p_\vartheta(\theta) \to \delta(\theta - 1)$. Hence, by the integral equation given by eq(3.8) and the sifting property of the Dirac delta function we should expect the pdf of the AML estimator to converge to the Gaussian pdf of the ML estimator. We prove this convergence in the next chapter.

Chapter 4

Statistical Analysis of AML Signal Estimator

4.1 Properties of AML Signal Estimator

The pdf of the AML given by eq(3.9) will allows us to develop some qualitative insights and intuitions about the effects of SCM's on the performance of the AML signal estimation procedure.

Note that $p_{S_{AML}}(\hat{S}_0)$ is a function of $|\hat{S}_0 - S|$. It is therefore circularly symmetric about the true value of the signal S; hence, $E\{S_{AML}\} = S$ and therefore S_{AML} is an unbiased estimate of S. This unbiased property could have likewise been deduced from eq(3.3) by first conditioning on $\hat{\mathbf{R}}$ and noting that the conditional expectation $E\{S_{AML}|\hat{\mathbf{R}}\} = S$, which is independent of $\hat{\mathbf{R}}$. Hence, removing the conditioning by integrating with respect to $\mathbf{R}$ via the CW density eq(2.8) yields unity times S.

Consider the following useful property of the confluent hypergeometric function known as *Kummer's Transformation* (see Appendix A)

$$ {}_1F_1(a;b;-z) = e^{-z}\,{}_1F_1(b-a;b;z). \tag{4.1} $$

This transformation allows a clear illustration to be given of the asymptotic behavior of the pdf $p_{S_{AML}}(\hat{S}_0)$ as the sample size of the SCM increases, *i.e.* as $Q \to \infty$. As the sample size of $\hat{\mathbf{R}}$ becomes arbitrarily large the SCM converges in probability to $\mathbf{R}$ [4, 12]. We should therefore expect that $p_{S_{AML}}(\hat{S}_0) \to p_{S_{ML}}(\hat{S}_0)$ as $Q \to \infty$, *i.e.* S_{AML} converges in distribution to S_{ML}. Using eq(3.9) and eq(4.1) observe that

$$ \lim_{Q\to\infty} p_{S_{AML}}(\hat{S}_0) = \frac{\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}}{\pi} e^{-|\hat{S}_0 - S|^2 \mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}} \times \tag{4.2} $$
$$ \lim_{Q\to\infty}\left[\frac{Q-B+2}{Q+1}\right] \times \lim_{Q\to\infty} {}_1F_1\left(B-1, Q+2; |\hat{S}_0 - S|^2\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}\right) $$

Making liberal use of the linearity, product, and quotient limit theorems of mathematical analysis, it can be shown indeed that for arguments of finite magnitude the following asymptotic behavior holds

$$ \lim_{Q\to\infty} p_{S_{AML}}(\hat{S}_0) = \frac{\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}}{\pi} e^{-|\hat{S}_0 - S|^2 \mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}} \triangleq CN_1\left(S, \frac{1}{\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}}\right). \tag{4.3} $$

Proof: Clearly the first limit of eq(4.2) approaches unity, *i.e.* $\lim_{Q\to\infty}(Q-B+2)/(Q+1) = 1$. Let $z = |\hat{S}_0 - S|^2\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}$. Recall from Appendix A that the second limit can be written as

$$\lim_{Q\to\infty} {}_1F_1(B-1, Q+2; z) = \lim_{Q\to\infty} 1 + \frac{(B-1)z}{(Q+2)} + \frac{(B-1)(B)}{(Q+2)(Q+3)}\frac{z^2}{2} + \cdots. \tag{4.4}$$

If z is finite then it is clear that this limit approaches unity as well; hence, the validity of eq(4.3). If the argument of the pdf has an infinite magnitude then $p_{S_{AML}}(\hat{S}_0)$ is zero by virtue of eq(3.7) and eq(3.8). The integral of zero is zero. Q.E.D.

Since the distribution of S_{AML} converges to eq(3.1) it is true that the error variance, *i.e.* the variance of S_{AML}, also converges to that of S_{ML}. Hence, S_{AML} is asymptotically efficient. In summary it has been demonstrated that

1. S_{AML} has a circularly symmetric distribution about S, *i.e.* its pdf depends only on the magnitude of $|\hat{S}_0 - S|$ and is therefore an unbiased estimator.
2. S_{AML} converges in distribution to S_{ML}.
3. S_{AML} is asymptotically efficient.

4.1.1 Convergence in Distribution

The usefulness of the AML pdf eq(3.9) includes the questions it allows one to probe quantitatively, and the intuition it gives qualitatively. For example, if the array size is fixed at some value $B = B_0$, then how does the pdf, *i.e.* the statistics of the estimator, behave as the sample size Q of the employed estimated noise covariance varies? This question has already been addressed somewhat. Specifically, it has been shown that as Q became much larger than the array size B the pdf of S_{AML} converges to the Gaussian pdf of S_{ML}. This convergence can in fact be illustrated graphically. Before illustrating this convergence property first note that the pdf is in fact a circularly symmetric molehill in the complex $\hat{S}_0$ plane centered symmetrically on the mean S, the shape of which is illustrated in fig(4.1).

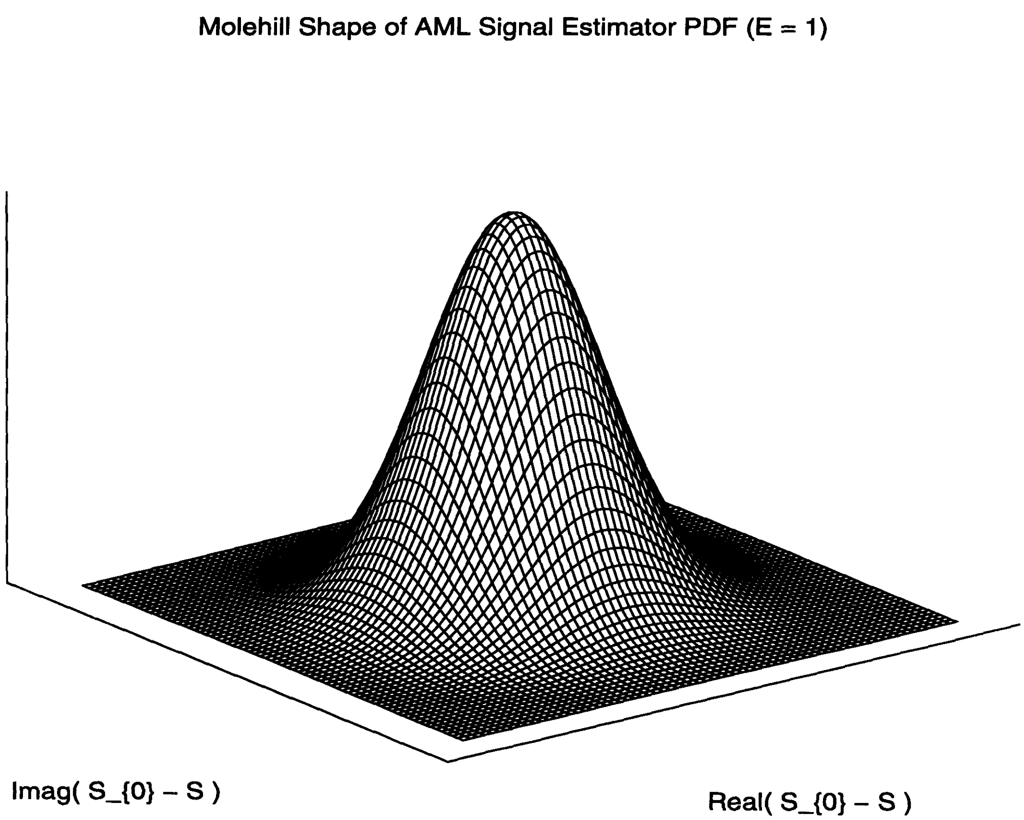

Figure 4.1: Molehill Shape of AML Signal Estimator PDF

To show the convergence, we slice this mound down the middle and observe the asymptotic behavior of a slice of the pdf in fig(4.2). In fig(4.2) a $B = 64$ element array is chosen, $|\hat{S}_0 - S|$ is in units of the standard deviation of the ML signal estimator $\sigma_{S_{ML}} = (\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g})^{-1/2}$, and the pdf constant of proportionality is $\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}/\pi$; hence, the area under each curve in fig(4.2) is $\pi/\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}$ to guarantee unity. Note that when

$Q \geq 3B$ the pdf of the AML estimator does not deviate too badly from the Gaussian. The larger the array the more samples $\hat{\mathbf{R}}$ needs to learn from before it is confident that it has a good estimate of $\mathbf{R}$. It is clear from fig(4.2) that for values of Q closer to B there is a significant increase in the spread of the pdf. This spread is undoubtedly the manifestation of the uncertainty present in the noise covariance estimator $\hat{\mathbf{R}}$. An understanding, therefore, of how these parameters affect the spread of the distribution should prove insightful. A moment analysis is therefore given in the next section.

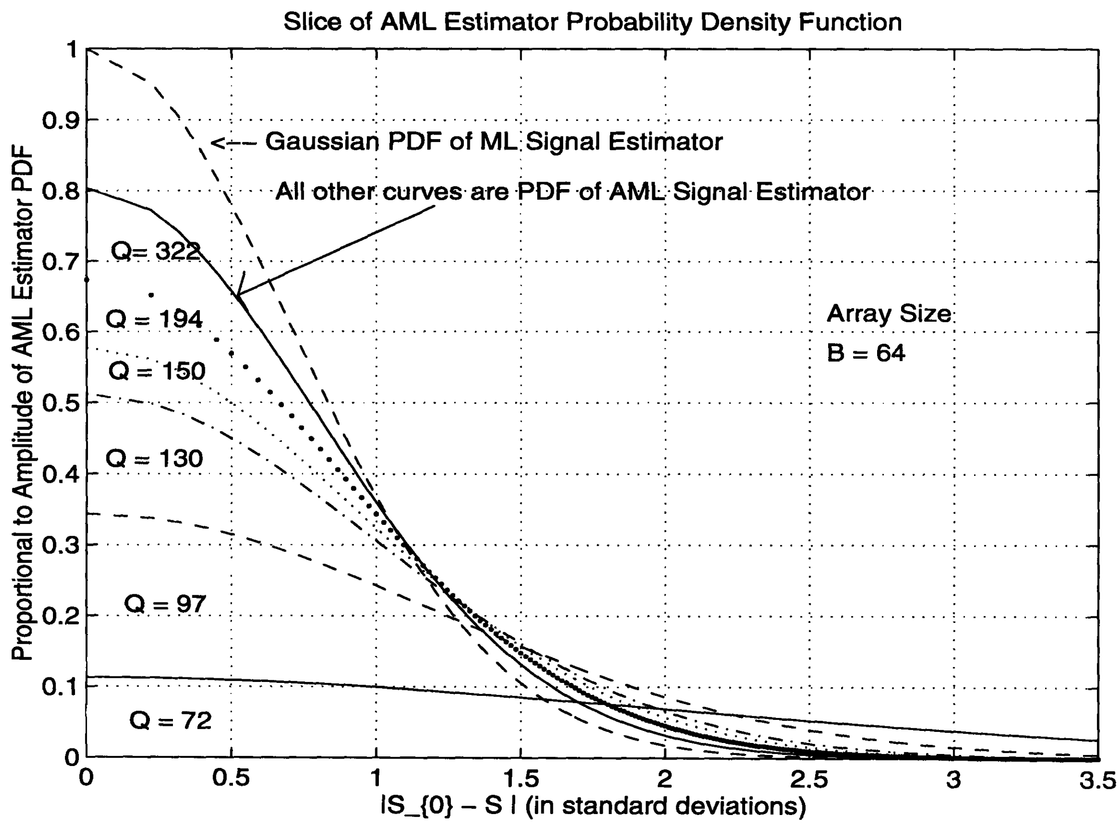

Figure 4.2: PDF of S_{AML} as Q varies; $|\hat{S}_0 - S|$ is in units of $(\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g})^{-1/2}$

4.2 Moments of the AML Signal Estimator

4.2.1 Second Moment of AML Estimator

It can be shown that the associated *moment generating function* of the AML estimator is

$$\Phi_{S_{AML}}(X,Y) = E\left\{\exp\left[X \cdot \mathrm{Re}(S_{AML}) + Y \cdot \mathrm{Im}(S_{AML})\right]\right\} = \\ e^{X \cdot S_R + Y \cdot S_I} \times {}_2F_2\left(1, -Q; 1, B-Q-1; \frac{X^2+Y^2}{4\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}}\right) \tag{4.5}$$

where the region of integral convergence is $X^2+Y^2 < 4\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}$ and where the true signal value is $S = S_R + jS_I$ (see Appendix B). From this moment generating function one can find the variance of the AML estimator $\sigma^2_{S_{AML}} = E\{|S_{AML}|^2\} - |E\{S_{AML}\}|^2$. Finding the second moment and all other moments and mixed moments of the AML estimator can be done by noting the following differentiation formula from [20, 21]

$$\frac{d^n}{dz^n}{}_2F_2(a_1, a_2; b_1, b_2; z) = \frac{(a_1)_n(a_2)_n}{(b_1)_n(b_2)_n}{}_2F_2(a_1+n, a_2+n; b_1+n, b_2+n; z), \tag{4.6}$$

and recalling from Appendix A that ${}_2F_2(a_1, a_2; b_1, b_2; 0) = 1$. The AML variance, however, can be more easily obtained from the work of [4] via the chain rule of expectations $E\{x\} = E\{E\{x|y\}\}$; namely,

$$\sigma^2_{S_{AML}} = E\left\{\frac{1}{\vartheta}\right\}\frac{1}{\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}} = \left[\frac{Q}{Q-B+1}\right]\frac{1}{\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}} = \left(\begin{array}{c}\text{Loss}\\ \text{Factor}\end{array}\right) \times \frac{1}{\text{SNR}} \tag{4.7}$$

Note that by definition $E\{1/\vartheta\} = \int_0^1 (1/\theta)p_\vartheta(\theta)d\theta$, which the reader can easily verify with the aid of eq(3.5), yields the above. Clearly the spread of the pdf is a function of the *ratio* of Q to B and the SNR $\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}$.

4.2.2 SNR and Noise Adaptivity Tradeoff

It is interesting at this point to observe from eq(4.7) that $\sigma^2_{S_{AML}} \to \sigma^2_{S_{ML}} = 1/\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}$ as $Q \to \infty$ (B fixed), but this is also true when $B \to 1$ (Q fixed). The rather peculiar asymptotic behavior of the latter begs the following question: If the total number of noise snapshots $\mathbf{n}_i$ available is constrained or limited (Q fixed), can the dimensionality of the array data B be varied in order to improve the statistics (increase confidence level for fixed region of confidence) of the AML estimator? The answer certainly appears to be yes in view of the variance eq(4.7). This predicted improvement is illustrated in fig(4.3).

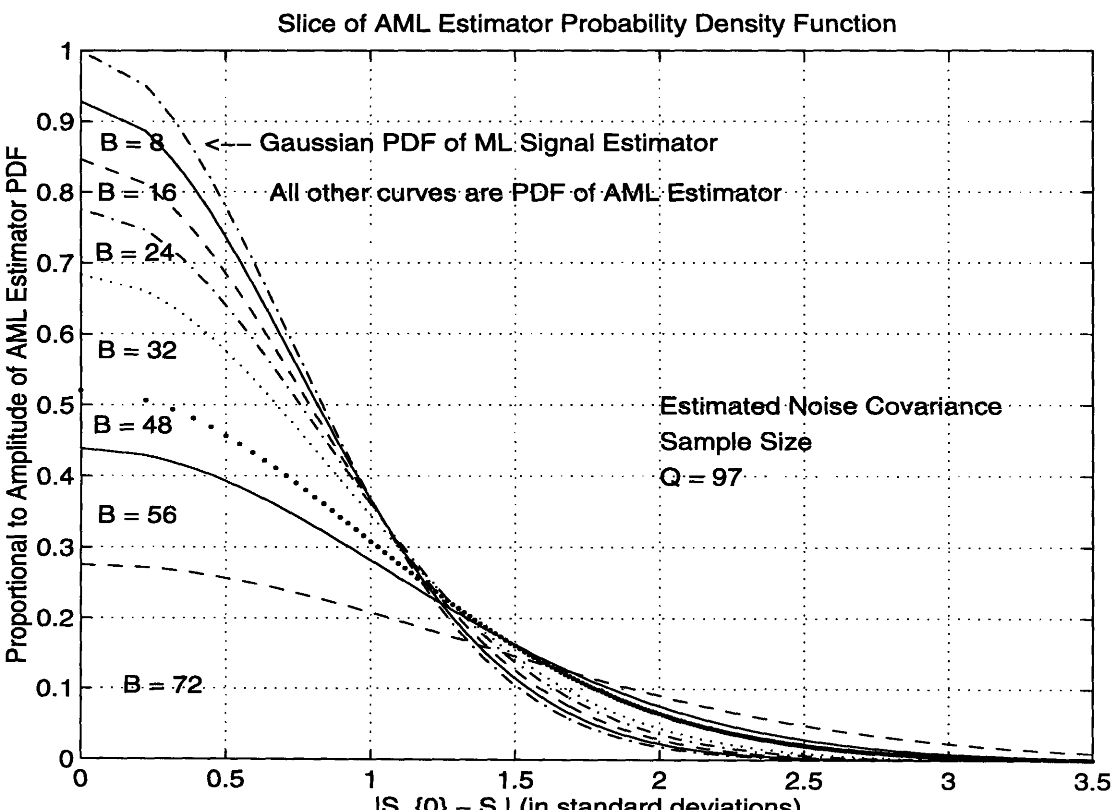

Figure 4.3: PDF of S_{AML} as B varies; $|\hat{S}_0 - S|$ is in units of $(\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g})^{-1/2}$

In this figure Q is arbitrarily fixed at 97 and the array data dimension B is varied. As B is reduced the covariance estimator $\hat{\mathbf{R}}$ reduces in dimension. Consequently, it has less to correlate and requires less information (fewer samples) to adapt. As a result $\hat{\mathbf{R}}$ converges faster in probability to $\mathbf{R}$ and yields the potentially better statistics shown in fig(4.3). Of course this sounds all too good to be true without any strings attached, and indeed it is. $B = 1$ would appear to be statistically the optimal choice; however, the array data dimension should not be reduced below a certain level or one will suffer a significant loss in SNR (or *array gain*). There's an apparent trade off between noise adaptivity and SNR as the dimensionality of the data B varies.

The first factor of eq(4.7), which is given by the ratio $Q/(Q - B + 1)$, is referred to as the *loss factor* [6, 2] because it is always ≥ 1 (because $Q \geq B$ to insure invertibility of $\hat{\mathbf{R}}$). Hence, there will *always* be an increase in the variance (a loss in performance) of the AML estimator S_{AML} over that of S_{ML} as a result of employing an estimated noise covariance. The second factor in eq(4.7) is inversely proportional to the SNR. It represents what is gained by using multiple sensors as opposed to a single sensor.

To reinforce these central ideas and illustrate the existing tradeoff consider the simplest case in which $\mathbf{R} = \sigma^2 \mathbf{I}$ and $||\mathbf{g}||^2 = B$, *i.e.* the same level of spatially uncorrelated (*white*) noise is experienced on each sensor, and each sensor has the same amplification of signal. This scenario is often encounter in beamforming applications (see [2] Ch. 4), and it yields

$$\sigma^2_{S_{ML}} = \frac{\sigma^2}{B}, \text{ and } \sigma^2_{S_{AML}} = \left[\frac{Q}{Q - B + 1}\right] \times \left[\frac{\sigma^2}{B}\right] = \frac{Q \times \sigma^2}{f(B)} \tag{4.8}$$

where $f(B) = -B[B - (Q+1)]$. This function $f(B)$ is plotted in fig(4.4) and is clearly parabolic in B (Q is assumed fixed).

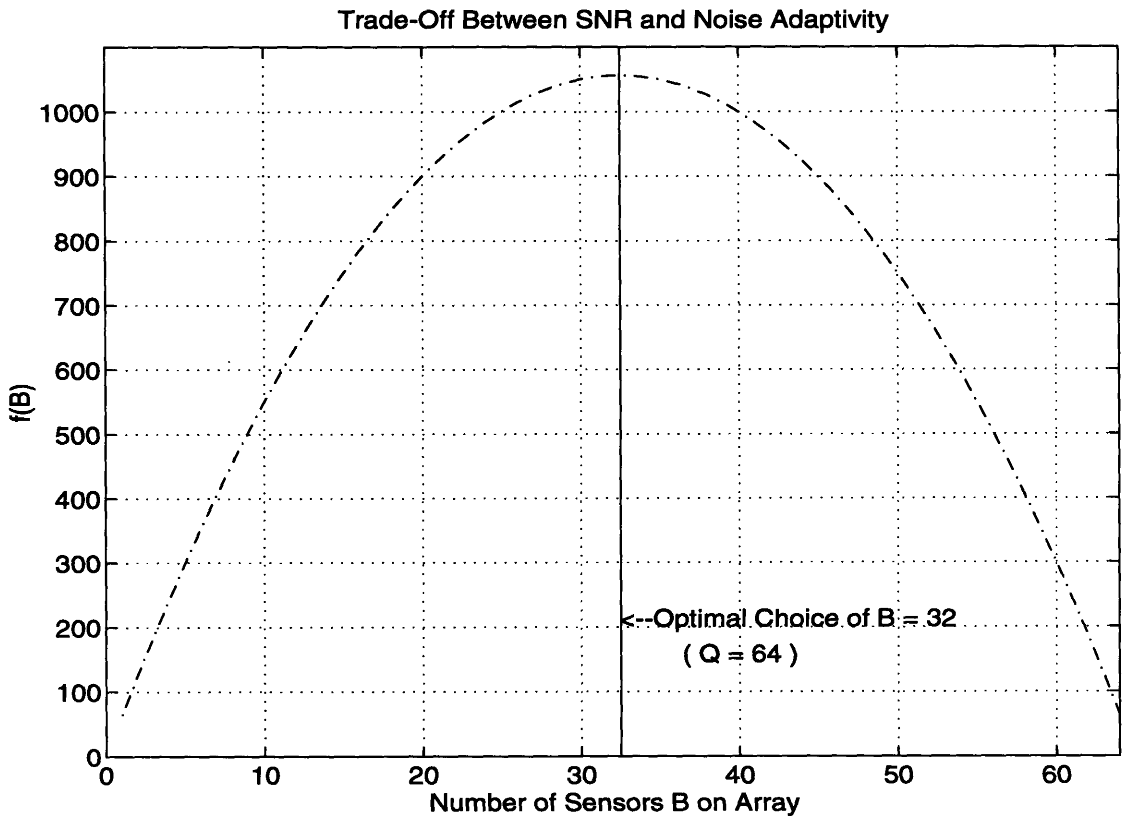

Figure 4.4: SNR and Noise Adaptivity Trade-off as B varies

Note that as B increases $\sigma^2_{S_{ML}}$ decreases monotonically. Hence, when **R** is known exactly and S_{ML} can therefore be used to estimate S then we will always do better by adding more sensors to the array (increasing B). In contrast note (with Q fixed at 64) from fig(4.4) that as the number of sensors B increases from 1 to 32 the AML variance $\sigma^2_{S_{AML}}$ decreases, and hence the AML performance improves. This performance improvement is a result of an increase in SNR and is in fact the benefit of employing an array of sensors as opposed to a single sensor. When B exceeds 32, however, the AML variance begins to increase, and hence the performance degrades. This is a consequence of the AML estimator's dependence upon $\hat{\mathbf{R}}$. When both (1) the sample size Q of the estimated noise covariance $\hat{\mathbf{R}}$ remains fixed, and (2) more sensors are added to the array (B increases and hence the dimensions of $\hat{\mathbf{R}}$ increase), then the confidence level of $\hat{\mathbf{R}}$ begins to fall. The declining confidence of the estimated noise covariance is manifested in the AML by an increase in variance. The reader can verify that the *optimal* choice of the parameter B which minimizes the AML variance is

$$\boxed{B_{OPT} = (Q+1)/2} \tag{4.9}$$

for the scenario $\mathbf{R} = \sigma^2\mathbf{I}$ and $||\mathbf{g}||^2 = B$, *i.e.* by setting $df(B)/dB = 0$ and solving for B. This optimal choice of parameter B_{OPT} yields a minimum variance of

$$\boxed{\min_{B} \quad \sigma^2_{S_{AML}} = \frac{4\sigma^2 Q}{(Q+1)^2}.} \tag{4.10}$$

Note that the for this choice of B the AML variances decreases approximately as $4\frac{\sigma^2}{Q}$. [1]

In the next section numerical simulations will support the theoretical results found in this section.

[1]This is not the most useful scenario in adaptive array processing. Indeed, there usually exist spatial correlation among sensors.

4.3 Empirical Data Analysis

The following presentation of statistical measures is made in order to develop an awareness and feel for the level of increase one can anticipate in the spread of the distribution of the AML estimator over that of the ML estimator. This presentation will in some sense develop a motivation and desire to improve upon the statistical performance of the AML signal estimator and validate empirically some of the theoretic results derived in the previous sections.

Consider the following scatter plot in fig(4.5) of many repeated ML signal estimations made from data consisting of spatially white noise, *i.e.* $\mathbf{R} \propto \mathbf{I}$:

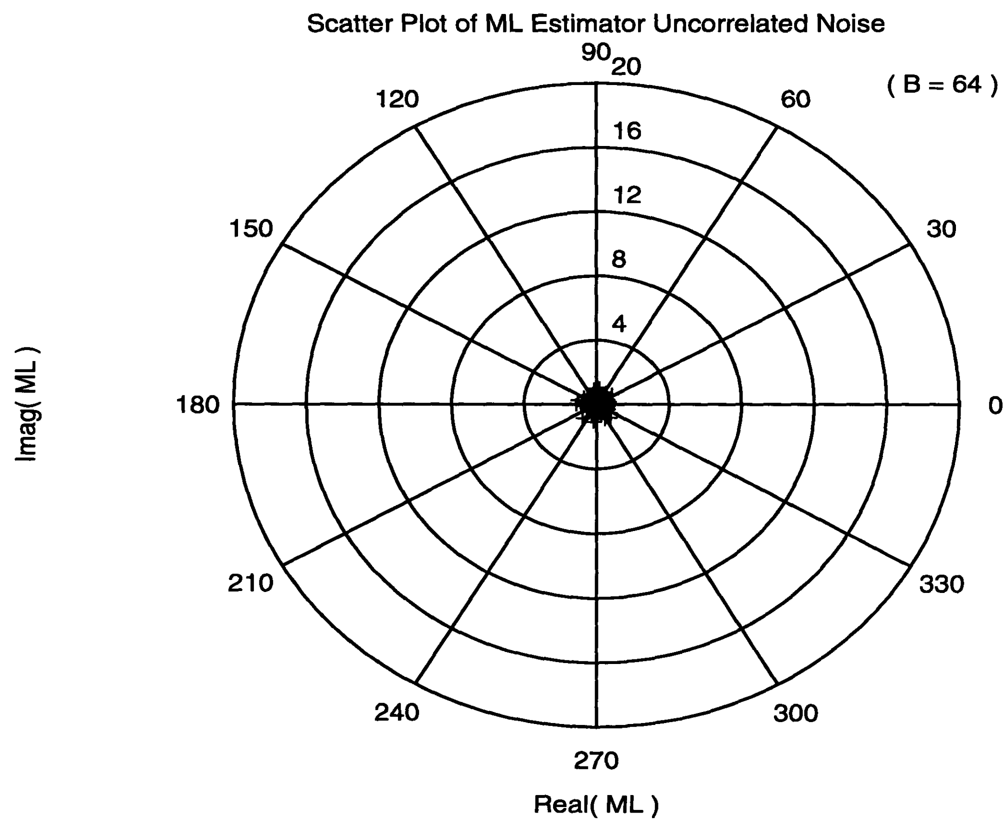

Figure 4.5: Scatter Plot of 300 samples of $S_{ML} - S$

This plot has been centered such that the true value of the signal S is the origin of the plot. The array size is $B = 64$. This scatter plot was generated by repeatedly (300 samples) simulating the received data and subsequently estimating the signal S via the estimator S_{ML}. This plot will in effect serve as a reference to which to compare the performance of the AML estimator.

Now consider the analogous scatter plot of the AML estimator based on the same received data:

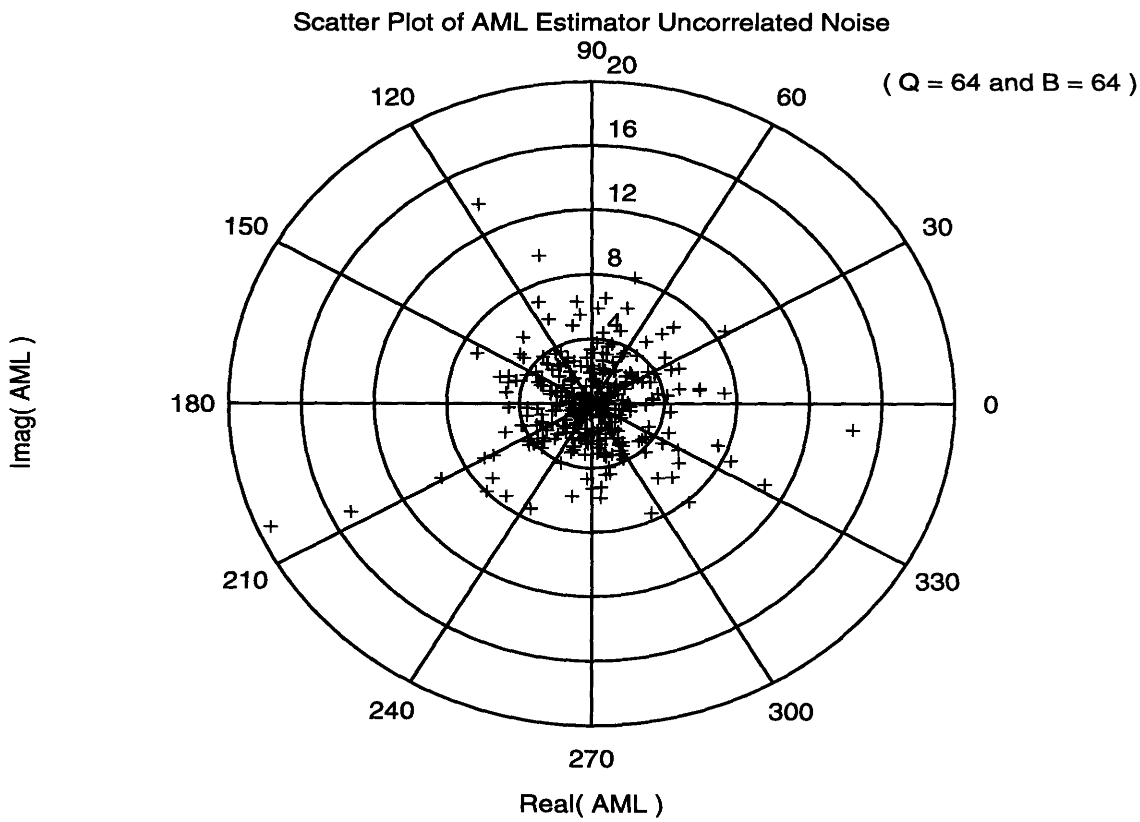

Figure 4.6: Scatter Plot of 300 samples of $S_{AML} - S$

Note that both appear to be circularly symmetric about S as required. The pdf of the phase of $S_{AML} - S$ is in fact uniform between $[0,\ 2\pi]$. There is, however, an apparent increase in the spread about S for the AML estimator. Indeed, there are several outliers apparent in the AML scatter plot. These outliers become more evident in the following boxplots of the magnitudes $|S_{AML} - S|$ and $|S_{ML} - S|$ in fig(4.7); namely, the third column of each plot corresponds to the same data in the scatter plots:

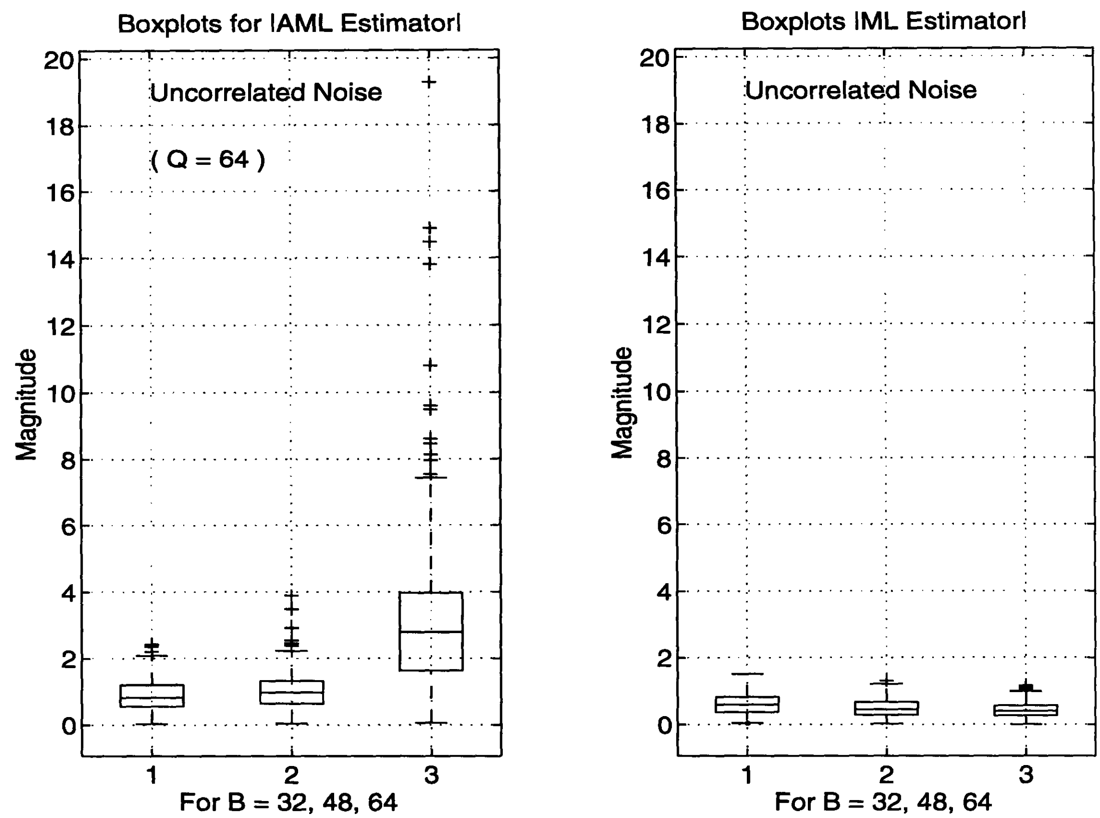

Figure 4.7: Boxplots of 300 samples of $|S_{AML} - S|$ and $|S_{ML} - S|$

Note especially from fig(4.7) that as B increases, the spread of the ML estimates decreases while on the other hand that of the AML increases! Clearly the SCM is the source of this performance degradation. The actual cause for this phenomena was given in the previous section on moments section 4.2. We saw that for a spatially white noise environment that the variance of the AML estimator was inversely parabolic in B. Fig(4.7) appears to support this theoretic result empirically.

Consider the following experiment in which many repeated (300) AML and ML signal estimations were made from simulated array data consisting of spatially white noise. In this experiment, however, the array size B was allowed to vary and the resulting sample variance was computed from the 300 sample estimates as a function of parameter B. The results appear in fig(4.8).

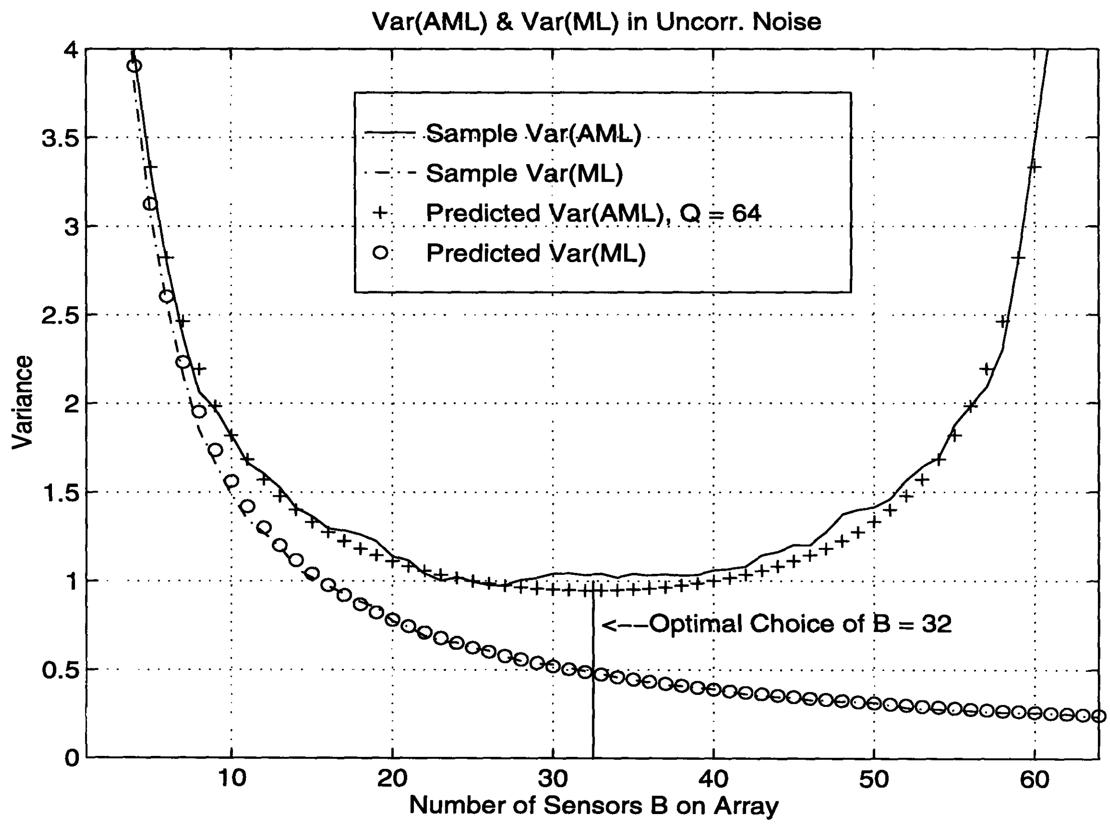

Figure 4.8: Sample Variance of Repeated Signal Estimations

To compare these empirical results to the predicted theoretic results a plot of eq(4.8) appears in fig(4.8) as the predicted variances. The match is extremely good! The variance of the ML estimator $\sigma^2_{S_{ML}}$ decreases monotonically as $1/B$ and that of the AML estimator is inversely parabolic in B with an optimal array data dimension of $B_{OPT} = (Q+1)/2 \approx 32$. This recondite behavior of the AML estimator suggests that lower dimensional data could potentially yield better performance than that obtained when *blindly* applying the AML estimator directly to the given data (sometimes referred to as the *fully adaptive scenario*).

Lastly, consider the q-q plot of the quantiles of $|S_{ML} - S|$ versus $|S_{AML} - S|$:

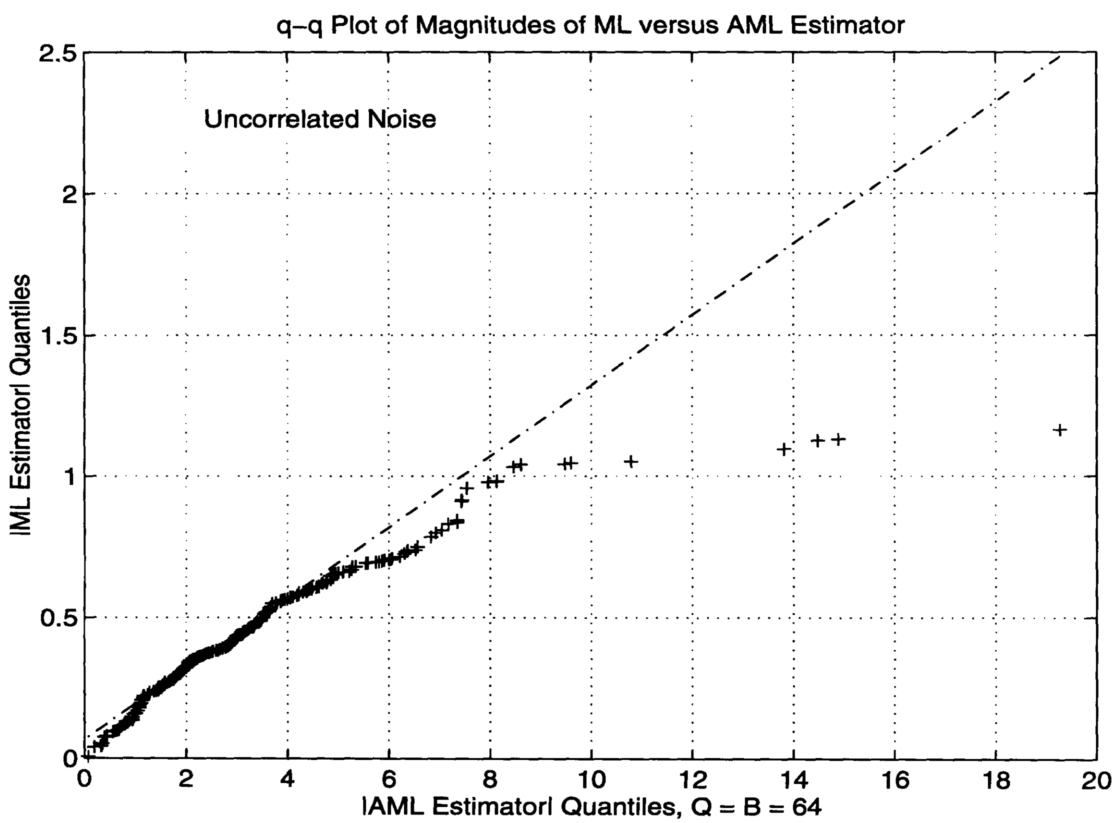

Figure 4.9: q-q plot of $|S_{ML} - S|$ versus $|S_{AML} - S|$

What's nice about q-q plots is that they allow one to see the difference in two distributions via samples taken from those distributions. Even if the exact distributions of the samples are unknown these plots convey their differences. If the samples come from the same distribution then the quantiles will form a linear pattern when plotted versus one another. In fig(4.9) a least squares best linear fit of the resulting pattern also appears. It is interesting to note that the lower percentage quantiles appear to form a fairly linear pattern, suggesting that a Gaussian distribution with an appropriately chosen variance would serve as a decent approximation for the AML pdf in this region. Note, however, that the higher percentage quantiles of $|S_{AML} - S|$ fall much further out than those of $|S_{ML} - S|$. Also, the q-q pattern is very non-linear for these higher percentage quantiles, suggesting that a Gaussian approximation to the AML pdf would be very poor in this region. Overall there is room for improvement in the AML signal estimation considering how drastically it can potentially differ statistically from the ML estimator in variation alone. Indeed, as in fig(4.9) the peak loss can be on the order of Q in the worst case scenario ($B = Q$, see eq(4.7)).

4.4 Confidence Intervals for the AML Signal Estimator

When estimations are made in practice it is useful to attach to them a number between zero and unity conveying the level of trust one should have in these estimations. This normally consists of quantifying for the user a region (in $\Re^n$−space generally) around the estimate, that is certain to contain the *true value* of the parameter at least $(1 - \alpha_0)\%$ of the time. This region is called the $(1 - \alpha_0)\%$ confidence region. Confidence regions for the AML signal parameter estimator can be obtained from the pdf eq(3.9). Note that the cumulative distribution function for the AML signal estimator can be shown to be (see Appendix C for details)

$$Pr\left(|S_{AML} - S| \leq R_{CR}\right) = 1 - {}_1F_1(Q - B + 2, Q + 1; -R_{CR}^2 \mathbf{g}^H \mathbf{R}^{-1} \mathbf{g}) \qquad (4.11)$$

from which the confidence regions of fig(4.10) are obtained. Note that from such a plot it is possible to determine the number of noise samples Q needed to ensure an $(1-\alpha_0)\%$ level of confidence for a fixed radius R_{CR} about an estimate S_{AML}. Since the pdf of S_{AML} is a function of $|S_{AML} - S|$, each confidence region consists of a circular region centered about the resulting signal estimate. In fig(4.10) the radius of this circular region is measured in units of the ML signal estimator's standard deviation $\sigma_{S_{ML}}$. As an example, to obtain a 70% confidence level for a circular region of radius 1.5 $\sigma_{S_{ML}}$ when a $B = 64$ sensor array is employed fig(4.10) indicates that Q must be no less than approximately 150.

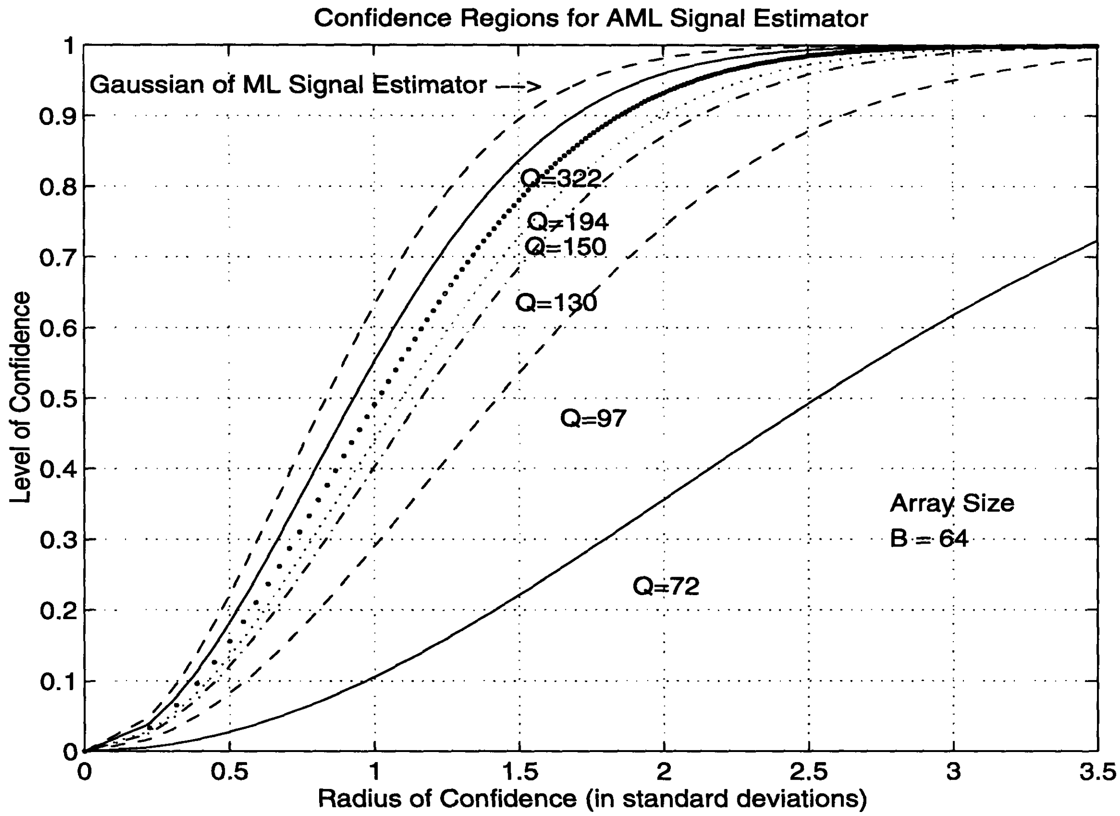

Figure 4.10: Confidence Regions for AML Estimator; R_{CR} is in units of $(\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g})^{-1/2}$

Chapter 5

Conclusions

5.1 Summary

We began by deriving the exact pdf for the sample covariance dependent AML signal estimator S_{AML}, discovering Kummer's function to be that distribution. With an exact pdf in hand standard statistical measures were found and estimator properties were deduced; namely, confidence intervals were evaluated, and the AML estimator was shown to remain unbiased, but only asymptotically efficient. S_{AML}, moreover, was shown to converge in distribution to the Gaussian non-adaptive beamformer output S_{ML}, which is an even stronger statement statistically speaking. A detailed moment analysis of the AML estimator revealed a not so obvious phenomena. Specifically, it was demonstrated that when the sample size of the estimated noise covariance matrix is fixed, there exist a dynamic tradeoff between Signal-to-Noise Ratio and noise adaptivity as the dimensionality of array data was varied. Such intelligence allowed for the design of a "truly" optimal processor for the scenario of a uniform linear array in a spatially white noise environment. This optimal dimension was shown to be $B_{OPT} = (Q+1)/2$. One major implication of these results concerns SCM's of "small" sample size. By small we mean approximately $Q < 2B$; namely, to suggest more spatial samples, *i.e.* adding more sensors to the array and consequently increasing B, as a means of compensating for the lack of noise snapshots $\mathbf{n}_i$ could be imprudent. This is quite evident in view of eq(4.7). As an example, say $\mathbf{R} = \mathbf{I}$, $Q = 60$, $B = 40$ and hence a loss factor of 2.8571 and a variance of 0.0714. Now lets add 20 more sensors to the array. Although this increases the SNR significantly it also increases the loss factor to 60 and hence the variance to 1!

Steinhardt [3] has derived *marginal* pdfs for the (filter) weights $\hat{w}_i$ of the AML signal parameter estimator; namely, the filter given by

$$\hat{\mathbf{w}}_{OPT} = [\hat{w}_1, \hat{w}_2, \ldots, \hat{w}_B]^T = \hat{\mathbf{R}}^{-1}\mathbf{g}/(\mathbf{g}^H\hat{\mathbf{R}}^{-1}\mathbf{g}), \tag{5.1}$$

where T denotes matrix transposition and the filter output is $S_{AML} = \hat{\mathbf{w}}_{OPT}^H\mathbf{x}$. In this thesis, however, the derivation of the pdf for the resulting signal estimate (output of the filter) is given.

5.2 Suggestions for Further Research

As for further research one can consider the following unresolved issues. One of the most interesting outcomes of this research was proposing the dimensionality of the array data as a parameter over which to optimize the AML processor's performance. Recall that when the optimal array dimension was computed we had to assume a structure for the true noise covariance $\mathbf{R}$; however, if we knew more about the structure of $\mathbf{R}$ then a SCM estimate may not be the optimal choice of covariance estimators [13]. For example, if we knew that $\mathbf{R}$ was proportional to the identity matrix $\mathbf{I}$ then we'd use $S_{ML} = \mathbf{g}^H\mathbf{x}/||\mathbf{g}||^2$ as a signal estimator since the ML estimator is invariant to that proportionality constant. So the problem is that we rarely know much about $\mathbf{R}$ and hence can not optimize the processor over B since it depends on the true noise covariance. I propose that a methodology be sought which initially attempts to whiten the data and subsequently seeks to optimize over B, *i.e.* try to reduce the problem to something we've already solved.

Appendix A

Kummer's Function

This appendix is not intended to be a detailed exposition of the origins and history of Kummer's function and the related hypergeometric functions. For such theoretic excursions the reader is referred to [20]. The following presentation provides essentially enough information to embrace the ideas presented in this thesis. Most of the following facts about the confluent hypergeometric function have been taken from [21].

Kummer's function is one of two linearly independent solutions of the following ordinary linear differential equation

$$z^2\frac{d^2u}{dz^2} + (b - z)\frac{du}{dz} - au = 0 \tag{A.1}$$

known as *Kummer's Equation.* The confluent hypergeometric function results when a power series form of the solution is assumed and the coefficients of the power series are evaluated as constrained by the differential equation. The power series representation of

Kummer's function that result is

$$ {}_1F_1(a,b;z) = \sum_{n=0}^{\infty} \frac{(a)_n}{(b)_n} \frac{z^n}{n!} = 1 + \frac{az}{b} + \frac{a(a+1)}{b(b+1)} \frac{z^2}{2} + \cdots \tag{A.2} $$

where the coefficients of the power series are defined to be

$$ (a)_n = \begin{cases} a(a+1)(a+2)\cdots(a+n-1) & n = 1,2,3,\ldots \\ 1 & n = 0. \end{cases} \tag{A.3} $$

This power series for Kummer's function is a special case of the following *generalized hypergeometric function*:

$$ {}_pF_q\left(a_1, a_2, \ldots, a_p; b_1, b_2, \ldots, b_q; z\right) = \sum_{n=0}^{\infty} \frac{(a_1)_n (a_2)_n \cdots (a_p)_n}{(b_1)_n (b_2)_n \cdots (b_q)_n} \frac{z^n}{n!}. \tag{A.4} $$

This series is known to converge when certain conditions are satisfied. One such relevant condition is that (i) $p \le q$ and (ii) the argument is finite in magnitude, *i.e.* $|z| < \infty$; hence, for finite values of the argument of eq(3.9) these conditions for convergence are satisfied. Its not difficult to verify these conditions via the ratio test for series convergence. To illustrate another useful fact note that since

$$ {}_pF_q\left(a_1, a_2, \ldots, a_p; b_1, b_2, \ldots, b_q; z\right) = \tag{A.5} $$
$$ 1 + \frac{a_1 a_2 \cdots a_p}{b_1 b_2 \cdots b_q} z + \frac{a_1(a_1+1) a_2(a_2+1) \cdots a_p(a_p+1)}{b_1(b_1+1) b_2(b_2+1) \cdots b_q(b_q+1)} \frac{z^2}{2} + \cdots $$

then it is true that ${}_pF_q\left(a_1, a_2, \ldots, a_p; b_1, b_2, \ldots, b_q; 0\right) = 1$. This will be useful for deriving all mixed moments of the AML estimator (see section 4.2).

Integrating Kummer's Equation will yield the following integral representation for the confluent hypergeometric function [21]:

$$ {}_1F_1(a; b; z) = [\beta(a, b-a)]^{-1} \int_0^1 e^{xz} x^{a-1} (1-x)^{b-a-1} dx \tag{A.6} $$

where $\beta(a,c) = \Gamma(a)\Gamma(c)/\Gamma(a+c)$. Note that eq(A.6) implies the following equivalence

$$ \begin{aligned} {}_1F_1(a; b; -z) &= [\beta(a, b-a)]^{-1} \int_0^1 e^{-xz} x^{a-1} (1-x)^{b-a-1} dx \\ &= [\beta(a, b-a)]^{-1} (-1) \int_1^0 e^{-xz} x^{a-1} (1-x)^{b-a-1} dx. \end{aligned} \tag{A.7} $$

Now make the change of variables $y = 1 - x$ and $-dx = dy$. The integral becomes

$$ \begin{aligned} {}_1F_1(a; b; -z) &= [\beta(a, b-a)]^{-1} \int_0^1 e^{z(y-1)} y^{b-a-1} (1-y)^{a-1} dy \\ &= e^{-z} [\beta(b-a, a)]^{-1} \int_0^1 e^{zy} y^{(b-a)-1} (1-y)^{a-1} dy \\ {}_1F_1(a; b; -z) &= e^{-z} {}_1F_1(b-a; b; z). \end{aligned} \tag{A.8} $$

This resulting equivalence is very useful in practice. Indeed, it allows one to relate Kummer's function of negative arguments those of positive ones. It is known as *Kummer's Transformation.*

Appendix B

Derivation of AML Moment Generating Function

The following is a derivation of the moment generating function for the pdf of the AML estimator S_{AML}.

Defining the following variables for notational convenience

$$\alpha = 1/\mathbf{g}^H \mathbf{R}^{-1} \mathbf{g}, \quad a = Q - B + 3, \quad b = Q + 2, \qquad \zeta = (a-1)/[\pi\alpha(b-1)]$$

$$S = S_R + jS_I, \quad \hat{S}_0 = \hat{S}_{0R} + j\hat{S}_{0I}, \quad \mathbf{z} = [\mathrm{Re}(S_{AML}),\ \mathrm{Im}(S_{AML})]^T,$$

the pdf in eq(3.9) is rewritten as the following real bivariate distribution

$$p_{S_{AML}}(\mathbf{z}_0) = p_{S_{AML}}(\hat{S}_{0R}, \hat{S}_{0I}) = \zeta \cdot {}_1F_1\left(a, b, -[(\hat{S}_{0R} - S_R)^2 + (\hat{S}_{0I} - S_I)^2]/\alpha\right). \quad \text{(B.1)}$$

Let $\mathbf{d} = [X, Y]^T$. The moment generating function of the AML estimator is by definition given by the following expectation

$$\Phi_{S_{AML}}(X, Y) = E\left\{e^{\mathbf{z}^T \mathbf{d}}\right\} = E\left\{\exp\left[X \cdot \mathrm{Re}(S_{AML}) + Y \cdot \mathrm{Im}(S_{AML})\right]\right\} = \quad \text{(B.2)}$$

$$\zeta \cdot \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} d\hat{S}_{0R}\, d\hat{S}_{0I}\; e^{X\cdot\hat{S}_{0R}+Y\cdot\hat{S}_{0I}}\,{}_1F_1\left(a, b, -[(\hat{S}_{0R}-S_R)^2+(\hat{S}_{0I}-S_I)^2]/\alpha\right). \quad \text{(B.3)}$$

Let $w = \hat{S}_{0R} - S_R$ and $v = \hat{S}_{0I} - S_I$. This yields

$$\Phi_{S_{AML}}(X,Y) = \zeta \cdot \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dw\; dv\; e^{X\cdot(w+S_R)+Y\cdot(v+S_I)}\,{}_1F_1(a, b, -(w^2+v^2)/\alpha). \quad \text{(B.4)}$$

Let $w = r cos\theta$ and $v = r sin\theta$. Thus, $dwdv = rdrd\theta$ and $w^2 + v^2 = r^2$ and the integral becomes

$$\Phi_{S_{AML}}(X,Y) = \zeta \cdot e^{X\cdot S_R+Y\cdot S_I} \cdot \int_{0}^{\infty}\int_{-\pi}^{\pi} r\; dr\; d\theta\; e^{r(X\cdot cos\theta+Y\cdot sin\theta)}\,{}_1F_1(a, b, -r^2/\alpha). \quad \text{(B.5)}$$

Note the following integral identity ([19] p. 310)

$$\int_{-\pi}^{\pi} d\theta\; e^{(f\cdot cos\theta+g\cdot sin\theta)} = 2\pi \cdot I_0\left(\sqrt{f^2+g^2}\right) = 2\pi \cdot {}_0F_1\left[1; (f^2+g^2)/4\right]. \quad \text{(B.6)}$$

Integrating with respect to θ, and using Kummer's Transformation

$${}_1F_1(a; b; -x) = e^{-x}\,{}_1F_1(b-a; b; x), \quad \text{(B.7)}$$

the integral becomes

$$\Phi_{S_{AML}}(X,Y) = 2\pi\zeta \cdot e^{X\cdot S_R+Y\cdot S_I} \int_{0}^{\infty} rdr\; {}_0F_1\left[1; \frac{r^2(X^2+Y^2)}{4}\right] e^{-r^2/\alpha}\,{}_1F_1\left(b-a; b; \frac{r^2}{\alpha}\right) \text{(B.8)}$$

Next we make the following change of variables $x = r^2$, $dx = 2rdr$. This yields

$$\Phi_{S_{AML}}(X,Y) = \pi\zeta \cdot e^{X \cdot S_R + Y \cdot S_I} \int_0^\infty dx \; e^{-x/\alpha} \; {}_0F_1\left[1; \frac{x \cdot (X^2+Y^2)}{4}\right] \; {}_1F_1\left(b-a; b; \frac{x}{\alpha}\right) \tag{B.9}$$

Making use of the following integral identity ([20] p. 54)

$$\begin{aligned} &\int_0^\infty dx \; e^{-cx} \; x^{d-1} \; {}_pF_q\left(a_1, a_2, \ldots, a_p; b_1, b_2, \ldots, b_q; kx\right) \; {}_1F_1(a^{'}; b^{'}; cx) = \\ &\qquad c^{-d} \frac{\Gamma(d)\Gamma(b^{'})\Gamma(b^{'}-a^{'}-d)}{\Gamma(b^{'}-a^{'})\Gamma(b^{'}-d)} \times \\ &{}_{p+2}F_{q+1}\left(a_1, a_2, \ldots, a_p, d, 1+d-b^{'}; b_1, b_2, \ldots, b_q, 1+d+a^{'}-b^{'}; k/c\right) \end{aligned} \tag{B.10}$$

where $\text{Re}(d) > 0$, $\text{Re}(c) > 0$, $p \le q$, and $|c| > |k|$ the reader can verify that indeed the moment generating function is given by

$$\Phi_{S_{AML}}(X,Y) = e^{X \cdot S_R + Y \cdot S_I} \times {}_2F_2\left(1, -Q; 1, B-Q-1; \frac{X^2+Y^2}{4\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}}\right)$$

where the region of integral convergence is $X^2 + Y^2 < 4\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}$. Q.E.D

Appendix C

Derivation of AML Estimator's CDF

The following integral identity from [19] will aid in deriving the cdf of the AML signal estimator:

$$\int_0^x {}_1F_1(a, b, -t)dt = \left(\frac{b-1}{a-1}\right)\left[1 - {}_1F_1(a-1, b-1, x)\right]. \qquad \text{(C.1)}$$

Define the following variables

$$\alpha = 1/\mathbf{g}^H\mathbf{R}^{-1}\mathbf{g}, \quad a = Q - B + 3, \quad b = Q + 2, \quad \zeta = (a-1)/[\pi\alpha(b-1)]$$
$$S = S_R + jS_I, \quad \hat{S}_0 = \hat{S}_{0R} + j\hat{S}_{0I}.$$

Let $rcos\theta = \hat{S}_{0R} - S_R$ and $rsin\theta = \hat{S}_{0I} - S_I$. This gives the differential area $d\hat{S}_{0R} \cdot d\hat{S}_{0I} = rdr \cdot d\theta$, and the equivalence $(\hat{S}_{0R} - S_R)^2 + (\hat{S}_{0I} - S_I)^2 = r^2$. Hence, the cdf is given by

$$Pr\left(|S_{AML} - S| \leq R_{CR}\right) = \zeta \cdot \int_0^{2\pi} d\theta \int_0^{R_{CR}} rdr \cdot {}_1F_1(a, b; -r^2/\alpha). \qquad \text{(C.2)}$$

Making the change of variables $t = r^2/\alpha$, $\alpha dt/2 = rdr$, and noting that the integrand of eq(C.2) is independent of θ yields

$$Pr\left(|S_{AML} - S| \leq R_{CR}\right) = 2\pi\zeta \cdot (\alpha/2) \cdot \int_0^{R_{CR}^2/\alpha} dt \cdot {}_1F_1(a, b; -t). \tag{C.3}$$

Using the integral identity given in eq(C.1) the reader can verify that the cdf of the AML is indeed given by eq(4.11). Q.E.D.

Bibliography

[1] C. D. Richmond, "Derived PDF of Maximum Likelihood (ML) Signal Estimator which Employs an Estimated Noise Covariance," to appear in *IEEE Trans. on Signal Processing 1995.*

[2] 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).

[3] A. O. Steinhardt, "The PDF of Adaptive Beamforming Weights," *IEEE Trans. Signal Processing*, Vol. 39, No. 5, 1232 – 1235 (1991).

[4] 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).

[5] R. Choudary Hanumara, "An Alternate Derivation of the Distribution of the Conditioned Signal-to-Noise Ratio," *IEEE Trans. Antennas and Propag.*, Vol. AP–34, No. 3, 463 – 464 (1986).

[6] E. Kelly and K. Forsythe, "Adaptive Detection and Parameter Estimation for Multidimensional Signal Models," MIT Lincoln Laboratory report TR848 (1989).

[7] E. J. Kelly, "An Adaptive Detection Algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, Vol. AES-22, 115 – 127 (1986).

[8] B. D. Carlson, "Covariance Matrix Estimation Errors and Diagonal Loading in Adaptive Arrays," *IEEE Trans. Aerospace and Electronic Systems*, Vol. 24, No. 4, 397 – 401 (1988).

[9] S. Kalson, "Adaptive Detection with Diagonal Loading," *Proc. of IEEE ICASSP Conf.* D4.7, 1369 –1372 (1990).

[10] S. M. Kay, *Fundamentals of Statistical Signal Processing*, PTR Prentice-Hall, Englewood Cliffs, NJ (1993).

[11] R. A. Monzingo, T. W. Miller, *Introduction to Adaptive Arrays*, John Wiley & Sons, Inc. (1980).

[12] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis 2nd Edition*, John Wiley & Sons, Inc. (1984).

[13] J. P. Burg, D. G. Luenberger, D. L. Wegner, "Estimation of Structured Covariance Matrices," Proc. IEEE, Vol. 70, No. 9, 963 – 974 (Sept 1982).

[14] J. Capon N.R. Goodman,"Probability Distributions for Estimators of Frequency Wavenumber Spectrum," *Proc. IEEE*, Vol. 58, No. 10, 1785 – 1786 (1970).

[15] N. R. Goodman,"Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution," *Ann. Math. Stat.*, Vol. 34, 152 – 177 (1963).

[16] H. Van Trees, *Detection, Estimation, and Modulation Theory Part I*, Wiley, New York.

[17] P. Mignerey, S. Finette, "Multichannel Deconvolution of an Acoustic Transient in an Oceanic Waveguide," *J. Acoustical Society of America* **92**(1), 351 – 364 (1992).

[18] S. Finette, P. Mignerey, J. Smith, Christ D. Richmond, "Broadband Source Signature Extraction Using Vertical Array," *J. Acoust. Soc. of Amer.*, **94**(1), 309 – 318 (1993).

[19] I. S. Gradshteyn, I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, Inc., San Diego, CA (1980).

[20] L. J. Slater, *Confluent Hypergeometric Functions*, Cambridge University Press, (1960).

[21] M. Abramowitz, I. A. Stegun, *Handbook of Mathematical Functions*, Dover Publications, Inc., New York, NY (1970).