1996

Blind adaptive near-far resistant receivers for DS/CDMA multi-user communication systems

Sang Chul Park
Iowa State University

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

Blind adaptive near-far resistant receivers for DS/CDMA multi-user communication systems

by

Sang Chul Park

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Electrical Engineering (Communications and Signal Processing)

Major Professor: John F. Doherty

Iowa State University

Ames, Iowa

1996

Graduate College
Iowa State University

This is to certify that the Doctoral dissertation of

Sang Chul Park

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Committee Member

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program

Signature was redacted for privacy.

For the Graduate College

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to express my gratitude to Dr. John F. Doherty, for his support, guidance and encouragement throughout my Ph.D. work. I am very grateful for his confidence in my work and the innumerable discussions we have had throughout this study.

I would like to extend my appreciation to the committee members, Dr. Satish S. Udpa, Dr. William Q. Meeker, Dr. Lalita Udpa, Dr. Hsiu C. Han and Dr. Julie A. Dickerson for sharing their time and knowledge towards the success of this work.

I would also like to acknowledge the support from the Republic of Korean Army and the U. S. Army Research Office (Grant number : DAAH04-95-1-0342), without which this Ph.D. research would not have been conceived.

Very special thanks go to my wife, Mi-Young, and my sons, Daehee and Gunhee, for their constant faith and sacrifice throughout my study in the United States. Words cannot express my gratitude to them. I dedicate this work to my beloved family.

ABSTRACT

Code-division multiple-access (CDMA) systems have multiple users that simultaneously share a common channel using pre-assigned signature waveforms. The conventional receiver suffers from the *near-far problem* when the received signal power of the desired user is weaker than those of the other users. Optimum and suboptimum multi-user detectors outperform the conventional single-user receiver at the expense of a significant increase in complexity and need for side-information about interfering users. Complexity of these detectors may not be acceptable for many practical applications and communication security may restrict the distribution of all users' signature waveforms to all the receivers.

For a single-user receiver, the multi-user detection problem can be viewed as an interference suppression problem. This dissertation presents a *cost-constraint strategy* to implement adaptive single-user receivers that suppress the multiple-access interference without using training sequences. A constrained LMS algorithm that converges to a near-optimum solution by using the received signal and some known properties of the desired signal is developed. The constrained LMS receiver can be used for static CDMA detection where the channel accessed by the desired user is time-invariant. The dissertation also develops an adaptive space-alternating generalized EM (SAGE) algorithm. This algorithm jointly updates estimates of filter weights and adaptive reference signal in a sequential manner. The SAGE receiver outperforms the existing blind receiver that employ the constrained output-power-minimizing algorithm while using the same amount of information. The SAGE receiver can be used for dynamic CDMA detec-

tion where the channel accessed by the desired signal is time-varying. The dissertation further generalizes the adaptive SAGE algorithm to an adaptive space-alternating generalized projection (SAGP) algorithm that uses the same amount of information as in the conventional receiver.

Proposed receivers are tested by Monte-Carlo simulations and compared with the existing receivers that use the same amount of information. Throughout the analytical analysis and simulations of the proposed receivers, the dissertation shows that, for realistic CDMA communications, achieving both the near-far resistance and the near-optimum performance is possible with the same or similar information required by the conventional receiver.

CHAPTER 1 INTRODUCTION

Overview

The electromagnetic spectrum is an invaluable but limited natural resource for wireless communications. Increasing demands of military, commercial, and private users on the electromagnetic spectrum requires an efficient spectrum channel sharing strategy. Ever since the second pair of the transceiver came into existence in the wireless channel, communication system designers have confronted the problem of multiple-access channel sharing with mutual interference. The system block diagram of a wireless multiple-access communication channel is shown in Figure 1.1. Static strategies, such as frequency-division multiple-accessing (FDMA) and time-division multiple-accessing (TDMA), by which the multiple-access channel is effectively partitioned into independent single-user subchannels, tend to be wasteful in applications where most users actively send information sporadically. Dynamic channel sharing strategies, which allow the active users a larger share of the channel while they are transmitting, fit into two categories: random-access communication and simultaneous transmission systems [1]. In random-access communication, it is assumed that the receiver cannot demodulate more than one simultaneous transmission, and so the problem is to design protocols to schedule channel access at non-overlapping times, and if collisions between messages occur to ensure that those messages are eventually retransmitted successfully. Simultaneous transmission systems differ from static strategies and random access protocols in that users are allowed to demodulate all (as in the satellite communications) or a subset (as in multipoint-to-

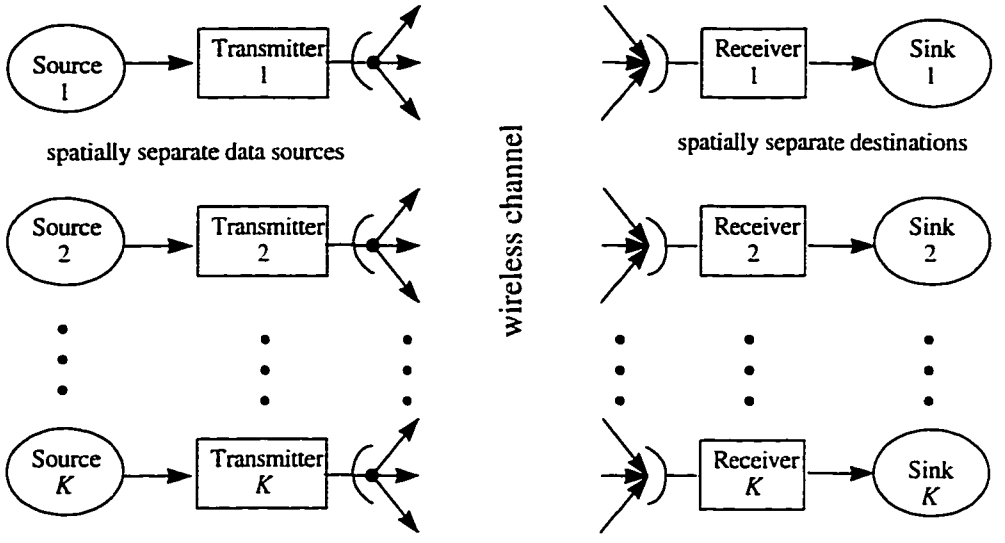

Figure 1.1 A wireless multiple-access channel description

multipoint topologies) of the transmitted messages.

A major multiple-access strategy using the simultaneous transmission philosophy is code-division multiple-accessing (CDMA) and it has become a main candidate for the next generation of mobile land and satellite communication systems. In CDMA communication systems, each transmitter generates a spread spectrum signal by modulating a data signal onto a pseudo-random signature waveform so that the resultant signal has a bandwidth much larger than the data signal bandwidth. Unlike FDMA or TDMA, CDMA has multiple users simultaneously sharing the same wide-band channel. If a CDMA system is viewed in either the frequency or time domain, the multiple-access signals appear to co-exist. Frequency and time domain representation of FDMA, TDMA and CDMA is described in Figure 1.2.

The conventional CDMA receiver recovers the information of the desired user by correlating the received signal with a replica of the signature waveform assigned to the desired user, i.e., a signature matched filtering. As is well-known, when the received signal is corrupted by only additive white Gaussian noise (AWGN), the conventional

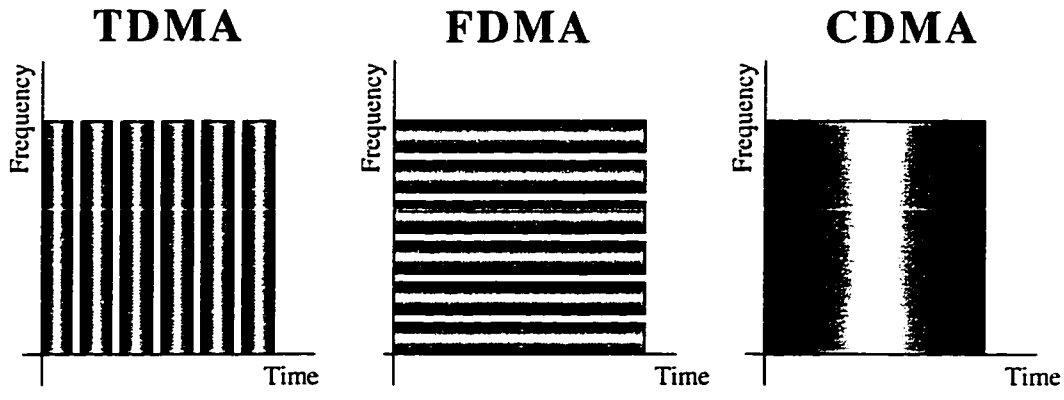

Figure 1.2 Graphical description of TDMA, FDMA and CDMA

matched filter receiver minimizes the error probability. This is not true in the conventional CDMA receiver, whose decision variables for the signal of a particular user are corrupted by *multiple-access interference* (MAI) in addition to AWGN. The MAI originates from crosscorrelations between the signature waveform of the desired signal and signals of other active users. When the received signal power of the desired user is relatively weaker than those of the other users, the conventional receiver is unable to reliably recover the information sent by the desired user, even if the signature waveforms have very low crosscorrelations. This is known as the *near-far problem* [2].

The current approach to dealing with the near-far problem is to use transmitter power control [3]. Another alternative is to use some form of a multi-user detector. Many different optimum/suboptimum structures of the multi-user detector have been proposed in the literature [2, 4, 5, 6, 7, 8, 9, 10, 11]. Multi-user detectors are generally characterized by centralized detection that demodulates all the users' signals at the output of a matched filter bank. Substantial performance gains can be achieved in coherent multi-user systems by using a multi-user detector that takes advantage of the structure of the CDMA signals [2]. The multi-user detectors outperform the conventional receiver at the expense of a significant increase in complexity. For example, the complexity of

the optimum multi-user detector in [2, 4] grows exponentially in the number of users. Less complex suboptimum multi-user detectors in [5, 6] linearly increase in complexity with the number of users. In addition to complexity, multi-user detectors also require large amounts of side-information about the received signal, which includes the number of users, the signature waveforms, associated time delays, and phase offsets of all active CDMA signals. Often the received amplitude of each CDMA signal is also needed. The complexity of such multi-user detectors may be unacceptably high for many practical applications or communication security restrict the distribution of all users' signature waveforms to all the receivers, or both. Furthermore, some information of relatively weak signals is likely to be more difficult to obtain due to the time-varying nature of the channel.

This dissertation considers fully decentralized single-user detection, in which the receiver is constrained to demodulate the signal of only one user, but unlike the conventional receiver, is optimized to take into account the structure of the CDMA signal. The decentralized detection approach views the multi-user detection problem as an interference suppression problem, where at a particular receiver one signal is considered the desired signal and the other signals are considered the interferences. One such single-user receiver is the linear minimum mean-squared-error (MMSE) receiver proposed in [12]. This MMSE receiver has been particularly attractive because it lends itself for adaptive implementation.

Several adaptive MMSE receivers have been proposed recently [12, 13, 14]. Although reducing the complexity and eliminating requirement of the information of the interfering signals, most of the adaptive MMSE receivers require training sequences for the implementation of the least-mean-square (LMS) algorithm both at the beginning and during data transmission. That is, adaptive MMSE receivers need to switch back and forth between a training mode and a decision-directed mode during actual data transmission as a new strong user accesses the system.

Use of a blind scheme, i.e., one that does not rely on a training sequence for adaptation, has been practically attractive for high-speed data transmission over a communication channel [15]. If the receiver yields a bit-error-rate (BER) less than 10^{-1} errors per bit, a decision-directed LMS algorithm may be an option for a blind receiver. However when the system experiences the near-far problem, detection capability of a non-optimum receiver is completely lost and the algorithm may suffer from the misconvergence to a local minimum associated with a strong interfering signal, i.e., the receiver may adapt its parameters to detect the signal of a strong interfering user instead of the signal of the desired user. Honig, Madhow and Verdu in [16] and Schodorf and Williams in [17] have proposed constrained output-power minimizing (OPM) receivers. They implement algorithms for a blind adaptive receiver through gradient projection (GP) algorithms. Algorithms in these receivers adjust the filter tap weights by minimizing the output power while constraining the gradient of the cost function to satisfy a prescribed constraint.

It should be noted that the inclusion of the desired signal component (e.g., as in the output power) during the adaptation process results in the signal cancellation phenomenon [18], which increases the steady state BER. This phenomenon occurs because the algorithm does not guarantee the zero filter tap gain increment even when the optimum weight vector is achieved. As a remedy, the algorithm in [16] has been suggested to switch from a blind mode to a decision-directed mode after convergence. As is well-known, this requirement of switching is unrealistic in practical applications. In addition, unlike the conventional or decision-directed LMS algorithm, the constrained adaptive algorithm in the decision-directed mode does not converge to near-optimum steady state without properly scaling the estimated reference signal.

Blind equalization has been used to implement a linear adaptive receiver that suppresses the intersymbol interference (ISI) caused by the unknown linear single-user channel [19]. Global convergence of the blind equalization algorithms is highly dependent

on good initialization, i.e., to start with a parameter setting that is within the region of attraction of a desirable minimum. Because of lack of good initialization procedures, existing blind equalization schemes have not been considered for the CDMA detection in the near-far communication environment.

This dissertation presents a *cost-constraint strategy* to implement a blind adaptive algorithm that avoids convergence to undesirable local minima and ensures near-optimum steady-state performance without using a training sequence. With one extra information on the desired signal other than those of existing blind algorithms, the dissertation introduces the constrained LMS (CLMS) algorithm that minimizes the mean-squared-error (MSE) between the modulus of the filter outputs and the amplitude of the desired signal. To address the sensitivity of the implementation error in the CLMS algorithm, a robust constraint in the form of a filter norm constraint is also incorporated. A linear receiver that implements the CLMS algorithm is useful for the single-user detection in a static CDMA multi-user communication channel where the channel accessed by the desired user is time-invariant. The dissertation also develops an adaptive space-alternating generalized expectation-maximization (SAGE) algorithm. This algorithm jointly updates estimates of filter weights and the adaptation reference in a sequential manner. The adaptive SAGE algorithm is applicable to single-user detection in non-stationary CDMA multi-user communication environments where the channel accessed by the desired signal is time-varying. The receiver that implements the adaptive SAGE algorithm uses the same amount of information required by the existing blind receivers of [16] and [17]. The dissertation further generalizes the adaptive SAGE algorithm to an adaptive space-alternating generalized-projection (SAGP) algorithm. The receiver that implements the adaptive SAGP algorithm requires only the information used in the conventional receiver. Performances of the proposed receivers are tested by an extensive series of Monte-Carlo simulations and compared with the receivers that use the same or similar amount of information.

Objective and Scope of Dissertation Study

The objectives of the dissertation study can be summarized as:

- Provide an overview of multi-user detection schemes and MAI suppression techniques that improve performance of CDMA systems,
- Analyze the characteristics of the CDMA signals and some fundamental limitations of the linear detection schemes and
- Develop practical adaptive algorithms for a linear single-user receiver that improve performance of asynchronous CDMA systems over the previous works.

In particular, the dissertation will focus on development of blind adaptive receivers that converge to a desirable optimum even in the near-far situation. Performances of the receivers in the transient mode are also investigated. The dissertation also analyzes the characteristics of CDMA signals and the performance limitations of a linear filtering scheme. From this analysis, the upper-bound of channel capacity, asymptotic efficiency and the optimum MMSE performance are formulated in terms of the number of CDMA signals, the crosscorrelations between the signature waveforms, the input SNR and the input signal-to-interference (SIR) for each interfering signal. In the implementation of blind adaptive receivers, it is assumed that the receiver is equipped with the knowledge on the signature sequence and the associated time delay of the desired user, which is the same as in the conventional receiver. In the constrained LMS approach, it is further assumed that the exact or a good estimate of the amplitude for the desired user's signal is available and used for the adaptation reference. The effect of implementation errors in the constrained adaptive algorithm will be investigated and a solution will be introduced. Assumption on the *a priori* known amplitude of the desired signal is released by the adaptive SAGE receiver, which jointly estimates both the filter weights and the

amplitude of the desired signal while performing the detection operation. The dissertation further generalizes the LMS stochastic gradient descent algorithm to a generalized projection algorithm, where the data detection relies only on the received signal and some properties of the desired signal. Therefore, the projection-based receiver is robust to the unknown or time-varying nature of the received signal. Throughout the analytical analysis and simulation of proposed receivers, the dissertation shows that, for realistic DS/CDMA communications, achieving near-optimum performance and near-far resistance are possible with the same or similar information required by the conventional receiver.

Dissertation Outline

The following chapter mathematically describes a standard DS/CDMA communication system and gives an overview of previous work on multi-user detection schemes and adaptive MAI suppression techniques. In particular, the overview reveals that existing detection schemes either are unrealistic due to computational complexity and limitation in obtaining the required information or should be improved for reliable data transmission.

Chapter 3 describes characteristics of DS/CDMA signal structure and presents several classes of linear optimum filters. The chapter also explores the fundamental limitations of the linear optimum filter through a geometric analysis of the linear MMSE optimum solution.

Chapter 4 presents a cost-constraint strategy that leads a nonconvex gradient descent algorithm to avoid convergence to undesirable local minima. A linearly-constrained LMS algorithm is introduced. Practical considerations on the constrained adaptive algorithms are also investigated. A filter gain maximization problem is formulated based on the cost-constraint strategy. The constrained LMS receiver is implemented for a blind adaptive

receiver. Convergence and the steady state performance of the receiver are analyzed. Error-correcting feature of the constrained LMS algorithm is also presented through the geometric interpretation of the constrained adaptive operation.

Chapter 5 explores an iterative joint parameter estimation approach for implementing a blind adaptive receiver in nonstationary communication environments. In this chapter, it is assumed that both the filter weights and the adaptation reference signal are unknown. The adaptive algorithm jointly updates estimates of the filter weights and the reference as a new received signal inputs to the receiver. The space-alternating generalized EM (SAGE) algorithm, which is used for the iterative maximum likelihood (ML) parameter estimation in the statistical estimation, are introduced. An adaptive SAGE algorithm, the stochastic approximation of the SAGE algorithm, is presented and applied for the implementation of a blind adaptive near-far resistant receiver.

Chapter 6 explores the projection-based approach to implement an algorithm for a blind adaptive near-far resistant receiver. This receiver generalizes the adaptive SAGE receiver and does not require any other information than that of the conventional receiver. The method of generalized projection for signal restoration applications is introduced. This chapter formulates an adaptive space-alternating generalized-projection (SAGP) algorithm by stochastically approximating and combining the generalized projection method and the SAGE algorithm. In the implementation of the algorithm , the dissertation views every received signal as data constraint sets whereas property constraint sets are obtained from the properties of the desired signal. The data constraint sets are used to specify that the response of the filter should be distortion-free with respect to the information sent by the desired user whereas the property constraint sets are used as remedial measures that continually refine the filter weights in the direction of the desirable optimum. Computational complexity and convergence property of the receiver are also analyzed.

Chapter 7 presents the performances of the proposed receivers in the near-far situ-

ations. The ensemble averaged output signal-to-interference ratio (SIR) curves in the transient mode are plotted. The steady state BER performances of the algorithms vs. various input parameters are plotted and discussed through extensive series of the Monte-Carlo simulations. Finally, Chapter 8 concludes the dissertation work.

CHAPTER 2 CDMA MULTI-USER COMMUNICATION SYSTEM

This chapter describes a code-division multiple-access (CDMA) communication system model and briefly overviews previous works on multi-user detection and adaptive MAI suppression techniques. A CDMA system allows multiple transmitters to share the same wide-band channel by use of different codes, or signature sequences, to distinguish the signals at the receiver. CDMA systems can be classified as frequency-hopped CDMA systems (FH/CDMA) or direct-sequence CDMA systems (DS/CDMA) [20]. An FH/CDMA transmitter varies its RF carrier at regular intervals as prescribed by a frequency-hopping pattern whereas a DS/CDMA transmitter phase-shift-keys its RF carrier with a signature sequence with very high pulse rate. This dissertation only considers DS/CDMA communication systems.

System Model

An asynchronous DS/CDMA communication system model is mathematically formalized to clarify problems in the DS/CDMA detection. Figure 2.1 shows a block diagram of a standard DS/CDMA communication system, where K users are transmitting individually spread spectrum signals, $u_k(t)$, $k = 1, \cdots, K$ simultaneously and in the same frequency band. This system can also be considered as the reverse link or uplink of a single cell in a cellular mobile communication system [21].

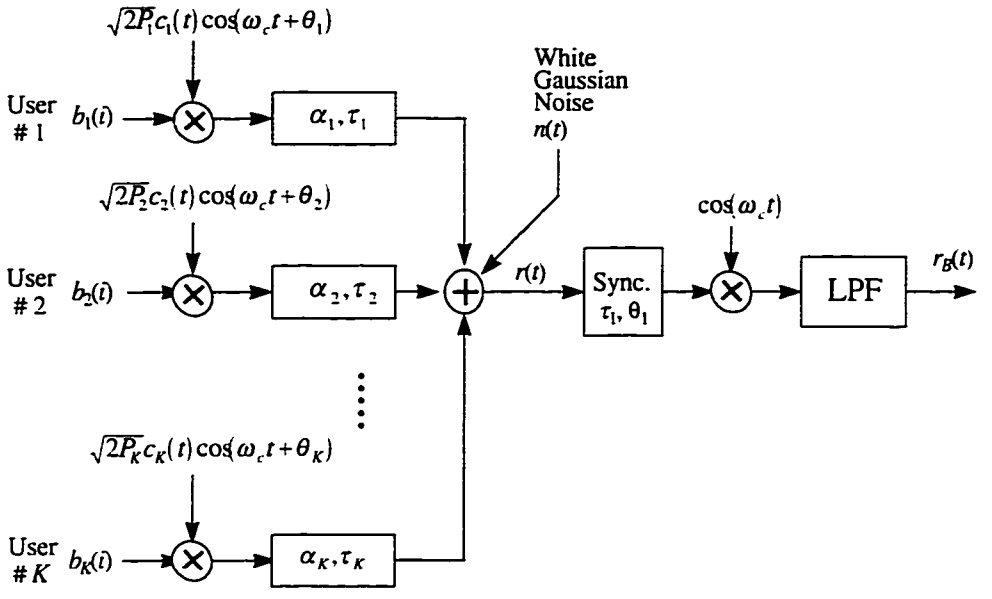

Figure 2.1 A standard DS/CDMA communication system

Transmitted Signals

Each DS/CDMA transmitter generates a spread spectrum signal by modulating data signal onto a wide-band carrier so that the resultant signal has a bandwidth much larger than the data signal bandwidth and is relatively insensitive to the data signal [22]. For simplicity, it is assumed that the data is binary and the signature waveform is BPSK modulated. The kth user's transmitted signal $u_k(t)$ can be written as

$$u_k(t) = \sum_{i=-\infty}^{\infty} \sqrt{2P_k} b_k(i)\, c_k(t - iT) \cos(\omega_c t + \theta_k), \quad k = 1, \cdots, K \tag{2.1}$$

where ω_c is the common carrier frequency, $P_k, b_k(i) \in \{+1, -1\}$, $c_k(t)$ and θ_k are the transmitted power, data bit, signature waveform, and phase offset of the kth user, respectively. The signature waveform $c_k(t)$ can be written as

$$c_k(t) = \sum_{n=0}^{N-1} c_k(n)\, \psi_{T_c}(t - nT_c), \quad k = 1, \cdots, K, \quad 0 \le t < T \tag{2.2}$$

where $c_k(n) \in \{+1, -1\}$ is the nth spreading code bit of the kth user, which has periodicity $N = T/T_c$ for all users, and $\psi_{T_c}(t)$ is the rectangular waveform with unity

energy defined on $[0, T_c)$. Signature waveforms are typically chosen to be linearly independent and to have good crosscorrelation properties so that it reduces multiple-access interference and unintended detection.

Received Signal

It is assumed that the location of each transmitter is different so that the receiving power and time delay of each signal are dissimilar to a particular receiver. Without loss of any generality, let us restrict our attention to the receiver for the signal of the user 1. It is assumed that the receiver is synchronized to the signal of the user 1 (i.e., $\tau_1 = 0$ and $\theta_1 = 0$) by employing a synchronization tracking algorithm. We further assume that the relative time delay $\tau_k \in [0, T)$ and relative phase offset $\phi_k \in [0, 2\pi)$ for the user $k = 1, \cdots, K$ are mutually independent. For simplicity, we also assume that the channel is distortionless, which implies that the kth user's signal at the receiver is proportional to a scaled and delayed version of its transmitted signal, i.e.,

$$r_k(t) = \alpha_k u_k(t - \tau_k), \quad k = 1, \cdots, K, \tag{2.3}$$

where α_k is a channel gain (which may be less than one) and τ_k is a transmission delay.

The received signal is a mixture of signals from K users embedded in additive white Gaussian noise (AWGN):

$$\begin{aligned} r(t) &= \sum_{k=1}^{K} r_k(t) + v(t) \\ &= \sum_{k=1}^{K} \sum_{i=-\infty}^{\infty} \alpha_k \sqrt{2P_k} b_k(i) c_k(t - iT - \tau_k) \cos(\omega_c t + \phi_k) + v(t) \end{aligned} \tag{2.4}$$

where $\phi_k \triangleq \theta_k - \omega_c \tau_k$, $v(t)$ is zero-mean AWGN. The baseband signal can be obtained by RF downconverting and lowpass filtering the received signal in Eq.(2.4):

$$r_B(t) = \sum_{k=1}^{K} \sum_{i=-\infty}^{\infty} A_k b_k(i) c_k(t - iT - \tau_k) + n(t). \tag{2.5}$$

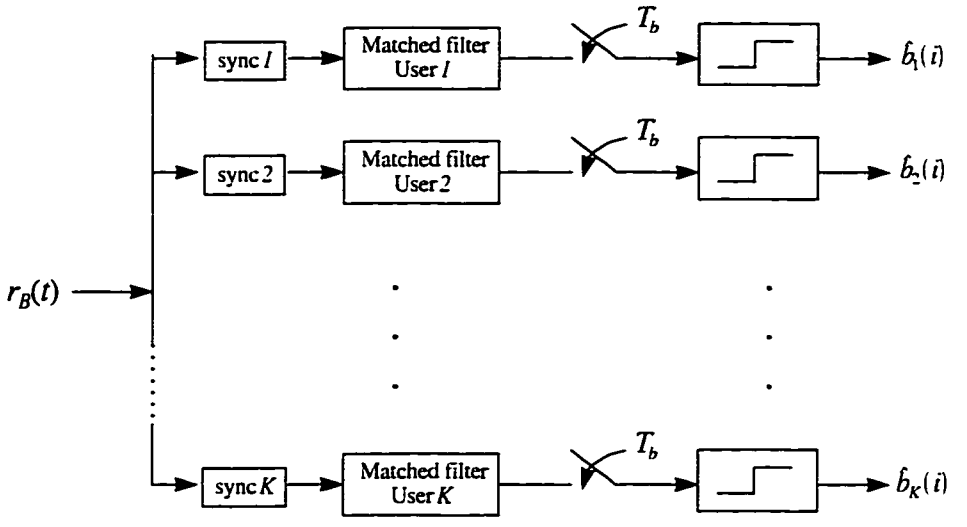

Figure 2.2 The conventional multi-user detector for the DS/CDMA system.

where $A_k = \alpha_k\sqrt{2P_k}\cos(\phi_k)$ for $k = 1, \cdots, K$ denotes the amplitude of the kth user's signal at the receiver and $n(t)$ denotes the baseband AWGN with two-sided power spectral density of σ^2. This simple model is sufficient to illustrate the demodulation difficulties that the receiver encounters.

Conventional Receiver

The conventional receiver demodulates the desired data by correlating the received signal with a replica of the signature waveform of the desired user. This conventional receiver demodulates each user's signal by matched filtering as if it were the only one present. The conventional multi-user detector consists of a bank of filters matched to each user's signature waveform in corresponding time and phase synchronism as shown in Figure 2.2. The decision scheme of the conventional receiver for the user 1 is given by

$$b_1(i) = \text{sgn}[z_1(i)] \tag{2.6}$$

where sgn $(\cdot)$ denotes a simple sign operator and the decision statistic $z_1(i)$ can be expressed by

$$z_1(i) = \int_{iT}^{iT+T} r_B(t)\, c_1(t)\, dt, \tag{2.7}$$

$$\begin{aligned} = {} & A_1 b_1(i) \int_{iT}^{iT+T} c_1^2(t)\, dt \\ & + \sum_{k=2}^{K} A_k b_k(i-1) \int_{iT}^{iT+\tau_k} c_k(t-\tau_k+T)\, c_1(t)\, dt \\ & + \sum_{k=2}^{K} A_k b_k(i) \int_{iT+\tau_k}^{iT+T} c_k(t-\tau_k)\, c_1(t)\, dt \\ & + \int_{iT}^{iT+T} n(t)\, c_1(t)\, dt. \end{aligned} \tag{2.8}$$

The first term of the right-hand side in Eq.(2.8) shows the desired signal that gives the exact information on $b_1(i)$, the second and third terms indicate MAI due to the signals of no interest, and the last term indicates AWGN. The magnitude of the MAI is determined by the amplitudes of interesting signals and the partial crosscorrelations between the signature waveforms of the desired user and the users of no interest. In the conventional receiver, the MAI prevents acquisition and tracking of synchronism for the desired signal and restricts the number of users simultaneously accessing the channel. This is because the MAI in the matched filter output hampers the detection of the desired signal and the MAI increases with the number of simultaneous users.

Previous Works on Multi-user Detection

From the discussion of the conventional receiver, it is noted that the MAI due to cross-correlations of signature waveforms hampers establishment of acquisition and tracking of synchronism and limits the system capacity. The MAI results in the *near-far problem* that relates to the problem of very strong signals from users of no interest at the receiver overwhelming the relatively weaker signals from the desired user and is the

main shortcoming of currently operational DS/CDMA systems [2]. Due to the reduced multiple-access capability and the increased vulnerability to interfering users caused by the near-far problem, solutions that overcome the near-far problem or that remove the effects of MAI have been an active research area in recent years.

Single-user Detection and Multi-user Detection

Generally, DS/CDMA detection can be classified as either centralized multi-user or decentralized single-user detection. The conventional matched filter receiver assigned to each user is an example of the single-user receiver. The performance of the conventional receiver is acceptable if the energies of the received signals from active users are not too dissimilar and that the signature waveforms are designed so that their crosscorrelations are low enough. In practice, low crosscorrelations are usually achieved employing complex constellation pseudo-random sequences of long periodicity in synchronous systems and are much more difficult to achieve in asynchronous systems.

Current approaches to address the near-far problem are to use the conventional receiver with a power control technique and/or the design of signature waveforms with more stringent crosscorrelation properties [3]. Unfortunately, power control dictates significant reductions in the transmitted powers of the strong users in order for the weak users to achieve reliable communications. Thus power control can become self-defeating strategy since it actually decreases the overall multiple-access and antijamming capabilities of the system. Furthermore, ever more complex signature waveforms lead to rapid increases in system cost and bandwidth, and do not eliminate the near-far problem [2]. A plausible reason for keeping the conventional detection scheme is the belief that the MAI can be accurately modeled as AWGN and thus the matched filter is essentially optimum.

The problem with such an idea was recognized by H. V. Poor [23], who proposed techniques from both minimax robustness and non-Gaussian signal detection to improve

the performance of the conventional receiver in multi-user channels. Schneider in [24] has also claimed that an appropriately chosen linear transformation of a bank of matched filter outputs results in optimum detections. Such a receiver and its generalization to the asynchronous case have very desirable properties and provide a motivation for multi-user detection approaches.

Multi-user detectors have centralized structures that process outputs of the matched filters associated with each user. Unlike in the single-user detection, the detection of all the users' data bits is an inter-dependent process in the multi-user detection. Error probability performances of multi-user detectors can greatly exceed that of the conventional receiver with expense of increased complexity. multi-user detectors vary in complexity from the optimum multi-user detector in [2] to the suboptimum detectors in [5], [7].[9] and [12].

Optimum Multi-user Detector

Verdu in [2] proposes the optimum multi-user detector that selects the data bit estimates that best explain the observations in a mean-squared-error sense. Although a data-synchronous DS/CDMA system is more the exception than the rule, the development of the optimum multi-user detector considers the special case of Eq.(2.5) where the users are data-synchronous:

$$r(t) = \sum_{k=1}^{K} A_k b_k(i)\, c_k(t - iT) + n(t), \quad t \in [iT, iT + T]. \tag{2.9}$$

The optimum multi-user detector selects the most likely hypothesis

$$\widehat{\mathbf{b}}^* = \left[\widehat{b}_1^*, \cdots, \widehat{b}_K^*\right] \tag{2.10}$$

from the maximum likelihood (ML) decision rule:

$$\begin{aligned} \hat{\mathbf{b}}^* &\in \arg\left\{ \min_{\mathbf{b}\in\{-1,1\}^K} \int_{iT}^{iT+T} \left[r(t) - \sum_{k=1}^{K} A_k b_k(i)\, c_k(t - iT) \right] dt \right\} \\ &= \arg\left\{ \max_{\mathbf{b}\in\{-1,1\}^K} \left[2\mathbf{z}^T\mathbf{b} - \mathbf{b}^T\mathbf{H}\mathbf{A}\mathbf{b} \right] \right\}. \end{aligned} \tag{2.11}$$

where

$$\mathbf{z} = [z_1, \cdots, z_K]^T, \tag{2.12}$$

$$z_k = \int_{iT}^{iT+T} r(t)\, c_k(t)\, dt, \quad k = 1, \cdots, K, \tag{2.13}$$

$$\mathbf{A} = diag\{A_1, \cdots, A_K\}, \tag{2.14}$$

$$\mathbf{b} = [b_1, \cdots, b_K]^T, \tag{2.15}$$

$$\mathbf{n} = [n_1, \cdots, n_K]^T, \tag{2.16}$$

and $\mathbf{H}$ is the crosscorrelation matrix whose ij coefficient is given by

$$\rho_{ij} = \int_0^T c_i(t)\, c_j(t)\, dt. \tag{2.17}$$

The optimum multi-user detector must solve the nondeterministic polynomial (NP)-complete combinatorial optimization problem [2]. Thus, no known algorithm in K exists for optimal multi-user detection. In an asynchronous case, the receiver consists of a matched filter front-end followed by a Viterbi algorithm that is well known as a maximum likelihood sequence estimator (MLSE) [4]. The number of states in the Viterbi algorithm is exponential in K with metrics computed in terms of the matched filter outputs and crosscorrelations.

The error probability of the optimum multi-user detector for a particular user is asymptotically equivalent to that of a binary test between the two closest hypotheses that differ in the data bit of the desired user. Therefore the asymptotic efficiency of the multi-user detector for single-user detection is defined as [6]

$$\eta_k^o = \frac{1}{A_k^2} \min_{\substack{\varepsilon \in \{-1,0,1\}^K \\ \varepsilon_k = 1}} \varepsilon^T \mathbf{H} \varepsilon, \quad k = 1, \cdots, K. \tag{2.18}$$

For example in the two user case, Eq.(2.18) for user 1 can be reduced to

$$\eta_1^o = \min\left\{1, 1 + \left(\frac{A_2}{A_1}\right)^2 - 2\,|\rho_{12}|\,\frac{A_2}{A_1}\right\}. \tag{2.19}$$

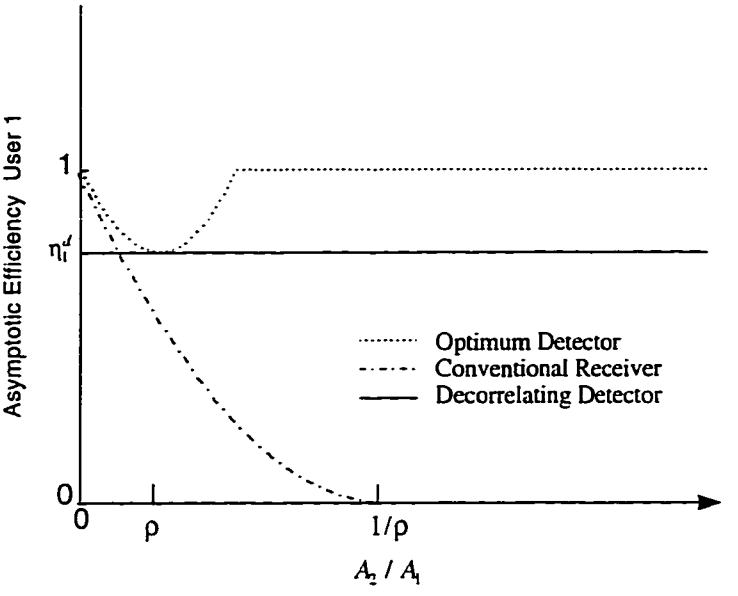

Figure 2.3 Asymptotic efficiencies for the two users with $\rho = 0.5$

In terms of the crosscorrelation matrix, the near-far resistance of the user k is equal to the reciprocal of the kth diagonal element of the inverse of the crosscorrelation matrix. That is, if the signal of the user 1 does not belong to the subspace spanned by the other signals, then

$$\overline{\eta}_k = \frac{1}{[\mathbf{H}^{-1}]_{kk}}, \quad k = 1, \cdots, K. \tag{2.20}$$

Otherwise, $\overline{\eta}_k = 0$. This implies that a huge performance gap exists between the conventional receiver and the optimum achievable performance. For example, while the near-far resistance of the conventional receiver is zero, the expected optimum near-far resistance using DS/CDMA signature waveforms with N chips per data bit is lower bounded by [25]

$$E[\overline{\eta}_k] \geq 1 - \frac{K-1}{N}, \quad k = 1, \cdots, K. \tag{2.21}$$

The asymptotic efficiency of two user case in terms of the relative power difference is shown in Figure 2.3. The implementation of the optimum detector not only requires the substantial amount of information such as the signature waveforms, received ampli-

tudes and timings of all users but also is not practical because its complexity increases exponentially in the number of users K.

Decorrelating Multi-user Detector

Lupas and Verdu in [5] propose the less complex suboptimum detector, called the decorrelating multi-user detector, which performs the linear transformation of the matched filter bank outputs in Eq.(2.12) by the generalized inverse of the crosscorrelation matrix $\mathbf{H}^I$, such that

$$\mathbf{H}^I\mathbf{z} = \mathbf{A}\mathbf{b} + \mathbf{H}^I\mathbf{n} \tag{2.22}$$

and takes the sign of the vector $\mathbf{H}^I\mathbf{z}$. Thus, in the absence of AWGN, the detector recovers the transmitted bits without corrupting by the MAI. In an asynchronous case, the decorrelator must generalize to an infinite impulse response (IIR) filter [5].

If the user 1 is linearly independent with other users, the error probability of the decorrelating detector for the user 1 can be expressed by using a linear transformation $\mathbf{H}^I$:

$$P_1^d = P\left[\left(\mathbf{A}\mathbf{b} + \mathbf{H}^I\mathbf{n}\right)_j > 0 \mid b_j = -1\right] \tag{2.23}$$

$$= Q\left(\frac{A_1}{\sigma \left\|[H^I]_{11}\right\|}\right) \tag{2.24}$$

Thus the asymptotic efficiency of the decorrelating detector for the user 1 is equal to

$$\eta_1^d = \frac{1}{[H^I]_{11}}. \tag{2.25}$$

The decorrelating detector is the maximum likelihood solution lacking any knowledge about the received amplitudes. Lupas and Verdu [5, 6] have shown that the decorrelating detector achieves optimum near-far resistance without AWGN. The bit-error-rate (BER) of the decorrelating detector is independent of the amplitudes of the interfering signals in the absence of AWGN. This is because the decorrelating linear transformation projects the received waveform on a subspace that is orthogonal to the space spanned by the

interfering signature waveforms. Main advantages of the decorrelating detector are the optimum near-far resistance property and the fact that it does not require knowledge of the received amplitudes, whereas the disadvantage is the computational complexity required to obtain the decorrelator coefficients from the crosscorrelations.

Multi-stage and Decision-Feedback Multi-user Detector

Another approach to multi-user detection is that of successive cancellation: detect the data of the strongest user first and then subtract the signal due to that user from the received waveform. If one can repeat this process with no error in its demodulation, the resulting waveform contains no signals due to a stronger user. This technique requires extremely accurate estimation of the received amplitudes, and unless the users can be ordered so that the received amplitudes satisfy

$$A_1 \gg A_2 \gg \cdots \gg A_K, \tag{2.26}$$

its performance may be actually worse than that of the decorrelating detector, which requires no knowledge of the received amplitudes.

The multi-stage detector of Varanasi and Aazhang [7, 8] is an example of such a detector. In this detector, the first stage consists of a bank of conventional receivers [7](or decorrelating detectors [8]) whereas the second stage assumes that the previous decisions are correct and simply subtracts the corresponding waveforms from the received waveform. By continuously repeating the above operation, a clear single-user channel is obtained when all the previous decisions are correct. The multi-stage detector is a successive implementation of the optimum detector, in which an estimate at the stage $(m+1)$ is made using estimates at the stage m.

Duel-Hallen in [9] also proposes the decorrelating decision-feedback multi-user detector. This detector incorporates features common to both successive cancellation and multi-stage detection with a decorrelating front-end. Application of a single-user

decision-feedback equalization scheme to the multi-user detection is proposed by Kohno, Imai and Hatori [26]. In these detection schemes, it is assumed that decisions made about earlier bits in an asynchronous system are correct and therefore they can be canceled. The idea of decision-feedback has been extended for implementation of an adaptive multi-user detector. Abdulrahman, Falconer and Sheikh in [27] use a fractionally-spaced decision-feedback equalization (DFE) detector whose coefficients are adapted to minimize mean-squared-error (MSE) using training sequences. Kohno, Imai and Hatori in [21] consider a CDMA channel with a limited bandwidth for which they design an adaptive MMSE detector that uses decision-feedback to remove intersymbol interference (ISI). The first stage of that detector performs preliminary decision based on the conventional detection scheme and then uses this decision in the adaptation stage. However, Rapajic and Vucetic in [13] find no improvement over the adaptive MMSE detector by incorporation the possibility of decision feedback. An adaptive version of the multi-stage detector has been proposed by Chen, Siveski and Bar-Ness [28]. The first stage of this detector is the decorrelating detector that requires knowledge of all the signature waveforms. The adaptation is carried out by gradient descent of the mean-squared error between the decorrelating detector output and the adaptive filter output and therefore it does not require training sequences.

Linear MMSE multi-user Detector

The decorrelating detector may have worse BER than conventional receiver when all the interfering signals are very weak. Madhow and Honig in [12] propose a linear MMSE multi-user detector that outperforms the decorrelating detector by incorporating knowledge of the received amplitudes. According to the minimum mean-squared-error (MMSE) criterion, the linear MMSE multi-user detector chooses the $K \times K$ matrix L

that achieves

$$\min_{L \in \mathcal{R}^{K \times K}} E\left[\|\mathbf{b} - \mathbf{Lz}\|^2\right] \tag{2.27}$$

where the expectation is with respect to the vector of transmitted bits $\mathbf{b}$ and the noise vector $\mathbf{n}$ that has zero mean and covariance matrix of $\sigma^2\mathbf{H}$.

Without invoking the Gaussian nature of $\mathbf{n}$, the linear MMSE detector replaces the inverse crosscorrelation matrix $\mathbf{H}^I$ by the matrix

$$\left[\mathbf{H} + \sigma^2\mathbf{A}^{-2}\right]^{-1} \tag{2.28}$$

where $\mathbf{A}$ is given in (2.14). Thus the linear MMSE detection has those aforementioned features of the decorrelating detector, but it requires the knowledge of received amplitudes of all the signals. If either the background noise level or the kth user received energy dominates, then the linear MMSE detector approaches the conventional receiver: on the other hand, as the background noise level vanishes $\sigma \to 0$, the linear MMSE detector approaches the decorrelating detector. Therefore, the asymptotic multi-user efficiency and the near-far resistance of the linear MMSE detector are the same as those of the decorrelator. Note that the great advantage of the linear MMSE detector is the ease with which it lends itself to adaptive implementation with a training sequence for fully decentralized single-user detection.

Previous Works on Linear MAI Suppression

Linear Adaptive MMSE Receiver

The linear MMSE single-user receiver was originally proposed by Xie, Short and Rushforth in [11]. Recently, several versions of the adaptive linear MMSE receiver were proposed in [12], [13] and [14]. Rapajic and Vucetic [13] present a linear fractionally spaced LMS-type adaptive receiver that suppresses MAI. The adaptive filter is used for both despreading the desired signal and suppressing the MAI. Miller [14] and Madhow

and Honig [12] present an N tap MMSE receiver, where N is the processing gain, also despreading the desired signal and suppressing the MAI with the use of a N tap adaptive filter. Madhow and Honig [12] also present two reduced complexity schemes, namely, the cyclically shifted filter bank (CSFB) and the over-sampling scheme. They show that the CSFB scheme outperforms the oversampling scheme at the expense of complexity.

The adaptive algorithm implementation for a single-user receiver can be briefly described as follows. When d_i is the ith data bit of the desired user, the cost function of the adaptive algorithm is given by

$$J(\mathbf{w}) = E\left[\left(d_i - \mathbf{w}^T \mathbf{r}_i\right)^2\right] \tag{2.29}$$

where $\mathbf{w}$ is the filter impulse response and $\mathbf{r}_i$ is the received waveform in the ith bit interval. The gradient of the cost function with respect to $\mathbf{w}$ is equal to

$$\nabla_{\mathbf{w}} = \frac{\partial J}{\partial \mathbf{w}} = \left(\mathbf{w}^T \mathbf{r}_i - d_i\right) \mathbf{r}_i. \tag{2.30}$$

The ith update equation of a stochastic gradient descent algorithm is given by

$$\begin{aligned} \mathbf{w}_{i+1} &= \mathbf{w}_i - \mu \nabla_{\mathbf{w}}(i) \\ &= \mathbf{w}_i - \mu \left(\mathbf{w}_i^T \mathbf{r}_i - b_i\right) \mathbf{r}_i \end{aligned} \tag{2.31}$$

where μ is the step size and $\nabla_{\mathbf{w}}(i)$ is the instantaneous gradient estimate from the input data. This algorithm is known as the least-mean-squares (LMS) algorithm.

Because of the convexity of the cost function in Eq.(2.29), Eq.(2.31) will converge to the argument that minimizes the cost function in Eq.(2.29). The update equation of the filter tap in Eq.(2.31) requires a training sequence for the desired user. It can be implemented in an asynchronous channel with the requirement of the timing of the desired user. Attractive features of the adaptive MMSE receiver are the near-far resistance, the optimum convergence without the knowledge of the desired signature sequence and the ability to suppress multipath and narrowband interference. However, although eliminating the need to know the side-information about the received signal, the requirement of

a training sequence both at the initialization and during data transmission makes the algorithm impractical for the application to the high-rate data communication systems. That is, as new users access the system, these receivers involve switching back and forth between a training mode and a decision-directed mode during data transmission.

Blind Adaptive OPM Receiver

While the linear adaptive MMSE receiver is an effective strategy to combat the MAI, the requirement of training sequences is cumbersome in practical applications. Constrained adaptive output-power minimizing (OPM) algorithms have been proposed by Honig, Madhow and Verdu [16] and Schodorf and Williams [17] for the blind adaptive receiver implementation. These receivers require the signature waveform and the timing information of the desired user but does not require a training sequence and any information of the interfering signals. The approach of the blind adaptive OPM receiver is similar to that of anchored minimum energy blind equalization proposed in [29].

In the blind adaptive receiver of [16], a linear filter $\mathbf{w}$ is orthogonally decomposed as

$$\mathbf{w} = \mathbf{s}_1 + \mathbf{x}, \quad \mathbf{s}_1^T \mathbf{x} = 0 \tag{2.32}$$

where $\mathbf{s}_1$ denotes the signature sequence of the user 1 and $\mathbf{x}$ denotes the adaptive component that will be continually adjusted to cancel MAI. The basic idea behind this decomposition is as follows; The outputs of the linear filter,

$$z = (\mathbf{s}_1 + \mathbf{x})^T \mathbf{r}, \tag{2.33}$$

can be decomposed by three additive components: the first is due to the desired user, the second is due to MAI, and the third is due to AWGN. The first component is transparent to the choice of $\mathbf{x}$ and varying $\mathbf{x}$ only changes the contribution of MAI and AWGN. Therefore adaptation is focused on $\mathbf{x}$, while preserving orthogonality to $\mathbf{s}_1$. That is, the problem is to adapt $\mathbf{x}$ so that it minimizes the output power while preserving the

orthogonality to $\mathbf{s}_1$:

$$\min_{\mathbf{x}} \quad E\left\{\left[(\mathbf{s}_1+\mathbf{x})^T\mathbf{r}\right]^2\right\}. \tag{2.34}$$

A gradient projection (GP) method is adopted to implement the following adaptation rule:

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \mu z_i\left(\mathbf{r}_i - z_i^{MF}\mathbf{s}_1\right) \tag{2.35}$$

where $\mathbf{x}_i$ is the adaptive component at ith update, μ is the step size, $z_i \triangleq (\mathbf{s}_1+\mathbf{x}_i)^T\mathbf{r}_i$and $z_i^{MF} \triangleq \mathbf{s}_1^T\mathbf{r}_i$ are the outputs of the adaptive filter and the matched filter at the ith update, respectively.

The blind adaptive OPM receiver of [17] is similar to that of [16] but the algorithm implementation that satisfies the constraint is different. The algorithm of [16] is based on the generalized sidelobe canceller (GSC) structure used in beamforming applications. The adaptation rule is given by

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \mu z_i\mathbf{B}\mathbf{r}_i \tag{2.36}$$

$$\mathbf{w}_{i+1} = \mathbf{s}_1 + \mathbf{x}_{i+1} \tag{2.37}$$

where $\mathbf{B}$ is a blocking matrix that ensures orthogonality to the signal of interest, i.e., $\mathbf{B}\mathbf{s}_1=\mathbf{0}$.

Both receivers in [16] and [17] are claimed *blind* because they do not employ any adaptation reference. However, the inclusion of the desired signal (as the output power) in the minimization process results in the signal cancellation phenomenon [18]. This phenomenon occurs because the algorithm does not guarantee the zero filter tap gain increment even when the perfect adaptation is achieved.

There have been other efforts in blind detection of a CDMA signal. Oda and Sato [30] propose a multi-dimensional generalization of the conventional single-user blind equalization methods that attempt to minimize a nonconvex cost function of the filter output. However the nonconvexity of the cost function does not guarantee the global

convergence and therefore the BER performance of the detector in [30] would be poor for week users. Singh [31] also proposes a decision-directed normalized LMS approach for blind implementation of a receiver. The adaptive receiver suppresses MAI reasonably well without the aid of training sequence only when the BER yields 10^{-1} errors per bit or lower. However this receiver cannot be employed in the near-far situation where the detection capability of the receiver is completely lost.

CHAPTER 3 CHARACTERISTICS OF SIGNALS AND LINEAR FILTERING OPERATION

multi-user detectors offer the great performance gain over the conventional (single-user) receiver, but some drawbacks of them are higher complexity (i.e., exponential or linear in the number of users) and the need to know substantial side-information about signals of no interest. For practical systems such as wireless personal, indoor and mobile radio communications, implementation complexity should be reduced to a feasible level even if the performance degrades slightly from optimum one. Furthermore, estimation of signal parameters in the CDMA channel is likely to be more difficult due to the time-varying nature of the channel. In fact, presently developed CDMA cellular radio networks does not consider optimum/suboptimum multi-user detectors [21].

A single-user receiver is constrained to demodulate the signal of only one user. In the design of a single-user receiver, the multi-user detection problem can be viewed as an interference suppression problem, where at the receiver a signal of interest is considered a signal and the signals of all other users are treated as interference. This chapter describes the characteristics of CDMA signals and some optimum linear filters that is suboptimum but implements more practical single-user receivers. The fundamental limitations of the linear optimum filter are also investigated.

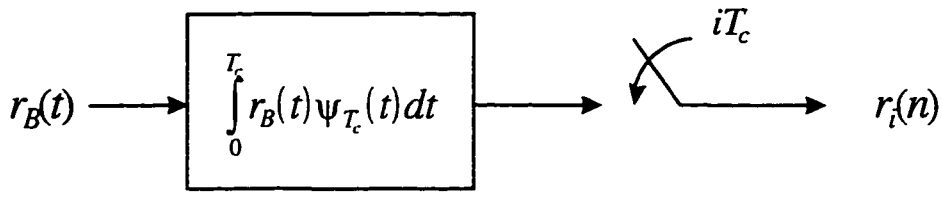

Figure 3.1 The single-user receiver front-end structure

Characteristics of CDMA Signals

A linear receiver considered is composed of a correlation-type linear filter and a zero-memory nonlinear decision device. It is assumed that the linear filter is realized by a tapped-delay-line (TDL) structure and the decision device is a simple sign operator. The receiver must obtain a sufficient statistic to detect the desired signal and to suppress the MAI by filtering operation.

For convenience, let us recall the baseband signal in the ith bit interval of the desired signal:

$$r_B(t) = A_1 b_1(i) c_1(t) + \sum_{k=2}^{K} A_k b_k(i) c_k(t - \tau_k) + n(t). \tag{3.1}$$

for $iT \leq t < (i+1)T$. It is assumed that this signal is passed through a chip-matched filter and sampled with sampling rate T_c. The front-end structure of the receiver that converts a continuous-time signal to a discrete-time signal is shown in Figure 3.1. Time delay of a DS/CDMA signal is defined by

$$\tau_k = l_k T_c + \varepsilon_k, k = 1, \cdots, K \tag{3.2}$$

where l_k is an integer in $[0, N-1]$ and ε_k is a remainder defined in $[0, T_c)$. Then the nth chip sample in the ith data bit interval can be written as

$$r_i(n) = A_1 b_1(i) c_1(n) + \sum_{k=2}^{K} A_k \left\{ b_k(i-1) c_k^L(n) + b_k(i) c_k^R(n) \right\} + n_i(n) \tag{3.3}$$

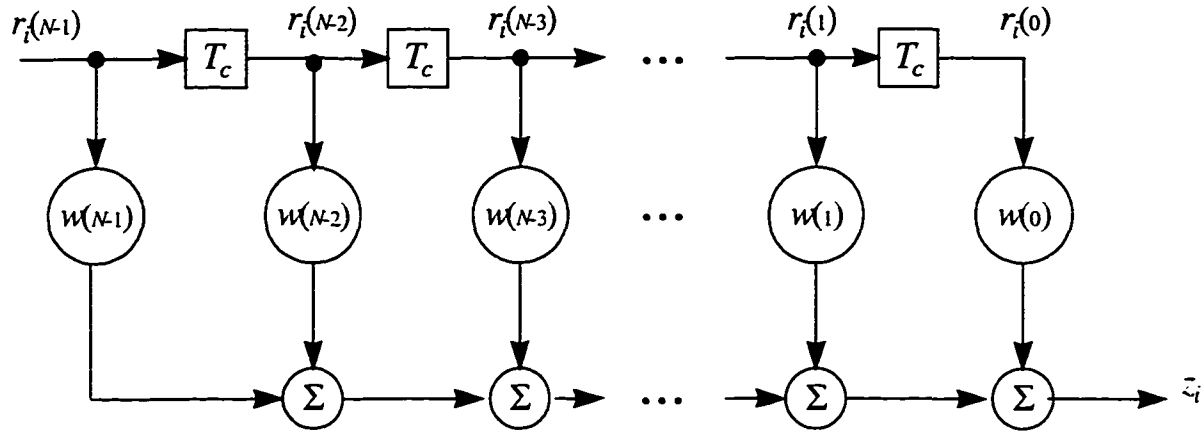

Figure 3.2 A tapped-delay-line filter structure

where

$$c_k^L(n) = \begin{cases} c_k(N-l_k-1+n)\,\phi_{k,0} + c_k(N-l_k+n)\,\phi_{k,1}, & \text{for } 0 \le n < l_k \\ c_k(N-l_k-1+n)\,\phi_{k,0}, & \text{for } n = l_k \\ 0, & \text{otherwise.} \end{cases}$$

$$c_k^R(n) = \begin{cases} 0, & \text{for } 0 < n < l_k \\ c_k(n-l_k)\,\phi_{k,1}, & \text{for } n = l_k \\ c_k(n-l_k-1)\,\phi_{k,0} + c_k(n-l_k)\,\phi_{k,1}, & \text{otherwise.} \end{cases}$$

$$\phi_{k,0} = \int_0^{T_c} \varphi_{T_c}(t)\,\varphi_{T_c}(t-\varepsilon_k)\,dt.$$

$$\phi_{k,1} = \int_0^{T_c} \varphi_{T_c}(t)\,\varphi_{T_c}(t-T_c+\varepsilon_k)\,dt$$

and

$$n_i(n) = \int_0^{T_c} n(t)\,\varphi_{T_c}(t)\,dt.$$

The sampled signal is passed through a linear TDL filter with a delay of T_c, in which the structure is shown in Figure 3.2. Since the detection of $b_1(i)$ depends on the received signal $r(t)$ for $t \in [iT, (i+1)T)$, or equivalently, the collection of the chip-rate

samples in the ith bit interval, the $(N \times 1)$ signal vector, $\mathbf{r}_i \in \mathcal{R}^N$, becomes *a sufficient statistic* for the detection of $b_1(i)$. The detection statistic can be written by

$$\mathbf{r}_i = A_1 b_1(i)\,\mathbf{c}_1 + \sum_{k=2}^{K} A_k \left\{ b_k(i-1)\,\mathbf{c}_k^L + b_k(i)\,\mathbf{c}_k^R \right\} + \mathbf{n}_i \tag{3.4}$$

where

$$\begin{aligned} \mathbf{r}_i &= [r_i(0), \cdots, r_i(N-1)]^T, \\ \mathbf{c}_k^L &= \left[c_k^L(0), \cdots, c_k^L(N-1)\right]^T, \\ \mathbf{c}_k^R &= \left[c_k^R(0), \cdots, c_k^R(N-1)\right]^T \end{aligned}$$

and $\mathbf{n}_i \in R^N$ is a while Gaussian noise vector with zero mean and covariance matrix $\sigma^2 \mathbf{I}_N$ where $\mathbf{I}_N$ denotes an $(N \times N)$ identity matrix.

Note that for $2 \le k \le K$, $\mathbf{c}_k^L$ and $\mathbf{c}_k^R$ are linearly independent and modulated by independent bits. Therefore, we can consider the asynchronous system as an equivalent synchronous system with additional interfering signals. Accordingly, the continuous baseband signal in Eq.(3.1) can be simply written as an *equivalent discrete-time synchronous signal*:

$$\mathbf{r}_i = \sum_{l=1}^{L} A_l b_l(i)\,\mathbf{s}_l + \mathbf{n}_i \tag{3.5}$$

where $L \le 2K-1$ is the number of equivalent signals, $\mathbf{s}_1 = \mathbf{c}_1$ and $\mathbf{s}_l \in \left\{\mathbf{c}_k^L, \mathbf{c}_k^R\right\}$. $2 \le l \le L$, $2 \le k \le K$ are the signature sequence vectors of hypothetical signals. Therefore it is noted that both the desired signal and the MAI preserve a *cyclostationary structure* even in the asynchronous case. Using matrix-vector notation and removing the dependence on the subscript i without loss of generality, the signal vector in (3.5) is compactly rewritten as a matrix-vector form;

$$\mathbf{r} = \mathbf{SAb} + \mathbf{v} \tag{3.6}$$

where $\mathbf{S} = [\mathbf{s}_1, \cdots, \mathbf{s}_L]$ denotes an $(N \times L)$ signature matrix, $\mathbf{A} = diag[A_1, \cdots, A_L]$ denotes an $(L \times L)$ amplitude matrix and $\mathbf{b} = [b_1, \cdots, b_L]^T$ denotes an $(L \times 1)$ data

vector. If the transmission is data-synchronous, $\mathbf{s}_k = \mathbf{c}_k, 1 \leq k \leq K$ and the size of the vectors and the matrices in Eq.(3.6) should be changed appropriately.

Performance Measures of Linear Filter

In the demodulation of CDMA signals, the performance is generally examined in the low background noise region, i.e., $\sigma \rightarrow 0$. The *asymptotic efficiency* [2] is defined as the loss relative to a conventional (matched filter) receiver operating in the presence of only additive white Gaussian noise in the limit as the signal to noise ratio (SNR) goes to zero. Denoting the BER of the kth user by $P_k(\sigma)$, the asymptotic efficiency of the kth user is formally defined by

$$\eta_k = \sup \left\{ 0 \leq \gamma \leq 1; \lim_{\sigma \rightarrow 0} P_k(\sigma) / Q\left(\frac{\sqrt{\gamma} A_k}{\sigma}\right) < +\infty \right\} \tag{3.7}$$

i.e., the log bit-error-rate of the kth user goes to zero with the same slope as that of a single-user with energy $\eta_k A_k^2$ where A_k^2 is the actual signal energy. The worst case asymptotic efficiency over all received interfering amplitudes is called the *near-far resistance* [2], denoted by $\overline{\eta}_k$.

For a single-user detector, the near-far resistance may be an attractive performance measure. Therefore, single-user asymptotic efficiency must be considered to design a linear receiver for upper bound. Let us define the linear transformation given by a linear filter as $\mathbf{f} : \mathcal{R}^N \rightarrow \mathcal{R}$. Then the kth user asymptotic efficiency of a linear filter is given by [6]

$$\eta_k^l = \max^2 \left\{ 0, \frac{1}{A_k} \frac{\langle \mathbf{f}, A_k \mathbf{s}_k \rangle - \sum_{j \neq k} |\langle \mathbf{f}, A_j \mathbf{s}_j \rangle|}{\sqrt{\langle \mathbf{f}, \mathbf{f} \rangle}} \right\} \tag{3.8}$$

where $\langle \mathbf{x}, \mathbf{y} \rangle \triangleq \mathbf{x}^T \mathbf{y}$ is the inner product of vectors $\mathbf{x}$ and $\mathbf{y}$, $A_k$and $\mathbf{s}_k$ are the received amplitude and the signature sequence of the kth signal, respectively. Another important performance measure is the *bit-error- rate (BER)* or the *probability of bit error*, which

is a monotonic decreasing function of the asymptotic efficiency for the near-far resistant receiver. The error probability of the kth user can be written as

$$P_k = Q\left(\sqrt{\frac{\eta_k^l A_k^2}{\sigma^2}}\right) \tag{3.9}$$

where $Q(\cdot)$ is the standard Q-function.

Performance of Conventional Receiver

Let us consider performance of the conventional receiver. The error probability of the conventional receiver for the user 1 is given by

$$\begin{aligned} P_1^c &= P[z > 0 \mid b_1 = -1] \\ &= 2^{1-L} \sum_{\substack{b\in\{-1,1\}^L \\ b_1=-1}} Q\left(\frac{A_1 + \sum_{k=2}^{L} A_k b_k \rho_{1k}}{\sigma}\right) \end{aligned} \tag{3.10}$$

where ρ_{1k} denotes the normalized crosscorrelation between the signature sequences of the user 1 and k. The asymptotic efficiency of the conventional receiver for the user 1 is equal to

$$\eta_1^c = \max^2\left\{0, 1 - \sum_{k=2}^{K} |\rho_{1k}| \frac{A_k}{A_1}\right\}. \tag{3.11}$$

It follows from Eq.(3.11) that the conventional receiver is near-far resistant only if $\rho_{1k} = 0$ for all k. Otherwise. $\overline{\eta}_1^c = 0$.

Linear Class of Optimum Filter

This section describes ideal optimum linear filters based on various optimizing criterion and available knowledge. It is assumed that the received signal $\mathbf{r} \in \mathcal{R}^N$ is absolutely summable. Let us define $\mathbf{w}$ as the filter tap vector to be designed. The filter linearly transforms the received signal into a decision statistic by an inner product operation:

$$z \triangleq \mathbf{w}^T \mathbf{r}. \tag{3.12}$$

The filtering operation is shown in Figure 3.2. The decision device decides b_1 from the filter output based on the sign operation:

$$\widehat{b}_1 = \text{sgn}\,[z]\,, \tag{3.13}$$

i.e.,

$$\widehat{b}_1 = \begin{cases} +1, & \text{if } z \geq 0 \\ -1, & \text{otherwize} \end{cases}\,. \tag{3.14}$$

The filter output can be decomposed into three additive components:

$$z = A_1 b_1 \left[\mathbf{w}^T \mathbf{s}_1\right] + \sum_{k=2}^{L} A_k b_k \left[\mathbf{w}^T \mathbf{s}_k\right] + \left[\mathbf{w}^T \mathbf{v}\right] \tag{3.15}$$

where the first term contains the desired information, the second term is the results of MAI, and the last is due to AWGN. The filter output is an element in the inner product space on $\mathcal{R}$. Note that the inner product space of finite-dimensional spaces is the Hilbert space. From the geometry of the Hilbert space, we can use a special linear operator, called *projections onto a Hilbert space*, to obtain a linear class of optimum filter. A brief summary of the *projection theory on a Hilbert space* is given in Appendix.

Near-Far Resistant Filter with Zero-Forcing Solution

Let us first consider a simple solution to address the near-far problem. If the knowledge of the signature sequences and time delays of all the users are available, we can design a near-far resistant filter that produces desired information by completely removing MAI in the absence of AWGN. The near-far solution requires the filter tap vector to satisfy two conditions. First, it should be perpendicular to the span of the set of all interference signature vectors, $\{\mathbf{s}_k : 2 \leq k \leq L\}$, so that linear transformation of the received signal by the filter eliminates the MAI components. Second, it should be parallel to the signature vector of the desired signal so that it preserves the transmitted information when the signal is passed through the filter. These conditions are satisfied

by a *zero-forcing* (ZF) solution, which is obtained by solving a set of linear equations when $L < N$:

$$\begin{bmatrix} s_1(0) & s_1(1) & \cdots & s_1(N-2) & s_1(N-1) \\ s_2(0) & s_2(1) & \cdots & s_2(N-2) & s_2(N-1) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ s_{L-1}(0) & s_{L-1}(1) & \cdots & s_{L-1}(N-2) & s_{L-1}(N-1) \\ s_L(0) & s_L(1) & \cdots & s_L(N-2) & s_L(N-1) \end{bmatrix} \begin{bmatrix} w(0) \\ w(1) \\ \vdots \\ w(N-1) \\ w(N) \end{bmatrix} = \begin{bmatrix} \zeta \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{3.16}$$

or compactly

$$\mathbf{Sw} = \mathbf{d} \tag{3.17}$$

where

$$\mathbf{S} \triangleq [\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_L]^T, \tag{3.18}$$

$$\mathbf{w} \triangleq [w(0), w(1), \cdots, w(N-1)]^T \tag{3.19}$$

$$\mathbf{d} \triangleq [\zeta, 0, \cdots, 0]^T \tag{3.20}$$

and ζ is an arbitrary chosen constant. In the hypothetical absence of AWGN, this filter perfectly recovers the transmitted bits. However when the AWGN presents, the near-far resistant filter, $\mathbf{w}_{NF}$, produces

$$\mathbf{w}_{ZF}^T \mathbf{r} = \zeta A_1 b_1 + v, \tag{3.21}$$

where $v = \mathbf{w}_{ZF}^T \mathbf{n}$ is zero mean AWGN with a variance of $\sigma^2 \|\mathbf{w}_{ZF}\|^2$ and $\|\mathbf{w}_{ZF}\|^2 \triangleq \mathbf{w}_{ZF}^T \mathbf{w}_{ZF}$. Since ζA_1 is a constant and the decision device estimates b_1 by employing the sign operator, the ZF or near-far resistance solution, which accounts for the perfect suppression of MAI, gives the output signal-to-interference ratio (OSIR):

$$OSIR_{ZF} = \frac{[\zeta A_1]^2}{\sigma^2 \|\mathbf{w}_{ZF}\|^2}. \tag{3.22}$$

The implementation of the zero-forcing filter is different from that of the decorrelating multi-user detector [5] in which the former linearly transforms the signal sampled at the

chip-rate in a single-user receiver, whereas the latter performs a linear transformation on the outputs of the matched filter bank at the bit-rate. However, both detectors have identical asymptotic efficiency for a particular user and require the same amount of information such as the signature sequences and the relative time delays of all users.

Minimax Filter with Peak Distortion Criterion

Let us assume that the ZF solution is not feasible. Then the ZF filter is no longer near-far resistant and its performance may be worse than that of the conventional receiver without incorporating the amplitude information of all the signals besides the information required for the ZF solution. Park and Doherty [32] presented a minimax optimization approach to suppress the sidelobes of the filter output in spread spectrum radar applications. Since peak distortion of the filter output in CDMA multi-user communications is related to the maximum value associated to a strong interfering user, we can design a filter that minimizes the l_∞ norm of the error vector between the desired vector and the filter output vector, which is represented by

$$\mathbf{d} = \mathbf{ASw} + \varepsilon \tag{3.23}$$

where

$$\mathbf{A} \triangleq diag\left([A_1, A_2, \cdots, A_L]\right), \tag{3.24}$$

$$\mathbf{S} \triangleq [\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_L]^T, \tag{3.25}$$

$$\varepsilon \triangleq [e_1, e_2, \cdots, e_L]^T, \tag{3.26}$$

$$\mathbf{d} \triangleq [d_1, d_2, \cdots, d_L]^T. \tag{3.27}$$

d_k is given by

$$d_k = \begin{cases} 1, & k = 1 \\ 0. & \text{otherwise} \end{cases} \tag{3.28}$$

and the constant 1 is arbitrary chosen. Let us define e to be the variable to be minimized. i.e.,

$$e = \min\{e_k\}, \quad k = 1, \cdots, L \tag{3.29}$$

so that the difference between the desired and actual filter response is bounded by e. which is given by

$$|(\mathbf{ASw})_k - d_k| \leq e, \quad \text{for all } k. \tag{3.30}$$

Accordingly, one can formulate a mathematical optimization problem that satisfies a set of inequality constraints, which is given by a compact matrix-vector form:

$$\begin{aligned} \min_{\mathbf{w}} \quad & e \\ \text{subject to} \quad & \mathbf{ASw} - \mathbf{d} \leq \mathbf{e} \\ & \mathbf{ASw} - \mathbf{d} \geq \mathbf{e} \end{aligned} \tag{3.31}$$

where $\mathbf{e} \triangleq [e, e, \cdots, e]^T$. This type of optimization problem can be easily solved by the method of linear programming [33]. When the ZF solution is not feasible, the minimax filter outperforms over the ZF filter and is linear optimum in suppressing the MAI in the absence of AWGN. In the presence of AWGN, the output of the minimax filter is given by

$$\mathbf{w}_{MX}^T \mathbf{r} = b_1 + \sum_{k=2}^{L} e_k b_k + v, \tag{3.32}$$

where $v \triangleq \mathbf{w}_{MX}^T \mathbf{n}$ denotes white Gaussian noise with a variance of $\sigma^2 \|\mathbf{w}_{MX}\|^2$. The minimax filter gives the following output SIR:

$$OSIR_{MX} = \frac{[\zeta A_1]^2}{\|\varepsilon\|^2 + \sigma^2 \|\mathbf{w}_{MX}\|^2}. \tag{3.33}$$

MMSE Filter with Least-Squares Criterion

Note that both the ZF and the minimax solutions are deterministic solutions, which accounts for the exact knowledge of the desired and interfering signals. In practice, the received signal is a stochastic process corrupted by AWGN in addition to MAI. When the

statistics of the received signal and the desired signal are known, the least-squares (LS), or linear minimum mean-squared-error (MMSE) optimization gives a linear optimum solution. For a random input signal, let us write a statistical optimization problem:

$$\min_{\mathbf{w}} \quad E\left[\left(\mathbf{r}^T\mathbf{w} - d\right)^2\right] \tag{3.34}$$

where d is the desired response . A MMSE filter can be designed so that the mean-squared-error (MSE) between the desired response and the actual filter response is minimized [34]. The linear MMSE solution is obtained by solving a set of linear algebraic equations known as the normal equations:

$$\mathbf{R}\mathbf{w} = \mathbf{p} \tag{3.35}$$

where $\mathbf{R} \triangleq E\left[\mathbf{r}\mathbf{r}^T\right]$ denotes the $(N \times N)$ autocorrelation matrix of the received signal and $\mathbf{p} \triangleq E\left[d\mathbf{r}\right]$ denotes the $(N \times 1)$ crosscorrelation vector between the desired signal and the received signal, respectively. For nonsingular $\mathbf{R}$, the MMSE filter, or simply the LS filter, is obtained by

$$\mathbf{w}_{MS} = \mathbf{R}^{-1}\mathbf{p}. \tag{3.36}$$

Therefore, based on the knowledge of $\mathbf{R}$ and $\mathbf{p}$, $\mathbf{w}_{MS}$ can be readily calculated. By solving the given optimization problem, the minimum value of the MSE, $J_{\min}$, is obtained by

$$J_{\min} = E\left[d^2\right] - \mathbf{w}_{MS}^T\mathbf{p} \tag{3.37}$$

and therefore the output SIR is calculated by

$$OSIR_{MS} = \frac{E\left[d^2\right]}{J_{\min}} = \frac{E\left[d^2\right]}{E\left[d^2\right] - \mathbf{w}_{MS}^T\mathbf{p}}. \tag{3.38}$$

Fundamental Limitations of Linear Filter

Geometric Interpretation of Linear MMSE Solution

In this section, the linear filtering operation is analyzed through the geometric interpretation of the linear MMSE solution by using the *Hilbert space projection theorem* [35] which is described in the Appendix. In fact, the MMSE solution is optimum when the filter input is a random process. Let us define the desired signal by $d \triangleq A_1 b_1$. Then the linear MMSE solution satisfies the normal equations:

$$\mathbf{R}\mathbf{w} = \mathbf{p} \tag{3.39}$$

where

$$\mathbf{R} = A_1^2 \mathbf{s}_1 \mathbf{s}_1^T + \sum_{k=2}^{K} A_k^2 \mathbf{s}_k \mathbf{s}_k^T + \sigma^2 \mathbf{I} \tag{3.40}$$

and

$$\mathbf{p} = E(d\mathbf{r}) = A_1^2 \mathbf{s}_1. \tag{3.41}$$

The MMSE performance of this solution is given by

$$MMSE = A_1^2 - \mathbf{w}^T \mathbf{p} = A_1^2 \left(1 - \mathbf{w}^T \mathbf{s}_1\right) \tag{3.42}$$

or the normalized MMSE is given by

$$J_{\min} = \left(1 - \mathbf{w}^T \mathbf{s}_1\right). \tag{3.43}$$

Without loss of any generality, the Eq.(3.39) can be written as an alternative form:

$$\mathbf{B}\mathbf{w} = \left(1 - \mathbf{w}^T \mathbf{s}_1\right) \mathbf{s}_1 \tag{3.44}$$

where the interference crosscorrelation matrix $\mathbf{B}$ is given by

$$\mathbf{B} = \sum_{k=2}^{K} \beta_k \mathbf{s}_k \mathbf{s}_k^T + \alpha \mathbf{I}, \tag{3.45}$$

β_k and α denote reciprocals of the input SIR for the interfering user k and the input SNR, respectively, which are defined as

$$\beta_k = \frac{A_k^2}{A_1^2}, \quad \alpha = \frac{\sigma^2}{A_1^2}. \tag{3.46}$$

Consider the signature sequences, $\mathbf{s}_k \in \mathcal{R}^N$, $k = 1, \cdots, K$, are finite-dimensional vectors. To explore the fundamental limitation of the linear filter, the desired signature vector and the filter impulse response vector are decomposed into two orthogonal components; one is in the interference space and the other is in the space orthogonal to the interference space. Define a Hilbert space $\mathcal{M} \in \mathcal{R}^N$ as the subspace spanned by the interference signature vector $\{\mathbf{s}_k, k = 2, \cdots, K\}$ and let a Hilbert space $\mathcal{M}^\perp \in \mathcal{R}^N$ denote the subspace orthogonal to $\mathcal{M}$. If a Hilbert space $\mathcal{S}$ is the direct sum of $\mathcal{M}$ and $\mathcal{M}^\perp$, which is given by

$$\mathcal{S} = \mathcal{M} \oplus \mathcal{M}^\perp, \tag{3.47}$$

then the desired signature vector $\mathbf{s}_1$ and the filter impulse response vector $\mathbf{w}$ must lie in the space $\mathcal{S}$. Let $\mathbf{s}_1^I$ and $\mathbf{s}_1^O$ denote the projection of $\mathbf{s}_1$ onto $\mathcal{M}$ and $\mathcal{M}^\perp$, respectively. Similarly, let $\mathbf{w}^I$ and $\mathbf{w}^O$ denote the projection of $\mathbf{w}$ onto $\mathcal{M}$ and $\mathcal{M}^\perp$, respectively. Then, $\mathbf{s}_1$ and $\mathbf{w}$ can be orthogonally decomposed as

$$\mathbf{s}_1 = \mathbf{s}_1^I + \mathbf{s}_1^O, \quad \mathbf{w} = \mathbf{w}^I + \mathbf{w}^O. \tag{3.48}$$

These equations can be conveniently rewritten as a linear combination of corresponding vectors:

$$\mathbf{s}_1 = \sum_{k=2}^{K} \chi_k \mathbf{s}_k + \mathbf{s}_1^O, \tag{3.49}$$

$$\mathbf{w} = \sum_{k=2}^{K} \zeta_k \mathbf{s}_k + \delta \mathbf{s}_1^O \tag{3.50}$$

where χ_k, ζ_k, $k = 2, \cdots, K$ and δ are the coefficients to be determined. Figure 3.3 illustrates the geometric relationship between the desired signature vector and the interference subspace. Projecting each side of Eq.(3.44) onto the subspace $\mathcal{M}$ and $\mathcal{M}^\perp$

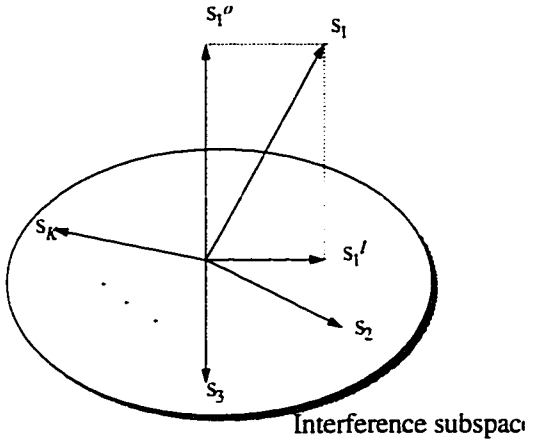

Figure 3.3 Orthogonal decomposition of signature sequences

gives following two equations:

$$\left(\sum_{k=2}^{K}\beta_k\mathbf{s}_k\mathbf{s}_k^T+\alpha\mathbf{I}\right)\left(\sum_{k=2}^{K}\zeta_k\mathbf{s}_k\right) = \left(1-\mathbf{w}^T\mathbf{s}_1\right)\left(\sum_{k=2}^{K}\chi_k\mathbf{s}_k\right) \tag{3.51}$$

$$\alpha\delta = \left(1-\mathbf{w}^T\mathbf{s}_1\right) \tag{3.52}$$

Solving Eq.(3.52) for the $J_{\min}$ in terms of Eq.(3.48) gives

$$J_{\min}=\frac{\alpha\left[1-\left(\mathbf{w}^I\right)^T\mathbf{s}_1^I\right]}{\alpha+\left\|\mathbf{s}_1^O\right\|^2}, \tag{3.53}$$

or, compactly

$$J_{\min}==\frac{\alpha\left[1-\mathbf{y}^T\mathbf{H}\mathbf{x}\right]}{\alpha+\left[1-\mathbf{x}^T\mathbf{H}\mathbf{x}\right]} \tag{3.54}$$

where $\mathbf{x}\triangleq[\chi_2,\cdots,\chi_K]^T$ is the coefficient vector for $\mathbf{s}_1$, $\mathbf{y}=[\zeta_2,\cdots.\zeta_K]^T$ is the coefficient vector for $\mathbf{w}$ and $\mathbf{H}$ is the signature correlation matrix of the interference signals. where ijth element of $\mathbf{H}$ is given by

$$[\mathbf{H}]_{ij}=\rho_{ij},\quad 2\le i,j\le L, \tag{3.55}$$

and $\rho_{ij}\triangleq\mathbf{s}_i^T\mathbf{s}_j$ is a crosscorrelation between the ith signature and the jth signature sequence.

Asymptotic Efficiency and System Capacity

The MMSE solution is near-far resistant in the sense that $\left\|\mathbf{s}_1^O\right\|^2 > 0$ as α goes to 0. From the definition, we can write the asymptotic efficiency of the linear optimum filter as

$$\eta_1^l = \left\|\mathbf{s}_1^O\right\|^2. \tag{3.56}$$

That is, the near-far resistance of the linear MMSE filter is the l_2 norm of the component of $\mathbf{s}_1$that is orthogonal to $\mathcal{M}$. As long as $\mathbf{s}_1$ is not fully contained in $\mathcal{M}$, the near-far resistance is nonzero. A necessary condition for this is to be true is that the dimension of $\mathcal{M}$ be strictly less than the dimension of the signature sequence N. Since the dimension of $\mathcal{M}$ is upper-bounded by the number of interfering signals, $L-1$, it is reasonable to expect nonzero near-far resistance when $L-1 < N$. In terms of the number of users K, it can be written as $K-1 < N$ for the synchronous case and when $2(K-1) < N$ for the asynchronous case.

Let us compute the coefficients of $\mathbf{s}_k$, $k = 2, \cdots, L$ in representing $\mathbf{s}_1$. The orthogonality condition in the Hilbert space satisfies

$$\mathbf{s}_j^T \mathbf{s}_1^O = \mathbf{s}_j^T \left(\mathbf{s}_1 - \sum_{k=2}^{K} \chi_k \mathbf{s}_k \right) = 0, \qquad 2 \le j \le K. \tag{3.57}$$

or compactly,

$$\mathbf{H}\mathbf{x} = \mathbf{h} \tag{3.58}$$

where

$$\mathbf{h} = [\rho_{12}, \cdots, \rho_{1K}]^T \tag{3.59}$$

Accordingly, the asymptotic efficiency of the linear filter can be expressed in terms of the crosscorrelations of signature sequences such as

$$\eta_1^l = 1 - \mathbf{h}^T \mathbf{x} \tag{3.60}$$

where taking $\mathbf{s}_1^T \mathbf{s}_1 = 1$. The vector $\mathbf{x}$ is obtained from Eq.(3.58) and need not be unique unless the interference signature vectors are linearly independent. Figure 3.4 shows the

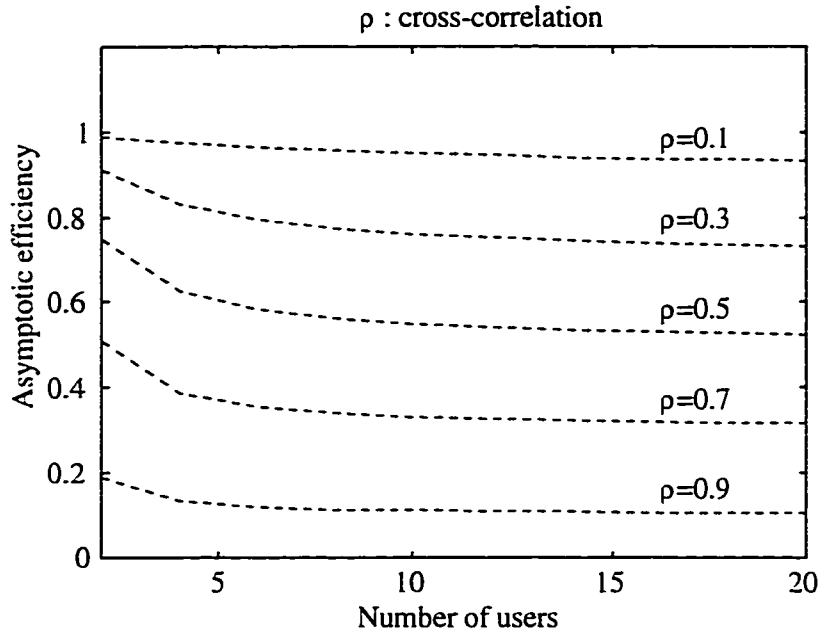

Figure 3.4 Near-far resistance in terms of number of users and crosscorrelations

asymptotic efficiency in terms of the number of interference and the crosscorrelations between the signature sequences.

MMSE in Terms of Crosscorrelations, Input SNR and Input SIR

To compute the MMSE in terms of parameters of the received signal, let us consider the MMSE solution in the interference subspace. Taking an inner product both side of (3.51) with $\mathbf{s}_m$, $2 \leq m \leq K$ gives

$$\sum_{j=2}^{K}\sum_{k=2}^{K}\beta_j\zeta_j\rho_{jk}\rho_{jm} + \alpha\sum_{j=2}^{K}\zeta_j\rho_{jm} = J_{\min}\sum_{j=2}^{K}\chi_j\rho_{jm}, \qquad 2 \leq m \leq K. \tag{3.61}$$

Let us write it compactly in a matrix-vector form:

$$(\mathbf{HWH} + \alpha\mathbf{H})\,\mathbf{y} = J_{\min}\mathbf{Hx} \tag{3.62}$$

where $\mathbf{W}$ is an input SIR matrix which has the kkth diagonal element:

$$[\mathbf{W}]_{kk} = \beta_k = \frac{A_k^2}{A_1^2} \triangleq \frac{1}{SIR_k}, \quad 2 \leq k \leq K \tag{3.63}$$

where SIR_k denotes the SIR for the interference k.

By defining a vector as $\mathbf{y} \triangleq J_{\min}\mathbf{v}$ and substituting $\mathbf{Hx} = \mathbf{h}$, we can rewrite (3.48) as

$$(\mathbf{HWH} + SNR \cdot \mathbf{H})\,\mathbf{v} = \mathbf{h}. \tag{3.64}$$

Accordingly, the normalized MMSE of the linear MMSE filter can be obtained by

$$J_{\min} = \left[1 + \frac{1 - \mathbf{h}^T\mathbf{x}}{SNR} + \mathbf{h}^T\mathbf{v}\right]^{-1} \tag{3.65}$$

and the minimum output signal-to-interference ratio (MSIR) is given by

$$MSIR_o = 1 + \frac{1 - \mathbf{h}^T\mathbf{x}}{SNR} + \mathbf{h}^T\mathbf{v}. \tag{3.66}$$

Accordingly, the normalized MMSE and the output MSIR performances are conveniently expressed in terms of the input SIR and crosscorrelation of each interference. and the input SNR. Note also that $\mathbf{h}$, $\mathbf{v}$ and $\mathbf{x}$ are obtained by Eq.(3.59), Eq.(3.49) and Eq.(3.58). respectively. The MMSE performance in terms of the number of users. input SIR. input SNR and crosscorrelations between the signature sequences are given in Figure 3.5 through Figure 3.8.

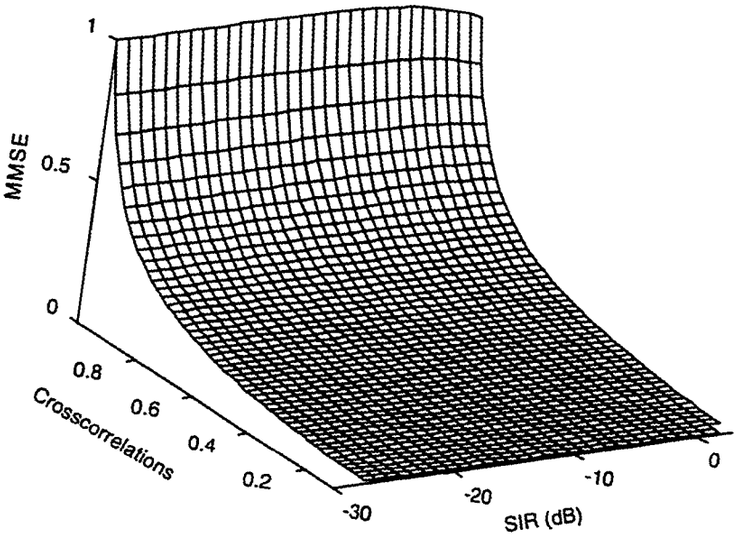

Figure 3.5 Plot of MMSE vs cross-correlations and input SIRs.

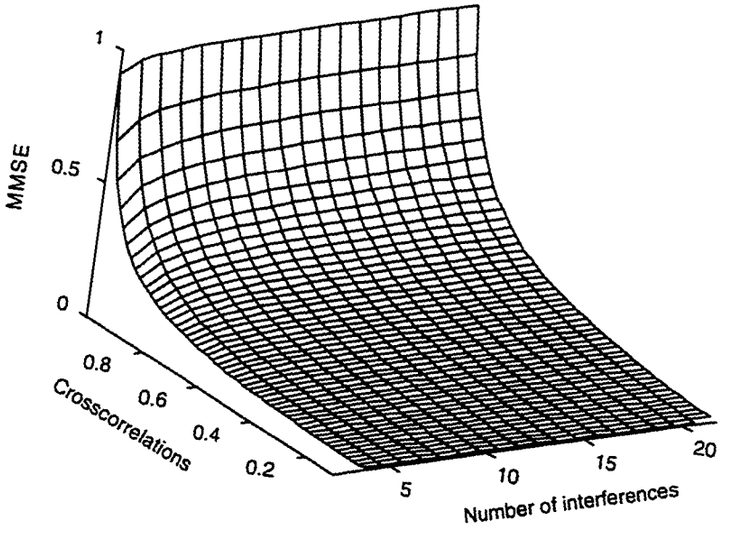

Figure 3.6 Plot of MMSE vs cross-correlations and number of interferences.

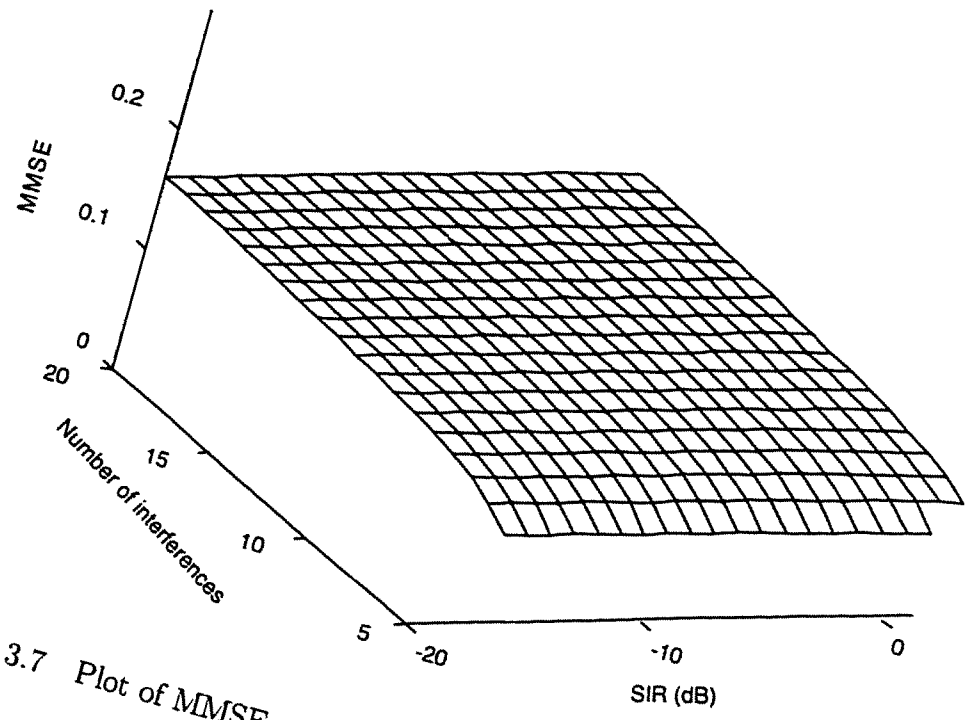

Figure 3.7 Plot of MMSE vs number of interferences and input SIRs

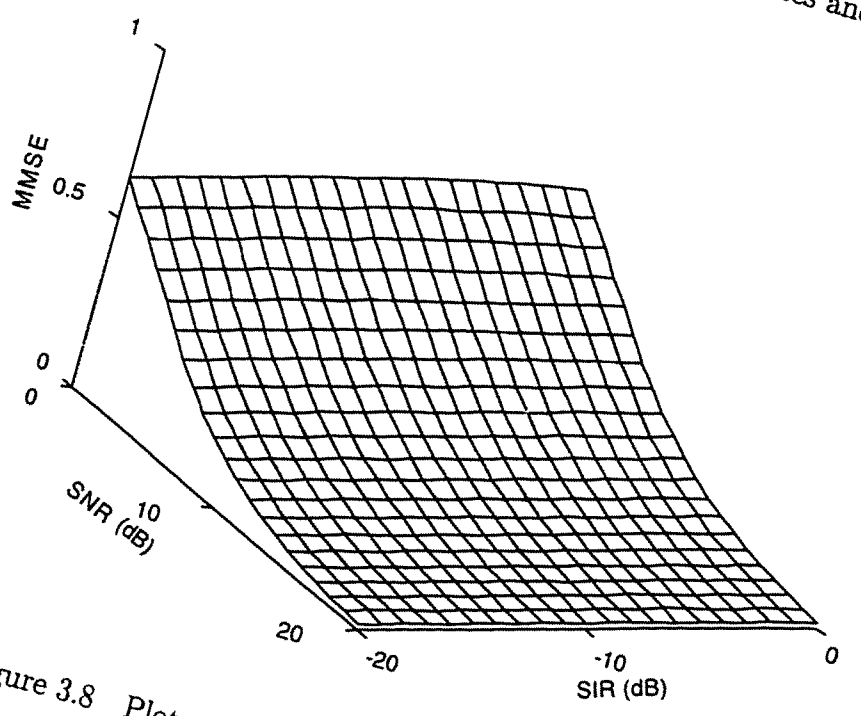

Figure 3.8 Plot of MMSE vs input SNR and input SIR

CHAPTER 4 CONSTRAINED ADAPTIVE MAI SUPPRESSION

Introduction

In practice, the information required to specify an optimum/suboptimum filter may not be available, or the operating environment may change with time, or both. Therefore we should autonomously assess the operating environment and then adjust the filter parameters to maintain and retain filter optimality. An *Adaptive filter* refers to a time-varying filter that adjusts its tap parameters based on only the received signal and some knowledge of the desired signal such that it can produce the desired information while suppressing the interference. Adaptive filtering to suppress the MAI is quite similar to adaptive equalization to suppress the intersymbol interference (ISI) with some minor difference. The difference is that the input to the MAI suppression filter is sampled at the chip-rate while the output is sampled at bit-rate whereas in the ISI equalizer the output is also sampled as the same rate as the input. It is assumed that all the users in the DS/CDMA multi-user channel use the same length signature sequences so that the *cyclostationary structure* of the signature sequence can be used for the filter design whereas in the ISI channel the stationary crosscorrelation property of the ISI is utilized for the filter design.

This chapter investigates the *cost-constraint strategy* for the implementation of a blind adaptive receiver or filter. A constrained LMS algorithm is formulated by combining a nonconvex cost function with a filter constraint that passes the signature sequence

of the desired signal with a prescribed response. A simple linear receiver implemented by the constrained adaptive LMS algorithm can be used to demodulate the desired information in near-far DS/CDMA multi-user communication environments without requiring any training sequence. It has been shown that a constrained adaptive algorithm is sensitive to the filter or signal mismatch due to implementation error or random noise effect. In order for the constrained LMS algorithm to be robust to the mismatch effect, the filter is further constrained to be bounded in a norm while being adjusted by a constrained LMS stochastic gradient descent algorithm. In this approach, all other users' signals in the filter input, which are independent of the desired signal, are considered interferences.

Brief Overview of Related Work

Recently proposed blind adaptive receiver in [16] is implemented by the gradient projection (GP) algorithm that adjusts the filter taps to minimize the output power while projecting the gradient onto the constraint space to satisfy the constraint. This algorithm considers blind because it does not have a reference in the cost function, i.e., the cost function is simply the filter output power. The receiver proposed in [17] is also implemented by an output power minimizing approach with a generalized sidelobe canceller structure which blocks the desired signal for the adaptive filter update. However, it can be found that both algorithms have two major drawbacks;

- Inclusion of a desired signal component (i.e., as the output power) in the adaptation process results in considerable performance loss in the steady state [18].
- The gradient projection method is susceptible to accumulate roundoff errors and is not suitable for long runs without an error-correction procedure [36].

A related work, known as *blind equalization*, has been widely applied to remove ISI caused by the unknown linear channel [19]. One of the drawbacks in the existing

blind equalization schemes is the convergence to a local minimum, which results in a closed eye-pattern. Without the open eye-pattern in the initialization stage, the adaptive DS/CDMA filter based on the blind equalization structure is likely to converge to a local minimum associated with a strong interference signal. Therefore, the direct application of a blind equalization algorithm has not been considered for the DS/CDMA demodulation. As a special case of blind equalization algorithm, a decision-directed LMS algorithm has been proposed in [31]. However, the simulation results of this algorithm shows that the algorithm works well only when the receiver can yield the BER of 10^{-1} or lower. This condition is never satisfied when the system is in the near-far situation.

Yamazaki and Kennedy in [37] apply a linearly-constrained adaptive algorithm to blind equalization. They interpret a linear constraint on the taps of an adaptive linear equalizer to use a cascade of a fixed linear prefilter and an adaptive equalizer with a single fixed-tap constraint. They suggest that a globally convergent blind equalization scheme can be formulated by either generalizing the tap constraint or generalizing the specific cost to a general convex cost and such a scheme may be applied to blind equalization. This idea gives a motivation for the constrained adaptive approach to address the blind MAI suppression problem.

Introduction to Linearly-Constrained LMS Algorithm

The linearly-constrained LMS optimization method has been proposed for constrained beamforming applications [36]. The linearly-constrained LMS algorithm is different from gradient projection (GP) algorithms in [16] and [17] in the algorithm implementation to satisfy the constraint such that the former projects the filter parameters onto the constraint space whereas latter projects the gradient onto the constraint space.

The constrained LMS algorithm minimizes the output power with a set of linear

equality constraints:

$$\min_{\mathbf{w}} E\left[\left(\mathbf{r}^T\mathbf{w}\right)^2\right] \qquad \text{subject to } \mathbf{C}^T\mathbf{w} = \mathbf{f} \tag{4.1}$$

where $\mathbf{r}$ is the filter input vector, $\mathbf{w}$ is the tap weight vector to be optimized. $\mathbf{C}$ is the constraint matrix and $\mathbf{f}$ is the constrained response vector. The objective function is formulated by adjoining the constraint function to the cost function by using the Lagrange multiplier technique:

$$J(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T\mathbf{R}\mathbf{w} + \lambda^T\left(\mathbf{C}^T\mathbf{w} - \mathbf{f}\right) \tag{4.2}$$

where $\mathbf{R} \triangleq E\left[\mathbf{r}\mathbf{r}^T\right]$ is the correlation matrix of the input vector $\mathbf{x}$ and λ is a Lagrange multiplier. Solving for the equation to satisfy the constraint, the constrained optimum weight vector is obtained by:

$$\mathbf{w}_{opt} = \mathbf{R}^{-1}\mathbf{C}\left[\mathbf{C}^T\mathbf{R}^{-1}\mathbf{C}\right]^{-1}\mathbf{f} \tag{4.3}$$

The adaptive algorithm is derived by temporarily assuming that the correlation matrix $\mathbf{R}$ is known. The initial weight vector is chosen to satisfy the constraint in Eq.(4.1). say

$$\mathbf{w}_0 = \mathbf{C}\left(\mathbf{C}^T\mathbf{C}\right)^{-1}\mathbf{f}. \tag{4.4}$$

and the weight vector is moved in the direction of the constrained gradient descent:

$$\begin{aligned} \mathbf{w}_{i+1} &= \mathbf{w}_i - \mu\nabla_{\mathbf{w}}(i) \\ &= \mathbf{w}_i - \mu\left(\mathbf{R}\mathbf{w}_i + \mathbf{C}\lambda_i\right). \end{aligned} \tag{4.5}$$

In the adaptive algorithm, the Lagrange multiplier is chosen to satisfy the constraint at each update stage:

$$\mathbf{C}^T\mathbf{w}_{i+1} = \mathbf{C}^T\mathbf{w}_i - \mu\mathbf{C}^T\mathbf{R}\mathbf{w}_i - \mu\mathbf{C}^T\mathbf{C}\lambda_i = \mathbf{f}. \tag{4.6}$$

Solving for λ_i and substituting into the Eq.(4.5), we have

$$\begin{aligned} \mathbf{w}_{i+1} &= \mathbf{w}_i - \mu \left[\mathbf{I} - \mathbf{C}\left(\mathbf{C}^T\mathbf{C}\right)^{-1}\mathbf{C}^T\right]\mathbf{R}\mathbf{w}_i \\ &\quad + \mathbf{C}\left(\mathbf{C}^T\mathbf{C}\right)^{-1}\left(\mathbf{f} - \mathbf{C}^T\mathbf{w}_i\right). \end{aligned} \tag{4.7}$$

Let us define the instantaneous estimate of $\mathbf{R}$ as $\mathbf{R}_i = \mathbf{r}_i\mathbf{r}_i^T$. Then we have the following stochastic LMS algorithm:

$$\begin{aligned} \mathbf{w}_0 &= \mathbf{F} \qquad (4.8) \\ \mathbf{w}_{i+1} &= \mathbf{P}\left[\mathbf{w}_i - \mu z_i \mathbf{r}_i\right] + \mathbf{F} \end{aligned}$$

where

$$\mathbf{P} \triangleq \mathbf{I} - \mathbf{C}\left[\mathbf{C}^T\mathbf{C}\right]^{-1}\mathbf{C}^T, \tag{4.9}$$

$$\mathbf{F} \triangleq \mathbf{C}\left[\mathbf{C}^T\mathbf{C}\right]^{-1}\mathbf{f}. \tag{4.10}$$

$z_i = \mathbf{w}_i^T\mathbf{r}_i$ is the array output and μ is the step size of the algorithm.

Practical Considerations in Constrained Adaptive Algorithm

Filter Output, Conditional Mean and Desired Signal

The adaptive algorithm adjusts the filter taps by minimizing the mean cost function. Generally, the conditional mean of the unconstrained MMSE solution depends only on the input signal, i.e.,

$$\widehat{d}_{UC} \triangleq E\left[d \mid \mathbf{r}\right]. \tag{4.11}$$

However, the conditional mean is also dependent on both the input signal and the prescribed constraint, if any constraint is incorporated.

The relationships among the filter output, the conditional mean and the desired signal in the constrained optimization algorithm can be described as follows. Let d, $\mathbf{r}$

and $\mathbf{w}$ denote a desired signal, an input signal, and a filter impulse response, respectively. The mean cost function of the MMSE solution can be written as

$$\begin{aligned} MMSE &= E\left(d-\mathbf{r}^T\mathbf{w}\right)^T\left(d-\mathbf{r}^T\mathbf{w}\right) \\ &= E\left(d-\widehat{d}+\widehat{d}-\mathbf{r}^T\mathbf{w}\right)^T\left(d-\widehat{d}+\widehat{d}-\mathbf{r}^T\mathbf{w}\right) \\ &= E\left(d-\widehat{d}\right)^T\left(d-\widehat{d}\right)+E\left(\widehat{d}-\mathbf{r}^T\mathbf{w}\right)^T\left(\widehat{d}-\mathbf{r}^T\mathbf{w}\right) \end{aligned} \tag{4.12}$$

where $\widehat{d}$ denotes the conditional mean estimator of d. The cross-terms on the right-hand side have vanished because the error $d-\widehat{d}$ is orthogonal to every measurable function of $\mathbf{r}$. Two points are evident in the Eq.(4.12) [38]:

- The MSE of a linear filter is never smaller than the MSE of the conditional mean estimator.
- The linear MMSE of d is also the linear MMSE estimator of the conditional mean $\widehat{d}$.

In the unconstrained MMSE solution, the conditional mean is equal to the desired output so that the MSE of the linear adaptive filter asymptotically approaches close to the MSE of the conditional mean estimator. However if any constraint is imposed on the MMSE solution, the conditional mean $\widehat{d}_{CO}$ is dependent on both the input signal and the constraint, i.e.,

$$\widehat{d}_{CO} \triangleq E\left[d \mid \mathbf{r}, \text{constraint}\right], \tag{4.13}$$

whereas the filter output $\mathbf{r}^T\mathbf{w}$ is forcefully to approach to the desired output d by descending the gradient of the cost function. The steady state performance of an adaptive algorithm is proportional to $J_{\min}$. Accordingly, the steady state performance of the constrained adaptive filter is highly dependent on whether the cost function is chosen appropriately to match the constraint function or not.

Constrained OPM Algorithm

Constrained adaptive algorithms proposed in [16] and [17] are investigated. These algorithms adjust the filter tap weights by minimizing the output power while the filter response to the signature sequence of the desired user is constrained to be unity. The optimization problem is given by

$$\min_{\mathbf{w}} E\left(\mathbf{w}^T\mathbf{r}\right)^2 \quad \text{subject to } \mathbf{w}^T\mathbf{s}_1 = \mathbf{1}. \tag{4.14}$$

Recall the representation of the filter output signal in the Eq.(3.15) and note that the linear optimum solution is the near-far resistant solution in the absence of AWGN. A conditional mean that is decided by both the filter input and the constraint is given by

$$\widehat{d} = A_1 b_1 \tag{4.15}$$

whereas the desired signal $d = 0$ because it is a trivial solution that minimizes the mean output power cost function.

Accordingly, if the perfect adaptation is achieved, the optimum cost of the output power minimizing (OPM) solution can be written as

$$\begin{aligned} J_{\min}(OPM) &= E(0 - A_1 b_1)^T (0 - A_1 b_1) + E\left(A_1 b_1 - \mathbf{r}^T\mathbf{w}\right)^T \left(A_1 b_1 - \mathbf{r}^T\mathbf{w}\right) && (4.16) \\ &= A_1^2 + J_{\min}(MMSE). && (4.17) \end{aligned}$$

where $J_{\min}(OPM)$ and $J_{\min}(MMSE)$ denote the optimum costs of the OPM solution and the MMSE solution, respectively. Similar interpretation is also given by Honig, Madhow and Verdu [16], who propose the OPM-based algorithm for the MAI suppression. Although they claim that the OPM solution is equivalent to the MMSE solution, this is no longer true. That is, $J_{\min}(OPM)$ becomes the signal output power plus the MSE between the filter output signal and the conditional mean estimator. This gives a motivation for an algorithm development to improve the filter performance, if any constraint is incorporated.

As a different standpoint, a fundamental drawback in the update rule of the OPM-based algorithm can be described as follows. Let us view each filter tap weight as having a nominal value plus a random variation. The nominal values of the filter tap weights produce the desired output without the random variation and AWGN. Since the algorithm is implemented for the filter to satisfy $\mathbf{s}_1^T \mathbf{w}_i = 1$ at ith update stage, the filter must produce the output of

$$z_i = A_1 b_1(i) \tag{4.18}$$

when the perfect adaptation is achieved. Accordingly, if the filter tap weights are optimum at ith update, the filter update rule becomes:

$$\mathbf{w}_{i+1} = \mathbf{w}_i - \mu A_1 b_{1,i} \tilde{\mathbf{r}}_i. \tag{4.19}$$

where $\tilde{\mathbf{r}}_i$ denotes the projected version of the received signal, becomes no longer optimum at $(i+1)$th update. In other words, adaptive filter coefficients in the OPM receiver fluctuate about the nominal optimum values according to the filter output power that is proportional to the strength of the desired signal power in the steady state. In order for the adaptive filter taps to stay in the nominal values in the steady state, the filter update rule should guarantee the zero tap gain increment when the perfect adaptation is achieved. Accordingly, a receiver equipped with the reference-free blind adaptive algorithms in [16] and [17] is problematic for the reliable data detection. In particular, the performance of those algorithms may be worse than that of the conventional receiver when the SNR is high.

Robust Filter Gain Maximization

Filter Gain and White Noise Gain

In the interference suppression problem, the improvement in signal-to-interference ratio (SIR) offered by filtering operation plays an important role. This improvement in

SIR is called the *filter gain*. For the design of a linear filter that maximizes the filter gain, the correlation matrix of the received signal can be decomposed into the desired signal and interference components:

$$\mathbf{R} = \sigma_S^2 \mathbf{P} + \sigma_I^2 \mathbf{Q}, \tag{4.20}$$

where $\mathbf{P}$ denotes the normalized correlation matrix of the desired signal component and $\mathbf{Q}$ denotes that of the interference components and σ_S^2/σ_I^2 is the input SIR. In CDMA multi-user communications, the interference is the MAI plus AWGN so that $\mathbf{P}$ and $\mathbf{Q}$ can be written by

$$\mathbf{P} = \mathbf{s}_1 \mathbf{s}_1^T, \tag{4.21}$$

$$\mathbf{Q} = \frac{1}{\sigma_I^2}\left(\sum_{k=2}^{K} A_k^2 \mathbf{s}_k \mathbf{s}_k^T + \sigma^2 \mathbf{I}\right) \tag{4.22}$$

and σ_S^2 is taken to be A_1^2. Then the *filter gain* for the MAI suppression can be expressed by

$$G_Q = \frac{\left|\mathbf{w}^T \mathbf{s}_1\right|^2}{\mathbf{w}^T \mathbf{Q} \mathbf{w}}. \tag{4.23}$$

and a linear filter that maximizes the filter gain should be in the form:

$$\mathbf{w}_{opt} = \kappa \mathbf{Q}^{-1} \mathbf{s}_1, \tag{4.24}$$

where κ is an arbitrary constant. These forms of the filter may be obtained, if $\mathbf{Q}$ and $\mathbf{s}_1$ are known *a priori*. It is also observed that the filter gain can be maximized by cascading an interference whitening filter $\mathbf{Q}^{-1}$ to the matched filter $\mathbf{s}_1$. In the absence of the MAI or any other structured noise, the filter gain becomes the *white noise gain*:

$$G_w = \frac{\left|\mathbf{w}^T \mathbf{s}_1\right|^2}{\mathbf{w}^T \mathbf{w}}. \tag{4.25}$$

The white noise gain is maximized by the matched filter, i.e., $\mathbf{w} = \mathbf{s}_1$.

In the filtering problem, the sensitivity of filter gain to a signal mismatch should be examined by considering the signal to be perturbed by small zero mean random errors

with a normalized covariance matrix Σ, so that the signal correlation matrix becomes

$$\tilde{\mathbf{P}} = \mathbf{s}_1\mathbf{s}_1^T + \zeta\boldsymbol{\Sigma} \tag{4.26}$$

where ζ is a strength parameter [39]. The sensitivity of filter gain to this random error is given by

$$S = \left(\frac{\partial G/\partial\zeta}{G}\right) = \frac{\mathbf{w}^T\boldsymbol{\Sigma}\mathbf{w}}{|\mathbf{w}^T\mathbf{s}_1|^2} = \frac{1}{G_\Sigma}. \tag{4.27}$$

When the errors are uncorrelated, the sensitivity is equal to the reciprocal of the white noise gain:

$$S_W = \frac{1}{G_W}. \tag{4.28}$$

Accordingly, the white noise gain can be used as a convenient measure of robustness in filtering operations [39].

Robust Constrained Filter Optimization

Since stability or convergence is a fundamental concern in adaptive algorithm behavior and since instability is manifested by large values of $\mathbf{w}$, one may penalize the norm of $\mathbf{w}$ during the iteration to avoid such instability. This section presents a robust optimization problem that maximizes filter gain subject to a constraint on the white noise gain. It has been shown that the constrained LMS algorithm in [36] maximizes the filter gain. An optimization problem can be formulated by

$$\max_{\mathbf{w}} G_Q \quad \text{subject to } S_W \leq \frac{1}{\delta^2} \tag{4.29}$$

where G_Q and G_W are defined in Eq.(4.23) and Eq.(4.25), respectively and the constraining value δ^2 is chosen to be less than or equal to the maximum allowable white noise gain. The maximum allowable white noise gain is obtained from the input signal-to-background noise ratio (SNR). The problem can be equivalently written as follows:

$$\min_{\mathbf{w}} \frac{\mathbf{w}^T\mathbf{Q}\mathbf{w}}{|\mathbf{w}^T\mathbf{s}_1|^2} + \lambda\frac{\mathbf{w}^T\mathbf{w}}{|\mathbf{w}^T\mathbf{s}_1|^2} = \min_{\mathbf{w}} \frac{\mathbf{w}^T\left(\mathbf{Q}+\lambda\mathbf{I}\right)\mathbf{w}}{|\mathbf{w}^T\mathbf{s}_1|^2} \tag{4.30}$$

where λ is a Lagrange multiplier. As long as $\mathbf{w}^T\mathbf{s}_1 \in \mathcal{R}$ is a nonzero constant, a robust optimum filter is given by

$$\mathbf{w}_{RO} = \kappa(\mathbf{Q} + \lambda\mathbf{I})^{-1}\mathbf{s}_1 \tag{4.31}$$

where κ is an arbitrary constant that related to the value of $\mathbf{w}^T\mathbf{s}_1$ and λ is an adjustable variable that satisfies the white noise gain constraint. One may realize that constraining that $\mathbf{w}^T\mathbf{s}_1 = 1$ does not lose any generality. With this constraint, the robust constrained optimum (RCO) solution is given by

$$\mathbf{w}_{RCO} = \frac{(\mathbf{Q} + \lambda\mathbf{I})^{-1}\mathbf{s}_1}{\mathbf{s}_1^T(\mathbf{Q} + \lambda\mathbf{I})^{-1}\mathbf{s}_1}. \tag{4.32}$$

In this solution, the Lagrange multiplier λ provides a continuous monotonic parameterization between the unconstrained optimum ($\lambda = 0$) solution and the conventional ($\lambda = \infty$) matched filter solution.

Robust Constrained LMS Algorithm

The constrained LMS algorithm presented in [36] may be useful in array beamforming applications. However, the constrained LMS algorithm that minimizes the output power cannot be applied for the data detection because the algorithm generates the relatively high self-noise even in the high SNR environment. This phenomenon occurs when the adaptive algorithm includes the signal component in the minimization process as the output power. To overcome this problem and to ensure near-optimum steady state performance, the author formulates a constrained LMS optimization problem that minimizes a modified non-MSE cost function based on the cost-constraint strategy. The strategy is to form a mean cost function that is a scaled version of the actual MSE but share a common point with the constraint function. It has been shown that a linear constraint from the signal properties can be used to ensure the desired convergence of the adaptive algorithm. For the improved performance of the filter, the author extends this

idea such that the constrained filter is stable in the steady state and robust to the signal mismatch and the filter implementation error. That is, the filter is further constrained to have a bounded norm based on the white noise gain constraint. Filter response constraint leads the filter to converge to the desired optimum whereas the white noise gain constraint gives robustness to the random variation of the system parameters. Since the interference correlation matrix $\mathbf{Q}$ in Eq.(4.32) is not readily available for the algorithm implementation, we develop a non-MSE cost function that is similar to the actual MSE but approaches close to zero in the steady state.

Interference Suppression Problem

The robust constrained LMS approach can be explained by applying the idea to the blind MAI suppression problem. Without loss of generality, let us constrain the filter response to the signature sequence of the desired signal to be $\mathbf{w}^T\mathbf{s}_1 = 1$. The near-far resistant filter should asymptotically produce $A_1 b_1(i)$ for all i, provided that the constraint, $\mathbf{w}^T\mathbf{s}_1 = 1$, is satisfied and without AWGN. Therefore, a plausible cost function that satisfies the constraint $\mathbf{w}^T\mathbf{s}_1 = 1$ is chosen by

$$E\left[\left(\mathbf{r}_i^T\mathbf{w} - A_1 b_1(i)\right)^2\right]. \tag{4.33}$$

By combining the filter response constraint and the white noise gain constraint, a constrained optimization problem is formulated by

$$\begin{aligned} \min_{w} \quad & E\left[\left(\mathbf{r}_i^T\mathbf{w} - A_1 b_1(i)\right)^2\right] \\ \text{subject to} \quad & \mathbf{w}^T\mathbf{s}_1 = 1 \text{ and } \mathbf{w}^T\mathbf{w} \le \delta^2. \end{aligned} \tag{4.34}$$

To solve this constrained optimization problem, one can write an objective function using the method of Lagrange multipliers:

$$J(\mathbf{w}) = \frac{1}{2}\left[\mathbf{w}^T\mathbf{R}\mathbf{w} - 2\mathbf{w}^T\mathbf{p} + A_1^2\right] + \lambda_1\left[\mathbf{w}^T\mathbf{w} - \delta^2\right] + \lambda_2\left[\mathbf{w}^T\mathbf{s}_1 - 1\right] \tag{4.35}$$

where $\mathbf{R} \triangleq E\left[\mathbf{r}_i\mathbf{r}_i^T\right]$, $\mathbf{p} \triangleq E\left[A_1 b_1(i)\,\mathbf{r}_i\right] = A_1^2\mathbf{s}_1$, λ_1 and λ_2 are Lagrange multipliers which are adjusted to satisfy the constraints. The robust constrained optimum filter that minimizes the objective function is

$$\mathbf{w}_{opt} = \frac{\left[\mathbf{R} + \lambda_1\mathbf{I}\right]^{-1}\mathbf{s}_1}{\mathbf{s}_1^T\left[\mathbf{R} + \lambda_1\mathbf{I}\right]^{-1}\mathbf{s}_1}. \tag{4.36}$$

Observe that the Lagrange multiplier λ_2 is absorbed into the denominator whereas the Lagrange multiplier λ_1 is added to each eigenvalue of the matrix $\mathbf{R}$ without modifying the eigenvectors. The Lagrange multiplier λ_1 is continuously parameterized to satisfy the given constraint condition.

Adaptation Rule

Constrained adaptive algorithms can be classified into two categories; soft-constrained one and hard-constrained one. The adaptive algorithm considered in this section is a hard-constrained one, which satisfies the constraints exactly at each adaptation stage. Previous works on blind adaptive techniques for DS/CDMA detection used the gradient projection technique of [40]; It has been shown that gradient-projection methods are susceptible for long runs without an additional error-correction procedure [36]. The constrained least-mean-squares (CLMS) algorithm [36] is designed to avoid error accumulation while maintaining a hard constraint; as a result, it can provide continual filtering for arbitrary large number of iterations.

In the near-far situation, the desired data bit $b_1(i)$ is not readily available for the blind adaptation nor cannot be estimated by the matched filter. It is assumed that the amplitude or the channel gain of the desired signal is available for the adaptation. Since $b_1(i)$ is known to be equally probable in $\{+1, -1\}$ and cannot be predictable, the maximum a posteriori (MAP) estimation rule for $b_1(i)$ can be given by

$$\widehat{b}_1(i) = \mathrm{sgn}\left(\mathbf{r}_i^T\mathbf{w}_i\right). \tag{4.37}$$

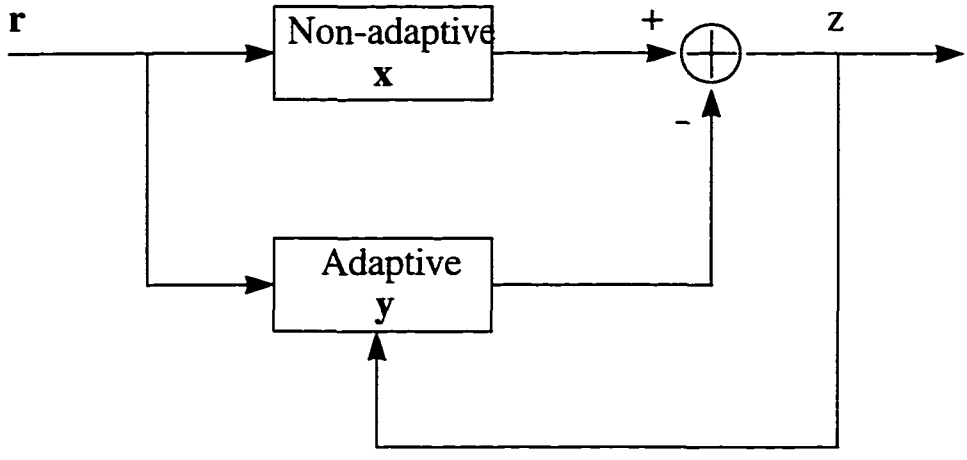

Figure 4.1 Decomposed form of the constrained adaptive filter.

Then the instantaneous gradient of the estimated cost function can be written as

$$\widehat{\nabla}_w(i) = \left[\mathbf{r}_i^T \mathbf{w}_i - A_1 \cdot \mathrm{sgn}\left(\mathbf{r}_i^T \mathbf{w}_i\right)\right] \mathbf{r}_i. \tag{4.38}$$

For the adaptive implementation of a constrained filter, we can decompose the filter impulse response into the sum of two orthogonal components associated with the constraint space and its orthogonal space [16]. It can be described by

$$\mathbf{w} = \mathbf{x} + \mathbf{y}, \tag{4.39}$$

where $\mathbf{x}$ is the non-adaptive component to satisfy the constraints and $\mathbf{y}$ is the adaptive component that is the projection of $\mathbf{w}$ onto the space orthogonal to $\mathbf{x}$. The realization of a filter by decomposing into two component is given by Figure 4.1.

The algorithm is implemented using three step procedures.

1. A tentative vector for the $(i+1)$th update, $\widetilde{\mathbf{w}}_{i+1}$, is obtained in a similar manner as a conventional LMS algorithm;

$$\widetilde{\mathbf{w}}_{i+1} = \mathbf{w}_i - \mu \left[\mathbf{r}_i^T \mathbf{w}_i - A_1 \cdot sgn\left(\mathbf{r}_i^T \mathbf{w}_i\right)\right] \mathbf{r}_i \tag{4.40}$$

where μ is the step size to control the convergence of the algorithm.

2. The adaptive component of the $(i+1)$th filter vector, $\mathbf{y}_{i+1}$, is refined by projecting the tentative vector, $\widetilde{\mathbf{w}}_{i+1}$, onto the set of constraint spaces, which can be described by

$$\mathbf{y}_{i+1} = \mathbf{P}_2 \cdot \mathbf{P}_1 \cdot \widetilde{\mathbf{w}}_i, \tag{4.41}$$

where $\mathbf{P}_1$ and $\mathbf{P}_2$ are the projection operators, which are based on the linear-constraint and the quadratic-constraint in Eq.(4.34), respectively.

3. The total filter impulse response for the $(i+1)$th update is obtained by

$$\begin{aligned} \mathbf{w}_{i+1} &= \mathbf{x} + \mathbf{y}_{i+1} && (4.42) \\ &= \mathbf{x} + \mathbf{P}_2 \cdot \mathbf{P}_1 \cdot \left\{ \mathbf{w}_i - \mu \left[\mathbf{r}_i^T \mathbf{w}_i - A_1 \cdot sgn\left(\mathbf{r}_i^T \mathbf{w}_i\right) \right] \mathbf{r}_i \right\} && (4.43) \end{aligned}$$

After some mathematical manipulation of Eq.(4.34), one can easily find the non-adaptive component of the filter impulse response and the corresponding projection operators which are given by

$$\begin{aligned} \mathbf{x} &= \mathbf{s}_1 && (4.44) \\ \mathbf{P}_1 \cdot \mathbf{w}_i &= \mathbf{w}_i - \langle \mathbf{s}_1, \mathbf{w}_i \rangle \mathbf{s}_1, && (4.45) \end{aligned}$$

and

$$\mathbf{P}_2 \cdot \mathbf{w}_i = \begin{cases} \mathbf{w}_i, & \text{if } \langle \mathbf{w}_i, \mathbf{w}_i \rangle \leq \delta^2 - \mathbf{s}_1^T \mathbf{s}_1 \\ \sqrt{\delta^2 - \mathbf{s}_1^T \mathbf{s}_1} \frac{\mathbf{w}_i}{\sqrt{\mathbf{w}_i^T \mathbf{w}_i}}, & \text{otherwise} \end{cases} . \tag{4.46}$$

Although the information obtained by $\text{sgn}\left(\mathbf{r}_i^T \mathbf{w}_i\right)$ is not reliable in the initial adaptation stage, the projection applied at each update stage leads the filter to converge to a desired minimum in the mean sense. Furthermore, as the filter approaches to the steady state, the filter output is likely to be correct so that the mean cost function asymptotically converges to zero.

This algorithm is different from blind adaptive algorithms of [16] and [17] in three major aspects: First, the algorithm satisfies all the constraints exactly at each step of

updates so that any implementation errors such as roundoff, computational error and quantization error will not accumulate in the iteration process. Second, the algorithm guarantees the zero tap gain increment when the perfect adaptation achieved so that the stochastic driving term of the former produces much smaller asymptotic MSE than that of the latter for the same rate of convergence. Third, the algorithm provides better tolerance to random variations of system parameters (i.e., the filter and the signal) by incorporating the white noise gain constraint.

Algorithm Analysis

Convergence in the Mean

The filter tap vector $\mathbf{w}$ obtained by the adaptive algorithm is a random vector. Convergence of the mean tap vector to the optimum is described by showing that the norm of the difference vector between the mean tap vector and the optimum asymptotically approaches zero. Proof of convergence in the mean is simplified by the independent assumption that successive samples of the input vector are statistically independent.

Let us define the tap weight error vector for the CLMS algorithm as the difference between the adaptive weight vector at the ith iteration and the optimum weight vector in Eq.(4.36)

$$\varepsilon_i \triangleq \mathbf{w}_i - \mathbf{w}_{opt} \tag{4.47}$$

Subtracting the optimum weight vector $\mathbf{w}_{opt}$ from both sides of Eq.(4.43), one may write the iterative equation for the constrained LMS algorithm in terms of the weight error vector $\mathbf{e}_i$ as follows:

$$\varepsilon_{i+1} = \mathbf{P}\left[\mathbf{I} - \mu \mathbf{r}_i \mathbf{r}_i^T\right] \varepsilon_i + \mu \mathbf{P} \epsilon_{\mathbf{opt},i} \mathbf{r}_i \tag{4.48}$$

where $\mathbf{P} \triangleq \mathbf{P}_2\mathbf{P}_1$, $\mathbf{I}$ is the identity matrix and

$$e_{opt,i} \triangleq A_1 \cdot \mathrm{sgn}\left(\mathbf{r}_i^T \mathbf{w}_{opt}\right) - \mathbf{r}_i^T \mathbf{w}_{opt} \tag{4.49}$$

is the estimation error produced in the constrained optimum MMSE solution.

Taking the mathematical expectation on both sides of the algorithm in Eq.(4.48) yields

$$E(\varepsilon_{i+1}) = \mathbf{P}E\left\{\left[\mathbf{I} - \mu\mathbf{r}_i\mathbf{r}_i^T\right]\varepsilon_i\right\} + \mu\mathbf{P}E(e_{opt,i}\mathbf{r}_i). \tag{4.50}$$

Using $\mathbf{R} = E\left(\mathbf{r}_i\mathbf{r}_i^T\right)$ and because of the independent assumption, one can write the first expectation term on the right side of Eq.(4.50) as

$$E\left\{\left[\mathbf{I} - \mu\mathbf{r}_i\mathbf{r}_i^T\right]\varepsilon_i\right\} = [\mathbf{I} - \mu\mathbf{R}]E(\varepsilon_i). \tag{4.51}$$

For the second expectation term of Eq.(4.50), let us invoke the principle of orthogonality:

$$E(e_{opt,i}\mathbf{r}_i) = \mathbf{0}. \tag{4.52}$$

Accordingly, the simplified form of Eq.(4.50) becomes

$$\begin{aligned} E(\varepsilon_{i+1}) &= \mathbf{P}[\mathbf{I} - \mu\mathbf{R}]E(\varepsilon_i) \\ &= \mathbf{P}^{i+1}[\mathbf{I} - \mu\mathbf{R}]^{i+1}E(\varepsilon_0). \end{aligned} \tag{4.53}$$

The idempotence of the projection operator $\mathbf{P}$ (i.e., $\mathbf{P}^{i+1} = \mathbf{P}$) yields that $\mathbf{P}E(\varepsilon_i) = E(\varepsilon_i)$ for all i, and Eq.(4.53) can be written as

$$E(\varepsilon_{i+1}) = [\mathbf{I} - \mu\mathbf{PRP}]^{i+1}E(\varepsilon_0). \tag{4.54}$$

Note that the projected crosscorrelation matrix, $\mathbf{PRP}$, determines both the convergence rate of the mean weight vector to the optimum and the steady state variance of the weight vector about the optimum.

Examination of $\varepsilon_0 = \mathbf{s}_1 - \mathbf{w}_{opt}$ shows that it can be expressed as a linear combination of the eigenvectors of $\mathbf{PRP}$ corresponding nonzero eigenvalues. If $\mathbf{e}_0$ is equal to an eigenvector of $\mathbf{PRP}$, say $\mathbf{u}_k$ with the eigenvalue $\lambda_k \neq 0$ then

$$E(\varepsilon_{i+1}) = [\mathbf{I} - \mu\mathbf{PRP}]^{i+1}\mathbf{u}_k \tag{4.55}$$

$$= [1 - \mu\lambda_k]^{i+1}\mathbf{u}_k. \tag{4.56}$$

Therefore, the convergence of the mean tap weight vector to the optimum is therefore geometric with a geometric ratio

$$(1 - \mu\lambda_k)\,. \tag{4.57}$$

This geometric series converges to zero when

$$-1 < 1 - \mu\lambda_k < 1. \tag{4.58}$$

That is, for the CLMS algorithm, $E(\varepsilon_i)$ converges to zero as $i \rightarrow \infty$, if the following condition is satisfied

$$0 < \mu < \frac{2}{\lambda_{\max}} \tag{4.59}$$

where $\lambda_{\max}$ is the largest eigenvalue of the matrix $\mathbf{PRP}$. Therefore the filter tap vector converges to the optimum, i.e.,

$$\lim_{i\rightarrow\infty} \|E(\mathbf{w}_i) - \mathbf{w}_{opt}\| = 0 \tag{4.60}$$

with properly chosen μ given by Eq.(4.59).

Steady State Performance

The constrained adaptive algorithm is designed to continually adjust the filter taps for coping with nonstationary noise environment. In the steady state of stationary environment, this adaptation causes the tap weight to have a random variation about the optimum and produces an additional component of noise to appear at the filter output. The difference between the ith MSE produced by the adaptive algorithm, $J(i)$ and the optimum cost $J_{\min}$ is defined by the excess MSE [34]. The asymptotic excess MSE, $J_{ex}(\infty)$, provides a measure of performance of the algorithm in the steady state.

Let us express the estimation error, e_i, produced by the constrained adaptive algorithm as

$$e_i \;=\; A_1 b_1(i) - \mathbf{w}_i^T \mathbf{r}_i \tag{4.61}$$

$$= A_1 b_1(i) - \mathbf{w}_{opt}^T \mathbf{r}_i - \varepsilon_i^T \mathbf{r}_i \tag{4.62}$$

$$= e_{opt,i} - \varepsilon_i^T \mathbf{r}_i \tag{4.63}$$

where $e_{opt,i}$ is the estimation error in the constrained optimum solution and assuming that $sgn\left(\mathbf{w}_i^T \mathbf{r}_i\right) = b_1(i)$. Let $J(i)$ denote the MSE due to the adaptive algorithm at the ith update, which is given by

$$J(i) = E\left(|e_i|^2\right) \tag{4.64}$$

$$= E\left[\left(e_{opt,i} - \varepsilon_i^T \mathbf{r}_i\right)\left(e_{opt,i} - \varepsilon_i^T \mathbf{r}_i\right)\right] \tag{4.65}$$

$$= J_{\min} + E\left(\varepsilon_i^T \mathbf{r}_i \mathbf{r}_i^T \varepsilon_i\right) \tag{4.66}$$

where $J_{\min}$ is the optimum cost produced by the optimum filter. The second term of right-hand side of Eq.(4.66) is the excess MSE, which is given by

$$J_{ex}(i) = E\left(\varepsilon_i^T \mathbf{r}_i \mathbf{r}_i^T \varepsilon_i\right). \tag{4.67}$$

Steady state excess MSE $J_{ex}(\infty)$ can be expressed in terms of the eigenvalues of the correlation matrix $\mathbf{R}$, that is given by [34]

$$J_{ex}(\infty) = J_{\min} \frac{\sum_{k=1}^{N} \mu\lambda_k / (2 - \mu\lambda_k)}{1 - \sum_{k=1}^{N} \mu\lambda_k / (2 - \mu\lambda_k)} \tag{4.68}$$

where μ is the step size and λ_k is the eigenvalue of the projected correlation matrix $\mathbf{PRP}$. Note that $J_{ex}(\infty)$ is directly proportional to the optimum cost $J_{\min}$.

Comparison with the OPM Algorithm

Let $J_{\min}(CLMS)$ and $J_{\min}(OPM)$ denote the optimum costs of the CLMS and the OPM algorithms. When the filter is near-far resistant, one can write the optimum cost of the LMS solution as

$$J_{\min}(CLMS) = 1 - A_1^2 s_1^T R^{-1} s_1 = \frac{\gamma\sigma^2}{\gamma\sigma^2 + \eta A_1^2} \tag{4.69}$$

whereas that of the OPM solution as

$$J_{\min}(OPM) = \left[s_1^T R^{-1} s_1\right]^{-1} = \eta A_1^2 + \gamma\sigma^2 \tag{4.70}$$

where γ is the filter white noise gain and η is the asymptotic efficiency.

Since the optimum cost of the OPM solution is much greater than that of the CLMS solution for small level of background noise, the asymptotic excess MSE of the OPM algorithm is significantly greater than that of the CLMS algorithm. As the background noise vanishes, the difference between $J_{\min}(CLMS)$ and $J_{\min}(OPM)$ is likely to be substantial. This implies that the steady state performance of the OPM algorithm is worse than that of the CLMS algorithm. In other words, each component of the stochastic driving term of the OPM algorithm has a variance on the order of $J_{\min}(OPM)$ whereas each component of the stochastic driving term for the CLMS algorithm has a variance on the order of $J_{\min}(CLMS)$.

Error Correcting Features

This section describes a simple geometric interpretation to visualize the error-correcting features of the CLMS algorithm. The linearly-constrained LMS algorithm is simply written by

$$\mathbf{w}_{i+1} = \mathbf{P}_1\left[\mathbf{w}_i - \mu e_i \mathbf{r}_i\right] + \mathbf{s}_1. \tag{4.71}$$

where P_1 is the projection operator that satisfies $\mathbf{w}^T\mathbf{s}_1 = 1$. It has been shown that a difficulty is raised in applying the gradient projection algorithm to the real-time array processing problem due to the error propagation effect from implementation errors such as truncation, roundoff, or quantization errors. The gradient projection (GP) algorithm [33] is named because that the unconstrained gradient estimate is projected onto the constrained subspace and added to the current weight vector. The GP algorithm is given by

$$\mathbf{w}_{i+1} = \mathbf{w}_i - \mu z_i \mathbf{P}_1 \mathbf{r}_i. \tag{4.72}$$

Comparison with the GP Algorithm

Let us define the constraint subspace Ω and the constraint hyperplane Φ as

$$\begin{aligned} \Omega &= \left\{\mathbf{w} : \mathbf{s}_1^T \mathbf{w} = 0\right\} \\ \Phi &= \left\{\mathbf{w} : \mathbf{s}_1^T \mathbf{w} = 1\right\}. \end{aligned}$$

A comparison between the effect of the computational errors on the linearly-constrained LMS algorithm and that of the GP algorithm is shown in Figure 4.2 and Figure 4.3. During the implementation of the algorithm, it is likely that small computational errors will occur at iteration. The GP algorithm has the problem of accumulating the error in the iteration process so that this algorithm is limited to problems requiring few enough iterations. On the other hand, the CLMS algorithm has the capability of continuously correcting such errors and preventing them from accumulating. Accumulation of errors in the GP algorithm can be expected to cause the weight vector to do a random work away from the constraint hyperplane with a variance increasing linearly with the number of iterations. However, the expected deviations of the CLMS algorithm from the constraint remains at its original value.

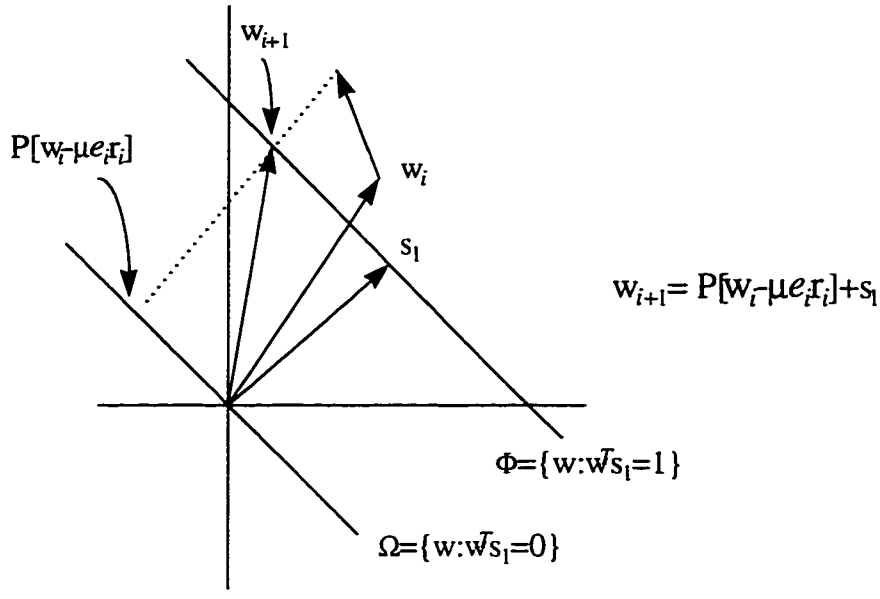

Figure 4.2 Error correction feature in the constrained LMS algorithm.

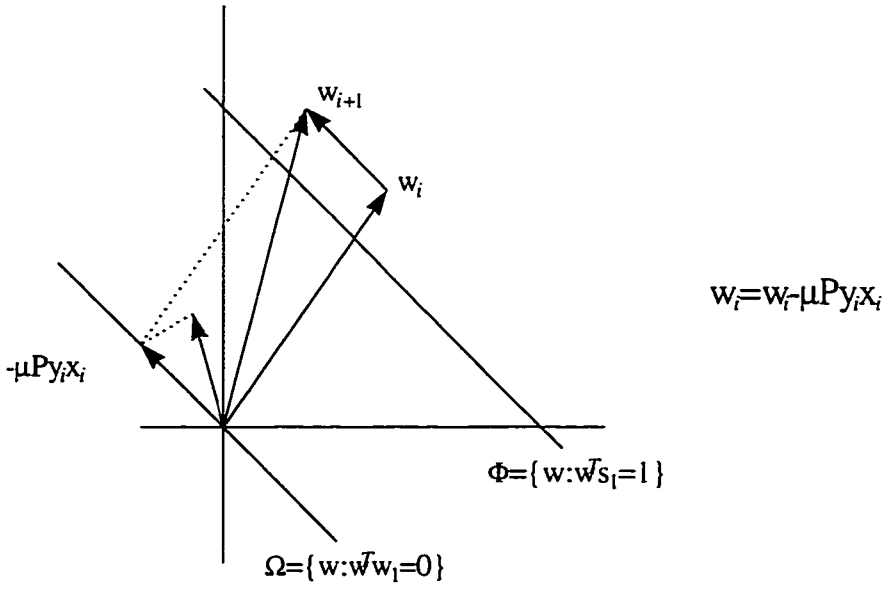

Figure 4.3 Error propagation effect in the gradient projection algorithm

CHAPTER 5 JOINT ADAPTIVE MAI SUPPRESSION

This chapter investigates a new approach to blind MAI suppression in which the receiver performs joint filter and reference updates in a sequential manner. The constrained LMS algorithm requires *a priori* knowledge or good estimate of the amplitude of the desired signal for the adaptation reference, which may be unknown to the receiver or is time-varying in a nonstationary environment. By assuming no knowledge on the amplitude of the desired signal, this chapter introduces the adaptive SAGE algorithm for the joint filter and amplitude estimation and applies it to the blind MAI suppression problem in a nonstationary environment. Let us consider a nonconvex cost function that has unknown parameters more than one element. The constrained LMS algorithm may be an example, in which the adaptation reference is unknown besides the filter tap weights. To address this problem, we divide the unknown parameters in the cost function into a number of indexed parameter sets and decompose the joint estimation problem of the unknown parameter sets into a number of decoupled parameter estimation problems. Then assuming that the other parameter set is optimal, each indexed parameter set is sequentially updated in the direction of minimizing the conditional cost function. That is, for each received signal vector, the algorithm iterates between estimating a conditional cost based on the current input vector and the previous estimates of the parameter sets (E-step) and updating an indexed parameter set by minimizing the conditional cost with respect to the indexed parameter set (M-step). This algorithm will be useful for the dynamic DS/CDMA detection in which the channel accessed by the desired signal is time-varying or nonstationary.

Introduction to SAGE Method

In a variety of signal processing applications, direct calculations of maximum-likelihood (ML) or maximum penalized-likelihood parameter estimates are intractable due to the complexity of the likelihood functions, to the coupling introduced by smoothness penalties, or both [41]. The EM algorithm [42] has been useful for iterative joint ML parameter estimation in the statistical estimation. However, the simultaneous update used by the EM algorithm leads to slow convergence and difficult maximization steps due to coupling when smoothness penalties are used [41]. The SAGE algorithm of [41] has been proposed to overcome these problems by sequentially updating the indexed parameter sets. Since the SAGE algorithm is a generalization of the EM algorithm, let us briefly investigate the EM algorithm.

Expectation-Maximization (EM) Algorithm

Let us denote by Y the data vector with the associated probability density function (pdf) $f_Y(y:\vartheta)$ indexed by the parameter vector ϑ where the possible parameter values are contained in a set Θ. Given an observed y, the joint ML estimate $\widehat{\vartheta}_{ML}$ is the value of ϑ that maximizes the log-likelihood, that is,

$$\widehat{\vartheta}_{ML} = \arg\left\{\max_{\vartheta\in\Theta} \log f_Y(y:\vartheta)\right\} \tag{5.1}$$

Suppose that the data vector Y can be viewed as incomplete and we can specify some complete data X related to Y by

$$H(X) = Y \tag{5.2}$$

where $H(\cdot)$ is a noninvertible transformation. The EM algorithm is an iterative algorithm to find the solution to Eq.(5.1) but it does so by making an essential use of complete data specification [43]. It starts with an initial guess ϑ_0 and ϑ_{k+1} is defined by

$$\vartheta_{k+1} = \arg\left\{\max_{\vartheta\in\Theta} E\left[\log f_X(x;\vartheta) \mid y;\vartheta_k\right]\right\} \tag{5.3}$$

where $f_X(x;\vartheta)$ is the pdf of X, and $E[\cdot \mid y;\vartheta_k]$ denotes the conditional expectation given y, computed using the parameter value ϑ_k.

An intuitive idea in the EM algorithm is that one would like to choose the ϑ that maximizes $\log f_X(x;\vartheta)$, the log likelihood of the complete data. However since $\log f_X(x;\vartheta)$ is not available, we maximize instead its expectation, given the observed data y. Since the conditional expectation is not exact, the algorithm iterates, using each new parameter estimate to improve the conditional expectation on the next iteration cycle (E-step) and then uses this conditional estimate to improve the next parameter estimate (M-step) [43].

Space-Alternating Generalized EM (SAGE) Algorithm

The simultaneous update used by the EM algorithm requires too informative complete-data spaces, which lead to slow convergence. The SAGE algorithm that trade-offs between convergence rate and complexity is suited to problems where one can sequentially update small groups of the elements of the parameter vector. Besides the convergence rate, the SAGE algorithm ensures monotonic increase in the objective function because the SAGE method is based on statistical considerations.

Let us define the log-likelihood by

$$\Phi(\vartheta) = \log f_Y(y:\vartheta). \tag{5.4}$$

Direct maximization of Φ is often intractable due to the complexity, the coupling, or both. The SAGE algorithm updates the subsets of the elements of the parameter vector ϑ. Following definitions formalize the idea.

Definition 1 *[41] A set S is defined to be an index set if it (a) is nonempty, (b) is a subset of the set $\{1,2,\cdots,p\}$, and (c) has no repeated entries. The set $\overline{S}$ denotes the complement of S intersected with $\{1,2,\cdots,p\}$.*

Let the number of elements in the set S be m. Then, one can use ϑ_S to denote the m-dimensional vector consisting of the m elements of ϑ indexed by the members of S. Similarly, define $\vartheta_{\overline{S}}$ to be the $(p-m)$-dimensional vector consisting of the remaining elements of ϑ. Then, one can define expressions such as the following to be equivalent:

$$\Phi(\vartheta) = \Phi(\vartheta_S, \vartheta_{\overline{S}}). \tag{5.5}$$

Another important terminology is an admissible hidden-data space which should be identified to generate the conditional likelihood function ϕ^S for each index set S of interest. It is defined in the following sense:

Definition 2 *[41] A random vector X^S with pdf $f(x;\vartheta)$ is an admissible hidden-data space with respect to ϑ_S for $f(y;\vartheta)$ if the joint density of X^S and Y satisfies*

$$f(y, x; \vartheta) = f(y \mid x, \vartheta_{\overline{S}})\, f(x; \vartheta). \tag{5.6}$$

In other words, the conditional distribution $f(y \mid x; \vartheta_{\overline{S}})$ must be independent of ϑ_S.

The basic idea behind the SAGE method is as follows. By introducing a hidden-data space for ϑ_S based on the statistical structure of the likelihood, one can replace the maximization of $\Phi\left(\vartheta_S, \vartheta_{\overline{S}}^k\right)$ over ϑ_S with the maximization of another likelihood function $\phi^S\left(\vartheta_S, \vartheta^k\right)$. If the hidden-data space is chosen appropriately, one can maximize the function $\phi^S\left(\cdot, \vartheta^k\right)$ analytically. Just as for the EM algorithm, ϕ^S is constructed to ensure that increases in ϕ^S yield increases in Φ. An essential ingredient of the SAGE algorithm is the conditional expectation of the log-likelihood of X^S, $\phi^S\left(\vartheta_S; \overline{\vartheta}\right)$:

$$\begin{aligned}
\phi^S\left(\vartheta_S; \overline{\vartheta}\right) &= \phi^S\left(\vartheta_S; \overline{\vartheta}_S, \overline{\vartheta}_{\overline{S}}\right) && (5.7)\\
&\triangleq E\left\{\log f\left(X^S; \vartheta_S, \overline{\vartheta}_{\overline{S}}\right) \mid Y = y; \overline{\vartheta}\right\} && (5.8)\\
&= \int f\left(x \mid Y = y; \overline{\vartheta}\right) \log f\left(x; \vartheta_S, \overline{\vartheta}_{\overline{S}}\right) dx. && (5.9)
\end{aligned}$$

A basic SAGE algorithm is given as follows. Let $\vartheta^0 \in \Theta$ be an initial parameter estimate. For $k = 0, 1, \cdots$

1. Choose an index set $S = S_k$.
2. Choose an admissible hidden-data space X^S for ϑ_{S_k}.
3. E step: Compute $\phi^{S_k}\left(\vartheta_{S_k}; \vartheta^k\right)$ using (4).
4. M step:

$$\begin{aligned} \vartheta_{S_k}^{k+1} &= \arg \max_{\vartheta_{S_k} \in \Theta^S\left(\vartheta^k\right)} \phi^{S_k}\left(\vartheta_{S_k}; \vartheta^k\right) \\ \vartheta_{\overline{S}_k}^{k+1} &= \vartheta_{\overline{S}_k}^{k} \end{aligned}$$

where

$$\Theta^S\left(\vartheta^k\right) = \left\{\vartheta_{S_k} : \left(\vartheta_{S_k}, \vartheta_{\overline{S}_k}^k\right) \in \Theta\right\}.$$

5. Repeat steps 3 and 4, if necessary.

The SAGE algorithm is a generalization of the EM algorithm in which one alternate between several hidden-data spaces rather than using just one complete-data space, and updates only a subset of the elements of the parameter vector at each iteration. By updating the parameters sequentially, rather than simultaneously, the SAGE algorithm is shown to yield faster convergence than the EM algorithm in several signal processing applications [41].

Adaptive Algorithm with Approximate SAGE Mapping

The SAGE algorithm is a block processing algorithm in each iteration, which requires maximizing a new likelihood function in each iteration. This cannot be performed in general, since it involves an extensive computational load and storage requirement as more data are processed. To overcome this problem and to achieve an adaptive implementation, a sequential algorithm with an approximate SAGE mapping is presented and called the *adaptive SAGE algorithm*. Although the algorithm is no longer the SAGE algorithm, the algorithm stochastically approximates the SAGE algorithm. Thus the convergence

results and the asymptotic behavior of the stochastic approximation method are applied to the algorithm.

The SAGE algorithm maximizes the log-likelihood function by sequentially updating the decoupled parameter sets. Without loss of any generality and for the adaptive implementation, one can replace the log-likelihood function to a cost function and minimize the cost function instead of maximizing the log-likelihood function with respect to the unknown parameters. Consider a minimization problem:

$$\min_{\theta \in \Omega} \quad J_i(\theta) \tag{5.10}$$

where $J_i(\theta)$ is a mean cost function to be minimized in the statistical sense, θ is a parameter vector to be optimized and Ω is a constraint set. Some definitions are formalized before presenting the adaptive SAGE algorithm.

For conceptual clarity, let us simplify the problem in which the number of indexed parameter sets is two. Then, define the parameter set one as $\theta_1 \in \Omega_1$ and the parameter set two as $\theta_2 \in \Omega_2$ so that $\theta \in \Omega$ where

$$\theta = [\theta_1, \theta_2], \tag{5.11}$$

$$\Omega = \{(a, b) : a \in \Omega_1, b \in \Omega_2\}. \tag{5.12}$$

Define k to be the index for the parameter considered and $\overline{k}$ to be the other index. i.e., if $k = 1$, then $\overline{k} = 2$ and vice versa. Define P_Ω to be the projection operator to satisfy the constraint set Ω, i.e.,

$$P_\Omega \mathbf{x} \in \Omega, \quad \text{for any } \mathbf{x}. \tag{5.13}$$

Define a conditional cost function for the parameter indexed k to be

$$J_i^k\left(\theta_k; \widehat{\theta}\right) = J_i^k\left(\theta_k; \widehat{\theta}_k, \widehat{\theta}_{\overline{k}}\right). \tag{5.14}$$

Assume that the input signal is wide-sense stationary or slowly time-varying. The adaptive SAGE algorithm is implemented as follows:

Let $\theta_1(0) = P_{\Omega_1} \cdot 0$ and $\theta_2(0) = P_{\Omega_2} \cdot 0$ be initial parameter estimate. For $i = 0, 1, \cdots$

1. Choose an index $k = k^i$.
2. Choose an admissible hidden-data space $X^k(i)$ for $\theta_k(i)$.
3. E step: Compute $J_k^i\left(\theta_k; \widehat{\theta}(i)\right)$ using (4).
4. M step:

$$\begin{aligned} \widehat{\theta}_k(i+1) &= \arg\min_{\theta_k \in \Omega_k} J_i^k\left(\theta_k; \widehat{\theta}(i)\right) \\ \widehat{\theta}_{\overline{k}}(i+1) &= \widehat{\theta}_{\overline{k}}(i) \end{aligned}$$

where the minimization is realized by a constrained LMS algorithm:

$$\widehat{\theta}_k(i+1) = P_{\Omega_k} \cdot \left[\widehat{\theta}_k(i) - \mu \nabla J_k(i)\right].$$

5. Repeat steps 1 - 4 for $\overline{k}^i$.

This algorithm monotonically minimizes the cost function such that

$$J_k^{i+1} \leq J_{\overline{k}}^i \leq J_k^i. \tag{5.15}$$

Adaptive SAGE Interference Suppression

Interference Suppression Problem

This section describes an interference suppression problem when the received signal and some a priori knowledge on the desired signal are available. The a priori knowledge can be obtained from the properties of the desired signal (i.e., time delay and signature sequence of the signal), the distribution of the data signal and so on. Equipped with all the knowledge available, the MSE optimization problem can be given by

$$\min_{\theta} \quad E\left[\left(\mathbf{r}_i^T \widehat{\mathbf{w}} - \widehat{A}_1 \widehat{b}_1(i)\right)^2\right] \tag{5.16}$$

subject to

$$\mathbf{w}^T \mathbf{s}_1 = 1. \tag{5.17}$$

$$\mathbf{w}^T\mathbf{w} \leq \delta^2. \tag{5.18}$$

$$A_1 \in R^+ \tag{5.19}$$

$$b_1(i) \in [+1, -1] \tag{5.20}$$

where

$$\theta \triangleq \left[\mathbf{w}^T, A_1, b_1(i)\right] \tag{5.21}$$

denotes the unknown parameter vector in the optimization problem and R^+ denotes a positive real constant. We assume that the exact information on w, A_1 and $b_1(i)$ are not available but the global minimum of Eq.(5.16) can be achieved by a certain parameter setting. Since the *a priori* knowledge on the distribution of $b_1(i)$ is equi-probable in $[+1, -1]$ and $b_1(i)$ is unpredictable, a maximum a posteriori (MAP) decision rule for $b_1(i)$ is simply obtained by

$$\widehat{b}_1(i) = \text{sgn}\left(\mathbf{r}_i^T\mathbf{w}\right) \tag{5.22}$$

where sgn$(\cdot)$ is the signum function. By substituting $\widehat{b}_1(i)$ into the minimization problem, one can rewrite the problem as a joint estimation problem for two unknown parameters:

$$\theta^* = \arg\left\{\min_{\theta\in\Theta} J_i(\theta)\right\}. \tag{5.23}$$

where

$$J_i(\theta) = E\left\{\left[\mathbf{r}_i^T\widehat{\mathbf{w}} - \widehat{A}_1 \cdot \text{sgn}\left(\mathbf{r}_i^T\mathbf{w}\right)\right]^2\right\}, \tag{5.24}$$

$$\Theta \triangleq \{(\mathbf{f}, A) : \mathbf{f} \in C_w, A \in C_A\}, \tag{5.25}$$

$$C_w \triangleq \left\{\mathbf{f} : \mathbf{f}^T\mathbf{s}_1 = 1 \cap \mathbf{f}^T\mathbf{f} \leq \delta^2\right\}. \tag{5.26}$$

$$C_A \triangleq \left\{A : A \in R^+\right\}. \tag{5.27}$$

The minimization of the problem is mathematically intractable due to the coupling effects between two parameters. The ith received signal must be simultaneously imposed on the unknown parameters for the minimization of $J_i(\theta)$ so that there may exist infinitely many solution sets. However, if $\theta \in \Theta$, then the solution pair may be in a feasible

region. Accordingly, one can force the solution set θ to belong to a feasible solution set Θ by projecting the parameters onto the constraint sets at each update.

Conditional Parameter Estimation

Let us now decompose the joint optimization problem in Eq.(5.23) into two decoupled parameter estimation problems. For each parameter estimation, it is temporarily assumed that the other parameter is optimal.

Filter Estimation

Let us assume that the amplitude A_1 is known *a priori*. Then the optimization problem is given by

$$\mathbf{w}^* = \arg\left\{\min_{\mathbf{w}\in C_w} J_i(\mathbf{w})\right\}. \tag{5.28}$$

To minimize $J_i(\mathbf{w})$, the constrained LMS stochastic gradient descent algorithm is given by

$$\mathbf{w}_{i+1} = \mathbf{P}_w \cdot [\mathbf{w}_i - \mu\nabla J_i(\mathbf{w})] \tag{5.29}$$

where $\mathbf{P}_w$ is the projector based on the constraint set C_w defined by

$$\mathbf{P}_w\mathbf{w} = \mathbf{s}_1 + \begin{cases} \mathbf{w} - (\mathbf{w}^T\mathbf{s}_1)\,\mathbf{s}_1, & \left\|\mathbf{w} - (\mathbf{w}^T\mathbf{s}_1)\,\mathbf{s}_1\right\|^2 \le \theta^2 \\ \frac{\theta[\mathbf{w}-(\mathbf{w}^T\mathbf{s}_1)\mathbf{s}_1]}{\|\mathbf{w}-(\mathbf{w}^T\mathbf{s}_1)\mathbf{s}_1\|}, & \text{otherwise} \end{cases} \tag{5.30}$$

and $\theta = \sqrt{\delta^2 - \mathbf{s}_1^T\mathbf{s}_1}$. The details of the implementation of this algorithm are described in Chapter 4.

Amplitude Estimation

Let us now assume that the filter tap weights are optimal such that the filter outputs can be modeled as $\mathbf{w}^T\mathbf{r}_i = A_1 b_1(i) + n_i$, $-\infty < i < \infty$, where A_1 is the amplitude. $b_1(i)$ is the ith data bit and n_i is zero mean Gaussian noise with a variance of σ_1^2. The

maximum likelihood (ML) estimate of A_1 is given by [44]

$$\widehat{A}_{1,ML} = \lim_{i\to\infty} \frac{1}{i} \sum_{m=1}^{i} \left|\mathbf{w}^T \mathbf{r}_m\right|. \tag{5.31}$$

Note that the algorithm for $\widehat{A}_{1,ML}$ cannot be performed recursively in general, since it involves computational load and storage requirement as more data are processed. Therefore, we consider a stochastic approximation of the ML estimate, which is implemented in recursive way:

$$A_{1,i+1} = A_{1,i} + \gamma \left(\left|\mathbf{w}^T \mathbf{r}_i\right| - A_{1,i}\right); \quad A_{1,0} = 0 \tag{5.32}$$

where γ is a forgetting factor bounded by $0 \leq \gamma \ll 1$. Note that $E(A_i)$ asymptotically converges to $E\left(\left|\mathbf{w}^T \mathbf{r}_i\right|\right) = A_1^*$ and $A_{1,i}$ has the form [45]

$$A_{1,i} = \gamma \sum_{m=0}^{i} (1-\gamma)^{i-m} \left|\mathbf{w}^T \mathbf{r}_m\right|. \tag{5.33}$$

An advantage of the weighted mean estimator over the arithmetic mean estimator is that it can track time-varying amplitudes [45].

Adaptation Rule

The adaptive SAGE algorithm for the blind MAI suppression is implemented as follows:

1. Guess the initial value of $(\mathbf{w}^*, A_1^*)$ and call this $(\mathbf{w}_0, A_{1,0})$. A reasonable choice is given by $\mathbf{w}_0 \triangleq P_w \cdot \mathbf{0} = \mathbf{s}_1$ and $A_{1,0} = 0$.
2. Start $i = 1$: For each new received signal, $\mathbf{r}_i$.
3. Solve for $A_{1,i}$

   E-step : compute $J_i(A_1 \mid \mathbf{w}_{i-1}, A_{1,i-1})$

   M-step : calculate $A_{1,i} = A_{1,i-1} + \gamma\left(\left|\mathbf{w}_{i-1}^T \mathbf{r}_i\right| - A_{1,i-1}\right)$
4. Solve for $\mathbf{w}_i$

   E-step : compute $J_i(\mathbf{w} \mid \mathbf{w}_{i-1}, A_{1,i})$

   M-step : calculate $\mathbf{w}_i = \mathbf{P}_w \cdot [\mathbf{w}_{i-1} - \mu \nabla J_i(\mathbf{w}_{i-1})]$

5. Record $\mathbf{w}_i$ and $A_{1,i}$. Set $i = i + 1$ and go to step 2.

In each iteration step of this algorithm, we implement the constrained LMS algorithm to update the filter parameters and the iterative ML estimation algorithm to update the amplitude estimate, and thus it monotonically minimizes the cost function such that

$$J_i\left(A_{1,i}, \mathbf{w}_i\right) \leq J_i\left(A_{1,i}, \mathbf{w}_{i-1}\right) \leq J_i\left(A_{1,i-1}, \mathbf{w}_{i-1}\right). \tag{5.34}$$

CHAPTER 6 PROJECTION-BASED MAI SUPPRESSION

The receiving power of the signal either may be unknown in practice or is varying with time in a nonstationary environment. In the LMS-type stochastic gradient algorithm, the correction $\mu e_i \mathbf{r}_i$ applied to the filter coefficient vector $\mathbf{w}_i$ at time $(i+1)$ is directly proportional to the received signal $\mathbf{r}_i$. Therefore, when $\mathbf{r}_i$ becomes large, the LMS stochastic gradient descent algorithm experience a *gradient noise amplification problem* [34]. The CLMS algorithm in Chapter 4 and the adaptive SAGE algorithm in Chapter 5 may also experience this problem.

To overcome this problem, the dissertation implements a fully robust adaptive algorithm (or filter) that uses only the information required by the conventional matched filter receiver but outperforms the conventional receiver by further suppressing the interference in the near-far situation. In this algorithm, the unknown parameters are iteratively projected onto the constraint spaces to satisfy the constraint sets. The constraint sets fall into two categories: *data sets* and *property sets*. Data sets are obtained from the received signal and used to specify that the filter should pass the desired signal without any distortion. Property sets are obtained from the prior knowledge of the desired signal and can be used as remedial measures that refine the parameters to converge to the desired solution. In practice, the exact data sets cannot be obtained without providing the desired data bit explicitly. Furthermore, the algorithm based on the estimated data sets experiences convergence to a local minimum. However, the method of generalized projection shows that an iterative projection algorithm that has one or more nonconvex sets can be converged by incorporating a measure to control the

performance of the algorithm.

This chapter presents adaptive space-alternating generalized projection (SAGP) algorithm that generalize the adaptive SAGE algorithm presented in Chapter 5. The application of the algorithm to the blind interference suppression problem in the DS/CDMA detection is also discussed. Let us briefly describe the projection method used in the signal restoration problem in sequel.

Introduction to Projection Method

A signal restoration problem addresses the problem of recovering a desired signal from noisy data that has been corrupted by an unknown distorting function. In the projection method for signal restoration, every *a priori* known property of the unknown signal, $f \in \mathcal{H}$, where $\mathcal{H}$ is a Hilbert space, is viewed as a constraint that restricts f to lie in a well-defined set C_i. If there are m known properties, there are m constraint sets C_i, $i = 1, \cdots, m$, so that f must lie in the intersection

$$C_0 \triangleq \bigcap_{i=1}^{m} C_i. \tag{6.1}$$

We shall call any restoration that satisfies all the constraints a solution to the problem. Then the problem is to find a point in C_0. If C_0 contains only a single point, the solution is unique; otherwise, C_0 may not be sufficient to define a unique solution. That is, the solution space may have more than one point. In any case, what is needed is an algorithm that converges to a signal that satisfies all the known constraints.

Theory of Projections Onto Convex Sets

To introduce the theory of projections onto convex sets (POCS), some definitions are formalized.

Definition 3 *[46] For all closed sets, the projection $P_i f$ of f onto the set C_i is defined by*

$$\|f - P_i f\| \triangleq \min_{x \in C_i} \|f - x\| . \tag{6.2}$$

If the set is closed and convex, the projection is unique. Otherwise, there may be a set of points to satisfy the definition of projection. Projections always exist for convex sets. For the proof, refer to [46].

Definition 4 *[46] The relaxed projection is defined by*

$$T_i \triangleq 1 + \lambda_i (P_i - 1), i = 1, 2, \cdots, m \tag{6.3}$$

where 1 is the identity operator and λ_i is a relaxation parameter (RP) bounded by the range $0 < \lambda_i < 2$.

Based on these definitions, Sezan and Stark [47] described the method of POCS. The fundamental theorem of the POCS is given as follows

Theorem 1 *[47] When all the constraint sets are of convex type, the general recursive algorithm given by*

$$f_{k+1} = T_1 T_2 \cdots T_m f_k, \quad f_0 \text{ arbitrary} \tag{6.4}$$

converges weakly to a point in C_0 under quite general conditions.

Method of Generalized Projection

If one or more sets of $C_i, i = 1, \cdots, m$ are of nonconvex type, the convergence property of POCS does not generally hold. In such a case, one can incorporate a performance measure that allows one to control the performance of the algorithm during the recursion process. The following definition formalizes this idea.

Definition 5 *[46] A performance of the recursive algorithm given in Eq.(6.4) is measured by a criterion called the summed distance error (SDE), which is given by*

$$J(f_k) \triangleq \sum_{i=1}^{m} \|P_i f_k - f_k\| . \tag{6.5}$$

Theorem 2 *[46] If $m = 2$, the recursive algorithm of Eq.(6.4) which contains nonconvex sets has a property that for all $k \geq 1$,*

$$J(f_{k+1}) \leq J(T_2 f_k) \leq J(f_k) \tag{6.6}$$

under certain circumstances.

This property is called the set distance reduction (SDR) property. This method has been successfully applied in several signal restoration problems [46]. Note that $J(f_k) \geq 0$ and that the equality holds if and only if $f_k \in C_0$. Note that the SDR property given in Eq.(6.6) cannot be generalized to more than two sets [46]. Nevertheless, the above theorem is not so restrictive in practice because it does not restrict the complexity of the sets. That is, a set can contain multiple constraints.

Adaptive SAGP Algorithm

This section presents an adaptive space-alternating generalized projection (SAGP) algorithm for a joint parameter estimation problem. A generalized projection algorithm can be used for the signal restoration problem that has more than one nonconvex constraint set. A SAGE algorithm has shown to be a useful tool for the joint parameter estimation problem. The constrained interference suppression problem for the DS/CDMA detection is an example that has a nonconvex cost set and requires joint parameter estimation. Therefore, we implement an algorithm that approximates the generalized projection and the SAGE algorithms by stochastic approximation at each iteration stage. This algorithm is also a generalization of the adaptive SAGE algorithm

presented in Chapter 5. As in the adaptive SAGE algorithm, this algorithm decomposes the joint parameter estimation problem in Eq.(5.23) into two decoupled parameter estimation problems. Each parameter is updated based on the method of generalized projection. For each parameter estimation, it is temporarily assumed that the other parameter is optimal.

Interference Suppression Problem

Consider a DS/CDMA communication system model. It is assumed that the received amplitude of the desired signal is unknown due to the time-varying nature of the channel but should be estimated to be used as a reference of the constrained adaptive algorithm. It is also assumed that the detection capability of the receiver is completely lost so that the decision-directed algorithm is not feasible. Under the assumed scenario, we may consider a cost functional,

$$J_i(\mathbf{w}, A_1) = E\left\|\left|\mathbf{w}^T\mathbf{r}_i\right| - A_1\right\| + \left\|\mathbf{w}^T\mathbf{s}_1 - 1\right\|. \tag{6.7}$$

as a performance measure to be minimized during the adaptation stage. In this cost function, $\mathbf{w}$ and A_1 are both unknown *a priori*. Suppose that the minimum of Eq.(6.7) can be achieved by a certain parameter setting. Then, we have a joint optimization problem:

$$\theta^* = \arg\left\{\min_{\theta\in\Theta} J_i(\theta)\right\}. \tag{6.8}$$

where $\theta \triangleq (\mathbf{w}, A_1)$ and $\Theta \triangleq \{(\mathbf{g}, A) : \mathbf{g} \in C_w, A \in C_A\}$

Let us consider a data set imposed by the ith received signal as a solution pair: $C_{R,i} \triangleq \left\{(\mathbf{w}, A_1) : E\left|\mathbf{w}^T\mathbf{r}_i\right| = A_1\right\}$. The ith received signal must be simultaneously imposed on two unknown parameters for the minimization of $J_i(\theta)$ so that there may exist infinitely many solution sets. However, if $\theta \in \Theta$, then the solution pair may be in a feasible region. Accordingly, we consider an algorithm that utilizes as much information about θ^* as possible to obtain a meaningful solution. In other words, let us force the solution set

θ to belong to a feasible solution set Θ by iterative minimization based on the constraint sets of the unknown parameters.

Constraint Sets and Projections

Filter Constraint Sets

Consider a filter estimation problem as a signal restoration problem. The filter can be adaptively implemented by the generalized projection method. Assume that the amplitude of the desired signal is known *a priori*. Then the filter constraint sets fall into two categories: data sets and property sets. The ith *data constraint set* imposed by the ith received signal is given by

$$C_{R,i} \triangleq \left\{ \mathbf{w} : E\left|\mathbf{w}^T \mathbf{r}_i\right| = A_1 \right\}. \tag{6.9}$$

Since $\mathbf{r}_i$ is noisy, we make a stochastic approximation by using a relaxed projector. Note that the instantaneous estimate of $E\left|\mathbf{w}^T \mathbf{r}_i\right|$ is $\mathbf{w}_i^T \mathbf{r}_i$. The relaxed projection $T_{R,i}$ of a filter $\mathbf{w}_i$ onto $C_{R,i}$ is

$$T_{R,i}\mathbf{w}_i = \begin{cases} \mathbf{w}_i - \xi \frac{\mathbf{w}_i^T \mathbf{r}_i - A_1}{\mathbf{r}_i^T \mathbf{r}_i} \mathbf{r}_i, & \text{if } \mathbf{w}_i^T \mathbf{r}_i \geq 0 \\ \mathbf{w}_i - \xi \frac{\mathbf{w}_i^T \mathbf{r}_i + A_1}{\mathbf{r}_i^T \mathbf{r}_i} \mathbf{r}_i, & \text{otherwise.} \end{cases} \tag{6.10}$$

where ξ denotes a relaxation parameter (RP) bounded in the range $0 < \xi < 2$. Typically, $\xi \ll 1$ for very noisy data.

If the data constraint set is given by $\widetilde{C}_R = \left\{ \mathbf{w} : \mathbf{w}^T \mathbf{r}_i = A_1 b_1(i) \right\}$ with the exact information on $b_1(i)$, the above set $\widetilde{C}_R$ becomes convex function and an algorithm implemented by this projection is referred the normalized LMS (NLMS) algorithm [34]. However, since the constraint set C_R is a nonconvex set and $\mathbf{r}_i$ contains strong interference components, the convergence of the above algorithm is not guaranteed without incorporating other constraint sets.

The *property constraint set* imposed by the signature sequence of the desired signal is given by

$$C_S \triangleq \left\{ \mathbf{w} : \mathbf{w}^T \mathbf{s}_1 = 1 \right\}, \tag{6.11}$$

where the unity response on the signature sequence is chosen arbitrarily. It is noted that C_R is chosen to satisfy C_S if the filter sets optimum. The projection P_S of an arbitrary vector $\mathbf{w}_i$ onto C_S is given by

$$P_S \cdot \mathbf{w}_i = \mathbf{w}_i - \left(\mathbf{s}_1^T \mathbf{w}_i - 1 \right) \mathbf{s}_1 \tag{6.12}$$

where the norm of the signature sequence is assumed to be $\mathbf{s}_1^T \mathbf{s}_1 = 1$.

Amplitude Constraint Sets

Let us assume that the filter is optimum so that the filter output is considered as a single-user transmission corrupted by only AWGN, i.e.,

$$z_i = A_1 b_1(i) + n_i, \quad i = 1, \cdots, \infty \tag{6.13}$$

where A_1 and $b_1(i)$ are the amplitude and the ith data bit of the user 1, respectively, and n_i is zero mean Gaussian noise with a variance of σ_1^2. The maximum likelihood (ML) estimate of A_1 is given by [44]

$$A_{1\,ML} = \lim_{i \to \infty} \frac{1}{i} \sum_{m=1}^{i} |z(m)| . \tag{6.14}$$

Note that the algorithm for $A_{1,ML}$ cannot be performed recursively in general, since it involves increasing computational load and storage requirements as more data are processed. Therefore, we consider a stochastic approximation of the ML estimate, which is implemented in recursive way:

$$A_{1,i+1} = A_{1,i} + \gamma \left(|z(i)| - A_{1,i} \right); \quad A_{1,0} = 0 \tag{6.15}$$

where γ is a forgetting factor bounded by $0 \leq \gamma < 1$. Note that $E(A_{1,i})$ asymptotically converges to $E(|z(i)|) = A_1^*$ and $A_{1,i}$ has the form of a weighted absolute mean estimator

$$A_{1,i} = \gamma \sum_{m=1}^{i} (1-\gamma)^{i-m} |z(m)| . \tag{6.16}$$

where an advantage of the weighted mean estimator over the arithmetic mean estimator is that it can track time-varying amplitudes [45].

We may consider the recursive estimation algorithm in Eq.(6.15) as a successive projection algorithm based on the data constraint sets. Let us define the *data constraint set* imposed by the ith filter output as $C_{A,R} = \{A : A = E(|z(i)|)\}$. Then the instantaneous approximation of A_1 is just z_i if $z_i \geq 0$, or $-z_i$ if $z_i < 0$. The stochastic approximate projection of any arbitrary $A_{i,i}$ onto $C_{A,R}$ is given by

$$T_{AR}A_{1,i} = \begin{cases} A_{1,i} - \gamma(A_{1,i} - z_i), & \text{if } z_i \geq 0 \\ A_{1,i} - \gamma(A_{1,i} + z_i), & \text{otherwise} \end{cases} \tag{6.17}$$

where the forgetting factor γ is considered a relaxation parameter (RP) bounded by $0 < \gamma \ll 1$.

Let us assume that some reasonable experimental data such as a support region for the amplitude is available. One can incorporate it to disregard an estimation outlier in the initial estimation stage. The *property constraint set* imposed by the support region is defined as $C_{AS} = \{A : A \in [\mu_L, \mu_H]\}$. The projection of any arbitrary $A_{1,i}$ onto C_{AS} is given by

$$P_{AS}A_{1,i} = \begin{cases} \mu_L, & \text{if } A_{1,i} < \mu_L \\ A_{1,i}, & \text{if } \mu_L \leq A_{1,i} \leq \mu_H \\ \mu_H, & \text{otherwise.} \end{cases} \tag{6.18}$$

Note that the iterative ML algorithm converges to near-optimum without the aid of the property constraint set but incorporating the property set can facilitate the convergence rate of the algorithm.

Conditional Parameter Update Algorithms

Let us now decompose the joint parameter estimation problem in Eq.(6.8) into two decoupled parameter estimation problems. For each parameter estimation, it is temporarily assumed that the other parameter set is optimal.

Filter Update Algorithm

Let us assume that the amplitude A_1 is known *a priori*. Then the optimization problem in Eq.(6.7) is reduced as

$$\mathbf{w}^* = \arg\left\{\min_{\mathbf{w}\in C_w} J_i(\mathbf{w})\right\}. \tag{6.19}$$

where

$$J_i(\mathbf{w}) = E\left\|\left|\mathbf{w}^T\mathbf{r}_i\right| - A_1\right\| + \left\|\mathbf{w}^T\mathbf{s}_1 - 1\right\|. \tag{6.20}$$

To minimize $J_i(\mathbf{w})$, the adaptive generalized projection algorithm can be used, which is given by

$$\mathbf{w}_{i+1} = P_S T_{R,i}\mathbf{w}_i, \tag{6.21}$$

where $\mathbf{w}_0 = P_S \cdot \underline{0}$, P_S and $T_{R,i}$ are projection operators, defined in Eq.(6.10) and Eq.(6.12), respectively.

The iterative projection algorithm in Eq.(6.21) satisfies the following property:

$$J(\mathbf{w}_{i+1}) \leq J(T_R\mathbf{w}_i) \leq J(\mathbf{w}_i) \tag{6.22}$$

so that the successive projections statistically converge to $\mathbf{w}^*$ in the mean for every choice of the starting point $\mathbf{w}_0$; i.e.,

$$\lim_{i\to\infty} E(P_S T_R\mathbf{w}_i) = \mathbf{w}^*. \tag{6.23}$$

This property is a statistical approximation of that of the generalized projection and has been also introduced in the *contraction mapping theorem* of functional analysis [48].

Amplitude Update Algorithm

Similarly, let us assume that the filter is optimal and its output is Gaussian distributed in the mean of $A_1 b_1(i)$ and the variance of σ_1^2. Then the ML estimation problem of the amplitude can be written as

$$A_1^* = \arg\left\{\min_{A \in C_A} J_i(A_1)\right\} \tag{6.24}$$

where

$$J_i(A_1) = E\left\|\mathbf{r}_i^T \mathbf{w}_{opt} - A_1 b_1(i)\right\|^2. \tag{6.25}$$

To minimize $J_i(A_1)$, an iterative amplitude estimation algorithm is given by

$$A_{1,i+1} = P_{AS} T_{AR} A_{1,i} \tag{6.26}$$

and it satisfies that

$$E\left\|A_1^* - A_{1,i+1}\right\| \leq E\left\|A_1^* - T_{AR} A_{1,i}\right\| \leq E\left\|A_1^* - A_{1,i}\right\|. \tag{6.27}$$

or

$$J(A_{1,i+1}) \leq J(T_{AR} A_{1,i}) \leq J(A_{1,i}). \tag{6.28}$$

and asymptotically, $E(A_{1,i}) \rightarrow A_1^*$ as $i \rightarrow \infty$.

Adaptation Rule

The adaptive SAGP algorithm is to alternate between using each new parameter estimate to improve the conditional performance measure on the next iteration cycle (E-step) and then using this conditional performance measure to improve the next parameter estimate (M-step). For the M-step of the algorithm, an adaptive generalized projection algorithm is implemented to update the parameters. The proposed algorithm is implemented as follows:

Table 6.1 Summary of projection-based algorithms

Algorithm	Adaptation rule	Assumption
NLMS	$\mathbf{w}_{i+1} = \mathbf{w}_i - \mu \left(\mathbf{w}_i^T \mathbf{r}_i - d_i\right) \frac{\mathbf{r}_i}{\mathbf{r}_i^T \mathbf{r}_i}$	d_i is known
NGP	$\mathbf{x}_{i+1} = \mathbf{x}_i - \mu \left(\mathbf{r}_i - \left[\mathbf{s}_1^T \mathbf{r}_i\right] \mathbf{s}_1\right) \frac{\mathbf{w}_i^T \mathbf{r}_i}{\mathbf{r}_i^T \mathbf{r}_i}$ $\mathbf{w}_{i+1} = \mathbf{s}_1 + \mathbf{x}_{i+1}$	$\mathbf{s}_1$ is known
SAGP	$\mathbf{w}_{i,1} = T_R \mathbf{w}_i = \mathbf{w}_i - \mu \left(\left\|\mathbf{w}_i^T \mathbf{r}_i\right\| - A_i\right) \frac{\mathbf{r}_i}{\mathbf{r}_i^T \mathbf{r}_i}$ $\mathbf{w}_{i+1} = P_S \mathbf{w}_{i,1} = \left[1 - \mathbf{s}_1^T \mathbf{w}_{i,1}\right] \mathbf{s}_1 + \mathbf{w}_{i,1}$ $A_{i+1} = A_i - \gamma \left[\left\|\mathbf{w}_i^T \mathbf{r}_i\right\| - A_i\right]$	$\mathbf{s}_1$ is known

1. Guess the initial value of $(\mathbf{w}^*, A_1^*)$ and call this $(\mathbf{w}_0, A_{1,0})$. Use $\mathbf{w}_0 \triangleq P_S \cdot \mathbf{0} = \mathbf{s}_1$ and $A_{1,0} \triangleq P_A \cdot 0 = \mu_L$.

2. Start, $i = 1$: For each new received signal, $\mathbf{r}_i$.

3. Solve for $A_{1,i}$

$$\text{E step :} \quad \text{compute } J_i\left(A_1 \mid \mathbf{w}_{i-1}, A_{1,i-1}\right)$$

$$\text{M step :} \quad \text{calculate } A_{1,i} = P_{AS} T_{AR} A_{1,i-1}$$

4. Solve for $\mathbf{w}_i$

$$\text{E step :} \quad \text{compute } J_i\left(\mathbf{w} \mid \mathbf{w}_{i-1}, A_{1,i}\right)$$

$$\text{M step :} \quad \text{calculate } \mathbf{w}_i = P_S T_R \mathbf{w}_{i-1}$$

5. Record $\mathbf{w}_i$ and $A_{1,i}$. Set $i = i + 1$.

Algorithm Analysis

Computational Complexity

Computational complexity of the SAGP algorithm is compared with that of the normalized LMS (NLMS), the normalized OPM-based gradient projection (NGP) algorithm. The summary of all the algorithms is given in Table 6.1.

It is assumed that all algorithms have the filter length equal to N. For comparison to be fair, we consider the normalized versions of the LMS and GP algorithms. Table 6.2 shows the computational complexity of the algorithms. It is shown that the added

Table 6.2 Computational complexity of projection-based algorithms

	NLMS	NGP	SAGP
Multiplication / Division	$3N+3$	$5N+1$	$5N+3$
Addition / Subtraction	$3N+1$	$6N$	$5N+4$
$\lvert\cdot\rvert$	0	0	1
Storage requirement	$N+1$	$2N+1$	$2N+3$

complexity in the implementation of the SAGP algorithm is negligible. It is noted that the clipping operation, P_{AS}, in the algorithm is not considered in the complexity comparison since it does not significantly contribute to the complexity of the algorithm.

Convergence

The convergence of the SAGP algorithm is achieved by a fundamental theorem, which is given as follows:

Theorem 3 *The iteration algorithm defined by the generalized projection algorithm will statistically converge to achieve the minimum of the mean cost function* $J_i(\mathbf{w}, A)$ *via the following contraction mapping :*

$$J_i(\mathbf{w}_i, A_i) \leq J_i(\mathbf{w}_{i-1}, A_i) \leq J_i(\mathbf{w}_{i-1}, A_{i-1}).$$

Proof $J_i(\mathbf{w}_i, A_i) \geq 0$ for all $(\mathbf{w}_i, A_i) \in \mathcal{H}$ and the projection based on the data sets is the contraction mapping in the direction of nonincreasing J_i. The contraction mapping with the projection based on the property sets is nonexpansive and will converge to a global minimum. The following remarks are in order.

1. If the channel is noiseless and the number of interferences is less than the dimension of the filter, then $J_i(\mathbf{w}_i, A_i)$, as $i \rightarrow \infty$. converges to zero. The resulting $\mathbf{w}_i$ will be the near-far solution and A_i will be the amplitude of the desired signal.

2. If the MAI is relatively weak and the background noise is dominant, $\mathbf{w}_i$ will converge to a matched filter solution $\mathbf{s}_1$.

3. The algorithm preserves a self-recovering error correction feature. That is, for every new input signal the parameters are projected onto the signature constraint spaces of the desired signal. In [16], an output power cost function is minimized using the gradient projection method with a signature constraint. It has been shown that the gradient projection method is susceptible to accumulation of round-off errors.

CHAPTER 7 PERFORMANCE RESULTS

Simulation Overview

Simulation Procedure

This chapter presents the results obtained by an extensive series of Monte-Carlo simulations to provide a basis for comparing the various algorithms discussed in the dissertation. The simulation procedure is given as follows. The multi-user DS/CDMA signals are generated by employing 31-length Gold sequences. Total number of 33 Gold sequences is generated by a proper decimation of two m-sequences. The signature sequence of each user at each realization is randomly chosen from those 33 Gold sequences in non-overlapping manner. Data bits for each user are realized at random from $[+1, -1]$ with probability of 0.5. The simulation is conducted with BPSK-modulated baseband signals. Since BPSK is equivalent to the modulo-2 addition of the signature code bit and the data bit, each transmitted signal is obtained by multiplying a Gold sequence with the data bit and scaling by a given SIR.

Transmission is assumed to be distortionless so that the received signal is formed by asynchronous reception of all the individual signals in AWGN where the time-delay and receiving power of each signal is randomly chosen at each realization. However, the signal-to-interference ratio (SIR) for each interfering signals and the signal-to-background noise ratio (SNR) remained constant during one simulation setup. Without loss of any generality, it is assumed that the received signals from multiple users are

chip-synchronous but data-asynchronous so that a chip sample of the received signal is formed by linear addition of chip samples from multiple users and a sample of AWGN. The chip sample of each signal is appropriately shifted by time-delay and scaled by the input SIR of signals. The receiver is assumed to be perfectly synchronous to the desired signal and equipped with a 31-tap adaptive linear tapped-delay-line filter followed by a decision device. Filter taps of the receiver are adjusted by the algorithms discussed in the Chapter 4, 5 and 6. The decision on the data bit is given by a simple sign operator.

Definition of Simulation Parameters

For conceptual clarity, some definitions of simulation parameters are formalized. The input SIR for the kth interfering signal is given by

$$SIR_k = 20\log_{10}\left(\frac{A_1}{A_k}\right) \text{ dB}, \quad \text{for } k \neq 1 \tag{7.1}$$

where A_1 is the amplitude of the desired signal and A_k, $k \neq 1$ is the amplitude of interfering signal. It is noted that the definition of the SIR in this simulation is different from that in the general sense. The SIR generally denotes the signal to total interference power ratio whereas the SIR in this scenario denotes the signal to each interference power ratio. The input SNR is given by

$$SNR = 10\log_{10}\left(\frac{A_1^2}{\sigma^2}\right) \text{ dB} \tag{7.2}$$

where σ^2 is the two-sided power spectral density of AWGN in the bit level. Performance of algorithms is investigated under varying the input SIR for each interfering signals. the input SNR, the number of interfering users and the signal to a filter mismatch (or implementation error) ratio.

Simulation parameters are chosen as follows: The step size is chosen by

$$\mu = \frac{0.2}{(L+1)\,P_I} \tag{7.3}$$

where L is the total number of interferences, $P_I = A_k^2$, $k \neq 1$ is the power of each interfering signal while assuming that each interference has equal power. The step size is used in the constrained LMS (CLMS) algorithm, the output-power minimizing gradient projection (OPM-GP) algorithm, the LMS algorithm and the adaptive SAGE (SAGE) algorithm. The relaxation parameter is chosen by $\xi = 0.2$ and used in the SAGP algorithm, the normalized versions of the OPM-GP and LMS algorithms. The forgetting factor is chosen by $\gamma = 0.005$ and used in the SAGE algorithm and the SAGP algorithm.

The ensemble-averaged output signal-to-interference ratio (OSIR) curves in the transient mode and the bit-error-rate (BER) curves in the steady-state are simulated to exhibit the performance of algorithms. For the OSIR curves, the OSIR at the ith iteration is obtained by

$$OSIR_i = \frac{\sum_{m=1}^{M} \left(\mathbf{w}_{i,m}^T \mathbf{s}_1\right)^2}{\sum_{m=1}^{M} \left[\left(\mathbf{w}_{i,m}^T \mathbf{y}_{i,m}\right)/A_{1,m}\right]^2} \tag{7.4}$$

where $\mathbf{w}_{i,m}$ and $\mathbf{y}_{i,m} = \mathbf{r}_{i,m} - A_{1,m} b_{1,m} \mathbf{s}_1$ are the filter taps and the interference plus noise component of the received signal at the ith iteration in the mth realization, respectively. For BER performance simulations, a total number of 2000 bits are transmitted for each realization and the last 1000 bits of 100 realizations are used to capture the steady state performance.

Simulation Scenarios

Performances of the proposed algorithms are investigated in three different scenarios. First scenario is the data demodulation in a static communication environment where the locations of the transmitter and the receiver are stationary so that the receiver knows the exact or a good estimate of signal power (or amplitude) of the desired signal *a priori*. For this scenario, the receiver equipped with the CLMS algorithm (or the CLMS receiver) is simulated by varying the input SNR, the input SIR and the number of

interferences and compared with the matched filter receiver (or the MF receiver) and the receiver having the OPM-GP algorithm (or the OPM-GP receiver). The CLMS receiver is further examined by assuming that the adaptive filter experiences the implementation error during the iteration. For such a scenario, performance of the CLMS receivers with a robust constraint and without a robust constraint are compared with the OPM-GP receiver. Second scenario is the data demodulation in a nonstationary communication environment where either the transmitter or the receiver, or both are dynamically moving so that the receiver does not have any knowledge on the received signal except for the signature sequence and time delay of the desired signal and the signal power of the received signal. For such a scenario, performance of the receiver equipped with the adaptive SAGE algorithm (or the SAGE receiver) is investigated by varying the input SNR, the input SIR and the number of interferences and compared with those of the MF, the LMS and the OPM-GP receivers. In the third scenario, a dynamic communication channel is assumed such that the variations in the signal power of the received signal is abrupt so that any stochastic gradient algorithm with a fixed step size experiences the gradient noise amplification. For such a scenario, a projection-based adaptive algorithm called the adaptive SAGP algorithm are simulated. Performance of the receiver equipped with the adaptive SAGP algorithm (or the SAGP receiver) is compared with those of the MF, the normalized versions of LMS and OPM-GP receivers.

CLMS Receiver Performance

SIR Improvement

In this section, the SIR improvement of the CLMS receiver is investigated. Figure 7.1 shows the OSIR curve of the CLMS receiver and compared with those of the LMS receiver in the training mode and the OPM-GP receiver. The OSIR curves of these receivers are obtained by using same data and averaging over 100 realizations. In this

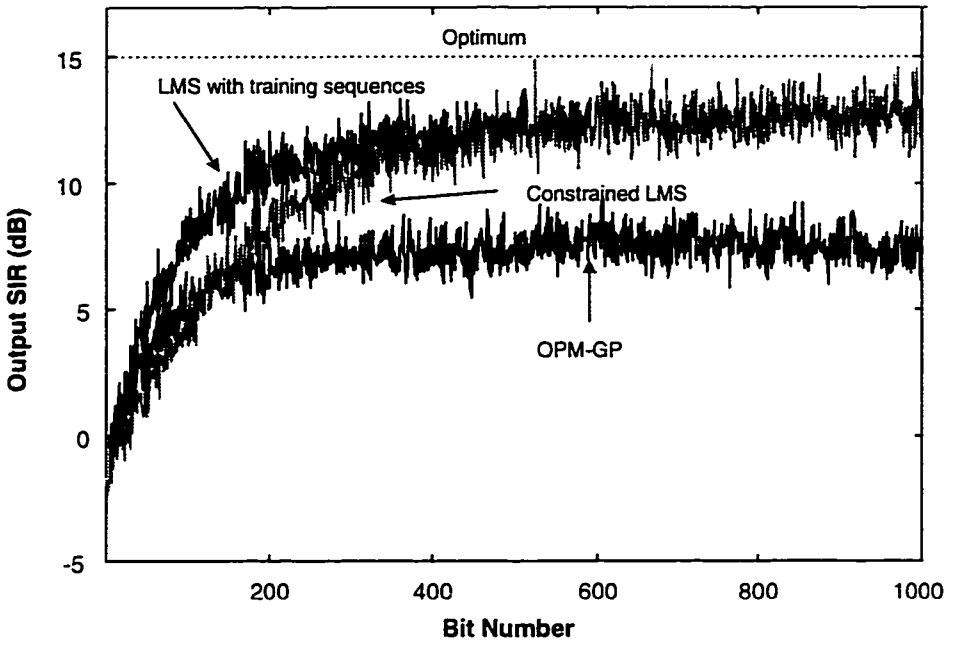

Figure 7.1 Plot of OSIR curves for the CLMS, OPM-GP and LMS receivers with 5 interferences and -10 dB SIR

simulation, the number of interference is given by 5 and the input SNR is given by 15 dB. For each realization, a different channel is assumed, i.e., A_k, τ_k and s_k, $k = 1, \cdots, 6$ are chosen at random but for each realization

$$A_k = \sqrt{10}A_1, \quad k = 2, \cdots, 6$$

remains constant. At the beginning of each realization, the initial tap weights are set to be the matched filter weights for fair comparison. The result shows that the CLMS receiver approaches to near-optimum steady state without using training sequences. The convergence rate of the CLMS receiver is approximately same as the LMS receiver. It can be seen that the OPM-GP receiver suffers from the large filter variance in the steady state and the LMS receiver requires training sequences to achieve the near-optimum steady-state convergence.

The effects of the time-varying interference power on the CLMS receiver is also investigated. In Figure 7.2, it is simulated that relatively strong interferences randomly access

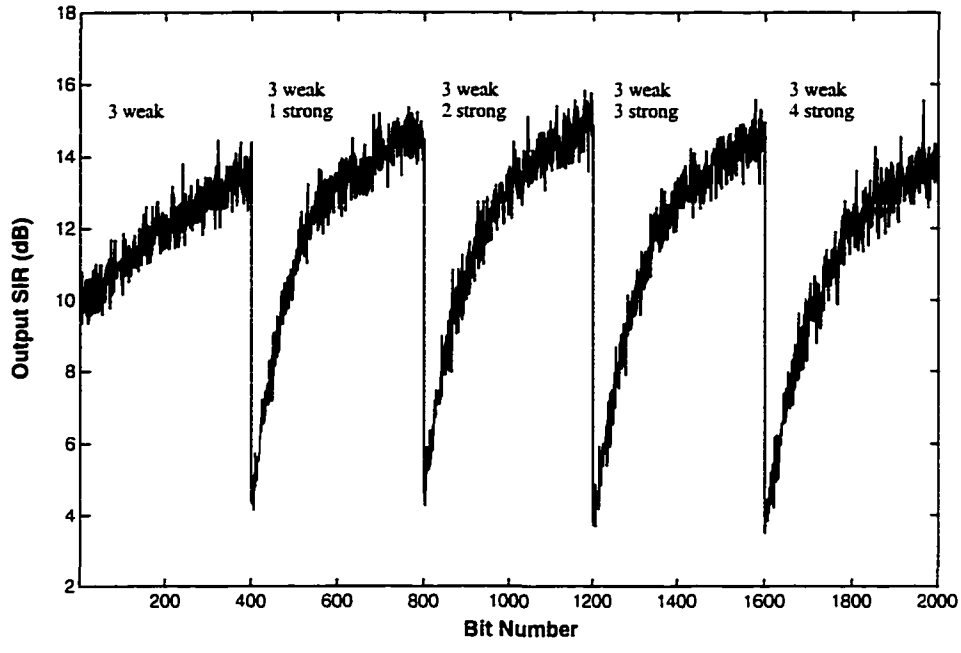

Figure 7.2 Plot of OSIR curves for the CLMS receiver with increasing number of interferences and -10 dB SIR

the system in the middle of the transmission of the user 1 and the receiver keeps adjusting the filter weights based on the received signal without using a training sequence. Received signal amplitudes of signals at each realization are given by

$$A_1 = A_2 = A_3 = A_4, \quad \text{for } i = 1, \cdots, 2000,$$

$$A_k = \begin{cases} 0, & \text{for } i = 1, \cdots, L_k \\ \sqrt{10}A_1, & \text{for } i = L_k + 1, \cdots, 2000 \end{cases}, \quad k = 5, \cdots, 8$$

where $L_k = 400\,(k-4)$, $k = 5, \cdots, 8$. The result shows that the performance of the CLMS receiver is relatively robust to the effect of the time-varying interference power and recovers near-optimum state of the filter weights while detecting the data bits.

BER vs. SNR

The graphs in Figure 7.3 to Figure 7.6 show plots of BER vs. input SNR for MAI consisting of 5 interferences with a particular SIR. The BER performance of the CLMS

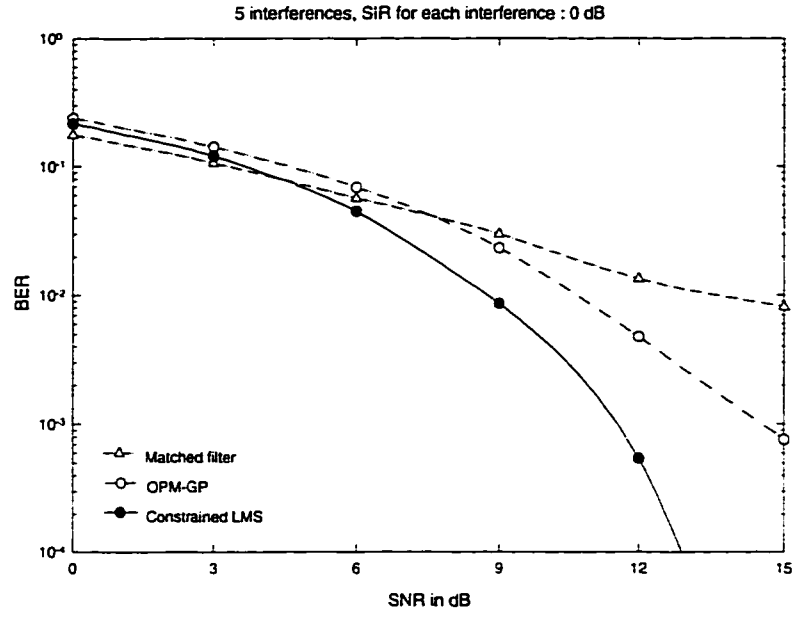

Figure 7.3 Plot of BER vs SNR for the CLMS, OPM-GP and MF receivers with 5 interferences and 0 dB SIR

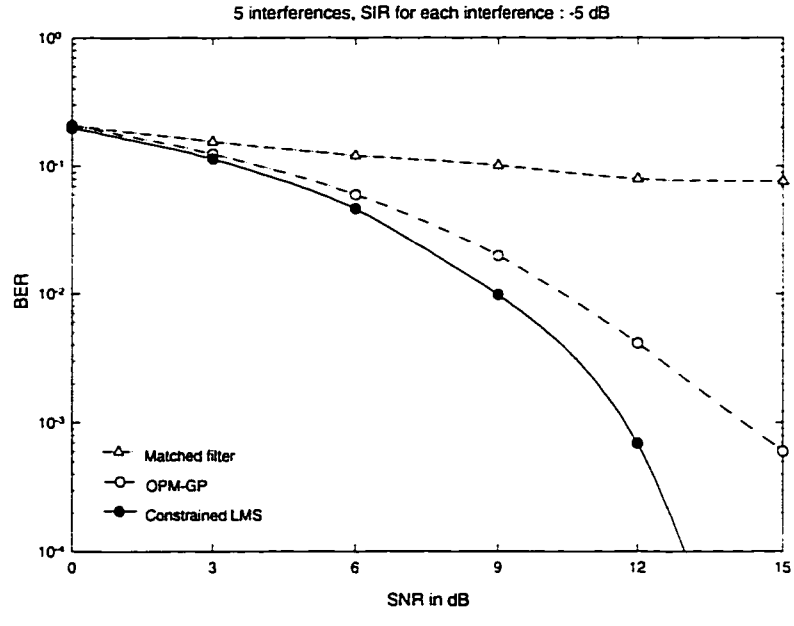

Figure 7.4 Plot of BER vs SNR for the CLMS, OPM-GP and MF receivers with 5 interferences and -5 dB SIR

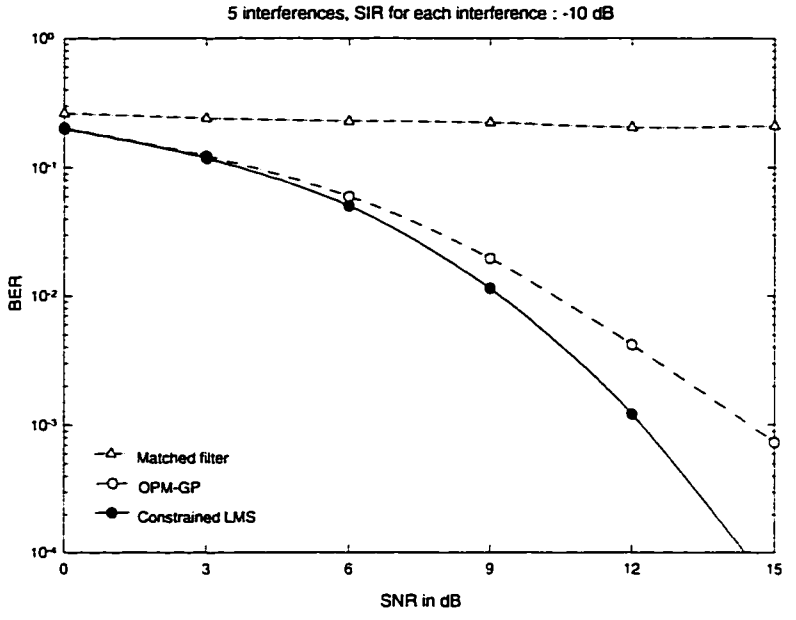

Figure 7.5 Plot of BER vs SNR for the CLMS, OPM-GP and MF receivers with 5 interferences and -10 dB SIR

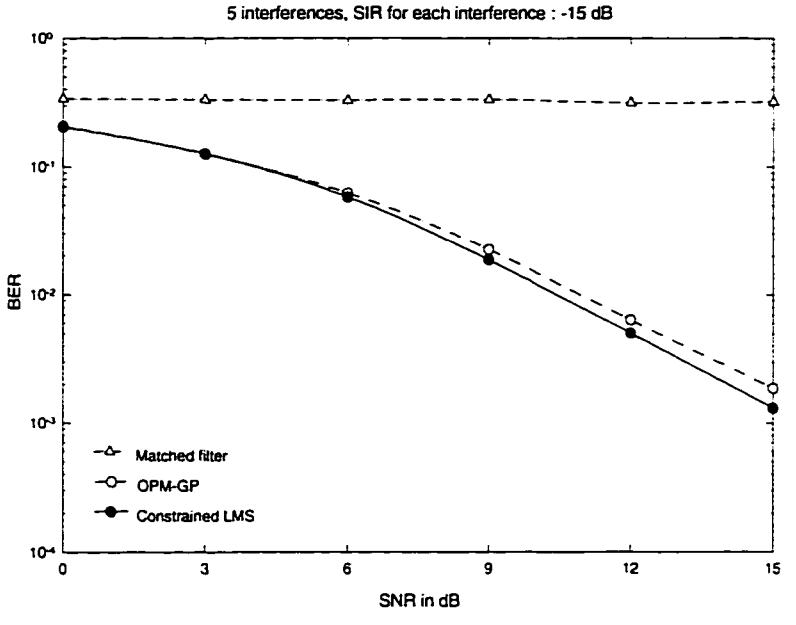

Figure 7.6 Plot of BER vs SNR for the CLMS, OPM-GP and MF receivers with 5 interferences and -15 dB SIR

receiver is compared with those of the MF and OPM-GP receivers. It can be seen that the CLMS receiver outperforms the MF and OPM-GP receivers. In particular, the performance gain of the CLMS receiver over the OPM-GP receiver is substantially improved when the input SNR increases. It is also observed that the performance of the CLMS receiver is similar to that of the OPM-GP receiver when the SNR is lower than 6 dB. This can be explained by the fact that the constrained filters in both the CLMS and OPM-GP receivers are sensitive to the random variation effect of the filter and signal characteristics in low SNR region. This phenomenon gives a motivation for incorporating a robust constraint to the CLMS algorithm.

BER vs. SIR

In Figure 7.7 to Figure 7.10, plots of BER vs. input SIR are shown by varying the input SNR. In this simulation, the MAI consists of 5 interferences. The BER performance of the CLMS receiver is compared with those of the MF and OPM-GP receivers. For a fixed SNR, it can be seen that the performance of the CLMS receiver is always better than that of the OPM-GP receiver. It can also be seen that the performances of both the CLMS receiver and the OPM-GP receiver remain approximately constant regardless of the input SIR when the input SNR is 12 dB or lower. This explains that both receivers are near-far resistant. When the input SIR is higher than -10 dB and the input SNR is 15 dB, the CLMS receiver performs error-free detection in the near-far situation whereas the MF receiver suffers from the near-far effect and the OPM-GP receiver allows errors in the detection as shown in Figure 7.10.

BER vs. Number of Interference

Figure 7.11 to Figure 7.14 shows plots of BER vs. the number of interferences with increasing SIRs. In these simulations, the input SNR is given by 15 dB. Figure 7.11 and Figure 7.12 show that the CLMS receiver performs error-free detection regardless

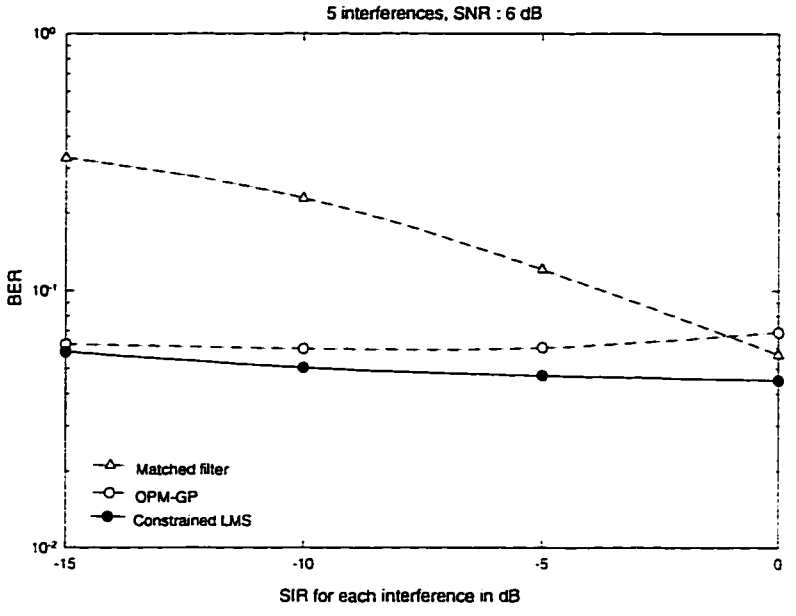

Figure 7.7 Plot of BER vs SIR for the CLMS, OPM-GP and MF receivers with 6 dB SNR

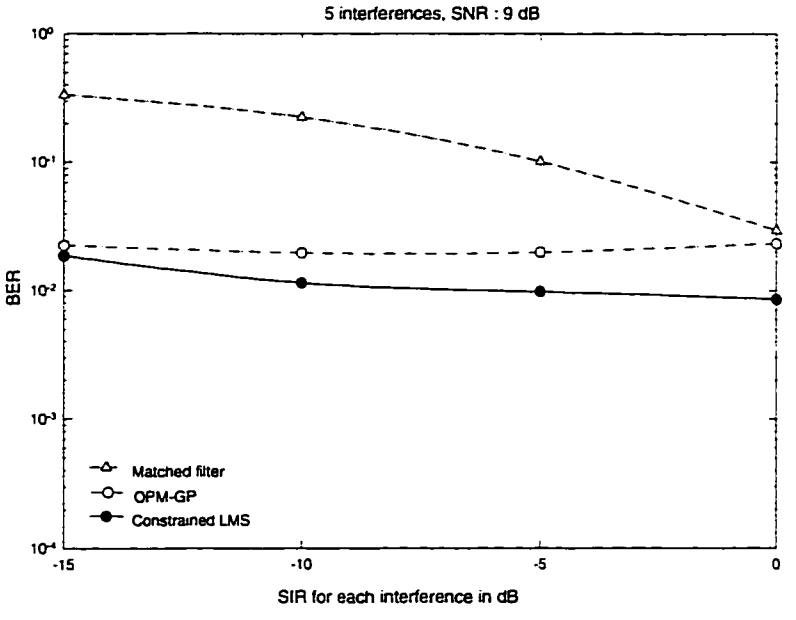

Figure 7.8 Plot of BER vs SIR for the CLMS, OPM-GP and MF receivers with 9 dB SNR

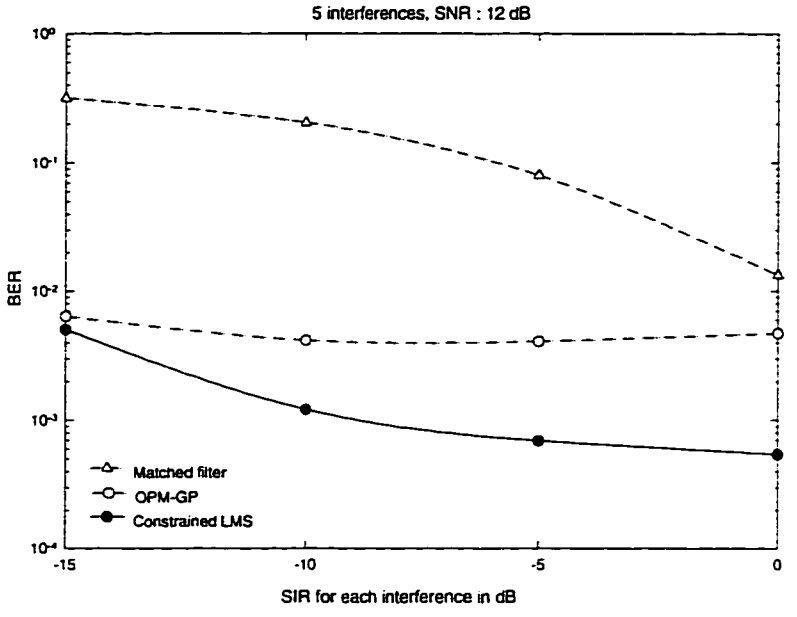

Figure 7.9 Plot of BER vs SIR for the CLMS, OPM-GP and MF receivers with 12 dB SNR

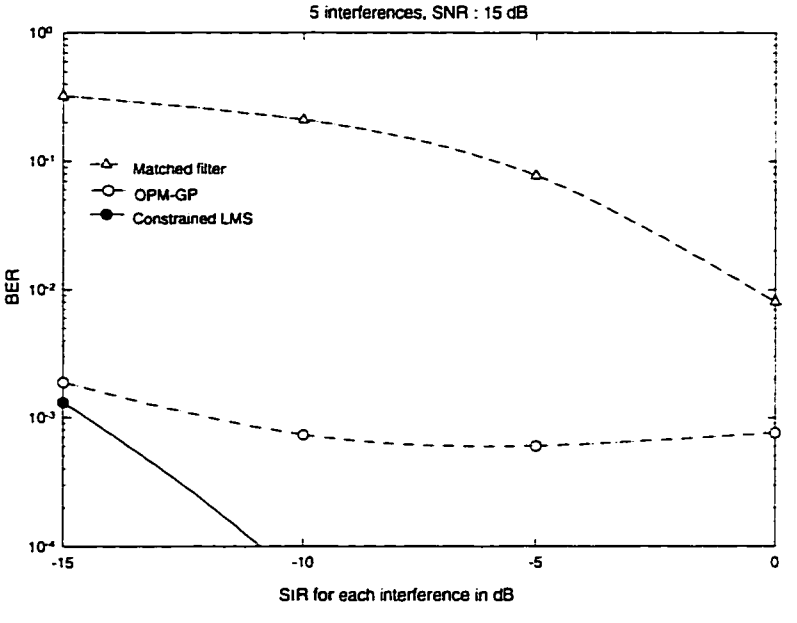

Figure 7.10 Plot of BER vs SIR for the CLMS, OPM-GP and MF receivers with 15 dB SNR

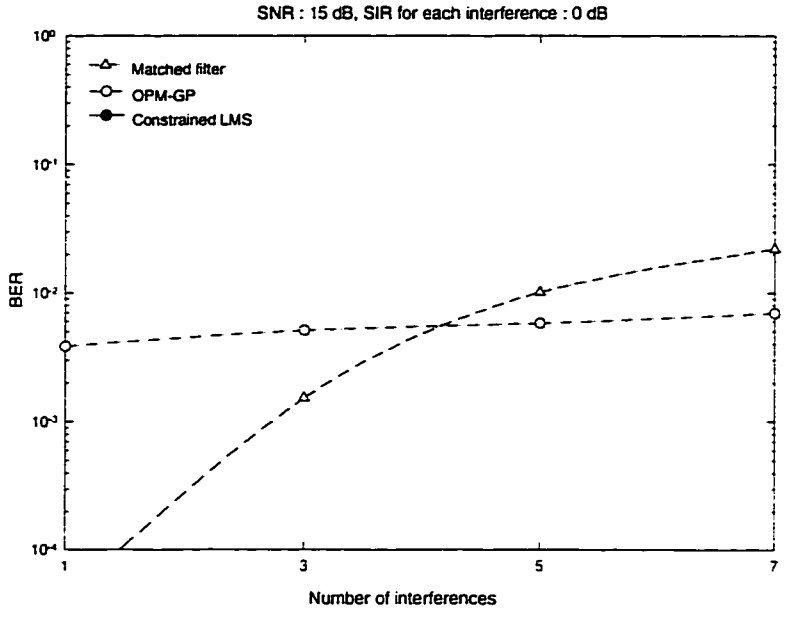

Figure 7.11 Plot of BER vs the number of interferences for the CLMS. OPM-GP and MF receivers with 0 dB SIR

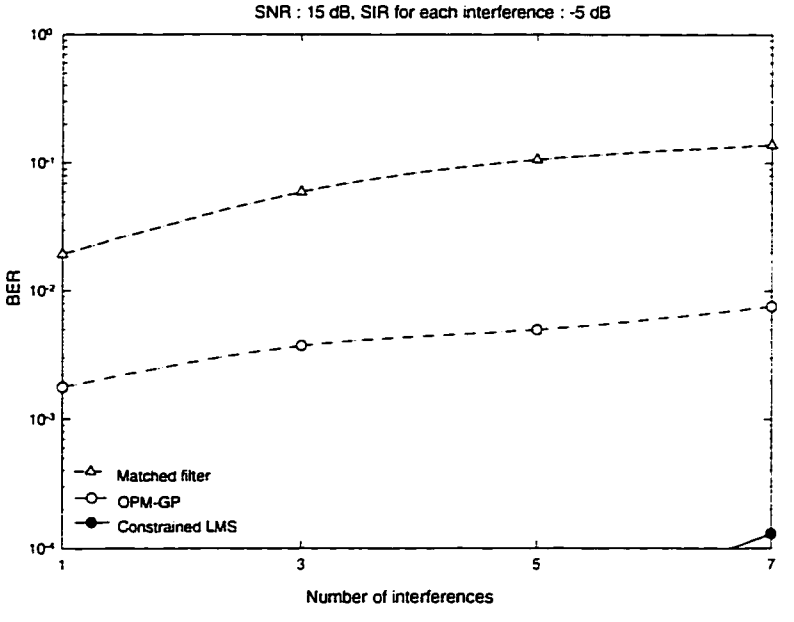

Figure 7.12 Plot of BER vs the number of interferences for CLMS. OPM-GP and MF receivers with -5 dB SIR

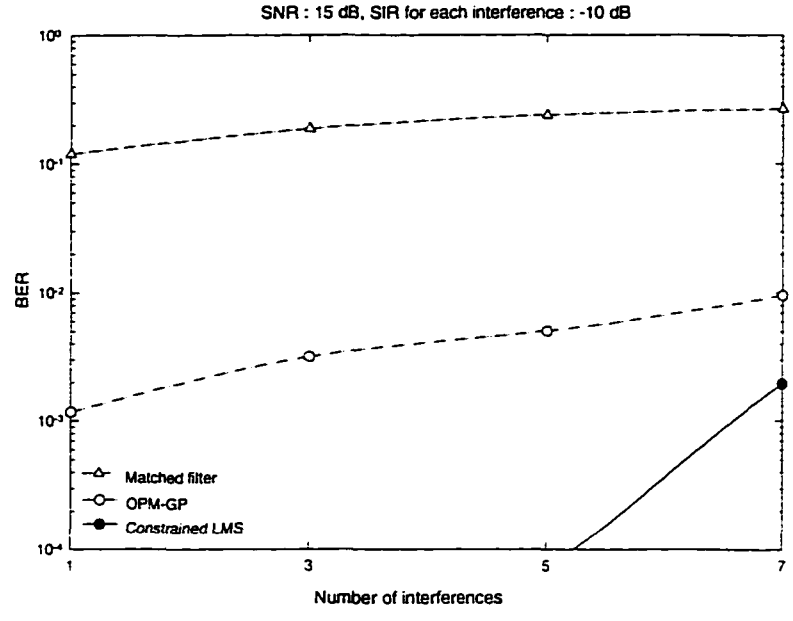

Figure 7.13 Plot of BER vs the number of interferences for the CLMS. OPM-GP and MF receivers with -10 dB SIR

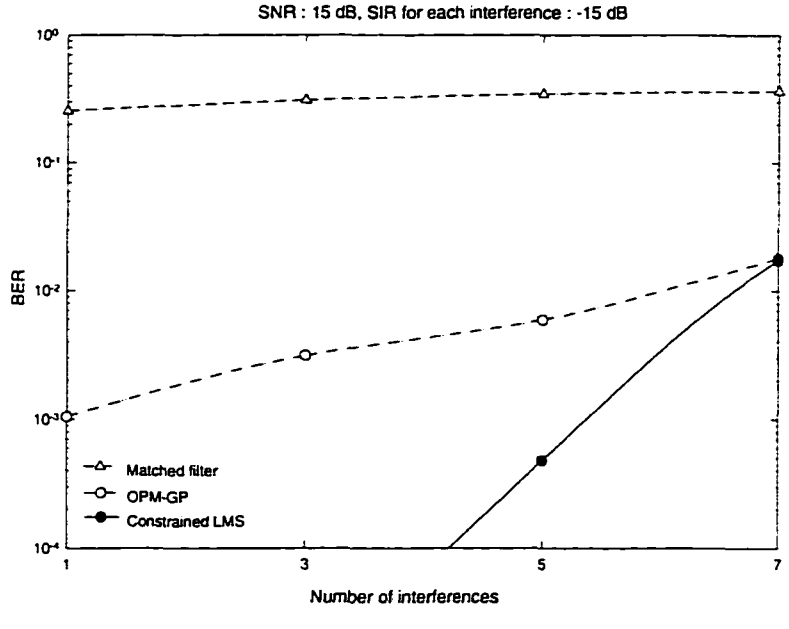

Figure 7.14 Plot of BER vs the number of interferences for the CLMS. OPM-GP and MF receivers with -15 dB SIR

of the number of interferences in the simulated range of 1 to 7 interferences when the input SIR is higher than -5 dB whereas the performance of the MF receiver degrades with increasing number of interferences. It can be seen that the OPM-GP receiver allows some errors when the CLMS receiver is error-free. Figure 7.13 and Figure 7.14 show that the performance of the CLMS receiver slowly degrades with the number of interference when the input SIR for each interfering signal is lower than -10 dB.

BER vs. Signal to Filter Mismatch Ratio

Figure 7.15 to Figure 7.18 show plots of BER vs. the signal-to-filter mismatch (or implementation error) ratio by varying the input SIR. In these simulation. the MAI consists of 5 interferences and the input SNR is given by 15 dB. The performances of the CLMS receiver with and without a robust constraint are compared with that of the OPM-GP receiver. It can be seen that the CLMS receiver is relatively robust to the filter implementation error even without employing the robust constraint but however its performance is further improved by incorporating a robust constraint. For the simulations of Figure 7.15 to Figure 7.18, the filter norm upper-bound is chosen by $\delta^2 = 1.8$ while assuming $\mathbf{s}_1^T\mathbf{s}_1 = 1$. It can be also seen that the performance gain of the CLMS receiver by incorporating the robust constraint is significantly improved when the input SIR is higher than -5 dB, whereas it is negligible for the CLMS algorithm when the SIR is relatively low as shown in Figure 7.18.

SAGE Receiver Performance

SIR Improvement

In the simulation of the SAGE receiver, it is assumed that the amplitude of the desired signal is unknown or time-varying so that the algorithm must jointly update the filter taps and the amplitude from the received signal. In Figure 7.19, the OSIR curve of

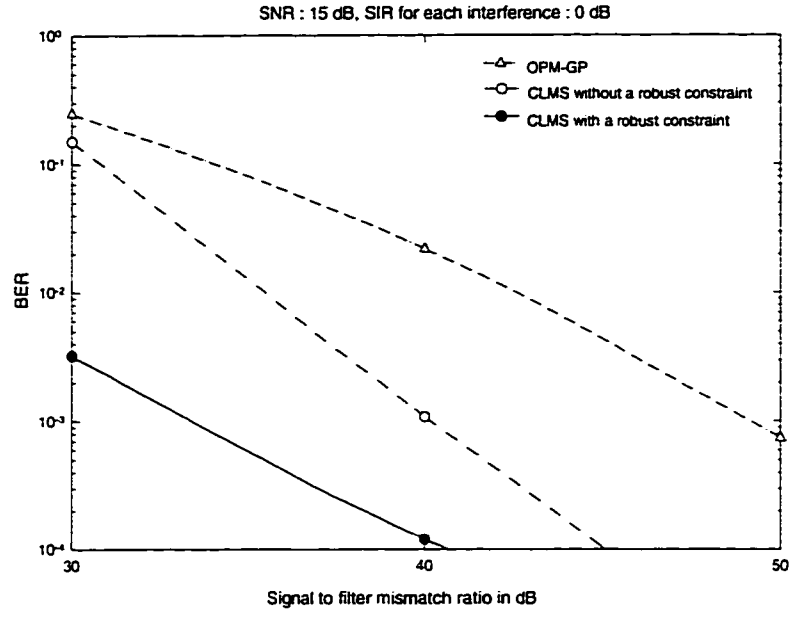

Figure 7.15 Plot of BER vs the signal-to-filter mismatch ratio for the CLMS, OPM-GP and MF receivers with 0 dB SIR

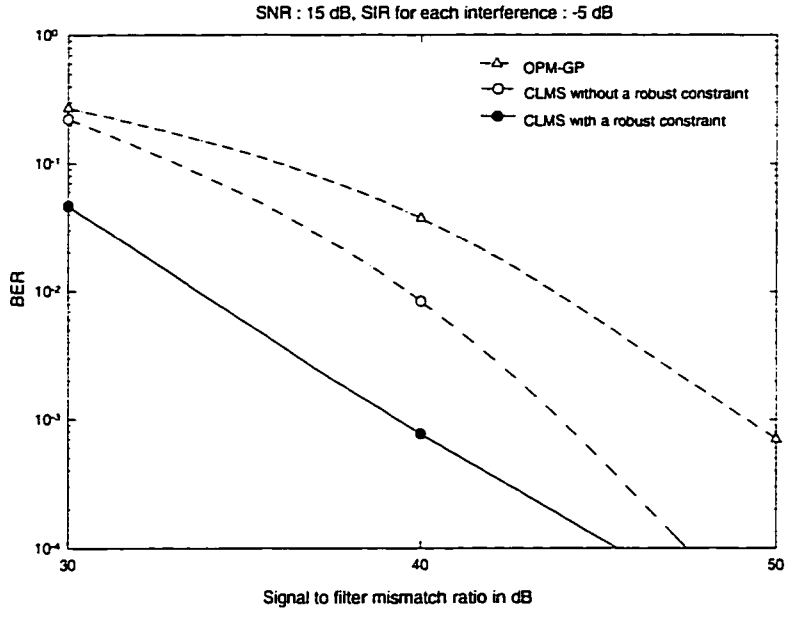

Figure 7.16 Plot of BER vs the signal-to-filter mismatch ratio for the CLMS, OPM-GP and MF receivers with -5 dB SIR

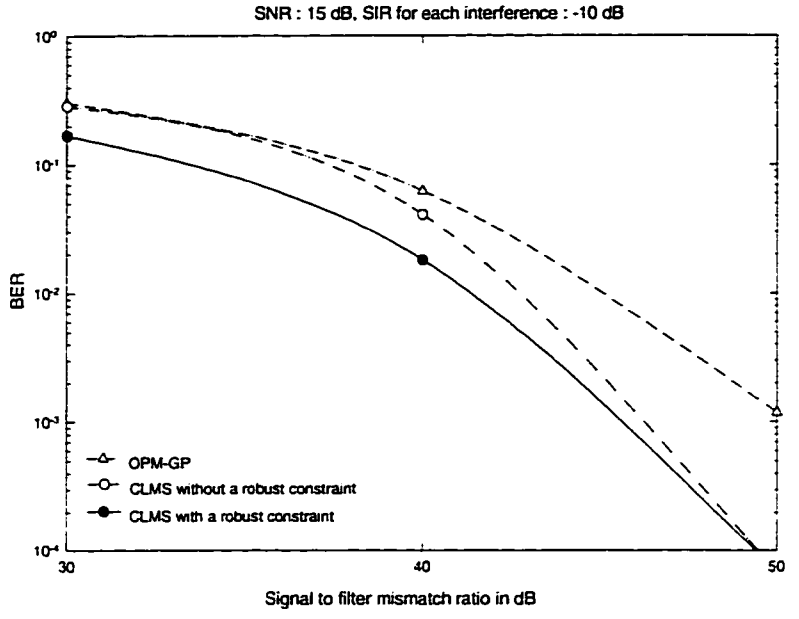

Figure 7.17 Plot of BER vs the signal-to-filter mismatch ratio for the CLMS, OPM-GP and MF receivers with -10 dB SIR

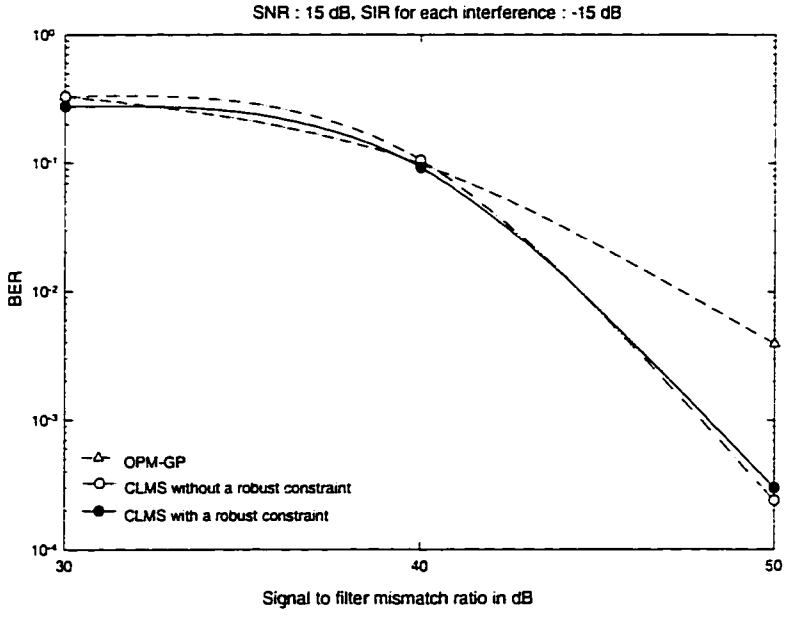

Figure 7.18 Plot of BER vs the signal-to-filter mismatch ratio for the CLMS, OPM-GP and MF receivers with -15 dB SIR

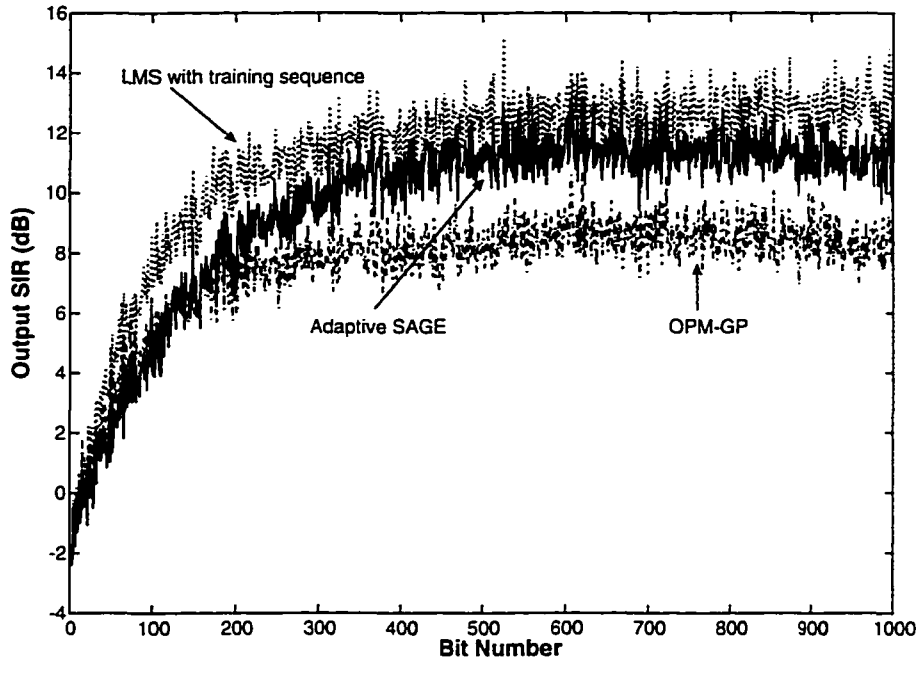

Figure 7.19 Plot of OSIR curves for the SAGE, LMS and OPM-GP receivers

the SAGE receiver is compared with those of the LMS receiver in the training mode and of the OPM-GP receiver. In this simulation, the MAI consists of 3 interferences having -10 dB SIR and the AWGN has 15 dB SNR. For each realization, a different channel is assumed, i.e., A_1, τ_k and $\mathbf{s}_k$, $k = 2, 3, 4$ are chosen at random but for each realization

$$A_2 = A_3 = A_4 = \sqrt{10}A_1$$

remains constant. At the beginning of each realization, the initial tap weights are set to the matched filter weights. The OSIR curves are obtained by averaging over 100 realizations. It can be seen that the OSIR curve of the SAGE receiver without using training sequences converges to the steady state close to that of the training-mode LMS receiver but that of the OPM-GP receiver converges poorly compared with the SAGE receiver. The small SIR loss of the SAGE receiver over the LMS receiver in the steady state is due to the estimation mismatch between the actual amplitude and the estimated amplitude. This loss can be further reduced by controlling the parameters of the algorithm at the expense of convergence rate.

BER Performance

The graphs in Figure 7.20 to Figure 7.23 present plots of BER vs. input SNR for increasing input SIRs. Typical plot of BER vs. SNR with the MAI consisting of 3 interferences is shown for the MF, the OPM-GP and the SAGE receivers. In Figure 7.20 with the 0 dB input SIR, it can be seen that the performance of three receivers are similar in low SNR region but the SAGE receiver outperforms the other receivers when the input SNR increases. Figure 7.21 to Figure 7.23 show that the SAGE receiver is robust to the near-far effect and outperforms the OPM-GP receiver regardless of the input SNR. Particularly when the SNR is high, the performance of the SAGE receiver is significantly better than that of the OPM-GP receiver.

In Figure 7.24 to 7.27, the BER performance of the SAGE receiver is compared with

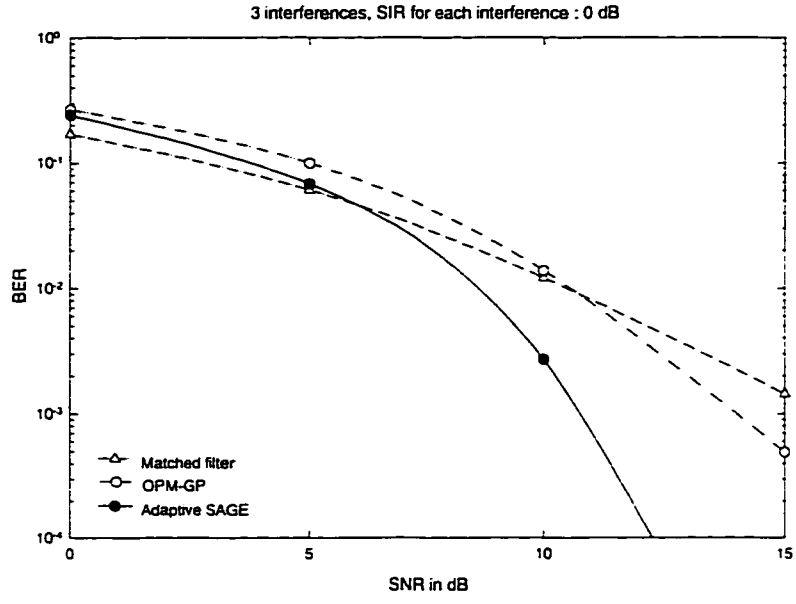

Figure 7.20 Plot of BER vs SNR for the SAGE, MF and OPM-GP receivers with 3 interferences and 0 dB SIR

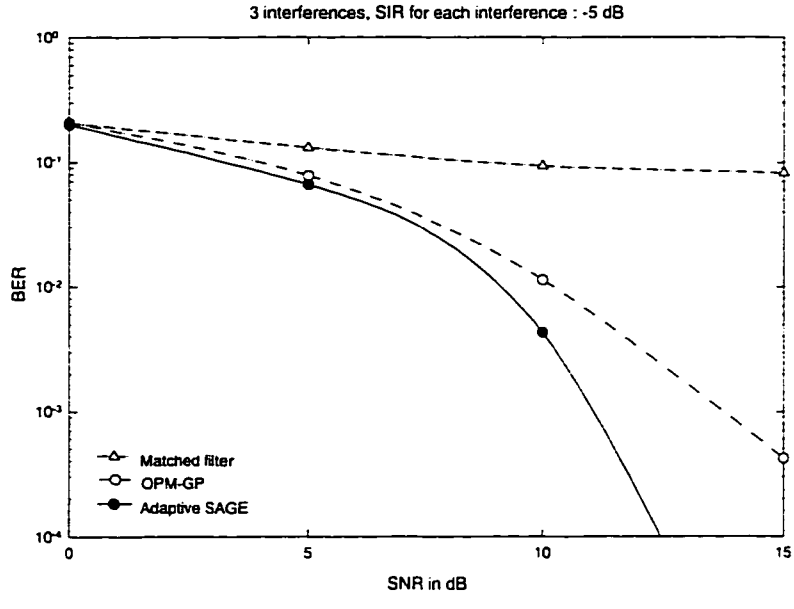

Figure 7.21 Plot of BER vs SNR for the SAGE, MF and OPM-GP receivers with 3 interferences and -5 dB SIR

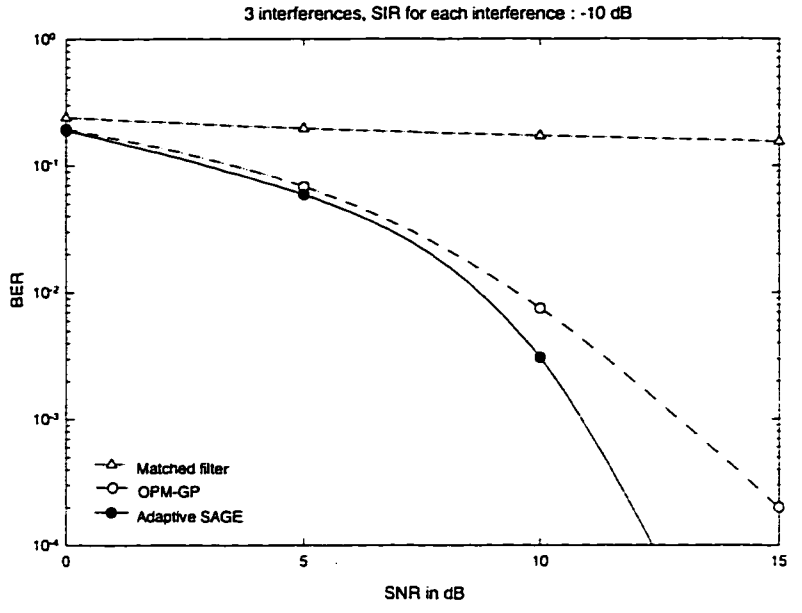

Figure 7.22 Plot of BER vs SNR for the SAGE, MF and OPM-GP receivers with 3 interferences and -10 dB SIR

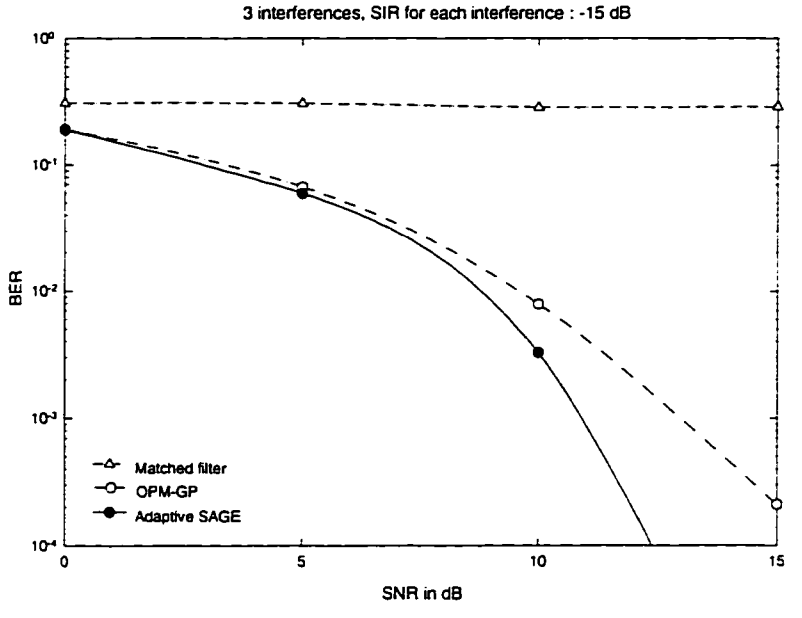

Figure 7.23 Plot of BER vs SNR for the SAGE, MF and OPM-GP receivers with 3 interferences and -15 dB SIR

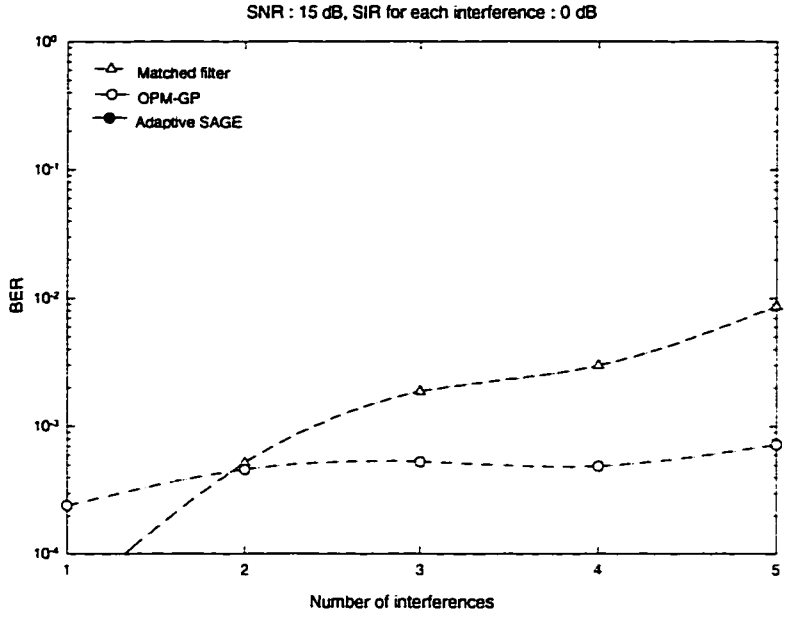

Figure 7.24 Plot of BER vs the number of interferences for the SAGE. MF and OPM-GP receivers with 0 dB SIR

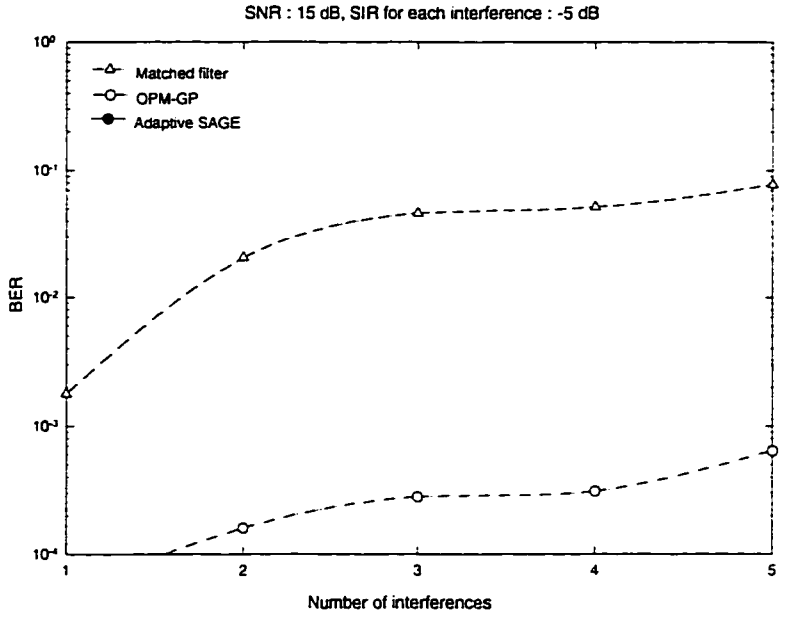

Figure 7.25 Plot of BER vs the number of interferences for the SAGE. MF and OPM-GP receivers with -5 dB SIR

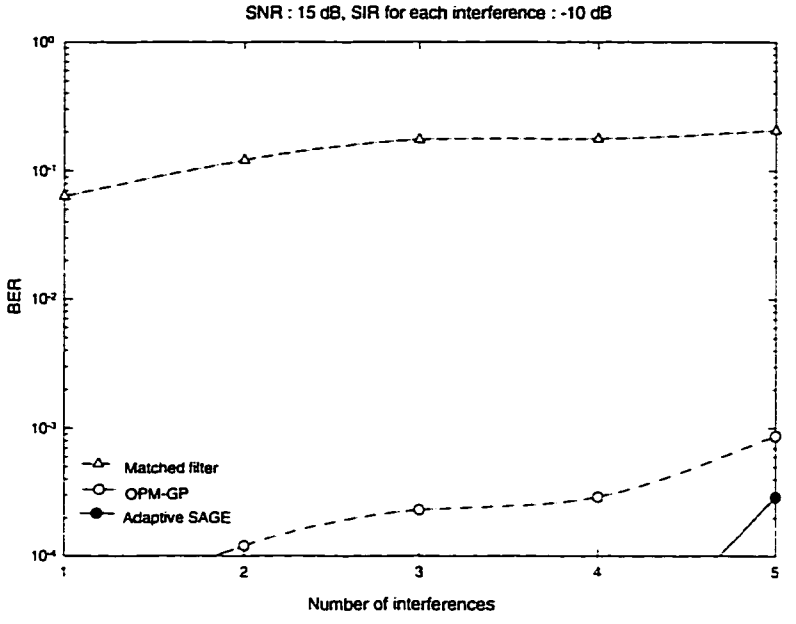

Figure 7.26 Plot of BER vs the number of interferences for the SAGE, MF and OPM-GP receivers with -10 dB SIR

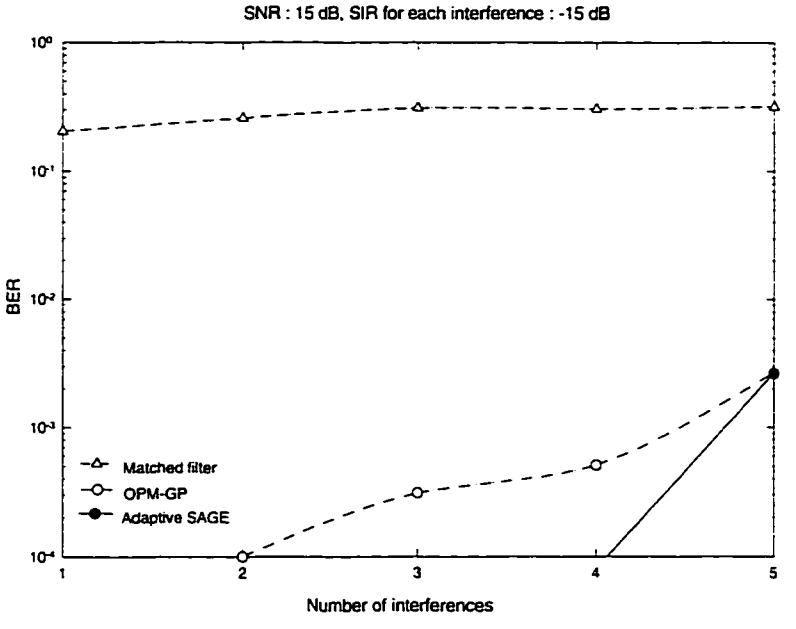

Figure 7.27 Plot of BER vs the number of interferences for the SAGE, MF and OPM-GP receivers with -15 dB SIR

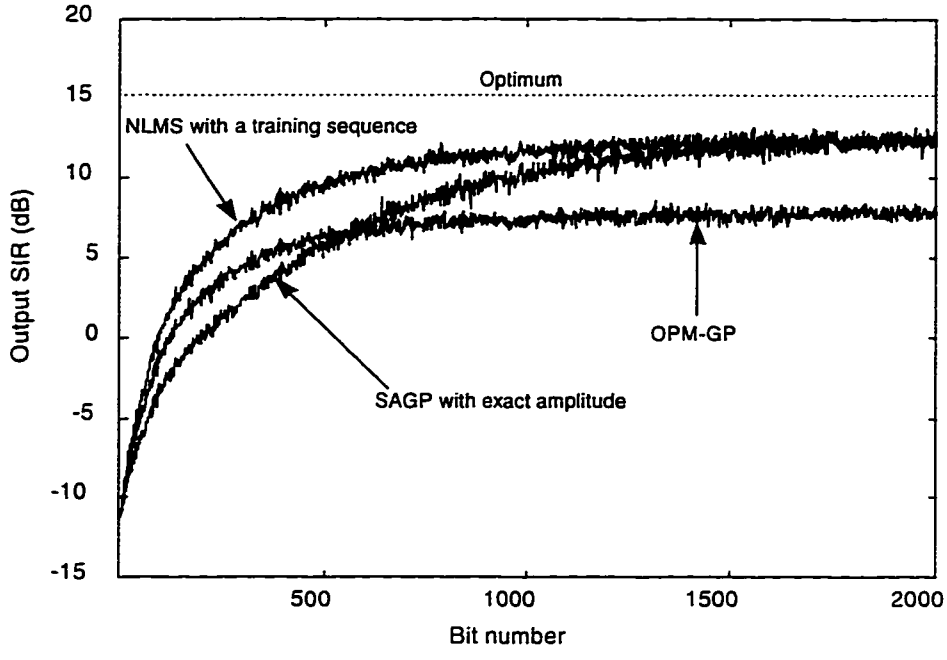

Figure 7.28 Plot of OSIR curves for the normalized LMS receiver and the SAGP receiver with exact amplitude

those of the MF and OPM-GP receivers by varying the number of interferences and the input SIR. In this simulation, the input SNR is given by 15 dB. It can be seen that the SAGE receiver performs error-free detection when the input SIR is higher than -10 dB. The performance of the SAGE receiver becomes similar to that of the OPM-GP receiver when the number of interference is greater than 5 and the input SIR is higher than -10 dB. The simulation results show that the performance of the SAGE receiver significantly outperforms the MF receiver regardless of the number of interference and the input SIR.

SAGP Receiver Performance

SIR Improvement

Figure 7.28 and Figure 7.29 show the SIR improvement of the SAGP receiver over

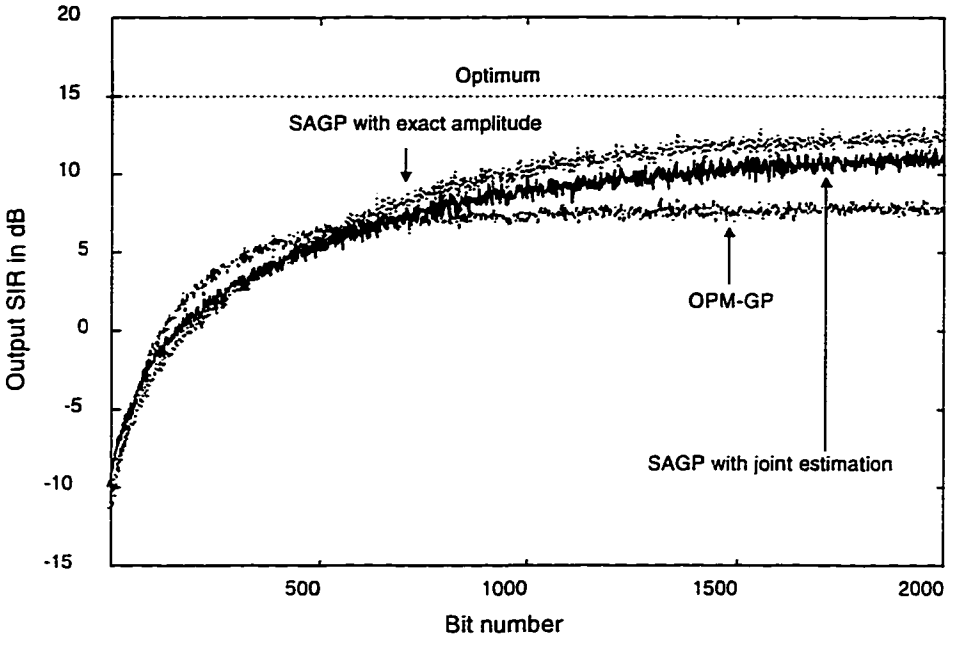

Figure 7.29 Plot of OSIR curves for the SAGP with joint estimation and normaized OPM-GP receivers

the normalized OPM-GP receivers. In these simulations, the received signal is generated with the MAI consisting of 5 interferences having -10 dB SIR and the AWGN having 15 dB SNR. For each realization, a different channel is assumed, i.e., A_1, τ_k and s_k. $k = 2, \cdots, 6$ are chosen at random but for each realization

$$A_k = \sqrt{10}A_1, \quad \text{for } k = 2, \cdots, 6$$

remains constant. At the beginning of each realization, the initial tap weights are set to the matched filter tap weights. The OSIR curves are obtained by averaging over 500 realizations. It can be seen that the SAGP receiver with exact amplitude of the desired signal converges to the near-optimum steady state without using a training sequence. Convergence rate of the SAGP algorithm without using a training sequence is shown to be little slower than that of the normalized LMS receiver. This loss in convergence rate is negligible whereas the improvement in the SIR performance is substantial. Figure

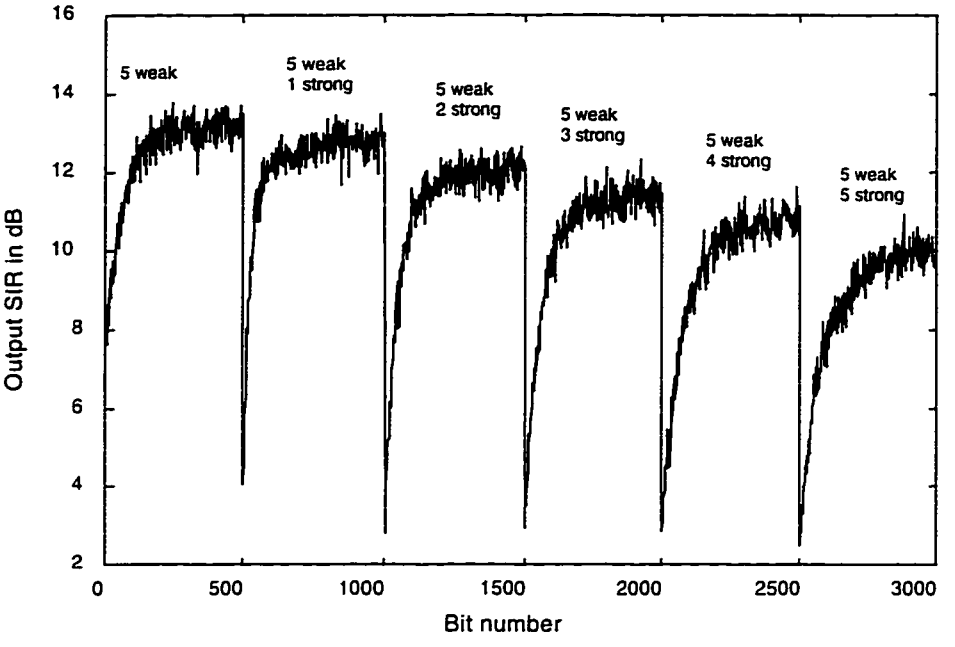

Figure 7.30 Plot of OSIR curve for the SAGP receiver with increasing number of interferences from 5 to 10

7.29 shows that the OSIR curve of the SAGP receiver with joint filter and amplitude estimation is close to that of the SAGP receiver with exact amplitude information. The SIR loss in the SAGP receiver with joint parameter estimation is due to the effect of the amplitude estimation error. Nevertheless, the SAGP receiver outperforms the normalized OPM-GP receiver while using the same amount of information.

Figure 7.30 shows the effect of time-varying interference power on the receiving system. In this simulation, some near-far interferences randomly access the system in the middle of the demodulation of the desired signal according to the following rule: Initially. 5 weak interferences are in the system and in every 500 bits, a new interference with -10 dB SIR accesses the system and keeps transmitting the signals. The amplitude of each DS/CDMA signal at each realization is given by

$$A_1 = A_2 = A_3 = A_4 = A_5 = A_6, \quad \text{for } i = 1, \cdots, 3000.$$

$$A_k = \begin{cases} 0, & \text{for } i = 1, \cdots, L_k \\ \sqrt{10}A_1, & \text{for } i = L_k + 1, \cdots, 2000 \end{cases}, \quad k = 7, \cdots, 11$$

where $L_k = 500\,(k-6)$, $k = 7, \cdots, 11$. It can be seen that, with increasing number of interferences, the SAGP receiver recovers the near-optimum state in a short duration. However, the result also shows that the near-optimum steady state level of the system slowly degrades with increasing number of interferences.

BER Performance

In Figure 7.31 to Figure 7.34, plots of BER vs. input SNR for the SAGP, MF and normalized OPM-GP receivers are shown by varying the input SIRs. In this simulation, the MAI consists of 5 interferences having equal power. It can be seen that the SAGP receiver outperforms the MF receiver while using the same amount of information. The performance gain over the normalized OPM-GP receiver is also substantial regardless of the input SIR when the input SNR is high.

Figure 7.35 to Figure 7.38 show that BER vs. SIR for the SAGP, MF and normalized OPM-GP receivers by varying the input SNR. It can be seen that the performance of the SAGP receiver and the normalized OPM-GP receiver is similar to that of the MF receiver when the input SNR is lower than 5 dB. It can however be seen that the performance gain of the SAGP algorithm over the other receivers is significantly improved when the input SNR is higher than 10 dB. In particular, it can be seen that the SAGP receiver performs error-free detection in the near-far situation when the input SNR is 15 dB.

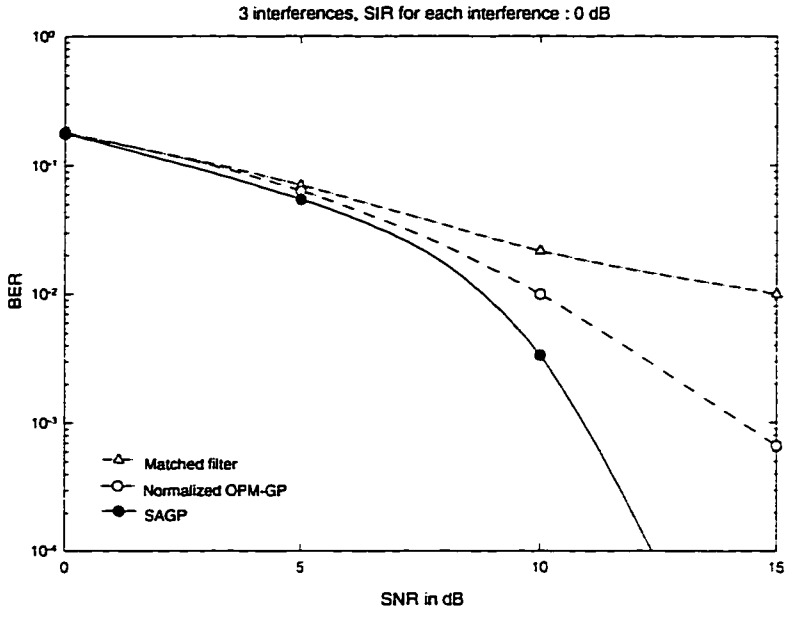

Figure 7.31 Plot of BER vs SNR for the SAGP. MF and normalized OPM-GP receivers with 5 interferences having 0 dB SIR

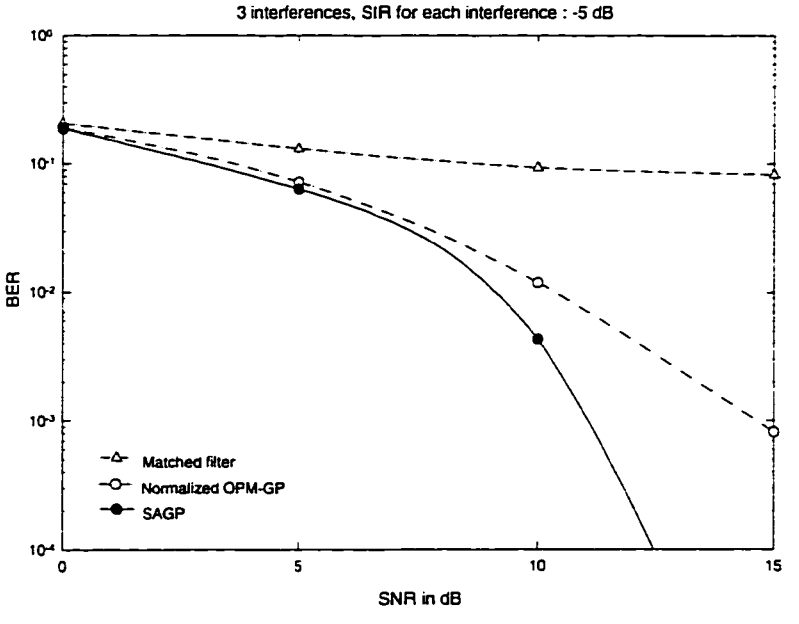

Figure 7.32 Plot of BER vs SNR for the SAGP. MF and normalized OPM-GP receivers with 5 interferences having -5 dB SIR

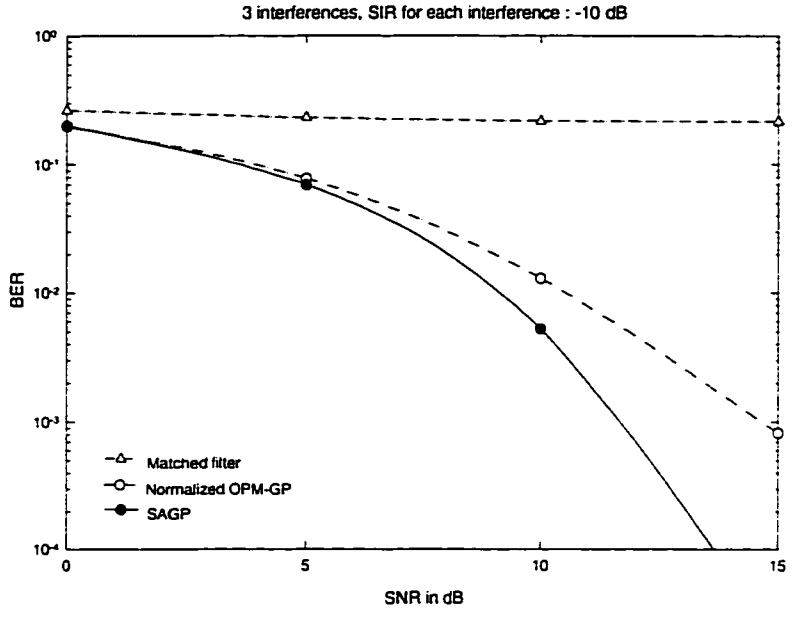

Figure 7.33 Plot of BER vs SNR for the SAGP, MF and normalized receivers with 5 interferences having -10 dB SIR

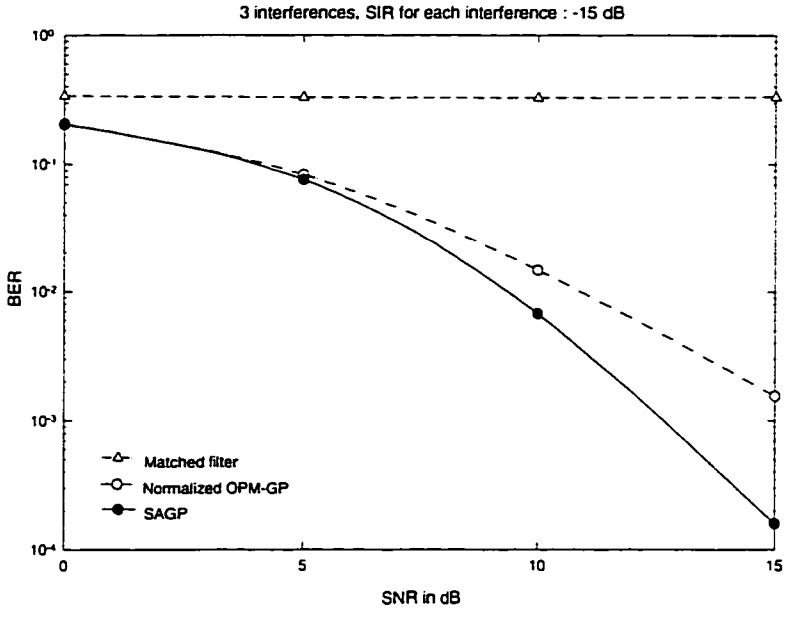

Figure 7.34 Plot of BER vs SNR for the SAGP, MF and normalized OPM-GP receivers with 5 interferences having -15 dB SIR

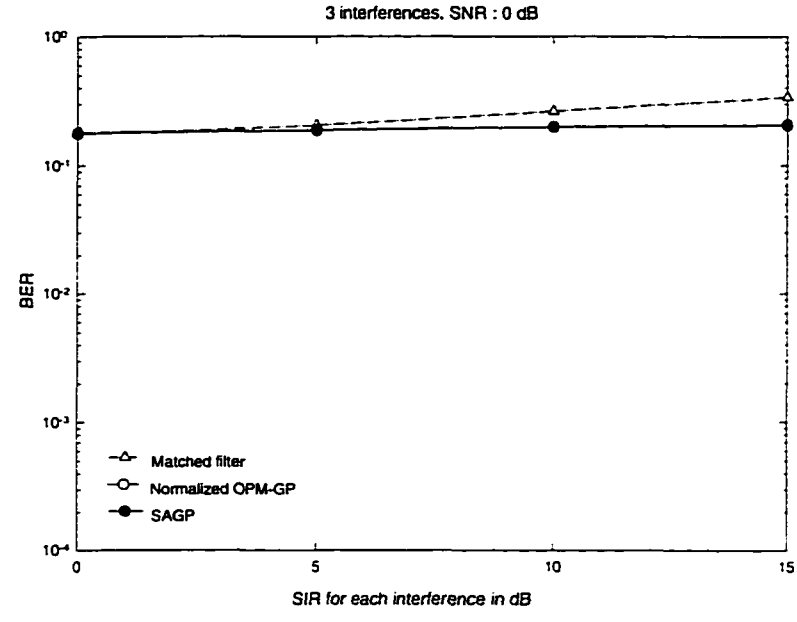

Figure 7.35 Plot of BER vs SIR for the SAGP, MF and normalized receivers with 5 interferences and 0 dB SNR

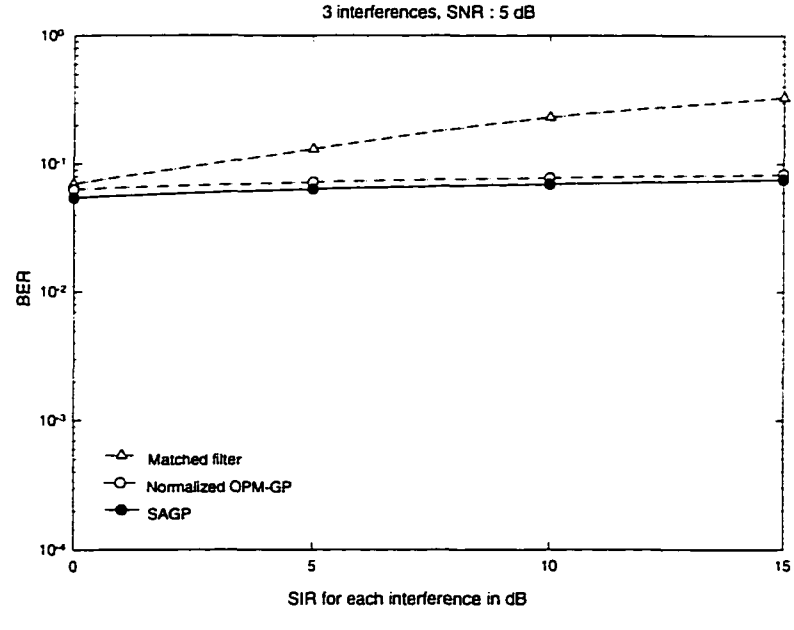

Figure 7.36 Plot of BER vs SIR for the SAGP, MF and normalized OPM-GP receivers with 5 interferences and 5 dB SNR

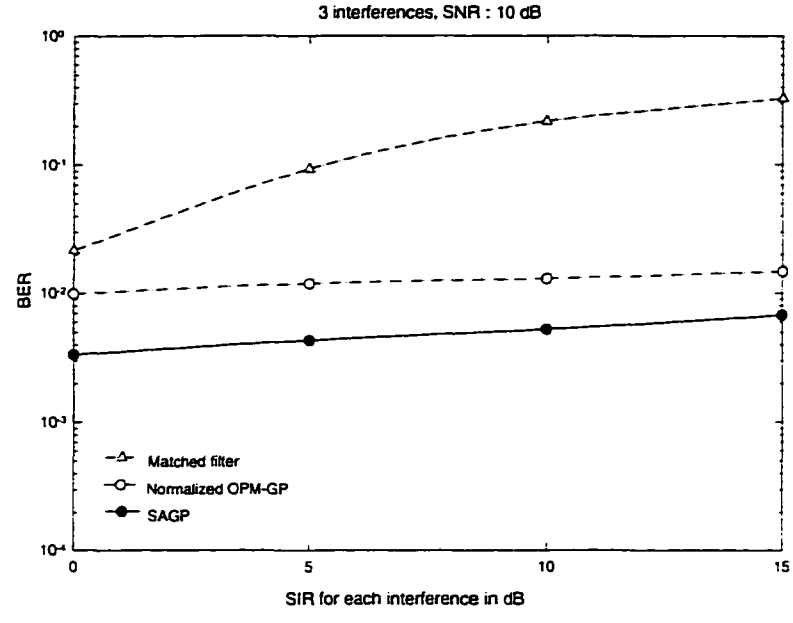

Figure 7.37 Plot of BER vs SIR for the SAGP, MF and normalized OPM-GP receivers with 5 interferences and 10 dB SNR

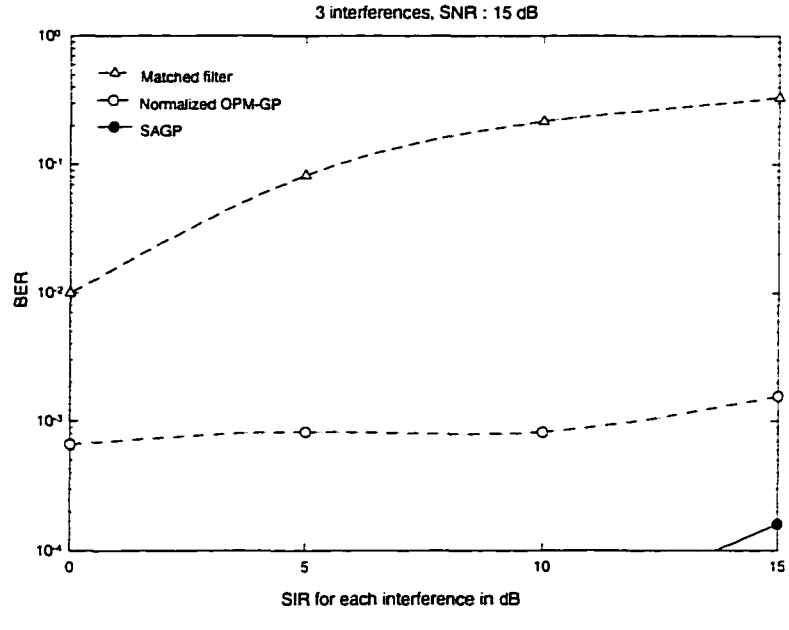

Figure 7.38 Plot of BER vs SIR for the SAGP, MF and normalized OPm-GP receivers with 5 interferences and 15 dB SNR

CHAPTER 8 CONCLUSIONS

New blind adaptive near-far resistant receivers that require neither training sequences nor any information of the interfering signals are proposed. Proposed receivers are tested in simulated stationary and nonstationary environments. The simulation results show that the proposed receivers significantly outperform the conventional matched filter receiver and the blind adaptive receiver employing the constrained output-power minimizing (OPM) algorithm. It is also shown that algorithms in the proposed receivers are computationally efficient in the implementation and robust to the implementation errors. Following conclusions are drawn from the discussion and simulation results:

- Proposed receivers achieve error-free detection regardless of the input SIR in the high SNR region. This proves that the proposed receivers are asymptotically near-far resistant and asymptotically converges to a zero-forcing solution.

- Performance gains of proposed receivers over the constrained OPM receiver is significant when the input SNR is high. This can be explained by the fact that the algorithms in the proposed receivers guarantee asymptotic zero tap gain increment when the perfect adaptation is achieved whereas the OPM algorithm experiences the variance proportional to the signal power in the same situation. In other words, the stochastic driving term of proposed algorithms in the steady state produces much smaller MSE than that of the OPM algorithm for same rate of convergence.

- The performance gain of the constrained LMS receiver over the receiver employing gradient projection (GP) algorithm is significant when the filter mismatches to

the desired signal characteristic in the mismatch ratio of -30 dB or lower. This can be explained by the fact that the constrained LMS algorithm provides error correction features whereas the GP algorithm suffers from error propagation effect during the iteration process.

- The performance of the constrained LMS receiver is further improved by incorporating a robust constraint in the presence of the filter mismatch effect. This can be explained by the fact that the robust constraint holds the filter weights in the region of attraction by properly scaling its norm.

- Even when the amplitude of the desired signal is not available, the detection performance of the receiver can be improved by incorporating an adaptive joint parameter estimation algorithm. This can be explained by the fact that the algorithm is implemented in the direction of ensuring the asymptotic zero tap gain inclement when the perfect adaptation is achieved.

- Without any other information on the received signal other than that of the conventional receiver, it is possible to achieve near-far resistance and near-optimum performance by adaptively learning the structure of the DS/CDMA signals.

APPENDIX PROJECTIONS ON A HILBERT SPACE

This appendix gives some definitions and properties in *projection theorem on a Hilbert space.* A close relationship between a projection operator and the linear transformation in a Hilbert space can be used to obtain an optimum filter in a finite dimensional space. Some definitions are in order:

Definition 6 *[49] Let U, V be two vector spaces over the field $\mathcal{R}$. A map from T from U into V (or briefly, $T : U \rightarrow V$) is said to be a linear transformation (or map) if*

$$T(x+y) = T(x) + T(y), \quad \text{for all } x, y \text{ in } V \tag{A.1}$$

and

$$T(\alpha x) = \alpha T(x), \quad \text{for all } \alpha \text{ in } \mathcal{R} \text{ and all } x \text{ in } V. \tag{A.2}$$

Definition 7 *[49] An inner product in a vector space is a numerically valued function of the ordered pair of vector u and v, such that*

$$\langle x, y\rangle = \langle x, y\rangle \tag{A.3}$$

$$\langle \alpha_1 x_1 + \alpha_2 x_2, y\rangle = \alpha_1 \langle x_1, y\rangle + \alpha_2 \langle x_2, y\rangle, \tag{A.4}$$

$$\langle x, x\rangle \geq 0; \quad \langle x, x\rangle = 0 \text{ if and only if } x = 0. \tag{A.5}$$

An inner product space is a vector space with an inner product.

Definition 8 *[49] Let $(H, \langle\rangle)$ be an inner product space and let $\|\cdot\|$ be its associated norm. If $(H, \|\cdot\|)$ is a complete normed space, then we shall say that $(H, \langle\rangle)$ is a Hilbert space.*

That is, the Hilbert space is an inner product space satisfying one extra condition, namely completeness. This condition is automatically satisfied in the finite-dimensional vector space [49].

Definition 9 *[49] Let U, V be two vector spaces over the same filter $\mathcal{R}$. The direct sum is the vector space W denoted by*

$$W = U \oplus V \tag{A.6}$$

whose elements are all the ordered pairs $\langle u, v\rangle$ with u in U and v in V, with the linear transformation defined by

$$\alpha_1 \langle u_1, v_1\rangle + \alpha_2 \langle u_2, v_2\rangle = \langle \alpha_1 u_1 + \alpha_2 u_2, \alpha_1 v_1 + \alpha_2 v_2\rangle . \tag{A.7}$$

Especially important for our purpose is connection between the direct sum and the linear transformation. The *projections on a Hilbert space* is defined as follows.

Definition 10 *[35] If $W = U \oplus V$, so that every w in W may be written, uniquely, in the form $w = u + v$, with u in U and v in V, the projection on U along V is the linear transformation P defined by*

$$Pw = u. \tag{A.8}$$

Here are some nice properties of projections.

Theorem 4 *[35] A linear transformation P is a projection on some subspace if and only if it is idempotent, that is,*

$$P^2 = P. \tag{A.9}$$

If P is the projection on U along V, and if $w = u + v$, with u in U and v in V, then the decomposition of u is $u + 0$, so that

$$P^2 w = PPw = Pu = u = Pw. \tag{A.10}$$

If w is in U, then $Pw = w$; if w is in V, then $Pw = 0$; hence if w is in both U and V, then $w = 0$. For an arbitrary w, it can be decomposed as

$$w = Pw + (1 - P)\, w. \tag{A.11}$$

If we write $Pw = u$ and $(1 - P)\, w = v$, then

$$Pu = P^2 w = Pw = u, \tag{A.12}$$

and

$$Pv = P\,(1 - P)\, w = Pw - P^2 w = 0, \tag{A.13}$$

so that u is in U and v is in V. This proves that

$$W = U \oplus V \tag{A.14}$$

and that the projection on U along V is precisely P. As an immediate consequence, one can obtain the following results.

Theorem 5 *[35] If P is the projection on U along V, then U and V are, respectively, the sets of all solutions $Pw = w$ and $Pw = 0$.*

Theorem 6 *[35] A linear transformation P is a projection if and only if $1 - P$ is a projection; if P is the projection on U along V, then $1 - P$ is the projection on V along U.*

BIBLIOGRAPHY

[1] S. Verdu, "Recent Progress in Multiuser Detection." in *Multiple Access Communications*, IEEE Press, Piscataway, NJ, 1993.

[2] S. Verdu, "Minimum Probability of Error for Asynchronous Gaussian Multiple-Access Channels," *IEEE Transactions on Information Theory, Vol.32, No.1*, pp. 85–96, Jan. 1986.

[3] QUALCOMM, *The CDMA Network Engineering Handbook Volume 1: Concepts in CDMA*. QUALCOMM Inc., San Diego, CA, February, 1993.

[4] S. Verdu, "Optimum Multiuser Asymptotic Efficiency," *IEEE Transactions on Communications, Vol.34*, pp. 890–897, Sep. 1986.

[5] R. Lupas and S. Verdu, "Near-Far Resistance of Multiuser Detector in Asynchronous Channels," *IEEE Transactions on Communications, Vol.38, No.4*, pp. 496–508, Apr. 1990.

[6] R. Lupas and S. Verdu, "Linear Multiuser Detectors for Synchronous Code-Division Multiple-Access Channels," *IEEE Transactions on Information Theory, Vol.35, No.1*, pp. 123–136, Jan. 1989.

[7] M. K. Varanasi and B. Aazhang, "Near-Optimum Detection in Synchronous Code-Division Multiple-Access Systems," *IEEE Transactions on Communications. Vol.39, No.5*, pp. 725–736, May 1991.

[8] M. K. Varanasi and B. Aazhang, "Multistage Detection in Asynchronous Code-Division Multiple-Access Communications," *IEEE Transactions on Communications, Vol.38, No.4*, pp. 509–519, Apr. 1990.

[9] A. Duel-Hallen, "Decorrelating Decision-Feedback Multiuser Detector for Synchronous Code-Division Multiple-Access Channel," *IEEE Transactions on Communications, Vol.41, No.2*, pp. 285–290, Feb. 1993.

[10] A. Duel-Hallen, "Equalizers for Multiple Input/Multiple Output Channels and PAM Systems with Cyclostationary Input Sequences," *IEEE Journal on Selected Areas in Communications, Vol.10, N0.3*, pp. 630–639, April 1992.

[11] Z. Xie, R. T. Short, and C. K. Rushforth, "A family pf suboptimum detectors for coherent multiuser communications," *IEEE Journal on Selected Areas in Communications. Vol.8. N0.4*, pp. 683–690, May 1990.

[12] U. Madhow and M. L. Honig, "MMSE Interference Suppression for Direct-Sequence Spread-Spectrum CDMA," *IEEE Transactions on Communications, Vol.42. N0.12.* pp. 3178–3188, Dec. 1994.

[13] P. B. Rapajic and B. S. Vucetic, "Adaptive Receiver Structures for Asynchronous CDMA Systems," *IEEE Journal on Selected Areas in Communications. Vol.12. No.4*, pp. 685–697, May 1994.

[14] S. L. Miller, "An Adaptive Direct-Sequence Code-Division Multiple-Access Receiver for Multiuser Interference Rejection," *IEEE Transactions on Communications. Vol.43. No.2.* pp. 1746–1754. 1995.

[15] S. Haykin, *Blind Deconvolution.* Prentice Hall, Inc., Englewood Cliffs, NJ, 1994.

[16] M. Honig, U. Madhow, and S. Verdu, "Blind Adaptive Multiuser Detection," *IEEE Transactions on Information Theory, Vol.41, N0.4,* pp. 944–960, July 1995.

[17] J. B. Schodorf and D. B. Williams, "A Blind Adaptive Interference Cancellation Scheme for CDMA Systems," *Proc. Asilomar Conf. Sig. Sys. and Comp.*, 1995.

[18] V. Widrow, K. M. Duvall, R. Gooch, and W. C. Newman, "Signal Cancellation Phenomena in Adaptive Antennas: Causes and Cures," *IEEE Transactions on Antenna and Propagation, Vol. AP-30, No. 3,* pp. 469–478, May 1982.

[19] Z. Ding and R. A. Kennedy, "On the Whereabouts of Local Minima for Blind Adaptive Equalizers," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing, Vol. 39, No. 2,* pp. 119–123, February 1992.

[20] R. E. Ziemer and R. L. Peterson, *Digital Communications and Spread Spectrum Systems.* Macmillan, New York, NY, 1985.

[21] R. Kohno, "Spatial and temporal filtering for co-channel interference in cdma," in *Code Division Multiple Access Communications* (S. Glisic and P. Leppanen, eds.). pp. 117–146, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1995.

[22] R. A. Scholtz, "The spread spectrum concept," in *Multiple Access Communications.* IEEE Press, New York, NY, 1993.

[23] H. V. Poor and S. Verdu, "Single-User Detectors for Multiuser Channels," *IEEE Transactions on Communications, Vol. 36, No. 1,* pp. 50–60, January 1988.

[24] K. S. Schneider, "Optimum Detection of Code Division Multiplexed Signals," *IEEE Transactions on Aerospace and Electronic Systems, Vol. 15, No. 1,* pp. 181–185, January 1979.

[25] S. Verdu, "Adaptive multiuser detection," *IEEE Third International Symposium on Spread Spectrum Techniques and Applications*, pp. 43–50, Jul. 1994.

[26] R. Kohno, H. Imai, and M. Hatori, "Cancellation Techniques of Co-Channel Interference in Asynchronous spread Spectrum Multiple Access Systems," *Trans. IECE (Electronics and Communications in Japan), Vol. 66*, pp. 416–423, May 1983.

[27] M. Abdulrahman and D. D. Falconer, "Cyclostationary Crosstalk Suppression by Decision Feedback Equalization on Digital Subscriber Loops," *IEEE Journal of Selected Areas in Communications*, pp. 640–649, April 1992.

[28] D. W. Chen, Z. Siveski, and Y. Bar-Ness, "Synchronous Multiuser CDMA Detector with Soft Decision Adaptive Canceller." *Proceedings of the 1994 Conference on Information Science and Systems*, march 1994.

[29] S. Verdu, B. D. O. Anderson, and R. Kennedy, "Blind Equalization without Gain Identification," *IEEE Transactions on Information Theory, Vol.39*, pp. 292–297. 1993.

[30] H. Oda and Y. Sato, "A Method of Multidimensional Blind Equalization," *IEEE International Symposium on Information Theory*, p. 327, January, 1993.

[31] R. Singh, *Adaptive direct-sequence spread-spectrum communications*. PhD thesis, University of Califonia, San Diego, CA, 1996.

[32] S. C. Park and J. F. Doherty, "A MINIMAX Optimization Approach to Sidelobe Suppression Filter Design," in *Proceedings of the 1996 International Conference on Acoustics, Speech, and Signal Processing*, 1996.

[33] D. G. Luenberger, *Optimization by Vector Space Methods*. John Wiley and Sons. Inc., New York, NY, 1969.

[34] S. Haykin, *Adaptive Filter Theory.* Prentice-Hall Inc., Englewood Cliffs, NJ, 1991.

[35] P. R. Halmos, *Finite-Dimensional Vector Spaces.* Springer-Verlag, New York, NY. 1993.

[36] O. L. Frost, "An Algorithm for Linearly Constrained Adaptive Array Processing," *Proceedings of the IEEE, Vol.60, No.8.* pp. 926–935, 1972.

[37] K. Yamazaki and R. A. Kennedy, "Reformulation of Linearly Constrained Adaptation and its Application to Blind Equalization," *IEEE Transactions on Signal Processing, Vol.42, No.7,* pp. 1837–1841, 1994.

[38] L. L. Scharf. *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis.* Addison-Wesley Publishing Company, New York, NY, 1991.

[39] H. Cox, R. Zeskind, and M. Owen, "Robust Adaptive Beamforming," *IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol.35, No.10,* pp. 1365–1375. Oct. 1987.

[40] J. B. Rosen, "The Gradient Projection Method For Nonlinear Programming. Part 1. Linear Constraints," *J. Soc. Indust. Appl. Math., Vol. 8, No. 1,* pp. 181–215. Mar. 1960.

[41] J. A. Fessler and A. O. Hero, "Space-Alternating Generalized Expectation-Maximization Algorithm," *IEEE Transactions on Signal Processing,* vol. 42. pp. 2664–2677, October 1994.

[42] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of Royal Statististical Society, Ser.B. Vol.37,* pp. 1–38, 1977.

[43] M. Feder, A. V. Oppenheim, and E. Weinstein, "Maximum Likelihood Noise Cancellation Using the EM Algorithm," *IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol.37, No.2*, pp. 204–216, Feb. 1989.

[44] H. V. Poor, "On Parameter Estimation in DS/SSMA Formats," in *Lecture Notes in Control and Information Sciences Vol. 129.*, Springer-Verlag, Berlin, Germany, 1989.

[45] H. V. Poor and R. Vijayan, "Analysis of Class of Adaptive Nonlinear Predictors," in *Advances in Communications and Signal Processing*, pp. 231–241, Springer-Verlag, Berlin, Germany, 1989.

[46] A. Levi and H. Stark, "Image Restoration by the Method of Generalized Projections with Application to Restoration from Magnitude," *Journal of Optical Society of America A, Vol.1, No.9*, pp. 932–943, 1984.

[47] M. Sezan and H. Stark, "Image Restoration by Convex Projections in the Presence of Noise," *Applied Optics*, vol. 22, pp. 2781–2789, September 1983.

[48] J. A. Cadzow, "An extrapolation procedure for band-limited signals," *IEEE Trans. Acoust., Speech, and Signal Processing, Vol. ASSP-27*, pp. 4–12, Feb. 1979.

[49] C. L. DeVito, *Functional Analysis and Linear Operator Theory*. Addison-Wesley Publishing Company, Redwood City, CA, 1990.