ABSTRACT

SPACE-TIME REDUCED RANK METHODS AND CFAR SIGNAL DETECTION ALGORITHMS WITH APPLICATIONS TO HPRF RADAR

by
Tareq F. Ayoub

In radar applications, the statistical properties (covariance matrix) of the interference are typically unknown *a priori* and are estimated from a dataset with limited sample support. Often, the limited sample support leads to numerically ill-conditioned radar detectors. Under such circumstances, classical interference cancellation methods such as sample matrix inversion (SMI) do not perform satisfactorily. In these cases, innovative reduced-rank space-time adaptive processing (STAP) techniques outperform full-rank techniques. The high pulse repetition frequency (HPRF) radar problem is analyzed and it is shown that it is in the class of adaptive radar with limited sample support. Reduced-rank methods are studied for the HPRF radar problem. In particular, the method known as diagonally loaded covariance matrix SMI (L-SMI) is closely investigated. Diagonal loading improves the numerical conditioning of the estimated covariance matrix, and hence, is well suited to be applied in a limited sample support environment. The performance of L-SMI is obtained through a theoretical distribution of the output conditioned signal-to-noise ratio of the space-time array. Reduced-rank techniques are extended to constant false alarm rate (CFAR) detectors based on the generalized likelihood ratio test (GLRT). Two new modified CFAR GLRT detectors are considered and analyzed. The first is a subspace-based GLRT detector where subspace-based transformations are applied to the data prior to detection. A subspace transformation adds statistical stability which tends to improve performance at the expense of an additional SNR loss. The second detector is a modified GLRT detector that

incorporates a diagonally loaded covariance matrix. Both detectors show improved performance over the traditional GLRT.

SPACE-TIME REDUCED RANK METHODS AND CFAR SIGNAL DETECTION ALGORITHMS WITH APPLICATIONS TO HPRF RADAR

by
Tareq F. Ayoub

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

Department of Electrical and Computer Engineering

May 1998

APPROVAL PAGE

SPACE-TIME REDUCED RANK METHODS AND CFAR SIGNAL DETECTION ALGORITHMS WITH APPLICATIONS TO HPRF RADAR

Tareq F. Ayoub

4-24-98

Dr. Alexander Haimovich, Dissertation Advisor — Date
Associate Professor of Electrical and Computer Engineering,
New Jersey Institute of Technology, NJ

4-24-98

Dr. Yeheskel Bar-Ness, Committee Member — Date
Distinguished Professor of Electrical and Computer Engineering,
New Jersey Institute of Technology, NJ

4-24-98

Dr. Stanley Reisman, Committee Member — Date
Professor of Electrical and Computer Engineering,
New Jersey Institute of Technology, NJ

4/24/98

Dr. Zoi-Heleni Michalopoulou, Committee Member — Date
Assistant Professor of Mathematics,
New Jersey Institute of Technology, NJ

4/24/98

Dr. Hong Wang, Committee Member — Date
Professor of Electrical and Computer Engineering,
Syracuse University, Syracuse, NY

BIOGRAPHICAL SKETCH

Author: Tareq F. Ayoub

Degree: Doctor of Philosophy

Date: May, 1998

Undergraduate and Graduate Education:

- Doctor of Philosophy in Electrical Engineering,
  New Jersey Institute of Technology, Newark, NJ, 1998
- Master of Science in Electrical Engineering,
  Manhattan College, Riverdale, NY, 1993
- Bachelor of Science in Electrical Engineering,
  Manhattan College, Riverdale, NY, 1992

Major: Electrical Engineering

Presentations and Publications:

A. M. Haimovich, C. Peckham, T. Ayoub, J. S. Goldstein, I. S. Reed
"Performance Analysis of Reduced Rank STAP," *Proceedings of the 1997 IEEE National Radar Conference*, Syracuse, New York, May 13-15, 1997, pp. 42-47.

T. F. Ayoub, M. L. Pugh, A. M. Haimovich
"Space-Time Adaptive Processing for High PRF Radar," *Proceedings of the IEE Radar97 Conference*, Edinburgh, UK, October 14-16, 1997, pp. 468-472.

T. F. Ayoub, M. L. Pugh, A. M. Haimovich
"Reduced Rank STAP for High PRF Radar," submitted to *IEEE Transactions on Aerospace and Electronic Systems*, February, 1998.

This dissertation is dedicated to my family.

ACKNOWLEDGMENT

I would like to express my deepest and most sincere appreciation and admiration to my dissertation advisor Dr. Alexander Haimovich for his very helpful guidance as well as encouragement throughout this work. I am very grateful to him for his advice and the constructive criticism that helped guide me and helped me not lose sight of real objectives. He is a great teacher.

Sincere thanks also go to Dr. Yeheskel Bar-Ness, Dr. Stanley Reisman, Dr. Zoi-Heleni Michalopoulou, and Dr. Hong Wang for serving on my dissertation committee. I appreciate their time as well as their comments on the dissertation. Special thanks are due to Dr. Bar-Ness, director of the Center for Communication and Signal Processing Research at NJIT, for his efforts to provide us with an ideal, intellectual and enjoyable place for research.

My gratitude is extended to Dr. Zoran Siveski who is currently a member of the technical staff at Lucent Technologies and to Dr. Radomir Bozovic who is the president and chief engineer at BNK Electronics. Dr. Siveski provided endless support and guidance especially at the beginning of my studies here at NJIT. Dr. Bozovic provided the initial incentive and push to pursue my Ph.D. degree and was always a source of technical and moral support whenever it seemed that I was at a dead end. He was also extremely generous with his books and papers.

Thanks are also due to Mr. Mark Pugh of Rome Air Force Laboratories for the valuable information and literature he provided throughout the research project. His comments and suggestions are extremely appreciated and acknowledged.

I would like to sincerely thank my father, mother, and two brothers for their constant love, tolerance, support, and encouragement. I am indebted to all of them and I truly believe that this work would not have been possible if it were not for them.

Last but not least, I would like to thank all my friends as well as my present and former colleagues at the Center for Communication and Signal Processing Research at NJIT for their friendship, help, and encouragement.

This work was supported in part by a grant from Rome Air Force Laboratories under grant number F30602-94-1-0012. This support is gratefully acknowledged.

TABLE OF CONTENTS

Chapter Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Adaptive antenna arrays are employed in communication and radar systems for the purpose of interference cancellation and desired signal reception or target detection. The main function of adaptive arrays is to dynamically place deep nulls in the antenna beam pattern in the direction of interference sources while preserving the desired signal or target. Adaptive arrays are used since conventional signal reception systems are susceptible to degradation in performance due to the presence of undesired signals in the signal environment that enter the system through the beam pattern mainlobe or sidelobes. The undesired signals may consist of deliberate interference, clutter scatterer returns, and noise sources. Adaptive antenna arrays have the ability to automatically sense the presence of undesired signals (interference signals) and suppress them while preserving the desired signal without the prior knowledge of the interference environment. Hence, adaptive arrays improve the probability of detection.

The theory of adaptive arrays and adaptive radar has been an active area of research for over three decades. Howells, Applebaum, Widrow and Frost were among the first to investigate the applications of adaptive array systems [1, 2, 3, 4]. Examples of their work are the Howells-Applebaum sidelobe canceler and Widrow's beamformer. Frost in particular suggested a linearly constrained array that preserves presumed targets while cancelling the interferences. Frost's architecture serves as the basis for the methods presented in this work.

Space-time adaptive processing (STAP) is an extension of adaptive arrays. A space-time adaptive array consists of an antenna array (spatial domain) with a tapped-delay line at each element (temporal domain). The receiver then combines the spatial and temporal samples of the received signal in order to obtain infor-

mation about interferences and subsequently cancel them. The adaptive architecture will automatically adjust the array beam pattern so that the performance is maximized. The STAP architecture is important in certain scenarios, in airborne radar for example, where interference extends in both range and Doppler simultaneously. In this case, a conventional moving target indicator (MTI) that uses only temporal degrees of freedom, is ineffective in cancelling the interference. As a result, spatial degrees of freedom must be added and hence, STAP is required. Spatial and temporal dimensions provide the processor with range and Doppler information, respectively. STAP is a multidimensional filtering approach that utilizes adaptive arrays to mitigate interference in range and Doppler simultaneously. The cost of improved performance offered by STAP architectures is increased computational overhead.

Brennan and Reed pioneered the theory of adaptive radar [5]. They showed that, in the maximum likelihood sense, the optimal detector for a target in Gaussian interference whose covariance matrix is known *a priori* is a linear detector where the output is a weighted sum of the input samples. This adaptive array architecture maximizes the output signal-to-colored noise ratio where colored noise is the aggregate of interference, clutter, and noise. This technique however, is limited in the sense that in most applications, the covariance matrix of the colored noise is not known *a priori* and has to be estimated. The covariance matrix is estimated from a set of returned data referred to as the secondary data set. This data is obtained from range cells around the range cell under test. Reed et al. suggested using the sample covariance matrix in place of the true covariance matrix in the linear detector, where the former is the maximum likelihood estimate of the latter. This method is known as the sample matrix inversion (SMI) method [6]. As a performance measure, they used the output conditioned signal-to-noise (actually colored noise) ratio (CSNR). The CSNR is defined as the ratio of the output signal-to-noise ratio

(SNR) to the optimal one which is obtained when the covariance matrix is known. Since the sample covariance matrix is random, the CSNR is also random and is bounded between 0 and 1 since it is normalized by the optimal SNR. Reed et al. derived an expression for the probability distribution of the CSNR which turned out to be a function of the array's dimension and of the size of the secondary data set used to obtain the sample covariance matrix. This distribution is not a function of the covariance matrix and hence, is independent of the interference power. It was also shown that for an array of dimension N, the SMI detector achieves an average CSNR of 0.5 or equivalently, 3 dB within the optimal SNR, if $K \approx 2N$ samples are used to estimate the covariance matrix. Obviously, the computational complexity increases with N. Other drawbacks of the SMI are the lack of robustness against calibration errors and that it is does not exhibit the desirable property of constant false alarm rate (CFAR). Various authors investigated various improvements to the SMI detector [7, 8, 9].

Kelly derived an adaptive detection algorithm for signal presence in colored noise with unknown covariance. This algorithm is based on the maximum likelihood ratio principle and is known as the generalized likelihood ratio test (GLRT) [10]. The generalization is done on the target model, where the likelihood test is maximized over all unknown target parameters. This test exhibits the CFAR property where the probability of false alarm is independent of the covariance matrix of the colored noise. The CFAR property is desirable in detectors since the colored noise environment is usually non-homogeneous and hence, for non CFAR detectors, the detection threshold has to be continuously adjusted. Kelly derived a maximum likelihood decision rule and also expressions for the probabilities of detection and false alarm. The probability of detection is dependent on the estimate of the covariance matrix. Kelly's algorithm is different from the SMI detector in that it is based on detecting a target rather than cancelling interference. This detector improves on the SMI in

exhibiting the CFAR property, however, it shows similar slow convergence properties where convergence refers to the number of samples needed to estimate the covariance matrix to result in a specified performance.

The need for a detector that combines the desirable properties of fast convergence and robustness gives way to an eigen-analysis based detector. The eigencanceler [11] and the principal component inverse (PCI) [12] are two forms of the eigen-based detector. Both of these methods compute their weight vectors based on a subspace of the signal space, and use the fact that the signal space can be partitioned into an interference subspace and a white noise subspace. The eigencanceler produces a minimum norm weight vector which is subject to a set of linear constraints and which is orthogonal to the interference subspace. The PCI produces a weight vector in the noise subspace. Under certain situations, the two methods produce similar weight vectors. Advantages of such eigen-analysis based detectors are faster convergence, higher output CSNR, and reduced complexity. It is shown in [11] that for a uniform array and fixed pulse repetition frequency (PRF), the space-time covariance matrix is low rank due to the inherent oversampling nature of the STAP architecture. Both the eigencanceler and the PCI based detectors exploit this low rank nature of the covariance matrix. Since the interference power is usually much higher than the signal and the white noise power, the large eigenvalues of the covariance matrix are dominated by the interference eigenvalues. It is shown in [11] that for an array of dimension N, the number of secondary samples needed to obtain an output SNR within 3 dB of the optimal SNR is $K \approx 2r$, where r is the interference subspace rank and $r \leq N$. The eigen-analysis based detector is a member of the class of *reduced-rank* detectors. Reduced-rank detectors take advantage of the nature of the STAP architecture from a reduced complexity standpoint. Other reduced-rank techniques were discussed and compared in [13, 14, 15].

As was mentioned earlier, the sample covariance matrix is the maximum likelihood estimate of the true covariance matrix and is obtained from a secondary data set from neighboring range cells to the range cell under test. The application of adaptive interference cancellation techniques such as the SMI for example, which utilize the inverse of the sample covariance matrix, requires the estimate to be non-singular. For this to be true, the number of secondary data samples K, must exceed the array's dimension N. It is also true that K must exceed N by a significant factor if the covariance matrix estimation is not to cause a major loss in performance [6]. This requirement of a non-singular estimate usually does not pose a problem if the adaptive array is operating in a radar that utilizes a low pulse repetition frequency (LPRF) waveform. In this case, the radar would not face any range ambiguities that would limit the number of available secondary data set samples. However, the radar sometimes operates in a limited sample support environment, such as in the case when an airborne radar is utilizing a high pulse repetition frequency (HPRF) waveform. In such an environment, due to range ambiguities, the unambiguous range is usually short and hence, the number of independent samples available is low. As a consequence, the sample covariance matrix will be singular. In this case, methods like the SMI cannot be applied. Reduced-rank methods like the eigencanceler and the PCI take advantage of such an environment.

Limited sample support relative to the large dimensionality of the adaptive array may also suggest a cascade STAP approach where spatial and temporal processing are done in cascade rather than jointly [16, 17, 18]. One such architecture is known as element-space post-Doppler. This technique reduces the STAP dimensionality by transforming the space-time data snapshot into a snapshot of array elements' data and Doppler bins. This is achieved by applying temporal filtering to the space-time data across the array elements. Consequently, the joint-STAP's N-dimensional problem is separated into q separate N/q-dimensional problems where

q is an integer. Although the receiver's performance might suffer because of reduced degrees of freedom, low sample support will not be a significant problem. This is so because, with this cascade architecture, the minimum number of samples needed to estimate the covariance matrix is $K = 2N/q$ instead of $K = 2N$ samples for the joint architecture. Hence, SMI can then be applied to the reduced dimension data. Another cascade architecture is referred to as beam-space pre-Doppler where spatial filtering is applied to the space-time data instead of temporal filtering prior to the adaptive array.

A modification of the traditional SMI technique was investigated in the early eighties by Abramovich and Cheremisin [19, 20]. They investigated the problem of detecting a signal in colored noise when the sample covariance matrix is singular due to reduced sample support as in the HPRF radar case mentioned above. They suggested that the singular sample covariance matrix can be regularized by loading it with a scaled identity matrix. Once this is done, SMI can be applied. This method is known as the loaded sample matrix inversion (L-SMI) technique. Loading is applied to the sample covariance matrix whenever it is ill-conditioned. They also found an expression for the probability distribution of the CSNR obtained by applying the L-SMI method. This CSNR distribution expression was shown to be a function of the interference subspace rank r rather than N, and the size of the sample support K. The L-SMI technique provides higher CSNR values at the output of the adaptive array than the SMI techniques for low rank interference ($r < N$). Also, L-SMI requires only $K \approx 2r$ samples to achieve a SNR within 3 dB of the optimal SNR. Abramovich and Cheremisin showed that applying loading to the sample covariance matrix, results in improved conditioning of the noise eigenvalues. The L-SMI technique, though obtains its weight vector by evaluating the inverse of the full size sample covariance matrix, behaves somewhat similar to reduced-rank

techniques where only $K \approx 2r$ samples are required for a CSNR that is within 3 dB of the optimal. This method was also investigated by other authors [21, 22].

The focus of this dissertation is to study and investigate the application of reduced-rank STAP techniques to interference cancellation in radar systems in a limited sample support environment such as the case when the radar is using a HPRF waveform, as well as to derive improved CFAR maximum likelihood detectors based on Kelly's GLRT detector. The following contributions are made:

1. The HPRF radar problem is analyzed. Due to the range ambiguous nature of the HPRF waveform, low Doppler targets are masked by strong sidelobe clutter located at relatively short ranges and steep grazing angles. Another issue associated with the HPRF waveform, is the limited sample support. This causes the covariance matrix estimate to be ill-conditioned and hence, classical SMI does not perform satisfactorily.

2. The advantages of applying reduced-rank STAP to address the two problems of range ambiguity and limited sample support are shown.

3. Due to practical limitation, errors resulting from including the desired target in the training set, are inevitable. The robustness of reduced-rank techniques against these errors is investigated and shown.

4. An expression for the probability distribution function of the CSNR for the LSMI detector in the case that the sample covariance matrix is non-singular ($K \geq N$) but ill-conditioned is derived and analyzed.

5. Kelly's GLRT CFAR detector is extended to the subspace-based GLRT detector where subspace-based transformations are applied to the data prior to detection. Expressions for false alarm and detection probabilities are obtained.

6. A CFAR generalized likelihood ratio test detector under the condition that the sample covariance matrix is diagonally loaded prior to detection, is derived and analyzed. Expressions for false alarm and detection probabilities are obtained.

This dissertation is organized as follows: Chapter 2 presents the signal model and reviews previous related work. Chapter 3 discusses the application of STAP techniques for interference cancellation in HPRF radar and analyses the robustness of these methods. The distribution of the CSNR for the LSMI technique for $K \geq N$ is presented in Chapter 4. Chapter 5 has the derivation and analysis of the subspace-based GLRT and the diagonally loaded sample covariance matrix GLRT. Chapter 6 has the conclusions.

CHAPTER 2

SIGNAL ENVIRONMENT

This chapter presents the mathematical model and definitions of the various signals used throughout this work. Also, a brief review of previous and related work is presented for reference in proceeding chapters.

2.1 Signal Model

The system considered is a pulsed Doppler airborne radar. The radar antenna is a uniformly spaced linear antenna array consisting of N_s elements. The radar transmits a coherent burst of N_t pulses at a constant pulse repetition frequency (PRF). The time interval over which N_t returns are collected is referred to as the coherent pulse interval. The resulting array is shown in Figure 2.1. After carrier demodulation, matched filtering and sampling, the data received at the array is organized into a N_sN_t-dimensional vector $\mathbf{x}(k)$, where the index represents a range gate. The vector $\mathbf{x}(k)$ contains samples of the complex envelope of a bandpass signal and hence it is complex-valued. Under hypothesis $\mathbf{H}_0$, no desired signal is present and $\mathbf{x}(k)$ is the sum of clutter $\mathbf{c}(k)$ (interference) and additive white Gaussian noise contributions $\mathbf{n}(k)$

$$\mathbf{x}(k) = \mathbf{c}(k) + \mathbf{n}(k). \tag{2.1}$$

The clutter and noise are assumed to be independent. The vector $\mathbf{x}(k)$, under hypothesis $\mathbf{H}_1$, is given by

$$\mathbf{x}(k) = a\mathbf{s} + \mathbf{c}(k) + \mathbf{n}(k), \tag{2.2}$$

where a is zero-mean, complex Gaussian random variable with variance σ_t^2 and $\mathbf{s}$ represents the space-time steering vector. The vector $\mathbf{x}(k)$ could also be defined to incorporate interference from jammers. From this point on, interference refers to

clutter contributions in the signal vector and colored noise refers to the aggregate of interference and white noise. The vector $\mathbf{x}(k)$ is made up of $N = N_s \times N_t$ samples of the complex envelope of the received signal from a specific range gate k for $k = 1, \cdots, L$. Considering returns from all possible L range gates, the vector $\mathbf{x}_k$ makes up one 'slice' in the $N_s \times N_t \times L$ data cube shown in Figure 2.2.

The basic airborne radar geometry is shown in Figure 2.3. Under the assumption that the spatial channels are co-linear, identical, omni-directional and equally spaced with spacing d, the components of vector $\mathbf{x}(k)$ due to a clutter point source can be written as

$$\begin{aligned} x_{st} &= e^{2\pi(s-1)u} e^{2\pi(t-1)\nu} \qquad s = 1, \cdots, N_s \\ & \qquad\qquad\qquad\qquad\quad t = 1, \cdots, N_t, \end{aligned} \tag{2.3}$$

where u and ν are the point source spatial and normalized Doppler frequencies and are given by

$$\begin{aligned} u &= \frac{d}{\lambda} \sin\phi \, \cos\theta \\ \nu &= \frac{2V}{\lambda \, \mathrm{PRF}} \sin\phi \, \cos\theta \end{aligned} \tag{2.4}$$

where d is the spacing between the antenna elements, ϕ is the point source azimuth angle, θ is the point source elevation angle, and V is relative speed of the point source as seen by the airborne radar.

The desired signal component of vector $\mathbf{x}(k)$, under the same conditions, can be written as

$$\mathbf{s} = \mathbf{s}_s \otimes \mathbf{s}_t \tag{2.5}$$

where $\otimes$ is the Kronecker product, $\mathbf{s}_s$ is a $N_s \times 1$ normalized spatial steering vector, and $\mathbf{s}_t$ is a $N_t \times 1$ normalized temporal steering vector. The two vectors $\mathbf{s}_s$ and $\mathbf{s}_t$

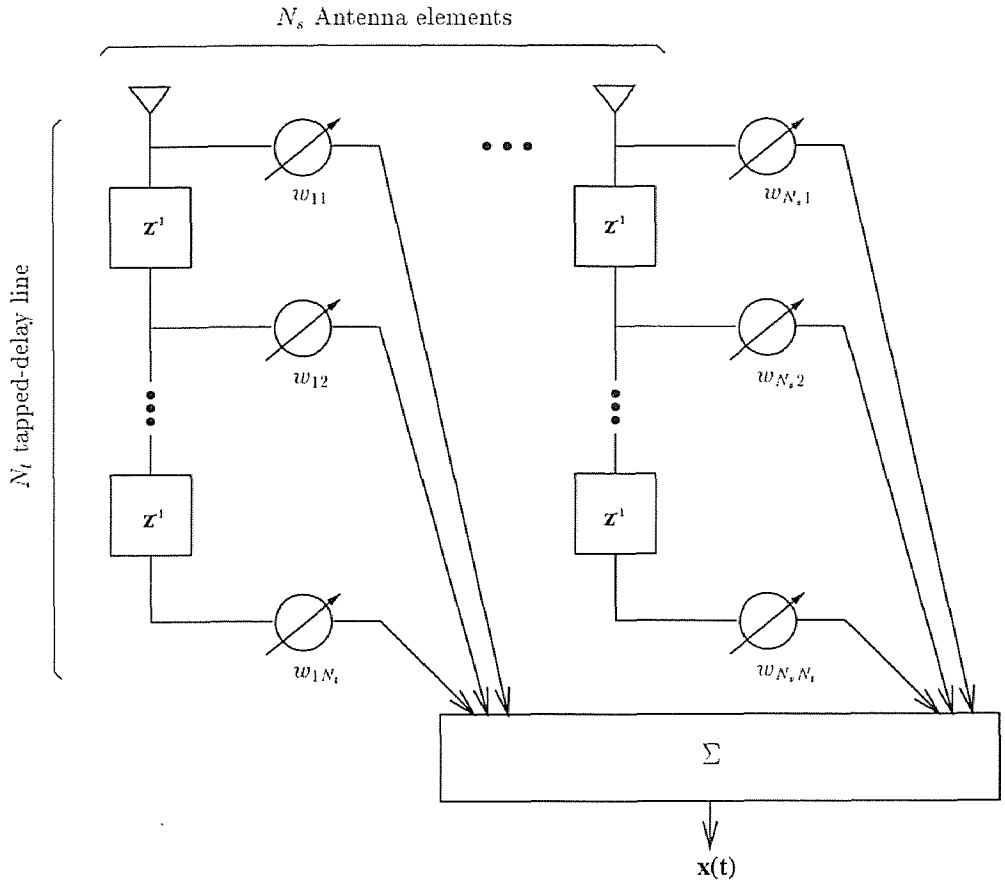

Figure 2.1 Space-time adaptive array architecture

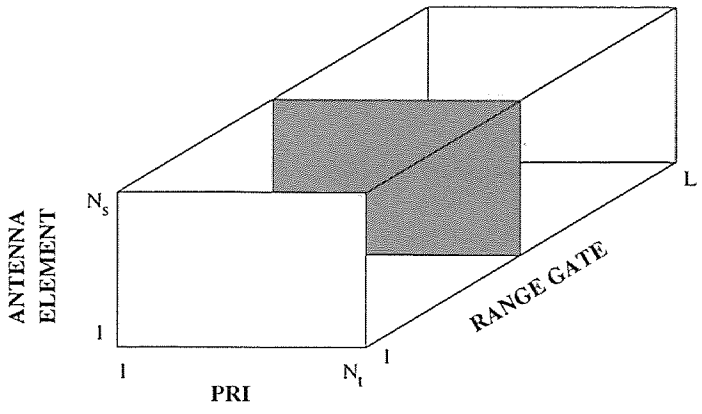

Figure 2.2 Data cube

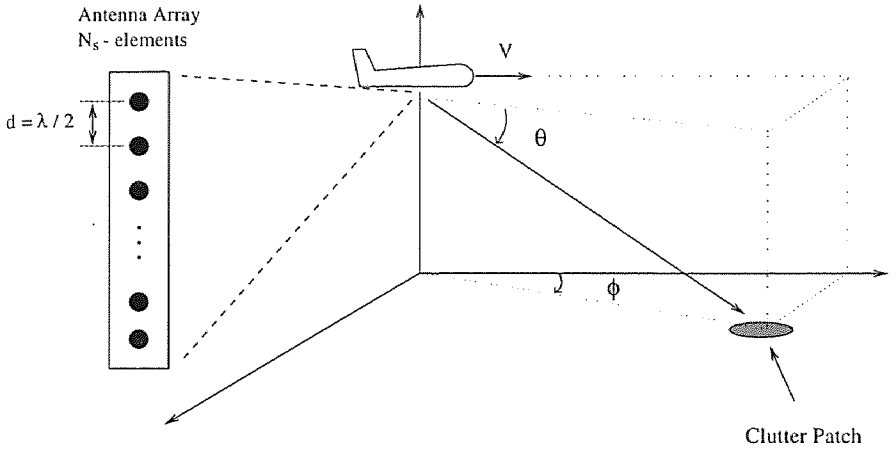

Figure 2.3 Airborne radar basic geometry.

are given by

$$\mathbf{s}_s = \begin{bmatrix} 1 \\ e^{j2\pi u_t} \\ \vdots \\ e^{j2\pi(N_s-1)u_t} \end{bmatrix} \tag{2.6}$$

and

$$\mathbf{s}_t = \begin{bmatrix} 1 \\ e^{j2\pi \nu_t} \\ \vdots \\ e^{j2\pi(N_t-1)\nu_t} \end{bmatrix} \tag{2.7}$$

where u_t and ν_t are the presumed target's spatial and normalized Doppler frequencies.

Since the colored noise true covariance matrix is usually not known *a priori*, a maximum likelihood estimate is obtained from a secondary data set from neighboring range cells around the cell under test. This estimate is referred to as the sample covariance matrix and is given as:

$$\widehat{\mathbf{R}} = \frac{1}{K}\sum_{k=1}^{K} \mathbf{x}(k)\mathbf{x}(k)^H \tag{2.8}$$

where $\mathbf{x}_k$ is defined in equation (2.1).

Equation (2.8) is an estimate of the true covariance matrix $\mathbf{R}$. For low rank interference, $\mathbf{R}$ can be decomposed into interference and white noise contributions as follows

$$\mathbf{R} = \mathbf{Q}_r \mathbf{\Lambda}_r \mathbf{Q}_r^H + \sigma_v^2 \mathbf{Q}_v \mathbf{Q}_v^H, \tag{2.9}$$

where the diagonal of the $r \times r$ matrix $\mathbf{\Lambda}_r$ consists of the r principal eigenvalues of $\mathbf{R}$, the columns of $\mathbf{Q}_r$ are the corresponding eigenvectors, σ_v^2 is the variance of the white noise, and the columns of $\mathbf{Q}_v$ are the remaining eigenvectors of $\mathbf{R}$. In [11] it was shown that the rank of the interference has an upper bound of $r_{max} = N_s + N_t - 1$. The spectral decomposition of $\mathbf{R}$ in equation (2.9) suggests the following decomposition for the estimate $\widehat{\mathbf{R}}$

$$\widehat{\mathbf{R}} = \widehat{\mathbf{Q}}_r \widehat{\mathbf{\Lambda}}_r \widehat{\mathbf{Q}}_r^H + \widehat{\mathbf{Q}}_v \widehat{\mathbf{\Lambda}}_v \widehat{\mathbf{Q}}_v^H, \tag{2.10}$$

where $\widehat{\mathbf{\Lambda}}_v$ consists of the $N - r$ eigenvalues of the noise subspace. The addition of the hat [$\widehat{\ }$] to the eigenvector and eigenvalue matrices denotes that they correspond to

the estimate $\widehat{\mathbf{R}}$. The noise eigenvalues of the true covariance matrix in equation (2.9) are all the same and given by the noise power σ_v^2. However, the noise eigenvalues obtained by decomposing the covariance matrix estimate $\widehat{\mathbf{R}}$ in equation (2.10) are going to be spread over a range of power which depends on the sample support size used to estimate the covariance matrix. Noise exists at all frequencies and hence, an infinite number of degrees of freedom is needed to fully estimate the noise subspace. As the sample support size K increases, $\widehat{\mathbf{R}}$ approaches the true covariance matrix $\mathbf{R}$ and the noise eigenvalues in $\hat{\mathbf{\Lambda}}_v$ all converge towards the noise power σ_v^2. This representation of $\mathbf{R}$ and $\widehat{\mathbf{R}}$ will be used in subsequent sections and chapters.

2.2 Background

The following sections present a brief review of previous and related work in adaptive processing for radar. This discussion is provided since it will be used in proceeding chapters of this work.

2.2.1 Optimum Signal Processing

In [5], Brennan and Reed developed the theory for optimum adaptive arrays which maximize the probability of detection. They showed that the optimal detector for a desired signal in Gaussian noise whose covariance matrix is known *a priori*, is achieved by using a Wiener filter. The weights of this filter are given by

$$\mathbf{w} = k\mathbf{R}^{-1}\mathbf{s} \tag{2.11}$$

where k is a gain constant and $\mathbf{R}$ is the colored noise true covariance matrix. The Wiener filter can be interpreted as a cascade of a whitening filter for the interference, followed by a matched filter for the desired signal. This solution requires the prior knowledge of the true covariance matrix $\mathbf{R}$.

2.2.2 Sample Matrix Inversion

In practice, the true covariance matrix of the interference, $\mathbf{R}$, is unknown *a priori*. The covariance matrix has to be estimated from a finite number of secondary data (interference and white noise) set as given in equation (2.8).

Substituting $\widehat{\mathbf{R}}$ for the true covariance matrix, $\mathbf{R}$, in equation (2.11) yields the following solution

$$\mathbf{w} = k\widehat{\mathbf{R}}^{-1}\mathbf{s}. \tag{2.12}$$

This solution for the array's weight coefficients is not optimal since the covariance matrix has to be estimated from a finite number of secondary samples.

The conditioned signal-to-colored noise (CSNR) is defined as the ratio of the actual signal-to-colored noise (SNR) ratio to the optimal one which results when the covariance is known. Again, colored noise refers to interference as well as white noise. Reed et al. showed that the solution in (2.12) achieves a CSNR at the output of the array that has a mean of 0.5 if $K \approx 2N$ snapshots are used to estimate $\widehat{\mathbf{R}}$ in equation (2.8). Since the covariance matrix is estimated using equation (2.8) from random secondary data, the resulting SNR is a random variable, and the CSNR is a random variable bounded between 0 and 1. A CSNR with a mean of 0.5 is equivalent to a 3 dB loss with respect to the optimal SNR.

The sample covariance matrix estimate given in equation (2.8) has to be continually updated since the environment maybe considered stationary for only short periods of time.

2.2.3 Pseudoinverse SMI

If the sample covariance matrix obtained using equation (2.8) is non-singular, then the SMI algorithm can be applied and the expression in equation (2.12) can be used to find the weight coefficients. A necessary condition for a non-singular sample covariance matrix is that $K \geq N$. However, in a limited sample support environment,

this condition may not be satisfied. The pseudoinverse SMI (P-SMI) technique suggests replacing the true inverse of the singular estimated covariance matrix $\widehat{\mathbf{R}}$, by the *pseudoinverse*. A low rank approximation to $\widehat{\mathbf{R}}$ is constructed from the singular value decomposition [23]:

$$\widehat{\mathbf{R}}_M = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_M & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^H \\ \mathbf{V}_2^H \end{bmatrix} = \mathbf{U}_1 \mathbf{\Sigma}_M \mathbf{V}_1^H \tag{2.13}$$

where $\mathbf{\Sigma}_M$ is the $M \times M$ diagonal matrix of the largest singular values of $\widehat{\mathbf{R}}$. The total number of singular values included in $\mathbf{\Sigma}_M$ is chosen equal to the rank of the original covariance matrix. The two unitary matrices $\mathbf{U}_1$ and $\mathbf{V}_1$ are the left singular vectors and right singular vectors of $\widehat{\mathbf{R}}_M$, respectively. This construction ensures that the generalized inverse $\widehat{\mathbf{R}}_M^{\#} = \left(\widehat{\mathbf{R}}_M^H \widehat{\mathbf{R}}_M\right)^{-1} \widehat{\mathbf{R}}_M^H$, exists. The low rank approximation weight vector is then given by:

$$\mathbf{w} = \widehat{\mathbf{R}}_M^{\#} \mathbf{s}. \tag{2.14}$$

The weight vector in equation (2.14) defines the P-SMI weight vector. The matrix $\widehat{\mathbf{R}}_M^{\#}$ is sometimes referred to as the *Moore-Penrose generalized inverse* of the covariance matrix $\widehat{\mathbf{R}}$ [23].

2.2.4 Diagonally Loaded Sample Matrix Inversion

Under similar conditions to these mentioned in the previous section where the sample covariance matrix is singular, Abramovich and Cheremisin suggested diagonally loading the matrix by a scaled identity matrix [19, 20]. This process was shown to improve the conditioning of the noise eigenvalues by reducing the eigenvalue spread. Then following equation (2.12), the weight vector is written as

$$\mathbf{w} = \left(\widehat{\mathbf{R}} + \alpha \mathbf{I}\right)^{-1} \mathbf{s} \tag{2.15}$$

where α is the loading factor.

The loading of the covariance matrix decreases the fluctuations of the small eigenvalues, which are predominantly white noise eigenvalues, and as a result

decreases fluctuations in w. It was suggested in [20] to use a loading factor such that $\sigma_v^2 < \alpha < \lambda_{min}$ where λ_{min} is the minimum interference eigenvalue.

2.2.5 Eigencanceler

The sample covariance can be represented as shown in equation (2.10) where the matrix is decomposed into two orthogonal subspaces $\widehat{\mathbf{Q}}_r$ and $\widehat{\mathbf{Q}}_v$ denoting the interference and noise subspaces, respectively. The eigencanceler selects its weight vector to lie in a space that is orthogonal to the interference subspace and hence, in the noise subspace. It is shown in [11] that the weight vector is written as

$$\begin{aligned} \mathbf{w} &= k\left(\mathbf{I} - \widehat{\mathbf{Q}}_r\widehat{\mathbf{Q}}_r^H\right)\mathbf{s} \\ &= k\widehat{\mathbf{Q}}_v\widehat{\mathbf{Q}}_v^H\mathbf{s} \end{aligned} \tag{2.16}$$

where k is a complex constant and $\mathbf{I}$ is the identity matrix of dimension N.

This choice for the weight vector produces higher values of the SNR as compared to the SMI. Also, to achieve a signal-to-noise ratio within 3 dB of the optimal, only $K \approx 2r$ samples are needed to obtain the sample covariance matrix.

This chapter presented a brief discussion of some adaptive processing techniques. Proceeding chapters will assume the knowledge of these techniques and their corresponding filter weight coefficients.

CHAPTER 3

SPACE-TIME PROCESSING FOR HPRF RADAR

Radar returns consist mainly of interference echoes from unwanted objects that often obscure echoes from targets of interest. In airborne radar, ground clutter returns from all ranges and angles appear to be moving relative to the platform. These mainlobe and sidelobe clutter returns exist in the Doppler region given by $\pm 2Vf_c/c$ Hz, where V is the platform velocity, f_c is the radar's carrier frequency, and c is the speed of propagation. The geometry for an airborne radar system is shown in Figure 3.1. A clutter patch seen by the radar at an azimuth angle ϕ and an elevation (depression) angle θ has a Doppler frequency associated with it which is given by $\nu = (2V/\lambda)\sin\phi\cos\theta$ Hz, where λ is the wavelength corresponding to the transmitted carrier frequency. Radar systems, for missions such as airborne early warning (AEW), may employ high pulse repetition frequency (HPRF) waveforms to enhance long-range detection of high closing-rate targets which appear in the clutter free region of the radar system's Doppler spectrum. However, due to the range-ambiguous nature of the HPRF waveform, strong near-range ground clutter returns received in the antenna sidelobes cannot be simply gated out and are, therefore, folded in with desired signal returns that fall within the Doppler bandwidth of the clutter. This chapter discusses the application of space time adaptive processing (STAP) techniques for clutter suppression in HPRF radar systems.

In recent years, STAP has been studied and applied mainly to low pulse repetition frequency (LPRF) radar [5, 24, 25, 18]. The application of STAP to the HPRF radar problem presents a unique set of challenges and differences from the traditional LPRF radar, which are discussed in this chapter.

Clutter seen by an airborne radar extends in both space and time domains. In order to cancel the clutter, temporal and spatial degrees of freedom are needed. As shown in Figure 3.2 and by observing the clutter Doppler frequency expression above,

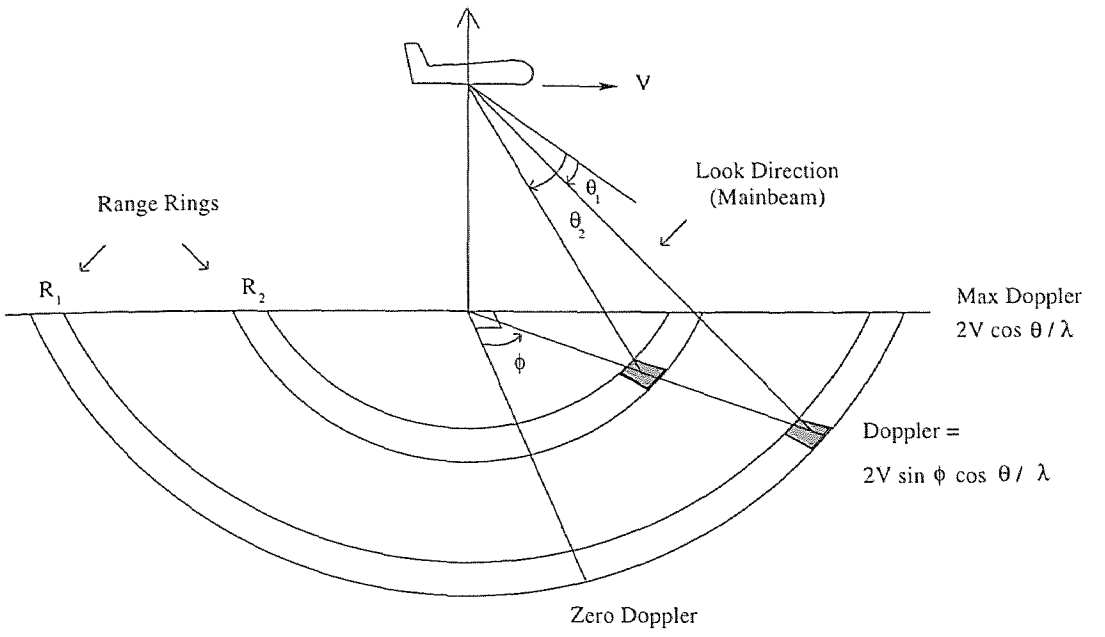

Figure 3.1 Geometry of the airborne radar problem.

it is evident that elevation sidelobe clutter returns are a function of both Doppler and elevation. If only temporal (Doppler) processing is used, and the Doppler 'mainbeam' points at frequency ν_1, then the clutter at elevation angle $\theta_1 = \cos^{-1}(k\nu)$, where k is a constant, will compete with the target (see Figure 3.2). Alternatively, if only elevation processing is employed, and the mainbeam points to elevation angle θ_1, then the clutter at Doppler frequency $\nu_1 = k\theta_1$ will compete with the target. STAP is therefore required to eliminate competing clutter in the mainlobe. Thus performing processing along a single domain will be ineffective in cancelling the clutter. STAP combines both spatial and temporal degrees of freedom needed in this case.

Three STAP approaches will be investigated. The first is the *pseudoinverse sample matrix inversion* (P-SMI) technique. With traditional SMI, the adaptive weight vector is computed by taking the inverse of the sample covariance matrix. In HPRF radar, the sample support is greatly reduced as compared to LPRF due to range ambiguity effects. Hence, the pseudoinverse of the covariance matrix is used.

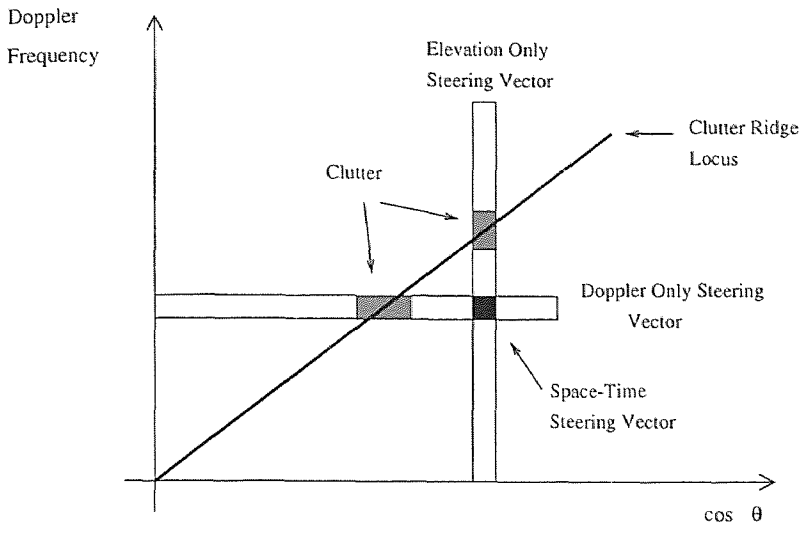

Figure 3.2 STAP for HPRF radar.

The second method investigated is the diagonally loaded SMI [26]. This method is traced back to the early 1980's with publications by Abramovich and Cheremisin [19, 20]. In this technique the singularity of the sample covariance matrix is overcome by diagonal loading. The last STAP technique investigated is the *eigencanceler* [11]. The different STAP techniques are evaluated through a number of performance measures such as output signal-to-colored noise (clutter and noise) ratio and probability of detection. In addition, we provide simulation results for interference cancellation using an element space post-Doppler processor where Doppler filtering is applied prior to the adaptive array to reduce the dimension of the data. The adaptive array's weight vector is then obtained using SMI.

3.1 Advantages of HPRF Operation

The radar's choice of the pulse repetition frequency (PRF) depends on its requirements in terms of range measurement, Doppler measurement, transmit power, and ambiguities. There are three PRF classes based on the sampling criteria for range and Doppler: LPRF, HPRF, and medium PRF. In HPRF systems, the sampling

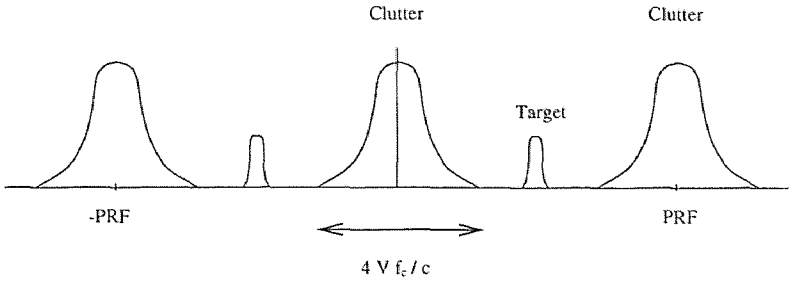

Figure 3.3 Clutter and target Doppler returns.

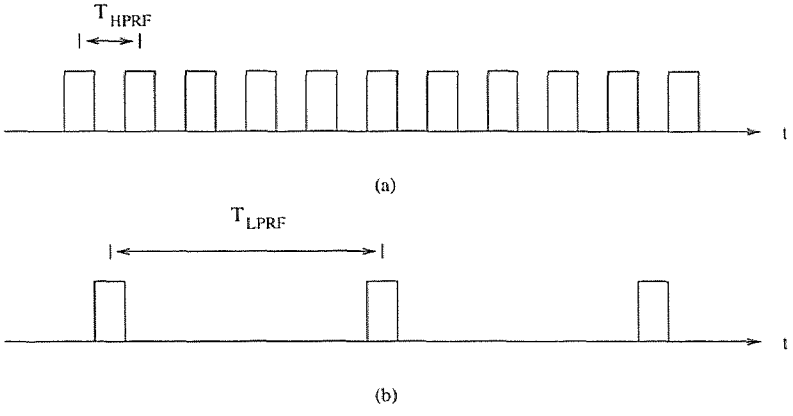

Figure 3.4 (a) HPRF and (b) LPRF waveforms.

rate is fast enough to meet the Nyquist rate for the target Doppler shift. There are advantages for choosing a HPRF waveform. These are summarized below.

1. High closing-rate targets appear in the clutter-free region. Since the PRF is high, high closing-rate targets will have a greater Doppler shift than any of the clutter. This is shown in Figure 3.3

2. High average transmit power. A high duty cycle is associated with a HPRF waveform. This makes HPRF inherently the longest range PRF class for the ability to detect targets at very far ranges. A HPRF waveform is shown in Figure 3.4 relative to a LPRF waveform.

3. Capability to reject high-power near-range clutter while achieving target detection. This is true for both nose and tail aspect targets. Tail aspect targets and near-range clutter exist at different elevation angles with respect to the radar's platform, hence clutter can be rejected while preserving the desired target. This can be seen from Figure 3.5.

The HPRF waveform, due to range ambiguities, obviously has some disadvantages as well. However, depending on the specific application of the radar and by the use of processing techniques, some of the disadvantages could be overcome. This discussion is deferred to the next section.

3.2 Problem Statement

As a result of the airborne radar platform motion, regions of ground clutter can compete in both range and Doppler with targets of interest. As illustrated in Figure 3.6, regions of competing clutter for LPRF airborne radars are at the intersections of the target range ring and iso-Dopplers ambiguous with the target Doppler. Classical STAP techniques have been shown to be effective in suppressing this competing clutter for LPRF radars [5, 18].

Unlike the LPRF problem, for HPRF airborne radars, regions of competing clutter lie along the iso-Doppler contour at points where the ambiguous range is the same as that of a target. This is shown schematically in Figure 3.7. Due to the range-ambiguous nature of the HPRF waveform, each range gate consists of the superposition of the returns from all visible ambiguous ranges. Therefore, regions of strong sidelobe clutter, located at relatively short ranges and steep grazing angles, cannot be gated out and are folded in with mainbeam target returns. Another issue of much importance is sample support. Since the clutter statistics are typically unknown a priori, an estimate has to be obtained from the secondary data. If the sample support is small, there will be a problem of obtaining a good enough estimate

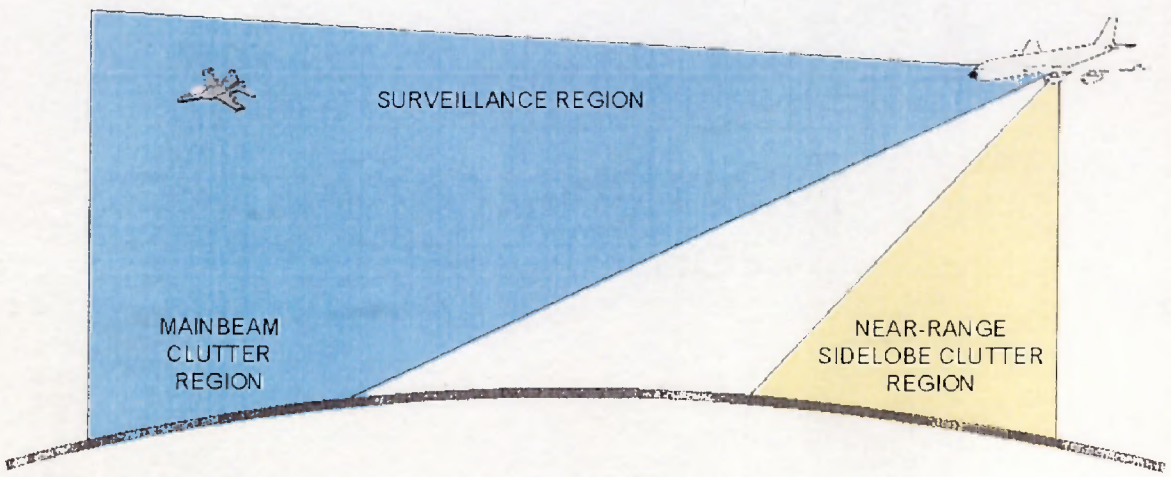

Figure 3.5 Surveillance geometry.

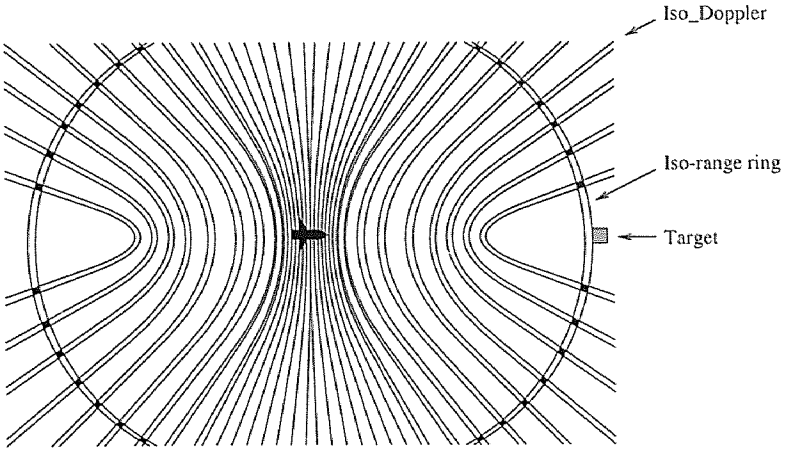

Figure 3.6 Iso-Doppler, Iso-range ring map for LPRF radar.

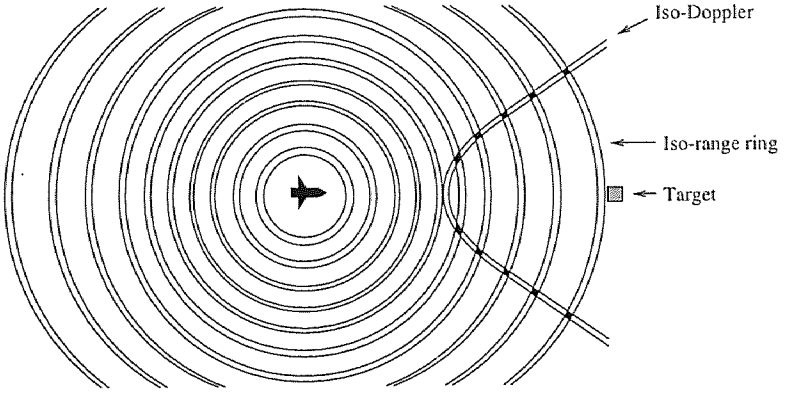

Figure 3.7 Iso-Doppler, Iso-range ring map for HPRF radar.

of the colored noise covariance matrix. The filter coefficients used to cancel the clutter are derived from the estimate of the covariance matrix. Due to the nature of the HPRF waveform, sample support for the estimation of the covariance matrix is limited to the number of range gates available in the radar system. This limited sample support can lead to ill-conditioning of the covariance matrix, and hence, poor performance is offered by classical STAP approaches, such as SMI.

In general, the covariance matrix estimate used by the various STAP techniques is obtained using target-free data. However, due to calibration errors, this is not always the case. If the sample support size is large, the calibration errors might be negligible since the target exists only in few range cells. In HPRF radar, where the sample support is limited, calibration errors can be significant. L-SMI and the eigencanceler are members of the class of *reduced-rank* techniques. Reduced-rank techniques [12, 27, 11] have been shown to outperform full-rank techniques such as SMI for limited sample support and have also been shown to be more robust to calibration errors [11].

The application of innovative reduced rank STAP solutions, shown to outperform conventional STAP techniques [12, 11, 15], is proposed to address the two problems of range ambiguity and limited sample support associated with HPRF radar systems.

3.3 Processing Strategy

The HPRF waveform is range ambiguous by nature. The reason for range ambiguity is explained by Figure 3.8. Unlike the LPRF waveform, with the HPRF waveform, it is not readily possible to identify which pulse caused a particular echo. This implies that discrimination between different returns in range is not possible. The far-range target and the near-range clutter however, can be discriminated in the elevation plane. Although the target and the near-range clutter appear to have the same range, they exist at different elevation angles. Hence, the adaptive array's

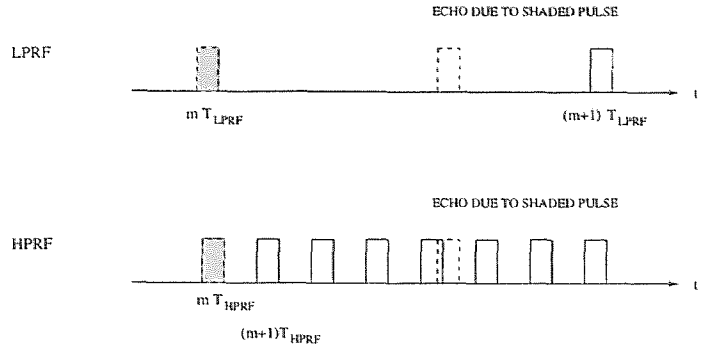

Figure 3.8 Range returns.

weight vector is chosen to cancel the near-range sidelobe clutter region shown in Figure 3.5. This is done while also protecting the target in the mainbeam clutter region (surveillance region). The mainbeam clutter at the same elevation angle as the target will have a different Doppler shift, hence we expect to isolate the target from near-range and mainbeam clutter once processing is applied.

3.4 HPRF System's Definitions and Properties

This section presents definitions and requirements relevant for the HPRF radar system discussed in this chapter. Also an explanation of the spectrum of the clutter seen by the airborne radar for such a system is provided.

3.4.1 System's Definitions

We use the same signal model presented in Chapter 2, however, we define a few additional parameters that are necessary for the discussion of HPRF radar.

The radar's unambiguous range is related to the PRF:

$$R_{\text{un}} = c/(2\text{PRF}). \tag{3.1}$$

If the target's range extends beyond R_{un}, the radar cannot measure the true range of the target. Consequently, target returns may be folded over close range clutter echoes.

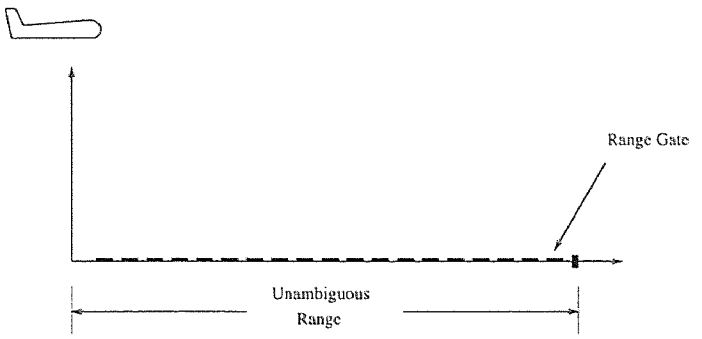

Figure 3.9 Available range gates.

The unambiguous range is divided into a number of range gates depending on the bandwidth of the radar system. The range resolution is given by

$$\delta_R = c/(2\,B_W) \tag{3.2}$$

where B_W is the radar system bandwidth. Hence using equations (3.1) and (3.2), the number of available range gates or equivalently, the number of independent sample returns is given by

$$K = R_{\text{un}}/\delta \tag{3.3}$$

Equation (3.3) can be explained by observing Figure 3.9. As shown in the figure, the unambiguous range is divided into a finite number of range gates governed by the resolution requirement which is in turn a function of the system's bandwidth. Hence only K secondary data samples are independent and can be used in estimating the covariance matrix.

As a consequence of equation (3.1), radar echoes will have an apparent range R_{app}, and a true range R_{true}, which are be related by

$$R_{\text{app}} = R_{\text{true}} - R_{\text{un}}[R_{\text{true}}/R_{\text{un}}], \tag{3.4}$$

where the brackets denote integer part. The data vector $\mathbf{x}(k)$ consists of the sum of contributions of all the range cells folded onto the cell corresponding to index k.

Also, the number of independent data samples equals the number of range cells in an unambiguous range interval. Let K be the number of range gates spanning the unambiguous range interval of the radar. Then the maximum likelihood estimate of the space-time covariance matrix is given again by

$$\widehat{\mathbf{R}} = \sum_{k=1}^{K} \mathbf{x}(k)\mathbf{x}(k)^H. \tag{3.5}$$

The matrix $\widehat{\mathbf{R}}$ is clearly singular if $K < N$. The parameter K is controlled by the radar's range resolution which is a direct function of the system's bandwidth.

3.4.2 Degrees of Freedom

The number of the degrees of freedom that is needed for effective clutter cancellation is determined by the rank of the clutter covariance matrix. In a typical airborne scenario and a calibrated radar, eigen-decomposition, when applied to the covariance matrix, yields a few large eigenvalues while the rest are relatively small. The number of significant eigenvalues or equivalently the rank, can be predicted by the Landau-Pollak relation $r \cong 2BT + 1$, where B is the clutter bandwidth and T is the time across the filter structure [11].

The rank of the space-time covariance matrix has been studied in [11, 18]. Here we provide a brief argument for illustrative purposes. For an array with elements at half-wavelength intervals, and a point clutter source at azimuth angle ϕ with respect to the array normal and at elevation angle θ, a space-time sample is given by,

$$x_{nk} = e^{j\pi(n-1)\sin\phi\cos\theta}\, e^{j2\pi(k-1)\nu}, \tag{3.6}$$

where ν is the point source Doppler frequency normalized with respect to the PRF. The maximum frequency space-time component is then

$$x_{nk} = e^{j\pi[(N_s-1)+2(N_t-1)\nu_{\max}]}, \tag{3.7}$$

where $\nu_{\max} = 2V/(\lambda\ \mathrm{PRF})$ is the highest normalized Doppler component of the clutter returns. It follows that the number of space-time samples required to

represent the clutter contributions is upper bounded

$$r \leq \quad N_s \; + \; 2(N_t - 1)\nu_{\max}. \tag{3.8}$$

This is also the highest approximate rank of the clutter and noise covariance matrix for high clutter-to-noise ratio. In the case of a HPRF radar, $\nu_{\max} \ll 1$ since the clutter occupies only a fraction of the Doppler spectrum. Thus, the HPRF problem is of lower rank than an equivalent LPRF problem.

3.4.3 'J-Hook' Clutter

In an airborne HPRF application, clutter enters the receiver primarily through the mainbeam and principal elevation sidelobes. This is illustrated in the clutter intensity plot, shown in Figure 3.10. It can be observed that at far ranges clutter returns are approximately parallel to iso-Doppler contours. This implies little variation in the Doppler frequency as a function of range. The clutter ridge crosses increasingly more iso-Doppler contours as it gets closer to Nadir. This results in the characteristic 'J-Hook' curvature of the clutter ridge in the range-Doppler domain. The 'J-Hook' is clearly visible on the range-Doppler plot shown later in Figure 3.16 in the Numerical Analysis section. As this plot illustrates, for the HPRF waveform, most of the Doppler band is clutter free with mainlobe and sidelobe ground clutter returns located only in the Doppler region given by $f_d = \pm 2V/\lambda$.

The 'J-Hook' formation could also be shown mathematically by analyzing the elevation angle-Doppler frequency relation. The clutter's Doppler frequency was given previously in section 2.1 as

$$\nu = (2V/\lambda \mathrm{PRF}) \; \sin\phi \; \cos\theta. \tag{3.9}$$

Then the elevation angle θ can be expressed in terms of the Doppler frequency as

$$\theta = \cos^{-1}\left(\frac{\lambda \; \mathrm{PRF}}{2V} \frac{\nu}{\sin\phi}\right), \tag{3.10}$$

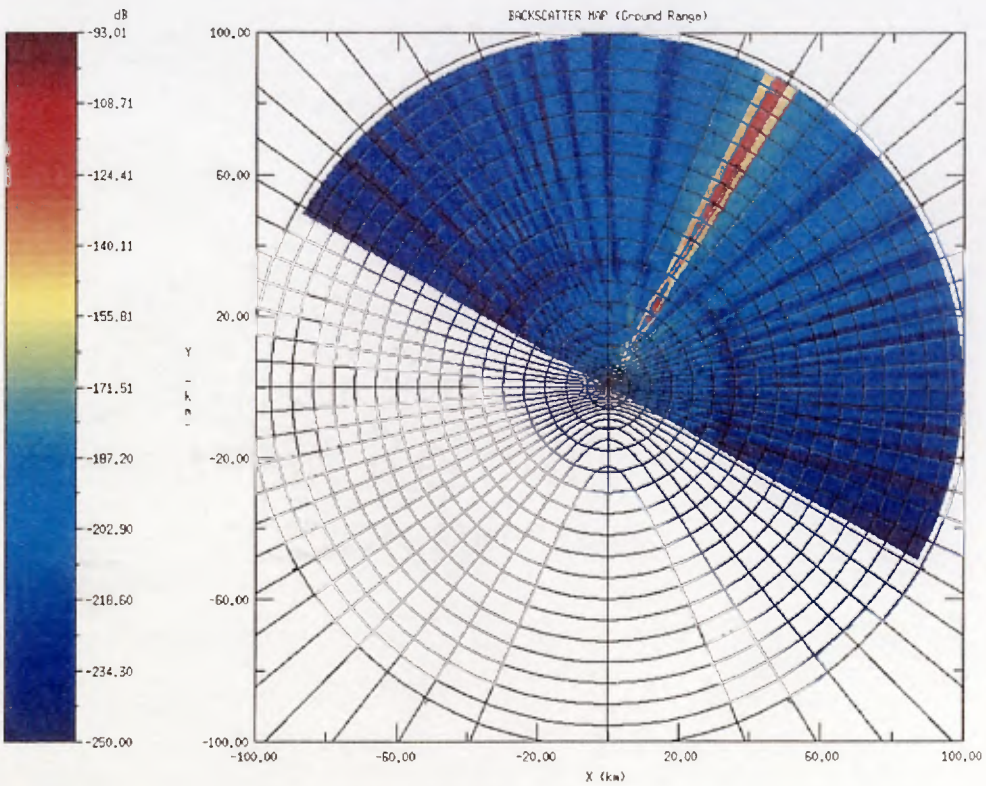

Figure 3.10 HPRF Clutter Intensity map.

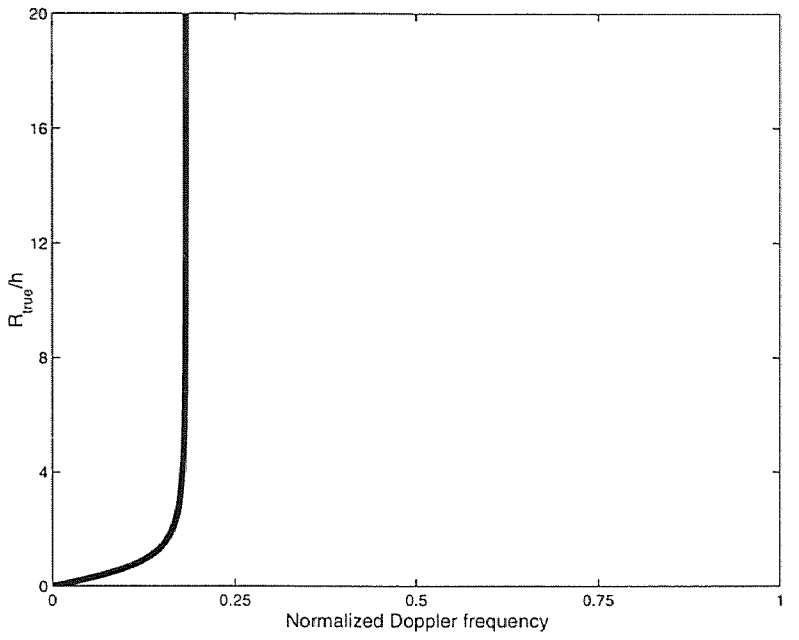

Figure 3.11 'J-Hook' Clutter.

or equivalently we can relate the range to the Doppler frequency where

$$R_{true} = h/\tan\theta, \tag{3.11}$$

where R_{true} is the true range and h is the altitude of the radar platform. By plotting this last equation for a fixed azimuth angle, the 'J-Hook' shape becomes apparent. This is shown in Figure 3.11.

3.5 Performance Measures

The P-SMI, L-SMI, eigencanceler and the element-space post-Doppler will be compared under two performance measures: signal-to-colored noise ratio and probability of detection.

3.5.1 Conditioned Signal-to-Noise Ratio (CSNR)

SNR is a very common performance measure in signal processing applications. It shows the effectiveness of the radar in cancelling the interference. A related

performance measure is the CSNR. The CSNR is defined as the ratio of the SNR obtained using the weight vector derived from a specified method to the optimal SNR when the covariance matrix is known,

$$\text{CSNR} = \rho = \frac{\text{SNR}}{\text{SNR}_o} \tag{3.12}$$

where the SNR is given by

$$\begin{aligned} \text{SNR} &= \frac{\left|\mathbf{w}^H\, a\mathbf{s}\right|^2}{\mathbf{w}^H\mathbf{R}\mathbf{w}} \\ &= \frac{\sigma_t^2\,\left|\mathbf{w}^H\,\mathbf{s}\right|^2}{\mathbf{w}^H\mathbf{R}\mathbf{w}} \end{aligned} \tag{3.13}$$

and the optimal SNR (SNR_o) can be shown to be

$$\text{SNR}_o = \sigma_t^2\,\mathbf{s}^H\mathbf{R}^{-1}\mathbf{s} \tag{3.14}$$

where $\mathbf{R}$ is the true covariance matrix of the colored noise. Then the CSNR in equation (3.12) can be written as

$$\rho = \frac{\left|\mathbf{w}^H\mathbf{s}\right|^2}{\mathbf{w}^H\mathbf{R}\mathbf{w}\;\mathbf{s}^H\mathbf{R}^{-1}\mathbf{s}}. \tag{3.15}$$

The CSNR in equation (3.15) is a random variable bounded by $0 \leq \rho \leq 1$. A probability density function of this random variable ρ, can be obtained and is given in [6, 20, 11] for SMI, L-SMI, and Eigencanceler respectively. For reasons of comparison between theoretical and simulation results, the CSNR distributions for L-SMI and the eigencanceler are provided next.

In [20], the distribution of the CSNR, ρ, for L-SMI is given as

$$f(\rho) = \frac{\Gamma(K+1)}{\Gamma(K-r+1)\Gamma(r)}\rho^{K-r}(1-\rho)^{r-1} \tag{3.16}$$

where $\Gamma(k) = (k-1)!$ is the Gamma function, K is the sample support dimension, and r is the interference subspace rank. In [20, 28], it is suggested to choose a loading factor α such that $\sigma_v^2 < \alpha < \lambda_{\min}$, where ${\sigma_v}^2$ is the noise power, and $\lambda_{\min}$ denotes the smallest interference eigenvalue.

In [29], the asymptotic distribution for the CSNR, ρ, was derived for the eigencanceler. It is shown that the CSNR can be expressed as

$$\rho = \frac{1}{1 + \frac{1}{K}\zeta} \tag{3.17}$$

where $\zeta = \sum_{i=1}^{r} \nu_i$, and ν_i are independent identically distributed random variables with exponential distribution and hence, ζ is a Gamma random variable with r degrees of freedom and parameter 1. This characterization results in the density

$$f(\rho) = K\rho^{-2} \sum_{i=1}^{r} \frac{\pi_i}{\overline{\nu}_i} \exp\left(\frac{-K\left(\frac{1}{\rho} - 1\right)}{\overline{\nu}_i}\right) \tag{3.18}$$

where $\pi_i = \prod_{j=1,,j\neq i}^{r} \overline{\nu}_i / (\overline{\nu}_j - \overline{\nu}_i)$. Also, [29] gives an approximation to the distribution in equation (3.18) for the case of a large interference-to-white noise ratio. This is given by

$$f(\rho) = \frac{K^r}{\Gamma(r)} e^{-K\rho} \rho^{r-1}, \qquad 0 \leq \rho \leq 1. \tag{3.19}$$

3.5.2 Probability of Detection

Probability of detection is the main measure for the radar system's performance. The processor decides on one of the two hypotheses, $\mathbf{H}_0$ or $\mathbf{H}_1$, by taking the instantaneous output power $\mathbf{w}^H\mathbf{x}(k)$, where $\mathbf{w}$ is the appropriate weight vector as

$$\gamma = \left|\mathbf{w}^H\mathbf{x}(k)\right|^2 \underset{\mathbf{H}_0}{\overset{\mathbf{H}_1}{\gtrless}} \eta_T \tag{3.20}$$

where η_T is a given threshold. Since the received vector $\mathbf{x}(k)$ has a Gaussian distribution with circular symmetry under both hypotheses, the statistic γ has a chi-squared distribution given as

$$f_\gamma(\gamma \mid \mathbf{H}_i) = \frac{1}{\gamma_i} e^{-\gamma/\gamma_i} \tag{3.21}$$

where i=0,1, and γ_0 and γ_1 are the average output powers under hypotheses $\mathbf{H}_0$ and $\mathbf{H}_1$, respectively. For a given threshold η_T in equation (3.20), the probability of false

alarm is given as

$$\begin{aligned} P_{fa} &= \int_{\eta_T}^{\infty} f_\gamma(\gamma \mid \mathbf{H}_0)\, d\gamma \\ &= e^{-\eta_T/\eta_0}, \end{aligned} \tag{3.22}$$

and the probability of detection as

$$\begin{aligned} P_D &= \int_{\eta_T}^{\infty} f_\gamma(\gamma \mid \mathbf{H}_1)\, d\gamma \\ &= e^{-\eta_T/\eta_1}. \end{aligned} \tag{3.23}$$

The average power under hypothesis $\mathbf{H}_0$ can be written as

$$\begin{aligned} \eta_0 &= E\{\eta \mid \mathbf{H}_0\} \\ &= \mathbf{w}^H \mathbf{R} \mathbf{w}. \end{aligned} \tag{3.24}$$

When the covariance matrix is known, the optimal weight vector is given by $\mathbf{w} = k\mathbf{R}^{-1}\mathbf{s}$. Define α as the optimum CSNR given in equation (3.15) which is achieved when the covariance matrix is known as

$$\begin{aligned} \alpha &= \frac{\left|\mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}\right|^2}{\mathbf{s}^H \mathbf{R}^{-1} \mathbf{R} \mathbf{R}^{-1} \mathbf{s}\; \mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}} \\ &= \mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}. \end{aligned} \tag{3.25}$$

Then, the average power η_0 under hypothesis $\mathbf{H}_0$ can be expressed in terms of α in equation (3.25) and ρ in equation (3.15) as

$$\eta_{0/\rho} = \frac{1}{\rho\alpha}. \tag{3.26}$$

The average power under hypothesis $\mathbf{H}_1$ under the a unity signal gain constraint ($\mathbf{w}^H \mathbf{s} = 1$) can be written as

$$\begin{aligned} \eta_1 &= E\{\eta \mid \mathbf{H}_1\} \\ &= \left|\mathbf{w}^H\, a\mathbf{s}\right|^2 + \mathbf{w}^H \mathbf{R} \mathbf{w} \\ &= \sigma_t^2 + \mathbf{w}^H \mathbf{R} \mathbf{w}. \end{aligned} \tag{3.27}$$

Using α in equation (3.25) and ρ in equation (3.15), the average power under hypothesis $\mathbf{H}_1$ can be written as

$$\eta_{1/\rho} = \sigma_t^2 + \frac{1}{\rho\alpha}. \tag{3.28}$$

The probability of false alarm for the test given in equation (3.20) can be written as

$$\begin{aligned} P_{fa/\rho} &= e^{-\eta_T/\eta_0} \\ &= e^{-\rho\alpha\eta_T}. \end{aligned} \tag{3.29}$$

Similarly, the probability of detection can be written as

$$\begin{aligned} P_{D/\rho} &= e^{-\eta_T/\eta_1} \\ &= e^{-\eta_T/(\sigma_t^2+\frac{1}{\rho\alpha})}. \end{aligned} \tag{3.30}$$

The average probabilities of false alarm and detection are obtained by averaging the conditional probabilities of false alarm and detection over all possible values of ρ. Hence, the unconditional probability expressions are given as

$$P_{fa} = \int_0^1 P_{fa/\rho} f(\rho) d\rho \tag{3.31}$$

and

$$P_D = \int_0^1 P_{D/\rho} f(\rho) d\rho. \tag{3.32}$$

The integral expressions in equations (3.31) and (3.32) are numerically evaluated for different values of the SNR.

3.6 Element-Space Post-Doppler

Element-space post-Doppler processor performs Doppler processing on the data from each array element prior to the adaptive array. Hence, the space-time data snapshot is transformed into a snapshot of array's elements and Doppler bins. This process is

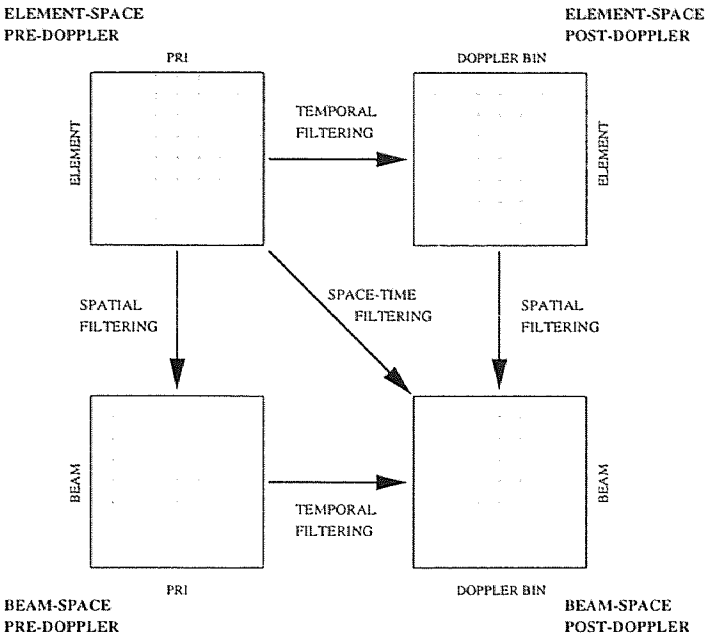

Figure 3.12 Possible reduced rank STAP algorithms.

shown in Figure 3.12 which is borrowed from [18]. The element-space post-Doppler is one class of possible algorithms. Once this algorithm is applied, a different adaptive problem is to be solved for each Doppler bin. For example, after Doppler filtering is applied to the space-time data, SMI can be used to calculate the adaptive weight vector as

$$\mathbf{w} = k\widehat{\mathbf{R}}_s^{-1}\mathbf{s}_s \tag{3.33}$$

where $\widehat{\mathbf{R}}_s$ is the estimated spatial covariance matrix and $\mathbf{s}_s$ is the spatial steering vector for a Doppler bin.

The advantages of such an algorithm are summarized in the following:

1. Reduced data dimension which is necessary in an environment of reduced sample support.

2. Faster conversion rate since the size of the matrix to be inverted is much smaller.

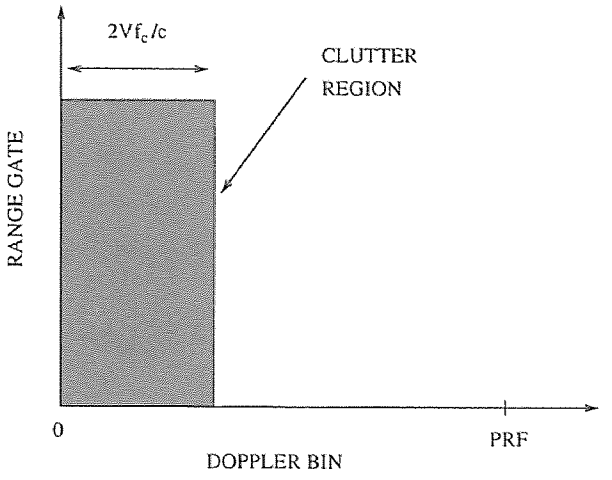

Figure 3.13 Clutter region for HPRF radar.

3. This algorithm is suitable to apply to HPRF radar where the clutter's Doppler shifts do not extend over the whole radar's bandwidth. Hence, processing and detection is applied only in a few, usually, about 10%-20% of the total number of Doppler bins depending on the radar's parameters. This is shown in Figure 3.13.

3.7 Robustness Analysis

As was mentioned in section 3.2, the sample support for HPRF radar is limited. The covariance matrix estimate given in (2.8) is for the interference and noise and it is obtained using data from range cells away from the cell under test and its immediate neighborhood. If the target is present during training, then mismatches between the target estimate (as incorporated in $\widehat{\mathbf{R}}$) and the steering vector will result in signal cancellation. The cancellation becomes more pronounced with the increase in the target power. This effect may be negligible if the estimate of the covariance matrix is done by averaging a very large number of range cells. In this case, the impact of the few range cells containing the target return is small. However, with HPRF radar, range cells containing the target form a larger share of the overall number of

snapshots and the effect is not negligible anymore. The estimate of the covariance matrix could then be written as

$$\widehat{\mathbf{R}}_t = \sum_{k=1}^{K} \mathrm{x}(k)\mathrm{x}(k)^H + \sum_{k=1}^{L} |a_k|^2 \mathrm{s}\mathrm{s}^H \tag{3.34}$$

where L is the number of range cells for which there is a target return and $|a_k|^2$ is the target power in each range cell. From (3.34) and (2.8), it can be seen that if $K \gg L$, $\widehat{\mathbf{R}}_t \simeq \widehat{\mathbf{R}}$.

The effect of having the desired signal present in the training region can be investigated theoretically. The expression for the CSNR in the presence of the target can be written as

$$\rho_t = \frac{|\mathrm{w}_t{}^H \mathrm{s}|^2}{\mathrm{w}_t{}^H \mathbf{R} \mathrm{w}_t \mathrm{s}^H \mathbf{R}^{-1} \mathrm{s}}, \tag{3.35}$$

where w_t is the weight vector obtained using $\widehat{\mathbf{R}}_t$ rather than $\widehat{\mathbf{R}}$.

It is shown in [30] that the following relation exists between ρ and ρ_t:

$$\rho_t = \frac{\rho}{\mathrm{SNR}_o(1-\rho)+1}, \tag{3.36}$$

where ρ is given in equation (3.15) and SNR_o is given as

$$\mathrm{SNR}_o = \mathrm{s}^H \mathbf{R}^{-1} \mathrm{s}. \tag{3.37}$$

The CSNR distributions for the eigencanceler and the L-SMI techniques are very similar with respect to the mean CSNR. The probability density function (PDF) of the CSNR for L-SMI was previously given in equation (3.16). If the target was included in the training set during training, the transformed PDF can be obtained by using equation (3.36) and by applying the Jacobian transformation. Then the mean CSNR can be found as

$$\begin{aligned} E[\rho_t] &= \int_0^1 \rho_t f(\rho_t) d\rho_t \\ &= \left(\frac{1+K-r}{1+K}\right) (1+\mathrm{SNR}_o)^{1+K-r} \, {}_2F_1[1+K; 2+K-r; 2+K; -\mathrm{SNR}_o], \\ & \qquad\qquad K > r-2 \end{aligned} \tag{3.38}$$

where ${}_2F_1$ is a hypergeometric function [31] that is defined as

$$2F1[a;b;c;x] = \sum_{j=0}^{\infty} \frac{\Gamma(a+j)\Gamma(b+j)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(c)}{\Gamma(c+j)} \frac{x^j}{j!}. \tag{3.39}$$

The CSNR mean, $E[\rho_l]$, is a function of the desired signal power. As the desired signal power increases, the mean will decrease and will eventually reach zero.

3.8 Numerical Analysis

The simulation model assumed a linear uniform antenna array with $N_s = 8$ elements spaced at half wavelength. Each array channel consisted of $N_t = 64$ tap FIR filter. The system under consideration is mounted on an airborne platform. The platform was assumed to be at an altitude of 30,000 ft and moving with a constant velocity of 250 m/s. The clutter was assumed to come from all elevation angles and was modeled to have a complex-valued Gaussian distribution, with zero mean, and variance equal to the clutter-to-noise ratio (CNR). We assumed the clutter returns are uncorrelated with each other and also uncorrelated between snapshots. The CNR was calculated from the total contributions of all clutter sources and was set to 60 dB. Attenuation due to free-space propagation was assumed proportional to R_{true}^{-3}, where R_{true} is true range of the cell under test. The radar PRF was chosen to be 25 KHz. The radar's transmitted frequency was 3.3 GHz. Using equation (3.1), the unambiguous range was calculated to be 6 km. The target signal-to-noise ratio was 48 dB and the target was chosen to be at an ambiguous range of 90 km with a Doppler frequency of 0.05×PRF.

The sample covariance matrix was obtained from equation (3.5) using K=112 independent identically distributed snapshots. Equations (2.14), (2.15), (2.16), and (3.33) are used to calculate the adaptive filter weight vectors for P-SMI, L-SMI, eigencanceler, and element-space post-Doppler, respectively. Equation (3.12) is then used to calculate the CSNR values for all methods by averaging 10,000 Monte-Carlo

runs. The resulting simulated probability density functions for all four methods are shown in Figure 3.14. The figure shows that the L-SMI and eigencanceler techniques produce high values of the CSNR. Equation (3.30) is used to calculate the probability of detection (PD) as a function of detection for all the methods. The probability of false alarm was assumed to be 10^{-5}. The PD curves were obtained by averaging 200 Monte-Carlo runs and are shown in Figure 3.15. The PD curves provide similar results to the CSNR distributions in terms of performance.

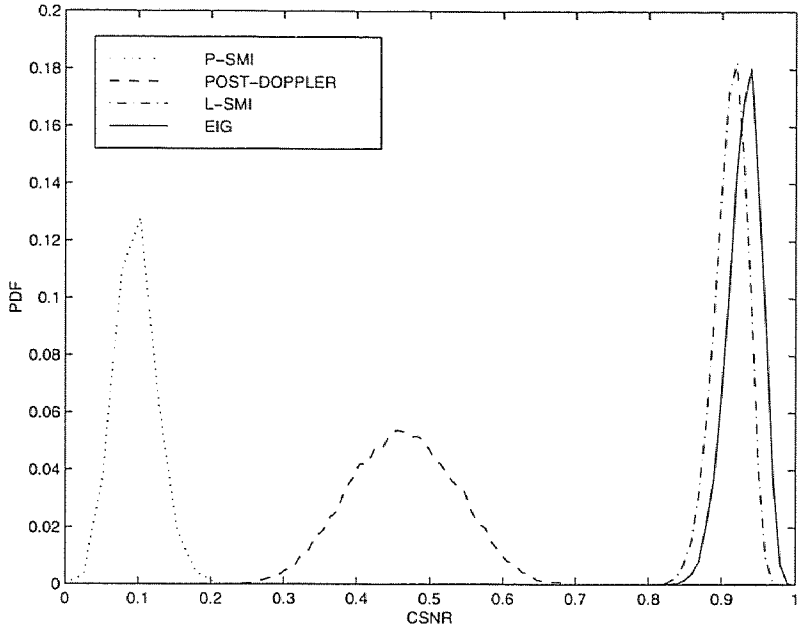

Figure 3.14 PDF of the CSNR.

The range ambiguous clutter map is shown in Figure 3.16. The clutter has Doppler returns in approximately 12 out the 64 Doppler bins used. The 'J-Hook' shape of the clutter ridge is apparent in the figure. The post processing clutter maps for L-SMI and the eigencanceler are provided in Figures 3.17 and 3.18 respectively. The two figures show that the near range clutter masking the target has been rejected and the target is now evident.

Finally, the issue of robustness of L-SMI and the eigencanceler is investigated. A target with a SNR of 10 dB was included in the training set used to obtain the

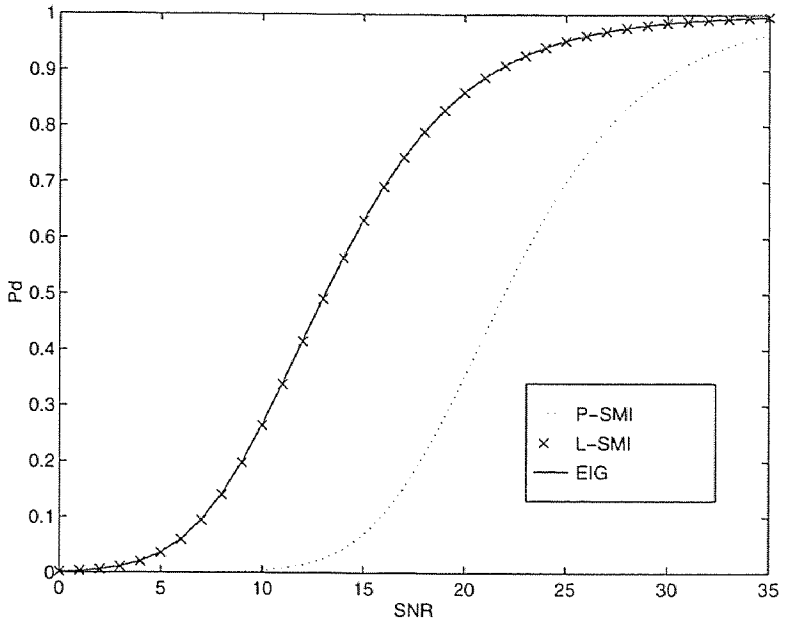

Figure 3.15 Probability of detection of the CSNR obtained by averaging 200 runs.

sample covariance matrix as in equation (3.34). The probability density functions of the CSNR for both methods were obtained by averaging 10,000 Monte-Carlo runs. The simulated CSNR distribution curves with and without the target included are given in Figure 3.19. The theoretical CSNR distributions are provided for comparison reasons in Figure 3.20. The simulation and theoretical results of the CSNR distributions are shown to be very similar.

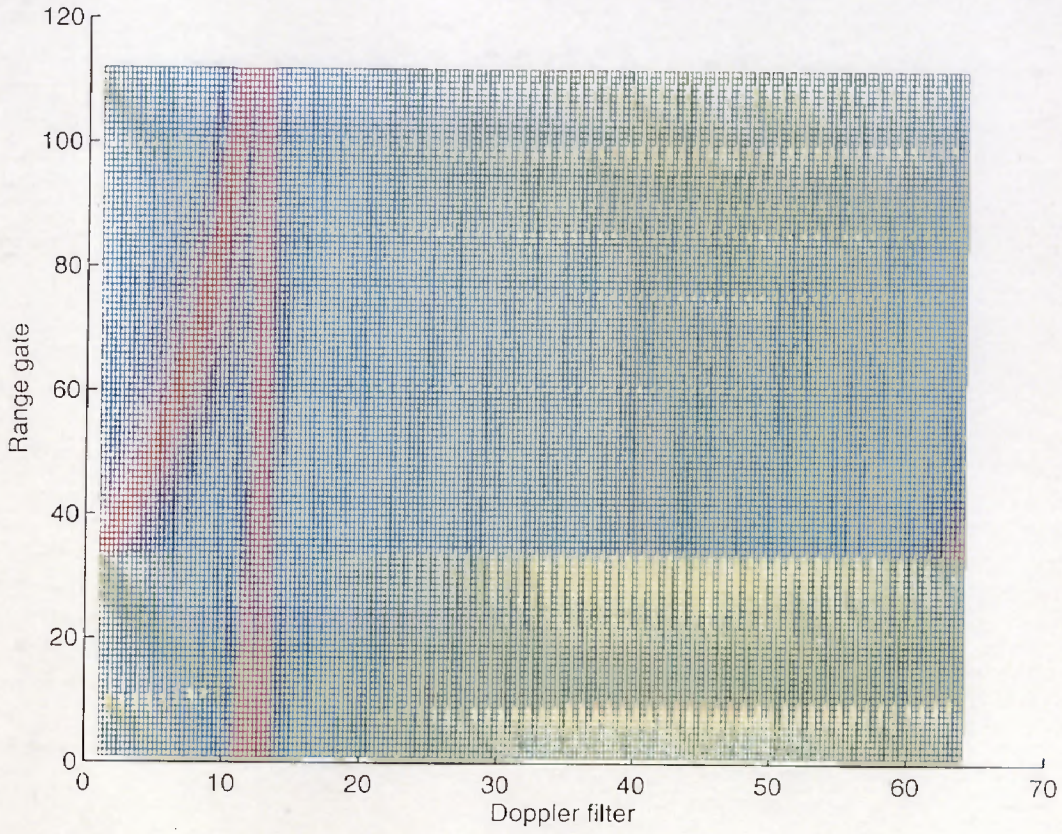

Figure 3.16 Clutter map for HPRF radar.

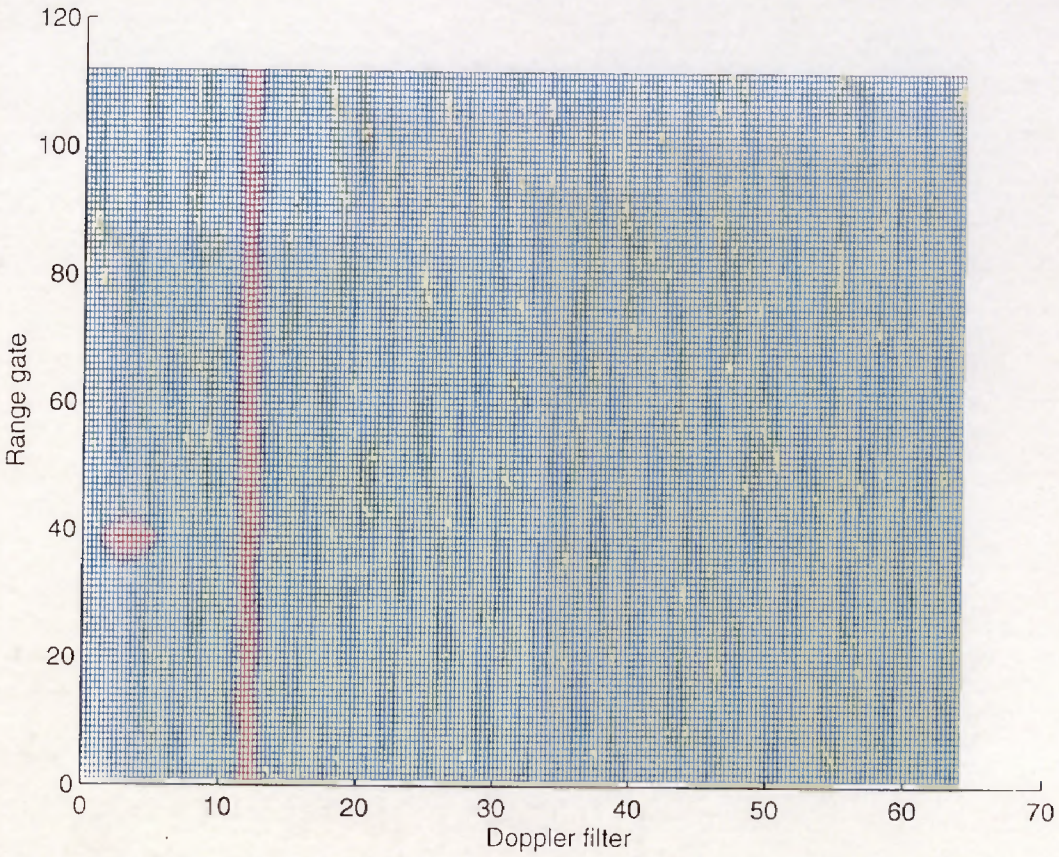

Figure 3.17 Post-processing clutter using the Loaded SMI.

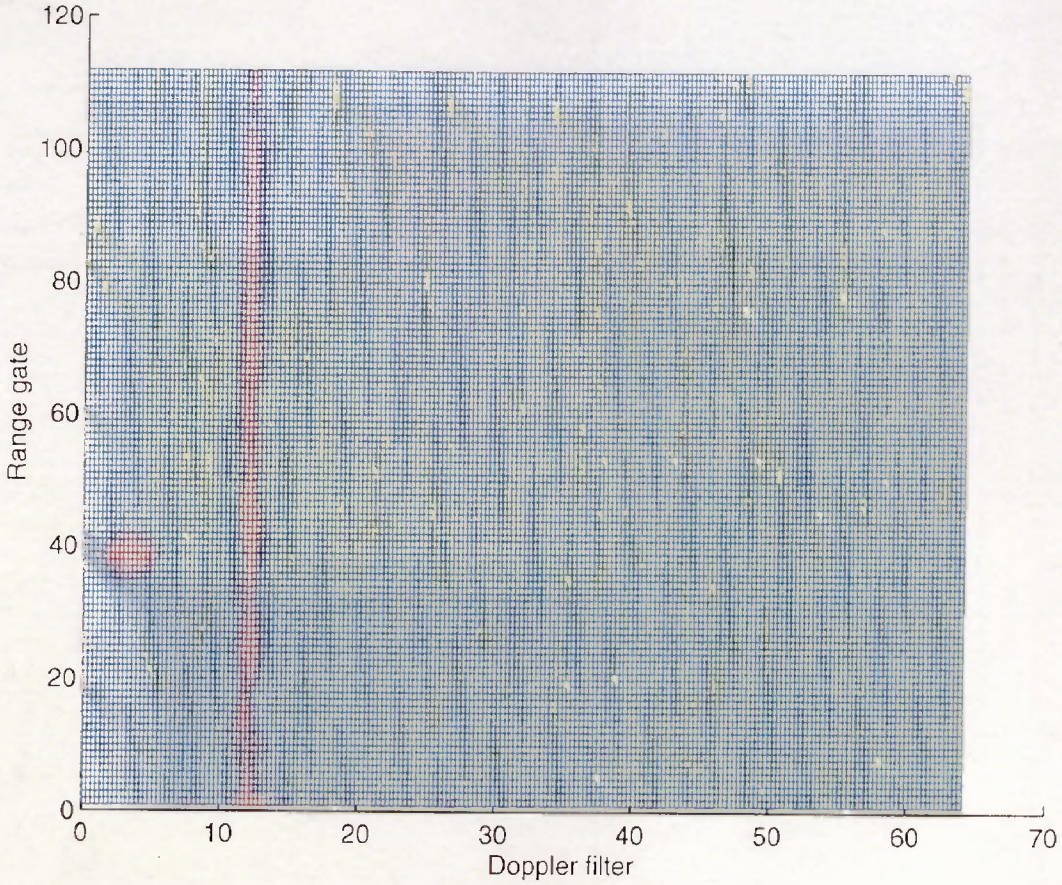

Figure 3.18 Post-processing clutter using the eigencanceler.

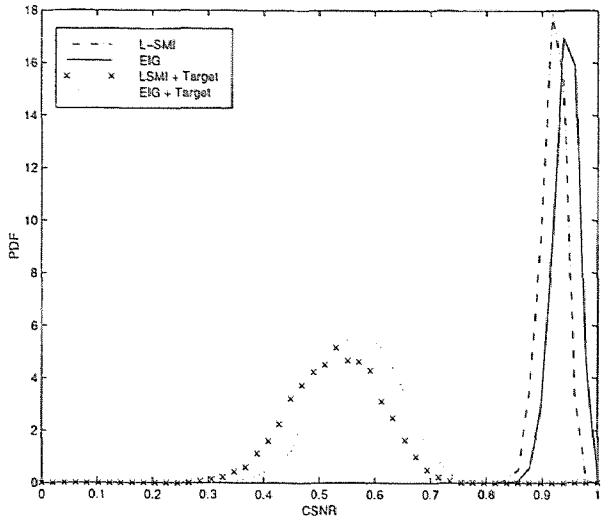

Figure 3.19 Simulation PDF of the CSNR with and without the desired signal present.

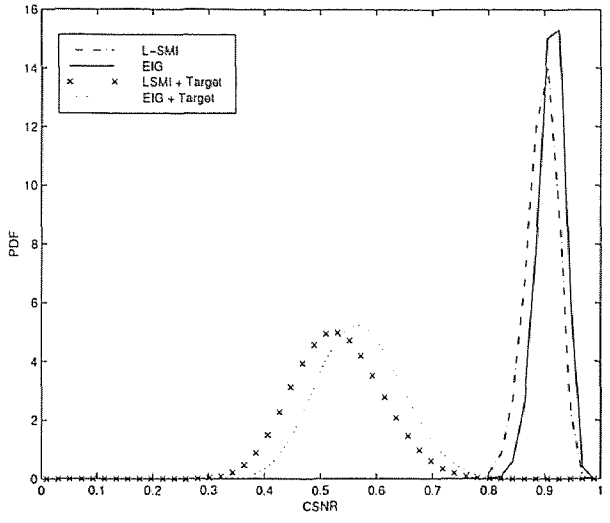

Figure 3.20 Theoretical PDF of the CSNR with and without the desired signal present.

3.9 Discussion

This chapter discussed the application of STAP techniques to suppress the interference effects of ground clutter in HPRF radars. It was shown that reduced-rank methods are necessary for clutter cancellation due to reduced sample support as a

result of range ambiguity. Three STAP methods were compared in this chapter: P-SMI, L-SMI, and the eigencanceler. The three techniques were evaluated for performance measures CSNR and probability of detection. It was shown that the eigencanceler slightly outperforms L-SMI. P-SMI does not perform satisfactorily. The element-space post-Doppler method resulted in a mediocre output CSNR as compared to the L-SMI and the eigencanceler. This is due to the fact that only spatial degrees of freedom are not enough to cancel the clutter. Although the size of the secondary data used to estimate the spatial covariance matrix in each Doppler bin exceeds the spatial dimension of the system, the performance still suffers.

CHAPTER 4

DISTRIBUTION OF THE CONDITIONAL SIGNAL-TO-NOISE RATIO FOR L-SMI

The loss incurred from estimating the covariance matrix of the interference sources can be measured from the conditioned signal-to-colored noise ratio (CSNR). The CSNR was previously defined in section 3.5.1 as the ratio of the actual SNR at the output of the adaptive array to the optimal SNR. The CSNR is a random variable with certain statistical properties. These statistical properties depend on the adaptive algorithm used by the adaptive array to obtain its weight vector. Reed et al. obtained an expression for the distribution of the CSNR when the sample matrix inversion (SMI) algorithm is used to obtain the array's weight vector [6]. Haimovich derived an asymptotic distribution of the CSNR for the eigencanceler algorithm in [11]. Kirsteins and Tufts obtained a distribution of the CSNR for the principal component inverse (PCI) algorithm in [12]. Cheremisin derived a distribution of the CSNR for the diagonally loaded SMI (L-SMI) algorithm in [20] where the sample covariance matrix is singular and diagonal loading is applied to improve the conditioning of the eigenvalues. The CSNR distribution gives a theoretical performance measure and a means of comparison between different interference cancellation methods in terms of the output SNR.

Cheremisin's distribution of the CSNR was derived for the case when the sample covariance matrix estimate is singular due to limited sample support. If the traditional SMI algorithm is to be applied to find the weight vector for an array, $K \approx 2N$ samples are needed to estimate the covariance matrix for a mean CSNR of 0.5 [32] where N is system's dimension. Hence, even if the sample covariance matrix is non-singular ($K \geq N$) but K is not much larger than N, the system's performance will suffer. Diagonal loading of the sample covariance matrix can improve the conditioning of the eigenvalues. In this chapter, the discussion is limited to the L-SMI

adaptive algorithm. The L-SMI technique is investigated and analyzed in terms of the benefits it adds to the system's performance. A probability distribution expression is obtained of the CSNR for the L-SMI algorithm when the covariance matrix is non-singular ($K \geq N$) but ill-conditioned.

4.1 Background

The sample covariance matrix is obtained from a set of secondary data vectors obtained from neighboring range cells to the one under test. The sample covariance matrix is the maximum likelihood estimate of the true covariance matrix and was given in equation (2.8) and is repeated here for convenience as

$$\widehat{\mathbf{R}} = \frac{1}{K}\sum_{k=1}^{K} \mathbf{x}(k)\mathbf{x}(k)^H \tag{4.1}$$

where $\mathbf{x}(k)$ is a space-time snapshot of dimension $(N = N_s N_t) \times 1$ and K is the number of secondary data vectors. It is emphasized that the true covariance matrix $\mathbf{R}$ consists of two subspaces; an interference subspace and a noise subspace. Matrix $\mathbf{R}$ obeys the model given by

$$\mathbf{R} = \mathbf{Q}_r \mathbf{\Lambda}_r \mathbf{Q}_r^H + \sigma_v^2 \mathbf{Q}_v \mathbf{Q}_v^H \tag{4.2}$$

where $\mathbf{Q}_r$ is the $N \times r$ matrix of principal eigenvectors, $\mathbf{\Lambda}_r$ is the $r \times r$ diagonal matrix of principal eigenvalues, σ_v^2 is the variance of the white noise, and $\mathbf{Q}_v$ is the $N \times p$ matrix of noise eigenvectors, where $p = N - r$. Hence, when the covariance matrix is said to be of low rank, it is implied that $r < N$.

As was briefly discussed in Chapter 2, the adaptive array that produces the highest SNR is based on the Wiener filter. Since the true covariance matrix is usually unknown *a priori*, the sample covariance matrix, $\widehat{\mathbf{R}}$, given in equation (4.1) is substituted for it (SMI) [6]. Hence, the system's performance is highly dependent on the size of the sample support K relative to N. The number of independent samples K used to generate the sample covariance matrix, affects the adapted beam

pattern generated by the array. The beam gain and sidelobe structure are directly related to the estimate of the covariance matrix. If a low number of samples is used, the array pattern will have a distorted beam shape and high sidelobes which will result in inaccurate target detection and poor performance. Kelly showed in [33] that the expected value of the array sidelobe pattern is given by

$$E(sidelobe) = \frac{1}{1+K}. \tag{4.3}$$

This shows that in order to achieve low gain in the sidelobe region, a very large number of secondary samples K should be used. In practical applications, the average clutter sidelobe level is attenuated to -40 to -80 dB.

The distorted mainbeam pattern is a result of widely spread eigenvalues of the sample covariance matrix which is in turn a result of a limited size of sample support K. The input to the radar receiver (returned data) is made up of interference contributions which include white noise. Noise exists at all Doppler frequencies, and hence, an infinite number of secondary samples are required to estimate the noise subspace. Strong interfering signals can be estimated using a low number of samples. As more samples are used, the estimation of the noise subspace improves and the noise eigenvalues converge to the expected value of the noise power. If the noise subspace is poorly estimated, an ill-conditioned covariance matrix estimate is obtained and hence, randomly shaped noise eigenbeams occur in the array's beam pattern. Applying the SMI algorithm using the ill-conditioned covariance matrix results in the subtraction of these eigenbeams and hence a distorted mainbeam pattern. The eigenvalues of a sample covariance matrix where the interference is low rank ($r < N$ in equation (4.2)) are shown for $K = N$ and $K = 10N$ in Figure 4.1. It is obvious from the Figure that the size of the sample support controls the noise eigenvalues spread. For this scenario, the noise eigenvalues are spread over more than 30 dB at $K = N$ as compared to only 3 or 4 dB at $K = 10N$. As K increases, the eigenvalues of the sample covariance matrix will converge towards the

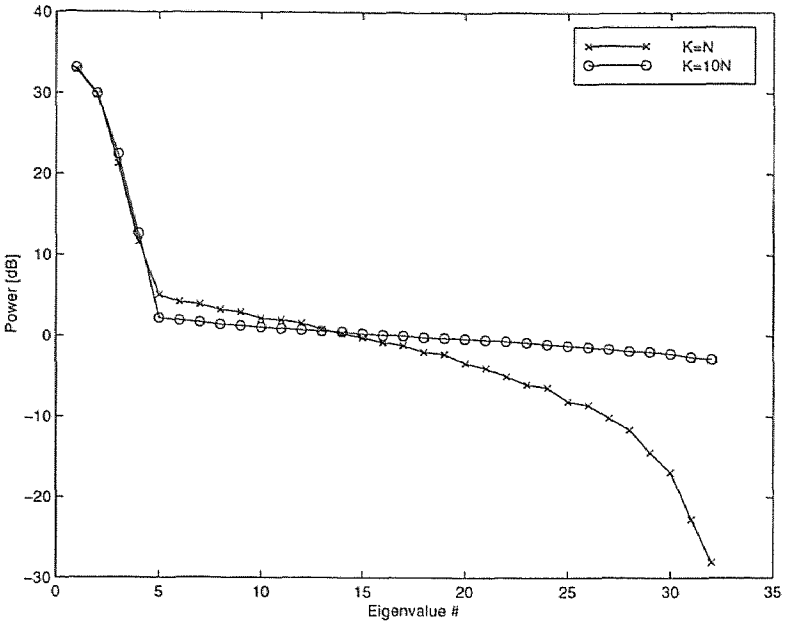

Figure 4.1 Eigenvalues of the sample covariance matrix.

eigenvalues of the true covariance matrix and all the noise eigenvalues will be at the noise power.

Diagonal loading is the process of replacing the sample covariance matrix, $\widehat{\mathbf{R}}$ by another matrix $\widehat{\mathbf{R}}_d$ where

$$\widehat{\mathbf{R}}_d = \widehat{\mathbf{R}} + \alpha \mathbf{I}. \tag{4.4}$$

The loading factor α is a scalar and $\mathbf{I}$ is the identity matrix. This process in effect reduces the spread of the noise eigenvalues. When an appropriate loading factor α is chosen, the strong interference eigenvalues will be minimally affected by the loading of the covariance matrix, but the widely spread noise eigenvalues will converge towards the loading factor α. This has the net effect of a better or less distorted beam pattern produced by the adaptive array. Hence, we can say that loading the covariance matrix is almost equivalent to using more secondary samples to obtain the sample covariance matrix from a performance standpoint. In Figure 4.2, the effect of loading is shown. For a particular radar scenario, a diagonal loading factor of 8 dB is used to load

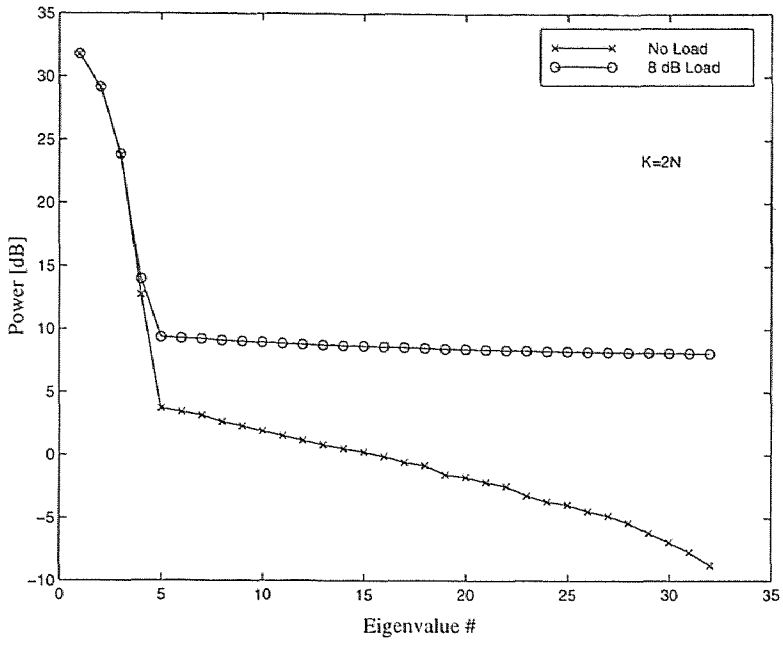

Figure 4.2 Eigenvalues of the sample covariance matrix under loading condition.

a sample covariance obtained by averaging $K = 2N$ secondary samples. The noise eigenvalues are not wide spread when diagonal loading is applied.

Hence the application of diagonal loading which reduces the spread of the noise eigenvalues, results in lower sidelobes in the mainbeam pattern and hence, results in higher values of the CSNR.

In the next section, we use the same approach used by Cheremisin and derive an expression for the distribution of the CSNR for L-SMI in the case that the sample covariance matrix is non-singular. We attempt to show that applying diagonal loading to the sample covariance matrix is equivalent to using a larger set of the secondary data for the estimation.

4.2 Distribution of the CSNR for $K > N$

In this section, the probability density function of the CSNR is developed. The CSNR expression was given in equation (3.12) as

$$\rho = \frac{\left|\mathbf{w}^H\mathbf{s}\right|^2}{\mathbf{w}^H\mathbf{R}\mathbf{w}\ \ \mathbf{s}^H\mathbf{R}^{-1}\mathbf{s}} \tag{4.5}$$

where $\mathbf{R}$ is the true covariance matrix of the colored noise and the loaded SMI weight vector $\mathbf{w}$ is given as

$$\mathbf{w} = k\widehat{\mathbf{R}}_d^{-1}\mathbf{s}. \tag{4.6}$$

The diagonally loaded sample covariance matrix $\widehat{\mathbf{R}}_d$ is give in equation (4.4). Substituting the expression of the weight vector into equation (4.5), the CSNR is written as

$$\rho = \frac{\left(\mathbf{s}^H\widehat{\mathbf{R}}_d^{-1}\mathbf{s}\right)^2}{\mathbf{s}^H\widehat{\mathbf{R}}_d^{-1}\mathbf{R}\widehat{\mathbf{R}}_d^{-1}\mathbf{s}\ \ \mathbf{s}^H\mathbf{R}^{-1}\mathbf{s}}. \tag{4.7}$$

The goal is to characterize the CSNR in equation (4.7) statistically. However, the CSNR expression is first simplified before the characterization can proceed.

Using equations (4.1) and (4.4), matrix $\widehat{\mathbf{R}}_d$ can be written as

$$\widehat{\mathbf{R}}_d = \frac{1}{K}\sum_{k=1}^{K}\mathbf{x}(k)\mathbf{x}(k)^H + \alpha\,\mathbf{I}. \tag{4.8}$$

A new matrix $\mathbf{X}$ is defined as a $N \times K$ matrix whose columns are made up of the statistically independent secondary data vectors $\mathbf{x}(k)$ for $k = 1, \cdots, K$. Under hypothesis $\mathbf{H}_1$ given in Chapter 2, $\mathbf{x}(k)$ for $k = 1, \cdots, K$ are zero mean, $N \times 1$ complex Gaussian random vector with the $N \times N$ covariance matrix $\mathbf{R}$. A distribution so defined is denoted as $\mathcal{CN}\ (\mathbf{0}, \mathbf{R})$. Hence, $\mathbf{X}$ is a matrix whose columns are distributed as $\mathcal{CN}\ (\mathbf{0}, \mathbf{R})$. Using this definition of the matrix $\mathbf{X}$, equation (4.8) is written as

$$\begin{aligned}\widehat{\mathbf{R}}_d &= \frac{1}{K}\mathbf{X}\mathbf{X}^H + \alpha\mathbf{I} \\ &= \alpha\left(\mathbf{I} + \frac{1}{K\alpha}\mathbf{X}\mathbf{X}^H\right).\end{aligned} \tag{4.9}$$

Substituting equation (4.9) into equation (4.7), the CSNR is written as

$$\rho = \frac{\left(\mathbf{s}^H \left(\mathbf{I} + \frac{1}{K\alpha}\mathbf{X}\mathbf{X}^H\right)^{-1} \mathbf{s}\right)^2}{\mathbf{s}^H \left(\mathbf{I} + \frac{1}{K\alpha}\mathbf{X}\mathbf{X}^H\right)^{-1} \mathbf{R} \left(\mathbf{I} + \frac{1}{K\alpha}\mathbf{X}\mathbf{X}^H\right)^{-1} \mathbf{s} \quad \mathbf{s}^H\mathbf{R}^{-1}\mathbf{s}}. \tag{4.10}$$

The true covariance matrix $\mathbf{R}$ is positive definite, and hence a positive definite square root matrix $\mathbf{R}^{1/2}$ can be defined. Then the matrix $\mathbf{R}^{-1/2}$ is obtained by taking the inverse of the matrix $\mathbf{R}^{1/2}$. Consider a matrix $\mathbf{Z}$ who is written in terms of matrix $\mathbf{X}$ as

$$\mathbf{Z} = \mathbf{R}^{-1/2}\mathbf{X}. \tag{4.11}$$

The columns of matrix $\mathbf{Z}$ are distributed as $\mathcal{CN}\ (\mathbf{0}, \mathbf{R}^{-1/2}\mathbf{R}\mathbf{R}^{-1/2})$ and equivalently as $\mathcal{CN}\ (\mathbf{0}, \mathbf{I})$ (see Appendix A) where

$$E[\mathbf{z}(l)\mathbf{z}(m)^H] = \mathbf{I}\,\delta(l, m) \tag{4.12}$$

where $\mathbf{z}_l$ and $\mathbf{z}_m$ are two columns of matrix $\mathbf{Z}$ and $\delta(l, m)$ is the Kronecker delta function defined as

$$\delta(l, m) = \begin{cases} 1 & \text{if } l = m \\ 0 & \text{otherwise} \end{cases}$$

Replacing matrix $\mathbf{X}$ in the CSNR expression in equation (4.10) by matrix $\mathbf{Z}$, the following expression is obtained

$$\rho = \frac{\left(\mathbf{s}^H\mathbf{R}^{-1/2}(\mathbf{R}^{-1} + \frac{1}{K\alpha}\mathbf{Z}\mathbf{Z}^H)^{-1}\mathbf{R}^{-1/2}\mathbf{s}\right)^2}{\mathbf{s}^H\mathbf{R}^{-1/2}(\mathbf{R}^{-1} + \frac{1}{K\alpha}\mathbf{Z}\mathbf{Z}^H)^{-1}\mathbf{R}^{-1/2}\mathbf{R}\mathbf{R}^{-1/2}(\mathbf{R}^{-1} + \frac{1}{K\alpha}\mathbf{Z}\mathbf{Z}^H)^{-1}\mathbf{R}^{-1/2}\mathbf{s} \quad \mathbf{s}^H\mathbf{R}^{-1}\mathbf{s}} \tag{4.13}$$

and equivalently

$$\rho = \frac{\left(\mathbf{s}^H\mathbf{R}^{-1/2}(\mathbf{R}^{-1} + \frac{1}{K\alpha}\mathbf{Z}\mathbf{Z}^H)^{-1}\mathbf{R}^{-1/2}\mathbf{s}\right)^2}{\mathbf{s}^H\mathbf{R}^{-1/2}(\mathbf{R}^{-1} + \frac{1}{K\alpha}\mathbf{Z}\mathbf{Z}^H)^{-2}\mathbf{R}^{-1/2}\mathbf{s} \quad \mathbf{s}^H\mathbf{R}^{-1}\mathbf{s}}. \tag{4.14}$$

Only matrix $\mathbf{Z}$ in the last CSNR expression is random.

The true covariance matrix $\mathbf{R}$ is assumed to obey the model given in equation (4.2). The inverse of this matrix is given as

$$\mathbf{R}^{-1} = \mathbf{Q}_r\mathbf{\Lambda}_r^{-1}\mathbf{Q}_r^H + \sigma_v^{-2}\mathbf{Q}_v\mathbf{Q}_v^H \tag{4.15}$$

where the $r \times r$ matrix $\mathbf{\Lambda}_r$ is diagonal and is given as

$$\mathbf{\Lambda}_r = \begin{bmatrix} \gamma_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \gamma_r \end{bmatrix} \tag{4.16}$$

and γ_l for $l = 1, \cdots, r$ are the interference eigenvalues. Assuming that the interference power is high relative to the noise power σ_v^2, such that for the minimum interference eigenvalue γ_{min} in the matrix $\mathbf{\Lambda}_r$, the following holds true

$$\gamma_{min}^{-1} \ll \sigma_v^{-2}. \tag{4.17}$$

Using this assumption on the interference power, the inverse of the covariance matrix in equation (4.15) can be written as

$$\mathbf{R}^{-1} \cong \sigma_v^{-2} \mathbf{Q}_v \mathbf{Q}_v^H \tag{4.18}$$

where $o(1/\gamma_l)$ for $l = 1, \cdots, r$ terms were neglected to obtain the last expression, where $o(\beta)$ refers to terms in the order of β. Using equation (4.18), $\mathbf{R}^{-1/2}$ is expressed as

$$\mathbf{R}^{-1/2} \cong \sigma_v^{-1} \mathbf{Q}_v \mathbf{Q}_v^H \tag{4.19}$$

under the same assumption of the high interference power. Setting the noise power σ_v^2 equal to unity, equations (4.18) and (4.19) are written as

$$\mathbf{R}^{-1} \cong \mathbf{Q}_v \mathbf{Q}_v^H \tag{4.20}$$

and

$$\mathbf{R}^{-1/2} \cong \mathbf{Q}_v \mathbf{Q}_v^H. \tag{4.21}$$

Substituting equations (4.20) and (4.21) for $\mathbf{R}^{-1}$ and $\mathbf{R}^{-1/2}$, respectively, in the CSNR expression in equation (4.14), the CSNR is written as

$$\rho = \frac{\left(\mathbf{s}^H \mathbf{Q}_v \mathbf{Q}_v^H (\mathbf{Q}_v \mathbf{Q}_v^H + \frac{1}{K\alpha} \mathbf{Z}\mathbf{Z}^H)^{-1} \mathbf{Q}_v \mathbf{Q}_v^H \mathbf{s}\right)^2}{\mathbf{s}^H \mathbf{Q}_v \mathbf{Q}_v^H (\mathbf{Q}_v \mathbf{Q}_v^H + \frac{1}{K\alpha} \mathbf{Z}\mathbf{Z}^H)^{-2} \mathbf{Q}_v \mathbf{Q}_v^H \mathbf{s}\mathbf{s}^H \mathbf{Q}_v \mathbf{Q}_v^H \mathbf{s}}. \tag{4.22}$$

Defining a matrix C such that

$$\mathbf{C} = (\mathbf{Q}_v\mathbf{Q}_v^H + \frac{1}{K\alpha}\mathbf{Z}\mathbf{Z}^H)^{-1}, \tag{4.23}$$

equation (4.22) is written as

$$\rho = \frac{\left(\mathbf{s}^H\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{C}\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{s}\right)^2}{\mathbf{s}^H\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{C}^2\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{s}\mathbf{s}^H\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{s}}. \tag{4.24}$$

To further simplify the last expression, we can define a normalized deterministic $(N-r)\times 1$ vector $\mathbf{b}$ which is related to the steering vector $\mathbf{s}$ as

$$\mathbf{b} = \frac{\mathbf{Q}_v^H\mathbf{s}}{\sqrt{\mathbf{s}^H\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{s}}}. \tag{4.25}$$

Using this definition of vector $\mathbf{b}$, equation (4.24) is written as

$$\rho = \frac{\left(\mathbf{b}^H\mathbf{Q}_v^H\mathbf{C}\mathbf{Q}_v\mathbf{b}\right)^2}{\mathbf{b}^H\mathbf{Q}_v^H\mathbf{C}^2\mathbf{Q}_v\mathbf{b}}. \tag{4.26}$$

At this point, since matrix $\mathbf{C}$ is not a simple matrix to analyze, we further proceed to obtain simplified expressions for the two matrices $\mathbf{Q}_v^H\mathbf{C}\mathbf{Q}_v$ and $\mathbf{Q}_v^H\mathbf{C}^2\mathbf{Q}_v$ appearing in the numerator and denominator of equation (4.26) respectively.

From the definition of matrix C in equation (4.23) and by pre-multiplying with $\mathbf{C}^{-1}$, we get

$$\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{C} + \frac{1}{K\alpha}\mathbf{Z}\mathbf{Z}^H\mathbf{C} = \mathbf{I}. \tag{4.27}$$

Using the fact that for the two orthogonal subspaces $\mathbf{Q}_r$ and $\mathbf{Q}_v$, it is true that $\mathbf{Q}_v\mathbf{Q}_v^H + \mathbf{Q}_r\mathbf{Q}_r^H = \mathbf{I}$, equation (4.27) can be written as

$$\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{C} + \frac{1}{K\alpha}\mathbf{Z}\mathbf{Z}^H\left(\mathbf{Q}_v\mathbf{Q}_v^H + \mathbf{Q}_r\mathbf{Q}_r^H\right)\mathbf{C} = \mathbf{I} \tag{4.28}$$

and equivalently

$$\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{C} + \frac{1}{K\alpha}\mathbf{Z}\mathbf{Z}^H\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{C} + \frac{1}{K\alpha}\mathbf{Z}\mathbf{Z}^H\mathbf{Q}_r\mathbf{Q}_r^H\mathbf{C} = \mathbf{I}. \tag{4.29}$$

By pre-multiplying equation (4.29) by $\mathbf{Q}_r^H$ and $\mathbf{Q}_v^H$, and by using the orthogonality between the two subspaces, the following two equations are obtained

$$\begin{aligned} \frac{1}{K\alpha}\mathbf{Q}_r^H\mathbf{Z}\mathbf{Z}^H\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{C} + \frac{1}{K\alpha}\mathbf{Q}_r^H\mathbf{Z}\mathbf{Z}^H\mathbf{Q}_r\mathbf{Q}_r^H\mathbf{C} &= \mathbf{Q}_r^H \\ \mathbf{Q}_v^H\mathbf{C} + \frac{1}{K\alpha}\mathbf{Q}_v^H\mathbf{Z}\mathbf{Z}^H\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{C} + \frac{1}{K\alpha}\mathbf{Q}_v^H\mathbf{Z}\mathbf{Z}^H\mathbf{Q}_r\mathbf{Q}_r^H\mathbf{C} &= \mathbf{Q}_v^H. \end{aligned} \quad (4.30)$$

Solving these two equations simultaneously for $\mathbf{Q}_v^H\mathbf{C}$, the following equation is obtained

$$\mathbf{Q}_v^H\mathbf{C} = \left(\mathbf{I} + \frac{1}{K\alpha}\mathbf{A}^H[I - \mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H]\mathbf{A}\right)^{-1}\left(\mathbf{Q}_v^H - \mathbf{A}^H\mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{Q}_r^H\right) \quad (4.31)$$

where matrices $\mathbf{A}$ and $\mathbf{B}$ are of sizes $K \times (N-r)$ and $K \times r$, respectively, and are defined as

$$\begin{aligned} \mathbf{A} &= \mathbf{Z}^H\mathbf{Q}_v \\ \mathbf{B} &= \mathbf{Z}^H\mathbf{Q}_r. \end{aligned} \quad (4.32)$$

Since the columns of matrix $\mathbf{Z}$ are distributed as $\mathcal{CN}(\mathbf{0},\mathbf{I})$ and since the two subspaces $\mathbf{Q}_v$ and $\mathbf{Q}_r$ are orthogonal, then the two matrices $\mathbf{A}$ and $\mathbf{B}$ are statistically independent. Moreover, the columns of each matrix $\mathbf{A}$ and $\mathbf{B}$ are independent and distributed as $\mathcal{CN}(\mathbf{0},\mathbf{I})$ (see Appendix A).

The matrix $\mathbf{P}_B = \mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H$ is a projection matrix, and so is the matrix $\mathbf{P}_D = \mathbf{I} - \mathbf{P}_B$. Projection matrices have unity eigenvalues such that $\mathbf{P}_B^2 = \mathbf{P}_B^H = \mathbf{P}_B$ and $\mathbf{P}_D^2 = \mathbf{P}_D^H = \mathbf{P}_D$ [34]. The expression in the square brackets in the first term of the right hand side of equation (4.31) represents a projection matrix and can be written as

$$[\mathbf{I} - \mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H] = \mathbf{D}\mathbf{D}^H \quad (4.33)$$

where matrix $\mathbf{D}$ is of dimension $K \times (K-r)$ and is orthogonal to the matrix $\mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1/2}$ where $\mathbf{D}^H\mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1/2} = \mathbf{0}$. Finally, $\mathbf{Q}_v^H\mathbf{C}$ in equation (4.31) can be written as

$$\mathbf{Q}_v^H\mathbf{C} = \left(\mathbf{I} + \frac{1}{K\alpha}\mathbf{A}^H\mathbf{D}\mathbf{D}^H\mathbf{A}\right)^{-1}\left(\mathbf{Q}_v^H - \mathbf{A}^H\mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{Q}_r^H\right). \quad (4.34)$$

At this point, as given in equation (4.34), the matrix $\mathbf{Q}_v^H\mathbf{C}$ is expressed in terms of matrices which are either deterministic or random with known distributions. Using equation (4.34), we proceed to find expressions for the two matrices $\mathbf{Q}_v^H\mathbf{C}\mathbf{Q}_v$ and $\mathbf{Q}_v^H\mathbf{C}^2\mathbf{Q}_v$ where both appear in the CSNR expression in equation (4.26).

There are two possible conditions under which equation (4.34) can be considered: a small loading factor where $\alpha \ll \frac{1}{K}$ and a large loading factor where $\alpha \gg \frac{1}{K}$. However, if the loading factor is very small and close to zero, the distribution of the CSNR expression in equation (4.26) seems to be intractable. Moreover, if the estimated covariance matrix is non-singular ($K > N$) and the loading factor α is very small, then L-SMI becomes SMI. The distribution of the CSNR when the weight vector is obtained using the SMI method was obtained by Reed et al. in [6]. Hence, in the proceeding derivation, only the case of a large loading factor is considered. The final obtained distribution of the CSNR for L-SMI will be compared with Reed's distribution of CSNR for SMI.

For a large loading factor α relative to the unity elements of the identity matrix appearing in right hand side of equation (4.34) where

$$\frac{1}{K\alpha} \ll 1 \tag{4.35}$$

or equivalently

$$\alpha \gg \frac{1}{K}, \tag{4.36}$$

equation (4.34) can be written as

$$\mathbf{Q}_v^H\mathbf{C} \cong \left(\mathbf{Q}_v^H - \mathbf{A}^H\mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{Q}_r^H\right) \tag{4.37}$$

where $o(1/K\alpha)$ terms were neglected to obtain equation (4.37). Using this expression for $\mathbf{Q}_v^H\mathbf{C}$, expressions for the two matrices $\mathbf{Q}_v^H\mathbf{C}\mathbf{Q}_v$ and $\mathbf{Q}_v^H\mathbf{C}^2\mathbf{Q}_v$ can be obtained. The first matrix $\mathbf{Q}_v^H\mathbf{C}\mathbf{Q}_v$ can be obtained from equation (4.37) by post-multiplying

$\mathbf{Q}_v^H\mathbf{C}$ by $\mathbf{Q}_v$ as

$$\begin{aligned}\mathbf{Q}_v^H\mathbf{C}\mathbf{Q}_v &= \left(\mathbf{Q}_v^H\mathbf{C}\right)\mathbf{Q}_v \\ &= \left(\mathbf{Q}_v^H - \mathbf{A}^H\mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{Q}_r^H\right)\mathbf{Q}_v \\ &= \mathbf{I} \end{aligned} \tag{4.38}$$

since the two subspaces $\mathbf{Q}_v$ and $\mathbf{Q}_r$ are orthogonal. The second matrix $\mathbf{Q}_v^H\mathbf{C}^2\mathbf{Q}_v$ is obtained as

$$\begin{aligned}\mathbf{Q}_v^H\mathbf{C}^2\mathbf{Q}_v &= \left(\mathbf{Q}_v^H\mathbf{C}\right)\left(\mathbf{Q}_v^H\mathbf{C}\right)^H \\ &= \left(\mathbf{Q}_v^H - \mathbf{A}^H\mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{Q}_r^H\right)\left(\mathbf{Q}_v^H - \mathbf{A}^H\mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{Q}_r^H\right)^H \\ &= \left(\mathbf{I} + \mathbf{A}^H\mathbf{B}(\mathbf{B}^H\mathbf{B})^{-2}\mathbf{B}^H\mathbf{A}\right). \end{aligned} \tag{4.39}$$

Substituting equations (4.38) and (4.39) into equation (4.26), the CSNR is written as

$$\rho = \frac{\left(\mathbf{b}^H\mathbf{b}\right)^2}{\mathbf{b}^H\left(\mathbf{I} + \mathbf{A}^H\mathbf{B}(\mathbf{B}^H\mathbf{B})^{-2}\mathbf{B}^H\mathbf{A}\right)\mathbf{b}} \tag{4.40}$$

and equivalently

$$\rho = \frac{\left(\mathbf{b}^H\mathbf{b}\right)^2}{\mathbf{b}^H\mathbf{b} + \mathbf{b}^H\mathbf{A}^H\mathbf{B}(\mathbf{B}^H\mathbf{B})^{-2}\mathbf{B}^H\mathbf{A}\mathbf{b}}. \tag{4.41}$$

From the definition of the deterministic vector $\mathbf{b}$ in equation (4.25), $\mathbf{b}^H\mathbf{b} = 1$ and hence, equation (4.41) is written as

$$\rho = \frac{1}{1 + \zeta^H(\mathbf{B}^H\mathbf{B})^{-1}\zeta} \tag{4.42}$$

where ζ is a $r \times 1$ vector that is expressed as

$$\zeta = (\mathbf{B}^H\mathbf{B})^{-1/2}\mathbf{B}^H\mathbf{A}\mathbf{b}. \tag{4.43}$$

As was previously shown, the columns of matrices $\mathbf{B}$ and $\mathbf{A}$ are statistically independent and are distributed as $\mathcal{CN}\ (\mathbf{0}, \mathbf{I})$. Now, the statistical properties of

the vector ζ are sought. Let $\mathbf{Ab} = \mathbf{c}$, then the mean of vector $\mathbf{c}$ is

$$\begin{aligned} E\{\mathbf{c}\} &= E\{\mathbf{Ab}\} \\ &= \mathbf{0}, \end{aligned} \tag{4.44}$$

since the columns of matrix $\mathbf{A}$ have zero mean. The covariance of vector $\mathbf{c}$ is

$$\begin{aligned} E\{\mathbf{cc}^H\} &= E\{(\mathbf{Ab})(\mathbf{Ab})^H\} \\ &= \sum_{k}^{N-r} \mathbf{A}_k b_k \sum_{m}^{N-r} b_m^H \mathbf{A}_m^H \\ &= \sum_{k}^{N-r} b_k^H b_k \mathbf{A}_k \mathbf{A}_k \\ &= \mathbf{I} \end{aligned} \tag{4.45}$$

where $\mathbf{A}_k$ for $k = 1, \cdots, (N-r)$ are columns of matrix $\mathbf{A}$ and b_m for $m = 1, \cdots, (N-r)$ are elements of vector $\mathbf{b}$. Hence vector $\mathbf{c}$ is distributed as $\mathcal{CN}(\mathbf{0}, \mathbf{I})$ and is statistically independent of matrix $\mathbf{B}$. The vector ζ is expressed in terms of vector $\mathbf{c}$ as

$$\zeta = (\mathbf{B}^H\mathbf{B})^{-1/2}\mathbf{B}^H\mathbf{c}. \tag{4.46}$$

Let the matrix $\mathbf{Q}^H = (\mathbf{B}^H\mathbf{B})^{-1/2}\mathbf{B}^H$. Since the vector $\mathbf{c}$ is statistically independent of matrix $\mathbf{B}$, matrix $\mathbf{Q}^H$ is statistically independent of vector $\mathbf{c}$. Also, $\mathbf{Q}^H\mathbf{Q} = \mathbf{I}$. Then the vector ζ is written in terms of vector $\mathbf{c}$ and matrix $\mathbf{Q}$ as

$$\zeta = \mathbf{Q}^H\mathbf{c}. \tag{4.47}$$

Since vector $\mathbf{c}$ is distributed as $\mathcal{CN}(\mathbf{0}, \mathbf{I})$, then by Appendix A, vector ζ is also $\mathcal{CN}(\mathbf{0}, \mathbf{I})$ and is statistically independent of matrix $\mathbf{Q}$ and hence, statistically independent of matrix $\mathbf{B}$. At this point, all the quantities appearing in the CSNR expression in equation (4.42) are of known distribution.

The distribution of the quantity $\xi = \zeta^H(\mathbf{B}^H\mathbf{B})^{-1}\zeta$ is obtained first and then by performing a transformation of variables, the distribution of ρ in equation (4.42)

is obtained. The quantity ξ obeys a multivariate F-distribution (Hotelling's T^2) [35] given as (see Appendix B)

$$f_\xi(\xi) = \frac{\Gamma(K+1)}{\Gamma(K-r+1)\Gamma(r)}\xi^{r-1}(1+\xi)^{-(K+1)}. \tag{4.48}$$

Performing a transformation of variables where ρ is related to ξ as

$$\rho = \frac{1}{1+\xi}, \tag{4.49}$$

the distribution of the CSNR in equation (4.42) is given as

$$f_\rho(\rho) = \frac{\Gamma(K+1)}{\Gamma(K-r+1)\Gamma(r)}\rho^{K-r}(1-\rho)^{r-1}. \tag{4.50}$$

This CSNR distribution has a mean value of

$$E\{\rho\} = \frac{K-r+1}{K+1} \tag{4.51}$$

where for $E\{\rho\}$ to equal 0.5, $K = 2r-1$ secondary data vectors are needed to estimate the covariance matrix. The final expression for the distribution of the CSNR for L-SMI given in equation (4.50) is the same as the distribution obtained by Cheremisin in [20] when the sample covariance matrix is singular.

4.3 Discussion

The expression for the distribution of the CSNR for L-SMI given in (4.50) is a function of the number of secondary data vectors used to estimate the covariance matrix and the number of interference (principal) eigenvalues of the covariance matrix. This distribution is similar in form to the expression obtained by Reed et al. in [6] for the SMI, which was given as

$$f_\rho(\rho) = \frac{\Gamma(K+1)}{\Gamma(K-N+2)\Gamma(N-1)}\rho^{K-N+1}(1-\rho)^{N-2}. \tag{4.52}$$

By comparing the distributions in equations (4.50) and (4.52) for the L-SMI and SMI respectively, it can be shown that for an average CSNR value of 0.5, $K \approx 2r$

samples are needed to estimate the covariance given in equation (4.4) for L-SMI, while $K \approx 2N$ samples are needed to estimate the covariance given in equation (4.1) for SMI (no loading). Assuming that the covariance matrix is of low rank and obeys the model given in equation (4.2) where $r < N$, less samples are needed to estimate the covariance matrix if loading is applied for an equivalent performance as compared to SMI. This is important when the radar is operating in a limited sample support environment. The trade off in using diagonal loading is that the system becomes less sensitive to weak interferences that are close to the noise power.

The next section provides comparison between the distribution of the CSNR for L-SMI and SMI theoretically and by using simulations.

4.4 Performance Analysis

The simulation model assumed a linear uniform antenna array with $N_s = 8$ elements spaced at half wavelength. Each array channel consisted of $N_t = 4$ tap FIR filter. The clutter was modeled to have a complex-valued Gaussian distribution, with zero mean, and variance equal to the clutter-to-noise ratio (CNR). We assumed the clutter returns are uncorrelated with each other and also uncorrelated between snapshots. The clutter consisted of 120 point sources randomly distributed in the azimuth sector 0 - 30 degrees with respect to the array's boresight. The CNR was calculated from the contributions of all clutter sources and was set to 10 dB. The clutter map in angle-Doppler coordinates is shown in Figure 4.3. The steering vector was pointed at a target assumed to be at 50 degrees and 0.5 normalized Doppler frequency with respect to the PRF. The steering vector is shown as the straight line on the clutter map in Figure 4.3.

The sample covariance matrix was obtained from equation (4.1) using $K = 2N_sN_t$ independent identically distributed snapshots. The eigenvalues of the sample covariance matrix are shown in Figure 4.4. The interference subspace has a rank of

$r = 4$. Equation (4.5) is used to calculate the CSNR values. The probability density functions of the CSNR are given in Figure 4.5 for the cases of small and large loading factors, obtained by averaging 2000 Monte-Carlo runs. The large loading factor used was 8 dB. The theoretical distributions of the CSNR are plotted and given in Figure 4.6. The theoretical distribution of the CSNR for a small loading factor is the distribution obtained by Reed et al. in [6] for the SMI adaptive detector which is also given in equation (4.52). The theoretical distribution of the CSNR for a large loading factor is obtained by plotting the expression given in equation (4.50). The simulation and theoretical CSNR distributions are very similar.

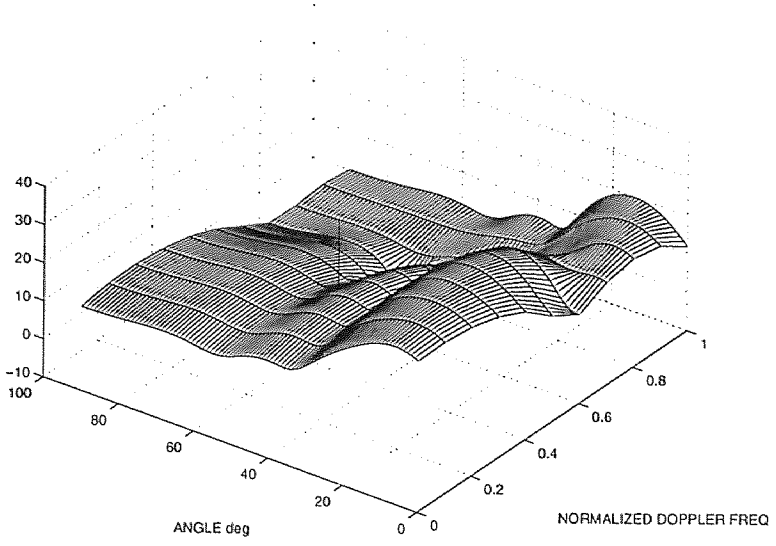

Figure 4.3 Simulation clutter map.

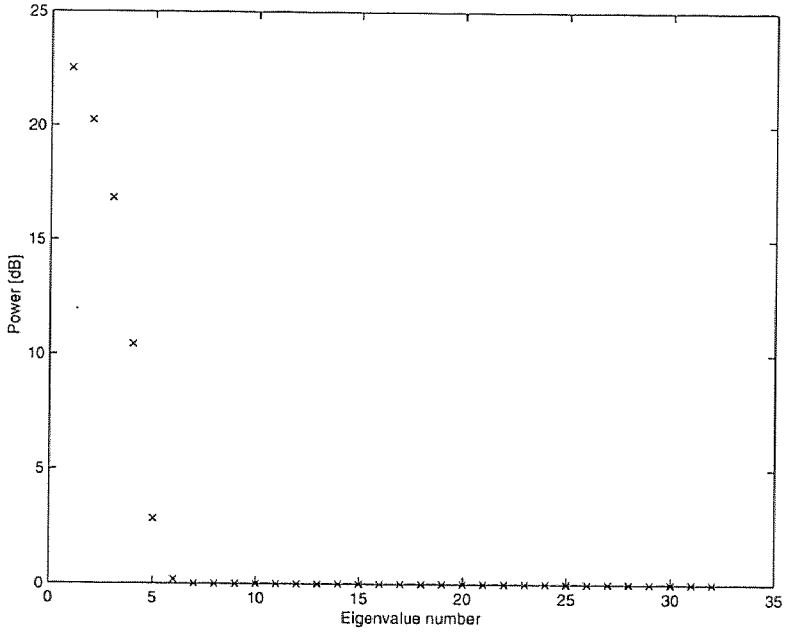

Figure 4.4 The power of the eigenvalues of the true covariance matrix.

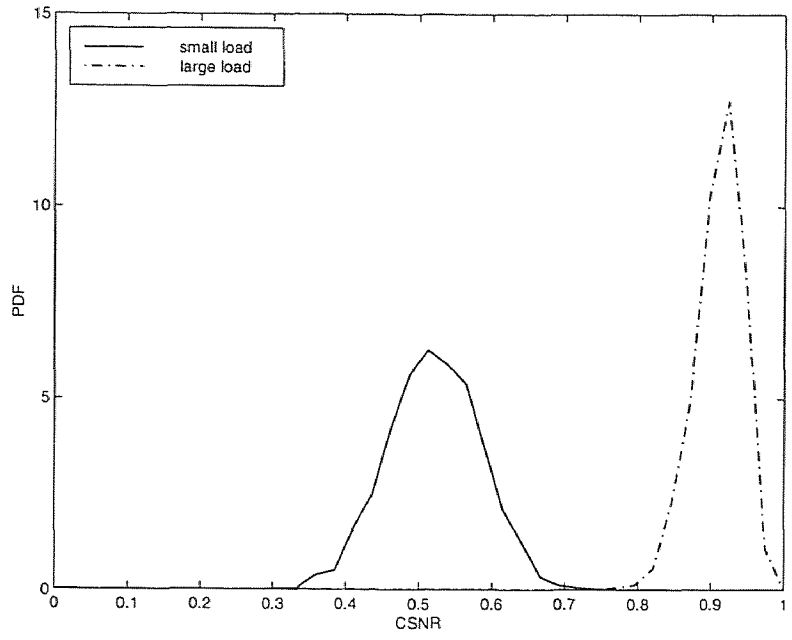

Figure 4.5 Simulated CSNR curves.

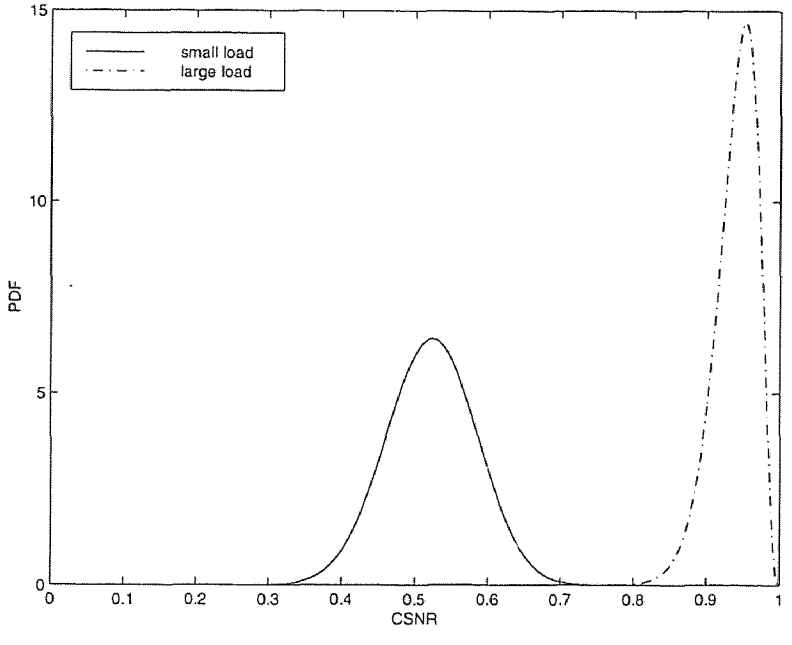

Figure 4.6 Theoretical CSNR curves.

CHAPTER 5

GENERALIZED LIKELIHOOD RATIO TEST

Kelly formulated and derived a decision rule for signal detection under the assumption of Gaussian interference with unknown covariance [10]. The algorithm is based on maximum likelihood ratio principle and is referred to as the generalized maximum likelihood ratio test (GLRT). The generalization is done on the target model, where the likelihood test is maximized over all unknown target parameters. The GLRT exhibits the property of constant false alarm rate (CFAR) where the probability of false alarm does not depend on the noise covariance matrix. The CFAR property is desirable in detectors since the interference is non-homogeneous and hence, for non CFAR detectors, the detection threshold has to be continuously adjusted. The adaptive detector of Reed, Mallet, and Brennan [6] used for interference rejection is not a CFAR detector. This detector is referred to as the sample matrix inversion (SMI) detector. A modification of this detector has been considered by various authors [7, 8, 9, 36, 37]. Robey et al. derived a SMI detector obeying the CFAR property which was called the adaptive matched filter (AMF) detector. The AMF detector is identical to Kelly's GLRT detector except for a missing term which vanishes as the size of the sample support used in estimating the covariance matrix becomes very large. Analysis and comparison of the GLRT and the AMF detectors showed that both lose in performance due to estimating the covariance matrix from limited sample support and also both detectors suffer a CFAR loss in SNR due to the CFAR normalization. These losses are relative to an optimum matched filter detector with known covariance matrix. Reed's non-CFAR SMI detector does not suffer a CFAR loss with the trade off being that it is non-CFAR. The GLRT outperforms the AMF at most SNR values. The AMF, depending on the signal vector and sample support dimensions, can outperform the GLRT at very high SNR ratio. Also, the

AMF detector suffers in performance in the case of any signal mismatch where the actual signal is not aligned with the presumed steering vector [37].

Both of these detectors, the GLRT and AMF, are dependent on the sample support size relative to the system's dimensions. When the interference is confined to a subspace of the signal space, the SNR loss can be reduced by mapping the data into that subspace prior to detection [38, 39, 15]. Hence, a reduced rank version of the GLRT detector (GLRT-RR) is considered in the next section. As a result of subspace transformation, the required sample support needed to estimate the noise unknown covariance matrix decreases because of the reduced dimensions. The trade off is that subspace processing results in an additional SNR loss which is a direct function of the specific transformation applied. The loss is referred to as a SNR bias [15]. This bias, however, can be minimized by designing optimum subspace transformations [38, 13]. The resulting probability of detection of the GLRT-RR detector is a function of the subspace transformation.

Recalling the results from Chapter 4, applying diagonal loading to the estimated covariance matrix improves the conditioning of the eigenvalues, and hence, emulates the effect of using a larger sample support. The trade off in using diagonal loading is that the system will become less sensitive to weak interferences. It is expected that if the estimated covariance matrix in the GLRT test statistic is diagonally loaded, the resulting probability of detection should improve. In other words, diagonal loading will reduce the SNR loss suffered by the detector, which is a direct function of the size of the sample support. Both subspace transformation and diagonal loading can improve the performance of the GLRT detector by reducing the size of the sample support required to achieve a certain performance or equivalently, by achieving a faster convergence rate.

The objective of this chapter is two fold. First, Kelly's GLRT detector is extended to obtain a general GLRT-RR test for fixed subspace transformations and

expressions for the probability of false alarm P_{fa} and the probability of detection P_D are obtained. It will be shown that the P_D is a function of the subspace transformation applied to the data prior to detection. Second, a GLRT test is derived when the estimated covariance matrix is diagonally loaded prior to detection, a test which is referred to as GLRT-LSMI.

This chapter is organized in the following order. Section 5.1 develops of the likelihood ratio test. Section 5.2 extends the GLRT to incorporate subspace transformation and the performance of the GLRT-RR is investigated. GLRT-LSMI is derived and analyzed in Section 5.3. Numerical results and the discussion are given in Section 5.4.

5.1 The Likelihood Ratio Test

Given an N-dimensional complex Gaussian primary data vector $\mathbf{z}$ from a range cell in which desired signal presence is being sought and a set of K additional N-dimensional complex Gaussian secondary data vectors $\mathbf{z}(k)$ for $k = 1, \cdots, K$ from neighboring range cells (desired signal free), two hypotheses are formulated. Under hypothesis $\mathbf{H}_0$, no desired signal is present and $\mathbf{z}(k)$ is the sum of clutter $\mathbf{c}(k)$ (interference) and additive white Gaussian noise contributions $\mathbf{n}(k)$:

$$\mathbf{H}_0 : \mathbf{z}(k) = \mathbf{c}(k) + \mathbf{n}(k). \tag{5.1}$$

The clutter and noise are assumed to be independent. The signal model under $\mathbf{H}_1$ is given by

$$\mathbf{H}_1 : \mathbf{z}(k) = b\mathbf{s} + \mathbf{c}(k) + \mathbf{n}(k), \tag{5.2}$$

where b is the signal amplitude and $\mathbf{s}$ represents the space-time steering vector. The colored noise (clutter and white noise) true covariance matrix $\mathbf{R}$ is typically unknown and hence, the sample covariance matrix is obtained as

$$\widehat{\mathbf{R}} = \frac{1}{K} \sum_{k=1}^{K} \mathbf{z}(k)\mathbf{z}(k)^H. \tag{5.3}$$

The secondary data vectors $\mathbf{z}(k)$ for $k = 1, \cdots, K$ have complex Gaussian distributions with zero mean and covariance matrix $\mathbf{R}$. In Goodman [40], it is shown that the distribution of the elements of $\widehat{\mathbf{R}}$ is given by the central complex Wishart distribution

$$f(K\widehat{\mathbf{R}}) = \frac{|K\widehat{\mathbf{R}}|^{K-N}}{\tilde{\Gamma}(N)|\mathbf{R}|^{K}} e^{-\mathrm{tr}\left(K\mathbf{R}^{-1}\widehat{\mathbf{R}}\right)} \tag{5.4}$$

where

$$\tilde{\Gamma}(N) = \pi^{\frac{N(N-1)}{2}} \Pi_{i=0}^{N-1} \Gamma(K-i). \tag{5.5}$$

The fundamental parameters for the central complex Wishart distribution are the number of secondary vectors K, the dimension of the vectors N, and the $N \times N$ true covariance matrix $\mathbf{R}$. Thus, the distribution of $\widehat{\mathbf{R}}$ is designated by $\mathcal{CW}(K, N; \mathbf{R})$. The matrix $\widehat{\mathbf{R}}$ is the maximum likelihood estimate of the covariance matrix and is highly dependent on the size of the secondary data set K, with the condition that $K \geq N$ for the estimate to be non-singular.

Under noise-alone hypothesis $\mathbf{H}_0$, the primary vector does not contain a desired signal and hence the primary vector along with the secondary vectors have complex normal distributions with zero mean and covariance matrix $\mathbf{R}$ and are denoted as $\mathcal{CN}\ (0, \mathbf{R})$. Under the same hypothesis, the joint probability density function (PDF) for all the input vectors is

$$f_0(\mathbf{z}, \mathbf{z}(1), \cdots, \mathbf{z}(K)) = \left\{ \frac{1}{\pi^N |\mathbf{R}|} e^{-\mathrm{tr}\ (\mathbf{R}^{-1}\mathbf{T}_0)} \right\}^{K+1} \tag{5.6}$$

where tr(.)is the trace operator and

$$\mathbf{T}_0 = \frac{1}{K+1} \left(\mathbf{z}\mathbf{z}^H + \sum_{k=1}^{K} \mathbf{z}(k)\mathbf{z}(k)^H \right). \tag{5.7}$$

Under signal presence hypothesis $\mathbf{H}_1$, the primary vector $\mathbf{z}$ will contain the desired signal and, hence, is $\mathcal{CN}\ (b\mathbf{s}, \mathbf{R})$ where b is the target's amplitude. The joint PDF in this case is

$$f_1(\mathbf{z}, \mathbf{z}(1), \cdots, \mathbf{z}(K)) = \left\{ \frac{1}{\pi^N |\mathbf{R}|} e^{-\mathrm{tr}\ (\mathbf{R}^{-1}\mathbf{T}_1)} \right\}^{K+1} \tag{5.8}$$

where

$$\mathbf{T}_1 = \frac{1}{K+1}\left((\mathbf{z}-b\mathbf{s})(\mathbf{z}-b\mathbf{s})^H + \sum_{k=1}^{K}\mathbf{z}(k)\mathbf{z}^H(k)\right). \tag{5.9}$$

The two PDFs are maximized over all unknown parameters (target's amplitude b) and the ratio of the maximas is obtained to form the detection statistic which can be written as

$$\frac{|\mathbf{s}^H\widehat{\mathbf{R}}^{-1}\mathbf{z}|^2}{\mathbf{s}^H\widehat{\mathbf{R}}^{-1}\mathbf{s}\left(1+\frac{1}{K}\mathbf{z}^H\widehat{\mathbf{R}}^{-1}\mathbf{z}\right)} \underset{\mathbf{H}_0}{\overset{\mathbf{H}_1}{\gtrless}} K\eta_o \tag{5.10}$$

where η_o is a threshold that satisfies the false alarm rate P_{fa}.

The AMF detector [37], which is a CFAR version of the SMI detector, has the same decision rule as that of the GLRT given in equation (5.10) except that the term $(1+\frac{1}{K}\mathbf{z}^H\widehat{\mathbf{R}}^{-1}\mathbf{z})$ is missing. Obviously, the two detectors will be equivalent if K is very large where $\frac{1}{K}\mathbf{z}^H\widehat{\mathbf{R}}^{-1}\mathbf{z}$ will become very close to zero.

The true covariance matrix $\mathbf{R}$ is positive definite and hence, $\mathbf{R}^{1/2}$ is also positive definite and can be defined. The matrix $\mathbf{R}^{-1/2}$ is the inverse of the matrix $\mathbf{R}^{1/2}$. Now consider the whitened steering vector

$$\tilde{\mathbf{s}} = \mathbf{R}^{-1/2}\mathbf{s}, \tag{5.11}$$

the whitened primary vector

$$\tilde{\mathbf{z}} = \mathbf{R}^{-1/2}\mathbf{z}, \tag{5.12}$$

and the whitened secondary vectors

$$\tilde{\mathbf{z}}(k) = \mathbf{R}^{-1/2}\mathbf{z}(k), \qquad k = 1,\cdots,K. \tag{5.13}$$

The primary and secondary vectors have Gaussian distributions with covariance

$$\begin{aligned} E\{\tilde{\mathbf{z}}(k)\tilde{\mathbf{z}}(k)^H\} &= \mathbf{R}^{-1/2}E\{\mathbf{z}(k)\mathbf{z}(k)^H\}\mathbf{R}^{-1/2} \\ &= \mathbf{R}^{-1/2}\mathbf{R}\mathbf{R}^{-1/2} \\ &= \mathbf{I}. \end{aligned} \tag{5.14}$$

The matrix $\mathbf{M}$ is defined as

$$\mathbf{M} = \mathbf{R}^{-1/2}\widehat{\mathbf{R}}\mathbf{R}^{-1/2}. \tag{5.15}$$

By using the expression for matrix $\widehat{\mathbf{R}}$ given in equation (5.3) and the definition in equation (5.13), matrix $\mathbf{M}$ is written as

$$\begin{aligned}
\mathbf{M} &= \frac{1}{K}\mathbf{R}^{-1/2}\left(\sum_{k=1}^{K}\mathbf{z}(k)\mathbf{z}(k)^H\right)\mathbf{R}^{-1/2} \\
&= \frac{1}{K}\sum_{k=1}^{K}\mathbf{R}^{-1/2}\mathbf{z}(k)\,\mathbf{z}(k)^H\mathbf{R}^{-1/2} \\
&= \frac{1}{K}\sum_{k=1}^{K}\tilde{\mathbf{z}}(k)\tilde{\mathbf{z}}(k)^H.
\end{aligned} \tag{5.16}$$

It was shown that matrix $\widehat{\mathbf{R}}$ is distributed as $\mathcal{CW}(K, N; \mathbf{R})$. The expectation of matrix $\mathbf{M}$ is

$$\begin{aligned}
E\{\mathbf{M}\} &= \mathbf{R}^{-1/2}\,E\{\widehat{\mathbf{R}}\}\,\mathbf{R}^{-1/2} \\
&= \mathbf{R}^{-1/2}\mathbf{R}\mathbf{R}^{-1/2} \\
&= \mathbf{I},
\end{aligned} \tag{5.17}$$

where $\mathbf{I}$ is a $N \times N$ identity matrix. Hence, $\mathbf{M}$ is distributed as $\mathcal{CW}(K, N; \mathbf{I})$.

Having defined the whitened vectors and matrix, the following can be written

$$\mathbf{s}^H\widehat{\mathbf{R}}^{-1}\mathbf{z} = \tilde{\mathbf{s}}^H\mathbf{M}^{-1}\tilde{\mathbf{z}} \tag{5.18}$$

$$\mathbf{s}^H\widehat{\mathbf{R}}^{-1}\mathbf{s} = \tilde{\mathbf{s}}^H\mathbf{M}^{-1}\tilde{\mathbf{s}} \tag{5.19}$$

$$\mathbf{z}^H\widehat{\mathbf{R}}^{-1}\mathbf{z} = \tilde{\mathbf{z}}^H\mathbf{M}^{-1}\tilde{\mathbf{z}}. \tag{5.20}$$

Incorporating these expressions into the GLRT test in equation (5.10), the test is written as

$$\frac{|\tilde{\mathbf{s}}^H\mathbf{M}^{-1}\tilde{\mathbf{z}}|^2}{\tilde{\mathbf{s}}^H\mathbf{M}^{-1}\tilde{\mathbf{s}}\left(1+\frac{1}{K}\tilde{\mathbf{z}}^H\mathbf{M}^{-1}\tilde{\mathbf{z}}\right)} \underset{\mathbf{H}_0}{\overset{\mathbf{H}_1}{\gtrless}} K\eta_o. \tag{5.21}$$

By defining the matrix $\mathbf{S}$ as

$$\begin{aligned} \mathbf{S} &= K\,\mathbf{M} \\ &= \sum_{k=1}^{K} \tilde{\mathbf{z}}(k)\tilde{\mathbf{z}}(k)^H, \end{aligned} \tag{5.22}$$

the GLRT is simply written as

$$\frac{|\tilde{\mathbf{s}}^H\mathbf{S}^{-1}\tilde{\mathbf{z}}|^2}{\tilde{\mathbf{s}}^H\mathbf{S}^{-1}\tilde{\mathbf{s}}\,(1+\tilde{\mathbf{z}}^H\mathbf{S}^{-1}\tilde{\mathbf{z}})} \underset{\mathbf{H}_0}{\overset{\mathbf{H}_1}{\gtrless}} \eta_o. \tag{5.23}$$

Note that the distributions of all random quantities in the GLRT expression are independent of true covariance matrix. The subspace transformation will be performed on the quantities in the GLRT expression given in equation (5.23) in the next section where the GLRT-RR test will be developed.

5.2 Subspace GLRT (GLRT-RR)

This section analyses a reduced rank version of the GLRT test which we refer to as GLRT-RR. With GLRT-RR, received vectors are pre-processed by a full column rank $N \times r$ matrix transformation $\mathbf{T}$ where $r \leq N$. Matrix $\mathbf{T}$ can be either a fixed transformation such as discrete Fourier transform (DFT) and discrete cosine transform (DCT) or a data dependent transformation such as Karhunen-Loeve transform (KLT).

By applying the subspace transformation to the quantities in the GLRT test given in equation (5.23), the GLRT-RR is written as

$$\frac{|(\tilde{\mathbf{s}}^H\mathbf{T})\,(\mathbf{T}^H\mathbf{S}\mathbf{T})^{-1}\,(\mathbf{T}^H\tilde{\mathbf{z}})|^2}{(\tilde{\mathbf{s}}^H\mathbf{T})\,(\mathbf{T}^H\mathbf{S}\mathbf{T})^{-1}\,(\mathbf{T}^H\tilde{\mathbf{s}})\,(1+(\tilde{\mathbf{z}}^H\mathbf{T})\,(\mathbf{T}^H\mathbf{S}\mathbf{T})^{-1}\,(\mathbf{T}^H\tilde{\mathbf{z}}))} \underset{\mathbf{H}_0}{\overset{\mathbf{H}_1}{\gtrless}} \eta_o. \tag{5.24}$$

By making the following definitions

$$\begin{aligned} \mathbf{t} &= \mathbf{T}^H \tilde{\mathbf{s}} \\ \mathbf{x} &= \mathbf{T}^H \tilde{\mathbf{z}} \\ \hat{\mathbf{\Sigma}} &= \mathbf{T}^H \mathbf{S} \mathbf{T}, \end{aligned} \tag{5.25}$$

the GLRT-RR is written as

$$\frac{|\mathbf{t}^H \hat{\mathbf{\Sigma}}^{-1} \mathbf{x}|^2}{\mathbf{t}^H \hat{\mathbf{\Sigma}}^{-1} \mathbf{t} \left(1 + \mathbf{x}^H \hat{\mathbf{\Sigma}}^{-1} \mathbf{x}\right)} \underset{\mathbf{H}_0}{\overset{\mathbf{H}_1}{\gtrless}} \eta_o. \tag{5.26}$$

The vectors $\mathbf{t}$ and $\mathbf{x}$ are of dimension $r \times 1$. The matrix $\hat{\mathbf{\Sigma}}$ is of dimension $r \times r$ and can be written using equation (5.22) as

$$\begin{aligned} \hat{\mathbf{\Sigma}} &= \sum_{k=1}^{K} \mathbf{T}^H \tilde{\mathbf{z}}(k)\, \tilde{\mathbf{z}}(k)^H \mathbf{T} \\ &= \sum_{k=1}^{K} \mathbf{x}(k)\, \mathbf{x}(k)^H \end{aligned} \tag{5.27}$$

where $\mathbf{x}(k)$ for $k = 1, \cdots, K$ are the transformed and whitened secondary vectors. The matrix $\mathbf{S}$ is distributed as $\mathcal{CW}(K, N; \mathbf{I})$. The expectation of matrix $\hat{\mathbf{\Sigma}}$ is

$$\begin{aligned} E\{\hat{\mathbf{\Sigma}}\} &= \mathbf{T}^H\, E\{\mathbf{S}\}\, \mathbf{T} \\ &= \mathbf{T}^H \mathbf{I} \mathbf{T} \\ &= \mathbf{I}. \end{aligned} \tag{5.28}$$

Thus, the matrix $\hat{\mathbf{\Sigma}}$ is distributed as $\mathcal{CW}(K, N; \mathbf{I})$.

The GLRT-RR in equation (5.26) is similar in form to Kelly's GLRT in equation (5.23) except that N is replaced by the reduced dimension r. Kelly analyzed the GLRT test and derived expressions for P_{fa} and P_D [10, 41, 42]. Following Kelly's results, under hypothesis $\mathbf{H}_0$, the test in equation (5.26) resembles a simple scalar CFAR test with P_{fa} given as

$$P_{fa} = \left(\frac{1}{\gamma}\right)^{K+1-r}, \tag{5.29}$$

where

$$\gamma = \frac{1}{1-\eta_o}. \tag{5.30}$$

The P_D for the GLRT [10] is found to be dependent a SNR expression which is given as

$$\text{SNR}_{eff} = \text{SNR}_{GLRT}\, r_l, \tag{5.31}$$

where r_l represents a loss factor due to estimating the covariance matrix and is bounded, $0 \leq r_l \leq 1$. The quantity SNR_{GLRT} represents the SNR obtained from

$$\begin{aligned} \text{SNR}_{GLRT} &= E\{\mathbf{z}\}^H \mathbf{R}^{-1} E\{\mathbf{z}\} \\ &= |b|^2\, \mathbf{s}^H \mathbf{R}^{-1} \mathbf{s} \end{aligned} \tag{5.32}$$

where $\mathbf{z}$ is the original primary vector, $\mathbf{s}$ is the original signal vector, $\mathbf{R}$ is the original true covariance matrix and b is the signal amplitude. By normalizing SNR_{eff} with $\mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}$, equation (5.31) can be written as

$$\text{SNR}_{eff} = |b|^2\, r_l. \tag{5.33}$$

However, by using subspace transformation, analogous to equation (5.32), the SNR obtained is

$$\begin{aligned} \text{SNR}_{GLRT-RR} &= E\{\mathbf{x}^H\} \boldsymbol{\Sigma}^{-1} E\{\mathbf{x}\} \\ &= |b|^2\, \mathbf{t}^H \boldsymbol{\Sigma}^{-1} \mathbf{t} \end{aligned} \tag{5.34}$$

where $\mathbf{x} = \mathbf{T}^H \mathbf{z}$, $\mathbf{t} = \mathbf{T}^H \mathbf{s}$, and $\boldsymbol{\Sigma} = \mathbf{T}^H \mathbf{R} \mathbf{T}$. Hence, for the GLRT-RR, the effective SNR is written as

$$\text{SNR-RR}_{eff} = |b|^2\, \mathbf{t}^H \boldsymbol{\Sigma}^{-1} \mathbf{t}\, r_l. \tag{5.35}$$

By performing a similar normalization as that done in equation (5.33), SNR-RR_{eff} is written as

$$\begin{aligned} \text{SNR-RR}_{eff} &= |b|^2\, r_l\, \frac{\mathbf{t}^H \boldsymbol{\Sigma}^{-1} \mathbf{t}}{\mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}} \\ &= |b|^2\, r_l\, r_b \end{aligned} \tag{5.36}$$

where the SNR bias, r_b, is defined as

$$r_b = \frac{\mathbf{t}^H \boldsymbol{\Sigma}^{-1} \mathbf{t}}{\mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}}. \tag{5.37}$$

For a $N \times r$ subspace transformation matrix $\mathbf{T}$ with $r \leq N$, we have $\mathbf{t}^H \boldsymbol{\Sigma}^{-1} \mathbf{t} \leq \mathbf{s}^H \mathbf{R}^{-1} \mathbf{s}$. Hence, the SNR bias takes on values $0 \leq r_b \leq 1$. This bias in the SNR is due to subspace transformation. Having defined the SNR expression, SNR_{eff}, as in equation (5.36), the P_D expression for a given value of r_l can be written as a function of SNR_{eff} as

$$P_{D/r_l} = 1 - \frac{1}{\gamma^{K+1-r}} \sum_{k=1}^{K+1-r} \left(\begin{array}{c} K+1-r \\ k \end{array} \right) (\gamma - 1)^k G_k \left(\frac{|b|^2 \, r_l \, r_b}{\gamma} \right) \tag{5.38}$$

where $G_k(y)$ is the incomplete Gamma function and is defined as

$$G_k(y) = e^{-y} \sum_{n=0}^{k-1} \frac{y^n}{n!}. \tag{5.39}$$

However, in order to get the final expression for P_D, the expression in equation (5.38) has to be averaged over the distribution of r_l. The loss factor r_l was derived in [10] for the GLRT and can be written for the GLRT-RR by replacing the vectors and matrices in the expression by their transformed counterparts and is given as

$$r_l = \frac{1}{1 + \Sigma_B} \tag{5.40}$$

where

$$\begin{aligned} \Sigma_B &= \mathbf{x}_B^H \hat{\boldsymbol{\Sigma}}_{BB}^{-1} \mathbf{x}_B \\ &= \mathbf{x}_B^H \Big(\sum_{k=1}^{K} \mathbf{x}(k)_B \, \mathbf{x}(k)_B^H \Big)^{-1} \mathbf{x}_B \end{aligned} \tag{5.41}$$

where $\mathbf{x}_B$, $\mathbf{x}(k)_B$ for $k = 1, \cdots, K$, and $\hat{\boldsymbol{\Sigma}}_{BB}$ are obtained from $\mathbf{x}$, $\mathbf{x}(k)$ for $k = 1, \cdots, K$, and $\hat{\boldsymbol{\Sigma}}$, respectively, by the following partitioning

$$\mathbf{x} = \left[\begin{array}{c} x_A \\ \mathbf{x}_B \end{array} \right] \tag{5.42}$$

$$\mathbf{x}(k) = \left[\begin{array}{c} x(k)_A \\ \mathbf{x}(\mathrm{k})_B \end{array} \right] \tag{5.43}$$

$$\hat{\boldsymbol{\Sigma}} = \left[\begin{array}{cc} \hat{\Sigma}_{AA} & \hat{\boldsymbol{\Sigma}}_{AB} \\ \hat{\boldsymbol{\Sigma}}_{BA} & \hat{\boldsymbol{\Sigma}}_{BB} \end{array} \right]. \tag{5.44}$$

All vectors appearing in equation (5.41) have complex Gaussian distributions and $\widehat{\Sigma}_{BB}$ has a complex Wishart distribution. As shown in Appendix B, the quantity Σ_B has a multivariate central F-distribution (Hotelling's T^2) distribution [35] given as

$$f_{\Sigma_B}(\Sigma_B) = \frac{\Gamma(K+1)}{\Gamma(r-1)\Gamma(K+2-r)} {\Sigma_B}^{r-2}(1-\Sigma_B)^{K+1}. \tag{5.45}$$

Then a simple change of variables yields the distribution of r_l as

$$f_{r_l}(r_l) = \frac{\Gamma(K+1)}{\Gamma(r-1)\Gamma(K+2-r)} \rho^{K+1-r}(1-\rho)^{r-2}. \tag{5.46}$$

The probability of detection P_D can then be expressed as

$$P_D = \int_0^1 P_{D/r_l} f_{r_l}(r_l)\, dr_l \tag{5.47}$$

giving the expression

$$P_D = 1 - \frac{1}{\gamma^{K+1-r}} \sum_{k=1}^{K+1-r} \binom{K+1-r}{k} (\gamma-1)^k H_k\left(\frac{|b|^2\, r_b}{\gamma}\right) \tag{5.48}$$

where

$$H_k(y) = \int_0^1 G_k(r_l y) f_{r_l}(r_l)\, dr_l. \tag{5.49}$$

If the desired signal is not present, where $|b|^2 = 0$, the expression for P_D gives back P_{fa}. This expression is similar to Kelly's P_D expression for the GLRT except that the subspace transformation rank r replaces the vector dimension N and the SNR bias ρ_b enters the expression. The expression for P_D is a function of SNR_{eff} which in turn is a function of the specific subspace transformation matrix $\mathbf{T}$ applied to the data. Hence, this expression can be used to evaluate the performance of the GLRT-RR using different transformations.

5.3 A Generalized Likelihood Ratio Test with Diagonal Loading (GLRT-LSMI)

This section will derive expressions for the probability of false alarm and probability of detection using the generalized likelihood ratio test that was defined by Kelly

in [10] under the condition that the sample covariance matrix is diagonally loaded (GLRT-LSMI) in order to improve the conditioning of the eigenvalues.

The maximum likelihood estimate of the covariance matrix is highly dependent on the size of the secondary data set K, with the condition that $K \geq N$ for the estimate to be non-singular. In order to improve the conditioning of the sample covariance matrix, loading is applied and the resulting estimate is

$$\widehat{\mathbf{R}}_d = \frac{1}{K}\sum_{k=1}^{K} \mathbf{z}(k)\mathbf{z}(k)^H + \alpha\mathbf{I}, \tag{5.50}$$

where the scalar α is the loading factor and $\mathbf{I}$ is the identity matrix. The subscript denotes diagonal loading. The GLRT-LSMI can then be written in a similar way to the GLRT in equation (5.10) as

$$\frac{|\mathbf{s}^H\mathbf{S}^{-1}\mathbf{z}|^2}{\mathbf{s}^H\mathbf{S}^{-1}\mathbf{s}\,(1+\mathbf{z}^H\mathbf{S}^{-1}\mathbf{z})} \begin{array}{c}\mathbf{H}_1 \\ > \\ < \\ \mathbf{H}_0\end{array} \eta_o \tag{5.51}$$

where η_o is a threshold that satisfies the false alarm rate P_{fa}. The matrix $\mathbf{S}$ is related to $\widehat{\mathbf{R}}_d$ as

$$\begin{aligned}\mathbf{S} &= K\,\widehat{\mathbf{R}}_d \\ &= \sum_{k=1}^{K} \mathbf{z}(k)\mathbf{z}(k)^H + K\alpha\mathbf{I}.\end{aligned} \tag{5.52}$$

This test will be shown to be CFAR where the P_{fa} will be shown to be independent of the true covariance matrix.

5.3.1 Analysis and Simplification of the Likelihood Ratio

In this section, the likelihood ratio of GLRT-LSMI in equation (5.51) is simplified so that it would be possible to find its distribution under the two hypotheses $\mathbf{H}_0$ and $\mathbf{H}_1$.

A new matrix $\mathbf{Z}$ is defined as a $N \times K$ matrix whose columns are made up of the statistically independent secondary data vectors $\mathbf{z}(k)$ for $k = 1, \cdots, K$. Under

hypothesis $\mathbf{H}_1$, $\mathbf{z}(k)$ for $k = 1, \cdots, K$ are distributed as $\mathcal{CN}(\mathbf{0}, \mathbf{R})$. Hence, $\mathbf{Z}$ is a matrix whose columns are distributed as $\mathcal{CN}(\mathbf{0}, \mathbf{R})$. Using this definition of matrix $\mathbf{Z}$, matrix $\mathbf{S}$ defined in equation (5.52) is rewritten as

$$\mathbf{S} = \mathbf{Z}\mathbf{Z}^H + K\alpha\mathbf{I}. \tag{5.53}$$

The true covariance matrix $\mathbf{R}$ is assumed to obey the low rank model

$$\mathbf{R} = \mathbf{Q}_r\mathbf{\Lambda}_r\mathbf{Q}_r^H + \sigma_v^2\mathbf{Q}_v\mathbf{Q}_v^H \tag{5.54}$$

where $\mathbf{Q}_r$ is the $N \times r$ matrix of principal eigenvectors, $\mathbf{\Lambda}_r$ is the $r \times r$ diagonal matrix of principal eigenvalues, σ_v^2 is the variance of the white noise, and $\mathbf{Q}_v$ is the $N \times p$ matrix of noise eigenvectors, where $p = N - r$. The inverse of $\mathbf{R}$ is give as

$$\mathbf{R}^{-1} = \mathbf{Q}_r\mathbf{\Lambda}_r^{-1}\mathbf{Q}_r^H + \sigma_v^{-2}\mathbf{Q}_v\mathbf{Q}_v^H. \tag{5.55}$$

As was the case in Chapter 4, the assumption of high interference power relative to the noise power is made such that

$$\gamma_{min}^{-1} \ll \sigma_v^{-2} \tag{5.56}$$

where γ_{min} in the smallest element of the diagonal matrix $\mathbf{\Lambda}_r$. Using this assumption on the interference power, the inverse of the covariance matrix can be written as

$$\mathbf{R}^{-1} \cong \sigma_v^{-2}\mathbf{Q}_v\mathbf{Q}_v^H \tag{5.57}$$

where $o(1/\gamma_l)$ for $l = 1, \cdots, r$ terms were neglected to obtain the last expression, where $o(\beta)$ refers to terms in the order of β. Using equation (5.57), $\mathbf{R}^{-1/2}$ is expressed as

$$\mathbf{R}^{-1/2} \cong \sigma_v^{-1}\mathbf{Q}_v\mathbf{Q}_v^H \tag{5.58}$$

under the same assumption of high interference power. Setting the noise power σ_v^2 equal to unity, equations (5.57) and (5.58) are written as

$$\mathbf{R}^{-1} \cong \mathbf{Q}_v\mathbf{Q}_v^H \tag{5.59}$$

and

$$\mathbf{R}^{-1/2} \cong \mathbf{Q}_v \mathbf{Q}_v^H. \tag{5.60}$$

The inverse of matrix $\mathbf{S}$ in equation (5.53) can be written as

$$\begin{aligned}
\mathbf{S}^{-1} &= (\mathbf{Z}\mathbf{Z}^H + K\alpha\mathbf{I})^{-1} \\
&= \mathbf{R}^{-1/2}\mathbf{R}^{1/2}(\mathbf{Z}\mathbf{Z}^H + K\alpha\mathbf{I})^{-1}\mathbf{R}^{1/2}\mathbf{R}^{-1/2} \\
&= \mathbf{R}^{-1/2}(\mathbf{R}^{-1/2}\mathbf{Z}\mathbf{Z}^H\mathbf{R}^{-1/2} + K\alpha\mathbf{R}^{-1})^{-1}\mathbf{R}^{-1/2} \\
&= \mathbf{R}^{-1/2}(\mathcal{Z}\mathcal{Z}^H + K\alpha\mathbf{R}^{-1})^{-1}\mathbf{R}^{-1/2}
\end{aligned} \tag{5.61}$$

where matrix $\mathcal{Z}$ is defined as

$$\mathcal{Z} = \mathbf{R}^{-1/2}\mathbf{Z} \tag{5.62}$$

which is a matrix whose columns are zero mean Gaussian vectors with a covariance matrix $\mathbf{I}$ and hence the columns are distributed as $\mathcal{CN}(\mathbf{0}, \mathbf{I})$ (Appendix A). Substituting equations (5.59) and (5.60) into equation (5.61) yields

$$\begin{aligned}
\mathbf{S}^{-1} &\cong \mathbf{Q}_v\mathbf{Q}_v^H(\mathcal{Z}\mathcal{Z}^H + K\alpha\mathbf{Q}_v\mathbf{Q}_v^H)^{-1}\mathbf{Q}_v\mathbf{Q}_v^H \\
&= \frac{1}{K\alpha}\mathbf{Q}_v\mathbf{Q}_v^H(\mathbf{Q}_v\mathbf{Q}_v^H + \frac{1}{K\alpha}\mathcal{Z}\mathcal{Z}^H)^{-1}\mathbf{Q}_v\mathbf{Q}_v^H \\
&= \frac{1}{K\alpha}\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{C}\mathbf{Q}_v\mathbf{Q}_v^H
\end{aligned} \tag{5.63}$$

where the matrix $\mathbf{C}$ is defined as

$$\mathbf{C} = (\mathbf{Q}_v\mathbf{Q}_v^H + \frac{1}{K\alpha}\mathcal{Z}\mathcal{Z}^H)^{-1}. \tag{5.64}$$

Substituting equation (5.63) into the GLRT-LSMI expression in equation (5.51), the test can be written as

$$\frac{|\frac{1}{K\alpha}\mathbf{s}^H\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{C}\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{z}|^2}{\frac{1}{K\alpha}\mathbf{s}^H\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{C}\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{s}\left(1 + \frac{1}{K\alpha}\mathbf{z}^H\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{C}\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{z}\right)} \begin{array}{c} \mathbf{H}_1 \\ > \\ < \\ \mathbf{H}_0 \end{array} \eta_o. \tag{5.65}$$

To further simplify the GLRT-LSMI expression, an expression for the quantity $\mathbf{Q}_v^H\mathbf{C}\mathbf{Q}_v$ using the definition of matrix $\mathbf{C}$ is obtained next. From equation (5.64)

and by pre-multiplying with $\mathbf{C}^{-1}$, we get

$$\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{C} + \frac{1}{K\alpha}\mathcal{Z}\mathcal{Z}^H\mathbf{C} = \mathbf{I}. \tag{5.66}$$

Using the fact that for the two orthogonal subspaces $\mathbf{Q}_r$ and $\mathbf{Q}_v$, it is true that $\mathbf{Q}_v\mathbf{Q}_v^H + \mathbf{Q}_r\mathbf{Q}_r^H = \mathbf{I}$, equation (5.66) can be written as

$$\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{C} + \frac{1}{K\alpha}\mathcal{Z}\mathcal{Z}^H\left(\mathbf{Q}_v\mathbf{Q}_v^H + \mathbf{Q}_r\mathbf{Q}_r^H\right)\mathbf{C} = \mathbf{I} \tag{5.67}$$

and equivalently

$$\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{C} + \frac{1}{K\alpha}\mathcal{Z}\mathcal{Z}^H\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{C} + \frac{1}{K\alpha}\mathcal{Z}\mathcal{Z}^H\mathbf{Q}_r\mathbf{Q}_r^H\mathbf{C} = \mathbf{I}. \tag{5.68}$$

By pre-multiplying equation (5.68) by $\mathbf{Q}_r^H$ and $\mathbf{Q}_v^H$, and by using the orthogonality between the two subspaces, the following two equations are obtained

$$\begin{aligned}
\frac{1}{K\alpha}\mathbf{Q}_r^H\mathcal{Z}\mathcal{Z}^H\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{C} + \frac{1}{K\alpha}\mathbf{Q}_r^H\mathcal{Z}\mathcal{Z}^H\mathbf{Q}_r\mathbf{Q}_r^H\mathbf{C} &= \mathbf{Q}_r^H \\
\mathbf{Q}_v^H\mathbf{C} + \frac{1}{K\alpha}\mathbf{Q}_v^H\mathcal{Z}\mathcal{Z}^H\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{C} + \frac{1}{K\alpha}\mathbf{Q}_v^H\mathcal{Z}\mathcal{Z}^H\mathbf{Q}_r\mathbf{Q}_r^H\mathbf{C} &= \mathbf{Q}_v^H.
\end{aligned} \tag{5.69}$$

Solving these two equations simultaneously for $\mathbf{Q}_v^H\mathbf{C}$, the following equation is obtained

$$\mathbf{Q}_v^H\mathbf{C} = \left(\mathbf{I} + \frac{1}{K\alpha}\mathbf{A}^H[I - \mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H]\mathbf{A}\right)^{-1}\left(\mathbf{Q}_v^H - \mathbf{A}^H\mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{Q}_r^H\right) \tag{5.70}$$

where matrices $\mathbf{A}$ and $\mathbf{B}$ are of sizes $K \times (N-r)$ and $K \times r$, respectively. They are given as

$$\begin{aligned}
\mathbf{A} &= \mathcal{Z}^H\mathbf{Q}_v \\
\mathbf{B} &= \mathcal{Z}^H\mathbf{Q}_r.
\end{aligned} \tag{5.71}$$

Since the columns of matrix $\mathcal{Z}$ are distributed as $\mathcal{CN}(\mathbf{0},\mathbf{I})$ and since the two subspaces $\mathbf{Q}_v$ and $\mathbf{Q}_r$ are orthogonal, then the two matrices $\mathbf{A}$ and $\mathbf{B}$ are statistically independent. Moreover, the columns of each matrix $\mathbf{A}$ and $\mathbf{B}$ are independent and distributed as $\mathcal{CN}(\mathbf{0},\mathbf{I})$ (see Appendix A).

The expression $[I - \mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H]$ represents an idempotent matrix [34], and hence can be expressed as

$$[\mathbf{I} - \mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{B}^H] = \mathbf{D}\mathbf{D}^H \tag{5.72}$$

where matrix $\mathbf{D}$ is of dimension $K \times (K - r)$ and is orthogonal to the matrix $\mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1/2}$ where $\mathbf{D}^H\mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1/2} = \mathbf{0}$. Using the definition in equation (5.72), equation (5.70) is written as

$$\mathbf{Q}_v^H\mathbf{C} = \left(\mathbf{I} + \frac{1}{K\alpha}\mathbf{A}^H\mathbf{D}\mathbf{D}^H\mathbf{A}\right)^{-1}\left(\mathbf{Q}_v^H - \mathbf{A}^H\mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{Q}_r^H\right). \tag{5.73}$$

For a large loading factor α relative to the unity elements of the identity matrix appearing in right hand side of equation (5.73), where

$$\frac{1}{K\alpha} \ll 1, \tag{5.74}$$

equation (5.73) can be written as

$$\mathbf{Q}_v^H\mathbf{C} \cong \left(\mathbf{Q}_v^H - \mathbf{A}^H\mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{Q}_r^H\right) \tag{5.75}$$

where $o(1/K\alpha)$ terms were neglected to obtain the last equation. Using the expression for $\mathbf{Q}_v^H\mathbf{C}$ and recalling the GLRT-LSMI test expression in equation (5.65), the matrix $\mathbf{Q}_v^H\mathbf{C}\mathbf{Q}_v$ is written as

$$\begin{aligned}\mathbf{Q}_v^H\mathbf{C}\mathbf{Q}_v &= \left(\mathbf{Q}_v^H\mathbf{C}\right)\mathbf{Q}_v \\ &= \left(\mathbf{Q}_v^H - \mathbf{A}^H\mathbf{B}(\mathbf{B}^H\mathbf{B})^{-1}\mathbf{Q}_r^H\right)\mathbf{Q}_v \\ &= \mathbf{I}\end{aligned} \tag{5.76}$$

since the two subspaces $\mathbf{Q}_r$ and $\mathbf{Q}_v$ are orthogonal. Substituting equation (5.76) into equation (5.65), the GLRT-LSMI is written as

$$\frac{|\frac{1}{K\alpha}\mathbf{s}^H\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{z}|^2}{\frac{1}{K\alpha}\mathbf{s}^H\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{s}\left(1 + \frac{1}{K\alpha}\mathbf{z}^H\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{z}\right)} \begin{matrix}\mathbf{H}_1 \\ > \\ < \\ \mathbf{H}_0\end{matrix} \eta_o. \tag{5.77}$$

To further simplify this expression, the $(N-r)\times 1$ deterministic vector $\mathbf{t}$ is defined as

$$\mathbf{t} = \mathbf{Q}_v^H \mathbf{s}. \tag{5.78}$$

In addition, the following definition is made

$$\begin{aligned} \mathbf{t}^H\mathbf{t} &= \mathbf{s}^H\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{s} \\ &\equiv 1. \end{aligned} \tag{5.79}$$

This definition will be useful later. Using this definition of vector $\mathbf{t}$, the GLRT-LSMI is written as

$$\frac{|\frac{1}{K\alpha}\mathbf{t}^H\mathbf{Q}_v^H\mathbf{z}|^2}{\frac{1}{K\alpha}\left(1+\frac{1}{K\alpha}\mathbf{z}^H\mathbf{Q}_v\mathbf{Q}_v^H\mathbf{z}\right)} \underset{\mathbf{H}_0}{\overset{\mathbf{H}_1}{\gtrless}} \eta_o. \tag{5.80}$$

The normalized signal vector $\mathbf{t}$ is then replaced by $\tilde{\mathbf{t}}$ where

$$\tilde{\mathbf{t}} = \mathbf{U}\mathbf{t} \tag{5.81}$$

where the matrix $\mathbf{U}$ is a unitary matrix that transforms the vector $\mathbf{t}$ into $\tilde{\mathbf{t}}$. By choosing a matrix $\mathbf{U}$ whose first row is the conjugate transpose (hermitian) of vector $\mathbf{t}$, and whose other rows are the conjugate transpose of unit vectors orthogonal to $\mathbf{t}$, the new signal vector $\mathbf{t}$ can be chosen to be a vector whose first element is unity and all other elements are zero. The vector $\tilde{\mathbf{t}}$ is written as

$$\tilde{\mathbf{t}} = \begin{bmatrix} 1 \\ \mathbf{0} \end{bmatrix}. \tag{5.82}$$

The $N\times 1$ primary vector $\mathbf{z}$ in the test statistic in equation (5.80) is partitioned in a similar way to vector $\tilde{\mathbf{t}}$ and is written as

$$\mathbf{z} = \begin{bmatrix} z_A \\ \mathbf{z}_B \end{bmatrix} \tag{5.83}$$

where z_A is a scalar and $\mathbf{z}_B$ is a $(N-1)\times 1$ vector. Using these definitions in (5.82) and (5.83), the GLRT-LSMI is written as

$$\frac{\left|\frac{1}{K\alpha}x_A\right|^2}{\frac{1}{K\alpha}\left(1+\frac{1}{K\alpha}\mathbf{x}^H\mathbf{x}\right)} \underset{\mathbf{H}_0}{\overset{\mathbf{H}_1}{\gtrless}} \eta_o \tag{5.84}$$

where the vector $\mathbf{x}$ is defined in terms of the primary vector $\mathbf{z}$ as

$$\begin{aligned}\mathbf{x} &= \mathbf{Q}_v^H\mathbf{z}\\ &= \begin{bmatrix} x_A \\ \mathbf{x}_B \end{bmatrix}.\end{aligned} \tag{5.85}$$

Under hypothesis $\mathbf{H}_0$, the original primary vector $\mathbf{z}$ is distributed as $\mathcal{CN}(\mathbf{0},\mathbf{R})$, hence vector $\mathbf{x}$ is distributed as $\mathcal{CN}(\mathbf{0},\mathbf{Q}_v^H\mathbf{R}\mathbf{Q}_v)$ or equivalently $\mathcal{CN}(\mathbf{0},\mathbf{I})$ where $\mathbf{R}$ is defined in terms of the subspaces in equation (5.54). Under hypothesis $\mathbf{H}_1$, the original primary vector $\mathbf{z}$ is distributed as $\mathcal{CN}(b\mathbf{s},\mathbf{I})$ and hence, vector $\mathbf{x}$ is distributed as $\mathcal{CN}(b\tilde{\mathbf{t}},\mathbf{I})$ where $\mathbf{t}$ is defined in equation (5.82).

The test ratio can be further simplified by noticing that

$$\mathbf{x}^H\mathbf{x} = |x_A|^2 + \mathbf{x}_B^H\mathbf{x}_B. \tag{5.86}$$

Using this last equality and by re-arranging the ratio in equation (5.84), the GLRT-LSMI can be finally written as

$$\frac{|x_A|^2}{K\alpha+\mathbf{x}_B^H\mathbf{x}_B} \underset{\mathbf{H}_0}{\overset{\mathbf{H}_1}{\gtrless}} \gamma \tag{5.87}$$

where

$$\gamma = \frac{\eta_0}{1-\eta_0}. \tag{5.88}$$

As was already shown, the elements of vector $\mathbf{x}$ are independent under both hypotheses. This is important because under this condition, the numerator and denominator of the GLRT-LSMI in equation (5.87) are independent. This makes

the statistical properties of the test easier to analyze. In the next section, the test statistic is analyzed under both hypotheses and expressions for P_{fa} and P_D are derived.

5.3.2 Derivation of Probabilities of False Alarm and Detection

Define the two random variables ζ and ξ as

$$\zeta = |x_A|^2 \tag{5.89}$$

$$\xi = \mathbf{x}_B^H \mathbf{x}_B \tag{5.90}$$

where the GLRT-LSMI in equation (5.87) is written as

$$\eta = \frac{\zeta}{K\alpha + \xi} \underset{\mathbf{H}_0}{\overset{\mathbf{H}_1}{\gtrless}} \gamma. \tag{5.91}$$

In order to obtain the distribution of η, the random variables ζ and ξ are characterized first. The distribution of η is considered under each hypothesis separately. Under hypothesis $\mathbf{H}_0$, the $(N-r) \times 1$ vector $\mathbf{x}$ is distributed as $\mathcal{CN}(\mathbf{0}, \mathbf{I})$, and hence, the random variable x_A is distributed as $\mathcal{CN}(0,1)$ and the $(N-r-1) \times 1$ vector $\mathbf{x}_B$ as $\mathcal{CN}(\mathbf{0}, \mathbf{I})$. The random variable ζ in equation (5.89) then has a central chi-squared distribution with two degrees of freedom given as

$$f_\zeta(\zeta) = e^{-\zeta}. \tag{5.92}$$

The random variable ξ in equation (5.90) has a central chi-squared distribution with $2(N-r-1)$ degrees of freedom given as

$$f_\xi(\xi) = \frac{\xi^{N-r-2}}{\Gamma(N-r-1)} e^{-\xi}. \tag{5.93}$$

Since ζ and ξ are independent, the distribution of η under hypothesis $\mathbf{H}_0$ can be found (see Appendix C) and is given as

$$f_\eta(\eta|\mathbf{H}_0) = [N-r-1+K\alpha(1+\eta)]\,(1+\eta)^{r-N} e^{-K\alpha\eta}. \tag{5.94}$$

The probability of false alarm can then found by integrating the distribution of the test statistic under hypothesis $\mathbf{H}_0$ given in equation (5.94) above the threshold γ, i.e.,

$$\begin{aligned} P_{fa} &= \int_{\gamma}^{\infty} f_{\eta}(u|\mathbf{H}_0)du \\ &= \int_{\gamma}^{\infty} [N-r-1+K\alpha(1+u)](1+u)^{r-N}e^{-K\alpha u}du \end{aligned} \tag{5.95}$$

where u is an integration variable. This integral can be solved in closed form and P_{fa} can be written as

$$P_{fa} = (1+\gamma)^{1-N+r}e^{-K\alpha\gamma}. \tag{5.96}$$

The probability of false alarm expression is a function of dimensional parameters as well as the loading factor. It is not a function of the covariance matrix. Hence, the GLRT-LSMI in equation (5.87) represents a CFAR test.

At this point, the GLRT-LSMI statistic in equation (5.87) is considered under hypothesis $\mathbf{H}_1$ where a signal is assumed to be present in order to find an expression for the probability of detection P_D.

The random variables ζ and ξ are defined as in equations (5.89) and (5.90). Under hypothesis $\mathbf{H}_1$, the $(N-r)\times 1$ vector $\mathbf{x}$ is distributed as $\mathcal{CN}(b\tilde{\mathbf{t}},\mathbf{I})$. Hence the random variable x_A is distributed as $\mathcal{CN}(b,1)$ and the $(N-r-1)\times 1$ vector $\mathbf{x}_B$ is distributed as $\mathcal{CN}(\mathbf{0},\mathbf{I})$ due to the definition of the signal vector $\tilde{\mathbf{t}}$ in equation (5.82) . The random variable ζ then has a non-central chi-squared distribution with two degrees of freedom given as

$$f_{\zeta}(\zeta) = e^{-\delta_1}{}_0F_1(1;\delta_1\zeta)e^{-\zeta}, \tag{5.97}$$

where δ_1 is its non-centrality parameter given as

$$\delta_1 = |b|^2. \tag{5.98}$$

The random variable ξ has a central chi-squared distribution with $2(N-r-1)$ degrees of freedom given as

$$f_\xi(\xi) = \frac{\xi^{N-r-2}}{\Gamma(N-r-1)} e^{-\xi}. \tag{5.99}$$

The generalized hypergeometric function ${}_0F_1$ is expressed as [31]

$${}_0F_1(;b;x) = \sum_{j=0}^{\infty} \frac{\Gamma(b+j)}{\Gamma(b)} \frac{x^j}{j!} \tag{5.100}$$

where $\Gamma(n+1) = n!$.

Using the distributions of ζ and ξ in equations (5.97) and (5.99), respectively, and recalling that the GLRT-LSMI is given as

$$\eta = \frac{\zeta}{K\alpha + \xi} \underset{\mathbf{H}_0}{\overset{\mathbf{H}_1}{\gtrless}} \gamma, \tag{5.101}$$

the distribution of η can be found (see Appendix D) and is given as

$$f_\eta(\eta|\mathbf{H}_1) = \frac{e^{-\delta_1}}{\Gamma(N-r-1)} \int_{-\infty}^{\infty} |u| \; {}_0F_1(1;\delta_1\eta u) e^{-\eta u} \; (u - K\alpha)^{N-r-2} \; e^{-(u-K\alpha)} du. \tag{5.102}$$

The probability of detection can then be found by integrating the distribution of the test statistic under hypothesis $\mathbf{H}_1$ given in equation (5.102) above the threshold γ, i.e.,

$$P_D = \int_\gamma^\infty f_\eta(u|\mathbf{H}_1) \, du \tag{5.103}$$

where the appropriate threshold γ is set using the P_{fa} expression given in equation (5.96). The integral in equation (5.103) cannot be solved in closed form and hence is numerically evaluated to give P_D. The distributions of the test statistic under both hypotheses are verified by simulations and are shown in Figures 5.1 and 5.2. The figures were obtained using the following parameters: $N = 32$, $K = 64$, $r = 4$, and a diagonal loading factor of 10 dB. In Figure 5.2, the SNR was set to 10 dB.

5.4 Numerical Results and Discussion

The analysis carried out in this chapter aimed at investigating two modified forms of the generalized likelihood ratio test (Kelly's GLRT) and at deriving expressions for their respective probabilities of false alarm and detection. The two detectors are the GLRT-RR and GLRT-LSMI. Both of these detectors were shown to be CFAR where their probability of false alarm expressions are not functions of the unknown covariance matrix. Kelly's GLRT detector was shown to be highly dependent on the size of the sample support used to estimate the unknown covariance matrix. The detector suffers a SNR loss due to this estimation. The need to reduce this SNR loss is the motivation behind a subspace-based GLRT; GLRT-RR, and a diagonally loaded sample covariance matrix GLRT; GLRT-LSMI. The former detector, because of the reduced dimension, requires substantially less SNR for an equivalent performance when compared with the GLRT. There is a penalty however. The GLRT-RR suffers an additional transformation dependent loss (bias) in the SNR as a direct result on the subspace transformation. This SNR bias can be controlled and minimized by designing optimum subspace transformations.

Another direction that was taken to reduce the SNR loss due to estimation is the use of diagonal loading of the covariance matrix. As was explained in Chapter 4, the conditioning of the eigenvalues of the estimated covariance matrix is improved by applying diagonal loading. Widely spread eigenvalues are equivalent to having high sidelobes in the antenna pattern. Hence, diagonal loading emulates the effect of using a larger size of sample support and hence improves the performance of the sample support dependent GLRT.

The performance of the two detectors is shown using the probability of detection. The results are compared with the optimum matched filter (MF) where the covariance matrix is known *a prior* and with the CFAR version of the SMI detector (AMF). The MF and AMF detectors should provide the upper and lower

bounds for most of the SNR interval. The MF decision statistic is

$$\eta = \left|\mathrm{w}^H \mathrm{x}(k)\right|^2 \underset{\mathbf{H}_0}{\overset{\mathbf{H}_1}{>}} \gamma. \tag{5.104}$$

The optimum weight vector is given as $\mathbf{w} = \mathbf{R}^{-1}\mathbf{s}$. Hence, the decision statistic can be written as

$$\eta = \left|\mathrm{s}^H \mathbf{R}^{-1} \mathrm{z}\right|^2 \underset{\mathbf{H}_0}{\overset{\mathbf{H}_1}{>}} \gamma. \tag{5.105}$$

Under hypothesis $\mathbf{H}_0$, the primary vector $\mathbf{z}$ is distributed as $\mathcal{CN}\ (\mathbf{0}, \mathbf{R})$ and hence the quantity $\mathrm{s}^H\mathbf{R}^{-1}\mathrm{z}$ is distributed as $\mathcal{CN}\ (0, \mathrm{s}^H\mathbf{R}^{-1}\mathrm{s})$. By normalizing the steering vector such that $\mathrm{s}^H\mathbf{R}^{-1}\mathrm{s} \equiv 1$, the test quantity $\mathrm{s}^H\mathbf{R}^{-1}\mathrm{z}$ is distributed as $\mathcal{CN}\ (0, 1)$. Hence, η will have a central chi-squared distribution with two degrees of freedom and P_{fa} will be given as

$$P_{fa} = e^{-\gamma}. \tag{5.106}$$

Under hypothesis $\mathbf{H}_1$, η will have a non-central chi-squared distribution with two degrees of freedom and P_D will be given as

$$P_D = e^{-|b|^2} \int_{\gamma}^{\infty} {}_0F_1(1, |b|^2 u) e^{-u}\, du. \tag{5.107}$$

The P_{fa} and P_D expressions for the AMF detector are given in [37] (equations 32 and 37). The model assumes a complex Gaussian primary and secondary vectors of size $N = 32$. The detectors will be analyzed for different sample support sizes K. The clutter is modeled to have a complex-valued Gaussian distribution, with zero mean, and variance equal to the clutter-to-noise ratio (CNR) of 10 dB. The clutter returns are assumed to be uncorrelated with each other and also uncorrelated between snapshots. The interference subspace has a rank of 4. Probability of false alarm is set to 10^{-5} for all detectors.

The P_D is plotted in Figure 5.3 for Kelly's GLRT using a sample support sizes of $K = 40$, 64, and 100 using the P_D expression in [10]. It is shown that when K

becomes close to N, the GLRT performance quickly deteriorates. The performance of the GLRT is highly dependent on K. For comparison, Figures 5.4 and 5.5 show the same plot for GLRT-RR and GLRT-LSMI. In Figure 5.4, a DCT transformation matrix of size $r \times N$ with $r = 4$ is applied to the data prior to detection and P_D is plotted using equation (5.48). It is shown that even when K gets close to N, substantially less SNR is required for an equivalent performance to the GLRT. Figure 5.5 shows a similar result for the GLRT-LSMI. In this figure, equation (5.103) is used to plot the P_D for a large diagonal loading factor. The GLRT-LSMI behaves similar to the GLRT-RR. For close comparison, Figures 5.6, 5.7, and 5.8 show the P_D curves for the AMF, GLRT, GLRT-RR using a DFT transformation of rank 4, GLRT-RR using a DCT transformation of rank 4, GLRT-LSMI with α=15 dB, and MF for $K = 40,\ 64$, and 100, respectively. The P_D for the AMF was plotted using the expression in [37]. As shown in the figures, the MF is the upper performance bound for all the detectors since the covariance matrix is known *a priori*. As K becomes very large, all the detectors approach the MF. However, since the covariance matrix is unknown and K is limited, the GLRT is not the optimal solution and is outperformed by sub-optimal solutions such as the GLRT-RR and GLRT-LSMI.

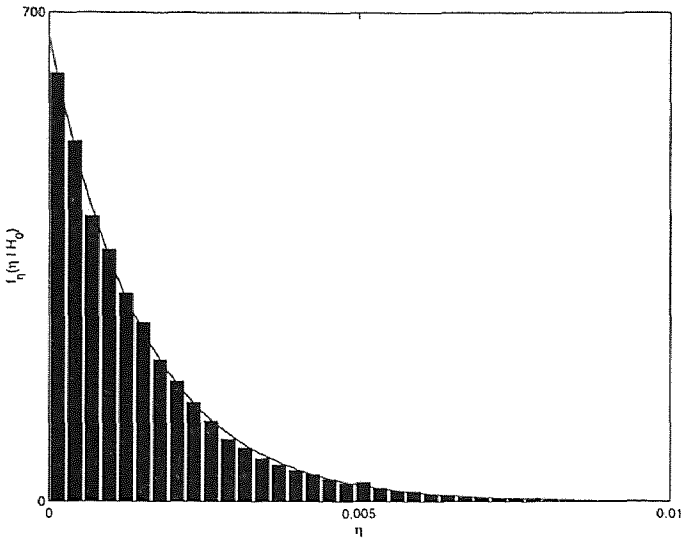

Figure 5.1 Distribution of the GLRT-LSMI test statistic under hypothesis $\mathbf{H}_0$.

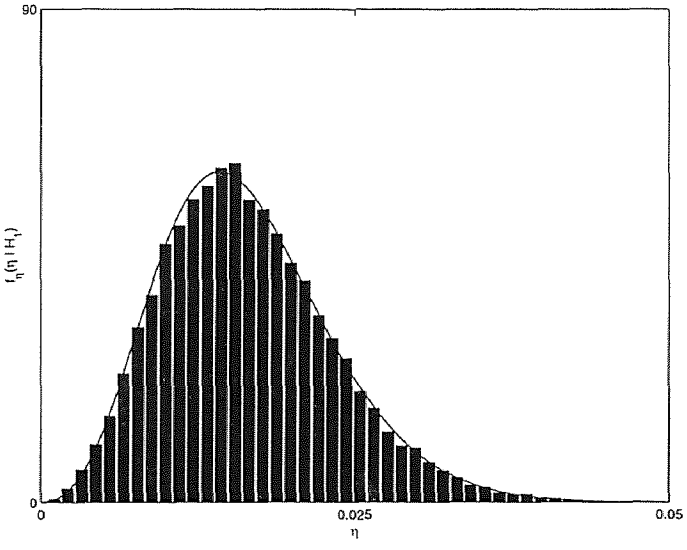

Figure 5.2 Distribution of the GLRT-LSMI test statistic under hypothesis $\mathbf{H}_1$.

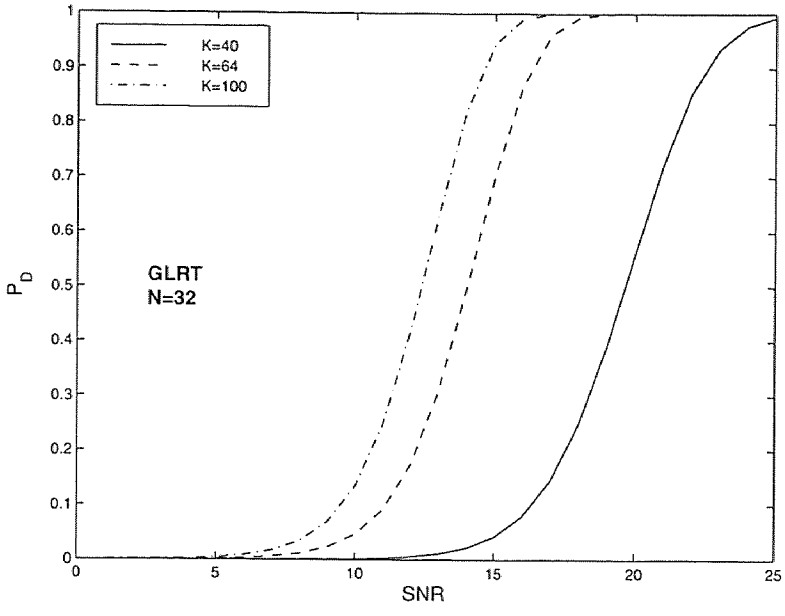

Figure 5.3 Probability of detection curves for Kelly's GLRT detector for different sample support K.

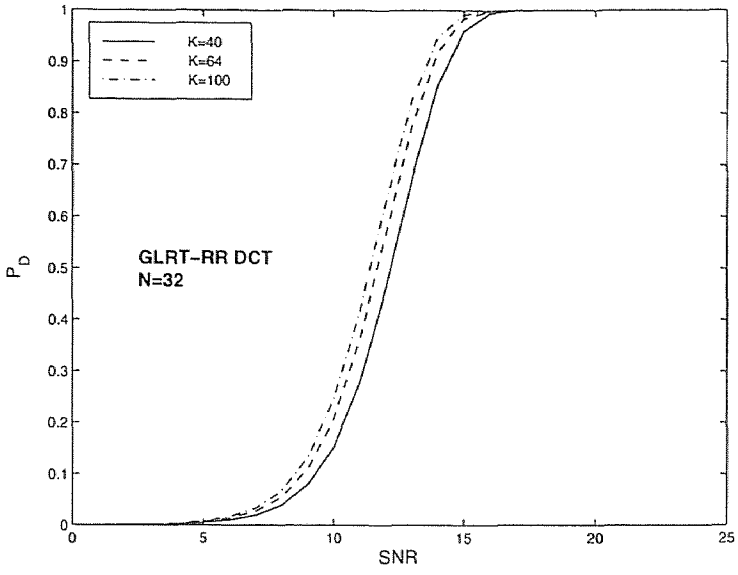

Figure 5.4 Probability of detection curves for GLRT-RR with DCT for different sample support K.

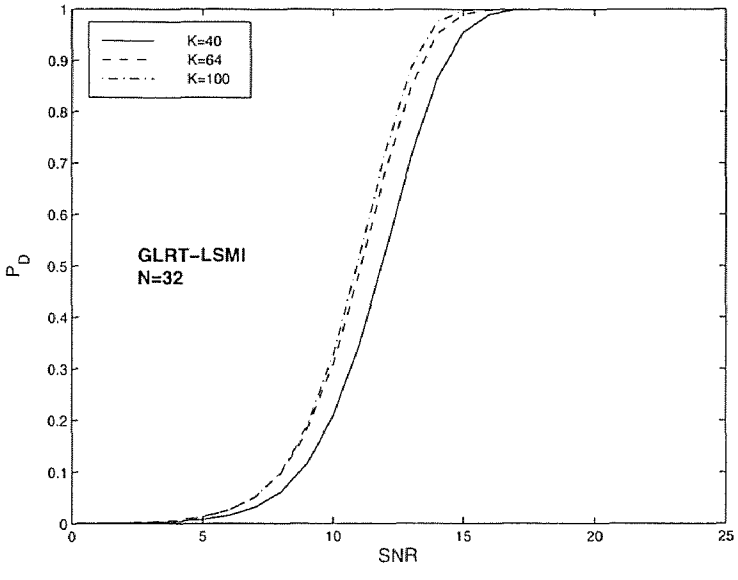

Figure 5.5 Probability of detection curves for GLRT-LSMI for different sample support K.

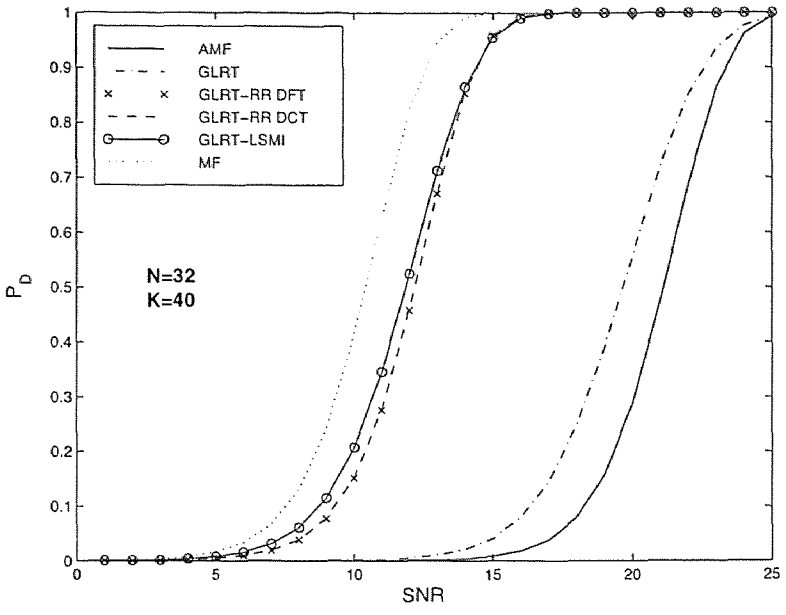

Figure 5.6 Comparison of methods for N=32 and K=40.

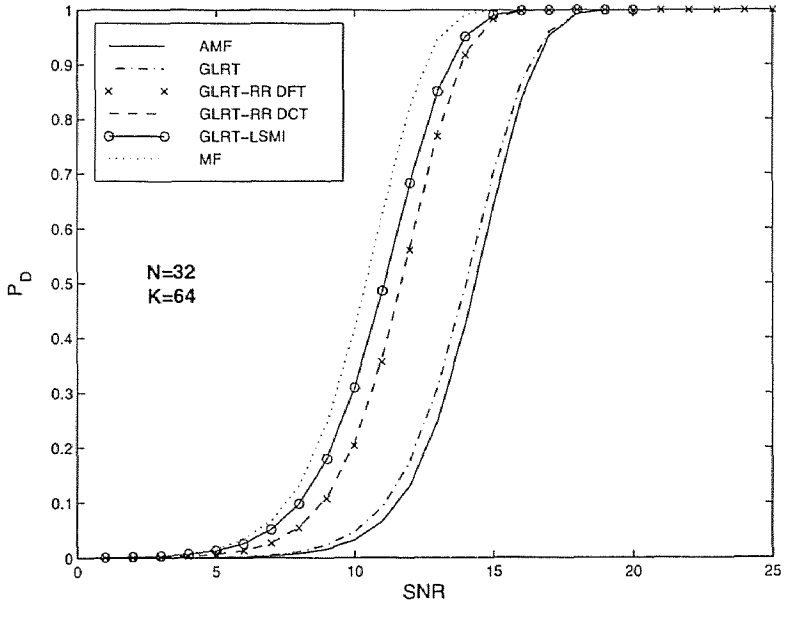

Figure 5.7 Comparison of methods for N=32 and K=64.

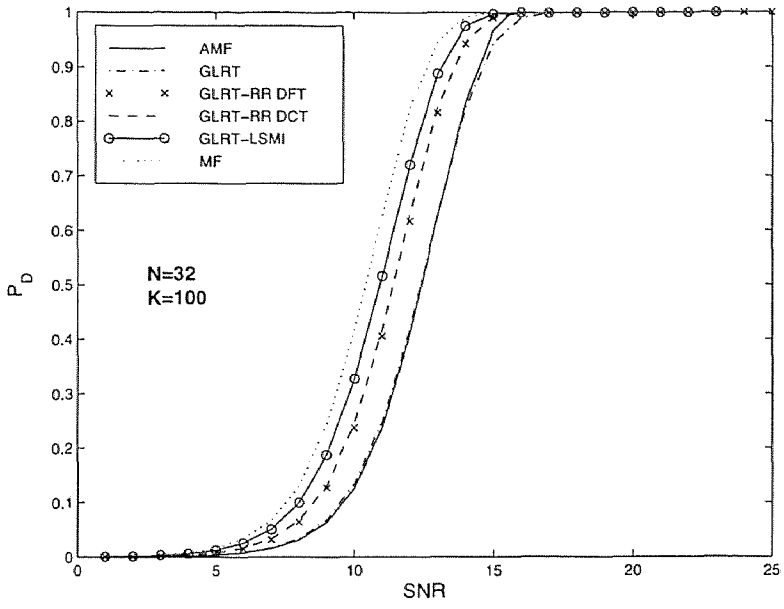

Figure 5.8 Comparison of methods for N=32 and K=100.

CHAPTER 6
CONCLUSIONS

In this dissertation, the high pulse repetition frequency (HPRF) airborne radar problem was analyzed and the application of innovative reduced-rank space time adaptive processing (STAP) techniques for clutter suppression in radar systems employing a HPRF waveform was investigated and discussed. The performance of the diagonally loaded sample matrix inversion (L-SMI) technique for clutter suppression in an adaptive array was derived and analyzed. Also, two constant false alarm rate (CFAR) detectors based on the generalized likelihood ratio (GLR) principle for signal detection in noise with unknown covariance matrix were derived and their performance was investigated.

In Chapter 3, the HPRF airborne radar problem was investigated. The application of STAP techniques to suppress the interference effects of ground clutter in HPRF radar were discussed. It was shown that reduced-rank STAP techniques are necessary due to reduced sample support resulting in a HPRF radar environment due to range ambiguous nature of the HPRF waveform. Three STAP techniques were applied to the clutter suppressing adaptive array and were compared; pseudoinverse sample matrix inversion (P-SMI), L-SMI, and eigencanceler. The three techniques were compared through the performance measures of conditioned signal-to-colored noise ratio (CSNR) and the probability of detection. It was shown that the eigencanceler along with the L-SMI techniques, both members of the class of reduced-rank techniques, perform well in clutter suppression where as P-SMI does not perform satisfactorily. It was also shown that due to the high output CSNR, reduced-rank techniques show better robustness against calibration errors in practical systems. Reduced-rank techniques are well suited to be used for clutter suppression in radar systems employing the HPRF waveform because of the reduced sample support where they require less degrees of freedom.

In Chapter 4, one of the previously mentioned STAP techniques, L-SMI, was closely investigated. An expression for the probability distribution of the CSNR for this technique was derived. The CSNR is an effective performance measure where it shows the effectiveness of the array in suppressing clutter and interferences. Through the distribution of the CSNR, the L-SMI technique was shown to outperform the traditional sample matrix inversion (SMI) technique for low rank interference. It was shown that L-SMI requires a substantially smaller size of size of sample support for an equivalent performance as that of the SMI. Hence, although a full rank technique that requires the inversion of the full size covariance matrix, L-SMI performs similar to reduced-rank techniques in terms of the required sample support. Diagonal loading was shown to enhance the conditioning of the estimated covariance matrix by reducing the spread of the noise eigenvalues of the matrix.

In Chapter 5, two GLR detectors exhibiting the desirable property of being CFAR were derived and analyzed. The first detector is a subspace-based GLR detector based on Kelly's generalized likelihood ratio test (GLRT) detector. It was shown that, because of the reduced dimension, the subspace-based detector requires substantially less signal-to-noise ratio (SNR) for an equivalent performance when compared with the GLRT. As a trade off for this gain in SNR, the subspace-based detector suffers an additional subspace transformation dependent loss in the SNR which is called a SNR bias. An expression for the probability of detection was derived and was shown to be a function of the specific subspace transformation used. This detector was shown to outperform the GLRT detector and to be less dependent on the size of the sample support. The second detector is a diagonally loaded sample matrix inversion GLR detector. This detector is a modification of the GLRT where the sample covariance matrix is diagonally loaded prior to detection. The resulting CFAR detector was shown to also require substantially less SNR for an equivalent performance when compared with the GLRT. This is due to the fact that

the diagonal loading improves the conditioning of the sample covariance matrix, a criterion on which the performance of the GLRT is highly dependent. This detector was shown to outperform the GLRT and to also be less dependent on the size of the sample support similar to the subspace-based GLR detector. Both detectors, the subspace-based GLR detector and the diagonally loaded sample covariance matrix GLR detector, outperform the GLRT for as long as the covariance matrix is unknown and has to be estimated from a limited number of samples.

Most of the work in this dissertation is based on investigating the application and performance of reduced-rank techniques for clutter and interference cancellation and signal detection. It was shown that when the covariance matrix is unknown, sub-optimal adaptive array STAP techniques such as the reduced-rank techniques, outperform classical techniques. Reduced-rank techniques were shown to be well-suited for applications with limited sample support. Moreover, these techniques result in reduced system's complexity. Finally, desired signal detectors show much improved performance when based on reduced-rank techniques.

APPENDIX A

SOME PROPERTIES OF MULTIVARIATE NORMAL DISTRIBUTION

Let $\mathbf{x}$ be a N-dimensional vector of independent complex normal random variables with mean $\mathrm{E}\{\mathbf{x}\} = \overline{\mathbf{x}}$ and covariance matrix $\mathbf{\Sigma}$. Hence the vector $\mathbf{x}$ is distributed as $\mathcal{CN}\ (\overline{\mathbf{x}}, \mathbf{\Sigma})$. Then for any $K \times N$ matrix $\mathbf{C}$, the matrix $\mathbf{Cx}$ will have the following distribution [43]

$$\mathbf{Cx} \sim \mathcal{CN}(\mathbf{C}\overline{\mathbf{x}}, \mathbf{C}\mathbf{\Sigma}\mathbf{C}^{\mathrm{H}}). \tag{A.1}$$

A.1 Whitening Transformation

Consider a matrix $\mathbf{X}$ of dimension $N \times K$ with its columns made up of statistically independent N-dimensional complex normally distributed vectors. The columns of matrix $\mathbf{X}$ are $\mathbf{x}_k$ for $k = 1, \cdots, K$ and are distributed as $\mathcal{CN}(\overline{\mathbf{x}}_k, \mathbf{\Sigma})$. Since the covariance matrix $\mathbf{\Sigma}$ is positive definite, a positive definite square root matrix $\mathbf{\Sigma}^{1/2}$ can be defined. The inverse of the matrix $\mathbf{\Sigma}^{1/2}$ is taken to obtain the matrix $\mathbf{\Sigma}^{-1/2}$. Applying a whitening transformation to matrix $\mathbf{X}$ such that

$$\mathbf{Z} = \mathbf{\Sigma}^{-1/2}\mathbf{X} \tag{A.2}$$

defines a new matrix $\mathbf{Z}$. The columns of this new matrix $\mathbf{Z}$, by the property given in equation (A.1), will therefore be distributed as $\mathcal{CN}\ (\mathbf{\Sigma}^{-1/2}\overline{\mathbf{x}}_k, \mathbf{I})$.

A.2 Distribution of the Matrix $\mathbf{A} = \mathbf{Z}^H\mathbf{Q}$

Consider the matrix $\mathbf{Z}$ which is $N \times K$ whose columns $\mathbf{z}_k$ for $k = 1, \cdots, K$ are independent and normally distributed as $\mathcal{CN}\ (\mathbf{0}, \mathbf{I})$. Also consider the two deterministic matrices $\mathbf{Q}_1$ and $\mathbf{Q}_2$ of dimensions $N \times r$ and $N \times N - r$, respectively, such that

$$\mathbf{Q}_1\mathbf{Q}_1^H + \mathbf{Q}_2\mathbf{Q}_2^H = \mathbf{I} \tag{A.3}$$

and

$$\mathbf{Q}_i^H \mathbf{Q}_j = \begin{cases} \mathbf{I} & \text{if } i = j \\ \mathbf{0} & \text{otherwise.} \end{cases}$$

Then the two matrices

$$\mathbf{A} = \mathbf{Z}^H \mathbf{Q}_1 \tag{A.4}$$

$$\mathbf{B} = \mathbf{Z}^H \mathbf{Q}_2 \tag{A.5}$$

are statistically independent. The columns of each matrix are independent and are distributed as $\mathcal{CN}(\mathbf{0}, \mathbf{I})$.

Proof: Matrices $\mathbf{A}$ and $\mathbf{B}$ are linear combinations of the columns of normally distributed matrix. Their expectations are:

$$\begin{aligned} E\{\mathbf{A}\} &= E\{\mathbf{Z}^H \mathbf{Q}_1\} \\ &= E\{\mathbf{Z}^H\}\, \mathbf{Q}_1 \\ &= \mathbf{0}, \end{aligned} \tag{A.6}$$

and

$$\begin{aligned} E\{\mathbf{B}\} &= E\{\mathbf{Z}^H \mathbf{Q}_2\} \\ &= E\{\mathbf{Z}^H\}\, \mathbf{Q}_2 \\ &= \mathbf{0}. \end{aligned} \tag{A.7}$$

The cross-correlation of the two matrices is

$$\begin{aligned} E\{\mathbf{A}\mathbf{B}^H\} &= E\{\mathbf{Z}^H \mathbf{Q}_1 \mathbf{Q}_2^H \mathbf{Z}\} \\ &= \mathbf{0}. \end{aligned} \tag{A.8}$$

A.3 Distribution of the Vector $\mathbf{v} = \mathbf{Q}^H \mathbf{c}$

For a random matrix $\mathbf{Q}$ of dimension $N \times K$ with the requirement that $\mathbf{Q}^H \mathbf{Q} = \mathbf{I}$ and an arbitrary distribution and which is statistically independent of the $\mathcal{CN}(\mathbf{0}, \mathbf{I})$

vector $\mathbf{c}$, the product vector

$$\mathbf{v} = \mathbf{Q}^H \mathbf{c} \tag{A.9}$$

is distributed as $\mathcal{CN}(\mathbf{0}, \mathbf{I})$ and is statistically independent of matrix $\mathbf{Q}$.

Proof: The joint PDF of $\mathbf{v}$ and $\mathbf{Q}$ is

$$f_{\mathbf{v},\mathbf{Q}} = f_{\mathbf{v}|\mathbf{Q}}\, f_{\mathbf{Q}} \tag{A.10}$$

where $f_{\mathbf{v}|\mathbf{Q}}$ is the PDF of $\mathbf{v}$ given $\mathbf{Q}$. Then the PDF of $\mathbf{v}$ as

$$f_{\mathbf{v}} = \int_{-\infty}^{\infty} f_{\mathbf{v}|\mathbf{Q}}\, f_{\mathbf{Q}}\, d\mathbf{Q}. \tag{A.11}$$

If $\mathbf{v}$ and $\mathbf{Q}$ are statistically independent

$$\begin{aligned} E\{\mathbf{v}|\mathbf{Q}\} &= \mathbf{Q}^H E\{\mathbf{c}|\mathbf{Q}\} \\ &= \mathbf{0}, \end{aligned} \tag{A.12}$$

and

$$\begin{aligned} E\{\mathbf{v}\mathbf{v}^H|\mathbf{Q}\} &= \mathbf{Q}^H E\{\mathbf{c}\mathbf{c}^H|\mathbf{Q}\}\mathbf{Q} \\ &= \mathbf{I}. \end{aligned} \tag{A.13}$$

The vector $(\mathbf{v}|\mathbf{Q}) = \mathbf{Q}^H\mathbf{c}$ is a linear combination of a complex normal vector and is itself complex normal with a PDF, $\Phi_{\mathbf{v}}$. Then (A.11) is written as

$$\begin{aligned} f_{\mathbf{v}} &= \Phi_{\mathbf{v}} \int_{-\infty}^{\infty} f_{\mathbf{Q}}\, d\mathbf{Q} \\ &= \Phi_{\mathbf{v}}. \end{aligned} \tag{A.14}$$

Then the two vectors $\mathbf{v}$ and $(\mathbf{v}|\mathbf{Q})$ are the same and hence, vector $\mathbf{v}$ is statistically independent of matrix $\mathbf{Q}$.

APPENDIX B

DISTRIBUTION OF THE QUANTITY $\mathbf{X}^H(\mathbf{Z}\mathbf{Z}^H)^{-1}\mathbf{X}$

Let $\mathbf{X}$ be a $N \times P$ complex matrix whose N-dimensional columns $\mathbf{x}_k$ for $k = 1, \cdots, P$, are distributed as $\mathcal{CN}\ (\overline{\mathbf{x}}_k, \mathbf{\Sigma})$. Also let $\mathbf{Z}$ be a $N \times K$ complex matrix whose N-dimensional columns $\mathbf{z}_k$ for $k = 1, \cdots, K$, are distributed as $\mathcal{CN}\ (0, \mathbf{\Sigma})$ with the condition that $K \geq N$ and $P \leq N$. Then the quantity

$$\eta = \mathbf{X}^H(\mathbf{Z}\mathbf{Z}^H)^{-1}\mathbf{X} \qquad K \geq N, \quad P \leq N \tag{B.1}$$

has a multivariate F-distribution [35] given as

$$\begin{aligned} f_\eta(\eta) &= e^{-tr\Omega}\ {}_1F_1(P+K; N; \Omega(I+\eta^{-1})^{-1}) \\ &\qquad \frac{\Gamma(P+K)}{\Gamma(N)\Gamma(P+K-N)} \frac{|\eta|^{N-P}}{\mathbf{I}+|\eta|^{P+K}} \end{aligned} \tag{B.2}$$

where

$$\overline{\mathbf{X}} = E\{\mathbf{X}\} \tag{B.3}$$

$$\Omega = \overline{\mathbf{X}}^H \mathbf{\Sigma}^{-1} \overline{\mathbf{X}} \tag{B.4}$$

$$\Gamma(n) = (n-1)!. \tag{B.5}$$

The ${}_1F_1$ is the confluent hypergeometric function [31] and defined as

$${}_1F_1(a; b; x) = \sum_{j=0}^{\infty} \frac{\Gamma(a+j)\Gamma(b)x^j}{\Gamma(a)\Gamma(b+j)j!}. \tag{B.6}$$

If instead of the $N \times P$ matrix $\mathbf{X}$, the $\mathcal{CN}\ (0, \mathbf{\Sigma})$ vector $\mathbf{x}$ is used in (B.1), then the distribution of η becomes a special case of the F-distribution which is called the Hotelling's T^2 distribution [44].

In the case where $\mathbf{x}$ is a vector of dimension $N \times 1$ with an expectation $E\{\mathbf{X}\} = 0$ such that $\mathbf{x}$ is $\mathcal{CN}\ (0, \mathbf{I})$ and matrix $\mathbf{Z}$ is $\mathcal{CN}(\mathbf{0}, \mathbf{I})$, the distribution of the quantity

$$\zeta = \mathbf{X}^H(\mathbf{Z}\mathbf{Z}^H)^{-1}\mathbf{X} \tag{B.7}$$

is Hotelling's T^2 distribution given as

$$f_\zeta(\zeta) = \frac{\Gamma(K+1)}{\Gamma(N)\Gamma(K-N+1)}\zeta^{N-1}(1-\zeta)^{K+1}. \tag{B.8}$$

APPENDIX C

DERIVATION OF THE DISTRIBUTION OF THE TEST STATISTIC UNDER HYPOTHESIS H_0, $f_\eta(\eta|H_0)$

The random variables ζ and ξ are defined as

$$\zeta = |x_A|^2 \tag{C.1}$$

$$\xi = \mathbf{x}_B^H \mathbf{x}_B. \tag{C.2}$$

The two random variables are independent and they have central chi-squared distributions which are given as

$$f_\zeta(\zeta) = e^{-\zeta} \tag{C.3}$$

$$f_\xi(\xi) = \frac{\xi^{N-r-2}}{\Gamma(N-r-1)} e^{-\xi}. \tag{C.4}$$

Define the random variable $\psi = K\alpha + \xi$. By performing a transformation of variables on the distribution given in (C.4), the distribution of ψ is given as

$$f_\psi(\psi) = \frac{(\psi - K\alpha)^{N-r-2}}{\Gamma(N-r-1)} e^{-\psi+K\alpha}. \tag{C.5}$$

The test statistic η is written as

$$\eta = \frac{|x_A|^2}{K\alpha + \mathbf{x}_B^H \mathbf{x}_B} \tag{C.6}$$

$$= \frac{\zeta}{K\alpha + \xi} \tag{C.7}$$

$$= \frac{\zeta}{\psi} \tag{C.8}$$

where the two random variables ζ and ψ are independent and their distributions are given in (C.4) and (C.5) respectively. At this point, the distribution of η is sought. Assume $\psi = a$, where a is a scalar, then $\eta = \zeta/a$ is a scaled version of ζ. Therefore

$$f_\eta(\eta|\psi = a) = |a| f_\zeta(a\eta|\psi = a). \tag{C.9}$$

The distribution of η is therefore

$$f_\eta(\eta) = \int_{-\infty}^{\infty} |u| f_\zeta(u\eta|\psi = u) f_\psi(u) du. \tag{C.10}$$

where u is an integration variable. Using the distribution of ζ and ψ in (C.4) and (C.5), the distribution of η is written as

$$f_\eta(\eta) = \int_{-\infty}^{\infty} |u| e^{-u\eta} \frac{(u - K\alpha)^{N-r-2}}{\Gamma(N-r-1)} e^{-u+K\alpha} du. \tag{C.11}$$

This integral can be solved in closed form to give

$$f_\eta(\eta) = [N - r - 1 + K\alpha(1+\eta)] (1+\eta)^{r-N} e^{-K\alpha\eta}. \tag{C.12}$$

APPENDIX D

DERIVATION OF THE DISTRIBUTION OF THE TEST STATISTIC UNDER HYPOTHESIS H_1, $f_\eta(\eta|H_1)$

The two independent random variables ζ and ξ are defined as

$$\zeta = |x_A|^2 \tag{D.1}$$

$$\xi = \mathbf{x}_B^H \mathbf{x}_B. \tag{D.2}$$

The random variable ζ has a non-central chi-squared distribution given as

$$f_\zeta(\zeta) = e^{-\delta_1} {}_0F_1(1; \delta_1\zeta) e^{-\zeta} \tag{D.3}$$

where δ_1 is its non-centrality parameter and is given as

$$\delta_1 = |b|^2. \tag{D.4}$$

The random variable ξ has a central chi-squared distribution given as

$$f_\xi(\xi) = \frac{\xi^{N-r-2}}{\Gamma(N-r-1)} e^{-\xi}. \tag{D.5}$$

Define the random variable $\psi = K\alpha + \xi$. By performing a transformation of variables on the distribution given in (D.5), the distribution of ψ is given as

$$f_\psi(\psi) = \frac{(\psi - K\alpha)^{N-r-2}}{\Gamma(N-r-1)} e^{-\psi+K\alpha}. \tag{D.6}$$

The test statistic η is written as

$$\eta = \frac{\zeta}{\psi} \tag{D.7}$$

where the two random variables ζ and ψ are independent and their distributions are given in (D.3) and (D.6) respectively. At this point, the distribution of η is sought. Assume $\psi = a$, where a is a scalar, then $\eta = \zeta/a$ is a scaled version of ζ. Therefore

$$f_\eta(\eta|\psi = a) = |a| f_\zeta(a\eta|\psi = a). \tag{D.8}$$

The distribution of η is therefore

$$f_\eta(\eta) = \int_{-\infty}^{\infty} |u| f_\zeta(u\eta|\psi = u) f_\psi(u) du \tag{D.9}$$

where u is an integration variable. Using the distribution of ζ and ψ in (C.4) and (C.5), the distribution of η is written as

$$f_\eta(\eta) = \int_{-\infty}^{\infty} |u| e^{-\delta_1} {}_0F_1(1; \delta_1 \eta u) e^{-\eta u} \frac{(u - K\alpha)^{N-r-2}}{\Gamma(N-r-1)} e^{-u+K\alpha} du \tag{D.10}$$

and equivalently as

$$f_\eta(\eta|\mathbf{H}_1) = \frac{e^{-\delta_1}}{\Gamma(N-r-1)} \int_{-\infty}^{\infty} |u| \, {}_0F_1(1; \delta_1 \eta u) e^{-\eta u} \, (u - K\alpha)^{N-r-2} \, e^{-(u-K\alpha)} du. \tag{D.11}$$

REFERENCES

1. P. W. Howells, "Explorations in fixed and adaptive resolution at GE and SURC," *IEEE Transactions on Antennas and Propagation*, vol. 24, pp. 575–584, September 1976.

2. S. P. Applebaum, "Adaptive arrays," *Syracuse University Reasearch Corporation*, vol. Rep. SPL TR 66-1, 1966.

3. B. Widrow et. al., "Adaptive antenna systems," *Proceedings of the IEEE*, vol. 55, pp. 2143–2159, December 1967.

4. O. L. Frost III, "An algorithm for linearly constrained adaptive array processing," *Proceedings of the IEEE*, vol. 60, pp. 926–935, August 1972.

5. L. E. Brennan and I. S. Reed, "Theory of adaptive radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 9, pp. 237–252, March 1973.

6. I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rates in adaptive arrays," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 10, pp. 853–863, November 1974.

7. M. H. Er and A. Cantoni, "A new set of linear constraints for broadband time domain element space processors," *IEEE Transactions on Antennas and Propagation*, vol. 34, pp. 320–329, March 1986.

8. D. Feldman and L. J. Griffith, "A projection approach for robust adaptive beamforming," *IEEE Transactions on Signal Processing*, vol. 42, pp. 867–876, Apr. 1994.

9. L. Cai and H. Wang, "Further results on adaptive filtering with embedded CFAR," *IEEE Transactions on Aerospace and Electronic Systems*, pp. 1009–1020, Oct. 1994.

10. E. J. Kelly, "An adaptive detection algorithm," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 22, pp. 115–127, March 1986.

11. A. M. Haimovich, "The eigencanceler: Adaptive radar by eigenanalysis methods," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 32, pp. 532–542, April 1996.

12. I. P. Kirsteins and D. W. Tufts, "Adaptive detection using low rank approximation to data matrix," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 30, pp. 55–67, Jan. 1994.

13. J. Goldstein and I. Reed, "Subspace selection for partially adaptive sensor array processing," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no. 2, pp. 539–543, April 1996.

14. J. Goldstein and I. Reed, "Reduced-rank adaptive filtering," *IEEE Transactions on Signal Processing*, vol. 45, no. 2, pp. 492–496, February 1996.

15. A. M. Haimovich, C. Peckham, T. Ayoub, J. S. Goldstein, I. S. Reed, "Performance Analysis of Reduced-Rank STAP," *1997 IEEE National Radar Conference*, pp. 42–47, May 13-15 1997.

16. D. Marshall, "A two step adaptive interference nulling algorithm for use with airborne sensor arrays," in *Seventh SP Workshop on Statistical Signal and Array Processing*, Quebec City, Canada, pp. 301–304, June 1994.

17. J. Ward, "Multiwindow post-doppler space-time adaptive processing," in *Seventh SP Workshop on Statistical Signal and Array Processing*, Quebec City, Canada, June 1994.

18. J. Ward, "Space-time adaptive processing for airborne radar," *Lincoln Laboratory Technical Report 1015*, Dec. 1994.

19. Y. I. Abramovich, "Controlled method for adaptive optimization of filters using the criterion of maximum signal to noise ratio," *Journal of Communication Technology and Electronics*, vol. 26, pp. 87–95, 1981, translated from Russian.

20. O. P. Cheremisin, "Efficiency of an adaptive algorithm with regularization of the sampled correlation matrix," *Journal of Communication Technology and Electronics*, vol. 27, pp. 69–77, 1982, translated from Russian.

21. B. D. Carlson, "Covariance matrix estimation errors and diagonal loading in adaptive arrays," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 24, pp. 397–401, July 1988.

22. S. Z. Kalson, "Adaptive array CFAR detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 31, pp. 534–541, April 1995.

23. S. Haykin, *Adaptive Filter Theory*, Prentice Hall, Englwood Cliffs, NJ, third ed., 1996.

24. H. Wang and L. Cai, "On adaptive spatial-temporal processing for airborne surveillance radar systems," *IEEE Transactions on Aerospace and Electrical Systems*, pp. 660–668, July 1994.

25. E. Barile, R. Fante, and J. Torres, "Some limitations on the effectiveness of airborne adaptive radar," *IEEE Transactions on Aerospace and Electrical Systems*, pp. 1015–1031, Oct. 1992.

26. R. L. Dilsavor and R. L. Moses, "Analysis of Modified SMI Method for Adaptive Array weight Control," *IFAC Identification and System Parameter Estimation*, pp. 417–422, 1991.

27. L. Chang and C. Yeh, "Performance of DMI and eigenspace-based beamformers," *IEEE Transactions on Antennas and Propagation*, vol. 40, pp. 1336–1347, Nov. 1992.

28. C. H. Gierull, "Performance analysis of fast projections of the hung turner type for adaptive beamforming," *Signal Processing, Elsevier*, vol. 50, pp. 17–28, April 1996.

29. A. Haimovich, "Asymptotic distribution of the conditioned signal-to-noise ratio in an eigenanalysis-based adaptive array," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no. 3, pp. 988–996, July 1996.

30. M. Miller, *Introduction to Adaptive Arrays*, John Wiley and Sons, NY, 1980.

31. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, National Bureau of Standards, Washington, DC, 1972. tenth printing.

32. D. M. Boroson, "Sample size considerations for adaptive arrays," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 16, pp. 446–451, July 1980.

33. E. J. Kelly, "Adaptive detection in non-stationary interference, part III," *Technical Report, Lincoln Laboratory*, vol. 761, August 1987.

34. L. Scharf, *Statistical Signal Processing; Detection, Estimation, and time series Analysis*, Addison-Wesley Publishing Company, MA, 1991.

35. A. James, "Distributions of matrix variates and latent roots derived from normal samples," *Annals of Mathematical Statistics*, vol. 35, pp. 475–501, 1964.

36. S. Kalson, "An adaptive array detector with mismatched signal rejection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, no. 1, pp. 195–206, January 1992.

37. F. Robey, D. Fuhrmann, E. Kelly, and R. Nitzberg, "A cfar adaptive matched filter detector," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 28, no. 1, pp. 208–216, 1992.

38. K. Burgess and B. V. Veen, "Subspace-based adaptive generalized likelihood ratio detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 44, no. 4, pp. 912–927, 1996.

39. I. Kirsteins, "Analysis and interpretation of the reduced-rank generalized likelihoo d ratio test," *Proceedings of EUSIPCO-94, Seventh European Processing Conference*, pp. 264–267, 1994.

40. N. Goodman, "Statistical analysis based on a certain multivariate complex gaussian distribution (an introduction)," *Annals of Mathematical Statistics*, vol. 34, pp. 152–177, 1963.

41. E. J. Kelly, "Finite-sum expressions for signal detection probabilities," *Technical Report, Lincoln Laboratory*, vol. 566, May 1981.

42. E. J. Kelly, "Adaptive detection in non-stationary interference, part I and part II," *Technical Report, Lincoln Laboratory*, vol. 724, June 1985.

43. R. J. Muirhead, *Aspects of Multivariate Statistical Theory*, John Wiley and Sons, Inc., New York, NY, 1982.

44. H. Hotelling, "The generalization of student's ratio," *Annals of Mathematical Statistics*, vol. 2, pp. 360–378, 1931.