THE UNIVERSITY OF CALGARY

# Multiuser Detection and Channel Estimation for Synchronous CDMA Systems

by

**Mohsen Hosseinian**

A DISSERTATION

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

CALGARY, ALBERTA

FEBRUARY, 1999

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

*Your file Votre référence*

*Our file Notre référence*



0-612-38474-8

Canada

# Abstract

In a code-division multiple-access (CDMA) system, several users transmit simultaneously over a common channel. Several different multiuser detectors for CDMA systems have been suggested by researchers. In this dissertation the most celebrated multiuser detectors are studied in terms of the bit-error-rate (BER) and the computational complexity. We also propose a multiuser detection scheme, namely the two-level threshold detection, for coherent demodulation in a synchronous CDMA system. The proposed method has a computational complexity that is linear in the number of users, and exhibits a performance that is close to that of the optimum detector whose complexity grows exponentially with the number of users.

Also a decorrelating-type detector is considered for multiuser detection in frequency selective synchronous CDMA channels. This detector operates on the outputs of matched filters which are matched to the original spreading codes of the system. A maximum likelihood estimation (MLE) to estimate parameters of the decorrelating-type filter is proposed and derived. The estimation method is based on inserting known training sequences into the information data by all users simultaneously. To achieve minimum mean-square-error (MMSE) in estimation, a criterion for selecting the training sequences is suggested. Orthogonal training sequences will be good candidates to approach MMSE. The estimation method requires a matrix inversion at the end of each training period. An iterative matrix inversion algorithm is introduced to distribute the computational load of the matrix inversion over the training period. The simulation results exhibit BER floor, which is due to the presence of intersymbol interference (ISI). The performance of the detector depends on the channel rms delay spread.

In order to cancel the existing ISI, a decorrelating-type decision-feedback detector is proposed. This detector consists of two filters: a forward filter to cancel the interchannel interference (ICI) and a feedback filter to cancel ISI. The MLE estimates of the coefficients of these two filters is proposed and derived. MMSE sequences are also obtained. The simulation results show no BER floor. The performance of the detector does not depend severely on the channel rms delay spread.

# Acknowledgements

I would like to acknowledge appreciation to a number of people who have contributed directly or indirectly to the writing of this dissertation.

My supervisors, Dr. M. Fattouche and Dr. A. B. Sesay, must be specially acknowledged for their great help, scholarly advice, editing of this manuscript and financial support.

I am deeply grateful to TRLabs for providing me a well-equipped, comfortable environment for graduate research. Special thanks go to George Squires and Leila Southwood for their role in fostering such an environment.

Thanks also due to all my friends who helped me in many ways during this endeavor.

The Iranian Ministry of Culture and Higher Education provided me with a financial support without which I might not even attempted pursuing my studies abroad.

My great thanks must go to my parents who did everything they could to educate me with all the difficulties they had.

Finally, this work would never have been completed without the encouragement, understanding and love of my dear wife, Souri. She has put up, uncomplainingly, with my seemingly unending periods in my study both during the evenings and at weekends.

To

my wife, Souri

and our dear son, Daniel

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

Over the last decade the interest in wireless communications has dramatically increased. Existing forms of wireless communications: cellular mobile telephones, wireless networks, cordless phones and radio pagers, continue to experience an explosive growth, showing a large increase in the number of users. Capacity, radio spectrum utilization efficiency and service quality are of primary concern, but also other factors have to be considered for wireless communication. An important issue is the complexity of the technology which will affect cost, power consumption, and complexity of the control functions required for proper network operation [3]. Digital modulation, detection and multiple access techniques are essential components in the design of any communication system. During the last decade, many known techniques have been analyzed and re-evaluated for mobile communications. Spread-spectrum communication is a well-known technique which has found a place in cellular mobile systems. In particular, direct-sequence code-division multiple-access (DS-CDMA) systems have attracted much interest. A variety of new techniques have also been investigated. The work in this dissertation concentrates on issues relating to multiuser modulation and detection schemes. The terminology

*multiuser detection* refers to techniques which detect the transmitted information of several users jointly. *Multiuser modulation* refers to multiple access schemes. The research work underlying this dissertation aims at achieving the following goals:

1. to study the most actively researched multiuser detectors, which have been proposed by researchers recently, and to compare their performances in terms of bit-error-rate and computational complexity.

2. to develop a low complexity multiuser detection structure that can achieve near-optimum performance for all users.

3. to design a channel estimation scheme that is simple and does not require extensive computational load.

CDMA has been widely applied in military communications. The original motivation for spread-spectrum systems was both to conceal transmission and to combat intentional jamming [4]. In commercial mobile communications, CDMA systems have many benefits such as interference suppression, a higher spectrum reuse factor, soft capacity, soft hand-over, wideband multipath diversity, and voice activity utilization [5]. For these reasons, CDMA has attracted much attention recently. One of the current second-generation cellular communications systems, IS-95, is based on narrowband DS-CDMA technology [5]. The most promising candidate for the new third-generation mobile communications systems, called International Mobile Telecommunications-2000 (IMT-2000), is wideband CDMA (W-CDMA) [6–10]. Recently the European Telecommunications Standards Institute (ETSI) decided to adopt W-CDMA technology for frequency-division duplex (FDD) bands [11]. In Japan, the Association of Radio Industries and Businesses (ARIB), the standardization body of the radio sector, is now developing a W-CDMA air interface stan-

dard [12].

In this dissertation the focus is mainly on DS-CDMA systems. The performance evaluation of CDMA systems is a complicated task, even on Gaussian channels [13]. CDMA systems have many benefits [5]. However, DS-CDMA systems with single-user detection suffer two major drawbacks: the *near-far problem* and an interference limitation on network capacity. The near-far problem is a situation in which users near the receiver are received at higher powers than those far away, and those further away suffer a degradation in performance, i.e. bit-error-rate. A single-user detector for each user consists of a matched filter and a decision device. Since the output of each matched filter contains a spurious component which is linear in the amplitude of each of the interfering users, the strongest user often severely interferes with the other users. Consequently, the bit-error-rate and the anti-jamming capability of the weakest user are degraded substantially. Thus, in order to maintain an acceptable level of bit-error-rate for all users, DS-CDMA often requires a strict control of the transmitter power for each user, which is often very difficult to realize. The interference limitation refers to the fact that the number of simultaneous users is limited to approximately 10% of the processing gain, even for the case of perfect power control [5]. These drawbacks, caused by the fact that single-user detection treats multiuser interference as noise, severely impair the performance of the CDMA system. It is not surprising that DS-CDMA does not provide a significantly higher spectrum utilization efficiency than TDMA [14].

The drawbacks of single-user detection have initiated recent interests into more sophisticated receiver structures such as multiuser joint detection, in which the multiuser interference is treated as a part of the information rather than noise. The work of Verdú [15–18] has shown that an optimum maximum likelihood multiuser

detector can achieve optimum near-far resistance and a significant performance improvement over the conventional (single-user) detector. The improvement, however, is obtained at the expense of a dramatic increase in computational complexity, which grows exponentially with the number of users. Thus, when the number of users is large, the optimum detector becomes infeasible. Many suboptimum multiuser detectors have been proposed [17,19–25]. Some of them can approach optimum bit-error-rate performance for the weakest user [19–21]. Some achieve near optimum performance for all users at the expense of more computational complexity [25]. In this dissertation, a suboptimum linear complexity detector is proposed, which can achieve near-optimum performance.

Multiuser detection can significantly improve the spectrum utilization efficiency and the performance of CDMA systems. The number of users jointly decoded in DS-CDMA systems could be very large, since each user occupies the entire bandwidth and is thus fully overlapped with all other users. Multiuser detection may then only be feasible at the base station. In general, it is easier to apply Multiuser detection into a system with short spreading codes since cross-correlation does not change every symbol as with long spreading codes. Due to implementation issues multiuser detection, at least currently, is not considered for W-CDMA downlink systems [10]. The dissertation is organized as follows.

Chapter 2 focuses on the multiuser detection in *additive white Gaussian noise* (AWGN) channels. Previous work on multiuser detection in the literature is presented. The significance of the multiuser detection and its advantages are pointed out. Six different multiuser detection schemes are studied and compared in terms of bit-error-rate and computational complexity. On one hand, the conventional detector is vulnerable to the near-far problem. On the other hand the computational

complexity of the optimum detector is so high that it is infeasible, at least with existing technology. Four other suboptimum multiuser detectors are studied. The computational complexity of these detectors is linear in terms of the number of users, while their performances are near-optimum. In an AWGN channel, provided that a set of spreading codes with good cross-correlation properties is used, all suboptimum multiuser detection schemes yield satisfactory bit-error-rate performances.

In Chapter 3 a new suboptimum multiuser detector, namely the two-level threshold detector, is introduced. Previous work in the area is summarized and the contribution of this present work is emphasized. In this work we develop and investigate the two-level threshold detector, whose complexity is linear in the number of users. The bit-error-rate performance of this detector is close to that of the optimum detector whose complexity grows exponentially with respect to the number of users. The two-level threshold detector uses a decorrelating filter to remove the multipath access interference (MAI) completely. If the output of the decorrelating filter is in the vicinity of decision regions' boundary, e.g. $\epsilon$, then the idea of the optimum detector with some simplifications is used. This way with some increase in the computational complexity, compared to the decorrelating detector, a near-optimum performance can be achieved. The radius $\epsilon$ has a crucial impact on the performance of this detector.

Chapter 4 includes a review of the radio propagation channels. Since we intend to investigate the performance of the multiuser detectors in the multipath fading channels, we need a simple and reliable model for such channels. The discrete-time channel impulse response is explained. Some useful parameters of multipath channels are put into perspective. It is shown that the two-ray Rayleigh model can achieve a large variety of multipath fading channels from flat fading to frequency

fading. The two-ray model has some drawbacks, but since our intention in this dissertation is to compare the behavior of several multiuser detectors, it seems appropriate to use this model.

In Chapter 5 an investigation of multiuser detection in multipath channels is presented. It concentrates on frequency selective channels. A two-ray model for such channels is used. The channel characteristics are assumed to be known perfectly. Two general strategies, namely *channel-matched* and *path-by-path* are explained. Due to the better bit-error-rate performance, the channel-matched approach is preferred. Using this strategy all multiuser detectors are modified and deployed in a multipath channel. The two-level threshold detector shows a better near-optimum performance compared to the other detectors. The optimum value for the radius $\epsilon$ is also obtained in this chapter.

Chapters 6 and 7 are concerned mostly with the estimation of channel parameter. In Chapter 6 a decorrelating-type detector is considered for multiuser detection. The decorrelating-type filter operates on the outputs of matched filters which are matched to the original spreading codes of system. A maximum likelihood estimation to estimate parameters of the decorrelating-type filter is proposed and derived. The estimation method is based on inserting known training sequences into the information data by all users simultaneously. To achieve *minimum mean-square-error* (MMSE) in estimation, a criterion for how to select the training sequences is suggested. Orthogonal training sequences are shown to be good candidates to approach MMSE. The estimation method requires a matrix inversion at the end of each training period. An iterative matrix inversion algorithm is introduced to distribute the computational load of the matrix inversion over the training period. The simulation results show some degradations in system performance compared to

the case where the perfect knowledge of the channel is assumed. This degradation is clearly due to errors in the channel estimation.

In Chapter 7 a decorrelating-type decision-feedback detector is considered for multiuser detection for frequency-selective channels. The feedback filter is used to remove *intersymbol interference* (ISI). A maximum likelihood estimation method is developed to estimate parameters of both forward filter and feedback filter. Similar to the estimation method suggested in Chapter 6, the estimation method is based on inserting known training sequences into the information data by all users simultaneously. The training sequences, however, are longer. MMSE training sequences are obtained and used.

In Chapter 8 a brief summary of the accomplished work, with an emphasis on the contributions to the area of detection and estimation methods for wireless communications, is presented. Some possible extensions of this work and directions for future research are also presented.

Four conference papers, [26–29], have been published based on this work.

# Chapter 2

# Multiuser Detection in AWGN Channel

The idea of interference cancellation arises in many contexts, e.g., noise cancellation in speech [30] and adaptive interference canceling [31, Chapter 12]. There are thus a number of non-CDMA references with ideas similar to those being currently studied for CDMA. We should differentiate between canceling noise or single-tone jammer, which have no useful purpose, from canceling interference which is due to other signals that are themselves to be detected. The CDMA case considered in this dissertation is of the second type, where the signals being canceled are of interest also. It should be mentioned, however, that the first type of cancelation is also of importance in CDMA systems, e.g., in suppressing narrowband interference (this is not discussed in this dissertation). Both types of interference cancelation have in common the goal of removing a noise-like interference from a desired signal. In the second type, however, the fact that the signals being removed are themselves information carrying, leads to a new viewpoint that of simultaneously detecting all the information carrying signals.

This chapter has the following outline. In Section 2.1 previous work in the area is summarized and in Section 2.2 the contributions of this chapter are put into perspective. In Section 2.3 the system model is presented. In Sections 2.4 we describe the idea behind multiuser detection. Section 2.5 clarifies the measures of the comparison that we have chosen as criteria to compare various multiuser detectors. The computer simulation model that we have used throughout this chapter is explained in Section 2.6. Sections 2.8, 2.9, 2.10, 2.11, 2.12 and 2.13 investigate six multiuser detection schemes: The conventional, the optimum, the decorrelating, the decision-feedback, the improved decision-feedback and the multistage detectors, respectively. Finally in Section 2.14 the results are discussed and summarized.

## 2.1 Previous Work

The first CDMA interference cancellation references we are aware of are [32, 33] and [34–36]. These papers describe a number of ideas that are present in much of the ongoing research. Estimates based on mean square error and maximum likelihood are discussed in [32]. Later on the detection scheme proposed in that paper was known as the decorrelating detector. In [32], however, it was erroneously shown that this detector is optimum in terms of bit-error-rate. Reference [36] shows how cancellation is implemented by solving simultaneous equations, in essence, by inverting a key matrix. There were subsequently a number of papers with variants of the ideas of [32] and [34]. Significant theoretical steps forward were taken in [15, 16, 37, 38] by Verdú, in analyzing the structure and complexity of optimal receivers. This work initiated a new research effort on suboptimal algorithms. The strong connection between multiple access interference (MAI) and intersymbol interference (ISI) was also made in [15].

The work of Verdú has shown that optimum near-far resistance and a significant performance improvement over the conventional single-user detector can be achieved by an optimum maximum likelihood multiuser detector. The substantial improvements, however, are obtained at the expense of a dramatic increase in computational complexity. Several low complexity multiuser detectors have been proposed. Lupas and Verdú considered a linear multiuser detector in [17,18,39]. The linear multiuser detector achieves optimum near-far resistance but cannot provide close to optimum performance. A multistage technique was proposed by Varanasi and Aazhang in [19,20], and the decision-feedback detector was suggested by Duel-Hallen in [21,22]. Due to error propagation, the multistage detector can achieve near-optimum performance only when the interfering users are significantly stronger than the user under consideration. The decision-feedback detector employs forward decorrelating and feedback filters to cancel MAI. This cancels MAI completely, provided that the feedback data are decoded correctly.

Tree-type maximum likelihood sequence detectors have also been studied for multiuser systems. Both sequential detection [23,40] and breadth-first algorithms [41] have been considered. The breadth-first algorithms are especially promising [42–44]. The improved decision-feedback detector in [25] is essentially the M-algorithm applied over all users in a given time slot [45]. Xie et al. also use the M-algorithm tree-search scheme in [41].

## 2.2 Contributions

This part of the dissertation is a precursor for a new suboptimum multiuser detection algorithm that we propose in [26]. In this chapter the most celebrated multiuser detection algorithms are compared in terms of both the bit-error-rate and the com-

putational complexity. Six different multiuser schemes are deployed as detectors for a 16-user DS-CDMA system which uses Gold sequences of length 31 as the spreading codes in an AWGN channel. Computer simulations are used to obtain the bit-error-rate of the detectors.

## 2.3 Multiuser System Description

A CDMA system with $K$ users is shown in Figure 2.1. All users share the same bandwidth. The signaling interval of each user is $T$ seconds, and the input alphabet is antipodal binary: $\{+1, -1\}$. During the $i$-th signaling interval, the input vector is $\boldsymbol{b}(i) = [b_1(i), b_2(i), \ldots, b_K(i)]^T$, where $b_k(i)$ is the input symbol of the $k$-th user. User $k$ is assigned a unit energy spreading code (or signature waveform, or spreading sequence) $s_k(t)$ which is zero outside $[0, T]$ and $\int_0^T |s(t)|^2 dt = 1$. Pulse amplitude modulation is employed at the transmitter. The baseband signal of the $k$-th user is

$$u_k(t) = \sum_{i=0}^{\infty} b_k(i) s_k(t - iT - \tau_k) \tag{2.1}$$

where $\tau_k$ is the transmission delay. For synchronous CDMA the delay $\tau_k = 0$ for all users while for asynchronous CDMA the delays can be different. Each transmitted signal goes through a channel and is attenuated. The channel attenuation is a complex number which can be different for different users and is denoted by

$$\alpha_k(i) = \sqrt{w_k(i)} e^{j\theta_k(i)} \tag{2.2}$$

where $w_k(i)$ and $\theta_k(i)$ are the received power and phase of the $k$-th user, respectively. At the receiver side the received baseband signal is the noisy sum of all users' signals.

Therefore the baseband received signal can be modeled as:

$$r(t) = \sum_{i=0}^{\infty} \sum_{k=1}^{K} \alpha_k(i) b_k(i) s_k(t - iT - \tau_k) + n(t) \tag{2.3}$$

where $n(t)$ is the complex additive white Gaussian noise (AWGN). According to (2.3), each user's signal travels along a single path, so this model does not illustrate multipath propagation. The effect of multipath is discussed in Chapter 5. For the

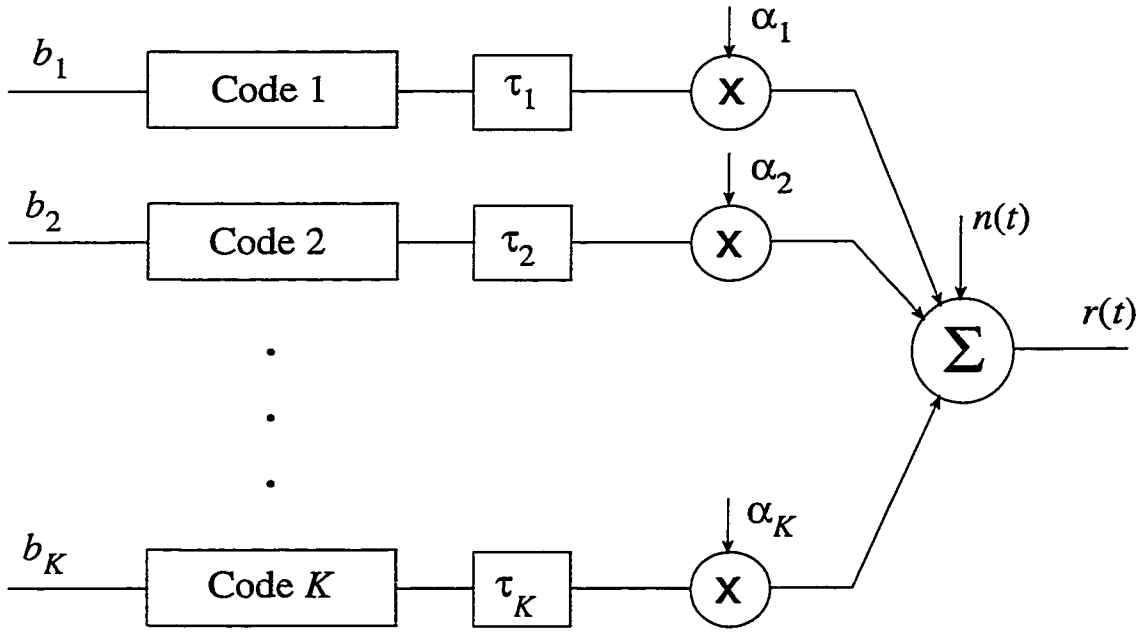


**Figure 2.1**: CDMA channel model.

rest of this section, we will consider a very simplified DS-CDMA system. A number of simplifications will be exposed in the rest of the dissertation. In fact, each relaxation of simplification will represent another factor to consider for multiuser detection system. The simplifying assumptions are as follows:

- *We consider a channel with real attenuation.* The real model is convenient for analyzing coherent methods, and can be easily generalized to the complex

case. In Chapter 5, we extend our treatment to multiuser detectors for fading channels, where complex attenuation need to be considered.

- *Multiuser detectors derivation is presented for synchronous CDMA systems.* The synchronous assumption considerably simplifies exposition and analysis and often permits the derivation of closed-form expressions for the desired performance measures. These are useful since similar trends are found in the analysis of the more complex asynchronous case. Furthermore, every asynchronous system can be viewed as an equivalent synchronous system with larger effective user population [18], which is often explored in burst CDMA communication. Moreover, synchronous systems are becoming more of practical interest since quasi-synchronous approaches have been proposed for satellite [46] and microcell applications [47]. It should be recognized, however, that the transition from synchronous to asynchronous can considerably increase the complexity of multiuser detection. Throughout the dissertation, we consider only synchronous transmissions.

- *Certain parameters are assumed to be known perfectly.* The multiuser detectors presented in this chapter and Chapters 3 and 5 take advantage of completely known channel parameters so that amplitudes , phases and delays do not appear in treatment at all. Chapters 6 and 7 propose some methods to estimate unknown channel parameters.

## 2.4 Multiuser Detection Concepts

The first step in the multiuser detection process is to pass the received signal $r(t)$ through a bank of matched filters (or correlators). The bank of matched filters

consists of $K$ filters matched to individual spreading codes of the system followed by samplers at instances $iT + \tau_k$, $(k = 1, \ldots, K)$, $(i = 1, 2, \ldots)$ (Figure 2.2). Recall that for a synchronous system $\tau_k = 0$. Van Etten [48] and Schneider [32] showed that the outputs of the matched filters form a set of sufficient statistics for demodulating the input sequence $b(i)$ from the given received signal $r(t)$.

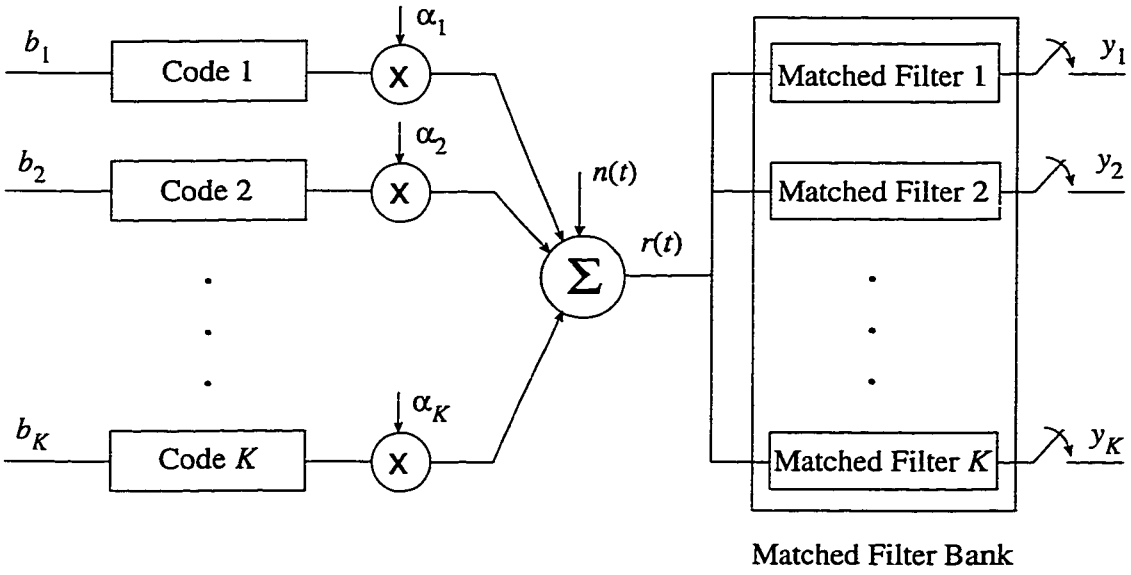


**Figure 2.2**: CDMA multiuser detection concept.

For synchronous CDMA, the received signal $r(t)$ for $iT \leq t \leq (i+1)T$ does not depend on the inputs of the users sent during past or future time intervals. Consequently, assuming that all possible information sequences are equally likely, it is sufficient to consider a one-shot system with input vector $\boldsymbol{b} = [b_1, b_2, \cdots, b_K]^T$, real positive channel attenuation $\alpha_1 = \sqrt{w_1}, \cdots, \alpha_K = \sqrt{w_K}$ and additive white Gaussian noise $n(t)$ with power spectral density $\frac{N_0}{2}$. The output of the $k$-th matched

filter (which is matched to the spreading code of user $k$) at the sampling time is

$$\begin{aligned} y_k &= \int_0^T r(t)s_k(t)dt \\ &= \int_0^T s_k(t)\Big[\sum_{i=1}^{K} \alpha_i b_i s_i(t) + n(t)\Big]dt \\ &= \alpha_k b_k + \sum_{\substack{i=1 \\ i\neq k}}^{K} \alpha_i b_i R_{k,i} + \int_0^T s_k(t)n(t)dt \end{aligned} \tag{2.4}$$

where:

$$R_{k,i} = \int_0^T s_k(t)s_i(t)dt, \qquad k,i = 1,\cdots,K \tag{2.5}$$

is the cross-correlation between the spreading codes assigned to user $k$ and user $i$. Note that $y_k$ consists of three terms. The first is the desired information which gives the sign of the information bit $b_k$ (which is exactly what is sought). The second term is the result of the multiple access interference (MAI), and the last is due to noise. The second term typically dominates the noise so that one would like to remove it. Its influence is felt through the cross-correlation between the chip sequences and the powers of users. If one knew the cross-correlation and the powers, then one could attempt to cancel the effects of one user upon another. This is, in fact, the intuitive motivation behind any interference cancellation scheme.

## 2.5 Measures of Comparison

The probability of error or bit-error-rate (BER), as a function of the Signal-to-Noise-Ratio (SNR), is a common and essential figure of merit for a communication system, indicating the feasibility of reliable data transfer across the channel. The BER can

be used as a metric to compare different communication systems. Throughout this dissertation we utilize the BER as a measure of performance for the multiuser detection schemes which we study.

As we investigate various multiuser detectors in the following sections, we notice that for most of them finding a closed-form expression for the probability of error is either impossible or very complicated. This problem is more evident especially when the number of users is large. In such a case computer simulations are most commonly used in determining the BER. In the following sections whenever it is easy to obtain a closed-form formula for the probability of error, we will do so. However, for the sake of consistency where we want to compare the studied detectors in terms of their respective probability of error, we use the results of the computer simulations. In the next section we review the model of the computer simulations used in the present and next chapters.

Another criterion that we use to compare multiuser detection schemes is the computational load. The computational load for the various multiuser algorithms to be studied in this chapter are different. One way to quantify the computational load is with the notation of a *flop*. A flop is a floating point operation. Throughout this dissertation, operations such as multiply, add and compare are considered as one flop[1]. The computational load is primarily a function of the number of users, $K$. We use two measures for computational load denoted by $\mathcal{O}(K)$ and $\mathcal{O}_i(K)$. The former denotes total required flops to detect transmitted bits from all users, whereas the latter denotes required flops to detect transmitted bits from user $i$. In a point to multipoint communication link, that uses purely CDMA, receiver $i$ is interested only in detecting $b_i$. Therefore $\mathcal{O}_i(K)$ is a correct measure of the computational

[1]In a conventional all-purpose DSP chip a multiply operation takes the same number of cycles as an add operation.

load for this type of structure. The reason that we have also considered $\mathcal{O}(K)$, is that the multiuser detection idea can be applied to a point to point communication makeup too. In such a case a receiver detects $(b_1, b_2, \cdots, b_K)$. Thus $\mathcal{O}(K)$ will be a correct measure of the computational load. In the computational load calculations we will count only the computational load involved in the multiuser algorithm and ignore the matched filters.

## 2.6 Computer Simulation Model

The simulation model used throughout this chapter and Chapter 3 is shown in Figure 2.3. Binary data are first modulated by BPSK signalling. Then each modulated symbol is spread by spreading a sequence assigned to its respective user. The scaling factor $\sqrt{w_k}$, $(k = 1, \cdots, K)$ represents the energy for each user. We assume all users have the same power. In other words, in the simulations we are not studying the effects of near-far problem on the detectors. The chip rate is set at 8 Mcps, and a bandwidth of 8 MHz is assumed for the system. The receiver observes the sum of the transmitted signals from all users that is embedded in noise. SNR is defined as follow (see Figure 2.3)

$$\mathrm{SNR}_{\mathrm{dB}} = 10 \log_{10} \frac{P_{av}(S(\boldsymbol{b}, t))}{P_{av}(z(t))} \tag{2.6}$$

where $P_{av}(\cdot)$ represents the average power of a signal, and $z(t)$ is the additive white Gaussian noise, which is filtered by an ideal low-pass filter with (two-sided low-pass) bandwidth equal to 8 MHz.

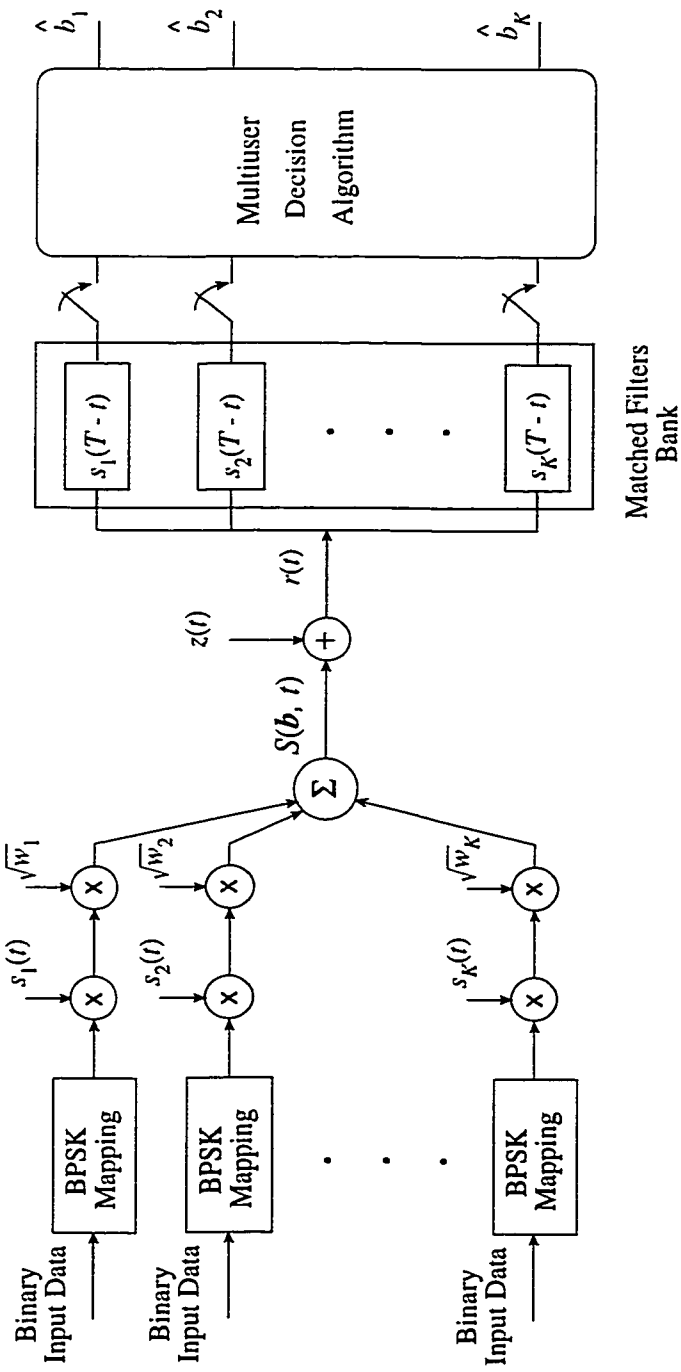


**Figure 2.3**: Simulation platform.

### Number of Users

In the literature mainly small CDMA systems with 2, 4 or at most 6 users have been considered. The reason is that such systems are easy to analyze or simulate. In this dissertation, however, we have selected a relatively large CDMA system with 16 users in order to be able to realistically investigate the feasibility of multiuser detection for CDMA systems.

### Spreading Codes

Throughout this dissertation we have chosen a set of 16 Gold sequences each of length 31 chips as the spreading codes assigned to the 16 users. Appendix A shows the Gold codes as well as their auto-correlation and cross-correlation properties. As shown in the Appendix, the length, $n$, of a Gold sequence is equal to $2^m - 1$, where $m$ is an integer. The total number of Gold sequences with length $n$ is always $n + 2$.

We choose the length of the spreading codes to be 31 in order to accommodate 16 users. Other choices are 15, 63, or higher. The choice of 63 and higher is not bandwidth efficient. The choice of length 15, although adequate for 16 users, does not offer good periodic cross-correlations.

## 2.7 Multiuser Detection Algorithms

In the following sections we briefly discuss several previously proposed multiuser detectors of interest. Each detector is first introduced, then its probability of error is derived and finally its computational load is analyzed. We begin our discussion with the simplest detector, namely, the conventional detector.

## 2.8 Conventional Detector

The conventional detector, which is sometimes referred to as the single-user detector, uses the same approach as the optimal receiver for the single user system. It detects the bit from user $k$ by correlating the received signal with the spreading sequence corresponding to user $k$. Thus, the conventional detector makes a decision at the output of the matched filter bank (Figure 2.4):

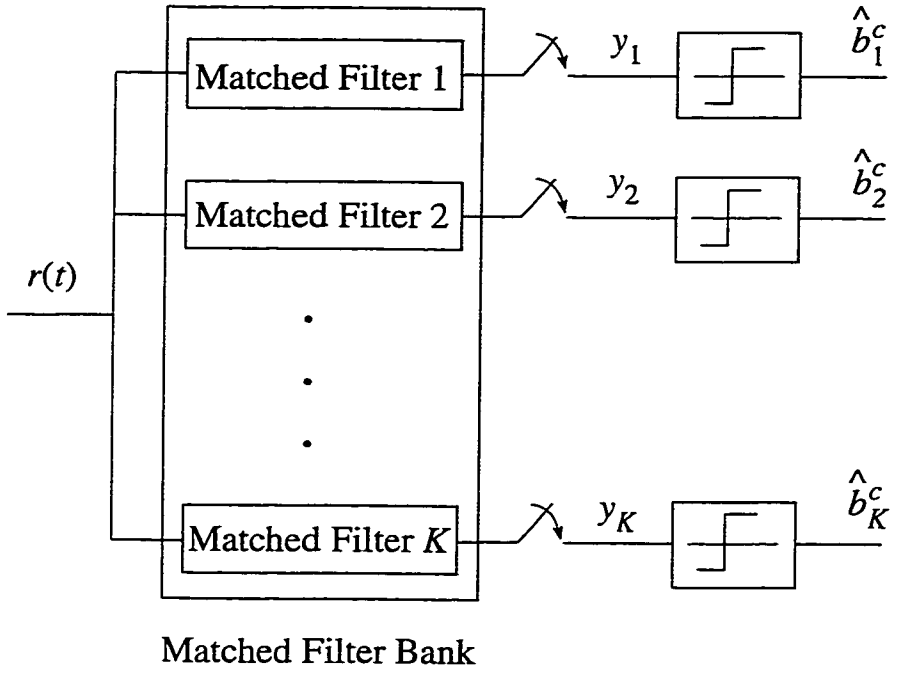


**Figure 2.4**: Conventional (Single-user) detector.

$$\hat{b}_k^c = \text{sgn}(y_k) \tag{2.7}$$

This method ignores MAI and treats it as noise. When MAI terms are significant,

as shown in (2.4), the bit-error-rate for this detector is high. Note that MAI depends both on the cross-correlation and the power of each user.

### 2.8.1 Probability of Error for Conventional Detector

It is quite straightforward to find the $k$-th user probability of error, $P_k^c$, for the conventional detector. Assuming that the $k$-th user transmits information symbol '-1', one should calculate the probability that $y_k$ is positive. If this is the case, the receiver fails to detect the transmitted information correctly.

$$P_k^c = P[y_k > 0 \mid b_k = -1] \tag{2.8}$$

Using the theorem of total probability [49];

$$P_k^c = \sum_{\substack{\boldsymbol{b}\in\{-1,1\}^K \\ b_k=-1}} P[y_k > 0 \mid \boldsymbol{b}]\, P[\boldsymbol{b} \mid b_k = -1] \tag{2.9}$$

Since $n(t)$ is Gaussian, from (2.4), $y_k$ is also a Gaussian random variable with a conditional mean $\mu_{y_k}$ and a variance $\sigma_{y_k}^2$ as follow:

$$\mu_{y_k} = -\alpha_k + \sum_{\substack{i=1 \\ i\neq k}}^{K} \alpha_i b_i R_{k,i} \tag{2.10}$$

$$\sigma_{y_k}^2 = \frac{N_0}{2} \tag{2.11}$$

Therefore

$$P[y_k > 0 \mid \boldsymbol{b}, b_k = -1] = Q\Big(\frac{\alpha_k - \sum_{\substack{i=1 \\ i\neq k}}^{K} \alpha_i b_i R_{k,i}}{\sqrt{N_0/2}}\Big) \tag{2.12}$$

where the $Q$-function is defined as

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\nu^2/2} d\nu$$

On the other hand, since the input alphabet is antipodal binary, i.e. $\boldsymbol{b} \in \{-1, 1\}^K$, the second conditional probability in (2.9) is obtained as

$$P[\boldsymbol{b} \mid b_k = -1] = \frac{1}{2^{K-1}}, \tag{2.13}$$

assuming that $b \in \{-1, 1\}^K$ with equal probability. Finally, by combining the results in (2.9), (2.12) and (2.13), we obtain the $k$-th user probability of error for the conventional detector as

$$P_k^c = \frac{1}{2^{K-1}} \sum_{\substack{b \in \{-1,1\}^K \\ b_k = -1}} Q\Big( \frac{\alpha_k - \sum_{\substack{i=1 \\ i \neq k}}^K \alpha_i b_i R_{k,i}}{\sqrt{N_0/2}} \Big) \tag{2.14}$$

In Figure 2.5 we have plotted the simulation results for the average BER versus SNR for the conventional detector for the DS-CDMA system described in Section 2.6. It is worthwhile to mention that although we have plotted only the average BER for all 16 users, the BER for each individual user is almost the same. This is because the energies, $\alpha_1, \alpha_2, \cdots, \alpha_K$, of all users have been assumed to be equal with equal cross-correlations between sequences (See Appendix A). We observe that a BER of around $10^{-5}$ is achievable even with a conventional detector .

When MAI terms in (2.4) are significant, the bit-error-rate for this detector can be high. This is due to the fact that MAI depends both on the cross-correlations between sequences and the energies , $\alpha_1, \alpha_2, \cdots, \alpha_K$. Suppose there are only two users in the system, i.e. $K = 2$. Let $r$ be the cross-correlation between the spreading

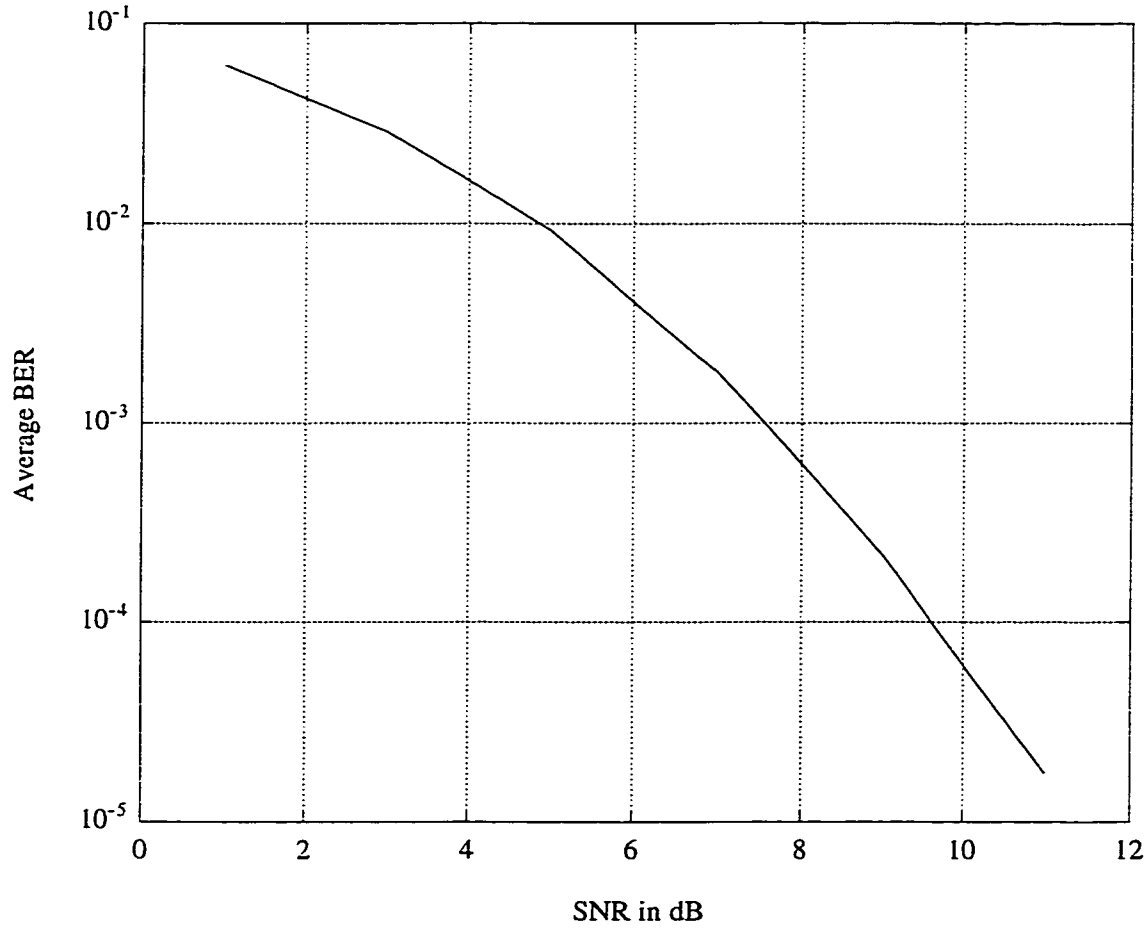


**Figure 2.5**: Average BER of the conventional detector vs. SNR in an AWGN channel.

codes of the two users

$$r = \int_0^T s_1(t)s_2(t)dt \tag{2.15}$$

In this case, the outputs of the matched filters are

$$y_1 = \alpha_1 b_1 + r\alpha_2 b_2 + n_1 \tag{2.16}$$

$$y_2 = \alpha_2 b_2 + r\alpha_1 b_1 + n_2 \tag{2.17}$$

The MAI terms for users 1 and 2 are $r\alpha_2 b_2$ and $r\alpha_1 b_1$, respectively. Without loss of generality, we consider the probability of error for user 1 . Using (2.14)

$$P_1^c = \frac{1}{2}\Big(Q\Big(\frac{\alpha_1 - \alpha_2 r}{\sqrt{N_0/2}}\Big) + Q\Big(\frac{\alpha_1 + \alpha_2 r}{\sqrt{N_0/2}}\Big)\Big) \tag{2.18}$$

Equation (2.18) implies that if the cross-correlation between the spreading codes, $r$, is zero, i.e. the spreading codes are orthogonal, then the probability of error is

$$P_1^c\Big|_{r=0} = Q\Big(\frac{\alpha_1}{\sqrt{N_0/2}}\Big) \tag{2.19}$$

It can be shown that (2.19) is the minimum of (2.18) in terms of $r$ [50]. Moreover (2.19) is exactly the same as the probability of error for the optimal detector for a single user system. Indeed the probability of error of the optimal detector for the single user system serves as a lower bound on the performance of any other detector.

Now suppose $r$ is nonzero and user 1 is much stronger than user 2 (i.e. $\alpha_1 \gg \alpha_2$, which is referred to as the near-far problem), the MAI term $r\alpha_1 b_1$, present in the signal of the second user, is very large, and can significantly degrade the performance of the conventional detector for that user as shown in (2.18). A multiuser detector called a successive interference canceller (decision directed) can remedy this problem as follows. First, a decision $\hat{b}_1$ is made for the stronger user 1 using the conventional detector. Since user 2 is much weaker than user 1, this decision is reliable from the point of view of user 2. So, this decision can be used to subtract the estimate of MAI from the signal of the weaker user. The decision for user 2 is given by

$$\hat{b}_2 = \mathrm{sgn}(y_2 - r\alpha_1 b_1) \tag{2.20}$$

$$= \mathrm{sgn}(\alpha_2 b_2 + r\alpha_1(b_1 - \hat{b}_1) + n_2) \tag{2.21}$$

Provided the decision of the first user is correct, all MAI can be subtracted from the signal of user 2. If we fix the energy of the second user, and let the energy of the first user grow, the error rate of the successive interference canceller for the second user will approach the single-user bound. Thus, this detector is successful in combating the near-far problem. This simple example motivates the use of multiuser detectors for CDMA channels.

### 2.8.2 Computational Load of Conventional Detector

According to the definition given for $\mathcal{O}(K)$ and $\mathcal{O}_i(K)$ in Section 2.5, it is easy to find the computational load for the conventional detector. By investigating the conventional detector in Figure 2.4, we find out easily that the detection of $b_i$ involves 1 flop (1 compare), whereas the detection of all $b_i$s involves $K$ flops. In other words

$$\begin{aligned} \mathcal{O}(K) &= K \\ \mathcal{O}_i(K) &= 1 \end{aligned} \tag{2.22}$$

## 2.9 Optimum Detector

The optimum multiuser detector, which is proposed by Verdú in [37], is defined as a detector that selects the set of symbols corresponding to that signal among the possible ones which resembles most closely, in the mean-square sense, the received signal. If the noise is Gaussian and white, then this rule is optimum in the maximum-likelihood sense. Furthermore, if all vectors $\boldsymbol{b}$ are a priori equiprobable, then the minimum distance rule gives the maximum-a-posteriori (MAP) decision.

The optimum multiuser detector, thus can be expressed mathematically as

$$\hat{\boldsymbol{b}}^o = \arg \min_{\boldsymbol{b}\in\{-1,+1\}^K} \|r(t) - \sum_{i=1}^{K} \alpha_i b_i s_i(t)\| \tag{2.23}$$

As it was mentioned earlier we are considering a one-shot system, hence (2.23) may be rewritten as

$$\begin{aligned}\hat{\boldsymbol{b}}^o &= \arg \min_{\boldsymbol{b}\in\{-1,+1\}^K} \int_0^T \Big[r(t) - \sum_{i=1}^{K} \alpha_i b_i s_i(t)\Big]^2 dt \\ &= \arg \min_{\boldsymbol{b}\in\{-1,+1\}^K} \Big[\int_0^T \Big(r(t)\Big)^2 dt - 2\sum_{i=1}^{K} \alpha_i b_i \int_0^T r(t)s_i(t)dt \\ &\quad + \sum_{i=1}^{K}\sum_{j=1}^{K} \alpha_i b_i \alpha_j b_j \int_0^T s_i(t)s_j(t)dt\Big]\end{aligned} \tag{2.24}$$

or equivalently in matrix notation as

$$\hat{\boldsymbol{b}}^o = \arg \max_{\boldsymbol{b}\in\{-1,+1\}^K} (2\boldsymbol{b}^T W\boldsymbol{y} - \boldsymbol{b}^T WRW\boldsymbol{b}) \tag{2.25}$$

where $W$ is a diagonal matrix whose diagonal entries are the energies of the users,

$$W = \begin{pmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \alpha_K \end{pmatrix} = \begin{pmatrix} \sqrt{w_1} & 0 & \cdots & 0 \\ 0 & \sqrt{w_2} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \sqrt{w_K} \end{pmatrix}, \tag{2.26}$$

and $\boldsymbol{y}$ is the vector output of the matched filters, $\boldsymbol{y} = [y_1, y_2, \cdots, y_K]^T$.

Equation (2.25) dictates an exhaustive search over the $2^K$ possible combinations of the components of the bit vector $\boldsymbol{b}$. Verdú showed that the above problem is NP-hard [51,52], i.e. there is no algorithm to solve (2.24) in polynomial time in the

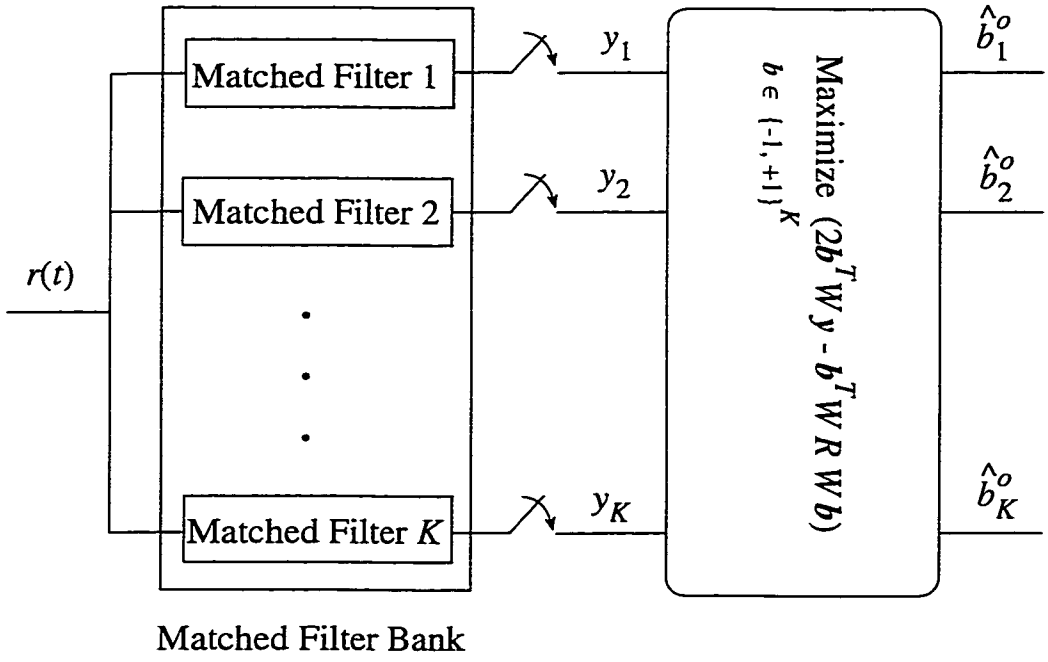


**Figure 2.6**: Optimum multiuser detector.

number of users $K$. Indeed the computational complexity of the optimum multiuser detector grows exponentially in the number of users. Since a CDMA system could potentially have a large number of users, this solution may prove, in a number of situations, to be impractical and too expensive to implement. Furthermore, most CDMA systems with a current need for only a small number of users may require an expansion capability without a steep increase in computational requirements.

### 2.9.1 Probability of Error for Optimum Detector

The derivation of a closed-form formula for the optimum detector is not quite as straightforward and easy as what it was for the conventional detector. An ap-

proach that yields the probability of error of the optimum detector can be found in Appendix B.

In this section we present the computer simulation results. Figure 2.7 shows the average BER of the optimum detector versus SNR. We observe that by using

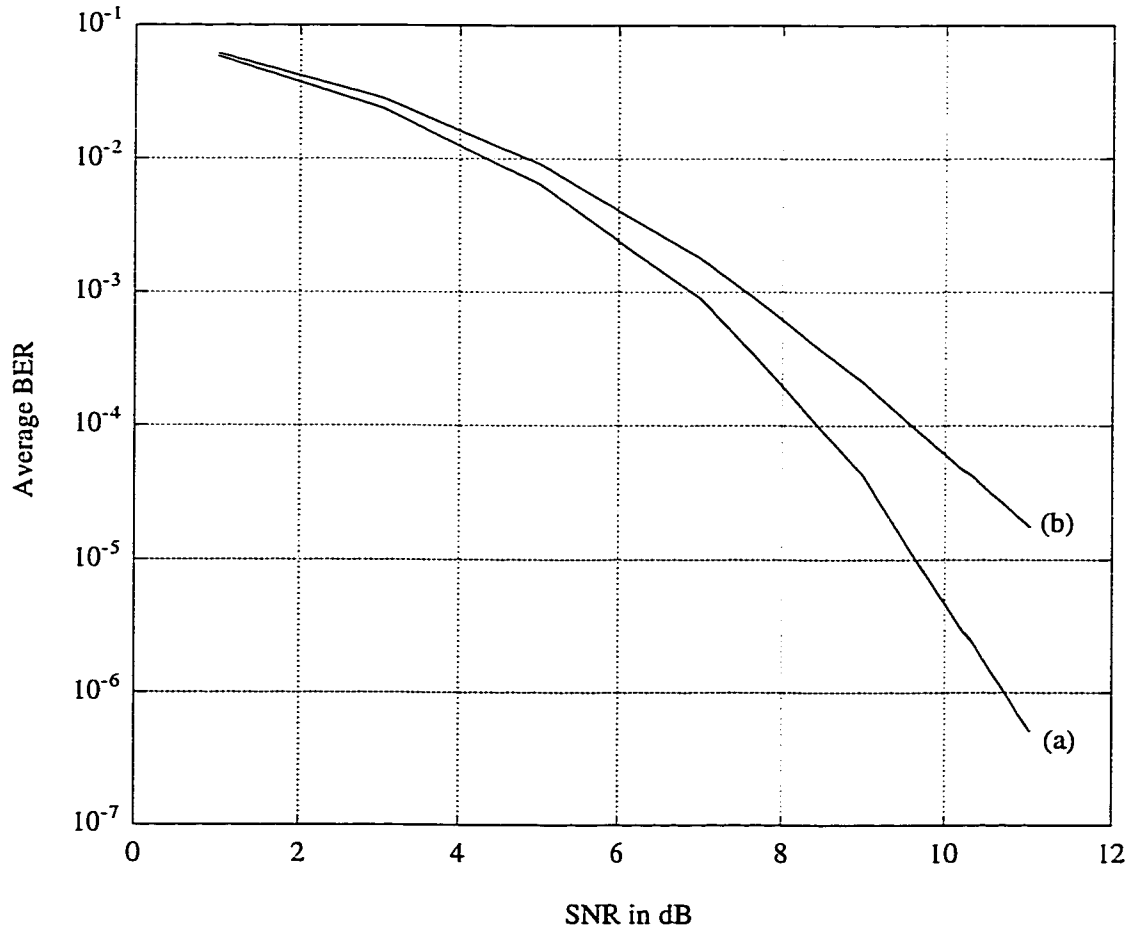


**Figure 2.7**: Comparison of the average bit-error-rates of the optimum and conventional detectors in an AWGN channel; (a) optimum, (b) conventional.

the optimum detector the BER becomes steeper and improves by almost 2 dB at high SNRs. However, as we will see in the next section, the optimum detector has a prohibitive computational complexity. In order to show that the other suboptimum detectors are near-optimum, we have simulated the optimum detector and

calculated its performance. Indeed, the optimum BER is used as a benchmark to compare suboptimum detectors. Any claim to near-optimality of a suboptimal detector will have to be based on how closely the BERs of the suboptimum schemes track optimum performance.

### 2.9.2 Computational Load of Optimum Detector

The optimum detector finds vector $\boldsymbol{b}$ that maximizes (2.25). This problem is known to be NP-hard. The straight forward method to solve this problem is an exhaustive search over $2^K$ possible candidates for $\boldsymbol{b}$. In this case, it is easy to show that the computation of the maximizing expression in (2.25) requires $2K^2 + 3K$ flops. We also notice that the optimum detection scheme finds $\hat{\boldsymbol{b}}^o$ at once, therefore $\mathcal{O}_i(K)$ and $\mathcal{O}(K)$ are the same for this detector. Hence,

$$\mathcal{O}(K) = \mathcal{O}_i(K) = (2K^2 + 3K) \cdot 2^K \tag{2.27}$$

From the above one can see that the computational load of the optimum multiuser detector grows exponentially with respect to $K$. In fact this huge computational load made researchers suggest other multiuser detectors which are categorized as suboptimum detectors. These detectors have good performances fairly close to that of the optimum detector while their computational complexity are reasonably low such that it makes them feasible. In the rest of this chapter we investigate some of the most celebrated suboptimal multiuser detectors.

## 2.10 Decorrelating Detector

As a step towards the most general formulation, consider the matrix version of the equivalent discrete time model (2.4). The output vector $\boldsymbol{y}$ can be expressed as

$$\boldsymbol{y} = RW\boldsymbol{b} + \boldsymbol{n} \tag{2.28}$$

where $R$ and $W$ are $K \times K$ matrices as defined in (2.5) and (2.26), respectively, and $\boldsymbol{n}$ is a colored Gaussian noise vector with the $i$-th element $n_i$ obtained as

$$n_i = \int_0^T s_i(t)n(t)dt \tag{2.29}$$

Inspection of (2.28) suggests a method to solve for $\boldsymbol{b}$, whose components $b_i$ contain the bit information sought. If $\boldsymbol{n}$ was identically zero, we would have a linear system of equations, $\boldsymbol{y} = RW\boldsymbol{b}$, the solution of which can be obtained by inverting $R$ (it is invertible in most cases of interest [17,53]). With a non-zero noise vector $\boldsymbol{n}$, inverting $R$ is still an effective procedure and actually optimal in certain circumstances, to be discussed later. This results in

$$\hat{\boldsymbol{y}} = R^{-1}\boldsymbol{y} = W\boldsymbol{b} + \boldsymbol{\eta} \tag{2.30}$$

where the information vector $\boldsymbol{b}$ is recovered but contaminated by a new noise term. From (2.30), the signal of the $i$-th user is

$$\hat{y}_i = \alpha_i b_i + \eta_i. \tag{2.31}$$

The decision is

$$\hat{b}_i^d = \text{sgn}(\hat{y}_i). \tag{2.32}$$

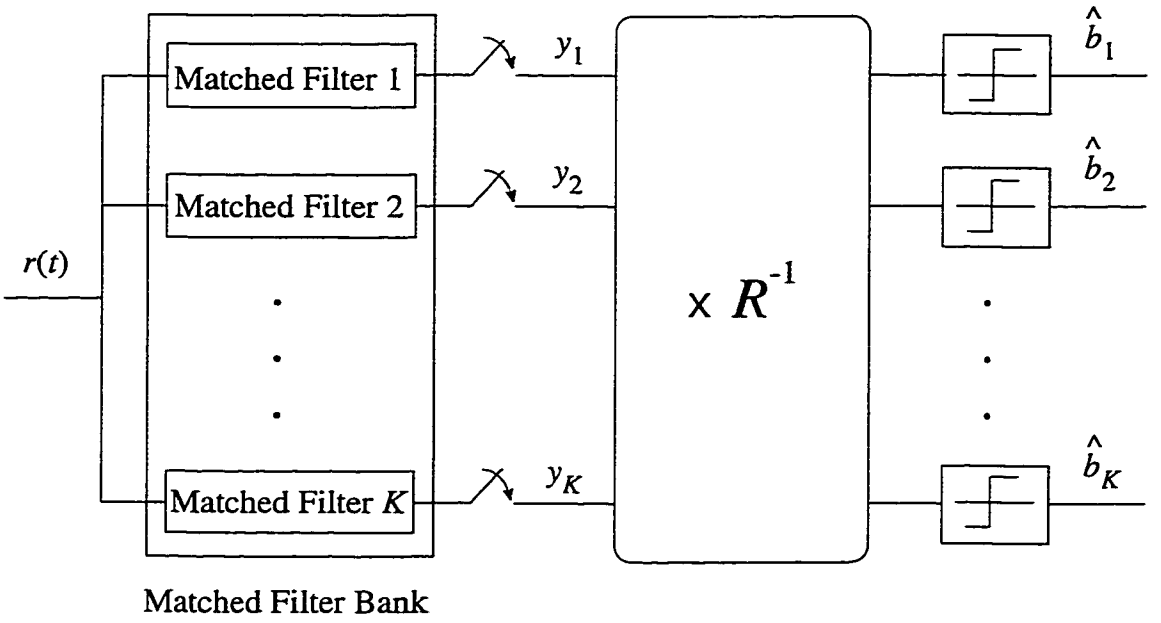


**Figure 2.8**: Decorrelating multiuser detector.

Note that the decorrelating detector completely eliminates MAI. However, the power of the noise $\eta_i$ is $\frac{N_0}{2}(R^{-1})_{i,i}$, which is greater than the noise power $\frac{N_0}{2}$ at the output of the matched filter (2.28). For example, for the two-user system with the cross-correlation $r$ (2.15), the noise power at the output of the decorrelating filter is $N_0/2(1-r^2)$. As it is proved in the next section, the error rate of the decorrelator

is given by

$$P_k^d = Q\Big(\sqrt{\frac{w_k}{(R^{-1})_{k,k}N_0/2}}\Big)$$

The performance of the decorrelating detector degrades as the cross-correlations between users increase.

The decorrelating detector has several desirable features. It does not require the knowledge of the users' energies, and thus its performance is independent of the energies of the interfering users. This can be seen from (2.31). The only requirement is the knowledge of timing which is anyway necessary for the code despreading at the centralized receiver. Observe that neither signal nor noise terms depend on the energies of interferers. In addition, when users' energies are not known, and the objective is to optimize performance for the worst case MAI scenario, the decorrelator is the optimal approach [17,38]. Moreover, the noncoherent version of the decorrelator has been developed [54]. These properties of the decorrelator make it very well suited for the near-far environment.

Multiuser detection is closely related to equalization for intersymbol interference (ISI) channels [55]. For example, the decorrelating detector is analogous to the zero-forcing equalizer. Similarly, the minimum mean-square-error linear multiuser detector [24] (also given by a matrix inverse) is the multidimensional version of the MMSE linear equalizer for the single-user ISI channel. The linear structure of these detectors often limits their performance. In the following sections, we will describe several non-linear approaches to multiuser detection. But first let us derive the probability of error and the computational load of the decorrelating detector.

### 2.10.1 Probability of Error for Decorrelating Detector

In this section we consider the bit error probability of the decorrelating detector. Assume that the $k$-th user transmits information symbol '-1'. We should calculate the probability that $\hat{y}_k$ is greater than zero, i.e. the case when the receiver fails to correctly detect the transmitted information,

$$P_k^d = P[\hat{y}_k > 0 \mid b_k = -1] \tag{2.33}$$

From (2.31) we note that provided that $b_k = -1$, $\hat{y}_k$ is a Gaussian random variable with conditional mean $\mu_{\hat{y}_k}$ and variance $\sigma^2_{\hat{y}_k}$ as follow;

$$\mu_{\hat{y}_k} = -\alpha_k = \sqrt{w_k} \tag{2.34}$$

$$\sigma^2_{\hat{y}_k} = (R^{-1})_{k,k} N_0/2 \tag{2.35}$$

and hence

$$P_k^d = \mathcal{Q}\Big(\sqrt{\frac{w_k}{(R^{-1})_{k,k} N_0/2}}\Big) \tag{2.36}$$

Although (2.36) gives a closed-form equation for the probability of error of the decorrelating detector, in order to be consistent with other detectors we use simulation results to portray the decorrelating detector BER. Figure 2.9 shows the decorrelating detector average BER as well as those of the conventional detector and the optimum detector. As expected, the performance of the decorrelating detector lies between the optimum detector and the conventional one. We observe indeed that using a decorrelating detector yields a satisfactory performance that is close to the optimum to a considerable extent. Of course this is achieved at

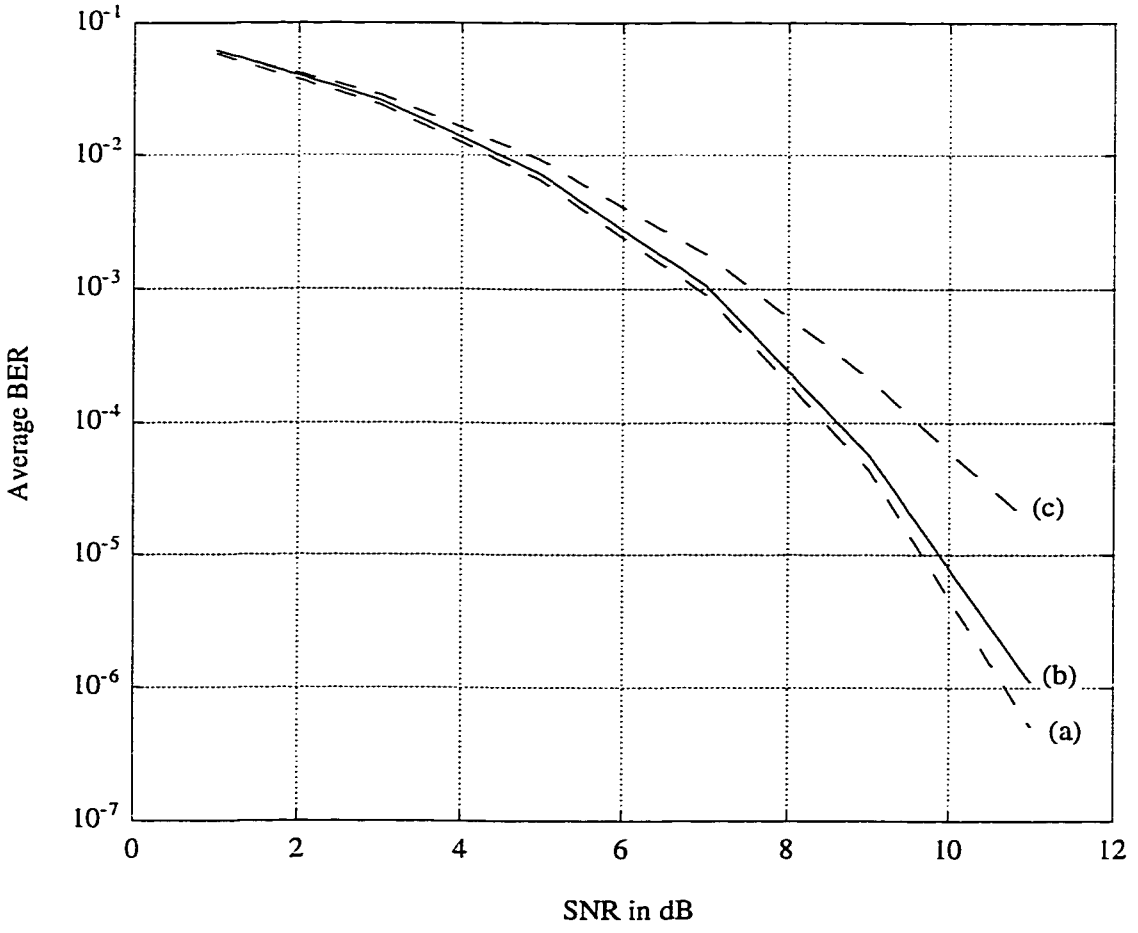


**Figure 2.9**: Comparison of bit-error-rates in an AWGN channel; (a) optimum, (b) decorrelating, (c) conventional.

the expense of a little increase in complexity. The computational complexity of the decorrelating detector is studied and compared with those of the optimum and conventional detectors in the next section.

### 2.10.2 Computational Load of Decorrelating Detector

From (2.30) and (2.32) it is seen that the decorrelating detector requires a matrix-vector multiplication followed by threshold devices, which are implemented using compare modules. It is easy to show that the total calculations involve $2K^2$ flops,

i.e.

$$\mathcal{O}(K) = 2K^2 \tag{2.37}$$

Computing $\mathcal{O}_i(K)$, on the other hand, involves a vector-vector multiplication and a compare that yields,

$$\mathcal{O}_i(K) = 2K \tag{2.38}$$

Figure 2.10 compares $\mathcal{O}(K)$ and $\mathcal{O}_i(K)$ of the decorrelating detector with those of the conventional detector and the optimum detector. The huge gap between the computational load of the decorrelating detector and that of the optimum detector is evident.

## 2.11 Decision-Feedback Detector

The decision-feedback detector, which was proposed in [21, 22, 56], is indeed multiuser decision-feedback equalization, characterized by two matrix transformations: a forward filter and a feedback filter. This detector is analogous to the decision-feedback equalizers employed in single user ISI channels [55]. However, in addition to equalization, the decision-feedback multiuser detector employs successive cancellation. In each time frame, decisions are made in the order of decreasing user's strength, i.e., the stronger user make decisions first, allowing the weaker users to utilize these decisions. The sorting is performed by any multiuser detector with successive MAI cancellation. We will explain the rationale for using this particular order later.

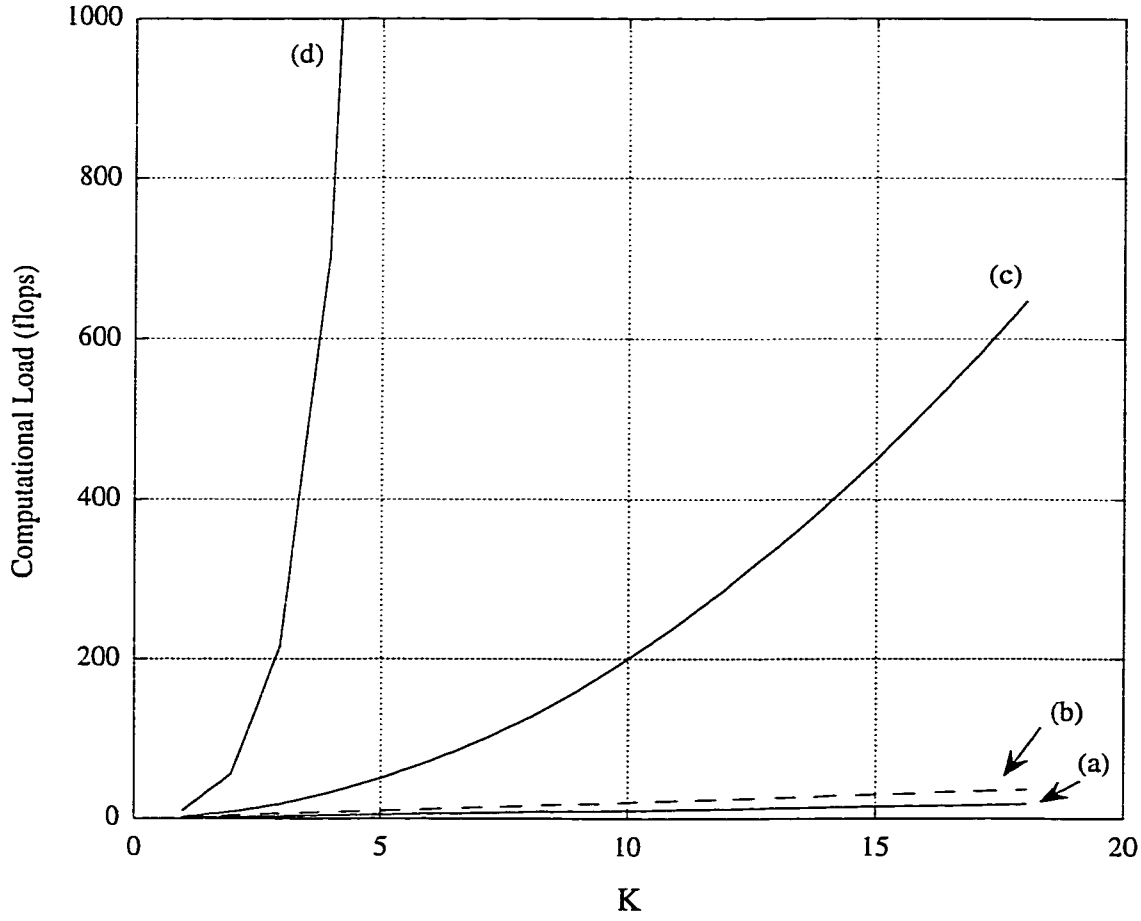


**Figure 2.10**: Computational loads of multiuser detectors; (a) $\mathcal{O}(K)$ of conventional, (b) $\mathcal{O}_i(K)$ of decorrelating, (c) $\mathcal{O}(K)$ of decorrelating, (d) $\mathcal{O}_i(K)$ and $\mathcal{O}(K)$ of optimum.

In synchronous CDMA, a white noise model can be obtained by factorizing the positive definite cross-correlation matrix as $R = F^T F$, where $F$ is a lower triangular matrix (see Cholesky decomposition algorithm [57, 58]). If the filter with response $(F^T)^{-1}$ is applied to the sampled output of the matched filters in (2.28), the resulting output vector is

$$\check{\boldsymbol{y}} = FW\boldsymbol{b} + \check{\boldsymbol{\eta}} \tag{2.39}$$

where $\check{\boldsymbol{\eta}}$ is a white Gaussian noise vector with the autocorrelation matrix $\mathcal{R}(\check{\boldsymbol{\eta}}) = \frac{N_0}{2}\mathcal{I}_K$ ($\mathcal{I}_K$ is the $K \times K$ identity matrix.). The discrete-time models (2.39) and (2.28) correspond to the outputs of the standard and whitened matched filters, respectively, in single user channels with ISI [55].

Since the components of the noise vector $\check{\boldsymbol{\eta}}$ in (2.39) are uncorrelated, the optimum (maximum likelihood) detector for synchronous CDMA has the Euclidean metric $\|\check{\boldsymbol{y}} - \hat{\boldsymbol{t}}\|^2 = \sum_{k=1}^{K}(\check{y}_k - \hat{t})^2$, where $\hat{\boldsymbol{t}}$ is the signal associated with an input $\hat{\boldsymbol{b}}$, i.e., $\hat{\boldsymbol{t}} = FW\hat{\boldsymbol{b}}$, [20]. Both the metric and the expression for the probability of error of the optimal detector have a simpler derivation when the model (2.39) is used instead of (2.28).

The model (2.39) also gives rise to the decorrelating decision-feedback detector (DF). The $k$-th component of $\check{\boldsymbol{y}}$ is given by

$$\check{y}_k = F_{k,k}\sqrt{w_k}b_k + \sum_{i=1}^{k-1} F_{k,i}\sqrt{w_i}b_i + \check{\eta}_k. \tag{2.40}$$

Since this expression does not contain a multiuser interference term for the strongest user ($k = 1$), a decision for this user is made first: $\hat{b}_1 = \mathrm{sgn}(\check{y}_1)$. Multiuser interference for the second user is $F_{2,1}\sqrt{w_1}b_1$. Since a decision for the first user is available, we can use it in a feedback loop to estimate the second user. Thus, the second decision is $\check{y}_2 - F_{2,1}\sqrt{w_1}\hat{b}_1$. Similarly, for the $k$-th user, multiuser interference depends on the stronger users ($i = 1, \cdots, k-1$). Decisions for these users have already been made, and they can be used to form a feedback term (Figure 2.11), i.e.,

$$\begin{aligned}\hat{b}_k^{df} &= \mathrm{sgn}\Big(\check{y}_k - \sum_{i=1}^{k-1} F_{k,i}\sqrt{w_i}\hat{b}_i^{df}\Big) \\ &= \mathrm{sgn}\Big(F_{k,k}\sqrt{w_k}b_k + \sum_{i=1}^{k-1} F_{k,i}\sqrt{w_i}\big(b_i - \hat{b}_i^{df}\big) + \check{\eta}_k\Big)\end{aligned} \tag{2.41}$$

To summarize, the decision-feedback detector is characterized by a feedback filter $\mathcal{B} = (F - F^d)W$, where $F^d$ is a diagonal matrix obtained from $F$ by setting all off-diagonal elements to zero. The filter is fed by the vector of decisions $\hat{\boldsymbol{b}}$. The vector input to the set of decision devices is $\check{\boldsymbol{y}} - \mathcal{B}\hat{\boldsymbol{b}} = F^d W \boldsymbol{b} + (F - F^d)W(\boldsymbol{b} - \hat{\boldsymbol{b}}_{df}) + \check{\boldsymbol{\eta}}$. Since $\mathcal{B}$ is lower triangular with zeros along the diagonal, only previously made decisions (i.e., $\hat{b}_{k-1}^{df}, \hat{b}_{k-2}^{df}, \cdots, \hat{b}_1^{df}$) are required for forming the input to the $k$-th quantizer. The decision-feedback detector corresponds to the zero-forcing decision-feedback equalizer for ISI channels [55], since it attempts to cancel all multiuser interference. The strictly lower triangular $\mathcal{B}$ corresponds to the purely causal feedback filter used in single-user systems.

An important measure of performance for a decision-feedback detector is the signal-to-noise ratio at the input to the decision device under the assumption of correct previous decisions. From (2.41), the signal-to-noise ratio for the decision-feedback is

$$\mathrm{SNR}_k = \frac{F_{k,k}^2 \, w_k}{N_0/2} \tag{2.42}$$

The performance of the decision-feedback detector is similar to that of the decorrelating detector for the strongest user, and gradually approaches the single user bound as the user's power decreases relative to the powers of the interferers. Thus, for the decision-feedback detector, performance advantages with respect to the conventional or the decorrelating detectors are greater for relatively weaker users.

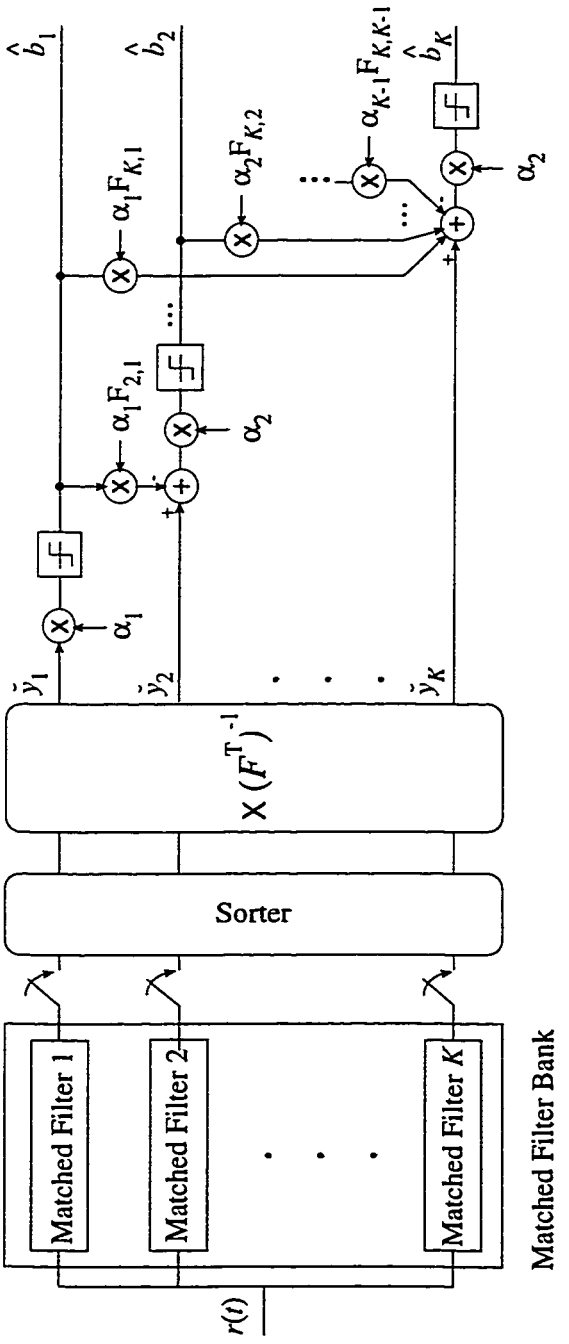


**Figure 2.11**: Decision-feedback multiuser detector.

### 2.11.1 Probability of Error for Decision-Feedback Detector

First, assume that the energies of the users are estimated correctly. The signal-to-noise ratio (2.42) gives rise to the probability of error of the decision-feedback detector for the $k$-th user under the assumption of correct previous decisions

$$\tilde{P}_k^{df} = Q(\frac{F_{k,k}\sqrt{w_k}}{\sqrt{N_0/2}}). \tag{2.43}$$

It is easy to show that $F_{k,k}^2 \geq 1/(R^{-1})_{k,k}$. Note that for the strongest user ($k = 1$), $F_{1,1}^2 = 1/(R^{-1})_{1,1}$ and the estimate (2.43) gives the probability of error since the receiver for this user does not utilize feedback. Therefore, the error rates for the decision-feedback detector (2.43) and the decorrelating detector (2.36) are the same for the strongest user. For $k \geq 2$, the inequality $F_{k,k}^2 \geq 1/(R^{-1})_{k,k}$ is tight provided that multiuser interference affects the $k$-th user. Thus, an improvement over the performance of the decorrelator is suggested by comparison of (2.43) and (2.36). (It is noteworthy to mention that when the decision-feedback method is compared to the decorrelating detector, the decorrelating detector has the following advantage: it does not require the knowledge of energies.) Finally, for the weakest user, $F_{K,K}^2 = 1$, and the ideal performance of the decision-feedback detector (2.43) agrees with the error probability of the single-user system given by

$$P_k^s = Q(\frac{\sqrt{w_k}}{\sqrt{N_0/2}}). \tag{2.44}$$

The above discussion is valid if we assume that the previous decisions are correct. To find the exact error rate of the decision-feedback detector for the $k$-th user, one has to average the conditional error probability given a particular error pattern for

the users $1, \cdots, k-1$ over all such error patterns:

$$P_k^{df} = \frac{1}{2} E_{\Delta b_1, \cdots, \Delta b_{k-1}} Q\left(\frac{F_{k,k}\sqrt{w_k} + \sum_{i=1}^{k-1} F_{k,i}\sqrt{w_i}\Delta b_i}{\sqrt{N_0/2}}\right) \tag{2.45}$$

where the error patten for the $i$-th user is $\Delta b_i = (b_i - \hat{b}_i^{df})$. For a large number of users, exact computation of error probability becomes complex. The problem is similar to computing the exact probability of a decision-feedback equalizer for a single-user system. Although this probability can be bounded analytically, as the author of this detector suggests [21], computer simulations are most commonly used in determining the error rate and the effects of error propagation [55]. In Figure 2.12 we have plotted the average BER of the decision-feedback detector versus SNR. One can see that the performance of the decision-feedback detector is slightly closer to optimum than the decorrelating detector (compare Figure 2.12 with Figure 2.9). This improvement, however, is not exactly as expected from the discussion on page 40. The reason is that in the above qualitative discussion we ignored the error propagation effects on the decision-feedback performance, which is an important degrader factor which is associated to any decision-feedback scheme.

### 2.11.2 Computational Load of Decision-Feedback Detector

As seen in (2.39) the decision-feedback detector primarily applies the whitening filter $(F^T)^{-1}$ to the outputs of the matched filters. Since $(F^T)^{-1}$ is an upper triangular matrix, the total computations for applying the whitening filter involves $K(K+1)/2$ multiplications and $K(K-1)/2$ adds, which are equivalent to $K^2$ flops. Also it can be shown that the needed computational load to calculate (2.41) for all bits involves $K^2$ flops. Therefore the total computational load required to detect $\hat{b}_i^{df}$

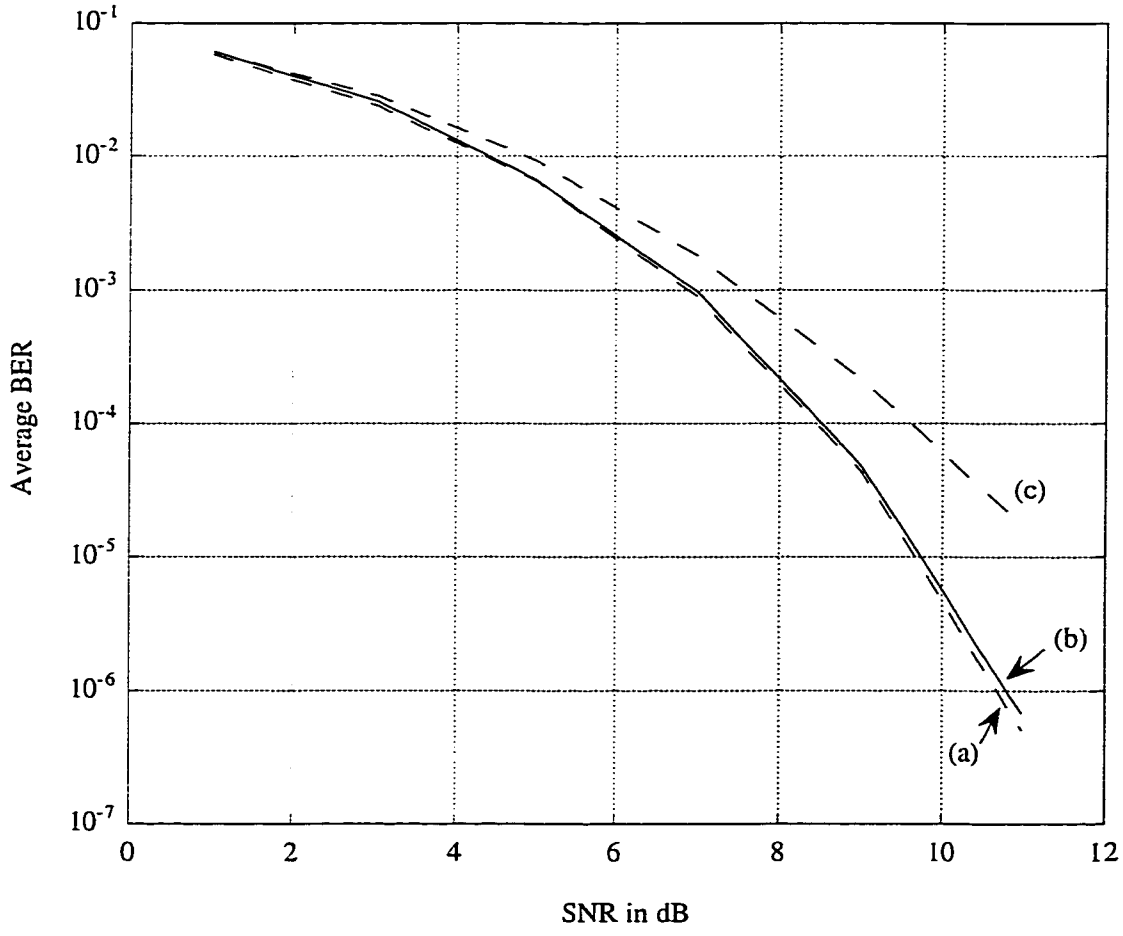


**Figure 2.12**: Comparison of bit-error-rates in an AWGN channel; (a) optimum, (b) decision-feedback, (c) conventional.

$(i = 1, \cdots, K)$ is

$$\mathcal{O}(K) = 2K^2 \tag{2.46}$$

From (2.41)it is apparent that in order to calculate $\hat{b}_i^{df}$ one has to already calculate the previous bits, $\hat{b}_1^{df}, \hat{b}_2^{df}, \cdots, \hat{b}_{i-1}^{df}$. This indicates that the required computational load varies for different users. In other words $\mathcal{O}_i(K)$ depends on $i$ as well as $K$. One can show that the required computational load to calculate $\hat{b}_i^{df}$, without

counting for previous bits, is $2K$. Hence, for user $i$, the computational load is $i$ times as many, i.e.

$$\mathcal{O}_i(K) = 2iK \tag{2.47}$$

Comparing (2.46) with (2.37) we find out that the computational load, $\mathcal{O}(K)$, for the decision-feedback detector is the same as that for the decorrelating detector. However it should be mentioned that when calculating $\mathcal{O}(K)$ for the decision-feedback detector, we assume that the energies of users are known. While in practice the receiver has to estimate the energies, and this requires some computational load as well.

Figure 2.13 depicts $\mathcal{O}(K)$ and $\mathcal{O}_i(K)$ $(i = 1, K)$ of the decision-feedback detector versus $K$. As the reference curves, $\mathcal{O}(K)$ and $\mathcal{O}_i(K)$ of the optimum and conventional detectors are plotted too.

## 2.12 Improved Decision-Feedback Detector

As shown in the previous section, if the whitening filter with response $(F^T)^{-1}$ is applied to the sampled output of the matched filters in (2.28), the resulting output vector is

$$\check{\boldsymbol{y}} = FW\boldsymbol{b} + \check{\boldsymbol{\eta}}$$

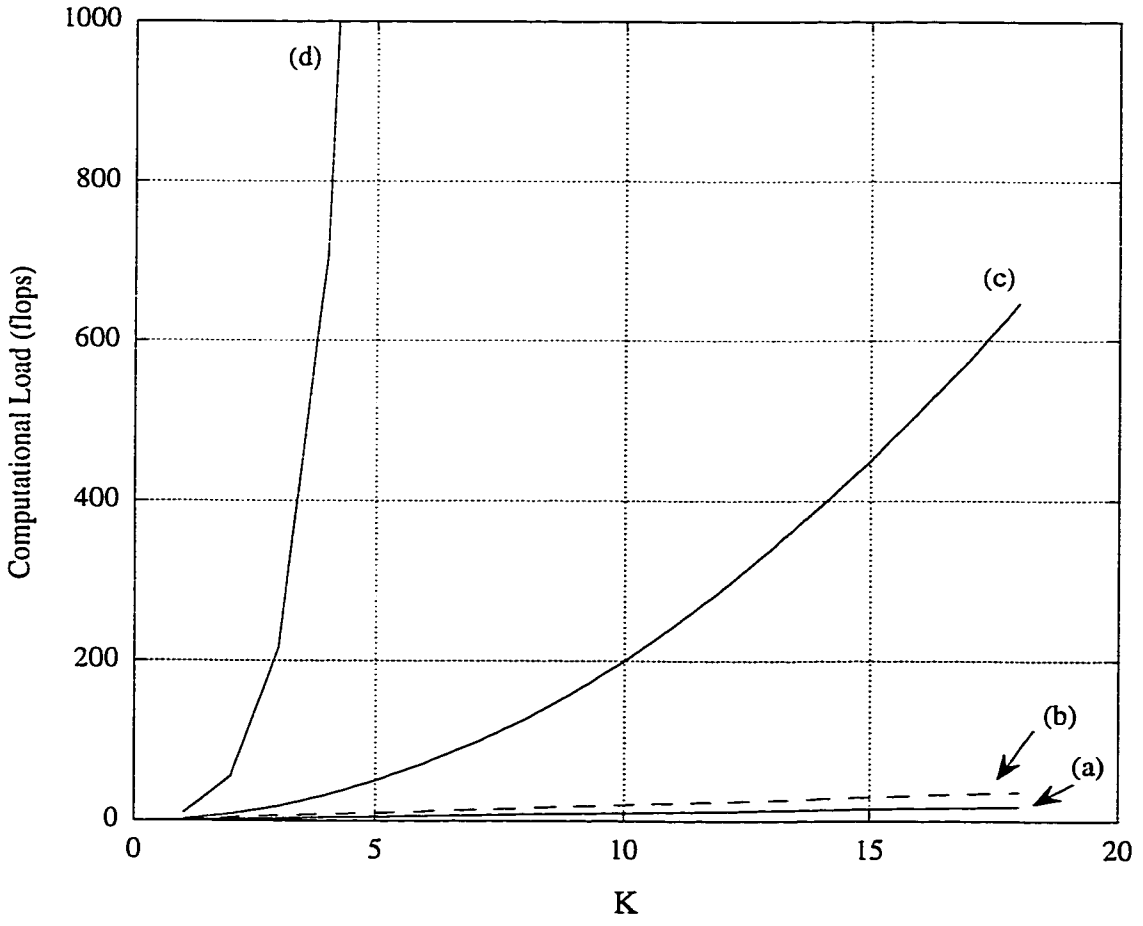


**Figure 2.13**: Computational loads of multiuser detectors; (a) $\mathcal{O}(K)$ of conventional, (b) $\mathcal{O}_1(K)$ of decision-feedback, (c) $\mathcal{O}(K)$ and $\mathcal{O}_K(K)$ of decision-feedback, (d) $\mathcal{O}_i(K)$ and $\mathcal{O}(K)$ of optimum.

The decision rule of the optimum detector is to select from a total of $2^K$ possible vectors the symbol vector $\boldsymbol{b}$ which minimizes the Euclidean metric

$$\|\tilde{\boldsymbol{y}} - FW\boldsymbol{b}\| = \sum_{k=1}^{K}[\tilde{y}_k - \sum_{j=1}^{k} F_{k,j}\sqrt{w_j}b_j]^2. \tag{2.48}$$

On the other hand, the idea behind the decision-feedback detector is that user

$k$ is demodulated by selecting $b_k$ to minimize

$$m'_k = [\tilde{y}_k - F_{k,k}\sqrt{w_k}b_k - \sum_{j=1}^{k-1} F_{k,j}\sqrt{w_j}\hat{b}_j]^2 \tag{2.49}$$

where $\{\hat{b}_1, \hat{b}_2, \cdots, \hat{b}_{k-1}\}$ are the decoded symbols of the previous $k-1$ users. Compared to the optimum detector, there are two differences: a) the decision-feedback detector feeds one vector of decoded symbols to user $k$, while the optimum detector feeds $2^{k-1}$ vectors and their metrics to user $k$, and b) the decision-feedback detector makes a decision for user $k$ based only on current and previous users, while the decision of the optimum detector is based on all users. These two differences cause the disadvantages for the decision-feedback detector, which, however, can be easily overcome if we modify the feedback strategy. This leads to the improved decision-feedback detector [25, 45].

The idea of the improved decision-feedback detector is to feed back $N_m(k)$ partial symbol vectors $\boldsymbol{b}' = [\hat{b}_1, \hat{b}_2, \cdots, \hat{b}_{k-1}]$ and their metrics at stage $k$ where $N_m(k) = \min\{2^k, N_f\}$ and $N_f$ is a maximum fixed value which is usually much less than $2^K$. For example assuming $N_f = 3$, the following steps are performed:

- *Step 1*: Find $N_m(1) = 2$ values of $b_1$ and their corresponding metrics $m'_1$.
- *Step 2*: Feed both these values ($\hat{b}_1 = +1, or -1$) with their metrics to stage 2 and find the $N_m(2) = N_f = 3$ vectors $[\hat{b}_1, \hat{b}_2]$ with the smallest metric sums $m'_1 + m'_2$ for all four possible vectors. Now let $k = 2$.
- *Step 3*: Feed the previous $N_f$ partial symbol vectors $[\hat{b}_1, \hat{b}_2, \cdots, \hat{b}_k]$ and their metrics to the next stage and find the $N_f = 3$ vectors with the smallest metric sums, $m'_1 + m'_2 + m'_{k+1}$, from the $2 \times N_f$ possible partial candidates $[\hat{b}_1, \hat{b}_2, \cdots, \hat{b}_k, b_{k+1}]$.

- *Step 4*: Repeat Step 3 for $k = 3, 4, \cdots, K-1$.
- *Step 5*: Select the symbol vector $[\hat{b}_1, \hat{b}_2, \cdots, \hat{b}_K]$ with smallest total metric sum, $m'_1 + m'_2 + \cdots + m'_K$ as the decoded vector.

When $N_f = 2^{K-1}$, the improved decision-feedback detector is the optimum detector, and when $N_f = 1$, then it is reduced to the decision-feedback detector. The complexity of the improved decision-feedback detector is $2 \times \sum_{k=1}^{K} N_m(k)$ metric computations. For small $N_f$ this is about $2 \times N_f \times K$ which is linear in the number of users.

### 2.12.1 Probability of Error for Improved Decision-Feedback Detector

The probability of error analysis for the improved decision-feedback detector is so complicated that even the author of this detector did not offer any solution for it [25]. So we are contented by computer simulation results. Figure 2.14 shows the BER of the improved decision-feedback detector, where we have chosen $N_f = 3$. We observe the performance of this detector is so close to that of the optimum detector that they are not distinguishable from each other. Of course this excellent suboptimum performance is achieved at the expense of a considerable increased computational complexity compared to the other suboptimum detectors. Next section will discuss this matter more.

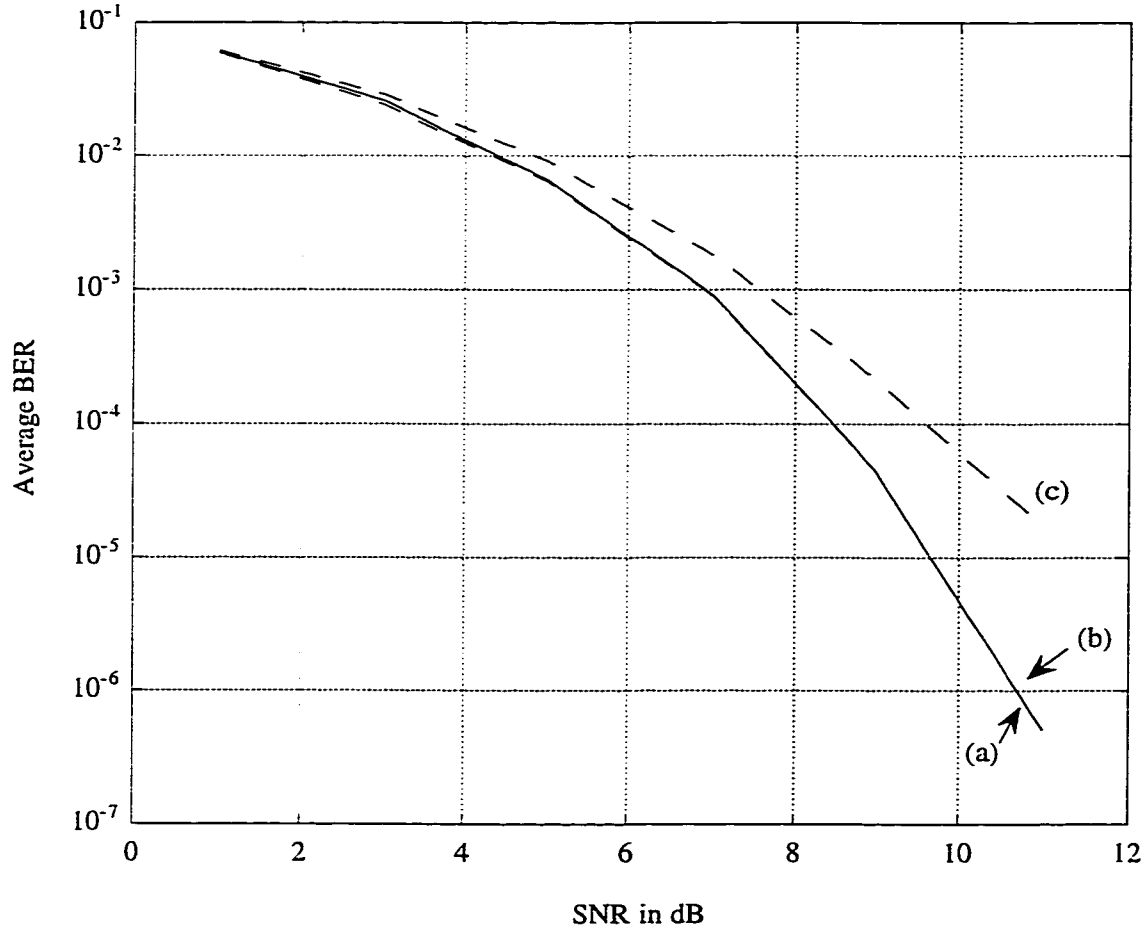


**Figure 2.14**: Comparison of bit-error-rates in an AWGN channel; (a) optimum, (b) improved decision-feedback, (c) conventional.

### 2.12.2 Computational Load of Improved Decision-Feedback Detector

As we mentioned when $N_f$ is chosen to be $2^{K-1}$, the improved decision-feedback detector is the same as the optimum detector. Also, when $N_f$ is chosen to be 1, it is reduced to the decision-feedback detector. The computational load of improved decision-feedback detector depends on $N_f$ as well as $K$. Throughout this dissertation we have assumed $N_f = 3$ in all simulations . Here for the computational load

calculations, we again assume $N_f = 3$. It can be seen that the computational load of the improved decision-feedback detector is $2 \times \sum_{k=1}^{K} N_m(k)$ metric calculations [25]. For $N_f = 3$ we have $\sum_{k=1}^{K} N_m(k) \approx 3K$. The metric computation in (2.49) requires $K^2$ flops. Counting the necessary computational load to apply the whitening filter to matched filters outputs, we will have

$$\mathcal{O}(K) = 6K^3 + K^2 \tag{2.50}$$

and equivalently

$$\mathcal{O}_i(K) = 6K^2 + K \tag{2.51}$$

In Figure 2.15 we have plotted $\mathcal{O}(K)$ and $\mathcal{O}_i(K)$ as functions of $K$.

## 2.13 Multistage Detector

A multistage detector (Figure 2.16), proposed in [19, 20], uses (2.52) instead of (2.25):

$$\hat{b}_k^{mu}(m+1) = \arg \max_{\substack{b_k \in \{-1,+1\} \\ b_l = \hat{b}_l(m), l \neq k}} (2\boldsymbol{b}^T W \boldsymbol{y} - \boldsymbol{b}^T W R W \boldsymbol{b}) \tag{2.52}$$

for $m \geq 1$. It is easily shown that

$$\hat{b}_k^{mu}(m+1) = \text{sgn}[z_k(m)] \tag{2.53}$$

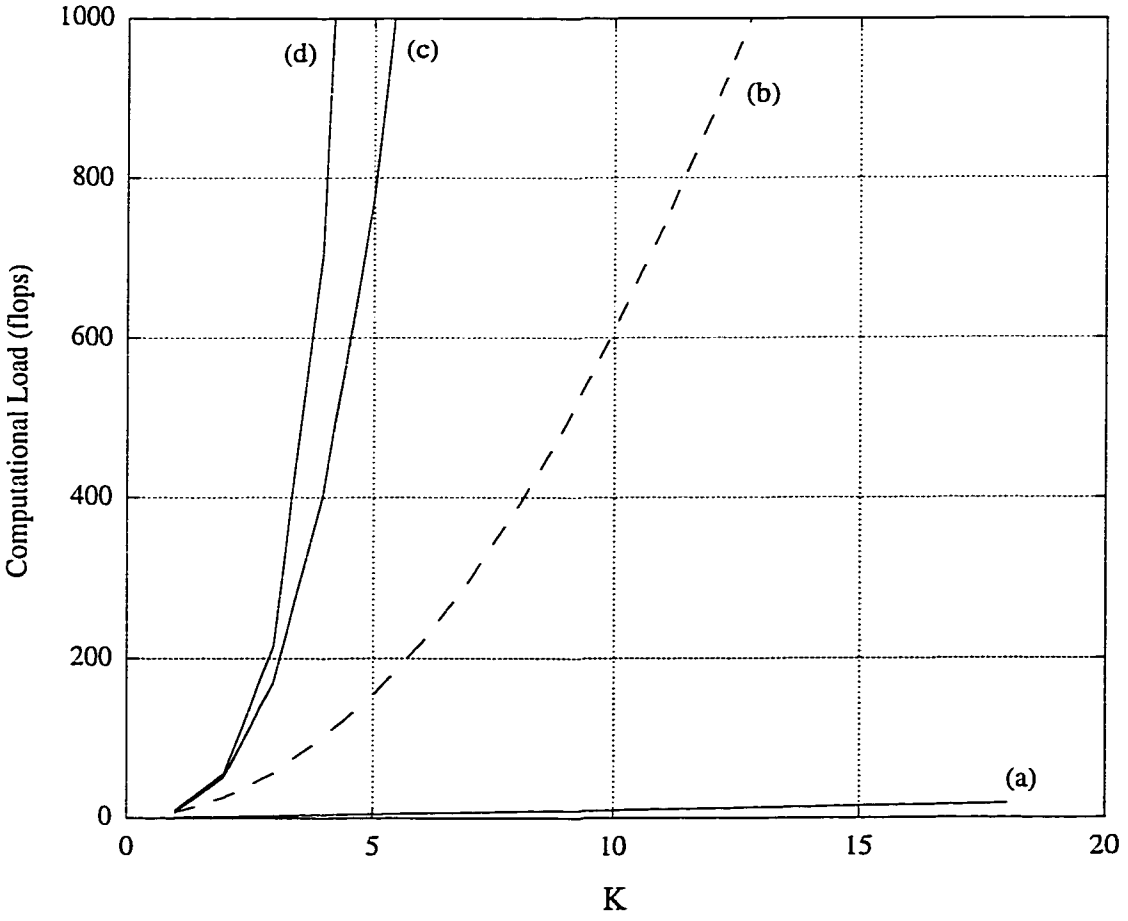


**Figure 2.15**: Computational loads of multiuser detectors; (a) $\mathcal{O}(K)$ of conventional, (b) $\mathcal{O}_i(K)$ of improved decision-feedback, (c) $\mathcal{O}(K)$ of improved decision-feedback, (d) $\mathcal{O}_i(K)$ and $\mathcal{O}(K)$ of optimum.

where $z_k(m)$ is the $m$-th stage statistic for the $k$-th user given as

$$z_k(m) = y_k - \sum_{\substack{j=1 \\ j \neq k}}^{K} \hat{b}_j^{mu}(m) \sqrt{w_j} R_{j,k}. \tag{2.54}$$

In demodulating the information bits of all the users, the maximization of (2.53) is performed for each $k = 1, 2, \cdots, K$. The $(m+1)$-th stage estimate of $\boldsymbol{b}$ can be written as the sign of the $m$-th stage vector of decision statistics $\boldsymbol{z}(m) =$

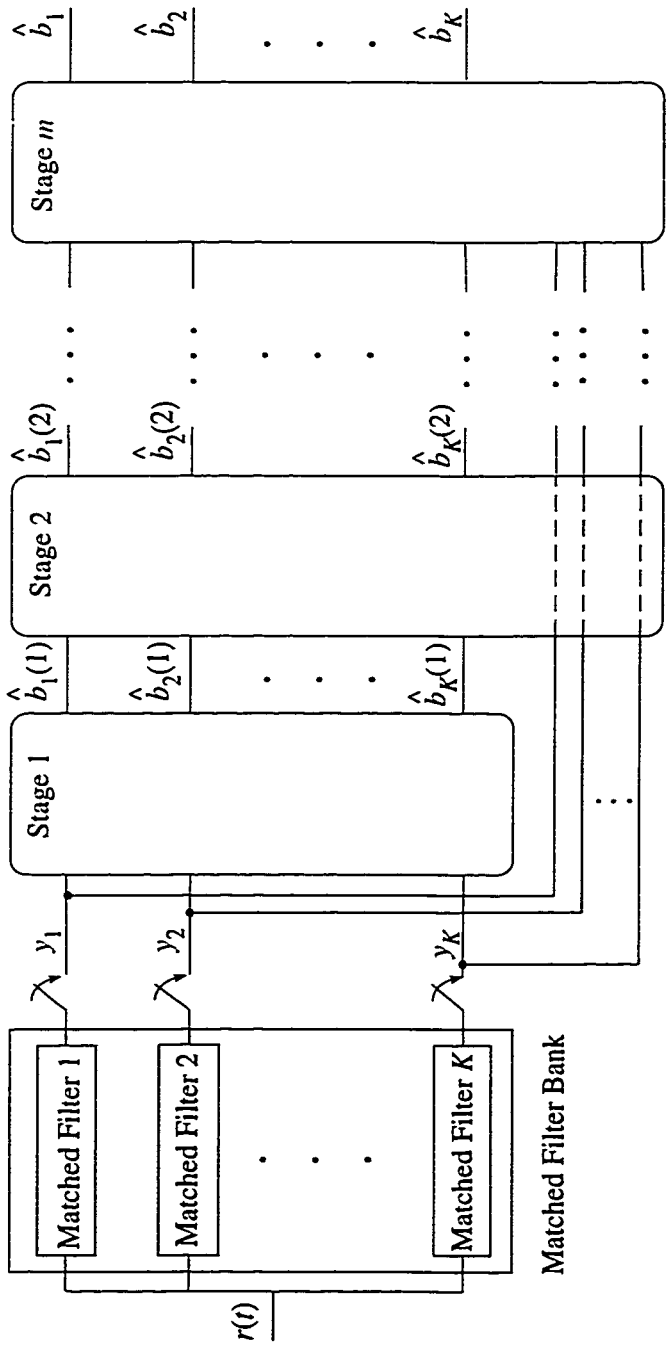


**Figure 2.16**: Multistage multiuser detector.

$[z_1(m), z_2(m), \cdots, z_K(m)]^T$ so that

$$\hat{\boldsymbol{b}}^{mu}(m+1) = \text{sgn}[\boldsymbol{z}(m)] = \text{sgn}[\boldsymbol{y} - (R - \mathcal{I}_K)W\hat{\boldsymbol{b}}^{mu}(m)] \tag{2.55}$$

From Equation (2.28), it is easily shown that

$$\boldsymbol{y} = RW\boldsymbol{b} + \boldsymbol{n} = W\boldsymbol{b} + M_I(\boldsymbol{b}) + \boldsymbol{n} \tag{2.56}$$

where $M_I(\boldsymbol{b}) \triangleq (R - \mathcal{I}_K)W\boldsymbol{b}$ represents the multiple-access interference vector. Substituting (2.56) in (2.55), the expression for the $(m+1)$-th stage estimate of $\boldsymbol{b}$ is given as

$$\begin{aligned} \hat{\boldsymbol{b}}^{mu}(m+1) &= \text{sgn}[\boldsymbol{z}(m)] \\ &= \text{sgn}[W\boldsymbol{b} + M_I(\boldsymbol{b}) - M_I(\hat{\boldsymbol{b}}^{mu}(m)) + \boldsymbol{n}]. \end{aligned} \tag{2.57}$$

The result in (2.57) has a simple interpretation. The $(m+1)$-th stage estimate of $\boldsymbol{b}$ is obtained as the sign of the $m$-th stage statistics which in turn is obtained by subtracting from the sufficient statistic $\boldsymbol{y}$, the estimate of the multiple-access interference based on the $m$-th stage estimate of $\boldsymbol{b}$.

It is noteworthy to mention that Equation (2.52) performs the maximization over one bit at a time, instead of over $K$ bits, as in (2.25). Due to delay constraints, it is desirable to limit the number of stages to two, hence a two-stage detector [20].

### 2.13.1 First Stage For Multistage Detector

The development of the multistage solution described in the last section does not specify the first stage which delivers the initial estimate of the bits $\hat{\boldsymbol{b}}^{mu}(1)$. The choice of the first stage will prove to be important not only in the performance

and the computational load of the multistage detector, but also in simplifying the error probability analysis [20]. For example, consider a two-stage detector with the conventional first stage for the synchronous two-user system with the cross-correlation $r$ (2.15). The decisions produced by the first stage (conventional) detector are $\hat{b}_1(1)$ and $\hat{b}_2(1)$ computed as in (2.7). The decisions of the second stage are $\hat{b}_1(2) = \text{sgn}(y_1 - r\alpha_2\hat{b}_2(1))$ and $\hat{b}_2(2) = \text{sgn}(y_2 - r\alpha_1\hat{b}_1(1))$. The performance of this two-stage detector depends on the relative energies of the users. Clearly, if the first user is stronger than the second, the decisions of the second stage for user 2 agree with those of the decision-directed successive interference canceller, described in the last paragraph of Section 2.8.1. Thus, for the weaker user, the second stage produces more reliable decisions than the first stage. However, for the stronger user, feedback might not be beneficial since the decision produced by the conventional detector for the weaker user is poor. A more reliable two-stage detector results if the conventional detector in the first stage is replaced by the decorrelating detector [20].

### 2.13.2 Probability of Error for Multistage Detector

As mentioned in the previous section, the choice of the first stage for the multistage detector has a twofold effect on the bit-error-rate as well as the computational complexity of this detector. The analysis of the probability of error of the multistage detector is not easy. In Appendix C we have presented the derivation of the error probability of the two-stage detector which employs a decorrelating-type linear detector as the first stage. The conventional detector and the decorrelating detector can be seen as special cases of the decorrelating-type linear detector. As shown in Appendix C, the calculation of error probability requires the evaluation of a $(K-1)$-dimension Gaussian distribution function. The average BER of the

two-stage detector is plotted in Figure 2.17 using a conventional detector as well as a decorrelating detector as the first stage. Both detectors yield BERs very close

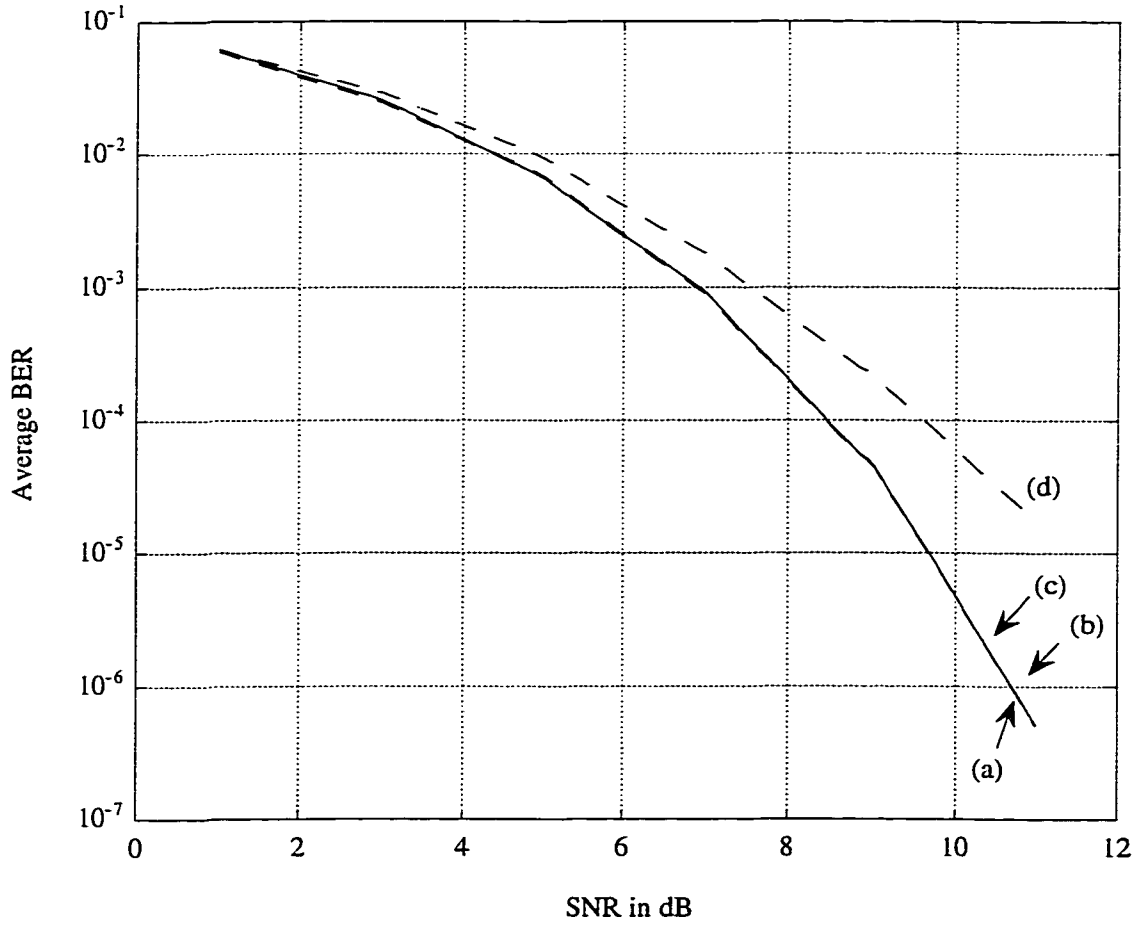


**Figure 2.17**: Comparison of bit-error-rates in an AWGN channel; (a) optimum, (b) two-stage (conventional 1st stage), (c) two-stage (decorrelating 1st stage) (d) conventional.

to optimum, so that the BER curves overlap. But as discussed before, we would expect the decorrelating first stage to outperform the conventional first stage. The fact that the conventional first stage is performing as well as the decorrelating first stage can be explained as follows. The good cross-correlation properties of the spreading codes cause the conventional detector to give acceptable estimates of the

transmitted bits. Therefore the MAI cancellation in the second stage performs well enough to improve the BER near the optimum BER which is the lower bound for all detectors. Thus the difference between the two first stage schemes, i.e. the conventional and the decorrelating, is not remarkable. In Chapter 5, where we study the performance of these two detectors in multipath channels, we find the decorrelating detector as the first stage to be more reliable than the conventional detector.

### 2.13.3 Computational Load of Multistage Detector

The computational load of the multistage detector depends on the choice of the first stage and the number of stages as well as the number of users $K$. Here we consider two different detectors, the conventional detector and the decorrelating detector, as the first stage. Also we assume a multistage detector with two stages. These assumptions are the same as the ones made for simulations purposes.

For a two-stage detector with a conventional detector as stage one, we can show that the detection of all bits involves $2K^2$ flops, i.e.

$$\mathcal{O}(K) = 2K^2 \tag{2.58}$$

Hence,

$$\mathcal{O}_i(K) = 2K \tag{2.59}$$

Similarly for a two-stage detector with a decorrelating detector as the first stage it

is easy to verify that

$$\begin{aligned} \mathcal{O}(K) &= 4K^2 - K \\ \mathcal{O}_i(K) &= 2K^2 + 2K - 1 \end{aligned} \tag{2.60}$$

Figure 2.18 depicts the computational load of the two-stage detector.

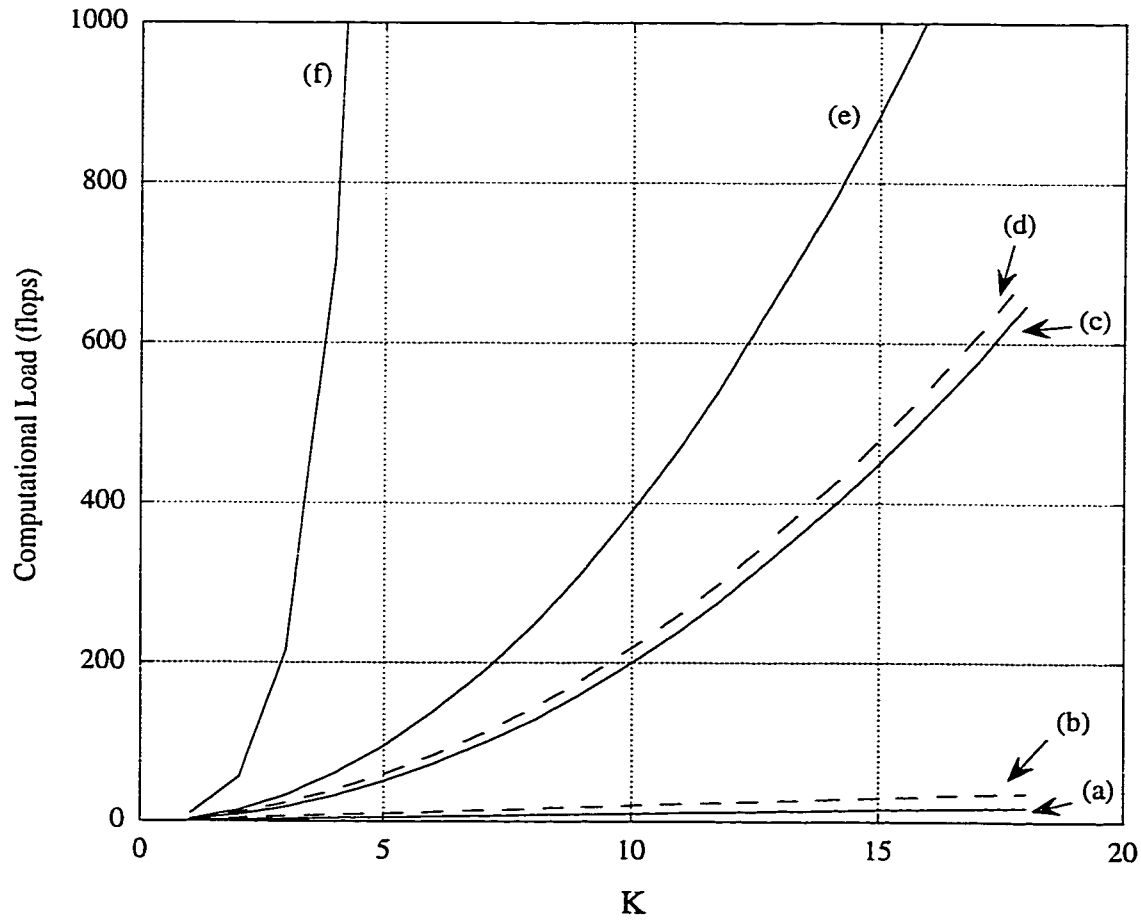


**Figure 2.18**: Computational loads of multiuser detectors; (a) $\mathcal{O}(K)$ of conventional, (b) $\mathcal{O}_i(K)$ of two-stage (conventional 1st stage), (c) $\mathcal{O}(K)$ of two-stage (conventional 1st stage), (d) $\mathcal{O}_i(K)$ of two-stage (decorrelating 1st stage), (e) $\mathcal{O}(K)$ of two-stage (decorrelating 1st stage), (f) $\mathcal{O}_i(K)$ and $\mathcal{O}(K)$ of optimum.

## 2.14 Discussion

In this chapter we studied the idea behind multiuser detection. Then six detection schemes were discussed and investigated with respect to both their bit-error-rate and computational complexity. In order to compare comprehensively the BER of all multiuser detectors presented in this chapter, we have plotted Figure 2.19. Also

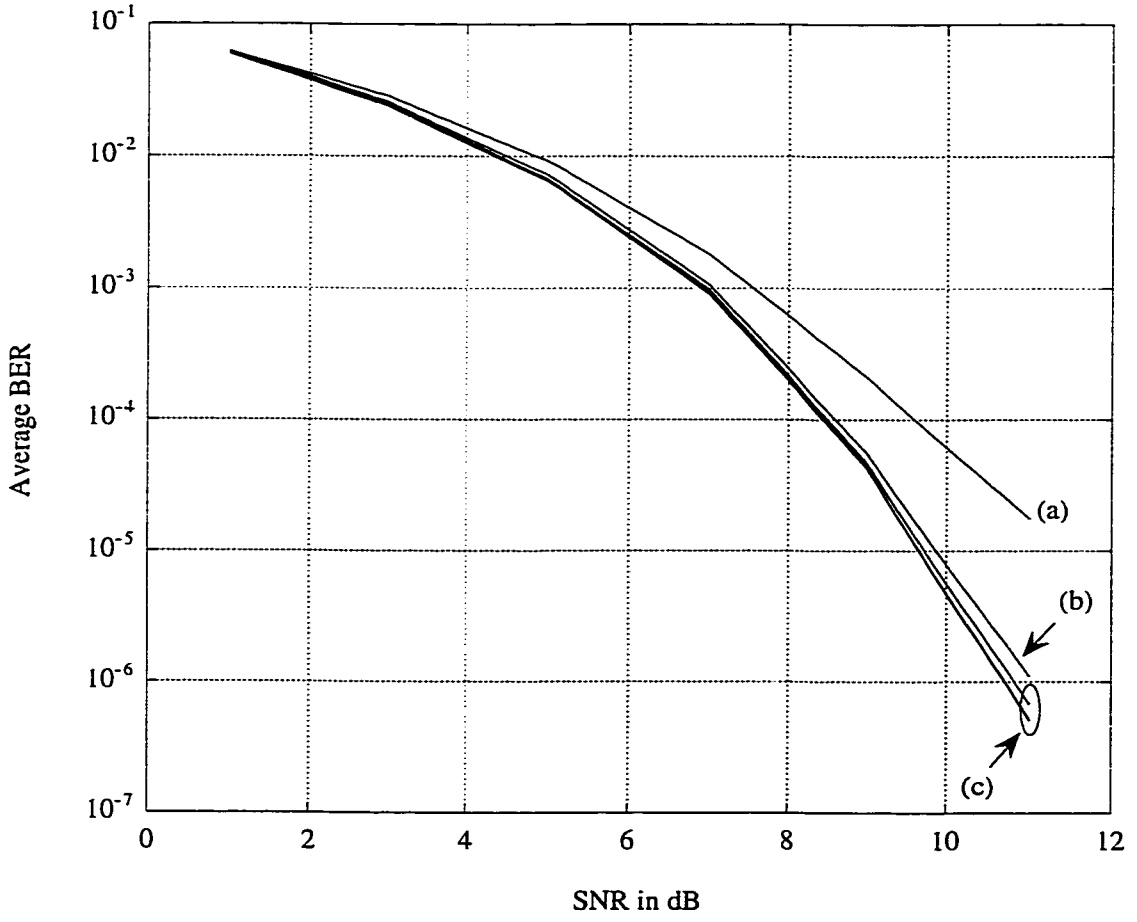


**Figure 2.19**: Average BERs of all multiuser detectors in AWGN channel; (a) conventional, (b) decorrelating, (c) decision-feedback, improved decision-feedback, multistage and optimum overlap.

in Table 2.1, we have sorted, in an ascending manner, the multiuser detectors in terms of the two presented definitions for the computational complexity, i.e. $\mathcal{O}(K)$

and $\mathcal{O}_i(K)$.

In Figure 2.19, we observe that except for the conventional detector, which indeed is not a multiuser detection scheme, the rest of the suboptimum detectors perform reasonably well compared to the optimum detector. Amongst them the decorrelating detector does not yield such a near-optimum BER. On the other hand we notice that amongst all suboptimum multiuser detectors the decorrelating detector is the only one that does not require the knowledge of energies of the users. As mentioned before, in the computer simulations we have assumed that the receivers have perfect knowledge of the energies. In practice, however, this is not the case and the receivers have to estimate the received signal energies or apply other techniques such as power control. This, of course, causes some degradation in BER due to imperfection in such estimates. Therefore if one wants to have a completely fair comparison between suboptimum detectors, one should take into account this fact (the study of methods of estimates of energies is out of the scope of this dissertation). The satisfactory performance of the decorrelating detector, as well as its low computational complexity besides its independence from the knowledge of energies, makes it a good candidate for multiuser detection in AWGN channels.

| Computational Load | Detector |
|---|---|
| lower | Conventional |
| | Decorrelating, Decision-Feedback, Two-Stage (Conventional 1st Stage) |
| | Two-Stage (Decorrelating 1st Stage) |
| | Improved Decision-Feedback |
| higher | Optimum |

**Table 2.1**: Computational loads in an ascending order.

# Chapter 3

# Two-Level Threshold Detection Scheme

From the discussion in the previous chapter, it is clear that DS-CDMA with the conventional detector suffers from the near-far problem, i.e. the bit-error-rate of the weak users are degraded substantially, if the received powers of the users are dissimilar. In the previous chapter, however, we assumed that the received powers from all users are identical, which requires perfect power control and is very difficult to realize in practice. When the power control is not perfect, the performance of the weakest user will be much worse than the results shown in the previous chapter for the conventional detector. The drawbacks of the conventional detector initiated recent efforts to develop more sophisticated receivers in which MAI is treated as a part of information rather than noise. The study of the optimum detector showed that while superior performance over the conventional detector is possible, it can be obtained only at a marked increase in computational complexity. Several suboptimum multiuser detectors were discussed in the previous chapter. The decorrelating detector while simple cannot provide close to optimum performance. A near-optimum

performance is achievable by a two-stage detector at the expense of a considerable increased computational complexity. In this chapter we introduce a new suboptimum multiuser detector. The new detector yields a near-optimum bit-error-rate, while its computational load is slightly more than that of the decorrelating detector.

The rest of this chapter is organized as follows. In Section 3.1 previous work in the area is summarized and in Section 3.2 the contributions of this chapter is reviewed. Section 3.3 thoroughly discusses the idea behind the proposed new detector. The probability of error for the new detector is derived and analyzed in Section 3.4. In Section 3.5 the computational load of this detector is obtained and then it is compared with the other suboptimum detectors. Section 3.6 points out some comments about the proposed new detector. Finally Section 3.7 presents a discussion and summarizes the results.

## 3.1 Previous Work

This chapter may be regarded as the continuation of the previous chapter. The background history mentioned in the previous chapter can be considered as the background history for this chapter too. However, the idea of the two-level threshold detector is mostly indebted to two suboptimum detectors. The first one is the decorrelating detector suggested in [32] and analyzed in [17]. The second detector is the two-stage detector in [19,20] which utilizes a decorrelating detector as its first stage.

## 3.2 Contribution

Based on the work discussed in this chapter we propose a new suboptimum multiuser detection algorithm [26], whose complexity is linear in the number of users. This new detector, namely the two-level threshold detector, has a performance close to that of the optimum detector whose complexity grows exponentially with the number of users. Simulation results have shown that in a DS-CDMA system using Gold sequences of length 31 with 16 users, the two-level threshold detector with a complexity slightly more than the decorrelating detector, can achieve a near-optimum performance.

In [26] we applied a simple inspection and modification method to the outputs of the decorrelating filter. This causes some augmentation in the computational complexity compared to the decorrelating detector. The computational complexity, however, depends on SNR (It also depends on other factors which will be discussed in Section 3.5). For the SNR of interest, the augmented computational complexity is ignorable. This leads to a new detector with a reasonable computational load, which has an excellent near-optimum performance in terms of bit-error-rate.

## 3.3 Two-Level Threshold Detector

In this section, we propose a new suboptimum multiuser detection algorithm based on a novel approach to the maximization problem in (2.25). The proposed algorithm exploits both the simplicity of a decorrelating detector and the novel approach of the multistage detector in maximizing (2.25).

As we saw in Chapter 2, the decorrelating filter, $R^{-1}$, removes MAI completely. Therefore each output of this filter consists of the desired signal and noise and does

not include the interference component. The decorrelating detector uses this fact and applies a simple threshold device to the outputs of $R^{-1}$. However, a more reliable decision criteria can be achieved as it is explained in the next Section.

### 3.3.1 Idea Behind Two-Level Threshold Detector

Figure 2.8 shows a decorrelating detector. The output of the decorrelating filter $R^{-1}$ can be written as:

$$\begin{aligned} \hat{\boldsymbol{y}} &= W\boldsymbol{b} + R^{-1}\boldsymbol{n} \\ &= W\boldsymbol{b} + \boldsymbol{\eta}, \end{aligned} \tag{3.1}$$

where $\boldsymbol{\eta}$ is a zero-mean Gaussian noise vector with covariance matrix equal to $\frac{N_0}{2}R^{-1}$. From (3.1) we may write

$$\hat{y}_k = \sqrt{w_k}b_k + \eta_k, \qquad k = 1, 2, \cdots, K \tag{3.2}$$

where,

$$\eta_k = \sum_{i=1}^{K} (R^{-1})_{k,i} n_i, \qquad k = 1, 2, \cdots, K \tag{3.3}$$

Note that the decorrelating detector completely removes multiple access interference. $\eta_k$ is a Gaussian zero-mean noise component with a variance of $\mathcal{E}\{\eta_k^2\} = (R^{-1})_{k,k}N_0/2$, which is denoted as $\sigma_k^2$. In practice, a signal set with good characteristics, in terms of orthogonality, is chosen as the spreading codes. If that is the case, since the off-diagonal entries of $R$ are the cross-correlations between the spreading code of the system, they are much smaller than 1 (In the ideal case, where codes are orthogonal, R is the identity matrix). Consequently, $R^{-1}$ is very close to the

identity matrix. This implies that $\sigma_k^2$ can be approximated by $N_0/2$.[1]

### 3.3.2 Decision Criteria

For the sake of simplicity we consider a binary modulation scheme. In this case $\hat{y}_k$ takes one of two choices. We refer to them as hypotheses and label them $H_{-1}$ and $H_1$ in the two-choice case. We consider BPSK signalling, wherein the transmitter transmits information by sending 1's and -1's. When '1' is sent, we call it $H_1$, and when '-1' is sent, we call it $H_{-1}$. Under the two hypotheses, we have:

$$\begin{aligned} H_1 &: \hat{y}_k = \sqrt{w_k} + \eta_k \qquad : 1 \text{ is sent} \\ H_{-1} &: \hat{y}_k = -\sqrt{w_k} + \eta_k \qquad : \text{-1 is sent} \end{aligned} \tag{3.4}$$

The probability densities of $\hat{y}_k$ on the two hypotheses are shown in Figure 3.1. As it has been assumed throughout this dissertation, transmitter $k$ sends information $b_k$ equiprobably. Now we consider a neighborhood around zero with radius $\epsilon$. Correspondingly, we investigate four different regions. We determine in each region what decision rule gives us a more reliable answer. What we will do is restating the *likelihood ratio test* explained in detection theory (see for instance [50]) in simple words.

- *Region 1* : $\hat{y}_k \leq -\epsilon$. If $\hat{y}_k$ lies in this region, the probability that $H_{-1}$ is true, i.e. $b_k = -1$, can be evaluated as follows. Since $H_{-1}$ and $H_1$ are mutually

[1] However as it was mentioned in the previous chapter $\mathcal{E}\{\eta_k^2\} \geq N_0/2$, which holds with equality if the spreading codes are orthogonal or in other words $R = R^{-1} = \mathcal{I}_K$.

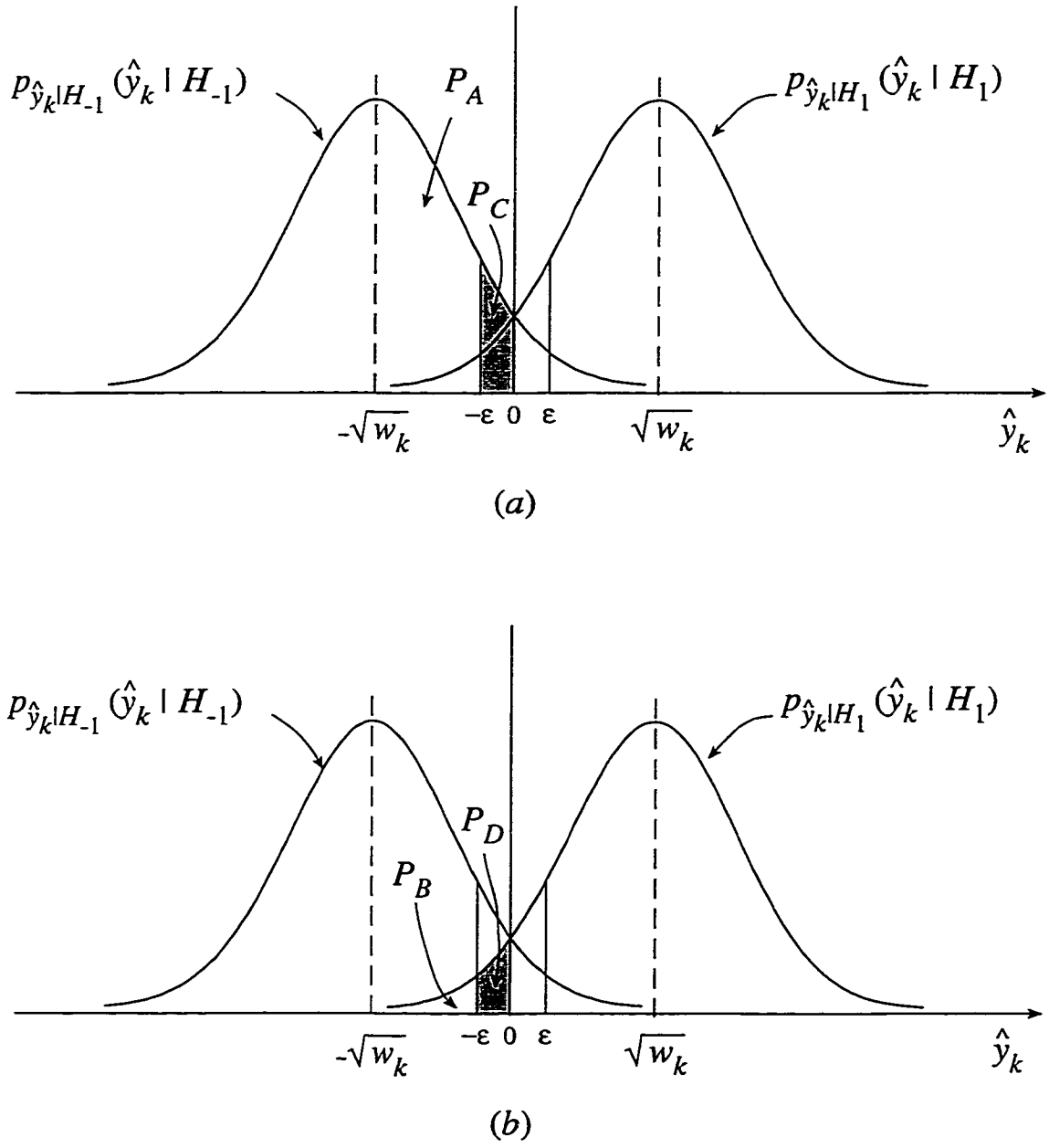


**Figure 3.1**: Error probabilities: (a) $P_A$ and $P_C$; (b) $P_B$ and $P_D$

exclusive hypotheses, using the Bayes' theorem [49] we may write:

$$P(b_k = -1 \mid \hat{y}_k \leq -\epsilon) = \frac{P(\hat{y}_k \leq -\epsilon \mid b_k = -1) \cdot P(b_k = -1)}{P(\hat{y}_k \leq -\epsilon \mid b_k = -1) \cdot P(b_k = -1) + P(\hat{y}_k \leq -\epsilon \mid b_k = 1) \cdot P(b_k = 1)} \quad (3.5)$$

We have assumed that $P(b_k = -1) = P(b_k = 1) = \frac{1}{2}$. Therefore (3.5) can be simplified as

$$P(b_k = -1 \mid \hat{y}_k \leq -\epsilon) = \frac{P(\hat{y}_k \leq -\epsilon \mid b_k = -1)}{P(\hat{y}_k \leq -\epsilon \mid b_k = -1) + P(\hat{y}_k \leq -\epsilon \mid b_k = 1)} \tag{3.6}$$

$P(\hat{y}_k \leq -\epsilon \mid b_k = -1)$ is equivalent to the area under $p_{\hat{y}_k|H_{-1}}(\hat{y}_k \mid H_{-1})$ from $-\infty$ to $-\epsilon$, which is simply the integral of $p_{\hat{y}_k|H_{-1}}(\hat{y}_k \mid H_{-1})$ to the left of the point $-\epsilon$ (Figure 3.1), where we denote it as $P_A$. Similarly $P(\hat{y}_k \leq -\epsilon \mid b_k = 1)$ is equivalent to the area under $p_{\hat{y}_k|H_1}(\hat{y}_k \mid H_1)$ from $-\infty$ to $-\epsilon$, and can be calculated as the integral of $p_{\hat{y}_k|H_1}(\hat{y}_k \mid H_1)$ to the left of the point $-\epsilon$. Notation $P_B$ is used for this value. Hence, the probability that the hypothesis $H_1$ is true, under the condition $\hat{y}_k \leq -\epsilon$, might be expressed as

$$P(b_k = -1 \mid \hat{y}_k \leq -\epsilon) = \frac{P_A}{P_A + P_B} \tag{3.7}$$

Similarly in the same manner we may find the probability that the hypothesis $H_1$ is true provided that $\hat{y}_k \leq -\epsilon$.

$$P(b_k = 1 \mid \hat{y}_k \leq -\epsilon) = \frac{P_B}{P_A + P_B} \tag{3.8}$$

Generally, by inspection of Figure 3.1, we may notice that $P_A \gg P_B$ depending on the value of $\epsilon$. Consequently $P(b_k = -1 \mid \hat{y}_k \leq -\epsilon)$ is close to 1 and $P(b_k = 1 \mid \hat{y}_k \leq -\epsilon)$ is small and close to zero. Let us try to show this fact in a more legitimate manner.

As we mentioned earlier $p_{\hat{y}_k|H_{-1}}(\hat{y}_k \mid H_{-1})$ and $p_{\hat{y}_k|H_1}(\hat{y}_k \mid H_1)$ are normal curves with means $-\sqrt{w_k}$ and $\sqrt{w_k}$, respectively, and both have the same variance $\mathcal{E}\{\eta_k^2\}$.

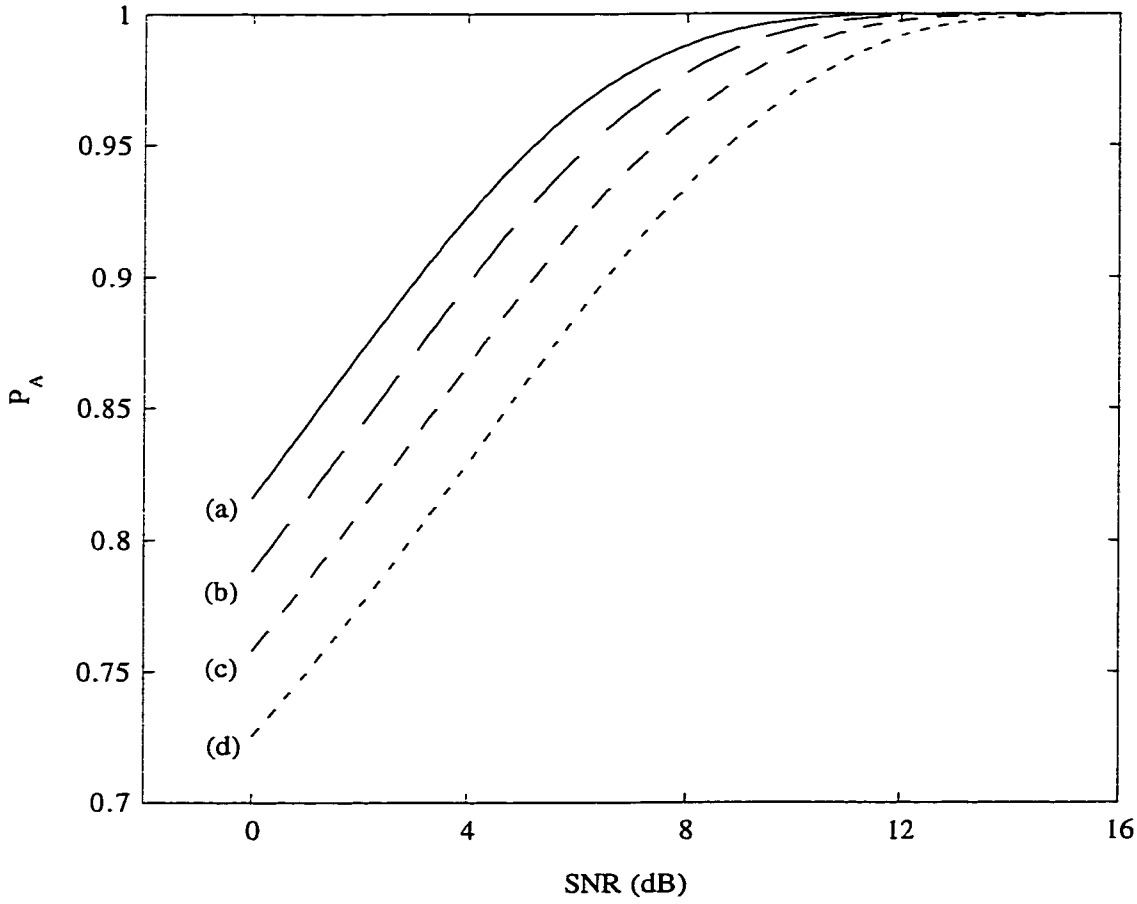


**Figure 3.2**: $P_A$ versus SNR; (a) $\epsilon = 0.1\sqrt{w_k}$, (b) $\epsilon = 0.2\sqrt{w_k}$, (c) $\epsilon = 0.3\sqrt{w_k}$, (d) $\epsilon = 0.4\sqrt{w_k}$

Therefore, the values $P_A$ and $P_B$ depend on $\sqrt{w_k}$ and $\mathcal{E}\{\eta_k^2\}$ and of course $\epsilon$. Since $p_{\hat{y}_k|H_{-1}}(\hat{y}_k \mid H_{-1})$ and $p_{\hat{y}_k|H_1}(\hat{y}_k \mid H_1)$ are normal curves, it is not possible to find a closed-form formula for $P_A$ or $P_B$. However, the area under a normal curve is tabulated in many references (for instance in [59]). In Figure 3.2 we have plotted $P_A$ versus SNR in decibel. Here SNR is defined as below:

$$\text{SNR}_{\text{dB}} = 10 \log_{10} \frac{w_k}{\mathcal{E}\{\eta_k^2\}} \tag{3.9}$$

We have chosen four various values as for $\epsilon$; $0.1\sqrt{w_k}$, $0.2\sqrt{w_k}$, $0.3\sqrt{w_k}$ and $0.4\sqrt{w_k}$.

We also consider positive SNRs which are of interest. We notice that $P_A$ is greater than 0.7 and it increases as SNR increases. Figure 3.3, on the other hand, depicts $P_B$ versus SNR for the same values of $\epsilon$ as in Figure 3.2. Here we observe that $P_B$

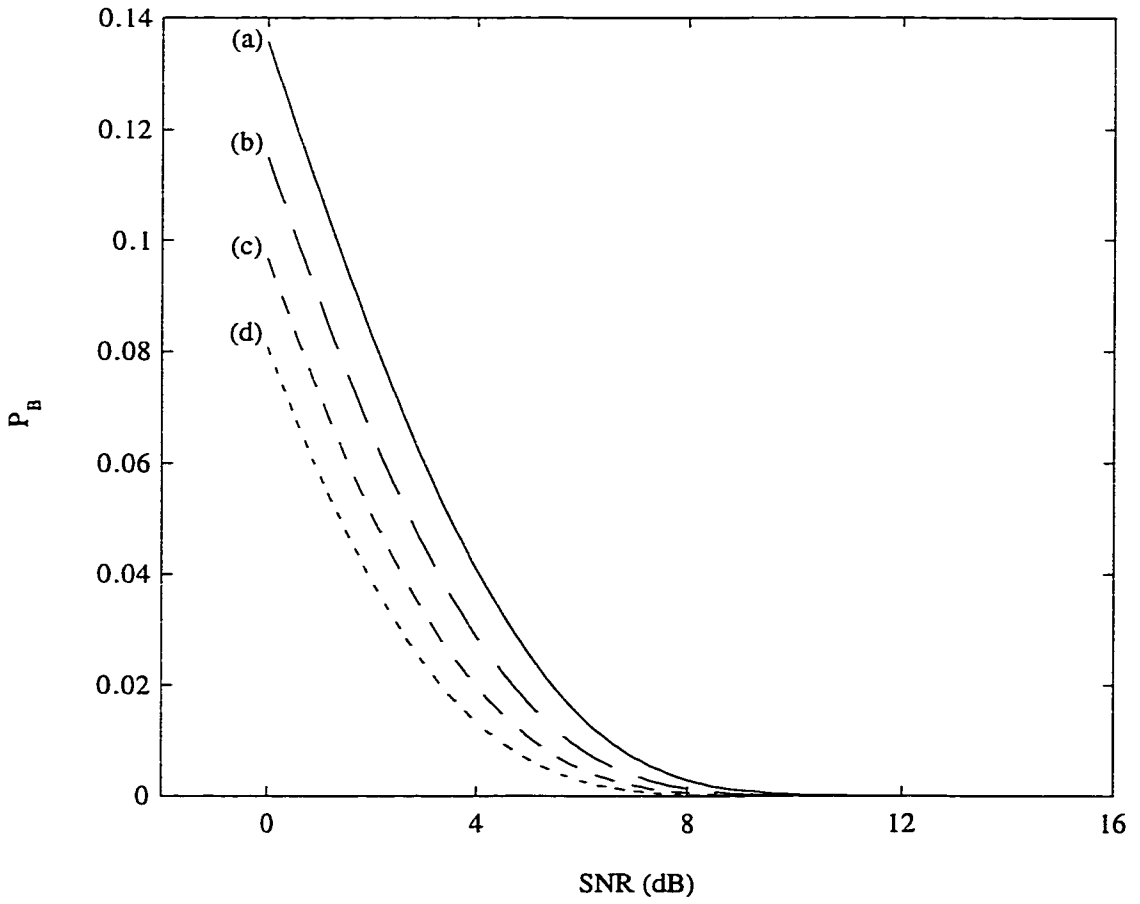


**Figure 3.3**: $P_B$ versus SNR; (a) $\epsilon = 0.1\sqrt{w_k}$, (b) $\epsilon = 0.2\sqrt{w_k}$, (c) $\epsilon = 0.3\sqrt{w_k}$, (d) $\epsilon = 0.4\sqrt{w_k}$

is very small and less than 0.14 for the experimented SNRs. In Figures 3.4 and 3.5 we have plotted $\frac{P_A}{P_A+P_B}$ and $\frac{P_B}{P_A+P_B}$ which are likeliness of $H_{-1}$ and $H_1$ being true, respectively. By inspection of Figures 3.4 and 3.5 we find that the difference between $P(y_k = -1 \mid \hat{y}_k \leq -\epsilon)$ and $P(y_k = 1 \mid \hat{y}_k \leq -\epsilon)$ is so obvious that choosing $H_{-1}$ as the dominant hypothesis seems quite reasonable. Therefore, the decision is

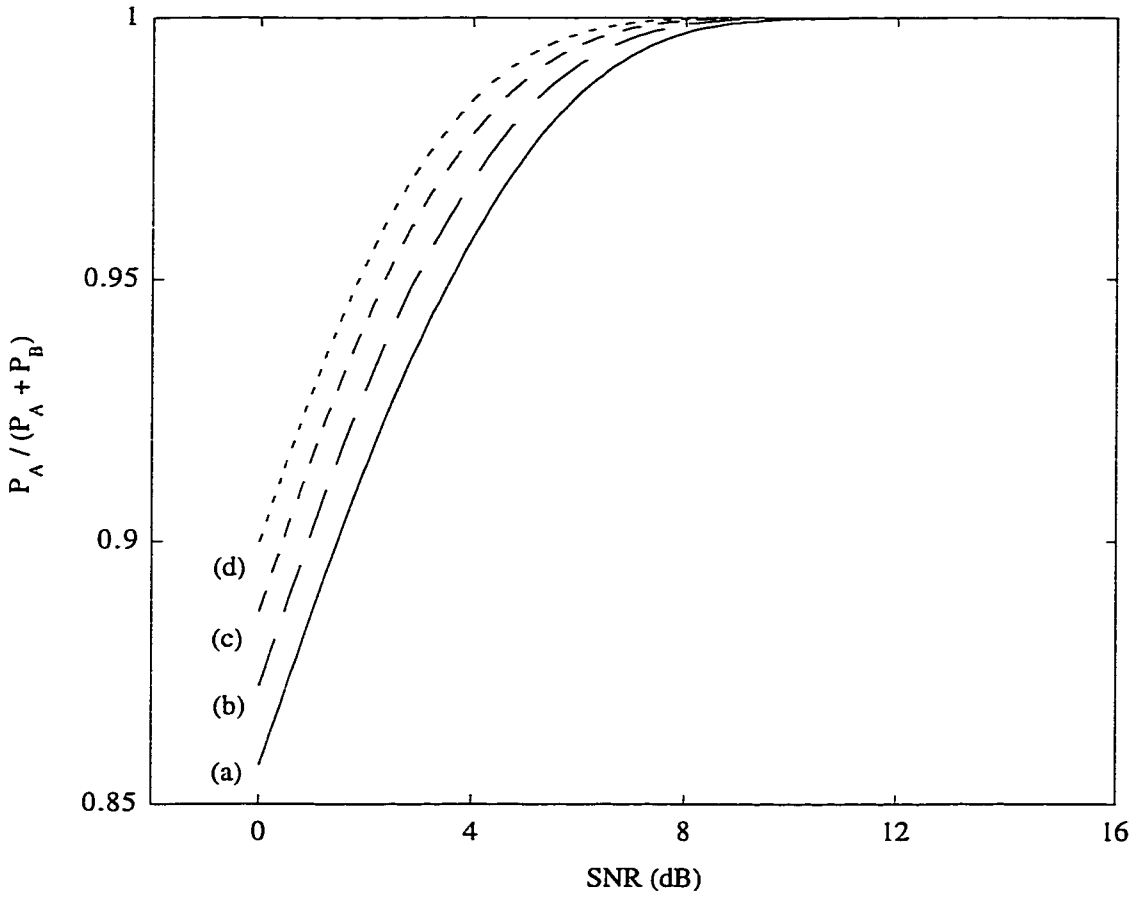


**Figure 3.4:** $\frac{P_A}{P_A+P_B}$ versus SNR; (a) $\epsilon = 0.1\sqrt{w_k}$, (b) $\epsilon = 0.2\sqrt{w_k}$, (c) $\epsilon = 0.3\sqrt{w_k}$, (d) $\epsilon = 0.4\sqrt{w_k}$

in favor of $H_{-1}$ being true, i.e. '-1' is sent. A simple threshold device as the decision algorithm would fulfill the requirements of the likelihood ratio test.

Thus we notice that if $\hat{y}_k \leq -\epsilon$, a decorrelating filter followed by a threshold device will yield a satisfactory detection performance.

- *Region 2* : $-\epsilon < \hat{y}_k \leq 0$. If $\hat{y}_k$ lies in this region, again since $H_1$ and $H_{-1}$

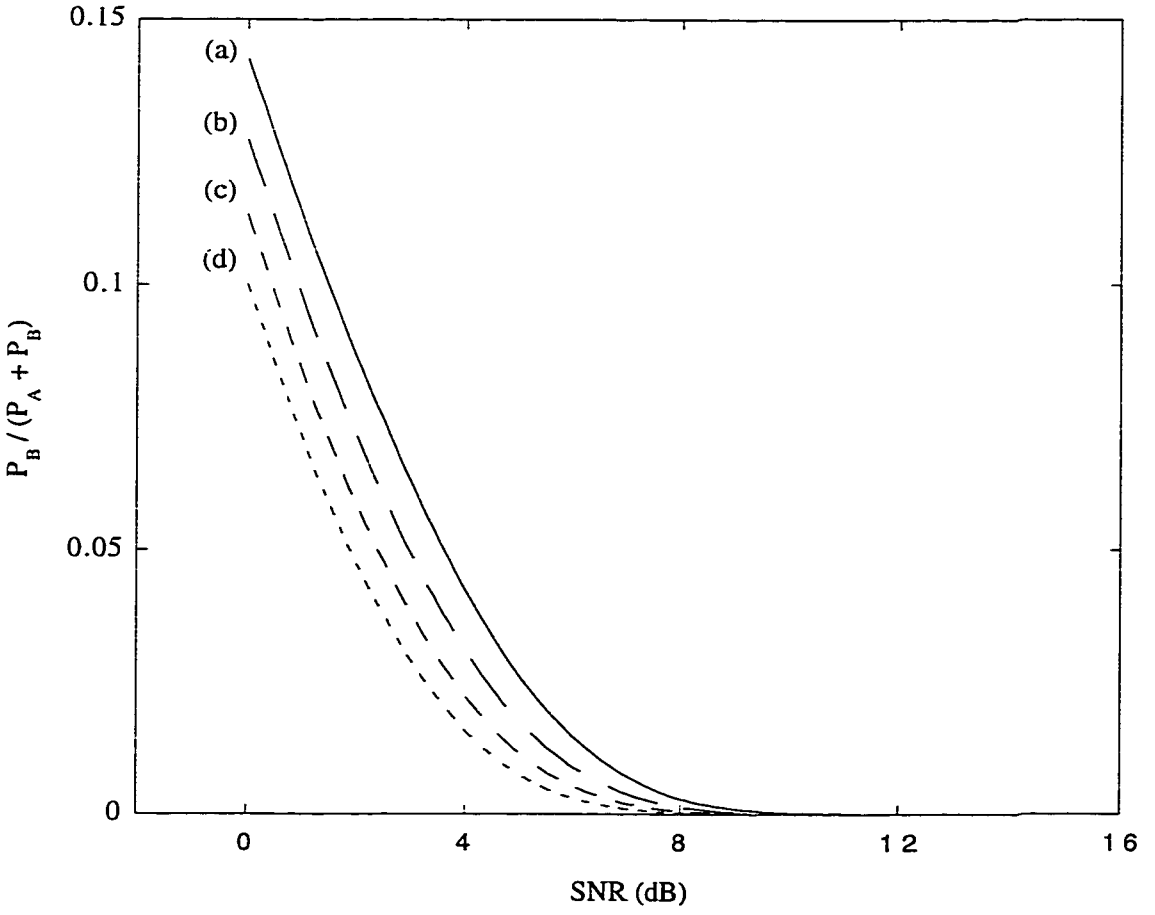


**Figure 3.5:** $\frac{P_B}{P_A+P_B}$ versus SNR; (a) $\epsilon = 0.1\sqrt{w_k}$, (b) $\epsilon = 0.2\sqrt{w_k}$, (c) $\epsilon = 0.3\sqrt{w_k}$, (d) $\epsilon = 0.4\sqrt{w_k}$

are mutually exclusive hypotheses, we may write:

$$P(b_k = -1 \mid -\epsilon < \hat{y}_k \leq 0) = \frac{P(-\epsilon < \hat{y}_k \leq 0 \mid b_k = -1)}{P(-\epsilon < \hat{y}_k \leq 0 \mid b_k = -1) + P(-\epsilon < \hat{y}_k \leq 0 \mid b_k = 1)} \tag{3.10}$$

Or in a simpler notation,

$$P(b_k = -1 \mid -\epsilon < \hat{y}_k \leq 0) = \frac{P_C}{P_C + P_D} \tag{3.11}$$

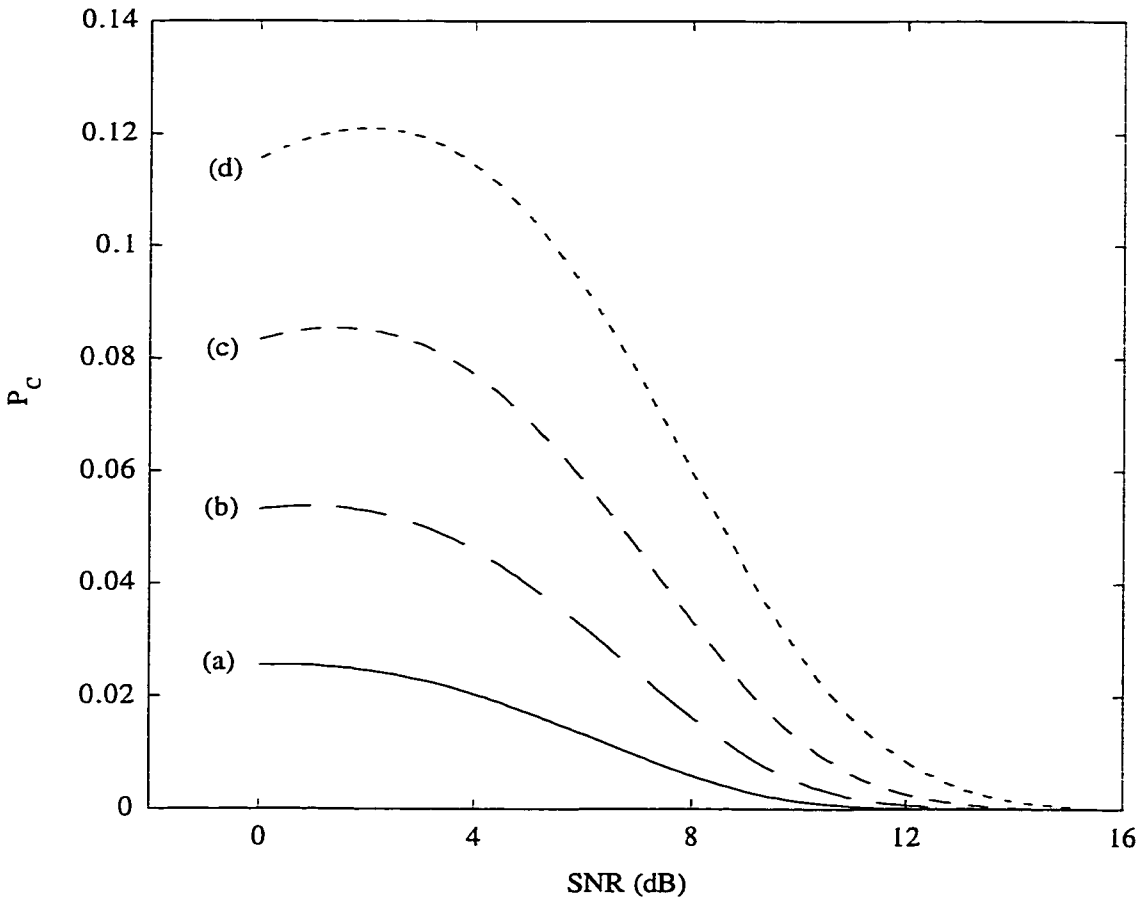


**Figure 3.6**: $P_C$ versus SNR; (a) $\epsilon = 0.1\sqrt{w_k}$, (b) $\epsilon = 0.2\sqrt{w_k}$, (c) $\epsilon = 0.3\sqrt{w_k}$, (d) $\epsilon = 0.4\sqrt{w_k}$

where $P_C$ is the area under $p_{\hat{y}_k \mid H_{-1}}(\hat{y}_k \mid H_{-1})$ from $-\epsilon$ to 0, and $P_D$ is the area under $p_{\hat{y}_k \mid H1}(\hat{y}_k \mid H1)$ also from $-\epsilon$ to 0 (see Figure 3.1). In the same fashion we can express the probability of $H_1$ being true provided that $-\epsilon < \hat{y}_k \leq 0$,

$$P(b_k = 1 \mid -\epsilon < \hat{y}_k \leq 0) = \frac{P_D}{P_C + P_D} \tag{3.12}$$

Figure 3.6 shows $P_C$ versus SNR as it is defined in (3.9). We see that $P_C$ is less than 0.12 for the examined $\epsilon$s. In Figure 3.7, we have plotted $P_D$. We observe that $P_D$ is also small and less than 0.08 for the examined values of $\epsilon$. Figures 3.8

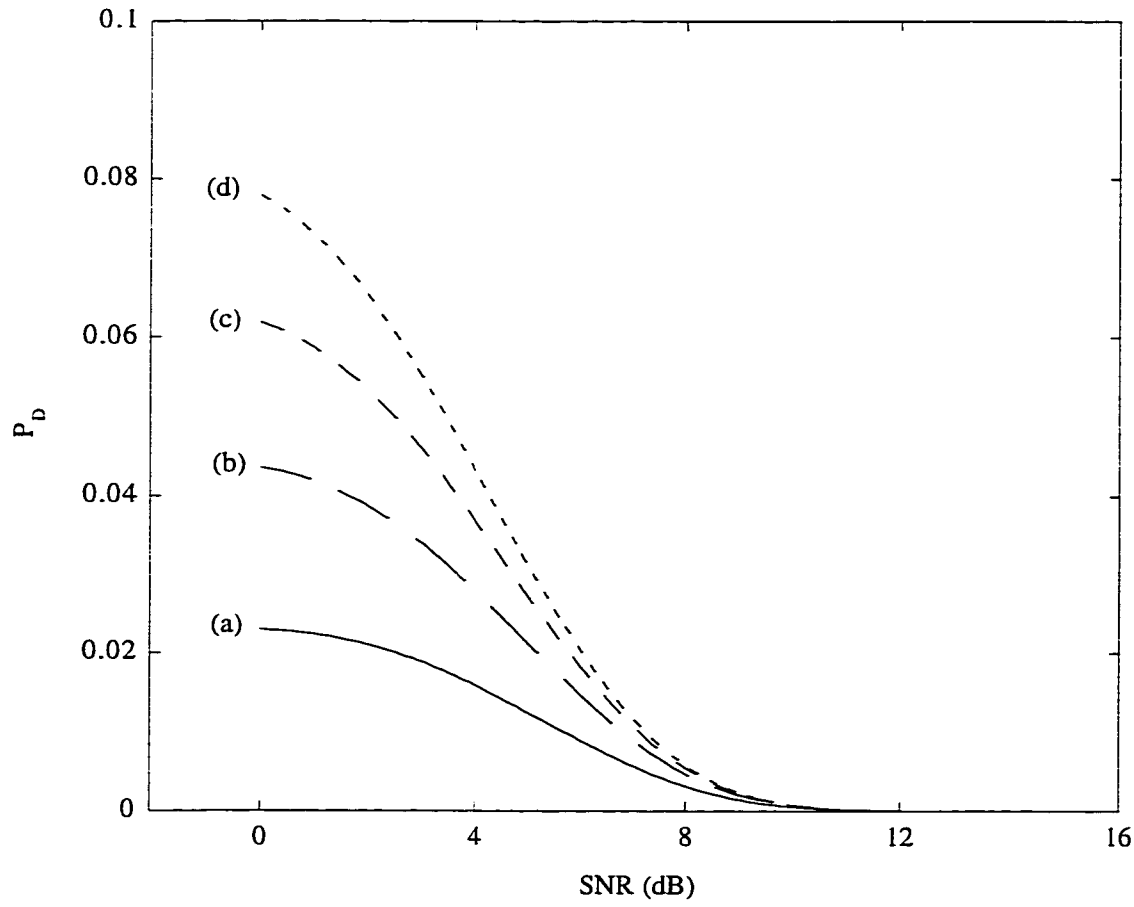


**Figure 3.7**: $P_D$ versus SNR; (a) $\epsilon = 0.1\sqrt{w_k}$, (b) $\epsilon = 0.2\sqrt{w_k}$, (c) $\epsilon = 0.3\sqrt{w_k}$, (d) $\epsilon = 0.4\sqrt{w_k}$

and 3.9 show $\frac{P_C}{P_C+P_D}$ and $\frac{P_D}{P_C+P_D}$, which are the likeliness of $H_{-1}$ and $H_1$ being true, respectively. In this case, unlike the previous case, choosing one of the hypotheses as the most likely one is not very much straightforward [2]. This fact is more evident especially for low SNRs, where $\frac{P_C}{P_C+P_D}$ and $\frac{P_D}{P_C+P_D}$ are both close to 0.5. Indeed what this fact implies is that the probability of $H_{-1}$ being true is almost the same as the probability of $H_1$ being true. In such a case if the decision algorithm is chosen to be just a simple threshold device we are very susceptible to making the wrong decision.

[2]However, the likelihood ratio test is still in favor of $H_{-1}$.

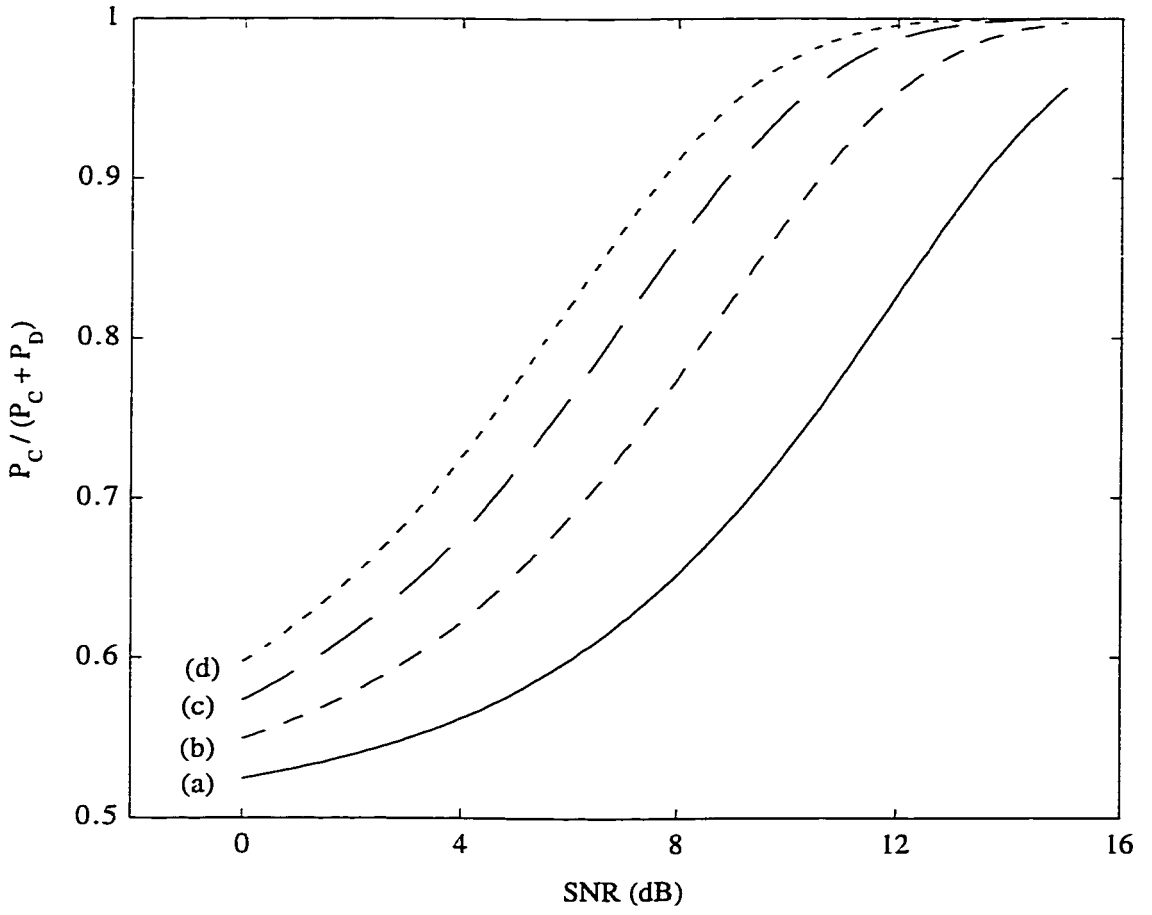


**Figure 3.8:** $\frac{P_C}{P_C+P_D}$ versus SNR; (a) $\epsilon = 0.1\sqrt{w_k}$, (b) $\epsilon = 0.2\sqrt{w_k}$, (c) $\epsilon = 0.3\sqrt{w_k}$, (d) $\epsilon = 0.4\sqrt{w_k}$

Thus a more sophisticated decision algorithm is needed. To achieve a higher level of accuracy we could use the idea of the local maximization of (2.25). In other words instead of simply using a threshold device, we find $\hat{b}_k$ so that

$$2\boldsymbol{b}^T W\boldsymbol{y} - \boldsymbol{b}^T WRW\boldsymbol{b} \tag{3.13}$$

is maximized. As discussed in Section 2.13, local maximization of (3.13) is in fact the main idea of the multistage detector too. Since it plays an important role in

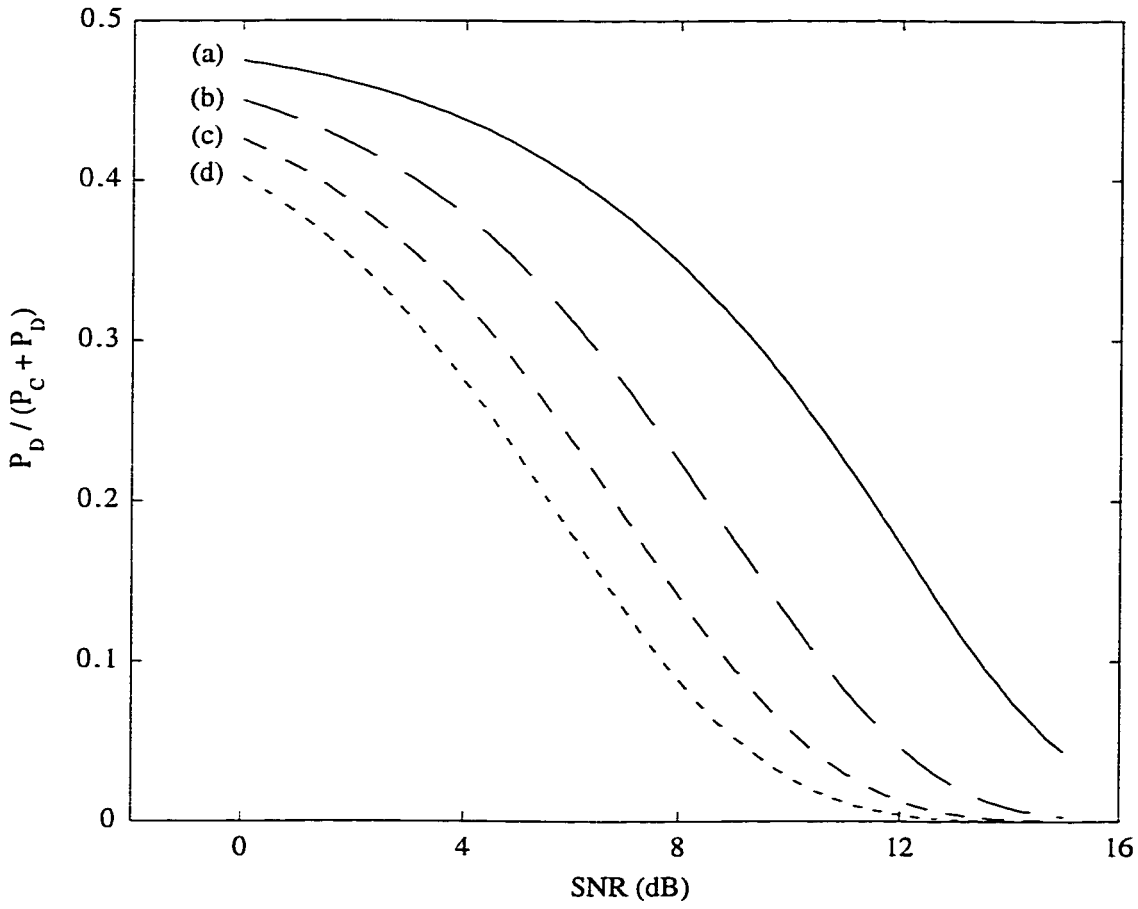


**Figure 3.9**: $\frac{P_D}{P_C+P_D}$ versus SNR; (a) $\epsilon = 0.1\sqrt{w_k}$, (b) $\epsilon = 0.2\sqrt{w_k}$, (c) $\epsilon = 0.3\sqrt{w_k}$, (d) $\epsilon = 0.4\sqrt{w_k}$

the suggested detector, we try to elaborate on that a little bit more. Our target is to find $\hat{b}_k$ that maximizes (3.13).

$$\hat{b}_k = \arg \max_{b_k \in \{-1,+1\}} (2\boldsymbol{b}^T W \boldsymbol{y} - \boldsymbol{b}^T W R W \boldsymbol{b}) \tag{3.14}$$

We start by expanding (3.13) (remember that $\boldsymbol{b} = [b_0, b_1, \cdots, b_K]^T$, $\boldsymbol{y} = [y_0, y_1, \cdots, y_K]^T$, $W$ is a diagonal $K \times K$ matrix, $\{W\}_{i,j} = \sqrt{w_i}$, $i = 1, 2, \cdots, K$ and R is a $K \times K$

matrix whose entries are denoted by $\{R\}_{i,j} = R_{i,j}$),

$$\begin{aligned} & 2\boldsymbol{b}^T W\boldsymbol{y} - \boldsymbol{b}^T WRW\boldsymbol{b} \\ & = 2\sum_{i=1}^{K} y_i\sqrt{w_i}b_i - \sum_{i=1}^{K}\sum_{i=1}^{K} b_i\sqrt{w_i}R_{i,j}\sqrt{w_j}b_j \\ & = 2y_k\sqrt{w_k}b_k - 2\sum_{\substack{i=1 \\ i\neq k}}^{K} b_i\sqrt{w_i}R_{i,k}\sqrt{w_k}b_k - R_{k,k}w_k \\ & \quad + 2\sum_{\substack{i=1 \\ i\neq k}}^{K} y_i\sqrt{w_i}b_i - \sum_{\substack{i=1 \\ i\neq k}}^{K}\sum_{\substack{j=1 \\ j\neq k}}^{K} b_i\sqrt{w_i}R_{i,j}\sqrt{w_j}b_j \end{aligned} \tag{3.15}$$

The above equation is organized in such a way that the first two terms depend on $b_k$ while the remaining terms do not. Since we are searching for $\hat{b}_k$ that maximizes the above expression, hereafter we consider only those terms which depend on $b_k$ and denote them by $\Gamma(b_k)$. We also ignore the multiplier factor $2\sqrt{w_k}$, since it does not affect the calculations.

$$\Gamma(b_k) = y_k b_k - \sum_{\substack{i=1 \\ i\neq k}}^{K} R_{i,k}\sqrt{w_i}b_i b_k \tag{3.16}$$

We notice that $b_k$ is chosen from the binary set $\{-1, +1\}$. Thus the problem of maximization of $\Gamma(b_k)$ in terms of $b_k$ will be simple. By factoring out $b_k$ in (3.16), we have

$$\Gamma(b_k) = \left(y_k - \sum_{\substack{i=1 \\ i\neq k}}^{K} R_{i,k}\sqrt{w_i}b_i\right) b_k. \tag{3.17}$$

Therefore, by inspection Equation (3.17), we simply find out that if we choose $b_k$ and $y_k - \sum_{\substack{i=1 \\ i\neq k}}^{K} R_{i,k}\sqrt{w_i}b_i$ in such a way that they have the same sign, indeed we

have maximized $\Gamma(b_k)$. From the above we can complete Equation (3.14)

$$\begin{aligned}\hat{b}_k &= \arg \max_{b_k \in \{-1,+1\}} (2\boldsymbol{b}^T W \boldsymbol{y} - \boldsymbol{b}^T W R W \boldsymbol{b}) \\ &= \operatorname{sgn}(y_k - \sum_{\substack{i=1 \\ i \neq k}}^{K} R_{i,k} \sqrt{w_i} b_i)\end{aligned} \tag{3.18}$$

As we can see Equation (3.18) is very similar to (2.53) and (2.54), i.e. the detection scheme of the multistage detector.

There is still one question remaining to be answered. We know $y_k$ as well as $\sqrt{w_i}$. $R_{i,k}, (i = 1, 2, \cdots, K)$ are also assumed to be constant and known. What about $b_i, (i = 1, 2, \cdots, K, i \neq k)$? We will return to answer this question in the next section where we summarize the two-level threshold algorithm.

- *Region 3* : $0 < \hat{y}_k \leq \epsilon$. This case is similar to the above case, i.e. $-\epsilon < \hat{y}_k \leq 0$, except the two hypotheses $H_{-1}$ and $H_1$ are swapped. In other words $H_1$ is little bit more likely than $H_{-1}$. However, the difference between them is not significant so that we could not simply select $H_1$ as the dominant one. The same discussion stands for this case as in the previous case, so we do not repeat it.

- *Region 4* : $\hat{y}_k > \epsilon$. This case is similar to the case where $\hat{y}_k < -\epsilon$, with $H_{-1}$ and $H_1$ being swapped. $H_1$ is much more likely than $H_{-1}$ and the difference between them is observable. The decision goes for $H_1$.

From the above discussion we may deduce the following facts: (i) when $\hat{y}_k$ is in the neighborhood of zero, i.e. $-\epsilon < \hat{y}_k < \epsilon$, where $\epsilon$ is a small positive value, the decorrelating detector is likely to fail; (ii) when $\hat{y}_k$ is not in the neighborhood of zero, the decorrelating detector gives a reasonable estimate of the transmitted

bit $b_k$. This leads us to develop a new detection algorithm, namely the two-level threshold algorithm.

### 3.3.3 Two-Level Threshold Algorithm

The two-level threshold algorithm consists of two steps as follows (Figure 3.10.

**Step 1** Use the decorrelating filter, $R^{-1}$, and threshold devices to find $\hat{\boldsymbol{y}}$ and $\hat{\boldsymbol{b}}^d$, respectively.

$$\begin{aligned} \hat{\boldsymbol{y}} &= R^{-1}\boldsymbol{y} \\ \hat{\boldsymbol{b}}^d &= \text{sgn}(\hat{\boldsymbol{y}}) \end{aligned} \tag{3.19}$$

**Step 2** Cycling through the outputs of the decorrelating filter, if $\hat{y}_k$ is in the neighborhood of zero, i.e. $-\epsilon < \hat{y}_k < \epsilon$, then

$$\hat{b}_k^\epsilon = \arg \max_{\substack{b_k \in \{-1,+1\} \\ b_l = \hat{b}_l^\epsilon,\, l<k \\ b_l = \hat{b}_l^d,\, l>k}} (2\boldsymbol{b}^T W \boldsymbol{y} - \boldsymbol{b}^T W R W \boldsymbol{b}), \tag{3.20}$$

and if $\hat{y}_k$ is not in the neighborhood of zero, then simply

$$\hat{b}_k^\epsilon = \hat{b}_k^d = \text{sgn}(\hat{y}_k). \tag{3.21}$$

As it was already shown, (3.20) reduces to:

$$\begin{aligned} \hat{b}_k^\epsilon &= \text{sgn}(q_k) \\ q_k &= y_k - \sum_{i=1}^{k-1} R_{i,k}\sqrt{w_i}\,\hat{b}_i^\epsilon - \sum_{i=k+1}^{K} R_{i,k}\sqrt{w_i}\,\hat{b}_i^d \end{aligned} \tag{3.22}$$

Now the answer to the question on page 74 is clear. To calculate $\hat{b}_k^\epsilon$ in (3.20), $\hat{b}_i^\epsilon$,

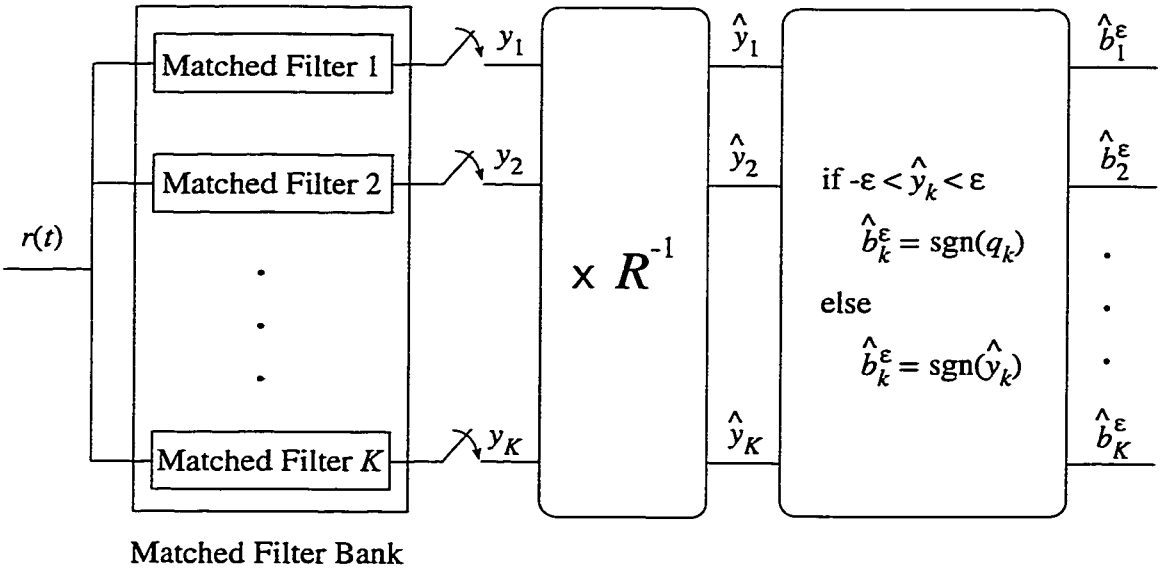


**Figure 3.10**: Two-level threshold detector

$(i = 1, \cdots, k-1)$ and $\hat{b}_i^d$, $(i = k+1, \cdots, K)$ are used.

In Step 2, we have proposed an inspection strategy which, in comparison with the two-stage detector, results in a lower computational complexity by decreasing the number of elements in $\hat{\boldsymbol{b}}$ that must be modified. We will discuss this in detail later. It is worthwhile to mention that the outputs of the decorrelating filter are compared with two levels, i.e. $-\epsilon$ and $\epsilon$, whereas in the decorrelating detector they are compared with one level, i.e. 0. Hence the name two-level threshold.

## 3.4 Probability of Error

In this section we try to derive the probability of error for the two-level threshold detector. Since this detector employs the latest estimates of the bits to estimate the

interfering signal, the probability of error is not necessarily the same for different users. First we consider user 1. Since $b_1$ takes only two equiprobable values $-1$ and $+1$, the probability of error for user 1 can be written as

$$P_1^\epsilon = \frac{1}{2}P[\hat{b}_1^\epsilon = 1 \mid b_1 = -1] + \frac{1}{2}P[\hat{b}_1^\epsilon = -1 \mid b_1 = 1] \tag{3.23}$$

Now we evaluate the first probability in (3.23). The new algorithm suggests the estimated bit $\hat{b}_1 = 1$ provided that; i) $\hat{y}_1 > \epsilon$ or ii) $-\epsilon < \hat{y}_1 < \epsilon$ and $q_1 > 0$. Therefore, the first probability in (3.23) can be written as

$$\begin{aligned} P[\hat{b}_1^\epsilon = 1 \mid b_1 = -1] =& P[\hat{y}_1 > \epsilon \mid b_1 = -1] + \\ & P[-\epsilon < \hat{y}_1 < \epsilon, q_1 > 0 \mid b_1 = -1] \end{aligned} \tag{3.24}$$

The first term in (3.24) can be easily calculated. As we saw in Section 2.10 $\hat{y}_1$ is a Gaussian random variable with conditional mean and variance as $-\sqrt{w_1}$ and $(R^{-1})_{1,1}N_0/2$, respectively. Consequently we can write

$$\begin{aligned} P[\hat{y}_1 > \epsilon \mid b_1 = -1] &= \mathcal{Q}\left(\frac{\sqrt{w_1} + \epsilon}{\sqrt{(R^{-1})_{1,1}N_0/2}}\right) \\ &= \mathcal{Q}\left(\frac{\theta^+}{\sqrt{(R^{-1})_{1,1}N_0/2}}\right) \end{aligned} \tag{3.25}$$

where we have defined $\theta^+ \triangleq \sqrt{w_1} + \epsilon$ for convenience.

Now we calculate the second probability in (3.24). First we expand $q_1$ as follows.

From (3.22)

$$\begin{aligned} q_1 &= y_1 - \sum_{j=2}^{K} R_{1,j}\sqrt{w_j}\hat{b}_j^d \\ &= \sum_{j=1}^{K} R_{1,j}(\sqrt{w_j}b_j + \eta_j) - \sum_{j=2}^{K} R_{1,j}\sqrt{w_j}\hat{b}_j^d \\ &= R_{1,1}(\sqrt{w_1}b_1 + \eta_1) + \sum_{j=2}^{K}[R_{1,j}\sqrt{w_j}(b_j - \hat{b}_j^d) + \eta_j] \end{aligned} \tag{3.26}$$

As we saw in the previous section, to estimate the multiuser interference component we use the latest updates of $\hat{b}_j$. For user 1, however, $\hat{b}_j, j = 2, \cdots, K$, are all obtained from the output of the decorrelating filter. In other words, to evaluate $q_1$ in (3.26), we use $\hat{b}_j^d$ which is calculated as

$$\begin{aligned} \hat{b}_j^d &= \text{sgn}(\hat{y}_j) \\ &= \text{sgn}(\sqrt{w_j}b_j + \eta_j). \end{aligned} \tag{3.27}$$

From (3.26) and (3.27), $q_1$ can be rewritten as

$$q_1 = R_{1,1}(\sqrt{w_1}b_1 + \eta_1) + \sum_{j=2}^{K}[R_{1,j}\sqrt{w_j}(b_j - \text{sgn}(\sqrt{w_j}b_j - \eta_j)) + \eta_j] \tag{3.28}$$

The second term in (3.28) obviously depends only on $\boldsymbol{\beta}_1 \triangleq [b_2, \cdots, b_K]^T$ and $\boldsymbol{\eta}_1 \triangleq [\eta_2, \cdots, \eta_K]^T$. We denote it as $\kappa(\boldsymbol{\beta}_1, \boldsymbol{\eta}_1)$ to explicitly show its dependence on both $\boldsymbol{\beta}_1$ and $\boldsymbol{\eta}_1$. Using this notation and (3.28), we express the second probability in (3.24) as

$$P[-\epsilon < \hat{y}_1 < \epsilon, q_1 > 0 \mid b_1 = -1] = \mathcal{E}_{\boldsymbol{\beta}_1}\{\mathcal{E}_{\boldsymbol{\eta}_1}\{\psi_{-1}(\boldsymbol{\beta}_1, \boldsymbol{\eta}_1)\}\} \tag{3.29}$$

where

$$\psi_{-1}(\boldsymbol{\beta}_1, \boldsymbol{\eta}_1) = P[\theta^- < \eta_1 < \theta^+, \eta_1 > \sqrt{w_1} - \frac{\kappa(\boldsymbol{\beta}_1, \boldsymbol{\eta}_1)}{R_{1,1}} \mid \boldsymbol{\beta}_1, \boldsymbol{\eta}_1] \tag{3.30}$$

and $\theta^-$ is defined as $\theta^- \triangleq \sqrt{w_i} - \epsilon$. $\mathcal{E}_{\boldsymbol{\beta}_1}$ denotes expectation over the ensemble of identical uniformly distributed $\boldsymbol{\beta}_1 \in \{-1, +1\}^{K-1}$. $\mathcal{E}_{\boldsymbol{\eta}_1}$ denotes the expectation over statistically correlated Gaussian random variables $\eta_i$, $(i = 2, \cdots, K)$. To evaluate this expectation, of course, the calculation of a joint $(K-1)$-dimensional Gaussian distribution function, $f(\eta_2, \cdots, \eta_K)$, is required. On page 61 we showed that the covariance matrix of vector $\boldsymbol{\eta}$ is $\text{Cov}(\boldsymbol{\eta}) = \frac{N_0}{2} R^{-1}$. From this we may notice that the covariance matrix of the random variables $\eta_i$, $(i = 2, \cdots, K)$ can be easily determined, $\text{Cov}(\boldsymbol{\eta}_1) = \frac{N_0}{2} \langle R^{-1} \rangle_{1,1}$, where $\langle R^{-1} \rangle_{1,1}$ is a $K-1 \times K-1$ matrix obtained by removing row 1 and column 1 of matrix $R^{-1}$. We know that once the covariance matrix of some statistically dependent Gaussian random variables is known their joint pdf is determined [49].

Now let us return to Equation (3.30). This equation suggests that to calculate the probability of error we need to know the conditional probability density function (pdf) of $\eta_1$, denoted as $f_{\eta_1 | \boldsymbol{\beta}_1, \boldsymbol{\eta}_1}$. Since the vector of transmitted information bits $\boldsymbol{\beta}_1$ is independent of the noise vector $\boldsymbol{\eta}_1$, henceforth the conditional pdf of $\eta_1$ is denoted as $f_{\eta_1 | \boldsymbol{\eta}_1}$. The conditional pdf $f_{\eta_1 | \boldsymbol{\eta}_1}$ may be expressed as the ratio between two joint probability distribution functions

$$f_{\eta_1 | \boldsymbol{\eta}_1}(\eta_1 \mid \boldsymbol{\eta}_1) = \frac{f(\eta_1, \eta_2, \cdots, \eta_K)}{f(\eta_2, \cdots, \eta_K)} = \frac{f(\boldsymbol{\eta})}{f(\boldsymbol{\eta}_1)} \tag{3.31}$$

We already mentioned that $f(\boldsymbol{\eta}_1)$ is easy to compute. On the other hand, as we know the covariance matrix of vector $\boldsymbol{\eta}$ is $\frac{N_0}{2} R^{-1}$, hence $f(\boldsymbol{\eta})$ is easy to calculate

too. Once we find $f_{\eta_1|\eta_1}(\eta_1 \mid \boldsymbol{\eta}_1)$, $\psi_{-1}(\boldsymbol{\beta}_1, \boldsymbol{\eta}_1)$ may be expressed as

$$\psi_{-1}(\boldsymbol{\beta}_1, \boldsymbol{\eta}_1) = \begin{cases} \int_{\theta^-}^{\theta^+} f_{\eta_1|\eta_1}(\eta_1 \mid \boldsymbol{\eta}_1) d\eta_1 & ; \sqrt{w_1} - \frac{\kappa(\boldsymbol{\beta}_1, \boldsymbol{\eta}_1)}{R_{1,1}} < \theta^-, \\ \int_{\sqrt{w_1} - \frac{\kappa(\boldsymbol{\beta}_1, \boldsymbol{\eta}_1)}{R_{1,1}}}^{\theta^+} f_{\eta_1|\eta_1}(\eta_1 \mid \boldsymbol{\eta}_1) d\eta_1 & ; \theta^- \leq \sqrt{w_1} - \frac{\kappa(\boldsymbol{\beta}_1, \boldsymbol{\eta}_1)}{R_{1,1}} < \theta^+, \\ 0 & ; \theta^+ \leq \sqrt{w_1} - \frac{\kappa(\boldsymbol{\beta}_1, \boldsymbol{\eta}_1)}{R_{1,1}}. \end{cases} \tag{3.32}$$

To simplify (3.32), a limiter function $\phi(\cdot)$ is defined,

$$\phi(x) \triangleq \begin{cases} \theta^- & ; x < \theta^- \\ x & ; \theta^- \leq x < \theta^+ \\ \theta^+ & ; \theta^+ \leq x \end{cases} \tag{3.33}$$

Using (3.33), (3.32) can be rewritten as

$$\psi_{-1}(\boldsymbol{\beta}_1, \boldsymbol{\eta}_1) = \int_{\phi(\sqrt{w_1} - \frac{\kappa(\boldsymbol{\beta}_1, \boldsymbol{\eta}_1)}{R_{1,1}})}^{\theta^+} f_{\eta_1|\eta_1}(\eta_1 \mid \boldsymbol{\eta}_1) d\eta_1 \tag{3.34}$$

Similarly it can be proved that the second probability in (3.23) is obtained as

$$P[\hat{b}_1^{\epsilon} = -1 \mid b_1 = 1] = Q\left(\frac{\theta^+}{\sqrt{(R^{-1})_{1,1} N_0/2}}\right) + \mathcal{E}_{\boldsymbol{\beta}_1}\{\mathcal{E}_{\boldsymbol{\eta}_1}\{\psi_1(\boldsymbol{\beta}_1, \boldsymbol{\eta}_1)\}\} \tag{3.35}$$

where

$$\psi_1(\boldsymbol{\beta}_1, \boldsymbol{\eta}_1) = \int_{-\theta^+}^{-\phi(\sqrt{w_1} + \frac{\kappa(\boldsymbol{\beta}_1, \boldsymbol{\eta}_1)}{R_{1,1}})} f_{\eta_1|\eta_1}(\eta_1 \mid \boldsymbol{\eta}_1) d\eta_1 \tag{3.36}$$

Finally from Equations (3.24), (3.25), (3.29) and (3.35) we can write an expression

for the probability of error of the new algorithm for user 1.

$$P_1^{\epsilon} = \mathcal{Q}\left(\frac{\theta^+}{\sqrt{(R^{-1})_{1,1}N_0/2}}\right) + \frac{1}{2}\mathcal{E}_{\boldsymbol{\beta}_1}\{\mathcal{E}_{\boldsymbol{\eta}_1}\{\psi_{-1}(\boldsymbol{\beta}_1,\boldsymbol{\eta}_1) + \psi_1(\boldsymbol{\beta}_1,\boldsymbol{\eta}_1)\}\} \tag{3.37}$$

where $\psi_{-1}(\boldsymbol{\beta}_1,\boldsymbol{\eta}_1)$ and $\psi_1(\boldsymbol{\beta}_1,\boldsymbol{\eta}_1)$ are defined as (3.34) and (3.36), respectively.

Using Equations (3.31), (3.34), (3.36) and (3.37) we may notice that the computation of $\mathcal{E}_{\boldsymbol{\eta}_1}\{\psi_{-1}(\boldsymbol{\beta}_1,\boldsymbol{\eta}_1) + \psi_1(\boldsymbol{\beta}_1,\boldsymbol{\eta}_1)\}$ requires two $K$-fold integration of the joint pdf $f(\boldsymbol{\eta})$. Since $f(\boldsymbol{\eta})$ is the pdf of statistically dependent Gaussian random variables, we have to use numerical methods to find the integrals. On the other hand, to take the expectation over $\boldsymbol{\beta}_1$ we need to calculate those integrals $2^{K-1}$ times.

The probability of error for other users, for instance user $k$, requires more caution. Basically the approach will be the same, except that the definition of function $\kappa(\cdot)$ (which in this case is a function of $\boldsymbol{\beta}_k \triangleq [b_1,\cdots,b_{k-1},b_{k+1},\cdots,b_K]^T$ and $\boldsymbol{\eta}_k \triangleq [\eta_1,\cdots,\eta_{k-1},\eta_{k+1},\cdots,\eta_K]^T$) will be different. Since this leads to a much more complicated calculations compared to calculations for user 1, we ignore it.

The large amount of numerical computations for the probability of error, especially when the number of users is large, is even more than the required calculations for computer simulations. Consequently, once again as we did in Chapter 2, we use computer simulations to obtain the bit-error-rate of the two-level threshold detector. Figure 3.11 shows the simulation results for a 16-user DS-CDMA system which we characterized in Section 2.6. For the two-level threshold detector, we have chosen $\epsilon = 0.1\sqrt{w_k}$. Of course this value affects the BER. In Chapter 5, where we deploy this detector in multipath fading channels, we investigate how the BER varies with $\epsilon$. In Figure 3.11 we have plotted the average BER of the two-level threshold detector as well as those of the decorrelating detector and the two-stage detector which utilizes a decorrelating detector as the first stage. The average BER

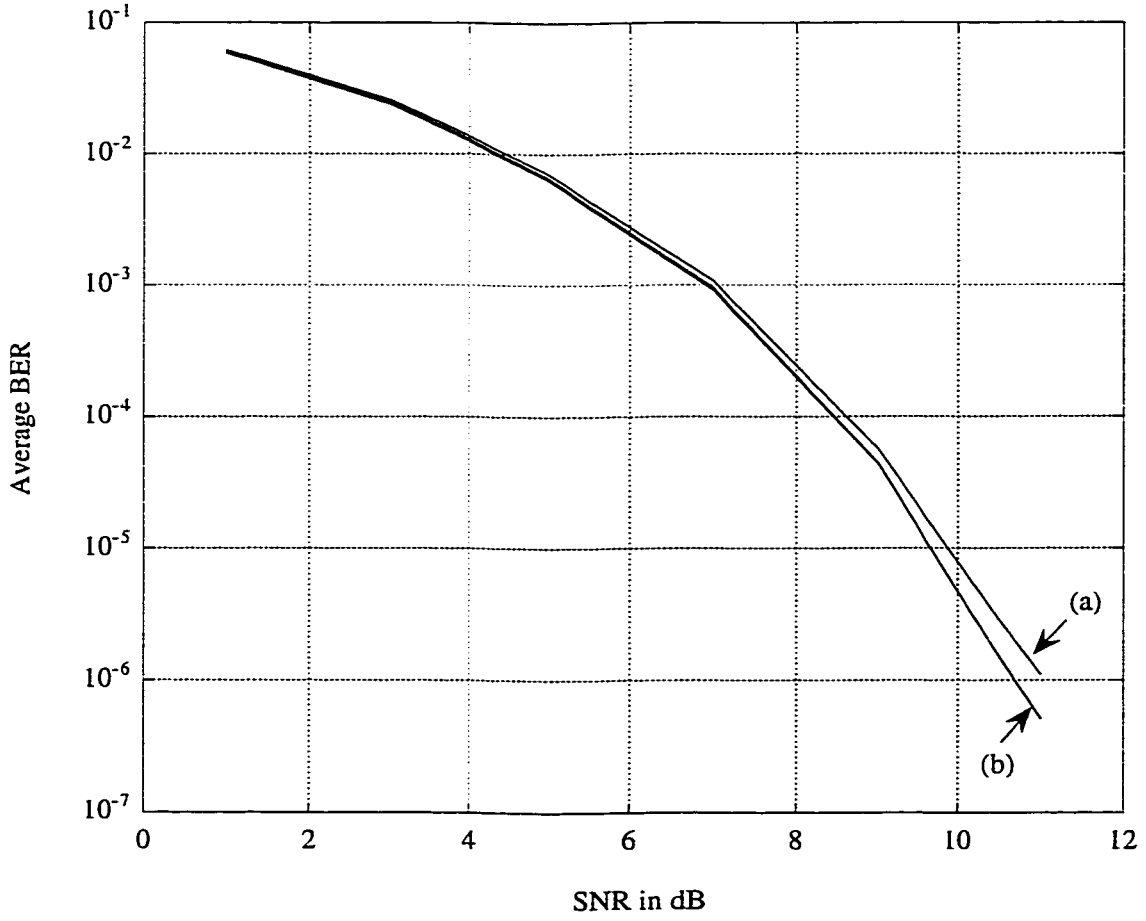


**Figure 3.11**: Comparison of bit-error-rates in an AWGN channel; (a) decorrelating, (b) optimum, two-level threshold, two-stage (decorrelating 1st stage)

of the optimum detector is plotted as a benchmark too.

## 3.5 Computational Load

In this section we find the computational complexity of the two-level threshold detector and compare it with the computational complexity of the two other detectors, the decorrelating detector and the two-stage detector which uses a decorrelating detector as its first stage.

As we saw the new detector consists of two steps. The first step is a decorrelating filter, and the second step is examining the outputs of the decorrelating filter so that if they are in a neighborhood of zero we use (3.22), otherwise we simply use a threshold device. In Chapter 2, Section 2.10, we derived the computational complexity of a decorrelating filter, i.e. Step 1 of the new detector. Here we derive the computational load of the Step 2. Notation $\mathcal{O}^2$ is used for this purpose.

First let us consider user 1. We calculate the computational load required to detect $\hat{b}_1$ and denote it as $\mathcal{O}_1^2$. Let us also review the function of the new detector. If the output of the decorrelating filter, $\hat{y}_1$, is less than $-\epsilon$ or greater than $\epsilon$, then the only required computation is a compare. Otherwise, if $\hat{y}_1$ is between $-\epsilon$ and $\epsilon$ then $\hat{b}_1$ is computed using (3.22), which obviously requires more than just one compare. We notice that the computational load depends on the magnitude of $\hat{y}_1$. Since $\hat{y}_1$ is a random variable, the computational load is a random variable too. In order to have a measure of computational load so that we compare this detector with other detectors in this respect, we find the average of the computational load. To accomplish this, first we calculate the probability of $\hat{y}_k$ being between $-\epsilon$ and $\epsilon$ and denote it as $p_k$. We recall that $\hat{y}_k$ is a Gaussian random variable with a mean of $\sqrt{w_k}$ or $-\sqrt{w_k}$ and with variance $\sigma_k^2$. Due to the symmetrical property of the Gaussian distribution, regardless of whether the mean of $\hat{y}_k$ is $\sqrt{w_k}$ or $-\sqrt{w_k}$, $p_k$ can be expressed by the $\mathcal{Q}$-function as,

$$p_k = \mathcal{Q}(\frac{-\epsilon + \sqrt{w_k}}{\sigma_k}) - \mathcal{Q}(\frac{\epsilon + \sqrt{w_k}}{\sigma_k}) \tag{3.38}$$

Now, the average computational load involving the detection of $b_1$, i.e. $\bar{\mathcal{O}}_1^2$, can be calculated as follow. There is a compare needed to check whether $\hat{y}_1$ is between $-\epsilon$ and $\epsilon$ or not. If $\hat{y}_1$ is not in this region there is another compare needed. This

event happens with a probability of $1 - p_1$. Otherwise, $\hat{y}_1$ lies between $-\epsilon$ and $\epsilon$. Then according to Equation (3.22), $K-1$ more compares are needed to calculate $\hat{b}_i$, $i = 2, \cdots, K$, and $2K-1$ multiplies and adds are needed to calculate (3.22) as well as a compare to find $\hat{b}_1$. This happens, of course, with probability $p_1$. Therefore, the average computational load required to calculate $\hat{b}_1$, denoted by $\bar{\mathcal{O}}_1^2(K)$, can be written as,

$$\bar{\mathcal{O}}_1^2(K) = 1 + (1 - p_1) \times 1 + p_1(3K - 2) \tag{3.39}$$

Similarly $\bar{\mathcal{O}}_k^2(K)$, $k = 2, \cdots, K$ can be calculated. However, we notice here that if $-\epsilon \leq \hat{y}_k < \epsilon$, there is no more a need to a compare to determine the sign of $\hat{y}_k$, since it has already been determined during the calculation of $\hat{b}_1$. Also there is no need to calculate $\hat{b}_i$, $i = 1, \cdots, K$, $i \neq k$, since they have been already calculated in the prior calculations. Thus the average computation load required to evaluate Equation (3.22) for users 2 and higher, assuming that $\hat{b}_i$, $i = 1, \cdots, k-1$ are already calculated, is

$$\bar{o}_k(K) = 1 + p_k(2K - 1) \tag{3.40}$$

The above equation, however, does not include the total calculations required to detect $\hat{b}_k$, since it does not count the computational load to calculate $\hat{b}_i$, $i = 1, \cdots, k-1$ (We notice that to calculate $\hat{b}_k$, according to (3.22), we need to know $\hat{b}_i$, $i = 1, \cdots, k-1$). Therefore the total computational load in Step 2 required to

detect $\hat{b}_k$ is,

$$\begin{aligned}\bar{\mathcal{O}}_k^2(K) &= \bar{\mathcal{O}}_1 + \sum_{i=2}^{k} \bar{o}_i \\ &= 1 + (1-p_1) + p_1(3K-2) + \sum_{i=2}^{k}[1 + p_i(2K-1)]\end{aligned} \tag{3.41}$$

Finally we add the computational load required to perform the decorrelating filter to (3.39) and (3.41). The general expression for the computational load of the new detector is

$$\bar{\mathcal{O}}_k(K) = 2K^2 - K + 1 + (1-p_1) + p_1(3K-2) + \sum_{i=2}^{k}[1 + p_i(2K-1)] \tag{3.42}$$

A particular case of interest is when $p_1 = p_2 = \cdots = p_K \triangleq p$. If this is the case, then,

$$\bar{\mathcal{O}}_k(K) = 2K^2 - K + k + (1-p) + p[(2k+1)K - (k+1)] \tag{3.43}$$

### 3.5.1 Comparison of Computational Loads

In this section we compare the computational load of the new detector with that of the decorrelating detector and the two-stage detector which uses a decorrelating detector as the first stage. Equation (3.41) shows the computational load required to calculate $\hat{b}_k$. We notice that detection of $\hat{b}_K$ involves detection of all other bits $\hat{b}_i$, $i = 1, \cdots, K-1$. Therefore if in (3.41) we set $k = K$, it will yield the computational load required to detect all transmitted $b_i$, $(i = 1, \cdots, K)$. By setting $k = K$ in

(3.41), we have

$$\bar{\mathcal{O}}_K(K) = \mathcal{O}(K) = 2K^2 + (1-p) + p[2K^2 - 1] \tag{3.44}$$

We compare (3.42) with (2.37) and (2.60), i.e. the total computational load of the decorrelating detector and the two-stage detector while detecting all transmitted bits. Of course, (3.44) depends on $p$, and the latter , itself, depends on $\epsilon$ and SNR as it was defined in (3.9).

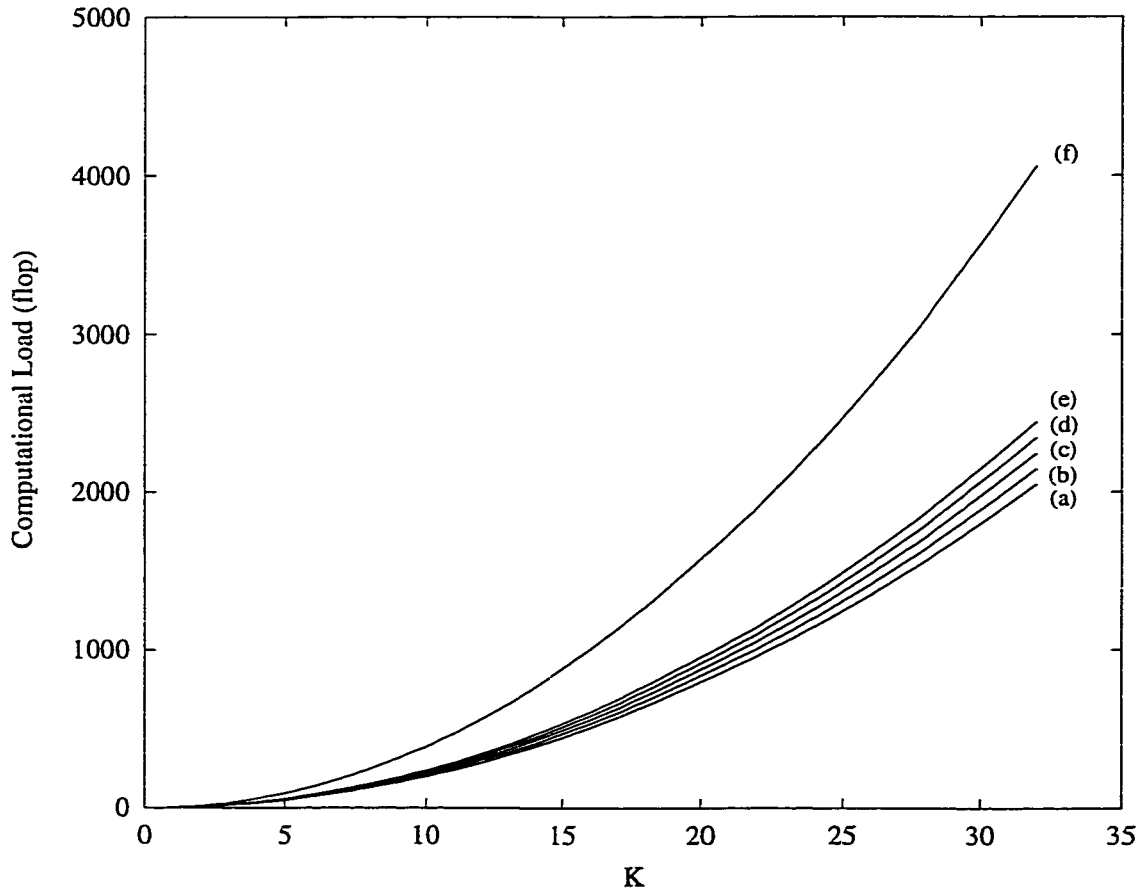


**Figure 3.12**: Comparison of computational loads at SNR = 0 dB; (a) decorrelating, (b) two-level threshold $\epsilon = 0.1\sqrt{w_k}$, (c) two-level threshold $\epsilon = 0.2\sqrt{w_k}$, (d) two-level threshold $\epsilon = 0.3\sqrt{w_k}$, (e) two-level threshold $\epsilon = 0.4\sqrt{w_k}$, (f) two-stage (decorrelating 1st stage)

First, we consider three different values for SNR; 0 dB, 5 dB and 10 dB. In Figures 3.12, 3.13 and 3.14, we have plotted the computational load of the two-level threshold detector, corresponding to four values for $\epsilon$; $0.1\sqrt{w_k}$, $0.2\sqrt{w_k}$, $0.3\sqrt{w_k}$ and $0.4\sqrt{w_k}$. The plots are versus $K$, the number of users. As it is clear from all

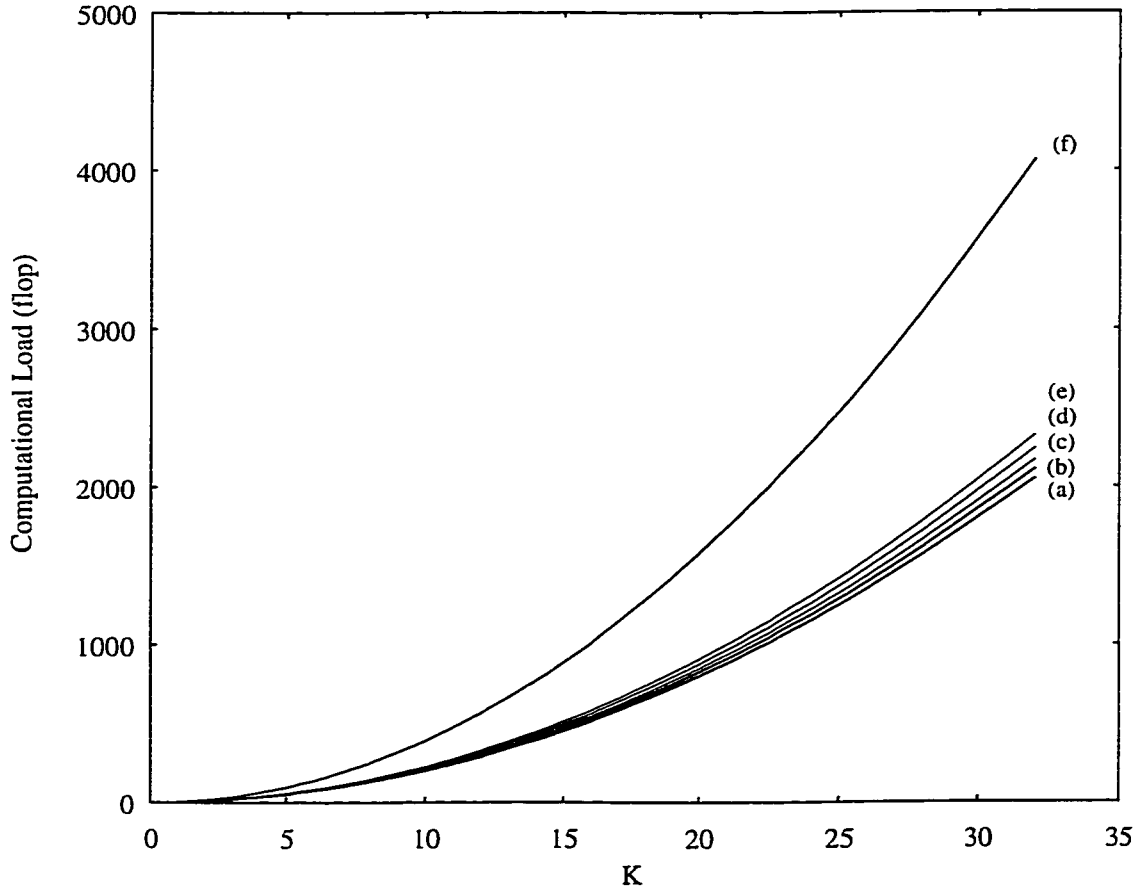


**Figure 3.13**: Comparison of computational loads at SNR = 5 dB; (a) decorrelating, (b) two-level threshold $\epsilon = 0.1\sqrt{w_k}$, (c) two-level threshold $\epsilon = 0.2\sqrt{w_k}$, (d) two-level threshold $\epsilon = 0.3\sqrt{w_k}$, (e) two-level threshold $\epsilon = 0.4\sqrt{w_k}$, (f) two-stage (decorrelating 1st stage)

figures, the computational load of the two-level threshold detector, for all examined SNRs and $\epsilon$ is a little bit more and very close to that of the decorrelating detector. The huge gap between the computational load of the two-stage detector on one side

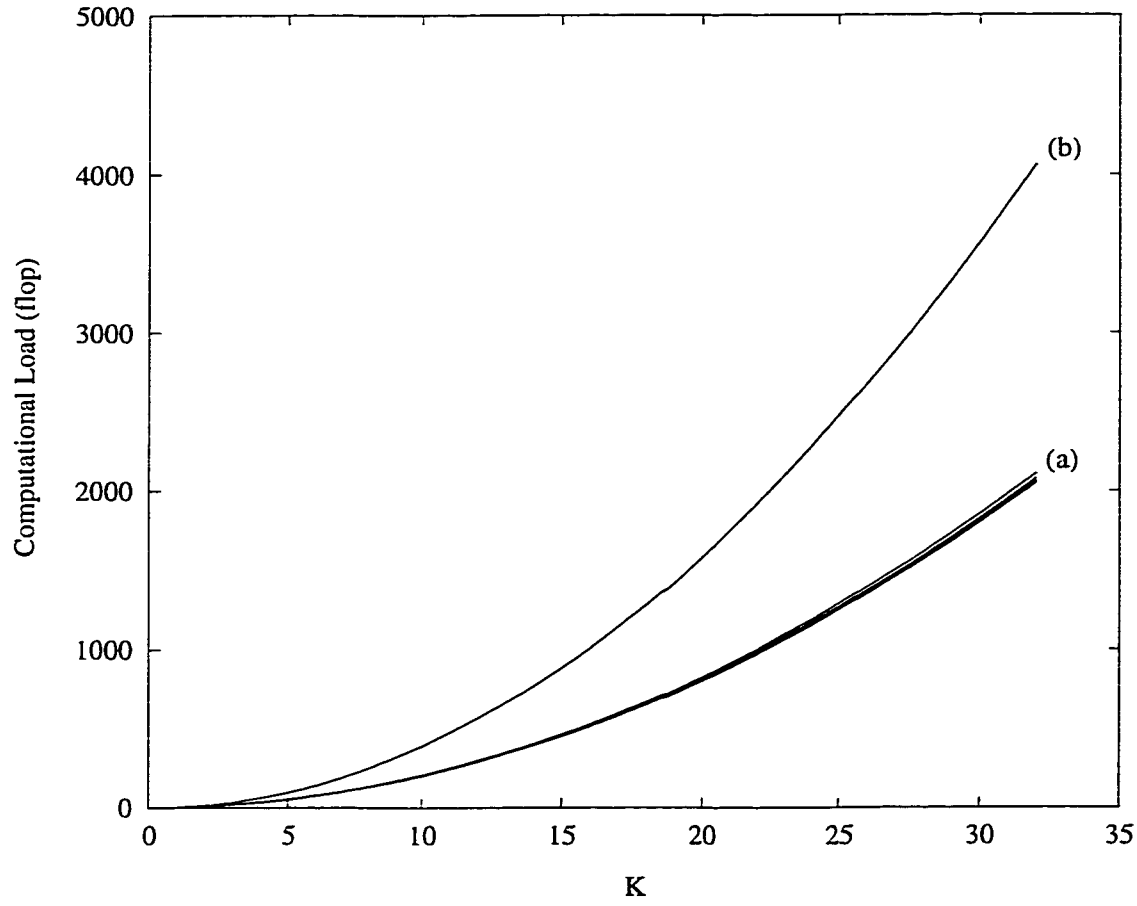


**Figure 3.14**: Comparison of computational loads at SNR = 10 dB; (a) decorrelating, two-level threshold ($\epsilon = 0.1\sqrt{w_k}$, $\epsilon = 0.2\sqrt{w_k}$, $\epsilon = 0.3\sqrt{w_k}$, $\epsilon = 0.4\sqrt{w_k}$), (b) two-stage (decorrelating 1st stage)

and the computational load of the new detector and the decorrelating detector on the other side is clear. This fact is more evident for higher SNRs, for instance 10 dB.

To show how the computational load of the new detector varies with respect to SNR, we have plotted $\bar{\mathcal{O}}_{16}(16)$ versus SNR in Figure 3.15. From Figure 3.15 one can see that as the SNR increases, the computational load the new detector decreases. Indeed for SNRs higher than 10 dB the computational load of the new detector is not differentiable from that of the decorrelating detector. It is worthwhile to mention

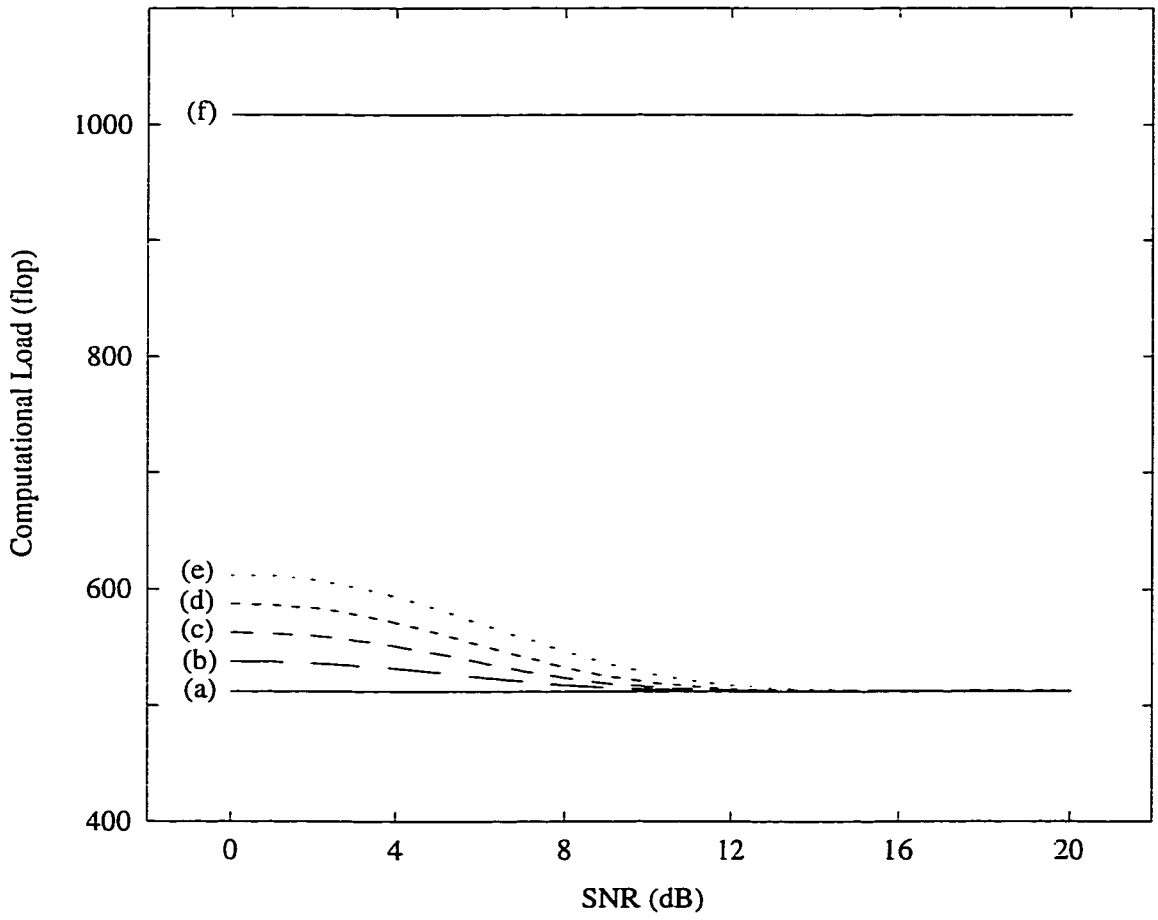


**Figure 3.15**: Comparison of computational loads where K = 16; (a) decorrelating, (b) two-level threshold $\epsilon = 0.1\sqrt{w_k}$, (c) two-level threshold $\epsilon = 0.2\sqrt{w_k}$, (d) two-level threshold $\epsilon = 0.3\sqrt{w_k}$, (e) two-level threshold $\epsilon = 0.4\sqrt{w_k}$, (f) two-stage (decorrelating 1st stage)

that as expected the computational load for both the decorrelating detector and the two-stage detector does not vary with SNR.

The above investigation suggests that generally the computational load of the two-level threshold detector is slightly more than that of the decorrelating detector, such that in most cases the difference is not considerable, whereas, the computational load of the two-stage detector is much more than both.

## 3.6 Some Comments

In this Chapter we have proposed a new algorithm for multiuser detection. Also, the probability of error and the computational load of the proposed detector are analyzed. The function of this detector can be described briefly as follows. If the magnitude of the output of the decorrelating filter is large enough, this detector performs exactly like a decorrelating detector. Otherwise, if the magnitude of the output of the decorrelating filter is small and in the vicinity of zero, it uses (3.22) to find $\hat{b}_k$. (3.22), however, follows almost the same approach that a multistage detector utilizes to maximize (3.13). Therefore, the position of the new detector amongst other multiuser detection schemes is somewhere between the decorrelating detector and the two-stage detector with a decorrelating detector as the first stage. The radius of the vicinity of zero, $\epsilon$, acts as a criterion to determine whether the outputs of the decorrelating filter are reliable or not. It also specifies whether the new detector performs similar to the decorrelating detector or to the two-stage detector. The smaller the $\epsilon$, the more the new detector acts like the decorrelating detector. On the other hand, the larger the $\epsilon$, the more the new detector acts like the two-stage detector.

There is still one fact remaining that is worthwhile to mention. Comparing Equation (3.22) with (2.54), one may notice that there is a slight difference between the maximization approach of the two-stage detector and that of the new detector. Let us consider the detection of $\hat{b}_k$. We assume $\hat{y}_k$ is in the neighborhood of zero, i.e. $\hat{y}_k \in [-\epsilon, \epsilon]$. Based on (2.54), to compute $\hat{b}_k$, the two-stage detector takes values for $\hat{b}_j$, $j = 1, \cdots, k-1, k+1, \cdots, K$, from the first stage, that is the outputs of the decorrelating detector. Whereas, according to (3.22), to compute $\hat{b}_k$, the two-level threshold detector takes values for $\hat{b}_j$, $j = 1, \cdots, k-1$, from the previous

detected bits, and for $\hat{b}_j$, $j = k+1, \cdots, K$, from the outputs of a decorrelating detector. This way we have tried to deploy the most recent information in hand to decrease the probability of error. This difference between the new detector and the two-stage detector is similar to the difference between the Jacobi method [60] and Seidel method [61] in solving the linear system $Ax = b$ iteratively. The former uses an iteration method by simultaneous displacement (same idea as the multi-stage detector), while the latter uses successive displacement (same idea as the new detector).

## 3.7 Discussion

In this chapter a suboptimum multiuser detection was introduced, whose complexity is linear in the number of users. The two-level threshold detector performs closely to the optimum detector, while its complexity is slightly more than that of the decorrelating detector.

The decorrelating filter removes completely the MAI, hence the outputs of the decorrelating filter contain only the desired signal embedded in noise. The likelihood ratio test in such a case suggests a simple threshold device with one level of threshold (assuming binary signaling). However, as shown the decision of the likelihood ratio test is not reliable when the outputs of the decorrelating filter are in the vicinity of zero. In such a case, we propose to use the multiuser maximum likelihood detector (optimum detector). However, to decrease the computational load we choose to maximize locally the likelihood expression.

The radius of the neighborhood of zero, i.e. $\epsilon$, plays an important role in the performance as well as in the computational complexity of the new detector. When $\epsilon = \infty$, the two-level threshold detector becomes the two-stage detector which uses

a decorrelating detector as the first stage. When $\epsilon = 0$, the two-level threshold detector is reduced to the standard decorrelating detector. Intuitively it seems that $\epsilon = \infty$ is an optimum value which yields the best bit-error-rate. In Chapter 5 we will see whether this is true or not.

Simulation results show that the two-level threshold detector can achieve a near-optimum BER. The increased computational load compared to the decorrelating detector is not that much and could be ignored at high SNRs. In Chapter 5, this detector is compared with the other multiuser detectors in a multipath channel. But before doing this, let us quickly characterize the multipath channel considered in this dissertation.

# Chapter 4

# Radio Propagation Channels

Thus far in this dissertation the multiuser detection for DS-CDMA systems over the AWGN channel has been considered. Given the fact that CDMA transmissions are frequently made over channels that exhibit fading and/or dispersion, it would seem appropriate to design receivers for such channels. Henceforth we consider multiuser detection in multipath channels. The mobile radio channel places important restrictions on the performance of wireless communication systems. The transmission path between the transmitter and the receiver may change from simple line-of-sight (LOS) to one that is obstructed by terrain, mountains or buildings. Unlike wired channels that are stationary and predictable, radio channels are random and are not easy to be analyzed. Modeling the radio channel has been one of the most complicated steps of mobile radio system design, and is usually done in statistical method, based on empirical measurements made particularly for a specific communication system or spectrum allocation.

This chapter has the following outline. Section 4.1 gives an introduction to radio wave propagation. Section 4.2 reviews the impulse response model of the multipath channels. In Section 4.3 we explain how one could obtain a discrete-time

impulse response model for a communication channel from its continuous impulse response. Section 4.4 introduces some useful parameters of multipath channels. Section 4.5 describes briefly different types of multipath fading. In Section 4.6 the two-ray fading model of the channel is explained, and finally in Section 4.7 a brief discussion is given.

## 4.1 Radio Wave Propagation

There are different mechanisms behind electromagnetic wave propagation, but generally they can be attributed to three phenomena, namely: reflection, diffraction, and scattering. Because of multiple reflections from various objects and obstacles, the electromagnetic waves travel along different paths with different lengths (Figure 4.1). The interaction between these waves causes multipath fading at a

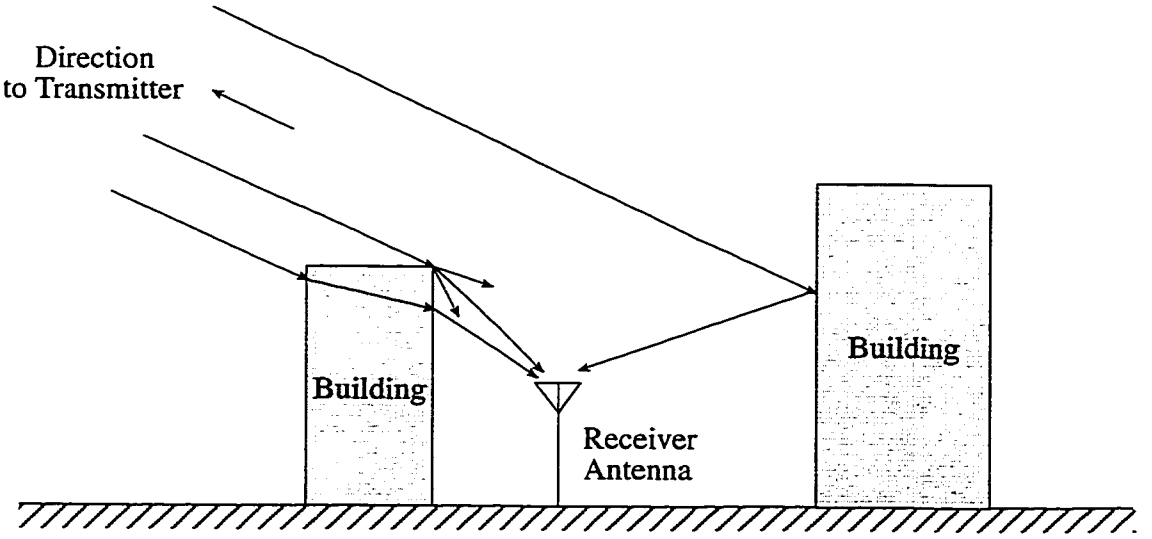


**Figure 4.1**: Illustration the mechanism of the multipath radio propagation.

specific location, and the strengths of the waves decreases as the distance between the transmitter and receiver increases. Propagation models have been considering

mobile channels from two standpoints:

- Predicting the average received power at a given distance from the transmitter; and
- Considering the variability of the signal power in close spatial proximity to the particular location.

Propagation models that predict the mean signal power for an arbitrary transmitter-receiver separation distance are useful in estimating the radio coverage area of a transmitter and are called large-scale propagation models. On the other hand, propagation models that characterize the rapid changes of the received signal strength over very short travel distances (a few wavelengths) or short time durations (on the order of seconds) are called small-scale or fading models (Figure 4.2).

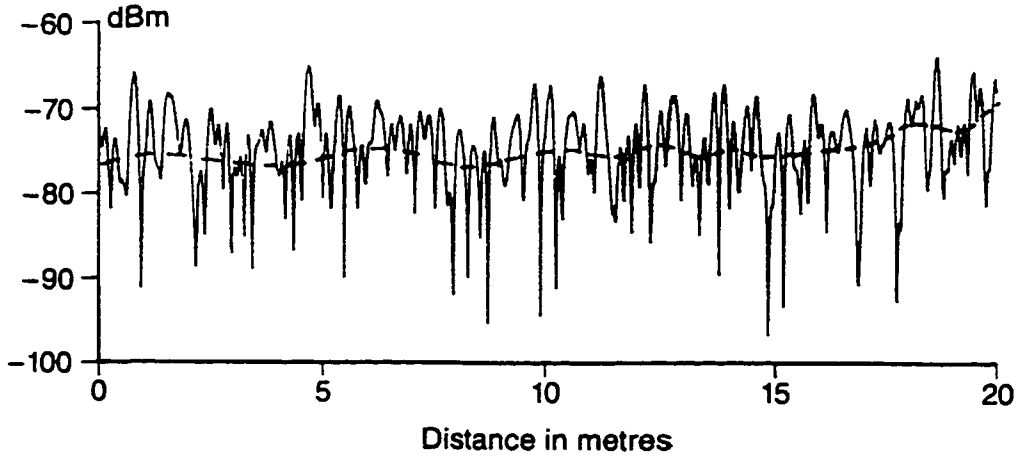


**Figure 4.2**: Experimental record of received signal envelope in an urban area which shows the small-scale (solid curve) and the large-scale (dashed curve) fading. (from [2])

Throughout this dissertation we have assumed that the average signal power is known, thus we can neglect the large-scale effects. In the following sections we will

concentrate on the small-scale or fading model of the channel.

### 4.1.1 Small-Scale Multipath Propagation

Small-scale fading explains the fast variation of the amplitude of a radio signal over a short period of time or travel distance. Fading is caused by interference between two or more versions of the transmitted signal which arrive at the receiver at slightly different times. These waves, called multipath waves, combine at the receiver antenna and result in a signal which may vary widely in amplitude and phase, depending on the distribution of the intensity and relative propagation time of the waves and the bandwidth of the transmitted signal.

In urban areas, fading occurs because the height of the mobile antennas are below the height of surrounding buildings, so there is no LOS path to the base station. Even when the LOS path exists, multipath still occurs due to reflections from the ground and other objects. The incoming radio waves arrive from different directions with different propagation delays. These multipath components, combine constructively or destructively at the receiver antenna, and cause the signal to distort or fade. Note that even when a mobile receiver is stationary, the received signal may fade because of movement of other objects in the radio channel. As a result of the relative motion between the mobile and the base station, each multipath wave experiences a shift in frequency. The shift in the received signal frequency due to motion is called a *Doppler* shift.

## 4.2 Impulse Response Model of a multipath Channel

The small-scale variations of a mobile radio signal can be directly related to the impulse response of the mobile radio channel. The impulse response contains all information necessary to simulate the channel. Indeed a mobile radio channel may be modeled as a linear filter with an impulse response that changes with respect to time [62]. The impulse response is a useful characterization of the channel, since it may be used to predict and compare the performance of many different mobile communication systems and transmission bandwidths for a particular mobile channel.

In many studies, where the multipath channel is researched (see e.g. [63]), it has been shown that due to the motion of a mobile receiver, or because of the motion of objects surrounding a stationary receiver, the received signal $r(t)$ can be expressed as a convolution of the transmitted signal $s(t)$ with a time-varying channel impulse response $h(t, \tau)$:

$$r(t) = \int_{-\infty}^{\infty} s(\tau) h(t, \tau) d\tau \tag{4.1}$$

The variable $t$ in (4.1) represents the time variations due to motion, whereas $\tau$ represents the channel multipath delay for a fixed value of $t$. If the multipath channel is assumed to be a bandlimited bandpass channel, then $h(t, \tau)$ may be equivalently described by a complex baseband impulse response $c(t, \tau)$ with the input and output being the complex envelope representations of the transmitted

and received signals, respectively [55,64]. That is,

$$v(t) = \int_{-\infty}^{\infty} u(\tau)c(t,\tau)d\tau \tag{4.2}$$

where $u(t)$ and $v(t)$ are lowpass equivalents of $s(t)$ and $r(t)$, respectively, where

$$s(t) = \Re\{u(t)e^{j2\pi f_c t}\} \tag{4.3}$$

$$r(t) = \Re\{v(t)e^{j2\pi f_c t}\} \tag{4.4}$$

$f_c$ is the carrier frequency and $\Re\{\cdot\}$ denotes a real part.

## 4.3 Discrete-Time Channel Impulse Response

It is useful to discretize the multipath delay axis $\tau$ of the channel impulse response into equal time delay segments called excess delay bins. In this case, each bin has a width equal to $\tau_{i+1} - \tau_i = \Delta\tau$. Let us set $\tau_i = i\Delta\tau$, for $i = 0$ to $L-1$, where $L$ represents the total number of possible equally-spaced multipath components, including the first arriving component. Any number of multipath signals received within the $i$th bin are represented by a single resolvable multipath component having a delay $\tau_i$. Note that $\tau_0 = 0$ is the excess time delay of the first arriving multipath component, and the propagation delay between the transmitter and receiver is ignored. The maximum excess delay of the channel is given by $L\Delta\tau$.

Since the received signal in a multipath channel consists of a series of attenuated, time-delayed, phase shifted replicas of the transmitted signal, the baseband impulse

response of a multipath channel can be expressed as

$$c(t,\tau) = \sum_{i=0}^{L-1} \alpha_i(t,\tau) \exp[j2\pi f_c \tau_i(t) + \phi_i(t,\tau)]\, \delta(\tau - \tau_i(t)), \tag{4.5}$$

where $\alpha_i(t,\tau)$ and $\tau_i(t)$ are the real amplitudes and excess delays, respectively, corresponding to the $i$th multipath component at time $t$. The phase term $2\pi f_c \tau_i(t) + \phi_i(t,\tau)$ in (4.5) represents the phase shift due to free space propagation of the $i$th multipath component. In general, the phase term is simply represented by a single variable $\theta_i(t,\tau)$. Note that some excess delay bins may have no multipath at some time $t$ and delay $\tau_i$, since $\alpha_i(t,\tau)$ may be zero. In other words there are no arrivals within some excess delay bins. In equation (4.5), $N$ is the total possible number of multipath components (bins), and $\delta(\cdot)$ is the unit impulse function. Figure 4.3 illustrates an example of different snapshots of $c(t,\tau)$.

If the channel impulse response is assumed to be time invariant, then the channel impulse response may be simplified as

$$c(\tau) = \sum_{i=0}^{L-1} \alpha_i e^{-j\theta_i} \delta(\tau - \tau_i) \tag{4.6}$$

For small-scale channel modeling, the *power delay profile* of the channel, $P(t;\tau)$, is defined as

$$P(t;\tau) \triangleq |c(t;\tau)|^2 \tag{4.7}$$

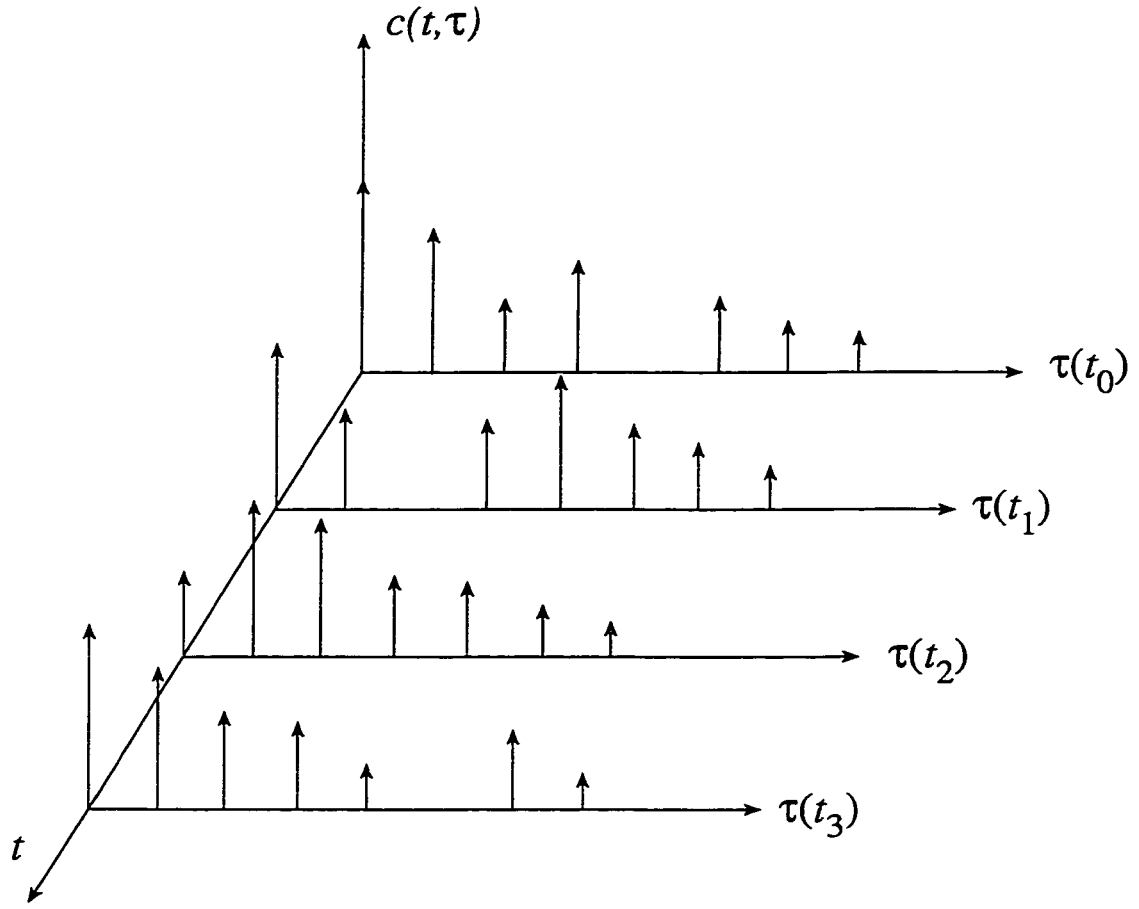


**Figure 4.3**: An example of a time varying discrete-time impulse response model for a multipath radio channel.

## 4.4 Some Parameters of Multipath Channels

Different multipath channel parameters are derived from the power delay profile, $P(\tau)$. Power delay profiles are found by averaging instantaneous power delay profile measurements over a local area in order to determine an average small-scale power delay profile. Figure 4.4 shows a typical power delay profile plot for an indoor channel, determined from a large number of sampled profiles.

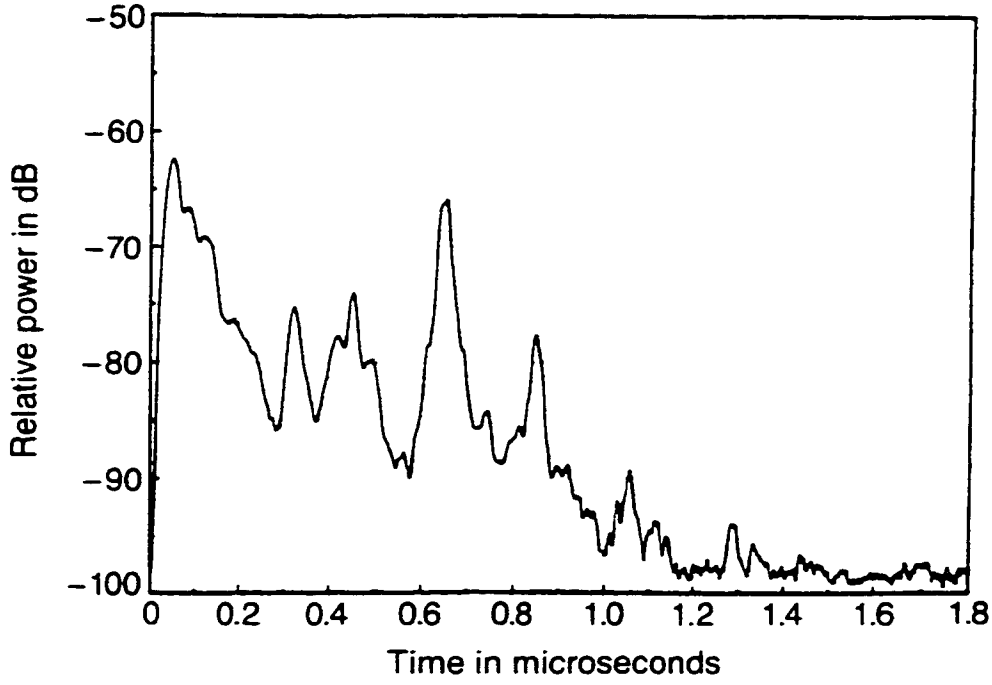


**Figure 4.4**: Measured power delay profile within a large six-story building [2].

### 4.4.1 Time Dispersion Parameters

To compare different multipath channels some parameters which quantify the multipath channel are used. The mean excess delay and rms delay spread are multipath channel parameters that can be determined from a power delay profile. The time dispersive properties of wideband multipath channels are most commonly quantified by their rms delay spread ($\tau_{rms}$) and mean excess delay ($\bar{\tau}$). The rms delay spread is the square root of the second central moment of the power delay spread and is defined by [65]

$$\tau_{rms} = \left( \frac{\sum_k (\tau_k - \bar{\tau})^2 \, \alpha_k^2}{\sum_k \alpha_k^2} \right)^{1/2} \tag{4.8}$$

| Environment | Frequency (MHz) | rms Delay Spread ($\tau_{rms}$) | Notes | Reference |
|---|---|---|---|---|
| Urban | 910 | 1300 ns avg.<br>600 ns st. dev.<br>3500 ns max. | New York City | [66] |
| Urban | 892 | 10-25 $\mu$s | San Francisco | [67] |
| Suburban | 910 | 200-310 ns | Averaged typical case | [68] |
| Suburban | 910 | 1960-2110 ns | Averaged extreme case | [68] |
| Indoor | 1500 | 10-20 ns | Office building | [69] |
| Indoor | 850 | 270 ns max. | Office building | [70] |
| Indoor | 1900 | 70-94 ns avg.<br>1470 ns max. | Three San Francisco buildings | [71] |

**Table 4.1**: Typical measured values of rms delay spread. (from [1, page 162])

Where $\bar{\tau}$ is the mean excess delay defined as:

$$\bar{\tau} = \frac{\sum_k a_k^2 \tau_k}{\sum_k a_k^2} = \frac{\sum_k P(\tau_k)\tau_k}{\sum_k P(\tau_k)} \tag{4.9}$$

The above expression shows that $\bar{\tau}$ is the first moment of the power delay profile. These delays are measured relative to the first detectable signal arriving at the receiver at $\tau_0 = 0$. Equations (4.8) and (4.9) do not rely on the absolute power level of $P(\tau)$, but only the relative amplitudes of the multipath components within $P(\tau)$. Typical values of rms delay spread are on the order of microseconds in outdoor mobile radio channels and on the order of nanoseconds in indoor radio channels. Table 4.1 shows the typical values of rms delay spread based on the empirical measurements.

It should be noted that the power delay profile and the magnitude frequency response of a mobile radio channel are related through the Fourier transform. Therefore, it is possible to obtain an equivalent description of the channel in the frequency

domain using its frequency response characteristics. Analogous to the delay spread parameters in the time domain, coherence bandwidth is used to characterize the channel in the frequency domain. The rms delay spread and coherence bandwidth are inversely proportional to one another, although their exact relationship is a function of the exact multipath structure.

### 4.4.2 Coherence Bandwidth

Coherence bandwidth, $B_c$ is a statistical measure of the range of frequencies over which the channel can be considered *flat*. In other words, coherence bandwidth is the range of frequencies over which two frequency components have a strong potential for amplitude correlation. Two sinusoids with frequency separation greater than $B_c$, are affected quite differently by the channel. The coherence bandwidth is usually defined as the reciprocal of the rms delay spread [72], that is,

$$B_c \approx \frac{1}{\tau_{rms}} \tag{4.10}$$

It is worthwhile to note that an exact relationship between coherence bandwidth and rms delay spread does not exist, and Equation (4.10) is a rough estimate. Generally, spectral analysis techniques and simulation are required to determine the exact impact that time varying multipath has on a particular transmitted signal [73, 74]. For this reason, accurate multipath channel models must be used in the design of specific modems for wireless applications [75, 76].

## 4.5 Various Small-Scale Fadings

Depending on various signal parameters (such as bandwidth, symbol period, etc.) and various channel parameters (such as rms delay spread and Doppler spread), different transmitted signals will undergo different types of fading. The time dispersion and frequency dispersion mechanisms in a mobile radio channel cause four possible effects. While multipath delay spread leads to time dispersion and frequency selective fading, Doppler spread leads to frequency dispersion and time selective fading. The two propagation mechanisms are independent of one another.

### 4.5.1 Time Delay Spread Fading Effects

Time dispersion due to multipath causes the transmitted signal to undergo either flat or frequency selective fading.

#### Flat Fading

If the mobile radio channel has a constant gain and a linear phase response over a bandwidth which is greater than the bandwidth of the transmitted signal, then the received signal will undergo flat fading. In flat fading, the multipath structure of the channel is such that the spectral characteristics of the transmitted signal are preserved at the receiver. However, the strength of the received signal changes with time, due to variations in the gain of the channel. The distribution of the instantaneous gain of flat fading channels is important in designing radio links, and the most common amplitude distribution is the Rayleigh distribution. The Rayleigh flat fading channel model assumes an amplitude which varies in time according to

the Rayleigh distribution. To summarize, a signal undergoes flat fading if

$$B_s \ll B_c \qquad \text{and} \qquad T_s \ll \tau_{rms} \tag{4.11}$$

where $T_s$ is the reciprocal bandwidth (e.g. symbol period) and $B_s$ is the bandwidth of the transmitted modulation.

### Frequency Selective Fading

If the channel has a constant-gain and a linear phase response over a bandwidth that is smaller than the bandwidth of the transmitted signal, then the channel is referred to as frequency selective fading. Under such conditions the channel impulse response has a multipath delay spread which is greater than the reciprocal bandwidth of the transmitted message waveform. When this occurs, the received signal includes multiple versions of the transmitted waveform which are faded and delayed in time. Frequency selective fading is due to time dispersion of the transmitted symbols within the channel. Thus the channel induces ISI. When analyzing mobile communication systems, statistical impulse response models or computer generated or measured impulse responses are generally used for analyzing frequency selective small-scale fading. An example of a statistical model is the two-ray Rayleigh fading model, which considers the impulse response to be made up of two delta functions. The delta functions independently fade and have sufficient time delay between them to induce frequency selective fading upon the applied signal.

To summarize, a signal undergoes frequency selective fading if

$$B_s > B_c \qquad \text{and} \qquad T_s < \tau_{rms} \tag{4.12}$$

A common rule of thumb is that a channel is frequency selective if $\tau_{rms} > 0.1T_s$, although this is dependent on the specific type of modulation used.

### 4.5.2 Doppler Spread Fading Effects

Depending on how rapidly the transmitted baseband signal changes as compared to the rate of change of the channel, a channel may be classified either as fast fading or slow fading.

#### Fast Fading

In a fast fading channel, the channel impulse response changes rapidly within one symbol duration. That is, the coherence time of the channel is smaller than the symbol period of the transmitted signal. Therefore, a signal undergoes fast fading if

$$T_s > T_c \qquad \text{and} \qquad B_s < B_d \tag{4.13}$$

where $T_c$ is the coherence time usually defined as

$$T_c \approx \frac{1}{B_d} \tag{4.14}$$

#### Slow Fading

In a slow fading channel, the channel impulse response changes at a rate much slower than the transmitted baseband signal $u(t)$. In this case, the channel may be assumed to be static over one or several reciprocal bandwidth intervals. Thus, a

signal undergoes slow fading if

$$T_s \ll T_c \qquad \text{and} \qquad B_s \gg B_d \tag{4.15}$$

It should be clear that the velocity of the mobile (or velocity of objects in the channel) and the baseband signaling determines whether a signal undergoes fast fading or slow fading.

## 4.6 Two-ray Rayleigh Fading Model

In modern mobile communication systems with high data rates, it has become necessary to model the effects of multipath delay spread as well as fading. A commonly used multipath model is an independent Rayleigh fading two-ray model. The two-ray model is often used in the theory of mobile communications, since it can represent most important properties of a channel while still being simple enough to allow analytic computations as well as simulations (see e.g. [77–81]). Despite these advantages, we have to keep in mind that the two-ray model is only a first order approximation and usually does not provide an exact physical description of an arbitrary mobile radio channel. Figures 4.5 and 4.6 show the impulse response and a block diagram of the two-ray independent Rayleigh fading channel model, respectively. The impulse response of the model is represented as

$$c(t) = \alpha_1 e^{j\phi_1}\delta(t) + \alpha_2 e^{j\phi_2}\delta(t-\tau_0) \tag{4.16}$$

where $\alpha_1$, and $\alpha_2$, are independent and Rayleigh distributed, $\phi_1$ and $\phi_2$ are independent and uniformly distributed over [0, $2\pi$],and $\tau$ is the time delay between the

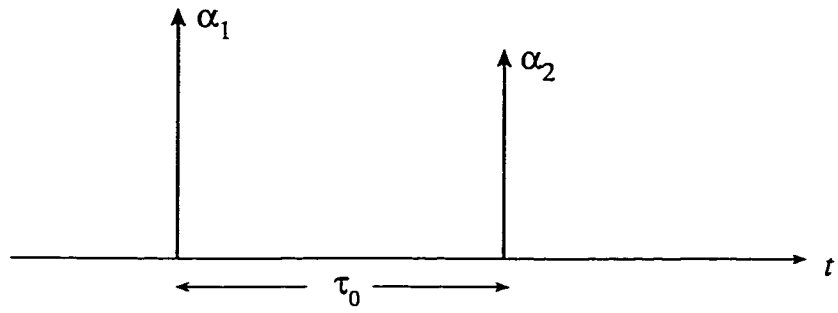


**Figure 4.5**: Impulse response of the two-ray Rayleigh fading model of channel

two rays. By setting $\alpha_2$ equal to zero, the special case of a flat Rayleigh fading channel can obtained as

$$c(t) = \alpha_1 e^{j\phi_1}\delta(t) \tag{4.17}$$

By varying $\tau_0$ as well as $\alpha_1$ and $\alpha_2$ in (4.16), it is possible to create a wide range

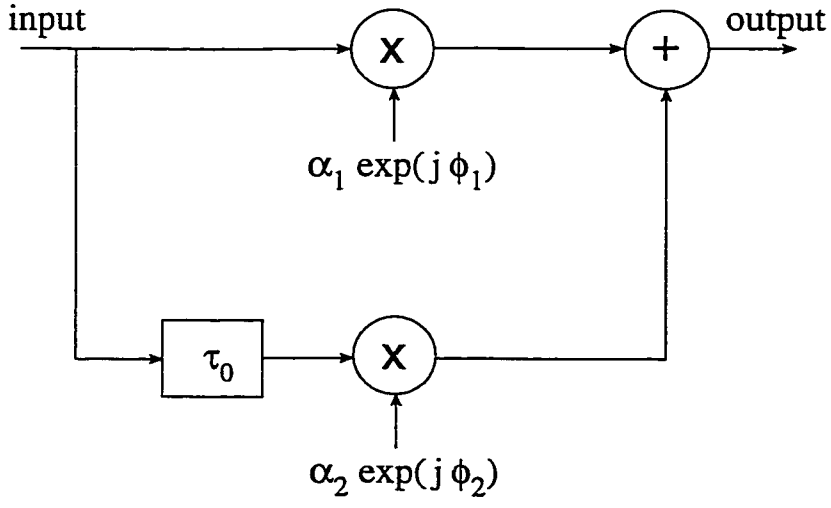


**Figure 4.6**: Two-ray Rayleigh fading model

of frequency selective fading effects. Figure 4.7 shows the instantaneous frequency response of the two-ray fading model. From (4.8) we see that the rms delay spread

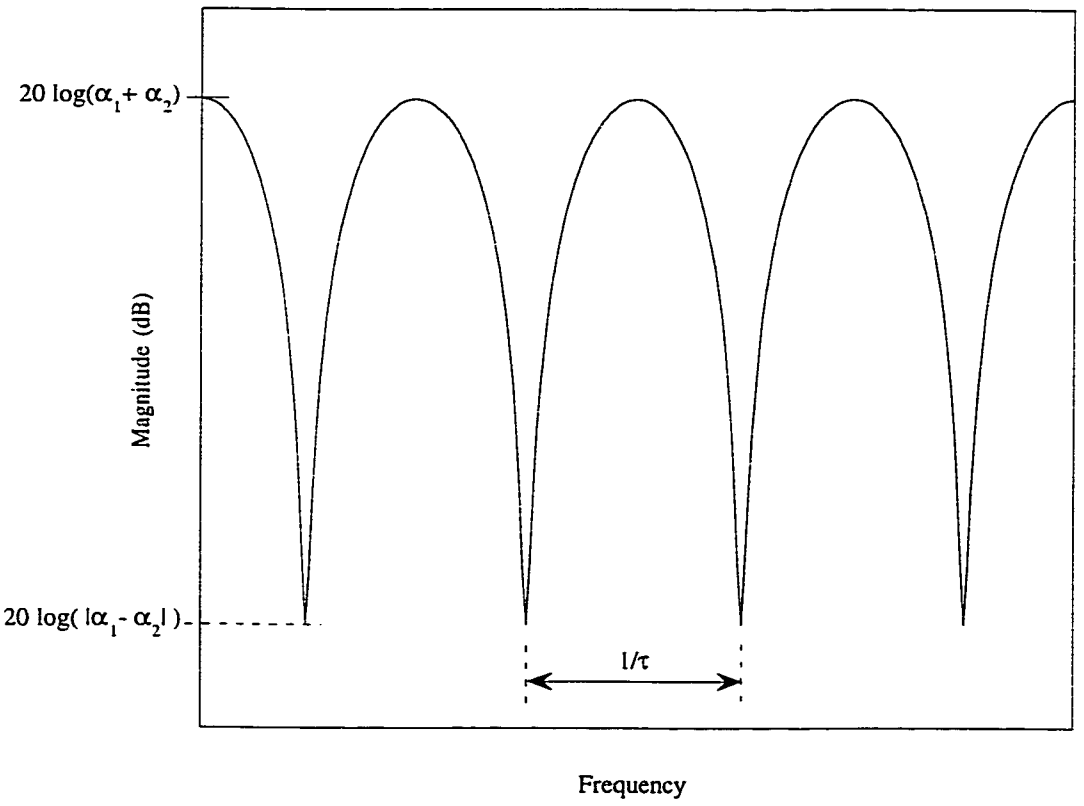


**Figure 4.7**: Frequency response of the two-ray fading model of channel

is proportional to $\tau_0$, i.e. the larger $\tau_0$, the larger $\tau_{rms}$. On the other hand, in Figure 4.7, we observe that the larger $\tau_0$, the smaller inter-fade bandwidth. Therefore, by increasing $\tau_0$, we can create a severe frequency selective channel. $\alpha_1$ and $\alpha_2$ may also be used to change the rms delay spread, i.e. the amount of frequency selectivity. But as we observe in Figure 4.7, these values have a more direct impact on the depth of the fades than on the frequency selectivity. Of course, this model has some drawbacks. The major drawback is that it cannot create fades with different depths. Also the fades are equally spaced in frequency during one snapshot of the frequency response.

Nonetheless, the simplicity of the two-ray fading model, in terms of analytic computations and computer simulations, motivates us to use it in our channel simulations. It is important to recall that our intention in this dissertation is to compare the performance of various multiuser detectors. Therefore, a very precise model of the channel does not seem necessarily crucial.

## 4.7 Discussion

In this chapter we reviewed quickly the radio propagation channels. The discrete-time impulse response seems an appropriate way to model a multipath fading channel, especially when the analyses and/or the simulations are performed in time domain. Despite its drawbacks, a two-ray Rayleigh model can achieve a large variety of multipath fading channels from flat fading to frequency selective fading. Since in this dissertation, wherever we deal with a multipath channel, our goal is to compare different multiuser detectors is such a environment, a two-ray fading model appears to be suitable for modeling multipath channels.

# Chapter 5

# Multiuser Detection in Multipath Environment

In Chapter 2 various types of suboptimum multiuser detection schemes as well as the optimum and the conventional detectors are studied. In Chapter 3 we introduced a new suboptimum multiuser detector, namely the two-level threshold detector, as well. We studied these detectors in the AWGN channel. Simulation results show that the BER performance of the detectors are close provided that the spreading sequences have good cross-correlation properties.

In the present chapter the multiuser detection in multipath channels is investigated. We concentrate on frequency selective channels and use a two-ray model for such channels as it was described in the previous chapter. The channel parameters are assumed to be known perfectly. A general strategy, namely channel-matched approach, is used to modify the multiuser detection schemes and then deploy them in multipath environment. In this strategy the bank of matched filters at the front-end of the multiuser detectors are matched to the channel. Correspondingly, the relevant parameters of the detectors are changed compared to the detectors for the

AWGN channel.

The rest of this chapter is organized as follows. In Section 5.2 previous work in this area is summarized. Section 5.3 describes the contributions of the chapter. In Section 5.3 the system model is presented. Section 5.4 provides us with two different general approaches to design multiuser detections for multipath channels. In Section 5.5 we compare these two approaches and choose the one with a better BER performance. Section 5.6 explains the model for the multipath channel that we have used for simulations. In Section 5.7 the simulation platform is reviewed. In Section 5.8 all multiuser detectors are deployed in a multipath channel and their performances are compared. Section 5.9 explains how we should choose the optimum radius of the neighborhood of zero for the two-level threshold detector to obtain a low BER while maintaining a low computational load. Section 5.10 presents a discussion and summarizes the results

## 5.1 Previous Work

In the past six years multiuser detection theory has been extended to handle multipath fading channels. Zvonar and Brady [82] have developed detectors for a slowly Rayleigh fading channel. They used a single-path fading model for the channel. Later on they proposed the optimum multiuser detection and a linear suboptimum multiuser detection scheme for frequency-selective Rayleigh fading channels [83–86]. Vesudevan and Varanasi [87–89] have developed detectors for a Rician fading channel. Fawer and Aazhang [90, 91] generalized the idea of the multistage detector to multipath channels. Zvonar used the suboptimal detector in [84] and proposed a combined multiuser detection and diversity reception scheme for frequency-selective channels [92, 93].

## 5.2 Contribution

This part of the dissertation provides a thorough study of different multiuser detectors in multipath fading channels. The channel-matched approach is applied to the previously proposed multiuser detectors besides the two-level threshold detector proposed in this dissertation. Simulation results are used to compare the performances of the multiuser detectors for a 16-user DS-CDMA system, which employs Gold sequences of length 31 as the spreading sequences.

## 5.3 System Model

As we saw in the previous chapter, in mobile communication environments, since there are many propagation paths with different delays between the transmitter and the receiver, the transmitted signal components corresponding to these multipath propagation paths arrive at different times.

We assume a multiuser system in a multipath environment. The impulse response of the channel for the $k$th user, as we saw in Chapter 4, Section 4.3, can be represented by a tapped delay line given by

$$c_k(t) = \sum_{l=1}^{L} c_{k,l}\delta(t - \tau_{k,l}) \tag{5.1}$$

Where $L$ is the number of the propagation paths and $c_{k,l}$ is the normalized fading complex envelope of the $k$th user encountered with path $l$. There are $K$ simultaneous users in the system being considered. Each of the simultaneous users transmits its information symbol using the DS-CDMA signalling scheme. Therefore, the base-

band received signal can be expressed as

$$r(t) = \sum_{i=-\infty}^{\infty} \sum_{k=1}^{K} \sum_{l=1}^{L} b_k(i) c_{k,l}(i) e^{j\psi_k} \sqrt{w_k(i)} s_k(t - iT - \tau_{k,l}) + n(t) \qquad (5.2)$$

where $n(t)$ is a zero-mean additive white Gaussian noise with power spectral density $\sigma^2$. $s_k(t)$ is the normalized spreading sequence and is zero outside the interval $[0, T]$, with $T$ being the information duration. $b_k(i)$ is the information symbol transmitted by user $k$ within epoch $i$, and $i \in \{-\infty, \infty\}$ is the time interval index. $\psi_k$ is the $k$th user's phase and $w_k(i)$ is the $k$th user's received energy. $w_k(i)$, indeed, reflects the normalization on $c_{k,l}(i)$'s, so that

$$\sum_{l=1}^{L} c_{k,l}(i) c_{k,l}^*(i) = 1, \qquad \text{for} \qquad 1 \leq k \leq K \qquad (5.3)$$

where $*$ denotes the complex conjugate. In Equation (5.2) $\tau_{k,l}$ is the delay on $c_{k,l}(i)$. Without loss of generality we assume that $\tau_{1,1} = 0$ and other $\tau_{k,l}$'s lie within the range of $[0, T]$.

## 5.4 Approaches towards the detection

There are various methods considering the multiuser detection in multipath fading environment. Generally, these methods follow two different approaches: i) Path-by-path approach, or ii) Channel-matched approach. In this section we explain these two approaches. In order to do so, we choose one of the multiuser detection schemes described in Chapter 2, e.g. the decorrelating detector. In the primary paper [94] that explicitly compared these two approaches, the decorrelating detector was investigated, and that is the reason why we are using this detector as an example

now.

### 5.4.1 Path-by-path

The path-by-path approach assumes that each of the received signal corresponding to the propagation paths is an independent interference, [84,95,96]. We consider the system model in (5.2), and assume a slowly fading channel, which implies that $c_{k,l}(i)$ does not change during a symbol transmission time. Also, we assume $T \gg \tau_{rms}$, i.e. any intersymbol interference (ISI) due to the channel dispersion may be neglected [55]. Having these assumptions we consider the received signal $r(t)$ over only one symbol duration, e.g. $i = 0$. The noiseless signal from $k$th user is given by

$$g_k(t) = \sum_{l=1}^{L} c_{k,l} e^{j\psi_k} \sqrt{w_k} s_k(t - \tau_{k,l}) = \boldsymbol{s}_k^T(t) W_k \Psi_k \boldsymbol{c}_k \tag{5.4}$$

Using matrix notation, the single-user channel vector, $\boldsymbol{c}_k$, is defined as

$$\boldsymbol{c}_k = [c_{k,1}, c_{k,2}, \cdots, c_{k,L}]^T \tag{5.5}$$

the single-user normalized spreading sequence vector , $\boldsymbol{s}_k(t)$, is defined as

$$\boldsymbol{s}_k(t) = [s_k(t), s_k(t - \tau_{k,2}), \cdots, s_k(t - \tau_{k,L})]^T \tag{5.6}$$

$W_k = \sqrt{w_k}\mathcal{I}_L$, and $\Psi_k = e^{j\psi_k}\mathcal{I}_L$ ($\mathcal{I}_L$ being the $L \times L$ identity matrix). Now we can simplify the received signal as

$$r(t) = \sum_{k=1}^{K} b_k g_k(t) + n(t) = \boldsymbol{g}^T(t)\boldsymbol{b} + n(t), \qquad 0 \le t \le T \tag{5.7}$$

where $\boldsymbol{b} = [b_1, b_2, \cdots, b_K]^T$ is a vector of data symbols from $K$ users, and $\boldsymbol{g}(t) = [g_1(t), g_2(t), \cdots, g_K(t)]^T$. The equivalent spreading waveform vector can be expressed as $\boldsymbol{g}^T(t) = \boldsymbol{s}^T(t) W_p \Psi_p C_p$ with the normalized spreading waveform vector, $\boldsymbol{s}(t)$, defined as

$$\boldsymbol{s}(t) = [\boldsymbol{s}_1^T(t), \boldsymbol{s}_2^T(t), \cdots, \boldsymbol{s}_K^T(t)]^T. \tag{5.8}$$

The composite matrices, $W_p$ and $\Psi_p$, defined as

$$W_p = diag(W_1, W_2, \cdots, W_K), \tag{5.9}$$

$$\Psi_p = diag(\Psi_1, \Psi_2, \cdots, \Psi_K), \tag{5.10}$$

and the $KL \times K$ multi-channel matrix, $C_p$, defined as

$$C_p = \begin{pmatrix} \boldsymbol{c}_1 & 0 & \cdots & 0 \\ 0 & \boldsymbol{c}_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \boldsymbol{c}_K \end{pmatrix}. \tag{5.11}$$

The path-by-path decorrelating detector passes the received signal $r(t)$ through a bank of $KL$ filters matched to the delayed, normalized spreading sequences of the users. The outputs are sampled at the symbol interval $T$ (Figure 5.1). It can be shown that the resulting $KL$ vector may be expressed as

$$\boldsymbol{y}_p = R_p W_p \Psi_p C_p \boldsymbol{b} + \boldsymbol{n}, \tag{5.12}$$

where the $KL \times KL$ cross-correlation matrix of normalized spreading sequences,

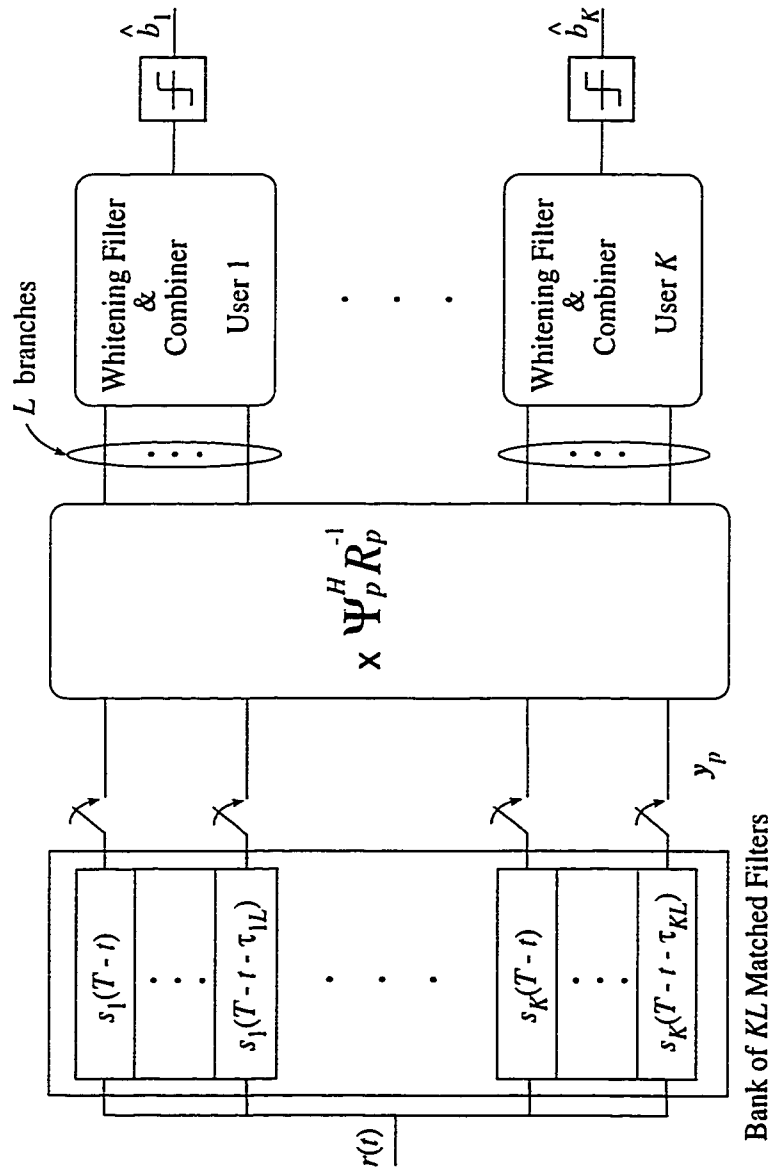


**Figure 5.1**: Path-by-path decorrelating detector structure.

$R_p$, is defined as

$$R_p = \int_0^T s(t)s^T(t)dt. \tag{5.13}$$

Let $z_p = \Psi_p^H R_p^{-1} y_p$ be the outputs of the multipath signal decorrelator (Figure 5.1), where the matrix superscript $H$ denotes conjugate transpose. For coherent BPSK signaling, we face the single-user problem of optimally combining independently fading signals in correlated noise. Path-by-path detector combines the different fading paths prior to decision making. To do this, we whiten the noise with the filter $(T_k^H)^{-1}$. This filter is obtained through the Cholesky decomposition of the $L \times L$ matrix $\bar{R}_{p[k,k]} = T_k^H T_k$, where $\bar{R}_{p[k,k]}$ is a partition of $R_p$ partitioned as follow

$$R_p = \begin{pmatrix} \bar{R}_{p[1,1]} & \bar{R}_{p[1,2]} & \cdots & \bar{R}_{p[1,K]} \\ \bar{R}_{p[2,1]} & \bar{R}_{p[2,2]} & & \vdots \\ \vdots & & \ddots & \bar{R}_{p[K-1,K]} \\ \bar{R}_{p[K,1]} & \cdots & \bar{R}_{p[K,K-1]} & \bar{R}_{p[K,K]} \end{pmatrix}. \tag{5.14}$$

Following the whitening filter, maximal-ratio combining is performed for each user by weighting signals proportionally to their strength and coherently combining them prior to threshold detection. A comprehensive study of the path-by-path approach is out of the scope of this dissertation. In [84,97] one can find detailed discussions regarding a path-by-path decorrelating detector.

### 5.4.2 Channel-matched

The other type of decorrelator combines all the received signal components transmitted from the identical users prior to decorrelation. Let us define the distorted spreading sequences, $h_k(t)$ $(k = 1, \cdots, K)$, as

$$h_k(t) = \sum_{l=1}^{L} c_{k,l} s_k(t - \tau_{k,l}) = \boldsymbol{s}_k^T(t)\boldsymbol{c}_k \tag{5.15}$$

If we follow the general arguments given in Section 2.10 for a BPSK linear multiuser detector but use $h_k(t)$ as the $k$th user's spreading sequence, we can derive the channel-matched decorrelating detector which uses a bank of $K$ matched filters. See Figure 5.2 for a block diagram of this receiver. Note that $h_k(t)$ are not necessarily

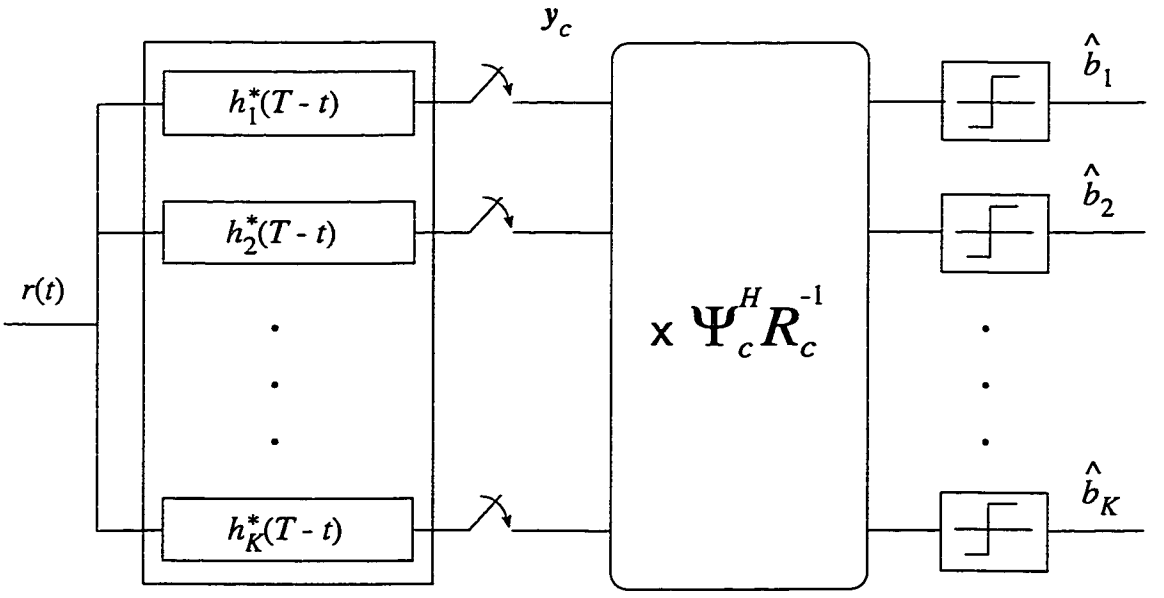


**Figure 5.2**: Channel-matched decorrelating detector structure.

normalized. The output of this bank is sampled every $T$ seconds which yields a $K$

vector $\boldsymbol{y}_c$ whose $k$-th element, $y_{c,k}$, is defined as

$$y_{c,k} = \int_0^T r(t)\, h_k^*(t)dt, \qquad k = 1, \cdots, K \tag{5.16}$$

We can alternatively write

$$\boldsymbol{y}_c = R_c \Psi_c W_c \boldsymbol{b} + \boldsymbol{n}_c \tag{5.17}$$

where $R_c$ is the cross-correlation matrix of the distorted spreading sequences $h_k(t)$, $k = 1, \cdots, K$, defined as

$$[R_c]_{i,j} = \int_0^T h_i^*(t) h_j(t)dt, \qquad i, j = 1, \cdots, K \tag{5.18}$$

and

$$\begin{aligned} W_c &= diag(\sqrt{w_1}, \sqrt{w_2}, \cdots, \sqrt{w_K}), \\ \Psi_c &= diag(e^{-j\psi_1}, e^{-j\psi_2}, \cdots, e^{-j\psi_K}), \end{aligned} \tag{5.19}$$

and $\boldsymbol{n}_c$ is a zero-mean complex Gaussian vector.

The phase matrix $\Psi_c$ represents the effects of the carrier phase. In Chapter 2 we had assumed that the carrier phases are known perfectly, thus we did not take them into consideration. In the present chapter we assume that $\Psi_c$ is known perfectly too. However, the reason that we consider it here is that in the next two chapters we study channel estimation and demonstrate that the carrier phase matrix, $\Psi_c$, can be considered as part of the channel. In other words, it can be estimated jointly along with the other channel parameters.

Having the above fact in mind, we may notice that the similarity between (5.17)

and (2.28) is evident. The channel-matched decorrelating detector in fact follows the same technique that was discussed in Chapter 2 for the decorrelating detector. It applies the decorrelating filter $\Psi_c^H R_c^{-1}$ to the outputs of the bank of matched filters and then takes the signum of the decorrelating filter outputs.

## 5.5 Comparison of two approaches

In [94] one can find an analytic comparison between the path-by-path approach and the channel-matched approach. However, [94] applies these two approaches to a decorrelating detector (as we did in the previous section) and compares them in terms of performance, i.e. probability of error. Here we briefly compare these two methods in term of both complexity and performance. Since $R_c$ is a $K \times K$ matrix the channel-matched approach requires the inversion of a $K \times K$ matrix which depends on the channel coefficients. Hence a matrix inversion will most likely be required within the coherence time of the channel, i.e. the time interval that the channel can be assumed to be stationary. We will discuss this problem in Chapter 6. On the other hand, the $KL \times KL$ matrix inversion for the path-by-path approach (note that $R_p$ is a $KL \times KL$ matrix) is not dependent on the channel coefficients and only need to be determined once the set of active spreading sequences changes. However, the maximal ratio combiners at the output of the whitening filters depend on the channel coefficients and needs to be updated within the coherence time of the channel. A comprehensive comparison in terms of complexity is out of the scope of this dissertation.

With respect to the performance, it can be shown that for a given channel realization the probability of error of the channel-matched decorrelating detector is smaller than that of the path-by-path one. While a mathematical proof can be

found in [94], we should be able to accept it with only the following reasoning: first, realize that the decorrelating detector discussed in Chapter 2 for single path signals provides optimum (in the maximum likelihood sense) linear multiuser detector performance when users' energies are unknown. Now, since both the path-by-path and channel-matched methods are linear detectors which do not require knowledge of the users' energies and since the channel-matched detector can be thought of as the multipath signal analog of the decorrelating detector (if we consider $h_k(t)$ to be the $k$th user's spreading sequence), the channel-matched approach should therefore provide the optimum performance over all linear multiuser detectors for multipath signals of unknown energy. Hence our claim immediately follows. Due to this fact we henceforth concentrate on the channel-matched approach. In the following sections, we apply the channel-matched method to the detectors which we investigated before in Chapter 2 for AWGN channels. For all such detectors we assume $h_k(t)$ as the $k$th user's spreading code and follow the same concept as in Chapter 2. Before presenting the simulation results, let us explain the channel model that we have considered throughout the simulations.

## 5.6 Multipath Channel Model

The multipath channel model, which we use in the following simulations, is the two-ray model described in Section 4.6. The impulse response of the two-ray model is shown in Figure 5.3.

Since the objective of the investigations and simulations in the current chapter is to study the detectors' performances and not their sensitivity to the channel variations with time, we consider channel rays of fixed amplitudes. This way indeed we have assumed that the receiver estimates the channel parameters perfectly and

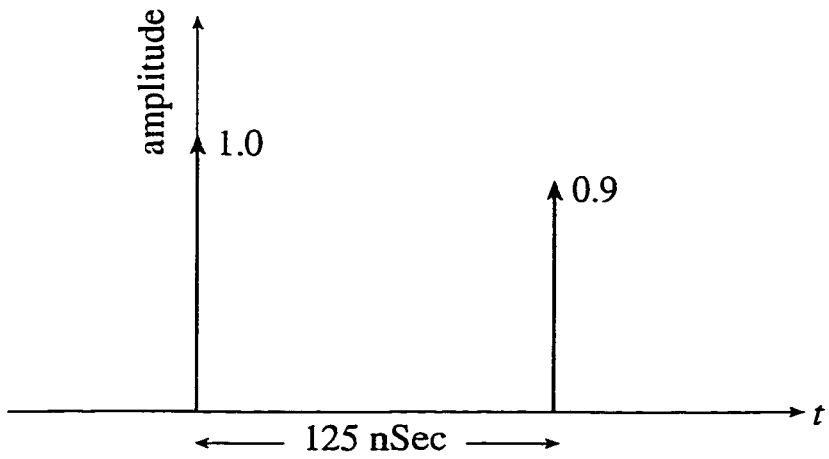


**Figure 5.3**: Impulse response of the 2-ray model of the channel ($\tau_{rms} \approx$ 62nSec).

also that the channel parameters do not change at all within two consecutive channel estimations ( We will discuss the channel estimation further in the Chapters 6 and 7).

The corresponding channel frequency response is shown in Figure 5.4 for a 16 MHz frequency band centered at 950 MHz. The center frequency, $f_c = 950$MHz, is arbitrary and does not affect the results of the simulations, since in fact the simulations are performed in baseband equivalent of the system. The 16MHz bandwidth is twice as much as the passband bandwidth of the system being considered.

## 5.7 Simulation Model

Figure 5.5 depicts the simulation model used in this Chapter. Binary data are modulated using BPSK signaling. Then the modulated symbols are spread by spreading codes assigned to each user. The scaling factor $\sqrt{w_k}$, $k = 1, \cdots, K$, merely reflects the energy of user $k$. The spread signal is passed through a multipath channel. As we discussed in Section 5.6, a two-ray model is used to model the

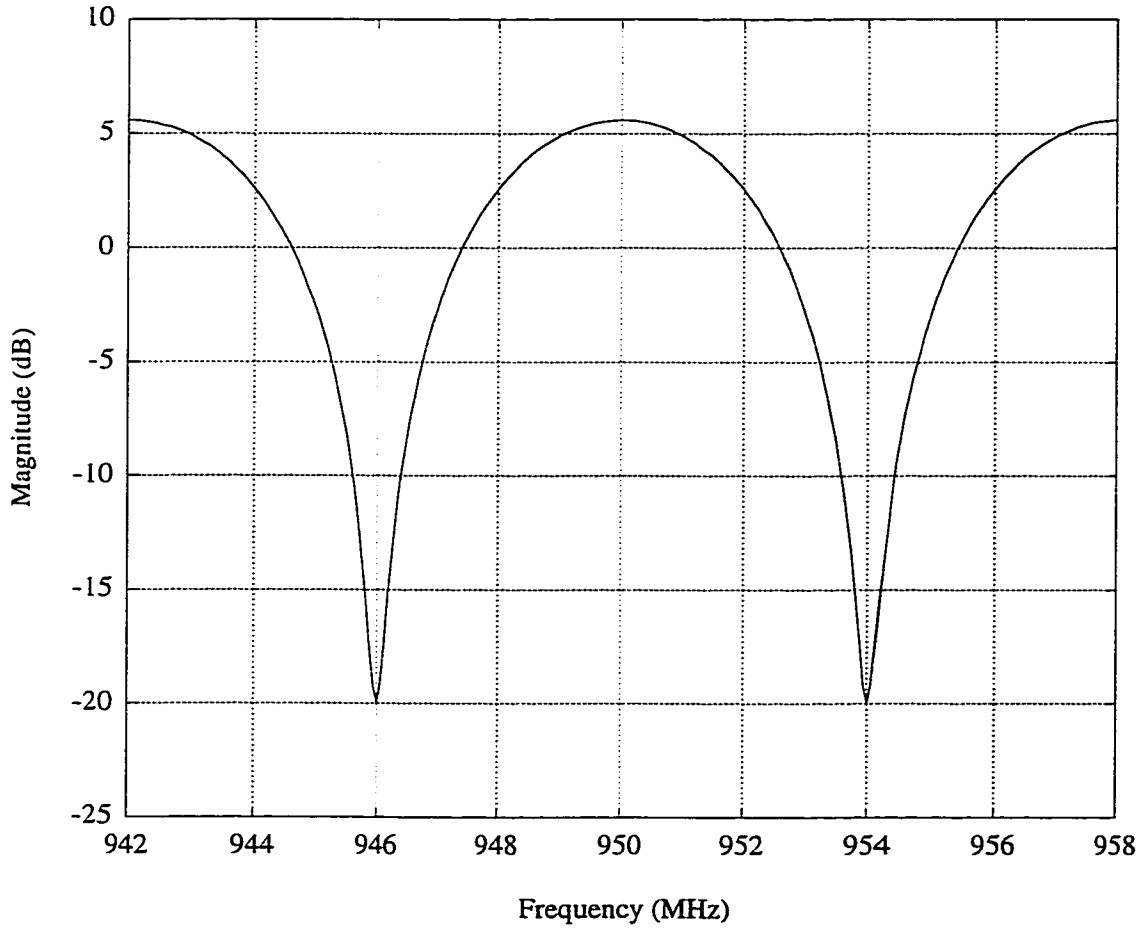


**Figure 5.4**: Passband frequency response of the channel ($\tau_{rms} \approx 62$nSec).

multipath channel. The receiver observes the sum of the distorted signals which is embedded in noise. Throughout the simulation we have used the following definition for Signal-to-Noise-Ratio, SNR (See Figure 5.5).

$$\mathrm{SNR}_{\mathrm{dB}} = 10 \log_{10} \frac{P_{av}(S(\boldsymbol{b}, t))}{P_{av}(z(t))} \tag{5.20}$$

where $P_{av}(\cdot)$ denotes the average power of a signal, and $z(t)$ is the additive white Gaussian noise, which is filtered by an ideal low-pass filter with (two-sided low-pass) bandwidth equal to 8 MHz.

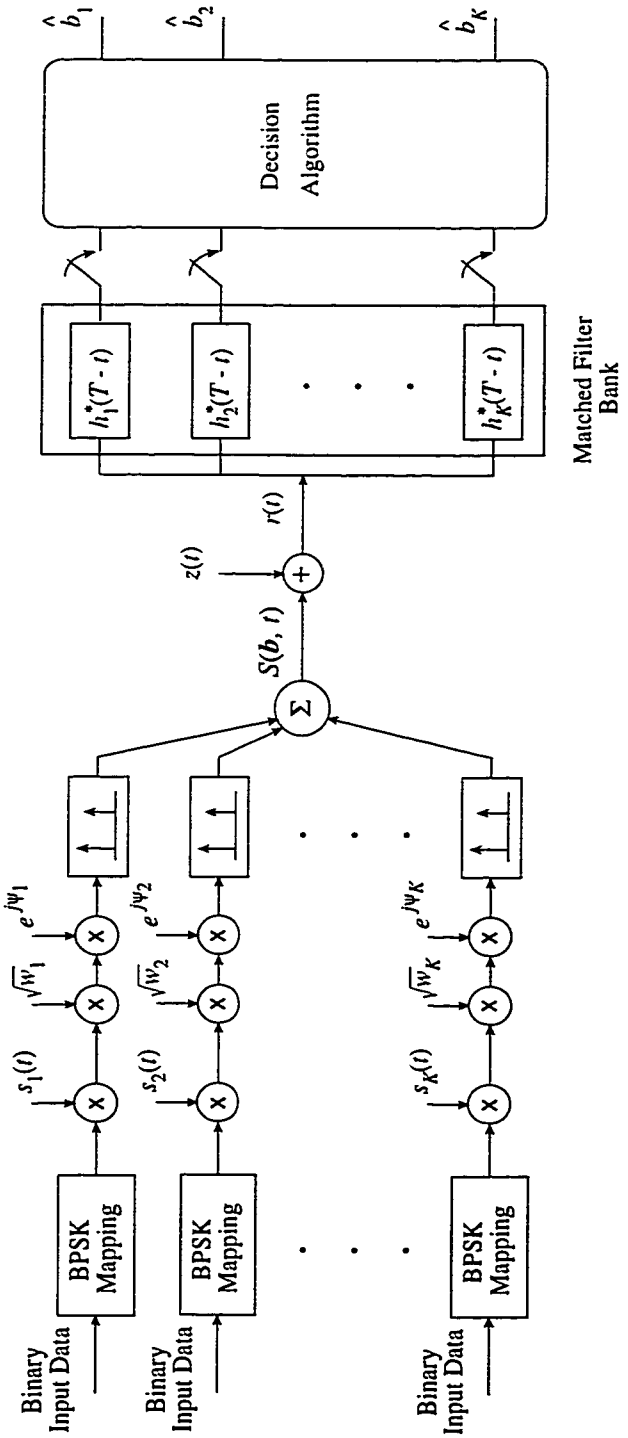


**Figure 5.5**: Diagram of simulation platform.

## 5.8 Multiuser Detectors in Multipath Channels

We generalize the idea behind the channel-matched decorrelating detector and apply it to all multiuser detectors that we have studied so far. Let us compare the discrete system model for AWGN channels in (2.28) with its counterpart system model for multipath channels in (5.17). We observe that (5.17) can be obtained by replacing $R$ in (2.28) by $R_c\Psi_c$. Consequently, if we substitute $R$ by $R_c\Psi_c$ in the multiuser detector algorithms, we will be able to devise new detectors that take into account multipath channels. In this section we consider the performance of the multiuser detectors which we studied in Chapter 2 and Chapter 3. For each case first we briefly study the detector in the multipath channel and then we present the simulation results.

### 5.8.1 Channel-Matched Conventional Detector

The conventional detector makes its decision at the output of the matched filter bank,

$$\hat{b}_k^c = \text{sgn}(y_{c,k}), \tag{5.21}$$

This method simply ignores MAI and treats it as noise. As we may notice the only difference between this detector and its counterpart in Chapter 2, i.e. the conventional detector designed for the AWGN channel, is that the matched filters are matched to the distorted spreading codes ($h_k(t)$) that have gone through the multipath channel, rather than the original spreading codes of the system. Figure 5.6 shows the simulations results, wherein the BER has been plotted versus SNR. The average BER is observed to be well above $10^{-2}$. There are two important issues

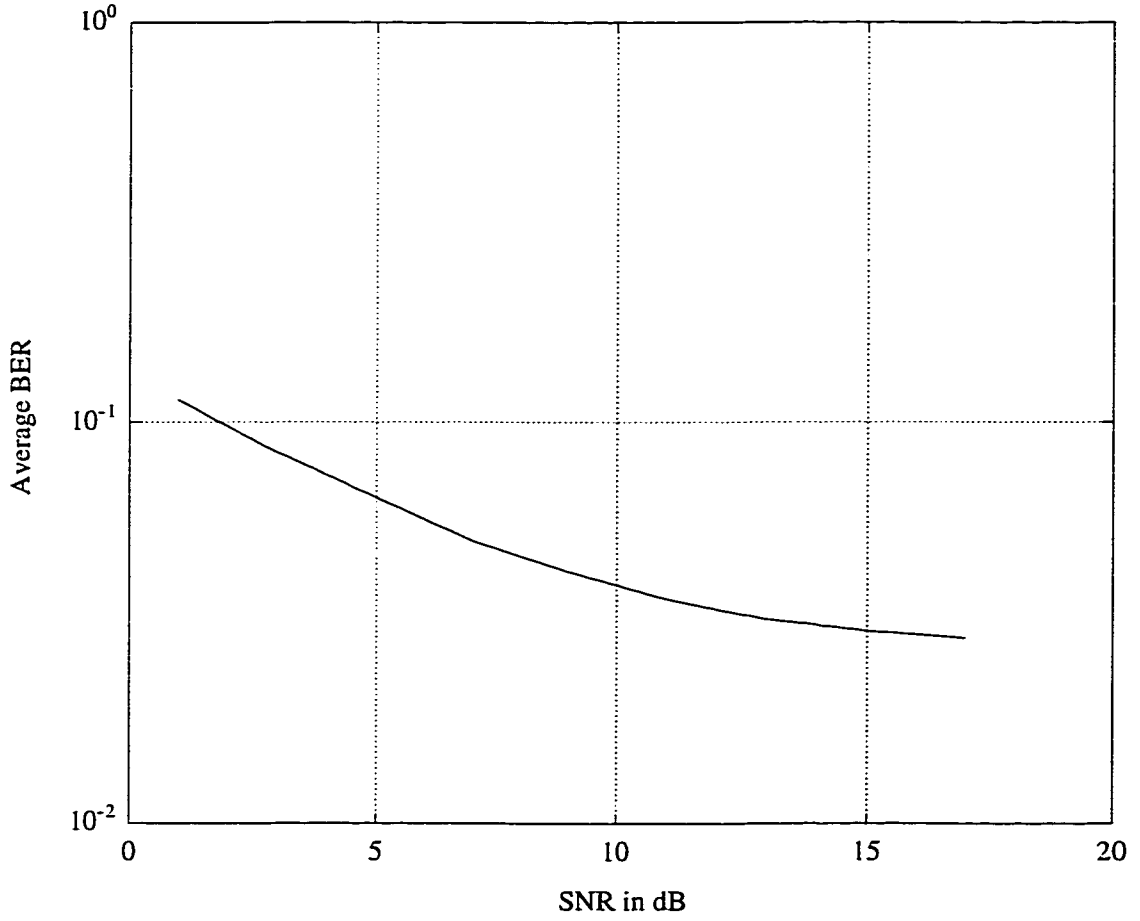


**Figure 5.6**: Average BER of the channel-matched conventional detector vs. SNR in a multipath channel.

that should be pointed out here.

First, although the matched filters are matched to the distorted spreading sequences, the BER is much worse than the BER of the conventional detector for the AWGN channel (compare Figure 5.6 and Figure 2.5). This can be explained as follows. Due to the random behavior of the multipath channel, the distorted spreading sequences are not guaranteed to have good cross-correlation properties anymore. Hence MAI components at the outputs of the matched filters are considerable compared to the desired signals. This leads to a very high BER for the

channel-matched conventional detector.

The second issue is the shape of the BER curve in Figure 5.6. In Chapter 2, we mentioned that the most important drawback of the conventional detector is the near-far problem. One consequence of the near-far problem is that the BER of the detector does not go to zero, even if the background noise goes to zero[1]. However, we were not able to show this fact explicitly in Figure 2.5, since the spreading sequences had very low cross-correlations and also we did not examine high SNR regions. Here, however, in Figure 5.6 we can see that the BER curve is asymptotically nonzero. In other words, the BER curve exhibits an *error floor*, i.e. the error performance is poor at medium to high SNR. Another factor that causes the error floor in BER is the presence of ISI. As it was mentioned in Section 5.3, we ignore the ISI and based on this assumption we design the multiuser detectors in this chapter. However, the multipath channel model in the computer simulation introduces some ISI. The impact of the presence of ISI, even though small, while the detector ignores it, appears as an error floor in the BER curve.

### 5.8.2 Channel-Matched Optimum Detector

The optimum multiuser detector in a multipath environment using channel-matched approach performs a maximum-likelihood sequence estimation (MLSE) scheme. The objective of MLSE is to find the input sequence which maximizes the conditional probability or likelihood of the given output sequence. The maximum-likelihood decision for vector $\boldsymbol{b}$ in Equation (5.17) is given by

$$\hat{\boldsymbol{b}}^{o} = \arg \max_{\boldsymbol{b}\in\{-1,1\}^{K}} (2\boldsymbol{b}^{T}\Psi_{c}W_{c}\boldsymbol{y}_{c} - \boldsymbol{b}^{T}\Psi_{c}W_{c}R_{c}W_{c}\Psi_{c}\boldsymbol{b}) \tag{5.22}$$

[1]For the detailed analytic study on this fact one could refer to [39]

The proof is similar to the discussion in Chapter 2.

Equation (5.22) requires an exhaustive search over all $2^K$ possible combinations of the entries of the bit vector $\boldsymbol{b}$. Figure 5.7 shows the average BER of the channel-matched optimum detector versus SNR. The huge difference between

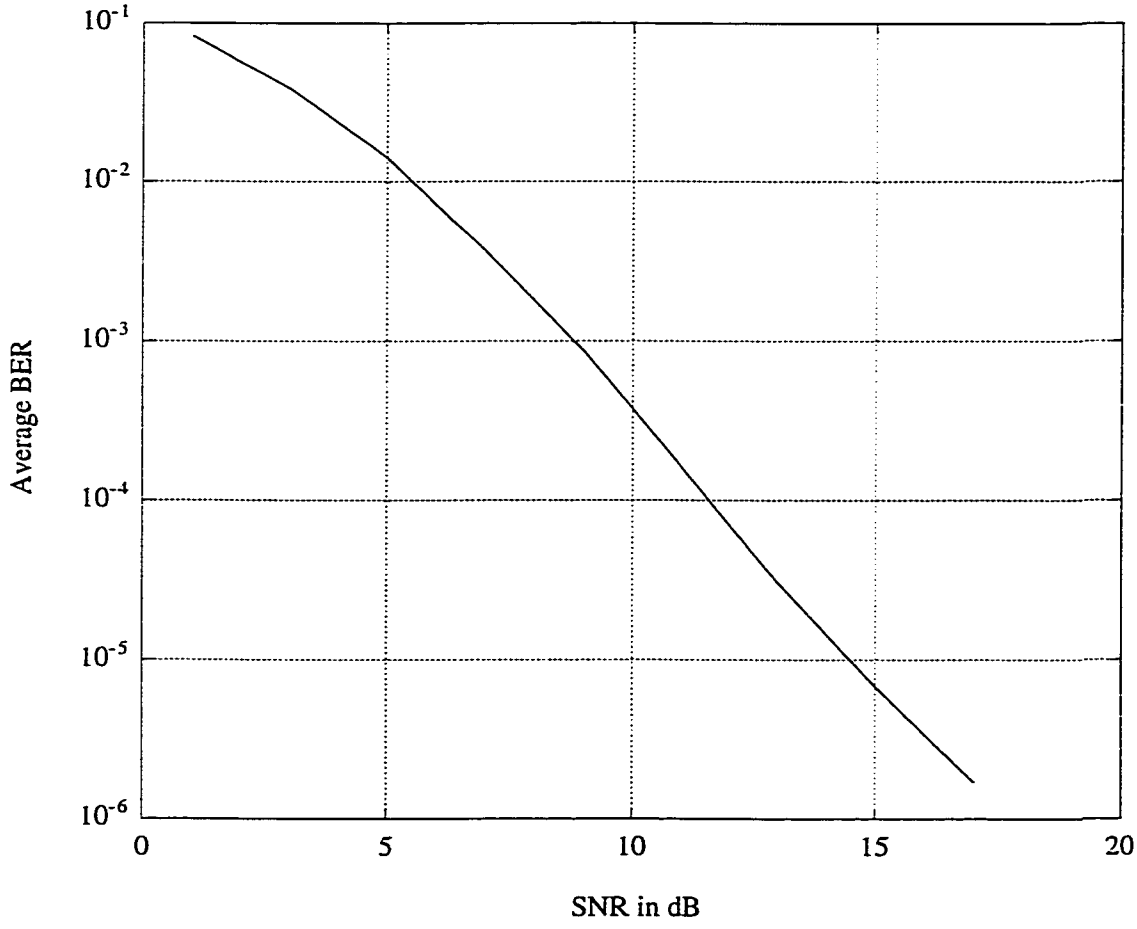


**Figure 5.7**: Average BER of the channel-matched optimum detector vs. SNR in a multipath channel.

the average BER for the channel-matched optimum detector and the average BER for the channel-matched conventional detector is evident (compare Figure 5.7 with Figure 5.6). The BER curve in Figure 5.7 starts exhibiting an error floor at SNR higher than 15 dB. As we explained in Chapter 2, the optimum multiuser detector

is near-far resistant and in the low background noise zone, it yields an error-free performance. The error floor in Figure 5.7 is due to ISI only. This fact is true for all other suboptimum detectors.

### 5.8.3 Channel-Matched Decorrelating Detector

In Section 5.4.2, when explaining the channel-matched approach, as an example we considered the channel-matched decorrelating detector. Having the same idea as the decorrelating detector described in Section 2.10, if we ignore the noise in (5.17), the transmitted information vector $\boldsymbol{b}$ may be obtained as

$$\hat{\boldsymbol{b}}^d = \text{sgn}(\Psi_C^H R_c^{-1} \boldsymbol{y}_c) \tag{5.23}$$

Since $W_c$ is a diagonal matrix with positive real diagonal components, we have ignored it in (5.23).

The average BER curve of the channel-matched decorrelating detector versus SNR is depicted in Figure 5.8. In order to compare the performance of this detector with those of the conventional and optimum detectors, we have plotted the BERs of these two detectors too. We observe that the BER performance of the channel-matched decorrelating detector follows almost the same trend as that of the channel-matched optimum detector, with a couple of decibels degradation.

### 5.8.4 Channel-Matched Decision Feedback Detector

A white noise model can be obtained by factorizing matrix $R_c$ using Cholesky decomposition as $R_c = F_c^T F_c$. Where $F_c$ is a lower triangular matrix. If the filter with response $\Psi_c^H (F_c^T)^{-1}$ is applied to the sampled outputs of the matched filters,

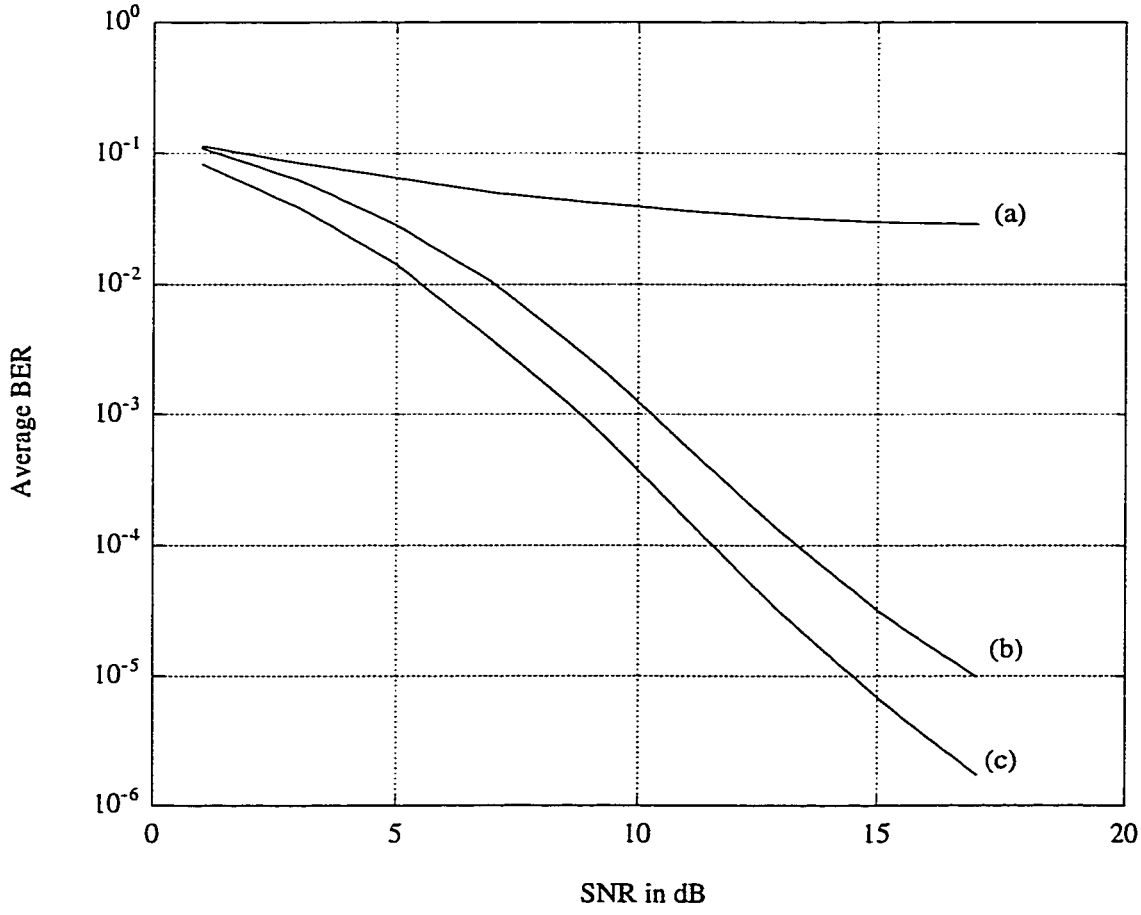


**Figure 5.8**: Comparison of bit-error-rates in a multipath channel; (a) conventional, (b) decorrelating, (c) optimum.

the resulting output vector is

$$\check{\boldsymbol{y}}_c = F_c W_c \boldsymbol{b} + \boldsymbol{\eta}_c \tag{5.24}$$

where $\boldsymbol{\eta}_c$ is a white Gaussian noise vector with the autocorrelation matrix $N_0\mathcal{I}$. In the same fashion as in decision-feedback detector for AWGN channels in Section 2.11

decisions can be made using

$$\hat{b}_k^{df} = \operatorname{sgn}\big(\check{y}_{c,k} - \sum_{i=1}^{k-1} [F_c]_{k,i}\, \sqrt{w_i}\, \hat{b}_i\big) \tag{5.25}$$

We notice that for the $k$th user, multiuser interference (the second term inside the signum function) depends on users 1 to $k-1$. Decisions for these users have already been made and they are used to form a feedback term, hence the name decision-feedback. In Figure 5.9 we have plotted the average BER of the channel-matched decision-feedback detector as well as the conventional and the optimum detectors.

### 5.8.5 Channel-Matched Improved Decision-Feedback Detector

As in Section 2.12 where we reviewed the improved decision-feedback detector, the idea behind this detector is to find the symbol vector $[\hat{b}_1, \hat{b}_2, \cdots, \hat{b}_K]$ with the smallest total metric sum, $m'_1 + m'_2 + \cdots + m'_K$, where $m'_k$ is the metric of user $k$ defined as

$$m'_k = \big(\check{y}_{c,k} - [F_c]_{k,k}\, \sqrt{w_k}\, b_k - \sum_{j=1}^{k-1} [F_c]_{k,j}\, \sqrt{w_j}\, \hat{b}_j\big)^2 \tag{5.26}$$

with $\check{y}_c$ and $F_c$ as defined in the previous section. The channel-matched improved decision-feedback detector follows the same algorithm explained on page 45. The only difference is that it uses metrics as defined in (5.26).

Figure 5.10 depicts the average BER of the channel-matched improved decision-feedback detector versus SNR. The performance of the conventional and optimum detectors are plotted as well. Here we have chosen $N_f = 3$ as we did in the AWGN channel. We notice from Figure 5.10 that the channel-matched improved decision-

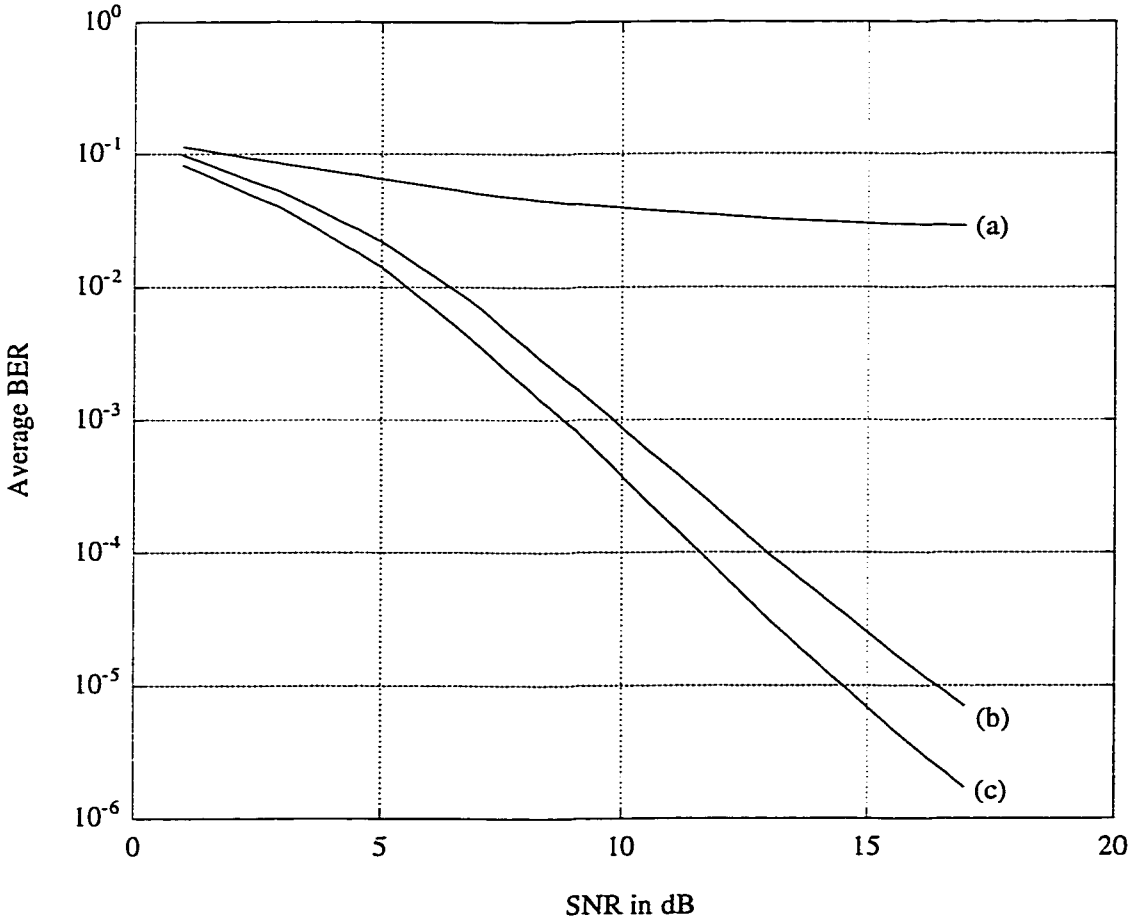


**Figure 5.9**: Comparison of bit-error-rates in a multipath channel; (a) conventional, (b) decision-feedback, (c) optimum.

feedback detector yields a BER performance very close to the optimum.

### 5.8.6 Channel-Matched Multistage Detector

A channel-matched multistage detector tries to remove the MAI in a stage-by-stage fashion using an iterative method. The estimated bit of the $k$-th user at stage $m+1$

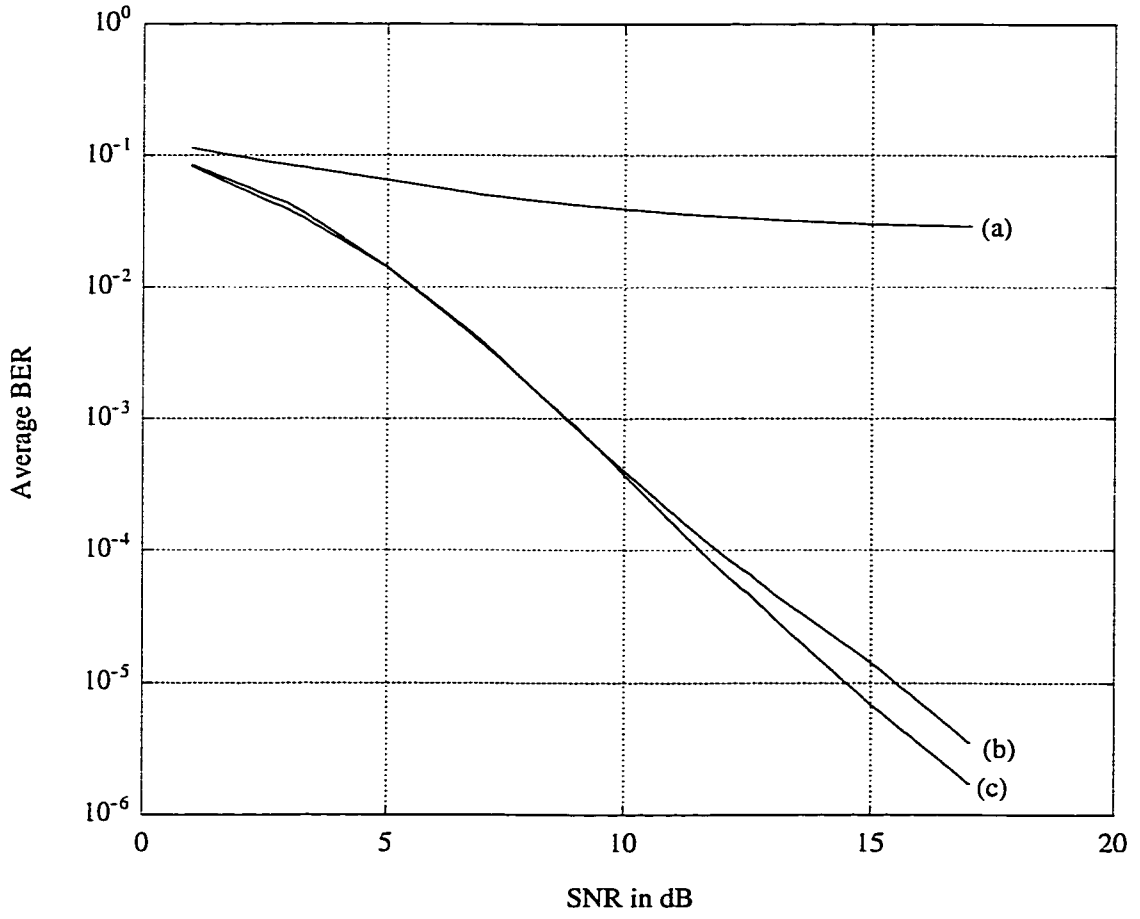


**Figure 5.10**: Comparison of bit-error-rates in a multipath channel; (a) conventional, (b) improved decision-feedback, (c) optimum.

is obtained by

$$\hat{b}_k^{mu}(m+1) = \mathrm{sgn}\big(y_{c,k} - \sum_{\substack{j=1 \\ j\neq k}}^{K} \hat{b}_j^{mu}(m) e^{-j\psi_j} \sqrt{w_j}\,[R_c]_{j,k}\big), \qquad m \geq 1 \qquad (5.27)$$

Due to delay limitations it is desirable to restrict the number of stages to two. We can use arbitrary values as the initial values for $\hat{b}_j(1)$'s. However, it is recommended to use a simple and effective multiuser detection scheme such that the iterative method in (5.27) yields an accurate detected information vector $\boldsymbol{b}$ as fast as possible.

As in Section 2.13, we use either a channel-matched conventional detector or a channel-matched decorrelating detector as the first stage.

Figure 5.11 shows the simulation results, wherein the average BER of two channel-matched 2-stage detectors are plotted. One detector uses a conventional detector as the first stage, while the other uses a decorrelating detector as the first stage. The BERs for the optimum and the conventional detectors are also plotted for the sake of comparison. We observe that when a decorrelating detector is used

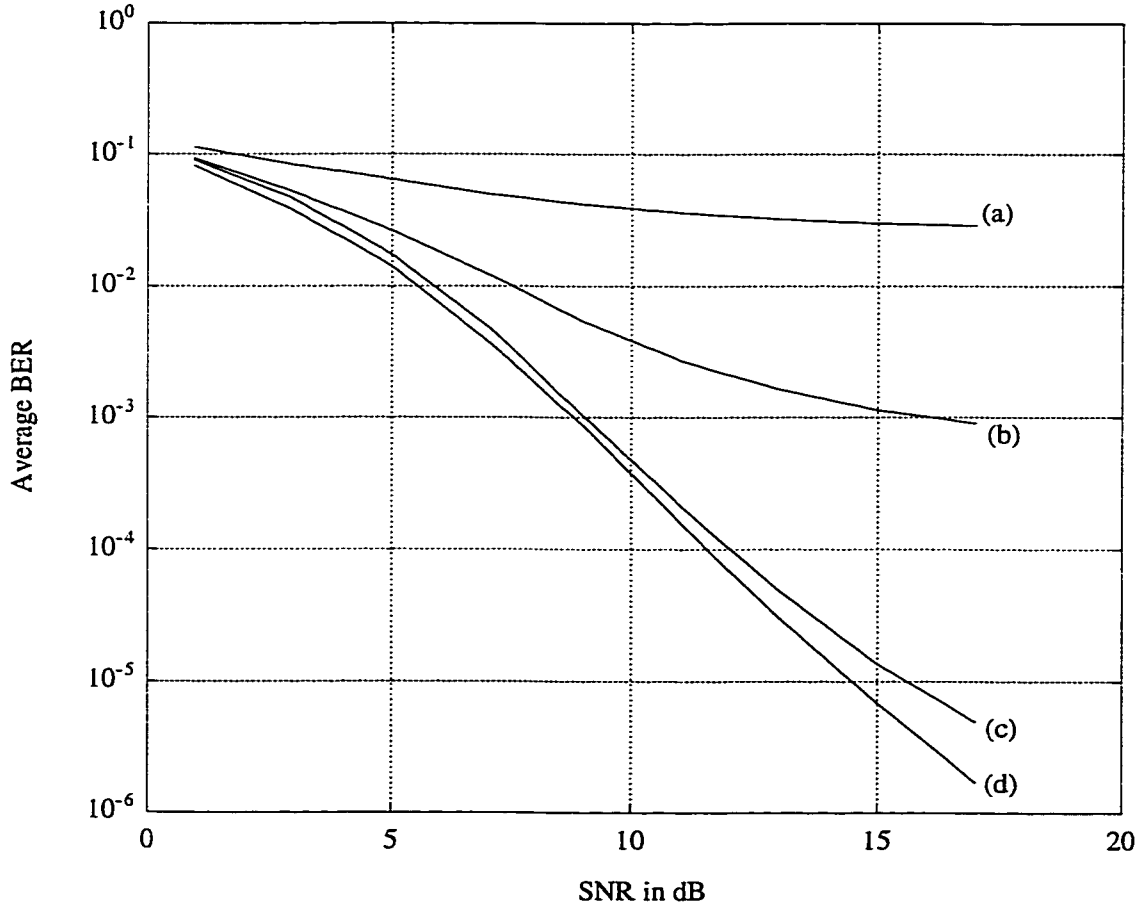


**Figure 5.11**: Comparison of bit-error-rates in a multipath channel; (a) conventional, (b) two-stage (conventional 1st stage), (c) two-stage (decorrelating 1st stage), (d) optimum.

as the first stage, the BER performance is much better than when a conventional

detector is used as the first stage. Indeed these results imply that the choice of the conventional detector as the first stage for a channel-matched two-stage detector is questionable.

### 5.8.7 Channel-Matched Two-Level Threshold Detector

The idea behind the channel-matched approach can be applied to the two-level threshold detector as well. This results in a multiuser detector for the multipath environment which works like its counterpart in Chapter 3. The outputs of the matched filters, i.e. the components of $\boldsymbol{y}_c$, are filtered using the channel-matched decorrelating filter. An algorithm that cycles through the outputs of the decorrelating filter determines whether it is necessary to remove the MAI from its outputs or not. The algorithm consists of two steps as follows:

**Step 1** Use the channel-matched decorrelating filter to find $\hat{\boldsymbol{y}}_c$,

$$\hat{\boldsymbol{y}}_c = \Psi_c^H R_c^{-1} \boldsymbol{y}_c \tag{5.28}$$

**Step 2** Cycle through the outputs of the channel-matched decorrelating filter, $\hat{\boldsymbol{y}}_c$ components, if $-\epsilon < \hat{y}_{c,k} < \epsilon$, where $\epsilon$ is a small number, then

$$\hat{b}_k = \text{sgn}(y_{c,k} - \sum_{i=1}^{k-1} \hat{b}_i e^{-j\psi_i} \sqrt{w_i}\,[R_c]_{i,k}) - \sum_{i=K+1}^{K} \hat{b}_i^d e^{-j\psi_i} \sqrt{w_i}\,[R_c]_{i,k}) \tag{5.29}$$

otherwise

$$\hat{b}_k = \text{sgn}(\hat{y}_{c,k}). \tag{5.30}$$

For a detailed description one is referred to Chapter 3, where a counterpart of this detector is studied for an AWGN channel.

Figure 5.12 shows the simulation results. The average BER of the channel-matched new detector is plotted versus SNR. We have chosen two values for $\epsilon$,

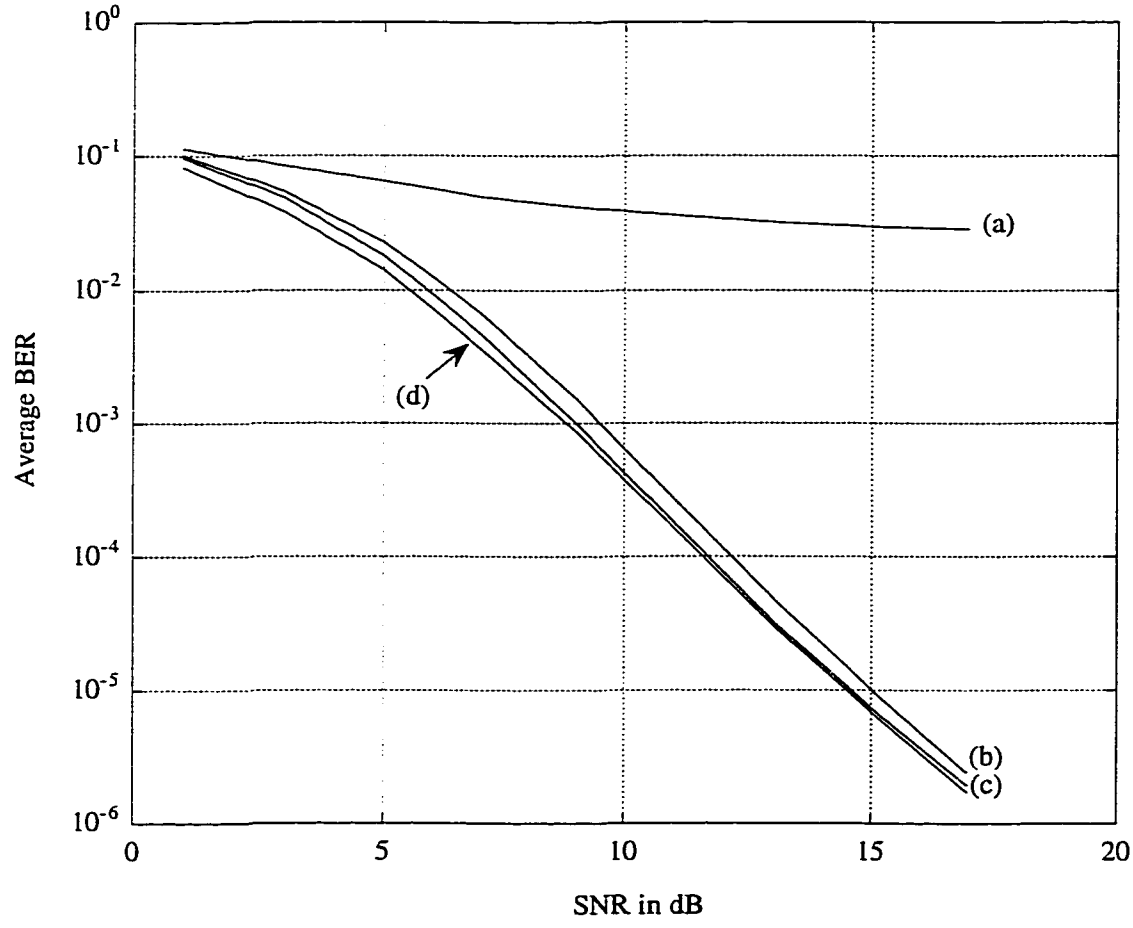


**Figure 5.12**: Comparison of bit-error-rates in a multipath channel; (a) conventional, (b) two-level threshold ($\epsilon = 0.1\sqrt{w_k}$), (c) two-level threshold ($\epsilon = 0.25\sqrt{w_k}$), (d) optimum.

$0.1\sqrt{w_k}$ and $0.25\sqrt{w_k}$. We may notice that by choosing an appropriate value for $\epsilon$, the channel-matched two-level threshold detector may achieve the near-optimum BER performance. For instance in Figure 5.12, choosing $\epsilon = 0.25\sqrt{w_k}$ instead of

$\epsilon = 0.1\sqrt{w_k}$ improves the BER. In the next section we study the effects of $\epsilon$ on the BER performance.

## 5.9 Optimum $\epsilon$

In Section 3.7 and in the previous section as well we pointed out that the radius of the neighborhood of zero, $\epsilon$, has a direct effect on the BER performance of the two-level threshold detector. By increasing $\epsilon$, the outputs of the decorrelating filter are more likely to be checked in Step 2 of the two-level threshold algorithm. Hence, the larger $\epsilon$ is, the more Equation (5.29) is used to detect the transmitted bit. If $\epsilon \longrightarrow \infty$, Equation (5.29) is always used and the two-level threshold detector becomes a counterpart of the two-stage detector. Therefore, intuitively it appears that the BER for the two-level threshold detector is a monotonic decreasing function with respect to $\epsilon$. In this section we try to find out how the BER performance depends on $\epsilon$.

In Chapter 3 we presented an approach to finding the probability of error for the two-level threshold detector. If we had a simple closed-form formula for the probability of error in terms of $\epsilon$, then finding a formula that determines the variations of the BER with $\epsilon$ would be convenient. However, as shown in Chapter 3 we failed to find such a closed-form formula. Consequently we choose to use simulation results once again. In Figures 5.13 and 5.14, we have plotted the average BER as a function of $\epsilon$, for two different values of SNR, 9 dB and 15 dB, which yield BER of about $10^{-3}$ and $10^{-5}$, respectively. The computational load of the two-level threshold detector is plotted as well. As we observe, unlike what we intuitively assumed, the BER is not a monotonic curve in terms of $\epsilon$. Increasing $\epsilon$ to some point improves the BER performance, but afterwards not only does it not help to enhance the BER

performance but also it can cause an adverse effect.

The existence of an optimum $\epsilon$ with respect to BER is evident in both curves. Fortunately, in both cases the optimum $\epsilon$ happens where the computational load is still small and slightly more than the computational load of the decorrelating detector [2].

The results plotted in Figures 5.13 and 5.14 show that the optimum value for $\epsilon$

[2]We recall that the computational load of the (channel-matched) two-level threshold detector varies between the computational load of the (channel-matched) decorrelating detector and that of the (channel-matched) two-stage detector with a decorrelating detector as the first stage.

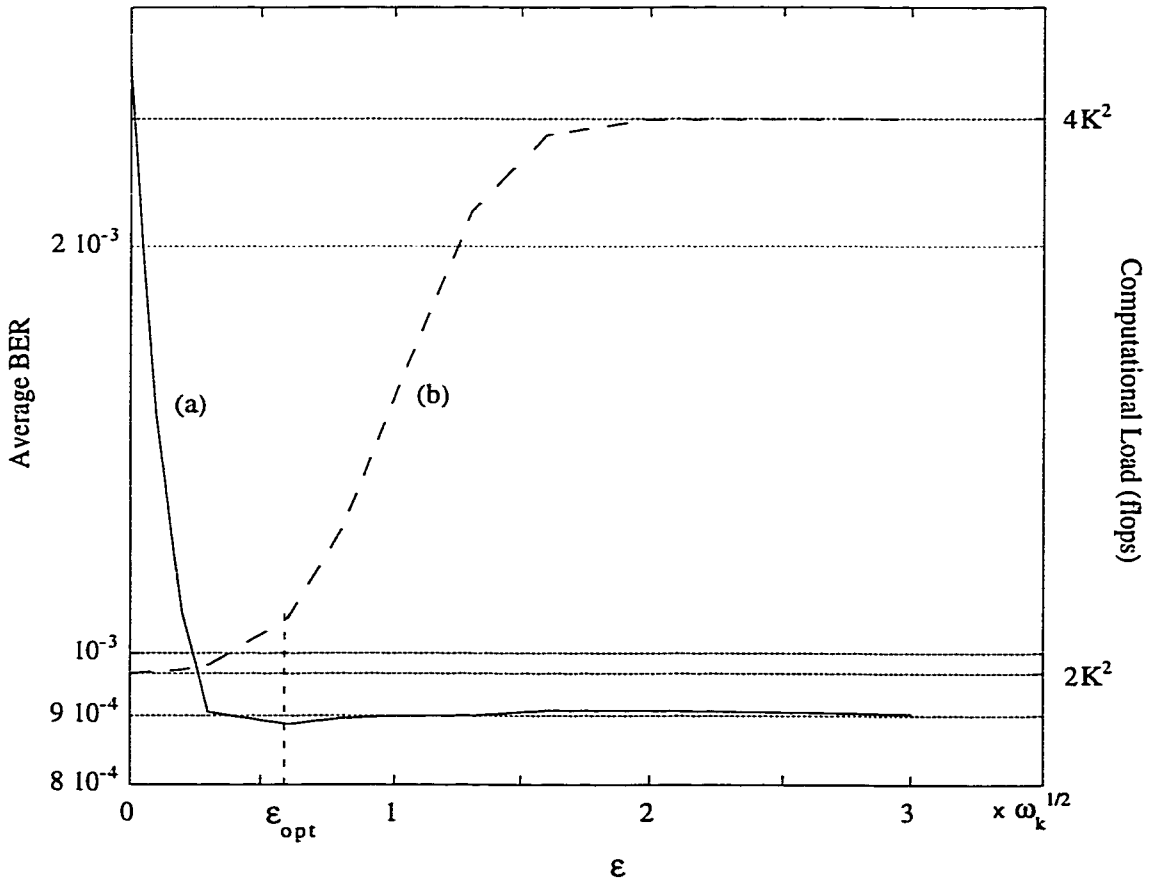


**Figure 5.13**: Finding the optimum $\epsilon$ for the two-level threshold detector at SNR = 9 dB; (a) average BER, (b) computational load.

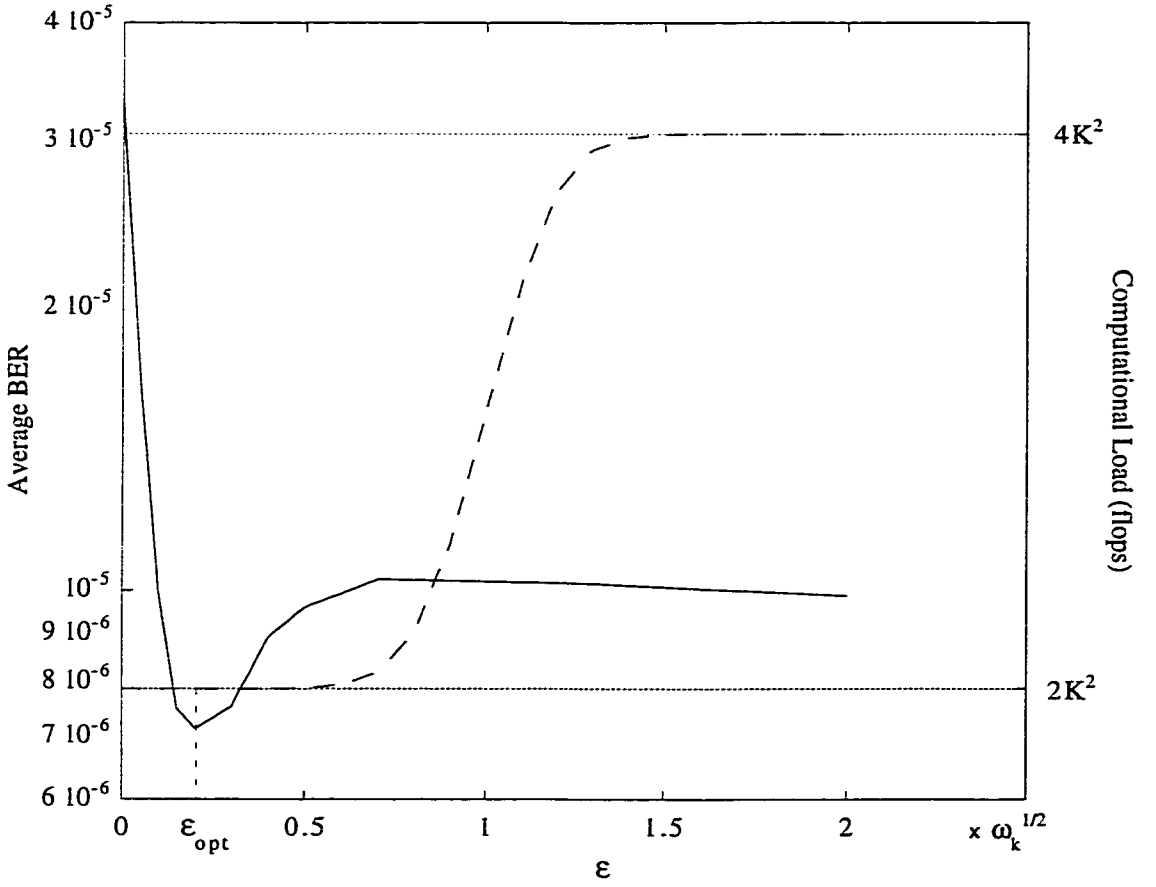


**Figure 5.14**: Finding the optimum $\epsilon$ for the two-level threshold detector at SNR = 15 dB; (a) average BER, (b) computational load.

depends on SNR, i.e. there is no unique optimum $\epsilon$. In practice, once the BER is determined, the operating SNR and hence the optimum value for $\epsilon$ is specified.

## 5.10 Discussion

In this chapter we studied the multiuser detectors presented in Chapters 2 and 3 in a multipath fading frequency-selective channel. There are two general strategy towards the design of multiuser detectors for multipath channels; i) path-by-path and ii) channel-matched. We chose the channel-matched approach, since it offers a

better BER performance. This fact is true at least for the decorrelating detector. That is, the channel-matched decorrelating detector outperforms the path-by-path decorrelating detector. A thorough study to find out whether this fact is true for other multiuser detectors or not, is out of the scope of this dissertation.

In order to compare the BER performance for all multiuser detectors presented in this chapter, we have plotted Figure 5.15. We observe that the conventional detector

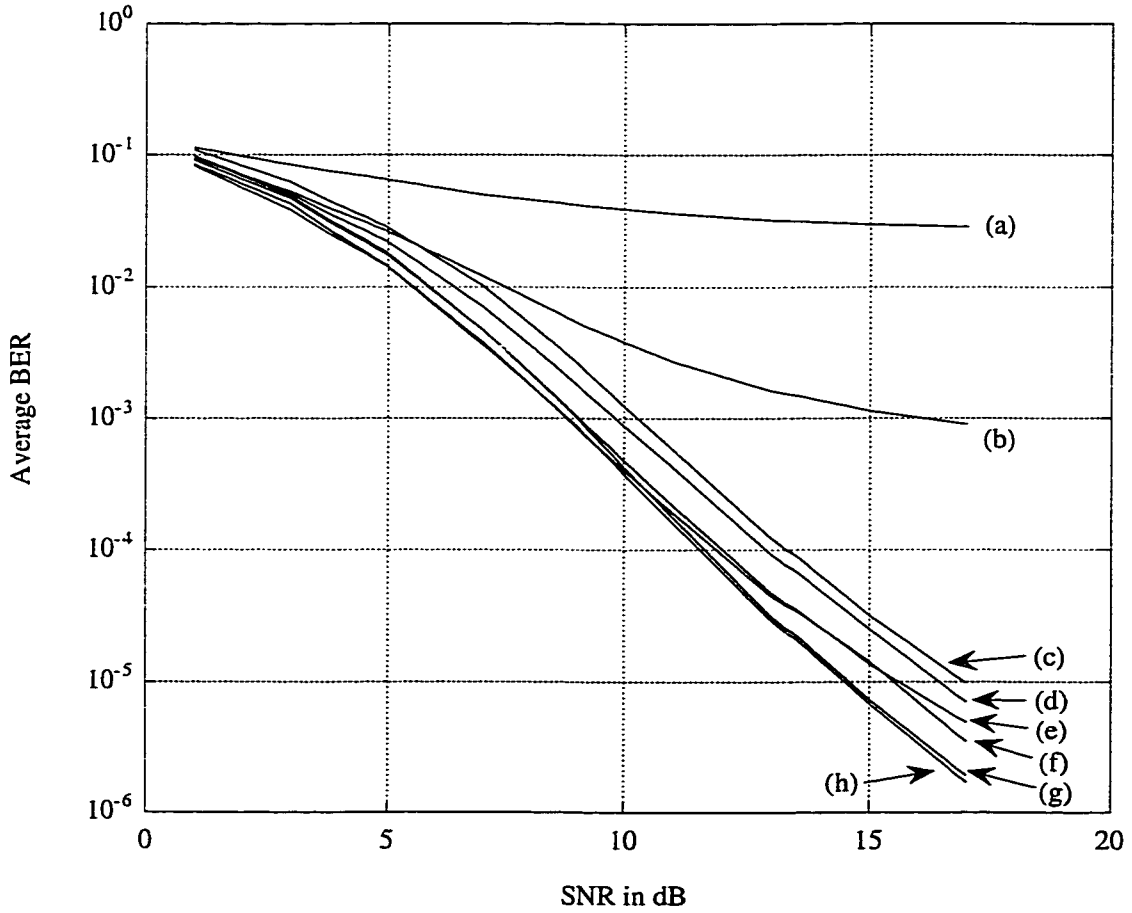


**Figure 5.15**: Comparison of bit-error-rates in a multipath channel; (a) conventional, (b) two-stage (conventional 1st stage), (c) decorrelating, (d) decision-feedback, (e) two-stage (decorrelating 1st stage), (f) improved decision-feedback, (g) two-level threshold ($\epsilon = 0.25\sqrt{w_k}$), (h) optimum.

performs poorly. Also the two-stage detector that uses the conventional detector

as the first stage does not offer a satisfactory BER performance either. This is obviously due to the poor performance of the conventional detector. The rest of the detectors perform well enough compared to the optimum detector. Amongst them, the two-level threshold detector, proposed in this dissertation, offers a near-optimum performance, provided that the value of $\epsilon$ is chosen optimally. Fortunately the optimum value of $\epsilon$ occurs where the computational load of the two-level threshold is slightly more than that of the decorrelating detector, i.e. $2K^2$.

# Chapter 6

# Symbol-Aided Channel Estimation and Multiuser Detection Using a Decorrelating-Type Detector

Thus far in this dissertation various multiuser detectors in both AWGN and fading channels have been considered. Almost all of the discussions and analyses have been based on a number of idealizations. For instance, the system and channel parameters are assumed to be perfectly known. While multiuser detection has been analyzed thoroughly for both AWGN and fading channels, less work has been reported on the underlying parameter estimation issues in CDMA channels [54]. This chapter is devoted to prospect this issue. A decorrelating-type filter detector is considered for multiuser detection. This filter operates on the outputs of filters that are matched to the original code waveforms of the system. A maximum likelihood estimation for the parameters of the decorrelating filter is proposed and derived.

This chapter has the following outline. Section 6.1 summarizes the previous work. In Section 6.2 the contributions of this chapter are reviewed. In Section 6.3

we briefly describe the system and the channel model that we consider in the present chapter. Section 6.4 describes the detection scheme which we choose for multiuser detection. In Section 6.5 we derive a maximum likelihood estimation technique for the channel and for the receiver decorrelating-type filter. Section 6.6 proposes a method to select training sequences to approach the minimum mean-square-error (MMSE) criterion. In Section 6.7 we introduce an iterative method for matrix inversion based on orthogonalization. In iteration $i$ this method requires only column $i$ of the matrix. Section 6.8 presents some simulation results for a 16-user CDMA system. Finally, Section 6.9 concludes this chapter with a summary of results.

## 6.1 Previous Work

Multiuser detection techniques have been considered to increase channel capacity and alleviate the near-far problem in CDMA systems. Various methods have been proposed for multiuser detection. Primary studies on multiuser detectors concerned proposing some systems and analyzing them for additive white Gaussian noise CDMA channels [17, 19–21]. However, most of the multiple-access channels of interest, e.g. the indoor wireless channel and the mobile radio channel, exhibit multipath fading. Consequently a large amount of research has been focusing on proposing multiuser detectors to take into account multipath fading [82–84, 87, 88, 92]. Recently, more realistic multiuser detectors have been studied in which both detection and channel estimation are considered. Some attention has been concentrated on adaptive multiuser detection (see e.g. [98–112]) and blind adaptive multiuser detection (see e.g. [113–119]). There are also some methods wherein the channel parameters are estimated explicitly [41, 120]. Work reported in [121] models the frequency nonselective fading channel as a second order Auto-Regressive (AR) pro-

cess, and uses a decision-directed Kalman filter to estimate the channel coefficients. In [97] a joint multiuser detection and channel estimation scheme is suggested which uses a path-by-path decorrelator to estimate the channel parameters and hence the coefficients of a channel-matched decorrelator. Decisions are then made on the output of the channel-matched decorrelator.

## 6.2 Contribution

Based on the work of this chapter Hosseinian et al. [27] have proposed a joint multiuser detection and channel estimation scheme. A decorrelating-type filter whose coefficients are estimated using the maximum likelihood estimation method is proposed and derived. The estimation method is based on inserting known training sequences into the information data by all users simultaneously. To achieve MMSE in estimation, a criterion for selecting the training sequences is suggested. Orthogonal training sequences will be good candidates to approach MMSE. The estimation method requires a matrix inversion at the end of each training period. An iterative matrix inversion method is introduced to distribute the computational load of the matrix inversion over the training period. The simulation results show some degradations in BER performance compared to the case where perfect knowledge of the channel is assumed. This degradation is clearly due to errors in the channel estimation.

## 6.3 System and Channel Model

We consider a synchronous CDMA system with $K$ users in the multipath channel. The $k$-th user is assigned a normalized spreading waveform $\{s_k(t), \quad t \in [0, T]\}$.

As it was shown in the previous section, at the receiver, the received signal can be written as:

$$r(t) = S(t, \boldsymbol{b}) + n(t) \tag{6.1}$$

where $n(t)$ is a zero-mean complex additive white Gaussian noise whose real and imaginary parts are independent and each have power spectral density $N_0/2$ and:

$$S(t, \boldsymbol{b}) = \sum_{i=-\infty}^{\infty} \sum_{k=1}^{K} \sum_{l=1}^{L} b_k(i) c_{k,l}(i) e^{j\psi_k} \sqrt{w_k(i)} s_k(t - iT - \tau_{k,l}) \tag{6.2}$$

where $i$ is the time interval index, $b_k(i)$ is the information transmitted by user $k$ in the epoch $i$, $L$ is the number of propagation paths, $c_{k,l}$ is the fading complex envelope of the $k$th user corresponding to path $l$, and $\tau_{k,l}$ is the delay of path $l$ of user $k$. We assume that the symbol duration $T$ is much longer than the multipath delay spread, i.e. $T \gg \tau_{rms}$. If this is the case, then the intersymbol interference (ISI) due to channel dispersion can be ignored [55]. We consider slowly fading channels, which implies that $c_{k,l}(i)$ does not change during a symbol transmission time. Under these assumptions, we consider the received signal $r(t)$ over only one symbol interval, e.g. $i = 0$. Consequently, we may consider a one shot system as follows

$$r(t) = \sum_{k=1}^{K} \sum_{l=1}^{L} b_k c_{k,l} e^{j\psi_k} \sqrt{w_k} s_k(t - \tau_{k,l}) + n(t) \tag{6.3}$$

## 6.4 Detection scheme

In Section 5.8.3 we considered the channel-matched decorrelating detection scheme. Therein we assumed the bank of matched filters at the front-end of the receiver are matched to the distorted spreading codes, i.e. $h_k(t)$ where

$$h_k(t) = \sum_{l=1}^{L} c_{k,l} e^{j\psi_k} \sqrt{w_k} s_k(t - \tau_{k,l}). \tag{6.4}$$

In this chapter we assume that the matched filters are matched to the spreading codes, i.e. to $s_k(t)$. We choose this approach for two reasons. Firstly, the structure of the matched filters are much simpler. The system that we are considering is a DS-CDMA with BPSK modulation scheme. The spreading codes for such a system are sequences of '-1' and '+1'. Implementation of digital matched filters (correlators) matched to the sequences of '-1' and '+1' is easy and does not require multipliers. While the implementation of digital matched filters which are matched to $h_k(t)$ requires complex multipliers. It is worthwhile to mention that, in fact, a filter with the response $h_k^*(T-t)$ functions like the RAKE receiver for the single user systems. Whereas a filter with response $s_k^*(T-t)$ performs as a single branch of the RAKE receiver with a unity tap weight. Therefore, using matched filters that are matched to $s_k(t)$ does not yield optimum BER performance.

Secondly, in this chapter we do not assume that the channel parameters are known. In the next section we investigate a scheme to estimate the channel parameters. By using matched filters that are matched to $s_k(t)$, as opposed to $h_k(t)$, we will not have to estimate the coefficients of the matched filters and the estimation algorithm will be much simpler.

Now, suppose $r(t)$ is applied to a bank of matched filters which are matched

to the set of spreading sequences and followed by samplers at time $T$, then the following discrete-time model results:

$$\boldsymbol{y} = \mathcal{X}\boldsymbol{b} + \boldsymbol{n} \tag{6.5}$$

where $\boldsymbol{b} = (b_1, b_2, \ldots, b_K)^T$, $\boldsymbol{y} = (y_1, y_2, \ldots, y_K)^T$ are the outputs of the samplers, $\mathcal{X}$ is a $K \times K$ matrix whose entries are defined as:

$$\mathcal{X}_{i,j} = \int_0^T s_i(t) \sum_{l=1}^{L} c_{j,l} e^{j\psi_j} \sqrt{w_j} s_j(t - \tau_{j,l}) dt, \tag{6.6}$$

and $\boldsymbol{n}$ is a zero-mean Gaussian $K$-vector with covariance matrix equal to $N_0 R$, where $R$ is a nonnegative definite matrix of cross-correlation between the spreading sequences as defined in Equation (2.5). We use a decorrelating-type scheme to detect the transmitted information $\boldsymbol{b}$. This method, while very simple, exhibits a reasonable performance in terms of both bit-error-rate and near-far resistance [17]. Our decorrelating-type detector makes decisions as:

$$\hat{\boldsymbol{b}} = \text{sgn}(\mathcal{X}^{-1}\boldsymbol{y}) \tag{6.7}$$

## 6.5 Estimation Scheme

A symbol-aided scheme is used to estimate $\mathcal{X}^{-1}$ in (6.7). A reference training sequence with length $K$ is inserted into the sequence of information data symbols by each transmitter. The training sequences are sent simultaneously by transmitters (Figure 6.1). We assume that the channel parameters do not change during the training period, i.e. $\mathcal{X}$ does not vary. Consequently after the training period,

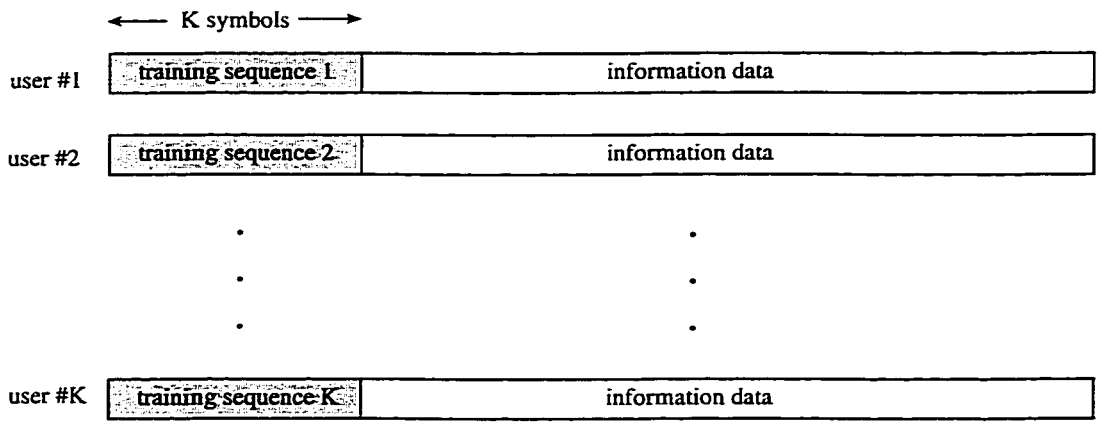


**Figure 6.1**: Synchronous transmission of training sequences.

considering Equation (6.7), we will have $K$ systems of equations:

$$\boldsymbol{y}(i) = \mathcal{X}\boldsymbol{b}(i) + \boldsymbol{n}(i), \qquad i = 1, \cdots, K \tag{6.8}$$

Merging these equations into one equation, we can form a matrix equation as:

$$\mathcal{Y} = \mathcal{X}\mathcal{B} + \mathcal{N} \tag{6.9}$$

where $\mathcal{B}$ is a $K \times K$ matrix whose column $i$ corresponds to the $i$th symbols of the training sequences, $\mathcal{Y}$ is a $K \times K$ matrix whose column $i$ is the outputs of the matched filters corresponding to the transmission of column $i$ of $\mathcal{B}$, and $\mathcal{N}$ is a $K \times K$ zero-mean Gaussian matrix. We notice that the correlation between each two entries of $\mathcal{N}$ can be expressed as follows:

$$\mathcal{E}\{\mathcal{N}^*_{i,j}\mathcal{N}_{k,l}\} = \begin{cases} 0, & j \neq l \\ R_{i,k}N_0/2, & j = l \end{cases} \tag{6.10}$$

where $\mathcal{E}\{\cdot\}$ denotes expectation. The maximum likelihood estimation (MLE) of matrix $\mathcal{X}^{-1}$ can easily be determined as follows. Rewriting (6.9) in terms of matrix transpositions, we have:

$$\mathcal{Y}^T = \mathcal{B}^T \mathcal{X}^T + \mathcal{N}^T \tag{6.11}$$

Or equivalently:

$$[\mathcal{Y}^T]_i = \mathcal{B}^T [\mathcal{X}^T]_i + [\mathcal{N}^T]_i, \qquad i = 1, \cdots, K \tag{6.12}$$

where $[\cdot]_i$ denotes column $i$ of a matrix. Now, we may notice that the elements of $[\mathcal{N}^T]_i$ are the output noise components of matched filter $i$ and therefore they are independent Gaussian random variables. Hence the maximum likelihood estimator of $[\mathcal{X}^T]_i$ is [122]:

$$[\hat{\mathcal{X}}^T]_i = (\mathcal{B}^T)^{-1} [\mathcal{Y}^T]_i, \qquad i = 1, \cdots, K \tag{6.13}$$

with the covariance matrix of $[\hat{\mathcal{X}}^T]_i$:

$$C_{[\hat{\mathcal{X}}^T]_i} = \frac{N_0}{2} (\mathcal{B}\mathcal{B}^T)^{-1} \tag{6.14}$$

Here we have assumed that the training sequences are chosen so that $\mathcal{B}$ is invertible. Merging the above $K$ equations into one matrix equation, we have:

$$\hat{\mathcal{X}}^T = (\mathcal{B}^T)^{-1} \mathcal{Y}^T \tag{6.15}$$

and hence the maximum likelihood estimator of $\mathcal{X}^{-1}$ is:

$$\hat{\mathcal{X}}^{-1} = \mathcal{B}\mathcal{Y}^{-1} \tag{6.16}$$

Equation (6.16) suggests that the outputs of the matched filters be stored column-wise as a matrix, $\mathcal{Y}$, during the training period. At the end of the training period, the inverse matrix $\mathcal{Y}^{-1}$ is calculated and premultiplied by matrix $\mathcal{B}$ to yield an MLE of the inverse filter $\mathcal{X}^{-1}$ (Figure 6.2). From the computational complexity point of

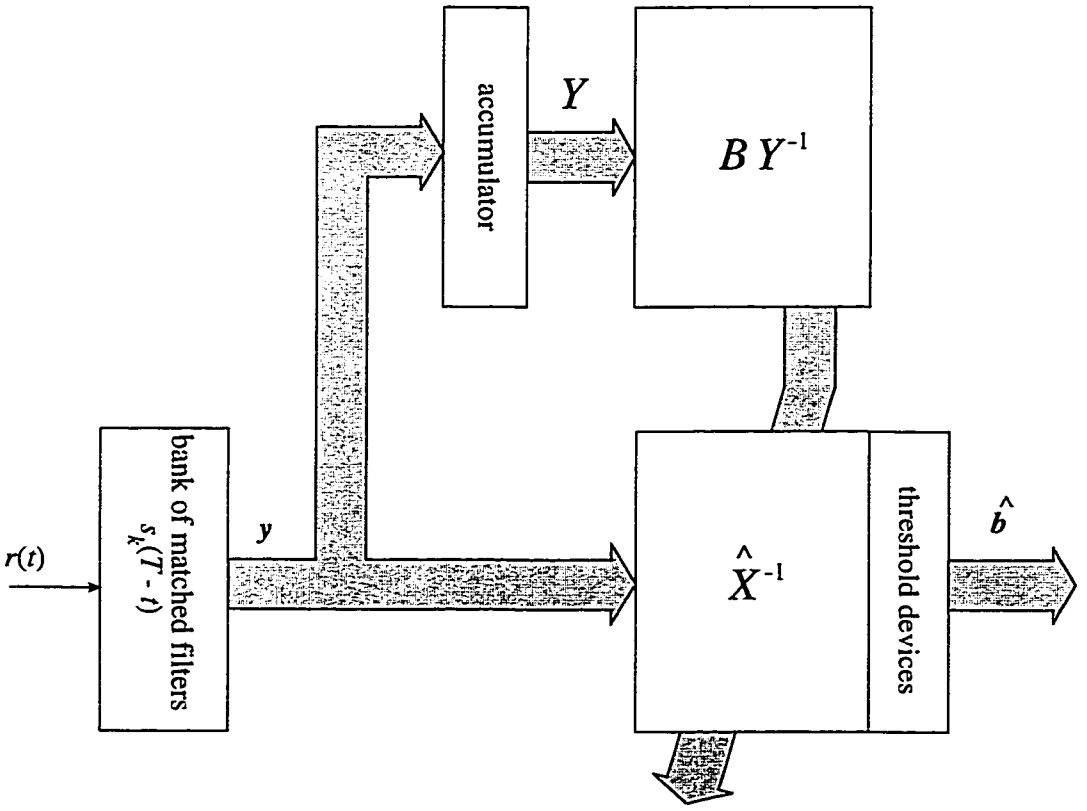


**Figure 6.2**: Decorrelating-type detector and estimator

view, this method is not homogeneous. In other words the computational load is not distributed fairly. In Section 6.7 we suggest a method for matrix inversion, which is based on matrix orthogonalization, to distribute the computational load of

matrix inversion over the training period.

## 6.6 Training Sequences

The MLE analysis in the previous section assumes that matrix $\mathcal{B}$ is known. This matrix, however, is one of the system attributes and must be chosen appropriately at the system design step. In this section we consider an approach to identify the training sequences which yield the MMSE criterion for the estimation of $\mathcal{X}^{-1}$ or $\mathcal{X}$. Define error matrix $E$ as the difference between the estimated matrix $\hat{\mathcal{X}}$ and $\mathcal{X}$. From (6.9) and (6.16) we may write:

$$E = \hat{\mathcal{X}} - \mathcal{X} = \mathcal{N}\mathcal{B}^{-1} \tag{6.17}$$

The mean-square-error, MSE, is defined as

$$\text{MSE} = \mathcal{E}\{\sum_{i=1}^{K}\sum_{j=1}^{K}|E_{i,j}|^2\}, \tag{6.18}$$

The MMSE sequences are obtained by minimizing (6.18) with respect to $\mathcal{B}$. Using equations (6.17) and (6.10) we can expand equation (6.18) and then simplify it as follows:

$$\begin{aligned}\text{MSE} &= \sum_{i=1}^{K}\sum_{j=1}^{K}\sum_{p=1}^{K}\sum_{q=1}^{K}\mathcal{E}\{\mathcal{N}_{i,p}^{*}\mathcal{N}_{i,q}\}b_{p,j}^{-1*}b_{q,j}^{-1} \\ &= \frac{N_0}{2}\sum_{i=1}^{K}R_{i,i}\sum_{j=1}^{K}\sum_{p=1}^{K}|b_{p,j}^{-1}|^2,\end{aligned} \tag{6.19}$$

where $b_{ij}^{-1} = \{\mathcal{B}^{-1}\}_{i,j}$. Finally, one can note that the first summation in (6.19) is the trace of $R$ and the second double summation is in fact the square of the Frobenius

norm of $\mathcal{B}^{-1}$. Therefore MSE may be expressed as:

$$\text{MSE} = \frac{N_0}{2} \text{tr}(R) \, \|\mathcal{B}^{-1}\|_F^2 \tag{6.20}$$

where $\text{tr}(R)$ denotes the trace of $R$.

Note that since $\text{tr}(R)$ depends on the assigned waveforms characteristics, it can be regarded as a constant. Hence the problem of minimization of the MSE reduces to the choice of $\mathcal{B}$ so that the Frobenius norm of its inverse is minimum. Considering the fact that the elements of the training sequences matrix $\mathcal{B}$ are selected from a finite set (signal constellation points), we are presented with a discrete optimization problem. A general approach for solving this problem is the exhaustive search method. When $K$, i.e. the dimension of $\mathcal{B}$, is small (e.g. less than 8) the search method is practical, but for larger values of $K$ the tedious job of numerous large matrix inversions appears to be impractical. In the rest of this section we attempt to find a lower bound for the Frobenius norm of $\mathcal{B}^{-1}$ and then we try to find those matrices for which the Frobenius norms of their inverses are close to this lower bound.

From linear algebra we know that [1]

$$\|\mathcal{A} \cdot \mathcal{B}\|_F \leq \|\mathcal{A}\|_F \cdot \|\mathcal{B}\|_F \tag{6.21}$$

where $\mathcal{A}$ is an arbitrary matrix. If we choose $\mathcal{A} = \mathcal{B}^{-1}$ we will have,

$$\|\mathcal{I}\|_F \leq \|\mathcal{B}^{-1}\|_F \|\mathcal{B}\|_F \tag{6.22}$$

[1] For some interesting matrix norm properties the reader is referred to [57].

where $\mathcal{I}$ denotes the identity matrix. As it was mentioned earlier, we consider a BPSK constellation in which the elements of $\mathcal{B}$ are chosen from the finite set $\{1, -1\}$. Hence $\|\mathcal{B}\|_F = \|\mathcal{B}^T\|_F = K$. (For other constellations, $\|\mathcal{B}\|_F$ can easily be determined as well.) On the other hand $\|\mathcal{I}\|_F = \sqrt{K}$. Substituting these two in (6.22), we obtain a lower bound for $\|\mathcal{B}^{-1}\|_F$ as

$$\|\mathcal{B}^{-1}\|_F \geq \frac{1}{\sqrt{K}} \tag{6.23}$$

Now we notice that if $\mathcal{B}$ is chosen so that

$$\mathcal{B}^{-1} = \frac{1}{K}\mathcal{B}^T \tag{6.24}$$

then

$$\begin{aligned} \|\mathcal{B}^{-1}\|_F^2 &= \frac{1}{K^2}\|\mathcal{B}^T\|_F^2 \\ &= 1 \end{aligned} \tag{6.25}$$

Equation (6.25) suggests that if we choose the training sequences to satisfy (6.24), then the MSE approaches its lower bound. For the case of $K$ being a power of 2, $\mathcal{B}$ can be a Hadamard matrix[2].

## 6.7 Distributed Matrix Inversion

As we saw in Equation (6.16), in order to estimate $\mathcal{X}^{-1}$ we need to calculate the inverse of $\mathcal{Y}$ after each training period. The scenario for the receiver to obtain MLE of the $\mathcal{X}^{-1}$ is as follows. During the training period, the receiver collects the columns

[2]To find out how to generate a Hadamard matrix one could refer to [55].

of $\mathcal{Y}$ one by one. Once matrix $\mathcal{Y}$ is formed , i.e. at the end of the training period, the receiver starts carrying out the calculations to invert $\mathcal{Y}$. We notice that this way the computational load is not distributed evenly. During receiving the training sequences, the receiver merely collects the received data, whereas afterwards at one point the tedious calculations of the matrix inversion have to be performed. In this section we introduce an iterative method for matrix inversion, with $K$ iterations, so that in iteration $i$ it only requires column $i$ of the inverting matrix. Using this method, the receiver is able to start the computation of the inversion of matrix $\mathcal{Y}$ as soon as it receives the first column of this matrix (see Figure 6.3). This algorithm is

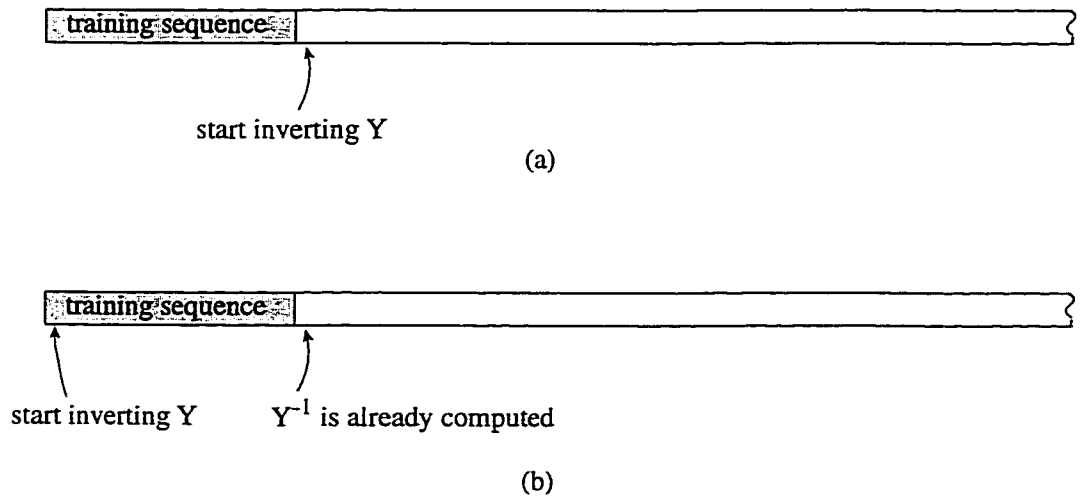


**Figure 6.3**: Comparison of matrix inversion methods; (a) conventional method, (b) proposed method.

based on the fact that the rows of $\mathcal{Y}^{-1}$ are orthonormal to the columns of $\mathcal{Y}$. Thus successive matrices $\mathcal{Y}^{-1}(0)$, $\mathcal{Y}^{-1}(1)$, $\cdots$, $\mathcal{Y}^{-1}(K)$ are computed, beginning with an arbitrary matrix $\mathcal{Y}^{-1}(0)$, so that the rows of $\mathcal{Y}^{-1}(i)$ are orthonormal to the first $i$

columns of $\mathcal{Y}$. Define scalars $\alpha_{ki}$ for $k \neq i$;

$$\begin{aligned} \alpha_{ki} &= \frac{(\text{row } k \text{ of } \mathcal{Y}^{-1}(i-1)) \cdot (\text{column } i \text{ of } \mathcal{Y})}{(\text{row } i \text{ of } \mathcal{Y}^{-1}(i-1)) \cdot (\text{column } i \text{ of } \mathcal{Y})} \\ &= \frac{\sum_{j=1}^{K} \mathcal{Y}_{k,j}^{-1}(i-1)\, \mathcal{Y}_{j,i}}{\sum_{j=1}^{K} \mathcal{Y}_{i,j}^{-1}(i-1)\, \mathcal{Y}_{j,i}} \end{aligned} \tag{6.26}$$

Then the matrices $\mathcal{Y}^{-1}(i)$ are defined recursively for $i = 1, \cdots, K$:

row $i$ of $\mathcal{Y}^{-1}(i)$ = row $i$ of $\mathcal{Y}^{-1}(i-1)$,

row $k$ of $\mathcal{Y}^{-1}(i)$ = row $k$ of $\mathcal{Y}^{-1}(i-1)$ - $\alpha_{ki}$· (row $i$ of $\mathcal{Y}^{-1}(i-1)$) for $k \neq i$.

In terms of individual matrix elements, the equivalent definition is

$$\begin{aligned} \mathcal{Y}_{i,j}^{-1}(i) &= \mathcal{Y}_{i,j}^{-1}(i-1) \\ \mathcal{Y}_{k,j}^{-1}(i) &= \mathcal{Y}_{k,j}^{-1}(i-1) - \alpha_{ki} \cdot \mathcal{Y}_{i,j}^{-1}(i-1) \qquad \text{for} \quad k \neq i, \end{aligned} \tag{6.27}$$

where $i = 1, \cdots, K$; $j = 1, \cdots, K$; $k = 1, \cdots, K$ and $k \neq i$. Then it may be shown that $\mathcal{Y}^{-1}(K) \cdot \mathcal{Y} = \mathcal{I}$. The idea of this algorithm is obtained from a method for matrix orthogonalization introduced in [123, page 25]. Figure 6.4 presents a pseudo-code for the algorithm, where we have used a stylized version of the Matlab language to express the algorithm.

The computational complexity of this algorithm has the same order as other regular matrix inversion methods, i.e. $\mathcal{O}(K^3)$.

## 6.8 Simulation Results

This section presents some simulation results. We consider a 16-user DS-CDMA system using BPSK modulation scheme and Gold codes as the spreading codes. The simulation platform is the same as what we explained in Section 5.7. A two-

```
Y^-1 = identity matrix
for i = 1 : K
    n = Y^-1(i,:) * Y(:,i)
    if n == 0
      Y is singular
    end
    y = Y(:,i)/n
    for j = 1 : K
      if (j ≠ i)
        α = Y^-1(j,:) * y
        Y^-1(j,:) = Y^-1(j,:) − α * Y^-1(i,:)
      end
    end
    Y^-1(i,:) = Y^-1(i,:)/n
end
```

**Figure 6.4:** Algorithm of distributed matrix inversion.

ray model for the channel is assumed as it was discussed in Section 5.6. A $16 \times 16$ Hadamard matrix is selected as $\mathcal{B}$, i.e. the training sequences matrix. We consider three different rms delay spreads: 62 *nSec*, 125 *nSec* and 500 *nSec*, corresponding to the second ray delays, 1, 2 and 8 chips, respectively. Figures 6.5 and 6.6 show the magnitude frequency response of the channel for $\tau_{rms} = 125$ *nSec* and $\tau_{rms} = 500$ *nSec*. The frequency response for $\tau_{rms} = 62$ *nSec* was shown in the previous chapter in Figure 5.4 on page 124 .

Figure 6.7 depicts the simulation results in which the average BER for the three different rms delay spreads are plotted versus SNR. In Figure 6.7 we have plotted two curves for each rms delay spread. In one we assume that we have perfect knowledge of the channel, and in the other we use the aforementioned MLE method to estimate the channel parameters. For $\tau_{rms} = 62$ *nSec*, which is a good approximation for indoor channels, the BER curves, have a linear shape. The almost 3 *dB* difference between the two curves is due to error in channel estimation, i.e. estimation matrix

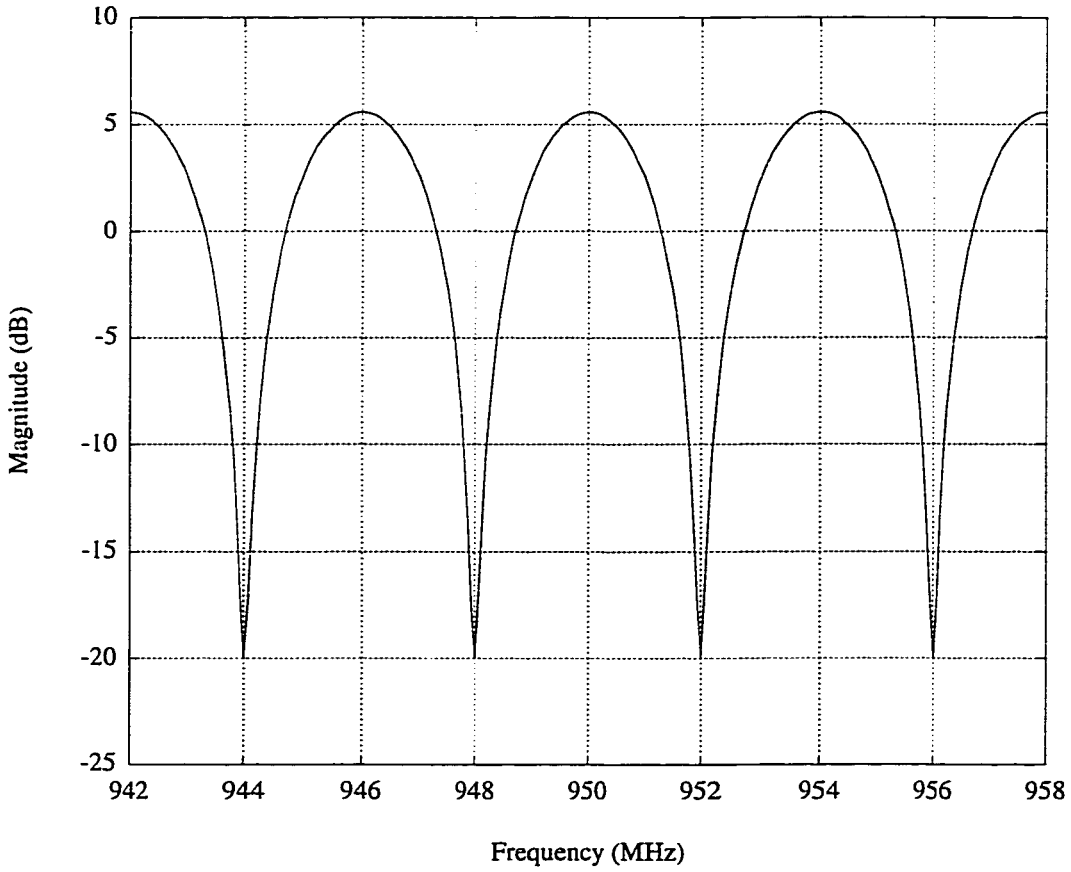


**Figure 6.5**: Frequency response of the two-ray model channel $\tau_{\text{rms}} = 125$ *nSec.*

$\mathcal{X}^{-1}$. We expect, for higher SNR, the two curves to come closer to each other, since in that case there is no error in estimating $\mathcal{X}^{-1}$. Of course the linear shapes of the curves will not be extended to higher SNR, because ISI has been ignored in the detection scheme. This is much clearer for rms delay spreads 125 *nSec* and 500 *nSec*. In these cases $\tau_{rms}$ is comparable to the symbol duration $T = 3.9$ $\mu Sec$.

## 6.9 Discussion

In this chapter, we have considered a decorrelating-type filter detector for multiuser detection in a multipath fading propagation environment. The decorrelating-type

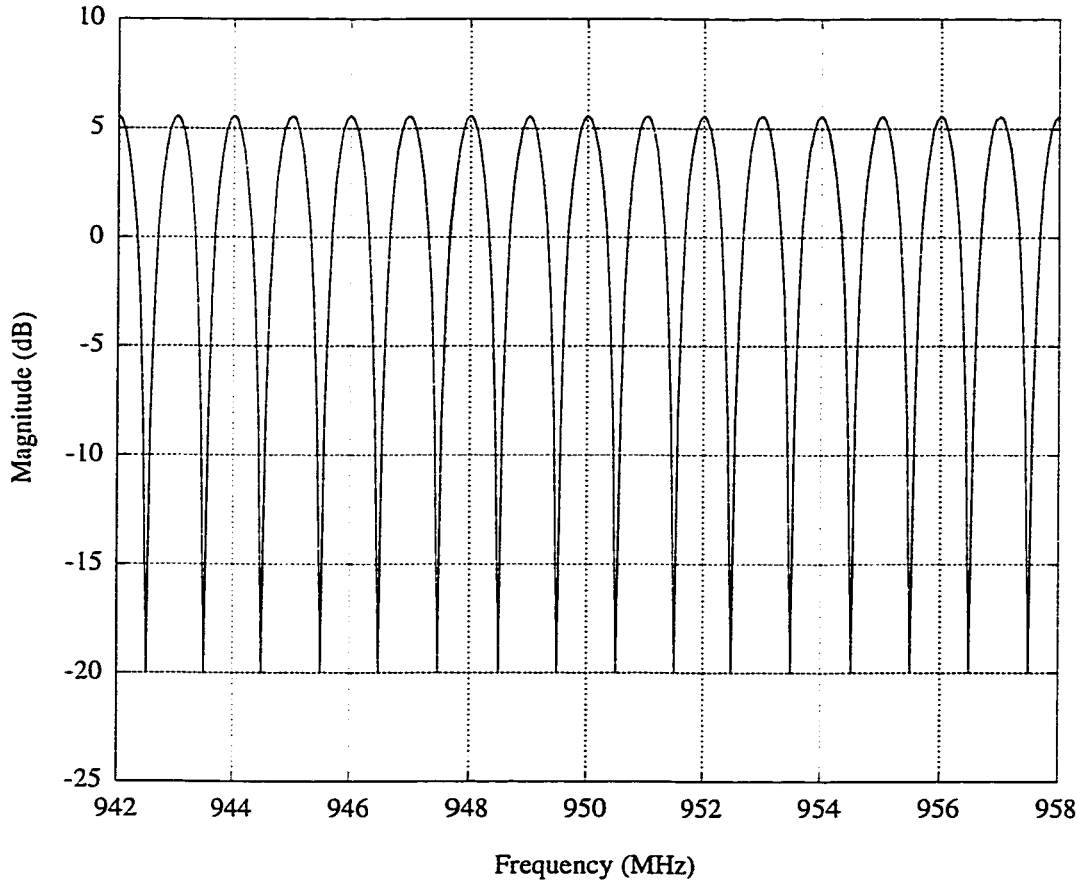


**Figure 6.6**: Frequency response of the two-ray model channel $\tau_{\text{rms}} = 500$ *nSec.*

filter processes the outputs of a bank of matched filters which are simply matched to the original spreading code of the system. We proposed a MLE scheme to estimate the parameters of the decorrelating filter. The estimation method inserts known training sequences into the information data by all users simultaneously. It was shown that the MLE of the decorrelating filter consists of multiplying the matrix of the training sequences by the inverse matrix of the outputs of matched filters. To achieve MMSE in estimation, we showed that the training sequences matrix must be chosen in a manner that the Frobenius norm of its inverse is minimized. We also showed that if we choose the matrix of the training sequences to be orthogonal, the

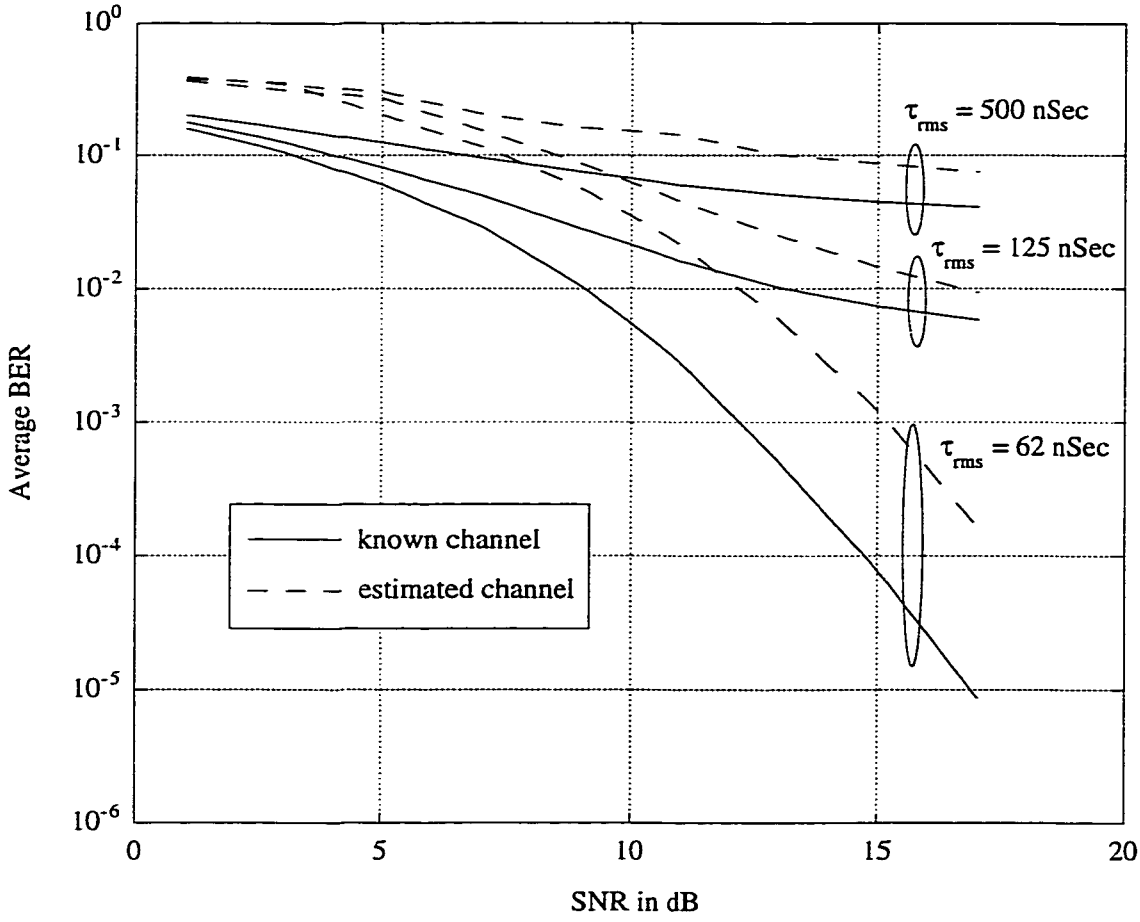


**Figure 6.7**: Average BER versus SNR

minimization of MSE is satisfied with a good approximation. When the number of users is a power of 2, a Hadamard matrix can be used for the training sequences matrix. We considered that the estimation method requires a matrix inversion at the end of each training period. An iterative matrix inversion method was introduced to distribute the computational load of the matrix inversion over the training period. This method is based on matrix orthogonalization, and in each iteration it uses only the corresponding column of the matrix.

# Chapter 7

# Symbol-Aided Channel Estimation and Multiuser Detection Using a Decorrelating-Type Decision-Feedback Detector

From the discussion in the previous chapter, it is clear that the existing error floor in the BER performance of the decorrelating-type detector is due to ISI. As it was mentioned there, the decorrelating-type detector does not attempt to combat ISI and simply ignores it. In this chapter we propose a multiuser detector that uses a feedback filter to cancel ISI. We consider a decorrelating-type decision-feedback detector for multiuser detection for frequency-selective synchronous CDMA channels. This detector operates on the outputs of matched filters which are matched to the original spreading codes of the system. We also propose and derive a maximum likelihood estimation to estimate the parameters of the forward and feedback filters.

The rest of this chapter is organized as follows. In Section 7.1 the contributions

of this chapter are put into perspective. In Section 7.2 we briefly describe the system that we consider in this chapter. Section 7.3 describes the detection scheme which we choose for multiuser detection. In Section 7.4 we derive a maximum likelihood estimation technique for the channel and for the receiver filters. Section 7.5 proposes a method to select the training sequences to approach the MMSE criterion. In Section 7.6 we present some simulation results for a 16-user CDMA system. Finally, Section 7.7 concludes the chapter with a summary of results.

## 7.1 Contribution

This chapter is a continuation of the previous chapter. The background history that we mentioned in the previous chapter may be considered as the background history for the current chapter too. On the basis of this chapter we have proposed a joint multiuser detection and channel estimation scheme [28,29]. A decorrelating-type decision-feedback detector for multiuser detection for frequency-selective synchronous CDMA channels is proposed and derived. The front-end matched filters of this detector are matched to the original spreading codes of the system, so that there is no need to estimate the matched filters impulse responses. A maximum likelihood estimation to estimate the parameters of the forward and feedback filters is derived. The estimation method is based on inserting known training sequences into the information data by all users simultaneously. To achieve MMSE in estimation, a criterion for how to select the training sequences is suggested.

## 7.2 System and Channel Model

The system model that we study in this chapter is the same as what we studied in Chapter 6. However, here we make different assumptions, which lead to a different analysis for the system. We consider a synchronous CDMA system with $K$ users. The $k$th user is assigned a finite energy code waveform $\{s_k(t), t \in [0,T]\}$ and it transmits information by modulating that waveform antipodally over one symbol interval $T$. At the receiver, the received signal can be written as:

$$r(t) = S(t, \boldsymbol{b}) + n(t) \tag{7.1}$$

where $n(t)$ is a zero-mean complex additive white Gaussian noise whose real and imaginary parts are independent and each have power spectral density $N_0/2$ and:

$$S(t, \boldsymbol{b}) = \sum_{i=-\infty}^{\infty} \sum_{k=1}^{K} \sum_{l=1}^{L} b_k(i) c_{k,l}(i) e^{j\psi_k} \sqrt{w_k(i)} s_k(t - iT - \tau_{k,l}) \tag{7.2}$$

where $i$ is the time interval index, $b_k$ is the information transmitted by user $k$, $L$ is the number of the propagation paths, $c_{k,l}$ is the fading complex envelope of the $k$th user corresponding to path $l$, and $\tau_{k,l}$ is the delay of path $l$ of user $k$. We consider slowly fading channels, which implies that $c_{k,l}(i)$ does not change during a symbol transmission time. We assume that $(\tau_{j,l})_{max} \leq T$, which suggests that the rms delay spread of the channel is comparable to the symbol duration T. In the previous section we assumed that the channel rms delay spread is small and can be ignored compared to T. Indeed we ignored ISI. That assumption lead to an error floor in the BER curves of the decorrelating-type detector. In this chapter, however, we take into account the ISI and design a detector that attempts to remove ISI as well as MAI.

## 7.3 Detection Scheme

Under the assumptions made in the previous section, we consider the received signal $r(t)$ over an arbitrary symbol interval, e.g. $i = m$. Now, suppose $r(t)$ is applied to a bank of matched filters which are matched to the set of code waveforms $\{s_k(t),\ k = 1, \cdots .K\}$, and followed by samplers at time $T$, then the following discrete-time model results:

$$\boldsymbol{y}(m) = \mathcal{U}\boldsymbol{b}(m) + \mathcal{V}\boldsymbol{b}(m-1) + \boldsymbol{n}(m) \tag{7.3}$$

where $\boldsymbol{b}(i) = [b_1(i), b_2(i), \ldots, b_K(i)]^T, (i = m, m-1)$, $\boldsymbol{y}(m) = [y_1(m), y_2(m), \ldots, y_K(m)]^T$ are the outputs of the samplers, $\mathcal{U}$ and $\mathcal{V}$ are $K \times K$ matrices whose entries are defined as:

$$\begin{aligned} \mathcal{U}_{i,j} &= \int_0^T s_i(t) \sum_{l=1}^{L} c_{j,l} s_j(t - \tau_{j,l}) dt, \\ \mathcal{V}_{i,j} &= \int_0^T s_i(t) \sum_{l=1}^{L} c_{j,l} s_j(t + T - \tau_{j,l}) dt, \end{aligned} \tag{7.4}$$

and $\boldsymbol{n}(m)$ is a zero-mean Gaussian K-vector with covariance matrix equal to $N_0\, R$, where $R$ is a nonnegative definite matrix of crosscorrelation between the assigned spreading codes as defined in Equation (2.5).

We suggest a decorrelating-type decision-feedback scheme to detect the transmitted information $\boldsymbol{b}(m)$. In (7.3) if we temporarily ignore the noise vector $\boldsymbol{n}(m)$, then, assuming that the information symbols are chosen from the finite set $\{-1, 1\}$, we have,

$$\hat{\boldsymbol{b}}(m) = \mathrm{sgn}(\mathcal{U}^{-1}(\boldsymbol{y}(m) - \mathcal{V}\hat{\boldsymbol{b}}(m-1))) \tag{7.5}$$

Figure 7.1 depicts the suggested detector. The forward filter $\mathcal{U}^{-1}$ is used to remove MAI, whereas the feedback filter $\mathcal{V}$ is used to remove ISI.

## 7.4 Estimation Scheme

A symbol-aided scheme is used to estimate $\mathcal{U}^{-1}$ and $\mathcal{V}$ in (7.5). A reference training sequence with length $2K + 1$ is inserted into the sequence of information data symbol by each transmitter. The training sequences are sent simultaneously by transmitters (Figure 7.2). We assume that the channel parameters do not change during two consecutive training sequences. This implies that the coherence time of the channel is larger than the time space between two consecutive trainings. Considering Equation (7.3), we have $2K$ systems of equations:

$$\boldsymbol{y}(i) = \mathcal{U}\boldsymbol{b}(i) + \mathcal{V}\boldsymbol{b}(i-1) + \boldsymbol{n}(i), \qquad i = 1, 2, \cdots, 2K \tag{7.6}$$

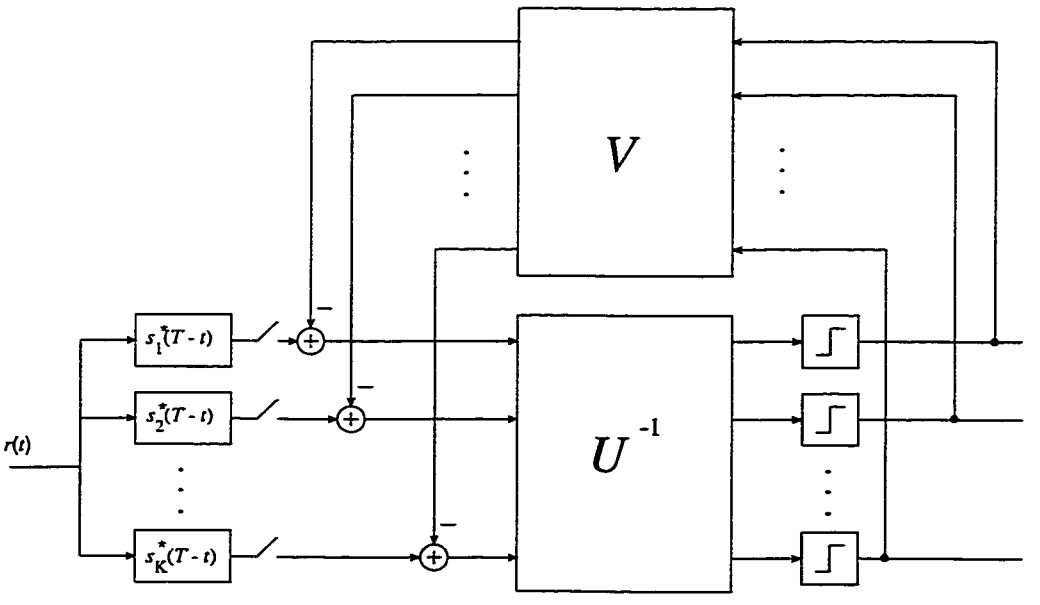


**Figure 7.1**: Decorrelating-type decision-feedback detector.

We will prove that the maximum likelihood estimations (MLE) of the two matrices $\mathcal{U}^{-1}$ and $\mathcal{V}$ can be obtained as,

$$\begin{aligned} \hat{\mathcal{U}}^{-1} &= (\mathcal{B}_{21} - \mathcal{B}_{11}\mathcal{B}_{12}^{-1}\mathcal{B}_{22})(\mathcal{Y}_2 - \mathcal{Y}_1\mathcal{B}_{12}^{-1}\mathcal{B}_{22})^{-1} \\ \hat{\mathcal{V}} &= (\mathcal{Y}_2 - \mathcal{Y}_1\mathcal{B}_{11}^{-1}\mathcal{B}_{21})(\mathcal{B}_{22} - \mathcal{B}_{12}\mathcal{B}_{11}^{-1}\mathcal{B}_{21})^{-1} \end{aligned} \tag{7.7}$$

where $\mathcal{Y}_1$ and $\mathcal{Y}_2$ are $K \times K$ matrices whose columns are made up of the vectors $\boldsymbol{y}(1), \boldsymbol{y}(2), \cdots \boldsymbol{y}(2K)$ as follows,

$$\begin{aligned} \mathcal{Y}_1 &= [\boldsymbol{y}(1), \boldsymbol{y}(2), \cdots, \boldsymbol{y}(K)] \\ \mathcal{Y}_2 &= [\boldsymbol{y}(K+1), \boldsymbol{y}(K+2), \cdots, \boldsymbol{y}(2K)] \end{aligned} \tag{7.8}$$

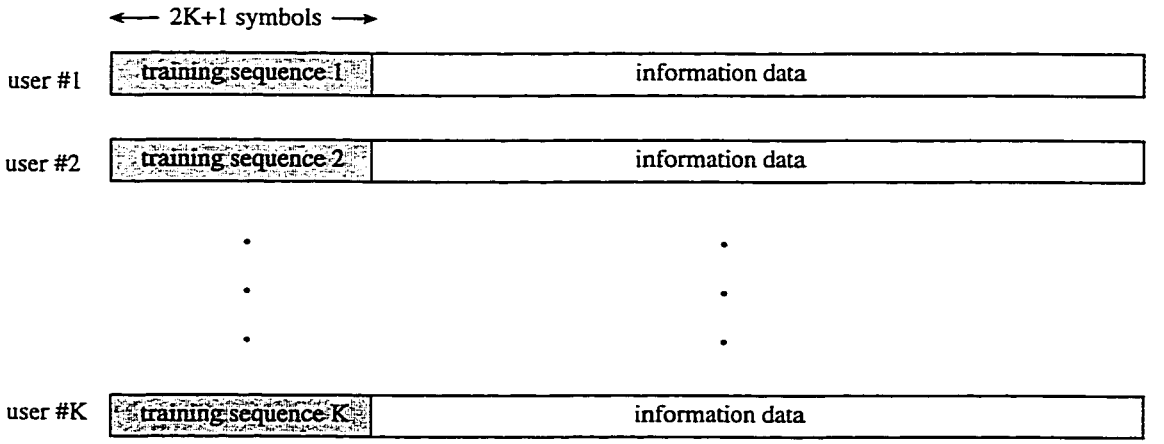


**Figure 7.2**: Synchronous transmission of training sequences

and also $\mathcal{B}_{11}, \mathcal{B}_{12}, \mathcal{B}_{21}$, and $\mathcal{B}_{22}$ are $K \times K$ matrices which are defined as,

$$\begin{aligned}
\mathcal{B}_{11} &= [\boldsymbol{b}(1), \boldsymbol{b}(2), \cdots, \boldsymbol{b}(K)], \\
\mathcal{B}_{12} &= [\boldsymbol{b}(0), \boldsymbol{b}(1), \cdots, \boldsymbol{b}(K-1)], \\
\mathcal{B}_{21} &= [\boldsymbol{b}(K+1), \boldsymbol{b}(K+2), \cdots, \boldsymbol{b}(2K)], \\
\mathcal{B}_{22} &= [\boldsymbol{b}(K), \boldsymbol{b}(K+1), \cdots, \boldsymbol{b}(2K-1)]
\end{aligned} \tag{7.9}$$

Notice that (7.6) consists of $2K$ system equations. We combine the first $K$ system equations ($i = 1, \cdots, K$) to form one matrix equation. Similarly the second $K$ system equations ($i = K+1, \cdots, 2K$) can be merged into another matrix equation. This results in the following equations:

$$\begin{aligned}
\mathcal{Y}_1 &= \mathcal{U}\mathcal{B}_{11} + \mathcal{V}\mathcal{B}_{12} + \mathcal{N}_1, \\
\mathcal{Y}_2 &= \mathcal{U}\mathcal{B}_{21} + \mathcal{V}\mathcal{B}_{22} + \mathcal{N}_2
\end{aligned} \tag{7.10}$$

where

$$\begin{aligned}
\mathcal{N}_1 &= [\boldsymbol{n}(1), \boldsymbol{n}(2), \cdots, \boldsymbol{n}(K)], \\
\mathcal{N}_2 &= [\boldsymbol{n}(K+1), \boldsymbol{n}(K+2), \cdots, \boldsymbol{n}(2K)]
\end{aligned} \tag{7.11}$$

The two equations in (7.10) can be rearranged as follows;

$$\begin{bmatrix} \mathcal{Y}_1^T \\ \mathcal{Y}_2^T \end{bmatrix} = \begin{bmatrix} \mathcal{B}_{11}^T & \mathcal{B}_{12}^T \\ \mathcal{B}_{21}^T & \mathcal{B}_{22}^T \end{bmatrix} \begin{bmatrix} \mathcal{U}^T \\ \mathcal{V}^T \end{bmatrix} + \begin{bmatrix} \mathcal{N}_1^T \\ \mathcal{N}_2^T \end{bmatrix} \tag{7.12}$$

Or equivalently:

$$\begin{pmatrix} \mathcal{Y}_1^T \\ \mathcal{Y}_2^T \end{pmatrix}_i = \begin{bmatrix} \mathcal{B}_{11}^T & \mathcal{B}_{12}^T \\ \mathcal{B}_{21}^T & \mathcal{B}_{22}^T \end{bmatrix} \begin{pmatrix} \mathcal{U}^T \\ \mathcal{V}^T \end{pmatrix}_i + \begin{pmatrix} \mathcal{N}_1^T \\ \mathcal{N}_2^T \end{pmatrix}_i , \tag{7.13}$$

$$i = 1, 2, \cdots, K$$

where $(\cdot)_i$ is used to denote column $i$ of a matrix. The elements of the noise vector $\begin{pmatrix} \mathcal{N}_1^T \\ \mathcal{N}_2^T \end{pmatrix}_i$ in equation (7.13) are the noise components of the output of the matched filter $i$ at different sampling times, hence they are independent Gaussian random variables. Consequently, the maximum likelihood estimation of the vector $\begin{pmatrix} \mathcal{U}^T \\ \mathcal{V}^T \end{pmatrix}_i$ is [122]:

$$\begin{pmatrix} \hat{\mathcal{U}}^T \\ \hat{\mathcal{V}}^T \end{pmatrix}_i = \begin{bmatrix} \mathcal{B}_{11}^T & \mathcal{B}_{12}^T \\ \mathcal{B}_{21}^T & \mathcal{B}_{22}^T \end{bmatrix}^{-1} \begin{pmatrix} \mathcal{Y}_1^T \\ \mathcal{Y}_2^T \end{pmatrix}_i , \quad i = 1, \cdots, K \tag{7.14}$$

After combining the above $K$ system equations and getting transposition we may write

$$\begin{bmatrix} \hat{\mathcal{U}} & \hat{\mathcal{V}} \end{bmatrix} = \begin{bmatrix} \mathcal{Y}_1 & \mathcal{Y}_2 \end{bmatrix} \begin{bmatrix} \mathcal{B}_{11} & \mathcal{B}_{21} \\ \mathcal{B}_{12} & \mathcal{B}_{22} \end{bmatrix}^{-1} \tag{7.15}$$

Finally, using the inversion rules for a partitioned matrix [58], we obtain (7.7). Note that the matrices $\mathcal{B}_{11}$, $\mathcal{B}_{12}$ and $(\mathcal{B}_{22} - \mathcal{B}_{12}\mathcal{B}_{11}^{-1}\mathcal{B}_{21})$ are assumed to be invertible.

Equation (7.7) suggests that the outputs of the matched filters be stored column-

wise as two matrices, $\mathcal{Y}_1$ and $\mathcal{Y}_2$, during the first half training period and the second half training period, respectively. After the first half training period, $\mathcal{Y}_1\mathcal{B}_{12}^{-1}\mathcal{B}_{22}$ and $\mathcal{Y}_1\mathcal{B}_{11}^{-1}\mathcal{B}_{21}$ can be evaluated. Finally, at the end of the training, $\mathcal{U}^{-1}$ and $\mathcal{V}$ are calculated (Figure 7.3). Evaluation of $\mathcal{U}^{-1}$ requires matrix inversion at the end of

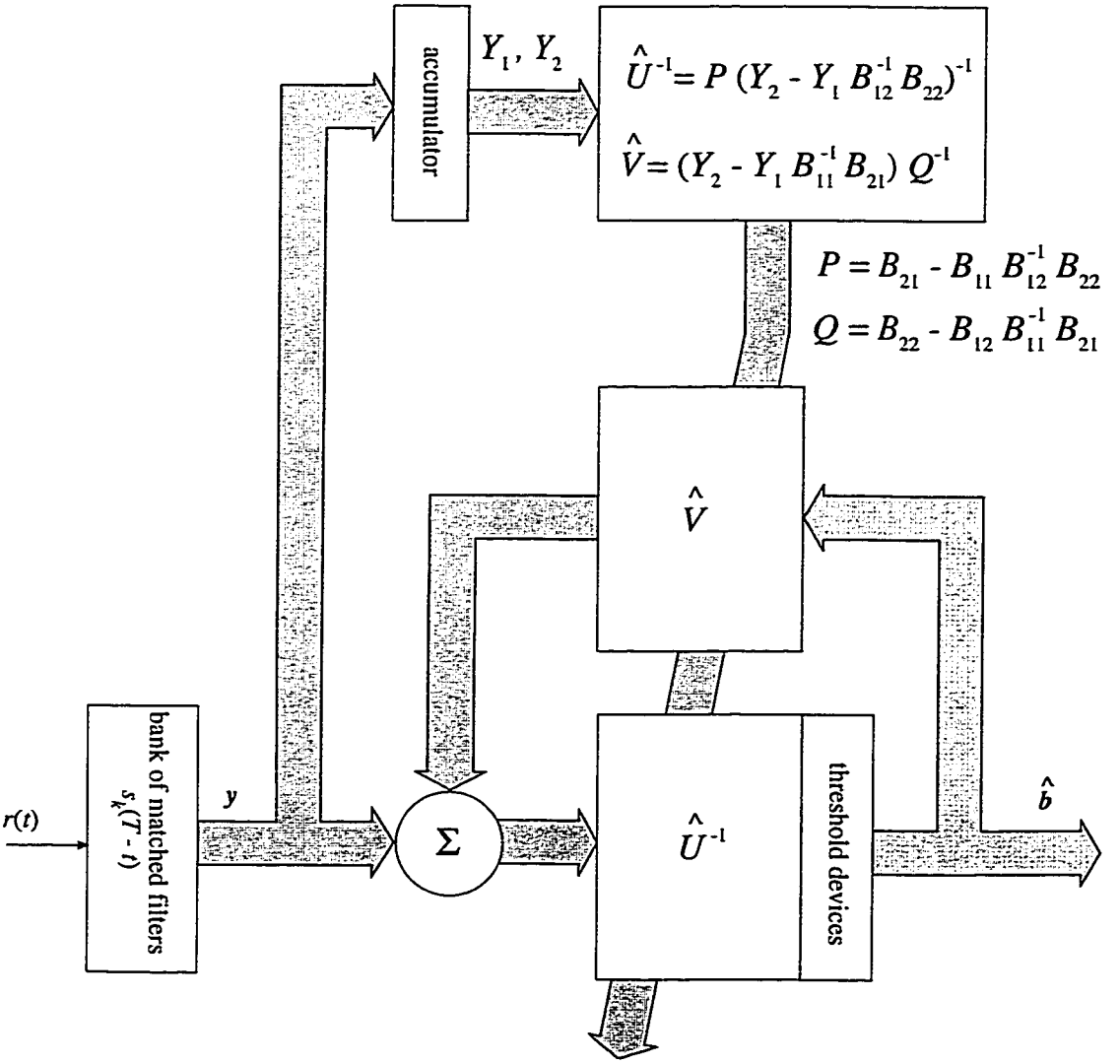


**Figure 7.3**: Decorrelating-type decision-feedback detector and estimator.

each training period. In terms of the computational complexity, this method is not homogeneous. In Section 6.7 we introduced an algorithm for matrix inversion,

which is based on matrix orthogonalization, to distribute the computational load of the matrix inversion calculations. Using this algorithm, the receiver is able to start the calculations to invert $\mathcal{Y}_2 - \mathcal{Y}_1\mathcal{B}_{12}^{-1}\mathcal{B}_{22}$ once $\mathcal{Y}_1\mathcal{B}_{12}^{-1}\mathcal{B}_{22}$ is formed, that is in the middle of the training period (see Figure 7.4).

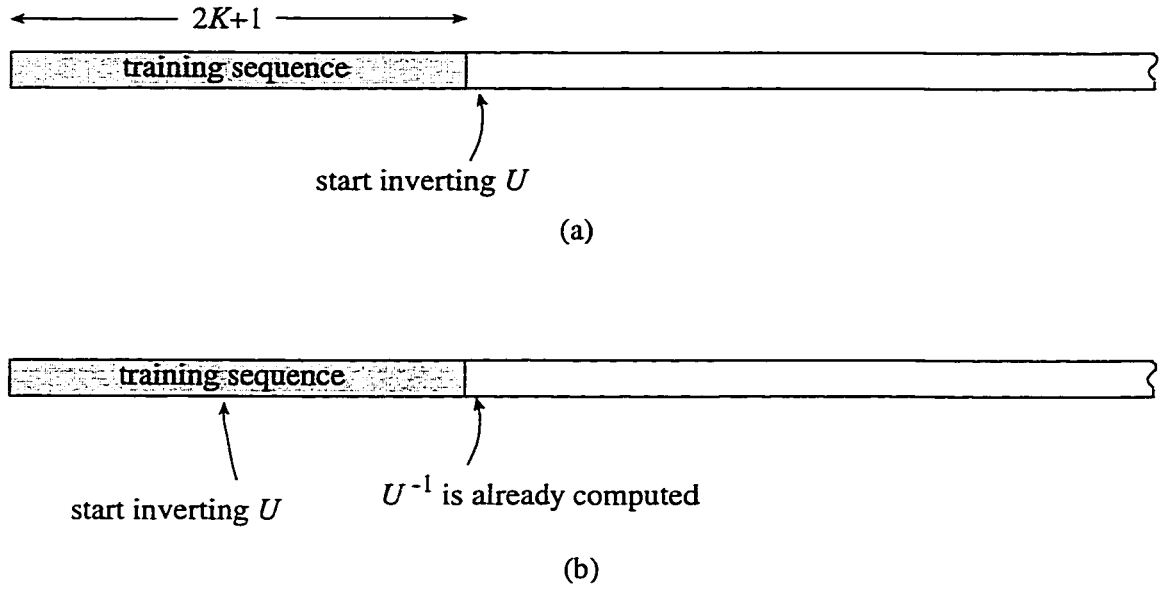


**Figure 7.4**: Comparison of matrix inversion methods; (a) conventional method, (b) proposed method.

## 7.5 Training Sequences

The MLE analysis in the previous section assumes that matrices $\mathcal{B}_{11}, \mathcal{B}_{12}, \mathcal{B}_{21}$ and $\mathcal{B}_{22}$ are known. These matrices, however, are determined by training sequences. In this section we will try to identify the training sequences which yield MMSE for estimation of both $\mathcal{U}^{-1}$ and $\mathcal{V}$. To simplify notation let us introduce the following

definitions:

$$\mathcal{W} \triangleq \begin{bmatrix} \mathcal{U} & \mathcal{V} \end{bmatrix}, \qquad \mathcal{Y} \triangleq \begin{bmatrix} \mathcal{Y}_1 & \mathcal{Y}_2 \end{bmatrix},$$

$$\mathcal{N} \triangleq \begin{bmatrix} \mathcal{N}_1 & \mathcal{N}_2 \end{bmatrix} \text{ and } \mathcal{B} \triangleq \begin{bmatrix} \mathcal{B}_{11} & \mathcal{B}_{21} \\ \mathcal{B}_{12} & \mathcal{B}_{22} \end{bmatrix}$$

Also let us define the error matrix $E$ as the difference between the estimated matrix $\hat{\mathcal{W}}$ and $\mathcal{W}$. From (7.12) and (7.15) we may write

$$E = \hat{\mathcal{W}} - \mathcal{W} = \mathcal{N}\mathcal{B}^{-1} \tag{7.16}$$

We define mean-square-error as

$$\text{MSE} = \mathcal{E}\{\sum_{i=1}^{K}\sum_{j=1}^{2K}|E_{i,j}|^2\}, \tag{7.17}$$

where $\mathcal{E}\{\cdot\}$ denotes expectation. By minimizing (7.17) with respect to $\mathcal{B}$ the MMSE sequences can be identified. We can expand equation (7.17) as follows:

$$\text{MSE} = \sum_{i=1}^{K}\sum_{j=1}^{2K}\sum_{p=1}^{2K}\sum_{q=1}^{2K}\mathcal{E}\{\mathcal{N}_{i,p}^{*}\,\mathcal{N}_{i,q}\}\; b_{p,j}^{-1^{*}}\; b_{q,j}^{-1} \tag{7.18}$$

where $b_{i,j}^{-1} = \{\mathcal{B}^{-1}\}_{i,j}$. On the other hand, we notice that the correlation between each two entries of $\mathcal{N}$ can be expressed as

$$\mathcal{E}\{n_{ij}^{*} n_{kl}\} = \begin{cases} 0, & j \neq l \\ \frac{N_0}{2} R_{i,k}, & j = l \end{cases} \tag{7.19}$$

Therefore, (7.18) can be simplified as

$$\begin{aligned} \text{MSE} &= \frac{N_0}{2} \sum_{i=1}^{K} h_{i,i} \sum_{j=1}^{2K} \sum_{p=1}^{2K} |b_{p,j}^{-1}|^2, \\ &= \frac{N_0}{2} \operatorname{tr}(R) \, \|\mathcal{B}^{-1}\|_F^2 \end{aligned} \tag{7.20}$$

where $\operatorname{tr}(R)$ is the trace of $R$ and $\| \cdot \|_F$ denotes the Frobenius norm. We notice that $\operatorname{tr}(R)$ depends on the assigned waveforms characteristics and it can be regarded as a constant. Hence, Equation (7.20) suggests that MMSE sequences are those which minimize the Frobenius norm of $\mathcal{B}^{-1}$. Considering the fact that the elements of the training sequences are chosen from a finite set, i.e. signal constellation points, we are presented with a discrete optimization problem. Moreover, matrix $\mathcal{B}$ consists of four matrices $\mathcal{B}_{11}, \mathcal{B}_{12}, \mathcal{B}_{21}$ and $\mathcal{B}_{22}$ which have some similar columns. This fact imposes a constraint on choosing $\mathcal{B}$. Indeed if we define a $K \times 2K-1$ matrix $\mathcal{F}$ as

$$\mathcal{F} = [\boldsymbol{b}(1), \boldsymbol{b}(2), \cdots, \boldsymbol{b}(2K-1)]$$

then matrix $\mathcal{B}$ can be repartitioned as

$$\mathcal{B} = \begin{bmatrix} \mathcal{F} & \boldsymbol{b}(2K) \\ \boldsymbol{b}(0) & \mathcal{F} \end{bmatrix}$$

which shows the upper part of $\mathcal{B}$ and the lower part of $\mathcal{B}$ having $(2K-1)$ similar columns.

A general approach for finding matrix $\mathcal{B}$ with the above constraint so that the Frobenius norm of its inverse is minimized is the search method. When $K$ is small (e.g. less than 8) the search method is practical, but for larger values of $K$ the tedious job of numerous large matrix inversions appears to be impractical. In this chapter, however, we use a random search method, in which we randomly generate a number of matrices as candidates for the matrix $\mathcal{B}$, and then we choose the one whose Frobenius norm of its inverse is minimum.

## 7.6 Simulation Results

In this section we present some simulation results. We consider a 16-user DS-CDMA system using BPSK modulation scheme and Gold codes with length 31 as the spreading waveforms. The simulation model is the same as what he had in Section 5.7. A two-ray model for the channel is assumed as it was explained in Section 5.6. We consider three different rms delay spreads: 62 *nSec*, 125 *nSec* and 500 *nSec*, corresponding to the second ray delays, 1, 2 and 8 chips, respectively. The magnitude frequency responses for these channels are shown in Figures 5.4, 6.5, and 6.6. One million $32 \times 32$ matrices are generated randomly as matrix $\mathcal{B}$, and the one whose Frobenius norm is the smallest is chosen as the training sequences matrix.

Figures 7.5, 7.6 and 7.7 depict the simulation results in which the average BER for the three different rms delay spreads are plotted versus SNR. Each Figure shows four curves. In one we assume that we have a perfect knowledge of the channel. This curve is in fact a lower bound for the performance of the estimation technique that we are using. To improve the performance of the estimates we can use the concepts

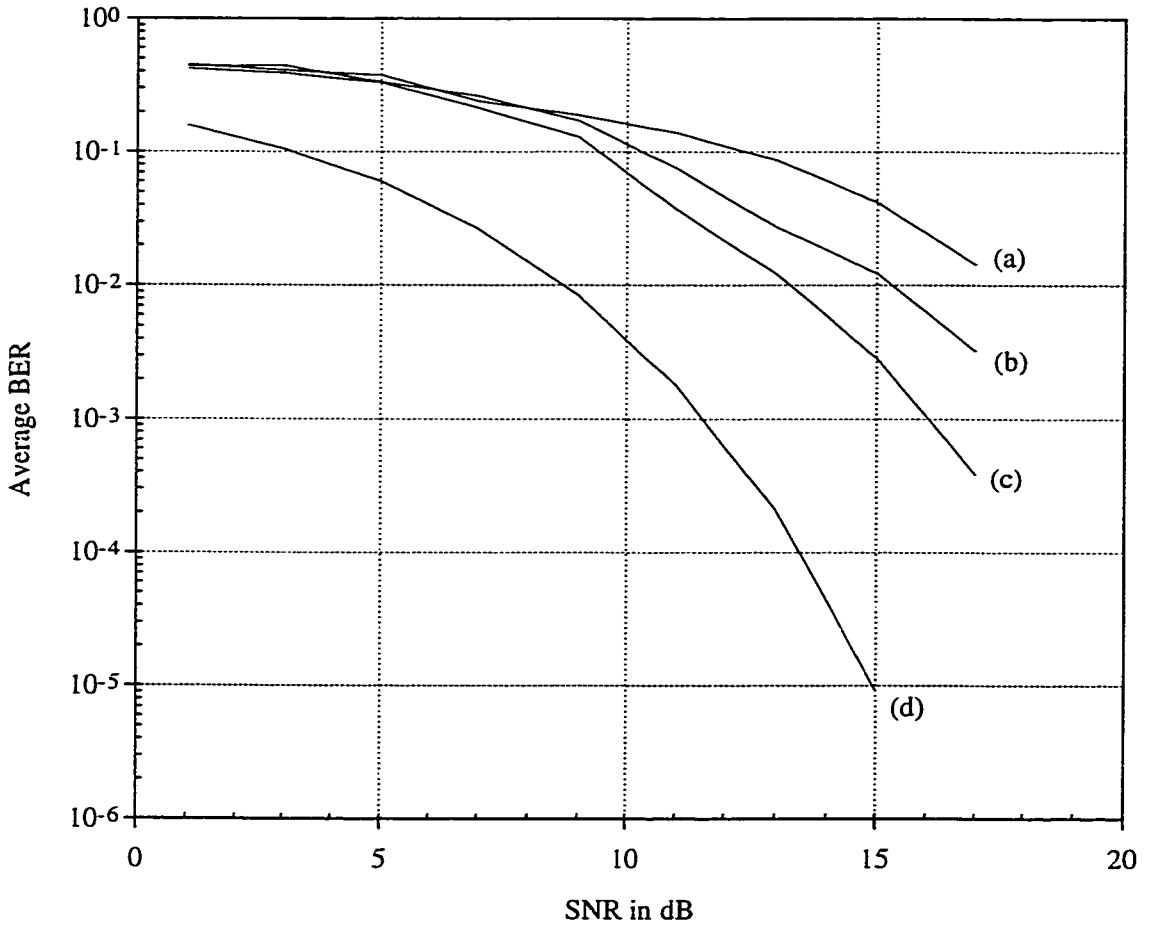


**Figure 7.5**: Average BER versus SNR for $\tau_{\text{rms}} = 62$ nSec; (a) 1 training sequence, (b) 2 training sequences, (c) 4 training sequences, (d) known channel.

of averaging, in which instead of sending only one training sequence during the training period, several training sequences are sent. Upon arrival of each training sequence, the coefficients of the filters $\mathcal{U}^{-1}$ and $\mathcal{V}$ are estimated and ultimately, after receiving all training sequences, an average of the estimates is computed and used as the filter coefficients. Figures 7.5, 7.6 and 7.7 show doubling the number of the training sequences corresponds to about 2 dB improvement in SNR. The accuracy of a decision-feedback scheme is very sensitive to the function of the feedback filter. In fact if the output of the feedback filter is not a good approximation of the

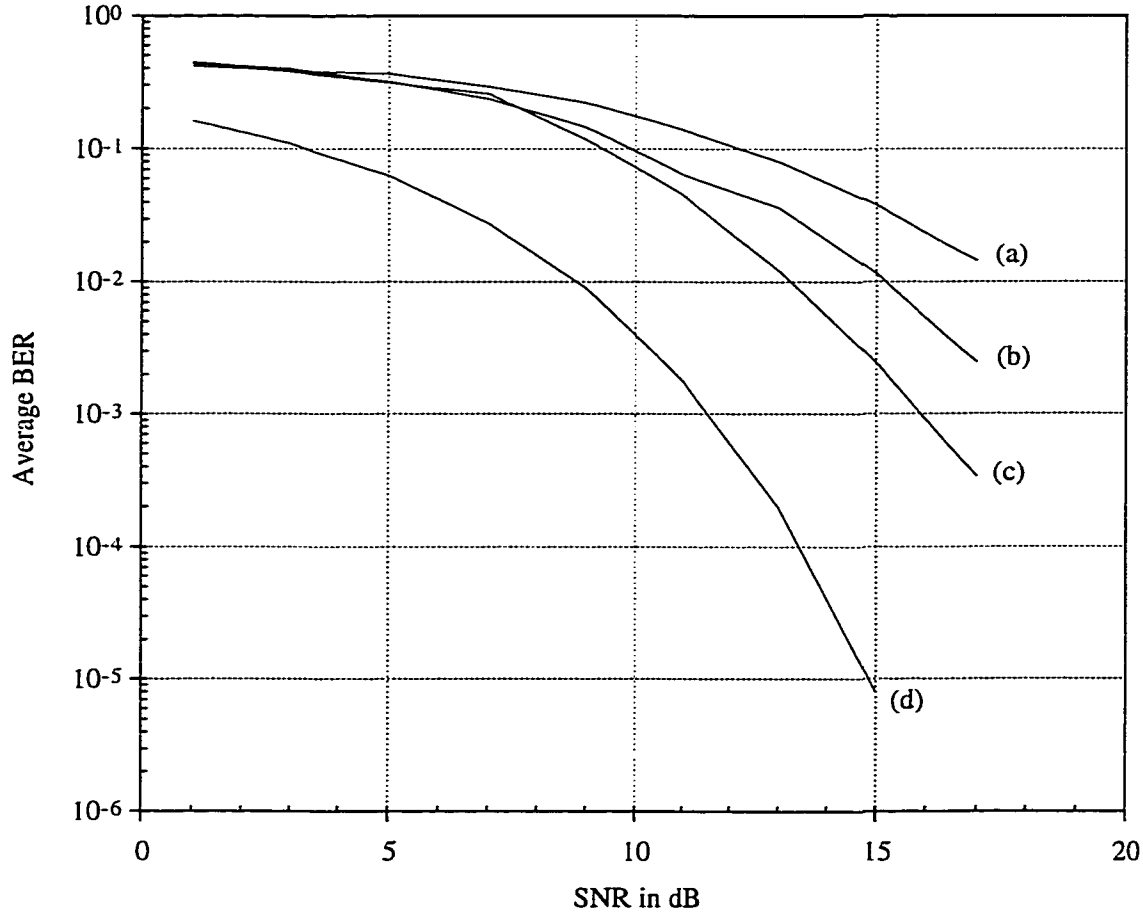


**Figure 7.6**: Average BER versus SNR for $\tau_{\text{rms}} = 125$ nSec; (a) 1 training sequence, (b) 2 training sequences, (c) 4 training sequences, (d) known channel.

interfering component from the previous symbols, then the feedback filter, not only does it not help the process of detection, but also causes some degradations in the forward filter performance. In the presented decision-feedback detector, due to errors in the estimation of the feedback filter coefficients, especially for low SNR, the average BER is higher compared to the case that only the forward filter is used (see Figure 6.7). However, the feedback filter $\mathcal{V}$ removes ISI and the BER curves, therefore, do not have error floors. Comparing Figure 7.5, 7.6 and 7.7, we find out that the BER performance of the decorrelating-type decision-feedback

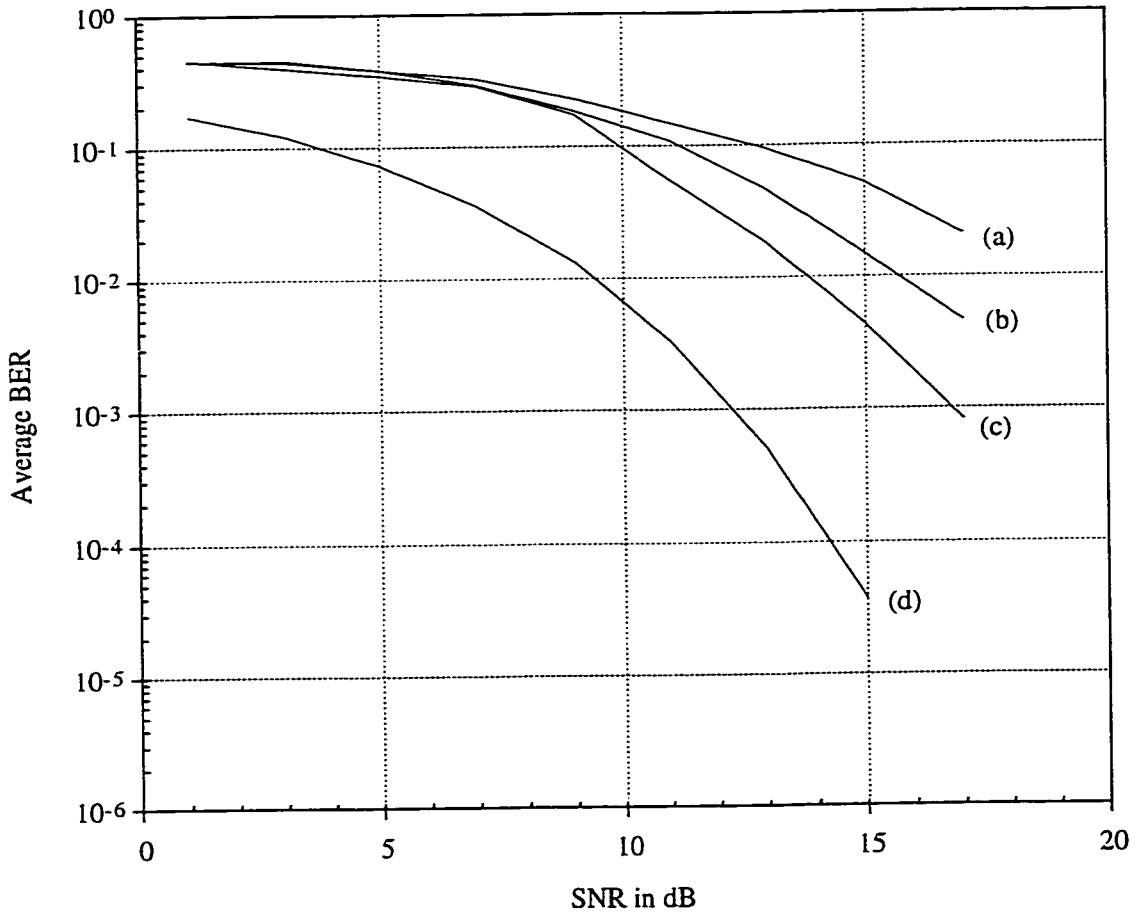


**Figure 7.7**: Average BER versus SNR for $\tau_{\text{rms}} = 500$ nSec; (a) 1 training sequence, (b) 2 training sequences, (c) 4 training sequences, (d) known channel.

detector does not depend very much on the channel rms delay spread. This is in contrast with the BER performance of the decorrelating-type detector, which was studied in the previous chapter, and which degrades dramatically with an increase in the channel rms delay spread. The reason is that the decorrelating-type decision-feedback detector, due to the ISI canceling property of the feedback filter, is not sensitive to the channel rms delay spread.

## 7.7 Discussion

In this chapter, a decorrelating-type decision-feedback detector for multiuser detection in the frequency-selective multipath propagation environment is considered. The decorrelating-type decision-feedback filter processes the outputs of a bank of matched filters which are simply matched to the original spreading codes of the system. The forward filter removes MAI, whereas the feedback filter removes ISI. We propose an MLE scheme to estimate the coefficients of both the feedback and forward filters. The estimation method inserts known training sequences into the information data by all users simultaneously. To achieve MMSE in the estimation, we showed that the training sequences matrix $\mathcal{B}$ must be chosen in a manner that the Frobenius norm of its inverse is minimized. The estimation method requires a matrix inversion at the end of each training period. Due to errors in estimation the coefficients of the feedback filter there is a degradation for low SNR in BER compared to the case when this filter is not used. However, when using this filter the BER curves have a linear shape, which is an indication of decreasing BER at high SNR. The other advantage in using the feedback filter is that the BER performance does not depend on the channel rms delay spread.

# Chapter 8

# Conclusions and Future Work

## 8.1 Conclusions

A brief summary of the accomplished work is presented in this chapter, with an emphasis on the contributions to the area of multiuser detection and channel estimation methods for wireless communications.

We studied the idea behind multiuser detection. Four suboptimum multiuser detectors as well as the conventional detector and the optimum detector were discussed and investigated with respect to both their bit-error-rate and computational complexity. The simulation results showed that in AWGN channels, these four suboptimum detectors perform fairly well compared to the optimum detector. Among them the decorrelating detector does not yield such a near-optimum bit-error-rate. The decorrelating detector, however, is the only suboptimum detector that does not require the knowledge of energies of the users. The computational load of this detector is reasonably low making it a good candidate for multiuser detection in AWGN channels.

In Chapter 3 a new suboptimum multiuser detector, namely the two-level thresh-

old detector, was introduced. The computational complexity of this detector is linear in the number of users. The performance of the two-level threshold detector is close to the optimum detector, while its computational complexity, for SNR values of interest, is a little bit more than that of the decorrelating detector. The two-level threshold detector consists of a decorrelating filter. This filter removes the MAI completely. Therefore, the outputs of the decorrelating filter contain only the desired signal embedded in noise. Assuming BPSK signaling, the likelihood ratio test in such a case suggests a simple threshold device with one level of threshold. However, the decision of the likelihood ratio test is not dependable, when the outputs of the decorrelating filter are in the neighborhood of zero. In such a case, we suggested to use the multiuser maximum likelihood detector. To avoid the high computational complexity of the multiuser maximum likelihood detector we chose local maximization of the likelihood expression.

The radius of the vicinity of zero, i.e. $\epsilon$, had a significant impact on both the bit-error-rate performance and the computational complexity of the two-level threshold detector. When $\epsilon = 0$, the two-level threshold detector becomes the decorrelating detector, and when $\epsilon = \infty$, this detector functions, although not exactly, very similar to the two-stage detector with a decorrelating detector as the first stage. The larger $\epsilon$, the higher computational complexity.

In Chapter 4 it is shown that the discrete-time channel impulse response is an appropriate way to model the multipath fading channel, since the simulations are performed in the time domain. A two-ray Rayleigh model, while simple, could achieve a large variety of multipath fading channels from flat fading to frequency-selective fading. Despite these advantages, the two-ray fading model is only a first-order approximation and does not yield a precise physical description of any

arbitrary radio channel. However, since our aim was to compare multiuser detectors, this model provided satisfactory results.

In Chapter 5 the multiuser detectors in a multipath fading frequency-selective channel were studied. There are two different approaches to design multiuser detectors for multipath channels, channel-matched and path-by-path. We chose the channel-matched approach, since it yields a better bit-error-rate performance. We observed that the conventional detector performs poorly. The two-stage detector with a conventional detector as the first stage did not offer a good performance either. The remaining detectors performed near-optimal. Amongst them, the two-level threshold detector offered the second best performance next to the optimum detector. The satisfactory performance of the two-level threshold detector was obtained by choosing the optimal value for $\epsilon$. Fortunately, for the SNRs of interest the optimal $\epsilon$ happened where the computational complexity was still low and close to that of the decorrelating detector.

In Chapters 6 and 7 we considered various issues in channel estimation for two multiuser detectors. In Chapter 6 we proposed a decorrelating-type detection scheme equipped with a symbol-aided channel estimation method. The detection scheme is similar to the channel-matched decorrelating detector except that the bank of matched filters at the front-end of the receiver are matched to the original spreading codes of the system. An MLE scheme was proposed to estimate the coefficients of the decorrelating-type filter. The estimation method inserts known training sequences in the information data sequences. It was shown that the MLE of the decorrelating-type filter is obtained by multiplying the matrix of the training sequences by the inverse matrix of the outputs of matched filters. We showed that in order to obtain MMSE in estimation, the training sequences matrix must

be chosen in such a way that the Frobenius norm of its inverse is minimized. We also showed that should we choose the training sequences matrix to be orthogonal, the MMSE would be satisfied with a good approximation. When the number of users is a power of 2, a Hadamard matrix could be used as the training sequences matrix. We noticed that the estimation method required a matrix inversion at the end of each training period. An iterative matrix inversion algorithm was introduced to distribute the computational load of the matrix inversion. The detection scheme had some drawbacks. The BER curves exhibited an error floor. This was due to the fact that while designing the decorrelating-type detector, we ignored the ISI imposed by the multipath channel. However, the existing ISI, even though small, could degrade the performance of the detection.

In Chapter 7 we proposed a decorrelating-type decision feedback detection scheme. This detector consists of two filters, forward filter and feedback filter. The forward filter is intended to remove the MAI, while the feedback filter is used to remove ISI. A symbol-aided estimation scheme based on inserting known sequences into the information data sequences was proposed. We investigated an MLE method to estimate the coefficients of both forward and feedback filters. To achieve MMSE in estimation, a criterion was derived. However, due to some constraints, the MMSE sequences had to be found using an exhaustive search method. The introduced algorithm for matrix inversion in Chapter 6 was applied to the estimation scheme to apportion the computational load. By using the feedback filter, BER curves became linear shape. Another advantage of the feedback filter was that the BER curves did not depend on the channel rms delay spread, as opposed to the decorrelating-type detector.

## 8.2 Future Work

Some topics for future study are now addressed. For numerical and simulation results presented in this dissertation, especially for the proposed two-level threshold detector, we had an optimistic condition; equal received energies are assumed (perfect power control). How the two-level threshold detector is affected by imperfect power control would be a good topic to investigate. Also a symbol-aided channel estimation is proposed and applied to two different multiuser detectors. The estimation method might be applied to other multiuser detectors presented in this dissertation too. The comparison of the estimates in terms of the bit-error-rate performance and the complexity would be another useful topic.

Most of the research conducted in the area of multiuser detection has so far focused on uncoded systems, and less work has been reported on the question of coding for multiuser systems [124]. One of the issues in this area is the design of both optimum and low complexity detectors/decoders when all users are using known error control codes [125–129]. Another issue is the design of new codes such that they match the characteristics of multipath channels and account for the MAI.

As we mentioned in Chapter 6, less work has been conducted in the parameter estimation issues in multiuser detection, and consequently, on the impact of imperfect parameter estimates on the performance of the multiuser receivers [130–132]. The problems of acquisition, timing synchronization, tracking and carrier offset sensitivity all need more attention [133–135].

# Appendix A

# Gold Sequences

Gold sequences are useful due to the large number of codes that they supply. They can be chosen so that over a set of codes available from a given genarator, the cross-correlation between the codes is uniform and bounded [136, 137]. In this appendix we describe how we generate a set of Gold sequences and then we present the Gold codes set that we have used in computer simulations as the users' spreading codes.

Let $u$ and $v$ represent a preferred [136] pair of m-sequences having period $N = 2^n - 1$. The family of codes defined by $\{u, v, u+v, u+Dv, u+D^2v, \cdots, u+D^{N-1}v\}$, where $D$ is the delay element, is called the set of Gold codes for this preferred pair of m-sequences. It can be proved that the $N+1$ elements of a Gold codes set have the property that the cross-correlation between any pair of codes in the set is three-valued [138], where those three values are $-t(n)$, $-1$, and $t(n)-2$, where

$$t(n) = \begin{cases} 1 + 2^{\frac{n+1}{2}} & \text{for } n \text{ odd}, \\ 1 + 2^{\frac{n+2}{2}} & \text{for } n \text{ even} \end{cases} \tag{A.1}$$

A set of 33 Gold codes with length 31 is generated by two preferred m-sequences

with generator polynomials (1 1 1 0 1 1) and (1 0 0 1 0 1). We choose 16 codes from the set of 33 codes so that the cross-correlation between each pair is -1. The autocorrelation of each code is obviously 31, i.e equal to the length of the code. The assigend codes are shown in Figure A.1.

user 1: 1 1 1 1 1 1 -1 1 1 1 -1 1 -1 -1 -1 1 -1 -1 -1 1 1 1 1 -1 1 1 -1
user 2: -1 1 1 -1 1 1 -1 -1 1 -1 -1 -1 1 -1 1 1 -1 1 -1 1 1 -1 -1 1 1 -1
user 3: -1 -1 1 1 -1 -1 1 -1 -1 1 1 -1 -1 1 -1 1 1 1 1 -1 -1 -1 1 -1 -1
user 4: -1 1 -1 1 -1 1 1 -1 1 -1 -1 1 1 1 1 -1 1 1 1 -1 1 -1 1 1 -1 1
user 5: 1 1 -1 -1 1 -1 1 1 -1 -1 -1 1 1 -1 1 -1 1 1 -1 -1 -1 1 -1 -1 -1
user 6: 1 -1 -1 -1 -1 1 -1 -1 -1 1 1 -1 1 1 1 -1 -1 -1 1 1 -1 -1 -1 -1 -1
user 7: 1 -1 1 1 -1 1 1 -1 1 -1 1 1 -1 -1 1 -1 1 -1 -1 1 1 1 1 1 -1 -1
user 8: -1 1 -1 -1 -1 1 1 1 -1 1 1 -1 -1 1 1 -1 -1 -1 -1 -1 -1 -1 1 1 1
user 9: 1 -1 -1 1 1 -1 -1 -1 1 -1 1 1 1 -1 -1 1 -1 -1 1 1 1 1 1 1 1 1
user 10: 1 1 1 1 -1 1 1 -1 -1 -1 1 1 1 1 -1 -1 -1 1 1 -1 -1 -1 -1 -1 1 1
user 11: -1 1 -1 -1 1 1 -1 -1 -1 -1 1 1 -1 -1 1 1 -1 1 1 1 -1 1 1 -1 -1 1
user 12: 1 -1 1 -1 -1 1 -1 -1 -1 -1 1 1 -1 -1 1 1 1 1 1 -1 -1 -1 1 1 -1 -1
user 13: -1 -1 -1 -1 1 -1 1 -1 -1 1 -1 1 1 -1 -1 -1 -1 1 1 1 -1 1 1 1 1 1
user 14: 1 -1 1 -1 1 1 -1 1 1 1 1 1 -1 -1 -1 1 1 1 1 1 -1 -1 -1 -1 -1 1
user 15: -1 -1 1 -1 1 -1 -1 1 1 1 -1 -1 1 1 1 -1 1 1 -1 -1 1 1 1 -1 1 1
user 16: -1 -1 1 1 1 -1 1 1 -1 -1 -1 -1 1 -1 -1 1 -1 -1 -1 1 1 -1 -1 -1 1

**Figure A.1**: Gold codes assigned to the system users.

# Appendix B

# The Probabilty of Error Analysis for the Optimum Detector

In this appendix, the bit-error probabilty of the optimum detector given by Equation (2.25) is derived. This derivation is adopted from [20]. Without loss of generality, consider the error probability of the first user. Denote the transmitted vector of bits as $\boldsymbol{b}^s = [b_1^s, \cdots, b_K^s]^T$ and $\boldsymbol{y}$, the vector of sufficient statistics as $\boldsymbol{y}(\boldsymbol{b}^s)$ to express its dependence on $\boldsymbol{b}^s$. Let the set of all decisions which are erroneous in the bit of the first user be denoted as $\mathcal{S}(\boldsymbol{b}^s)$. Conditioned on $\boldsymbol{b}$, the error probability of the first user, denoted as $P_1^o$, is

$$P_1^o = \mathcal{E}_{\boldsymbol{b}^s}\{P[\boldsymbol{b}^o \in \mathcal{S}(\boldsymbol{b}^s)]\} \tag{B.1}$$

where $\boldsymbol{b}^o$ is the solution of Equation (2.25) and $\mathcal{E}_{\boldsymbol{b}^s}$ denotes the expectation over the ensemble of identical, uniformly distributed $\boldsymbol{b}^s \in \{-1, +1\}^K$. Since each $\boldsymbol{b}'$ that belongs to $\mathcal{S}(\boldsymbol{b}^s)$ represents a disjoint region in $\boldsymbol{y}(\boldsymbol{b}^s)$-space, the probability that $\boldsymbol{b}^o$ is a member of $\mathcal{S}(\boldsymbol{b}^s)$ is equal to the sum of the $2^{K-1}$ probabilities that each member

of $\mathcal{S}(\boldsymbol{b}^s)$ is equal to the optimum decision. Therefore,

$$\begin{aligned} P_1^o &= \mathcal{E}_{\boldsymbol{b}^s}\{\sum_{\boldsymbol{b}^o \in \mathcal{S}(\boldsymbol{b}^s)} P[\boldsymbol{b}^o = \arg \max_{\boldsymbol{b}\in\{-1,+1\}^K} (2\boldsymbol{b}^T W \boldsymbol{y}(\boldsymbol{b}^s) - \boldsymbol{b}^T WRW\boldsymbol{b})]\} \\ &= \mathcal{E}_{\boldsymbol{b}^s}\{\sum_{\boldsymbol{b}^o \in \mathcal{S}(\boldsymbol{b}^s)} P[(2\boldsymbol{b}^{oT} W \boldsymbol{y}(\boldsymbol{b}^s) - \boldsymbol{b}^{oT} WRW\boldsymbol{b}^o) \\ &\qquad \geq (2\boldsymbol{b}^T W \boldsymbol{y}(\boldsymbol{b}^s) - \boldsymbol{b}^T WRW\boldsymbol{b}, \forall \boldsymbol{b} \neq \boldsymbol{b}^o]\} \end{aligned} \tag{B.2}$$

Noting from (2.28) that $\boldsymbol{y}(\boldsymbol{b}^s) = RW\boldsymbol{b}^s + \boldsymbol{n}$, we have

$$\begin{aligned} P_1^o = \mathcal{E}_{\boldsymbol{b}^s}\{&\sum_{\boldsymbol{b}^o \in \mathcal{S}(\boldsymbol{b}^s)} P[(\boldsymbol{b}^o - \boldsymbol{b})^T W\boldsymbol{n} \\ &\geq \frac{1}{2}(\boldsymbol{b}^{oT} WRW\boldsymbol{b}^o - \boldsymbol{b}^T WRW\boldsymbol{b}) - (\boldsymbol{b}^o - \boldsymbol{b})^T WRW\boldsymbol{b}^s, \forall \boldsymbol{b} \neq \boldsymbol{b}^o]\} \end{aligned} \tag{B.3}$$

Let $A$ denote a $2^K - 1 \times K$ matrix whose elements are either 0 or 1 so that the elements of the $i$th row represent a binary number, the decimal equivalent of which is $i$. Let $B^o$ represent a diagonal matrix with $i$th diagonal element $b_i^o$. Define $\boldsymbol{\nu} \triangleq B^o W\boldsymbol{n}$ which is therefore a zero-mean Gaussian random vector with covariance matrix $N_0 B^o WRWB^o$. Writing the $2^K - 1$ inequalities in the probability expression in (B.3) in vector notation, we have

$$P_1^o = \mathcal{E}_{\boldsymbol{b}^s}\{\sum_{\boldsymbol{b}^o \in \mathcal{S}(\boldsymbol{b}^s)} P[A\boldsymbol{\nu} \geq \boldsymbol{c}(\boldsymbol{b}^o)]\} \tag{B.4}$$

where we adopt the convention that the symbol $\geq$ between vectors denotes "greater than or equal to" component-wise. It can be verified that the $i$th element of the $2^K - 1$ column vector $\boldsymbol{c}(\boldsymbol{b}^o)$ is written as

$$c_i(\boldsymbol{b}^o) = C(\boldsymbol{b}^o, D^{(i)}\boldsymbol{b}^o) \tag{B.5}$$

where the $K \times K$ diagonal matrices $D^{(i)}$ for $i = 1, 2, \ldots, 2^K - 1$ are defined such that the $j$-th diagonal element of $D^{(i)}$ is given as $(-1)^{A_{i,j}}$, $A_{i,j}$ being the $(i,j)$th element of the matrix $A$. Further, the scalar $C(\boldsymbol{b}^o, \boldsymbol{x})$ for $\boldsymbol{x} \in \{-1, +1\}^K$ is defined as

$$C(\boldsymbol{b}^o, \boldsymbol{x}) = \frac{1}{4}(\boldsymbol{b}^{oT} W R W \boldsymbol{b}^o - \boldsymbol{x}^T W R W \boldsymbol{x}) - \frac{1}{2}(\boldsymbol{b}^o - \boldsymbol{x})^T W R W \boldsymbol{b}^s. \tag{B.6}$$

Now, note that for given $\boldsymbol{b}^s$ and $\boldsymbol{b}^o$, the probability $P[A\boldsymbol{\nu} \geq \boldsymbol{c}(\cdot)]$ is equivalent to the computation of the integral of the multivariate normal density function of $\boldsymbol{\nu}$ over the polytope $A\boldsymbol{\nu} \geq \boldsymbol{c}(\cdot)$. It is not difficult to show that this probability can be expressed in terms of sums and differences of multivariate normal distribution functions of dimension less than or equal to $K$. This fact is illustrated by an example when $K = 2$. In this case, the matrix $A$ will be $\left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$ and the probability in the curly bracket in (B.4) can be written as

$$P[A\boldsymbol{\nu} \geq \boldsymbol{c}(\cdot)] = \begin{cases} P[\nu_1 \geq c_2, \nu_2 \geq c_1]; & \text{if} \quad c_1 + c_2 \geq c_3, \\ P[\nu_1 \geq c_2, \nu_1 + \nu_2 \geq c_3] + & \\ P[\nu_2 \geq c_1, \nu_1 + \nu_2 \geq c_3] - P[\nu_1 + \nu_2 \geq c_3]; & \text{if} \quad c_1 + c_2 < c_3. \end{cases} \tag{B.7}$$

Clearly, each probability term in the above equation can be written as a bivariate or a univariate normal distribution function.

# Appendix C

# The Probabilty of Error Analysis for the Multistage Detector

This appendix consists of the derivation of the error probabilty of the two-stage detector in AWGN channels, which employs a decorrelator-type linear detector $V$, as the first stage. The derivation is adopted from [20]. Without loss of generality consider again the demodulation of $b_1$. The two-stage detector decision from (2.57) can be written as

$$\hat{b}_1(2) = \text{sgn}[z_1(1)]$$

where the decision statistic $z_1(1)$ is also obtained from (2.57) as

$$z_1(1) = \sqrt{w_1} b_1 + \sum_{k=2}^{K} \sqrt{w_k} R_{1,k}[b_k - \hat{b}_k(1)] + n_1. \tag{C.1}$$

The first, second, and the third terms will be referred to as the desired component, residual interference and additive noise, respectively. Note that the residual

interference depends on $\boldsymbol{\delta} = \boldsymbol{\beta} - \hat{\boldsymbol{\beta}}(1)$, where $\boldsymbol{\beta} \triangleq [b_2, b_3, \cdots, b_K]^T$ is the $(K-1)$-dimensional column vector that represents the bits that interfere with $b_1$ and $\hat{\boldsymbol{\beta}}(1)$ denotes the first stage estimate of $\boldsymbol{\beta}$ obtained by the linear detector $V$. Therefore, we denote the residual interference as $M_I(\boldsymbol{\beta}, \boldsymbol{\delta})$ to explicitly show its dependence on $\boldsymbol{\delta}$ and hence on $\boldsymbol{\beta}$. The decision statistic can be written as

$$z_1(1) = \sqrt{w_1} b_1 + n_1 + M_I(\boldsymbol{\beta}, \boldsymbol{\delta}) \tag{C.2}$$

The general strategy will now be outlined. The error probability is expressed as the expectation of the conditional error probability, conditioned on the bits $\boldsymbol{b} = [b_1, \boldsymbol{\beta}^T]^T$ and on the error vector $\boldsymbol{\delta}$, to obtain

$$P(\text{error}) = \mathcal{E}_{b_1,\boldsymbol{\beta}}\{\mathcal{E}_{\boldsymbol{\delta}}\{P(\text{error} \mid b_1, \boldsymbol{\beta}, \boldsymbol{\delta})\}\}. \tag{C.3}$$

Next, using (C.2) in the probability expression on the righthand side of (C.3), we have

$$P(\text{error} \mid b_1, \boldsymbol{\beta}, \boldsymbol{\delta}) = \begin{cases} P[n_1 > \sqrt{w_1} - M_I(\boldsymbol{\beta}, \boldsymbol{\delta}) \mid b_1, \boldsymbol{\beta}, \boldsymbol{\delta}] & \text{if} \quad b_1 = -1 \\ P[n_1 < -\sqrt{w_1} - M_I(\boldsymbol{\beta}, \boldsymbol{\delta}) \mid b_1, \boldsymbol{\beta}, \boldsymbol{\delta}] & \text{if} \quad b_1 = 1 \end{cases} \tag{C.4}$$

The rest of the derivation deals first with the evaluation of the probabilities in (C.4). Then the expectation over $\boldsymbol{\delta}$ in (C.3) is evaluated by specifying the probability density function of the error vector $\boldsymbol{\delta}$. Finally, removing the conditioning on the bits is simply a matter of averaging simple random variables.

In order to compute the probability in (C.4), we require the probability density function of $n_1$ conditioned on $\boldsymbol{\delta}$. Since $\boldsymbol{\delta} = \boldsymbol{\beta} - \hat{\boldsymbol{\beta}}(1)$ and $\hat{\boldsymbol{\beta}}(1) = [\hat{b}_2(1), \cdots, \hat{b}_K(1)]$

with each first stage estimate given by the linear detector $V$, we have from Equation (2.56),

$$\hat{b}_k(1) = \text{sgn}[\boldsymbol{v}_k \cdot \boldsymbol{y}] = \text{sgn}[\boldsymbol{f}_k \cdot \boldsymbol{b} + \gamma_k]. \tag{C.5}$$

Therefore, given $\boldsymbol{b}$, the error vector $\boldsymbol{\delta}$ is entirely determined by $\boldsymbol{\xi} \triangleq [\gamma_2, \cdots, \gamma_K]^T$. Therefore, we need to determine the density function of $n_1$ conditioned on $\boldsymbol{\xi}$. Consider the $K$-dimensional zero-mean Gaussian random vector $[\boldsymbol{\xi}^T n_1]^T$ which has a covariance matrix given as

$$\mathcal{E}\left[\begin{pmatrix}\boldsymbol{\xi}\\ n_1\end{pmatrix}\begin{pmatrix}\boldsymbol{\xi}^T & n_1\end{pmatrix}\right] = N_0\begin{bmatrix}\tilde{G} & \tilde{\boldsymbol{f}}_{c1}\\ \tilde{\boldsymbol{f}}_{c1}^T & H_{1,1}\end{bmatrix} \tag{C.6}$$

where each of the blocks is derived from the decomposition of $F$ and $G$ given by

$$F = VR = \begin{bmatrix}F_{1,1} & \tilde{\boldsymbol{f}}_{r1}\\ \tilde{\boldsymbol{f}}_{c1} & \tilde{F}\end{bmatrix}, \quad \text{and} \quad G = VRV^T = \begin{bmatrix}G_{1,1} & \tilde{\boldsymbol{g}}_{r1}\\ \tilde{\boldsymbol{g}}_{c1} & \tilde{G}\end{bmatrix} \tag{C.7}$$

where $\tilde{\boldsymbol{f}}_{c1}$ and $\tilde{\boldsymbol{g}}_{c1}$ are $(K-1)\times 1$ column vectors and $\tilde{\boldsymbol{f}}_{r1}$ and $\tilde{\boldsymbol{g}}_{r1}$ are $1\times(K-1)$ row vectors and $\tilde{F}$ and $\tilde{G}$ are $(K-1)\times(K-1)$ matrices. In particular, the cross-correlation between $\boldsymbol{\xi}$ and $n_1$ is given as $\mathcal{E}[\boldsymbol{\xi} n_1] = \tilde{\boldsymbol{f}}_{c1}$. However, it can be shown that, $\boldsymbol{\xi}$ is independent of $n_1$ (see [20] for details). Therefore, for a given $\boldsymbol{b}$, since $\boldsymbol{\delta}$ is a function of $\boldsymbol{\xi}$ only, the additive noise $n_1$ is independent of $\boldsymbol{\delta}$ and hence of the residual interference. As a result, the probability density function of $n_1$ conditioned on $\boldsymbol{\delta}$ is the same as its marginal zero-mean normal density with variance $R_{1,1}N_0/2 = N_0/2$.

Therefore,

$$P[n_1 > \sqrt{w_1} - M_I(\boldsymbol{\beta}, \boldsymbol{\delta}) \mid b_1 = -1, \boldsymbol{\beta}, \boldsymbol{\delta}] = \mathcal{Q}(\frac{\sqrt{w_1} - M_I(\boldsymbol{\beta}, \boldsymbol{\delta})}{\sqrt{N_0/2}}) \tag{C.8}$$

and

$$P[n_1 < -\sqrt{w_1} - M_I(\boldsymbol{\beta}, \boldsymbol{\delta}) \mid b_1 = +1, \boldsymbol{\beta}, \boldsymbol{\delta}] = \mathcal{Q}(\frac{\sqrt{w_1} + M_I(\boldsymbol{\beta}, \boldsymbol{\delta})}{\sqrt{N_0/2}}) \tag{C.9}$$

We now have to evaluate the expectation over $\boldsymbol{\delta}$, in (C.3). Note that $\boldsymbol{\delta}$ takes on $2^{K-1}$ possible values corresponding to $\delta_j \in \{2\beta_j, 0\}$ depending on whether $\hat{\beta}_j(1)$ is an error or not. Denote the set of all possible values of $\boldsymbol{\delta}$ as $\mathcal{S}(\boldsymbol{\beta})$ to show its dependence on $\boldsymbol{\beta}$. For a given $\boldsymbol{b}$, each realization of $\boldsymbol{\delta} \in \mathcal{S}(\boldsymbol{\beta})$ corresponds to $\hat{\beta}_1(1)$ being equal to $\boldsymbol{\beta} - \boldsymbol{\delta}$. Therefore, the probability of such an event is equal to the joint density function of the first stage decisions $\hat{\boldsymbol{\beta}}(1)$ evaluated at $\boldsymbol{\beta} - \boldsymbol{\delta}$. From (C.5), it is easily seen that this joint density is equivalent to the probability that $[\boldsymbol{v}_2 \cdot \boldsymbol{y}, \cdots, \boldsymbol{v}_K \cdot \boldsymbol{y}]$ belongs to a certain hyperquadrant which depends on the specific value of $\boldsymbol{\beta} - \boldsymbol{\delta}$. It is clear from (C.5) that this probability can be expressed equivalently as the $(K-1)$-dimensional normal distribution function of $\boldsymbol{\xi}$. Let us denote it as

$$P_{b_1}(\boldsymbol{\beta}, \boldsymbol{\delta}) \triangleq P[\hat{\boldsymbol{\beta}}(1) = \boldsymbol{\beta} - \boldsymbol{\delta} \mid \boldsymbol{b}, \boldsymbol{\delta}]. \tag{C.10}$$

Finally, we substitute (C.8) and (C.9) in (C.3), and write the expectation over $\boldsymbol{\delta}$ explicitly using (C.10). Denoting the error probability as $P(V)$, we have from

(C.3),

$$P(V) = \frac{1}{2}\mathcal{E}_{\boldsymbol{\beta}}\{\sum_{\boldsymbol{\delta}\in\mathcal{S}(\boldsymbol{\beta})} P_1(\boldsymbol{\beta},\boldsymbol{\delta})\mathcal{Q}(\frac{\sqrt{w_1}+M_I(\boldsymbol{\beta},\boldsymbol{\delta})}{\sqrt{N_0/2}}) + \sum_{\boldsymbol{\delta}\in\mathcal{S}(\boldsymbol{\beta})} P_{-1}(\boldsymbol{\beta},\boldsymbol{\delta})\mathcal{Q}(\frac{\sqrt{w_1}-M_I(\boldsymbol{\beta},\boldsymbol{\delta})}{\sqrt{N_0/2}})\}. \tag{C.11}$$

Clearly, the evaluation of this expression in (C.11) does not involve any numerical integrations. The evaluation of a $(K-1)$-dimensional Gaussian distribution function is, however, still required. For details on the efficient computation of the multivariate normal distribution function see [139].

# Bibliography


[1] T. Rappaport, *Wireless Communications: Principles and Practice*, Printice-Hall, New Jersey, 1996.

[2] D. Parsons, *The Mobile Radio Propagation Channel*, John Wiley, New york, 1992.

[3] A. J. Viterbi, "Wireless digital communication: A view based on three lessons learned," *IEEE Communications Magazine*, pp. 33–36, Sept. 1991.

[4] R. Pickholtz, D. Schilling, and L. Milstein, "Theory of spread-spectrum communications - A tutorial," *IEEE Trans. Communications*, pp. 855–884, May 1982.

[5] QUALCOMM, Inc., *Proposed EIA/TIA Interim Standard*, Apr. 1992.

[6] U. C. Fiebig et al., "Design study for a CDMA-based third generation mobile radio system," *IEEE Journal on Selected Areas in Communications*, vol. 12, pp. 733–743, May 1994.

[7] F. Adachi et al., "Coherent multicode DS-CDMA mobile radio access," *IEICE Trans. communications*, vol. E79-B, pp. 1316–1325, Sept. 1996.

[8] E. Dahlman, B. Gudmundson, and J. Sköld, "UMTS/IMT-2000 based on

wideband CDMA," *IEEE Communications Magazine*, vol. 36, pp. 70–80, Sept. 1998.

[9] T. Ojanperä, K. Rikkinen, H. Hakkinen, K. Pehkonen, A. Hottinen, and J. Lilleberg, "Design of a 3rd generation multirate CDMA system with multiuser detection, MUD-CDMA," in *Proc. 1996 IEEE Int. Symp. on Spread Spectrum Techniques and Applications*, Sept. 1996, pp. 334–338.

[10] T. Ojanperä and R. Prasad, "An overview of air interface multiple access for IMT-2000/UMTS," *IEEE Communications Magazine*, vol. 36, pp. 82–95, Sept. 1998.

[11] F. Adachi, M. Sawahashi, and H. Suda, "Wideband DS-CDMA for next-generation mobile communications systems," *IEEE Communications Magazine*, vol. 36, pp. 56–69, Sept. 1998.

[12] Y. Ishida, "Recent study on candidate radio transmission technology for IMT-2000," in *1st Annual CDMA Eyro. Cong.*, Oct. 1997.

[13] T. Rappaport, "The wireless revolution," *IEEE Communications Magazine*, pp. 52–71, Nov. 1991.

[14] D. Cox, "Wireless network access for personal communications," *IEEE Communications Magazine*, pp. 96–115, Dec. 1992.

[15] S. Verdú, "Minimum probability of error for asynchronous gaussian multiple-access channels," *IEEE Trans. Information Theory*, vol. 32, pp. 85–96, Jan. 1986.

[16] S. Verdú, "Optimum multiuser asymptotic efficiency," *IEEE Trans. Communications*, vol. 34, pp. 890–897, Sept. 1986.

[17] R. Lupas and S. Verdú, "Linear multiuser detectors for synchronous code-division multiple-access channels," *IEEE Trans. Information Theory*, vol. 35, pp. 123–136, Jan. 1989.

[18] R. Lupas and S. Verdú, "Near-far resistance of multiuser detectors in asynchronous channels," *IEEE Trans. Communications*, vol. 38, pp. 496–508, Apr. 1990.

[19] M. Varanasi and B. Aazhang, "Multistage detection in asynchronous code-division multiple-access communications," *IEEE Trans. Communications*, vol. 38, pp. 509–519, Apr. 1990.

[20] M. Varanasi and B. Aazhang, "Near-optimum detection in synchronous code-division multiple-access systems," *IEEE Trans. Communications*, vol. 39, pp. 725–736, May 1991.

[21] A. Duel-Hallen, "Decorrelating decision-feedback multiuser detector for synchronous code-division multiple-access channel," *IEEE Trans. Communications*, vol. 41, pp. 285–290, Feb. 1993.

[22] A. Duel-Hallen, "A family of multiuser decision-feedback detectors for synchronous code division multiple access channels," *IEEE Trans. Communications*, vol. 43, pp. 421–434, Feb./Mar./Apr. 1995.

[23] Z. Xie, C. Rushforth, and R. Short, "Multiuser signal detection using sequential decoding," *IEEE Trans. Communications*, vol. 38, pp. 578–583, 1990.

[24] Z. Xie, R. Short, and C. Rushforth, "A family of suboptimum detectors for coherent multiuser communications," *IEEE Journal on Selected Areas in Communications*, vol. 8, pp. 683–690, May 1990.

[25] L. Wei and C. Schlegel, "Synchronous DS-CDMA system with improved decorrelating decision-feedback multiuser detection," *IEEE Trans. Vehicular Technology*, vol. 43, pp. 767–772, Aug. 1994.

[26] M. Hosseinian, M. Fattouche, and A. B. Sesay, "A new multiuser detection algorithm for synchronous CDMA systems," in *Proc. 7th International Conference on Wireless Communications*, July 1995, vol. 1, pp. 191–198.

[27] M. Hosseinian, M. Fattouche, and A. B. Sesay, "Symbol-aided channel estimation and multiuser detection for CDMA systems using a decorrelating detector," in *Proc. 1997 IEEE pacific Rim Conference on Communications, Computers and Signal Processing*, Aug. 1997, vol. 2, pp. 615–619.

[28] M. Hosseinian, M. Fattouche, and A. B. Sesay, "A multiuser detection scheme with pilot symbol-aided channel estimation for synchronous CDMA systems," in *Proc. 1998 IEEE Vehicular Technology Conference*, May 1998, vol. 2, pp. 796–800.

[29] M. Hosseinian, M. Fattouche, and A. B. Sesay, "Symbol-aided channel estimation and multiuser detection for CDMA systems using a decorrelating decision-feedback detector," in *Proc. 1998 IEEE International Conference on Communications*, June 1998, vol. 2.

[30] J. S. Lim, Ed., *Speech Enhancement*, Prentice-Hall, 1983.

[31] B. widrow and S. D. Stearns, *Adaptive Signal Processing*, Prentice-Hall, 1985.

[32] K. Schneider, "Optimum detection of code division multiplexed signals," *IEEE Trans. Aerospace and Electronic Systems*, vol. AES-15, pp. 181–185, Jan. 1979.

[33] K. Schneider, "Crosstalk resistant receiver for M-ary multiplexed communications," *IEEE Trans. Aerospace and Electronic Systems*, vol. AES-16, pp. 426–433, July. 1980.

[34] R. Kohno, H. Imai, and M. Hatori, "Optimum receiver using canceller of co-channel interference in SSMA," in *1982 IECE National Conf. (Spring)*, 1982, pp. 197–201, In Japanese.

[35] R. Kohno, H. Imai, and M. Hatori, "Cancellation techniques of co-channel interference and application of viterbi algorithm in asynchronous spread spectrum multiple access systems," in *Proc.1982 Symp. on Information Theory and Its Applications*, Oct. 1982, pp. 659–666, In Japanese.

[36] R. Kohno, M. Hatori, and H. Imai, "Cancellation techniques of co-channel interference in asynchronous spread spectrum multiple access systems," *Electronics and Communications*, vol. 66-A, pp. 20–29, 1983.

[37] S. Verdú, "Minimum probability of error for asynchronous multiple-access communication systems," in *Proc. 1983 IEEE Military Communications Conf.*, Nov. 1983, vol. 1, pp. 213–219.

[38] S. Verdú, *Optimum Multiuser Signal Detection*, Ph.D. thesis, University of Illinois at Urbana-Champaign, Aug. 1984.

[39] R. Lupas-Golaszewski and S. Verdú, "Asymptotic efficiency of linear multiuser detectors," in *Proc. 25th IEEE Conf. on Decision and Control*, Dec. 1986, pp. 2094–2100.

[40] F. Xiong and E. Shwedyk, "Sequential sequence estimation for multiple chan-

nel systems with intersymbol and interchannel interference," *IEEE Trans. Communications*, vol. 41, pp. 322–331, Feb. 1993.

[41] Z. Xie, C. Rushforth, R. Short, and T. Moon, "Joint signal detection and parameter estimation in multi-user communications," *IEEE Trans. Communications*, vol. 41, pp. 1208–1216, Aug. 1993.

[42] S. T. Simmons, "Breadth-first trellis decoding with adaptive effort," *IEEE Trans. Communications*, vol. 38, pp. 3–12, Jan. 1990.

[43] S. T. Simmons and P. Senyshyn, "Reduced-search trellis decoding of coded modulations over ISI channels," in *Proc. IEEE Global Telecommunication Conf.*, 1990, pp. 393–396.

[44] J. B. Anderson, "Limited search trellis decoding of convolutional codes," *IEEE Trans. Information Theory*, vol. 35, pp. 944–955, Sept. 1989.

[45] L. Wei, *Single-user and Multiuser Modulation and Detection Methods for Mobile Communications*, Ph.D. thesis, University of South Australia, 1994.

[46] R. De guadenzi, C. Elia, and R. Viola, "Bandlimited quasi-synchronous CDMA: A novel satellite access technique for mobile and personal communication systems," *IEEE Journal on Selected Areas in Communications*, vol. 10, pp. 328–343, Feb. 1992.

[47] A. Kajiwara and M. Nakagawa, "Microcellular CDMA with a linear multiuser interference canceller," *IEEE Journal on Selected Areas in Communications*, vol. 12, pp. 605–611, May 1994.

[48] W. Van Etten, "An optimum linear receiver for multiple channel digital

transmission systems," *IEEE Trans. Communications*, vol. 23, pp. 828–834, Aug. 1975.

[49] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, 3rd edition, 1991.

[50] H. Van Trees, *Detection, Estimation and Modulation Theory, Part I*, John Wiley, New York, 1968.

[51] S. Verdú, "Computational complexity of optimum multiuser detection," *Algoritmica*, vol. 4, pp. 303–312, 1989.

[52] M. Garey and D. Johnson, *Computers and Intractability: A guide to the Theory of NP-completeness*, Freeman, San Francisco, 1979.

[53] J. Kahn, J. komlós, and E. Szemerédi, "On the probability that a random matrix is singular," *Journal of the American Mathematical Society*, vol. 8, pp. 223–240, Jan. 1995.

[54] A. Duel-Hallen, J. Holtzman, and Z. Zvonar, "Multiuser detection for CDMA systems," *IEEE Personal Communications*, vol. 2, pp. 46–58, Apr. 1995.

[55] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, third edition, 1995.

[56] H. Wu and A. Duel-Hallen, "Performance of multiuser decision-feedback detectors for flat fading synchronous CDMA channels," in *Proc. 28th Conf. on Information, Sciences and Systems*, Mar. 1994, pp. 133–138.

[57] Gene H. Golub and Charles F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, second edition, 1989.

[58] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, first edition, 1985.

[59] E. W. Ng and M. Geller, "A table of integrals of the error functions," *Journal of Research of the National Bureau of Standards-B. Mathematical Sciences*, vol. 73B, pp. 1–20, Jan.-Mar. 1969.

[60] L. Fox, *An Introduction to Numerical Linear Algebra*, Clarendon, Oxford, 1964.

[61] R. Van Norton, "The solution of linear equations by the gauss-seidel method," in *Mathematical Methods for Digital Computers*, A. Ralston and H. Wilf, Eds., pp. 56–67. John Wiley, New York, 1960.

[62] G. L. Turin et al., "A statistical model of urban multipath propagation," *IEEE Trans. Vehicular Technology*, vol. 21, pp. 1–9, Feb. 1972.

[63] T. Rappaport and S. Seidel, "Multipath propagation models for in-building communications," in *Proc. IEE 5th Int. Conf. Mobile Radio and Personal Communications*, Dec. 1989, pp. 69–74.

[64] L. Couch, *Digital and Analog Communication Systems*, Macmillan, New York, 4th edition, 1993.

[65] H. Hashemi, "The indoor radio propagation channel," *Proc. of the IEEE*, vol. 81, pp. 943–968, July 1993.

[66] D. Cox and R. Leck, "Delay Doppler characteristics of multipath delay spread and average excess delay for 910 MHz urban mobile radio paths," *IEEE Trans. Antennas and Propagation*, vol. 20, pp. 625–635, Sept. 1972.

[67] T. rappaport, S. seidel, and R. Singh, "900 MHz multipath propagation measurements for U.S. digital cellular radiotelephone," *IEEE Trans. Vehicular Technology*, pp. 132–139, May 1990.

[68] D. Cox, "Distributions of multipath delay spread and average excess delay for 910 MHz urban mobile radio paths," *IEEE Trans. Antennas and Propagation*, vol. 23, pp. 206–213, Mar. 1975.

[69] A. Saleh and R. Valenzuela, "A statistical model for indoor multipath propagation," *IEEE Journal on Selected Areas in Communications*, vol. 5, pp. 128–137, Feb. 1987.

[70] D. Devasirvatham, M. Krain, and D. Rappaport, "Radio propagation measurements at 850 MHz, 1.7 GHz, and 4.0 GHz inside two dissimilar office buildings," *Electronics Letters*, vol. 26, pp. 445–447, 1990.

[71] S. Seidel et al., "The impact of surrounding buildings on propagation for wireless in-building personal communications system design," in *1992 IEEE Vehicular Technology Conference*, May 1992, pp. 814–818.

[72] W. C. Lee, *Mobile Cellular Telecommunications Systems*, McGraw Hill, New York, 1989.

[73] J. Chuang, "The effects of time delay spread on portable communications channels with digital modulation," *IEEE Journal on Selected Areas in Communications*, vol. 5, pp. 879–889, June 1987.

[74] R. Steele, Ed., *Mobile Radioo Communications*, IEEE Press, 1994.

[75] T. Rappaport et al., "Statistical channel impulse response models for factory

and open plan building radio communication system design," *IEEE Trans. Communications*, vol. 39, pp. 794–806, May 1991.

[76] B. Woerner, J. Reed, and T. Rappaport, "Simulation issues for future wireless modems," *IEEE Communications Magazine*, pp. 19–35, July 1994.

[77] P. A. Bello and B. D. Nelin, "The effects of frequency selective fading on the binary error probabilities of incoherent and differentially coherent matched filter receivers," *IEEE Trans. Communications Systems*, vol. 11, pp. 170–186, 1963.

[78] I. Korn, "Differential phase shift keying on two-path rayleigh channel with adjacent channel interference," *IEEE Trans. Vehicular Technology*, vol. 40, pp. 461–471, 1989.

[79] C. L. Liu and K. Feher, "An asymmetrical pulse shaping technique to combat rms delay spread," *IEEE Trans. Vehicular Technology*, vol. 42, pp. 425–433, 1993.

[80] S. Yoshida, S. Onoe, and F. Ikegama, "The effect of sample timing on bit error rate performance in a multipath fading channel," *IEEE Trans. Vehicular Technology*, vol. 35, pp. 168–174, 1986.

[81] A. F. Molisch, "Statistical properties of the rms delay-spread of mobile radio channels with independent Rayleigh-fading paths," *IEEE Trans. Vehicular Technology*, vol. 45, pp. 201–205, Feb. 1996.

[82] Z. Zvonar and D. Brady, "Multiuser detection in single-path fading channels," *IEEE Trans. Communications*, vol. 42, pp. 1729–1738, Feb./March/Apr. 1994.

[83] Z. Zvonar and D. Brady, "Optimum detection in asynchronous multiple-access multipath Rayleigh fading channels," in *Proc. 26th Conf. on Information, Sciences and Systems*, Mar. 1992, pp. 826–831.

[84] Z. Zvonar and D. Brady, "Suboptimal multiuser detector for frequency-selective Rayleigh fading synchronous CDMA channels," *IEEE Trans. Communications*, vol. 43, pp. 154–157, Feb./March/Apr. 1995.

[85] Z. Zvonar, "Multiuser detection in asynchronous CDMA frequency-selective fading channels," *Wireless Personal Communications*, vol. 2, pp. 373–392, 1995/1996.

[86] Z. Zvonar and D. Brady, "Linear multipath decorrelating-receivers for CDMA frequency-selective fading channels," *IEEE Trans. Communications*, vol. 44, pp. 650–653, June 1996.

[87] M. Varanasi and S. Vesudevan, "Multiuser detectors for synchronous CDMA communication over non-selective Rician fading channels," *IEEE Trans. Communications*, vol. 42, no. 2/3/4, pp. 711–722, Feb./March/Apr. 1994.

[88] S. Vesudevan and M. Varanasi, "Optimum diversity combiner based multiuser detection for time-dispersive Rician fading channels," *IEEE Journal on Selected Areas in Communications*, vol. 12, pp. 580–591, May 1994.

[89] S. Vesudevan and M. Varanasi, "Achieving near-optimum asymptotic efficiency and fading resistance over the time-varying Rayleigh-faded CDMA channel," *IEEE Trans. Communications*, vol. 44, pp. 1130–1143, Sept. 1996.

[90] U. Fawer and B. Aazhang, "A multiuser receiver for code division multiple

access communications over multipath channels," *IEEE Trans. Communications*, vol. 43, pp. 1556–1565, Feb. 1995.

[91] U. Fawer and B. Aazhang, "Multiuser receivers for code-division multiple-access systems with trellis-based modulation," *IEEE Journal on Selected Areas in Communications*, vol. 14, pp. 1602–1609, Oct. 1996.

[92] Z. Zvonar, "Combined multiuser detection and diversity reception for wireless CDMA systems," *IEEE Trans. Vehicular Technology*, vol. 45, pp. 205–211, Feb. 1996.

[93] Z. Zvonar, "Multiuser detection and diversity combining for wireless CDMA systems," in *Wireless and mobile Communications*, J. M. Holtzman and D. J. Goodman, Eds., pp. 51–65. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1994.

[94] H. Huang and S. Schwartz, "A comparative analysis of linear multiuser detectors for fading multipath channels," in *Proc. 1994 IEEE Globecom*, Nov.-Dec. 1994, vol. 1, pp. 11–15.

[95] S. Wijayasuriya, J. McGeehan, and G. Norton, "RAKE decorrelating receiver for DS-CDMA mobile radio networks," *IEE Electronics Letters*, vol. 29, pp. 395–396, Aug. 1993.

[96] Z. Zvonar and D. Brady, "Differentially coherent multiuser detection in synchronous CDMA flat Rayleigh fading channels," *IEEE Trans. Communications*, vol. 43, pp. 1251–1255, Apr. 1995.

[97] T. Kawahara and T. Matsumoto, "Joint decorrelating multiuser detection

and channel estimation in asynchronous CDMA mobile communications channels," *IEEE Trans. Vehicular Technology*, vol. 44, pp. 506–515, Aug. 1995.

[98] Y. Bar-Ness and J. Punt, "Adaptive bootstrap CDMA multiuser detector," *Wireless Personal Communications*, , no. 1-2, pp. 55–71, 1996.

[99] D. Brady and J. Catipovic, "Adaptive multiuser detection for underwater acoustical channels," *IEEE Journal of Oceanic Engineering*, vol. 19, pp. 158–165, Apr. 1994.

[100] R. Buehrer and B. Woerner, "Analysis of adaptive multistage interference cancellation for CDMA using an improved Gaussian approximation," *IEEE Trans. Communications*, vol. 44, pp. 1308–1321, Oct. 1996.

[101] D. Chen and S. Roy, "An adaptive multiuser receiver for CDMA systems," *IEEE Journal on Selected Areas in Communications*, vol. 12, pp. 808–816, June 1994.

[102] S. Hosur, A. Tawfik, and V. Moghadam, "Adaptive multiuser receiver schemes for antenna arrays," in *Proc. 6th IEEE Int. Symp. on Personal, Indoor and Mobile Radio Communications*, Sept. 1995, vol. 3, pp. 940–944.

[103] T. Lim and L. Rasmussen, "Adaptive symbol and parameter estimation in synchronous multiuser CDMA detectors," *IEEE Trans. Communications*, vol. 45, pp. 213–220, Feb. 1997.

[104] T. Lim and S. Roy, "Adaptive detectors for multiuser CDMA," to appear in Wireless Networks, 1998.

[105] S. Miller, "An adaptive direct-sequence code-division multiple-access receiver

for multi-user interference rejection," *IEEE Trans. Communications*, vol. 43, pp. 1746–1755, Apr. 1995.

[106] S. Miller, "Training analysis of adaptive interference suppression for direct-sequence code-division multiple-access systems," *IEEE Trans. Communications*, vol. 44, pp. 488–495, Apr. 1996.

[107] U. Mitra and H. V. Poor, "Adaptive receiver algorithms for near-far resistant CDMA," *IEEE Trans. Communications*, pp. 1713–1724, Feb. 1995.

[108] U. Mitra and H. V. Poor, "Adaptive decorrelating detectors for CDMA systems," *Wireless Personal Communications*, pp. 415–440, 1996.

[109] U. Mitra and H. V. Poor, "Analysis of an adaptive decorrelating detector for synchronous CDMA channels," *IEEE Trans. Communications*, pp. 257–268, Feb. 1996.

[110] S. Verdú, "Adaptive multiuser detection," in *Code Division Multiple Access Communications*, S. G. Glisic and P. A. Leppänen, Eds., pp. 97–116. Kluwer Academic, Dordrecht, The Netherlands, 1995.

[111] X. Wang and H. V. Poor, "Adaptive multiuser detection in non-gaussian channels," in *Proc. 35th Allerton Conf. Communications, Control and Computing*, Sept. 1997.

[112] X. Wang and H. V. Poor, "Adaptive joint multiuser detection and channel estimation for multipath fading CDMA channels," *Wireless Networks*, 1998, to appear.

[113] C. Haro, J. Fonollosa, Z. Zvonar, and J. Fonollosa, "Probabilistic algorithms

for blind adaptive multiuser detection," Submitted to IEEE Trans. Signal Processing, 1997.

[114] M. L. Honig, "Rapid detection and suppression of multi-user interference in DS-CDMA," in *Proc. 1995 Int. Conf. Acoustics, Speech and Signal Processing*, May 1995, pp. 1057–1060.

[115] E. del Re and L. Ronga, "Blind adaptive MUD with silence listening," *Signal Processing*, vol. 61, pp. 91–100, Sept. 1997.

[116] S. Ulukus and R. Yates, "A blind adaptive decorrelating detector for CDMA systems," in *Proc. 1997 IEEE Global Telecommunications Conf.*, Nov. 1997, pp. 664–668.

[117] J. Fonollosa, J. Fonollosa, Z. Zvonar, and J. Catalá, "Blind multiuser detection with array observations," *Wireless Personal Communications*, vol. 6, pp. 179–196, Jan. 1998.

[118] J. Fonollosa, J. Fonollosa, Z. Zvonar, and J. Vidal, "Blind multiuser identification and detection in CDMA systems," in *Proc. 1995 IEEE Int. Conf. on Acoustic, Speech and Signal Processing*, May 1995, pp. 1876–1879.

[119] J. Fonollosa, J. Fonollosa, Z. Zvonar, and J. Vidal, "Blind multiuser deconvolution in fading and dispersive channels," in *Proc. 1995 IEEE Int. Symp. on Information Theory*, Sept. 1995, p. 384.

[120] T. Moon, Z. Xie, C. K. Rushforth, and R. Short, "Parameter estimation in a multi-user communication system," *IEEE Trans. Communications*, vol. 42, pp. 2553–2560, Aug. 1994.

[121] H. Wu and A. Duel-Hallen, "Channel estimation and multiuser detection for frequency nonselective fading synchronous CDMA channels," in *Proc. 32nd Allerton Conf. on Communications, Control and Computing.* Monticello, IL, Sept. 1994, pp. 335–344.

[122] S. M. Kay, *Fundamentals of Statistical Signal processing: Estimation Theory*, Prentice-Hall, 1993.

[123] J. R. Westlake, *A Handbook of Numerical Matrix Inversion and Solution of Linear Equations*, John Wiley, New York, 1968.

[124] P. Hoeher, "On channel coding and multiuser detection for DS/CDMA," in *Proc. 2nd Int. Conf. on Universal Personal Communications*, 1993, pp. 641–646.

[125] A. Hafeez and W. Stark, "Combined decision-feedback multiuser detection/soft-decision decoding for CDMA channels," in *Proc. 1996 IEEE Vehicular Technology Conf.*, May 1996, pp. 382–386.

[126] K. Shaheen and S. Gupta, "Adaptive combination of cancelling co-channel interference and decoding for error-correcting codes for DS-SS CDMA mobile communication system," in *Proc. 6th IEEE Int. Symp. on Personal. Indoor and Mobile Radio Communications Conf.*, Sept. 1995, vol. 2, pp. 737–741.

[127] A. El-Ezabi and A. Duel-Hallen, "Combined error correction and multiuser detection for synchronous CDMA channels," in *Proc. 33rd Allerton Conf. on Communications, Control and Computing*, Oct. 1995, pp. 1–10.

[128] T. Giallorenzi and S. Wilson, "Multiuser ml sequence estimator for convolu-

tionally coded asynchronous DS-CDMA systems," *IEEE Trans. Communications*, vol. 44, pp. 997–1008, Aug. 1996.

[129] T. Giallorenzi and S. Wilson, "Suboptimum multiuser receivers for convolutionally coded asynchronous DS-CDMA systems," *IEEE Trans. Communications*, vol. 44, pp. 1184–1196, Sept. 1996.

[130] E. Kudoh and T. Matsumoto, "Effects of power control error on the system user capacity of DS/SSMA cellular mobile radios," *IEICE Trans. Communications Japan*, vol. E75-B, pp. 524–529, June 1992.

[131] M. Varanasi, "Power control for multiuser detection," in *Proc. 1996 Conf. on Information Sciences and Systems*, Mar. 1996, pp. 866–874.

[132] H. Wu and A. Duel-Hallen, "Multiuser detection with differentially encoded data for mismatched flat Rayleigh fading CDMA channels," in *Proc. 30th Conf. on Information, Sciences and Systems*, Mar. 1996, pp. 332–337.

[133] R. Buehrer, A. Kaul, S. Striglis, and B. Woerner, "Analysis of DS-CDMA parallel interference cancellation with phase and timing errors," *IEEE Journal on Selected Areas in Communications*, vol. 14, pp. 1522–1535, Oct. 1996.

[134] S. Gray, M. Kocic, and D. Brady, "Multiuser detection in mismatched multiple-access channels," *IEEE Trans. Communications*, vol. 43, pp. 3080–3089, Dec. 1995.

[135] R. Iltis and L. Mailaender, "An adaptive multiuser detector with joint amplitude and delay estimation," *IEEE Journal on Selected Areas in Communications*, vol. 12, pp. 774–785, June 1994.

[136] P. Fan and M. Darnell, *Sequence Design for Communication Application*, Research Studies Press, U.K., 1996.

[137] R. Gold, "Optimal binary sequences for spread spectrum multiplexing," *IEEE Trans. Information Theory*, vol. 14, pp. 619–621, Oct. 1967.

[138] R. Gold, "Maximal recursive sequences with 3-valued recursive cross-correlation functions," *IEEE Trans. Information Theory*, vol. 4, pp. 154–156, Jan. 1968.

[139] R. L. Plackett, "A reduction formula for normal multivariate integrals," *Biometrika*, 1954.