COMPUTATIONALLY EFFICIENT ALGORITHMS FOR ROBUST ADAPTIVE BEAMFORMING AND MULTIUSER DETECTION

By

Amr ElKeyi, B.Sc., M.Sc.

2006

A Thesis

Submitted to the Department of Electrical & Computer Engineering

and the School of Graduate Studies

in Partial Fulfilment of the Requirements

for the Degree of

Doctor of Philosophy

McMaster University

© Copyright by Amr ElKeyi, 2006

Canada

COMPUTATIONALLY EFFICIENT ALGORITHMS FOR ROBUST ADAPTIVE BEAMFORMING AND MULTIUSER DETECTION

DOCTOR OF PHILOSOPHY (2006)   McMaster University

(Electrical & Computer Engineering)   Hamilton, Ontario

TITLE: Computationally Efficient Algorithms for Robust Adaptive Beamforming and Multiuser Detection

AUTHOR: Amr ElKeyi

                        B.Sc. (Communications and Electrophysics),

                        M.Sc. (Communications Engineering),

                        Alexandria University, Alexandria, Egypt.

SUPERVISORS: Alex B. Gershman, Professor

                        Thia Kirubarajan, Professor

NUMBER OF PAGES: xxii, 180

Dedications

To my parents, my family,

and my beloved country, Egypt.

iv

Abstract

Classical adaptive beamforming techniques are based on idealistic assumptions that are not valid in many practical situations. One of the serious problems that occur in practical adaptive array systems is the mismatch between the presumed and actual array manifolds. This mismatch can arise due to look direction errors, sensor location errors, or wavefront distortions due to propagation in inhomogeneous media. A similar problem occurs in multiuser detection where the desired user signature is not precisely known at the receiver because of the effect of the communication channel, e.g., frequency selective fading. Recently, there has been much interest in developing mathematically rigourous adaptive array processing and multiuser detection algorithms that provide an amount of robustness directly related to the amount of uncertainty in the signal environment. In this thesis, we aim to develop computationally efficient robust adaptive beamforming and multiuser detection algorithms that are applicable to practical nonstationary environments where the array and signal characteristics are not precisely known.

First, we consider the narrowband signal environment. We develop a novel computationally efficient implementation of the robust minimum variance distortionless response (MVDR) beamformer. This beamformer is based on worst-case performance optimization and has been shown to provide excellent robustness against arbitrary but norm-bounded mismatches in the desired signal steering vector. However, existing algorithms that solve this problem do not have direct online implementations.

v

We develop an extended Kalman filter (EKF)-based algorithm to estimate the beamformer weight vector adaptively with lower computational cost per iteration than that of previous existing algorithms. This makes it very attractive in practical nonstationary scenarios where the beamformer weight vector has to be updated whenever a new data snapshot is received. We also present two improved modifications of our algorithm to additionally account for abruptly changing nonstationary environments.

Next, we develop a robust wideband beamformer that extends previous narrowband approaches while avoiding suboptimal subband decomposition. Our beamformer provides high gain not only to the presumed desired signal but also to all signals within a pre-specified uncertainty set. The phase response of the array is controlled through additional linear phase constraints on each of the finite impulse response (FIR) filters of the array processor. The problem is formulated as a convex optimization problem and two implementations are presented which can be solved efficiently in polynomial time using well-established interior point methods. We also present an adaptive implementation for our robust wideband beamformer and a first- or second-order EKF can be used to estimate the beamformer weight vector adaptively with low computational complexity per iteration. Simulation results are presented showing a superior performance of the proposed algorithms and an improved robustness against various types of mismatches compared to earlier approaches.

Furthermore, we develop a state-space approach to the blind multiuser detection problem with robustness against arbitrary mismatches in the desired user signature. Our solution is obtained adaptively using a second-order EKF and requires the same computational complexity as that of the previous non-robust recursive least squares (RLS)-based algorithms. We also develop a state-space approach to the decision-directed multiuser detection problem and an algorithm for switching between robust blind and decision-directed detectors. The proposed switching algorithm combines

vi

the advantages of both these detection schemes. It can achieve an output signal-to-interference-plus-noise ratio (SINR) close to that of the minimum mean square error (MMSE) detector, even in the presence of mismatches in the desired user signature and without any need for training. Therefore, it is well-suited to nonstationary environments where users repeatedly enter and leave the system making the cost of training unaffordable.

viii

Acknowledgements

All deepest thanks are due to Almighty God,

the Most Merciful, the Most Compassionate.

My first acknowledgment goes to my advisor, Prof. A. B. Gershman, for his valuable guidance, encouragement, and support. His enthusiasm, approachability, and patience were essential to the completion of this dissertation. I am also grateful to Prof. T. Kirubarajan, my co-advisor, for his insightful comments, and for his valuable advice and suggestions. In addition, I would like to thank my supervisory committee members, Prof. D. Pelinovsky and Prof. S. Hranilovic, for their help and support.

I owe a huge debt of gratitude to my professors at Alexandria University and McMaster University for all that I have learned from them during my B.Sc., M.Sc., and Ph.D. studies. Specifically, I would like to express my gratitude to Prof. Said ElKhamy, my M.Sc. supervisor, for his parental guidance both on academic and non-academic matters, and his endless effort which contributed greatly towards building my background as a researcher in the field of signal processing. I would also like to thank Prof. Z.-Q. Luo for the valuable experience I have gained from him throughout the courses he has taught at McMaster University.

I would like to give thanks to all my colleagues at McMaster University who have provided a supportive and exciting research atmosphere. In addition, I would like to thank my best friend Amr Nour-Eldin, and my friend Dr. Aboulnasr Hassanien, for proof reading this thesis and for their helpful comments.

ix

I would like to sincerely thank my mother for her uninterrupted love and support, my father, and my family in Egypt for their continuous encouragement throughout my life. I would also like to thank the family and friends I have found in Hamilton. Their help and support has made Canada my second home.

Finally, I would like to dedicate sincere feelings of devotion to my wonderful home country: Egypt — I really owe you a lot.

x

List of Acronyms

BER	Bit Error Rate
BM	Blind Mode
CBRMV	Constant Beamwidth Robust Minimum Variance
CDMA	Code Division Multiple Access
CPRMV	Constant Powerwidth Robust Minimum Variance
DDKF	Decision Directed Kalman Filter
DDM	Decision Directed Mode
DDRLS	Decision Directed Recursive Least Squares
DFT	Discrete Fourier Transform
DOA	Direction of Arrival
EKF	Extended Kalman Filter
FIR	Finite Impulse Response
GSC	Generalized Sidelobe Canceller
IDFT	Inverse Discrete Fourier Transform
IMM	Interacting Multiple Model
INR	Interference−to−Noise Ratio
ISR	Interference−to−Signal Ratio
LCMV	Linearly Constrained Minimum Variance
LMI	Linear Matrix Inequality
LMS	Least Mean Squares

MAI	Multiple Access Interference
MMSE	Minimum Mean Square Error
MOE	Minimum Output Energy
MSE	Mean Square Error
MVDR	Minimum Variance Distortionless Response
NIS	Normalized Innovation Square
NMSE	Normalized Mean Square Error
NS	Necessary and Sufficient
NS1−PS	Necessary and Sufficient First−Order Presteering
OFDM	Orthogonal Frequency Division Multiplexing
pdf	Probability Density Function
RBKF	Robust Blind Kalman Filter
RLS	Recursive Least Squares
SDP	Semi−Definite Programming
SINR	Signal−to−Interference−plus−Noise Ratio
SMI	Sample Matrix Inversion
SNR	Signal−to−Noise Ratio
SOC	Second−Order Cone
SOCP	Second−Order Cone Programming
ULA	Uniform Linear Array

List of Notations

Boldface lowercase letters are used to denote column vectors.

Boldface uppercase letters are used to denote matrices.

j	the unit imaginary number $j = \sqrt{-1}$
$(\cdot)^*$	the conjugate operator
$(\cdot)^T$	the transpose of a vector or a matrix
$(\cdot)^H$	the Hermitian transpose of a vector or a matrix
$(\cdot)^{-1}$	the inversion of a matrix
a_i	the ith element of the vector $\boldsymbol{a}$
$[\cdot]_{m,n}$	the mnth element of a matrix
$\lvert \cdot \rvert$	the determinant of a matrix
$\lVert \cdot \rVert$	the Euclidean norm of a vector
$\lVert \cdot \rVert_i$	the ith norm of a vector
$\lVert \cdot \rVert_F$	the Frobenius norm of a matrix
$\langle \cdot, \cdot \rangle$	the inner product of two matrices
$\lceil x \rceil$	the integer greater than or equal to x
$\otimes$	the Kronecker matrix product
∇	the gradient operator
∇^2	the Laplacian operator
λ	the wavelength

xiii

σ_n^2	the noise power
σ_s^2	the desired signal power
$\mathbf{0}_{m \times n}$	the $m \times n$–dimensional matrix containing all zeros
$\mathbf{1}_k$	the k–dimensional vector containing all ones
$\boldsymbol{e}_i$	the vector containing all zeros except for 1 in the ith position
$\boldsymbol{I}_k$	the identity matrix of dimension $k \times k$
$\boldsymbol{R}_{i+n}$	the interference$-$plus$-$noise covariance matrix
$\boldsymbol{R}_x$	the true convariance matrix
$\hat{\boldsymbol{R}}_x$	the sample convariance matrix
$\mathrm{E}\{\cdot\}$	the statistical expectation operator
$\mathrm{Im}\{a\}$	the imaginary part of the complex number a
$\mathrm{angle}\{a\}$	the phase of the complex number a
$\mathrm{diag}\{\mathbf{a}\}$	the diagonal matrix with the ith diagonal element a_i
$\mathrm{sign}\{\cdot\}$	the signum operator which returns ± 1
$\mathrm{tr}\{\cdot\}$	the trace of a square matrix
L	the number of taps of an FIR filter
L_c	the central tap index of an FIR filter
L_A	the length of the array aperture
M	the number of sensors
N	the number of points in the discretization grid
T_s	the sampling frequency

xiv

Contents

xx

List of Figures

Chapter 1

Introduction

In the early days of development of statistical signal processing techniques, the emphasis was on the derivation of *optimum* schemes designed to maximize the output signal-to-interference-plus-noise ratio (SINR) or minimize the bit error rate (BER) in a specific signal environment. However, optimum signal processing algorithms often suffer from drastic performance degradation in the presence of even small deviations from the nominal model assumptions upon which they were derived. This observation was the motive for searching for *robust* signal processing techniques; that is, techniques with good performance under nominal conditions and acceptable performance for any signal environment within a certain pre-specified uncertainty set [61]. Thus, in seeking robust signal processing techniques it is recognized that a single precise characterization of the operating environment is an unrealistic assumption, and therefore a set of operating conditions has to be considered and taken into account while designing practical signal processing schemes.

One of the frequently used assumptions for the signal environment is that it is stationary and ergodic, i.e., with fixed statistics over time. This assumption has simplified the derivation of optimum statistical signal processing techniques as all the signal statistics can be evaluated from a single realization of the process. However,

1

in practical applications this assumption is not valid. Therefore, there is a need for adaptive signal processing algorithms that can be updated in real-time with low computational complexity and, hence, are suitable for nonstationary environments.

The aim of this thesis is to derive computationally efficient robust algorithms for adaptive beamforming and multiuser detection. In this chapter, we will give a brief introduction to the array signal processing and multiuser detection areas. We will also motivate the objectives of this thesis. At the end of the chapter, we provide an outline of the thesis and summarize our main contributions.

1.1 Adaptive Array Processing

Array processing is a branch of signal processing that uses sensor arrays to extract information contained in spatially-propagating signals. This information content can be either the location parameters of the source or the transmitted signal waveform itself. The two main fields of array processing are direction of arrival (DOA) estimation and beamforming. The goal of beamforming is to perform spatial filtering in order to separate signals that originate from different spatial locations. This is achieved by providing different gains to different spatially-propagating signals based on their DOAs.

Beamformers can be classified into two categories depending on the way they are designed; either data-independent or adaptive. Data-independent beamformers are designed such that they approximate a desired spatial response that is independent of the operating signal environment. This design objective is similar to that of classical finite impulse response (FIR) filter design [86]. On the other hand, adaptive beamformers are designed based on the received sensor data or their statistics. The design objective of adaptive beamformers is to maximize the output SINR of the array which is equivalent to minimizing the output power due to noise and signals arriving from

2

directions other than the desired signal direction while preserving the desired signal.

Adaptive beamforming dates back to the work of Howells and Applebaum that started in mid-1950s [4], [58]. In 1957, Howells developed a two-channel adaptive sidelobe canceller system which can suppress a single sidelobe jamming source. In 1966, Applebaum devised the control logic for computing the weights of a general Howells adaptive array. Since then, adaptive beamforming has received significant interest in the literature and has been the subject of numerous research papers and books (see [21], [49], [50], [65], [109] and the references therein).

Adaptive beamforming has played an important role in many areas such as radar, sonar, and wireless communications. In what follows, we briefly describe some adaptive beamforming applications.

Radar

Radar systems are used in many contexts, including air traffic control, police detection of speeding traffic, and military reconnaissance. The term "RADAR" was coined in 1941 as an acronym for **RA**dio **D**etection **A**nd **R**anging, and has since entered the English language as a standard word losing the capitalization in the process. Most radar systems are active, i.e., antenna arrays are used to transmit electromagnetic energy and to listen for the echo of the transmitted signal. The transmitted signal can either be a continuous wave or a pulse modulated signal. Adaptive beamforming is widely used in radar systems to enhance the detection of target echoes through interference/jamming cancellation and ground clutter suppression. Examples can be found in surface wave over-the-horizon radars used for coastal surveillance [35], [100] and airborne moving target indication radars used to detect moving targets [10].

3

Sonar

The use of sonar systems to detect underwater targets dates back to the World War I [111]. The word "SONAR" was first used in the World War II as an acronym for **SO**und, **NA**vigation, and **R**anging. Sonar systems use hydrophones that are capable of converting underwater acoustic energy to electric signals. They can be either passive or active. The theory of active sonar has much in common with radar. However, the design of sonar systems is more sophisticated than that of radars due to the complicated propagation properties of the underwater acoustic wavefield [13]. Adaptive beamforming is usually applied to sonar systems as a preprocessing step to suppress strong interferers [46]. Hydrophone array-based sonar has been successfully applied to real data as reported in [46], [66], [75], and [96].

Wireless communications

The increasing demand for mobile communications in the last two decades has motivated the search for new techniques to increase the capacity, reliability, and quality of service of wireless communication systems. One approach that has shown real promise of substantial improvements is the use of spatial processing with adaptive antenna arrays [50]. Adaptive beamforming is capable of separating signals transmitted at the same time and in the same frequency band, provided that they are separated in the spatial domain [72]. In addition, the use of antenna arrays in mobile communications offers the ability to extend the range of coverage, combat fading, and provide more robustness against co-channel interference by focusing the beam in the direction of the desired user and nulling-out the interference [21], [49], [95].

4

1.2 Multiuser Detection

Multiuser detection (also known as co-channel interference suppression) deals with the demodulation of mutually interfering digital streams of information. Cellular telephony, satellite communications, and high-speed data transmission lines are examples of communication systems that suffer from multiple access interference (MAI) [114]. The interference between the received signals may originate from the unideal characteristics of the transmission medium as in the case of orthogonal code division multiple access (CDMA) systems, time division multiple access systems, and orthogonal frequency division multiplexing (OFDM), or it may be an integral part of the multiplexing method as in the case of non-orthogonal CDMA.

When CDMA was first commercially launched in 1995, the presence of MAI at the receiver was simply neglected on the grounds that its statistical properties would be similar to additive white Gaussian noise, and thus the use of a single-user matched filter to combat such interference would be nearly optimal. In order to limit the multiple access interference, very accurate and fast power control is required in the IS-95 system to keep the levels of the received signals within a fraction of one decibel [71]. Unfortunately, ultra-fast ultra-accurate power control is not always feasible or cost-effective. In the last two decades, a large number of results under the name "Multiuser Detection" has shown that the philosophy of combatting MAI with power control only is inefficient and that full channel utilization can only be achieved by the additional use of signal processing techniques such as multiuser detection and beamforming [60], [122]. Multiuser detection exploits the structure of the MAI to increase the efficiency with which channel resources are utilized, hence increasing the capacity, coverage, and quality of service of multiple access communication systems [113].

1.3   Overview of the Thesis

1.3.1   Motivation and objectives of the thesis

Most signal processing techniques for adaptive beamforming and multiuser detection are based on the following assumptions:

1. The received signal environment is precisely known at the receiver.

2. The signal and noise processes are stationary over the observation interval.

However, in practical operating environments these assumptions are rarely valid, and existing conventional signal processing algorithms suffer from significant performance degradation due to their high sensitivity to any deviations from these assumptions. In this section, we will briefly discuss a few problems that occur in practical adaptive beamforming and multiuser detection systems.

Mismatches in the signal environment

Classical adaptive beamforming techniques are based on the assumption that the signal propagation model and array characteristics are precisely known. However, in practical situations many mismatches exist between the *presumed* signal environment and sensor array characteristics, and the *actual* ones [40]. Examples of such mismatches include imperfect knowledge of the DOA of the desired signal (look direction errors) [47], imperfect array calibration and distorted antenna shape [59], and unideal channel propagation effects such as wavefront distortions due to propagation in inhomogeneous media and source local scattering [42]. The performance of classical adaptive beamforming techniques is known to degrade severely in the presence of even slight mismatches of any of these types as the beamformer interprets the desired signal as interference and accordingly suppresses it by adaptive nulling instead of preserving it [36], [41]. Hence, robust approaches to adaptive beamforming are of

6

great importance. Most existing approaches to robust adaptive beamforming use ad-hoc modifications of the optimum adaptive beamforming techniques to guard against performance degradation under certain mismatch scenarios such as look direction errors or sensor location errors [12], [14], [34], [38], [106], [124]. There is a need for mathematically rigourous adaptive beamforming algorithms that can combat multiple mismatches in the operating environment and provide an amount of robustness directly related to the amount of environmental mismatches.

Quite a similar problem is typical for multiuser detection. In CDMA systems, each user is assigned a specific signature that uniquely identifies him from other users. There are two main approaches to multiuser detection; training-based and blind detection. In training-based techniques, a data sequence (which is already known by the receiver) is sent by the transmitter and is used to estimate the receiver weight vector [81]. This causes a waste of the transmission bandwidth. On the other hand, blind multiuser detection does not need any training. It requires only the knowledge of the desired user signature and precise timing [57]. However, it suffers from severe performance degradation for even slight mismatches between the presumed and actual signatures of the desired user. These mismatches frequently arise due to the effect of propagation through the physical channel [24]. For example, the receiver might assume that the desired user signature is the nominal transmitted one, whereas the actual received signature may contain additional multipath components or other types of channel-related signal distortion. Existing approaches that provide robustness against these mismatches are ad-hoc and frequently do not yield satisfactory performance [24], [45], [57]. More work is required to develop multiuser detection algorithms that do not require training and can maintain high output SINR even in the presence of uncertainties in the desired user signature.

7

Nonstationary environments

A typical cause of performance degradation in adaptive signal processing algorithms is the nonstationary nature of the operating environment which can be attributed mainly to the effect of multipath propagation. Even small source motion can lead to different path delays and thus the overall channel response can change dramatically over time. This effect is more pronounced for rapidly moving sources as the propagation loss along each path will vary and induce variations in the received signal power. Note that moving sources are quite typical for radar, sonar, and wireless communications [44], [55], [94]. The same situation occurs if the receiver is moving, e.g., towed hydrophone arrays in sonar [75], or moving antenna platforms in airborne radar [10], [55]. Moreover, these systems are liable to abrupt environmental changes such as the powering up of a new user or interference signal which can cause drastic performance degradation.

This nonstationary behaviour of practical operating environments limits the number of available data samples required for "training" signal processing algorithms [91]. Furthermore, it may happen that these algorithms are not able to adapt fast enough to follow the environmental changes, and as a result a considerable amount of interference is not filtered out causing severe performance degradation [43], [53]. Hence, computationally efficient and rapidly convergent algorithms are needed that can be implemented in real-time and easily updated every time a new data sample is received.

1.3.2   Thesis outline

The main objective of this thesis is to derive mathematically rigourous and computationally efficient algorithms for adaptive beamforming and multiuser detection that are robust against mismatches between the presumed and actual signal environments. The main body of the thesis consists of Chapters 3 to 6. In what follows, we give a

8

brief overview of our work.

In Chapter 2, we overview the conventional techniques for adaptive beamforming. We present the signal model that has been widely used in array signal processing over the last few decades. We discuss several popular adaptive beamforming techniques which are optimal when the array and signal characteristics are precisely known. Next, we review some classical robust beamforming approaches which can maintain satisfactory performance in the presence of certain mismatches between the presumed and actual signal environments. At the end of the chapter, we discuss real-time implementations of adaptive beamforming algorithms that are suitable for nonstationary environments.

We start Chapter 3 by revising the recently proposed narrowband robust minimum variance distortionless response (MVDR) beamformer [116] which is one of the theoretically rigourous and most powerful approaches to robust beamforming in the presence of arbitrary mismatches in the desired signal steering vector. Next, we develop a Kalman filter-based technique to implement the robust MVDR beamformer online. The proposed beamformer can be updated whenever a new data snapshot is received and requires less computations than those required for earlier implementations of the MVDR beamformer. Furthermore, we propose two modifications of our Kalman filter-based beamformer which additionally account for abruptly changing nonstationary environments. At the end of the chapter, we provide simulation results that compare our Kalman filter-based implementation of the robust MVDR beamformer to its original implementation of [116].

In Chapter 4, we present a novel wideband beamformer with robustness directly related to the amount of uncertainties in the array manifold. Our beamformer extends the work in [116] to the more general wideband case and avoids the suboptimal subband decomposition approach of [118]. The proposed beamformer prevents the

9

cancellation of the desired signal components even in the presence of array manifold mismatches and maintains a linear phase response towards the desired signal. The resulting problem is solved using worst-case performance optimization and is shown to be convex. We also provide two implementations for our beamformer which can both be solved with polynomial complexity using well-established interior point methods [83]. Simulation results are presented showing an improved robustness of our beamformer against various mismatches compared to the earlier techniques.

In Chapter 5, we modify the proposed wideband beamformer to derive an online implementation suitable for nonstationary environments. We use a state-space modelling approach similar to that we adopted in Chapter 3, and an extended Kalman filter (EKF) is used to solve for the optimal weight vector adaptively with a reduced computational cost compared to the previous interior point methods-based implementations. Simulation results are presented validating a superior performance of our algorithm both in stationary and nonstationary environments.

In Chapter 6, we consider the problem of robust multiuser detection. First, we develop a state-space approach to the blind multiuser detection problem with robustness against arbitrary mismatches in the desired user signature. Our solution is obtained using a second-order EKF and has the same computational complexity as that of previous non-robust online algorithms. We also present a state-space approach to the training-based decision-directed multiuser detection problem. Furthermore, we develop an algorithm for switching between robust blind and decision-directed detection. Simulation results are presented showing an improved performance of our algorithms compared to earlier multiuser detection techniques.

In Chapter 7, we give our concluding remarks and point out some important future research directions.

1.3.3   Contributions and related publications

The contributions of this thesis can be summarized in the following developments:

1. A novel computationally efficient online implementation for the robust MVDR beamformer which is well-suited to nonstationary environments.

2. A new mathematically rigourous wideband beamforming algorithm with robustness against multiple mismatches in the array manifold.

3. A computationally efficient real-time robust wideband beamforming algorithm.

4. A computationally efficient recursive implementation of the bind multiuser detector with robustness against desired user signature mismatches.

5. A novel algorithm for switching between robust blind and decision-directed multiuser detection.

The contents of Chapter 3 have been published in one full *IEEE Transactions on Signal Processing* paper [27]. An earlier exposition of some parts of Chapter 3 was also presented at the *IEEE Workshop on Statistical Signal Processing* [29]. The contents of Chapter 4 have been presented in part at the *IEEE Workshop on Statistical Signal Processing* [28] and the full contents of the Chapter have been submitted to the *IEEE Transactions on Signal Processing*. Some of the contents of Chapter 5 have been presented at the *IEEE Workshop on Sensor Array and Multi-channel Processing* [31], and a full journal version is currently under preparation. The contents of Chapter 6 have been presented at the *IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing* [30], and have been also accepted for publication in the *IEEE Transactions on Wireless Communications* [32].

11

Chapter 2

Conventional Adaptive Beamforming

2.1 Introduction

Adaptive beamforming has received considerable attention during the last decades, particularly in the fields of speech acquisition, sonar, radar and, more recently, in wireless communications [65], [110], [112]. In all these applications, beamforming (spatial filtering) is used to distinguish between the desired signal, interference, and noise based on the spatial properties of each. In order to enable the beamformer to respond to an unknown interference environment, it has to be *adaptive*, i.e., it can extract the information about the environment from the received data and filter out the interference and noise while preserving the desired signal.

We start this chapter with an overview of array processing fundamentals, including spatially-propagating waves and the array signal model widely used in the literature. Next, we discuss main adaptive beamforming techniques which are optimum when the array and signal characteristics are precisely known. Both narrowband and wideband optimum beamforming approaches are considered. We also review the most important

13

classical robust beamforming approaches that can maintain satisfactory performance in the presence of mismatches between the presumed and actual signal environments. At the end of the chapter, we discuss online implementations of adaptive beamforming algorithms that are suitable for nonstationary environments.

2.2   Array Signal Model

Modern approaches to array signal processing are model-based, in the sense that they rely on certain assumptions about the observed data [65], [109], [112]. In this section, we present the array signal model commonly used in adaptive beamforming. We start with discussing the wave equation and its solution in the case of a spherical wave propagating in a homogeneous medium. Next, we introduce the concept of sensor arrays that spatially sample the wavefield at discrete locations. In most practical beamforming applications, the array sensors are placed far enough from the wavefield source and, hence, the propagating wave appears as a plane wave. We discuss the plane wave solution to the wave equation which will be used throughout this thesis. Finally, we present the narrowband array signal model which is applicable when the bandwidth of the desired signal is small compared to its carrier frequency.

2.2.1   Propagating waves

A wave is defined as a disturbance that propagates from one location to another. This disturbance might occur in the electric and magnetic fields as in the case of electro-magnetic waves, in the air pressure as in the case of acoustic waves propagating in air [33], or in the ground displacement as in the case of seismic waves [99]. Propagating waves are functions of space and time. Their propagation can be characterized

14

by the wave equation [33, p. 240], [99, p. 27]

$$\nabla^2 s(t, \boldsymbol{r}) = \frac{1}{c^2} \frac{\partial^2 s(t, \boldsymbol{r})}{\partial t^2} \tag{2.1}$$

where ∇^2 is the Laplacian operator [64, p. 493], $s(t, \boldsymbol{r})$ is the intensity of the disturbance (signal) observed at time t and in the location specified by the vector $\boldsymbol{r}$, and c is the wave propagation velocity in the medium.

In the case of a spherically symmetric sinusoidal wave with frequency f_0 propagating from the origin in a homogeneous medium without any reflectors or boundaries, the solution of the wave equation is given by [33, p. 152]

$$s(t, r) = \frac{a_0}{r} e^{j(2\pi f_0 t - \kappa r)} \tag{2.2}$$

where $j = \sqrt{-1}$, a_0 is the amplitude of the wave at unit radius, r is the distance from the origin, and κ is the wavenumber. The distance travelled by the wave during one period $T_0 = 1/f_0$ is denoted by the wavelength λ, where

$$\lambda = cT_0 = \frac{c}{f_0} = \frac{2\pi}{\kappa}. \tag{2.3}$$

On the other hand, when the propagating wave contains multiple frequencies, the solution of the wave equation can be expressed as [33]

$$s(t, r) = \int_{-\infty}^{\infty} \frac{A_0(f)}{r} e^{j(2\pi f t - \kappa r)} df. \tag{2.4}$$

where $A_0(f)$ is the Fourier transform of the disturbance signal at unit radius.

2.2.2   Sensor arrays

Sensor arrays are composed of multiple sensors that are located at different points in space. They can be used to filter the received signals by exploiting their spatial characteristics. This is achieved by linearly combining the outputs of the sensors using complex-valued weights so that the desired signal arriving from a particular

direction, or a set of directions, can be enhanced by constructive combination while attenuating the remaining signals by destructive combination.

Let us consider an array of M sensors. The mth sensor location is defined by the position vector

$$\boldsymbol{r}_m = [\mathrm{x}_m, \mathrm{y}_m, \mathrm{z}_m]^T \tag{2.5}$$

where $(\cdot)^T$ denotes the vector transpose. The sensors spatially sample the propagating wave at locations $\{\boldsymbol{r}_m\}_{m=1}^{M}$. The M–dimensional vector containing the output signal of the array sensors is given by

$$\boldsymbol{x}(t) = [x_1(t, \boldsymbol{r}_1), x_2(t, \boldsymbol{r}_2), \ldots, x_M(t, \boldsymbol{r}_M)]^T \tag{2.6}$$

where $x_m(t, \boldsymbol{r}_m)$ is the output of the mth sensor. If there are $J+1$ emitting sources, then we can write the output of the mth sensor as

$$x_m(t, \boldsymbol{r}_m) = \sum_{i=1}^{J+1} \mathit{\text{\ss}}_m(\boldsymbol{\Theta}_i) s_i(t, \boldsymbol{r}_m) + n_m(t) \tag{2.7}$$

where $s_i(t, \boldsymbol{r}_m)$ is the amplitude of the ith wave at the location $\boldsymbol{r}_m$, $\text{\ss}_m(\boldsymbol{\Theta}_i)$ is the response of the mth sensor towards the signal received from the ith source, $\boldsymbol{\Theta}_i$ is a parameter vector of the ith wave which can include its frequency, DOA, polarization, etc., and $n_m(t)$ is an additive noise term that represents the background and/or thermal noise associated with the mth sensor.

Usually, we assume that the sensors are identical and omnidirectional and that their response is constant over the frequency band of the received signal [82, p. 11]. Thus, $\text{\ss}_m(\boldsymbol{\Theta}_i)$ are constant and they can be omitted from (2.7). Therefore, we can write the output signal of the mth sensor as

$$x_m(t, \boldsymbol{r}_m) = \sum_{i=1}^{J+1} s_i(t, \boldsymbol{r}_m) + n_m(t). \tag{2.8}$$

16

2.2.3   Plane waves

A plane wave is characterized by the condition that the disturbance is constant in magnitude and phase on planes perpendicular to the direction of propagation. If the array sensors are placed at sufficiently large distance from the emitting source, then the wavefront of the propagating wave becomes flat enough that a plane wave approximation becomes locally valid. This *far-field* plane wave approximation is accurate when $r_0 \gg L_A^2/\lambda$ [101, p. 229], where r_0 is the distance between the source and the array center, and L_A is the array aperture, i.e., the maximum distance between any two sensors. Using this plane wave approximation, the solution to the wave equation in (2.1) at time t and location r for a sinusoidal wave with frequency f_0 is given by [99, p. 31]

$$s(t, r) = ae^{j\left(2\pi f_0 t - \kappa^T r\right)} \tag{2.9}$$

where a is the amplitude of the wave and $\kappa = [\kappa_x, \kappa_y, \kappa_z]^T$ is the wavenumber vector that determines the direction of propagation of the plane wave and whose magnitude is equal to the wavenumber. Therefore, we can write the wavenumber vector as

$$\kappa = \frac{2\pi}{\lambda} u \tag{2.10}$$

where u is a unit vector in the direction of propagation of the wave.

Assume that a plane wave impinges on an M-sensor array. Let $s(t)$ denote the signal that would be received by a virtual noiseless sensor placed at the origin. The output signal of the mth sensor can then be written as

$$x_m(t, r_m) = s(t - \tau_m) + n_m(t) \tag{2.11}$$

where

$$\tau_m = \frac{u^T r_m}{c} \tag{2.12}$$

is the time required for the plane wave to travel from the origin of the coordinate system to the location of the mth sensor.

17
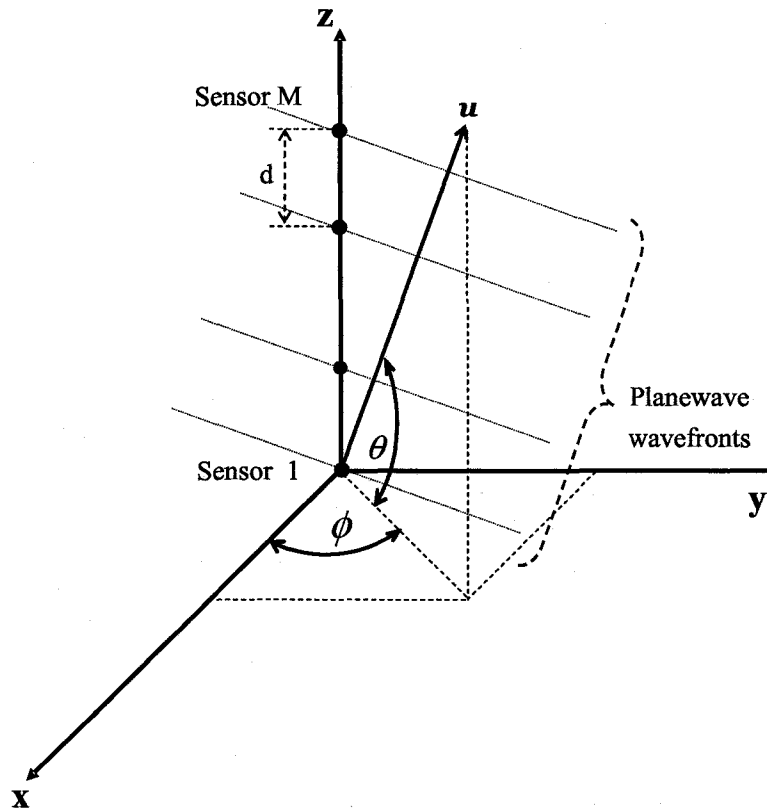
Figure 2.1: Plane wave incident on a ULA.

To illustrate the above model, let us consider the uniform linear array (ULA) shown in Figure 2.1. The sensors are placed along the z-axis at locations $z_m = (m-1)d$, where d is the array inter-element spacing. Thus, the location of the mth sensor is given by the vector

$$r_m = [0, 0, (m-1)d]^T. \qquad (2.13)$$

Assume there is only one plane wave arriving from the direction specified by the vector $u = [\cos(\phi)\cos(\theta), \sin(\phi)\cos(\theta), \sin(\theta)]$, where ϕ and θ are the azimuth and elevation angles, respectively. Therefore, the propagation delay incurred by the wave

18

to travel from the first sensor to the mth sensor is given by

$$\tau_m = \frac{[\cos(\phi)\cos(\theta), \sin(\phi)\cos(\theta), \sin(\theta)] [0, 0, (m-1)d]^T}{c}$$

$$= \frac{(m-1)d\sin(\theta)}{c},\qquad (2.14)$$

and the output of the mth sensor is

$$x_m(t) = s\left(t - \frac{(m-1)d\sin(\theta)}{c}\right) + n_m(t). \qquad (2.15)$$

2.2.4 Narrowband signal model

In many practical array processing applications, a typical assumption is that the information bearing signal is narrowband, i.e., the signal bandwidth is much smaller than the carrier frequency. Let us consider a single carrier-modulated plane wave impinging on an M-sensor array. We can write the noise-free complex bandpass signal observed at the mth sensor as

$$\bar{s}_m(t) = \bar{s}(t - \tau_m)$$

$$= s(t - \tau_m)e^{j2\pi f_0(t-\tau_m)} \qquad (2.16)$$

where $\bar{s}(t)$ is the complex bandpass signal received at the origin of the coordinate system, $s(t)$ is its informative (baseband) complex envelope, and f_0 is the carrier frequency. The Fourier transform of $\bar{s}_m(t)$ is given by

$$\bar{S}_m(f) = S(f - f_0)e^{-j2\pi f \tau_m} \qquad (2.17)$$

where $S(f)$ is the Fourier transform of $s(t)$.

Following the reception of the spatially-propagating signal by the mth sensor, the signal is down-converted to baseband before sampling [65]. We can write the Fourier transform of the complex-valued down-converted signal $s_m(t) = \bar{s}_m(t)e^{-j2\pi f_0 t}$ as

$$S_m(f) = S(f)e^{-j2\pi(f+f_0)\tau_m}. \qquad (2.18)$$

19

Using the narrowband assumption $B \ll f_0$ where B is the signal bandwidth, we can approximate $S_m(f)$ as

$$S_m(f) \approx S(f)e^{-j2\pi f_0 \tau_m}. \tag{2.19}$$

Therefore, taking the inverse Fourier transform of (2.19) we can write

$$s_m(t) \approx s(t)e^{-j2\pi f_0 \tau_m}. \tag{2.20}$$

Comparing (2.16) and (2.20) we can notice that the narrowband assumption implies that the baseband signal $s(t)$ remains almost constant during the time required by the propagating wave to travel across the array aperture, i.e., $B\tau_m \ll 1$ and $s(t) \approx s(t - \tau_m)$.

Therefore, using (2.20) we can write the vector containing the output of the M sensors as

$$
\begin{aligned}
\boldsymbol{x}(t) &= s(t)\begin{bmatrix} e^{-j2\pi f_0 \tau_1} \\ e^{-j2\pi f_0 \tau_2} \\ \vdots \\ e^{-j2\pi f_0 \tau_M} \end{bmatrix} + \boldsymbol{n}(t) \\
&= s(t)\boldsymbol{a}(f_0, \boldsymbol{\tau}) + \boldsymbol{n}(t)
\end{aligned} \tag{2.21}
$$

where $\boldsymbol{n}(t) = [n_1(t), n_2(t), \ldots, n_M(t)]^T$ is the vector of complex baseband noise, and $\boldsymbol{a}(f_0, \boldsymbol{\tau}) = \left[e^{-j2\pi f_0 \tau_1}, e^{-j2\pi f_0 \tau_2}, \ldots, e^{-j2\pi f_0 \tau_M}\right]^T$ is the steering vector (also called the array manifold vector) associated with the frequency f_0 and the delay vector $\boldsymbol{\tau} = [\tau_1, \tau_2, \ldots, \tau_M]^T$.

Let us consider the ULA described in Section 2.2.3. If there are $J+1$ narrowband waves impinging on the array from the directions specified by the elevation angles $\{\theta_i\}_{i=1}^{J+1}$, then using the principle of superposition we can write the array output vector as

$$\boldsymbol{x}(t) = \sum_{i=1}^{J+1} s_i(t)\boldsymbol{a}(\theta_i) + \boldsymbol{n}(t) \tag{2.22}$$

20

where $\boldsymbol{a}(\theta_i) = \left[1, e^{-j2\pi f_0 \frac{d}{c} \sin(\theta_i)}, e^{-j2\pi f_0 \frac{2d}{c} \sin(\theta_i)}, \ldots, e^{-j2\pi f_0 \frac{(M-1)d}{c} \sin(\theta_i)}\right]^T$ is the array steering vector corresponding to the ith wave where we have dropped its dependence on the frequency due to the narrowband approximation used.

We can also write (2.22) in matrix form as [65]

$$\boldsymbol{x}(t) = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{s}(t) + \boldsymbol{n}(t) \tag{2.23}$$

where

$$\boldsymbol{s}(t) = [s_1(t), s_2(t), \ldots, s_{J+1}(t)]^T \tag{2.24}$$

is the vector containing the waveforms of the $J + 1$ baseband signals,

$$\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots \theta_{J+1}]^T \tag{2.25}$$

is the vector containing the DOAs of the $J + 1$ signals, and

$$\boldsymbol{A}(\boldsymbol{\theta}) = [\boldsymbol{a}(\theta_1), \boldsymbol{a}(\theta_2), \ldots, \boldsymbol{a}(\theta_{J+1})] \tag{2.26}$$

is the $M \times L$ array response matrix which is usually full rank given that the steering vectors correspond to different DOAs [77].

2.3   Optimum Adaptive Beamforming

The main objective of adaptive beamforming is to extract the information content of the signal-of-interest in the presence of (strong) interferers and background noise [79]. This information content can be a message contained in the signal, such as in wireless communication applications [94], or some informative parameters of the signal such as in radar and sonar [78]. Adaptive beamforming, in its most general form, is done by combining the weighted outputs of different sensors. The weights are selected adaptively according to the received sensor data such that the desired signal is enhanced by constructive combination and the interference signals are suppressed

21

by destructive suppression. In this way, the beamformer acts as a spatial filter and can be thought of as casting a beam in the direction of the desired signal with its nulls in the directions of the interferers [112]. In this section we will review some classical adaptive beamforming approaches for narrowband and wideband signals.

2.3.1 Narrowband beamforming

Figure 2.2 shows a narrowband beamformer. The signal received by the ith sensor is sampled with the sampling frequency $f_s = 1/T_s$ where T_s is the sampling interval. We denote the kth sample of the received signal vector by $x(k)$. The signal received by the ith sensor is weighted by w_i^* where $(\cdot)^*$ denotes the complex conjugate operator. The output of the beamformer at the kth time instant is given by

$$y(k) = w^H x(k) \tag{2.27}$$

where $w = [w_1, w_2, \ldots, w_M]^T$ is the M–dimensional vector containing the weights of the beamformer and $(\cdot)^H$ denotes the Hermitian transpose.

The observation vector $x(k)$ typically contains one desired signal, J interference signals, and white noise. Therefore, we can write

$$x(k) = s_0(k) a(\theta_s) + \sum_{i=1}^{J} s_i(k) a(\theta_i) + n(k) \tag{2.28}$$

where θ_s is the DOA of the desired signal, θ_i is the DOA of the ith interference signal, $s_0(k)$ is the desired signal waveform, and $s_i(k)$ is the ith interference waveform. The desired and the interference signals are assumed to have zero mean and to be stationary and mutually independent [7], [112].

The total output power of the beamformer is given by $\mathrm{E}\left\{\left|w^H x(k)\right|^2\right\} = w^H R_x w$, where $\mathrm{E}\left\{\cdot\right\}$ denotes the statistical expectation and $R_x = \mathrm{E}\left\{x(k) x^H(k)\right\}$ is the covariance matrix of the received signal vector.

22
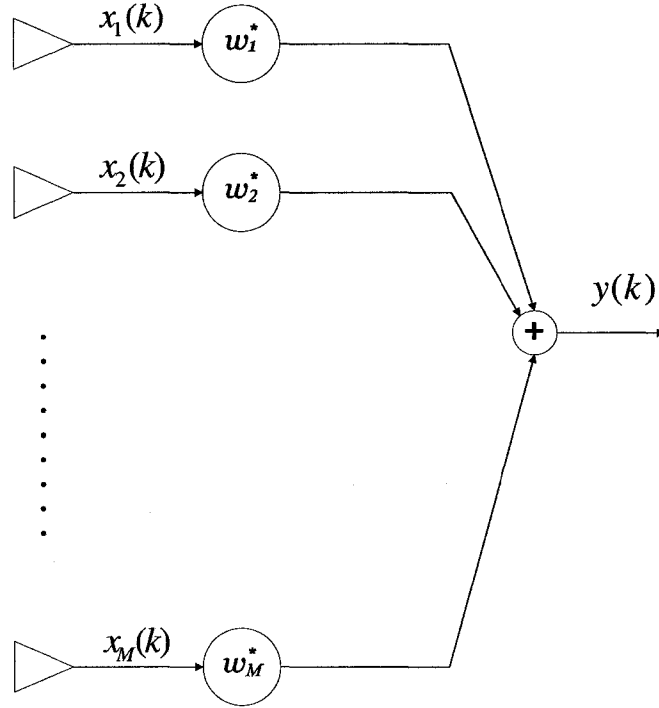
Figure 2.2: Narrowband beamformer.

The output SINR of the beamformer is defined as the ratio between the output desired signal power and the output power due to interference-plus-noise. It is given by

$$
\begin{aligned}
\text{SINR} &= \frac{\text{E}\left\{\left|s_0(k)\boldsymbol{w}^H\boldsymbol{a}(\theta_s)\right|^2\right\}}{\text{E}\left\{\left|\sum_{i=1}^{J} s_i(k)\boldsymbol{w}^H\boldsymbol{a}(\theta_i) + \boldsymbol{w}^H\boldsymbol{n}(k)\right|^2\right\}} \\
&= \frac{\sigma_s^2 \left|\boldsymbol{w}^H\boldsymbol{a}(\theta_s)\right|^2}{\boldsymbol{w}^H\boldsymbol{R}_{i+n}\boldsymbol{w}}
\end{aligned}
\tag{2.29}
$$

where σ_s^2 is the received power of the desired signal at each sensor, and

$$
\boldsymbol{R}_{i+n} = \text{E}\left\{\left(\sum_{i=1}^{J} s_i(k)\boldsymbol{a}(\theta_i) + \boldsymbol{n}(k)\right)\left(\sum_{i=1}^{J} s_i(k)\boldsymbol{a}(\theta_i) + \boldsymbol{n}(k)\right)^H\right\}
\tag{2.30}
$$

is the covariance matrix of the interference-plus-noise.

23

Several criteria exist for selecting the beamformer optimal weight vector [79]. One of the well-known adaptive beamforming algorithms for narrowband signals is the MVDR algorithm [7], [23], [40], [79]. This algorithm minimizes the output power of the beamformer due to interference-plus-noise subject to a constraint that provides a distortionless response towards the desired signal. The problem can be formulated as [116]

$$\min_{\boldsymbol{w}} \ \boldsymbol{w}^H \boldsymbol{R}_{i+n} \boldsymbol{w} \qquad \text{s.t.} \quad \boldsymbol{w}^H \boldsymbol{a}(\theta_s) = 1. \tag{2.31}$$

The solution to the above optimization problem can be simply evaluated using the method of Lagrange multipliers as [40]

$$\boldsymbol{w} = \frac{\boldsymbol{R}_{i+n}^{-1} \boldsymbol{a}(\theta_s)}{\boldsymbol{a}^H(\theta_s) \boldsymbol{R}_{i+n}^{-1} \boldsymbol{a}(\theta_s)}. \tag{2.32}$$

An equivalent formulation for the MVDR beamformer can be derived by minimizing the total output power of the array subject to the same distortionless response constraint in (2.31). The problem can be written as

$$\min_{\boldsymbol{w}} \ \boldsymbol{w}^H \boldsymbol{R}_x \boldsymbol{w} \qquad \text{s.t.} \quad \boldsymbol{w}^H \boldsymbol{a}(\theta_s) = 1. \tag{2.33}$$

Minimizing the total output power of the beamformer while preserving the desired signal is equivalent to minimizing the output power due to interference-plus-noise. This can be seen by writing the cost function of (2.33) as

$$\begin{aligned}
\boldsymbol{w}^H \boldsymbol{R}_x \boldsymbol{w} &= \sigma_s^2 \boldsymbol{w}^H \boldsymbol{a}(\theta_s) \boldsymbol{a}^H(\theta_s) \boldsymbol{w} + \boldsymbol{w}^H \boldsymbol{R}_{i+n} \boldsymbol{w} \\
&= \sigma_s^2 + \boldsymbol{w}^H \boldsymbol{R}_{i+n} \boldsymbol{w}
\end{aligned} \tag{2.34}$$

and thus the two optimization problems in (2.31) and (2.33) are equivalent and have the same solution in (2.32).

In practical applications, neither the interference-plus-noise covariance matrix nor the data covariance matrix is known at the receiver. Therefore, the sample covariance

matrix

$$\hat{\boldsymbol{R}}_x = \frac{1}{N_s} \sum_{k=1}^{N_s} \boldsymbol{x}(k)\boldsymbol{x}^H(k) \tag{2.35}$$

is used instead of the true covariance matrix, where N_s is the number of available training snapshots [40]. The resulting beamformer is known as the sample matrix inversion (SMI) beamformer [41]. It is well-known that in the absence of the desired signal in the beamformer training data, the SMI beamformer converges rapidly to the optimal SINR of the MVDR beamformer which is given by

$$\text{SINR}_{\text{opt}} = \sigma_s^2 \boldsymbol{a}^H(\theta_s) \boldsymbol{R}_{i+n}^{-1} \boldsymbol{a}(\theta_s) \tag{2.36}$$

so that if the number of training snapshots $N_s \geq 2M$, then the average loss relative to (2.36) is less than 3 dB [23]. However, if the desired signal is present in the training data, the performance of the SMI beamformer can degrade severely [41].

Moreover, in practical scenarios, the actual desired signal steering vector may not be precisely known. In such cases, an estimate $\hat{\boldsymbol{a}}(\theta_0)$ is used instead of the actual steering vector $\boldsymbol{a}(\theta_s)$, where θ_0 is the presumed DOA of the desired signal. This will be further discussed in Section 2.4.

The beamformer response (also termed the beam pattern [109, p. 33]) is defined as the magnitude and phase response to a plane wave as a function of its DOA. We can write the beamformer response to a narrowband plane wave arriving from the elevation angle θ as

$$H(\theta) = \boldsymbol{w}^H \boldsymbol{a}(\theta). \tag{2.37}$$

The magnitude response (in decibels) of a beamformer,

$$G(\theta) = 20 \log_{10}\left(|H(\theta)|\right) \tag{2.38}$$

is usually plotted against the DOA to demonstrate the spatial selectivity of the beamformer.

25

2.3.2   Wideband beamforming

If the desired signal bandwidth is greater than 1% of its center frequency, then the dependence of the array manifold vector on frequency can not be ignored. Thus, the approximations in (2.19) and (2.20) are no longer valid and the array signal model in (2.11) has to be used. Moreover, the beam pattern of the narrowband beamformer in Figure 2.2 changes with frequency and, therefore, more sophisticated beamforming schemes are required to produce frequency-invariant beam patterns [98]. In this section, we will review two of the most common adaptive beamforming approaches for wideband signals.
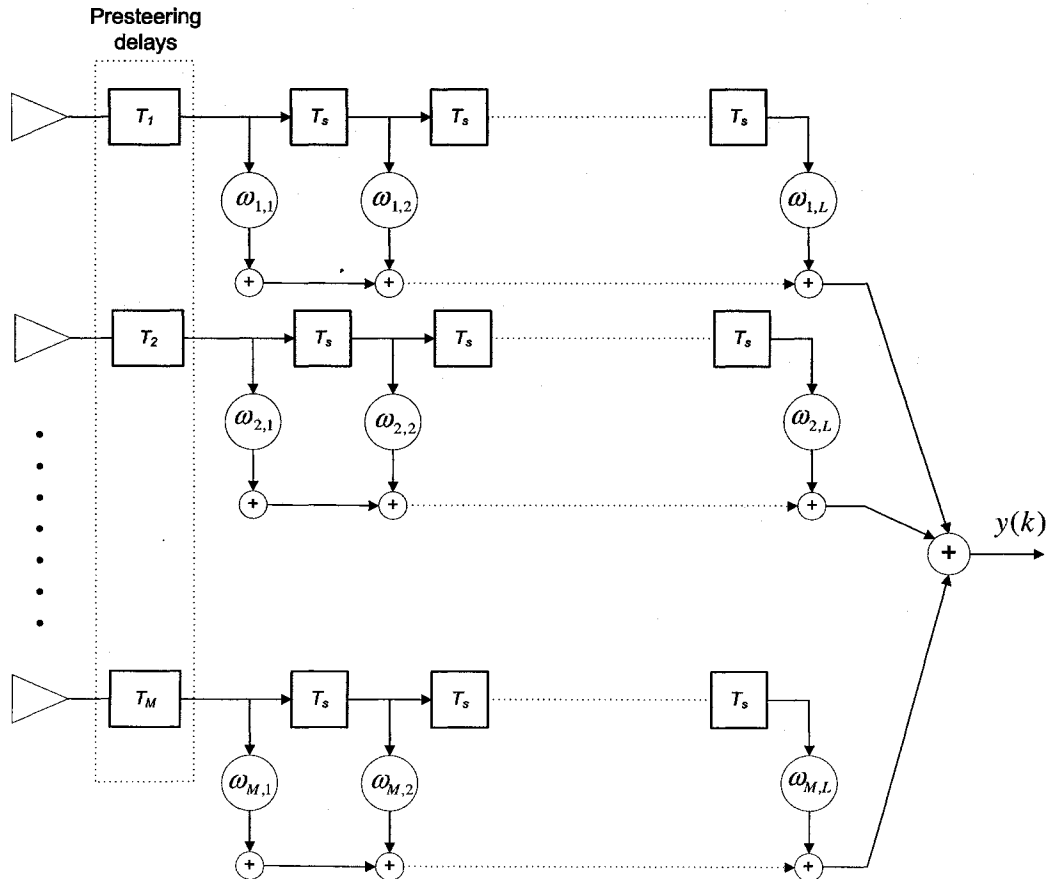
Figure 2.3: Presteered broadband array processor.

26

2.3.2.1   Time domain beamformers

Figure 2.3 shows a time domain presteered wideband array processor, commonly referred to as Frost's beamformer [84]. The front-end of the processor consists of wideband presteering delay filters with the delays $\{T_i\}_{i=1}^M$ whose function is to steer the array towards the desired signal direction so that all its frequency components appear in-phase across the array after the presteering filters. The output signal of the presteering delays is sampled at the rate $f_s = 1/T_s$ samples/second, where f_s is selected to be higher than or equal to $2f_u$ to avoid aliasing in the frequency domain and f_u is the maximum frequency of the desired signal [37, p. 14]. Each presteering delay filter is followed by an FIR filter of length L taps, and the output of the beamformer is produced by adding the outputs of the M FIR filters.

We will denote the M–dimensional data vector at the output of the presteering delays at the kth time instant as $\boldsymbol{\chi}(k) = [\chi_1(k), \chi_2(k), \ldots, \chi_M(k)]^T$. The stacked data vector defined as $\boldsymbol{x}(k) = \left[\boldsymbol{\chi}^T(k), \boldsymbol{\chi}^T(k-1), \ldots, \boldsymbol{\chi}^T(k-L+1)\right]^T$ contains L delayed presteered data vectors. We assume that processing is performed in the passband (without down-conversion to baseband), and therefore, the array data are real-valued [12], [52]. The kth sample of the beamformer output $y(k)$ is given by the inner product of the stacked data vector $\boldsymbol{x}(k)$ and the real valued ML–dimensional stacked weight vector $\boldsymbol{w}$ with $(\boldsymbol{w})_{M(l-1)+m} = \omega_{m,l}$ where $\omega_{m,l}$ denotes the weight at the lth tap of the mth FIR filter, i.e.,

$$y(k) = \boldsymbol{w}^T \boldsymbol{x}(k) \tag{2.39}$$

where $\boldsymbol{w}^T = \left[\boldsymbol{\omega}_1^T, \boldsymbol{\omega}_2^T, \ldots, \boldsymbol{\omega}_L^T\right]$. The individual weight vectors $\boldsymbol{\omega}_l$ are termed the temporal weight sets as they multiply time delayed versions of the presteered data vector. Their scalar components are given by $\boldsymbol{\omega}_l = [\omega_{1,l}, \omega_{2,l}, \ldots, \omega_{M,l}]^T$.

Let us consider a linear array whose sensors are placed along the z-axis at locations $\{z_i\}_{i=1}^M$. The number of sensors is chosen to meet the Nyquist criterion for spatial

sampling, i.e., the sensors are placed densely enough so that the received signal is spatially sampled at a minimum of $2L_A f_u/c$ positions. For a uniformly-spaced linear array, this corresponds to the condition that the sensors are placed at a distance less than or equal to $c/(2f_u)$ [11]. The response of the array to a complex sinusoid with the frequency f and the arrival angle θ is given by

$$H(f,\theta) = \boldsymbol{w}^T\big(\boldsymbol{d}(f) \otimes (\boldsymbol{T}(f)\boldsymbol{a}(f,\theta))\big) \tag{2.40}$$

where $\otimes$ denotes the Kronecker product,

$$\boldsymbol{d}(f) = \big[1, e^{-j2\pi f T_s}, \dots, e^{-j2\pi f(L-1)T_s}\big]^T \tag{2.41}$$

$$\boldsymbol{T}(f) = \mathrm{diag}\big\{e^{-j2\pi f T_1}, \dots, e^{-j2\pi f T_M}\big\} \tag{2.42}$$

$$\boldsymbol{a}(f,\theta) = \big[e^{j2\pi f\tau_1(\theta)}, \dots, e^{j2\pi f\tau_M(\theta)}\big]^T \tag{2.43}$$

$$\tau_i(\theta) = \frac{z_i \sin(\theta)}{c}. \tag{2.44}$$

Note that the array response is a 2–dimensional function of the frequency and DOA. We can define the power response (in decibels) of the beamformer as a function of the arrival angle only as

$$P(\theta) = 10\log_{10}\left(\int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \rho_\theta(f)\left|\boldsymbol{w}^T\big(\boldsymbol{d}(f) \otimes (\boldsymbol{T}(f)\boldsymbol{a}(f,\theta))\big)\right|^2 df\right) \tag{2.45}$$

where $\rho_\theta(f)$ is the power spectral density of the wideband signal arriving from the direction θ.

The received signal vector consists of the desired signal, J interference signals, and white noise. Therefore, we can write

$$\boldsymbol{x}(k) = \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} A_{\theta_s}(f)\big(\boldsymbol{d}(f) \otimes (\boldsymbol{T}(f)\boldsymbol{a}(f,\theta_s))\big) df$$

$$+ \sum_{i=1}^{J}\int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} A_{\theta_i}(f)\big(\boldsymbol{d}(f) \otimes (\boldsymbol{T}(f)\boldsymbol{a}(f,\theta_i))\big) df + \boldsymbol{n}(k) \tag{2.46}$$

28

where $A_\theta(f)$ is the Fourier transform of the signal arriving from the direction θ. Following the definition of the output SINR in Section 2.3.1, we can write the output SINR of the wideband beamformer as

$$\text{SINR} = \frac{w^T R_s w}{w^T R_{i+n} w} \tag{2.47}$$

where R_s is the covariance matrix of the desired signal which is given by

$$R_s = \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \rho_{\theta_s}(f) \big(d(f) \otimes (T(f)a(f,\theta_s))\big) \big(d(f) \otimes (T(f)a(f,\theta_s))\big)^H df. \tag{2.48}$$

Using the assumption that the desired signal, interference signals, and noise are mutually independent, we can write the interference-plus-noise covariance matrix as

$$R_{i+n} = \sum_{i=1}^{J} \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \rho_{\theta_i}(f) \big(d(f) \otimes (T(f)a(f,\theta_i))\big) \big(d(f) \otimes (T(f)a(f,\theta_i))\big)^H df + \sigma_n^2 I_{ML} \tag{2.49}$$

where σ_n^2 is noise power at each sensor, and the data covariance matrix is given by

$$R_x = R_s + R_{i+n}. \tag{2.50}$$

The function of the presteering delays is to enable the beamformer to identify the desired signal. They are selected so that the signal arriving from the presumed look direction produces identical (phase-aligned) signals at the output of the M presteering delay filters, i.e.,

$$T(f)a(f,\theta_0) = 1_M \qquad \forall f \in [f_l, f_u] \tag{2.51}$$

where f_l is the lowest frequency of the desired signal.

Therefore, the array response towards the look direction signal becomes

$$H(f,\theta_0) = w^T\big(d(f) \otimes 1_M\big) = w^T C_0 d(f) \tag{2.52}$$

where $C_0 \in \mathbb{R}^{ML \times L} = I_L \otimes 1_M$, I_L is the $L \times L$ identity matrix, and 1_M is the M–dimensional vector containing all ones.

29

The weights of the beamformer are selected so that the total output power of the beamformer is minimized while simultaneously maintaining a prescribed frequency response towards the desired signal. The problem can be written as

$$w = \arg\min_{w} w^T R_x w \qquad \text{s.t.} \qquad C_0^T w = h_0 \tag{2.53}$$

where the L–dimensional vector h_0 specifies the frequency response of the beamformer towards the look direction signal. For example, a distortionless (all-pass) response can be provided by selecting the vector h_0 as e_{L_c}, where e_i is a vector of appropriate dimension containing all zeros except for one in the ith position and L_c is the central tap index, i.e., $L_c = (L+1)/2$. The vector of the beamformer coefficients that satisfy (2.53) is given by

$$w = R_x^{-1} C_0 \left(C_0^T R_x^{-1} C_0 \right)^{-1} h_0. \tag{2.54}$$

This beamformer is widely known as the linearly constrained minimum variance (LCMV) beamformer with gain-only constraints and has been studied extensively in the literature [4], [12], [67], [84]. Note that the MVDR beamformer in (2.32) can be considered as a special case of the LCMV beamformer when the desired signal is narrowband, i.e., $L = 1$, $C_0 = a(\theta_s)$ and $h_0 = 1$.

An equivalent formulation of the LCMV problem was developed by Griffths and Jim [52], [90]. This formulation, which is known as the generalized sidelobe canceller (GSC) beamformer, replaces the constrained minimization problem in (2.53) with an unconstrained one, hence, providing more insight into the problem and simplifying its implementation [12]. Figure 2.4 shows a block diagram of the GSC beamformer. The GSC formulation decomposes the weight vector w into two orthogonal components that lie in the range and null space of the constraint matrix C_0, i.e,

$$w = w_0 - C_0^\perp w_a \tag{2.55}$$

where $w_a \in \mathbb{C}^{ML-L}$ is the adaptive part of w, and the columns of the full rank $ML \times (ML - L)$–dimensional matrix $C_0^\perp$ form a basis for the null space of C_0, i.e.,

30
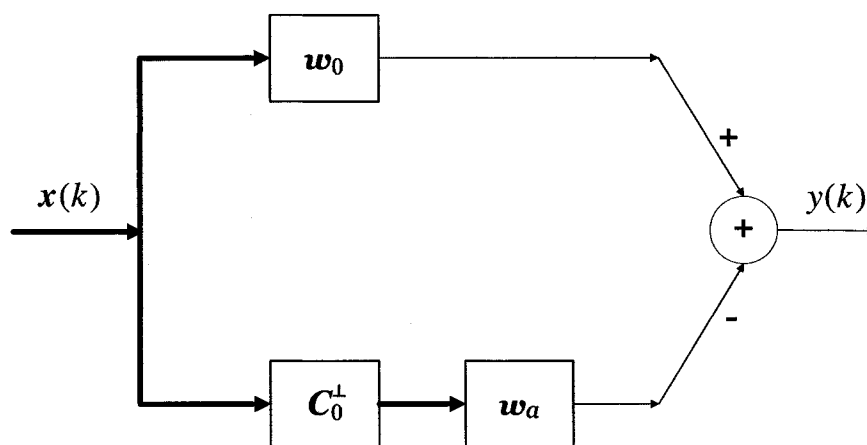
Figure 2.4: The GSC formulation of the LCMV algorithm.

$C_0^T C_0^\perp = 0_{L \times (ML-L)}$ where $0_{m \times n}$ is the $m \times n$–dimensional matrix containing zeros. In order for the weight vector w to satisfy the constraint in (2.53), we have

$$C_0^T w_0 = h_0 \tag{2.56}$$

and therefore the nonadaptive component w_0 of the weight vector w is given by

$$w_0 = C_0 \left(C_0^T C_0 \right)^{-1} h_0 \tag{2.57}$$

which obviously lies in the range space of the constraint matrix.

The use of the above weight vector decomposition eliminates the need for the constraint in (2.53). The unconstrained minimization problem can be written as

$$\min_{w_a} \left(w_0 - C_0^\perp w_a \right)^T R_x \left(w_0 - C_0^\perp w_a \right) \tag{2.58}$$

whose solution is given by [52]

$$w_a = \left(C_0^{\perp T} R_x C_0^\perp \right)^{-1} C_0^{\perp T} R_x w_0. \tag{2.59}$$

The advantage of the GSC formulation comes from the fact that the weight vector w_a is unconstrained which permits the use of very simple online algorithms to update

31
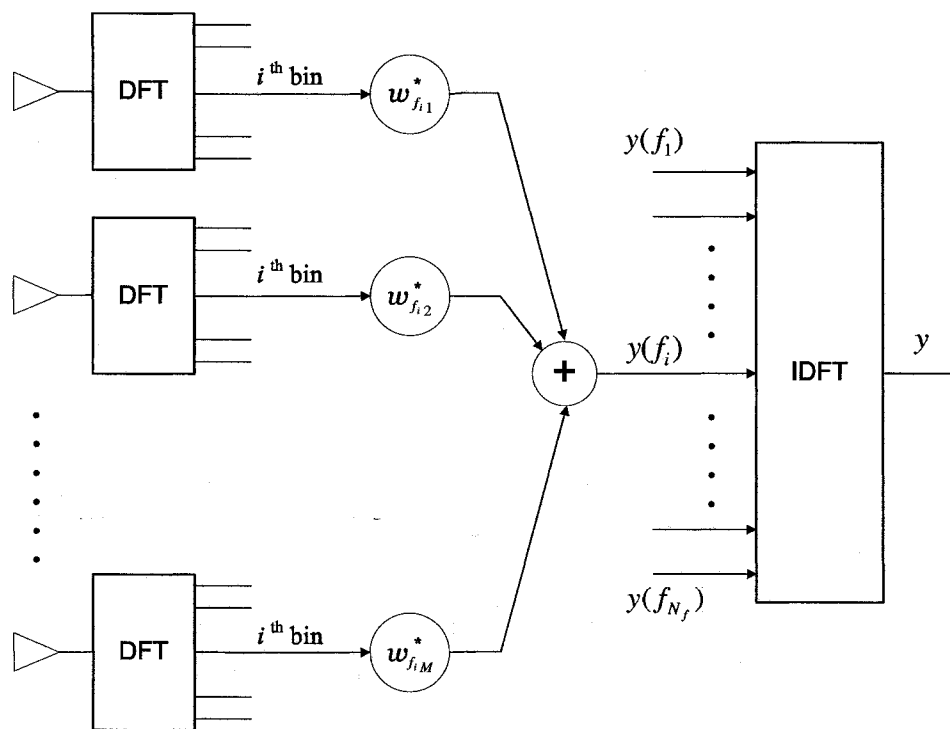
Figure 2.5: Frequency domain wideband beamformer.

the beamformer weight vector once any new data snapshot is received [12], [20], [119]. This will be further discussed in Section 2.5.

However, the LCMV wideband beamformer with gain-only constraints and its GSC formulation suffer from their sensitivity to array manifold mismatches. Even slight array calibration errors can lead to a substantial degradation of the output SINR of this beamformer [22], [40], [41]. Hence, robust adaptive beamforming techniques are needed that can maintain satisfactory performance even in the presence of array manifold mismatches.

2.3.2.2 Frequency domain beamformers

An alternative approach to the time domain-based spatial processing of wideband signals is to decompose the signal into multiple narrowband components through

32

filter banks or the discrete Fourier transform (DFT). This is commonly referred to as subband decomposition [26], [118]. Each subband signal can then be modelled as a narrowband signal and thus narrowband array processing techniques can be applied to each subband. This frequency domain processing approach is more suited to signals with very large bandwidths [48].

Figure 2.5 shows a block diagram of the frequency domain wideband beamformer. The data received by each of the M sensors is first transformed to the frequency domain through the DFT. The data across the array at each of the N_f subband frequencies are then processed by their own narrowband beamformer to produce the subband output $y(f_i) = \boldsymbol{w}_{f_i}^H \boldsymbol{x}(f_i)$. The total array output is then formed by taking the inverse DFT (IDFT) of the subband filtered data.

Note that there is an equivalence between the wideband time domain and frequency domain beamformers [112]. The outputs of the subband beamformers $\{y(f_i)\}_{i=1}^{N_f}$ can be made equivalent to the DFT of the output of the time domain beamformer in (2.39) with a proper selection of the beamformer weights and appropriate data partitioning [48].

2.4 Classical Robust Beamforming Techniques

The optimal adaptive beamforming techniques discussed in Section 2.3 are based on the assumption that the signal model and the array characteristics are precisely known. However, in practical situations many types of mismatches can exist between the presumed and actual signal environment and array characteristics. For instance, the actual steering vector of the desired signal might be different from the presumed one. This mismatch is commonly referred to as "array pointing error" and it can happen due to look direction errors or imperfect knowledge of the sensor locations or radiation characteristics [62]. Another form of model mismatch is the wavefront

33

distortions that occur due to propagation in inhomogeneous media or signal multipath propagation and scattering [87]. Adaptive beamforming techniques are known to be sensitive to such mismatches as the beamformer, in fact, interprets the desired signal as interference which results in its suppression [63]. Moreover, even if the signal environment and array characteristics are perfectly known at the receiver, the performance of adaptive beamforming algorithms can still degrade severely due to the finite sample support [36].

Hence, a strong need arises for adaptive beamforming techniques that can preserve the desired signal even in the presence of the aforementioned mismatches [41]. In this section, we will briefly review some classical algorithms that provide robustness against array calibration and manifold errors.

2.4.1   Directional derivative constraints

One of the very early approaches to allow for some margin of directional mismatch between the presumed and the actual desired signal is the use of directional derivative constraints. These constraints impose a maximally flat response up to some order N_d in the look direction. They are derived by forcing the directional derivatives of the power response of the array $(\rho(f, \theta) = H(f, \theta)H^*(f, \theta))$ up to the N_d-th order to be equal to zero in the look direction for all frequencies, that is,

$$\left. \frac{\partial^n}{\partial \theta^n} \rho(f, \theta) \right|_{\theta = \theta_0} = 0 \qquad \forall f \in \left[-\frac{f_s}{2}, \frac{f_s}{2} \right]; n = 1, \ldots, N_d. \tag{2.60}$$

In practice, only the first- and second-order directional derivative constraints are of practical use, as the use of additional higher-order constraints degrades the beamformer interference rejection capability. In [34], Er and Cantoni derived sufficient

34

conditions on the array weight vector for satisfying (2.60). For the linear array structure described before, the first- and second-order sufficient directional derivative constraints are given, respectively, by

$$\left(\boldsymbol{I}_L \otimes \left. \frac{\partial \boldsymbol{\tau}^T(\theta)}{\partial \theta} \right|_{\theta=\theta_0} \right) \boldsymbol{w} \;=\; \boldsymbol{0}_{L \times 1} \tag{2.61}$$

$$\left(\boldsymbol{I}_L \otimes \left. \frac{\partial^2 \boldsymbol{\tau}^T(\theta)}{\partial \theta^2} \right|_{\theta=\theta_0} \right) \boldsymbol{w} \;=\; \boldsymbol{0}_{L \times 1} \tag{2.62}$$

where $\boldsymbol{\tau}(\theta) = [\tau_1(\theta), \tau_2(\theta), \ldots, \tau_M(\theta)]^T$.

In [105], necessary and sufficient (NS) first- and second-order directional derivative constraints were derived. They differ from the sufficient constraints in (2.61) and (2.62) in that they do not constrain the phase response of the array and are therefore insensitive to the array phase-center location [12]. The set of NS first-order derivative constraints is given by

$$\boldsymbol{h}_0^T \left((\boldsymbol{J}_l - \boldsymbol{J}_l^T) \otimes \boldsymbol{1}_M^T \left. \frac{\partial \boldsymbol{\Lambda}_\tau(\theta)}{\partial \theta} \right|_{\theta=\theta_0} \right) \boldsymbol{w} = 0 \qquad l = 1, \ldots, L-1 \tag{2.63}$$

where $\boldsymbol{J}_l$ is the $L \times L$ matrix containing all zeros except for a diagonal of ones located on the lth diagonal above the main diagonal and $\boldsymbol{\Lambda}_\tau(\theta) = \mathrm{diag}\,\{\tau_1(\theta), \ldots, \tau_M(\theta)\}$.

We can see from (2.61), (2.62), and (2.63) that the sets of sufficient first- and second-order derivative constraints and NS first-order derivative constraints are linear and can therefore be easily implemented within the LCMV algorithm. On the other hand, the NS second-order derivative constraints are nonlinear and nonconvex which poses many problems in solving for the beamformer optimal weight vector [105], [106].

2.4.2   Presteering delays derivative constraints

The use of directional derivative constraints provides robustness against look direction errors, yet it does not provide enough robustness against other types of array manifold

mismatches. In [124], a set of constraints referred to as presteering delays derivative constraints was proposed. These constraints can preserve the desired signal despite multiple mismatches in the array manifold. They were derived by forcing a maximally-flat power response in the presteered direction, i.e., all the partial derivatives of the power response of the array with respect to $\{T_i\}_{i=1}^M$ up to the N_pth order are equal to zero at the selected presteering delay values $\left\{\hat{T}_i\right\}_{i=1}^M$ for all frequencies. This constraint can be written as

$$\left.\frac{\partial^n \rho(f,\theta)}{\partial T_1^{n_1} \partial T_2^{n_2} \dots \partial T_M^{n_M}}\right|_{T_i=\hat{T}_i} = 0 \quad \forall f \in \left[-\frac{f_s}{2}, \frac{f_s}{2}\right]; n = 1, \dots, N_p; n_m = 0, \dots, n \left| \sum_{m=1}^M n_m = n. \right.$$

(2.64)

For the case of first-order presteering delays derivative constraints, only the first-order derivatives of the power response are of concern, namely,

$$\left.\frac{\partial \rho(f,\theta)}{\partial T_i}\right|_{T_i=\hat{T}_i} = 0 \quad \forall f \in \left[-\frac{f_s}{2}, \frac{f_s}{2}\right]; i = 1, \dots, M. \tag{2.65}$$

NS conditions for satisfying (2.65) were derived in [124] and simplified into a set of linear constraints on the weight vector $\boldsymbol{w}$. These constraints are given by

$$\left(\left(\boldsymbol{h}_0^T (\boldsymbol{J}_l - \boldsymbol{J}_l^T)\right) \otimes \left(\mathbf{1}_M^T \boldsymbol{\Omega}_m\right)\right)\boldsymbol{w} = 0 \quad \forall m = 1, \dots, M; l = 1, \dots, L-1 \tag{2.66}$$

where $\boldsymbol{\Omega}_m$ is the $M \times M$ diagonal matrix containing all zeros except for a single unity-valued element located at the mth row and mth column.

For the case of an all-pass response, i.e., $\boldsymbol{h}_0 = \boldsymbol{e}_{L_c}$, the set of NS first-order presteering (NS1-PS) derivative constraints reduces to the following set [124]:

$$\left(\left(\boldsymbol{e}_{L_c}^T (\boldsymbol{J}_l - \boldsymbol{J}_l^T)\right) \otimes \left(\mathbf{1}_M^T \boldsymbol{\Omega}_m\right)\right)\boldsymbol{w} = 0 \quad \forall m = 1, \dots, M; l = 1, \dots, L_c-1. \tag{2.67}$$

Note that these constraints are equivalent to linear phase constraints for each of the M FIR filters of the array processor. This can be easily seen by rewriting (2.67) for

36

all $m = 1, \ldots, M$ and $l = 1, \ldots, L_c - 1$ as:

$$
\begin{aligned}
\left(\left(e_{L_c}^T J_l \right) \otimes \left(1_M^T \Omega_m \right) \right) w &= \left(\left(e_{L_c}^T J_l^T \right) \otimes \left(1_M^T \Omega_m \right) \right) w \\
\left(e_{L_c+l}^T \otimes e_m^T \right) w &= \left(e_{L_c-l}^T \otimes e_m^T \right) w \\
e_{M(L_c+l-1)+m}^T w &= e_{M(L_c-l-1)+m}^T w \\
\omega_{m,L_c+l} &= \omega_{m,L_c-l}
\end{aligned}
\tag{2.68}
$$

which is the Type 1 linear phase constraint on each of the M FIR filters of the array processor [86].

2.4.3   Diagonal loading

One of the popular approaches to add robustness to the LCMV beamformer against various types of array manifold mismatches is the diagonal loading technique [1], [2], [14], [38]. Its key idea is to impose an inequality constraint on the weight vector norm [23]. We can write the diagonally loaded LCMV problem as

$$
\begin{aligned}
\min_{w} \quad & w^T \hat{R}_x w \\
\text{s.t.} \quad & C_0^T w = h_0 \\
& w^T w \leq \gamma_{\mathrm{dl}}
\end{aligned}
\tag{2.69}
$$

where we have used the sample covariance matrix instead of the true covariance matrix. The solution to the above problem has the same form as the solution to the LCMV problem in (2.54), but with a diagonal loading term added to the covariance matrix, that is,

$$
w = \left(\hat{R}_x + \eta_{\mathrm{dl}} I \right)^{-1} C_0 \left(C_0^T \left(\hat{R}_x + \eta_{\mathrm{dl}} I \right)^{-1} C_0 \right)^{-1} h_0.
\tag{2.70}
$$

where the diagonal loading factor η_{dl} is selected iteratively so that the weight vector norm constraint is satisfied [23].

We can see from (2.70) that diagonal loading is equivalent to injecting white noise with the power η_{dl} into the data covariance matrix, or, equivalently, adding an omnidirectional interferer [15]. It is also clear that diagonal loading guarantees the invertibility of the sample covariance matrix even if the number of training snapshots is smaller than the dimension of the covariance matrix, and improves the robustness of the beamformer against finite sample size. Moreover, it is well-known that diagonal loading can improve the performance of the beamformer in scenarios with arbitrary mismatches in the array manifold while maintaining a low sidelobe level which makes the beamformer robust against unexpected interferers [23].

However, the main disadvantage of the diagonal loading approach is that the amount of diagonal loading can not be directly related to the amount of mismatches in the array manifold or even to the weight vector norm constraint in (2.69). It is often set in an ad-hoc way and typically chosen as $10 \, \sigma_n^2$ [23].

2.4.4   Eigenspace beamformer

One of the powerful approaches to robust adaptive beamforming for the case of narrowband signals is the eigenspace beamforming technique [18], [36]. This approach uses a more accurate estimate of the desired signal steering vector than the classical presumed vector $\hat{a}(\theta_0)$ used in the SMI beamformer. This is achieved by projecting the presumed desired signal steering vector onto the signal-plus-interference subspace evaluated from the sample covariance matrix through eigen decomposition. Hence, we can write the sample covariance matrix as

$$\hat{R}_x = E_s \Lambda_s E_s^H + E_n \Lambda_n E_n^H \tag{2.71}$$

where E_s and E_n are the $M \times (J+1)$ and $M \times (M-J-1)$-dimensional matrices containing the eigenvectors of the signal-plus-interference and noise subspaces, respectively, and Λ_s and Λ_n are the $(J+1) \times (J+1)$ and $(M-J-1) \times (M-J-1)$ diagonal

matrices containing the corresponding eigenvalues of the signal-plus-interference and noise subspaces, respectively.

The weight vector of the eigenspace beamformer can be written as the solution to the following problem

$$\min_{\boldsymbol{w}} \ \boldsymbol{w}^H \hat{\boldsymbol{R}}_x \boldsymbol{w} \qquad \text{s.t.} \quad \boldsymbol{w}^H \boldsymbol{E}_s \boldsymbol{E}_s^H \hat{\boldsymbol{a}}(\theta_0) = 1 \qquad (2.72)$$

where $\boldsymbol{E}_s \boldsymbol{E}_s^H$ is the projection matrix on the signal-plus-interference subspace. The solution to this problem is given by [18]

$$\boldsymbol{w} = \frac{\boldsymbol{E}_s \boldsymbol{\Lambda}_s^{-1} \boldsymbol{E}_s^H \hat{\boldsymbol{a}}(\theta_0)}{\hat{\boldsymbol{a}}^H(\theta_0) \boldsymbol{E}_s \boldsymbol{\Lambda}_s^{-1} \boldsymbol{E}_s^H \hat{\boldsymbol{a}}(\theta_0)}. \qquad (2.73)$$

However, the main disadvantage of this technique is that it is not applicable in low signal-to-noise ratio (SNR) scenarios where there is a high probability of subspace swapping which can lead to a poor estimate of the desired signal steering vector [107]. Moreover, this technique is limited to the narrowband case where the dimension of the signal-plus-interference subspace is low and known and can not be directly applied to the case of large number of interferers or to the wideband case.

2.5 Online Implementation of Adaptive Beamformers

As described earlier, the calculation of the optimum weight vector of various adaptive beamformers requires knowledge of the second-order statistics of the received data. If we assume that the underlying signal processes are stationary and ergodic, then these statistics can be calculated from the received data through time-averaging. However, in many practical scenarios the received data are nonstationary, i.e., their statistics may change over time. This might occur due to the nonstationary behaviour of the

propagation channel, or due to the motion of the desired signal or interference which is typical in radar, sonar, and wireless communications [53], [94].

Two main approaches exist to deal with this problem. The first approach is based on repeatedly estimating the required statistics from temporal blocks of array data and subsequently using these statistics to recalculate the optimum weight vector. This approach is suitable for slowly varying environments only. On the other hand, when the environment is rapidly changing or the dimension of the weight vector is moderate to large, then the weight vector has to be continuously updated *online* every time a new data sample is received [112].

We consider the more general case of complex-valued array data. We can write the cost function of the GSC optimization problem in (2.58) as

$$J(\boldsymbol{w}_a) = \left(\boldsymbol{w}_0 - \boldsymbol{C}_0^{\perp}\boldsymbol{w}_a\right)^H \mathrm{E}\left\{\boldsymbol{x}(k)\boldsymbol{x}^H(k)\right\}\left(\boldsymbol{w}_0 - \boldsymbol{C}_0^{\perp}\boldsymbol{w}_a\right) \tag{2.74}$$

$$= \mathrm{E}\left\{\left|\boldsymbol{w}_0^H\boldsymbol{x}(k) - \boldsymbol{w}_a^H\boldsymbol{C}_0^{\perp H}\boldsymbol{x}(k)\right|^2\right\} \tag{2.75}$$

$$= \mathrm{E}\left\{\left|y_d(k) - \boldsymbol{w}_a^H\boldsymbol{x}_C(k)\right|^2\right\} \tag{2.76}$$

where $y_d(k) = \boldsymbol{w}_0^H\boldsymbol{x}(k)$ is the desired output or reference signal and $\boldsymbol{x}_C(k) = \boldsymbol{C}_0^{\perp H}\boldsymbol{x}(k)$ is the input data vector of the adaptive filter. The objective of the online algorithm is to minimize the mean square difference between the desired output and the weighted sum of the input vector. Note that the cost function $J(\boldsymbol{w}_a)$ is a bowl-shaped surface whose dimension is equal to the dimension of the weight vector $\boldsymbol{w}_a$. This surface has a unique minimum which is the optimal solution of the adaptive filtering problem [54].

We will briefly review two categories of online algorithms which are based on the GSC formulation. Alternate online algorithms for the LCMV problem can be found in [12], [90], [92], and [119].

40

2.5.1   LMS-based beamformers

The least mean squares (LMS) algorithm is one of the most widely used adaptive filtering algorithms due to its simplicity. First, the error signal $e(k)$ is computed as the difference between the output of the adaptive filter and the desired output, i.e., $e(k) = y_d(k) - w_a^H(k)x_C(k)$ where $w_a(k)$ is the weight vector estimate at the kth time instant. Based on this measure, the adaptive filter will change its weight vector coefficients attempting to reduce the error. This is done by moving the estimate in the negative direction of the *instantaneous* gradient of the error surface. The weight vector update relation is a function of the squared error signal and is given by

$$
\begin{aligned}
w_a(k+1) &= w_a(k) + \mu\left(-\nabla|e(k)|^2\right) \\
&= w_a(k) + \mu e^*(k)x_C(k)
\end{aligned}
\tag{2.77}
$$

where $\nabla|e(k)|^2$ is the gradient of the squared error with respect to the weight vector $w_a(k)$. The gradient is a vector pointing in the direction of change in filter coefficients that will cause the greatest increase in the error signal. Because the goal is to minimize the error, the filter weight vector coefficients are updated in the direction opposite to the gradient, and hence the negation of the gradient term. The constant μ is the step-size, which determines the amount of gradient information used to update the weight vector and controls the convergence speed of the algorithm. After repeatedly adjusting the weight vector coefficients in the direction opposite to the gradient of the error, the adaptive filter should converge.

The main virtue of the LMS algorithm is its computational simplicity. However, its main disadvantage is that its convergence speed depends on the shape of the error surface and the eigenvalue distribution of the input data covariance matrix, and can be very slow if the covariance matrix is ill-conditioned [54].

41

2.5.2 RLS-based beamformers

An alternate approach to the LMS technique is the exponentially weighted recursive least squares (RLS) algorithm. At the kth time step, the algorithm chooses the weight vector estimate so that the weighted sum of past squared errors is minimized. The RLS optimization problem can be written as

$$\min_{\boldsymbol{w}_a(k)} \sum_{i=0}^{k} \lambda_{\text{RLS}}^{k-i} \left| y_d(i) - \boldsymbol{w}_a^H(k)\boldsymbol{x}_C(i) \right|^2 \tag{2.78}$$

where λ_{RLS} is the RLS forgetting factor which is positive, slightly less than 1, and determines how fast previous data are deemphasized [54].

One of the key features of the RLS algorithm is that its convergence is independent of the condition number of the data covariance matrix and is typically much faster than the LMS algorithm [54], [112]. However, its computational complexity is $\mathcal{O}\left(M_a^2\right)$ where M_a is the dimension of the weight vector $\boldsymbol{w}_a(k)$. This computational complexity is much higher than that of the LMS algorithm which is $\mathcal{O}\left(M_a\right)$ [112].

2.6 Conclusion

In this chapter, we have presented a brief overview of classical adaptive beamforming techniques. We have started by describing the array models for wideband and narrowband signals. Optimum adaptive beamforming techniques that assume perfect knowledge of the array and signal environment have been reviewed. We have also discussed some classical beamforming approaches that can provide robustness against various mismatches in the array manifold and highlighted the problems with each of these approaches. The next three chapters will focus on the development of computationally efficient algorithms for robust adaptive beamforming that can remedy the shortcomings of the existing approaches.

Chapter 3

Narrowband Robust Adaptive Beamforming

3.1 Introduction

The need for robust adaptive beamforming arises in many practical applications where the desired signal is present in the beamformer training data and where the assumptions on the nature of the desired signal and/or interference are violated. This need may also arise as a consequence of finite sample support, the inherent non-stationary nature of the underlying environmental processes, array manifold errors, etc. [1], [23], [40], [41], [43].

One of the main problems that occur in practical adaptive array systems is the mismatch between the presumed desired signal steering vector and the actual one [62], [116]. Adaptive array techniques are known to be very sensitive to even slight mismatches of such type that may occur due to look direction errors, imperfect array calibration, source local scattering, wavefront distortions, etc. [41].

In Section 2.4, we have revised several classical approaches to provide robustness against steering vector mismatches. One of these techniques is the directional

43

derivative constrained LCMV beamformer which provides robustness against look direction errors [79], [84]. However, this technique is not sufficiently robust against other types of mismatches [41]. We have also revised several ad-hoc approaches to overcome arbitrary desired signal steering vector mismatches, such as the diagonal loading technique [1], [2], [14], [23] and the eigenspace-based beamformer [18], [36].

We start this chapter by revising one of the recent theoretically rigourous and powerful approaches to robust beamforming in the presence of an arbitrary unknown desired signal steering vector mismatch [116]. This approach is based on worst-case performance optimization. It obtains the weight vector by minimizing the output interference-plus-noise power while maintaining a distortionless response for the worst-case (mismatched) signal steering vector. The robust MVDR problem is formulated in [116] as a second-order cone programming (SOCP) problem that can be solved in polynomial time using interior point methods. It is also shown in [116] that the robust MVDR beamformer can be interpreted as a diagonally loaded beamformer with an *adaptive* diagonal loading factor that is optimally matched to the amount of uncertainty in the desired signal steering vector. This work has been extended in [70], [74], [115] and [123], in particular, alternative Newton-type iterative algorithms have been developed in [70], [74], and [123] for this beamformer and its modifications. In [115], an SOCP-based generalization of the robust MVDR beamformer of [116] has been proposed which additionally accounts for nonstationary environments. However, the algorithms presented in [70], [74], [115], [116], and [123] do not have direct online implementations and their robustness against interference nonstationarity and abrupt environmental changes is insufficient.

In this chapter, we develop a Kalman filter-based technique for implementing the robust MVDR beamformer of [116], where the robustness constraint is incorporated in the measurement equation of the state-space model. The proposed beamformer can be updated online and requires only $\mathcal{O}\left(M^2\right)$ computations per iteration compared

to $\mathcal{O}\left(M^3\right)$ computations required for the SOCP-based algorithm of [116] and for the algorithms proposed in [70], [74], and [123]. Next, we study the convergence properties of the proposed beamformer in a stationary environment and prove that it has a convergence rate close to that of the SMI beamformer. Furthermore, we propose two modifications of our Kalman filter-based beamformer which additionally account for abruptly changing nonstationary environments. The first modification is based on the detection of environmental changes through hypothesis testing, whereas the second is based on hypotheses merging using the interacting multiple model (IMM) estimation technique [8]. At the end of the chapter, we provide simulation results comparing our Kalman filter-based implementation of the robust MVDR beamformer to its original SOCP-based implementation of [116].

3.2 Worst-Case Performance Optimization-based Beamforming

Let us consider a linear M-sensor array. In Section 2.3.1, we have presented the narrowband signal model. According to (2.28), the observation vector is given by

$$x(k) = s_0(k)a(\theta_s) + \sum_{i=1}^{J} s_i(k)a(\theta_i) + n(k). \tag{3.1}$$

The well-known MVDR beamformer discussed in Section 2.3.1 minimizes the output interference-plus-noise power while maintaining a distortionless response to the desired signal [79]. The MVDR problem is given by

$$\min_{w} \; w^H \hat{R}_x w \qquad \text{s.t.} \quad w^H \hat{a} = 1 \tag{3.2}$$

where $\hat{a}$ is the presumed desired signal steering vector. Note that the presumed steering vector may be an erroneous (mismatched) copy of the actual steering vector $a(\theta_s)$, yet the norm of the steering vector distortion can be usually bounded by some

45

known constant $\varepsilon > 0$ [116]. Therefore, the actual desired signal steering vector belongs to the set

$$\mathcal{A}(\varepsilon) \triangleq \left\{ \tilde{a} \,|\, \tilde{a} = \hat{a} + e, \|e\| \leq \varepsilon \right\}. \tag{3.3}$$

The robust MVDR beamforming problem in [116] minimizes the total output power of the beamformer while requiring a distortionless response not only for the presumed steering vector $\hat{a}$, but for all the vectors that belong to $\mathcal{A}(\varepsilon)$. Thus, the weight vector of the robust MVDR beamformer can be obtained as the solution to the following problem [116]

$$\min_{\boldsymbol{w}} \boldsymbol{w}^H \hat{\boldsymbol{R}}_x \boldsymbol{w} \qquad \text{s.t.} \quad |\boldsymbol{w}^H \tilde{a}| \geq 1 \qquad \forall \tilde{a} \in \mathcal{A}(\varepsilon). \tag{3.4}$$

Since there is an infinite number of vectors $\tilde{a} \in \mathcal{A}(\varepsilon)$, the constraint set in (3.4) represents an infinite number of nonconvex constraints. These constraints can be converted into a single constraint that corresponds to the worst-case (mismatched) vector $\tilde{a}$ over the set $\mathcal{A}(\varepsilon)$. Thus, we can write (3.4) as

$$\min_{\boldsymbol{w}} \boldsymbol{w}^H \hat{\boldsymbol{R}}_x \boldsymbol{w} \qquad \text{s.t.} \quad \min_{\tilde{a} \in \mathcal{A}(\varepsilon)} |\boldsymbol{w}^H \tilde{a}| \geq 1. \tag{3.5}$$

Applying the triangle and Cauchy-Shwartz inequalities along with the inequality $\|e\| \leq \varepsilon$, we can write

$$|\boldsymbol{w}^H \tilde{a}| = |\boldsymbol{w}^H \hat{a} + \boldsymbol{w}^H e| \geq |\boldsymbol{w}^H \hat{a}| - |\boldsymbol{w}^H e| \geq |\boldsymbol{w}^H \hat{a}| - \varepsilon \|\boldsymbol{w}\|. \tag{3.6}$$

The worst-case error vector that satisfies (3.6) with equality is given by

$$e = -\varepsilon \frac{\boldsymbol{w}}{\|\boldsymbol{w}\|} e^{j\phi_e} \tag{3.7}$$

where

$$\phi_e = \text{angle} \left\{ \boldsymbol{w}^H \hat{a} \right\}. \tag{3.8}$$

Thus, using (3.6) we can write

$$\min_{\tilde{a} \in \mathcal{A}(\varepsilon)} |\boldsymbol{w}^H \tilde{a}| = |\boldsymbol{w}^H \hat{a}| - \varepsilon \|\boldsymbol{w}\| \tag{3.9}$$

46

and therefore the problem can be written as

$$\min_{w} w^H \hat{R}_x w \quad \text{s.t.} \quad |w^H \hat{a}| - \varepsilon \|w\| \geq 1. \tag{3.10}$$

The problem is still nonconvex due to the absolute value operator in the constraint. Yet, we can notice that the cost function in (3.10) does not change when w undergoes an arbitrary phase rotation. Thus, we can always rotate w such that $w^H \hat{a}$ is real-valued, without affecting the value of the cost function. Therefore, without any loss of optimality, the problem in (3.10) can be written as [116]

$$\begin{aligned}
\min_{w} \quad & w^H \hat{R}_x w \\
\text{s.t.} \quad & w^H \hat{a} - \varepsilon \|w\| \geq 1 \\
& \text{Im}\{w^H \hat{a}\} = 0
\end{aligned} \tag{3.11}$$

which is a convex optimization problem.

This problem was solved in [116] using SOCP. It was also proved in [116] that the robustness constraint is satisfied with equality by the optimal solution. This can be seen by noticing that if the constraint was not active at the optimal point, i.e., if $w^H \hat{a} - \varepsilon \|w\| = \beta > 1$, then we could further decrease the cost function simply by dividing the weight vector by β, while still satisfying the constraint. Thus, we can replace the inequality constraint in (3.12) with an equality constraint without any loss of optimality. Based on this fact, different Newton-type iterative procedures have been proposed in [70], [74], and [123] to solve (3.12) and some of its extensions.

We further note that the second constraint in (3.11) is redundant as it is implied by the first constraint $w^H \hat{a} = \varepsilon \|w\| + 1$ where $\varepsilon \|w\|$ is pure real. Therefore we can write (3.11) as

$$\min_{w} w^H \hat{R}_x w \quad \text{s.t.} \quad w^H \hat{a} - \varepsilon \|w\| = 1. \tag{3.12}$$

3.3   Kalman Filter-based Robust Beamformer

The use of the constrained Kalman filter to solve the linearly constrained MVDR beamforming problem was suggested in [20]. However, perfect knowledge of the array manifold was assumed throughout [20] and the issue of robustness against mismatches in the desired signal steering vector was not considered. In this section, we derive our Kalman filter-based implementation for the robust MVDR beamformer of [116].

3.3.1   Derivation

For the sake of convenience of subsequent derivations, let us introduce the mean square error (MSE) between the zero signal and the beamformer output as

$$\text{MSE} = \text{E}\left\{\left|0 - \boldsymbol{x}^H(k)\boldsymbol{w}(k)\right|^2\right\} = \boldsymbol{w}^H \boldsymbol{R}_x \boldsymbol{w}. \tag{3.13}$$

Therefore, minimizing the beamformer output power is equivalent to minimizing the MSE in (3.13). Also, the constraint in (3.12) can be expressed as

$$h(\boldsymbol{w}(k)) = 1 \tag{3.14}$$

where

$$h(\boldsymbol{w}(k)) = \varepsilon^2 \boldsymbol{w}^H(k)\boldsymbol{w}(k) - \boldsymbol{w}^H(k)\hat{\boldsymbol{a}}\hat{\boldsymbol{a}}^H\boldsymbol{w}(k) + \boldsymbol{w}^H(k)\hat{\boldsymbol{a}} + \hat{\boldsymbol{a}}^H\boldsymbol{w}(k). \tag{3.15}$$

The robust beamforming problem can be written as

$$\min_{\boldsymbol{w}(k)} \text{MSE} \qquad \text{s.t.} \quad h(\boldsymbol{w}(k)) = 1. \tag{3.16}$$

The Kalman filter is a minimum mean square error (MMSE) estimator [6, p. 207] and, hence, it can be used to solve (3.16). An unknown dynamic system can be modelled as a filter whose state vector $\boldsymbol{w}(k)$ undergoes a first-order Markov process [54], i.e.,

$$\boldsymbol{w}(k+1) = \gamma\boldsymbol{w}(k) + \boldsymbol{n}_w(k) \tag{3.17}$$

48

where γ is a fixed parameter of the model, and $\boldsymbol{n}_w(k)$ is the process noise vector which is assumed to be white Gaussian with zero mean and covariance matrix $\boldsymbol{Q}_w = \sigma_w^2 \boldsymbol{I}$. Thus, the process equation of the optimal weight vector $\boldsymbol{w}$ is given by (3.17), whereas the measurement equation is given by

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}^H(k)\boldsymbol{w}(k) \\ h(\boldsymbol{w}(k)) \end{bmatrix} + \begin{bmatrix} n_{m,0}(k) \\ n_{m,1}(k) \end{bmatrix} \tag{3.18}$$

which can be written in matrix notation as

$$\boldsymbol{z} = \boldsymbol{h}(\boldsymbol{w}(k)) + \boldsymbol{n}_m(k) \tag{3.19}$$

where $n_{m,0}(k)$ and $n_{m,1}(k)$ are the residual and constraint errors, respectively. They can be modelled as independent zero mean white sequences with covariance matrix

$$\boldsymbol{R} = \begin{bmatrix} \sigma_{m,0}^2 & 0 \\ 0 & \sigma_{m,1}^2 \end{bmatrix} . \tag{3.20}$$

The use of the MMSE estimator to estimate the state vector $\boldsymbol{w}(k)$ will yield a solution that minimizes the uncertainties due to the state and measurement noises. Minimizing the mean square value of $n_{m,0}(k)$ will lead to minimizing the output power of the beamformer, while minimizing the mean square value of $n_{m,1}(k)$ will minimize the MSE in satisfying the robustness constraint and, hence, the resulting estimate will solve (3.12).

Due to the nonlinearity of the measurement equation in (3.19), we will use the second-order EKF to find a recursion for the weight vector estimate $\hat{\boldsymbol{w}}(k)$ [6, p. 381]. We start by evaluating the Jacobian $\boldsymbol{H}_w(k, \boldsymbol{w}(k))$ of $\boldsymbol{h}(\boldsymbol{w}(k))$, and the Hessian

49

matrices $H^{(0)}_{ww}$ and $H^{(1)}_{ww}$ of the two components of $h(w(k))$

$$
\begin{aligned}
H_w(k, w(k)) &= \left(\nabla h^T(w(k))\right)^T \\
&= \begin{bmatrix} x^H(k) \\ \varepsilon^2 w^H(k) - (\hat{a}\hat{a}^H w(k))^H + \hat{a}^H \end{bmatrix}
\end{aligned}
\tag{3.21}
$$

$$
H^{(0)}_{ww} = \nabla\nabla^H\left\{x^H(k)w(k)\right\} = 0
\tag{3.22}
$$

$$
H^{(1)}_{ww} = \nabla\nabla^H\left\{h(w(k))\right\} = \varepsilon^2 I - \hat{a}\hat{a}^H.
\tag{3.23}
$$

The recursion for the estimated weight vector starts with an initial weight vector estimate $\hat{w}(0)$ with the associated covariance matrix $P(0|0)$ and updates the weight vector estimate as

$$
\hat{w}(k) = \hat{w}(k-1) + G(k)\big(z - \hat{z}(k|k-1)\big)
\tag{3.24}
$$

where the predicted measurement $\hat{z}(k|k-1)$ and the filter gain $G(k)$ are given by

$$
\hat{z}(k|k-1) = \begin{bmatrix} \gamma x^H(k)\hat{w}(k-1) \\ h(\gamma\hat{w}(k-1)) + \frac{1}{2}\mathrm{tr}\left\{H^{(1)}_{ww}P(k|k-1)\right\} \end{bmatrix}
\tag{3.25}
$$

$$
G(k) = P(k|k-1)H^H_w(k, \gamma\hat{w}(k-1))S^{-1}(k)
\tag{3.26}
$$

respectively, where $\mathrm{tr}\{\cdot\}$ denotes the trace of a matrix. Here, the innovation covariance $S(k)$ and the predicted weight vector covariance $P(k|k-1)$ are given by

$$
\begin{aligned}
S(k) &= H_w(k, \gamma\hat{w}(k-1))P(k|k-1)H^H_w(k, \gamma\hat{w}(k-1)) \\
&\quad + \frac{1}{2}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\mathrm{tr}\left\{H^{(1)}_{ww}P(k|k-1)H^{(1)}_{ww}P(k|k-1)\right\} + R
\end{aligned}
\tag{3.27}
$$

$$
P(k|k-1) = \gamma^2 P(k-1|k-1) + Q_w.
\tag{3.28}
$$

The updated weight vector covariance can be expressed as

$$
\begin{aligned}
P(k|k) &= \big(I - G(k)H_w(k, \gamma\hat{w}(k-1))\big)P(k|k-1)\big(I - G(k)H_w(k, \gamma\hat{w}(k-1))\big)^H \\
&\quad + G(k)RG^H(k).
\end{aligned}
\tag{3.29}
$$

50

The parameters γ and σ_w^2 of the state equation should be chosen such that the model can track the optimal weight vector as the environment changes. Their effect can be seen from (3.28). From this equation, it is clear that when increasing these parameters, the Kalman filter tends to assign more weight to recent data enabling better tracking of the environment. For a nonstationary environment, a typical choice of γ is slightly greater than one. Although this choice makes the state equation unstable, the stability of the filter can be guaranteed from the observability condition [19, p. 192]. The value of σ_w^2 is chosen according to the degree of nonstationarity of the operating environment. For example, a choice of σ_w^2 equal to 10^{-4} indicates that each component of the optimal weight vector can change independently during one time step by the order of 10^{-2}. Note that even higher values of σ_w^2 can be chosen for more rapidly changing environments. On the other hand, for stationary environments the optimal weight vector does not change with time, and therefore, $\gamma = 1$ and $\sigma_w^2 = 0$.

From the first line of the measurement equation it follows that the parameter $\sigma_{m,0}^2$ should be chosen in the same order as the optimal output power of the beamformer. The latter power can be roughly approximated as $\|w\|^2(M\sigma_s^2 + \sigma_n^2)$ [79]. It is worth noting that our beamformer is not sensitive to the choice of $\sigma_{m,0}^2$ because the norm of the weight vector estimate is chosen by the filter so that the output power of the beamformer matches the value of $\sigma_{m,0}^2$. This will be illustrated by means of simulations. The value of the parameter $\sigma_{m,1}^2$ should be chosen very small, for example, $\sigma_{m,1}^2 = 10^{-12}$, so that the robustness constraint is satisfied with high accuracy without leading to any numerical errors [39].

The consistency of the beamformer can be checked online through the normalized innovation square (NIS) test [6, p. 236]. Under the Gaussian assumption for the state-space model, the NIS

$$\epsilon(k) = \big(z - \hat{z}(k|k-1)\big)^H S^{-1}(k)\big(z - \hat{z}(k|k-1)\big), \qquad (3.30)$$

51

is chi-square distributed. From (3.25), we can see that the first component of the innovation $z - \hat{z}(k|k-1)$ is generally complex-valued, whereas the second component is guaranteed to be real-valued because both $h(\gamma \hat{w}(k-1))$ and the trace of the Hermitian matrix $\boldsymbol{H}_{ww}^{(1)} \boldsymbol{P}(k|k-1)$ are real. Thus, the NIS is chi-square distributed with three degrees of freedom, and should be within acceptable limits with a certain probability if the beamformer is consistent. For example, using a 95% confidence region, the NIS should be less than 7.815 with probability 0.95.

An exponential window can be used to evaluate the NIS of the beamformer, i.e.,

$$\epsilon_\alpha(k) = \alpha\epsilon_\alpha(k-1) + \big(z - \hat{z}(k|k-1)\big)^H \boldsymbol{S}^{-1}(k)\big(z - \hat{z}(k|k-1)\big) \qquad (3.31)$$

where the effective window length is given by $\frac{1}{1-\alpha}$. Therefore, under the above mentioned Gaussian assumption, $\epsilon_\alpha(k)$ can be approximated using first-order moment matching as chi-square distributed with $\lceil \frac{3}{1-\alpha} \rceil$ degrees of freedom, where $\lceil x \rceil$ denotes the integer greater than or equal to x.

For initialization of the iterative algorithm in (3.24)–(3.29), a Gaussian random weight vector estimate $\hat{w}(0)$ can be used together with an initial covariance matrix estimate $\boldsymbol{P}(1|0) = \beta_0 \boldsymbol{I}$, where β_0 is selected so that the NIS of the first iteration is acceptable. Using (3.30) at $k = 1$, and approximating $\boldsymbol{S}(1)$ by ignoring the second-order term and the measurement noise covariance matrix, we can write

$$\epsilon(1) \approx \frac{1}{\beta_0} \big(z - \hat{z}(1|0)\big)^H \big(\boldsymbol{H}_w(1, \hat{w}(0)) \boldsymbol{H}_w^H(1, \hat{w}(0))\big)^{-1} \big(z - \hat{z}(1|0)\big). \qquad (3.32)$$

The value of β_0 can be calculated by choosing $\epsilon(1)$ such that it lies within acceptable limits, i.e., it lies with in the 90% double-sided acceptance region of the chi-square distribution with 3 degrees of freedom. For example $\epsilon(1)$ can be selected equal to 3 which is the expected value of the distribution.

3.3.2   Computationally efficient implementation

An important issue for the Kalman filter-based robust beamformer is its computational complexity. The evaluation of the Jacobian matrix in (3.21) has a complexity of $\mathcal{O}(M^2)$, whereas the Hessian matrix in (3.23) is constant for a fixed presumed steering vector $\hat{a}$. The weight vector update in (3.24) has a complexity of $\mathcal{O}(M)$, and the complexity of evaluating the predicted measurement in (3.25) is of $\mathcal{O}(M^2)$. Additionally, the computation of the filter gain in (3.26) has a complexity of $\mathcal{O}(M^2)$.

The computation of the innovation covariance matrix through (3.27) can be simplified by substituting with (3.23) and rewriting the second term as

$$
\begin{aligned}
\operatorname{tr} &\left\{ H_{ww}^{(1)} P(k|k-1) H_{ww}^{(1)} P(k|k-1) \right\} \\
&= \operatorname{tr}\left\{ \left(\varepsilon^2 P(k|k-1) - \hat{a}\hat{a}^H P(k|k-1) \right)\left(\varepsilon^2 P(k|k-1) - \hat{a}\hat{a}^H P(k|k-1) \right) \right\} \\
&= \varepsilon^4 \operatorname{tr}\left\{ P(k|k-1)^2 \right\} - 2\varepsilon^2 \operatorname{tr}\left\{ P(k|k-1)\hat{a}\hat{a}^H P(k|k-1) \right\} \\
&\quad + \operatorname{tr}\left\{ \hat{a}\hat{a}^H P(k|k-1)\hat{a}\hat{a}^H P(k|k-1) \right\} \\
&= \varepsilon^4 \| P(k|k-1) \|_F^2 - 2\varepsilon^2 \| P(k|k-1)\hat{a} \|^2 + \left(\hat{a}^H P(k|k-1)\hat{a} \right)^2
\end{aligned}
\tag{3.33}
$$

where $\| \cdot \|_F$ is the matrix Frobenius norm.

Therefore, the innovation covariance matrix update can be written as

$$
\begin{aligned}
S(k) = &\frac{1}{2}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\left(\varepsilon^4 \| P(k|k-1) \|_F^2 - 2\varepsilon^2 \| P(k|k-1)\hat{a} \|^2 + (\hat{a}^H P(k|k-1)\hat{a})^2 \right) \\
&+ H_w(k, \hat{w}(k-1)) P(k|k-1) H_w^H(k, \hat{w}(k-1)) + R
\end{aligned}
\tag{3.34}
$$

which requires only $\mathcal{O}(M^2)$ operations.

The computational bottleneck of the Kalman filter-based robust beamformer is equation (3.29) which requires $\mathcal{O}(M^3)$ multiplications. However, this equation can be replaced by [6, p. 206]

$$
P(k|k) = P(k|k-1) - G(k)S(k)G^H(k)
\tag{3.35}
$$

whose complexity is of $\mathcal{O}\left(M^2\right)$.

It is important to stress that although using the covariance update equation (3.35) requires less computational effort than (3.29), it was noted in [6, p. 206] that it is more prone to numerical errors (loss of positive semidefiniteness of the covariance matrix). However, this is not likely to happen for our state-space model in (3.17) and (3.19) due to the nonzero value of the diagonal entries of the process noise covariance matrix. We note that in our simulations there was no single run in which the use of (3.35) instead of (3.29) has caused any numerical problems.

In summary, the proposed Kalman filter-based implementation of the robust MVDR beamformer of [116] has a computational complexity of $\mathcal{O}\left(M^2\right)$ per iteration, whereas its original SOCP-based implementation in [116] has a complexity of $\mathcal{O}\left(M^3\right)$. Moreover, an important advantage of the proposed Kalman filter-based algorithm is that it can be easily implemented without any need of a specific built-in optimization software.

3.3.3 Convergence analysis

The convergence rate of the EKF-based robust beamformer can be analyzed in terms of its MSE. For a stationary environment, the MSE at the kth iteration is given by

$$
\begin{aligned}
\mathrm{MSE}(k) &= \mathrm{E}\left\{\left|0 - \boldsymbol{x}^H(k)\hat{\boldsymbol{w}}(k)\right|^2\right\} \\
&= \mathrm{E}\left\{\left|n_{m,0}(k) + \boldsymbol{x}^H(k)\left(\boldsymbol{w}(k) - \hat{\boldsymbol{w}}(k)\right)\right|^2\right\} \\
&= \sigma_{m,0}^2 + \mathrm{E}\left\{\boldsymbol{x}^H(k)\left(\boldsymbol{w}(k) - \hat{\boldsymbol{w}}(k)\right)\left(\boldsymbol{w}(k) - \hat{\boldsymbol{w}}(k)\right)^H \boldsymbol{x}(k)\right\} \\
&= \sigma_{m,0}^2 + \mathrm{tr}\left\{\mathrm{E}\{\left(\boldsymbol{w}(k) - \hat{\boldsymbol{w}}(k)\right)\left(\boldsymbol{w}(k) - \hat{\boldsymbol{w}}(k)\right)^H \boldsymbol{x}(k)\boldsymbol{x}^H(k)\}\right\} & (3.36) \\
&= \sigma_{m,0}^2 + \mathrm{tr}\left\{\mathrm{E}\{\left(\boldsymbol{w}(k) - \hat{\boldsymbol{w}}(k)\right)\left(\boldsymbol{w}(k) - \hat{\boldsymbol{w}}(k)\right)^H\}\mathrm{E}\{\boldsymbol{x}(k)\boldsymbol{x}^H(k)\}\right\} & (3.37) \\
&= \sigma_{m,0}^2 + \mathrm{tr}\left\{\boldsymbol{P}(k)\boldsymbol{R}_x\right\} & (3.38)
\end{aligned}
$$

where (3.37) follows from (3.36) due to the fact that the error in the weight vector estimate is independent of the input data [6, p. 123].

Next, we find an expression for the weight vector covariance matrix $\boldsymbol{P}(k)$. Using the recursion for the inverse covariance (information) [6, p. 303], we can write

$$
\begin{aligned}
\boldsymbol{P}^{-1}(k) &= \boldsymbol{P}^{-1}(k-1) + \boldsymbol{h}_w^H(k, \boldsymbol{w}(k))\boldsymbol{R}^{-1}\boldsymbol{h}_w(k, \boldsymbol{w}(k)) \\
&= \boldsymbol{P}^{-1}(k-1) + \frac{1}{\sigma_{m,0}^2}\boldsymbol{x}(k)\boldsymbol{x}^H(k) \\
&\quad + \frac{1}{\sigma_{m,1}^2}\big((\varepsilon^2\boldsymbol{I} - \hat{\boldsymbol{a}}\hat{\boldsymbol{a}}^H)\hat{\boldsymbol{w}}(k-1) + \hat{\boldsymbol{a}}\big)\big((\varepsilon^2\boldsymbol{I} - \hat{\boldsymbol{a}}\hat{\boldsymbol{a}}^H)\hat{\boldsymbol{w}}(k-1) + \hat{\boldsymbol{a}}\big)^H \\
&= \boldsymbol{P}^{-1}(0) + \frac{1}{\sigma_{m,0}^2}\sum_{i=1}^{k}\boldsymbol{x}(i)\boldsymbol{x}^H(i) \\
&\quad + \frac{1}{\sigma_{m,1}^2}\sum_{i=1}^{k}\big((\varepsilon^2\boldsymbol{I} - \hat{\boldsymbol{a}}\hat{\boldsymbol{a}}^H)\hat{\boldsymbol{w}}(i-1) + \hat{\boldsymbol{a}}\big)\big((\varepsilon^2\boldsymbol{I} - \hat{\boldsymbol{a}}\hat{\boldsymbol{a}}^H)\hat{\boldsymbol{w}}(i-1) + \hat{\boldsymbol{a}}\big)^H. \quad (3.39)
\end{aligned}
$$

In a stationary environment and for k sufficiently large, the initial inverse covariance $\boldsymbol{P}^{-1}(0)$ can be neglected and the weight vector estimate $\hat{\boldsymbol{w}}(k)$ can be replaced by the steady state weight vector estimate $\hat{\boldsymbol{w}}$. Thus, we can write

$$
\boldsymbol{P}^{-1}(k) = \frac{k}{\sigma_{m,0}^2}\hat{\boldsymbol{R}}_x(k) + \frac{k}{\sigma_{m,1}^2}\big((\varepsilon^2\boldsymbol{I} - \hat{\boldsymbol{a}}\hat{\boldsymbol{a}}^H)\hat{\boldsymbol{w}} + \hat{\boldsymbol{a}}\big)\big((\varepsilon^2\boldsymbol{I} - \hat{\boldsymbol{a}}\hat{\boldsymbol{a}}^H)\hat{\boldsymbol{w}} + \hat{\boldsymbol{a}}\big)^H \quad (3.40)
$$

where $\hat{\boldsymbol{R}}_x(k)$ is the k-snapshot sample covariance matrix given by

$$
\hat{\boldsymbol{R}}_x(k) = \frac{1}{k}\sum_{i=1}^{k}\boldsymbol{x}(i)\boldsymbol{x}^H(i). \quad (3.41)
$$

Using the matrix inversion lemma [51], we can write

$$
\boldsymbol{P}(k) = \frac{\sigma_{m,0}^2}{k}\hat{\boldsymbol{R}}_x^{-1}(k) - \frac{\sigma_{m,0}^2}{k\beta_1}\hat{\boldsymbol{R}}_x^{-1}(k)\big((\varepsilon^2\boldsymbol{I} - \hat{\boldsymbol{a}}\hat{\boldsymbol{a}}^H)\hat{\boldsymbol{w}} + \hat{\boldsymbol{a}}\big)\big((\varepsilon^2\boldsymbol{I} - \hat{\boldsymbol{a}}\hat{\boldsymbol{a}}^H)\hat{\boldsymbol{w}} + \hat{\boldsymbol{a}}\big)^H\hat{\boldsymbol{R}}_x^{-1}
$$

$$
(3.42)
$$

where the scalar β_1 is given by

$$
\begin{aligned}
\beta_1 &= \big((\varepsilon^2\boldsymbol{I} - \hat{\boldsymbol{a}}\hat{\boldsymbol{a}}^H)\hat{\boldsymbol{w}} + \hat{\boldsymbol{a}}\big)^H\hat{\boldsymbol{R}}_x^{-1}(k)\big((\varepsilon^2\boldsymbol{I} - \hat{\boldsymbol{a}}\hat{\boldsymbol{a}}^H)\hat{\boldsymbol{w}} + \hat{\boldsymbol{a}}\big) + \frac{\sigma_{m,1}^2}{\sigma_{m,0}^2} \\
&\approx \big((\varepsilon^2\boldsymbol{I} - \hat{\boldsymbol{a}}\hat{\boldsymbol{a}}^H)\hat{\boldsymbol{w}} + \hat{\boldsymbol{a}}\big)^H\hat{\boldsymbol{R}}_x^{-1}(k)\big((\varepsilon^2\boldsymbol{I} - \hat{\boldsymbol{a}}\hat{\boldsymbol{a}}^H)\hat{\boldsymbol{w}} + \hat{\boldsymbol{a}}\big), \quad (3.43)
\end{aligned}
$$

and the approximation in (3.43) is due to $\sigma_{m,1}^2 \ll \sigma_{m,0}^2$ which follows from the guidelines of choosing these parameters outlined in Subsection 3.3.1.

Substituting with (3.42) in (3.38), the MSE recursion can be approximated by

$$\text{MSE}(k) = \sigma_{m,0}^2 + \frac{\sigma_{m,0}^2}{k-1}\text{tr}\{\hat{\boldsymbol{R}}_x^{-1}(k)\boldsymbol{R}_x\} - \frac{\sigma_{m,0}^2}{(k-1)\beta_1}\text{tr}\left\{\hat{\boldsymbol{R}}_x^{-1}(k)\big((\varepsilon^2\boldsymbol{I} - \hat{\boldsymbol{a}}\hat{\boldsymbol{a}}^H)\hat{\boldsymbol{w}} + \hat{\boldsymbol{a}}\big)\right.$$
$$\left.\big((\varepsilon^2\boldsymbol{I} - \hat{\boldsymbol{a}}\hat{\boldsymbol{a}}^H)\hat{\boldsymbol{w}} + \hat{\boldsymbol{a}}\big)^H \hat{\boldsymbol{R}}_x^{-1}(k)\boldsymbol{R}_x\right\}. \tag{3.44}$$

For k sufficiently large, the sample covariance matrix provides a good approximation of the covariance matrix $\boldsymbol{R}_x$, thus, we can write,

$$\text{MSE}(k) = \sigma_{m,0}^2 + \frac{M\sigma_{m,0}^2}{k-1}$$
$$- \frac{\sigma_{m,0}^2}{(k-1)\beta_1}\left(\big((\varepsilon^2\boldsymbol{I} - \hat{\boldsymbol{a}}\hat{\boldsymbol{a}}^H)\hat{\boldsymbol{w}} + \hat{\boldsymbol{a}}\big)^H \hat{\boldsymbol{R}}_x^{-1}(k)\big((\varepsilon^2\boldsymbol{I} - \hat{\boldsymbol{a}}\hat{\boldsymbol{a}}^H)\hat{\boldsymbol{w}} + \hat{\boldsymbol{a}}\big)\right)$$
$$= \sigma_{m,0}^2\left(1 + \frac{M-1}{k-1}\right). \tag{3.45}$$

From the above equation we conclude that the steady state residual output power is given by $\sigma_{m,0}^2$, and that the beamformer converges to within 3 dB of its steady state output power in M iterations.

3.4   Multiple Model Beamformers

In this section, we introduce two beamformers that extend the robust beamformer design to abruptly changing nonstationary environments. Our first beamformer is based on hard decision detection of nonstationarities, while the other is based on hypotheses merging and IMM estimation techniques [6, p. 453]. We develop these techniques for two state-space models, namely, the stationary model with $\gamma = 1$ and $\sigma_w^2 = 0$ (that does not assume that the optimal weight vector changes in time and has the advantage of a lower misadjustment in the optimal weight vector estimate and, hence, a high steady state output SINR); and the nonstationary model, which has the advantage of a high capability of tracking changes in the operating environment.

56

3.4.1   Hard decision switching

If the statistical properties of the environment change abruptly, the difference between the measurement vector z and the expected measurement vector $\hat{z}(k|k-1)$ will be large. This will lead to a large value of the NIS of the stationary-modelled filter and this fact can be used to detect environmental changes. Therefore, the filter should be initialized at the beginning of operation in the stationary mode and the NIS should be monitored. If it exceeds a certain pre-specified threshold (for example, the upper 90% tail point of the chi-square distribution with the corresponding degrees of freedom), then nonstationarity is declared and the dynamic state-space model should be used to account for such nonstationarity. The effect of switching to the dynamic process equation leads to an increase in the uncertainty in the current weight vector estimate (reflected by the weight vector covariance matrix $P(k|k)$) due to multiplication by γ^2 and the addition of the process noise covariance matrix Q_w in (3.28). Hence, the beamformer will assign more weight to the latest measurements corresponding to the latest changes in the environment. This will allow the beamformer to efficiently track any environmental changes.

After switching to the nonstationary process model, the NIS should be monitored again, and when it drops below a certain pre-specified threshold (for example, the expected value of the chi-square distribution with the corresponding degrees of freedom), it is assumed that the beamformer has adapted to the new environment and the filter is switched back to the stationary model.

3.4.2   Interacting multiple model techniques

In multiple model environments, the optimal estimate of the state vector is known to be a function of the elemental state estimates obtained via estimators tuned to all possible parameter histories [17]. Thus, with time, an exponentially increasing

number of filters is required to keep track of all possible model histories which is computationally impractical. Many suboptimal techniques have been proposed [6], among which the IMM is the most cost-effective in implementation [8].

In this approach, the state estimate is computed at the kth time instant under two (or more) models, e.g., a stationary model M_1 and a nonstationary model M_2 with each filter using a different combination of the previous model conditioned estimates. The model switching process is assumed to be a Markov chain with the known transition matrix

$$\boldsymbol{P}_{mn} = \begin{bmatrix} p_{11} & 1 - p_{11} \\ 1 - p_{22} & p_{22} \end{bmatrix} \tag{3.46}$$

where p_{mn} is the probability that the model M_n is in effect at time k given that the model M_m was in effect at time $k - 1$. The value of p_{mm} can be estimated as

$$p_{mm} = 1 - \frac{1}{\zeta_m} \tag{3.47}$$

where ζ_m is the expected sojourn time in state m.

One cycle of the algorithm consists of the following steps [6, p. 455]:

1. **Calculation of the mixing probabilities:** The probability that the model M_m was in effect at time $k - 1$ given that M_n is in effect at time k, conditioned on the measured data up to and including time $k - 1$ is given by

$$\mu_{m|n}(k - 1|k - 1) = \frac{1}{\bar{c}_n} p_{mn} \mu_m(k - 1) \qquad \text{for } m, n = 1, 2 \tag{3.48}$$

where $\mu_m(k - 1)$ is the probability that the system is in the mode M_m at time $k - 1$ and the normalizing constants are given by

$$\bar{c}_n = \sum_{m=1}^{2} p_{mn} \mu_m(k - 1). \tag{3.49}$$

2. **Mixing:** Starting with the weight vector estimates of the two filters and their associated covariances at time $k - 1$, the mixed initial conditions $\hat{\boldsymbol{w}}_0^{(n)}(k - 1)$

58
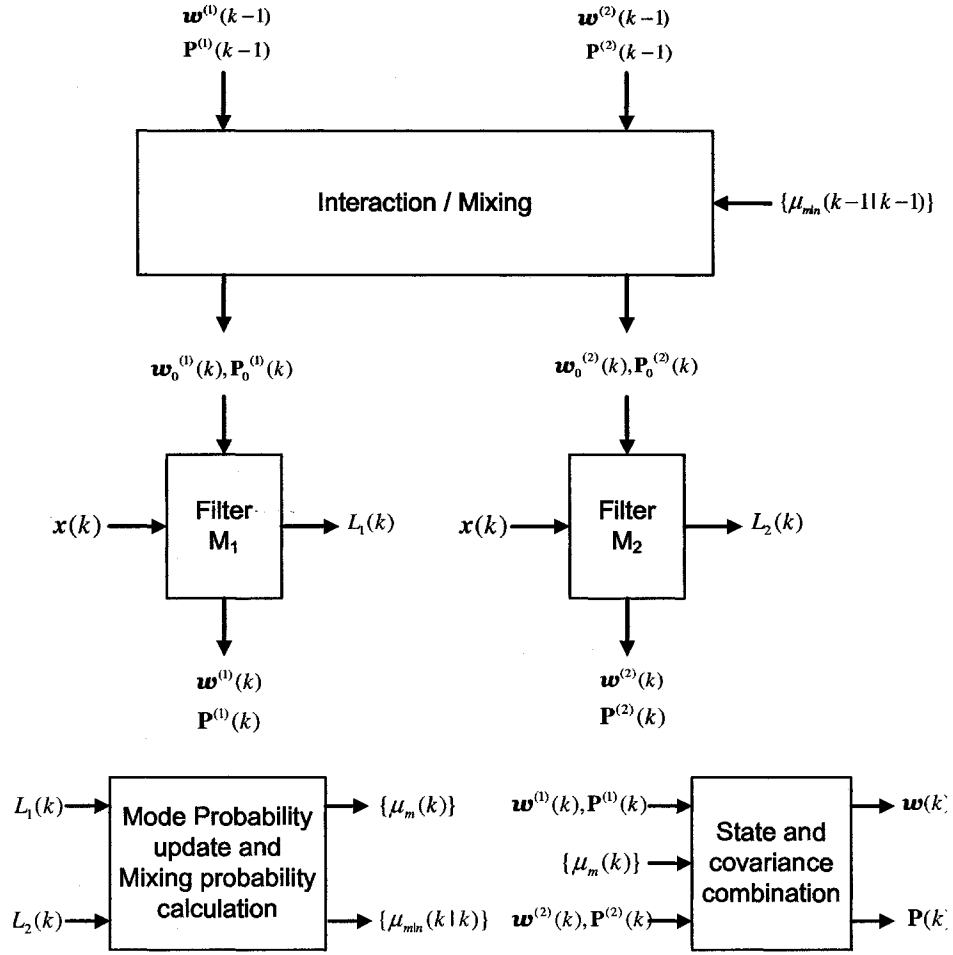
Figure 3.1: The IMM estimator.

and $\boldsymbol{P}_0^{(n)}(k-1|k-1)$ for the filter matched to M_n are computed as

$$\hat{\boldsymbol{w}}_0^{(n)}(k-1) = \sum_{m=1}^{2} \hat{\boldsymbol{w}}^{(m)}(k-1)\mu_{m|n}(k-1|k-1) \tag{3.50}$$

$$\boldsymbol{P}_0^{(n)}(k-1|k-1) = \sum_{m=1}^{2} \mu_{m|n}(k-1|k-1)\Big(\boldsymbol{P}^{(m)}(k-1|k-1)$$

$$+ \big(\hat{\boldsymbol{w}}^{(m)}(k-1) - \hat{\boldsymbol{w}}_0^{(n)}(k-1)\big)\big(\hat{\boldsymbol{w}}^{(m)}(k-1) - \hat{\boldsymbol{w}}_0^{(n)}(k-1)\big)^H\Big) \tag{3.51}$$

59

3. **Mode matched filtering:** The weight vector estimate from (3.50) and its associated covariance from (3.51) are used as inputs to the filter matched to M_n to yield $\hat{\boldsymbol{w}}^{(n)}(k)$ and $\hat{\boldsymbol{P}}^{(n)}(k|k)$.

The likelihood function corresponding to the nth filter, i.e., the probability of the measurement $\boldsymbol{z}(k)$ conditioned on the model M_n being the correct model at time k and given all the data history up to time $k - 1$, is calculated as

$$L_n(k) = \frac{1}{\left|2\pi \boldsymbol{S}^{(n)}(k)\right|} \exp\left(-\frac{1}{2}\big(\boldsymbol{z} - \hat{\boldsymbol{z}}^{(n)}(k|k-1)\big)^H \big(\boldsymbol{S}^{(n)}(k)\big)^{-1}\big(\boldsymbol{z} - \hat{\boldsymbol{z}}^{(n)}(k|k-1)\big)\right)$$

(3.52)

where $|\boldsymbol{S}|$ is the determinant of the matrix $\boldsymbol{S}$.

4. **Mode probability update:** The probability of each mode is updated according to its likelihood by

$$\mu_n(k) = \frac{1}{\bar{\bar{c}}} L_n(k) \bar{c}_n$$

(3.53)

where $\bar{\bar{c}}$ is a normalization constant given by

$$\bar{\bar{c}} = \sum_{n=1}^{2} L_n(k) \bar{c}_n.$$

(3.54)

5. **Estimate and covariance combination:** The model conditioned estimates and covariances are combined to give the beamformer weight vector estimate and its associated covariance as

$$\hat{\boldsymbol{w}}(k) = \sum_{n=1}^{2} \hat{\boldsymbol{w}}^{(n)}(k)\mu_n(k)$$

(3.55)

$$\boldsymbol{P}(k|k) = \sum_{n=1}^{2} \mu_n(k)\left(\boldsymbol{P}^{(n)}(k|k) + \big(\hat{\boldsymbol{w}}^{(n)}(k) - \hat{\boldsymbol{w}}(k)\big)\big(\hat{\boldsymbol{w}}^{(n)}(k) - \hat{\boldsymbol{w}}(k)\big)^H\right).$$

(3.56)

Figure 3.1 shows one cycle of the IMM estimator.

60

3.5 Simulations

3.5.1 Sensitivity to the choice of filter parameters

We study the effect of choosing the parameter $\sigma_{m,0}^2$ on the output SINR of the proposed Kalman filter-based beamformer. In this example, a stationary environment is assumed with a ULA of ten sensors spaced half wavelength apart. The array is steered towards the direction $\theta_0 = 3°$. The desired signal with SNR = 0 dB impinges on the array from the direction $\theta_s = 5°$. Also, two interference signals arrive from the directions $\theta_1 = 30°$ and $\theta_2 = 50°$, both with an interference-to-noise ratio (INR) equal to 30 dB. The desired signal is assumed to be always present in the test cell. The spatial signature of the desired signal is distorted by wave propagation effects in an inhomogeneous medium which are modelled as independent random phase increments drawn from a Gaussian random generator with variance 0.04. The phase distortions remain constant for all snapshots and are changed independently in each run [116]. The presumed signal steering vector is assumed to be normed as $\hat{a}^H \hat{a} = M$ and, following the guidelines of [116], the robustness parameter $\varepsilon = 3$ is used both in the Kalman filter-based beamformer and the robust MVDR beamformer of [116]. The parameters of the Kalman filter are selected as $\sigma_{m,1}^2 = 10^{-12}$, $\gamma = 1$, $\sigma_w^2 = 0$, and a random Gaussian vector is selected for the initialization of the Kalman filter-based beamformer, with the associated initial covariance calculated from (3.32) using $\epsilon(1) = 3$. Simulation results are averaged over 100 Monte Carlo runs.

Figure 3.2 shows the average output SINR after 100 iterations versus different choices of the parameter $\rho = \sigma_{m,0}^2 / (M\sigma_s^2 + \sigma_n^2)$. It can be seen from this figure that the Kalman filter is insensitive to the choice of $\sigma_{m,0}^2$ as the output SINR remains close to the optimal SINR for a wide range of the values of ρ. This behaviour can be attributed to the change in the norm of the weight vector estimate to yield an output power matching the value of $\sigma_{m,0}^2$. This can be seen from Figure 3.3 that displays

61
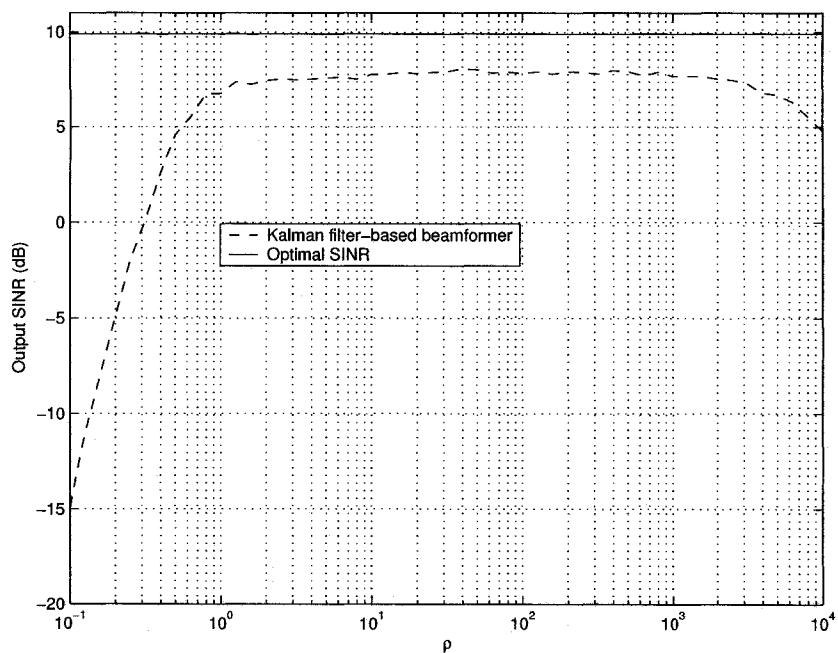
Figure 3.2: Average output SINR versus $\rho = \sigma_{m,0}^2/(M\sigma_s^2 + \sigma_n^2)$.
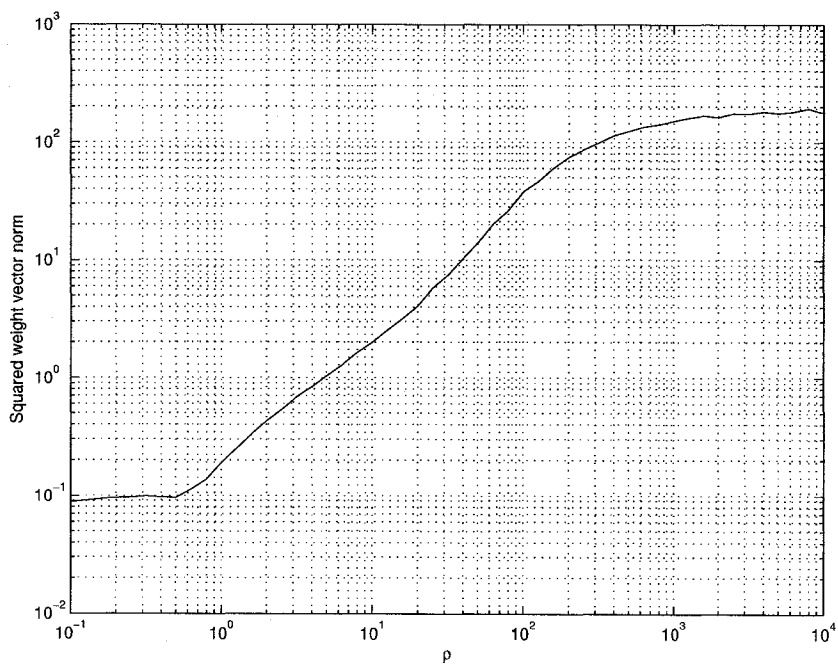
Figure 3.3: Squared norm of weight vector estimate versus $\rho = \sigma_{m,0}^2/(M\sigma_s^2 + \sigma_n^2)$.

62
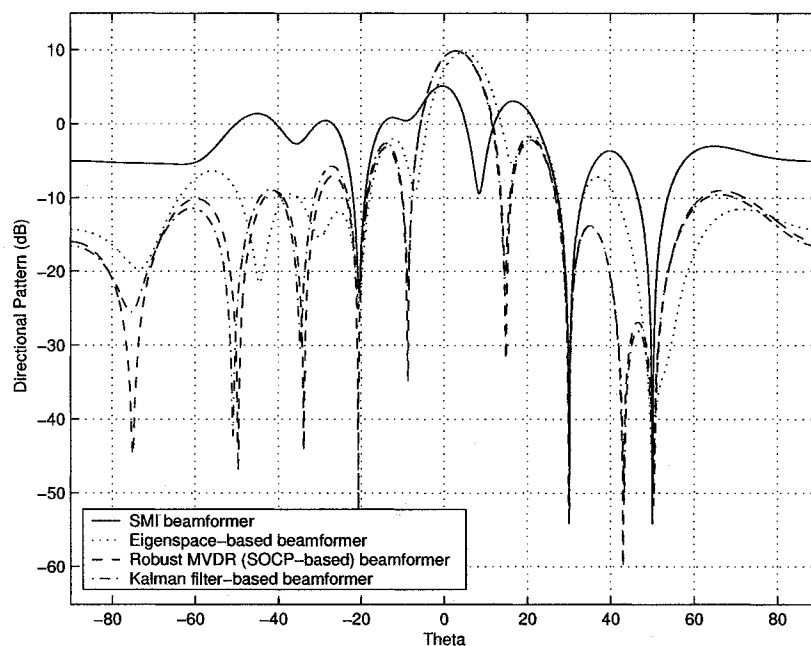
Figure 3.4: Array beampatterns of various beamformers.

the value of the squared norm of the final weight vector estimate versus ρ. From this figure, we can see how the squared norm of the weight vector changes over the range of ρ so that the output power of the beamformer matches the value of $\sigma_{m,0}^2$. At extremely high or low values of ρ, the weight vector norm can not match the value of ρ anymore due to the robustness constraint that controls the norm of this vector. In such extreme cases, the SINR becomes substantially degraded.

3.5.2   Performance comparison with other beamformers

In the next simulation examples, the performance of the proposed Kalman filter-based beamformer is compared with that of the SMI, the eigenspace-based, and the robust MVDR beamformers. First, we compare the performances of these techniques in the same stationary scenario described in the previous example. In particular, we use the same parameters of the Kalman filter and choose $\sigma_{m,0}^2 = 50(M\sigma_s^2 + \sigma_n^2)$.

63
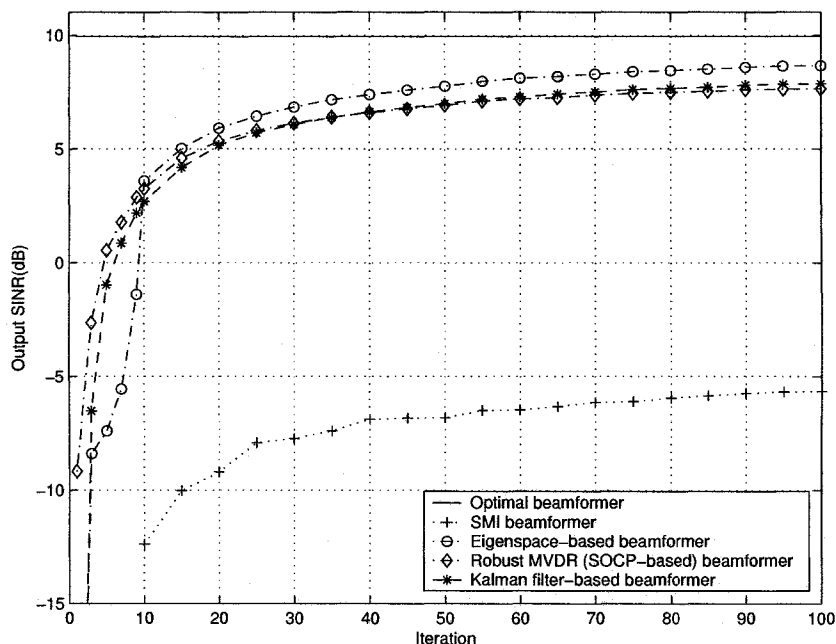
Figure 3.5: Average output SINR versus the iteration number.

Figure 3.4 shows the directional patterns of the four beamformers tested (the proposed beamformer is computed using 100 iterations). From this figure, we can see that although all these beamformers place nulls in the directions of the interference signals, self-nulling occurs only in the SMI beamformer whereas the directional patterns of the other three robust beamformers do not suffer from self-nulling. We can also observe a high degree of similarity in the directional patterns of the robust MVDR and the Kalman filter-based beamformers. This similarity can be explained by the fact that the two beamformers solve essentially the same problem but using different optimization approaches.

Figure 3.5 shows the average output SINR of the beamformers tested over 100 Monte Carlo runs versus the iteration number (snapshot index). Additionally, the optimal SINR curve is shown in this figure. From this figure, it can be observed that the eigenspace-based, Kalman filter-based, and robust MVDR beamformers have the best performances among the techniques tested. However, the performance of

64
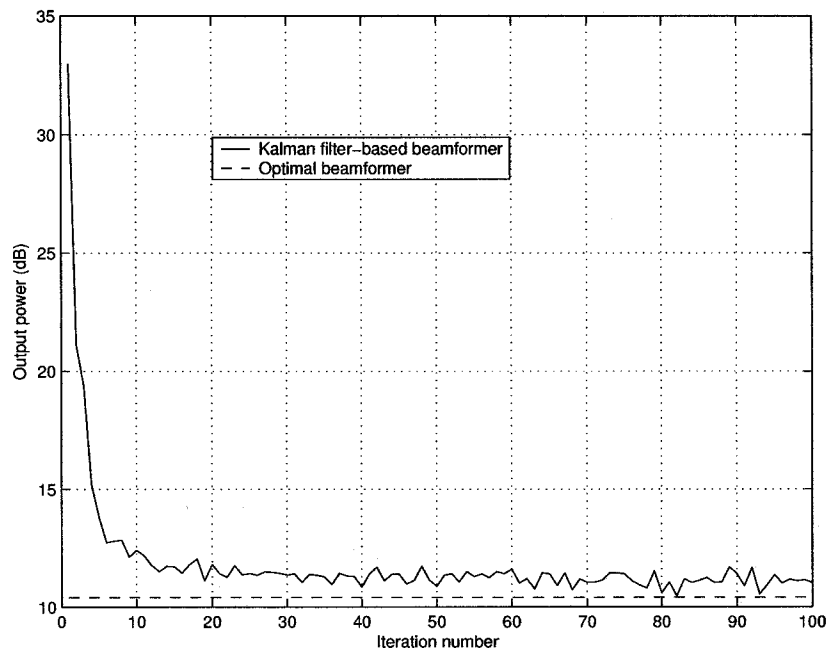
Figure 3.6: Output power versus the iteration number.

the eigenspace-based beamformer significantly drops if the number of snapshots is small. As expected, the performance of the Kalman filter-based and robust MVDR beamformers is nearly identical.

Figure 3.6 displays the instantaneous output power $|x^H(k)\hat{w}(k)|^2$ of the Kalman filter-based beamformer versus the iteration number (with the weight vector estimate normalized to have unit norm after each iteration) along with the optimal output power under the unit norm weight vector constraint. From this figure, we can see that the convergence rate of the Kalman filter beamformer to the optimal output power is quite fast and that the steady-state misadjustment is reasonably small. The beamformer converges to within 3 dB of its steady state output in around 10 iterations which agrees with our convergence analysis in (3.45).

Finally, we investigate the effect of the desired signal SNR on the performance of the four beamformers tested for the same scenario as in the previous example and

65
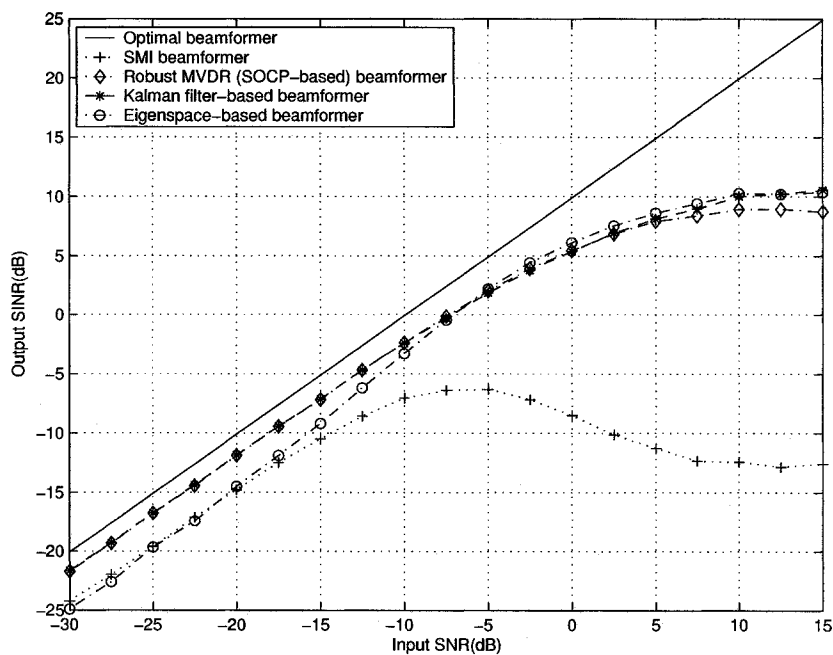
Figure 3.7: Average output SINR versus SNR.

the same Kalman filter parameters except for a sample size of $N = 20$ snapshots (which is enough for the convergence of our beamformer as follows from the previous simulation). Figure 3.7 shows the average output SINR of various beamformers versus the SNR. It can be seen from this figure that the performance of the Kalman filter-based beamformer is very similar to that of the robust MVDR beamformer. Both techniques outperform the eigenspace-based beamformer (at low SNR) and the SMI beamformer (at all SNR).

3.5.3   Robustness against rapid environmental changes

In our next simulation example, we consider a nonstationary scenario with only one interferer which changes its direction and power abruptly, while the desired signal is stationary and has the same look direction error and wavefront distortions as in the previous two examples. In particular, the interference signal arrives from $\theta_1 = 30°$

66

with INR = 30 dB for the first 625 snapshots, which is then abruptly changed to $\theta_1 = 60°$ with INR = 40 dB. We compare the stationary Kalman filter-based beamformer, the nonstationary beamformer with $\gamma = 1$ and $\sigma_w^2 = 10^{-4}$, the Kalman filter-based beamformer with hard decision switching between the previous two models, the IMM-based beamformer with the same previous two models, and the robust MVDR beamformer of [116] with a sliding window of 40 snapshots. For the hard decision switching algorithm, the NIS is calculated using a sliding window of length 5, and switching to the nonstationary model is performed if the NIS exceeds the upper 90% tail of the chi-square distribution with 15 degrees of freedom, whereas switching back to the stationary model occurs when the NIS drops back to its expected value which is equal to 15. For the IMM-based beamformer, the stationary model sojourn time is selected to be equal to 50 snapshots, whereas the nonstationary model sojourn time is selected to be equal to 10 snapshots. The other parameters of all the Kalman filters used are selected as $\sigma_{m,0}^2 = 50(M\sigma_s^2 + \sigma_n^2)$ and $\sigma_{m,1}^2 = 10^{-12}$. Simulation results are averaged over 500 Monte Carlo runs.

Figure 3.8 shows the average output SINR of the beamformers tested versus the iteration number. We can clearly see that the nonstationary Kalman filter beamformer can follow the environmental changes at the expense of a reduced steady-state SINR, whereas the stationary beamformer can not follow the rapid change in the environment. With the proposed modifications of the Kalman filter-based beamformer, the environmental changes can be followed and, at the same time, a high steady-state SINR can be maintained. It can be also observed from this figure that the IMM-based beamformer converges to the optimal solution faster than the hard decision switching-based beamformer.

It can be observed from Figure 3.8 that, because we have chosen a nearly optimal sliding window length for the robust MVDR beamformer in this example, its
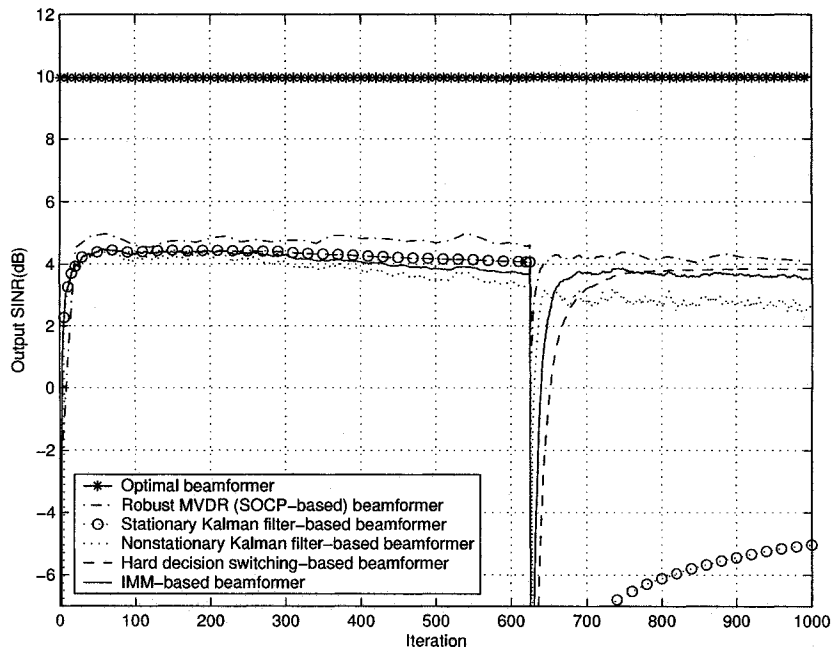
Figure 3.8: Average output SINR versus the iteration number.
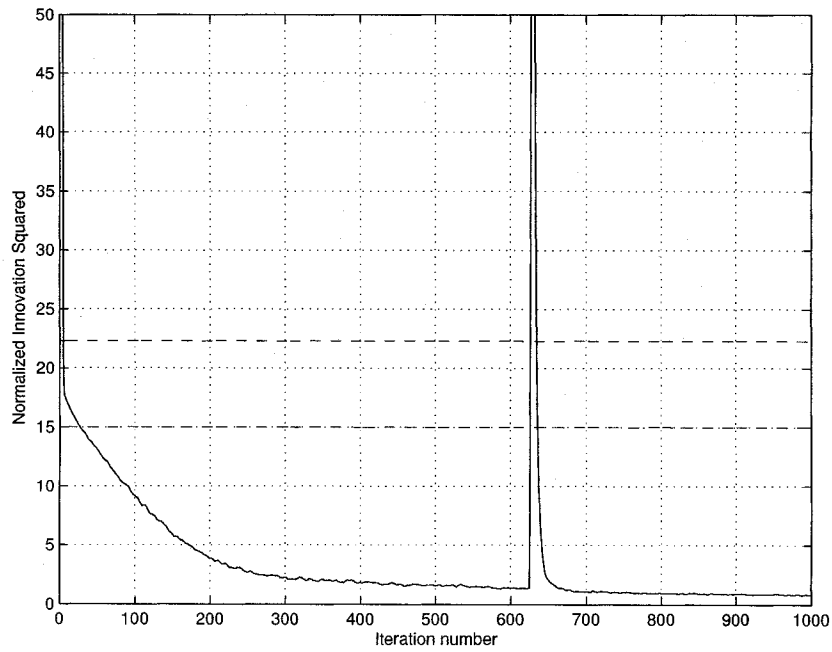
Figure 3.9: NIS for the hard decision-based beamformer.

68
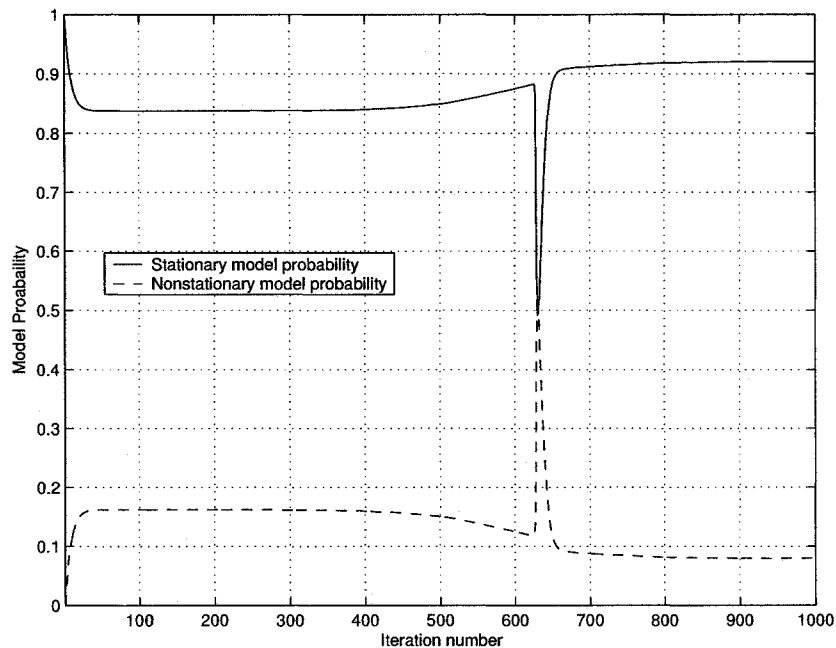
Figure 3.10: Model probabilities for the IMM-based beamformer.

performance is better than that of the proposed algorithms (both in terms of the convergence speed and the steady-state SINR). However, this advantage is achieved at the expense of higher computational complexity. Furthermore, in practical situations it may not be possible to optimize the window length.

Figure 3.9 shows the NIS for the hard decision-based beamformer together with the decision thresholds for a single run. From this figure, the effect of the abrupt environmental change on the NIS can be seen.

Figure 3.10 displays the average probability of the stationary and nonstationary models of the IMM-based beamformer. We notice from this figure that this beamformer successfully detects the nonstationarity, as can be seen from the increased probability of the nonstationary model at the transition time.

69

3.6   Conclusion

We have presented a novel computationally efficient Kalman filter-based algorithm for the narrowband robust MVDR beamformer of [116]. This algorithm is suitable for online implementation. We have also presented two modifications of the proposed algorithm to additionally account for nonstationary environments with abrupt changes using model switching and hypotheses merging techniques. Simulation results have validated an excellent robustness and a superior performance of the proposed Kalman filter-based beamformers compared to the existing robust beamforming algorithms.

The next two chapters will focus on deriving computationally efficient algorithms that extend the narrowband robust MVDR beamformer of [116] to the wideband signal environment.

Chapter 4

Wideband Robust Adaptive Beamforming: A Worst-Case Performance Optimization Approach

4.1 Introduction

The LCMV algorithm revised in Section 2.3.2.1 is one of the earliest adaptive beamforming algorithms for wideband signals. It minimizes the array output power subject to look direction constraints. These constraints preserve the signals arriving from the desired look direction that appear in phase after the presteering delay filters. However, in practical scenarios the array performance can be severely degraded due to mismatches in the presumed array manifold. These mismatches might result from look direction errors, array sensor location errors, presteering delays quantization

71

effects, and/or wavefront distortions. The effect of all these mismatches can be described as presteering errors. Hence, the received desired signal components can not be perfectly phase-aligned by the presteering delays and, thus, they are "interpreted" as interference and erroneously suppressed. In Section 2.4, we have reviewed several classical approaches to add robustness to the LCMV algorithm against various mismatches in the array manifold. However, all of these approaches either provide a fixed non-optimal amount of robustness that is not linked to the amount of mismatches in the array manifold, e.g., through the use of derivative constraints, or use an ad-hoc choice of robustness parameters as in the diagonal loading technique.

Recently, several theoretically-rigourous algorithms have been proposed to add robustness to the narrowband minimum variance beamformer so that it is matched to the amount of uncertainty in the array manifold. In Chapter 3 we have revised one of these recent approaches to robust narrowband beamforming. In this approach a spherical uncertainty set was used to describe the mismatches in the desired signal steering vector. Ellipsoidal and polyhedral uncertainty sets were used in [70], [74], and [121]. In all these approaches, the solution to the robust minimum variance beamforming problem was found through worst-case performance optimization.

Two extensions of the robust narrowband beamformer of [70] to the wideband case were presented in [118] that are based on subband decomposition. The first uses a frequency dependent robustness parameter for each subband and is termed the constant powerwidth robust minimum variance (CPRMV) beamformer. Whereas, the second combines the robust narrowband beamforming algorithm with a shading scheme designed for the delay and sum beamformer and is termed the constant beamwidth robust minimum variance (CBRMV) beamformer. However, the CBRMV beamformer has limited applicability as it requires a special underlying array structure due to the particular shading scheme employed. In addition, this shading scheme tends to deactivate some elements of the array at each subband which reduces the

interference suppression capability of the beamformer. Moreover, both the CPRMV and CBRMV beamformers are suboptimal as they optimize each subband weight vector independently and ignore constraining the phase response of the beamformer which causes distortions in the output signal.

In this chapter, we present a novel wideband beamformer with robustness directly related to the level of uncertainties in the array manifold. Our beamformer extends the work in [70], [74], and [116] to the more complicated wideband case and avoids the suboptimal subband decomposition approach of [118]. The proposed beamformer maintains a high gain not only to the signals appearing in phase after the presteering delays but also to all the signals with an additional norm-bounded phase error vector. Therefore, it prevents the cancellation of the desired signal components even in the presence of mismatches in the array manifold. The phase response of the array is controlled through additional linear phase constraints imposed on each of the FIR filters of the array processor. This has the effect of limiting phase distortions in the output signal. Moreover, these constraints are equivalent to NS1-PS derivative constraints and thus provide additional robustness to the beamformer. The resulting problem is solved through a worst-case performance optimization approach and is shown to be convex.

Next, we provide two implementations for the proposed beamformer; a lower computational complexity implementation that is based on discretizing the resulting *spectral* constraints over a finite grid of frequency points, and a higher complexity implementation using a technique recently developed in [25] that ensures that the spectral constraints are satisfied at each frequency point without discretization. The resulting optimization problems can both be solved with polynomial complexity using well-established interior point methods [83]. Simulation results show that the proposed beamformer has an improved robustness against various mismatches compared to earlier wideband beamforming techniques. Moreover, it is not sensitive to the

73

choice of the uncertainty set and requires only coarse knowledge about the amount of mismatches in the array manifold.

4.2 Robust Wideband Beamformer

As mentioned above, the effect of mismatches in the array manifold can be grouped into errors in the presteering delay values. As a result, the desired signal components appear incoherently at the output of the presteering filters and therefore they are considered as interference by the array processor. In this section, we will derive a novel wideband beamformer that is robust against presteering delay mismatches and provides an adjustable amount of robustness depending on the mismatch norm. The problem will be formulated as a convex optimization problem that can be solved in polynomial time using interior point methods [9].

4.2.1 Uncertainty set

Let us assume that the phase incoherency in the presteered desired signal components can be represented by a norm-bounded vector $\boldsymbol{\Delta}(f) \in \mathcal{A}_{\varepsilon,1}(f)$, i.e., the desired signal component with frequency f arriving from the direction θ_s appears after the presteering delays as

$$\boldsymbol{T}(f)\boldsymbol{a}(f,\theta_s) = e^{j2\pi f\varsigma}\mathbf{1}_M + \boldsymbol{\Delta}(f), \qquad \boldsymbol{\Delta}(f) \in \mathcal{A}_{\varepsilon,1}(f) \tag{4.1}$$

where ς is a common time delay at each of the M sensors and we have introduced the phase term $e^{j2\pi f\varsigma}$ to avoid dependency of the phase error vector on the phase center location of the array [12]. We define the mismatch set $\mathcal{A}_{\varepsilon,1}(f)$ as

$$\mathcal{A}_{\varepsilon,1}(f) \triangleq \left\{ \boldsymbol{\Delta}(f) \in \mathbb{C}^M \ \middle| \ \|\boldsymbol{\Delta}(f)\|_1 \leq \varepsilon(f) \right\} \tag{4.2}$$

where $\| \cdot \|_i$ is the vector i-norm.

74

For example, we can form the robustness set $\mathcal{A}_{\varepsilon,1}(f)$ for the linear array structure described in Section 2.2.3 by considering a look direction error $\Delta\theta$, a maximum error in sensor location along the array line given by $\Delta z = \alpha_z c/(2f_u)$ where α_z is the relative sensor displacement with respect to the presumed array inter-element spacing, and a quantization step in presteering delays given by ΔT. The M-dimensional vector containing the maximum error in the values of the presteering delays is given by

$$\boldsymbol{\Upsilon} = \frac{1}{c}\left((\hat{\boldsymbol{z}} + \Delta z\boldsymbol{\zeta})\sin(\theta_0 + \Delta\theta) - \hat{\boldsymbol{z}}\sin\theta_0\right) + \frac{1}{2}\Delta T\boldsymbol{\xi} \qquad (4.3)$$

where $\hat{\boldsymbol{z}} = [\hat{z}_1, \ldots, \hat{z}_M]^T$ is the M-dimensional vector containing the presumed locations of the M sensors along the z-axis, $\boldsymbol{\zeta}$ is a vector of length M with alternating 1's and -1's that corresponds to the maximum deformation in the array structure along the array line, and each component of the M-dimensional vector $\boldsymbol{\xi}$ is ± 1 with the same sign of the corresponding component of the first term of (4.3).

Therefore, the M-dimensional vector containing the maximum phase error in the presteered desired signal is given by

$$\boldsymbol{\Delta}_{\max}(f) = \left[e^{j2\pi f\Upsilon_1} - e^{j2\pi f\varsigma}, \ldots, e^{j2\pi f\Upsilon_M} - e^{j2\pi f\varsigma}\right]^T. \qquad (4.4)$$

The 1-norm of the maximum phase error vector at frequency f is therefore given by

$$\varepsilon(f) = \min_{-\pi \le 2\pi f\varsigma \le \pi} \sum_{i=1}^{M} \sqrt{2 - 2\cos(2\pi f(\Upsilon_i - \varsigma))}. \qquad (4.5)$$

4.2.2   Derivation

In order to prevent cancellation of the desired signal components, we impose a high gain constraint over the whole frequency band of the desired signal for all the presteered received signal vectors with phase errors that belong to the set $\mathcal{A}_{\varepsilon,1}(f)$. This constraint can be written as

$$|H(f, \theta_s)| \ge 1 \quad \forall \boldsymbol{\Delta}(f) \in \mathcal{A}_{\varepsilon,1}(f); f \in [f_l, f_u]. \qquad (4.6)$$

The optimal weight vector of the robust wideband beamformer can be obtained by minimizing the beamformer output power subject to the set of constraints in (4.6). This constraint set can be represented via infinite number of single constraints. Each constraint is imposed at a certain frequency f and corresponds to the worst-case mismatch over $\mathcal{A}_{\varepsilon,1}(f)$ at this frequency. Thus, we can write (4.6) as

$$\min_{\Delta(f) \in \mathcal{A}_{\varepsilon,1}(f)} |H(f, \theta_s)| \geq 1 \quad \forall f \in [f_l, f_u]. \tag{4.7}$$

In the presence of phase errors of the form (4.1), the array response towards the desired signal is given by

$$
\begin{aligned}
H(f, \theta_s) &= e^{j2\pi f_\varsigma} \boldsymbol{w}^T \big(\boldsymbol{d}(f) \otimes \mathbf{1}_M \big) + \boldsymbol{w}^T \big(\boldsymbol{d}(f) \otimes \boldsymbol{\Delta}(f) \big) \\
&= e^{j2\pi f_\varsigma} \boldsymbol{w}^T \boldsymbol{C}_0 \boldsymbol{d}(f) + \boldsymbol{w}^T \boldsymbol{Q}(f) \boldsymbol{\Delta}(f)
\end{aligned} \tag{4.8}
$$

where

$$\boldsymbol{Q}(f) \triangleq \boldsymbol{d}(f) \otimes \boldsymbol{I}_M \in \mathbb{C}^{ML \times M}. \tag{4.9}$$

We will consider each frequency component independently and find the minimum value of $|H(f, \theta_s)|$ that corresponds to the worst-case mismatch at this frequency. Using the triangle inequality we can write

$$\left| e^{j2\pi f_\varsigma} \boldsymbol{w}^T \boldsymbol{C}_0 \boldsymbol{d}(f) + \boldsymbol{w}^T \boldsymbol{Q}(f) \boldsymbol{\Delta}(f) \right| \geq \left| \boldsymbol{w}^T \boldsymbol{C}_0 \boldsymbol{d}(f) \right| - \left| \boldsymbol{w}^T \boldsymbol{Q}(f) \boldsymbol{\Delta}(f) \right|. \tag{4.10}$$

Maximizing the second term in the R.H.S. of (4.10) over the mismatch set $\mathcal{A}_{\varepsilon,1}(f)$ yields

$$\max_{\Delta(f) \in \mathcal{A}_{\varepsilon,1}(f)} \left| \boldsymbol{w}^T \boldsymbol{Q}(f) \boldsymbol{\Delta}(f) \right| = \varepsilon(f) \left\| \boldsymbol{Q}^T(f) \boldsymbol{w} \right\|_\infty. \tag{4.11}$$

Combining (4.10) and (4.11), the minimum value of $|H(f, \theta_s)|$ over the mismatch set at frequency f is given by

$$\min_{\Delta(f) \in \mathcal{A}_{\varepsilon,1}(f)} |H(f, \theta_s)| = \left| \boldsymbol{w}^T \boldsymbol{C}_0 \boldsymbol{d}(f) \right| - \varepsilon(f) \left\| \boldsymbol{Q}^T(f) \boldsymbol{w} \right\|_\infty. \tag{4.12}$$
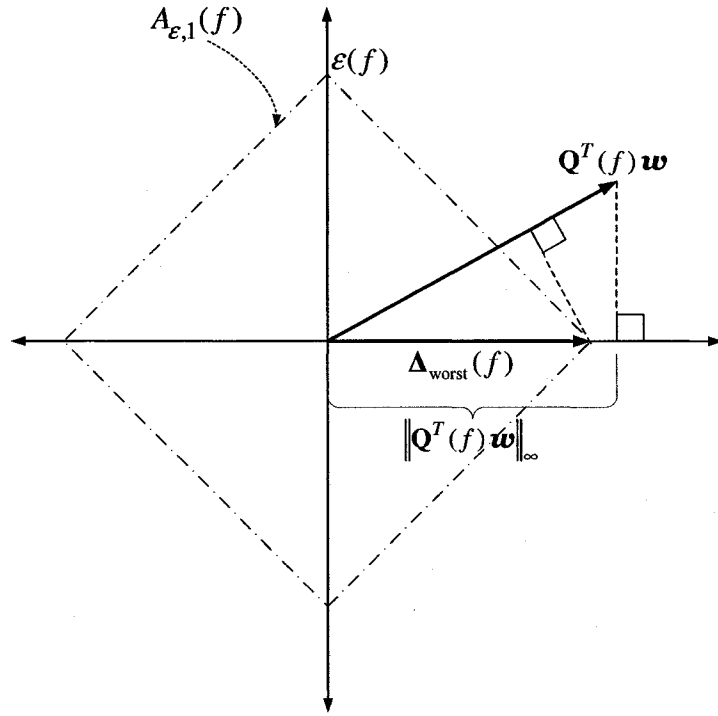
Figure 4.1: Worst-case mismatch (2–dimensional case).

The worst-case mismatch vector that yields this minimum is given by

$$\boldsymbol{\Delta}_{\text{worst}}(f) = -\varepsilon(f)e^{j(\phi_1(f)-\psi_1(f))}\boldsymbol{e}_p \tag{4.13}$$

where $\phi_1(f) = \text{angle}\left\{e^{j2\pi f_s}\boldsymbol{w}^T\boldsymbol{C}_0\boldsymbol{d}(f)\right\}$, $\psi_1(f) = \text{angle}\left\{(\boldsymbol{Q}^T(f)\boldsymbol{w})_p\right\}$, and p is the index of the component of $\boldsymbol{Q}^T(f)\boldsymbol{w}$ with the largest absolute value. This is demonstrated in Figure 4.1 for the case of a two sensor array.

The robust wideband beamforming problem can therefore be written as

$$\min_{\boldsymbol{w}} \quad \boldsymbol{w}^T\boldsymbol{R}_x\boldsymbol{w}$$

$$\text{s.t.} \quad \left|\boldsymbol{w}^T\boldsymbol{C}_0\boldsymbol{d}(f)\right| - \varepsilon(f)\left\|\boldsymbol{Q}^T(f)\boldsymbol{w}\right\|_\infty \geq 1 \quad \forall f \in [f_l, f_u]. \tag{4.14}$$

Note that $\boldsymbol{w}^T\boldsymbol{C}_0\boldsymbol{d}(f)$ and $\boldsymbol{Q}^T(f)\boldsymbol{w}$ can be expressed as

$$\boldsymbol{w}^T\boldsymbol{C}_0\boldsymbol{d}(f) = \mathcal{W}_1(f) + \ldots + \mathcal{W}_M(f) \tag{4.15}$$

$$\boldsymbol{Q}^T(f)\boldsymbol{w} = [\mathcal{W}_1(f), \ldots, \mathcal{W}_M(f)]^T \tag{4.16}$$

where $\mathcal{W}_m(f)$ is the frequency response of the mth FIR filter of the array processor, i.e., $\mathcal{W}_m(f) = \boldsymbol{\mathcal{W}}_m^T \boldsymbol{d}(f)$ where $\boldsymbol{\mathcal{W}}_m = [\omega_{m,1}, \omega_{m,2}, \ldots, \omega_{m,L}]^T$ is the $L \times 1$ vector containing the coefficients of the mth FIR filter.

Substituting with (4.15) and (4.16) in (4.14), we can write the robust beamforming problem with the constraints expressed in the frequency domain as

$$
\begin{aligned}
\min_{\boldsymbol{w}, v} \quad & \boldsymbol{w}^T \boldsymbol{R}_x \boldsymbol{w} \\
\text{s.t.} \quad & |\mathcal{W}_m(f)| \leq v && \forall m = 1, \ldots, M; f \in [f_l, f_u] \\
& \left| \sum_{m=1}^{M} \mathcal{W}_m(f) \right| - \varepsilon(f) v \geq 1 && \forall f \in [f_l, f_u].
\end{aligned}
\tag{4.17}
$$

The gain constraint in (4.6) can prevent the cancellation of the desired signal components, yet it does not guarantee a distortionless response because the phase response of the array towards the desired signal is completely unconstrained. To alleviate this problem, we impose a Type 1 linear phase constraint on each of the M FIR filters (i.e., L must be odd) [86]. Although this constraint does not guarantee an overall linear phase response towards all the received signals, the overall response has nearly linear phase for the desired signal components with norm-bounded phase errors. In addition, we have shown in (2.68) that the imposed linear phase constraints are equivalent to NS1-PS derivative constraints in the case of an all-pass response, and therefore, they add more robustness to the beamformer against presteering errors.

Under the linear phase constraint, we can write the frequency response of the mth FIR filter as

$$
\mathcal{W}_m(f) = e^{-j\phi_{L_c}(f)} \left(\omega_{m,L_c} + 2 \sum_{k=L_c+1}^{L} \omega_{m,k} \cos(2\pi f(k - L_c) T_s) \right)
\tag{4.18}
$$

where $\phi_{L_c}(f) = 2\pi f(L_c - 1) T_s$.

We define the vector $\boldsymbol{g}(f) \in \mathbb{R}^{L_c}$, and the matrices $\boldsymbol{C}_m \in \mathbb{R}^{ML \times L}$ and $\boldsymbol{B} \in \mathbb{R}^{L \times L_c}$ as

$$\boldsymbol{g}(f) \triangleq [1, \cos(2\pi f T_s), \ldots, \cos(2\pi f (L_c - 1) T_s)]^T \tag{4.19}$$

$$\boldsymbol{C}_m \triangleq \boldsymbol{I}_L \otimes \boldsymbol{e}_m \qquad \text{for } m = 1, \ldots, M \tag{4.20}$$

$$\boldsymbol{B}^T \triangleq \begin{bmatrix} 0 & \cdots & 0 & 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & \cdot^{\cdot^{\cdot}} & & & & \ddots & & \vdots \\ 0 & 1 & 0 & & & & 0 & 1 & 0 \\ 1 & 0 & 0 & & \cdots & & 0 & 0 & 1 \end{bmatrix}. \tag{4.21}$$

Note that $\boldsymbol{C}_m^T \boldsymbol{w} = \boldsymbol{\mathcal{W}}_m$. Thus, we can rewrite (4.18) as

$$\mathcal{W}_m(f) = e^{-j\phi_{L_c}(f)} \boldsymbol{w}^T \boldsymbol{C}_m \boldsymbol{B} \boldsymbol{g}(f). \tag{4.22}$$

Also, noticing that $\sum_{i=1}^{M} \boldsymbol{C}_i = \boldsymbol{C}_0$, and therefore, $\sum_{m=1}^{M} \boldsymbol{\mathcal{W}}_m = \boldsymbol{C}_0^T \boldsymbol{w}$, we can write

$$\sum_{m=1}^{M} \mathcal{W}_m(f) = e^{-j\phi_{L_c}(f)} \boldsymbol{w}^T \boldsymbol{C}_0 \boldsymbol{B} \boldsymbol{g}(f) \tag{4.23}$$

where $\boldsymbol{w}^T \boldsymbol{C}_m \boldsymbol{B} \boldsymbol{g}(f)$ and $\boldsymbol{w}^T \boldsymbol{C}_0 \boldsymbol{B} \boldsymbol{g}(f)$ are real.

Therefore, the robust beamforming problem in (4.14) with the additional linear phase constraints can be written as

$$\begin{aligned} \min_{\boldsymbol{w}, v} \quad & \boldsymbol{w}^T \boldsymbol{R}_x \boldsymbol{w} \\ \text{s.t.} \quad & \Big(\big(\boldsymbol{e}_{L_c}^T (\boldsymbol{J}_l - \boldsymbol{J}_l^T) \big) \otimes \big(\boldsymbol{1}_M^T \boldsymbol{\Omega}_m \big) \Big) \boldsymbol{w} = 0 && \forall m = 1, \ldots, M; l = 1, \ldots, L_c - 1 \\ & -v \le \boldsymbol{w}^T \boldsymbol{C}_m \boldsymbol{B} \boldsymbol{g}(f) \le v && \forall m = 1, \ldots, M; f \in [f_l, f_u] \\ & \big| \boldsymbol{w}^T \boldsymbol{C}_0 \boldsymbol{B} \boldsymbol{g}(f) \big| - \varepsilon(f) v \ge 1 && \forall f \in [f_l, f_u]. \end{aligned} \tag{4.24}$$

The problem is still non-convex due to the absolute value operator in the last constraint. However, we can see that if $\{\boldsymbol{w}, v\}$ is an optimal solution to the above optimization problem, then $\{-\boldsymbol{w}, v\}$ is also an optimal solution. Also, examining the

constraint $\left|\sum_{m=1}^{M} \mathcal{W}_m(f)\right| \geq 1 + \varepsilon(f)v$ and from the continuity of the Fourier transform, it is clear that for a feasible weight vector $\boldsymbol{w}$, $e^{j\phi_{L_c}(f)} \sum_{m=1}^{M} \mathcal{W}_m(f)$ can either be positive or negative over the entire frequency band $[f_l, f_u]$ and can not change sign at any frequency. From these two remarks, it follows that imposing a nonnegativity constraint on the phase-rotated real-valued sum of the Fourier transforms of the M FIR filters does not lead to any loss of optimality and yields the same global optimum solution of the nonconvex problem. Therefore, we can replace $\left|\boldsymbol{w}^T \boldsymbol{C}_0 \boldsymbol{Bg}(f)\right|$ in the last constraint of (4.24) by $\boldsymbol{w}^T \boldsymbol{C}_0 \boldsymbol{Bg}(f)$, and rewrite the constraint as $\boldsymbol{w}^T \boldsymbol{C}_0 \boldsymbol{Bg}(f)$ $-\varepsilon(f)v \geq 1$. The resulting problem is a convex optimization problem given by:

$$
\begin{aligned}
\min_{\boldsymbol{w},v} \quad & \boldsymbol{w}^T \boldsymbol{R}_x \boldsymbol{w} \\
\text{s.t.} \quad & \left(\left(\boldsymbol{e}_{L_c}^T (\boldsymbol{J}_l - \boldsymbol{J}_l^T)\right) \otimes \left(\boldsymbol{1}_M^T \boldsymbol{\Omega}_m\right)\right)\boldsymbol{w} = 0 && \forall m = 1,\ldots,M; l = 1,\ldots,L_c - 1 \\
& \boldsymbol{w}^T \boldsymbol{C}_m \boldsymbol{Bg}(f) \leq v && \forall m = 1,\ldots,M; f \in [f_l, f_u] \\
& \boldsymbol{w}^T \boldsymbol{C}_m \boldsymbol{Bg}(f) \geq -v && \forall m = 1,\ldots,M; f \in [f_l, f_u] \\
& \boldsymbol{w}^T \boldsymbol{C}_0 \boldsymbol{Bg}(f) - \varepsilon(f)v \geq 1 && \forall f \in [f_l, f_u] \quad (4.25)
\end{aligned}
$$

Note that each of the last three constraint sets in (4.25) is an infinite number of linear inequality constraints imposed over the frequency band $[f_l, f_u]$. These constraints are commonly referred to as spectral constraints [25]. For the case of temporal sampling at the Nyquist rate, i.e., $f_s = 2f_u$ and the case of $f_l = 0$, the frequency interval $[-f_u, -f_l] \bigcup [f_l, f_u]$ is a convex set. The spectral constraints can then be formulated using the results of [25] as a finite number of linear matrix inequalities (LMIs) and linear equality constraints. This will be further explained in the next section.

4.3   Robust Beamformer Algorithms

In this section, we discuss two algorithms for implementing the robust beamforming problem in (4.25). First, we will convert the quadratic objective function into a linear

80

one. Let

$$\boldsymbol{R}_x = \boldsymbol{U}^T \boldsymbol{U} \tag{4.26}$$

be the Cholesky factorization of $\boldsymbol{R}_x$ [51]. Thus, we can write

$$\boldsymbol{w}^T \boldsymbol{R}_x \boldsymbol{w} = \|\boldsymbol{U}\boldsymbol{w}\|^2 . \tag{4.27}$$

Obviously, minimizing $\|\boldsymbol{U}\boldsymbol{w}\|$ is equivalent to minimizing $\boldsymbol{w}^T \boldsymbol{R}_x \boldsymbol{w}$. Hence, introducing an auxiliary variable t, and imposing the additional constraint $\|\boldsymbol{U}\boldsymbol{w}\| \leq t$, we can convert (4.25) into the following problem:

$$
\begin{aligned}
\min_{\boldsymbol{w}, v, t} \quad & t \\
\text{s.t.} \quad & \|\boldsymbol{U}\boldsymbol{w}\| \leq t \\
& \left(\left(\boldsymbol{e}_{L_c}^T (\boldsymbol{J}_l - \boldsymbol{J}_l^T) \right) \otimes \left(\mathbf{1}_M^T \boldsymbol{\Omega}_m \right) \right) \boldsymbol{w} = 0 && \forall m = 1, \ldots, M; l = 1, \ldots, L_c - 1 \\
& \boldsymbol{w}^T \boldsymbol{C}_m \boldsymbol{B} \boldsymbol{g}(f) \leq v && \forall m = 1, \ldots, M; f \in [f_l, f_u] \\
& \boldsymbol{w}^T \boldsymbol{C}_m \boldsymbol{B} \boldsymbol{g}(f) \geq -v && \forall m = 1, \ldots, M; f \in [f_l, f_u] \\
& \boldsymbol{w}^T \boldsymbol{C}_0 \boldsymbol{B} \boldsymbol{g}(f) - \varepsilon(f) v \geq 1 && \forall f \in [f_l, f_u] \tag{4.28}
\end{aligned}
$$

4.3.1 SOCP-based algorithm

The problem of representing the infinite number of spectral constraints in a finite manner can be handled through frequency discretization. Several discretization techniques are available in the literature [56], [80]. A common, but ad-hoc, discretization approach selects a uniformly spaced grid $\{f_i\}_{i=1}^N$ over the frequency interval $[f_l, f_u]$.

81

Therefore, we can write the discretized implementation of (4.28) as

$$
\begin{aligned}
&\min_{\boldsymbol{w},v,t} && t \\
&\text{s.t.} && \|\boldsymbol{U}\boldsymbol{w}\| \leq t \\
& && \left(\left(\boldsymbol{e}_{L_c}^T (\boldsymbol{J}_l - \boldsymbol{J}_l^T) \right) \otimes \left(\boldsymbol{1}_M^T \boldsymbol{\Omega}_m \right) \right) \boldsymbol{w} = 0 && \forall m = 1, \ldots, M; l = 1, \ldots, L_c - 1 \\
& && \boldsymbol{w}^T \boldsymbol{C}_m \boldsymbol{B}\boldsymbol{g}(f_i) \leq v - \delta && \forall m = 1, \ldots, M; i = 1, \ldots, N \\
& && \boldsymbol{w}^T \boldsymbol{C}_m \boldsymbol{B}\boldsymbol{g}(f_i) \geq -v + \delta && \forall m = 1, \ldots, M; i = 1, \ldots, N \\
& && \boldsymbol{w}^T \boldsymbol{C}_0 \boldsymbol{B}\boldsymbol{g}(f_i) - \varepsilon(f_i)v \geq 1 + \delta && \forall i = 1, \ldots, N && (4.29)
\end{aligned}
$$

where N and δ are chosen as follows. For a fixed N, one must choose δ to be small enough so that the over-constraining of the problem at the frequencies $\{f_i\}_{i=1}^N$ does not result in significant performance loss. Yet, one must choose δ large enough to guarantee the satisfaction of the spectral constraints for all frequency components not included in the selected grid.

The number of design variables is $n_v = ML + 2$, and the first constraint is an $(ML + 1)$-dimensional second-order cone (SOC) constraint followed by $M(L_c - 1)$ linear equality constraints and $(2M + 1)N$ linear inequality constraints. The discretized implementation of the robust beamformer in (4.29) is therefore an SOCP problem [102]. The number of iterations required to solve an SOCP problem using interior point methods is bounded by the square root of the number of constraints [73]. The computational complexity associated with each iteration is of $\mathcal{O}\left(n_v^2 \sum_i q_i \right)$, where n_v is the number of optimization variables and q_i is the dimension of the ith constraint [73]. Therefore, the worst-case computational load of (4.29) is of $\mathcal{O}\left(M^{3.5} L^2 \zeta^{1.5} \right)$ where $\zeta = \max\{N, L\}$. This computational complexity is comparable to the complexity of the classical LCMV algorithm which requires $\mathcal{O}\left(M^3 L^3 \right)$ operations to compute the inverse of the data covariance matrix in (2.54) [84].

82

4.3.2   SDP-based algorithm

In [25], Davidson et. al. proposed a technique to precisely enforce piecewise constant and piecewise trigonometric polynomial spectral constraints in a finite and convex manner. This technique avoids any heuristic approximations involved in discretization techniques. However, it requires the spectral domain $[-f_u, -f_l] \bigcup [f_l, f_u]$ to be convex, which is true for the common case of temporal sampling at the Nyquist frequency. In the following, we will briefly review several results of [25] and show how they can be applied to our robust beamforming problem with $f_u = f_s/2$.

First, we approximate $\varepsilon(f)$ over the frequency interval $[-f_s/2, f_s/2]$ as a series expansion using the basis functions $\{\cos(2\pi f k T_s)\}_{k=0}^{L_c-1}$ as [93, p. 313]

$$\varepsilon(f) = \sum_{i=0}^{L_c-1} b_i \cos(2\pi f i T_s) \tag{4.30}$$

where

$$b_0 = \frac{1}{f_s} \int_{-f_s/2}^{f_s/2} \varepsilon(f) df \tag{4.31}$$

$$b_i = \frac{2}{f_s} \int_{-f_s/2}^{f_s/2} \varepsilon(f) \cos(2\pi f i T_s) df, \qquad i = 1, \dots, L_c - 1. \tag{4.32}$$

The robust beamforming problem with the approximate mismatch norm is given by

$$\min_{w,v,t} \quad t$$

$$\text{s.t.} \quad \|Uw\| \leq t$$

$$\left((e_{L_c}^T (J_l - J_l^T)) \otimes (\mathbf{1}_M^T \Omega_m) \right) w = 0 \qquad \forall m = 1, \dots, M; l = 1, \dots, L_c - 1$$

$$w^T C_m B g(f) \leq v \qquad \forall m = 1, \dots, M; f \in [f_l, f_u]$$

$$w^T C_m B g(f) \geq -v \qquad \forall m = 1, \dots, M; f \in [f_l, f_u]$$

$$w^T C_0 B g(f) - v b^T g(f) \geq 1 \qquad \forall f \in [f_l, f_u] \tag{4.33}$$

where $b = [b_0, b_1, \dots, b_{L_c-1}]^T$.

We define the lower triangular $(n+1) \times (n+1)$ Toeplitz matrices $\{\boldsymbol{P}_{0,n}\}_{i=0}^n$ as

$$[\boldsymbol{P}_{k,n}]_{h,l} = \begin{cases} 1 & \text{if } h = k+l \\ 0 & \text{otherwise} \end{cases} \qquad \text{for } h, l = 1, \ldots, n+1 \qquad (4.34)$$

Note that $\langle \boldsymbol{P}_{k,n}, \boldsymbol{Y} \rangle = \sum_{l=1}^{n-k+1} [\boldsymbol{Y}]_{l+k,l}$ for all $\boldsymbol{Y} \in \mathbb{R}^{(n+1) \times (n+1)}$ where the inner product of two real matrices $\langle \boldsymbol{A}, \boldsymbol{B} \rangle$ is defined as the trace of $\boldsymbol{A}^T \boldsymbol{B}$ [51]. Thus, $\langle \boldsymbol{P}_{k,n}, \boldsymbol{Y} \rangle$ is the sum of the elements on the kth lower diagonal of $\boldsymbol{Y}$. Also, we define the linear operator $\boldsymbol{\Gamma}(\boldsymbol{y}) : \mathbb{R}^{n+1} \mapsto \mathbb{R}^{(n+1) \times (n+1)}$ as

$$\boldsymbol{\Gamma}(\boldsymbol{y}) = y_1 \boldsymbol{P}_{0,n} + 2 \sum_{i=1}^n y_{i+1} \boldsymbol{P}_{i,n}, \qquad (4.35)$$

and thus the adjoint operator $\boldsymbol{y} = \boldsymbol{\Gamma}^\star(\boldsymbol{Y}) : \mathbb{R}^{(n+1) \times (n+1)} \mapsto \mathbb{R}^{n+1}$ is given by

$$y_1 = \langle \boldsymbol{P}_{0,n}, \boldsymbol{Y} \rangle \qquad (4.36)$$

$$y_i = 2\langle \boldsymbol{P}_{i-1,n}, \boldsymbol{Y} \rangle \qquad \text{for } i = 2, \ldots, n+1. \qquad (4.37)$$

Next, let us define the two scalars $\gamma_0(\vartheta_0)$ and $\gamma_1(\vartheta_0)$ and the family of operators $\boldsymbol{\Xi}(\boldsymbol{y}; \vartheta_0) : \mathbb{R}^{n+1} \mapsto \mathbb{R}^{n \times n}$ as

$$\gamma_0(\vartheta_0) = 2\cos(\vartheta_0)(1 - \cos(\vartheta_0)) \qquad (4.38)$$

$$\gamma_1(\vartheta_0) = -2(1 - \cos(\vartheta_0)) \qquad (4.39)$$

$$\begin{aligned} \boldsymbol{\Xi}(\boldsymbol{y}; \vartheta_0) &= \gamma_0(\vartheta_0) \left(y_1 \boldsymbol{P}_{0,n-1} + 2 \sum_{i=1}^{n-1} y_{i+1} \boldsymbol{P}_{i,n-1} \right) \\ &+ \gamma_1(\vartheta_0) \left(\sum_{i=1}^n y_{i+1} \boldsymbol{P}_{i-1,n-1} + \sum_{i=0}^{n-2} y_{i+1} \boldsymbol{P}_{i+1,n-1} \right). \end{aligned} \qquad (4.40)$$

The adjoint operator $\boldsymbol{y} = \boldsymbol{\Xi}^\star(\boldsymbol{Y}; \vartheta_0) : \mathbb{R}^{n \times n} \mapsto \mathbb{R}^{n+1}$ is defined such that

$$y_1 = \gamma_0(\vartheta_0) \langle \boldsymbol{P}_{0,n-1}, \boldsymbol{Y} \rangle + \gamma_1(\vartheta_0) \langle \boldsymbol{P}_{1,n-1}, \boldsymbol{Y} \rangle$$

$$y_k = 2\gamma_0(\vartheta_0) \langle \boldsymbol{P}_{k-1,n-1}, \boldsymbol{Y} \rangle + \gamma_1(\vartheta_0) \left(\langle \boldsymbol{P}_{k-2,n-1}, \boldsymbol{Y} \rangle + \langle \boldsymbol{P}_{k,n-1}, \boldsymbol{Y} \rangle \right) \quad k = 2, \ldots, n-1$$

$$y_n = 2\gamma_0(\vartheta_0) \langle \boldsymbol{P}_{n-1,n-1}, \boldsymbol{Y} \rangle + \gamma_1(\vartheta_0) \langle \boldsymbol{P}_{n-2,n-1}, \boldsymbol{Y} \rangle$$

$$y_{n+1} = \gamma_1(\vartheta_0) \langle \boldsymbol{P}_{n-1,n-1}, \boldsymbol{Y} \rangle. \qquad (4.41)$$

84

Let $\mathcal{K}(\vartheta_0)$ describe the set of coefficients of real trigonometric polynomials with the form $\sum_{k=1}^{n+1} p_k \cos\left((k-1)\vartheta\right)$ that are nonnegative on the interval $\vartheta \in [\vartheta_0, \pi]$,

$$\mathcal{K}(\vartheta_0) = \left\{ p \in \mathbb{R}^{n+1} \,\middle|\, \sum_{k=1}^{n+1} p_k \cos((k-1)\vartheta) \geq 0, \quad \text{for all } \vartheta \in [\vartheta_0, \pi] \right\} \quad (4.42)$$

The set $\mathcal{K}(\vartheta_0)$ can be interpreted as a convex cone in $\mathbb{R}^{n+1}$. In [25], Davidson et al. presented an alternative LMI description of $\mathcal{K}(\vartheta_0)$ given by

$$\mathcal{K}(\vartheta_0) = \left\{ p \in \mathbb{R}^{n+1} \,\middle|\, p = \Gamma^\star(Y) + \Xi^\star(X; \vartheta_0) \quad \text{for some } Y \in \mathcal{S}_+^{(n+1)\times(n+1)}, X \in \mathcal{S}_+^{n \times n} \right\} \quad (4.43)$$

where $\mathcal{S}_+^{n \times n}$ is the set of $n \times n$ real symmetric positive semidefinite matrices.

Note that each of the $2M + 1$ spectral constraints in (4.28) is in the form of a trigonometric polynomial (with L_c terms) inequality. Therefore, using the above descriptions of the set of real nonnegative trigonometric polynomials in (4.42) and (4.43) with $\vartheta = 2\pi f T_s$, and thus $f \in [f_l, f_s/2] \mapsto \vartheta \in [2\pi f_l T_s, \pi]$, i.e., $\vartheta_0 = 2\pi f_l T_s$, and introducing the auxiliary matrices $\left\{ Y_i \in \mathbb{R}^{L_c \times L_c} \right\}_{i=1}^{2M+1}$ and $\left\{ X_i \in \mathbb{R}^{(L_c-1)\times(L_c-1)} \right\}_{i=1}^{2M+1}$, we can reformulate the ith spectral constraint as a linear equality constraint and two positive semidefiniteness constraints on the two auxiliary matrices X_i and Y_i. Therefore, the LMI implementation of the robust beamformer is given by

$$\begin{aligned}
\min \quad & t \\
\text{s.t.} \quad & \|Uw\| \leq t \\
& \left(\left(e_{L_c}^T (J_l - J_l^T) \right) \otimes \left(\mathbf{1}_M^T \Omega_m \right) \right) w = 0 \qquad \forall m = 1, \ldots, M; l = 1, \ldots, L_c - 1 \\
& Y_i, X_i \succeq 0 \qquad\qquad\qquad\qquad\qquad\qquad \forall i = 1, \ldots, 2M + 1 \\
& v e_1 - B^T C_m^T w = \Gamma^\star(Y_m) + \Xi^\star(X_m; 2\pi f_l T_s) \qquad \forall m = 1, \ldots, M \\
& v e_1 + B^T C_m^T w = \Gamma^\star(Y_{M+m}) + \Xi^\star(X_{M+m}; 2\pi f_l T_s) \qquad \forall m = 1, \ldots, M \\
& B^T C_0^T w - e_1 - v b = \Gamma^\star(Y_{2M+1}) + \Xi^\star(X_{2M+1}; 2\pi f_l T_s) \qquad (4.44)
\end{aligned}$$

where the optimization variables are $t, v, w, \{X_i\}_{i=1}^{2M+1}$, and $\{Y_i\}_{i=1}^{2M+1}$. Note that

each of the last three constraint sets is an L_c–dimensional linear equality constraint. The optimization problem in (4.44) is a semi-definite programming (SDP) problem that can be solved efficiently using interior point methods [102]. The computational complexity of one iteration of SDP is given by $\mathcal{O}\left(n_v^2 \sum_i q_i^2\right)$ where n_v is the number of optimization variables and q_i is the dimension of the ith constraint. The upper bound for the number of required iterations is given by $\mathcal{O}\left(\sqrt{\sum_i q_i}\right)$ [73]. Thus, the worst-case complexity of (4.44) is of $\mathcal{O}\left(M^{4.5}L^{6.5}\right)$.

Note that if temporal sampling is done at a rate higher than the Nyquist rate we can still apply the above LMI implementation of the spectral constraints over the whole frequency interval $[f_l, f_s/2]$. The excess constrained frequency band does not result in any significant performance degradation as will be demonstrated through our numerical examples.

4.4   Simulations

We consider a linear microphone array with $M = 10$ and $L = 9$. The sensors are assumed to be equi-spaced with spacing c/f_s where $f_s = 8000$ Hz. The array is presteered to the direction $\theta_0 = -30°$ and the presteering delays are quantized to 256 levels. The desired signal is a wideband signal with SNR equal to 10 dB and constant power spectral density over the frequency band 800–4000 Hz. A wideband interference signal is received from the direction $\theta_1 = 30°$ with a constant spectral density over the frequency band of interest and interference-to-signal ratio (ISR) equal to 20 dB. A uniformly-spaced frequency grid with $N = 11$ points is selected to discretize the spectral constraints and a value of $\delta = 10^{-3}$ is chosen. A sample size of $N_s = 16384$ is used for calculating the sample covariance matrix. For the CPRMV beamformer, 512 subbands are considered and the samples are divided into 32 batches. The covariance matrix of each subband is formed using the available 32 samples for that subband.
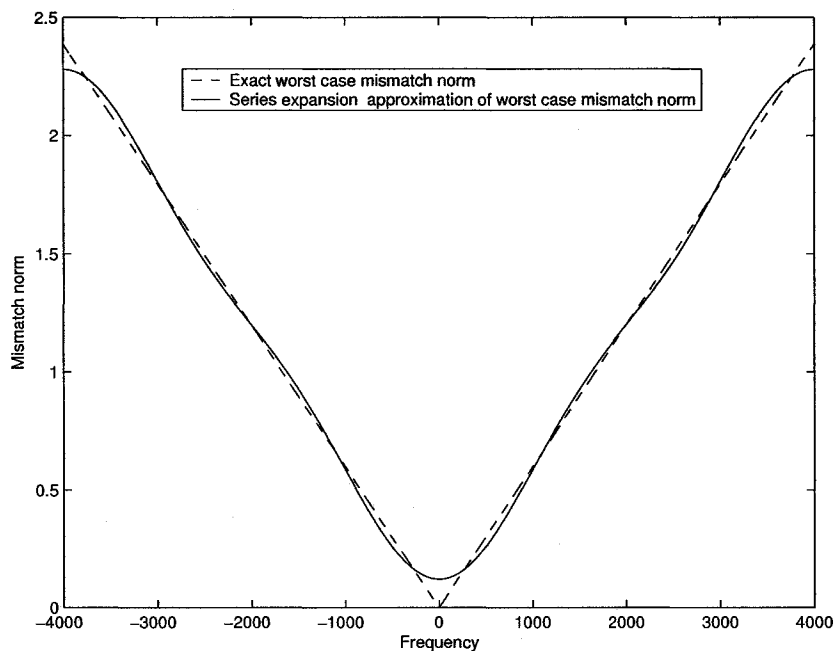
86

Figure 4.2: Mismatch norm and its series expansion approximation.

The desired signal is chosen as a chirp signal $s(k) = \cos(2\pi f_l k T_s + \pi \beta_c k^2 T_s^2)$ with the chirp rate $\beta_c = (f_u - f_l)/(N_s T_s)$, whereas the interference signal is generated as the sum of sinusoidal harmonics with random phases and equal powers occupying the band between 800 and 4000 Hz. The simulation results are averaged over 200 Monte Carlo runs.

4.4.1   Robustness against array calibration errors

We first consider the effect of the look direction error $\theta_0 - \theta_s$ on various beamformers. We assume perfect knowledge of the sensor locations and no quantization effects on presteering delays. Figure 4.2 shows the mismatch norm calculated from (4.5) with $\Delta\theta = 2°$, $\alpha_z = 0$, and $\Delta T = 0$, and its approximation by the series expansion of (4.30). We can clearly see that the expansion approximates $\varepsilon(f)$ over the desired signal frequency band very well. Figure 4.3 shows the average output SINRs versus
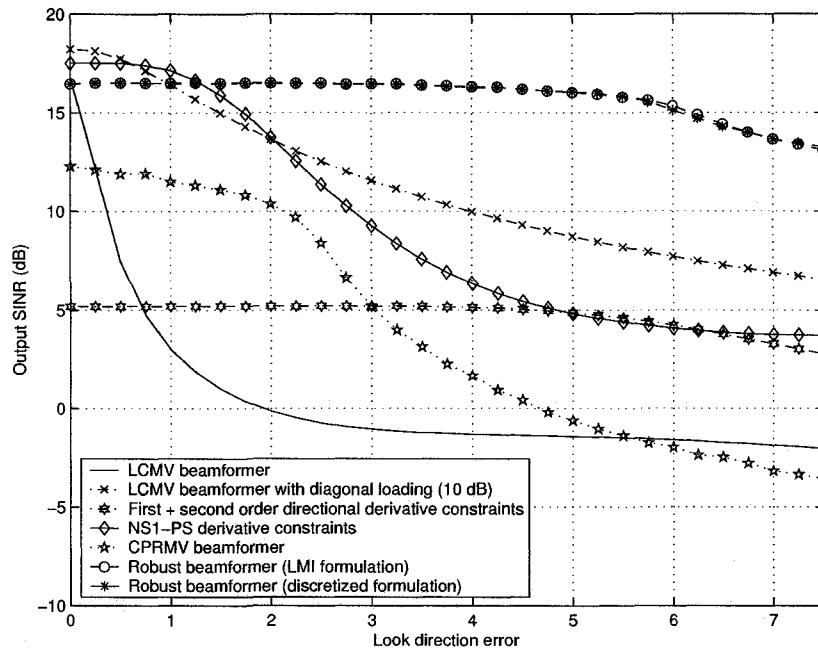
Figure 4.3: Average output SINR versus look direction error.

the look direction error. We can clearly see the improved performance of the proposed algorithm compared to all other algorithms (even those designed specifically to combat look direction errors). We can also notice a close resemblance of the two proposed implementations of our robust beamformer. For the case of zero look direction error, the decrease in the output SINR of our beamformer (due to the additional robustness constraint) compared to the NS1-PS derivative constrained beamformer is less than 1 dB. On the other hand, for the case of classical second-order directional derivative constraints the decrease in the output SINR is more than 10 dB.

Next, we investigate the effect of sensor location errors on our beamformer. We also consider a look direction error of 2° in addition to presteering delay quantization effects. The sensor location errors are selected such that they correspond to the maximum array deformation, i.e., $(\Delta z = \alpha_z c/2f_u)\zeta$. The robustness set is formed using $\Delta\theta = 2°$, $\alpha_z = 0.075$, and $\Delta T = \max T_i/256$. Figure 4.4 shows the average
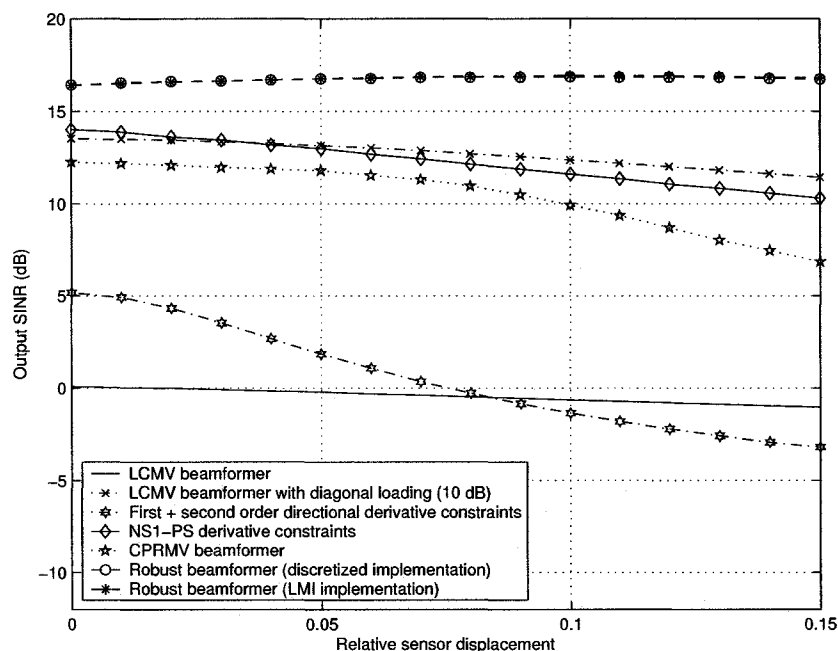
88

Figure 4.4: Average output SINR versus relative sensor location error.

output SINR versus the relative sensor displacement. We can clearly see from this figure the improvements achieved by our beamformer compared to the other tested beamformers.

Finally, we display the array response patterns for a look direction error of $4°$, sensor location errors of 3.75% with respect to the presumed inter-element spacing, and presteering delay quantization effects. Figures 4.5–4.8 show the magnitude and phase responses in the presumed look direction and the actual desired signal direction for the NS1-PS derivative constrained, the diagonally loaded LCMV, the CPRMV, and our robust beamformer, respectively. We can notice from Figure 4.8 that our robust beamformer has a nearly distortionless response towards the desired signal despite the mismatches, whereas other beamformers attenuate the high and interme-diate frequencies of the desired signal leading to a substantial decrease in the output
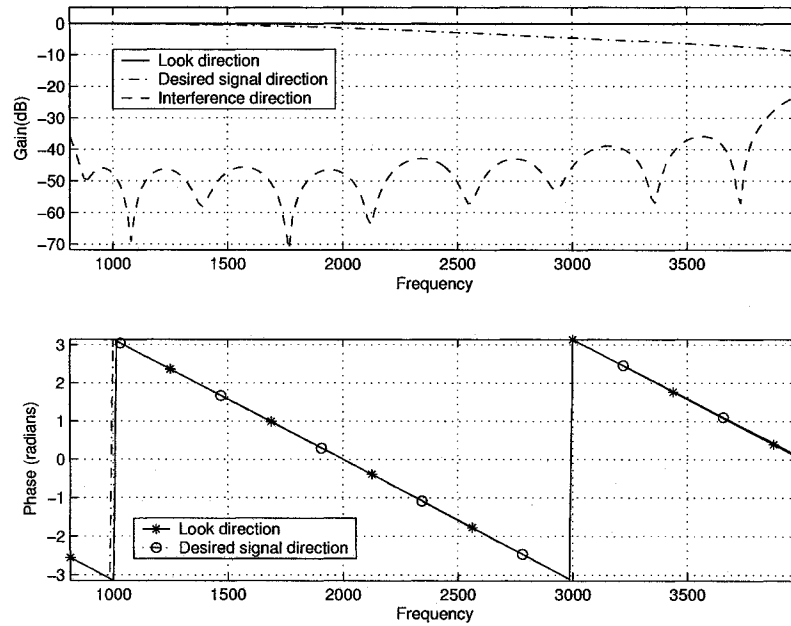
89

Figure 4.5: Frequency response of the NS1-PS derivative constrained LCMV beam-former.

SINR. It is also clear from these figures how the linear phase constraints can maintain a linear phase response towards the desired signal components in spite of the mismatches in the array manifold and the finite sample support. On the other hand, the CPRMV beamformer suffers from relatively-high phase nonlinearities. Figure 4.9 shows the power response of different beamformers versus the arrival angle. We can clearly notice an improved performance of our robust beamformer with respect to the other techniques; our beamformer yields the highest gain towards the desired signal while suppressing the interference and maintaining a relatively low sidelobe level.

4.4.2 Sensitivity to the choice of the uncertainty set

In this example, we investigate the effect of the amount of provided robustness on the performance of our beamformer. The look direction error is $4°$, worst-case sensor
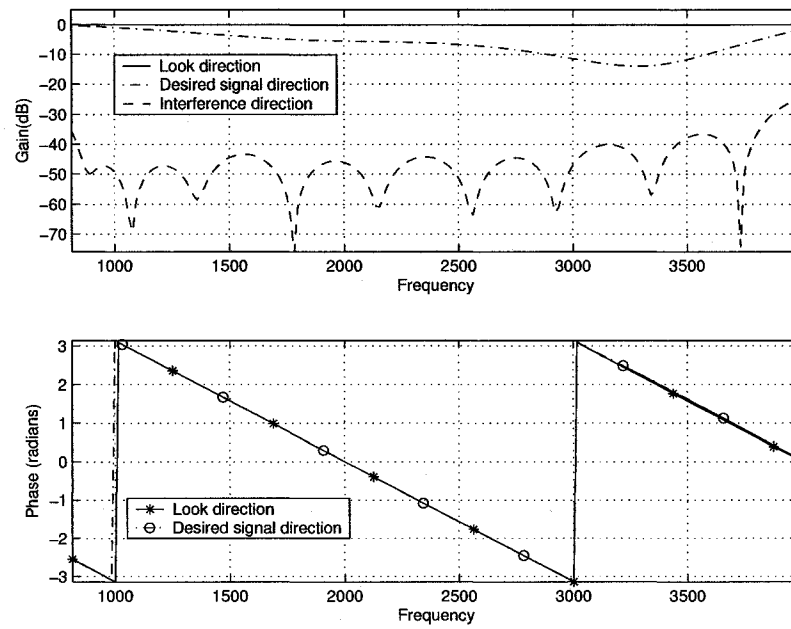
90

Figure 4.6: Frequency response of the diagonally-loaded LCMV beamformer.
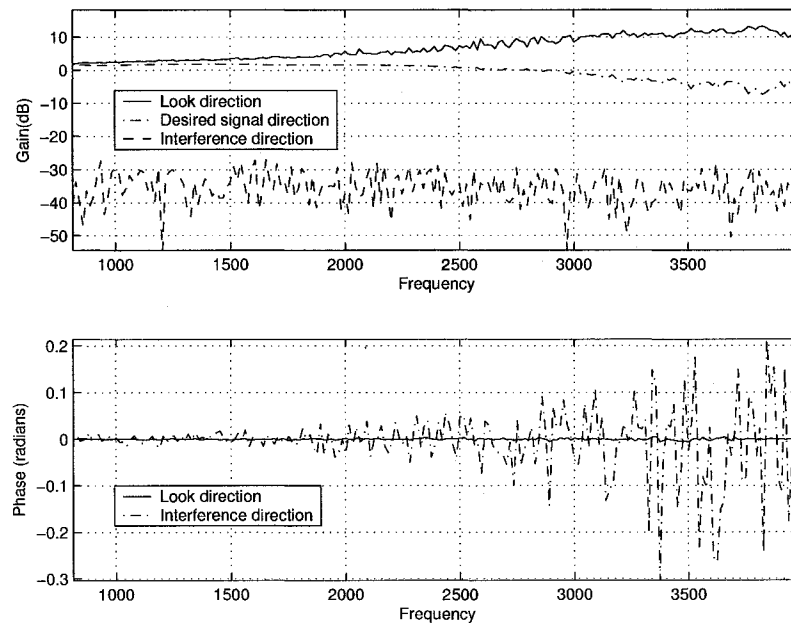

Figure 4.7: Frequency response of the CPRMV beamformer.
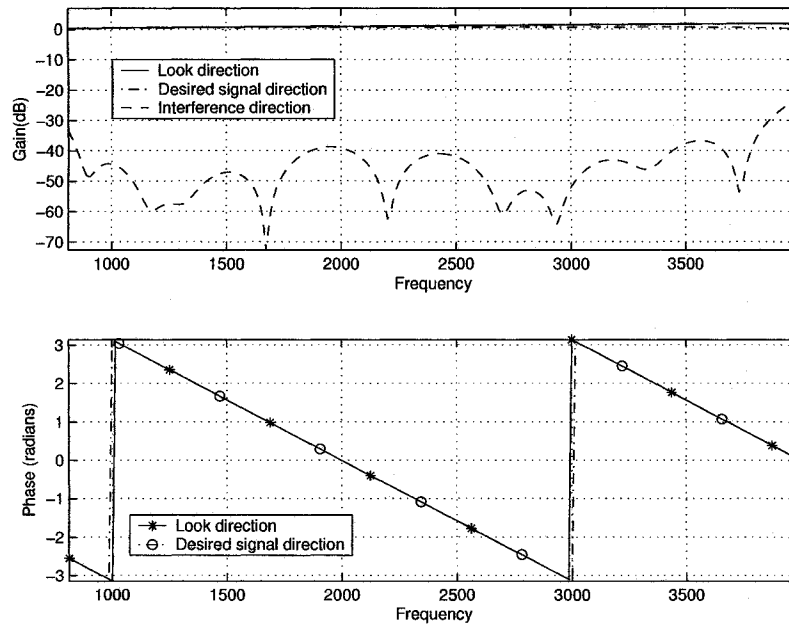
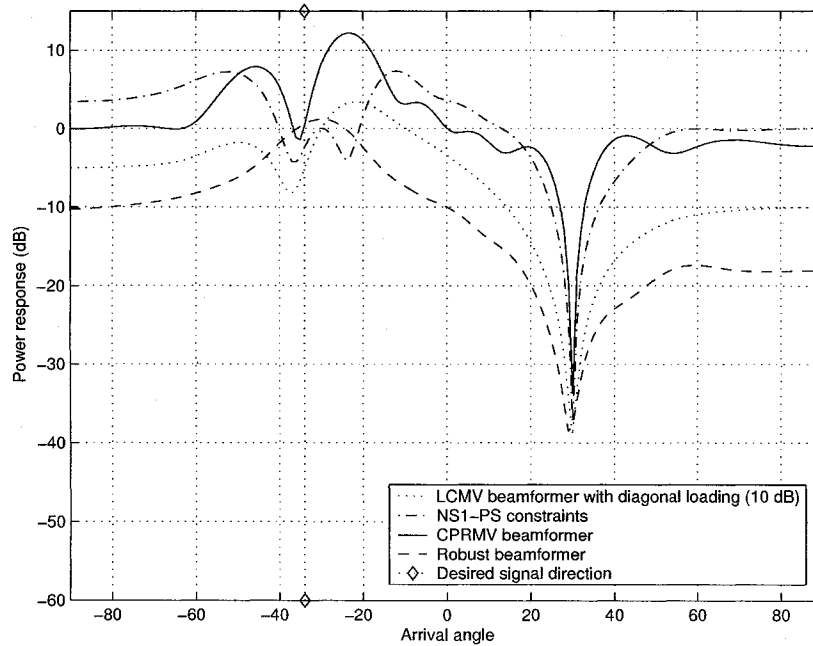Figure 4.8: Frequency response of the robust beamformer (LMI formulation).

Figure 4.9: Power response versus arrival angle.

92
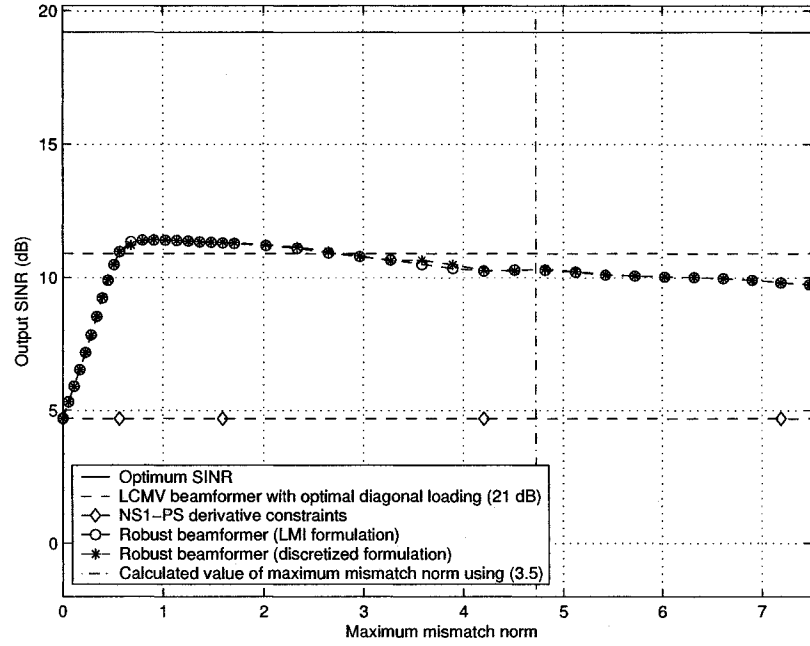
Figure 4.10: Output SINR versus maximum mismatch norm.

displacements are 7.5% of the sensor spacing, and the presteering delays are quantized to 256 levels. The data covariance matrix is computed assuming an infinite sample size. The infinite sample size covariance matrix is formed as

$$R_x = R_s + R_1 + \sigma_n^2 I \tag{4.45}$$

where the desired signal covariance matrix R_s and the interference covariance matrix R_1 are given, respectively, by

$$R_s = \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \rho_{\theta_s}(f) \Big(d(f) \otimes \big(T(f)a(f,\theta_s)\big) \Big) \Big(d(f) \otimes \big(T(f)a(f,\theta_s)\big) \Big)^H df \tag{4.46}$$

$$R_1 = \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \rho_{\theta_1}(f) \Big(d(f) \otimes \big(T(f)a(f,\theta_1)\big) \Big) \Big(d(f) \otimes \big(T(f)a(f,\theta_1)\big) \Big)^H df \tag{4.47}$$

and the integrals are evaluated through discretization over a grid of 201 points.

Figure 4.10 shows the output SINR of our proposed robust beamformer against different values of $\varepsilon(f_s/2)$ that correspond to different choices of the uncertainty set,

93

as well as the output SINRs for the ideal LCMV beamformer with perfect choice of presteering delays, the LCMV beamformer with NS1-PS derivative constraints and the diagonally loaded LCMV beamformer with an optimal choice of the diagonal loading factor. In addition, it also shows the value of $\varepsilon(f_s/2)$ computed from (4.5). Note that in practical scenarios the optimum amount of diagonal loading can not be easily determined. In this simulation, it was evaluated through multiple trials with different values of diagonal loading. We can notice from Figure 4.10 that the use of the uncertainty set corresponding to the true mismatches is rather excessive and does not yield the highest output SINR. This can be attributed to the effect of the linear phase constraints which provide additional robustness that is not taken into account in the robustness set. However, we can clearly see that the proposed beamformer is not very sensitive to the exact choice of the uncertainty set, and the over-estimation or under-estimation of the amount of mismatches does not lead to significant degradation in the output SINR of the beamformer.

4.4.3 Comparison between the proposed algorithms

In this example, we compare the performance of the two proposed beamformer implementations for different desired signal bandwidths. In particular, we investigate the effect of extending the frequency range of the spectral constraints beyond the highest frequency of the desired signal on our LMI implementation. We consider the same scenario as in the previous simulation. The highest frequency f_u is changed between 3000 and 4000 Hz while the lowest frequency of the desired signal and the sampling frequency are kept constant at 800 and 8000 Hz, respectively. For the discretized implementation of the robust beamformer a frequency grid with spacing 100 Hz that extends between f_l and f_u is used. For the LMI implementation, the spectral constraints are imposed over the frequency band $[f_l, f_s/2]$.

94
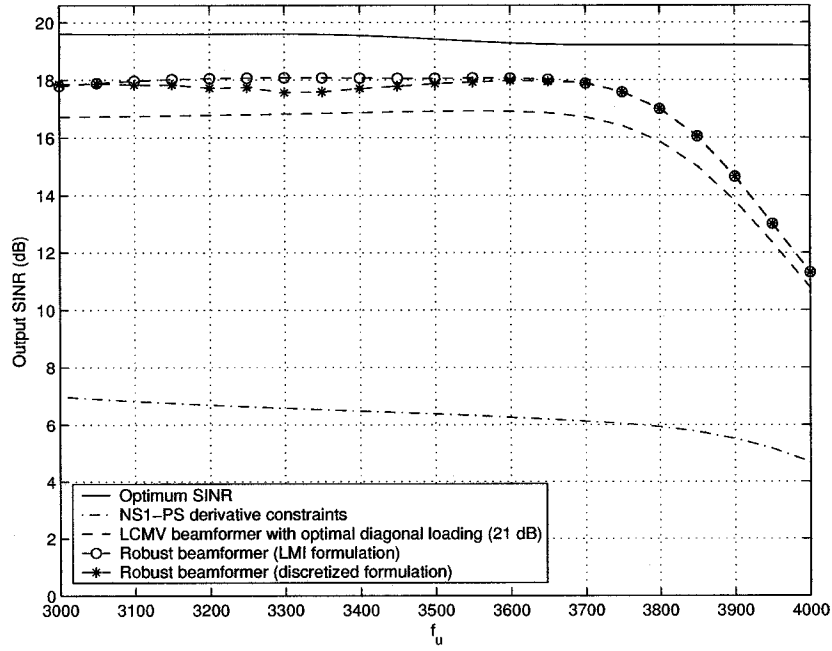
Figure 4.11: Output SINR versus the desired signal highest frequency.
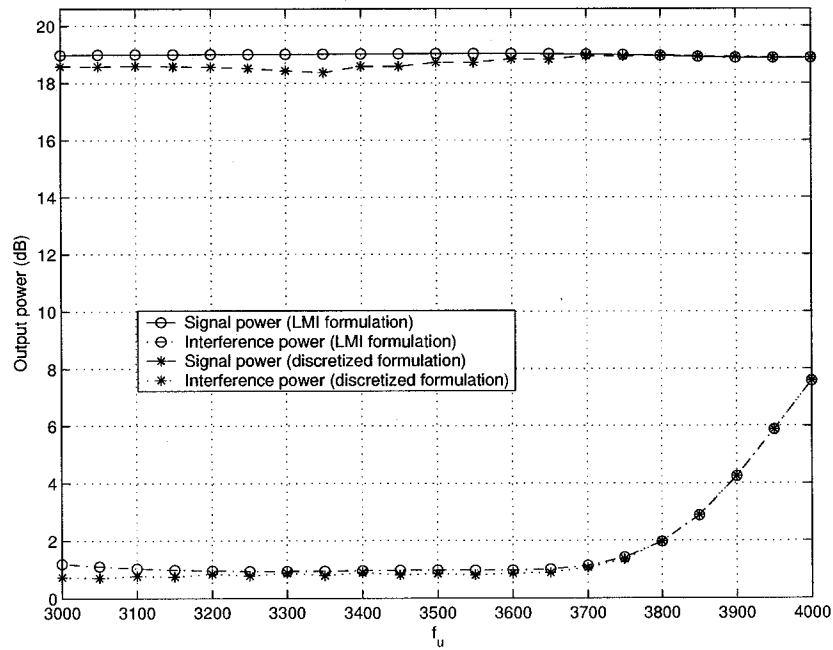
Figure 4.12: Output power versus the desired signal highest frequency.
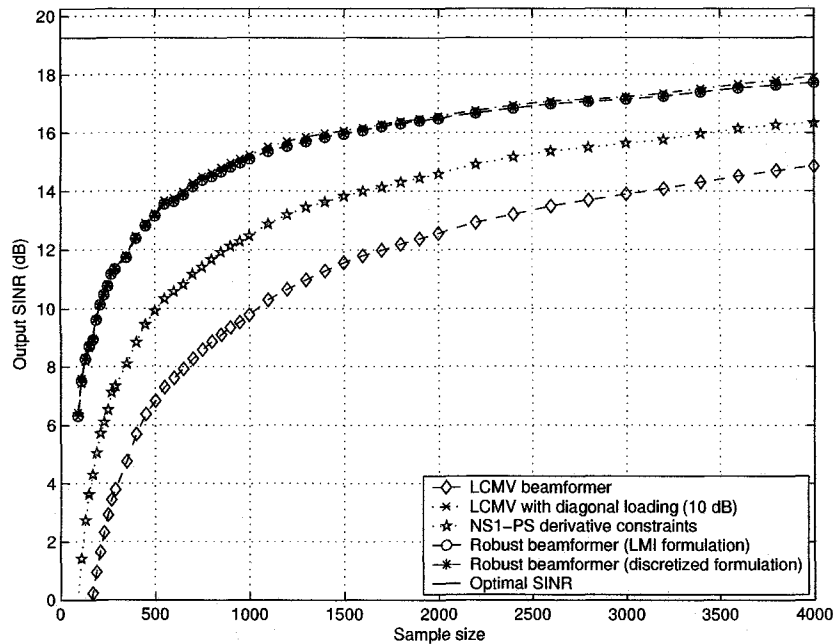
Figure 4.13: Average output SINR versus the sample size.

Figure 4.11 shows the output SINR versus f_u. We can clearly see that the performance of our two beamformer implementations is nearly identical. We can also see that this performance is not degraded by extending the spectral constraints beyond the desired signal bandwidth. This is evident from the quite close values of the output SINR of both formulations. Figure 4.12 displays the output desired signal and interference powers versus f_u. We can notice the increase in the output interference power as f_u approaches $f_s/2$. This can be attributed to the increase in the jamming signal bandwidth which increases the dimension of the interference subspace and, therefore, the beamformer can not place very deep nulls in the direction of the interference over the whole frequency band of interest [11]. On the other hand, the output desired signal power is nearly constant as the robustness constraint maintains a high gain for the desired signal.

96

4.4.4   Robustness against finite sample size

In this example, we investigate the effect of finite sample size on different robust beamforming algorithms. Perfect knowledge of the array manifold is assumed when selecting the presteering delay values and no quantization effects are considered. Therefore, the only mismatches considered are due to the finite sample size. A diagonal loading value of 10 dB is selected for the diagonally loaded LCMV beamformer, and our robustness set $\mathcal{A}_{\varepsilon,1}(f)$ is formed with $\Delta\theta = 1°$, $\alpha_z = 0$, and $\Delta T = 0$, yielding a value of $\varepsilon(f_s/2) \approx 1.2$. All simulation results are averaged over 100 Monte Carlo runs.

Figure 4.13 shows the average output SINR versus the sample size. We can clearly see that the robustness constraint used in our beamformer can combat finite sample size effects although it was originally proposed to provide robustness against uncertainties in the array manifold. This is due to the fact that errors in the choice of the presteering delays lead to errors in the data covariance matrix. This can be seen from (4.46) and (4.47). Therefore, providing robustness against presteering errors leads to robustness against finite sample support. This fact also explains the robustness of the NS1-PS derivative constrained LCMV beamformer against finite sample size as compared to the LCMV algorithm with gain-only constraints.

4.4.5   Sensitivity to the desired signal power

We investigate the performance of our beamformer for different values of the desired signal SNR for a look direction error of $4°$, worst-case sensor displacements that are 7.5% of the sensor spacing, and presteering delay quantization effects. The desired and interference signals are chosen as in Subsection 4.4.1 but with ISR equal to 0 dB. All simulation results are averaged over 100 Monte Carlo runs. Figure 4.14 shows the average output SINR for different beamformers versus the desired signal SNR. We can see from this figure the performance improvements of the proposed beamformer

97
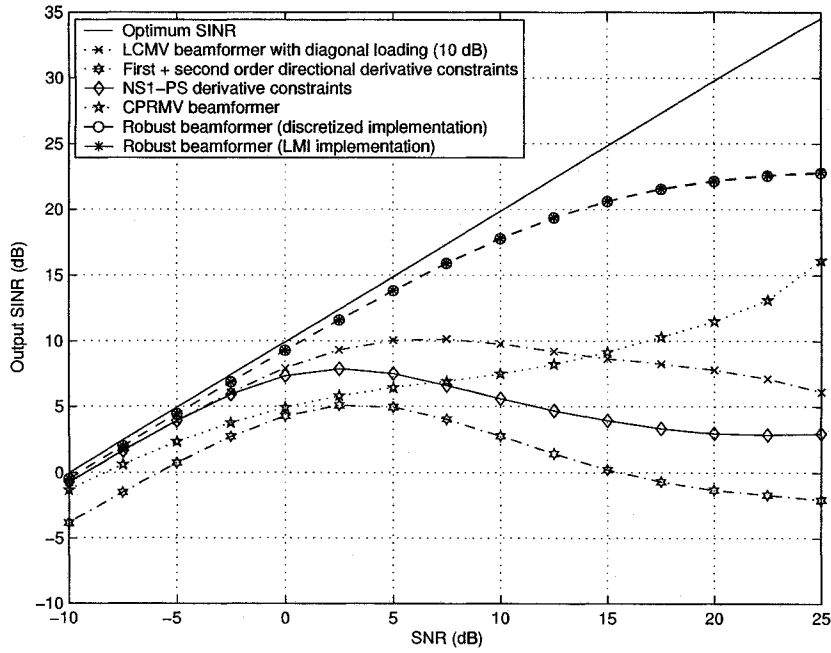
Figure 4.14: Average output SINR versus the desired signal SNR.

as compared to the other beamforming algorithms tested. These improvements are quite visible at high SNRs as the robustness constraint can effectively preserve the desired signal and prevent its cancellation despite the mismatches. At the same time, other beamformers suppress the desired signal which results in a decreased output SINR.

For the same scenario, we evaluate the normalized mean square error (NMSE) which is defined as

$$\text{NMSE} \triangleq \frac{1}{\sigma_s^2 N_s} \sum_{k=1}^{N_s} \big(\alpha_s s_0(kT_s - T_d) - y(k) \big)^2 \tag{4.48}$$

where α_s and T_d are the average beamformer gain and time delay towards the desired signal. Figure 4.15 shows the average NMSE versus the desired signal SNR for different beamformers. We can notice the effect of the phase distortions of the CPRMV beamformer on its NMSE where the increase in the output SINR in Figure 4.14 is not translated into a corresponding decrease in the NMSE. The improved performance of

98
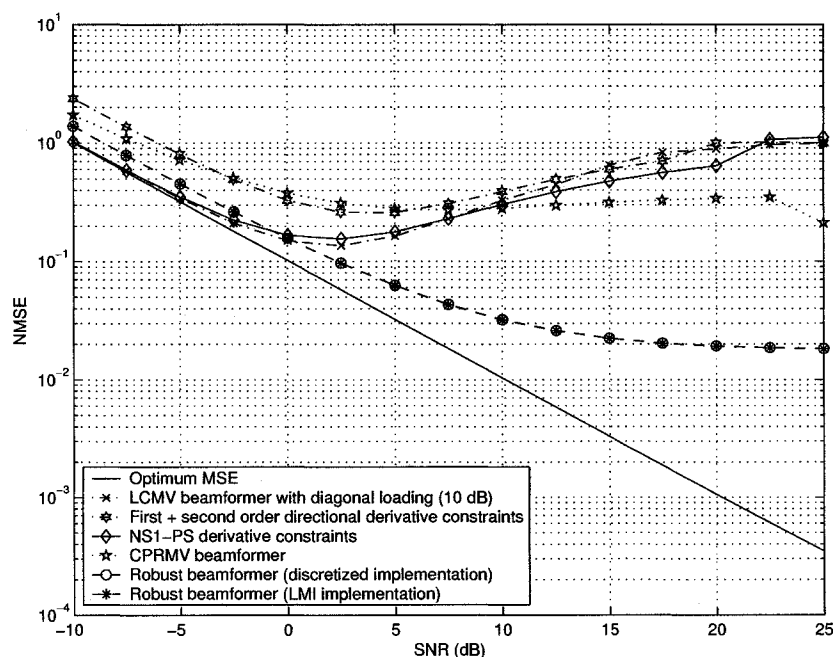
Figure 4.15: Average NMSE versus the desired signal SNR.

our beamformer in terms of its output SINR, coupled with its distortionless response and linear phase results in a superior performance in terms of the NMSE.

4.5   Conclusion

In this chapter, we have presented a novel wideband beamformer that in contrast with earlier approaches provides an amount of robustness directly linked to the amount of uncertainty in the array manifold. The problem is solved through worst-case performance optimization and is formulated as a convex optimization problem. We have provided two algorithms for our proposed beamformer; an SOCP- and an SDP-based algorithm. Both algorithms can be implemented efficiently with polynomial complexity using well-established interior point methods. Simulation results have been

presented that show an improved performance of our beamformer compared to earlier robust beamforming algorithms. Simulation results also indicate that the two proposed algorithms yield nearly the same weight vector estimate with an appropriate choice of the discretization grid used in the SOCP algorithm. This makes the SOCP algorithm more favourable due to its lower computational complexity which is comparable to the complexity of the classical LCMV beamforming algorithm.

100

Chapter 5

Computationally Efficient Online Algorithms for Robust Wideband Beamforming

5.1 Introduction

In Chapter 4, we have presented a wideband beamformer with robustness against various array calibration errors. This beamformer is similar to the narrowband beamformers proposed in [74] and [116] in that its robustness is directly related to the amount of uncertainties in the array manifold. We have also presented two algorithms for our beamformer. The first algorithm is based on discretization of the spectral constraints imposed over the whole frequency band of the desired signal, while the second ensures that these constraints are satisfied at every frequency point at the expense of increased computational cost. We have shown through numerical examples that with a proper choice of the discretization grid, the performance of both algorithms is nearly identical. However, both algorithms are not convenient for

101

nonstationary environments where the beamformer weight vector has to be updated whenever a new snapshot is received.

In this chapter, we modify our wideband beamformer in order to derive an online algorithm suitable for nonstationary scenarios that are often encountered in radar, sonar, speech processing, and wireless communications. Our new beamformer is derived using the same guidelines as those in Chapter 4 but using a 2-norm uncertainty set to describe the phase errors in the output of the presteering delays instead of the 1-norm uncertainty set used in Chapter 4. This yields a formulation that is more convenient for adaptive implementation. Our beamformer replaces the distortionless response constraint of the LCMV algorithm with an SOC robustness constraint imposed over the whole frequency band of the desired signal. A discretization grid is used to convert the robustness constraint to a finite number of constraints. The resulting problem is an SOCP optimization problem that can be solved with polynomial complexity using interior point methods [83].

We show that the SOC robustness constraints are always active , i.e., are satisfied with equality, by the optimal solution of the beamforming problem. This enables us to develop a state-space model for the robust wideband beamforming problem similar to that we proposed in Chapter 3, where the discretized robustness constraints are incorporated within the measurement equation of the model. Due to the nonlinearity of these constraints, a first- or second-order EKF is used to solve for the optimal weight vector iteratively with reduced computational cost compared to earlier interior point methods-based algorithms of the robust beamformer. Simulation results are presented illustrating an improved performance of the proposed beamformer and its advantages over earlier methods in nonstationary environments.

102

5.2 Alternative Formulation of the Robust Wideband Beamforming Problem

5.2.1 Derivation

Instead of using the uncertainty set $\mathcal{A}_{\varepsilon,1}(f)$ defined in (4.2) to describe the phase errors in the presteered desired signal components, we define another uncertainty set

$$\mathcal{A}_{\varepsilon,2}(f) \triangleq \left\{ \Delta(f) \in \mathbb{C}^M \mid \|\Delta(f)\| \leq \varepsilon(f) \right\}. \tag{5.1}$$

We impose a gain constraint over the whole frequency band of the desired signal for all the presteered data vectors with phase errors belonging to the set $\mathcal{A}_{\varepsilon,2}(f)$. This constraint can be written as

$$|H(f,\theta)| \geq 1 \qquad \forall \Delta(f) \in \mathcal{A}_{\varepsilon,2}(f); f \in [f_l, f_u]. \tag{5.2}$$

The optimal weight vector of the robust wideband beamformer can be found by minimizing the array output power subject to the set of constraints in (5.2). We use the worst-case performance optimization approach similar to that we used in (4.7). The minimum value of $|H(f,\theta)|$ over the mismatch set $\mathcal{A}_{\varepsilon,2}(f)$ can be obtained by noticing that

$$
\begin{aligned}
|H(f,\theta)| &= |\boldsymbol{w}^T(\boldsymbol{d}(f) \otimes e^{j2\pi f\varsigma}\boldsymbol{1}_M) + \boldsymbol{w}^T(\boldsymbol{d}(f) \otimes \Delta(f))| \\
&= |e^{j2\pi f\varsigma}\boldsymbol{w}^T\boldsymbol{C}_0\boldsymbol{d}(f) + \boldsymbol{w}^T\boldsymbol{Q}(f)\Delta(f)| \\
&\geq |\boldsymbol{w}^T\boldsymbol{C}_0\boldsymbol{d}(f)| - |\boldsymbol{w}^T\boldsymbol{Q}(f)\Delta(f)| \qquad (5.3) \\
&\geq |\boldsymbol{w}^T\boldsymbol{C}_0\boldsymbol{d}(f)| - \varepsilon(f)\|\boldsymbol{Q}^T(f)\boldsymbol{w}\| \qquad (5.4)
\end{aligned}
$$

where $\boldsymbol{Q}(f)$ is defined in (4.9), and (5.3) and (5.4) follow from the triangle inequality and Schwartz inequality, respectively [51]. The worst-case mismatch vector which

satisfies (5.3) and (5.4) with equality is given by

$$\boldsymbol{\Delta}_{\text{worst}}(f) = -\varepsilon(f)\frac{\boldsymbol{Q}^T(f)\boldsymbol{w}}{\|\boldsymbol{Q}^T(f)\boldsymbol{w}\|}e^{j(\phi_2 - \psi_2)} \tag{5.5}$$

where $\phi_2(f) = \text{angle}\left\{\boldsymbol{w}^T\boldsymbol{C}_0\boldsymbol{d}(f)\right\}$ and $\psi_2(f) = \text{angle}\left\{\boldsymbol{Q}^T(f)\boldsymbol{w}\right\}$.

Therefore, we can write the robust wideband beamforming problem as

$$\min_{\boldsymbol{w}} \quad \boldsymbol{w}^T\boldsymbol{R}_x\boldsymbol{w}$$
$$\text{s.t.} \quad \left|\boldsymbol{w}^T\boldsymbol{C}_0\boldsymbol{d}(f)\right| - \varepsilon(f)\left\|\boldsymbol{Q}^T(f)\boldsymbol{w}\right\| \geq 1 \quad \forall f \in [f_l, f_u]. \tag{5.6}$$

Next, we impose the linear phase constraint of (2.68) on each of the M FIR filters of the array processor. Using this constraint and (4.16), we can write

$$\left\|\boldsymbol{Q}^T(f)\boldsymbol{w}\right\| = \left\|\boldsymbol{\Lambda}^T(f)\boldsymbol{C}^T\boldsymbol{w}\right\| \tag{5.7}$$

where $\boldsymbol{\Lambda}(f) = \boldsymbol{I}_M \otimes (\boldsymbol{B}\boldsymbol{g}(f)) \in \mathbb{R}^{ML \times M}$, and $\boldsymbol{g}(f)$ and $\boldsymbol{B}$ are defined in (4.19) and (4.21), respectively. The matrix $\boldsymbol{C} \in \mathbb{R}^{ML \times ML}$ is defined as

$$\boldsymbol{C} \triangleq [\boldsymbol{C}_1, \boldsymbol{C}_2, \ldots, \boldsymbol{C}_M] \tag{5.8}$$

where the set of matrices $\{\boldsymbol{C}_m\}_{m=1}^M$ is defined in (4.20).

Therefore, using (4.23) and (5.7), we can rewrite the robust beamforming problem in (5.6) with the additional linear phase constraints as

$$\min_{\boldsymbol{w}} \quad \boldsymbol{w}^T\boldsymbol{R}_x\boldsymbol{w}$$
$$\text{s.t.} \quad \left(\left(\boldsymbol{e}_{L_c}^T(\boldsymbol{J}_l - \boldsymbol{J}_l^T)\right) \otimes \left(\boldsymbol{1}_M^T\boldsymbol{\Omega}_m\right)\right)\boldsymbol{w} = 0 \quad \forall m = 1, \ldots, M; l = 1, \ldots, L_c - 1$$
$$\left|\boldsymbol{g}^T(f)\boldsymbol{B}^T\boldsymbol{C}_0^T\boldsymbol{w}\right| - \varepsilon(f)\left\|\boldsymbol{\Lambda}^T(f)\boldsymbol{C}^T\boldsymbol{w}\right\| \geq 1 \quad \forall f \in [f_l, f_u] \tag{5.9}$$

The problem is still nonconvex due to the absolute value operator in the last constraint. Following the same argument that is used in Chapter 4, we can see that if $\boldsymbol{w}$ is an optimal solution of the above optimization problem, then $-\boldsymbol{w}$ is also an

104

optimal solution. Also, examining the second constraint in (5.9), $\left|\sum_{m=1}^{M} \mathcal{W}_m(f)\right| \geq$
$1 + \varepsilon(f) \left\| \boldsymbol{\Lambda}^T(f) \boldsymbol{C}^T \boldsymbol{w} \right\|$, and from the continuity of the Fourier transform, it is clear
that for a feasible weight vector $\boldsymbol{w}$, $\boldsymbol{g}^T(f) \boldsymbol{B}^T \boldsymbol{C}_0^T \boldsymbol{w}$ can either be positive or negative
over the whole frequency band $[f_l, f_u]$ and can not change its sign at any frequency
point. From these two remarks, replacing $\left| \boldsymbol{g}^T(f) \boldsymbol{B}^T \boldsymbol{C}_0^T \boldsymbol{w} \right|$ in the last constraint of
(5.9) by $\boldsymbol{g}^T(f) \boldsymbol{B}^T \boldsymbol{C}_0^T \boldsymbol{w}$ does not cause any loss of optimality. The resulting problem
is convex and is given by

$$\min_{\boldsymbol{w}} \quad \boldsymbol{w}^T \boldsymbol{R}_x \boldsymbol{w}$$

$$\text{s.t.} \quad \left(\left(\boldsymbol{e}_{L_c}^T (\boldsymbol{J}_l - \boldsymbol{J}_l^T) \right) \otimes \left(\mathbf{1}_M^T \boldsymbol{\Omega}_m \right) \right) \boldsymbol{w} = 0 \qquad \forall m = 1, \ldots, M; l = 1, \ldots, L_c - 1$$

$$\boldsymbol{g}^T(f) \boldsymbol{B}^T \boldsymbol{C}_0^T \boldsymbol{w} - \varepsilon(f) \left\| \boldsymbol{\Lambda}^T(f) \boldsymbol{C}^T \boldsymbol{w} \right\| \geq 1 \qquad \forall f \in [f_l, f_u] \qquad (5.10)$$

Theorem 5.1 *In the case of uncorrelated desired and interference signals, the robustness constraint is satisfied with equality at the optimal point of (5.10) for every frequency $f \in [f_l, f_u]$.*

Proof. See Appendix A.

5.2.2 SOCP-based algorithm

In this section, we present a batch implementation for the robust beamforming problem similar to the implementations presented in Section 4.3. First, we note that the robustness constraint in (5.10) is imposed over an infinite number of frequency points in the frequency band of the desired signal. This constraint can be represented in a finite manner through discretization over a uniformly spaced grid $\{f_i\}_{i=1}^{N}$ over the range of the continuous parameter [80]. We have shown in Section 4.3 that with a good choice of the discretization grid we can obtain a performance similar to that obtained when the constraints are enforced at every frequency point. The discretized

robust beamforming problem can therefore be written as

$$
\begin{aligned}
\min_{\boldsymbol{w}} \quad & \boldsymbol{w}^T \boldsymbol{R}_x \boldsymbol{w} \\
\text{s.t.} \quad & \left(\left(\boldsymbol{e}_{L_c}^T (\boldsymbol{J}_l - \boldsymbol{J}_l^T) \right) \otimes \left(\boldsymbol{1}_M^T \boldsymbol{\Omega}_m \right) \right) \boldsymbol{w} = 0 \qquad \forall m = 1, \dots, M; l = 1, \dots, L_c - 1 \\
& \varepsilon(f_i) \| \boldsymbol{\Lambda}^T(f_i) \boldsymbol{C}^T \boldsymbol{w} \| \leq \boldsymbol{g}^T(f_i) \boldsymbol{B}^T \boldsymbol{C}_0^T \boldsymbol{w} - 1 - \delta \qquad \forall i = 1, \dots, N
\end{aligned}
\tag{5.11}
$$

Using the Cholesky factorization of the covariance matrix in (4.26) and introducing the nonnegative scalar variable t, we can convert (5.11) into the following problem

$$
\begin{aligned}
\min_{\boldsymbol{w},t} \quad & t \\
\text{s.t.} \quad & \| \boldsymbol{U} \boldsymbol{w} \| \leq t \\
& \left(\left(\boldsymbol{e}_{L_c}^T (\boldsymbol{J}_l - \boldsymbol{J}_l^T) \right) \otimes \left(\boldsymbol{1}_M^T \boldsymbol{\Omega}_m \right) \right) \boldsymbol{w} = 0 \qquad \forall m = 1, \dots, M; l = 1, \dots, L_c - 1 \\
& \varepsilon(f_i) \| \boldsymbol{\Lambda}^T(f_i) \boldsymbol{C}^T \boldsymbol{w} \| \leq \boldsymbol{g}^T(f_i) \boldsymbol{B}^T \boldsymbol{C}_0^T \boldsymbol{w} - 1 - \delta \qquad \forall i = 1, \dots, N
\end{aligned}
\tag{5.12}
$$

The discretized problem in (5.12) can be solved as an SOCP convex optimization problem [102]. The number of optimization variables is equal to $ML + 1$. The first constraint is an $(ML+1)$–dimensional SOC constraint, followed by $M(L-1)/2$ linear equality constraints and N SOC constraints, each of dimension $M + 1$. The computational complexity of (5.12) is, therefore, of $\mathcal{O}\left(M^3 L^2 N^{\frac{1}{2}} \zeta \right)$ where $\zeta = \max(N, L)$.

5.3 State-Space Model for Robust Wideband Beamforming

The main disadvantage of all the earlier robust beamforming techniques presented in (4.29), (4.44), and (5.12) is that they are not iterative. Therefore, whenever a new sample is received, the sample covariance matrix has to be updated and the whole optimization procedure is repeated to estimate the new weight vector. This represents a major drawback in nonstationary environments where the weight vector has to be

repeatedly updated with low computational cost. In this section, we derive an online implementation of the robust wideband beamformer based on state-space modelling of the optimization problem in (5.11).

We start by eliminating the linear phase equality constraints in (5.11). We define the matrix $D \in \mathbb{R}^{ML \times ML_c}$ as

$$D = B \otimes I_M. \tag{5.13}$$

Any vector $w \in \mathbb{R}^{ML \times 1}$ satisfying the linear phase constraints in (5.10) can be expressed as $w = Dv$ where $v \in \mathbb{R}^{ML_c \times 1}$ contains the last ML_c elements of the weight vector w that can be chosen independently while satisfying the linear phase constraints. Using Theorem 5.1, we can write the robust beamforming problem as

$$\min_{v} \quad v^T D^T R_x D v$$
$$\text{s.t.} \quad h_i(v) = 1 + \delta \quad \forall i = 1, \ldots, N \tag{5.14}$$

where

$$h_i(v) = g^T(f_i) B^T C_0^T D v - \varepsilon(f_i) \| \Lambda^T(f_i) C^T D v \| \tag{5.15}$$

$$= r^T(f_i) v - \varepsilon(f_i) \| \tilde{\Lambda}^T(f_i) v \|, \tag{5.16}$$

$r(f_i) \in \mathbb{R}^{ML_c \times 1} = \tilde{g}(f_i) \otimes 1_M$, $\tilde{g}(f_i) = [1, 2\cos(2\pi f_i T), \ldots, 2\cos(2\pi f_i (L_c - 1)T)]^T$, and $\tilde{\Lambda}(f_i) = \tilde{g}(f_i) \otimes I_M$ (see Appendix B).

A state-space model describing the above optimization problem is given by the following process equation

$$v(k + 1) = v(k) + n_v(k) \tag{5.17}$$

where $n_v(k)$ is the process noise that allows tracking the optimal solution in non-stationary environments, and is assumed to be white Gaussian with zero mean and covariance matrix $Q_v = \sigma_v^2 I_{ML_c}$.

107

The associated measurement equation is given by

$$
\begin{bmatrix} 0 \\ 1+\delta \\ \vdots \\ 1+\delta \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}^T(k)\boldsymbol{D}\boldsymbol{v}(k) \\ h_1(\boldsymbol{v}(k)) \\ \vdots \\ h_N(\boldsymbol{v}(k)) \end{bmatrix} + \begin{bmatrix} n_{m,0}(k) \\ n_{m,1}(k) \\ \vdots \\ n_{m,N}(k) \end{bmatrix}
\tag{5.18}
$$

which can be written in matrix form as

$$
\boldsymbol{z} = \boldsymbol{h}(\boldsymbol{x}(k), \boldsymbol{v}(k)) + \boldsymbol{n}_m(k)
\tag{5.19}
$$

where $n_{m,0}(k)$ is the negative output of the beamformer and $\{n_{m,i}(k)\}_{i=1}^N$ represent the errors in satisfying the discretized robustness constraint. They can be modelled as zero mean independent white Gaussian sequences with covariance matrix $\boldsymbol{Q}_m = \mathrm{diag}\{\sigma_{m,0}^2, \sigma_{m,1}^2, \ldots, \sigma_{m,N}^2\}$.

Running a state estimator on the model given by (5.17) and (5.19) yields an estimate of $\boldsymbol{v}(k)$ that minimizes the uncertainties due to the process and measurement noises. Minimizing the mean square value of $n_{m,0}(k)$ is equivalent to minimizing the output power of the beamformer, while minimizing the mean square values of $\{n_{m,i}(k)\}_{i=1}^N$ minimizes the MSE in satisfying the robustness constraint. Thus, the resulting weight vector estimate $\hat{\boldsymbol{v}}(k)$ solves the optimization problem in (5.11) with the covariance matrix $\boldsymbol{R}_x$ replaced by the k-snapshot sample covariance matrix.

Due to the nonlinearity of the measurement equation, the optimal state estimator has to compute and iteratively update the probability density function (pdf) of the state vector $p\left(\boldsymbol{v}(k) \,\middle|\, \{\boldsymbol{x}(i)\}_{i=0}^k\right)$. If the initial pdf $p\left(\boldsymbol{v}(0) \,\middle|\, \boldsymbol{x}(0)\right)$ is available, then $p\left(\boldsymbol{v}(k) \,\middle|\, \{\boldsymbol{x}(i)\}_{i=0}^k\right)$ can be obtained recursively in two stages: prediction and update [5]. Thus, given $p\left(\boldsymbol{v}(k-1) \,\middle|\, \{\boldsymbol{x}(i)\}_{i=0}^{k-1}\right)$, the prediction step propagates the pdf one step in time using the state equation to estimate $p\left(\boldsymbol{v}(k) \,\middle|\, \{\boldsymbol{x}(i)\}_{i=0}^{k-1}\right)$. The update step corrects the predicted pdf using the newest measurement $\boldsymbol{x}(k)$ to yield $p\left(\boldsymbol{v}(k) \,\middle|\, \{\boldsymbol{x}(i)\}_{i=0}^k\right)$.

5.3.1   EKF-based algorithm

The optimal estimator is merely a conceptual solution and can not be evaluated analytically in our case, and therefore, we have to resort to a suboptimal approach. A recursive estimate $\hat{\boldsymbol{v}}(k)$ of the state vector can be obtained using the EKF [6, p. 381]. The EKF is based on local linearization of the measurement equation around the current estimate using the first-order term in the Taylor expansion of the measurement equation. Thus, the measurement equation can be approximated as

$$\boldsymbol{z} \approx \boldsymbol{h}(\boldsymbol{x}(k), \hat{\boldsymbol{v}}(k-1)) + \boldsymbol{H}_v(\hat{\boldsymbol{v}}(k-1))\big(\boldsymbol{v}(k) - \hat{\boldsymbol{v}}(k-1)\big) + \boldsymbol{n}_m(k) \tag{5.20}$$

where $\boldsymbol{H}_v(\hat{\boldsymbol{v}}(k-1))$ is the Jacobian of the vector $\boldsymbol{h}(\boldsymbol{x}(k), \boldsymbol{v}(k))$ evaluated at the latest estimate of the state. It is given by

$$\boldsymbol{H}_v(\hat{\boldsymbol{v}}(k-1)) = \begin{bmatrix} \boldsymbol{x}^T(k)\boldsymbol{D} \\ \boldsymbol{r}^T(f_1) - \varepsilon(f_1)\frac{\hat{\boldsymbol{v}}^T(k-1)\tilde{\boldsymbol{\Lambda}}(f_1)\tilde{\boldsymbol{\Lambda}}^T(f_1)}{\left\|\tilde{\boldsymbol{\Lambda}}^T(f_1)\hat{\boldsymbol{v}}(k-1)\right\|} \\ \vdots \\ \boldsymbol{r}^T(f_N) - \varepsilon(f_N)\frac{\hat{\boldsymbol{v}}^T(k-1)\tilde{\boldsymbol{\Lambda}}(f_N)\tilde{\boldsymbol{\Lambda}}^T(f_N)}{\left\|\tilde{\boldsymbol{\Lambda}}^T(f_N)\hat{\boldsymbol{v}}(k-1)\right\|} \end{bmatrix}. \tag{5.21}$$

Starting with an initial random weight vector estimate $\hat{\boldsymbol{v}}(0)$ and its associated initial covariance $\boldsymbol{P}(0|0)$, the estimate is updated recursively through

$$\hat{\boldsymbol{v}}(k) = \hat{\boldsymbol{v}}(k-1) + \boldsymbol{G}(k)\big(\boldsymbol{z} - \hat{\boldsymbol{z}}(k|k-1)\big) \tag{5.22}$$

where the predicted measurement and the filter gain are given, respectively, by

$$\hat{\boldsymbol{z}}(k|k-1) = \boldsymbol{h}(\boldsymbol{x}(k), \hat{\boldsymbol{v}}(k-1)) \tag{5.23}$$

$$\boldsymbol{G}(k) = \boldsymbol{P}(k|k-1)\boldsymbol{H}_v^T(\hat{\boldsymbol{v}}(k-1))\boldsymbol{S}^{-1}(k). \tag{5.24}$$

Here, the innovation covariance matrix $\boldsymbol{S}(k)$ and the predicted weight vector covariance matrix $\boldsymbol{P}(k|k-1)$ are given by

$$\boldsymbol{S}(k) = \boldsymbol{H}_v(\hat{\boldsymbol{v}}(k-1))\boldsymbol{P}(k|k-1)\boldsymbol{H}_v^T(\hat{\boldsymbol{v}}(k-1)) + \boldsymbol{Q}_m \tag{5.25}$$

$$\boldsymbol{P}(k|k-1) = \boldsymbol{P}(k-1|k-1) + \boldsymbol{Q}_v. \tag{5.26}$$

The updated weight vector covariance can be expressed as

$$\boldsymbol{P}(k|k) \;=\; \boldsymbol{P}(k|k-1) - \boldsymbol{G}(k)\boldsymbol{S}(k)\boldsymbol{G}^T(k). \tag{5.27}$$

Typically the value of N is much lower than ML_c and, thus, the computational complexity of one iteration of the EKF-based beamformer is of $\mathcal{O}\left(M^2 L_c^2 N\right)$ which is much lower than the complexity of the SOCP formulation in (5.12). This makes it more suitable for nonstationary scenarios.

The NIS of the EKF is defined as

$$\epsilon(k) = \left(\boldsymbol{z} - \hat{\boldsymbol{z}}(k|k-1)\right)^T \boldsymbol{S}^{-1}(k)\left(\boldsymbol{z} - \hat{\boldsymbol{z}}(k|k-1)\right). \tag{5.28}$$

Under the linearized approximation of the measurement equation in (5.20) and the Gaussian assumptions of the model, the NIS is chi-square distributed with $N+1$ degrees of freedom. The consistency of the beamformer can be checked through a chi-square test on the NIS [6, p. 236]. An exponential window can be used to evaluate the NIS of the beamformer as

$$\epsilon_\alpha(k) = \alpha\epsilon_\alpha(k-1) + \left(\boldsymbol{z} - \hat{\boldsymbol{z}}(k|k-1)\right)^H \boldsymbol{S}^{-1}(k)\left(\boldsymbol{z} - \hat{\boldsymbol{z}}(k|k-1)\right) \tag{5.29}$$

where the effective window length is given by $\frac{1}{1-\alpha}$. Therefore, under the above mentioned linear Gaussian assumptions, $\epsilon_\alpha(k)$ can be approximated using first-order moment matching as chi-square distributed with $\lceil \frac{N+1}{1-\alpha} \rceil$ degrees of freedom.

A random weight vector can be used to initialize the adaptive algorithm. The associated covariance $\boldsymbol{P}(1|0)$ can be selected as $\boldsymbol{P}(1|0) = \beta_0\boldsymbol{I}$ with β_0 chosen so that the NIS of the first iteration is close to the expected value of the chi-square distribution with $N+1$ degrees of freedom. Therefore, ignoring the measurement noise covariance, we can write

$$\beta_0 = \frac{1}{N+1}\left(\boldsymbol{z} - \hat{\boldsymbol{z}}(1|0)\right)^T \left(\boldsymbol{H}_v\big(\boldsymbol{x}(1), \hat{\boldsymbol{v}}(0)\big)\boldsymbol{H}_v^T\big(\boldsymbol{x}(1), \hat{\boldsymbol{v}}(0)\big)\right)^{-1}\left(\boldsymbol{z} - \hat{\boldsymbol{z}}(1|0)\right). \tag{5.30}$$

110

The parameters of the Kalman filter can be chosen as follows. The parameter σ_v^2 which controls the process noise covariance is selected according to the operating environment as discussed in Section 3.3. The parameter $\sigma_{m,0}^2$ should be chosen in the range of the output power of the beamformer. Assuming that the interference signals are completely nulled-out by the beamformer and that the desired signal passes without any distortion, the output power is given by $\sigma_s^2 + \sigma_n^2 \|\boldsymbol{D}\boldsymbol{v}\|^2$. However, it will be demonstrated through simulations that the filter is not very sensitive to the exact choice of $\sigma_{m,0}^2$, and that a satisfactory performance can be obtained within a wide range of selection of this parameter.

On the other hand, the values of $\{\sigma_{m,i}^2\}_{i=1}^N$ determine the MSE in satisfying the discretized robustness constraint. Note that treating the robustness constraints as perfect observations, i.e., $\{\sigma_{m,i}^2 = 0\}_{i=1}^N$, can lead to convergence problems [39]. This can be attributed to the errors induced by the truncation of the Taylor series expansion after the first term. Moreover, the Jacobian in (5.20) is evaluated at the estimated value of the state and not at its true value which results in additional errors. Therefore, using very small values for $\{\sigma_{m,i}^2\}_{i=1}^N$ might prevent the estimate from moving away from a badly linearized constraint surface resulting in convergence problems. Simulation results indicate that a typical choice of $\{\sigma_{m,i}^2\}_{i=1}^N$ in the order of 10^{-3} ensures that the constraints are satisfied with enough accuracy without causing numerical problems.

5.3.2　Second-order EKF-based algorithm

A second-order EKF can be used to estimate the beamformer weight vector. This reduces the truncation error in the Taylor series expansion of the measurement equation and thus reduces the sensitivity of the algorithm to the choice of its parameters [6, p. 387]. The second-order EKF approximates the nonlinear function in the measurement equation by its second-order Taylor expansion around the latest state estimate. The

approximate measurement equation is given by

$$z \approx h(x(k), \hat{v}(k-1)) + H_v(\hat{v}(k-1))(v(k) - \hat{v}(k-1))$$

$$+ \frac{1}{2} \sum_{n=0}^{N} e_{n+1}(v(k) - \hat{v}(k-1))^T H_{vv}^{(n)}(\hat{v}(k-1))(v(k) - \hat{v}(k-1)) + n_m(k) \quad (5.31)$$

where $\left\{ H_{vv}^{(n)}(\hat{v}(k-1)) \right\}_{n=0}^{N}$ are the Hessian matrices of the $N+1$ components of the vector $h(x(k), v(k))$ computed at the latest estimate $\hat{v}(k-1)$. They are given by

$$H_{vv}^{(0)} = 0 \quad (5.32)$$

$$H_{vv}^{(n)}(\hat{v}(k-1)) = \varepsilon(f_n) \frac{\tilde{\Lambda}(f_n)\tilde{\Lambda}^T(f_n)\hat{v}(k-1)\hat{v}^T(k-1)\tilde{\Lambda}(f_n)\tilde{\Lambda}^T(f_n)}{\|\tilde{\Lambda}^T(f_n)\hat{v}(k-1)\|^3}$$

$$- \varepsilon(f_n) \frac{\tilde{\Lambda}(f_n)\tilde{\Lambda}^T(f_n)}{\|\tilde{\Lambda}^T(f_n)\hat{v}(k-1)\|} \qquad \forall n = 1, \ldots, N \quad (5.33)$$

The second-order EKF implementation of the robust beamformer is identical to the EKF algorithm in (5.22)–(5.27) except for two modifications where the second-order term of the Taylor expansion of the measurement vector appears in the computation of the predicted measurement and the innovation covariance matrix. Thus, equation (5.23) is replaced by

$$\hat{z}(k|k-1) = \begin{bmatrix} x^T(k)D\hat{v}(k-1) \\ h_1(\hat{v}(k-1)) + 0.5 \operatorname{tr}\left\{ H_{vv}^{(1)}(\hat{v}(k-1))P(k|k-1) \right\} \\ \vdots \\ h_N(\hat{v}(k-1)) + 0.5 \operatorname{tr}\left\{ H_{vv}^{(n)}(\hat{v}(k-1))P(k|k-1) \right\} \end{bmatrix} \quad (5.34)$$

where $\operatorname{tr}\left\{ H_{vv}^{(i)}(\hat{v}(k-1))P(k|k-1) \right\}$ can be simplified as

$$\operatorname{tr}\left\{ H_{vv}^{(i)}(\hat{v}(k-1))P(k|k-1) \right\}$$

$$= \frac{\varepsilon(f_i)}{\|\tilde{\Lambda}^T(f_i)v(k-1)\|^3} \left(v^T(k-1)\tilde{\Lambda}(f_i)\tilde{\Lambda}^T(f_i)P(k|k-1)\tilde{\Lambda}(f_i)\tilde{\Lambda}^T(f_i)v(k) \right.$$

$$\left. - \|\tilde{\Lambda}^T(f_i)v(k-1)\|^2 \operatorname{tr}\left\{ \tilde{\Lambda}^T(f_i)P(k|k-1)\tilde{\Lambda}(f_i) \right\} \right). \quad (5.35)$$

112

Also, the innovation covariance matrix is computed through the following equation instead of (5.25)

$$S(k) = H_v(\hat{v}(k-1))P(k|k-1)H_v^T(\hat{v}(k-1)) + Q_m$$
$$+ \sum_{m=0}^{N}\sum_{n=0}^{N} e_{m+1}e_{n+1}^T \text{tr}\Big\{H_{vv}^{(m)}(\hat{v}(k-1))P(k|k-1)H_{vv}^{(n)}(\hat{v}(k-1))P(k|k-1)\Big\}$$

(5.36)

where $\text{tr}\{H_{vv}^{(m)}(\hat{v}(k-1))P(k|k-1)H_{vv}^{(n)}(\hat{v}(k-1))P(k|k-1)\}$ can be written as

$$\text{tr}\{H_{vv}^{(m)}(\hat{v}(k-1))P(k|k-1)H_{vv}^{(n)}(\hat{v}(k-1))P(k|k-1)\}$$
$$= \frac{\varepsilon(f_m)\varepsilon(f_n)\big(v^T(k-1)\tilde{\Lambda}(f_m)\tilde{\Lambda}^T(f_m)P(k|k-1)\tilde{\Lambda}(f_n)\tilde{\Lambda}^T(f_n)v(k-1)\big)^2}{\|\tilde{\Lambda}^T(f_m)v(k-1)\|^3\|\tilde{\Lambda}^T(f_n)v(k-1)\|^3}$$
$$+ \frac{\varepsilon(f_m)\varepsilon(f_n)\|\tilde{\Lambda}^T(f_m)P(k|k-1)\tilde{\Lambda}(f_n)\|_F}{\|\tilde{\Lambda}^T(f_m)v(k-1)\|\|\tilde{\Lambda}^T(f_n)v(k-1)\|}$$
$$+ \frac{\varepsilon(f_m)\varepsilon(f_n)\|\tilde{\Lambda}^T(f_m)P(k|k-1)\tilde{\Lambda}(f_n)\tilde{\Lambda}^T(f_n)v(k-1)\|}{\|\tilde{\Lambda}^T(f_m)v(k-1)\|\|\tilde{\Lambda}^T(f_n)v(k-1)\|^3}$$
$$+ \frac{\varepsilon(f_m)\varepsilon(f_n)\|\tilde{\Lambda}^T(f_n)P(k|k-1)\tilde{\Lambda}(f_m)\tilde{\Lambda}^T(f_m)v(k-1)\|}{\|\tilde{\Lambda}^T(f_m)v(k-1)\|^3\|\tilde{\Lambda}^T(f_n)v(k-1)\|}$$

(5.37)

Due to the sparsity of the matrices $\big\{\tilde{\Lambda}(f_i)\big\}_{i=1}^N$, the computational complexity of one iteration of the second-order EKF algorithm is of $\mathcal{O}(M^2L_c^2N^2)$. This complexity is higher than that of the EKF algorithm but still lower than that of the SOCP-based implementation in (5.12).

5.4 Simulations

We consider a broadband linear microphone array with $M = 10$ and $L = 9$. The sensors are assumed to be equi-spaced with spacing c/f_s where $f_s = 8000$ Hz. The actual sensor locations are selected such that they correspond to the maximum deformation
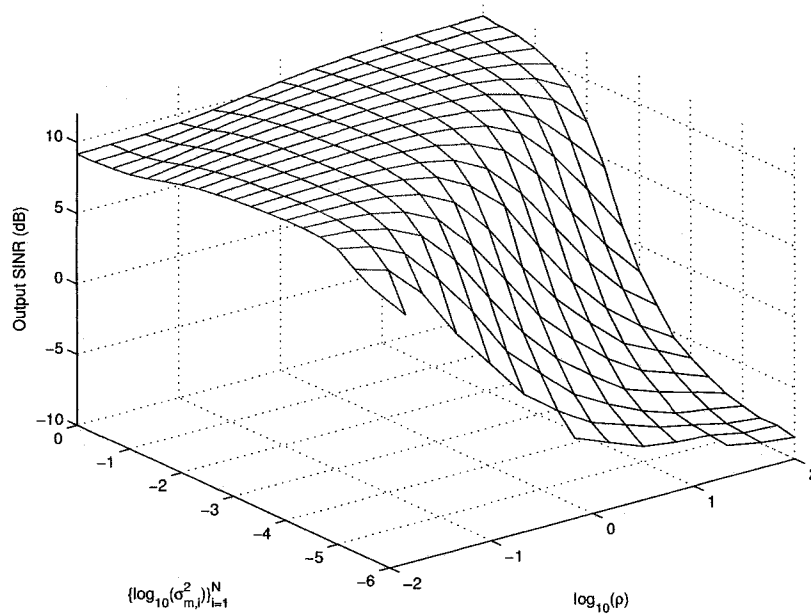
113

Figure 5.1: Average output SINR versus ρ and $\{\sigma_{m,i}^2\}_{i=1}^N$ (EKF algorithm).

in the array geometry along the array line. The sensor displacements are 7.5% relative to the presumed array inter-element spacing. The steering delays are quantized to 256 levels. The array is presteered to the direction $\theta_0 = -30°$, whereas the desired signal arrives from the direction $\theta_s = -34°$ with SNR= 10 dB. An interference signal arrives from $\theta_1 = 0°$ with INR= 30 dB. The desired and jamming signals are generated as the sum of sinusoidal harmonics with random phases and equal powers occupying the band between 800 and 4000 Hz. A uniformly-spaced frequency grid with $N = 5$ points is selected to discretize the robustness constraint together with a value of δ equal to 10^{-4}. The values of $\{\sigma_{m,i}^2\}_{i=1}^N$ are all chosen to be equal.

5.4.1   Sensitivity to the choice of filter parameters

We start by investigating the effect of the choice of the parameters $\rho = \sigma_{m,0}^2/(\sigma_s^2 + \sigma_n^2)$ and $\{\sigma_{m,i}^2\}_{i=1}^N$ on the performance of our EKF-based beamformers. The simulation
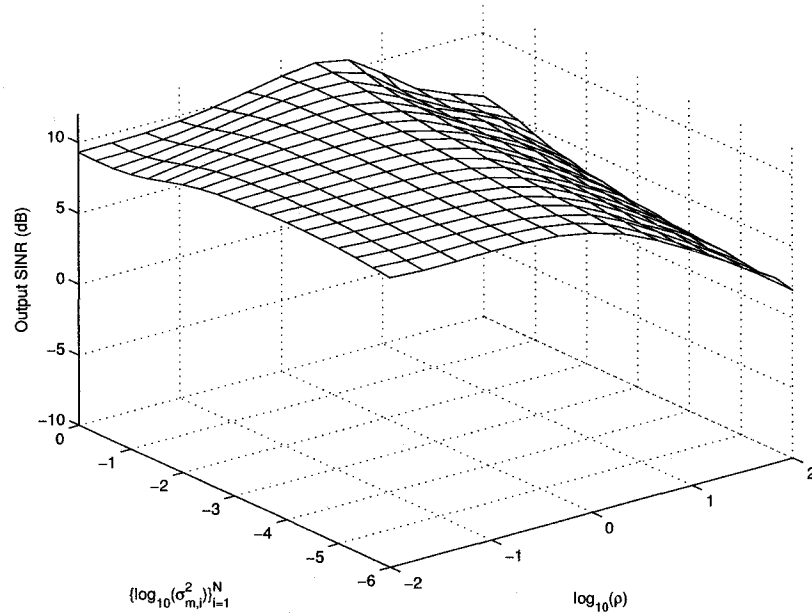
114

Figure 5.2: Average output SINR versus ρ and $\{\sigma_{m,i}^2\}_{i=1}^N$ (second-order EKF algorithm).

results have been averaged over 100 Monte Carlo runs. Figures 5.1 and 5.2 show the average output SINR after 2000 iterations versus different choices of the two parameters for the EKF-based and second-order EKF-based algorithms of the robust beamformer, respectively. The average SINR of the SOCP-based algorithm is equal to 11.02 dB. We can notice from Figure 5.1 that for a wide range of the two parameters there is no noticeable degradation in the output SINR of the EKF-based beamformer. Thus, our beamformer does not require exact knowledge of the desired signal or noise powers to estimate a exact pre-specified value of $\sigma_{m,0}^2$. However, we can notice that extremely small values of $\{\sigma_{m,i}^2\}_{i=1}^N$ cause degradation in the output SINR as well as numerical problems which is shown in the missing parts of the curve for some small values of $\{\sigma_{m,i}^2\}_{i=1}^N$ where the filter did not converge. On the other hand, we can clearly see from Figure 5.2 that with the use of the second-order EKF, the sensitivity of the beamformer towards the selection of its parameters is greatly decreased. This can be
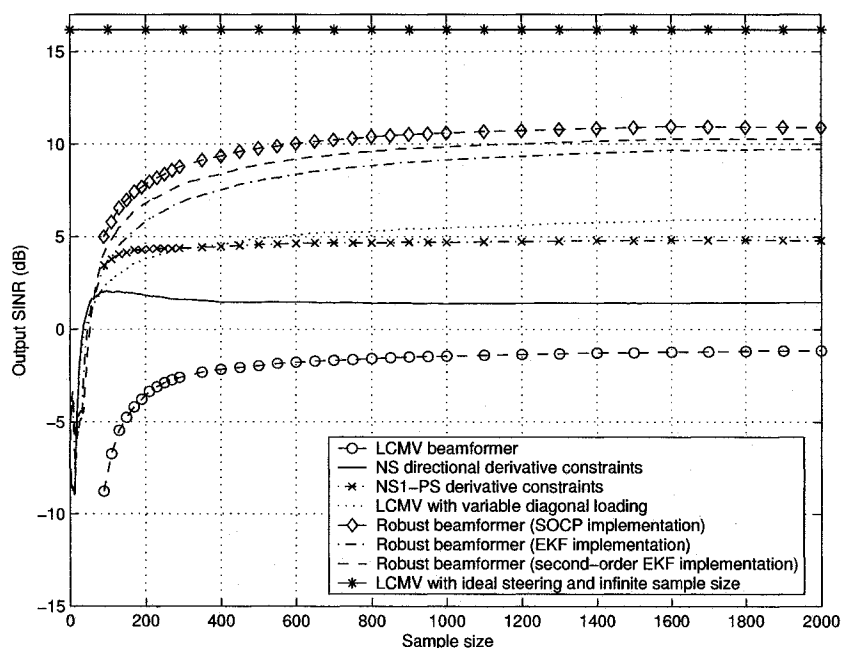
115

Figure 5.3: Average output SINR versus the iteration number.

explained by the better approximation of the robustness constraint by including the second-order terms in the Taylor expansion of the measurement equation.

5.4.2 Performance comparison with other beamformers

In this example, we compare the performance of various implementations of our beamformer versus the RLS-based adaptive implementations of the classical LCMV beamformer and its robust versions using directional derivative constraints, NS1-PS derivative constraints, and variable diagonal loading. The parameters of the EKF-based beamformers are chosen as $\rho = 1$ and $\left\{\sigma_{m,i}^2\right\}_{i=1}^N = 10^{-4}$. Figure 5.3 shows the output SINRs averaged over 100 Monte Carlo runs for different beamformers versus the iteration number for the same scenario as that considered in the previous example. This figure also shows the output SINR of the LCMV beamformer with perfect knowledge of the array manifold and infinite sample size. We can clearly notice that the EKF-
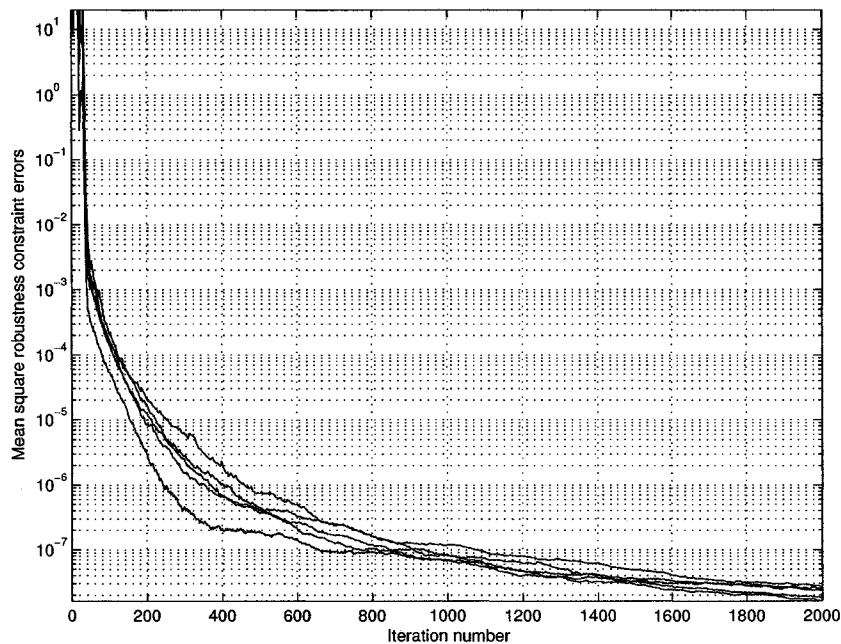
116

Figure 5.4: Robustness constraint MSE versus the iteration number (EKF algorithm).

and SOCP-based implementations of our beamformer have very close performance. Our beamformers clearly outperform the existing wideband robust beamforming algorithms. Figures 5.4 and 5.5 show the MSE in satisfying each of the N discretized robustness constraints for the EKF-based and second-order EKF-based algorithms, respectively. We can see from these figures that the chosen values for $\{\sigma_{m,i}^2\}_{i=1}^N$ can satisfy the robustness constraint with a high accuracy in spite of the approximations used in the EKF.

Figures 5.6–5.8 show the array response towards the desired and interference signals after 2000 iterations for the SOCP-based, EKF-based, and second-order EKF-based algorithms of our robust beamformer, respectively. We can notice that all the beamformers can maintain a distortionless response towards the desired signal in spite of the mismatches in the array manifold while suppressing the interference signal. We can also notice the close resemblance of the array response of the beamformers as they

117
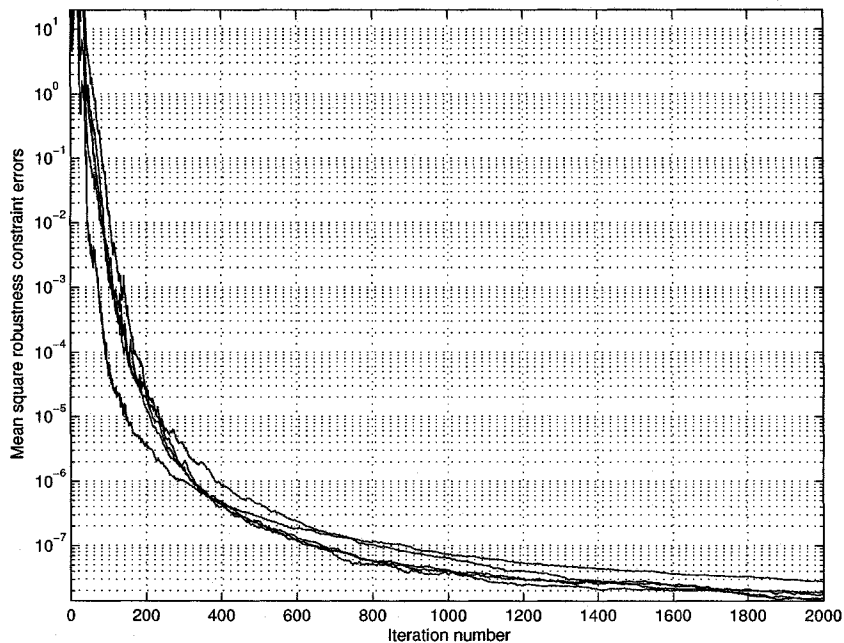
Figure 5.5: Robustness constraint MSE versus the iteration number (second-order EKF algorithm).

all solve the same problem. Figure 5.9 compares the power response of different robust beamformers for different arrival angles. We can clearly see the performance improvements achieved by our beamformers as they yield a high gain in the direction of the desired signal while rejecting the interference signal and maintaining a low sidelobe level.

5.4.3 Robustness against rapid environmental changes

Finally, we consider a nonstationary scenario in which the interference signal changes its DOA after 1500 snapshots from $\theta_1 = 0°$ to $\theta_1 = 10°$ while keeping the same INR of 30 dB. The desired signal and array manifold mismatches are identical to the previous scenario. A nonstationary model with σ_v^2 equal to 10^{-7} is selected for the EKF-based

118

Figure 5.6:   Frequency response of the robust beamformer (SOCP algorithm).
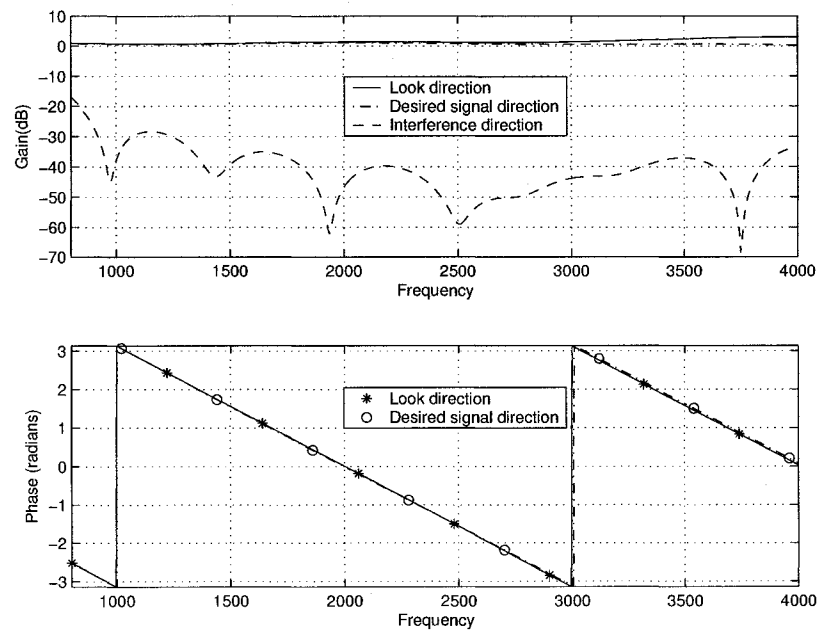
Figure 5.7: Frequency response of the robust beamformer (EKF algorithm).

119
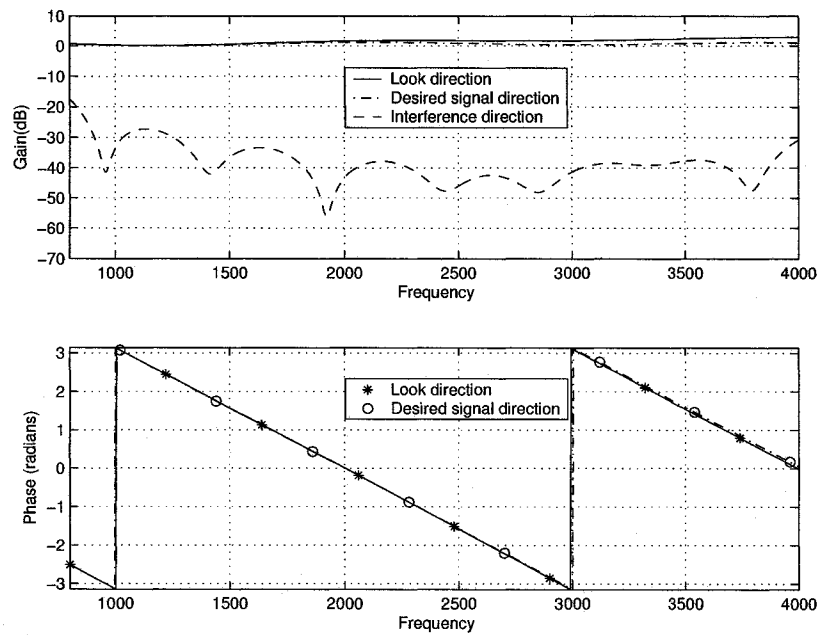
Figure 5.8: Frequency response of the robust beamformer (second-order EKF algorithm).
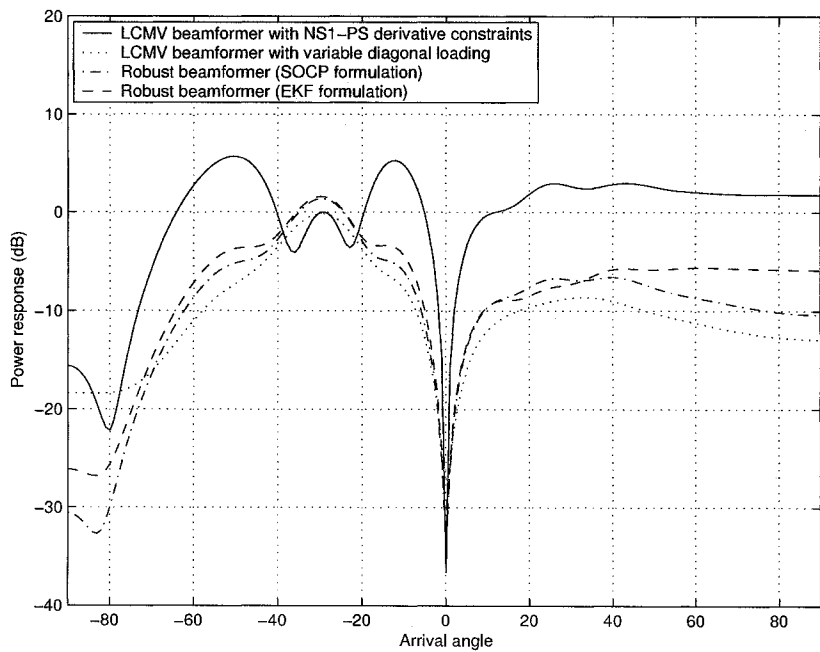
Figure 5.9: Power response versus the arrival angle.
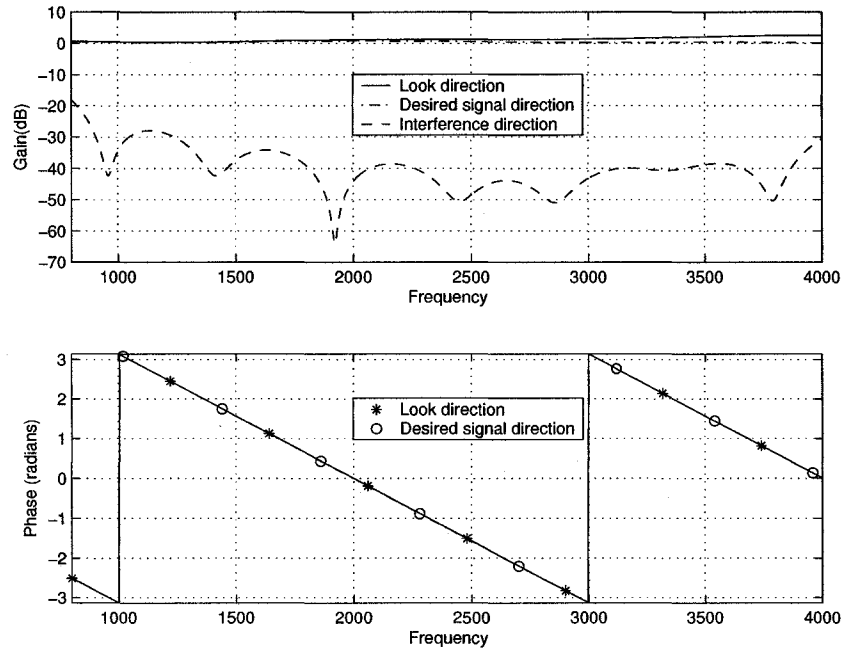
120
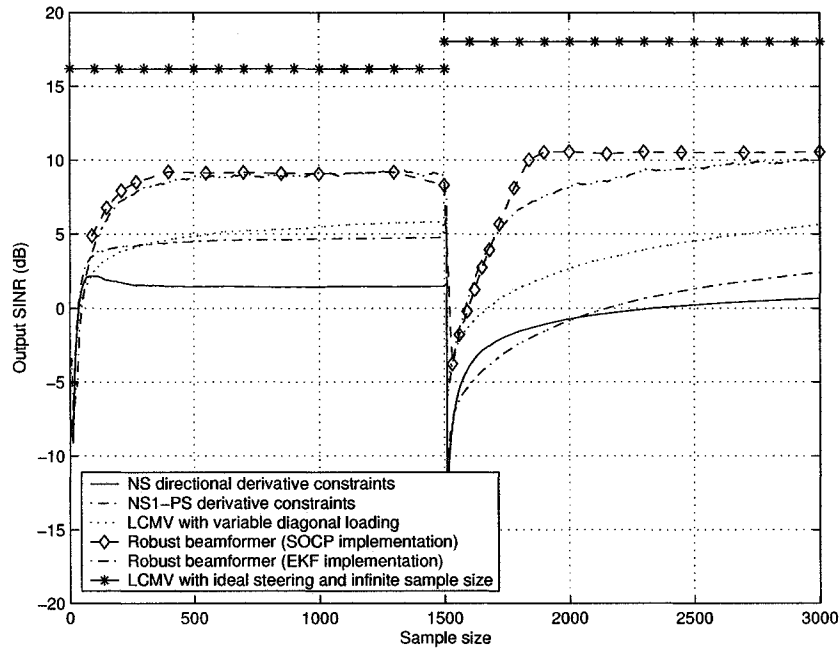
Figure 5.10: Average output SINR versus the iteration number.

algorithm of the robust beamformer. The remaining parameters of the filter are selected as $\rho = 1$ and $\{\sigma_{m,i}^2\}_{i=1}^N = 10^{-4}$. For the SOCP-based algorithm, a sliding window of maximum length 360 snapshots is used to compute the sample covariance matrix, and the whole optimization procedure is repeated every time a new snapshot arrives. We compare the performance of our proposed beamformers against that of the RLS-based implementation of the LCMV beamformer with NS first-order directional derivative constraints, NS1-PS derivative constraints, and variable diagonal loading [108]. The RLS forgetting factor is selected as 0.995. Simulation results are averaged over 100 Monte Carlo runs.

Figure 5.10 shows the average output SINR versus the iteration number for various beamformers. We can clearly see that our robust beamformers can maintain a high output SINR in spite of environmental changes and the mismatches in the array manifold, whereas the other beamformers suffer from decreased output SINR due to

121

insufficient robustness against array manifold mismatches. Although the SOCP-based implementation of our beamformer yields a higher SINR compared to the EKF-based implementation, yet, the computational complexity of our EKF-based beamformer is much lower and, thus, is more suitable for online implementation.

5.5 Conclusion

In this chapter, we have presented online algorithms for the wideband beamformer with robustness against multiple errors in the array manifold. Our algorithms are based on state-space modelling of the beamforming problem. A first- or second-order EKF can be used to estimate the weight vector of the beamformer with low computational cost per iteration compared to the previous LMI and SOCP-based algorithms presented in Chapter 4. This makes it favourable for nonstationary environments which are often encountered in radar, sonar, and mobile communications. Simulation results have illustrated an improved performance of our beamformer both in stationary and nonstationary environments.

Chapter 6

Robust Multiuser Detection

6.1 Introduction

Multiuser detection is a well-established field that has found numerous applications in wireless communications. It is capable of suppressing the interference due to multiple users in a communication system, and is known to be near-far resistant as opposed to conventional matched filtering schemes [114]. This eliminates the need for strict power control algorithms used in early CDMA systems such as IS-95, thus improving the bandwidth-efficiency and decreasing the complexity of the mobile transceiver [85].

The optimal multiuser detector (in terms of the BER) has a complexity that grows exponentially in the number of users, which limits its practical applicability [114]. On the other hand, the linear MMSE detector has much reduced complexity, in addition to the advantage of its convenient online implementation [76]. Typical adaptive implementations of the MMSE detector require that each user sends a training sequence that is already known by the receiver. Tracking minor changes in the operating environment can be done after that through decision-directed mode (DDM) adaptation, in which the detector output symbols are used instead of the training sequence as an estimate of the desired user data symbols. As a rule of thumb, DDM can work

123

properly most of the time if the error probability in the symbol estimate is less than 10% [54]. However, when there is a drastic change in the operating environment such as an addition of a powerful user or a sudden deep fade, DDM detection becomes unreliable and the training sequence has to be retransmitted to the receiver. This may cause a significant decrease in the data throughput rate of the system.

On the other hand, blind mode (BM) multiuser detection does not require any training. It has been proposed by Honig et al. who developed the minimum output energy (MOE) algorithm [57]. This algorithm is similar to the MVDR beamforming algorithm discussed in Chapter 2. It minimizes the output energy of the receiver subject to a linear constraint that prevents the cancellation of the desired user signal. It requires only knowledge of the desired user signature and timing, which are the requirements needed by the conventional matched filter detector as well [57]. Minimizing the output energy of the detector was shown to be equivalent to minimizing the MSE [57]. However, a major drawback of the MOE algorithm is its sensitivity to mismatches in the desired user signature. This problem was treated in an ad-hoc manner in [57] by constraining the surplus energy of the detector to achieve robustness against these mismatches.

A more mathematically rigourous way of achieving robustness against signature mismatches was reported in [24], where a constraint was proposed that maintains a distortionless response towards all the vectors differing from the presumed desired user signature vector by a norm-bounded error vector. The resulting optimization problem was shown to be convex. It was formulated in [24] as an SOCP problem that can be solved in polynomial time using interior point methods [83]. Further extensions of this approach to provide additional robustness against short data length can be found in [45] and [123]. However, a major disadvantage of the algorithms in [24], [45], and [123] is that they can not be implemented online, and therefore, are unsuitable for nonstationary environments.

124

Another disadvantage of robust blind detection schemes is their degraded output SINR in high SNR scenarios compared to training-based MMSE detection. This problem is similar to the one reported in [88] where BM detection using the RLS algorithm suffers from low steady-state SINR compared to decision-directed RLS-based (DDRLS) detection. In [104], the detection-aided RLS algorithm was proposed to solve this problem. This algorithm starts with BM adaptation until reaching a good starting point for DDM detection to continue the adaptation process, thus achieving a high steady-state SINR without the need for training. The detector is switched back to BM adaptation upon detection of environmental changes. The disadvantage of the detection-aided RLS algorithm is that it sets the switching thresholds in an ad-hoc manner based on empirical observations. Moreover, perfect knowledge of the desired user signature is assumed in [104], and therefore, the issue of robustness against desired user signature mismatches may arise.

The use of the Kalman filter as a blind multiuser detector has been originally proposed in [125]. However, perfect knowledge of the desired user signature was also assumed in the later work and the issue of robustness against desired user signature mismatches was not considered. In this chapter, we propose a novel implementation of the robust blind detector of [24]. Our implementation is based on state-space modelling of the robust blind detection problem where a second-order EKF is used to estimate the detector weight vector iteratively. Our robust blind Kalman filter (RBKF) detector has the same computational complexity per iteration as the RLS implementation of the MOE algorithm [88], which is much lower than the complexity of the algorithms reported in [24], [45], and [123].

We also present a state-space model for the DDM detection problem, and propose an algorithm for switching between robust blind and decision-directed detection. Our switching algorithm uses the NIS of the blind detector to test for convergence and accordingly switches to DDM detection. The NIS of the decision-directed detector is

125

used to detect its failure and then switching back to BM detection occurs. The proposed algorithm is therefore capable of detecting and tracking environmental changes (e.g., the addition of new users to the system), and yields an output SINR that is very close to the optimum output SINR of the MMSE detector without any need for training even in the presence of mismatches in the desired user signature.

6.2 Background

Let us consider a K-user synchronous CDMA system.[1] The received signal can be written as [114]

$$y(t) = \sum_{i=1}^{K} A_i b_i s_i(t) + n(t), \qquad t \in [0, T] \tag{6.1}$$

where A_i and b_i are the amplitude and transmitted bit of the ith user, respectively, $s_i(t)$ is the unit energy signature waveform of the ith user over the bit interval T, and $n(t)$ is white Gaussian noise with variance σ_n^2. The data bits $\{b_i\}_{i=1}^{K}$ are assumed to be independent and to take the values ± 1 with equal probabilities.

The signals $\{s_i(t)\}_{i=1}^{K}$ can be represented as finite-dimensional vectors using the orthonormal basis functions $\{\psi_m(t)\}_{m=1}^{M_s}$ defined on $[0, T]$. For example, for a direct-sequence CDMA system, $\psi_m(t) = g(t - (m-1)T_c)$ where $g(t)$ is the chip waveform and T_c is the chip interval. Therefore, we can write

$$s_k(t) = \sum_{m=1}^{M_s} s_{m,k} \psi_m(t). \tag{6.2}$$

Accordingly, we can express the kth M_s-dimensional received vector as

$$\boldsymbol{y}(k) = \sum_{i=1}^{K} A_i b_i(k) \boldsymbol{s}_i + \boldsymbol{n}(k) \tag{6.3}$$

---

[1]All the algorithms presented below can be readily extended to the more practical case of asynchronous CDMA by increasing the length of the observation window beyond one bit interval [57].

where $s_i = [s_{1,i}, s_{2,i}, \ldots, s_{M_s,i}]^T$ and $n(k)$ is the M_s–dimensional white noise vector with independent components each of variance σ_n^2.

Without any loss of generality, we assume that the first user is the desired user whose signature vector is s_1. A linear receiver for this user is given by the vector c and the detector output is given by $\hat{b}_1(k) = \text{sign}\{c^T y(k)\}$. If the transmitted symbols of the desired user are known either through training or estimation, then the MMSE receiver is given by [81]

$$c = \arg\min_c \ E\left\{(A_1 b_1(k) - y^T(k)c)^2\right\} \tag{6.4}$$

Adaptive techniques for the solution of (6.4) which are based on the RLS algorithm were proposed in [88]. For the kth received signal vector, the DDRLS receiver weight vector is given by

$$c(k) = \arg\min_c \sum_{i=1}^{k} \lambda_{\text{RLS}}^{k-i}(A_1 b_1(i) - y^T(i)c)^2. \tag{6.5}$$

On the other hand, the blind MOE receiver does not require any knowledge of the desired user symbols. It can be formulated as the following optimization problem [57]

$$\min_c \ E\left\{(c^T y)^2\right\} = c^T R_y c \qquad \text{s.t.} \quad c^T \hat{s}_1 = 1 \tag{6.6}$$

where $\hat{s}_1$ is the presumed desired signal signature vector and $R_y = E\{y(k)y^T(k)\}$ is the covariance matrix of the received vector. The solution to this problem is

$$c = \frac{R_y^{-1}\hat{s}_1}{\hat{s}_1^T R_y^{-1}\hat{s}_1}, \tag{6.7}$$

and yields the same SINR and BER as the MMSE detector for $\hat{s}_1 = s_1$ [57].

An equivalent formulation for the MOE detector known as the partitioned linear interference canceller was presented in [97]. This formulation is similar to the GSC formulation of the LCMV algorithm. It decomposes the receiver weight vector c as

$$c = \hat{s}_1 - C_1^\perp c_a \tag{6.8}$$

127

where c_a is the adaptive part of c, and the columns of the full rank $M_s \times (M_s - 1)$-dimensional matrix $C_1^\perp$ form an orthonormal basis for the nullspace of $\hat{s}_1$. The use of the above weight vector decomposition eliminates the need for the constraint in (6.6), and the MOE receiver can be written as

$$c_a = \arg\min_{c_a} \ \mathrm{E}\left\{(y^T(k)\hat{s}_1 - y^T(k)C_1^\perp c_a)^2\right\}. \tag{6.9}$$

Adaptive techniques to solve (6.9) were also proposed in [88]. For example, the RLS-based MOE algorithm chooses the weight vector $c(k)$ adaptively such that

$$c_a(k) = \arg\min_{c_a} \sum_{i=1}^{k} \lambda_{\mathrm{RLS}}^{k-i} \left(y^T(i)\hat{s}_1 - y^T(i)C_1^\perp c_a\right)^2 \tag{6.10}$$

$$c(k) = \hat{s}_1 - C_1^\perp c_a(k). \tag{6.11}$$

In practical situations, mismatches may exist between the presumed and actual desired user signature vectors, i.e., $\hat{s}_1 \neq s_1$. These mismatches arise due to the effect of propagation through the channel (e.g., multipath fading) and/or timing asynchronism. The performance of the MOE detector is known to degrade severely for even slight mismatches of such type. Several approaches have been considered to overcome this problem; see [24], [45], [57], and [117]. For example, in [57] Honig et al. proposed to constrain the norm of the weight vector to prevent the cancellation of the desired user signal. The corresponding modified problem can be formulated as

$$\min_{c} \quad c^T R_y c$$

$$\mathrm{s.t.} \quad c^T \hat{s}_1 = 1$$

$$\|c\|^2 \leq 1 + \gamma_c \tag{6.12}$$

where γ_c is the maximum allowable value of the so-called surplus energy of the detector. The solution to this problem can be found using the method of Lagrange multipliers and is given by [9]

$$c_{\mathrm{dl}} = \frac{(R_y + \eta_c I)^{-1}\hat{s}_1}{\hat{s}_1^T(R_y + \eta_c I)^{-1}\hat{s}_1} \tag{6.13}$$

128

where the parameter η_c is chosen so that the second constraint in (6.12) is satisfied. It is clear from (6.13) that constraining the surplus energy of the detector results in diagonal loading of the received vector covariance matrix. However the amount of diagonal loading is often set in an ad-hoc manner, independent of the amount of mismatches in the desired user signature.

In many scenarios, we can model the actual received signature vector of the desired user as

$$s_1 = \hat{s}_1 + e \tag{6.14}$$

where e is the signature error vector whose norm $\|e\|$ is a measure of the magnitude of the signature distortion. For example, in an unknown frequency selective fading environment, the actual received signature vector is given by $s_1 = \tilde{s}_1 * h_c$ where h_c is the M_s-dimensional channel response vector, $\tilde{s}_1$ is the transmitted signature vector of the desired user, and $*$ denotes the convolution operator [24]. The norm of the signature error vector can therefore be bounded by

$$
\begin{aligned}
\|e\| = \|s_1 - \hat{s}_1\| &= \|\tilde{s}_1 * (h_c - h_i)\| \\
&\leq \sqrt{M_s}\, \|\tilde{s}_1\|\, \|h_c - h_i\|
\end{aligned} \tag{6.15}
$$

where $\hat{s}_1$ is equal to $\tilde{s}_1$ augmented by $M_s - 1$ zeros and h_i is the ideal channel response (Kronecker delta). Another example, in the case of timing asynchronism, we can use Taylor approximation to bound $s_1(t + \tau_d) - s_1(t)$, where τ_d is the timing offset. Hence, the signature error vector is bounded by $\|e\| \leq B_d M_s \tau_d$ [24], where B_d is the upper bound for the derivative of the signature waveform $s_1(t)$.

One of the recent mathematically rigourous robust approaches to blind multiuser detection is based on providing a distortionless response for all the vectors that differ from the presumed desired user signature vector by a norm-bounded vector e [24]. The problem can be formulated as

$$\min_{c} c^T R_y c \quad\quad \text{s.t.} \quad c^T \tilde{s}_1 \geq 1 \quad \forall \tilde{s}_1 \in \mathcal{S}_1(\varepsilon) \tag{6.16}$$

129

where $\mathcal{S}_1(\varepsilon) = \left\{ \tilde{s}_1 \,|\, \tilde{s}_1 = \hat{s}_1 + e, \|e\| \leq \varepsilon \right\}$.

This problem is equivalent to the following SOCP optimization problem [24]

$$\min_{c} c^T R_y c \qquad \text{s.t.} \quad c^T \hat{s}_1 - \varepsilon \|c\| \geq 1 \qquad (6.17)$$

which can be solved efficiently using well-established interior point methods with a complexity of $\mathcal{O}\left(M_s^3\right)$ [83]. The disadvantage of this formulation is that it does not yield any closed form solution, and thus can not be readily implemented online.

6.3   The Robust Blind Kalman Filter Detector

In this section, we derive our new RBKF detector that alleviates most of the shortcomings of the previous approaches. First, we note that at the optimal point of (6.17), the robustness constraint has to be satisfied with equality. This can be easily seen by noticing that if the constraint was not active at the optimal point, i.e., if $c^T \hat{s}_1 - \varepsilon \|c\| = \beta > 1$, then we could further decrease the cost function simply by dividing the receiver weight vector estimate by β, while still satisfying the constraint. Thus, without any loss of generality we can replace the inequality in the robustness constraint with equality. Therefore, the robust blind detection problem in (6.17) can be written as

$$\min_{c_B(k)} \mathrm{E}\left\{ (y^T(k)c_B(k))^2 \right\} \qquad \text{s.t.} \quad h(c_B(k)) = 1 \qquad (6.18)$$

where $h(c_B(k)) = c_B^T(k)\hat{s}_1 - \varepsilon \|c_B(k)\|$, and the subscript $(\cdot)_B$ stands for "blind".

We will use a state-space modelling approach to solve the optimization problem in (6.18) similar to that we have used in Chapters 3 and 5. The process equation of the model is given by

$$c_B(k) = c_B(k-1) + w_B(k) \qquad (6.19)$$

130

where $c_B(k)$ is the state vector that represents the optimal robust blind detector weight vector, and $w_B(k)$ is the process noise vector that enables tracking the optimal solution of (6.18) in nonstationary environments. The process noise models the changes in the operating environment that can be attributed to a large number of random variables, e.g., the channel coefficients, the distortion in the signature, the number of users, etc. Hence, $w_B(k)$ is assumed to be Gaussian (by the Central Limit Theorem) with zero mean and covariance matrix $Q_B = \sigma_B^2 I$. The parameter σ_B^2 is selected such that it matches the degree of nonstationarity of the operating environment. For example a value of σ_B^2 equal to 10^{-6} indicates that each component of the optimal weight vector is expected to change independently by an order of 10^{-3} every time step.

The corresponding measurement equation can be written as

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} y^T(k)c_B(k) \\ h(c_B(k)) \end{bmatrix} + \begin{bmatrix} n_{B,0}(k) \\ n_{B,1}(k) \end{bmatrix} \tag{6.20}$$

where $n_{B,0}(k)$ and $n_{B,1}(k)$ are modelled as zero mean white sequences, independent of each other and of the process noise. In matrix notation, the latter equation can be expressed as

$$z_B = h(c_B(k)) + n_B(k) \tag{6.21}$$

where $z_B = [0, 1]^T$, $h(c_B(k)) = \left[y^T(k)c_B(k), h(c_B(k))\right]^T$, and $n_B(k) = [n_{B,0}(k), n_{B,1}(k)]^T$. The covariance matrix of the measurement noise is given by

$$R = E\{n_B(k)n_B^T(k)\} = \begin{bmatrix} \sigma_{B,0}^2 & 0 \\ 0 & \sigma_{B,1}^2 \end{bmatrix}. \tag{6.22}$$

Applying a state estimator to the above model will yield an estimate for the optimal weight vector $c_B(k)$ that minimizes the uncertainties due to the process and measurement noises. Note that minimizing the mean square value of $n_{B,0}(k)$ is equivalent to minimizing the output power of the detector, while minimizing the mean square value of $n_{B,1}(k)$ will minimize the MSE in satisfying the robustness constraint.

131

The mean square residual power $\sigma_{B,0}^2$ should be chosen of the same order as that of the optimal output power of the receiver, which is given roughly by $(A_1^2 + \sigma_n^2)\|c_B(k)\|^2$. It should be stressed that our RBKF detector does not require perfect knowledge of any of these parameters, and the subsequent simulation results show that it performs well for a wide range of the parameter $\sigma_{B,0}^2(k)$. At the same time, to satisfy the robustness constraint with good accuracy, the value of $\sigma_{B,1}^2$ should be as small as possible (for example, $\sigma_2^2 = 10^{-12}$, but a few orders of magnitude higher than the machine "epsilon" to avoid numerical problems.

Due to nonlinearity of the measurement equation, we will use the second-order EKF to find a recursion for the estimated vector $\hat{c}_B(k)$ [6]. We start by evaluating the Jacobian $H_c(k, c_B(k))$ of $h(c_B(k))$ and the Hessian matrices $H_{cc}^{(0)}$ and $H_{cc}^{(1)}(c_B(k))$ of its components, which are given by

$$H_c(k, c_B(k)) = \left(\nabla h^T(c_B(k))\right)^T = \left[y^T(k), \hat{s}_1^T - \varepsilon \frac{c_B^T(k)}{\|c_B(k)\|}\right]^T \tag{6.23}$$

$$H_{cc}^{(0)} = \nabla\nabla^T\left\{c_B^T(k)y(k)\right\} = 0 \tag{6.24}$$

$$H_{cc}^{(1)}(c_B(k)) = \nabla\nabla^T\left\{h(c_B(k))\right\} = \frac{\varepsilon c_B(k)c_B^T(k)}{\|c_B(k)\|^3} - \frac{\varepsilon I}{\|c_B(k)\|}. \tag{6.25}$$

The recursion for the estimated weight vector starts with an initial weight vector estimate $\hat{c}_B(0)$ together with its associated covariance matrix $P_B(0|0)$, and updates the weight vector estimate as

$$\hat{c}_B(k) = \hat{c}_B(k-1) + G_B(k)\left(z_B - \hat{z}_B(k|k-1)\right) \tag{6.26}$$

where the filter gain and the predicted measurement are given, respectively, by

$$G_B(k) = P_B(k|k-1)H_c^T(k, \hat{c}_B(k-1))S_B^{-1}(k) \tag{6.27}$$

$$\hat{z}_B(k|k-1) = \begin{bmatrix} \hat{c}_B^T(k-1)y(k) \\ h(\hat{c}_B(k-1)) + \frac{1}{2}\mathrm{tr}\left\{H_{cc}^{(1)}(\hat{c}_B(k-1))P_B(k|k-1)\right\} \end{bmatrix}. \tag{6.28}$$

In the latter equation, $\mathrm{tr}\left\{ \boldsymbol{H}_{cc}^{(1)}(\hat{\boldsymbol{c}}_{\mathrm{B}}(k-1))\boldsymbol{P}_{\mathrm{B}}(k|k-1) \right\}$ can be simplified as

$$
\begin{aligned}
&\mathrm{tr}\left\{ \boldsymbol{H}_{cc}^{(1)}(\hat{\boldsymbol{c}}_{\mathrm{B}}(k-1))\boldsymbol{P}_{\mathrm{B}}(k|i-1) \right\} \\
&= \frac{\varepsilon}{\|\hat{\boldsymbol{c}}_{\mathrm{B}}(k-1)\|^3}\left(\hat{\boldsymbol{c}}_{\mathrm{B}}^{T}(k-1)\boldsymbol{P}_{\mathrm{B}}(k|k-1)\hat{\boldsymbol{c}}_{\mathrm{B}}(k-1) - \|\hat{\boldsymbol{c}}_{\mathrm{B}}(k-1)\|^2\mathrm{tr}\{\boldsymbol{P}_{\mathrm{B}}(k|k-1)\} \right).
\end{aligned}
$$

$$(6.29)$$

The covariance matrix of the predicted weight vector $\boldsymbol{P}_{\mathrm{B}}(k|k-1)$, and the covariance matrix of the updated weight vector $\boldsymbol{P}_{\mathrm{B}}(k|k)$ are given by

$$
\boldsymbol{P}_{\mathrm{B}}(k|k-1) = \boldsymbol{P}_{\mathrm{B}}(k-1|k-1) + \boldsymbol{Q}_{\mathrm{B}} \tag{6.30}
$$

$$
\boldsymbol{P}_{\mathrm{B}}(k|k) = \boldsymbol{P}_{\mathrm{B}}(k|k-1) - \boldsymbol{G}_{\mathrm{B}}(k)\boldsymbol{S}_{\mathrm{B}}(k)\boldsymbol{G}_{\mathrm{B}}^{T}(k) \tag{6.31}
$$

The innovation covariance matrix $\boldsymbol{S}_{\mathrm{B}}(k)$ is given by

$$
\begin{aligned}
\boldsymbol{S}_{\mathrm{B}}(k) &= \boldsymbol{H}_{c}(k,\hat{\boldsymbol{c}}(k-1))\boldsymbol{P}_{\mathrm{B}}(k|k-1)\boldsymbol{H}_{c}^{T}(k,\hat{\boldsymbol{c}}(k-1)) + \boldsymbol{R} \\
&+ \frac{1}{2}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\mathrm{tr}\left\{ \boldsymbol{H}_{cc}^{(1)}(\hat{\boldsymbol{c}}_{\mathrm{B}}(k-1))\boldsymbol{P}_{\mathrm{B}}(k|k-1)\boldsymbol{H}_{cc}^{(1)}(\hat{\boldsymbol{c}}_{\mathrm{B}}(k-1))\boldsymbol{P}_{\mathrm{B}}(k|k-1) \right\}.
\end{aligned}
$$

$$(6.32)$$

Also, $\mathrm{tr}\left\{ \boldsymbol{H}_{cc}^{(1)}(\hat{\boldsymbol{c}}_{\mathrm{B}}(k-1))\boldsymbol{P}_{\mathrm{B}}(k|k-1)\boldsymbol{H}_{cc}^{(1)}(\hat{\boldsymbol{c}}_{\mathrm{B}}(k-1))\boldsymbol{P}_{\mathrm{B}}(k|k-1) \right\}$ can be written as

$$
\begin{aligned}
&\mathrm{tr}\left\{ \boldsymbol{H}_{cc}^{(1)}(\hat{\boldsymbol{c}}_{\mathrm{B}}(k-1))\boldsymbol{P}_{\mathrm{B}}(k|k-1)\boldsymbol{H}_{cc}^{(1)}(\hat{\boldsymbol{c}}_{\mathrm{B}}(k-1))\boldsymbol{P}_{\mathrm{B}}(k|k-1) \right\} \\
&= \frac{\varepsilon^2}{\|\hat{\boldsymbol{c}}_{\mathrm{B}}(k-1)\|^6}\left((\hat{\boldsymbol{c}}_{\mathrm{B}}^{T}(k-1)\boldsymbol{P}_{\mathrm{B}}(k|k-1)\hat{\boldsymbol{c}}_{\mathrm{B}}(k-1))^2 + \|\hat{\boldsymbol{c}}_{\mathrm{B}}(k-1)\|^4\mathrm{tr}\{\boldsymbol{P}_{\mathrm{B}}^2(k|k-1)\} \right. \\
&\left. - 2\|\hat{\boldsymbol{c}}_{\mathrm{B}}(k-1)\|^2\|\boldsymbol{P}_{\mathrm{B}}(k|k-1)\hat{\boldsymbol{c}}_{\mathrm{B}}(k-1)\|^2 \right).
\end{aligned}
$$

$$(6.33)$$

The consistency of the filter can be checked through testing its NIS, which is defined as [6]

$$
\epsilon_{\mathrm{B}}(k) \triangleq (\boldsymbol{z}_{\mathrm{B}} - \hat{\boldsymbol{z}}_{\mathrm{B}}(k|k-1))^{T}\boldsymbol{S}_{\mathrm{B}}^{-1}(k)(\boldsymbol{z}_{\mathrm{B}} - \hat{\boldsymbol{z}}_{\mathrm{B}}(k|k-1)), \tag{6.34}
$$

and is chi-square distributed with two degrees of freedom [6]. The NIS should lie within a certain region of acceptance if the filter is consistent. For example, the acceptance region can be selected as the double-sided interval containing 90% of the probability of the distribution. An exponential or moving average window can also be used for the calculation of the NIS to decrease the variability, and therefore, to increase the power of the test.

For initialization of the iterative algorithm in (6.26)–(6.32), a random Gaussian weight vector estimate $\hat{c}_B(0)$, together with an initial covariance matrix estimate $P_B(1|0) = \beta_0 I$ are used, where β_0 is a constant that reflects high uncertainty in our initial random weight vector estimate. One way of choosing β_0 is so that $\epsilon_B(1)$ lies in the center of the double-sided 90% acceptance region for the chi-square distribution with two degrees of freedom. That is, ignoring the second-order term and the measurement noise covariance matrix in (6.32), and substituting with the approximated $S_B(1)$ in (6.34), we obtain

$$\beta_0 = \frac{1}{\epsilon_B(1)}(z_B - \hat{z}_B(1|0))^T \left(H_c(1, \hat{c}(0))H_c^T(1, \hat{c}(0))\right)^{-1} (z_B - \hat{z}_B(1|0)). \qquad (6.35)$$

We stress here that the computational complexity of all the update equations of the RBKF detector is of $\mathcal{O}(M_s^2)$ operations per iteration, which is the same order of computational complexity as the RLS implementation of the non-robust blind MOE detector in [88]. This computational complexity is much lower than the $\mathcal{O}(M_s^3)$ complexity of the interior point methods-based implementation of the robust blind detector that can find the exact optimal solution of (6.18). Moreover, the RBKF detector has the advantage of being an iterative algorithm, which is more suited to nonstationary environments than the algorithms in [24] and [123].

6.4 State-Space Model for Decision-Directed Detection

Following the same guidelines of the previous section, we can easily derive a state-space model for the optimal weight vector $c_D(k)$ of the DDM detector that solves the optimization problem in (6.4). The process and measurement equations for the DDM receiver are given by

$$c_D(k) = c_D(k-1) + w_D(k) \tag{6.36}$$

$$A_1 \hat{b}_1(k) = y^T(k) c_D(k) + n_D(k) \tag{6.37}$$

where the subscript $(\cdot)_D$ stands for "decision-directed", $w_D(k)$ represents the process noise that is assumed to be Gaussian with zero mean and covariance matrix $Q_D = \sigma_D^2 I$, and $n_D(k)$ represents the measurement noise that is assumed to be white with zero mean and variance $\sigma_{D,0}^2$.

For a stationary environment, the value of $\sigma_{D,0}^2$ can be roughly determined by performing a steady-state analysis of (6.37), and assuming that at convergence to the optimal point the interferers have been completely nulled out and, hence, the detector always makes correct decisions, i.e., $\hat{b}_1(k) = b_1(k)$. Then,

$$
\begin{aligned}
\sigma_{D,0}^2 &= \lim_{i \to \infty} E\left\{ \left(A_1 \hat{b}_1(k) - y^T(k) c_D(k) \right)^2 \right\} \\
&= \lim_{i \to \infty} E\left\{ \left(A_1 b_1(k)\left(1 - s_1^T c_D(k)\right) - \sigma_n n^T(k) c_D(k) \right)^2 \right\} \\
&= \lim_{i \to \infty} A_1^2 (1 - s_1^T c_D(k))^2 + \sigma_n^2 \|c_D(k))\|^2 \\
&= A_1^2 (1 - s_1^T c_D)^2 + \sigma_n^2 \|c_D\|^2
\end{aligned}
\tag{6.38}
$$

where c_D is the optimal weight vector for the stationary environment. In practical situations, the optimal weight vector is unknown. However, subsequent simulation results show that the performance of the decision-directed Kalman filter (DDKF)

135

detector is not very sensitive to the choice of the value of $\sigma_{D,0}^2$, i.e., there is no significant degradation in the output SINR for a large range of selection of $\sigma_{D,0}^2$.

The recursive algorithm for estimating the weight vector $\hat{c}_D(k)$ of the DDKF detector is given by

$$\boldsymbol{P}_D(k|k-1) = \boldsymbol{P}_D(k-1|k-1) + \boldsymbol{Q}_D \tag{6.39}$$

$$S_D(k) = \sigma_{D,0}^2 + \boldsymbol{y}^T(k)\boldsymbol{P}_D(k|k-1)\boldsymbol{y}(k) \tag{6.40}$$

$$\boldsymbol{G}_D(k) = \boldsymbol{P}_D(k|k-1)\boldsymbol{y}(k)S_D^{-1}(k) \tag{6.41}$$

$$\hat{\boldsymbol{c}}_D(k) = \hat{\boldsymbol{c}}_D(k-1) + \boldsymbol{G}_D(k)\left(A_1\hat{b}_1(k) - \boldsymbol{y}^T(k)\hat{\boldsymbol{c}}_D(k-1)\right) \tag{6.42}$$

$$\boldsymbol{P}_D(k|k) = \boldsymbol{P}_D(k|k-1) - \boldsymbol{G}_D(k)S_D(k)\boldsymbol{G}_D^T(k). \tag{6.43}$$

In the above, $\hat{b}_1(k)$ is the estimate of the desired user symbol, which may be obtained using a training sequence. Alternatively, we can use some weight vector $\tilde{\boldsymbol{c}}(k)$ to estimate $\hat{b}_1(k)$ as $\hat{b}_1(k) = \text{sign}\left\{\boldsymbol{y}^T(k)\tilde{\boldsymbol{c}}(k)\right\}$. In the standard decision-directed algorithm, the weight vector $\tilde{\boldsymbol{c}}(k)$ is given by the previous weight vector estimate of the algorithm, that is, $\tilde{\boldsymbol{c}}(k) = \hat{\boldsymbol{c}}_D(k-1)$.

The NIS of the DDKF detector is defined as

$$\epsilon_D(k) = \frac{1}{S_D(k)}\left(A_1\hat{b}_1(k) - \boldsymbol{y}^T(k)\boldsymbol{c}_D(k)\right)^2. \tag{6.44}$$

It is chi-square distributed with one degree of freedom and should lie within acceptable limits if the detector is consistent [6]. Also, a sliding or exponential window can be used for the computation of the NIS.

Initialization of the DDKF detector can be done in a similar way to the RBKF detector, that is, the initial detector weight vector can be chosen as a random Gaussian vector and its associated covariance can be selected as $\boldsymbol{P}_D(1|0) = \beta_0\boldsymbol{I}$, where

$$\beta_0 = \frac{1}{\epsilon_D(1)\|\boldsymbol{y}(1)\|^2}\left(A_1\hat{b}_1(1) - \boldsymbol{y}^T(1)\hat{\boldsymbol{c}}_D(0)\right)^2. \tag{6.45}$$

136
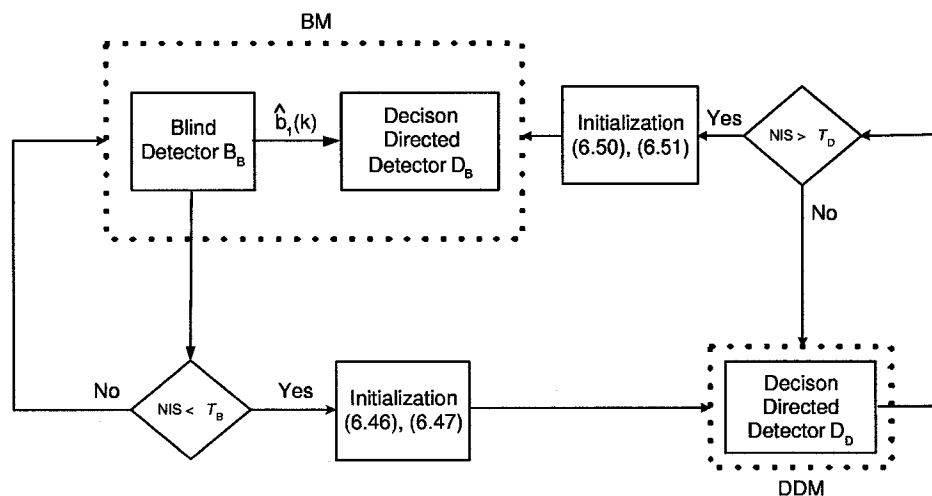
Figure 6.1: Combined blind and decision-directed detection algorithm.

6.5 Combined Blind and Decision-Directed Detection

As mentioned before, decision-directed detection can achieve a high steady-state SINR. However, it requires a good initial point for successful adaptation and can suffer from catastrophic failure in the case of abrupt changes in the environment [81], [104]. On the other hand, robust blind detection does not require any training and can combat mismatches in the desired user signature, but suffers from a decreased steady-state output SINR.

In this section, we present an algorithm for switching between the two proposed Kalman filter-based detection techniques. Our algorithm uses BM detection to provide a good starting point for DDM detection. Switching to DDM occurs when the convergence of the blind detector is detected through a chi-square test on its NIS. After switching to DDM, the NIS of the decision-directed detector is tested to detect possible failure of DDM due to nonstationarities, and subsequently switch back to BM detection. Figure 6.1 shows a block diagram of our algorithm.

137

6.5.1   Blind mode

The BM detector contains two filters: a robust blind filter denoted by B_B and given by (6.23)–(6.31), and a decision-directed filter denoted by D_B and given by (6.39)–(6.43). The symbol estimates for the filter D_B are provided using the weight vector estimate of the blind detector, i.e., $\hat{b}_1(k) = \text{sign}\left\{ \boldsymbol{y}^T(k)\hat{\boldsymbol{c}}_{B_B}(k) \right\}$. The BM receiver weight vector is given by the weight vector estimate of the B_B filter, i.e., $\hat{\boldsymbol{c}}(k) = \hat{\boldsymbol{c}}_{B_B}(k)$, where the subscript $(\cdot)_{X_Y}$ refers to the filter X (RBKF or DDKF) operating in the Y mode.

An exponentially-windowed NIS can be used to detect convergence of the blind filter. It is calculated through

$$\epsilon_{B_B}(k) = \alpha\, \epsilon_{B_B}(k-1) + \left(\boldsymbol{z}_{B_B} - \hat{\boldsymbol{z}}_{B_B}(k|k-1)\right)^T \boldsymbol{S}_{B_B}^{-1}(k)\left(\boldsymbol{z}_{B_B} - \hat{\boldsymbol{z}}_{B_B}(k|i-1)\right) \quad (6.46)$$

where α is a constant close to but less than unity, resulting in an effective window length of $\frac{1}{1-\alpha}$. Under the above mentioned Gaussian assumption, $\epsilon_{B_B}(k)$ can be approximated using first-order moment matching as chi-square distributed with $\lceil\frac{2}{1-\alpha}\rceil$ degrees of freedom [6].

After the ith iteration, $\epsilon_{B_B}(k)$ is compared against a threshold $T_B(k)$ and if $\epsilon_{B_B}(k) \leq T_B(k)$, convergence is declared and detection is switched to DDM adaptation. If the detector has been operating in BM for m iterations, the threshold $T_B(k)$ can be taken as the expected value of the chi-square distribution with the corresponding degrees of freedom, and can be approximated by

$$T_B(k) \approx \begin{cases} m\,\mu_{\chi_2^2} & \text{if } m < \lceil\frac{1}{1-\alpha}\rceil \\ \mu_{\chi_2^2}/(1-\alpha) & \text{if } m \geq \lceil\frac{1}{1-\alpha}\rceil \end{cases} \quad (6.47)$$

where $\mu_{\chi_2^2}$ is the mean value of the chi-square distributed random variable with two degrees of freedom.

6.5.2 Decision-directed mode

The DDM detector contains a single decision-directed filter D_D given by (6.39)–(6.43), which uses its previous weight estimate to obtain the estimate for the current symbol, i.e., $\hat{b}_1(k) = \text{sign}\left\{ \boldsymbol{y}^T(k)\hat{\boldsymbol{c}}_{D_D}(k-1) \right\}$.

Upon switching to DDM, the filter D_D is initialized using the estimates of the BM decision-directed filter. Therefore, if switching to DDM has occurred after the ith iteration, we set

$$\hat{\boldsymbol{c}}_{D_D}(i) = \hat{\boldsymbol{c}}_{D_B}(i) \tag{6.48}$$

$$\boldsymbol{P}_{D_D}(i|i) = \boldsymbol{P}_{D_B}(i|i). \tag{6.49}$$

If the operating environment changes and the filter fails to adapt to the new environment, the difference between the measurement $A_1\hat{b}_1(k)$ and its predicted value $\boldsymbol{y}^T(k)\hat{\boldsymbol{c}}_{D_D}(k-1)$ will increase, leading to a high value of the NIS of the filter D_D [6]. Therefore, a chi-square test can be used to detect nonstationarities to decide whether or not to switch back to BM detection. Similar to the BM detector, the exponentially-windowed NIS can be evaluated recursively as

$$\epsilon_{D_D}(k) = \alpha\,\epsilon_{D_D}(k-1) + \frac{1}{S_{D_D}(k)}\left(A_1\hat{b}_1(k) - \boldsymbol{y}^T(k)\hat{\boldsymbol{c}}_{D_D}(k-1) \right)^2. \tag{6.50}$$

If $\epsilon_{D_D}(k)$ exceeds a certain pre-specified threshold $T_D(k)$ then detection is switched to BM. This threshold can be selected as the γ_u upper-tail probability point of the chi-square distribution with $\lceil \frac{1}{1-\alpha} \rceil$ degrees of freedom. If the detector has been operating in DDM for m iterations, $T_D(k)$ can be approximated by

$$T_D(k) \approx \begin{cases} m\,\chi_1^2(\gamma_u) & \text{if } m < \lceil \frac{1}{1-\alpha} \rceil \\ \chi_1^2(\gamma_u)/(1-\alpha) & \text{if } m \geq \lceil \frac{1}{1-\alpha} \rceil \end{cases} \tag{6.51}$$

where $P\left\{ \chi_1^2 \geq \chi_1^2(\gamma_u) \right\} = 1 - \gamma_u$, and χ_1^2 is a chi-square distributed random variable with one degree of freedom.

139
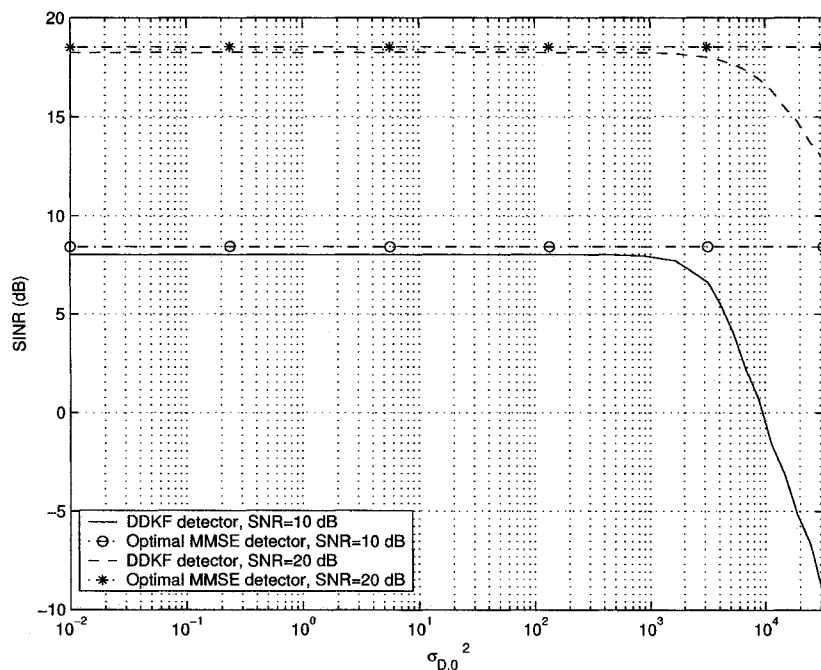
Figure 6.2: Average output SINR versus $\sigma_{D,0}^2$.

Upon switching back to BM detection, the weight vector estimate of the filter D_D is used to initialize the two BM filters B_B and D_B, and the corresponding covariance matrices are *re-initialized* again. That is, if switching to the BM occurred after the ith iteration, we have

$$\hat{c}_{B_B}(i) = \hat{c}_{D_B}(i) = \hat{c}_{D_D}(i) \tag{6.52}$$

$$P_{B_B}(i|i) = P_{D_B}(i|i) = \beta_0 I \tag{6.53}$$

where β_0 is chosen according to (6.35).

140
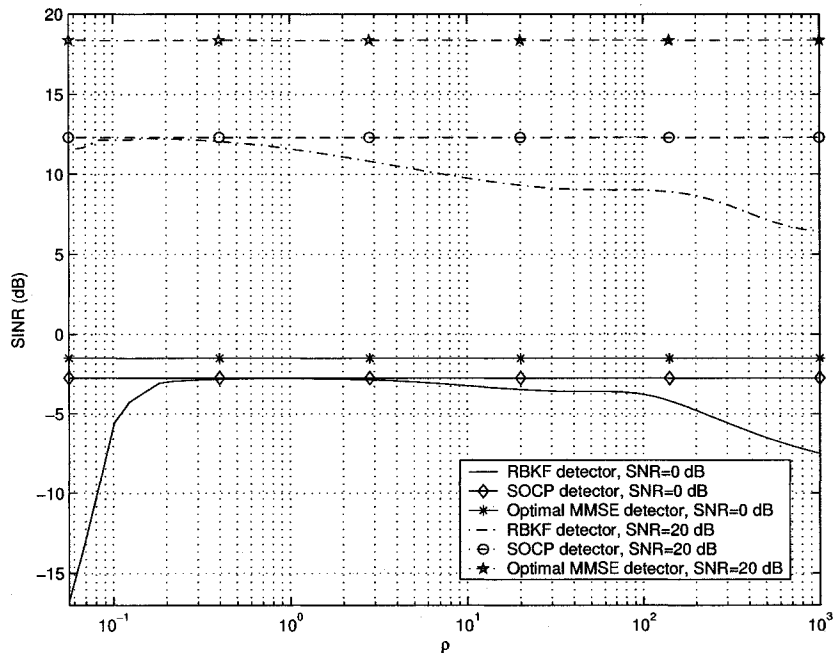
Figure 6.3: Average output SINR versus $\rho = \sigma_{B,0}^2/(A_1^2 + \sigma_n^2)$.

6.6 Simulations

6.6.1 Sensitivity to filter parameters

We start our numerical examples by examining the sensitivity of our Kalman filter-based detectors to the choice of their model parameters. We consider a synchronous CDMA system with randomly generated Gaussian signature vectors of length $M_s = 31$ [57], [104]. The signature vectors are affected by additive random Gaussian distortions with magnitude 0.4 relative to the magnitude of the original signature vectors. The norm of the distortion is assumed to be known at the receiver, i.e., $\varepsilon = 0.4$. The system has 9 interfering users with different ISR: 5 users with ISR = 10 dB, 2 users with ISR = 20 dB, and 2 users with ISR = 30 dB. The simulation results are averaged over 200 Monte Carlo runs.

First, we consider the performance of the DDKF detector (with training for the

141

first 100 symbols) for different values of the parameter $\sigma_{D,0}^2$. Figure 6.2 shows the average output SINR of the detector after 500 iterations versus $\sigma_{D,0}^2$ at two different SNRs; 10 dB and 20 dB. From this figure, we can notice that at both SNRs the DDKF receiver can achieve nearly the optimal SINR of the MMSE detector over a wide range of selection of the parameter $\sigma_{D,0}^2$.

Figure 6.3 shows the average output SINR of the RBKF detector after 1000 iterations for different choices of the parameter $\rho = \sigma_{B,0}^2/(A_1^2 + \sigma_n^2)$ and at two different SNRs; 0 dB and 20 dB. The other parameters of the detector are chosen as $\sigma_B^2 = 0$ and $\sigma_{B,1}^2 = 10^{-12}$. We can notice from Figure 6.3 that the RBKF detector has a satisfactory performance over a wide range of $\sigma_{B,0}^2$. Note that although our RBKF detector may be expected to perform worse than the SOCP detector (due to the approximations in the use of the EKF), its performance does not degrade significantly over a wide range of selection of the parameter $\sigma_{B,0}^2$ and can still provide a good starting point for DDM adaptation in our proposed switching algorithm.

6.6.2 Performance comparison in stationary environment

In this numerical example, we compare the proposed detectors with the existing ones for a CDMA system similar to that described in the previous example, with the desired user SNR equal to 20 dB. The signature vector of the desired user is distorted due to the effect of propagation through a frequency selective fading channel. The channel is modelled as a 3-tap filter with coefficients $[1, \varepsilon_1, \varepsilon_2]^T$ where ε_1 and ε_2 are Gaussian random variables normalized such that $\varepsilon_1^2 + \varepsilon_2^2 = 0.1$. The value of δ is selected equal to 0.4. The stationary signal model is chosen for all algorithms, i.e., $\sigma_B^2 = \sigma_D^2 = 0$, and $\lambda_{\text{RLS}} = 1$, and the values $\sigma_{B,0}^2 = 2(A_1^2 + \sigma_n^2)$, $\sigma_{B,1}^2 = 10^{-12}$, and $\sigma_{D,0}^2 = 2$ are selected. Training is used for the first 100 symbols of the DDRLS and DDKF detectors, followed by decision-directed adaptation. The simulation results are averaged over 200 Monte Carlo runs.
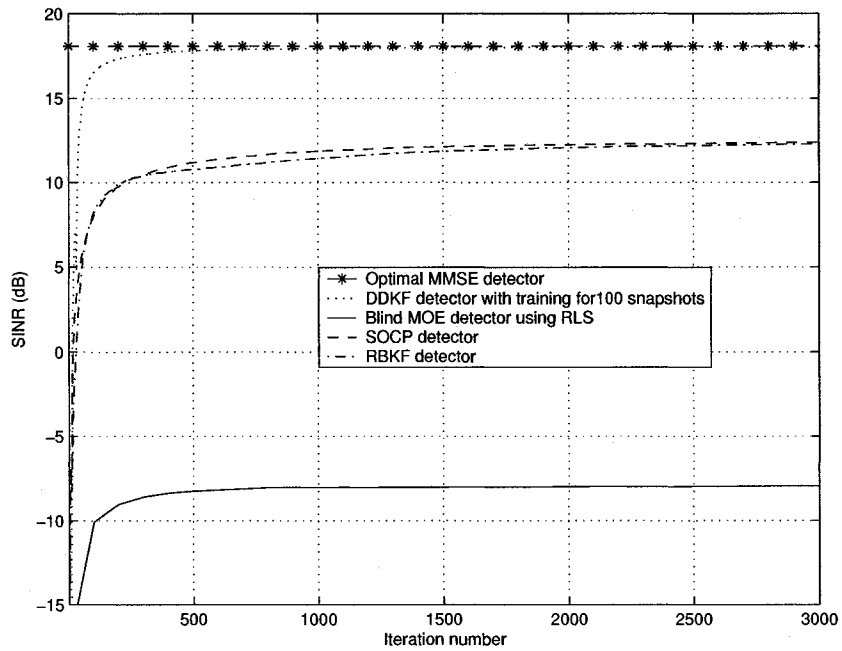
Figure 6.4: Average output SINR versus the iteration number.
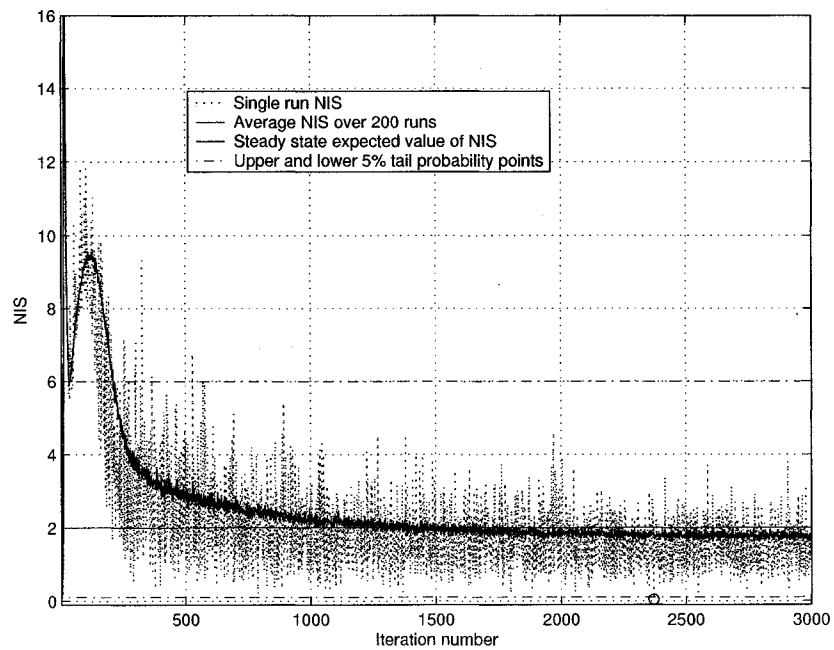
Figure 6.5: NIS versus the iteration number.

143
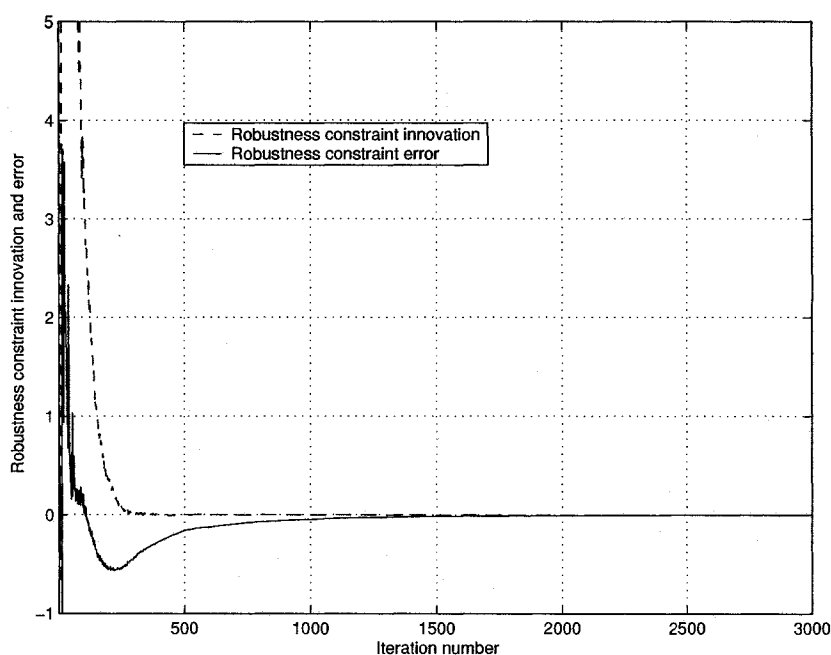
Figure 6.6: Robustness constraint innovation and error versus the iteration number.

Figure 6.4 shows the average output SINR of the proposed Kalman filter-based detectors, the SOCP-based detector, and the blind MOE detector versus the iteration number. It also shows the optimum SINR of the MMSE detector computed in the case when the exact signature vector of the desired user is known. From this figure, we can clearly see that the DDKF detector can achieve the optimum SINR with the help of training. We can also see that the robust detectors can combat the distortion in the desired user signature without any use of training but at the price of a decreased steady-state SINR, whereas the RLS-based MOE detector can be seen to suffer from a reduced output SINR due to the signal self-nulling phenomenon.

Figure 6.5 shows the NIS of the RBKF detector for a single run and the average NIS over 200 runs versus the iteration number. From this figure, we can see that the NIS of the detector converges to its steady-state estimate in approximately 1200 iterations. Fig. 5 also shows the upper and lower 5% tail probability points for the

144
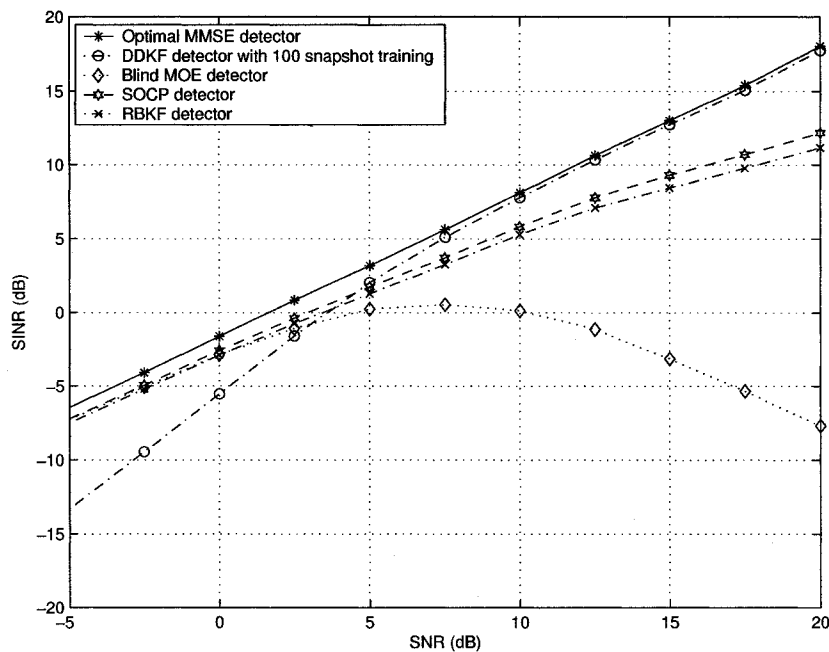
Figure 6.7: Average output SINR versus SNR.

Chi-square distribution with two degrees of freedom. It is clear from this figure that our RBKF detector is consistent on convergence to its steady-state estimate as the NIS lies within the 90% probability region of the distribution.

Figure 6.6 shows the innovation component corresponding to the robustness constraint and the error in satisfying the robustness constraint versus the iteration number for the same run whose NIS is shown in Figure 6.5. We can see from Figure 6.6 that the robustness constraint is satisfied by the weight vector estimate of the RBKF detector at convergence, in spite of the approximations involved in the use of the second-order EKF.

For the same scenario, Figure 6.7 shows the average output SINR versus SNR. From this figure, the advantage of decision directed adaptation for scenarios with high SNRs can be seen. The similarity in the performance of the RBKF and SOCP detectors is also apparent. This similarity translates into a very close BER for a wide

145
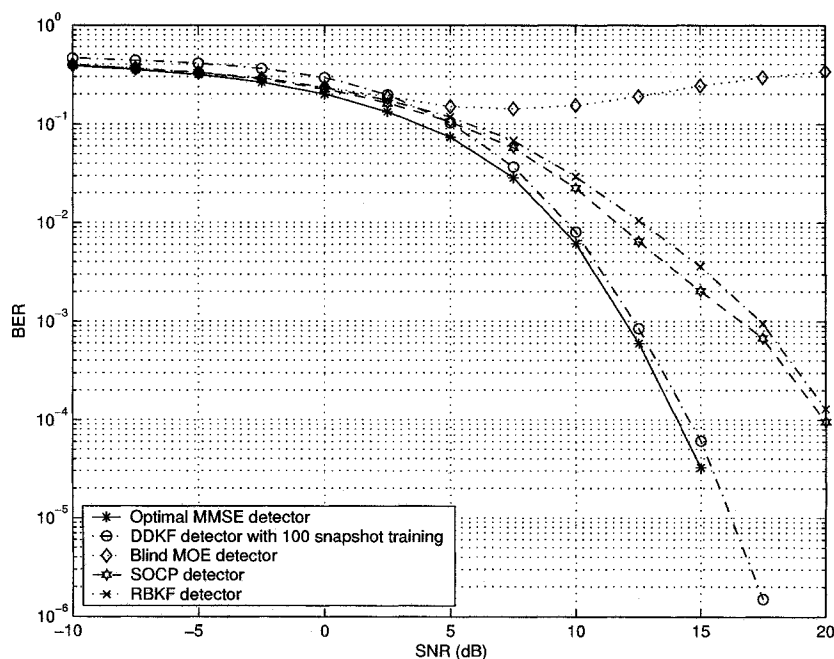
Figure 6.8: Average BER versus SNR.

range of SNR in Figure 6.8. We can also notice the advantage of robust detection schemes over classical MOE detection which fails completely at all values of SNR.

6.6.3   Performance in nonstationary environment

In this example, we consider a CDMA system with nine interferers having identical ISR of 10 dB, and three stronger users added at the 1500th time instant. These three users have randomly generated signature vectors and one of them has ISR = 20 dB, while the other two have ISR = 30 dB. The desired user has an SNR equal to 20 dB. The desired user signature distortion is similar to that considered in the previous simulation and the value of δ is also selected equal to 0.4. For the DDRLS algorithm, initial training is used for the first 100 iterations and the value of λ_{RLS} for all RLS-based detectors is selected as 0.995. For the RBKF detector, a nonstationary model is selected for the process equation with $\sigma_{\mathrm{B}}^2 = \times 10^{-6}$, whereas for all Kalman

146
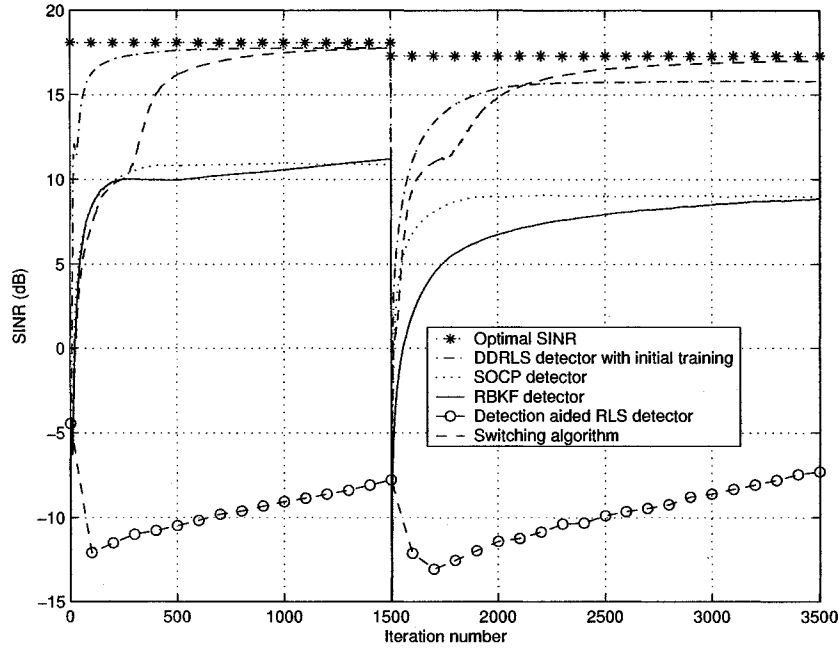
Figure 6.9: Average output SINR versus the iteration number.

filter-based detectors used in the proposed switching algorithm the stationary model is selected with $\sigma_B^2 = \sigma_D^2 = 0$ (as the algorithm can account for nonstationarities and, therefore, there is no need to use a nonstationary model for the process equation). The other parameters of the Kalman filters used in this example are selected as $\sigma_{B,0}^2 = 10(A_1^2 + \sigma_n^2)$, $\sigma_{B,1}^2 = 10^{-12}$, and $\sigma_{D,0}^2 = 2$. A sliding window of 400 snapshots is used to estimate the sample covariance matrix used in the SOCP-based detector. The exponential window parameter for calculating the NIS of the BM and DDM filters is selected as $\alpha = 0.98$, and the threshold parameter as $\gamma_u = 99.9\%$. For the detection-aided RLS algorithm in [104], the lower and upper switching SINRs are given by 0 dB and 7 dB, and both the probability of an early switch to DDM and the probability of false nonstationarity alarm are given by 5%. The simulation results are averaged over 500 Monte Carlo runs.

Figure 6.9 shows the average output SINR versus the iteration number. The

147
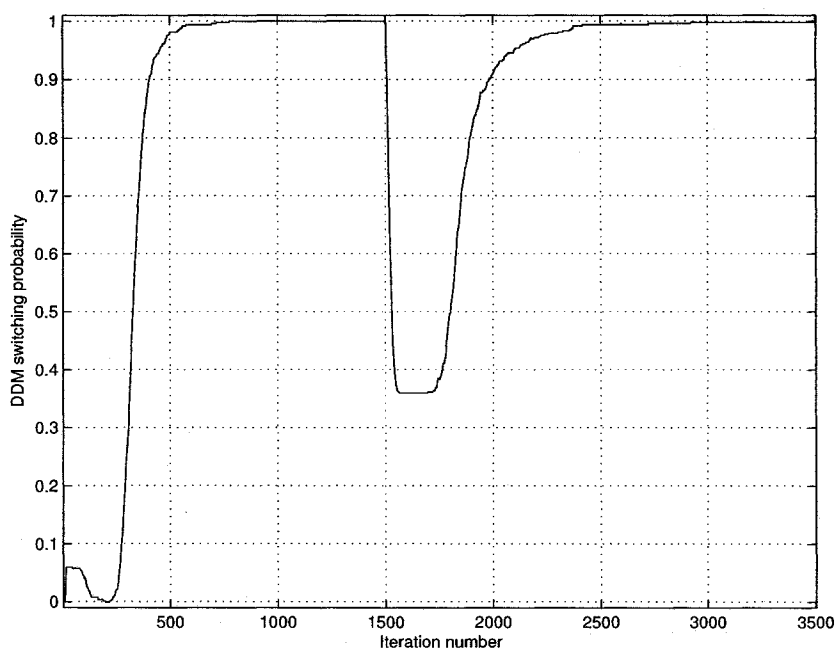
Figure 6.10: Probability of switching to DDM versus the iteration number.

decreased value of the steady-state SINR of the DDRLS algorithm can be attributed to its catastrophic failure in some scenarios. Our switching algorithm does not suffer from this phenomenon. It can automatically re-converge after the addition of the new users without retraining and yields an output SINR that is very close to the SINR of the optimal MMSE detector. From Figure 6.9, it also follows that our detector is capable of detecting changes in the environment rapidly and adapting to the new environment accordingly. The key to this feature is the re-initialization of the blind filter covariance matrix upon switching to BM, which erases all the memory of the filter. Without this re-initialization step, the blind filter will suffer from decreased adaptation rate to sudden environmental changes. The observed complete failure of the detection-aided RLS algorithm can be explained by the fact that its MOE detector is sensitive to mismatches in the desired user signature.

Figure 6.10 shows the DDM probability versus the iteration number. Note that in some runs, DDM detection can adapt to the additional users in the system and

148
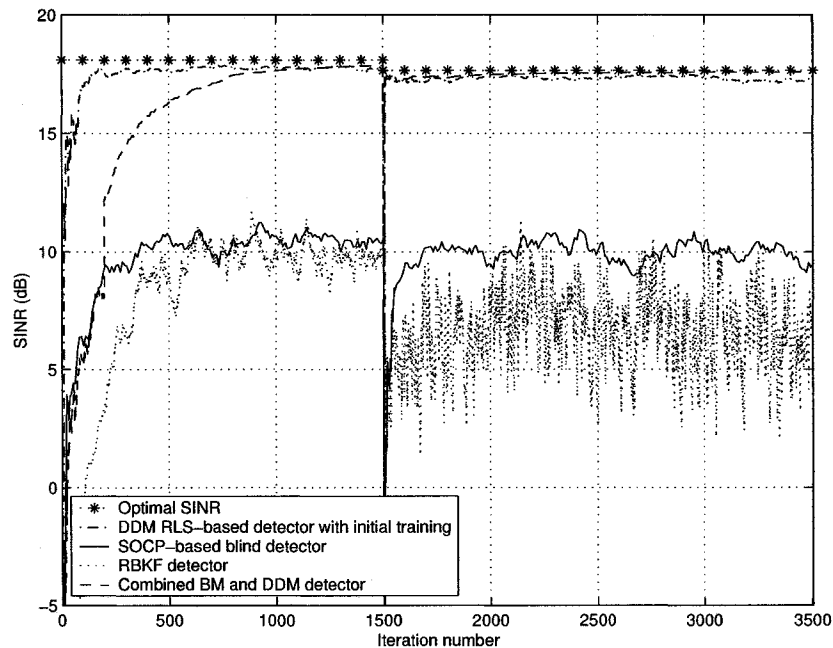
Figure 6.11: Single-run output SINR versus the iteration number.

therefore there is no need to switch to BM. This is reflected in Figure 6.10 by the non-zero probability of DDM after the addition of the three new users to the system. Figure 6.11 shows the single-run output SINR for one of these scenarios in which the decision-directed algorithms can readapt to the new environment without retraining. Figure 6.12 displays the corresponding NIS of both the BM and DDM detectors, from which we can see that in this scenario the NIS of the DDM filter does not exceed the upper threshold $T_D(k)$, and therefore, no switching to BM occurs. Figures 6.13 and 6.14 show, respectively, the single-run output SINR and the NIS for a scenario in which the DDRLS algorithm fails to readapt to the environmental change. It is clear from Figure 6.14 that our switching algorithm can detect the failure of the DDM adaptation as the NIS of the DDKF detector exceeds the acceptable threshold, and thus switching to BM occurs. We can also see that detection is switched back to DDM after convergence of the RBKF detector.

149
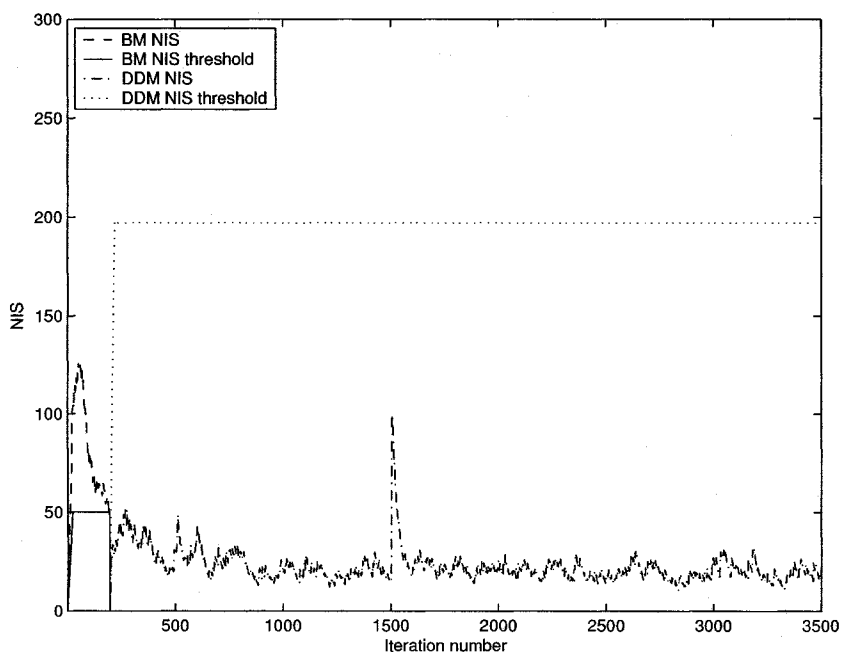
Figure 6.12: Exponentially-windowed NIS versus the iteration number.
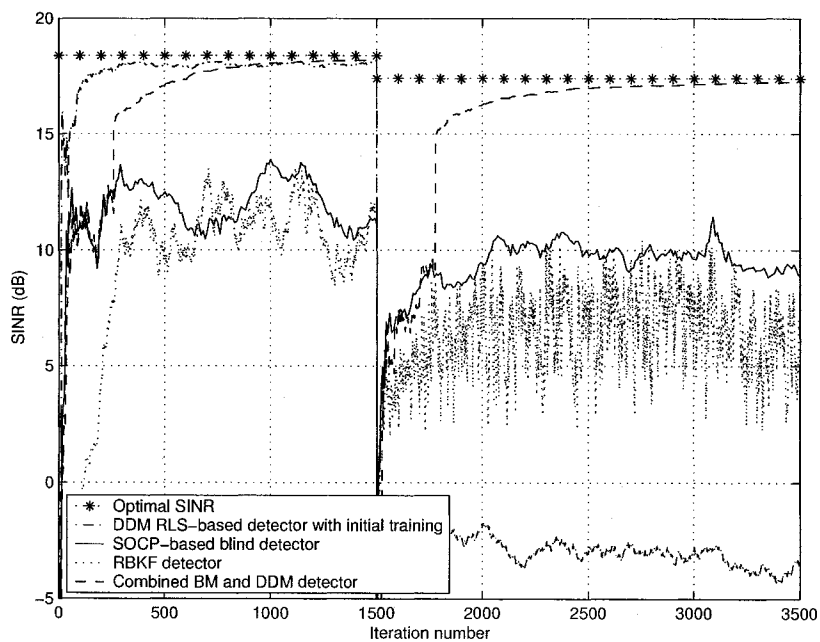
Figure 6.13: Single-run output SINR versus the iteration number.
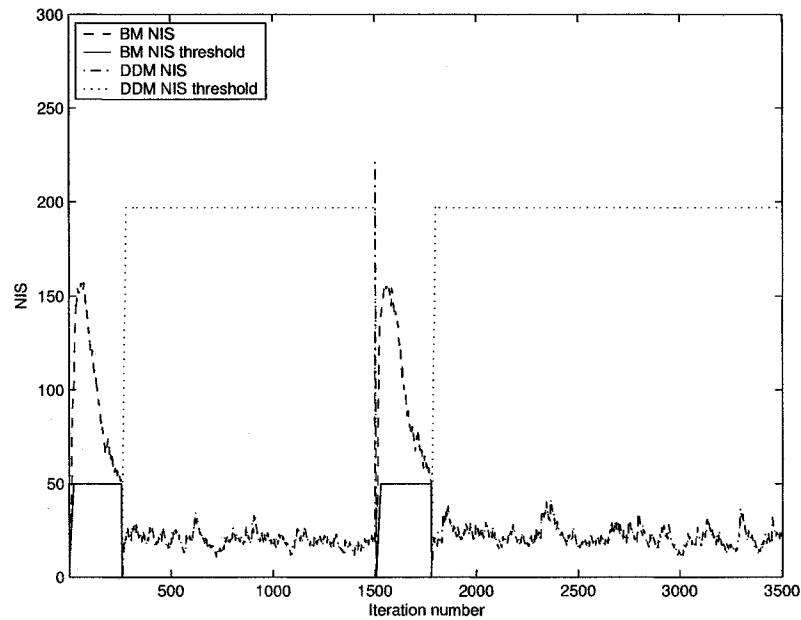
Figure 6.14: Exponentially-windowed NIS versus the iteration number.

6.7   Conclusion

We have presented a robust blind multiuser detector that provides an improved robustness against arbitrary mismatches in the desired user signature vector. In contrast to the previous SOCP-based implementation of this detector, our technique can be implemented online using a second-order EKF, and requires only $\mathcal{O}(M_s^2)$ operations per iteration. A state-space approach has been presented for the decision directed detection problem, and a new switching algorithm between BM and DDM detection has been introduced that is capable of detecting and tracking environmental non-stationarities. The new algorithm can achieve an output SINR that is very close to that of the MMSE detector even in the presence of mismatches in the desired user signature and without any need for training. Simulation results have been presented that show an improved performance of the proposed algorithms compared to earlier techniques.

151

152

Chapter 7

Concluding Remarks and Future Directions

In this chapter we highlight the main contributions in this thesis. We also discuss possible extensions for future work.

7.1 Conclusions

In this thesis, we have developed several computationally efficient adaptive beamforming and multiuser detection algorithms that are applicable in practical operating environments where the signal characteristics are not precisely known.

In Chapter 3, we have presented an iterative algorithm for the narrowband robust MVDR beamformer. Our algorithm has a reduced computational complexity compared to previous implementations of the robust MVDR beamformer. Therefore, it is well-suited to practical nonstationary environments where the beamformer weight vector has to be updated with a low computational cost whenever a new data sample is received. We have also presented two modifications of our proposed algorithm to account for abrupt environmental changes that are often encountered in practical

153

systems. Simulation results were presented illustrating an improved performance of the proposed algorithms.

The problem of wideband beamforming in the presence of multiple mismatches in the array manifold has been addressed in Chapter 4. We have presented a wideband time domain-based beamformer that can be considered as an extension of the robust MVDR beamformer. Our beamformer avoids previous heuristic constraints and provides an amount of robustness directly related to the amount of uncertainties in the array manifold. We have presented two algorithms for our wideband beamformer that can be implemented efficiently with polynomial complexity using interior point methods. Simulation results validated an improved performance of our proposed beamformer compared to the existing algorithms and its improved robustness against mismatches in the array manifold.

In Chapter 5, we have developed a computationally efficient online algorithm for our robust wideband beamformer presented in Chapter 4. Our approach is based on state-space modelling of the robust beamforming problem and has much lower computational complexity than the interior point-based implementations presented in Chapter 4. A first- or second-order EKF can be used to estimate the beamformer weight vector recursively which makes our technique suitable for nonstationary environments. Simulation results have illustrated an improved performance of the proposed beamformer both in stationary and nonstationary environments.

In Chapter 6, we have presented a new algorithm for blind multiuser detection with an improved robustness against arbitrary mismatches in the desired user signature. In contrast to the previous SOCP-based implementation of this detector, our technique can be implemented online using a second-order EKF with a reduced computational complexity. A state-space modelling approach has been also presented for the decision-directed detection problem, and a new switching algorithm between BM and DDM detection schemes has been developed. The new algorithm can achieve

154

an output SINR that is very close to that of the MMSE detector without any need for training even in the presence of mismatches in the desired user signature. Simulation results were presented illustrating an improved performance of the proposed algorithms compared to earlier multiuser detection techniques.

7.2   Future Work

Several extensions of the work presented in this thesis are of interest. In what follows, we will describe some of these future research directions.

7.2.1   Robust narrowband beamforming

It was assumed throughout the thesis that the desired and interference signals are mutually independent. However, either partially or completely correlated interference may be present in practical scenarios because of multipath propagation. In this case, traditional adaptive beamforming algorithms not only fail to place nulls in the directions of the interference but also tend to cancel the desired signal at the beamformer output [89], [120]. Further modification of the robust MVDR beamformer is required to handle correlated interference scenarios.

7.2.2   Robust wideband beamforming

Based on our work in Chapters 4 and 5 on robust wideband beamforming, several issues arise that are worth future study.

- It was shown in [116] that the narrowband robust MVDR beamformer can be interpreted in terms of diagonal loading where the optimal value of the diagonal loading factor is computed based on the known level of uncertainty in the desired signal steering vector. It would be interesting to explore the relationship

155

between our wideband beamformer and the diagonally loaded LCMV algorithm and study whether the former can be considered a special case of the latter with a coloured diagonal loading, i.e., with an optimum choice of the diagonal loading factor for each frequency.

- It would be useful if the computational complexity of the SOCP- and SDP-based implementations of our wideband beamformer could be further reduced. Recall that the matrices B, C_0, and $\{C_m\}_{m=1}^M$ in (4.29) and (4.44) are sparse. It is possible that specially tailored solvers for these problems can be developed.

- Simulation results have shown that the convergence rate of our Kalman filter-based robust wideband beamformer in Chapter 5 is of the same order as that of the LCMV algorithm. An analytical derivation of the convergence rate of the beamformer similar to the one we derived for the narrowband robust MVDR beamformer in Section 3.3.3 would be an interesting task.

7.2.3   Multiuser detection

In Chapter 6, we have presented an algorithm that switches between blind and decision-directed multiuser detection. A possible extension of this work is the development of multiuser detection algorithms that optimally *combine* the estimates of both detection schemes without hard-decision switching. This can be done using multiple model estimation techniques based on the state-space models we have developed in Sections 6.3 and 6.4.

7.2.4   Semi-Blind robust adaptive beamforming

In this thesis, we have developed several algorithms that utilize some partial knowledge available about the array manifold. Significant improvements can still be obtained by further exploitation of the statistical properties of communication signals. Cyclostationarity is one of the important properties of communication signals [3]. Almost all common modulation formats exhibit spectral self-coherence or conjugate self-coherence at discrete multiples of the time periodicities of their waveform statistics. This property has been used during the last decade to develop *blind* adaptive beamforming algorithms that do not require any knowledge about the array manifold [3], [16], [103]. Several research areas exist in this field, including:

- The development of computationally efficient algorithms for blind adaptive beamforming that are robust against cycle frequency errors. Blind adaptive beamforming algorithms suffer from severe performance degradation for even a small mismatch in the cycle frequency of the desired signal [68], [69], [103]. This mismatch frequently arises due to the Doppler shift phenomenon which is common in radar, sonar, and communication systems. There is a need for mathematically rigourous and computationally efficient blind adaptive array processing algorithms that provide a suitable amount of robustness against cycle frequency errors and prevent the degradation in performance that occurs in earlier approaches due to excessive or insufficient constraining of the problem.

- The development of robust semi-blind adaptive array processing algorithms that exploit both the available partial knowledge about the array manifold and the statistical properties of the signals used in practical communication systems. The development of such algorithms will link the blind and non-blind approaches that have dominated the field of adaptive beamforming for the last few decades.

157

158

Appendix A

Proof of Theorem 5.1

Let us consider an operating environment with white noise, one desired signal and J jamming signals that are mutually uncorrelated. Therefore, the data covariance matrix can be written as $\boldsymbol{R_x} = \sigma_n^2 \boldsymbol{I} + \boldsymbol{R_s} + \sum_{i=1}^{J} \boldsymbol{R_i}$, where the desired signal covariance matrix $\boldsymbol{R_s}$ is given by (2.48) and the ith interference covariance matrix is given by

$$\boldsymbol{R_i} = \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \rho_{\theta_i}(f) \Big(\boldsymbol{d}(f) \otimes \big(\boldsymbol{T}(f)\boldsymbol{a}(f,\theta_i) \big) \Big) \Big(\boldsymbol{d}(f) \otimes \big(\boldsymbol{T}(f)\boldsymbol{a}(f,\theta_i) \big) \Big)^{H} df. \quad \text{(A.1)}$$

Let $\boldsymbol{T}(f)\boldsymbol{a}(f,\theta_s) = \boldsymbol{\Psi_s}(f) \in \mathbb{C}^M$ and $\boldsymbol{T}(f)\boldsymbol{a}(f,\theta_i) = \boldsymbol{\Psi_i}(f) \in \mathbb{C}^M$. Therefore, we can write the cost function of the robust beamforming problem in (5.10) as

$$\boldsymbol{w}^T \boldsymbol{R_x} \boldsymbol{w} = \sigma_n^2 \|\boldsymbol{w}\|^2 + \boldsymbol{w}^T \boldsymbol{R_s} \boldsymbol{w} + \sum_{i=1}^{J} \boldsymbol{w}^T \boldsymbol{R_i} \boldsymbol{w}$$

$$= \sigma_n^2 \sum_{m=1}^{M} \|\boldsymbol{\mathcal{W}}_m\|^2 + \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \rho_s(f) \left| \boldsymbol{w}^T \big(\boldsymbol{d}(f) \otimes \boldsymbol{\Psi_s}(f) \big) \right|^2 df$$

$$+ \sum_{i=1}^{J} \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \rho_i(f) \left| \boldsymbol{w}^T \big(\boldsymbol{d}(f) \otimes \boldsymbol{\Psi_i}(f) \big) \right|^2 df \quad \text{(A.2)}$$

We can express this cost function in terms of the frequency responses of the M FIR

159

filters of the array processor as

$$
\boldsymbol{w}^T \boldsymbol{R_x} \boldsymbol{w} = \frac{\sigma_n^2}{f_s} \sum_{m=1}^{M} \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} |\mathcal{W}_m(f)|^2 \, df + \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \rho_s(f) \left| \sum_{m=1}^{M} \Psi_{s_m}(f) \sum_{l=1}^{L} \omega_{m,l} e^{j2\pi f(l-1)T_s} \right|^2 df
$$

$$
+ \sum_{i=1}^{J} \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \rho_i(f) \left| \sum_{m=1}^{M} \Psi_{i_m}(f) \sum_{l=1}^{L} \omega_{m,l} e^{j2\pi f(l-1)T_s} \right|^2 df \tag{A.3}
$$

$$
= \frac{\sigma_n^2}{f_s} \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \sum_{m=1}^{M} |\mathcal{W}_m(f)|^2 df + \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \rho_s(f) \left| \sum_{m=1}^{M} \Psi_{s_m}(f) \mathcal{W}_m(f) \right|^2 df
$$

$$
+ \sum_{i=1}^{J} \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \rho_i(f) \left| \sum_{m=1}^{M} \Psi_{i_m}(f) \mathcal{W}_m(f) \right|^2 df \tag{A.4}
$$

where the first term in (A.3) has been derived using Parseval's theorem.

We can also express the linear phase constraint in (5.10) in the frequency domain as

$$
\mathrm{Im}\left\{ \mathcal{W}_m(f) e^{j2\pi f(L_c-1)T_s} \right\} = 0 \qquad \forall m = 1, \ldots, M; f \in \left[-\frac{f_s}{2}, \frac{f_s}{2} \right]. \tag{A.5}
$$

Therefore, using (A.4) and (A.5), we can write the robust beamforming problem (5.10) as

$$
\min_{\{\mathcal{W}_m(f)\}_{m=1}^{M}} \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \rho_s(f) \left| \sum_{m=1}^{M} \Psi_{s_m}(f) \mathcal{W}_m(f) \right|^2 df + \frac{\sigma_n^2}{f_s} \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \sum_{m=1}^{M} |\mathcal{W}_m(f)|^2 \, df
$$

$$
+ \sum_{i=1}^{J} \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \rho_i(f) \left| \sum_{m=1}^{M} \Psi_{i_m}(f) \mathcal{W}_m(f) \right|^2 df
$$

$$
\text{s.t.} \quad e^{j2\pi f(L_c-1)T_s} \sum_{m=1}^{M} \mathcal{W}_m(f) - \varepsilon(f) \left\| [\mathcal{W}_1(f), \ldots, \mathcal{W}_M(f)]^T \right\| \geq 1 \quad \forall\, f \in [f_l, f_u]
$$

$$
\mathrm{Im}\left\{ \mathcal{W}_m(f) e^{j2\pi f(L_c-1)T_s} \right\} = 0 \qquad \forall m = 1, \ldots, M; f \in \left[-\frac{f_s}{2}, \frac{f_s}{2} \right] \tag{A.6}
$$

where the optimization variables are now the set of Fourier transforms of the M FIR filters of the array processor.

160

In what follows, we will proof theorem 5.1 by contradiction. Let $\{\mathcal{W}_m(f)\}_{m=1}^{M}$ be the optimal solution to the above problem such that

$$e^{j2\pi f(L_c-1)T_s} \sum_{m=1}^{M} \mathcal{W}_m(f) - \varepsilon(f) \left\| [\mathcal{W}_1(f), \ldots, \mathcal{W}_M(f)]^T \right\| = N(f) \qquad (A.7)$$

where $N(f)$ is real, continuous, and greater than or equal to 1 for all $f \in \left[-\frac{f_s}{2}, \frac{f_s}{2}\right]$.

We define the set of FIR filter coefficients $\{\boldsymbol{\mathcal{V}_m}\}_{m=1}^{M}$ such that their Fourier transforms are given by

$$\mathcal{V}_m(f) = \mathcal{W}_m(f)/N(f) \qquad (A.8)$$

where $\mathcal{V}_m(f)$ is the Fourier transform of $\boldsymbol{\mathcal{V}_m}$. First, we note that $\{\mathcal{V}_m(f)\}_{m=1}^{M}$ is feasible, that is,

$$\begin{aligned}
e^{j2\pi f(L_c-1)T_s} \sum_{m=1}^{M} &\mathcal{V}_m(f) - \varepsilon(f) \left\| [\mathcal{V}_1(f), \ldots, \mathcal{V}_M(f)]^T \right\| \\
&= \frac{1}{N(f)} \left(e^{j2\pi f(L_c-1)T_s} \sum_{m=1}^{M} \mathcal{W}_m(f) - \varepsilon(f) \left\| [\mathcal{W}_1(f), \ldots, \mathcal{W}_M(f)]^T \right\| \right) \\
&= 1 \qquad\qquad \forall f \in [f_l, f_u] \qquad (A.9)
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{Im}\left\{ \mathcal{V}_m(f) e^{j2\pi f(L_c-1)T_s} \right\} &= \frac{1}{N(f)} \mathrm{Im}\left\{ \mathcal{W}_m(f) e^{j2\pi f(L_c-1)T_s} \right\} \\
&= 0 \qquad\qquad \forall f \in \left[-\frac{f_s}{2}, \frac{f_s}{2} \right] \qquad (A.10)
\end{aligned}$$

161

Next, we evaluate the cost function for the solution set $\{\mathcal{V}_m\}_{m=1}^M$

$$
\begin{aligned}
\boldsymbol{v}^T \boldsymbol{R}_x \boldsymbol{v} &= \frac{\sigma_n^2}{f_s} \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \sum_{m=1}^M \left| \frac{\mathcal{W}_m(f)}{N(f)} \right|^2 df + \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \rho_s(f) \left| \sum_{m=1}^M \Psi_{s_m}(f) \frac{\mathcal{W}_m(f)}{N(f)} \right|^2 df \\
&\quad + \sum_{i=1}^J \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \rho_i(f) \left| \sum_{m=1}^M \Psi_{i_m}(f) \frac{\mathcal{W}_m(f)}{N(f)} \right|^2 df \\
&= \frac{\sigma_n^2}{f_s} \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \frac{1}{N^2(f)} \sum_{m=1}^M |\mathcal{W}_m(f)|^2 df + \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \frac{\rho_s(f)}{N^2(f)} \left| \sum_{m=1}^M \Psi_{s_m}(f) \mathcal{W}_m(f) \right|^2 df \\
&\quad + \sum_{i=1}^J \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \frac{\rho_i(f)}{N^2(f)} \left| \sum_{m=1}^M \Psi_{i_m}(f) \mathcal{W}_m(f) \right|^2 df \qquad\qquad\qquad (A.11) \\
&\leq \frac{\sigma_n^2}{f_s} \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \sum_{m=1}^M |\mathcal{W}_m(f)|^2 df + \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \rho_s(f) \left| \sum_{m=1}^M \Psi_{s_m}(f) \mathcal{W}_m(f) \right|^2 df \\
&\quad + \sum_{i=1}^J \int_{-\frac{f_s}{2}}^{\frac{f_s}{2}} \rho_i(f) \left| \sum_{m=1}^M \Psi_{i_m}(f) \mathcal{W}_m(f) \right|^2 df \qquad\qquad\qquad (A.12)
\end{aligned}
$$

where the inequality in (A.12) is due to the fact that $\rho_s(f)$ and $\{\rho_i(f)\}_{i=1}^J$ are non-negative for all $f \in \left[-\frac{f_s}{2}, \frac{f_s}{2}\right]$.

Note that (A.12) is satisfied with equality if and only if $N(f) = 1$, $\forall f \in \left[-\frac{f_s}{2}, \frac{f_s}{2}\right]$. Therefore, the robustness constraint has to be satisfied with equality for all frequencies, otherwise (A.12) contradicts with the assumption that $\boldsymbol{w}$ is the optimal solution for the robust beamforming problem.

Therefore, at the optimal point of the robust beamforming problem, the robustness constraint has to be satisfied with equality at all frequency points.

162

Appendix B

Simplification of (5.15)

Expanding the matrix $\Lambda(f_i) = \boldsymbol{I}_M \otimes (\boldsymbol{B}\boldsymbol{g}(f_i))$ and the matrix $\boldsymbol{C}$ defined in (5.8), we can write $\Lambda^T(f_i)\boldsymbol{C}^T\boldsymbol{D}$ as

$$\Lambda^T(f_i)\boldsymbol{C}^T\boldsymbol{D} = \begin{bmatrix} \boldsymbol{g}^T(f_i)\boldsymbol{B}^T & \boldsymbol{0}_{1\times L} & & \\ \boldsymbol{0}_{1\times L} & \boldsymbol{g}^T(f_i)\boldsymbol{B}^T & & \\ & & \ddots & \boldsymbol{0}_{1\times L} \\ & & \boldsymbol{0}_{1\times L} & \boldsymbol{g}^T(f_i)\boldsymbol{B}^T \end{bmatrix} \begin{bmatrix} \boldsymbol{C}_1^T\boldsymbol{D} \\ \boldsymbol{C}_2^T\boldsymbol{D} \\ \vdots \\ \boldsymbol{C}_M^T\boldsymbol{D} \end{bmatrix}. \quad \text{(B.1)}$$

Using the definitions of $\boldsymbol{C}_m$ and $\boldsymbol{D}$ in (4.20) and (5.13), respectively, we can write

$$\boldsymbol{C}_m^T\boldsymbol{D} = (\boldsymbol{I}_L \otimes \boldsymbol{e}_m)(\boldsymbol{B} \otimes \boldsymbol{I}_M) = \boldsymbol{B} \otimes \boldsymbol{e}_m^T. \quad \text{(B.2)}$$

Substituting with (B.2) in (B.1), we obtain

$$\Lambda^T(f_i)\boldsymbol{C}^T\boldsymbol{D} = \begin{bmatrix} (\boldsymbol{g}^T(f_i)\boldsymbol{B}^T\boldsymbol{B}) \otimes \boldsymbol{e}_1^T \\ (\boldsymbol{g}^T(f_i)\boldsymbol{B}^T\boldsymbol{B}) \otimes \boldsymbol{e}_2^T \\ \vdots \\ (\boldsymbol{g}^T(f_i)\boldsymbol{B}^T\boldsymbol{B}) \otimes \boldsymbol{e}_M^T \end{bmatrix} \quad \text{(B.3)}$$

$$= (\boldsymbol{g}^T(f_i)\boldsymbol{B}^T\boldsymbol{B}) \otimes \boldsymbol{I}_M. \quad \text{(B.4)}$$

163

Noting that $\boldsymbol{B}^T \boldsymbol{B} \in \mathbb{R}^{L_c \times L_c} = \mathrm{diag}\{1, 2, 2, \ldots, 2\}$, we can write

$$\boldsymbol{\Lambda}^T(f_i) \boldsymbol{C}^T \boldsymbol{D} = \tilde{\boldsymbol{g}}(f_i) \otimes \boldsymbol{I}_M = \tilde{\boldsymbol{\Lambda}}^T(f_i) \tag{B.5}$$

where $\tilde{\boldsymbol{g}}^T(f_i) = \boldsymbol{g}^T(f_i) \boldsymbol{B}^T \boldsymbol{B} = [1, 2\cos(2\pi f_i T), \ldots, 2\cos(2\pi f_i (L_c - 1)T)]$.

Similarly, substituting with $\boldsymbol{C}_0 = \boldsymbol{I}_L \otimes \boldsymbol{1}_M$ and $\boldsymbol{D} = \boldsymbol{B} \otimes \boldsymbol{I}_M$, we can write

$$
\begin{aligned}
\boldsymbol{g}^T(f_i) \boldsymbol{B}^T \boldsymbol{C}_0^T \boldsymbol{D} &= \boldsymbol{g}^T(f_i) \boldsymbol{B}^T \left(\boldsymbol{I}_L \otimes \boldsymbol{1}_M^T \right) \left(\boldsymbol{B} \otimes \boldsymbol{I}_M \right) \\
&= \boldsymbol{g}^T(f_i) \boldsymbol{B}^T (\boldsymbol{B} \otimes \boldsymbol{1}_M^T) \tag{B.6} \\
&= (\boldsymbol{g}^T(f_i) \boldsymbol{B}^T) \otimes 1)(\boldsymbol{B} \otimes \boldsymbol{1}_M^T) \\
&= (\boldsymbol{g}^T(f_i) \boldsymbol{B}^T \boldsymbol{B}) \otimes \boldsymbol{1}_M^T = \tilde{\boldsymbol{g}}^T(f_i) \otimes \boldsymbol{1}_M^T = \boldsymbol{r}^T(f_i). \tag{B.7}
\end{aligned}
$$

Using (B.5) and (B.7) we arrive at the simplified form of (5.15) in (5.16).

164

Bibliography

[1] Y. I. Abramovich, "Controlled method for adaptive optimization of filters using the criterion of maximum SNR," *Radio Engineering and Electronic Physics*, vol. 26, pp. 87–95, Mar. 1981.

[2] Y. I. Abramovich and A. I. Nevrev, "An analysis of effectiveness of adaptive maximization of the signal-to-noise ratio which utilizes the inversion of the estimated correlation matrix," *Radio Engineering and Electronic Physics*, vol. 26, pp. 67–74, Dec. 1981.

[3] B. G. Agee, S. Schell, and W. A. Gardner, "Spectral self-coherence restoral: A new approach to blind adaptive signal extraction using antenna arrays," *Proceedings of the IEEE*, vol. 78, pp. 753–767, Apr. 1990.

[4] S. P. Applebaum and D. J. Chapman, "Adaptive arrays with main beam constraints," *IEEE Transactions on Antennas and Propagation*, vol. 24, pp. 650–662, Sept. 1976.

[5] A. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, pp. 174–188, Feb. 2002.

[6] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Applications to Tracking and Navigation*. New York: John Wiley & Sons, 2001.

[7] K. L. Bell, Y. Ephraim, and H. L. Van Trees, "A Bayesian approach to robust adaptive beamforming," *IEEE Transactions on Signal Processing*, vol. 48, pp. 386–398, Feb. 2000.

[8] H. P. Blom and Y. Bar-Shalom, "The interacting multiple model algorithm for systems with Markovian switching coefficients," *IEEE Transactions on Automatic Control*, vol. 33, pp. 780–783, Aug. 1988.

[9] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.

[10] L. Brennan, J. Mallett, and I. Reed, "Adaptive arrays in airborne MTI radar," *IEEE Transactions on Antennas and Propagation*, vol. 24, pp. 607–615, Sept. 1976.

[11] K. M. Buckley, "Spatial/spectral filtering with linearly constrained minimum variance beamformers," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, pp. 249–266, Mar. 1987.

[12] K. M. Buckley and L. J. Griffths, "An adaptive generalized sidelobe canceller with derivative constraints," *IEEE Transactions on Antennas and Propagation*, vol. 34, pp. 311–319, Mar. 1986.

[13] W. S. Burdic, *Underwater Acoustic System Analysis*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1991.

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

[15] ——, "Equivalence of adaptive array diagonal loading and omnidirectional jamming," *IEEE Transactions on Antennas and Propagation*, vol. 43, pp. 540–541, May 1995.

[16] L. Castedo and A. R. Figueiras-Vidal, "An adaptive beamforming technique based on cyclostationary signal properties," *IEEE Transactions on Signal Processing*, vol. 43, pp. 1637–1650, July 1995.

[17] C. B. Chang and M. Athans, "State estimation for discrete systems with switching parameters," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 14, pp. 418–424, May 1978.

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

[19] C. T. Chen, *Linear System Theory and Design*. New York: CBS College Publishing, 1984.

[20] Y. H. Chen and C. T. Chiang, "Adaptive beamforming using the constrained Kalman filter," *IEEE Transactions on Antennas and Propagation*, vol. 41, pp. 1576–1580, Nov. 1993.

[21] M. Chryssomomallis, "Smart antennas," *IEEE Antennas and Propagation Magazine*, vol. 42, pp. 129–136, June 2000.

[22] H. Cox, "Resolving power and senstivity to mismatch of optimum array processors," *Journal of Acoustic Society America*, vol. 54, pp. 771–785, Sept. 1973.

[23] H. Cox, R. M. Zeskind, and M. H. Owen, "Robust adaptive beamforming," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 35, pp. 1365–1376, Oct. 1987.

[24] S. Cui, M. Kisialiou, Z. Q. Luo, and Z. Ding, "Robust blind multiuser detection against signature waveform mismatch based on second-order cone programming," *IEEE Transactions on Wireless Communications*, vol. 4, pp. 1285–1291, July 2005.

[25] T. N. Davidson, Z. Q. Luo, and J. F. Strum, "Linear matrix inequality formulation of spectral mask constraints with applications to FIR filter design," *IEEE Transactions on Signal Processing*, vol. 50, pp. 2702–2715, Nov. 2002.

[26] T. Do-Hong and P. Russer, "Signal processing for wideband smart antenna array applications," *IEEE Microwave Magazine*, vol. 5, pp. 57–67, Mar. 2004.

[27] A. El-Keyi, T. Kirubarajan, and A. B. Gershman, "Robust adaptive beamforming based on the Kalman filter," *IEEE Transactions on Signal Processing*, vol. 53, pp. 3032–3041, Aug. 2005.

[28] ——, "Robust wideband beamforming based on worst-case performance optimization," in *Proceedings of the IEEE Workshop on Statistical Signal Processing*, Bordeaux, France, July 2005.

[29] ——, "A state-space approach to robust adaptive beamforming," in *Proceedings of the IEEE Workshop on Statistical Signal Processing*, Bordeaux, France, July 2005.

[30] ——, "A state-space approach to robust multiuser detection," in *Proceedings of the IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP'05)*, Peurto Vallarta, Mexico, Dec. 2005.

[31] ——, "Adaptive wideband beamforming with robustness against steering errors," in *Proceedings of the IEEE Workshop on Sensor Arrays and Multichannel Processing (SAM'06)*, Waltham, MA, July 2006.

[32] ——, "A state-space approach to robust multiuser detection," *IEEE Transactions on Wireless Communications*, submitted.

[33] W. C. Elmore and M. E. Heald, *Physics of Waves*. New York: Dover Publications, 1985.

[34] M. H. Er and A. Cantoni, "Derivative constraints for broad-band element space antenna array processors," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 31, pp. 1378–1393, Dec. 1983.

[35] G. Fabrizio, A. B. Gershman, and M. Turley, "Robust adaptive beamforming for HF surface wave over-the-horizon radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 40, pp. 510–525, Apr. 2004.

[36] D. D. Feldman and L. J. Griffiths, "A projection approach to robust adaptive beamforming," *IEEE Transactions on Signal Processing*, vol. 42, pp. 867–876, Apr. 1994.

[37] M. E. Frerking, *Digital Signal Processing in Communication Systems*. Norwell, MA: Kluwer Academic Publishers, 1994.

[38] W. F. Gabriel, "Spectral analysis and adaptive array superresolution techniques," *Proceedings of the IEEE*, vol. 68, pp. 654–666, June 1980.

[39] J. D. Geeter, H. V. Brussel, and J. D. Schutter, "A smoothly constrained Kalman filter," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, pp. 1171–1177, Oct. 1997.

[40] A. B. Gershman, "Robust adaptive beamforming in sensor arrays," *AEU – International Journal of Electronics and Communications*, vol. 53, pp. 305–314, Dec. 1999.

169

[41] ——, "High-resolution and robust signal processing," Y. Hua, A. B. Gershman, and Q. Cheng, Eds. Marcel Dekker, 2003, ch. 2: Robustness issues in adaptive beamforming and high-resolution direction finding, pp. 63–110.

[42] A. B. Gershman, C. F. Mecklenbräuker, and J. F. Böhme, "Matrix fitting approach to direction of arrival estimation with imperfect spatial coherence of wavefronts," *IEEE Transactions on Signal Processing*, vol. 45, pp. 1894–1899, July 1997.

[43] A. B. Gershman, U. Nickel, and J. F. Bohme, "Adaptive beamforming algorithms with robustness against jammer motion," *IEEE Transactions on Signal Processing*, vol. 45, pp. 1878–1885, July 1997.

[44] A. B. Gershman, G. V. Serebryakov, and J. F. Böhme, "Constrained Hung-Turner adaptive beamforming algorithm with additional robustness to wideband and moving jammers," *IEEE Transactions on Antennas and Propagation*, vol. 44, pp. 361–367, Mar. 1996.

[45] A. B. Gershman and S. Shahbazpanahi, "Robust blind multiuser detection for synchronous CDMA systems using worst-case performance optimization," *IEEE Transactions on Wireless Communications*, vol. 3, pp. 2232–2245, Nov. 2004.

[46] A. B. Gershman, V. I. Turchin, and V. A. Zverev, "Experimental results of localization of moving underwater signal by adaptive beamforming," *IEEE Transactions on Signal Processing*, vol. 43, pp. 2249–2257, Oct. 1995.

[47] L. C. Godara, "The effect of phase-shift errors on the performance of an antenna-array beamformer," *IEEE Journal of Oceanic Engineering*, vol. 10, pp. 395–409, July 1985.

[48] ——, "Application of the fast Fourier transform to broadband beamforming," *Journal of Acoustic Society America*, vol. 98, pp. 230–240, July 1995.

[49] ——, "Application of antenna arrays to mobile communications, part I: Performance improvement, feasibility, and system considerations," *Proceedings of the IEEE*, vol. 85, pp. 1031–1060, July 1997.

[50] ——, "Application of antenna arrays to mobile communications, part II: Beam-forming and direction-of-arrival considerations," *Proceedings of the IEEE*, vol. 85, pp. 1195–1245, Aug. 1997.

[51] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: The Johns Hopkins University Press, 1996.

[52] L. J. Griffths and C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Transactions on Antennas and Propagation*, vol. 30, pp. 27–34, Jan. 1982.

[53] J. R. Guerci, "Theory and application of covariance matrix tapers to robust adaptive beamforming," *IEEE Transactions on Signal Processing*, vol. 47, pp. 977–985, Apr. 2000.

[54] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[55] S. D. Hayward, "Effects of motion on adaptive arrays," *IEEE Proceedings on Radar, Sonar, and Navigation*, vol. 144, pp. 15–20, Feb. 1997.

[56] R. Hettich and K. O. Kortanek, "Semi-infinite programming: Theory, methods, and applications," *SIAM Review*, vol. 35, pp. 380–429, Sept. 1993.

[57] M. Honig, U. Madhow, and S. Verdú, "Blind adaptive multiuser detection," *IEEE Transactions on Information Theory*, vol. 41, pp. 944–960, July 1995.

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

[59] N. K. Jablon, "Adaptive beamforming with the generalized sidelobe canceller in the presence of array imperfections," *IEEE Transactions on Antennas and Propagation*, vol. 334, pp. 996–1012, Aug. 1986.

[60] S. Kapoor, S. Gollamudi, S. Nagaraj, and Y. F. Huang, "Adaptive multiuser detection and beamforming for interference suppression in CDMA mobile radio systems," *IEEE Transactions on Vehicular Technology*, vol. 48, pp. 1341–1355, Sept. 1999.

[61] S. A. Kassam and H. V. Poor, "Robust techniques for signal processing: A survey," *Proceedings of the IEEE*, vol. 73, pp. 433–481, Mar. 1985.

[62] J. W. Kim and C. K. Un, "An adaptive array robust to beam pointing error," *IEEE Transactions on Signal Processing*, vol. 40, pp. 1582–1584, June 1992.

[63] S. M. Kogon, "Robust adaptive beamforming for passive sonar using eigenvector/beam association and excision," in *Proceedings of the IEEE Workshop on Sensor Array and Multichannel Signal Processing (SAM'02)*, Rosslyn, VA, Aug. 2002.

[64] E. Kreyszig, *Advanced Engineering Mathematics*, 7th ed. New York: John Wiley & Sons, 1993.

[65] H. Krim and M. Viberg, "Two decades of array signal processing research: The parametric approach," *IEEE Signal Processing Magazine*, vol. 13, pp. 67–94, July 1996.

[66] J. Krolik, "The performance of matched-field beamformers with mediterranean vertical array data," *IEEE Transactions on Signal Processing*, vol. 44, pp. 2605–2611, Oct. 1996.

[67] R. T. Lacoss, "Adaptive combining of wideband array data for optimal reception," *Transactions on Geoscience Electronics*, vol. GE-6, pp. 78–86, May 1968.

[68] J. H. Lee and Y. T. Lee, "Robust adaptive array beamforming for cyclostationary signals under cycle frequency error," *IEEE Transactions on Antennas and Propagation*, vol. 47, pp. 233–241, Feb. 1999.

[69] J. H. Lee, Y. T. Lee, and W. H. Shih, "Efficient robust adaptive beamforming for cyclostationary signals," *IEEE Transactions on Signal Processing*, vol. 48, pp. 1893–1901, July 2000.

[70] J. Li, P. Stoica, and Z. Wang, "On robust Capon beamforming and diagonal loading," *IEEE Transactions on Signal Processing*, vol. 51, pp. 1702–1715, July 2003.

[71] F. Ling, R. Love, M. M. Wang, T. Brown, P. Fleming, and H. Xu, "Behavior and performance of power controlled IS-95 reverse-link under soft handoff," *IEEE Transactions on Vehicular Technology*, vol. 49, pp. 1697–1704, Sept. 2000.

[72] J. Litva and T. K. Y. Lo, *Digital Beamforming in Wireless Communications*. London, UK: Artech House, 1996.

[73] M. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear Algebra and its Applications*, vol. 284, pp. 193–228, 1998.

[74] R. Lorenz and S. P. Boyd, "Robust minimum variance beamforming," *IEEE Transactions on Signal Processing*, vol. 53, pp. 1684–1696, May 2005.

[75] F. Lu, E. Milios, S. Stergiopoulos, and A. Dhanantwari, "New towed-array shape-estimation scheme for real-time sonar systems," *IEEE Journal of Oceanic Engineering*, vol. 28, pp. 552–563, July 2003.

[76] U. Madhow and M. Honig, "MMSE interference suppression for direct-sequence spread-spectrum CDMA," *IEEE Transactions on Communications*, vol. 42, pp. 3178–3188, Dec. 1994.

[77] A. Manikas and C. Proukakis, "Modeling and estimation of ambiguities in linear arrays," *IEEE Transactions on Signal Processing*, vol. 46, pp. 2166–2179, Aug. 1998.

[78] D. G. Manolakis, V. K. Ingle, and S. M. Kogon, *Statistical and Adaptive Signal Processing*. Boston, MA: McGraw Hill, 2000.

[79] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays*. New York: John Wiley & Sons, 1980.

[80] P. Moulin, M. Anitescu, K. O. Kortanek, and F. A. Porta, "The role of linear semi-infinite programming in signal adapted QMF bank design," *IEEE Transactions on Signal Processing*, vol. 45, pp. 2160–2174, Sept. 1997.

[81] G. V. Moustakides, "Decision directed algorithms for multiuser detection," in *Proceedings of the IEEE International Symposium on Information Theory*, Washington, D.C., June 2001, p. 8.

[82] P. S. Naidu, *Sensor Array Signal Processing.* Boca Raton, FL: CRC Press, 2001.

[83] Y. Nesterov and A. Nemirovsky, *Interior-Point Polynomial Methods in Convex Programming.* SIAM, Philadelphia, PA: volume 13 of *Studies in Applied Mathematics*, 1994.

[84] O. L. Frost III, "An algorithm for linearly constrained adaptive antenna array processing," *Proceedings of the IEEE*, vol. 60, pp. 926–935, Aug. 1972.

[85] K. Pahlavan and A. Levesque, *Wireless Information Networks.* New York: John Wiley & Sons, 1995.

[86] T. W. Parks and C. S. Burrus, *Digital Filter Design.* New York: John Wiley & Sons, 1987.

[87] K. I. Pederson, P. E. Mogensen, and B. H. Fleury, "Spatial channel characterstics in outdoor environments and their impact on BS antenna system performance," in *Proceedings of the Vehicular Technology Conference (VTC'98)*, Ottawa, Canada, May 1998, pp. 719–723.

[88] H. V. Poor and G. W. Wang, "Code-aided interference suppression for DS/CDMA communications–Part II: Parallel blind adaptive implementations," *IEEE Transactions on Communications*, vol. 45, pp. 1112–1122, Sept. 1997.

[89] V. U. Reddy, A. Paulraj, and T. Kailath, "Adaptive beamforming for coherent signals and interference," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, pp. 527–536, June 1985.

[90] B. R. Reed and J. Strauss, "A short proof of the equivalence of LCMV and GSC beamforming," *IEEE Signal Processing Letters*, vol. 9, pp. 168–169, June 2002.

175

[91] I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 10, pp. 853–863, Nov. 1972.

[92] L. S. Resende, J. M. T. Romano, and M. G. Bellanger, "A fast least-squares algorithm for linearly constrained adaptive filtering," *IEEE Transactions on Signal Processing*, vol. 44, pp. 1168–1174, May 1996.

[93] F. M. Reza, *An Introduction to Information Theory*. New York: McGraw Hill, 1961.

[94] J. Riba, J. Goldberg, and G. Vazquez, "Robust beamforming for interference rejection in mobile communications," *IEEE Transactions on Signal Processing*, vol. 45, pp. 271–275, Jan. 1997.

[95] Y. Rong, S. Vorobyov, A. B. Gershman, and N. D. Sidiropoulos, "Blind spatial signature estimation via time-varying user power loading and parallel factor analysis," *IEEE Transactions on Signal Processing*, vol. 53, pp. 1697–1710, May 2005.

[96] S. G. Schock, "A method for estimating the physical and acoustic properties of the sea bed using chirp sonar data," *IEEE Journal of Oceanic Engineering*, vol. 29, pp. 1200–1217, Oct. 2004.

[97] J. B. Schodorf and D. B. Williams, "A constrained optimization approach to multiuser detection," *IEEE Transactions on Signal Processing*, vol. 45, pp. 259–262, Jan. 1997.

[98] D. P. Scholnik and J. O. Coleman, "Optimal design of wideband array patterns," in *Proceedings of the IEEE International Radar Conference*, Washington, DC, May 2000, pp. 172–177.

[99] P. M. Shearer, *Introduction to Seismology*. Cambridge, UK: Cambridge University Press, 1999.

[100] I. Solomon, D. Gray, Y. Abramovich, and S. Anderson, "Over-the-horizon radar array calibration using echoes from ionised meteor trails," *IEE Proceedings - Radar, Sonar and Navigation*, vol. 145, pp. 173–180, June 1998.

[101] S. Strauss, *Distributed Sensor Networks*. Boca Raton, FL: CRC Press, 2004.

[102] J. F. Strum, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimzation Methods and Software*, vol. 11-12, pp. 625–653, Aug. 1999.

[103] H. Tang, K. M. Wong, A. B. Gershman, and S. Vorobyov, "Blind adaptive beamforming for cyclostationary signals with robustness against cycle frequency mismatch," in *Proceedings of the IEEE Workshop on Sensor Array and Multichannel Signal Processing (SAM'02)*, Rosslyn, VA, Aug. 2002.

[104] B. N. Thanh, V. Kirshnamurthy, and R. J. Evans, "Detection-aided recursive least squares adaptive multiuser detection," *IEEE Signal Processing Letters*, vol. 9, pp. 229–232, Aug. 2002.

[105] I. Thng, A. Cantoni, and Y. Leung, "Derivative constrained optimum broadband antenna arrays," *IEEE Transactions on Signal Processing*, vol. 41, pp. 2376–2388, July 1993.

[106] ——, "Constraints for maximally flat optimum broadband antenna arrays," *IEEE Transactions on Signal Processing*, vol. 43, pp. 1334–1347, June 1995.

[107] J. K. Thomas, L. L. Scharf, and D. W. Tufts, "The probability of a subspace swap in the SVD," *IEEE Transactions on Signal Processing*, vol. 43, pp. 730–736, Mar. 1995.

[108] Z. Tian, K. L. Bell, and H. L. V. Trees, "A recursive least squares implementation for LCMP beamforming under quadratic constraint," *IEEE Transactions on Signal Processing*, vol. 49, pp. 1138–1145, June 2001.

[109] H. L. V. Trees, *Optimum Array Processing*. New York: John Wiley & Sons, 2002.

[110] G. V. Tsoulos, Ed., *Adaptive Antennas for Wireless Communications*. Piscataway, NJ: IEEE Press, 2001.

[111] R. J. Vaccaro, "The past, present, and future of underwater acoustic signal processing," *IEEE Signal Processing Magazine*, vol. 15, pp. 21–51, July 1998.

[112] B. D. V. Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE Acoustics, Speech, and Signal Processing Magazine*, vol. 5, pp. 4–24, Apr. 1988.

[113] S. Verdú, *Advances in Statistical Signal Processing Vol. 2: Signal Detection*. Greenwich, CT: JAI Press, 1993, ch. Multiuser detection, pp. 369–409.

[114] ——, *Multiuser Detection*. New York: Cambridge University Press, 1998.

[115] S. Vorobyov, A. B. Gershman, Z.-Q. Luo, and N. Ma, "Adaptive beamforming with joint robustness against mismatched signal steering vector and interference nonstationarity," *IEEE Signal Processing Letters*, vol. 11, pp. 108–111, Feb. 2004.

[116] S. A. Vorobyov, A. B. Gershman, and Z. Q. Luo, "Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem," *IEEE Transactions on Signal Processing*, vol. 51, pp. 313–324, Feb. 2003.

[117] X. Wang and H. V. Poor, "Blind multiuser detection: A subspace approach," *IEEE Transactions on Information Theory*, vol. 44, pp. 677–690, Mar. 1998.

[118] Z. Wang, J. Li, P. Stoica, T. Nishida, and M. Sheplak, "Constant-beamwidth and constant-powerwidth wideband robust Capon beamformers for acoustic imaging," *Journal of Acoustic Society America*, vol. 116, pp. 1621–1631, Sept. 2004.

[119] S. Werner, J. A. Apolinário, and M. L. R. de Campos, "On equivalence of RLS implementaitons of LCMV and GSC processors," *IEEE Signal Processing Letters*, vol. 10, pp. 356–359, Dec. 2003.

[120] B. Widrow, K. M. Duvall, R. P. Gooch, and W. C. Newman, "Signal cancellation phenomena in adaptive antennas: causes and cures," *IEEE Transactions on Antennas and Propagation*, vol. 30, pp. 469–478, May 1982.

[121] S. Q. Wu and J. Y. Zhang, "A new robust beamforming method with antennae calibration errors," in *Proceedings of the IEEE Wireless Communications and Networking Conference*, vol. 2, New Orleans, LA, Sept. 1999, pp. 869–872.

[122] A. Yener, R. D. Yates, and S. Ulukus, "Interference management for CDMA systems through power control, multiuser detection, and beamforming," *IEEE Transactions on Communications*, vol. 49, pp. 1227–1239, July 2001.

[123] K. Zarifi, S. Shahbazpanahi, A. B. Gershman, and Z. Q. Luo, "Robust blind multiuser detection based on the worst-case performance optimization of the MMSE receiver," *IEEE Transactions on Signal Processing*, vol. 53, pp. 295–305, Jan. 2005.

[124] S. Zhang and I. Thng, "Robust presteering derivative constraints for broadband antenna arrays," *IEEE Transactions on Signal Processing*, vol. 50, pp. 1–10, Jan. 2002.

[125] X. D. Zhang and W. Wei, "Blind adaptive multiuser detection based on Kalman filtering," *IEEE Transactions on Signal Processing*, vol. 50, pp. 87–95, Jan. 2002.