ADAPTIVE TRANSMIT AND RECEIVE DESIGN FOR ARRAY RADAR

by

Zhu Chen

A DISSERTATION

Submitted to the Faculty of the Stevens Institute of Technology
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

______________________________
Zhu Chen, Candidate

ADVISORY COMMITTEE

______________________________
Hongbin Li, Chairman Date

______________________________
Yudong Yao Date

______________________________
Hong Man Date

______________________________
Wendy Wang Date

STEVENS INSTITUTE OF TECHNOLOGY
Castle Point on Hudson
Hoboken, NJ 07030
2013
This dissertation aims to apply adaptive processing for both the receiver and transmitter of a radar system, to address challenges such as the requirement of excessive training signals and high computational complexity in the conventional adaptive processing techniques for practical applications.

The first part of this dissertation deals with an adaptive reduced-rank detector, referred to as the CG-AMF detector, at the receiver of radar systems. The CG-AMF detector is obtained by using the conjugate gradient algorithm to solve for the weight vector of the adaptive matched filter (AMF). The CG is a computationally efficient iterative algorithm which finds the projection of the AMF weight vector to the Krylov subspace with a dimension growing with the CG iterations. This effectively leads to a family of reduced-rank detectors indexed by the number of CG iterations. We examine the output signal-to-interference-and-noise ratio (SINR) of the CG-AMF detector in the presence of strong clutter/interference. Specifically, by exploiting a connection between the CG algorithm and the Lanczos algorithm, we show the output SINR can be asymptotically expressed in a simple form involving a Ritz vector of the sample covariance. The probability density function (PDF) of the output SINR is then obtained based on this approximation. Our theoretical analysis of the CG-AMF detector is verified by computation simulation. Numerical comparisons are also made with several popular reduced-rank detectors using either data-independent or data-dependent rank reduction approaches. Our results show that for a fixed training size, the CG-AMF detector often reaches its peak output SINR with a lower rank compared with the other reduced-rank detectors, which implies that the CG-AMF detector has lower computational complexity requirement.
In the second part, we consider adaptive transmit and receive beampattern design for array radar systems. While adaptive processing is primarily employed at the receiver in conventional beamforming techniques, we also use it for adaptive transmit beamforming, which involves adaptively selecting the transmit correlation matrix by maximizing the output SINR at the receiver. One motivation of utilizing adaptive processing at the transmitter is that with imprecise knowledge of the interferences, only relying on adaptive receive beamforming may be inadequate for interference cancellation, whereas joint adaptive transmit and receive beamforming can potentially afford a stronger ability to handle the interferences. Simulations are provided to demonstrate the performance of the proposed joint adaptive transmit and receive beamforming approach.

Given a correlation matrix obtained in the previous optimal beampattern design stage, the next problem becomes that of determining the probing waveforms which is the ultimate goal of the designing exercise. The third part of the dissertation deals with the synthesis of transmit waveform matrix whose correlation matrix is equal or close to the pre-specified correlation matrix, and which also satisfies some practically motivated constraints. We consider the signal synthesis problem with a desired correlation matrix and a constant modulus constraint as well as a similarity constraint involving a known radar waveform with some desire properties. The proposed optimization algorithm based on cyclic algorithm and iterative algorithm yields solutions with good accuracy. Numerical simulations are also presented to demonstrate the effectiveness of the proposed algorithm.

Author: Zhu Chen
Advisor: Hongbin Li
Date: November 20, 2013
Department: Electrical and Computer Engineering
Degree: Doctor of Philosophy
Acknowledgments

First and foremost, I would like to gratefully and sincerely thank my advisor Prof. Hongbin Li for his kind supports, understanding, patience, encouragements as well as the insightful advices throughout my study. It was my great honor to purchase Ph.D degree under his guidance. He taught me how to conduct outstanding research, how to think critically and systematically and how to write concisely and clearly. Without his supervision and constant help this dissertation would not have been possible. Sincere thanks also extend to my Ph.D. committee members, Prof. Yudong Yao, Prof. Hong Man and Prof. Wendy Hui Wang. They provided the valuable suggestions, insights, and high standards to better my dissertation.

I am also indebted to Prof. Muralidhar Rangaswamy for his countless encouragements and tremendous support of my research works. He suggested me a promising research topic in the radar systems, reviewed and commented each of my research papers. Special thanks to Dr. Zafer Sahinoglu, Dr. Rui Ma and many other outstanding researchers and interns at Mitsubishi Electric Research Laboratories (MERL) when I worked there as an intern. I’ve been so fortunate to have their friendship during the joyful life in MERL.

Moreover, I would like to acknowledge my lab-mates in Signal Processing and Communication (SPAC) Laboratory at Stevens, especially, Dr. Pu Wang, Dr. Kil Woo Chung, Dr. Guolong Cui, Kiran Sampath, Ning Han, Kuang Cai, Zhe Wang and Xin Zhang. Thank you for all the supports and sharing the enjoyable experiences with me while pursuing the degree.

This thesis is attributed to my beloved parents, Zizhi Chen and Yulian Bai, who have always believed in me without reservation and given me their endless love. Finally, I would like to thank my wife, Lei Zhou, a lovely young lady who gives me a very special feeling to know that I’m sharing my life with someone who is not just my wonderful wife, but
also the greatest friend a person could ever have. Thank you for all your love and support!
Without you, I could never be able to reach this far.
Table of Contents

Abstract iii

Acknowledgments v

List of Tables xi

List of Figures xii

Chapter 1 Introduction 1
  1.1 Background 1
  1.2 Contribution of the Work 5

Chapter 2 Space-time Adaptive Processing Fundamentals 9
  2.1 Phased Array Radar Basics 9
  2.2 Full-rank STAP Solutions 11
    2.2.1 Sample Matrix Inversion 11
    2.2.2 Adaptive Matched Filter Detector 12
    2.2.3 Kelly’s GLRT 12
    2.2.4 Adaptive Coherence Estimator Detector 13
    2.2.5 Drawbacks of Full-rank STAP Detectors 13
  2.3 Reduced-rank STAP Solutions 14
    2.3.1 Data-independent RR Solutions 14
    2.3.2 Eigencanceler 15
    2.3.3 Cross Spectral Metric 15
    2.3.4 Drawbacks of the Existing RR Solutions 15
Chapter 3  The Conjugate Gradient Method 17
  3.1 Steepest Descent 17
  3.2 Conjugate Gradients 19
    3.2.1 General Search Directions 19
    3.2.2 A-Conjugate Search Directions 20
  3.3 The Lanczos Connection 23
  3.4 Equivalence of Multistage Wiener Filter and CG 24

Chapter 4  Conjugate Gradient Adaptive Matched Filter 26
  4.1 Data Model 26
  4.2 Conjugate Gradient Matched Filter 27
    4.2.1 General Covariance Matrix 29
    4.2.2 Structured Covariance Matrix 30
  4.3 Conjugate Gradient Adaptive Matched Filter 31
    4.3.1 Implementation of CG-AMF 32
  4.4 Performance of the CG-AMF Detector 34
    4.4.1 Output SINR of the CG-AMF Detector 34
    4.4.2 Low-Rank Approximation 36
    4.4.3 Statistical Analysis 39
  4.5 Numerical Results 46
  4.6 Conclusions 52
  4.7 Appendix 52
    4.7.1 Proof of Lemma 1 52
    4.7.2 Proof of Lemma 2 53
    4.7.3 Proof of Theorem 1 55
    4.7.4 Mean of the output SINR 57
Chapter 5  Adaptive Transmit and Receive Beamforming for Interference Mitigation

5.1 Data Model 58
5.2 Proposed Approach 61
   5.2.1 Transmit and Receive Beamforming Design 61
   5.2.2 Adaptive Estimation 64
5.3 Numerical Results 67
5.4 Conclusions 68
5.5 Appendix 70
   5.5.1 Proof of (5.12) and (5.13) 70
   5.5.2 Proof of (5.15) 71
   5.5.3 Derivation of the optimal solution of (5.21) 72

Chapter 6  Waveform Synthesis for MIMO Radar

6.1 Data Model 76
6.2 Cyclic Method 77
6.3 Waveform Synthesis with Transmit Power Constraint 78
6.4 Waveform Synthesis with Constant Modulus and Similarity Constraints 80
   6.4.1 Constant Modulus and Similarity Constraints 80
   6.4.2 Iterative Optimization Algorithm 81
   6.4.3 Summarization of the Proposed Algorithm 85
6.5 Numerical Results 85
6.6 Conclusions 87

Chapter 7  Conclusion

Bibliography 91
Vita
List of Tables

3.1 Steepest Descent 18
3.2 Conjugate Gradients 22
4.1 Conjugate-Gradient MF Detector 28
4.2 Conjugate-Gradient AMF Detector 33
6.1 Waveform Synthesis Algorithm I 79
6.2 Waveform Synthesis Algorithm II 86
List of Figures

4.1 Output SINR of CG-AMF with rank k=11 versus the training size $L$. 47
4.2 Output SINR versus detection rank $k$ with $L = 64$ training signals. 48
4.3 Probability of detection for the clairvoyant MF, fully adaptive AMF, and reduced-rank CG-AMF detectors with $L = 64$ and $P_f = 10^{-2}$. 49
4.4 Probability of detection for the reduced-rank EIG, CSM and CG-AMF detectors with $L = 64$ and $P_f = 10^{-2}$. 49
4.5 Output SINR versus detection rank $k$ with $L = 128$ training signals. 51
4.6 Probability of detection for the reduced-rank EIG, CSM and CG-AMF detectors with $L = 128$ and $P_f = 10^{-2}$. 51

5.1 Mean of the output SINR. 69
5.2 Beampatterns of the beamformers. 69

6.1 Square error between $S^H S$ and $R$ with the similarity parameter $\epsilon = 0, 0.5, 1.5, 1.8, 2$ respectively 87
6.2 Amplitude and phase of the reference waveform and the designed waveforms with the similarity parameter $\epsilon = 0.5, 1.5, 1.8, 2$ respectively 88
Chapter 1

Introduction

1.1 Background

The problem of detecting a multichannel signal from temporally and spatially correlated disturbance is encountered in a variety of applications including radar, sonar, wireless communications and others [34, 43, 58]. The problem has been extensively studied under the framework of space-time adaptive processing (STAP) in phased-array radar, which employs multi-element antennas and multi-pulse waveforms to probe and observe the radar scene. Numerous STAP detectors have been proposed. Among them, the covariance matrix based detectors, which need knowledge of the space-time covariance matrix of the disturbance signal to suppress the interferences, are the most widely used multichannel signal detectors. Examples of such detectors include the Reed, Mallett, and Brennan detector [44], Kelly’s generalized likelihood ratio test (GLRT) [31], the adaptive matched filter (AMF) detector [12], [45], the adaptive coherence estimator detector [35], among others. All of them involve estimating and inverting a space-time covariance matrix of the disturbance signal for each test cell using target-free training data, which may impose excessive training and computational burdens when the joint space-time dimension is large.

In order to mitigate the training and computation requirements of the full-rank covariance matrix based STAP detectors, reduced-rank techniques have been proposed to reduce the dimension of the input signal in advance of detection [34, 58]. Specifically, reduced-rank methods employ rank reduction mechanisms to reduce the degrees of freedom (DoFs) of the detectors so that to alleviate the estimation burden. For a fixed training size, reduced-rank methods often offer a better estimation accuracy than their full-rank
counterparts and, despite the loss of DoFs, may actually attain a better detection performance.

The rank reduction of the input signal space can be achieved by a linear transformation matrix $\mathbf{T}$. There are two broad families of reduced-rank transformations, namely, data-independent transformations and data-dependent transformations. The former include using sub-matrices formed from the discrete Fourier transform (DFT), discrete cosine transform (DCT), or other data-independent linear matrices. Examples of data-dependent reduced-rank transformations include the eigencanceler [26], the cross-spectral metric (CSM) [21], and others. The main advantage of data-independent transformations is their low computational cost. Data-dependent schemes normally require more complex computation. For example, the eigencanceler and CSM require the eigen-decomposition of the covariance or transformed covariance matrix. Despite the higher complexity, better performance is often obtained by using data-dependent transformations.

The conjugate gradient (CG) algorithm is a computationally efficient method for solving a linear system. It is guaranteed to converge in a fixed number of iterations, and able to provide a series of approximations to the solution in an expanding Krylov subspace. Due to these properties, the CG algorithm has been investigated for estimation and detection in a number of recent studies [8, 29, 30, 33, 48, 59]. Specifically, [8] considered the use of CG for adaptive filtering. Krylov subspace beamforming was studied in [33]. Parametric adaptive detection exploiting the computational efficiency of the CG algorithm for linear prediction was proposed in [29]. The CG algorithm was used to approximate the non-adaptive matched filter in [30] and a number of convergence properties were established. Meanwhile, [48, 59] reveals some interesting connections between the CG and the multi-stage Weiner filter (MWF).

Optimal linear beamformers [1, 7, 56] employ linear weights to optimize the receive beamformer response based on the statistics of the data. Specifically, the covariance matrix
of the disturbance signal (i.e., interferences and noise) is used to place nulls in the directions of interfering sources in an attempt to maximize the signal-to-interference-plus-noise ratio (SINR) at the output of the beamformer. In practice, data statistics are often unknown and may change with time. To cope with the problem, adaptive algorithms are used to obtain weights that converge to the statistically optimal solution. An adaptive beamformer requires training data to estimate the unknown disturbance covariance matrix. However, the challenge is that training data are often limited in many practical scenarios, which may cause significant performance loss due to lack of sufficient training data that are needed to form a reliable covariance matrix estimate. In an effort to improve the performance under these conditions, we propose to use training data not only for adaptive reception as traditional beamformers do, but also to adaptively control the transmit beampattern for radiation.

Transmit beampattern and waveform design has been a subject of recent interest for array radars, e.g., [10, 14, 16, 18, 20, 51]. However, most of them do not consider adaptive processing for radiation. In [51] and [20], the authors employ a matched-illumination design criteria where the correlation matrix of the probing signals is designed such that the power can be transmitted to a desired range of angles, without considering the mitigation of interference sources. Instead of designing the correlation matrix, [18] and [10] address the waveform design problem of array radars by maximizing the output SINR and, therefore, interference mitigation is explicitly accounted for. However, the limitation is that these approaches assume full knowledge of the disturbance covariance matrix and, as a result, are non-adaptive. Meanwhile, [14] and [16] examine radar phase code design under several constraints, e.g., constraints on a similarity to known radar codes, the peak-to-average power ratio, and estimation accuracy, etc. Their designs are also non-adaptive.

Waveform design has been a subject of recent interest for array radars [2, 37]. In traditional phased-array radars, directionality of the transmit beam is achieved by the ap-
lication of phase shifts to each of many transmit elements, with transmitted waveforms that are otherwise equal at each element and thus perfectly coherent. In contrast, the use of mutually uncorrelated waveforms has been proposed for multiple-input multiple-output (MIMO) radar systems. Uncorrelated waveforms achieve broad spatial patterns, nominally omnidirectional or more realistically the beampattern.

The problem of waveform design for MIMO radar can be classified into two categories. The first category thinks of designing directly the probing signals by optimizing a given performance measure with respect to the waveforms [10, 11, 18, 38, 42, 47]. In these work, [10, 18] address the waveform design problem of array radars by maximizing the output SINR and, therefore, interference mitigation is explicitly accounted for. Ambiguity function design was considered in [11, 47] to improve the radar performance in the spatial, range, and Doppler domains by optimizing the entire waveforms. In [42], MIMO waveform was devised by minimizing the estimation error of the Minimum Mean Squared Error (MMSE) estimators for uncorrelated and correlated targets. In [38], transmit waveforms were optimized for multiple targets in the presence of spatially colored interference and noise, based on several design criteria, including minimizing the trace, determinant, and the largest eigenvalue of the Cramer-Rao Bound (CRB) matrix.

The second category approaches the waveform design problem by first considering the transmit beampattern design and then synthesizing the associated probing waveform. Recently proposed techniques for transmit beampattern design for MIMO radar system have focused on the optimization of the correlation matrix of the waveforms [19, 51, 57]. In [19], the correlation matrix was devised to achieve or approximate a desired spatial beam pattern. In [51], the waveform correlation matrix was designed to attain a desired beam pattern as well as to minimize the cross-correlation between the probing signals at a number of given target locations. In [57], the authors considered constant modulus signal design to approximate a desired beam pattern while minimizing the levels of both the autocorrelation
and cross-correlation side lobes at given spatial angles. The advantage of optimizing the same performance measure with respect to the correlation matrix $\Omega$ of the transmitted waveforms rather than optimizing directly with respect to the waveform matrix $S$ is its less design complexity. This is so because $S$ has more unknowns than its correlation matrix $\Omega$ and the dependence of various performance measures on $S$ is more intricate than the dependence on $\Omega$. As the ultimate goal of the designing exercise is the probing waveforms, with a pre-selected or optimized correlation matrix $\Omega$, the next step is to determine an optimal signal waveform matrix $S$.

Several practical design constraints, such as the constant modulus constraint [57], the similarity constraint [15], [17], and the peak-to-average-power ratio (PAPR) constraint [16], are often considered in waveform design. The constant modulus or PAPR constraint is needed because radar amplifiers usually work in a saturation condition, which prohibits amplitude modulation in radar waveforms. Meanwhile, a similarity constraint uses a known waveform as a benchmark, which allows the designed waveform to share some of the good ambiguity properties of the known waveform. However, [15–17, 57] only applied these constraints in the directly waveform design i.e., they all belong to the aforementioned first category of work. For the second category, i.e., with a given correlation matrix, [19] introduced a method to synthesize constant modulus waveforms and [52] took the PAPR constraints into account during the waveform synthesis, while the waveform synthesis with similarity constraints appears not available in the open literature.

1.2 Contribution of the Work

This dissertation examines the application of adaptive processing at the receiver and transmitter of array Radar systems. The main research objectives are to design an efficient adaptive matched filter at the receiver, and an adaptive transmit and receive beampattern as
well as the associate probing waveforms.

First, we explore using the CG algorithm along with the AMF test for adaptive reduced-rank detection. Specifically, the computationally efficient CG algorithm is employed to iteratively calculate the weight vector of the AMF detector. This in turn produces a group of CG-AMF detectors, each having a different rank determined by the number of CG iterations. The performances of these CG-AMF detectors are considered in terms of their output SINRs. It is natural to ask which CG-AMF detector yields the highest output SINR in average. It should be noted that due to a trade-off between DoFs and estimation accuracy when the training size is fixed, the output SINR does not always grow with the rank or the number of CG iterations. An optimal rank is usually achieved well before the CG reaches its full iterations (at which point it yields a full-rank solution). Any additional CG iterations beyond the optimal rank is therefore not only wasteful of the computational resource, but also at the cost of reducing the detection performance. To this end, we examine the statistical behavior of the output SINR of the CG-AMF detector. Under the condition that the signal is contaminated by strong low-rank interference, the an asymptotic expression of the probability density function (PDF) of the output SINR is obtained by using a connection between the CG and the Lanczos algorithm.

Second, we investigate jointly adaptive transmit and receive beamforming in the presence of interferences. A closed-form expression of the transmit beamforming correlation matrix is obtained by maximizing a lower bound of the output SINR at the receiver. Our solutions of the transmit beamforming correlation matrix and the associated receive beamformer require some knowledge (i.e., locations and strengths) of the interferences, which are adaptively estimated from the training data. The advantage of employing adaptive processing at both the transmitter and receiver in the presence of interferences with uncertainties is demonstrated by numerical results.

Third, we consider the MIMO radar waveform synthesis for transmit beampattern
design, by taking into account the constant modulus constraint as well as a similarity constraint between the designed signal and a reference radar waveform. With a pre-selected or optimized correlation matrix $\Omega$ in the beampattern design stage, our problem is to determine a probing waveform matrix $S$ whose correlation matrix is equal or close to $\Omega$, and which also satisfies the practically motivated constraints. We propose an optimization algorithm to iteratively update the waveform matrix $S$. An NP-hard optimization problem is involved during the iteration, and we employ a relaxation and randomization approach which is known to yield approximate solutions with good accuracy [17]. The effectiveness of proposed methodology is evaluated by numerical results.

The rest of this dissertation is organized as follows. Chapter 2 briefly reviews the STAP for airborne radar detection. Both the full-rank and reduced-rank STAP detectors are introduced here. Chapter 3 introduces the conjugate gradient method, where we examine its convergence property and the connection between the Lanczos algorithm and multistage Wiener filter. Chapter 4 presents the conjugate-gradient matched filter and conjugate-gradient adaptive matched filter detectors. The asymptotic performance analysis of the proposed CG-AMF detector is also included in this chapter. Chapter 5 presents the adaptive transmit and receive beampattern design for array radar systems. Chapter 6 solves the remaining task of Chapter 5, synthesizing the probing waveforms for the pre-specified correlation matrix with practical constrains, i.e., transmit power constraint, constant modulus and similarity constraints. Finally, some conclusions are drawn in Chapter 7.

**Notation:** Vectors and matrices are denoted by boldface lower-case and upper-case letters, respectively. Transpose, complex conjugate and complex conjugate transpose are respectively represented by $(\cdot)^T$, $(\cdot)^*$ and $(\cdot)^H$. $\mathbb{C}$ and $\mathbb{R}$ denote the complex and real number fields. $\mathcal{CN}(\mu, R)$ denotes the multivariate complex Gaussian distribution with mean $\mu$ and covariance matrix $R$, $\Re \{ \cdot \}$ denotes the real part of a complex variable, $\text{tr}(\cdot)$ denotes the trace of a matrix, $\text{diag}(Z)$ denotes a diagonal matrix whose diagonal elements are the same
as $\mathbf{Z}$, $\|\cdot\|$ takes the Frobenius norm of a matrix/vector, $\otimes$ denotes the matrix Kronecker product, $\odot$ denotes Hadamard product and $\circ$ denotes the column-wise or Khatri-Rao product, vec$(\cdot)$ denotes the operation of stacking the columns of a matrix on top of each other and matrix$(\cdot)$ is the reverse operation of vec$(\cdot)$, and finally, $(\cdot)^\dagger$ denotes the Moore-Penrose pseudo-inverse.
Chapter 2
Space-time Adaptive Processing Fundamentals

Detecting targets in an interference background comprised of clutter and jamming is required by airborne radar system. STAP which is a multidimensional adaptive filtering algorithm simultaneously combines and processes the signals received from the elements of an array antenna (the spatial domain) and the multiple pulse repetition periods of a coherent radar waveform (the temporal domain). Therefore, it can be used to suppress interference and provide low-velocity target detection in strong interference environment.

In this chapter, we will review a variety of STAP approaches for airborne radar detection problem. However, it is worth to note that STAP has broad applications in radar, sonar, remote sensing, and communication systems [1-5]. Therefore, the proposed method in this proposal can also be applicable to those areas.

2.1 Phased Array Radar Basics

The conventional phase array radar system is described in [58], where the radar antenna is a uniformly spaced linear array antenna (ULA) consisting of $J$ elements. The radar transmits a coherent burst of $N$ pulses at a constant pulse repetition frequency (PRF) $f_r = 1/T_r$, where $T_r$ is the pulse repetition interval (PRI). The transmitter carrier frequency is $f_0 = c/\lambda_0$, where $c$ is the propagation velocity and $\lambda_0$ is the radar operating wavelength. The time interval over which the waveform returns are collected is commonly referred to as the coherent processing interval (CPI). For each PRI, $L_t$ time(range) samples are collected to cover the range interval. With $N$ pulses and $J$ receiver channels, the received data for one CPI comprises $JNK$ samples and is often referred as the $J \times N \times L_t$ CPI datacube.

Let the $J \times N$ vector $\mathbf{x}$ refer to a space-time snapshot at the range of interest,
the function of a surveillance radar is to ascertain whether targets are preset in the data, therefore, given the test data, the target detection problem can be expressed as a binary hypothesis testing:

\begin{align*}
    H_0 : & \quad x = d \\
    H_1 : & \quad x = \alpha s + d
\end{align*}

(2.1)

where \( s \) is the space-time steering vector, \( \alpha \) is an unknown complex amplitude, and \( d \) a disturbance signal that is assumed to be spatially and temporally correlated. The disturbance \( d \) is often modeled as a Gaussian random vector with zero-mean and space-time covariance matrix \( R \in \mathbb{C}^{JN \times JN} \). As a result, \( x \sim \mathcal{CN}(\alpha s, R) \), where \( \alpha = 0 \) under \( H_0 \) and \( \alpha \neq 0 \) under \( H_1 \).

For a uniform equal-distant linear array, the steering vector is given by [46]:

\[ s = s_t \otimes s_s \quad (2.2) \]

where the \( N \times 1 \) temporal steering vector takes the form as

\[ s_t = \left( \frac{1}{\sqrt{N}} \right) \begin{bmatrix} 1 & e^{i2\pi f_d} & \ldots & e^{i2\pi(N-1)f_d} \end{bmatrix}^T \quad (2.3) \]

with a normalized Doppler frequency \( f_d \). The temporal steering vector is given by

\[ s_s = \left( \frac{1}{\sqrt{J}} \right) \begin{bmatrix} 1 & e^{i2\pi f_s} & \ldots & e^{i2\pi(J-1)f_s} \end{bmatrix}^T \quad (2.4) \]

where \( f_s \) is a normalized spatial frequency.
2.2 Full-rank STAP Solutions

The space-time processor employs a linear filter that combines all the samples from the range gate of interest to produce a scalar output. Hence, the space-time processor can be represented by an $JN$-dimensional weight vector $w$, and the detector is given by,

$$T = |w^H x|^2 H_1 \geq \eta$$  \hspace{1cm} (2.5)

where $\eta$ is the test threshold.

The optimum detector that maximizes the output SINR is the matched filter (MF) [12] with the weight vector as

$$w_{MF} = R^{-1} s.$$  \hspace{1cm} (2.6)

Ideally, the space-time processor provides coherent gain on target while forming angle and Doppler response nulls to suppress clutter/interference. As the clutter scenario is not known in advance, the weight vector must be determined in a data-adaptive way from the training data.

2.2.1 Sample Matrix Inversion

The weight vector of MF is derived with assumed knowledge of the covariance matrix $R$. In practice, $R$ must be estimated from training data. Sample matrix inversion(SMI) algorithms are considered where the secondary data $\{x_l\}$ are used to form the sample covariance matrix estimate of $R$:

$$\hat{R} = \frac{1}{L} \sum_{l=1}^{L} x_l x_l^H.$$  \hspace{1cm} (2.7)
Typically, the training samples \( \{ x_l \} \) cover a range interval surrounding but not including the range gate of interest. The SMI weight vector is then cast from

\[
\hat{w} = \hat{R}^{-1}s.
\]  

It is also known as the RMB detector [44]. Note that this detector is suboptimum due to the covariance estimate.

### 2.2.2 Adaptive Matched Filter Detector

The AMF is a modification to the RMB detector in order to achieve the CFAR property [6, 12, 45], it takes the form as:

\[
T_{AMF} = \frac{| s^H \hat{R}^{-1} x |^2}{s^H \hat{R}^{-1} s} \begin{cases} H_1 \gtrless \eta_{AMF} \\ H_0 \end{cases}
\]  

where \( \eta_{AMF} \) is the AMF detector threshold.

### 2.2.3 Kelly’s GLRT

The amplitude \( \alpha \) in the detection model could also be unknown, and Kelly provide a GLRT approach to estimate \( \alpha \) and \( R \) successively by ML [31]. This GLRT is given as

\[
T_{Kelly} = \frac{| s^H \hat{R}^{-1} x |^2}{(s^H \hat{R}^{-1} s) \left( L + x^H \hat{R}^{-1} x \right)} \begin{cases} H_1 \gtrless \eta_{Kelly} \\ H_0 \end{cases}
\]  

where \( \eta_{Kelly} \) is the detector threshold.
2.2.4 Adaptive Coherence Estimator Detector

The ACE detector [13, 40] taking the form as the normalized AMF detector is also a GLRT based detection scheme,

\[
T_{ACE} = \frac{\left| s^H \hat{R}^{-1} x \right|^2}{\left( s^H \hat{R}^{-1} s \right) \left( x^H \hat{R}^{-1} x \right)} \xrightarrow{H_0 \rightarrow H_1} \eta_{ACE}
\]  

(2.11)

where \( \eta_{ACE} \) is the associated threshold.

2.2.5 Drawbacks of Full-rank STAP Detectors

There are three major shortages of these fully adaptive STAP detectors. First, it requires large training data. In particular, they all need to perform the sample covariance matrix inverse, which imposes a constraint on the training size.

\[
L \geq JN
\]  

(2.12)

to ensure a full-rank \( \hat{R} \). The RMB rule [44] suggests that at least

\[
L \geq (2JN - 3)
\]  

(2.13)

target-free secondary data vectors are need to obtain an expected performance within 3 dB from the optimum MF detector. Such a training requirement is often difficult to meet, especially in non-homogeneous or dense-target environments. Second, taking the inverse of a large sample covariance matrix will increase the computational complexity. Finally, in training-limited scenarios, the fully STAP will also incur substantial performance loss due to covariance estimation error.
2.3 Reduced-rank STAP Solutions

Reduced-rank detection offers an alternative approach when training data is limited. It can mitigate the training and computational burden by reducing the dimension of the input signal before detection. There are two broad families of reduced-rank detectors: data-independent methods and data-dependent methods. The later methods employ some rank reduction approach that is derived from the observed signal.

2.3.1 Data-independent RR Solutions

These methods apply a data-independent transformation matrix $T \in \mathbb{C}^{k \times M}$, $k < M$ to the received signal $x$ followed by a standard detection. For example, a reduced-rank version of the AMF has a weight vector given by

$$\hat{w} = T \left( T \hat{R} T^H \right)^{-1} T s.$$  \hspace{1cm} (2.14)

Among others, the discrete Fourier transform (DFT) and discrete cosine transform (DCT) matrices are popular choices for data-independent RR. However, directly applying the DFT or DCT as $T$ may bring a problem, since the steering vector has a chance to be (nearly) orthogonal to $T$, i.e., $Ts = 0$ which whittle the output SINR. This problem can be avoided by constructing the transforms as

$$T = [s \quad T_0]^H$$ \hspace{1cm} (2.15)

where $T_0$ contains $k - 1$ columns of the full-size DFT or DCT matrix. We can pick the $k - 1$ columns which are least correlated with $s$ to minimized signal cancellation.
2.3.2 Eigencanceler

The eigencanceler [26] utilizes the noise eigenvectors \( \hat{E}_n \in \mathbb{C}^{M \times k} \) of the sample covariance matrix \( \hat{R} \), where \( k \) is any integer between 1 and the maximum number of noise eigenvectors of \( \hat{R} \). \( \hat{E}_n \) spans a noise subspace which is orthogonal to the dominant interference contained in the observed signal. It also plays the same role of rank reduction as the \( T \) matrix for data-independent reduced-rank methods. The weight vector of the eigencanceler can be expressed as

\[
\hat{w}_E = \hat{E}_n \hat{E}_n^H s.
\]  

(2.16)

2.3.3 Cross Spectral Metric

Another well-known example of data-dependent reduced-rank detector is the cross spectral metric (CSM) based detector [21], whose weight vector is given by

\[
\hat{w}_{CSM} = \left[ I_M - \hat{A}^H U \left( \hat{U}^H \hat{A} \hat{A}^H U \right)^{-1} \hat{U}^H \hat{A} \right] s
\]

(2.17)

where \( \hat{A} \) is a signal blocking matrix satisfying

\[
A s = 0
\]

(2.18)

and \( U \) is formed from the \( k \) eigenvectors of \( \hat{A} \hat{R} \hat{A}^H \) that maximize the quantity \( \left| q_i^H \hat{A} \hat{R} s \right|^2 / \lambda_i \), with \( q_i \) and \( \lambda_i \) denoting the eigenvectors and eigenvalues of \( \hat{A} \hat{R} \hat{A}^H \), respectively.

2.3.4 Drawbacks of the Existing RR Solutions

In general, the data-independent RR solutions are easily to implement with low computational cost. However, since they don’t utilize the training data, the corresponding detection performances are often poor at low rank. For the data-dependent RR detectors such as the
eigencanceler and CSM, they outperform the data-independent RR detectors, especially in training limited cases ($L$ is small). However, these detectors often require high complex algorithm such as eigen-decomposition and therefore increase the computational complexity.
Chapter 3
The Conjugate Gradient Method

CG is the most popular iterative method for solving large systems of linear equations. CG is effective for systems of the form

$$Ax = b$$  

(3.1)

where $x$ is an unknown vector, $b$ is a known vector, and $A$ is a known, square, symmetric, positive-definite (or positive-indefinite) matrix. These systems arise in many important settings, such as finite difference and finite element methods for solving partial differential equations, structural analysis, circuit analysis.

The quadratic form is simply a scalar, quadratic function of a vector with the form

$$\phi(x) = \frac{1}{2}x^HAx - x^Hb$$  

(3.2)

The derivation of CG method is starting with how we minimize the above quadratic function. The minimum value of $\phi(x)$ is $-b^HA^{-1}b/2$, achieved by setting

$$x = A^{-1}b.$$  

(3.3)

Thus, minimizing $\phi$ and solving $Ax = b$ are equivalent problems if $A$ is symmetric positive definite.

3.1 Steepest Descent

One of the simplest strategies for minimizing $\phi$ is the method of steepest descent(SD) [22]. At a current point $x_k$ the function $\phi$ decreases most rapidly in the direction of the negative
gradient:
\[ - \nabla \phi(x_k) = b - Ax_k \]  \hspace{1cm} (3.4)

and the residual of \( x_k \) is given as:
\[ \gamma_k = b - Ax_k. \]  \hspace{1cm} (3.5)

If the residual is nonzero, then there exists a positive \( \alpha \) such that

\[ \phi(\gamma_k + \alpha \gamma_k) < \phi(\gamma_k). \]  \hspace{1cm} (3.6)

SD set the \( \alpha \) as
\[ \alpha = \frac{\gamma_k^H \gamma_k}{\gamma_k^H A \gamma_k} \]  \hspace{1cm} (3.7)

which results in the following minimization problem

\[ \phi(x_k + \alpha \gamma_k) = \phi(x_k) - \alpha \gamma_k^H \gamma_k + \frac{1}{2} \alpha^2 \gamma_k^H A \gamma_k. \]  \hspace{1cm} (3.8)

Therefore, the SD algorithm can be summarized as below:

<table>
<thead>
<tr>
<th>Steepest Descent</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_0 = ) initial guess</td>
</tr>
<tr>
<td>( \gamma_0 = b - Ax_0 )</td>
</tr>
<tr>
<td>( k = 0 )</td>
</tr>
<tr>
<td><strong>while</strong> ( \gamma_k \neq 0 )</td>
</tr>
<tr>
<td>( k = k + 1 )</td>
</tr>
<tr>
<td>( \alpha = \frac{\gamma_k^H \gamma_k}{\gamma_k^H A \gamma_k} )</td>
</tr>
<tr>
<td>( x_k = x_{k-1} + \alpha \gamma_{k-1} )</td>
</tr>
<tr>
<td>( \gamma_k = b - Ax_k )</td>
</tr>
<tr>
<td><strong>end</strong></td>
</tr>
</tbody>
</table>

Table 3.1: Steepest Descent
Moreover, it can be shown that

\[
\left( \phi(x_k) + \frac{1}{2} d^H A^{-1} d \right) \leq \left( 1 - \frac{1}{\kappa_2(A)} \right) \left( \phi(x_{k-1}) + \frac{1}{2} d^H A^{-1} d \right)
\]

(3.9)

which implies global convergence. The condition number is given as

\[
\kappa_2(A) = \frac{\lambda_1(A)}{\lambda_n(A)}
\]

(3.10)

Unfortunately, the rate of convergence may be prohibitively slow if the condition is large. Stated another way, in this case, the gradient directions that arise during the steepest descent iteration are not different enough.

### 3.2 Conjugate Gradients

Steepest Descent often finds itself taking steps in the same direction as earlier steps. To avoid the pitfalls of steepest descent, the method of Conjugate Gradients (CG) employs the A-conjugate directions which are constructed by conjugation of the residuals. The reason to apply residuals based A-conjugate directions is that, first, it keep the positive aspects of steepest descent by using residuals, second, the residual has the nice property that it is orthogonal to the previous search directions, so it is guaranteed always to produce a new, linearly independent search direction unless the residual is zero.

#### 3.2.1 General Search Directions

To avoid the shortcomings of steepest descent, it is assumed that the successive minimization of \( \phi \) along a set of directions \( \{d_1, d_2, \cdots \} \) that do not necessarily correspond to the
residuals \( \{ \gamma_0, \gamma_1, \cdots \} \), then \( \phi(x_{k-1} + \alpha d_k) \) is minimized by setting

\[
\alpha = \alpha_k = \frac{d_k^H \gamma_{k-1}}{d_k^H A d_k}.
\] (3.11)

With this choice it can be shown that

\[
\phi(x_k + \alpha \gamma_k) = \phi(x_k) - \frac{1}{2} \frac{(d_k^H \gamma_{k-1})^2}{d_k^H A d_k}.
\] (3.12)

To ensure a reduction in the size of \( \phi \), \( d_k \) should not be orthogonal to \( \gamma_{k-1} \). The conjugate-gradient algorithm choose A-conjugate search directions that guarantees convergence without the drawbacks of SD.

### 3.2.2 A-Conjugate Search Directions

If the search directions are linearly independent and \( x_k \) solves the problem

\[
\min \phi(x)
\] (3.13)

where

\[
x \in x_0 + \text{span} \{d_1, \cdots, d_k\}
\] (3.14)

for \( k = 1, 2, \cdots \), then convergence is guaranteed in at most \( n \) steps.

Give the expression about the determination of \( d_k \) as below,

\[
x_k = x_0 + D_{k-1} y + \alpha d_k
\] (3.15)

where

\[
D_{k-1} = [d_1, \cdots, d_{k-1}]
\] (3.16)
then

\[ \phi(x_k) = \phi(x_0 + D_{k-1}y) + \alpha y^H D_{k-1}^H A d_k + \frac{\alpha^2}{2} d_k^H A D_k - \alpha d_k^H \gamma_0. \]  

(3.17)

If

\[ d_k \in \text{span}\{A d_1, \cdots, A d_{k-1}\}^\perp \]  

(3.18)

then we have

\[ D_{k-1}^H A d_k = 0 \]  

(3.19)

and the search for the minimizing \( x_k \) splits into a pair of uncouple minimizations:

\[
\min \phi(x_k) = \min_{y, \alpha} \phi(x_0 + D_{k-1}y + \alpha d_k)
\]

\[
= \min_y \phi(x_0 + D_{k-1}y) + \min_\alpha \left( \frac{\alpha^2}{2} d_k^H A D_k - \alpha d_k^H \gamma_0 \right).
\]

(3.20)

If \( y_{k-1} \) solves the first min problem then

\[ x_{k-1} = x_0 + D_{k-1}y_{k-1} \]  

(3.21)

minimizes \( \phi \) over \( x \in x_0 + \text{span}\{d_1, \cdots, d_k\} \), and the solution to the \( \alpha \) min problem is given by

\[ \alpha_k = \frac{d_k^H \gamma_0}{d_k^H A d_k} \]  

(3.22)

Because of A-conjugacy, it also has

\[ d_k^H \gamma_0 = d_k^H \gamma_{k-1}. \]  

(3.23)

The conjugate gradient method combines the positive aspects of SD and A-conjugate searching by choosing \( d_k \) to be the closet vector to \( \gamma_{k-1} \) that is A-conjugate to \( d_1, d_2, \cdots, d_{k-1} \).
If \( A \in \mathbb{C}^{n \times n} \) is symmetric positive definite, \( b \in \mathbb{C}^n \), and \( x_0 \in \mathbb{C}^n \) is an initial guess, then the following conjugate gradients algorithm computes : \( x \in \mathbb{C}^n \) so \( Ax = b \)

<table>
<thead>
<tr>
<th>Conjugate Gradients</th>
</tr>
</thead>
<tbody>
<tr>
<td>( x_0 ) = initial guess</td>
</tr>
<tr>
<td>( \gamma_0 = b - Ax_0 )</td>
</tr>
<tr>
<td>( k = 0 )</td>
</tr>
<tr>
<td><strong>while</strong> ( \gamma_k \neq 0 )</td>
</tr>
<tr>
<td>( k = k + 1 )</td>
</tr>
<tr>
<td><strong>if</strong> ( k = 1 )</td>
</tr>
<tr>
<td>( d_1 = \gamma_0 )</td>
</tr>
<tr>
<td><strong>else</strong> ( k = 1 )</td>
</tr>
<tr>
<td>( \beta_k = \frac{\gamma_k^H \gamma_{k-1}}{\gamma_{k-2}^H \gamma_{k-2}} )</td>
</tr>
<tr>
<td>( d_k = \gamma_{k-1} + \beta_k d_{k-1} )</td>
</tr>
<tr>
<td><strong>end</strong></td>
</tr>
<tr>
<td>( \alpha_k = \frac{\gamma_k^H \gamma_k}{d_k^H A d_k} )</td>
</tr>
<tr>
<td>( x_k = x_{k-1} + \alpha_k d_k )</td>
</tr>
<tr>
<td>( \gamma_k = \gamma_{k-1} - \alpha_k A d_k )</td>
</tr>
<tr>
<td><strong>end</strong></td>
</tr>
<tr>
<td>( x = x_k )</td>
</tr>
</tbody>
</table>

Table 3.2: Conjugate Gradients

It is shown in [22] that, CG algorithm produces iterates \( \{x_k\} \) satisfies

\[
\| x - x_k \|_A \leq 2 \| x - x_0 \|_A \left( \frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right) ^\kappa
\]

(3.24)

where \( \kappa = \kappa_2(A) \). Therefore, the CG method converges very fast in the A-norm if

\[
\kappa_2(A) \approx 1.
\]

(3.25)

Moreover, when \( A = I + B \) is an n-by-n symmetric positive definite matrix and

\[
A = I + B
\]

(3.26)
where \( \text{rank}(B) = r \), the CG algorithm converges in at most \( r + 1 \) steps.

3.3 The Lanczos Connection

The Lanczos method is a technique that can be used to solve certain large, sparse, symmetric eigenproblems

\[
Ax = \lambda x. \tag{3.27}
\]

The method involves partial tridiagonalizations of the given matrix \( A \). This section explores the connection between the conjugate gradient and Lanczos method.

Define the matrix of residuals by:

\[
\gamma = [\gamma_1, \gamma_2, \cdots, \gamma_{k-1}] \tag{3.28}
\]

and the upper bidiagonal matrix \( B_k \) with all one and \([-\beta_2, -\beta_3, \cdots, -\beta_k]\) on its diagonal and subdiagonal elements respectively. Since the directional vectors are \( A \)-conjugate, it shows that

\[
\gamma^H A \gamma = B_k^H \text{diag}(d_1^H A d_1, \cdots, d_k^H A d_k) B_k \tag{3.29}
\]

is tridiagonal matrix. Let

\[
\Delta = \text{diag} (\rho_0, \cdots, \rho_{k-1}) \tag{3.30}
\]

where

\[
\rho_i = \left\| \gamma_i \right\|_2 \tag{3.31}
\]

then, the columns of the matrix \( \gamma \Delta^{-1} \) are the Lanczos vectors

\[
q_i = \pm \gamma_{i-1}/\rho_{i-1} \quad i = 1 : k \tag{3.32}
\]
and the tridiagonal matrix associated with these Lanczos vectors is given by

\[ T_k = \Delta^{-1} B_k^H \text{diag}(d_i^H A d_i) B_k \Delta^{-1} \]  \hspace{1cm} (3.33)

The diagonal and subdiagonal of this matrix involve quantities that are readily available during the CG iteration. Thus, it can provide good estimates of \( A \)'s extremal eigenvalues.

### 3.4 Equivalence of Multistage Wiener Filter and CG

The multistage Wiener filter (MWF) is a novel structure for reduced rank filtering, in which the reduced rank subspace is determined by maximizing cross correlations between the desired signal and the observed data. Under certain conditions the MWF and the method of conjugate gradients are equivalent in the sense that they produce the same weight vectors at each stage. That is

\[ w^{(cg)}_k = w^{(mwf)}_k \]  \hspace{1cm} (3.34)

for all \( k \) when applied to the Wiener-Hopf equations

\[ R w = s. \]  \hspace{1cm} (3.35)

Both algorithms make the same reduced rank estimates of the full rank Wiener filter

\[ w = R^{-1} s. \]  \hspace{1cm} (3.36)

The general equivalences between conjugate direction and MWF was established in [48], wherein the direction vectors of a conjugate direction filter and the stage wise vectors of a multistage filter are related through a one-term autoregressive recursion. For every MWF there is a family of equivalent conjugate direction Wiener filters, with identical
subspaces, gradients, and MSEs. And vice-versa. If the conjugate direction filter is a conjugate gradient filter, then the equivalent stage-wise filter is an orthogonal multistage filter, and vice-versa.
Chapter 4
Conjugate Gradient Adaptive Matched Filter

The conjugate gradient (CG) algorithm offers a computationally efficient approach to reduced-rank detection. A recent study of the CG algorithm used with the non-adaptive matched-filter (MF) for reduced-rank detection leads to not only computational complexity reduction but also detection performance benefits [30]. The problem of interest to this work is to examine the performance of the CG for adaptive reduced-rank detection relative to other data-dependent reduced-rank detectors.

4.1 Data Model

Consider the following signal detection problem [12, 21, 26, 29–31, 34, 35, 44, 45, 48, 58]:

\[ H_0 : \quad x = d \]
\[ H_1 : \quad x = \alpha s + d \]  \hspace{1cm} (4.1)

where \( x \in \mathbb{C}^{M \times 1} \) denotes the observation, \( s \) a deterministic signal of an unknown complex amplitude \( \alpha \), and \( d \) a disturbance signal that is assumed to be spatially and temporally correlated. The disturbance \( d \) is often modeled as a Gaussian random vector with zero-mean and space-time covariance matrix \( R \in \mathbb{C}^{M \times M} \). As a result, \( x \sim \mathcal{CN}(\alpha s, R) \), where \( \alpha = 0 \) under \( H_0 \) and \( \alpha \neq 0 \) under \( H_1 \).

The above problem is referred to as the space-time adaptive processing (STAP) problem in array radar literature. Specifically, the \( M \times 1 \) vector \( x \) contains the samples obtained with an array of \( J \) elements and over \( N \) temporal pulses, \( M = JN \) denotes the joint space-time dimension, \( s \) is the space-time steering vector, and \( d \) contains clutter and
noise. For a side-looking uniform linear array, the space-time steering vector $s$ is given by $s = s_t \otimes s_s$ where $s_t = (1/\sqrt{N}) \begin{bmatrix} e^{i2\pi f_d} & \ldots & e^{i2\pi(N-1)f_d} \end{bmatrix}^T$ is the spatial steering vector with a normalized Doppler frequency $f_d$, $s_s = (1/\sqrt{J}) \begin{bmatrix} e^{i2\pi f_s} & \ldots & e^{i2\pi(J-1)f_s} \end{bmatrix}^T$ is the spatial steering vector with a normalized spatial frequency $f_s$ and $\otimes$ denotes the Kronecker product.

The optimum detector for this problem is the matched filter (MF):

$$T_{MF} = \frac{|s^H R^{-1} x|^2}{s^H R^{-1} s} \overset{H_1}{\gtrsim} \eta_{MF}$$

where $\eta_{MF}$ is the MF detector threshold. The weight vector of the MF is

$$w = R^{-1} s$$

The MF test statistic can be alternatively expressed as

$$T_{MF} = \frac{|w_k^H x|^2}{w_k^H R w_k} \overset{H_1}{\gtrsim} \eta_{MF}$$

4.2 Conjugate Gradient Matched Filter

The authors in [30] uses CG algorithm to solve the Wiener-Hopf equation underlying the MF, which leads to a family of linear CG-MF detector that converge to the MF in a fixed number of iterations. The CG-MF test can be present as:

$$T_{CG-MF} = \frac{|w_k^H x|^2}{w_k^H R w_k} \overset{H_1}{\gtrsim} \eta_{CG-MF}$$

where $\eta_{CG-MF}$ is the threshold, and the weigh vector can be solved by the following algorithm:
Conjugate-Gradient MF Detector

Input:
- \( M \times M \) sample covariance matrix \( R \)
- \( M \times 1 \) signal vector \( s \)

Output:
- \( M \times 1 \) CG-MF weight vectors \( w_k, k = 1, 2, \cdots \)

Algorithm:
1) Initialization:
   - Initial conjugate-direction vector: \( d_1 = s \)
   - Initial gradient vector: \( \gamma_1 = -s \)
   - Initial weight vector: \( w_0 = 0 \)

2) Iterations:
   - for \( k = 1, 2, \cdots \), till convergence \( (k \leq M) \)
   do
     (a) Update the step size \( \alpha_k \):
     \[
     \alpha_k = \frac{||\gamma_k||^2}{d_k^H R d_k} \tag{4.6}
     \]
     (b) Update the solution \( w_k \):
     \[
     w_k = w_{k-1} + \alpha_k d_k \tag{4.7}
     \]
     (c) Update the gradient vector \( \gamma_{k+1} \):
     \[
     \gamma_{k+1} = \gamma_k + \alpha_k R d_k \tag{4.8}
     \]
     (d) Update the conjugate-direction vector \( d_{k+1} \):
     \[
     d_{k+1} = d_k - \frac{||\gamma_{k+1}||^2}{||\gamma_k||^2} \gamma_{k+1} \tag{4.9}
     \]
   end for

Table 4.1: Conjugate-Gradient MF Detector
4.2.1 General Covariance Matrix

The CG-MF detector can be present as

\[ w_k = D_k \alpha_k \]  \hspace{1cm} (4.10)

where

\[ \alpha_k = [\alpha_1, \alpha_2, \ldots, \alpha_k]^T \]  \hspace{1cm} (4.11)

contains the step sizes and

\[ D_k = [d_1, d_2, \ldots, d_k] \]  \hspace{1cm} (4.12)

consists of the first \( k \) conjugate direction directors. Moreover, \( D_k \) diagonalizes the covariance matrix

\[ D_k^H R D_k = \Lambda_k \]  \hspace{1cm} (4.13)

where

\[ \Lambda_k = \text{diag}(u_1^2, u_2^2, \ldots, u_k^2) \]  \hspace{1cm} (4.14)

and

\[ u_k = (d_k^H R d_k)^{\frac{1}{2}}. \]  \hspace{1cm} (4.15)

This allows \( \alpha_k \) to be compactly expressed as

\[ \alpha_k = \Lambda_k^{-1} D_k^H s \]  \hspace{1cm} (4.16)

which gives the following close-form expression for \( w_k \):

\[ w_k = D_k \Lambda_k^{-1} D_k^H s. \]  \hspace{1cm} (4.17)
The following are true for the $k$th CG-MF detector.

a) $w_k$ is a linear minimum mean square estimator that minimizes

$$w_k = \arg \min_{w \in \mathcal{K}(R,s,k)} \left\{ E \left\| a - w^H x \right\| \right\}$$

among all linear estimators within the $k$-dimensional Krylov subspace:

$$\mathcal{K}(R,s,k) = \text{span}\{q_1, q_2, \cdots, q_k\}.$$  \hspace{1cm} (4.19)

b) $w_k$ yields the largest output SINR

$$\rho_k = \frac{|a|^2 |w_k^H s|^2}{w_k^H R w_k}$$

among all linear detectors within $\mathcal{K}(R,s,k)$.

### 4.2.2 Structured Covariance Matrix

The disturbance covariance matrix can be assumed to have a low-rank structure:

$$R = R_i + \sigma_n^2 I$$

where the rank-$r$ ($r \leq M$) positive semi-definite matrix is due to the interference, and the scaled identity matrix $\sigma_n^2 I$ due to the noise subspace.

It considers the perturbed version of $R$ as

$$R_p = R_i + \sigma_n^2 I + \Delta_0$$

(4.22)
where, $\Delta_0$ is a Hermitian perturbation matrix assumed to be small, i.e.,

$$\|\Delta_0\| \ll \|R\|. \quad (4.23)$$

With this structure, it shows that the output SINR of the MF detector $w_{MF}$ and the CG-MF detector $w_{r+1}$ are identical within a first-order approximation:

$$\rho_{MF} - \rho_{r+1} = o(\|\Delta\|) \quad (4.24)$$

where

$$\rho_{r+1} = |\alpha|^2 \frac{|w_{r+1}^Hs|^2}{w_{r+1}^HRw_{r+1}} \quad (4.25)$$

and $\rho_{MF}$ is similarly defined by replacing $w_{r+1}$ with $w_{MF}$.

Therefore, if the covariance matrix has a low-rank structure disrupted by a small perturbation component, the SINR of the $(r+1)th$ CG-MF detector is nearly identical to that of the MF detector.

### 4.3 Conjugate Gradient Adaptive Matched Filter

We represent the AMF test which is a popular solution to the detection problem (1) as below,

$$\frac{|\hat{w}^Hx|^2}{\hat{w}^Hs} \overset{H_1}{\gtrless} \eta_{AMF} \quad (4.26)$$

where $\eta_{AMF}$ denotes the test threshold, and the linear weight vector of the AMF detector is

$$\hat{w} = \hat{R}^{-1}s \quad (4.27)$$
where $\hat{\mathbf{R}}$ denotes the sample covariance matrix which is an estimate of $\mathbf{R}$ obtained from the secondary data $\{x_i\}$:

$$
\hat{\mathbf{R}} = \frac{1}{L} \sum_{l=1}^{L} x_l x_l^H.
$$

(4.28)

4.3.1 Implementation of CG-AMF

The CG algorithm was employed in [30] along with the non-adaptive matched filter (MF) for reduced-rank detection. It is straightforward to extend it for adaptive reduced-rank detection. In particular, The CG algorithm can iteratively find a sequence of approximations $\hat{\mathbf{w}}_k$, $k = 1, 2, \ldots$, to the AMF weight vector $\hat{\mathbf{w}}$, each of which can be used to form a detector as in (4.26). As such, the CG iterations yield a family of detectors, referred to as the CG-AMF detectors. For the sake of discussion and also to introduce necessary notation, we summarize the iterative algorithm that can be used to compute the CG-AMF detector in Table 4.2.

The CG algorithm converges to the full-rank AMF solution in no more than $M$ iterations [23]. Even faster convergence is possible if the covariance matrix is structured. For example, it is known that if the covariance matrix $\hat{\mathbf{R}}$ contains a rank-$r$ component plus an identity matrix, then the CG algorithm converges in no more than $r + 1$ iterations [23].

In training-limited environment, however, convergence to the full-rank adaptive solution is often not the desired objective, since the full-rank solution may suffer considerable estimation error induced performance loss while a reduced-rank solution may yield better detection performance. Since the CG iteration yields a family of reduced-rank CG-AMF detectors $\hat{\mathbf{w}}_k$, it is naturally of interest to investigate the statistical behaviors of these CG-AMF detectors and determine how the detection performance changes as CG iterates.
Conjugate-Gradient AMF Detector

**Input:**

- $M \times M$ sample covariance matrix $\hat{R}$
- $M \times 1$ signal vector $s$

**Output:**

- $M \times 1$ CG-AMF weight vectors $\hat{w}_k$, $k = 1, 2, \cdots$

**Algorithm:**

1) **Initialization:**

   Initial conjugate-direction vector: $\hat{d}_1 = s$

   Initial gradient vector: $\hat{\gamma}_1 = -s$

   Initial weight vector: $\hat{w}_0 = 0$

2) **Iterations:**

   for $k = 1, 2, \cdots$, till convergence ($k \leq M$)

   do

   (a) Update the step size $\hat{\alpha}_k$:

   $$\hat{\alpha}_k = \frac{\|\hat{\gamma}_k\|^2}{\hat{d}_k^H \hat{R} \hat{d}_k} \quad (4.29)$$

   (b) Update the solution $\hat{w}_k$:

   $$\hat{w}_k = \hat{w}_{k-1} + \hat{\alpha}_k \hat{d}_k \quad (4.30)$$

   (c) Update the gradient vector $\hat{\gamma}_{k+1}$:

   $$\hat{\gamma}_{k+1} = \hat{\gamma}_k + \hat{\alpha}_k \hat{R} \hat{d}_k \quad (4.31)$$

   (d) Update the conjugate-direction vector $\hat{d}_{k+1}$:

   $$\hat{d}_{k+1} = \hat{d}_k \frac{\|\hat{\gamma}_{k+1}\|^2}{\|\hat{\gamma}_k\|^2} - \hat{\gamma}_{k+1} \quad (4.32)$$

   end for

---

Table 4.2: Conjugate-Gradient AMF Detector
4.4 Performance of the CG-AMF Detector

In this section, we consider the performance of the CG-AMF detector in terms of its output SINR. Since in adaptive detection, the output SINR is a random variable, it is necessary to consider the statistical distribution of the output SINR.

4.4.1 Output SINR of the CG-AMF Detector

The CG-AMF weight vector $\hat{w}_k$ after $k$ iterations is in the $k$-dimensional Krylov subspace $\mathcal{K}(\hat{R}, s, k)$:

$$\mathcal{K}(\hat{R}, s, k) \triangleq \text{span}\{s, \hat{R}s, \hat{R}^2s, \cdots , \hat{R}^{k-1}s\}. \quad (4.33)$$

The normalized gradient vectors $\hat{q}_1, \hat{q}_2, \cdots , \hat{q}_k$, where

$$\hat{q}_i = \frac{\hat{\gamma}_i}{\|\hat{\gamma}_i\|_2}, \quad 1 \leq i \leq k, \quad (4.34)$$

also span the same $k$-dimensional Krylov subspace:

$$\mathcal{K}(\hat{R}, s, k) = \text{span}\{\hat{q}_1, \hat{q}_2, \cdots , \hat{q}_k\}. \quad (4.35)$$

Therefore, by defining the matrix of residuals $\hat{Q}_k \in \mathbb{C}^{M \times k}$ as:

$$\hat{Q}_k = \begin{bmatrix} \hat{q}_1 & \hat{q}_2 & \cdots & \hat{q}_k \end{bmatrix} \quad (4.36)$$

$\hat{w}_k$ can be compactly expressed as

$$\hat{w}_k = \hat{Q}_k \hat{a} \quad (4.37)$$

where

$$\hat{a} = \begin{bmatrix} \hat{a}_1 & \hat{a}_2 & \cdots & \hat{a}_k \end{bmatrix}^T \quad (4.38)$$
contains the coefficients which can be determined as follows.

Specifically, \( \hat{w}_k \) can be considered as the \( \hat{R} \)-orthogonal projection of \( \hat{w} = \hat{R}^{-1}s \) onto the Krylov subspace \( K(\hat{R}, s, k) \) [23], which means that the \( \hat{R} \)-norm of the approximation error is minimized over all vectors in the Krylov subspace or, equivalently, the column space of \( \hat{Q}_k \). That is,

\[
\| \hat{w} - \hat{w}_k \|_{\hat{R}} = \min_{\hat{a}} \| \hat{w} - \hat{w}_k \|_{\hat{R}}
\]

\[
= \min_{\hat{a}} \| \hat{R}^{\frac{1}{2}} \hat{w} - \hat{R}^{\frac{1}{2}} \hat{w}_k \|_{\hat{R}}
\]  

where the \( \hat{R} \)-norm is defined as

\[
\| \cdot \|_{\hat{R}} = \| \hat{R}^{\frac{1}{2}} (\cdot) \|.
\]  

Denote by \( \epsilon \) the approximation error

\[
\epsilon \triangleq \hat{R}^{\frac{1}{2}} \hat{w} - \hat{R}^{\frac{1}{2}} \hat{w}_k = \hat{R}^{-\frac{1}{2}}s - \hat{R}^{\frac{1}{2}} \hat{Q}_k \hat{a}.
\]  

Since \( \hat{R}^{\frac{1}{2}} \hat{w}_k \) is the orthogonal projection of the vector \( \hat{R}^{\frac{1}{2}} \hat{w} \) onto \( \hat{R}^{\frac{1}{2}} K(\hat{R}, s, k) \), we have

\[
\epsilon \perp \hat{R}^{\frac{1}{2}} \hat{Q}_k
\]  

namely,

\[
\left( \hat{R}^{-\frac{1}{2}} s - \hat{R}^{\frac{1}{2}} \hat{Q}_k \hat{a} \right) \hat{R}^{\frac{1}{2}} \hat{Q}_k = 0
\]  

from which we obtain

\[
\hat{a} = \left( \hat{Q}_k^H \hat{R} \hat{Q}_k \right)^{-1} \hat{Q}_k^H s.
\]
Therefore, the CG-AMF weight vector can be written as
\[
\hat{w}_k = \hat{Q}_k \left( \hat{Q}_k^H \hat{R} \hat{Q}_k \right)^{-1} \hat{Q}_k^H s
\]  (4.45)

where \( \hat{Q}_k \) plays the same role as the data-dependent reduced-rank transform matrix \( T \) in equation (2.15).

For the full-rank AMF detector, the output SINR is given by
\[
\rho \triangleq \frac{|\alpha|^2 |\hat{w}^H s|^2}{\hat{w}^H R \hat{w}}.
\]  (4.46)

Replacing \( \hat{w} \) in (4.46) by \( \hat{w}_k \) of (4.45), we have the output SINR of the CG-AMF detector with \( k \) CG iterations as
\[
\rho_k = \frac{|\alpha|^2 \left| s^H \hat{Q}_k \left( \hat{Q}_k^H \hat{R} \hat{Q}_k \right)^{-1} \hat{Q}_k^H s \right|^2}{s^H \hat{Q}_k \left( \hat{Q}_k^H \hat{R} \hat{Q}_k \right)^{-1} \hat{Q}_k^H \hat{R} \hat{Q}_k \left( \hat{Q}_k^H \hat{R} \hat{Q}_k \right)^{-1} \hat{Q}_k^H s}.
\]  (4.47)

### 4.4.2 Low-Rank Approximation

Henceforth, we consider the case when the disturbance covariance matrix has some low-rank structure. Specifically, we assume that \( R \) has the following structure:
\[
R = R_i + \sigma_n^2 I
\]  (4.48)

where \( R_i \) is a rank-\( r \) (\( r < M \)) positive semi-definite matrix and \( I \) is an identify matrix.

The above structure of the covariance matrix is frequently encountered in practice. For example, in airborne radar applications, \( R \) may consist of two components, namely a low-rank \( R_i \) due to the presence of clutter and jamming and a scaled identity \( \sigma_n^2 I \) due to the presence of a thermal noise with variance \( \sigma_n^2 \). The rank \( r \) is typically much smaller than the
joint spatial-temporal dimension $M = JN$ where $J$ is the number of array elements and $N$ is the number of pulses. Specifically, if the disturbance is primarily due to ground clutter and thermal noise, then according to Brennan’s rule [5], the rank of the covariance matrix for the full-dimensional MF is approximately

$$r \approx \lceil J + (N - 1)\beta \rceil$$

(4.49)

where $\beta = 2v_g T_r / d$, $v_g$ is the platform velocity, $T_r$ is the pulse repetition period, $d$ is the antenna element spacing, and $\lceil \cdot \rceil$ rounds a real-valued number towards infinity.

In airborne radar detection, the clutter is often the dominating factor compared with the noise. Let $\lambda(\mathbf{R}) = \{\lambda_1, \cdots, \lambda_M\}$ denote the spectrum of $\mathbf{R}$ with the eigenvalues in descending order: $\lambda_1 \geq \cdots \geq \lambda_r \gg \lambda_{r+1} = \cdots = \lambda_M = \sigma_n^2$, where the first $r$ dominant eigenvalues are due to the clutter while the rest are due to the noise. Similarly, let $\lambda(\hat{\mathbf{R}}) = \{\hat{\lambda}_1, \cdots, \hat{\lambda}_M\}$ denote the spectrum of the sample covariance matrix $\hat{\mathbf{R}}$ with descending order: $\hat{\lambda}_1 \geq \cdots \geq \hat{\lambda}_r \gg \hat{\lambda}_{r+1} \geq \cdots \geq \hat{\lambda}_M$. For a sufficiently large $L$, the number of training signals [cf. (4.28)], the eigenvalues of $\hat{\mathbf{R}}$, like $\mathbf{R}$, are clustered with $r$ dominant ones and the rest close to the noise variance $\sigma_n^2$.

Our goal here is to obtain a useful approximation of the output SINR (4.47) of the CG-AFM detector for analysis, when the covariance matrix is low rank and the clutter is dominating. Our approach is based on a connection of the CG iterations and the Lanczos tridiagonalization. Specifically, the normalized gradient vectors $\hat{\mathbf{q}}_l$ in $\hat{\mathbf{Q}}_k$ are also called Lanczos vectors which tridiagonalize $\hat{\mathbf{R}}$ [23],

$$\hat{\mathbf{Q}}_k^H \hat{\mathbf{R}} \hat{\mathbf{Q}}_k = \hat{\mathbf{T}}_k$$

(4.50)
where $\hat{T}_k$ is a tridiagonal matrix. Consider the eigenvalue decomposition of $\hat{T}_k \in \mathbb{C}^{k \times k}$

$$\hat{T}_k = \hat{U}\hat{\Lambda}\hat{U}^H$$

(4.51)

where the diagonal matrix $\hat{\Lambda}$ consists of the $k$ eigenvalues of $\hat{T}_k$ in descending order: $\hat{\theta}_1 \geq \hat{\theta}_2 \geq \cdots \geq \hat{\theta}_k$, and the unitary matrix $\hat{U} = \begin{bmatrix} \hat{u}_1 & \hat{u}_2 & \cdots & \hat{u}_k \end{bmatrix}$ contains the corresponding eigenvectors. Based on (4.51) and the Lanczos decomposition (4.50), we can diagonalize $\hat{R}$ as:

$$\hat{P}^H\hat{R}\hat{P} = \hat{\Lambda}$$

(4.52)

in which the column orthogonal matrix $\hat{P} = \hat{Q}_k\hat{U} = \begin{bmatrix} \hat{p}_1 & \hat{p}_2 & \cdots & \hat{p}_k \end{bmatrix}$ contains the Ritz vectors of $\hat{R}$, and $\hat{\theta}_i$ are the Ritz values of $\hat{R}$ or the “Lanczos estimates” of the eigenvalues of $\hat{R}$. The Ritz values are known to converge rapidly (with respect to $k$) to the extremal eigenvalues, i.e., eigenvalues at the edges of the spectrum $\lambda(\hat{R})$ [23]. For the considered case, the spectrum $\lambda(\hat{R})$ has two clustered groups. One contains the $M - r$ noise eigenvalues which are uniformly spread around the noise variance $\sigma_n^2$, and the one contains the $r$ clutter eigenvalues which are significantly larger and may be considered as “outliers” compared to the noise eigenvalues. For such a case, the Ritz values are known to converge to the outlying eigenvalues first [53]. In particular, under the condition $k \leq r + 1$, where $k$ denotes the number of CG iterations, $\hat{\theta}_k$ converges to the smallest (i.e., extremal) eigenvalue $\hat{\lambda}_M$, $\hat{\theta}_1$ converges to the largest (i.e., extremal) $\hat{\lambda}_1$, whereas each of the $k - 2$ Ritz values $\{\hat{\theta}_2, \cdots, \hat{\theta}_{k-1}\}$ converges to one of the outlying eigenvalues in $\{\hat{\lambda}_2, \cdots, \hat{\lambda}_r\}$. The rapid convergence of the Ritz values implies that with a few iterations,
we have $\hat{\theta}_1 \geq \hat{\theta}_2 \geq \cdots \geq \hat{\theta}_{k-1} \gg \hat{\theta}_k$. Therefore, we have

$$(\hat{Q}_k^H \hat{R} \hat{Q}_k)^{-1} = \hat{T}_k^{-1} = \hat{U} \hat{\Lambda}^{-1} \hat{U}^H$$

$$= \sum_{i=1}^{k} \hat{u}_i \hat{\theta}_i^{-1} \hat{u}_i^H \approx \hat{u}_k \hat{\theta}_k^{-1} \hat{u}_k^H.$$ (4.53)

It follows that the output SINR (4.47) of the CG-AMF detector can be simplified as

$$\rho_k \approx |\alpha|^2 \left| s^H \hat{Q}_k \hat{u}_k \hat{\theta}_k^{-1} \hat{u}_k^H \hat{Q}_k^H s \right|^2$$

$$= \frac{|\alpha|^2 \hat{p}_k^H \hat{R} \hat{p}_k \hat{Q}_k^H s}{\hat{p}_k^H \hat{R} \hat{p}_k}$$ (4.54)

where $\hat{p}_k = \hat{Q}_k \hat{u}_k$ is the Ritz vector corresponding to the minimum Ritz value of $\hat{R}$.

### 4.4.3 Statistical Analysis

The output SINR $\rho_k$ is a random variable. To understand its statistical behavior, we derive its probability density function (PDF) by using (4.54). Our approach is similar to standard subspace perturbation analysis, e.g., as in [27]), but we have to cater to the particular structure of (4.54) which involves CG iterations and Lanczos transform. To begin with, we consider the following eigendecompositions:

$$T_k = Q_k^H R Q_k = U \Lambda U^H$$ (4.55)

$$\hat{T}_k = \hat{Q}_k^H \hat{R} \hat{Q}_k = \hat{U} \hat{\Lambda} \hat{U}^H$$ (4.56)

and

$$\hat{T}_k = Q_k^H \hat{R} Q_k = \hat{U} \hat{\Lambda} \hat{U}^H$$ (4.57)
where \( Q_k \) is the Lanczos transform matrix obtained from the true covariance matrix \( R \), the diagonal matrices \( \Lambda = \text{diag} \left[ \theta_1, \theta_2, \ldots, \theta_k \right] \), \( \hat{\Lambda} = \text{diag} \left[ \hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_k \right] \) and \( \tilde{\Lambda} = \text{diag} \left[ \tilde{\theta}_1, \tilde{\theta}_2, \ldots, \tilde{\theta}_k \right] \) contain the \( k \) eigenvalues of \( T_k \), \( \hat{T}_k \) and \( \tilde{T}_k \) in descending order, and finally the unitary matrices \( U, \hat{U} \) and \( \tilde{U} \) contain the corresponding eigenvectors of \( T_k \), \( \hat{T}_k \) and \( \tilde{T}_k \), respectively.

Note that both the numerator and denominator of (4.54) involve a quadratic form of the Ritz vector \( \hat{p}_k \) which consists of a data-dependent Lanczos transform matrix \( \hat{Q}_k \) and a data-dependent eigenvector \( \hat{u}_k \) and, therefore, is rather complicated. To simplify the problem, we consider the approximation \( \hat{p}_k \approx Q_k \tilde{u}_k \) with \( \tilde{u}_k \) denoting the eigenvector corresponding to the minimum eigenvalue of \( \tilde{T}_k \). This leads to

\[
\rho_k \approx \frac{|\alpha|^2 \tilde{u}_k^H Q_k^H s s^H Q_k \tilde{u}_k}{\tilde{u}_k^H Q_k^H R Q_k \tilde{u}_k} = \frac{|\alpha|^2 \tilde{u}_k^H s_1 s_1^H \tilde{u}_k}{\tilde{u}_k^H T_k \tilde{u}_k} \tag{4.58}
\]

where \( s_1 = Q_k^H s \) is the Lanczos transformed steering vector.

Next, we represent \( \tilde{u}_k \) in terms of the subspace of \( T_k \). Define the \( k \times k \) matrix

\[
\tilde{\Lambda} \triangleq U^H \hat{T}_k U \tag{4.59}
\]

which is in general complex-valued and non-diagonal. It can be considered a perturbed version of the diagonal \( \Lambda \). Let us define the perturbation matrix

\[
H \triangleq \sqrt{L} (\tilde{\Lambda} - \Lambda) \tag{4.60}
\]

From (4.57) and (4.59), we have

\[
\tilde{\Lambda} = U^H \tilde{U} \tilde{\Lambda} \tilde{U} U = Y \bar{\Lambda} \bar{Y}^H \tag{4.61}
\]
where the unitary matrix $Y \triangleq U^H\tilde{U}$ is the product of two unitary matrices: $U^H$ and $\tilde{U}$. Let $Z \triangleq \sqrt{L}(Y - I_k)$ and then

$$Y = I_k + \frac{1}{\sqrt{L}}Z. \quad (4.62)$$

Substituting (4.62) back into (4.61), we have

$$\tilde{\Lambda} = \left(I_k + \frac{1}{\sqrt{L}}Z\right) \tilde{\Lambda} \left(I_k + \frac{1}{\sqrt{L}}Z\right)^H$$

$$= \left(I_k + \frac{1}{\sqrt{L}}Z\right) \left(\Lambda + \frac{1}{\sqrt{L}}D\right) \left(I_k + \frac{1}{\sqrt{L}}Z\right)^H$$

$$= \Lambda + \frac{1}{\sqrt{L}} (Z\Lambda + D + \Lambda Z^H) + M \quad (4.63)$$

where $D \triangleq \sqrt{L} (\tilde{\Lambda} - \Lambda)$ and $M \triangleq \frac{1}{L} (ZD + DZ^H + Z\Lambda Z^H) + \frac{1}{\sqrt{L^3}}ZDZ^H$. Since $M \sim O(1/L)$ which contains quantities that are in the order of $1/L$ and higher order terms, it can be neglected for large $L$. Thus we have

$$\tilde{\Lambda} \approx \Lambda + \frac{1}{\sqrt{L}} (Z\Lambda + D + \Lambda Z^H). \quad (4.64)$$

It follows from (4.64) and (4.60) that

$$H \approx Z\Lambda + D + \Lambda Z^H. \quad (4.65)$$

Henceforth, we assume $k > 1$. The case of $k = 1$ is of little interest, since the detector reduces to a non-adaptive detector whose weight vector is $s$. For $k > 1$, we partition the
\[ k \times k \text{ matrix of eigenvalues } \Lambda \text{ as:} \]
\[
\Lambda = \begin{bmatrix}
\Lambda_1 & 0 \\
0 & \theta_k
\end{bmatrix}
\]

(4.66)

where \( \theta_k \) is the minimum eigenvalue. Similarly, \( Z \) can be partitioned as

\[
Z = \begin{bmatrix}
Z_{1,1} & z_{1,2} \\
Z_{2,1} & z_{2,2}
\end{bmatrix}
\]

(4.67)

where \( Z_{1,1} \in \mathbb{C}^{(k-1) \times (k-1)} \), \( z_{1,2} \in \mathbb{C}^{(k-1) \times 1} \), \( z_{2,1} \in \mathbb{C}^{1 \times (k-1)} \), and \( z_{2,2} \) is a scalar. Let the matrix \( H \) be partitioned in a similar way. It is easy to see from (4.65)–(4.67) and the fact that \( D = \sqrt{L} (\tilde{\Lambda} - \Lambda) \) is a diagonal matrix that

\[
h_{1,2} = \theta_k z_{1,2} + \Lambda_1 z_{2,1}^H.
\]

(4.68)

The above expression can be further simplified by using the following result.

**Lemma 1** Within the approximation of \( O(1/L) \),

\[
z_{1,2} \approx -z_{2,1}^H.
\]

(4.69)

**Proof:** See Appendix 4.7.1.

Applying the relationship (4.69) to (4.68), we obtain that

\[
z_{1,2} = (\theta_k I_{k-1} - \Lambda_1)^{-1} h_{1,2}.
\]

(4.70)

The next result establishes the asymptotic distribution of \( z_{1,2} \).
Lemma 2 The limiting distribution of the \((k-1) \times 1\) vector \(z_{1,2} = \begin{bmatrix} z_{1,k} & z_{2,k} & \cdots & z_{k-1,k} \end{bmatrix}^T\)
is normal with zero-mean and

\[
E[z_{m,k}z_{n,k}^*] = \begin{cases} 
\frac{\theta_k \theta_m}{(\theta_k - \theta_m)^2}, & m = n \\
0, & m \neq n
\end{cases} \tag{4.71}
\]

where \(m, n = 1, 2, \cdots, k - 1\).

Proof: See Appendix 4.7.2.

Partition the eigenvector matrix of \(T_k\) as \(U = \begin{bmatrix} U_1 & u_k \end{bmatrix}\), where \(u_k \in \mathbb{C}^{k \times 1}\) is the eigenvector corresponding to the minimum eigenvalue \(\theta_k\) of \(T_k\), and \(U_1 \in \mathbb{C}^{k \times (k-1)}\) contains the other eigenvectors. Note that \(U_1\) and \(u_k\) are orthogonal. i.e., \(U_1^H u_k = 0\). Similarly, let \(\tilde{U} = \begin{bmatrix} \tilde{U}_1 & \tilde{u}_k \end{bmatrix}\). Recall \(\tilde{U} = UY\) and (4.62). It follows that

\[
\tilde{u}_k = U_1 y_{1,2} + u_k y_{2,2} = \frac{1}{\sqrt{L}} U_1 z_{1,2} + u_k (1 + \frac{z_{2,2}}{\sqrt{L}}) \tag{4.72}
\]

where \(y_{1,2}\) and \(y_{2,2}\) are the partitioning components of \(Y\).

We now consider the numerator of the output SINR (4.58). Expanding \(\tilde{u}_k\) using (4.72), we can write the numerator as:

\[
|\alpha|^2 \tilde{u}_k^H s_1 s_1^H \tilde{u}_k = |\alpha|^2 s_1^H \left( \frac{2}{\sqrt{L}} \Re \{ U_1 z_{1,2} u_k^H \} + u_k u_k^H s_1 \right) + O(1/L). \tag{4.73}
\]

Suppose that the interference is mainly located in the sidelobe directions (i.e., no major mainlobe interference). Then, the projection of the transformed steering vector \(s_1\) to the interference subspace spanned by \(U_1\) is small compared to the projection to the noise sub-
space spanned by \( u_k \). That is, 
\[
|u_k^H s_1|^2 >> s_1^H U_1 z_{1,2} u_k^H s_1.
\]
This leads to the following approximation for the numerator of \( \rho_k \):
\[
|\alpha|^2 \tilde{u}_k^H s_1 s_1^H \tilde{u}_k \approx |\alpha|^2 u_k^H s_1 s_1^H u_k.
\] (4.74)

Similarly, the denominator of \( \rho_k \) can be simplified by using (4.72):
\[
\tilde{u}_k^H T_k \tilde{u}_k = \left[ \frac{1}{\sqrt{L}} U_1 z_{1,2} + u_k \left( 1 + \frac{z_{2,2}}{\sqrt{L}} \right) \right]^H T_k
\]
\[
= \frac{1}{L} z_{1,2}^H \Lambda_1 z_{1,2} + \theta_k \left( 1 + \frac{z_{2,2}}{\sqrt{L}} \right)^* \left( 1 + \frac{z_{2,2}}{\sqrt{L}} \right)
\] (4.75)

where the second equality is obtained by using the eigen-pair relations: \( U_1^H T_k U_1 = \Lambda_1 \), \( u_k^H T_k u_k = \theta_k \) and the orthogonality relation: \( U_1^H u_k = 0 \). Since \((1 + z_{2,2}/\sqrt{L})^*(1 + z_{2,2}/\sqrt{L}) \approx 1 \) [see (4.89) in Appendix 4.7.1], the denominator reduces to
\[
\tilde{u}_k^H T_k \tilde{u}_k \approx \frac{1}{L} z_{1,2}^H \Lambda_1 z_{1,2} + \theta_k.
\] (4.76)

Finally, the output SINR after applying (4.74) and (4.76) is given by
\[
\rho_k \approx \frac{L |\alpha|^2 u_k^H s_1 s_1^H u_k}{z_{1,2}^H \Lambda_1 z_{1,2} + L \theta_k}.
\] (4.77)

Define
\[
\mu \triangleq z_{1,2}^H \Lambda_1 z_{1,2} = \sum_{m=1}^{k-1} \theta_m z_{m,k}^* z_{m,k}.
\] (4.78)

From Lemma 2, \( \{ z_{m,k} \} \) are independent zero-mean Gaussian random variables. Therefore, \( \mu \) is a weighted complex central Chi-square random variable. The output SINR \( \rho_k \) is a one-to-one function of \( \mu \) with a distribution summarized in the following result.
Theorem 1  Under the condition that the covariance matrix $\mathbf{R}$ has a rank-$r$ component as specified in (4.48), the asymptotic PDF of the output SINR $\rho_k$ of the CG-AMF detector for $1 < k \leq r + 1$ is given by

$$f(\rho_k) = \frac{\kappa \left( \frac{\kappa}{\rho_k \theta_k} - L \right)^{k-2} \exp \left( L - \frac{\kappa}{\rho_k \theta_k} \right)}{\rho_k^2 \theta_k \Gamma(k-1)}, \quad 0 < \rho_k \leq \frac{\kappa}{\theta_k L}$$ (4.79)

where $\kappa = L |\alpha|^2 |u_{1,k}|^2 \|s\|^2$ with $u_{1,k}$ denoting the first element of the eigenvector $\mathbf{u}_k$.

Proof:  See Appendix 4.7.3

Following Theorem 1, it is shown in Appendix 4.7.4 that the statistical mean of the output SINR $\rho_k$ is given by:

$$E[\rho_k] = \begin{cases} \frac{\kappa e^L E_1(L)}{\theta_k}, & k = 2 \\ \sum_{j_1=0}^{k_1-1} \sum_{j_2=0}^{k_1-j_1-1} \frac{\kappa (-1)^{i_1} L^{j_1+j_2}}{\theta_k (k_1-j_1)! j_1! j_2!} + \frac{\kappa e^L (-L)^{k_1} E_1(L)}{\theta_k k_1!}, & k > 2. \end{cases}$$ (4.80)

where $k_1 = k - 2$ and the exponential integral $E_1(L)$ is defined as

$$E_1(L) \triangleq \int_L^\infty v^{-1} e^{-v} dv.$$ (4.81)

Some comments on the calculation of $E_1(L)$ are in order. The series representation of the exponential integral is given by [50]:

$$E_1(L) = -a - \ln L + \sum_{i_1=0}^{\infty} \frac{(-1)^{i_1+1} L^{i_1}}{i_1! i_1!}$$ (4.82)

where $a$ is the Euler constant. A truncated version of this formula can be used to approxi-
mate $E_1(L)$ with good accuracy for small $L$. For large $L$, a better approximation is [41]:

$$E_1(L) \approx \frac{e^{-L}}{L} \sum_{i_2=0}^{I_2-1} \frac{i_2!}{(-L)^{i_2}}$$

where $I_2$ is an integer that is used to control the precision of $E_1(L)$. It has an error of order $O(I_2!L^{-I_2})$ and is only valid for large value of $L$. Our simulation results in Section 4.5 use $I_2 = 20$ for training size $L > 64$, which yields good accuracy.

4.5 Numerical Results

Computer simulation is employed to verify the analytical results presented in the previous section. In addition, we compare the proposed CG-AMF detector and several conventional reduced-rank STAP detectors including the eigencanceler (EIG) [26], CSM [21] and DCT based detectors. The simulation uses a normalized spatial frequency $f_s = 0.3$ and a normalized Doppler frequency $f_d = 0.3$ for the target. The performance is assessed at different interference/noise level specified by the reference SINR defined by

$$\text{SINR} = |\alpha|^2 s^H R^{-1} s$$

which coincides with the output SINR obtained by the clairvoyant matched filter (MF), which requires knowledge of $R$. For testing we use the KASSPER data set [3], which is a simulated data set that includes practical airborne radar parameters and issues found in a real-world clutter environment. The radar platform considered in this data set has 11 horizontal antenna elements. For simplicity, we use only the outputs of the first $J = 4$ channels for processing. The number of pulses is $N = 16$.

For the CG-AMF detector, we verify our analysis in Section 4.4 with computer simulation. Note that our analysis is based on the assumption that the covariance matrix $R$. 
Figure 4.1: Output SINR of CG-AMF with rank $k=11$ versus the training size $L$.

has a rank-$r$ component as in (4.48). In many real-world scenarios including the KASSPER data, although $R$ usually does not have the exact low-rank structure, (4.48) is a useful approximation of $R$, in which the rank-$r$ component $R_i$ contains the effect of dominant interference or clutter sources that have to be effectively mitigated for detection. We would like assess the performance of the CG-AMF detector and verify our analysis in such cases.

We first examine the accuracy of our analysis. Figure 4.1 shows the mean of the output SINR versus the training size for the CG-AMF detector with rank $k = 11$, and reference SINR = 20 dB. We compare the asymptotic mean output SINR (4.80) with the output SINR computed by Monte Carlo simulation. When $L > 64$, we use (4.80) along with (4.83) to calculate the mean output SINR, with $I_2$ being set as 20. The \texttt{vpa} function in MATLAB is employed to ensure the calculation accuracy in this case. It can be seen that as the training size increased, the theoretical results converge to the numerical results rapidly. For instance, when the training size is 128, the gap between asymptotic analysis and simulation is less than 0.15 dB.

The size of the training data used in Figures 4.2 to 4.4 is $L = JN = 64$. Figure
Figure 4.2: Output SINR versus detection rank $k$ with $L = 64$ training signals.

4.2 depicts the output SINR of the various reduced-rank detectors versus the rank used in detection when the reference SINR = 20 dB. The theoretical results are computed based on (4.80). The CG iteration yields an estimate of the rank of $R_i$ component in (4.48) to be $r = 11$. Noted that our analysis is valid up to $k = r + 1$, which is why the analytical result is shown only up to $k = 12$. Meanwhile, It is seen that our analysis is able to predict the optimal rank of the CG-AMF detector which yields the highest output SINR. The gap between the analysis and the simulation is due to the relative small $L$ which does not meet the asymptotic condition assumed in our analysis. Meanwhile, it should be noted that the DCT based reduced-rank detector exhibits significant performance degradation compared with the other three data-dependent reduced-rank detectors at low rank. As for the data-dependent methods, they all have a similar maximum SINR, and the CG-AMF detector reaches its maximum output SINR at $k = 11$. Since both the EIG and CSM require a full eigen-decomposition which has a complexity of $O(J^3N^3)$, while the complexity of each CG iteration is just $O(J^2N^2)$, the CG-AMF scheme is computationally more efficient than the EIG and CSM.
Figure 4.3: Probability of detection for the clairvoyant MF, fully adaptive AMF, and reduced-rank CG-AMF detectors with $L = 64$ and $P_f = 10^{-2}$.

Figure 4.4: Probability of detection for the reduced-rank EIG, CSM and CG-AMF detectors with $L = 64$ and $P_f = 10^{-2}$. 
The probability of detection of the MF, AMF and CG-AMF detectors with 7, 11, 24 and 64 CG iterations, respectively, are shown in Figure 4.3 as a function of the reference SINR. The probability of false alarm is set as $P_f = 0.01$. It can be seen that the optimal CG-AMF detector is obtained with 11 CG iterations as predicted in Figure 4.2. Additional iterations should not be pursued since it leads to deteriorated detection performance and higher complexity. Normally, with full iterations $k = 64$, the CG-AMF should converge to the AMF detector; there is a gap between the two, since with $L = JN = 64$ as considered here, the sample covariance matrix $\hat{R}$ is poorly conditioned, which causes some numerical errors.

Figure 4.4 presents the performance comparisons among the CG, EIG, and CSM based reduced-rank detectors with different ranks. It shows that by setting the optimal rank for each detector, namely, $k = 11$ for CG-AMF, $k = 13$ for EIG, and $k = 12$ for CSM, respectively, these detectors yield similar detection performance. When we reduce the rank for EIG and CSM to $k = 11$, their performance degrades significantly.

In Figure 4.5, the training size is increased to $L = 2NJ = 128$, which leads to improvements on the output SINR of all four reduced-rank methods, and the difference between the analysis and the simulation result of the CG-AMF detector decreases. Interestingly, the optimal rank of the computationally efficient CG-AMF detector which yields the maximum output SINR is smaller than that of the CSM and EIG based detectors. We also increase the size of training data to $L = 2JN$ for Figure 4.6. It shows that the CG-AMF detector with the optimal rank $k = 12$ offers the best detection performance. Note that the performance of the CG-AMF detector with full rank $k = 64$ is identical that of the AMF detector.
Figure 4.5: Output SINR versus detection rank $k$ with $L = 128$ training signals.

Figure 4.6: Probability of detection for the reduced-rank EIG, CSM and CG-AMF detectors with $L = 128$ and $P_f = 10^{-2}$. 
4.6 Conclusions

We considered an adaptive reduced-rank approach by employing the iterative conjugate gradient (CG) algorithm and the adaptive matched filter (AMF). The resulting reduced-rank CG-AMF detector is a projection of the full-rank AMF to a Krylov subspace with a dimension determined by the number of CG iterations. Asymptotic analysis of the output SINR of the CG-AMF detector was carried out by exploiting a relation between the CG algorithm and the Lanczos method. Our results show that, in the training limited cases, not only is the CG-AMF detector computationally more efficient (thanks to the efficiency of the CG), but it also often enjoys the benefit of reaching the peak output SINR with a lower rank, when compared with several other popular reduced-rank solutions.

4.7 Appendix

4.7.1 Proof of Lemma 1

Partition $Y$ similarly as $Z$ in (4.67):

$$Y = \begin{bmatrix} Y_{1,1} & Y_{1,2} \\ Y_{2,1} & Y_{2,2} \end{bmatrix}$$

(4.85)

It follows from (4.62) that

$$Y = \begin{bmatrix} I_{k-1} + \frac{1}{\sqrt{L}}Z_{1,1} & \frac{1}{\sqrt{L}}Z_{1,2} \\ \frac{1}{\sqrt{L}}Z_{2,1} & 1 + \frac{z_{2,2}}{\sqrt{L}} \end{bmatrix}.$$  

(4.86)

Since $Y$ is unitary, i.e., $Y^H Y = I_k$, we have
\[ 1 = y_{2,1}y_{2,1}^H + y_{2,2}y_{2,2}^H = \frac{1}{L} z_{2,1}z_{1,2}^H + \left( 1 + \frac{z_{2,2}}{\sqrt{L}} \right)^* \left( 1 + \frac{z_{2,2}}{\sqrt{L}} \right) \]  

and

\[ 0 = Y_{1,1}y_{2,1}^H + y_{1,2}y_{2,2}^H = \frac{1}{\sqrt{L}} (z_{1,2} + z_{2,1}^H) + \frac{1}{L} (Z_{1,1}z_{2,1}^H + z_{1,2}z_{2,2}^H). \]  

By neglecting the terms in the order of \( O(1/L) \), we have

\[ \left( 1 + \frac{z_{2,2}}{\sqrt{L}} \right)^* \left( 1 + \frac{z_{2,2}}{\sqrt{L}} \right) \approx 1 \]  

and

\[ \frac{1}{\sqrt{L}} (z_{1,2} + z_{2,1}^H) \approx 0. \]  

Then, (4.69) follows immediately from (4.90).

4.7.2 Proof of Lemma 2

Let \( h_{1,2} = \begin{bmatrix} h_{1,k} & h_{2,k} & \cdots & h_{k-1,k} \end{bmatrix}^T \). From (4.70), we have

\[ z_{m,k} = \frac{h_{m,k}}{\theta_k - \theta_m} \quad m = 1, 2, \cdots, k - 1. \]  

Thus, we have

\[ E[z_{m,k}z_{n,k}^*] = E[h_{m,k}h_{n,k}^*] \frac{1}{(\theta_k - \theta_m)^2}. \]
Using (4.55), (4.57), (4.59) and (4.60), we can rewrite the eigenvalue perturbation matrix as

\[ H = \sqrt{L}(\bar{\Lambda} - \Lambda) = \sqrt{L}(U^H \tilde{T}_k U - U^H T_k U) \]

\[ = (UQ_k)^H BQ_k U \approx C^H BC \] (4.93)

where \( B = \sqrt{L} (\hat{R} - R) \). It was shown in [25] and [27] that the asymptotic distribution of \( B \) is Gaussian with zero mean and covariance:

\[ E[b_{i,j}^* b_{m,n}^*] = r_{i,m} r_{j,n}^* \] (4.94)

where \( b_{i,j} \) and \( r_{i,m} \) are the elements of \( B \) and \( R \), respectively.

Since \( H \) is a linear transform of \( B \), it follows that \( H \) is also asymptotically Gaussian with zero mean. Furthermore, it is clear from (4.91) that \( z_{m,k} \) is a Gaussian random variable with zero mean. Let \( b \triangleq \text{vec}(B) \in \mathbb{C}^{M^2 \times 1} \), where \( \text{vec}(.) \) denotes the operation of stacking the columns of a matrix on top of one another, an equivalent expression of (4.94) is given as

\[ E[bb^H] = R^* \otimes R. \] (4.95)

Applying the matrix Kronecker product result [24]

\[ \text{vec}(ABC) = (C^T \otimes A) \text{vec}(B) \] (4.96)

we have

\[ h \triangleq \text{vec}(H) = (C^T \otimes C^H) b. \] (4.97)
Therefore,

\[
E[hh^H] = (C^T \otimes C^H) \ E[bb^H] \ (C^T \otimes C^H)^H \\
= (C^T \otimes C^H) \ (R^* \otimes R) \ (C^T \otimes C^H)^H \\
= (C^H RC)^* \otimes (C^H RC) = \Lambda^* \otimes \Lambda
\]

(4.98)

which indicates that \( E[h_{i,j}h_{m,n}^*] = \beta_{i,m}\beta_{j,n}^* \), where \( \beta_{i,m} \) are the elements of \( \Lambda \). Moreover, as \( \Lambda \) is a diagonal matrix containing the eigenvalues of \( T_k \), we have \( \beta_{i,m} = \theta_i \delta_{i,m} \) and \( \beta_{j,n} = \theta_j \delta_{j,n} \), where \( \delta_{i,n} \) denotes the Kronecker delta. Therefore,

\[
E[h_{m,k}h_{n,k}^*] = \beta_{m,n}\beta_{k,k}^* = \theta_k \theta_m \delta_{m,n}.
\]

(4.99)

Substituting (4.99) in (4.92) leads to (4.71). This completes the proof.

4.7.3 Proof of Theorem 1

Since \( \theta_k << \theta_m \), (4.71) can be approximated as

\[
E[z_{m,k}z_{m,k}^*] \approx \frac{\theta_k}{\theta_m}.
\]

(4.100)

Let \( v_m = \sqrt{\frac{\theta_m}{\theta_k}} z_{m,k}, \ m = 1, \cdots, k - 1 \). These new variables are independent, Gaussian random variables with zero mean and unit variance. Then we can rewrite \( \mu \) as defined in (4.78) in terms of \( v_m \) as

\[
\mu = \theta_k \sum_{m=1}^{k-1} v_m v_m^*.
\]

(4.101)
Thus, $\zeta = \mu/\theta_k$ is a complex central Chi-squared random variable with $k - 1$ degrees of freedom [32], i.e., $\zeta \sim \chi^2_{k-1}$, and

$$f_{\zeta}(\zeta) = \frac{\zeta^{k-2}e^{-\zeta}}{\Gamma(k-1)}, \quad \zeta \geq 0. \quad (4.102)$$

The output SINR in (4.77) can be expressed in terms of $\zeta$ by

$$\rho_k = g(\zeta) = \frac{\kappa}{(\zeta + L)\theta_k}, \quad \zeta \geq 0 \quad (4.103)$$

where $\kappa = L |\alpha|^2 u_k^H s_1 s_1^H u_k$. Recall that $s_1 = Q_k^H s$ and the first column of $Q_k$ is a normalized $s$ with unit norm (i.e., normalized $\hat{\gamma}_1$ in Table 1), $s_1$ is a $k \times 1$ vector with the 1st element given by $\|s\|^2$, while all other k-1 elements are zeros. Therefore, $\kappa$ can be further simplified as $\kappa = L |\alpha|^2 |u_{1,k}|^2 \|s\|^2$.

Since the function $g(\zeta)$ is monotonic with the inverse function:

$$g^{-1}(\rho_k) = \zeta = \frac{\kappa}{\rho_k \theta_k} - L, \quad 0 < \rho_k \leq \frac{\kappa}{\theta_k L}. \quad (4.104)$$

The PDF of $\rho_k$ can be calculated in terms of the following equation

$$f_{\rho_k}(\rho_k) = \left| \frac{dg^{-1}(\rho_k)}{d\rho_k} \right| f_{\zeta}(g^{-1}(\rho_k)), \quad 0 < \rho_k \leq \frac{\kappa}{\theta_k L}. \quad (4.105)$$

It follows that the PDF of the SINR is given by

$$f(\rho_k) = \frac{\kappa \left(\frac{\kappa}{\rho_k \theta_k} - L\right)^{k-2} \exp \left(L - \frac{\kappa}{\rho_k \theta_k}\right)}{\rho_k^2 \theta_k \Gamma(k-1)}, \quad 0 < \rho_k \leq \frac{\kappa}{\theta_k L}. \quad (4.106)$$
4.7.4 Mean of the output SINR

The mean of the output SINR can be computed by using (4.102) and (4.103):

\[
E[\rho_k] = \int_0^\infty g(\zeta) f_\zeta(\zeta) d\zeta = \frac{\kappa e^L}{\theta_k k_1!} \int_L^\infty \frac{(v-L)^{k_1} e^{-v}}{v} dv 
\]

(4.107)

where \( k_1 = k - 2 \). For \( k = 2 \), (4.107) reduces to

\[
E[\rho_2] = \frac{\kappa e^L}{\theta_2} E_1(L) 
\]

(4.108)

where \( E_1(L) \) is the exponential integral defined in (4.81). For \( k > 2 \), (4.107) can be computed as follows:

\[
E[\rho_k] = \frac{\kappa e^L}{\theta_k k_1!} \int_L^\infty \sum_{j_1=0}^{k_1} \binom{k_1}{j_1} v^{k_1-1-j_1} (-L)^{j_1} e^{-v} dv 
\]

\[
= \frac{\kappa e^L}{\theta_k k_1!} \sum_{j_1=0}^{k_1-1} (-L)^{j_1} \binom{k_1}{j_1} \Gamma(k_1 - j_1, L) 
\]

\[
+ \frac{\kappa e^L}{\theta_k k_1!} (-L)^{k_1} E_1(L) 
\]

(4.109)

where \( \Gamma(k_1 - j_1, L) \) denotes the incomplete gamma function and is defined as \( \Gamma(k_1 - j_1, L) \triangleq \int_L^\infty v^{k_1-j_1-1} e^{-v} dv \). Alternatively, \( \Gamma(k_1 - j_1, L) \) can be expressed as

\[
\Gamma(k_1 - j_1, L) = (k_1 - j_1 - 1)! e^{-L} \sum_{j_2=0}^{k_1-j_1-1} \frac{L^{j_2}}{j_2!}. 
\]

(4.110)

Substituting (4.110) into (4.107) yields the closed-form solution in (4.80).
Chapter 5
Adaptive Transmit and Receive Beamforming for Interference Mitigation

In this section, we investigate jointly adaptive transmit and receive beamforming in the presence of interferences. A closed-form expression of the transmit beamforming correlation matrix is obtained by maximizing a lower bound of the output SINR at the receiver. Our solutions of the transmit beamforming correlation matrix and the associated receive beamformer require some knowledge (i.e., locations and strengths) of the interferences, which are adaptively estimated from the training data. The advantage of employing adaptive processing at both the transmitter and receiver in the presence of interferences with uncertainties is demonstrated by numerical results.

5.1 Data Model

Consider a monostatic radar system with $M_t$ transmit antennas and $M_r$ receive antennas. Let $s_m(n)$ denote the signal transmitted by the $m$th antenna, and $\theta$ denote the location parameter of a scatterer (target or interference). When the transmitted signals are narrow band, the baseband signal at a specific scatterer location can be described as $\mathbf{a}_t^H(\theta)s(n), n = 1, \cdots, N$, where $\mathbf{s}(n) = [s_1(n) \ s_2(n) \ \cdots \ s_{M_t}(n)]^T$ contains the transmitted signal samples and $\mathbf{a}_t(\theta)$ denotes the $M_t \times 1$ transmit steering vector containing complex-valued elements with unit amplitude and phase determined by the look angle $\theta$. As an example, for a uniform linear array with a half-wavelength separation between two adjacent array elements, the steering vector is given by $\mathbf{a}_t(\theta) = [1 \ e^{j\pi \sin \theta} \ \cdots \ e^{j(M_t-1)\pi \sin (\theta)}]^T$.

Suppose there is a target with complex amplitude $\alpha_0$ located at the look direction $\theta_0$ along with $K$ interferences located at $\mathbf{\theta}_i = [\theta_1 \ \theta_2 \ \cdots \ \theta_K], \ \theta_k \neq \theta_0$ for $k = 1, \cdots, K$. 
Then, the received signal reflected from the target and the interferences are given by [36]

\[
y_t(n) = \alpha_0 a_r^*(\theta_0) a_t^H(\theta_0) s(n) \quad (5.1)
\]

and

\[
y_i(n) = \sum_{k=1}^{K} \alpha_k a_r^*(\theta_k) a_t^H(\theta_k) s(n) \quad \triangleq A_r^*(\theta_i) \text{diag}(\alpha) A_t^H(\theta_i) s(n) \quad (5.2)
\]

\[= H(\theta_i) s(n) \]

respectively, where \(\alpha_k\) denotes the complex amplitude of the \(k\)th source, the vector \(a_r(\theta) \in \mathbb{C}^{M_r \times 1}\), which is similarly defined as \(a_t(\theta)\), is usually referred to as the receive steering vector. Moreover, the matrices \(A_r(\theta_i), A_t(\theta_i)\) and the amplitude vector \(\alpha\) are given as

\[
A_r(\theta_i) = \begin{bmatrix} a_r(\theta_1) & a_r(\theta_2) & \cdots & a_r(\theta_K) \end{bmatrix} ; \quad (5.3)
\]

\[
A_t(\theta_i) = \begin{bmatrix} a_t(\theta_1) & a_t(\theta_2) & \cdots & a_t(\theta_K) \end{bmatrix} \quad (5.4)
\]

and

\[
\alpha = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_K \end{bmatrix}^T , \quad (5.5)
\]

respectively.

Therefore, the received radar signal in the presence of both target and interferences is given by:

\[
y(n) = \alpha_0 a_r^*(\theta_0) a_t^H(\theta_0) s(n) + H(\theta_i) s(n) + \epsilon(n) \quad (5.6)
\]

where \(\epsilon(n)\) is a noise vector with zero mean and covariance matrix \(\sigma_n^2 I_{M_r}\). Define the
disturbance (interferences and noise) as

\[ d(n) \triangleq H(\theta_i)s(n) + \epsilon(n). \] (5.7)

Assume the complex amplitudes \( \alpha_k \) of the interference sources are uncorrelated with zero mean and variance \( \sigma_k^2 \). Then the covariance matrix of the disturbance is given by:

\[
R_d = E\{d(n)d^H(n)\} = \sum_{k=1}^{K} \sigma_k^2 a_t^*(\theta_k)a_t^H(\theta_k)R_s a_t(\theta_k)a_t^T(\theta_k) + \sigma_n^2 I_M \\
\triangleq \sum_{k=1}^{K} \sigma_k^2 g(\theta_k)a_t^*(\theta_k)a_t^T(\theta_k) + \sigma_n^2 I_M, \] (5.8)

where \( R_s \) is the correlation matrix of \( s(n) \), i.e., \( R_s = E\{s(n)s^H(n)\} \), and \( g(\theta_k) \triangleq a_t^H(\theta_k)R_s a_t(\theta_k) \) is the transmit beamforming gain at direction \( \theta_k \). At the receiver side, a linear beamformer \( w \in \mathbb{C}^{M_r \times 1} \) is applied to \( y(n) \) for interference mitigation, yielding the output \( z(n) = w^H y(n) \).

The problem of interest is to jointly optimize \( R_s \) for transmit beamforming and \( w \) for receive beamforming.
5.2 Proposed Approach

5.2.1 Transmit and Receive Beamforming Design

Assume the complex amplitude \( \alpha_0 \) of the target has zero mean and variance \( \sigma_0^2 \). The output SINR of the beamformer is given by

\[
\rho(w, R_s) \triangleq \frac{E \left\{ \left| \alpha_0 w^H a_t^*(\theta_0) a_t^H(\theta_0)s(n) \right|^2 \right\}}{E \left\{ \left| w^H [H(\theta_i)s(n) + \epsilon(n)] \right|^2 \right\}} = \sigma_0^2 \frac{w^H a_r^*(\theta_0) a_r^H(\theta_0) R_s a_t(\theta_0) a_t^T(\theta_0) w}{w^H R_d w}. \tag{5.9}
\]

Our approach is to maximize the SINR jointly with respect to (w.r.t.) the receive beamformer \( w \) and the transmit beamforming correlation matrix \( R_s \), subject to constraints on the transmit power and positive semi-definitiveness of \( R_s \), i.e.,

\[
\max_{w, R_s} \rho(w, R_s) \quad \text{s.t.} \quad \text{tr} \{ R_s \} \leq p, \quad R_s \succeq 0. \tag{5.10}
\]

To solve (5.10), we can first solve \( w \) in terms of \( R_s \). For a given \( R_s \), \( w \) can be obtained by:

\[
\max_w \frac{w^H a_r^*(\theta_0) a_r^H(\theta_0) R_s a_t(\theta_0) a_t^T(\theta_0) w}{w^H R_d w}. \tag{5.11}
\]

The optimal \( w \) and the associated maximum value \( \lambda_1 \) of (5.11) are given as (see Appendix 5.5.1 for a proof)

\[
w = \tau R_{d}^{-1} a_r^*(\theta_0) \tag{5.12}
\]

where \( \tau \neq 0 \) is an arbitrary constant, and

\[
\lambda_1 = a_t^H(\theta_0) R_s a_t(\theta_0) a_r^T(\theta_0) R_{d}^{-1} a_r^*(\theta_0). \tag{5.13}
\]
With (5.12), we have the output SINR $\rho(R_s) = \sigma_0^2 \lambda_1(R_s)$. The remaining step is to search for the $R_s$ which results in the maximum output SINR. Namely, we have the following optimization problem:

$$\max_{R_s} \lambda_1(R_s) \quad \text{s.t. } \text{tr} \{R_s\} \leq p, \ R_s \succeq 0.$$  \hfill (5.14)

A main difficulty associated with (5.14) is that the underlying objective function involves a non-diagonal matrix inverse. To circumvent this difficulty, we propose to maximize a lower bound of the objective function in (5.14). Specifically, by using the Cauchy-Schwarz inequality, we obtain the following lower bound (see Appendix 5.5.2 for a proof)

$$\lambda_1 \geq \frac{M_r^2 \text{tr} \{R_s B_t(\theta_0)\}}{\text{tr} \{R_s C\} + M_r \sigma_n^2}$$ \hfill (5.15)

where

$$B_t(\theta_0) \triangleq a_t(\theta_0)a_t^H(\theta_0)$$ \hfill (5.16)

and

$$C \triangleq \sum_{k=1}^{K} \sigma_k^2 a_t(\theta_k)a_t^T(\theta_k)B_r(\theta_0)a_r^*(\theta_k)a_r^H(\theta_k)$$ \hfill (5.17)

with

$$B_r(\theta_0) \triangleq a_r^*(\theta_0)a_r^T(\theta_0).$$ \hfill (5.18)

Thus, we have the new optimization problem w.r.t. $R_s$ as:

$$\max_{R_s} \frac{\text{tr} \{R_s B_t(\theta_0)\}}{\text{tr} \{R_s C\} + M_r \sigma_n^2} \quad \text{s.t. } \text{tr} \{R_s\} \leq p, \ R_s \succeq 0.$$ \hfill (5.19)

This is a constrained fractional semidefinite programming (SDP) problem whose solution can be obtained by solving its equivalent SDP via the so-called Charnes-Cooper transform.
mation [9]. Specifically, since the denominator of the fractional SDP is strictly positive [see (5.51)], we can define $Q = q R_s$, where $q > 0$ is a scaling parameter which makes $\text{tr} \{ q R_s C \} + q M_r \sigma_n^2 = 1$. Hence, multiplying by $q$ the numerator and the denominator of the objective function in (5.19), we obtain the equivalent SDP problem as

$$\max_{Q,q} \qquad \text{tr} \{ Q B_t(\theta_0) \}$$

$$\text{s.t.} \qquad \text{tr} \{ QC \} = 1 - q M_r \sigma_n^2$$

$$\qquad \qquad \qquad \text{tr} \{ Q \} = qp$$

$$\qquad \qquad \qquad Q \succeq 0, \quad q > 0. \quad (5.20)$$

The optimal solution $(Q_*, q_*)$ of (5.20) can be found by using standard convex optimization software. In turn, the solution of (5.19) can be obtained as $R_{s*} = \frac{1}{q_*} Q_*$. In fact, a closed-form solution to (5.19) can be derived as shown next. Let the eigenvalue decomposition of $R_s$ be $R_s = U \Gamma U^H$, where $\Gamma$ contains the eigenvalues $\{ \gamma_m \}_{m=1}^{M_t}$ on its diagonal. The optimization problem can be rewritten as

$$\max_{\Gamma, U} \qquad \frac{\text{tr} \{ U^H B_t(\theta_0) U \}}{\text{tr} \{ U^H C U \} + M_r \sigma_n^2}$$

$$\text{s.t.} \qquad \sum_{m=1}^{M_t} \gamma_m \leq p, \quad \gamma_m \geq 0. \quad (5.21)$$

Let $\eta_*$ denote the maximum value of the above objective function within the feasible set. Then, for any eigen-pair $\Gamma$ and $U$

$$\frac{\text{tr} \{ U^H B_t(\theta_0) U \}}{\text{tr} \{ U^H C U \} + M_r \sigma_n^2} \leq \eta_*.$$

$$\quad (5.22)$$

The problem is to construct $U_*$ and $\Gamma_*$ which satisfy the constraints in (5.21) and the following relation:

$$\text{tr} \{ \Gamma_* U_*^H (B_t(\theta_0) - \eta_* C) U_* \} = M_r \sigma_n^2 \eta_*.$$

$$\quad (5.23)$$
Let
\[
C_0 \triangleq C + \frac{M_r \sigma_n^2}{p} I_{M_t}.
\] (5.24)

It is shown in Appendix 5.5.3 that
\[
\eta_\star = a_t^H(\theta_0) C_0^{-1} a_t(\theta_0)
\] (5.25)
and
\[
R_{ss\star} = pvv^H
\] (5.26)
with the vector \(v\) given by:
\[
v = \frac{C_0^{-1} a_t(\theta_0)}{[a_t^H(\theta_0) C_0^{-2} a_t(\theta_0)]^{1/2}}.
\] (5.27)

The optimal transmit beamforming matrix \(R_{ss\star}\) is a rank-one matrix, which we found is consistent with the numerical result obtained by the aforementioned SDP approach.

### 5.2.2 Adaptive Estimation

It appears that the proposed method requires knowledge of the target direction \(\theta_0\). It does not since like all other beamformers, our method has to scan all possible directions of the target. That is, for each possible target direction, a beamformer is computed and used to filter the observed signal, and then an estimate of the target direction can be obtained from the filter outputs. On the other hand, our method does require some knowledge of the interferences, i.e., \(\{\theta_k\}_{k=1}^K\) and \(\{\sigma_k^2\}_{k=1}^K\). We discuss how to adaptively estimated these parameters from training signals.

Specifically, training signals are obtained by sending a selected waveform \(s_0(n), n = \ldots \)
1, 2, \ldots, N, to probe the environment when the target is absent (prior to target sensing). Let

\[ Y_0 = \begin{bmatrix} y_0(1) & y_0(2) & \cdots & y_0(N) \end{bmatrix} \in \mathbb{C}^{M_r \times N} \tag{5.28} \]

contains the corresponding received signal which can be expressed as

\[ Y_0 = A_r^*(\theta_i) \text{diag}(\alpha) A_t^H(\theta_i) S_0 + E_0 \tag{5.29} \]

where

\[ S_0 = \begin{bmatrix} s_0(1) & s_0(2) & \cdots & s_0(N) \end{bmatrix} \in \mathbb{C}^{M_t \times N} \tag{5.30} \]

and

\[ E_0 = \begin{bmatrix} \epsilon(1) & \epsilon(2) & \cdots & \epsilon(N) \end{bmatrix} \in \mathbb{C}^{M_r \times N} \tag{5.31} \]

contains the noise vectors. For simplicity, we use orthogonal waveforms, e.g., by choosing \( S_0 \) to be a submatrix of the discrete Fourier transform matrix, and omnidirectional transmission to learn an environment without any prior knowledge.

Given the training data \( Y_0 \), there is a multiple of methods that can be employed for interference parameter estimation. Here, we use the multiple signal classification (MUSIC) algorithm [49] to obtain the estimates \( \{\hat{\theta}_k\}_{k=1}^K \) of the interference locations. Moreover, from (5.29) we have

\[ Y_0 S_0^\dagger = A_r^*(\theta_i) \text{diag}(\alpha) A_t^H(\theta_i) + E_0 S_0^\dagger \]

\[ = \sum_{k=1}^K \alpha_k a_r^*(\theta_k) a_t^H(\theta_k) + E_0 S_0^\dagger \tag{5.32} \]
where $S_0^\dagger$ denotes the pseudo inverse of the matrix $S_0$ i.e.,

$$S_0^\dagger = S_0^H(S_0 S_0^H)^{-1}. \quad (5.33)$$

Let $\text{vec}(.)$ denotes the operation of stacking the columns of a matrix on top of one another, an equivalent expression of (5.32) is given as

$$\text{vec} \left( Y_0 S_0^\dagger \right) = \sum_{k=1}^{K} \alpha_k \left[ a_t^*(\theta_k) \otimes a_r^*(\theta_k) \right] + \text{vec} \left( E_0 S_0^\dagger \right) = [A_t^*(\hat{\theta}_i) \circ A_r^*(\hat{\theta}_i)] \alpha + \text{vec} \left( E_0 S_0^\dagger \right) \quad (5.34)$$

where $\otimes$ denotes the matrix Kronecker product while $\circ$ denotes the column-wise or Khatri-Rao product, i.e.,

$$A_t^*(\theta_i) \circ A_r^*(\theta_i) = [a_t^*(\theta_1) \otimes a_r^*(\theta_1), \ldots, a_t^*(\theta_K) \otimes a_r^*(\theta_K)]. \quad (5.35)$$

By defining

$$A_{t,r} \triangleq A_t^*(\hat{\theta}_i) \circ A_r^*(\hat{\theta}_i), \quad (5.36)$$

it follows an estimate of $\alpha$ is given by

$$\hat{\alpha} = \left( A_{t,r}^H A_{t,r} \right)^{-1} A_{t,r}^H \text{vec} \left( Y_0 S_0^\dagger \right). \quad (5.37)$$

Then, the variance of the amplitude of the interference can be simply estimated as $\hat{\sigma}_k^2 = \hat{\alpha}_k^2$.

Finally, we use the estimates $\hat{\theta}_k$ and $\hat{\sigma}_k^2$ to replace $\theta_k$ and $\sigma_k^2$ used in (5.27). The resulting correlation matrix of the probing waveform is then given as $\hat{R}_{ss} = p \hat{\nu} \hat{\nu}^H$, and its associated beamformer weight vector is
\[
\tilde{w} = \tau \hat{R}_d^{-1} a_r^*(\theta_0)
\] (5.38)

where \( \hat{R}_d \) is obtained in terms of (5.8) by replacing \( \theta_k, \sigma_k^2 \) and \( \mathbf{R}_s \) with their estimates \( \hat{\theta}_k, \hat{\sigma}_k^2 \) and \( \hat{\mathbf{R}}_{s*} \), respectively. Substituting (5.38) into (5.9), we have the output SINR as

\[
\hat{\rho}_* = \frac{\sigma_0^2 \tilde{w}^H a_r^*(\theta_0) a_t^H(\theta_0) \mathbf{R}_s a_t(\theta_0) a_r^T(\theta_0) \tilde{w}}{\tilde{w}^H \hat{R}_d \tilde{w}}.
\] (5.39)

In Section 5.3, beampattern is used to compare different adaptive beamforming schemes. We consider the joint transmit-receive beampattern given as

\[
\hat{P}(\theta) = \left[ a_t^H(\theta) \hat{R}_{s*,a}(\theta) \right] \left| \tilde{w}^H a_r^*(\theta) \right|^2
\] (5.40)

which includes the contribution from the transmit beamforming \( a_t^H(\theta) \hat{R}_{s*,a}(\theta) \) and the receive beamforming \( \left| \tilde{w}^H a_r^*(\theta) \right|^2 \), respectively. Additionally, the non-zero scalar \( \tau \) in (5.38) is selected as

\[
\tau \equiv \left[ a_r^T(\theta_0) \hat{R}_d^{-1} a_r^*(\theta_0) a_t^H(\theta_0) \hat{R}_{s*,a}(\theta_0) a_r^T(\theta_0) \hat{R}_d^{-1} a_r^*(\theta_0) \right]^{-\frac{1}{2}}
\] (5.41)

to normalize the beampattern (5.40) such that the beamforming gain at the target direction is one. Note the normalization does not change the shape of the beampattern.

### 5.3 Numerical Results

We present numerical results to demonstrate the merits of the proposed beamforming scheme. We compare it with a phased-array (PA) scheme that points at the target location \( \theta_0 \) at transmission. The transmit correlation matrix of the PA is \( \mathbf{R}_{s,PA} = \frac{p}{||a_t(\theta_0)||^2} a_t(\theta_0) a_t^H(\theta_0) \) [20], whereas the receive beamforming vector is similarly given by (5.12), except that its
disturbance covariance matrix \( \mathbf{R}_d \) depends on \( \mathbf{R}_{s,PA} \). Moreover, for adaptive interference cancellation, the PA scheme also requires knowledge of the interferences in order to compute \( \mathbf{R}_d \). We assume that \( \mathbf{R}_d \) used in the PA system is estimated in a similar approach as described in Section 5.2.2.

Consider an array radar system where the transmitter and receiver share a uniform linear array (ULA) of \( M_t = M_r = M = 6 \) elements with half-wavelength inter-element separation. The total transmit power is set to \( p = M \), a target is located at \( 0.1 \) and five interferences are at \((-0.38, -0.1, 0.15, 0.31, 0.46)\) in normalized spatial frequency \( f \triangleq (\sin \theta)/2 \). The overall power for the interferences is 1 and the noise variance is \( \sigma_n^2 = 0.01 \). The target power is either \( \sigma_0^2 = 1 \) or varied over a range of values as specified. We consider a training-limited scenario where the number of training data used for adaptive estimation is \( N = M = 6 \).

Fig.5.1 depicts the mean of the output SINR for the proposed and the PA schemes based on adaptive estimation. The output SINR for the proposed scheme with known interferences is also shown as a benchmark. We note that the proposed scheme has a similar SINR with known or adaptive estimated interferences. Moreover, it outperforms the PA scheme by 4.3dB. The joint beampatterns of the two adaptive approaches are shown in Fig.5.2. It can be seen that the proposed adaptive design is able to suppress all the five interferences, while the PA scheme cannot effectively mitigate the interferences at \((-0.38, 0.46)\). Therefore, our proposed adaptive approach has a stronger ability to handle the interferences in the training-limited situation.

5.4 Conclusions

We have proposed a jointly adaptive transmit and receive beamforming for array radars. The transmit and receive beampattern is obtained by jointly designing the transmit beam-
Figure 5.1: Mean of the output SINR.

Figure 5.2: Beampatterns of the beamformers.
forming correlation matrix and receive beamforming vector in terms of maximizing the output SINR. The adaptive processing is achieved in two stages. In the initial probing stage, we first transmit multiple uncorrelated waveforms with omnidirectional transmission to probe a target-free environment. In the next stage, the training data obtained from initial probing are then used for interference estimation and beamformer design. The numerical results show that by applying adaptive processing for both radiation and receiving in a training-limited situation, we can achieve a better beampattern, a stronger ability to handle interference, and a higher output SINR.

5.5 Appendix

5.5.1 Proof of (5.12) and (5.13)

Define

\[ w_1 \triangleq R_d^{1/2} w. \]  
(5.42)

The problem (5.11) becomes

\[
\max_{w_1} \frac{w_1^H R_d^{-\frac{1}{2}} a_r^*(\theta_0) a_t^H(\theta_0) R_s a_t(\theta_0) a_r^T(\theta_0) R_d^{-\frac{1}{2}} w_1}{\|w_1\|^2}
\]  
(5.43)

or equivalently,

\[
\max_{w_1} \frac{w_1^H R_d^{-\frac{1}{2}} a_r^*(\theta_0) a_t^H(\theta_0) R_s a_t(\theta_0) a_r^T(\theta_0) R_d^{-\frac{1}{2}} w_1}{\|w_1\|^2}
\]  
(5.44)

\[ \text{s.t. } \|w_1\| = 1. \]

The maximum of the objective function is the largest eigenvalue \( \lambda_1 \) of \( R_d^{-\frac{1}{2}} a_r^*(\theta_0) a_t^H(\theta_0) \times R_s a_t(\theta_0) a_r^T(\theta_0) R_d^{-\frac{1}{2}} \), and the solution of \( w_1 \) is the associated principal eigenvector. Since \( a_r^*(\theta_0) a_t^H(\theta_0) \) is a rank one matrix, there is only one non-zero eigenvalue of \( R_d^{-\frac{1}{2}} a_r^*(\theta_0) a_t^H(\theta_0) \)
\[
\times R_s a_t(\theta_0) a_t^T(\theta_0) R_d^{-\frac{1}{2}}, \text{ which is}
\]

\[
\lambda_1 = a_t^H(\theta_0) R_s a_t(\theta_0) a_t^T(\theta_0) R_d^{-1} a_t^*(\theta_0) \geq 0. \quad (5.45)
\]

The associated eigenvector \( w_1 \) is

\[
w_1 = \frac{R_d^{-\frac{1}{2}} a_t^*(\theta_0)}{[a_t^T(\theta_0) R_d^{-1} a_t^*(\theta_0)]^{\frac{1}{2}}}. \quad (5.46)
\]

In turn, we can write

\[
w = R_d^{-\frac{1}{2}} w_1 = \tau R_d^{-1} a_t^*(\theta_0) \quad (5.47)
\]

where \( \tau \) can be any non-zero constant since scaling does not change the value of \((5.43)\).

5.5.2 **Proof of** (5.15)

Let \( E = R_d^{\frac{1}{2}} a_t(\theta_0) a_t^T(\theta_0) R_d^{-\frac{1}{2}} \) and \( F = R_d^{\frac{1}{2}} a_t(\theta_0) a_t^T(\theta_0) R_d^{\frac{1}{2}} \). Then we have

\[
\lambda_1 = \text{tr} \left[ a_t^*(\theta_0) a_t^H(\theta_0) R_s a_t(\theta_0) a_t^T(\theta_0) R_d^{-1} \right] = \text{tr} \left( EE^H \right). \quad (5.48)
\]

By the Cauchy-Schwarz inequality:

\[
\text{tr} \left( EE^H \right) \text{tr} \left( FF^H \right) \geq \left[ \text{tr} \left( EF^H \right) \right]^2 \quad (5.49)
\]
a lower bound is given as

$$
\lambda_1 \geq \frac{\left[ \text{tr} \left( \mathbf{E} \mathbf{F}^H \right) \right]^2}{\text{tr} (\mathbf{F} \mathbf{F}^H)} = \frac{\left\{ \text{tr} \left[ \mathbf{R}_s \mathbf{a}_t(\theta_0) \mathbf{a}_t^H(\theta_0) \mathbf{a}_s^*(\theta_0) \mathbf{a}_s^H(\theta_0) \mathbf{R}_s \right] \right\}^2}{\text{tr} \left[ \mathbf{R}_s \mathbf{a}_t(\theta_0) \mathbf{a}_t^H(\theta_0) \mathbf{R}_d \mathbf{a}_s^*(\theta_0) \mathbf{a}_s^H(\theta_0) \mathbf{R}_s \right]} = \frac{M_r^2 \mathbf{a}_l^H(\theta_0) \mathbf{R}_s \mathbf{a}_l(\theta_0)}{\mathbf{a}_l^H(\theta_0) \mathbf{R}_d \mathbf{a}_l^*(\theta_0)}. \quad (5.50)
$$

The lower bound is tight if $\mathbf{E} = \alpha \mathbf{F}$, where $\alpha$ is a non-zero constant. It is easy to see that this condition is met if the interference is (approximately) spectrally white, or $\mathbf{R}_d \propto \mathbf{I}$.

Based on (5.8), the denominator of (5.50) can be written as

$$
\mathbf{a}_l^T(\theta_0) \mathbf{R}_d \mathbf{a}_l^*(\theta_0) = \sum_{k=1}^{K} \sigma_k^2 \mathbf{a}_s^H(\theta_k) \mathbf{B}_s(\theta_0) \mathbf{a}_s^*(\theta_k) + M_r \sigma_n^2
$$

$$
= \text{tr} \left[ \sum_{k=1}^{K} \sigma_k^2 \mathbf{a}_s(\theta_k) \mathbf{a}_s^T(\theta_k) \mathbf{B}_s(\theta_0) \mathbf{a}_s^*(\theta_k) \mathbf{a}_s^H(\theta_k) \mathbf{R}_s \right] + M_r \sigma_n^2
$$

$$
= \text{tr}(\mathbf{R}_s \mathbf{C}) + M_r \sigma_n^2 \quad (5.51)
$$

and the nominator as

$$
M_r^2 \mathbf{a}_l^H(\theta_0) \mathbf{R}_s \mathbf{a}_l(\theta_0) = M_r^2 \text{tr} [\mathbf{R}_s \mathbf{B}_s(\theta_0)]. \quad (5.52)
$$

### 5.5.3 Derivation of the optimal solution of (5.21)

A solution to the problem is obtained by construction. Let the rank of $\mathbf{R}_{ss}$ be $M$ with $M \leq M_t$, and $\mathbf{u}_m$ be the eigenvector of $\mathbf{R}_{ss}$ corresponding to the eigenvalue $\gamma_m$. Then,
(5.23) can be written as:

\[
\sum_{m=1}^{M} \left[ \gamma_m u_m^H (B_t(\theta_0) - \eta_\star C) u_m - b_m M_r \sigma_n^2 \eta_\star \right] = 0
\]  

(5.53)

where \( \{b_m\}_{m=1}^{M} \) are a set of coefficients which satisfy \( \sum_{m=1}^{M} b_m = 1 \). By selecting

\[
b_m \triangleq \gamma_m u_m^H (B_t(\theta_0) - \eta_\star C) u_m / M_r \sigma_n^2 \eta_\star,
\]

(5.54)

we have

\[
\gamma_m u_m^H (B_t(\theta_0) - \eta_\star C) u_m - b_m M_r \sigma_n^2 \eta_\star = 0
\]  

(5.55)

or equivalently,

\[
u_m^H \left[ B_t(\theta_0) - \eta_\star \left( C + \frac{b_m M_r \sigma_n^2}{\gamma_m} I_{M_t} \right) \right] u_m = 0
\]  

(5.56)

for \( m = 1, \ldots, M \). Hence, \( \eta_\star \) is a non-zero generalized eigenvalue of \( \{B_t(\theta_0), C_m\} \) with

\[
C_m \triangleq C + \frac{b_m M_r \sigma_n^2}{\gamma_m} I_{M_t},
\]

(5.57)

and \( u_m \) is the generalized eigenvector corresponding to \( \eta_\star \). That is,

\[
B_t(\theta_0) u_m = \eta_\star C_m u_m.
\]  

(5.58)

Since \( B_t(\theta_0) = a_t(\theta_0) a_t^H(\theta_0) \) is rank-one, it is easy to show

\[
\eta_\star = a_t^H(\theta_0) C_m^{-1} a_t(\theta_0).
\]  

(5.59)
The above solution obtained for different $m$ should be identical, i.e.,
\begin{equation}
\mathbf{a}_t^H(\theta_0)\mathbf{C}_j^{-1}\mathbf{a}_t(\theta_0) - \mathbf{a}_t^H(\theta_0)\mathbf{C}_l^{-1}\mathbf{a}_t(\theta_0) = 0,
\end{equation}
\forall j, l \in \{1, 2, \ldots, M\}.

Denote the eigen-decomposition of $\mathbf{C}$ as $\mathbf{C} = \mathbf{E}\mathbf{D}\mathbf{E}^H$, where the unitary matrix $\mathbf{E}$ contains the eigenvectors while the diagonal matrix $\mathbf{D}$ contains the eigenvalues. Let $\mu_j = \frac{b_j M \sigma^2}{\gamma_j}, j = 1, 2, \ldots, M$. Then, (5.60) can be expressed as
\begin{equation}
\mathbf{a}_t^H(\theta_0)\mathbf{E}\left[(\mathbf{D} + \mu_j \mathbf{I})^{-1} - (\mathbf{D} + \mu_l \mathbf{I})^{-1}\right]\mathbf{E}^H\mathbf{a}_t(\theta_0) = 0
\end{equation}
which implies $\frac{b_1}{\gamma_1} = \frac{b_2}{\gamma_2} = \cdots = \frac{b_M}{\gamma_M}$. In addition, from $\sum_{m=1}^{M} \gamma_m = p$ and $\sum_{m=1}^{M} b_m = 1$, we conclude $\frac{b_m}{\gamma_m} = \frac{1}{p}$. As such,
\begin{equation}
\mathbf{C}_m = \mathbf{C} + \frac{M \sigma^2}{p} \mathbf{I}_M \triangleq \mathbf{C}_0, \ m = 1, \ldots, M.
\end{equation}

It remains to determine the rank $M$ of $\mathbf{R}_{s^*}$. From (5.58) and (5.62), we have
\begin{equation}
\mathbf{C}_0^{-\frac{1}{2}}\mathbf{a}_t(\theta_0)\mathbf{a}_t^H(\theta_0)\mathbf{C}_0^{-\frac{1}{2}}\mathbf{C}_0^{-\frac{1}{2}}\mathbf{u}_m = \eta_s \mathbf{C}_0^\frac{1}{2}\mathbf{u}_m
\end{equation}
which indicates that the normalized form of $\mathbf{C}_0^\frac{1}{2}\mathbf{u}_m$ is the eigenvector of the rank-one matrix $\mathbf{C}_0^{-\frac{1}{2}}\mathbf{a}_t(\theta_0)\mathbf{a}_t^H(\theta_0)\mathbf{C}_0^{-\frac{1}{2}}$. Hence, we have $\mathbf{C}_0^\frac{1}{2}\mathbf{u}_m \propto \mathbf{C}_0^{-\frac{1}{2}}\mathbf{a}_t(\theta_0)$ and more specifically, we can write
\begin{equation}
\mathbf{C}_0^\frac{1}{2}\mathbf{u}_m = \frac{\mathbf{C}_0^{-\frac{1}{2}}\mathbf{a}_t(\theta_0)}{\left[\mathbf{a}_t^H(\theta_0)\mathbf{C}_0^{-\frac{1}{2}}\mathbf{a}_t(\theta_0)\right]^{1/2}}
\end{equation}
and it follows that all $\{\mathbf{u}_m\}$ are identical, given by
\begin{equation}
\mathbf{u}_m = \frac{\mathbf{C}_0^{-1}\mathbf{a}_t(\theta_0)}{\left[\mathbf{a}_t^H(\theta_0)\mathbf{C}_0^{-\frac{1}{2}}\mathbf{a}_t(\theta_0)\right]^{1/2}} \triangleq \mathbf{v}.
\end{equation}
Moreover, since \{u_m\} are by definition the eigenvectors of the semidefinite matrix \( R_{ss} \), they must be different. Therefore, we must have \( M = 1 \) and \( R_{ss} \) is rank-one, given by

\[
R_{ss} = \gamma_1 vv^H = p vv^H.
\]  

(5.66)
Chapter 6
Waveform Synthesis for MIMO Radar

6.1 Data Model

Let the columns of $S \in \mathbb{C}^{N \times N_t}$ be the normalized transmitted waveforms and satisfy $\|S\| = 1$, where $N_t$ is the number of the transmitters and $N$ is the number of samples in each waveform. The class of unconstrained waveform matrices $S$ that realize a given transmitting correlation matrix $\Omega = R^H R \in \mathbb{C}^{N_t \times N_t}$ is given by

$$S = UR$$  \hspace{1cm} (6.1)

where $U \in \mathbb{C}^{N \times N_t}$ with $N \geq N_t$ is an arbitrary semi-unitary matrix and satisfies $U^H U = I$. In addition to realizing or approaching to $\Omega$, the waveform matrix must also satisfy a number of practical constraints. Let $C$ denote the set of waveform matrices $S$ that satisfy these constraints. Then a possible mathematical formulation of the problem of synthesizing the waveform matrix $S$ is as follows:

$$\min_{S \in C; U} \| S - UR \|^2.$$  \hspace{1cm} (6.2)

The solution $S$ to (6.2) may realize $\Omega$ exactly or only approximately depend on the constraint set $C$.

We begin the proposed algorithm with a cyclic method which alternatively updates the $U$ and $S$ for the optimization problem (6.2). An iterative algorithm is then employed to find the optimal $S$ with constant modulus and similarity constraints in each update.
6.2 Cyclic Method

The minimization problem (6.2) is nonconvex due to the nonconvexity of the constraint $U^H U = I$ and possibly of the set $C$, as well. With this fact in mind, we prefer a cyclic method for solving (6.2), as suggested in a related context in [52, 54, 55].

The cyclic method can be summarized as following steps.

1. Set $S$ to an initial value.

2. Release the constraint on $S$, obtain the $U \ (U^H U = I)$ that minimizes (6.2) for $S$ fixed at its most recent value.

3. Release the constraint on $U$, determine the matrix $S \ (S \in C)$ for $U$ fixed at its most recent value.

Iterate Step 2 and Step 3 until a given stop criterion is satisfied.

In Step 2, the following minimization problem need to be solved w.r.t. $U$:

$$\min_U \|S - UR\|^2 \quad (6.3)$$

s.t. $U^H U = I$.

Let

$$RS^H = D\Sigma V^H \quad (6.4)$$

denote singular value decomposition (SVD) of $RS^H$, the optimal closed-form solution is (see, e.g., [28, 52])

$$U = VD^H. \quad (6.5)$$
In Step 3, we face the optimization problem as

\[
\min_S \|S - UR\|^2 \quad (6.6)
\]

s.t. \( S \in C. \)

The optimal solution depends on the constraint set \( C. \)

### 6.3 Waveform Synthesis with Transmit Power Constraint

In this section, we first consider the transmitted power constraint on \( S, \) the optimization problem (6.6) then becomes

\[
\min_S \|S - UR\|^2 \quad (6.7)
\]

s.t. \( \|S\| = 1, \)

which is equivalent to

\[
\min_s \|s - b\|^2 \quad (6.8)
\]

s.t. \( \|s\| = 1 \)

where \( s = \text{vec}(S) \) and \( b = \text{vec}(UR). \) Its closed-form solution is computed as \( s = b/\|b\|. \)

We summarize the optimization algorithm as below,
Waveform Synthesis Algorithm I:

**Input:**
- Waveform correlation matrix $\Omega$
- Initial waveform matrix $S_0$

**Output:**
- A solution $S^*$ of (6.6) with the constraint in (6.7)

1) when $m = 1$, initialize the transmit waveform matrix $S_1 = S_0$

2) for $m = 2, 3, \cdots$, till convergence
   
   do
   
   (a) Solve the following minimization problem w.r.t $U_m$:
   
   $\min_{U_m} \|S_{m-1} - U_m R\|^2$ \hspace{1cm} (6.9)
   
   s.t. $U_m^H U_m = I$

   the solution of (6.9) is $U^* = VD^H$ where the SVD is $RS_m^H = D\Sigma V^H$.

   (b) Given $U_m^*$, solve the following minimization problem w.r.t $S_m$:
   
   $\min_{S_m} \|S_m - U_m^* R\|^2$ \hspace{1cm} (6.10)
   
   s.t. $\|S_m\| = 1$

   which is equivalent to
   
   $\min_{s_m} \|s_m - b_m\|^2$ \hspace{1cm} (6.11)
   
   s.t. $\|s_m\| = 1$

   where $s_m = \text{vec}(S_m)$ and $b_m = \text{vec}(U_m^* R)$, and the solution of (6.11) is
   
   $s^*_m = b_m / \|b_m\|$ \hspace{1cm} (6.12)

   the associate matrix form is $S^*_m = \text{matrix}(s^*_m)$

   end for

Table 6.1: Waveform Synthesis Algorithm I
6.4 Waveform Synthesis with Constant Modulus and Similarity Constraints

6.4.1 Constant Modulus and Similarity Constraints

In practical applications, constant modulus waveforms are often required due to the limit of nonlinear radar amplifiers. Moreover, as noted in [10, 18], the resulting waveforms generally do not exhibit good pulse compression and ambiguity function properties. Therefore, additional constraints are necessary in the waveform design problem. Here, we focus on the constant modulus and similarity constraints.

The vector form \( s \) of the waveform matrix contains \( N_tN \) elements. Constant modulus constraint is to enforce the modulus of each element of the waveform \( s \) to be a constant. Specifically, considering the normalized transmitted power (i.e., \( \|s\|^2 = 1 \)), the element of \( s \) can be expressed as

\[
s(k) = \frac{1}{\sqrt{N_tN}} e^{j\varphi_k}, k = 1, \ldots, N_tN, \tag{6.13}
\]

where \( \varphi_k \) denotes the phase of each element of \( s \), which is to be determined in the waveform design problem.

Enforcing a similarity constraint on the waveform \( s \) allows a tradeoff between realizing \( \Omega \) and controlling other desired waveform properties (i.e., pulse compression and ambiguity) [17]. We assume that \( C_0 \) is the reference waveform matrix with the associate vector form as \( c_0 = \text{vec}(C_0) \), then the similarity constraint can be expressed as

\[
|s(k) - c_0(k)| \leq \epsilon, k = 1, \ldots, N_tN, \tag{6.14}
\]

where \( \epsilon \) is a real parameter controlling the degree of the similarity. Accounting for the
constant modulus constraint, (6.14) can be further recast as [17].

\[ \varphi_k = \arg s(k) \in [\gamma_k, \gamma_k + \delta_k], \quad k = 1, \cdots, N_t N \]  

(6.15)

where \( \gamma_k \) and \( \delta_k \) are respectively given by

\[ \gamma_k = \arg (c_0(k)) - \arccos(1 - \epsilon^2/2) \]  

(6.16)

\[ \delta_k = 2\arccos(1 - \epsilon^2/2) \]  

(6.17)

with \( 0 \leq \epsilon \leq 2 \). Notice that for \( \epsilon = 0 \), the waveform \( s \) is identical to the reference waveform \( c_0 \), whereas the similarity constraint vanishes and only the constant modulus constraint is in effect when \( \epsilon = 2 \).

### 6.4.2 Iterative Optimization Algorithm

We consider the constant modulus constraint as well as a similarity constraint between the designed waveform and a known radar waveform in this section. Based on the aforementioned discussions, the optimization problem (6.6) with enforcing a similarity constraint with a reference waveform \( c_0 \) is formulated as:

\[
\begin{align*}
\min_s & \quad \|s - b\|^2 \\
\text{s.t.} & \quad |s(k)| = \frac{1}{\sqrt{N_t N}}, \quad k = 1, \cdots, N_t N \\
& \quad |s(k) - c_0(k)| \leq \epsilon
\end{align*}
\]  

(6.18)
which is equivalent to

\[
\min_{s} \ s^H \Phi(s) s \tag{6.19}
\]

s.t. \[ |s(k)| = \frac{1}{\sqrt{N_t N}}, \quad k = 1, \ldots, N_t N \]

arg\(s(k)\) \(\in\) \(\gamma_k, \gamma_k + \delta_k\]

where

\[
\Phi(s) = (1 + b^H b) I - bs^H - sb^H. \tag{6.20}
\]

An iterative optimization algorithm can be used to solve the problem (6.19), and the waveform \(s\) can be found in an iterative fashion. Specifically, at the \(i\)th iteration, we first compute the matrix \(\Phi(s_{i-1})\), where \(s_{i-1}\) denotes the waveform obtained from the \(i - 1\)th iteration. Next, we solve the following problem w.r.t \(s_i\)

\[
\min_{s_i} \ s_i^H \Phi(s_{i-1}) s_i \tag{6.21}
\]

s.t. \[ |s_i(k)| = \frac{1}{\sqrt{N_t N}}, \quad k = 1, \ldots, N_t N \]

arg\(s_i(k)\) \(\in\) \(\gamma_k, \gamma_k + \delta_k\]

and the optimal solution \(s_i^*\) is then used to update \(\Phi(s_i)\) for next-round iteration. This process is repeated until the minimal value of (6.19) is convergent.

We resort to the relaxation and randomization method [39] to solve the problem (6.21). To be specific, first let \(\Phi(s_{i-1}) = \Phi_0\), which means \(\Phi(s_{i-1})\) is a constant matrix and independent of the waveform \(s_i\) in the \(i\)th iteration, we relax the problem (6.21) by
dropping the similarity constraint and the scaling factor \( \frac{1}{\sqrt{NN_t}} \) on \( s_i(k) \):

\[
\begin{align*}
\min_Z & \quad \text{tr} (\Phi_0 Z) \\
\text{s.t.} & \quad \text{diag} (Z) = I \\
& \quad Z = s_i s_i^H,
\end{align*}
\]

where \( \text{tr}(\cdot) \) denotes the trace of a matrix, \( \text{diag}(Z) \) denotes a diagonal matrix whose diagonal elements are the same as \( Z \). Furthermore, (6.22) reduces to a Semi-Definite Programming (SDP) problem by neglecting the rank-one constraint:

\[
\begin{align*}
\min_Z & \quad \text{tr} (\Phi_0 Z) \\
\text{s.t.} & \quad \text{diag} (Z) = I \\
& \quad Z \succeq 0,
\end{align*}
\]

where \( Z \succeq 0 \) indicates that \( Z \) is a semi-definite matrix. The above SDP problem can be effective solved by using the convex optimization toolbox CVX [4] in MATLAB. Then, we can use a randomization method [17, 39] to obtain an approximate solution of \( s_i \) from an Semi-Definite Relaxation (SDR) solution of (6.23). To illustrate the main idea, let \( \mu \) be a random vector with zero mean and covariance \( Z = E \{ \mu \mu^H \} \), and we consider the following stochastic optimization problem

\[
\begin{align*}
\min_{Z = E\{\mu \mu^H \} \succeq 0} & \quad E \{ \mu^H \Phi_0 \mu \} \\
\text{s.t.} & \quad E \{ \text{diag} (\mu \mu^H) \} = I.
\end{align*}
\]

One can see that the problem (6.24) is equivalent to the problem (6.23). Hence, the stochastic interpretation (6.24) of the SDR (6.23) allows us to obtain approximate rank-one solu-
tions. Next, we describe how the randomization method can be integrated with the similarity constraint. First, let $Z^*$ be an optimal solution to (6.23), and generate independent identically distributed Gaussian random vectors $\mu_l$, i.e., $\mu_l \sim \mathcal{N}(0, B)$ for $l = 1, \cdots, L$, where $L$ is the number of the randomization trials. The covariance matrix $B$ is constructed as

$$B = Z^* \odot [pp^H]$$

(6.25)

where $\odot$ denotes Hadamard product, $p$ is given by

$$p = \frac{1}{\sqrt{N_tN}} [e^{-j\gamma_1}, \cdots, e^{-j\gamma_{N_tN}}]^T,$$

(6.26)

and $\gamma_k$ is given in (6.16). Then, for the $l$th randomization trial, we assign

$$s_{i,l}(k) = p^*(k)f(\mu_l(k))$$

(6.27)

for $k = 1, \cdots, N_tN$, where

$$f(\mu_l(k)) = \exp\left(j \frac{\text{arg}(\mu_l(k))}{2\pi} \delta_k\right)$$

(6.28)

and $\delta_k$ is given in (6.17). As $0 \leq \frac{\text{arg}(\mu_l(k))}{2\pi} \leq 1$, it is clear that (6.27) ensures the similarity constraint (6.15) is met. Finally, the best solution among the $L$ randomization is selected as the one which minimizes the objective function

$$s_i^* = \arg\min_{s_i,l} s_i^H \Phi_0 s_i,l$$

(6.29)

The randomization method is known to yield to a good approximation provided that a solution for sufficient number of randomization trials is employed [39].

Note that the problem in [17] is similar with (6.21), which ignores the dependence
of $\Phi(s)$ on the waveform $s$, i.e., $\Phi(s) = \Phi_0$. However, solving our problem (6.19) with the waveform dependent $\Phi(s)$ defined in (6.20), the aforementioned iterative algorithm is required.

### 6.4.3 Summarization of the Proposed Algorithm

We summarize the proposed algorithm used to solve the problem (6.18) with the constant modulus constraint (6.13) and similarity constraint (6.14) as in the following table,

The stop criteria of the proposed algorithm are justified next. Note that at the $i$th iteration of the second loop, the object function in (6.18) and (6.19) can be calculated as

$$\rho_i = s_i^H \Phi(s_{i-1}) s_i.$$  \hspace{1cm} (6.31)

One stop criteria of the second loop can hence be given as $|\rho_i - \rho_{i-1}| < \tau$, where $\tau$ is a user selected parameter to control convergence. Similarly, we can also determine the iteration number $M$ in the first loop based on the cost function of (6.6).

### 6.5 Numerical Results

In the simulations we let $N_t = 4, N = 16$ and randomly select a full-rank symmetric positive-definite matrix as the desired transmit covariance matrix $\Omega$. The overall transmit power is set to 1, i.e., $\text{tr} (\Omega) = 1$. We consider the orthogonal linear frequency modulation (LFM) as the reference waveform matrix $C_0$, and $c_0 = \text{vec}(C_0)$ with $c_0^H c_0 = 1$. We also use the reference waveform matrix as the initial value for the proposed algorithm, namely, $S_0 = C_0$. The iteration numbers used to update the optimal solution $S_m$ are $M = 20$ for the first loop and $I = 10$ for the second loop, respectively. For each iteration of the second loop, the number of randomization trials is $L = 1000$.

The square of the difference between $S^H S$ and $\Omega$ versus the iteration number $m$ for
Waveform Synthesis Algorithm II:

**Input:**
- Waveform correlation matrix $\Omega$; Reference waveform matrix $C_0$; Initial waveform matrix $S_0$

**Output:**
- A solution $S^*$ of (6.18) with constraints (6.13) and (6.14)

1) When $m = 1$, initialize the transmit waveform matrix $S_1 = S_0$

2) for $m = 2, 3, \ldots, M$
   do
     (a) Update $U_m$ by using (6.5) with pre-calculated $S_{m-1}$
     (b) Update $S_m$ by solving the problem (6.18) with an iterative algorithm:
        i. When $i = 1$, initialize the transmit waveform $s_1 = \text{vec}(S_{m-1})$.
        ii. for $i = 2, 3, \ldots, I$
            do
                A. Compute $\Phi(s_{i-1})$ in terms of (6.20)
                B. Solve the SDP problem (6.23) with $\Phi(s_{i-1})$, and denote by $Z^*$ a solution
                C. Generate a random vector $\mu_l \sim \mathcal{N}(0, B)$ where $l = 1, \ldots, L$ is the number of randomization trials, and $B$ is given as (6.25).
                D. For the $l$th randomization trial, let $s_{i,l}(k) = p^*(k)f(\mu_l(k))$, where $k = 1, \ldots, N_tN$, and $f(\mu_l(k))$ is defined in (6.28). Choose $s_i$ from \{s_{i,l}\}_{l=1}^L$, such that
                    \[
                    s_i = \arg\min_{s_{i,l}} s_{i,l}^H \Phi(s_{i-1}) s_{i,l}
                    \] (6.30)
            end for
        iii. $S_m = \text{matrix}(s_I)$
   end for

3) $S^* = S_M$

Table 6.2: Waveform Synthesis Algorithm II
Figure 6.1: Square error between $S^H S$ and $R$ with the similarity parameter $\epsilon = 0, 0.5, 1.5, 1.8, 2$ respectively.

Various similarity parameters are given in Fig 6.1. It shows that the square of the difference is decreased as the similarity become loosen i.e., $\epsilon$ is increased. When $\epsilon = 0$, it’s curve is overlapped with that of $C_0$. When $\epsilon = 2$, the similarity constraint vanishes.

Fig 6.2 shows the amplitude and phase of the reference LFM waveform and the designed waveforms, where several levels of similarity is considered. It can be seen that the amplitude of all waveforms are constant and same, and as $\epsilon$ decreases, the designed waveforms become more and more similar to the LFM. This behavior agrees with the fact that the smaller the value of $\epsilon$, the stronger the phase constraint on the designed waveforms.

### 6.6 Conclusions

We have addressed the problem of waveform synthesis for MIMO radar. As optimization of a performance metric directly with respect to the signal matrix can lead to an intractable problem even under a relatively simple constraint, the correlation matrix $\Omega$ of the probing waveforms are preferred to be optimized during the beampattern design first. With a pre-
Figure 6.2: Amplitude and phase of the reference waveform and the designed waveforms with the similarity parameter $\epsilon = 0, 0.5, 1.5, 1.8, 2$ respectively.

specified $\Omega$, we synthesize a signal waveform matrix $S$ that, under the constant modulus constraint as well as a similarity constraint, realizes (at least approximately) the desired correlation matrix. We have presented an optimization algorithm employing a cyclic method and an iterative approach to iteratively update the optimal solution. The similarity parameter $\epsilon$ is used to control the degree of the similarity between the synthesized waveform and a reference waveform. Numerical results show that the larger $\epsilon$ (i.e., the weaker the similarity constraint), the smaller of the difference between the correlation matrix of the $S$ and the desired $\Omega$. This suggests a suitable tradeoff between the correlation matrix realization and the similarity to a benign reference waveform.
Chapter 7

Conclusion

The adaptive algorithm is a practical technique which allows utilizing the training data to estimate some unknown parameters for further processing. In this dissertation, we apply the adaptive processing to both the transmitter and receiver design of array radar systems. At the receiver side, we propose an adaptive matched filter detector which can mitigate the training and computation requirements. We then extend the adaptive processing to the transmitter side by jointly designing the transmit and receive beamforming for array radar. To complete the design exercise, a waveform synthesis algorithm by considering some practical constraints is also investigated in this dissertation.

In Chapter 4, we considered an adaptive reduced-rank approach by employing the iterative conjugate gradient (CG) algorithm and the adaptive matched filter (AMF). The resulting reduced-rank CG-AMF detector is a projection of the full-rank AMF to a Krylov subspace with a dimension determined by the number of CG iterations. Asymptotic analysis of the output SINR of the CG-AMF detector was carried out by exploiting a relation between the CG algorithm and the Lanczos method. Our results show that, in the training limited cases, not only is the CG-AMF detector computationally more efficient (thanks to the efficiency of the CG), but it also often enjoys the benefit of reaching the peak output SINR with a lower rank, when compared with several other popular reduced-rank solutions.

In Chapter 5, we have proposed a jointly adaptive transmit and receive beamforming for array radars. The transmit and receive beampattern is obtained by jointly designing the transmit beamforming correlation matrix and receive beamforming vector in terms of maximizing the output SINR. The adaptive processing is achieved in two stages. In the initial probing stage, we first transmit multiple uncorrelated waveforms with omnidirectional
transmission to probe a target-free environment. In the next stage, the training data obtained from initial probing are then used for interference estimation and beamformer design. The numerical results show that by applying adaptive processing for both radiation and receiving in a training-limited situation, we can achieve a better beampattern, a stronger ability to handle interference, and a higher output SINR.

In Chapter 6, we have addressed the problem of waveform synthesis for MIMO radar. As optimization of a performance metric directly with respect to the signal matrix can lead to an intractable problem even under a relatively simple constraint, the correlation matrix $\Omega$ of the probing waveforms are preferred to be optimized during the beampattern design first. With a pre-specified $\Omega$, we synthesize a signal waveform matrix $S$ that, under the constant modulus constraint as well as a similarity constraint, realizes (at least approximately) the desired correlation matrix. We have presented an optimization algorithm employing a cyclic method and an iterative approach to iteratively update the optimal solution. The similarity parameter $\epsilon$ is used to control the degree of the similarity between the synthesized waveform and a reference waveform. Numerical results show that the larger $\epsilon$ (i.e., the weaker the similarity constraint), the smaller of the difference between the correlation matrix of the $S$ and the desired $\Omega$. This suggests a suitable tradeoff between the correlation matrix realization and the similarity to a benign reference waveform.
Bibliography


Parametric adaptive matched filter for airborne radar applications. *IEEE Transactions


[48] L. L. Scharf, E. K. P. Chong, M. D. Zoltowski, J. S. Goldstein, and I. S. Reed. Sub-
space expansion and the equivalence of conjugate direction and multistage wiener


exponential integrals $E_1(x), x > 0$. *Journal of Computational Physics*, 25:199–204,
1977.

[51] Petre Stoica, Jian Li, and Yao Xie. On probing signal design for MIMO radar. *IEEE

[52] Petre Stoica, Jian Li, and Xumin Zhu. Waveform synthesis for diversity-based trans-
2008.

[53] L. N. Trefethen and David Bau,III. *Numerical Linear Algebra*. Society for Industrial


Vita

Zhu Chen

Address 53 Poplar Street, Jersey City, NJ 07307

Place of birth Nanyang, Henan, People’s Republic of China

Education

Stevens Institute of Technology, Hoboken, NJ
Doctor of Philosophy, Electrical Engineering
Sep. 2009 - Des. 2013

Chonbuk National University, Jeonju, South Korea
Master of Engineering, Electrical Engineering
Aug. 2008

South-Central University for Nationalities, Wuhan, China
Bachelor of Engineering, Automation
Jun. 2006

Industrial Experience

Broadcom Inc., Matawan, NJ
System Design Engineer (Sep. 2013 - present)

Mitsubishi Electric Research Laboratories, Cambridge, Boston
Research Intern (May 2012 - Sep. 2012)

Publications


Computation for OFDM in Cognitive Radio System.

*Journal of the Institute of Electronics Engineers of Korea, Vol.45-TC, No.2*


Yan, Y., Kim, J., Chen, Z., & Lee, M.(2008), Modified Clipping for Iterative Decoding of Superposition Coding. *Journal of the Institute of Electronics Engineers of Korea, Vol.45-TC, No.2*


**Patents**

“Method for estimating frequencies and phases in three phase power system”, Patent application filed, 2013, joint inventor with Sahinoglu, Z., MERL, USA.

**Academic Awards**

Best Student Paper Finalist, 2013 IEEE Radar Conference