Blind data detection and channel estimation for CDMA systems over fading channels

Yu, Qian

2005

Blind Data Detection and Channel Estimation for CDMA Systems over Fading Channels

Yu Qian

School of Electrical & Electronic Engineering

A thesis submitted to the Nanyang Technological University
in fulfillment of the requirement for the degree of
Doctor of Philosophy

2005

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research done by me and has not been submitted for a higher degree to any other University or Institute.

. . . .
Date

. . . .
Yu Qian

To my family,
for their love and support.

Acknowledgements

First and foremost, I would like to express my deepest appreciation and sincere gratitude to my supervisor, Associate Professor Guoan Bi, for his precious time, invaluable instruction and great support. It is he who gave me strong support and sincere encouragement throughout the tough times of the research. Within these three years, I not only learned specialized knowledge from him, but also learned the spirit of exploring science from him. This work would not have been possible without his excellent guidance and persistent encouragement.

I would like to express my sincere thanks to Miss Yongmei Wei, Mr. Gaonan Zhang, Mr. Kaixun Zhong and Mr. Hua Zhang, for their encouragement and friendship. Their useful advices on my work and sincere care gave me much help.

I would also like to acknowledge the School of EEE, Nanyang Technological University, Singapore, for awarding me the research scholarship and providing me with excellent research facilities.

I would like to sincerely thank my family for their understanding, encouragement and sacrifice. Their love and support are always the greatest inspiration to me.

Finally, I would like to thank all of my friends, graduate students, the faculty and technicians in Information Systems Research Laboratory. All of my experiences and interactions with these people have shaped me and my work, and helped making my graduate experience worthwhile and unforgettable.

Summary

Code-division multiple access (CDMA) technology is attractive as a popular multiple access scheme for personal, cellular and satellite communication services. Direct-sequence CDMA (DS-CDMA) is the most popular CDMA technique, which has been proposed for cellular telecommunication services due to numerous advantages over time-division multiple access and frequency-division multiple access. This thesis is concerned with DS-CDMA systems operating over flat fading channel, frequency-selective fading channel with white Gaussian background noise. The focus of this thesis is on blind channel estimation and blind multiuser detection.

Based on a combination of the Toeplitz displacement and the correlation matching estimation, a novel adaptive blind channel estimation method is developed for the DS-CDMA system employing long spreading codes. Since the Toeplitz displacement on the correlation matrix of the output vector removes the effects of the channel noise and other interfering users from the estimation scheme, only the knowledge of the desired user's signature waveform is required. Compared with the conventional correlation matching method and the subspace Toeplitz method, this method provides better MSE performance and higher robustness to the near-far resistance.

Bayesian Markov Chain Monte Carlo (MCMC) methodologies have recently emerged as low cost signal processing techniques with performance approaching to the theoretical optimum for wireless communication systems. In order to improve the convergence speed of the MCMC detection method, adaptive sampling is introduced to substitute the popular sampling process, Gibbs sampler. An efficient blind MCMC multiuser receiver is proposed for long code DS-CDMA system, which employs the adaptive sampling method for the Bayesian inference procedure to detect the data symbols. Compared to the Gibbs sampler receiver, for the same long code system, the proposed receiver achieves a faster convergence, lower computational complexity and comparable performance.

The Sequential Monte Carlo (SMC) methodologies are the other category of Monte Carlo signal processing methods, which provide better performance achieved by parallel processing and are better-suited to practical applications. However, it is not efficient for application of the multiuser detection of the DS-CDMA systems, since the computational complexity exponentially grows with the number of users. Two schemes are proposed to reduce the high complexity of the SMC detection. In the first one, the EM algorithm is adopted to decompose the multiuser estimation problem into a series of single user problems, then the symbol detection and channel estimation for every user are performed in parallel by SMC processing and Kalman filter. In the other scheme, a different solution is presented to decompose the superimposed observation signal, which utilizes the Cholesky factorization to decouple the signal model into separate components according to the number of users. Then under the decision-feedback framework, the channel parameters and the symbols of each user are estimated by SMC processing and Kalman filter sequentially. According to these two schemes, the EM-SMC receiver and the DF-SMC receiver are developed with the substantially reduced computational complexity which is linear to the number of users.

List of Abbreviations

AS	adaptive sampling
AWGN	additive white Gaussian noise
BER	bit-error-rate
BPSK	binary phase-shift keying
CDMA	code division multiple access
CM	correlation matching
DD	decorrelating detector
DF	decision-feedback
DS-CDMA	direct sequence code division multiple access
EM	expectation maximization
FDMA	frequency division multiple access
ISI	intersymbol interference
KF	Kalman filter
LMS	least mean square
MAI	multiple access interference

MAP	maximum a posteriori
MC	Monte Carlo
MCMC	Markov chain Monte Carlo
MF	matched filter
MIMO	multiple-input-multiple-output
ML	maximum likelihood
MMSE	minimum mean square error
MSE	mean square error
MUD	multiuser detection
PF	particle filtering
PN	pseudonoise
RLS	recursive leat square
SIS	sequential importance sampling
SMC	sequential Monte Carlo
SNR	signal-to-noise ratio
SVD	singular value decomposition
TD	Toeplitz displacement
TDMA	time division multiple access
WMF	whitened matched filter

List of Symbols

$\hat{x}$, $\hat{\mathbf{x}}$	The estimates of x and $\mathbf{x}$, respectively.
$\|a\|$	The magnitude of a.
a^*, $\mathbf{a}^*$, $\mathbf{A}^*$	The complex conjugate of a, $\mathbf{a}$ and $\mathbf{A}$, respectively.
$\mathbf{A}^{-1}$	The inverse of matrix $\mathbf{A}$.
$\mathbf{A}^T$	The transpose of matrix $\mathbf{A}$.
$\mathbf{A}^H$	The Hermitian transpose of matrix $\mathbf{A}$.
$\\|\mathbf{a}\\|$	The 2-norm of the vector $\mathbf{a}$.
$\\|\mathbf{A}\\|$	The Frobenius norm of the matrix $\mathbf{A}$.
$\mathbf{a}(i:j)$	A sub-vector formed by taking elements i through j from $\mathbf{a}(i:j)$.
$\mathbf{A}_{i,j}$	The element on the ith row and lth column of matrix $\mathbf{A}$.
$\mathbf{A}(i:j,k)$	A vector formed by taking rows i through j of the kth column from $\mathbf{A}$.
$\mathbf{A}(i:j,m:n)$	A sub-matrix formed by taking rows i through j

	and columns m through n from $\mathbf{A}$.
$\mathbf{I}_N$	A $n \times n$ identity matrix.
$\mathbf{0}_{m \times n}$	A $m \times n$ zero matrix.
$diag(a_1, \cdots, a_K)$	A square diagonal matrix with diagonal elements a_1 through a_K.
$diag(\mathbf{a}_1, \cdots, \mathbf{a}_k, \cdots, \mathbf{a}_K)$	A $KL \times K$ matrix obtained by the $L \times 1$ vectors $\mathbf{a}_1, \cdots, \mathbf{a}_K$, which satisfies $diag(\mathbf{a}_1, \cdots, \mathbf{a}_k, \cdots, \mathbf{a}_K)$ $((k-1)L+1 : kL, k) = \mathbf{a}_k$, with the other elements as 0.
$tr[\mathbf{A}]$	The trace of the matrix $\mathbf{A}$.
$vec(\mathbf{A})$	The operation of stacking the columns of $\mathbf{A}$ on top of one another. If $m \times n$ matrix $\mathbf{A} = [\mathbf{a}_1, \cdots, \mathbf{a}_n]$, then $vec(\mathbf{A}) = [\mathbf{a}_1^T, \cdots, \mathbf{a}_n^T]^T$ is a $mn \times 1$ vector.
$\mathbf{A} \otimes \mathbf{B}$	The Kronecher product between $\mathbf{A}$ and $\mathbf{B}$.
$\mathbb{E}\{\cdot\}$	Mathematical expectation.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Wireless CDMA system

Code division multiple access (CDMA) is among the most promising multiple access techniques for many emerging wireless applications [1]. In a CDMA system, users are assigned different *signature waveforms* or *codes* rather than orthogonal frequency bands, as in frequency division multiple access (FDMA), or orthogonal time slots, as in time division multiple access (TDMA). The signals of different users completely overlap in both time and frequency domains. All users transmit at the same time and each is allocated the entire available frequency bandwidth. The demodulation and separation of the signals at the receiver are achieved by using the pseudo-random code sequence that is uniquely assigned to each user's signal.

The direct-sequence code division multiple access (DS-CDMA) technique is of increasingly importance in wireless applications for cellular telecommunications services, such as the personal communications, mobile telephony, and indoor wireless networks [2][3]. The advantageous properties of DS-CDMA for these services include potentials of being implemented with lighter protocols and higher frequency-reuse capabilities, easily supporting the transmission of multi-rate information streams, being operated asynchronously over multi-path fading channels, having greater flexibility in the allocation of the channels and sharing bandwidth with narrow band services.

In a DS-CDMA system, signals are modulated with a binary pseudonoise (PN) sequence having a nearly flat spectrum before transmission, so that the transmission bandwidth is much wider than the message bandwidth. The received signal is composed of the sum of all users' signals, and is despreaded by multiplying it with the same PN sequence. This procedure is illustrated in Fig 1.1. The binary pulses comprising the PN sequences are known as chips to distinguish them from the binary bits of the data signal. The number of chips per data bit is called the *spreading factor*, *spreading gain* or *processing gain* of the system. Every user is assigned a distinct spreading PN code (also called *spreading sequence*) which is employed to distinguish it from other users at the receiver.

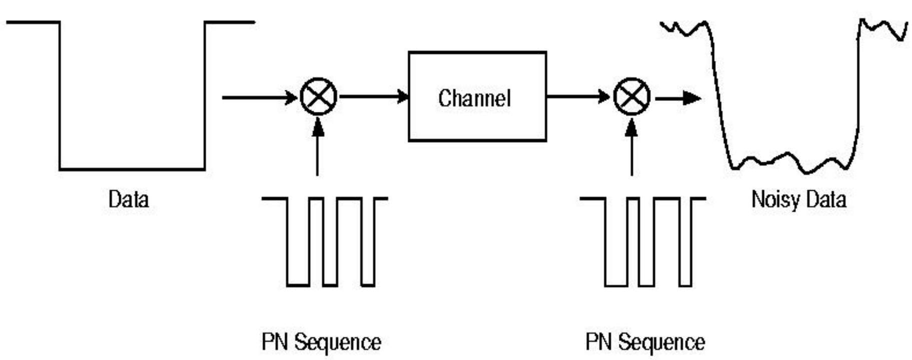


Figure 1.1: Spreading/despreading procedure

The properties of DS-CDMA systems have led it to expectations of large capacity increases over TDMA and FDMA systems. Notably, DS/CDMA techniques have been selected as the basic technology for the realization of the air-interface of third-generation (3G) cellular communication networks [4] [5].

1.2 CDMA system model

1.2.1 Basic CDMA signal model

The basic CDMA channel model of K users is described as the sum of modulated synchronous signature waveforms embedded in additive white Gaussian noise:

$$y(t) = \sum_{k=1}^{K} A_k b_k s_k(t) + \sigma w(t), \quad t \in [0, T]. \tag{1.1}$$

The notations introduced in (1.1) are defined as follows.

- T is the inverse of the data rate, or the symbol duration for the transmitted data.
- $s_k(t)$ is the deterministic signature waveform assigned to the kth user.
- A_k is the received amplitude of the kth user's signal, and A_k^2 is referred to as the energy of the kth user.
- b_k is the symbol transmitted by the kth user, and may be encoded from the source data information by different encoding schemes.
- $w(t)$ is white Gaussian noise with unit power spectral density. It models thermal noise plus other noise source unrelated to the transmitted signals.

In the synchronous model considered above, the closed-loop timing control is required to keep bit epochs aligned at the receiver. In cellular systems, the design of the reverse link is considerably simplified if the synchronization is not needed for users. To model the asynchronous CDMA system, the offsets should be introduced to model the lack of alignment of the bit epochs at the receiver, i.e., $\tau_k \in [0, T), k = 1, \ldots, K$. Then (1.1) is generalized to the asynchronous model as

$$y(t) = \sum_{k=1}^{K} \sum_{n=0}^{N-1} A_k b_k(n) s_k(t - nT - \tau_k) + \sigma w(t) \tag{1.2}$$

where $b_k(0), \ldots, b_k(n), \ldots, b_k(N-1)$ is a stream of bits sent by user, and N is assumed as the length of the frame.

1.2.2 Data streams

For the design and the analysis in this thesis, all the possible data streams $b_k(0), \ldots, b_k(n), \ldots, b_k(N-1), k = 1, \ldots, K$ are assumed to be equiprobable.

For some systems of interest in the thesis, differential encoding is made on the data stream to resolve the problem of phase uncertainties. In differential encoding [6], the transmitted stream is not the data stream itself but its transitions, i.e.,

$$d_k(n) = b_k(n) d_k^*(n-1). \tag{1.3}$$

The differential encoder is initialized with a bit $d_k(0)$ known to the receiver, which makes decisions on the consecutive products as

$$d_k(n) d_k^*(n-1) = b_k(n). \tag{1.4}$$

Differential encoding enables differentially coherent demodulation whereby the signal received in the previous interval is used to provide a carrier-phase reference in phase with the received waveform. Therefore, the differential encoding is used to facilite demodulation in the presence of phase ambiguity.

1.2.3 Signature waveform

It is not necessary to place any specific structure on the signature waveforms for many CDMA systems. In the thesis, however, the most popular structure of the signature waveform is employed, namely, *direct-sequence spread spectrum*. Direct-sequence refers to a specific approach to construct spread spectrum waveforms, characterized by chip waveform $\phi(t)$, the number of chips per bit P, and the binary sequence of length P, $c(p)$ $(c(1), \ldots, c(P))$. The direct-sequence spread-spectrum waveform with duration T is obtained by modulating the chip waveform antipodally with the binary sequence, i.e.,

$$s(t) = \sum_{p=0}^{P-1} c(p)\phi(t - pT_c), \tag{1.5}$$

where $T_c = T/P$ is the chip duration.

1.2.3.1 Spreading factor and spreading codes

The number of chips per symbol, i.e., P, is named as spreading factor, spreading gain, or processing gain. For a fixed duration of the chip waveform, the bandwidth of the signature waveform is proportional to P. Large values of P contribute to the privacy of the system because they hinder unintended receivers to unveil signature waveform and eavesdrop on the transmitted information. Large values of P also contribute to reduce the interference caused on coexisting narrow band transmissions.

In a DS-CDMA system, both the spreading factor and the chip waveform are same to all users. The different signature waveforms are distinguished by the assignment of the binary "sequence" or "code" $(c_1, \ldots, c_P)$. A wealth of combinatorial techniques exists for constructing *pseudo noise* (PN) signature sequences which achieve low cross-correlations for all possible offsets.

Assume the period of the pseudonoise waveform as P_0, two cases are considered in the practical CDMA systems. When the period of the pseudonoise waveform equals to the symbol period, i.e., $P_0/P = 1$, the corresponding signature sequence is called short spreading code; When the period of the pseudonoise waveform is larger than the symbol period, i.e., $P_0/P > 1$, the corresponding signature sequence is called long spreading code. It is noted that the spreading factor is determined by the number of chips per bit rather than the periodicity of the pseudonoise sequence. Large values of P_0/P enable the approximation of signature codes as to be more random so that the privacy is enhanced. However the use of some demodulation strategies become more difficult when $P_0/P > 1$, because the crosscorrelations of the signature waveforms vary at the data rate. Since the privacy of the system is relevant to the large value of P_0/P, the random sequences can be used to modulate the chip waveform. In the random signature model, the spreading sequences are independent and equally likely.

1.2.3.2 Long codes vs. short codes

Both short and long spreading codes are used for the wireless networks [7]. The major difference between them is that the short spreading code is periodic for every data bit and the long spreading code is aperiodic. The use of long codes guarantees that all of the active users achieve the same performance on average, which avoids the unfair situation that there exit some preferred users. With a period much larger than that of a data bit, however, the long sequences appear essentially random and destroy the bit-interval cyclostationary properties of the CDMA signals. The statistics of the multiple-access interference (MAI) change randomly from bit to bit and the performance is determined by the average interference level [8]. Most of the previous reported algorithms are based on the crucial assumption that the DS/CDMA is adopting the short codes. Many procedures of the parameter estimation so far proposed cannot be applied to systems with aperiodic long codes.

1.2.4 Channel fading

Fading refers to time-varying channel conditions. Any system with mobile transmitters and /or receivers is subject to fading. Even if the receivers and transmitters are not mobile, fading may be present in many wireless communication systems. According to the effects on the signals, the fading is divided to flat fading and frequency-selective fading. Flat fading affects the received amplitudes but does not introduce signature waveform distortion. Frequency-selective fading affects the received signals in both strength and shape of the waveform.

1.2.4.1 Flat fading

When propagation conditions change, for example, due to mobility, the received amplitudes vary with time. This feature is easily incorporated in the CDMA model:

$$y(t) = \sum_{k=1}^{K} \sum_{n=0}^{N-1} A_k(n) b_k(n) s_k(t - nT) + \sigma w(t), \tag{1.6}$$

where N is the number of the transmitted data. Whether or not the receiver is able to track the time-varying coefficients $A_k(n)$, its performance will depend on the statistical properties of those random processes. It is convenient to express (1.6) as the product between a deterministic component and a random component which contains the channel state information, that is

$$y(t) = \sum_{k=1}^{K} \sum_{n=0}^{N-1} A_k g_k(n) b_k(n) s_k(t - nT) + \sigma w(t). \tag{1.7}$$

In most systems, it is safe to assume that the random processes $\{g_k(n)\}$ are independent from user to user. Furthermore, it is usually assumed that $\{g_k(n)\}$ are wide-sense stationary processes.

1.2.4.2 Frequency-selective fading

In many multiuser systems, not only the received amplitudes vary with time but so do the received signature waveforms due to channel distortion. The additive multiple-access channel is invalidated by nonlinearities, so that the received waveform no longer comprises the noisy superposition of the various users' waveforms. Fortunately, channel distortion is often accurately modeled by a linear transformation.

The signature waveform of the kth user undergoes a linear time-varying transformation fully characterized by the complex-valued impulse response:

$$h_k(t, \tau), \tag{1.8}$$

which denotes the response of the system at time t due to a delta function at time τ. The effect of frequency-selective fading on the basic CDMA model is that the signature waveform at the receiver is not $s_k(t)$ but the convolution

$$s_k(t) * h_k(t,\tau) = \int_0^t h_k(t,\tau)s_k(\tau)d\tau. \tag{1.9}$$

Time-varying linear distortion is particularly prevalent in mobile communication systems, and the discussion of frequency-selective fading is focused on this scenario in the thesis.

1.3 Multiple access interference

One of the major features of modern wireless communication channels is the significant amount of multiple access interference which must be contented within such channels. Due to the simultaneous transmissions, the multiple access interference is inherent in many multiple access systems in which multiple transmitter-receiver pairs are communicating through the same physical channel using non-orthogonal multiplexing.

In DS-CDMA systems, the signals from different users cannot be kept orthogonal because of the random time offsets between signals. Although the spreading waveforms are designed with low-correlation, the signature waveforms are not truly orthogonal. As a result, the interference between direct-sequence users is inevitable, and denoted as the *multiple access interference* (MAI).

The existence of the MAI has a significant impact on the capacity and performance of DS-CDMA systems. As the number of interfering users increases, the amount of the MAI increases. In particular, if the users' signals have widely varying power levels, the weak users may be disturbed by the strong ones, and the effects of MAI on system performance become more substantial. Such a situation arises when the transmitters have different geographical locations relative to the receiver, since the signals of the transmitting users near the receiver undergo less amplitude attenuation than the signals of users that are further away. This is known as the near-far problem, which is illustrated in Fig 1.2. Note that, due to the different propagation effects, this problem also occurs even if all users are in the same distance from the receiver. For these reasons, MAI cancellation schemes are of great interest to CDMA systems.

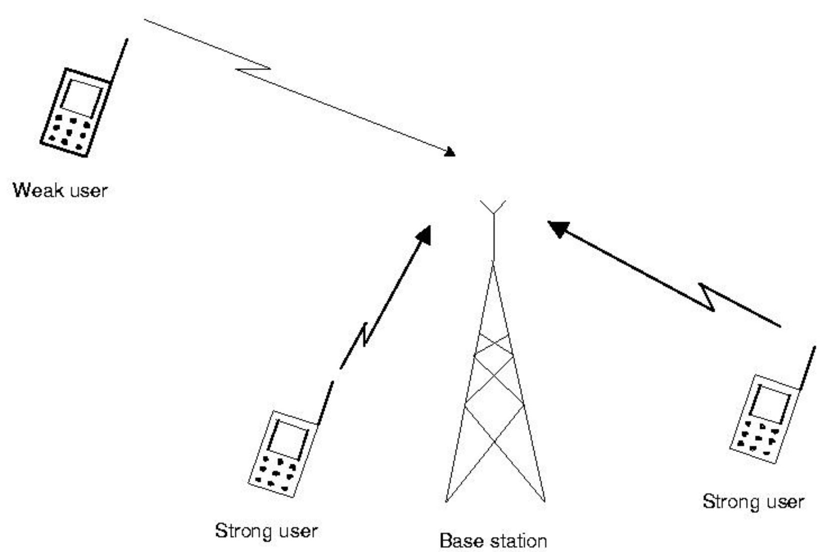


Figure 1.2: Near-far problem

All early works [9] [10] [11] approximate the MAI as a white Gaussian process, and the conventional solution to the problem of MAI is the standard single-user detection with power control. The use of power control ensures that all users reach base station with the same power, and thus no user is unfairly disadvantaged relative to others [12]. However, strict power control on mobile is not a simple task for the whole system. Moreover, power control limits the performance of the users with good channels for their transmitted powers are restricted by weak users [13]. The multiuser detection is currently of increasing attention because of the advantages in many performance capability. The single-user receiver treats the MAI as noise and has limitations in combating the effects of MAI. On the other hand, multiuser receiver deals with the MAI as a part of the information rather than noise. By processing this additional information, advanced multiuser techniques can enhance the capacity of DS-CDMA system significantly. A key issue in the design of a multiuser receiver is the synthesis of high-performance algorithms for parameter estimation. As to near-far resistance, conventional CDMA communication systems either ignore the near-far problem or try to limit it with power control. However, even a small amount of the near-far effect can drastically degrade the performance of conventional receivers. For many years, this was thought to be inherent limitation of CDMA systems until Verdu developed the optimum multiuser detector [14]. Since then, many suboptimal schemes that are near-far resistant and with lower computational complexity have been reported.

1.4 Signal detection and channel estimation

For the demodulation of DS-CDMA signals, the despreaded signal can be processed by using all kinds of detection techniques. Signal detection means to detect the transmitted source data from the observation of the received signals.

A channel is a transmission path from the transmitter to receiver. Channel estimation is important in digital communication system, especially when there is a little or no knowledge about the transmission channel. The detection problem will be solved more easily if the channel state information is known to the detector.

The blind channel estimation problem can be stated as follows: given samples of the received signal, the channel estimation is to determine the channel impulse response without the knowledge of the input signal.

In wireless CDMA communication applications, the channels exhibit different fading, which limits the system performance. Therefore the signal detection and channel estimation need to be performed to cope with fading problems.

Conventional approaches for channel estimation and signal detection in a possible non-stationary environment rely on a periodic transmission of a pilot signal (training sequence), which undesirably reduces the utilization of the available channel bandwidth. As a result, the blind approaches to channel estimation and signal detection have received much attention recently.

There are two modes for the channel estimation and signal detection. In the first approach, channel parameters are estimated firstly and detectors are then constructed based on the estimated channel parameters. These approaches simplify the design of the detectors. However, due to the separation of channel estimation and signal detection, they are suboptimal in performance and inefficient for implementation. On the other hand, the approaches for joint estimation of data and channel parameters iterate between the data detection and channel estimation to successively improve the receiver performance.

1.5 Objective and contributions

The main objectives of the research work in this thesis are to study the channel estimation and multiuser detection schemes for DS-CDMA systems in fading channels.

It is known from the previous related work of the DS-CDMA channel estimation algorithms that the channel estimation with random long spreading codes is a less explored but very valuable and attractive area. And the design of blind channel estimation algorithms that are simultaneously near-far resistant and computationally efficient is an interesting and challenging research direction. Therefore, one of our objectives is to develop blind channel estimation techniques for long code DS-CDMA system in fading channels. The major problems we will attempt to deal with are near-far problem caused by MAI and the fading problem of the single-path channels and multipath channels. At the same time, computation complexity issue will be also taken into account because it is important to the practical applications.

Multiuser detection has become one of the most active research areas in recent years, because of the capability to mitigate multiple access interference (MAI) and enhance channel capacity. The conventional deterministic methods for multiuser detection always face the tradeoff between the performance and the computational complexity, and at the same time, most of these methods are not suitable to the DS-CDMA systems with long spreading codes.

Bayesian Monte Carlo (MC) methodologies have recently emerged as low cost signal processing techniques with performance approaching to the theoretical optimum for wireless communication systems. Most MC techniques fall into one of the two categories - Markov chain Monte Carlo (MCMC) methods for batch signal processing and sequential Monte Carlo (SMC) methods for adaptive signal processing. The popular MCMC methods have one obvious problem, that is, it needs the burn-in period to achieve the convergence which results in the slow converging speed. Therefore, another objective of the thesis is to design an efficient blind Bayesian MCMC multiuser detection method for long code DS-CDMA system with a faster convergence speed. Both the cases of flat-fading channels and frequency-selective fading channels are considered in our research work.

Comparing to the MCMC methods, the SMC methods provide a better performance achieved by parallel processing and are well-suited to practical applications. However, the SMC methods are inefficient in application to multiuser detection of DS-CDMA systems, since the resulted computational complexity grows exponentially with the number of users. Thus, developing a low complexity SMC-based multiuser detection algorithm is an interesting and important research topic. Therefore it is also one objective of the research work of the thesis to develop blind SMC-based multiuser detection schemes for the DS-CDMA systems in the presence of the fading channels, which have the complexity linear to the number of users. Both short spreading codes and long spreading codes are concerned in the research work.

The contributions of this thesis are summarized as follows:

A new blind channel estimation scheme suitable for long code DS-CDMA systems is presented in the thesis. The approach is based on second-order statistics, and the spreading code of the user of interest is exploited directly via matched filtering. We exploit the statistical properties of the correlation matrix obtained from the received signal after matched filtering, and operate Toeplitz displacement to correlation matrix to remove the effects of the channel noise and other users' interferences. Then an adaptive algorithm based on correlation matching estimation is developed to yield the desired channel estimator. It is shown that the proposed estimation algorithm is computationally efficient, and offers better performance and higher robustness against the near-far problem.

A blind Bayesian MCMC multiuser detection algorithm is proposed for long code DS-CDMA systems in the presence of the unknown fading channels. In order to take advantage of the maximum a posterior (MAP) optimality of Bayesian inference and at the same time avoid the burn-in period which typically encumbers the convergence rates of MC techniques, the application of adaptive sampling to long code DS-CDMA system is considered. An efficient blind MC receiver based on the adaptive sampling algorithm is proposed for the joint of data detection and channel estimation. It is shown that the desirable improvements on convergence speed are achieved by the proposed blind MC Bayesian receiver.

A blind SMC-based formulation of multiuser detection is presented for DS-CDMA system with unknown fading channels. Firstly, the multiuser system is decoupled into separated single user systems by the EM decomposition algorithm, and then the *sequential importance sampling* (SIS) and the *Kalman filter* are combined to perform the data detection and channel estimation for every single user system in parallel. With the decomposition of the superimposed observation signals, the total computational complexity of the proposed SMC-based method is linear with the number of the users. Based on these concepts, a novel iterative receiver EM-SMC is developed for both channel estimation and data detection of DS-CDMA system. It is shown that the receiver performs well for the flat fading channels and frequency-selective fading channels with a significantly reduced computational complexity.

A different blind SMC-based multiuser detection algorithm is proposed for DS-CDMA systems in both flat fading and frequency selective fading channels. Unlike the previous SMC-based detector, we use the Cholesky factorization algorithm to decompose the observed data into separate signals according to the number of users. Then under the decision-feedback framework, the parameters of each user are estimated by SMC method and Kalman filter sequentially. Based on these concepts, a novel blind Cholesky-SMC receiver is developed for the joint channel estimation and data detection. Because it does not need iterations like EM-SMC receiver, this receiver achieves more reduction in computational complexity. It is shown that the receiver achieves comparable performance to the EM-SMC receiver with a more reduced computational complexity.

1.6 Organization of the thesis

The thesis is organized as follows:

Chapter 2 presents an overview of the design and analysis for multiuser detection is briefly described. Starting with the simplest matched filter, the optimum Maximum Likelihood detection, the major sub-optimum detection and recently emerged Bayesian Monte Carlo techniques are introduced to provide readers the basis of understanding.

The channel estimation problem of the long code DS-CDMA is considered in Chapter 3. A signal model is presented for the long code DS-CDMA system over the fading channels. A blind channel estimation algorithm based on the techniques of Toeplitz displacement and correlation matching is developed for both flat fading channels and frequency-selective fading channels. Some issues on the practical implementation are also discussed.

Based on the adaptive sampling technique, an blind Bayesian MCMC multiuser detection algorithm is proposed in Chapter 4 for long code DS-CDMA systems. Combined with the Bayesian channel estimation, an efficient blind multiuser receiver is derived for joint channel estimation and data detection. The convergence and the performance of the receiver are discussed in this Chapter.

A blind SMC-based formulation of multiuser detection is presented in Chapter 5 for DS-CDMA system with unknown fading channels. The formulation is based on the EM decomposition algorithm and the sequential importance sampling method. Incorporated with Kalman filter technique, a blind multiuser receiver is developed for joint channel estimation and data detection. The receiver is developed for short code DS-CDMA system, and the extension to long code system is also provided.

By taking the cholesky factorization algorithm, another blind SMC-based multiuser receiver is developed in Chapter 6 for channel estimation and data detection.

Chapter 7 makes comparisons between the proposed methods, which include the aspects of the requirements of coefficient knowledge, the implementation complexity, and the performance of the channel estimation and the data detection.

In Chapter 8, conclusions are presented and future research directions related to this work are recommended.

Chapter 2

Multiuser detection for CDMA systems

Multiuser detection (MUD) deals with the demodulation of mutually interfering digital streams of information. Multiuser detection exploits the considerable structure of the multiuser interference in order to increase the efficiency with which channel resources are employed. This chapter provides an overview of the major strategies of the multiuser detection for CDMA systems. It begins from the conventional matched filter, then introduces the optimum Maximum Likelihood detection, covers the major sub-optimum detection techniques, and finally describes the Bayesian Monte Carlo detection.

2.1 Matched filter

The matched filter (MF) is the simplest method to demodulate CDMA signals, and was first adopted in the implementation of CDMA receivers. Therefore it is frequently referred to as the conventional detector. The conventional detector is a bank of matched filters, as shown in Fig 2.1. Each spreading waveform is regenerated and correlated with the received signal in a separate detector branch. The outputs of the matched filters are sampled at the data bit rate, which yields "soft" estimates of the transmitted data. The final "hard" data decisions are made according to the signs of the "soft" estimates. It is shown in Fig 2.1 that the conventional detector follows a single-user detector strategy; each branch detects one user without regards to the existence of other users. The success of this detector depends on the properties of the correlation between the PN codes. We require that the autocorrelation of the spreading waveforms is much larger than the correlation between different spreading codes, i.e., the cross-correlation.

The matched filter was originally found in [15] as the maximal signal-to-noise ratio solution to a radar problem. The single-user matched filter receiver was first used for CDMA demodulation in [16]. An analysis of the capabilities of the single-user matched filter with direct-sequence signature waveforms dates back to [17].

The matched filter is computationally simple and requires no knowledge beyond the signature waveforms and timing of the desired users. However, the matched filter is optimal only for a single user in the presence of white Gaussian noise because of its single-user detection strategy. It can be highly suboptimal in the presence of MAI, especially when a significant near-far problem exists.

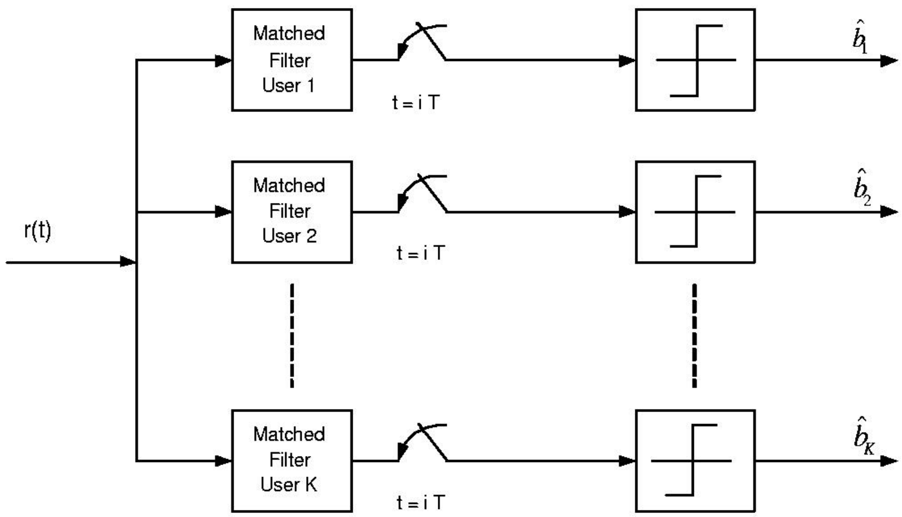


Figure 2.1: Matched filter

As discussed in Chapter 1, the CDMA systems are subject to serious multiple access interference, and the superposition of transmitted signals originates from non-ideal characteristics of the transmission medium. To alleviate the effects of the near-far problem on the matched filter, the spreading codes need to be selected with low cross-correlation properties between users, and power control techniques should be employed. The power control is used in the current CDMA systems to ensure that all users arrive at the receiver with about the same power. Very accurate and fast power control is required in the IS-95 system with single-user matched filter [3] [18]. However, strict power control on mobiles is not a simple task for the whole system. Moreover, power control limits the performance of the users with good channels for their transmitted powers are restricted by weak users [13].

2.2 Optimum multiuser detection

There has been great interest in the development and the improvement of the multiuser detection, which aims to mitigate multiple access interference (MAI) and enhance channel capacity. Unlike the single-user detection, the multiuser detection attempts to exploit the structure of the MAI, other than considering the MAI as the Gaussian noise.

The first multiuser detection was derived and analyzed by Verdu in [14][19], which is the optimum maximum likelihood (ML) detection. The optimum ML detection is based on the maximum likelihood rule, i.e. the estimator value can be found by maximizing the log-likelihood function. The derivation and analysis of the optimum multiuser detector in the presence of frequency-flat fading is due to [20], and frequency-selective fading is considered in [21]. The power of the optimum ML detection was demonstrated in [22], where it is shown, under a mild condition, that the near-far problem does not occur if the optimum maximum likelihood (ML) detection is used. However, compared to the matched filter detector, the optimal ML detector has an increased computational complexity that is exponentially increasing with the number of users [23], and the optimal ML detector requires knowledge of the received amplitudes, cross-correlations and the noise level.

2.3 Suboptimum linear detection

Since the large gaps in performance and complexity exist between the conventional matched filter and the optimum multiuser detector, the development of suboptimum multiuser detectors are motivated to achieve good performance/complexity tradeoffs. Among the developed suboptimum multiuser detectors, there are linear and non-linear detectors [24].

2.3.1 Decorrelating detection

The decorrelating detector consists of a bank of matched filters (one matched filter for one user), followed by the inverted cross-correlation matrix. After processing the output of the matched filter bank with the invertion of the cross-correlation, each output of the decorrelating detector is free from interference caused by any of other users. The only source of interference is the background noise. That is why this detector is called a decorrelating detector since it decorrelates the received multiple user signals into their individual components.

The derivation of decorrelating detector (DD) for synchronous channels was presented in [14]. The asynchronous decorrelating detctor was obtained in [25] [26]. Decorrelation for multiuser channels subject to intersymbol interference was also considered in [27] [28] and [29]. Differentially-coherent decorrelating detectors were studied in [30] for synchronous channels and in [31] [32] [33] for asynchronous channels. The adaptive decorrelating detectors were provided in [34] [35] [36] [37] and [38]. The capacity achievable with the decorrelating detectors in a frequency-selective channel was considered in [39]. In [40], a decorrelating detector which is robust against non-Gaussian noise was proposed.

The decorrelating detector does not require the knowledge of the received amplitudes. This negligence of information means that the performance of the DD cannot be better than the optimum detector studied previously. The DD can be implemented in a decentralized manner since demodulation of each user can be implemented independently.

The decorrelaring detector is not only a simple and natural strategy but also an optimal method according to three different criteria: least-squares, near-far resistance, and maximum-likelihood when the received amplitudes are unknown. However, the DD solution can be seen to eliminate all multiuser interference at the expense of increasing noise, the noise enhancement is present in the decorrelaring detector.

2.3.2 Suboptimum maximum likelihood detection

The optimal ML estimator has a computational complexity that is exponentially increasing with the number of users. Hence, numerous suboptimal ML algorithms were proposed, by sacrificing performance for the sake of a reduced complexity. In [41], a maximum-likelihood procedure is proposed to estimate the relevant parameters from the active users. A single-user ML method was proposed to decompose the multiuser problem into a series of single-user problems with reduced complexity [42]. In [43], a large-sample ML single-user delay estimator was proposed to model the MAI as non-white Gaussian process. Based on an ML criterion, a new data-aided detector over frequency-selective fading channels was developed in [44].

2.3.3 Minimum mean squared error detection

Since the decorrelating detector leads to severe noise enhancement, it is natural to consider another linear detector that trades off noise enhancement with some small but tolerable multiuser interference. Therefore a minimum mean square error (MMSE) detector was formulated in [45] [46] and [47]. Based on the mean square error (MSE) criterion, the estimators were found by minimizing the MSE of the received vector. The MMSE MUD strikes a balance in mitigating both additive noise and multiple access interference. However, the MSE detector requires knowledge of the received amplitudes and the noise level.

The MMSE receiver has received considerable attention due to its simplicity of implementation, better performance, and amenability to adaptive implementation. It has been shown that the MMSE receiver can be used to suppress multiple-access interference. The MMSE receiver can be implemented adaptively by using a training sequence of symbols for the desired transmission, or by blind adaptation in which the knowledge of the desired transmission's timing and spreading waveform is used instead of training sequence. Most adaptive multiuser detectors are based on the MMSE criterion. The early adaptive linear MMSE multiuser detection can be found in [48]. The training-sequence based adaptive MMSE linear multiuser detection was proposed in [46][49][50]. And the blind adaptive MMSE multiuser detector was developed in [51]. Differentially-coherent versions of the blind MMSE multiuser detector were developed in [52] for frequency-flat fading channels and in [53][54] for multipath channels. The surveys of adaptive multiuser detection techniques can be found in [55]. The design of MMSE receiver faces the tradeoff between robustness and excess mean square error, i.e., blind algorithms tend towards robustness while training sequence based algorithms can offer good excess mean-squared error. The blind MMSE receivers in [22][51] are robust to deep fading at the cost of higher excess mean square error. On the other hand, training sequence based MMSE receivers [46][49] suffer from the problem of robustness to deep fading. An improved correlation matrix estimation scheme [56] was proposed for blind adaptive MMSE receivers which achieved performance comparable to the training sequence based adaptive MMSE receivers for flat fading case.

The training sequence based adaptive multiuser detectors can be implemented by using least mean squares (LMS) or recursive least squares (RLS) algorithm. The blind adaptive detectors can be implemented by using LMS, RLS or the subspace tracking algorithms.

2.3.3.1 Least mean squares and recursive least squares

LMS and RLS are the estimation algorithms based on the least-squares criterion, i.e., the estimate vector is obtained as solution to the square minimization problem [57].

The LMS algorithm requires a few hundred information bits to converge to a steady state and the convergence time grows exponentially with the number of users even though it has $O(P)$ complexity, where P is the spreading gain.

Recursive least squares adaptive algorithms are well known for its invariably fast convergence at the expense of a lower robustness and a higher complexity [49] [58]. This algorithm relies on the transmission of known training symbols, but without requiring any prior timing acquisition. Many extension versions of RLS algorithm were presented to solve the two problems, such as fast RLS and QRD-RLS etc [59].

2.3.3.2 Subspace approach

It has been shown that the blind technique based on LMS or RLS suffers from a saturation effect in the steady state, which causes a significant gap between its steady state performance and the performance of the true linear MMSE detector [51].

In a subspace-based blind approach, the vector space of the receiver vector is decomposed into a single space (consisting of the subspace spanned by all the CDMA signals) and a noise subspace (which is orthogonal complement to the signal subspace). Since the desired signal is part of the signal subspace, it must be orthogonal to the noise subspace. Hence, the estimator is taken to be the value for which the code sequence is nearest to be orthogonal to the noise subspace. Based on this principle, several methods for the identification and tracking of the directions of the received data have been used for multiuser detection [60] [61] [62] [63] [64] [65]. It is seen in [29] [66] that the subspace-based detector outperforms the blind LMS or RLS detector in the steady state. However, the subspace-based algorithm is more complex due to the computational complexity of the eigen-decomposition of the sample autocorrelation matrix. Also the subspace-based approach has the disadvantage that it needs to know the number of users and that it will not function if $K \geq P/2$, where K is the number of users and P is the processing gain. This problem could be overcome in practice by identifying the most dominant users and including them in the signal subspace and lumping the remaining users within the noise. The subspace-based algorithm is amenable to an adaptive decentralized implementation, i.e., it can be adopted when only one user's signal is to be estimated based on the knowledge of the corresponding spreading code only.

2.4 Decision-driven detection

A number of literatures in multiuser detection have proposed nonlinear detectors that use decisions on the bits of interfering users in the demodulation of the bit of the interest. These are typically decision-driven in that they make use of bit decisions on interfering users' bits to subtract off the received signal in a repetitive manner much like the concept of successive decoding. It is hoped that the correct bit decisions will lead to gradually fewer interfering users and ultimately better performance. Such schemes work best in good SNR environment and when users are well separated in power levels. The family of the decision-driven detection includes the successive cancellation detector, multistage detector and decision-feedback detector. Successive cancellation has been studied in [67] [68] [69] [70] [71] for the multiuser detection of CDMA channels. The multistage detectors were proposed in [72] for synchronous channels and in [73] [74] for asynchronous channels. The decorrealting decision-feedback detectors for synchronous channels were presented in [75]. Various forms of asynchronous decision-feedback multiuser detectors have been proposed and analyzed in [45] [50] [76] [77] [78].

2.5 Bayesian multiuser detection

Bayesian Monte Carlo (MC) methodologies have recently emerged as low cost signal processing techniques with performance approaching to the theoretical optimum for wireless communication systems [79]. As a graphical modeling tool, a Bayesian framework can intuitively capture the relationship among the contributing factors in a complex system. When applied to digital wireless communication systems, a Bayesian detector can naturally exploit the structure of the coded signals. Bayesian detection is based on the Bayesian inference of all unknown quantities [80], all the recent Bayesian detectors use stochastic Monte Carlo sampling methods for Bayesian inference. The family of stochastic Monte carlo (MC) sampling algorithms is a well-developed and widely-used subclass of approximate inference algorithms [81]. These methods have an important advantage that is not subject to any linearity or Gaussianity constraints on the model, and also possess appealing convergence properties. Monte Carlo methods are relative slow, but provide a much better accuracy of the estimation results. Therefore, the convergence speed of the MC sampling process has become a major issue to be improved.

Most MC techniques fall into one of the two categories - Markov chain Monte Carlo (MCMC) methods for batch signal processing and sequential Monte Carlo (SMC) methods for adaptive signal processing.

The MCMC methods have been well developed and widely used to deal with optimal signal processing problems encountered in wireless communications. Some simulation results were reported in [82] [83] [84] [85] to demonstrate the desirable performance. The application in adaptive multiuser detection for synchronous CDMA with Gaussian and impulsive noise was presented in [82]. A Bayesian MCMC receiver was proposed in [83] for the space-time coded multi-carrier CDMA systems. Blind Bayesian MCMC multiuser receiver was presented in [84] for long code multipath CDMA system. All these receivers used the popular Monte carlo sampling method - Gibbs sampler. Another efficient Monte carlo sampling method, adaptive sampling algorithm, was recently developed in [85].

The SMC methods in the other category have also begun to show a great potential for solutions to a wide range of statistical inference problems [86]. Comparing to the MCMC methods, the SMC methods provide a better performance achieved by parallel processing and are well-suited to practical applications. By iteratively generating Monte Carlo samples of the state variables or other latent variables, the posterior distribution of any system parameter of interest can be approximated. The complete theoretical framework using SMC methods is described in [87]. The SMC methods have been successfully applied to solve a few problems in communications, such as blind equalization and detection in fading channels [79][88][89]. The solution to the detection problem for general MIMO systems was presented in [90]. For applications to CDMA system, the SMC-KF reported in [91] combines the conventional Kalman filter and importance sampling technique to approximate the multi-access interference (MAI) as circular Gaussian for the problem of single-user detection. The particle filtering methods, one category of SMC, were developed for the multiuser framework of CDMA system [92] [93]. However, the required computational complexity of all these methods grows exponentially with the number of users. The work reported in [94] [95] developed the methods based on the particle filtering with non-exponential complexity but also with the obvious performance limitation.

2.6 Summary

Multiuser detection for CDMA systems is one of the most active research areas of digital communications in recent years. All multiuser detection algorithms aim to mitigate multiple access interference (MAI) and enhance channel capacity. The conventional deterministic methods always face the tradeoffs between the requirements of the prior knowledge, performance and the computational complexity. Most of these methods were developed based on the short spreading codes, so that they have the limitations for the systems with long spreading codes. So far the techniques of multiuser detection are still far from maturity. In this thesis, the emphasis of the research about multiuser detection is on the statistical algorithms based on the Bayesian Monte Carlo theories.

Chapter 3

Blind channel estimation for long code DS-CDMA

3.1 Introduction

Both short and long spreading codes are used for the wireless networks. The major difference between long codes and short codes is that the short spreading code is periodic for every data bit and the long spreading code is aperiodic. Though short codes are an option, however, the majority of third generation CDMA-based wireless networks will be employing long (aperiodic) spreading sequences [7] [96]. The rationale for such a choice lies in the fact that, the use of long codes guarantees that all of the active users achieve the same performance on average, thus avoids the unfair situation that there exit some preferred users. With a period much larger than a data bit, however, the long sequences appear essentially random and destroy the bit-interval cyclostationarity properties of the CDMA signals. The statistics of the multiple-access interference (MAI) change randomly from bit to bit and the performance is determined by the average interference level [8]. Most of the previous algorithms are based on the crucial assumption that the DS-CDMA is adopting the short codes. Many reported procedures of the parameter estimation cannot be applied to systems with aperiodic long codes. For example, the MMSE receiver relies on the cyclostationarity of the interference statistics and requires short spreading sequence, which is the same as the subspace based algorithms. As a consequence, the design of intelligent signal processing techniques for both channel estimation and multiuser detection in DS-CDMA systems with aperiodic spreading codes poses new challenges in this research area.

A few algorithms were reported for long-code CDMA systems, the results in this area can be found in [97] [98] [99] [100]. Based on subspace algorithm in [97], both blind and pilot-assisted procedures were proposed for channel estimation in a synchronous CDMA. Blind channel estimation procedures based on array observations were reported in [98]. The correlation-matching techniques were employed to estimate multipath parameters blindly in [99]. A Toeplitz displacement method for multipath channel estimation was developed in [100]. The channel acquisition problem in a single-rate reverse link cellular scenario was considered in [101]. All these algorithms have some problems such as the high computational complexity, near-far sensitive and the need of assumptions on some known parameters. The least-squares criterion based algorithms for long-code DS-CDMA systems were proposed in [102] and [103]. Estimation procedures over a frequency-selective fading channel were proposed in [102] for both single-rate and multi-rate systems. The procedures rely on the transmission of training symbols, without requiring any prior timing acquisition, and can be recursively implemented with a computational complexity that is quadratic with the processing gain. However, the algorithms are near-far sensitive and based on the assumptions that the known training symbols are available. The procedures in [103] considered both multiuser estimation and single user estimation, and treated the problem of joint estimation of nearly all the signal parameters and channel noise variance simultaneously. Both algorithms are based on the crucial fact that, even if aperiodic codes are employed, the received effective signature waveform can be written as the product of a known time-varying code matrix times an unknown time-invariant vector, which contains the needed information on the estimated parameters.

From the results presented thereinbefore, it can be seen that the blind channel estimation algorithms that simultaneously are near-far resistant and have quadratic complexity are still far from being solved for long code DS-CDMA systems.

In this chapter, a new blind channel estimation method is developed by combining the advantages from both Toeplitz displacement and correlation matching techniques. The conventional correlation matching estimation method was developed in [99] to explore the output covariance matrix to match the approximations based on the received data. Compared to the subspace-based approach, the correlation matching estimation offers a better performance for loaded systems with only some mild assumptions. The basic idea of the proposed method is to remove the effects of the channel noise and other users' interferences by applying the Toeplitz displacement operation before the estimation of channel parameters is performed with the correlation matching method. Simulation results are presented to compare the performances of the proposed method with those of the conventional correlation matching and the subspace based Toeplitz estimation. The comparison shows that the proposed method offers better MSE performance and more robust near-far resistance.

3.2 Signal model

3.2.1 Flat fading channels

Consider a coherent DS-CDMA system over a flat fading channel with K active users. The representation of the received signal after coherent reception is given by

$$x(t) = \sum_{n=-\infty}^{\infty} \sum_{k=1}^{K} A_k c_k^n (t - nT - \tau_k) b_k(n) g_k + w(t) \tag{3.1}$$

where $w(t)$ is the additive and circularly symmetric Gaussian noise process with variance σ_w^2, A_k and b_k are, respectively, the amplitude of the signal and the transmitted bit for user k. The amplitude of each user's signal is modeled as a fixed, but unknown quantity. The delay τ_k for user k is assumed to be the integral multiples of a chip duration, and g_k is the fading coefficient of the kth user's channel. It is assumed that the fading coefficients remain constant over the entire data collection block. For randomized long code DS-CDMA, the spreading waveform $c_k^n(t)$, for user k, is formed by

$$c_k^n(t) = \sum_{p=0}^{P-1} c_k^n(p) \phi(t - pT_c) \tag{3.2}$$

where $\phi(t)$ is the shape of the chip with a duration T_c. In our case, the rectangular pulse is assumed for simplicity. The spreading waveform $c_k^n(p)$ for user k changes from symbol to symbol and takes values of $(\pm 1/\sqrt{P})$ with equal probability, where P is the spreading gain or the number of chips per symbol, i.e., the symbol duration $T = PT_c$.

The received signal is sampled at the chip rate and chip-matched by a filtering process. An observation vector $\mathbf{x}(n)$ is formed by concatenating aP samples, where a represents the number of symbols contained in the observation vector.

For the convenience, let us consider the expression of the observation vector for a synchronized system ($\tau_k = 0, \ \forall\ k$). The extension to asynchronous interfaces is straight forward once the relevant matrices have been defined. The observation vector of aP samples at the chip rate is given by

$$\mathbf{x}(n) = \sum_{k=1}^{K} A_k \mathbf{C}_k(n)\mathbf{H}_k\mathbf{b}_k(n) + \mathbf{w}(n) \tag{3.3}$$

where $\mathbf{x}(n) = [x(n), \cdots, x(n+aP-1)]^T$ and $\mathbf{w}(n) = [w(n), \cdots, w(n+aP-1)]^T$ are vectors of the received samples and noise samples of size $aP \times 1$, and $\mathbf{b}(n) = [b(\lfloor n/P \rfloor - 1), \cdots, b(\lfloor n/P \rfloor + a)]^T$ is a $(a+2) \times 1$ vector of data bits. The operator $\lfloor \cdot \rfloor$ returns the largest integer smaller than its argument. The channel matrix $\mathbf{H}_k$ for user k is given by

$$\mathbf{H}_k = g_k \mathbf{I}_{a+2}, \tag{3.4}$$

and $\mathbf{C}_k(n)$ is the spreading code matrix for user k with dimension $aP \times (a+2)$, which is denoted as

$$\mathbf{C}_k(n) = \begin{bmatrix} \mathbf{0} & \mathbf{c}_k(n+P) & & & \mathbf{0} \\ \mathbf{0} & & \mathbf{c}_k(n+2P) & \mathbf{0} & \mathbf{0} \\ \vdots & \mathbf{0} & \ddots & & \vdots \\ \mathbf{0} & & & \mathbf{c}_k(n+aP) & \mathbf{0} \end{bmatrix}.$$

where $\mathbf{c}_k(n)$ is defined as $\mathbf{c}_k(n) = [c_k(n), c_k(n+1), \ldots, c_k(n+P-1)]^T$.

Matched filters are used to fully exploit the properties of the received signals. Without loss of generality, user 1 is assumed to be the desired user. The $a \times 1$ observation vector $\mathbf{y}(n)$ is given by

$$\begin{aligned} \mathbf{y}(n) &= \mathbf{S}_1(n)\mathbf{x}(n) \\ &= \mathbf{S}_1(n)\left(\sum_{k=1}^{K} A_k \mathbf{C}_k(n)\mathbf{H}_k\mathbf{b}_k(n)\right) + \mathbf{S}_1(n)\mathbf{w}(n) \end{aligned} \tag{3.5}$$

where the matched filtering matrix $\mathbf{S}_1(n)$ is given by

$$\mathbf{S}_1^T(n) = \begin{bmatrix} \mathbf{c}_1(n+P) & & & \\ & \mathbf{c}_1(n+2P) & \mathbf{0} & \\ \mathbf{0} & & \ddots & \\ & & & \mathbf{c}_1(n+aP) \end{bmatrix}. \tag{3.6}$$

That is, $\mathbf{S}_1^T(n)$ is formed by truncating the first and the last column from $\mathbf{C}_1(n)$. Since the long spreading sequence changes from symbol to symbol, therefore, the parameters of the matched filter should be update from symbol to symbol.

3.2.2 Frequency-selective fading channels

Now we consider that the same DS-CDMA signals transmit over a frequency-selective fading (multipath) channel. Then the signal after coherent reception is given by

$$x(t) = \sum_{n=-\infty}^{\infty} \sum_{k=1}^{K} A_k \tilde{c}_k^n(t - nT - \tau_k) b_k(n) + w(t). \tag{3.7}$$

The assumptions about transmitted data, noise, and spreading waveform are the same as above. The spreading effective waveform $\tilde{c}_k^n(t)$ in (3.7) is constructed by the convolution of the original spreading waveform with the channel response, that is, $\tilde{c}_k^n(t) = c_k^n(t) * h_k(t)$, where $h_k(t)$ is the channel impulse response for user k. The channel length for each user is the same as L ($L < P$) chips, and the multipath delay spread is less than a symbol interval. It is also assumed that the fading coefficients remain constant over the entire data collection block.

The received signal is sampled at the chip rate and chip-matched by a filtering process. An observation vector $\mathbf{x}(n)$ is formed by concatenating $aP + L - 1$ samples, where a represents the number of symbols contained in the observation vector. The filtered and sampled complex channel impulse response is denoted by $\mathbf{h}_k = [h_k(0), \cdots, h_k(L-1)]^T$.

An observation vector is formed to contain a symbols and $L - 1$ bits which belong to a fraction of a symbol. Then the observation vector of $aP + L - 1$ samples at the chip rate is given by

$$\mathbf{x}(n) = \sum_{k=1}^{K} A_k \mathbf{C}_k(n) \mathbf{H}_k \mathbf{b}_k(n) + \mathbf{w}(n) \tag{3.8}$$

where $\mathbf{x}(n) = [x(n), \cdots, x(n + aP + L - 2)]^T$ and $\mathbf{w}(n) = [w(n), \cdots, w(n +$

$aP + L - 2)]^T$ are vectors of the received samples and noise samples of size $(aP+L-1)\times 1$, and $\mathbf{b}(n) = [b(\lfloor n/P \rfloor - 1), \cdots, b(\lfloor n/P \rfloor + a)]^T$ is a $(a+2)\times 1$ vector of data bits. The channel matrix $\mathbf{H}_k$ for user k is given by

$$\mathbf{H}_k = \mathbf{I}_{a+2} \otimes \mathbf{h}_k, \tag{3.9}$$

and the $(aP + L - 1) \times (a + 2)L$ matrix $\mathbf{C}_k(n)$ is the spreading code matrix for user k.

To derive the spreading code matrix, an $(P+L-1)\times L$ matrix $\mathbf{C}(\mathbf{c}_k(n), M)$ is defined as

$$\mathbf{C}(\mathbf{c}_k(n), L) = \begin{bmatrix} c_k(n) & 0 & \cdots & 0 \\ c_k(n+1) & c_k(n) & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ c_k(n+L-1) & \cdots & \cdots & c_k(n) \\ \vdots & & & \vdots \\ c_k(n+P-1) & \cdots & \cdots & c_k(n+P-L) \\ 0 & c_k(n+P-1) & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & c_k(n+P-1) \end{bmatrix}.$$

If $\mathbf{C}^1_{k,L}(n)$ is defined to be the first P rows of $\mathbf{C}(\mathbf{c}_k(n), L)$ and $\mathbf{C}^2_{k,L}(n)$ to be the last $L-1$ rows of $\mathbf{C}(\mathbf{c}_k(n), L)$, the spreading code matrix for user k, $\mathbf{C}_k(n)$ is given by

$$\begin{bmatrix} \mathbf{C}_{k,L}^{2}(n) & \mathbf{C}_{k,L}^{1}(n+P) & & \mathbf{0} & \\ & \mathbf{C}_{k,L}^{2}(n+P) & \mathbf{C}_{k,L}^{1}(n+2P) & & \\ \mathbf{0} & & \ddots & \ddots & \\ & & & \mathbf{C}_{k,L}^{2}(n+aP) & \tilde{\mathbf{C}}_{k,L}^{1}(n+(a+1)P) \end{bmatrix}$$

where $\tilde{\mathbf{C}}_{k,L}^{1}(n)$ is composed of the first $L-1$ rows of $\mathbf{C}_{k,L}^{1}(n)$.

Similarly, L matched filters per received symbol are used to fully exploit the properties of the received signals. The $aL \times 1$ observation vector $\mathbf{y}(n)$ is given by

$$\begin{aligned} \mathbf{y}(n) &= \mathbf{S}_1(n)\mathbf{x}(n) \\ &= \mathbf{S}_1(n)\left(\sum_{k=1}^{K} A_k \mathbf{C}_k(n)\mathbf{H}_k\mathbf{b}_k(n)\right) + \mathbf{S}_1(n)\mathbf{w}(n) \end{aligned} \tag{3.10}$$

where the matched filtering matrix $\mathbf{S}_1(n)$ is given by

$$\mathbf{S}_1^T(n) = \begin{bmatrix} \mathbf{C}_{1,L}^{1}(n+P) & & \\ \mathbf{C}_{1,L}^{2}(n+P) & \cdots & \mathbf{0} \\ & \ddots \quad \ddots & \\ \mathbf{0} & & \mathbf{C}_{1,L}^{1}(n+aP) \\ & & \mathbf{C}_{1,L}^{2}(n+aP) \end{bmatrix}. \tag{3.11}$$

The matrices $\mathbf{S}_1^T(n)$ and $\mathbf{C}_1(n)$ are related by

$$\mathbf{C}_1(n) = \begin{bmatrix} \mathbf{C}_{1,L}^2(n) & \vdots & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_1^T(n) & \mathbf{0} \\ \mathbf{0} & \vdots & \tilde{\mathbf{C}}_{1,L}^1(n+(a+1)P) \end{bmatrix}.$$

That is, $\mathbf{C}_1(n)$ is formed by augmenting $\mathbf{S}_1^T(n)$ by $2L$ appropriate columns.

It is seen from (3.5) and (3.10), the system over flat fading channel and the system over frequency-selective fading channel are described by similar equations, only with different forms of the corresponding variables.

3.3 Channel estimation

In this section, an efficient channel estimation method based on Toeplitz displacement and correlation-matching estimation is developed for both the systems over flat fading channel (3.5) and over frequency-selective fading channel (3.10).

3.3.1 Correlation-matching estimation

The conventional correlation-matching channel estimation is proposed in [99]. Compared with the subspace-based approaches, this method requires only mild identifiability assumptions and offers better performance for loaded systems. On the other hand, the correlation-matching method can avoid the disadvantage of the subspace-based method since it will not function if the number of users is larger than the spreading gain.

The basic idea is to match the output covariance matrix (parameterized by the unknown channel vectors) with the instantaneous approximations based on the received data. This method is briefly described as follows.

Let us consider the matched filter output vectors $\mathbf{y}(n)$ given in (3.5) and (3.10). For convenience, we present the unified expressions for both models in the following derivation.

The covariance matrix of this observation vector is obtained as

$$\begin{aligned}\mathbf{R}_y(n) &= \mathbb{E}\{\mathbf{y}(n)\mathbf{y}^H(n)\} \\ &= \sigma_1^2\mathbf{S}_1(n)\mathbf{C}_1(n)\mathbf{H}_1\mathbf{H}_1^H\mathbf{C}_1^H(n)\mathbf{S}_1^T(n) + \mathbf{R}_I(n) + \mathbf{R}_w(n) \quad (3.12)\end{aligned}$$

where $\sigma_1^2 = A_1^2\mathbb{E}\{b_1^2(n)\}$ and $\mathbf{R}_w(n) = \sigma_w^2\mathbf{S}_1(n)\mathbf{S}_1^T(n)$ is noise autocorrelation matrix. The contribution of other users' interferences is

$$\mathbf{R}_I(n) = \sum_{k=2}^{K}\sigma_k^2\mathbf{S}_1(n)\mathbf{C}_k(n)\mathbf{H}_k\mathbf{H}_k^H\mathbf{C}_k^H(n)\mathbf{S}_1^T(n). \quad (3.13)$$

Let $\hat{\mathbf{R}}_y(n)$ denote some estimator of $\mathbf{R}_y(n)$, and

$$\mathbf{E}(n) = \mathbf{R}_y(n) - \hat{\mathbf{R}}_y(n) \quad (3.14)$$

be the estimation error matrix. The cost function is defined as:

$$\mathbf{J} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{J}(n) = \frac{1}{N}\sum_{n=1}^{N}\|\mathbf{E}(n)\|_F^2 \quad (3.15)$$

where N is the number of transmitted symbols. By minimizing this cost function, all channel parameters can be obtained. It is the general framework of the correlation-matching technique.

3.3.2 Toeplitz displacement

Now we use the Toeplitz displacement to make improvements on the conventional correlation-matching channel estimation.

Toeplitz matrix is a special kind of matrix where each descending diagonal elements from left to right are constant, e.g., $m \times n$ matrix $\mathbf{F}$

$$\mathbf{F} = \begin{bmatrix} f_0 & f_{-1} & \cdots & \cdots & f_{-(n-1)} \\ f_1 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & f_{-1} \\ f_{m-1} & \cdots & \cdots & f_1 & f_0 \end{bmatrix}.$$

As seen from above Toeplitz matrix, the submatrix $\mathbf{F}(1 : m-1, 1 : n-1)$ is equal to the submatrix $\mathbf{F}(2 : m, 2 : n)$. Here the matrix notation $\mathbf{F}(i : j, k : l)$ denotes the sub-matrix formed by taking rows from i to j and columns from k to l of matrix $\mathbf{F}$. Thus, if we make such a displacement operation to a Toeplitz matrix, the obtained result is the zero matrix, i.e.

$$\mathbf{F}(2 : m, 2 : n) - \mathbf{F}(1 : m-1, 1 : n-1) = \mathbf{0}_{(m-1)\times(n-1)} \tag{3.16}$$

Here, we denote this displacement operation as Toeplitz displacement.

3.3.3 Channel estimation algorithm

Before operating the autocorrelation of the observation vector as described in (3.14), the Toeplitz displacement is to be applied to remove the effects of the channel noise and other users' interferences from the observation vector.

Let us define

$$\overline{\mathbf{SC}}_1 = \frac{1}{N}\sum_{n=1}^{N}\mathbf{S}_1(n)\mathbf{C}_1(n) \tag{3.17}$$

and

$$\overline{\mathbf{SC}}_k = \frac{1}{N}\sum_{n=1}^{N}\mathbf{S}_1(n)\mathbf{C}_k(n), k = 2, \cdots, K. \tag{3.18}$$

It is noted that the asymptotic approximation below follows from key assumptions made about the randomized spreading codes. That is, the components of the code sequence are independently and identically distributed, and are stationary at the chip rate. Therefore we have $\mathbf{S}_1(n)\mathbf{C}_1(n) = \mathbf{SC}_1 + \mathbf{A}(n)$, and $\mathbf{S}_1(n)\mathbf{C}_k(n) = \mathbf{SC}_k + \mathbf{B}_k(n), k = 2, \cdots, K$, where $\mathbf{A}(n)$ and $\mathbf{B}_k(n)$ are, respectively, time varying perturbation matrices, $\mathbf{SC}_1 = lim_{N\to\infty}\overline{\mathbf{SC}}_1 = \begin{bmatrix} \mathbf{0} & \mathbf{I}_{aM} & \mathbf{0} \end{bmatrix}$ and $\mathbf{SC}_k = lim_{N\to\infty}\overline{\mathbf{SC}}_k = \begin{bmatrix} \mathbf{0} & \mathbf{0}_{aM} & \mathbf{0} \end{bmatrix}$. As P increases, the perturbations $\mathbf{A}(n)$ and $\mathbf{B}_k(n)$ decrease. When $P \to \infty$, the effects of the perturbations can be negligible and the effects of the imperfect spreading autocorrelation are captured in $\overline{\mathbf{SC}}_1$. Hence,

$$\begin{aligned}
\mathbf{R}_y(n) &= \sigma_1^2 \mathbf{S}_1(n)\mathbf{C}_1(n)\mathbf{H}_1\mathbf{H}_1^H\mathbf{C}_1^H(n)\mathbf{S}_1^T(n) + \mathbf{R}_I(n) + \mathbf{R}_w(n) \\
&\approx |_{P\to\infty} \sigma_1^2 \mathbf{SC}_1\mathbf{H}_1\mathbf{H}_1^H\mathbf{SC}_1^H + \sum_{k=2}^{K} \sigma_k^2 \mathbf{SC}_k\mathbf{H}_k\mathbf{H}_k^H\mathbf{SC}_k^H + \sigma_w^2\mathbf{I} \\
&\approx |_{N\to\infty} \sigma_1^2 \overline{\mathbf{SC}}_1\mathbf{H}_1\mathbf{H}_1^H\overline{\mathbf{SC}}_1^H + \sum_{k=2}^{K} \sigma_k^2 \overline{\mathbf{SC}}_k\mathbf{H}_k\mathbf{H}_k^H\overline{\mathbf{SC}}_k^H + \sigma_w^2\mathbf{I}. \quad (3.19)
\end{aligned}$$

The last two terms in above expression are Toeplitz matrices, therefore, if the Toeplitz displacement is performed in the correlation matrix of the observation vector $\mathbf{R}_y(n)$, we can obtain

$$\begin{aligned}
\mathbf{R}_h(n) &= \mathbf{R}_y(n)(2:aM, 2:aM) - \mathbf{R}_y(n)(1:aM-1, 1:aM-1) \\
&= \mathbf{R}_y^+(n) - \mathbf{R}_y^-(n) \\
&= \sigma_1^2 \mathbf{SC}_1^+\mathbf{H}_1\mathbf{H}_1^H\mathbf{SC}_1^{+H} - \sigma_1^2 \mathbf{SC}_1^-\mathbf{H}_1\mathbf{H}_1^H\mathbf{SC}_1^{-H} \quad (3.20)
\end{aligned}$$

where $\mathbf{SC}_1^+$ and $\mathbf{SC}_1^-$ are formed by removing the first row and the last row of $\mathbf{SC}_1$, respectively. The updated observation vector in (3.20) contains only the information of the desired user without the information of other interfered users and noise, thus the contributions from $\mathbf{R}_I$ and $\mathbf{R}_w$ are removed. Then the channel estimation can be performed effectively without the effects of the noise and other users' interference.

Let $\hat{\mathbf{R}}_y(n)$ denote some estimator of $\mathbf{R}_y(n)$ and $\hat{\mathbf{R}}_h(n)$ be the corresponding estimator of $\mathbf{R}_h(n)$, then

$$\begin{aligned}
\hat{\mathbf{R}}_h(n) &= \hat{\mathbf{R}}_y^+(n) - \hat{\mathbf{R}}_y^-(n) \\
&= \hat{\mathbf{R}}_y(n)(2:aM, 2:aM) - \hat{\mathbf{R}}_y(n)(1:aM-1, 1:aM-1). \quad (3.21)
\end{aligned}$$

The estimation error matrix becomes

$$\begin{aligned} \mathbf{E}_h(n) &= \mathbf{R}_h(n) - \hat{\mathbf{R}}_h(n) \\ &= \sigma_1^2 \mathbf{S}\mathbf{C}_1^+ \mathbf{H}_1 \mathbf{H}_1^H \mathbf{S}\mathbf{C}_1^{+H} - \sigma_1^2 \mathbf{S}\mathbf{C}_1^- \mathbf{H}_1 \mathbf{H}_1^H \mathbf{S}\mathbf{C}_1^{-H} - \hat{\mathbf{R}}_h(n). \end{aligned} \tag{3.22}$$

The new estimation error can be defined with the squared Frobenius norm of $\mathbf{E}_h$

$$\mathbf{J}_h(n) = \|\mathbf{E}_h(n)\|_F^2 = tr[\mathbf{E}_h(n)\mathbf{E}_h^H(n)] \tag{3.23}$$

The cost function (3.23) can be built as the cumulative error

$$\begin{aligned} \mathbf{J}_h &= \frac{1}{N}\sum_{n=1}^{N} \mathbf{J}_h(n) = \frac{1}{N}\sum_{n=1}^{N} tr[\mathbf{E}_h(n)\mathbf{E}_h^H(n)] \\ &= \frac{1}{N}\sum_{n=1}^{N} vec^H[\mathbf{E}_h(n)]vec[\mathbf{E}_h(n)]. \end{aligned} \tag{3.24}$$

The channel parameters can be obtained by minimizing this cost function. In practice, the average correlation matrix $\hat{\mathbf{R}}_y$ is sampled and formed by

$$\hat{\mathbf{R}}_y = \frac{1}{N}\sum_{n=1}^{N} \hat{\mathbf{R}}_y(n) = \frac{1}{N}\sum_{n=1}^{N} \mathbf{y}(n)\mathbf{y}^H(n). \tag{3.25}$$

The estimated $\hat{\mathbf{R}}_h$ can be formed as

$$\hat{\mathbf{R}}_h = \hat{\mathbf{R}}_y^+ - \hat{\mathbf{R}}_y^-. \tag{3.26}$$

We define new unknowns by

$$\mathbf{D}_1 = \sigma_1^2 \mathbf{H}_1 \mathbf{H}_1^H. \tag{3.27}$$

The error matrix (3.22) becomes

$$\mathbf{E}_h(n) = \mathbf{SC}_1^{+}\mathbf{D}_1\mathbf{SC}_1^{+H} - \mathbf{SC}_1^{-}\mathbf{D}_1\mathbf{SC}_1^{-H} - \hat{\mathbf{R}}_h(n) \tag{3.28}$$

and

$$vec(\mathbf{E}_h(n)) = (\mathbf{SC}_1^{+*} \otimes \mathbf{SC}_1^{+} - \mathbf{SC}_1^{-*} \otimes \mathbf{SC}_1^{-})vec(\mathbf{D}_1) - vec(\hat{\mathbf{R}}_h(n)). \tag{3.29}$$

Then let

$$\mathbf{d}_1 = vec(\mathbf{D}_1), \tag{3.30}$$

$$\mathbf{Q} = \mathbf{SC}_1^{+*} \otimes \mathbf{SC}_1^{+} - \mathbf{SC}_1^{-*} \otimes \mathbf{SC}_1^{-}. \tag{3.31}$$

We have

$$\mathbf{J}_h(n) = \{\mathbf{Q}\mathbf{d}_1 - vec(\hat{\mathbf{R}}_h(n))\}^H\{\mathbf{Q}\mathbf{d}_1 - vec(\hat{\mathbf{R}}_h(n))\}. \tag{3.32}$$

Therefore, the cost function becomes

$$\mathbf{J}_h = \frac{1}{N}\sum_{n=1}^{N}\{\mathbf{Q}\mathbf{d}_1 - vec(\hat{\mathbf{R}}_h(n))\}^H\{\mathbf{Q}\mathbf{d}_1 - vec(\hat{\mathbf{R}}_h(n))\}. \tag{3.33}$$

Thus, a quadratic cost function of new unknowns is obtained by over-parameterizing the problem by (3.27).

3.3.4 Adaptive estimation

Based on the cost function (3.33), an adaptive algorithm can be derived by considering $\mathbf{J}_h(n)$ at time n. The least mean square (LMS) recursion can be formulated for $\mathbf{d}_1$ with step size μ

$$\mathbf{d}_1^{(n+1)} = \mathbf{d}_1^{(n)} - \mu \nabla_{\mathbf{d}_1^H} \mathbf{J}_h(n) \tag{3.34}$$

where $\mu \nabla_{\mathbf{d}_1^H} \mathbf{J}_h(n)$ is a function of $\mathbf{d}_1^{(n)}$ and computed by

$$\mu \nabla_{\mathbf{d}_1^H} \mathbf{J}_h(n) = \mathbf{Q}^H \mathbf{Q} \mathbf{d}_1^{(n)} - \mathbf{Q}^H vec[\hat{\mathbf{R}}_h(n)]. \tag{3.35}$$

Here, $\mathbf{Q}$ is approximately approached by

$$\hat{\mathbf{Q}} = \hat{\mathbf{SC}}_1^{+*} \otimes \hat{\mathbf{SC}}_1^{+} - \hat{\mathbf{SC}}_1^{-*} \otimes \hat{\mathbf{SC}}_1^{-} \tag{3.36}$$

where $\hat{\mathbf{SC}}_1^{+}$ and $\hat{\mathbf{SC}}_1^{-}$ are formed by removing the first and the last rows of $\hat{\mathbf{SC}}_1$, respectively, and

$$\hat{\mathbf{SC}}_1 = \frac{1}{N} \sum_{n=1}^{N} \mathbf{S}_1(n) \mathbf{C}_1(n). \tag{3.37}$$

Based on (3.34) and (3.35), $\mathbf{d}_1$ can be updated by

$$\mathbf{d}_1^{(n+1)} = \mathbf{d}_1^{(n)} - \mu \hat{\mathbf{Q}}^H \hat{\mathbf{Q}} \mathbf{d}_1^{(n)} + \mu \hat{\mathbf{Q}}^H vec[\hat{\mathbf{R}}_h(n)], \tag{3.38}$$

and consequently, $\mathbf{D}_1^{(n+1)}$ can be reconstructed from $\mathbf{d}_1^{(n+1)}$. Once $\mathbf{D}_1$ is found by the adaptive implementation, singular value decomposition (SVD) on $\mathbf{D}_1$ can be performed to obtain its eigenvector corresponding to the unique maximum eigenvalue, which is the estimated and normalized channel vector for desired user within a phase ambiguity.

3.4 Considerations on implementation

Some issues have to be considered for practical implementation. The considerations are as follows.

- The choice of the step size μ in the adaptive algorithm described by (3.38);
- A cleaning operation is needed for the estimated difference covariance matrix to improve performance;
- The computational complexity of the proposed method.

3.4.1 Choice of the step size

The asymptotic behavior of the adaptive algorithm is considered for the step size μ. Let $\Delta\mathbf{d}_1^{(n)} = \mathbb{E}\{\mathbf{d}_1^{(n)}\} - \mathbf{d}_1$ be the bias at time n and assume the i.i.d. processes of codes, inputs, and noise. By subtracting $\mathbf{d}_1$ on both sides of (3.38) and taking expectation, we obtain

$$\begin{aligned}\Delta\mathbf{d}_1^{(n+1)} &= [\mathbf{I} - \mu\mathbb{E}\{\hat{\mathbf{Q}}^H\hat{\mathbf{Q}}\}]\Delta\mathbf{d}_1^{(n)} + \mu\mathbb{E}\{\hat{\mathbf{Q}}^H vec[\hat{\mathbf{R}}_h(n)]\} - \mu\mathbb{E}\{\hat{\mathbf{Q}}^H\hat{\mathbf{Q}}\}\mathbf{d}_1 \\ &\approx [\mathbf{I} - \mu\mathbb{E}\{\hat{\mathbf{Q}}^H\hat{\mathbf{Q}}\}]\Delta\mathbf{d}_1^{(n)} = (\mathbf{I} - \mu\mathbf{U})\Delta\mathbf{d}_1^{(n)} \qquad (3.39)\end{aligned}$$

where $\mathbf{U}$ is a constant matrix characterized by the given system parameters. This equation implies that the convergence of the proposed adaptive method depends on the eigenvalue of matrix $\mathbf{I} - \mu\mathbf{U}$. The necessary condition on the step size is then $|1 - \mu\lambda_i| < 1 \; \forall \, i$, where λ_i's are the eigenvalues of $\mathbf{U}$. Equivalently, $0 < \mu < 1/\lambda_{max}$.

3.4.2 Cleaning operation

To further improve the channel estimate, we apply a cleaning operation to the sample covariance matrix. Under the assumption that the Toeplitz displacement results in a matrix free of the contribution of any interference. In practice, $\hat{\mathbf{R}}_h$ will not be block diagonal. In order to impose a block diagonal structure and remove the effects of nonideal correlation functions, we modify the construction of $\hat{\mathbf{R}}_h$

$$\hat{\mathbf{R}}_h = (\hat{\mathbf{R}}_y \odot \Psi)^+ - (\hat{\mathbf{R}}_y \odot \Psi)^- \tag{3.40}$$

where

$$\Psi(i,j) = \begin{cases} 1, & if\ \{(m-1)M+1 \le i \le mM \\ & and\ (m-1)M+1 \le j \le mM \\ & for\ \ m \in [1, \cdots, a]\} \\ 0, & else \end{cases} \tag{3.41}$$

and $\odot$ indicates the Schur product. Then, before applying the Toeplitz displacement to obtain $\hat{\mathbf{R}}_y$, we will exploit the cleaning operation by replacing the $\hat{\mathbf{R}}_h$ with $\hat{\mathbf{R}}_y \odot \Psi$

3.4.3 Computational complexity

In the method thereinbefore, we first perform the adaptive algorithm to obtain the estimation of D_j, then the SVD of D_j can be implemented to obtain its eigenvector corresponding to the normalized channel vector.

For the system over the flat fading channel, in the first step, the computational complexity of the adaptive method is on the order of $K(a-1)^2(a+2)^2$.

Once D is found by the adaptive implementation, the computational demand for the SVD operation is $K(a+2)^3$. Therefore the computation complexity of the proposed estimation method is the summation of $K(a-1)^2(a+2)^2$ multiplied by recursion number and $K(a+2)^3$ for SVD operation.

As for the system over the frequency-selective fading channel, the computational complexity of the adaptive method is on the order of $K(aL-1)^2(a+2)^2L^2$ and the computational demand for the SVD operation is $K(a+2)^3L^3$. The overall computation complexity is the summation of $K(aL-1)^2(a+2)^2L^2$ multiplied by the number of recursions and $K(a+2)^3L^3$ for SVD operation.

3.5 Simulation results

In this section, we provide some computer simulation examples to demonstrate the performance of the proposed blind channel estimation method for both the flat fading channel and the frequency-selective (multipath) channel. The comparisons with other related works are also made for the multipath channel.

During the simulations, long spreading codes of transmitted bits of all users are assumed to take values from independent equiprobable random variables +1 and -1. The data are regenerated randomly for each run of the simulation and the channel coefficients for all users are also randomly produced from independent complex Gaussian random variables.

It is noted that the estimator for channel vector has a complex scalar ambiguity. To simplify the presentation and avoid the norm ambiguity, the following

MSE is used as the performance measure for flat fading channel:

$$\mathbf{MSE} = \frac{1}{N_r}\sum_{i=1}^{N_r}\left\|\hat{g}_1^i - (g_1/\|g_1\|)\right\|^2 \tag{3.42}$$

where N_r is the number of the runs in the simulation. The true channel is denoted by g_1 and the channel estimate for run i is represented by $\hat{g}_1^i$.

Similarly, the MSE measure expression for frequency-selective fading channel is chosen as

$$\mathbf{MSE} = \frac{1}{N_r}\sum_{i=1}^{N_r}\left\|\hat{\mathbf{h}}_1^i - (\mathbf{h}_1/\|\mathbf{h}_1\|)\right\|^2 \tag{3.43}$$

where the true channel is denoted by $\mathbf{h}_1$ and the channel estimate for run i is represented by $\hat{\mathbf{h}}_1^i$, and $N_r = 50$ is selected for each simulation in our following examples.

3.5.1 MSE performance of the channel estimation

Let us first consider the MSE performance of the proposed channel estimation method by testing the adaptive method with variations of the number of users, the number of symbols, the spreading factor and the overall number of symbols in the observation vector. The results are discussed as follows.

3.5.1.1 MSE versus number of symbols

The convergence of the proposed algorithm is studied firstly. The simulations are made for $a = 2$, $SNR = 15\ dB$, and the spreading gain is chosen as $P = 35$. The MSEs are plotted as a function of the number of symbols in Fig 3.1 and Fig 3.2. Fig 3.1 shows the results for the system considering the flat fading channels, while Fig 3.2 shows the results for the system considering the frequency-selective fading channels with length $L = 5$. As illustrated in both Fig 3.1 and Fig 3.2, the performance improves by increasing the number of symbols. It is also seen from figures that convergence is achieved nearly after 160 symbols for flat-fading channel and about 200 symbols for frequency-selective fading channel.

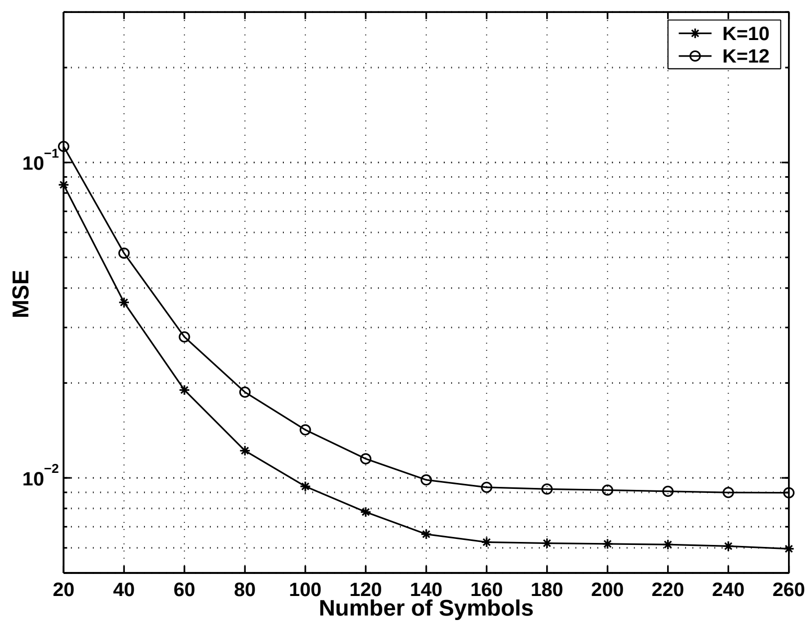


Figure 3.1: MSE versus the number of the symbols (flat fading channel)

3.5.1.2 MSE versus parameter a

Next let us consider the MSE performance as a function of a which is the number of the whole symbols used in the observation vector. The environment parameters are $SNR = 15\ dB$, spreading gain $P = 35$, and the number of the transmitted symbols is $N = 200$.

For the system over flat fading channel, it is clear that a must satisfy $a \geqslant 2$ in order to implement the Toeplitz displacement. Fig 3.3 shows the MSE values versus a which varies from 2 to 5 for the flat fading channel. As illustrated in

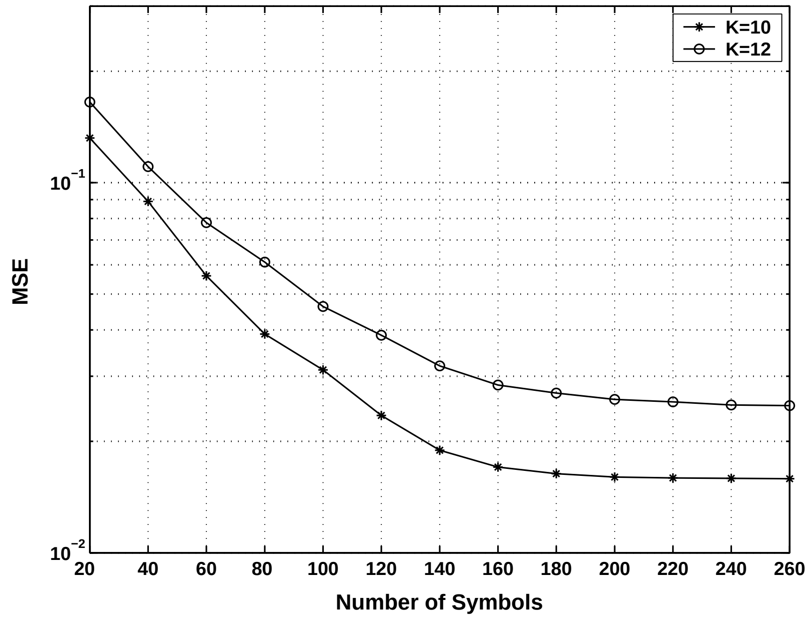


Figure 3.2: MSE versus the number of the symbols (frequency-selective fading channel)

Fig 3.3, the MSE has no obvious improvement by increasing a, that is, $a = 2$ is an adequate value for the channel estimation in such a system.

Fig 3.4 demonstrates the MSE values versus a for the frequency-selective fading channels. Two different channel lengths $L = 4$ and $L = 5$ are plotted respectively. Since the length of channel is less than the processing gain, i.e., $L < P$, the intersymbol interference to the current symbol affects less than $2P + L - 1$ bits. Therefore, when $a = 2$, the observation vector constructed by $aP + L - 1$ samples contains enough information of the current symbol for the channel identification. It can be seen that the MSE performance is not good enough when one complete symbol is used. However, significant improvements can be achieved after two complete symbols are used in the estimation. No obvious improvement is made when $a > 2$, therefore $a = 2$ is chosen as a suitable value for the following simulations.

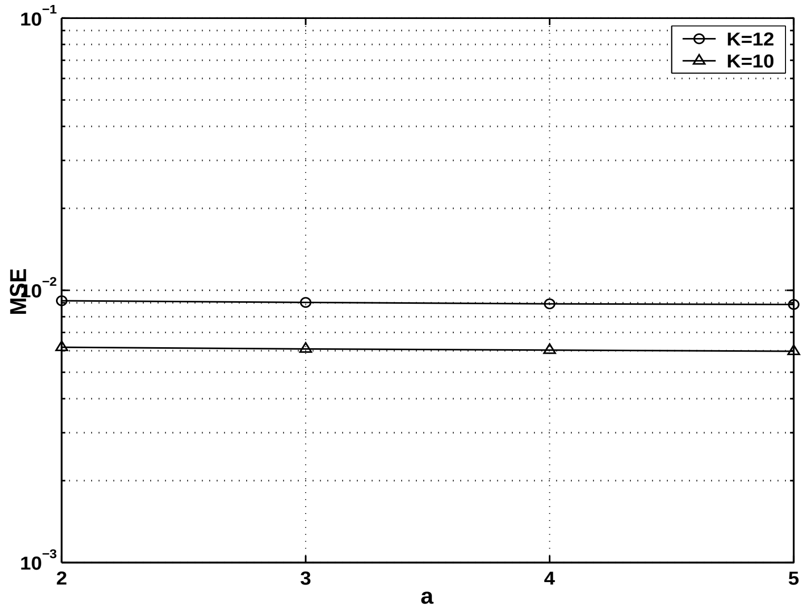


Figure 3.3: MSE versus a (flat fading channel)

3.5.1.3 MSE versus spreading gain

We now consider the effects of the spreading gain on the estimation performance. The conventional correlation matching (CM) method in [99] is developed for multipath fading channel. And as discussed in [100], the subspace-based Toeplitz displacement (TD) method cannot be used in the flat fading (single path) channel. Therefore, we plot only the results of the proposed method for the flat fading channel, whereas the results of three methods are compared for the frequency-selective fading channel.

Given $SNR = 15\ dB$, and 200 transmitted symbols for all users are used in the simulations. Fig 3.5 shows the MSE versus the spreading gain for the system over flat fading channel. Fig 3.6 shows the MSE versus the spreading gain for the system over flat fading channel with length $L = 5$, which are obtained by the proposed method, the CM method and the TD method.

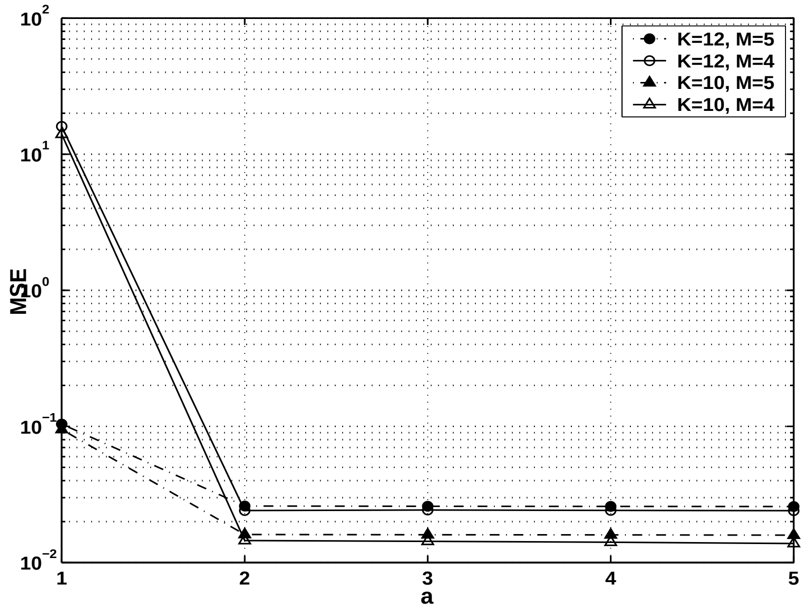


Figure 3.4: MSE versus a (frequency-selective fading channel)

Fig 3.5 and Fig 3.6 show that the MSE reduces as the spreading gain P increases. The reason is that the spreading sequences become increasingly orthogonal with each other as the increase of the spreading gain.

It is seen from Fig 3.6, the proposed method provides better performance than the subspace Toeplitz method, due to the advantages of the correlation matching estimation compared to subspace estimation. It is also observed from Fig 3.6, for small values of the spreading gain, the proposed algorithm is not superior to the conventional correlation matching method, but when the spreading gain $P > 25$, it achieves a better MSE performance compared to the CM method. This is because that the displacement is based on the approximation: spreading gain $P \to \infty$. When P is small, the approximation is not accurate enough and therefore the performance is not improved. When P is large to achieve more accurate approximation, the proposed method can suppress the effects of channel noise and interference by using the Toeplitz displacement op-

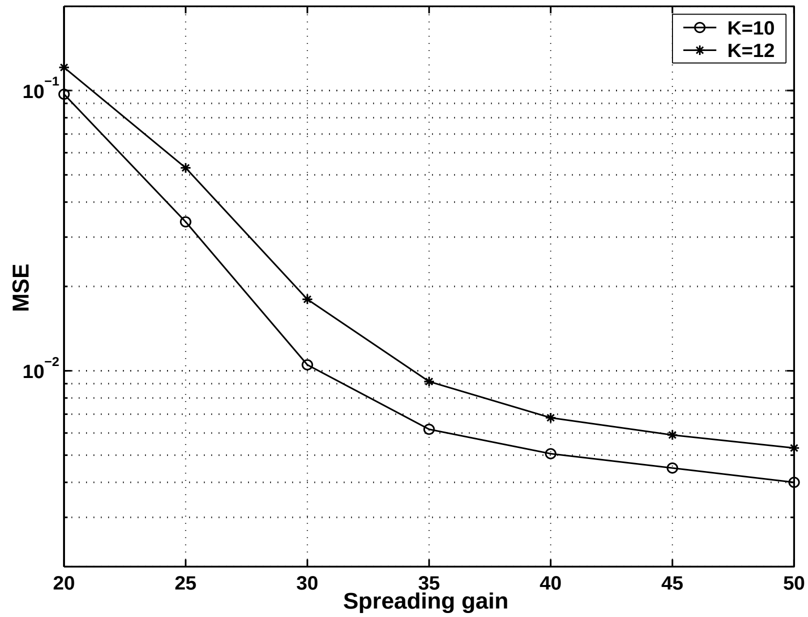


Figure 3.5: MSE versus spreading gain (flat fading channel)

eration.

3.5.2 Robustness of near-far resistance

The capability of near-far resistance is tested for the proposed method now. The simulated system is given as $P = 40$, $SNR = 15\ dB$, $a = 2$, and 200 transmitted symbols.

The near-far ratio is defined as 20log(A_1/A_k) dB, where A_1 is the received amplitude of the desired user and A_k is the received amplitude of other interfering users. Let us fix the power of the desired user and change the power of interfering users. It is assumed that all interfering users have the same power.

Similar to the discussion before, only the results of the proposed method are shown for the system over the flat fading channel. Fig 3.7 shows the MSE performance as the function of near-far ratio, which demonstrates the good

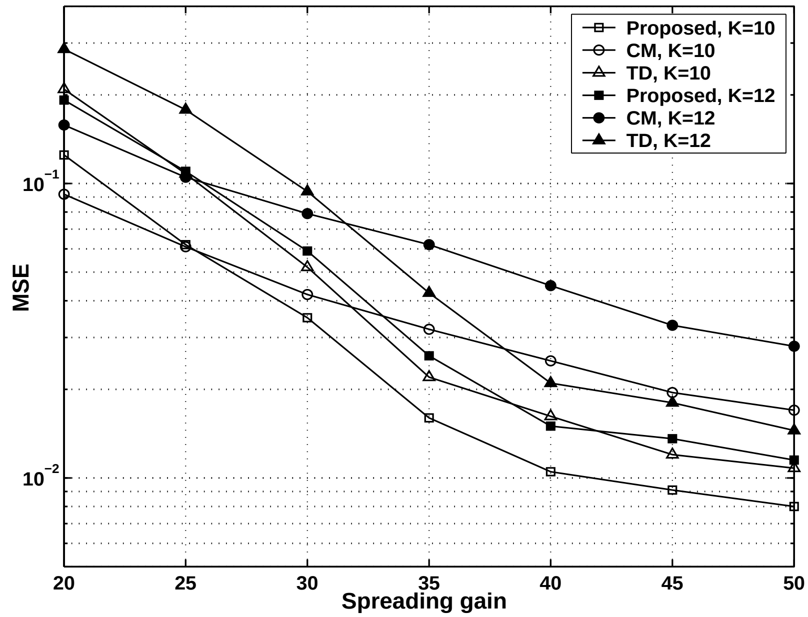


Figure 3.6: MSE versus spreading gain (frequency-selective fading channel)

capability of near-far resistance in such a system.

For the system over frequency-selective fading channels, Fig 3.8 tests the performance as the function of near-far ratio and compares the results with the CM and TD method. Fig 3.8 shows that as the near-far ratio increases, the proposed method achieves substantially better performance in suppressing the strong interference. Since the MSE changes slowly as the increase of the near-far ratio, the proposed method is very robust against near-far problem.

3.5.3 BER performance of the symbol detection

Finally, the bit-error-rate (BER) performance of symbol detection is obtained by using the estimated channel for a RAKE receiver [104].

Fig 3.9 plots the BER versus SNR for the system over flat fading channel, where $P = 35$, $a = 2$ and $K = 10, 12$. Fig 3.10 compares the BERs obtained

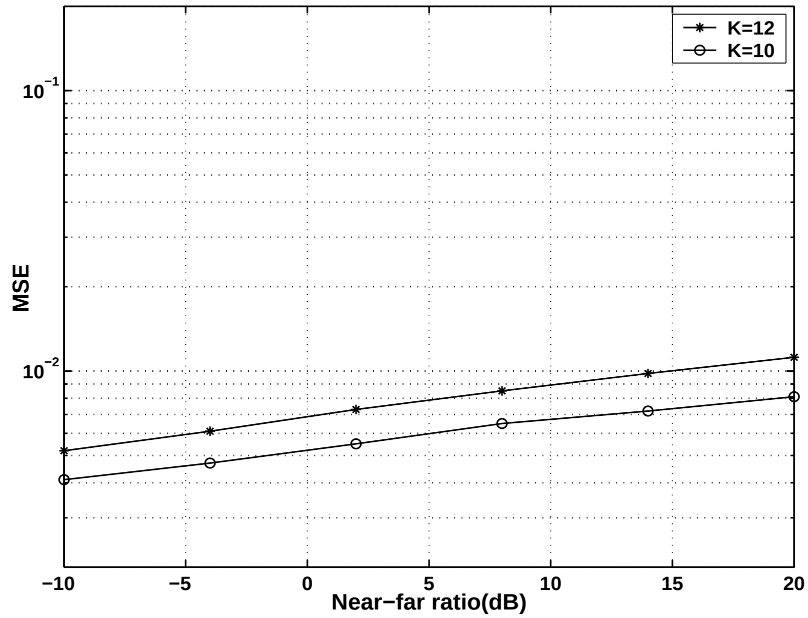


Figure 3.7: MSE evolution for a near-far environment (flat fading channel)

by using three estimation methods for the system for frequency selective fading channels, where $P = 35$, $a = 2$, $L = 5$ and $K = 10$, and the curve for perfect knowledge of the channel is also shown in the figure as a lower bound. It can be seen that the RAKE receivers exhibit poor performance. However, it is also illustrated that the proposed estimation method achieves the best BER performance among the compared methods.

3.6 Conclusions

In this chapter, we have developed an efficient blind adaptive channel estimation method for long code DS-CDMA systems. In such systems, the users' spreading sequences have the periods that are much longer than the symbol duration. Because the cross-correlation functions of the random spreading sequences in such systems vary with the time, the use of the asymptotic statistics of such spread-

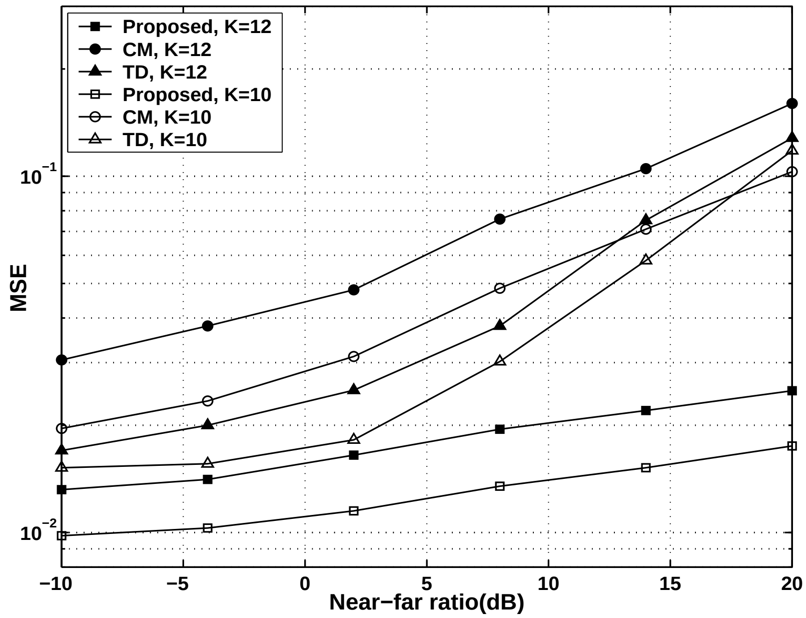


Figure 3.8: MSE evolution for a near-far environment (frequency-selective fading channel)

ing codes is made to deal with this difficult problem. By operating the Toeplitz displacement to the second order statistics of the output vector after matched filter, the effects of the channel noise and interference can be removed. Then the correlation matching method is explored to obtain the channel estimates. We have also discussed the considerations in the practical implementation. Simulation results for demonstration have been shown that the proposed technique has the promising performance and the analytical approximation is to be quite tight. In addition, the proposed algorithm is compared with related methods reported in [99] and [100] for the robustness against the near-far problem and the performance of channel estimation and resulted symbol detection. Our theory analysis and experimental results demonstrate the proposed method substantially improves performance in the interferences suppression, and the near-far resistance.

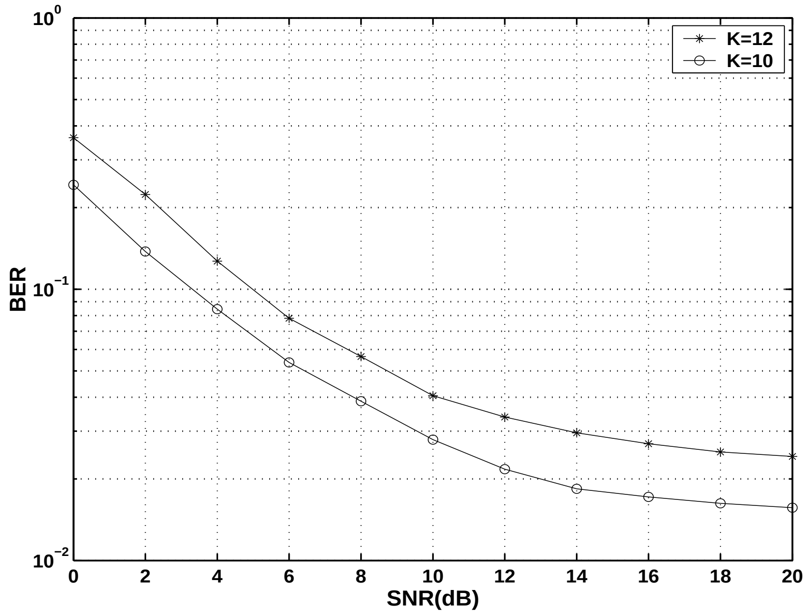


Figure 3.9: BER versus SNR (flat fading channel)

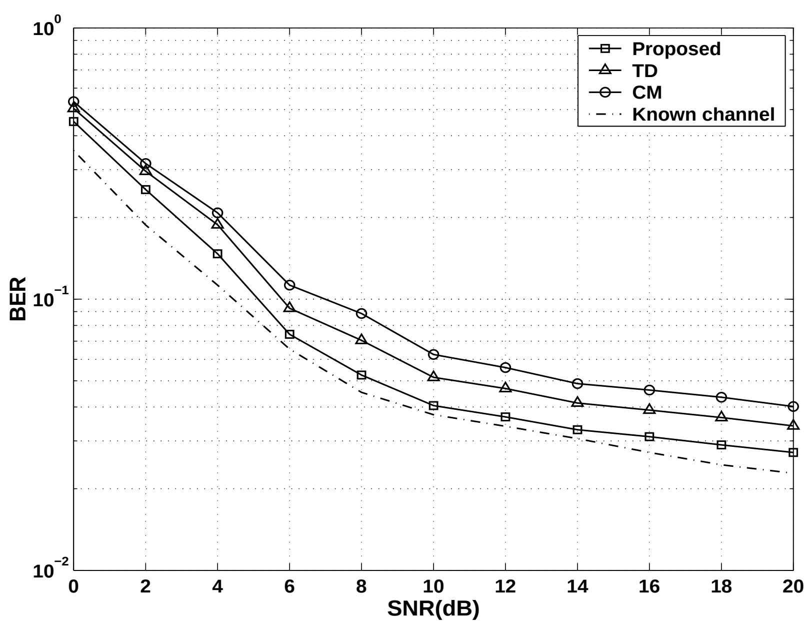


Figure 3.10: BER versus SNR (frequency-selective fading channel)

Chapter 4

Blind multiuser detection based on Bayesian MCMC inference

4.1 Introduction

Since long codes have been considered to be a main option for the next generation of CDMA based wireless networks [7], some methods for blind channel estimation were recently reported for systems based on long codes [99][100]. In the previous chapter, a new blind channel estimation scheme suitable for long code DS-CDMA systems is proposed to achieve improved performance. In most reported methods, channel parameters are estimated firstly and then the detectors are built relying on the estimated channels. However, due to the separation of channel estimation and data detection, these methods are suboptimal compared to the joint of channel estimation and data detection.

In this chapter we address the problem of the joint of channel estimation and data detection for the CDMA system employing the long spreading codes. It is assumed that the blind receiver has only the knowledge of the spreading sequences and the initial delays of the active users. The fading channels are unknown to the receiver, and no pilot symbols are employed.

Bayesian Monte Carlo (MC) methodologies have recently emerged as low cost signal processing techniques with performance approaching to the theoretical optimum for wireless communication systems [79]. Compared to the conventional techniques used for symbol detection and channel estimation applications, the MC methods are relatively slow, but provide a much better accuracy of the estimation results. Therefore, the convergence speed of the MC sampling process has become a major issue to be improved.

Bayesian MC techniques fall into one of the two categories - Markov chain Monte Carlo (MCMC) methods for batch signal processing and sequential Monte Carlo (SMC) methods for adaptive signal processing.

Most reported MCMC methods (for example, [82][83][84]) are based on a popular MCMC procedure - Gibbs sampler. The Gibbs sampler has been successfully applied for the optimal receiver design in various communication systems. However one obvious problem of the Gibbs sampler is that it needs the burn-in period to achieve the convergence and the samples generated during the burn-in period can not be used for calculating the estimation results. These characteristics lead to a relatively slow convergence speed and inefficiency in the sampling operation.

The adaptive sampling algorithm is a recently developed MC sampling method [85], and it is able to avoid searching the entire sample space so that the convergence speed can be increased by using the feedback of the available observations. Some results have demonstrated that adaptive sampling offers highly efficient Bayesian inference in the short code CDMA system without considering the fading problem. However, the channel fading problems can not be ignored for practical CDMA systems. With the consideration of the fading channels, the procedures for adaptive sampling will become complicated. It is also beneficial that the channel estimation should be jointly considered with data detection for the receiver design.

In order to take advantage of the maximum a posteriori (MAP) optimality of Bayesian inference, and at the same time avoid the burn-in period which typically encumbers the convergence rates of MCMC techniques, the application of adaptive sampling to long code fading CDMA system is considered in this chapter. An efficient blind MCMC receiver based on the adaptive sampling algorithm is proposed for the joint data detection and channel estimation. Simulation results show the desirable improvements on convergence speed and the BER performance achieved by the proposed blind Bayesian MCMC receiver.

4.2 Signal model

4.2.1 Flat fading channels

Let us consider a coherent long code system that has K active users. The signals are transmitted over the flat fading channel with unknown additive white

Gaussian noise. The received signal is expressed as

$$r(t) = \sum_{k=1}^{K}\sum_{n=0}^{N-1} A_k b_k(n) g_k \bar{c}_k^n(t - nT - d_k) + \omega(t) \tag{4.1}$$

where $b_k(n)$ is the transmitted data symbol for user k, and for each user $k = 1, \cdots, K$, the transmitted symbols $\{b_k(n)\}_{n=0}^{N-1}$ are differentially modulated from the source information symbols $\{s_k(n)\}_{n=1}^{N-1}$ with $b_k(0) = 1$. Such a different encoding scheme is necessary to resolve the phase ambiguity inherent to any blind receiver.

In (4.1), N is the length of the transmitted data frame, T is the symbol duration, A_k is the transmitted amplitude of user k and d_k denotes the delay of the kth user's signal. The spreading waveform $\bar{c}_k^n(t)$ is formed by the pulse shape $\phi(t)$ and spreading sequence $\bar{c}_k^n(p)$ for user k is defined as

$$\bar{c}_k^n(t) = \sum_{p=0}^{P-1} \bar{c}_k^n(p)\phi(t - pT_c) \tag{4.2}$$

where P is the spreading gain, i.e., the number of chips per symbol and $T_c = T/P$ is the chip duration. For long code DS-CDMA signals, the spreading waveform for every user changes from symbol to symbol. The g_k is the coefficient of fading channel for user k, and $\omega(t)$ is the received zero mean additive complex white Gaussian noise with variance σ^2. Assume the initial delay $0 < d_k < T$, hence the maximum symbol delay satisfies $\{d_k/T\}_{max} \leq 1$.

The received signal is processed by a chip-matched filter and sampled at the chip rate. Since the maximum symbol delay is not larger than one symbol, when the signal is sampled at the chip rate, at most two symbols' information can be contained in very chip duration. The signal sample at the output of the matched filter at time $t = nT + qT_c$ is expressed as

$$
\begin{aligned}
r_q(n) &= \int_{nT+qT_c}^{nT+(q+1)T_c} r(t)\phi(t-nT-qT_c)dt \\
&= \int_{nT+qT_c}^{nT+(q+1)T_c} \phi(t-nT-qT_c)\sum_{k=1}^{K}\sum_{m=0}^{N-1} A_k b_k(m) g_k \sum_{p=0}^{P-1} \bar{c}_k^m(p)\phi(t-mT-pT_c-d_k)dt + \omega_q(n) \\
&= \sum_{k=1}^{K}\sum_{m=n-1}^{n} A_k b_k(m) g_k \sum_{p=0}^{P-1} \bar{c}_k^m(p)\int_0^{T_c} \phi(t)\phi(t+nT-mT+qT_c-pT_c-d_k)dt + \omega_q(n) \\
&= \sum_{k=1}^{K}\sum_{m=0}^{1}\sum_{p=0}^{P-1} A_k b_k(n-m) g_k \bar{c}_k^{n-m}(q-p)\int_0^{T_c} \phi(t)\phi(t+mT+pT_c-d_k)dt + \omega_q(n) \\
&= \sum_{k=1}^{K}\sum_{m=0}^{1}\sum_{p=0}^{P-1} A_k b_k(n-m) g_k \bar{c}_k^{n-m}(q-p) + \omega_q(n).
\end{aligned} \tag{4.3}
$$

In the above derivation the noise sample $\omega_q(n) = \int_{nT+qT_c}^{nT+(q+1)T_c} \omega(t)\ \phi(t-nT-qT_c)dt$, and the set of noise samples $\{\omega_q(n)\}$ are i.i.d. zero mean complex Gaussian random variables with variance σ^2.

Let us define $\zeta_k = \lfloor (d_k/T_c) \rfloor - 1$ as the initial delay for user k in terms of the number of chips. Then the observation vector can be expressed as

$$
\begin{aligned}
\mathbf{r}(n) &= \sum_{k=1}^{K} A_k b_k(n)\mathbf{C}_{k,0}(n) g_k + \sum_{k=1}^{K} A_k b_k(n-1)\mathbf{C}_{k,1}(n-1) g_k + \mathbf{w}(n) \\
&= \sum_{k=1}^{K} A_k [b_k(n)\mathbf{C}_{k,0}(n) + b_k(n-1)\mathbf{C}_{k,1}(n-1)] g_k + \mathbf{w}(n)
\end{aligned} \tag{4.4}
$$

where $\mathbf{r}(n) = [r_1(n), r_2(n), \ldots, r_P(n)]^T$, $\mathbf{w}(n) = [\omega_1(n), \omega_2(n), \ldots, \omega_P(n)]^T$ for $n = 0, 1, \ldots, N-1$, and

$$
\begin{aligned}
\mathbf{C}_{k,0}(n) &= \mathbf{C}_k(n)[1:P] \\
\mathbf{C}_{k,1}(n) &= \mathbf{C}_k(n)[P+1:2P]
\end{aligned}
$$

where $\mathbf{C}_k(n)$ is defined as a $2P \times 1$ vector

$$\mathbf{C}_k(n) = [\mathbf{0}^T_{\zeta_k \times 1}, \bar{c}^n_k(0), \bar{c}^n_k(1), \cdots, \bar{c}^n_k(P-1), \mathbf{0}^T_{(P-\zeta_k)\times 1}]^T. \tag{4.5}$$

Let $\mathbf{u}[i:j]$ denote the subvector of $\mathbf{u}$ obtained by taking ith to jth elements of $\mathbf{u}$. For simplicity, the noise term $\{\mathbf{w}(n)\}$ is assumed to be an i.i.d complex white Gaussian vectors with a zero-mean and a variance σ^2, i.e., $\mathbf{w}(n)$ satisfies $\mathbf{w}(n) \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$.

4.2.2 Frequency selective fading channels

Now consider the long code system with the same signals over the frequency-selective fading channels. Assumptions on transmitted symbols, amplitudes and spreading waveforms are all the same as previous descriptions. The transmitted signal for user k is expressed as

$$x_k(t) = \sum_{n=0}^{N-1} A_k b_k(n) \bar{c}^n_k(t - nT - d_k) \tag{4.6}$$

The impulse response of multipath channel $h_k(t)$ for user k is defined as

$$h_k(t) = \sum_{l=1}^{L} \alpha_{k,l}\delta(t - \tau_{k,l}) \tag{4.7}$$

where L is the total number of resolvable paths in the channel, $\alpha_{k,l}$ is the complex path gain and $\tau_{k,l}$, $\tau_{k,1} < \tau_{k,2} < \cdots < \tau_{k,L}$, is the delay of the lth path for user k. Then the representation of the received signal is given by

$$r(t) = \sum_{k=1}^{K} x_k(t) * h_k(t) + \omega(t) \tag{4.8}$$

where $*$ denotes convolution, and $\omega(t)$ is the received zero mean additive complex white Gaussian noise with variance σ^2. Assume the maximum channel delay L is smaller than P, i.e., $\tau_{k,L} < T$. Hence, the maximum symbol delay satisfies $\{(d_k + \tau_{k,L})/T\}_{max} \leq 2$.

Similarly, the received signal is processed by a chip-matched filter and sampled at the chip rate. Since the maximum symbol delay is not larger than 2 symbols, when the signal is sampled at the chip rate, at most three symbols' information can be contained in very chip duration. The signal sample at the output of the matched filter at time $t = nT + qT_c$ is expressed as

$$\begin{aligned}
r_q(n) =& \int_{nT+qT_c}^{nT+(q+1)T_c} r(t)\phi(t - nT - qT_c)dt \\
=& \int_{nT+qT_c}^{nT+(q+1)T_c} \phi(t - nT - qT_c) \sum_{k=1}^{K} x_k(t) * h_k(t) + \omega_q(n) \\
=& \int_{nT+qT_c}^{nT+(q+1)T_c} \phi(t-nT-qT_c) \sum_{k=1}^{K} \sum_{m=0}^{N-1} A_k b_k(m) \\
& \times \sum_{l=1}^{L} \sum_{p=0}^{P-1} \alpha_{k,l} \bar{c}_k^m(p) \phi(t-mT-pT_c-d_k-\tau_{k,l})dt + \omega_q(n) \\
=& \sum_{k=1}^{K} \sum_{m=n-2}^{n} A_k b_k(m) \sum_{p=0}^{P-1} \bar{c}_k^m(p) \\
& \times \int_0^{T_c} \sum_{l=1}^{L} \alpha_{k,l} \phi(t)\phi(t+nT-mT+qT_c-pT_c-d_k-\tau_{k,l})dt + \omega_q(n) \\
=& \sum_{k=1}^{K} \sum_{m=0}^{2} \sum_{p=0}^{P-1} A_k b_k(n-m) \bar{c}_k^{n-m}(q-p) \\
& \times \int_0^{T_c} \sum_{l=1}^{L} \alpha_{k,l} \phi(t)\phi(t+mT+pT_c-d_k-\tau_{k,l})dt + \omega_q(n).
\end{aligned} \tag{4.9}$$

Denote

$$g_k(x) = \int_0^{T_c} \sum_{l=1}^{L} \alpha_{k,l} \phi(t)\phi(t + xT_c - d_k - \tau_{k,l})dt, \tag{4.10}$$

then (4.9) is equal to

$$r_q(n) = \sum_{k=1}^{K}\sum_{m=0}^{2}\sum_{p=0}^{P-1} A_k b_k(n-m)\bar{c}_k^{n-m}(q-p)g_k(mP+p) + \omega_q(n). \tag{4.11}$$

In the above derivation for (4.9) and (4.11), the noise sample $\omega_q(n) = \int_{nT+qT_c}^{nT+(q+1)T_c} \omega(t)\phi(t-nT-qT_c)dt$, and the set of noise samples $\{\omega_q(n)\}$ are i.i.d. zero mean complex Gaussian random variables with variance σ^2.

For convenience, $\zeta_k = \lfloor(d_k + \tau_{k,1}/T_c)\rfloor - 1 < P$ is defined as the initial delay for user k in terms of the number of chips. Let $\mathbf{h}_k = [g_k(\zeta_k+1), \cdots, g_k(\zeta_k+L)]$ define the channel response for the kth user, then the observation vector can be expressed as

$$\begin{aligned}\mathbf{r}(n) =& \sum_{k=1}^{K} A_k b_k(n)\mathbf{C}_{k,0}(n)\mathbf{h}_k + \sum_{k=1}^{K} A_k b_k(n-1)\mathbf{C}_{k,1}(n-1)\mathbf{h}_k \\ &+ \sum_{k=1}^{K} A_k b_k(n-2)\mathbf{C}_{k,2}(n-2)\mathbf{h}_k + \mathbf{w}(n) \\ =& \sum_{k=1}^{K} A_k[b_k(n)\mathbf{C}_{k,0}(n) + b_k(n-1)\mathbf{C}_{k,1}(n-1) + b_k(n-2)\mathbf{C}_{k,2}(n-2)]\mathbf{h}_k + \mathbf{w}(n) \end{aligned} \tag{4.12}$$

where $\mathbf{r}(n) = [r_1(n), r_2(n), \ldots, r_P(n)]^T$, $\mathbf{w}(n) = [\omega_1(n), \omega_2(n), \ldots, \omega_P(n)]^T$ for $n = 0, 1, \ldots, N-1$, and

$$\begin{aligned}\mathbf{C}_{k,0}(n) &= \mathbf{C}_k(n)[1:P,:] \\ \mathbf{C}_{k,1}(n) &= \mathbf{C}_k(n)[P+1:2P,:] \\ \mathbf{C}_{k,2}(n) &= \mathbf{C}_k(n)[2P+1:3P,:]\end{aligned} \tag{4.13}$$

where $\mathbf{C}_k(n)$ is defined as a $3P \times L$ matrix

$$\mathbf{C}_k(n) = \begin{bmatrix} \mathbf{0}_{\zeta_k \times 1} & \mathbf{0} & \cdots & \mathbf{0} \\ \bar{c}_k^n(0) & 0 & & \vdots \\ \bar{c}_k^n(1) & \bar{c}_k^n(0) & \ddots & \\ \vdots & \bar{c}_k^n(1) & \ddots & 0 \\ \bar{c}_k^n(P-1) & \vdots & & \bar{c}_k^n(0) \\ 0 & \bar{c}_k^n(P-1) & & \bar{c}_k^n(1) \\ & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \bar{c}_k^n(P-1) \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} \tag{4.14}$$

Let $\mathbf{U}[i:j,:]$ denote the submatrix of $\mathbf{U}$ obtained by appropriately taking row i to row j of $\mathbf{U}$. For simplicity, the noise term $\{\mathbf{w}(n)\}$ is assumed to be an i.i.d complex white Gaussian vectors with a zero-mean and a variance σ^2, i.e., $\mathbf{w}(n)$ satisfies $\mathbf{w}(n) \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$.

Based on the above signal models, the joint solution to blind data detection and channel estimation for long code CDMA systems is considered in the following sections.

4.3 Background

In this section, we provide a simple overview of the Bayesian inference with MCMC methods.

4.3.1 Bayesian inference framework

As a modeling tool, Bayesian framework can intuitively capture the relationship among the contributing factors in a complex system. The framework of the Bayesian MCMC inference is described as follows.

Let us assume $\theta = [\theta_1, \cdots, \theta_i, \cdots, \theta_d]$ be a vector of unknown parameters and $\mathbf{Y}$ be the observation data. In Bayesian approach, all unknown quantities are treated as random variables with some prior distribution $p(\theta)$. Then the Bayesian inference is made by computing the joint posterior distribution of the unknown parameters:

$$p(\theta|\mathbf{Y}) \propto p(\mathbf{Y}|\theta)p(\theta). \tag{4.15}$$

Now if the *a posteriori* marginal distribution of some unknown parameters is to be found from the observation data $\mathbf{Y}$, i.e., $p(\theta_i|\mathbf{Y})$, the direct evaluation is to integrate the joint *a posteriori* density with all other parameters, i.e.,

$$p(\theta_i|\mathbf{Y}) = \int\int\cdots\int p(\theta|\mathbf{Y})d\theta_1\cdots d\theta_{i-1}d\theta_{i+1}\cdots d\theta_d. \tag{4.16}$$

In most cases, the direct computation of (4.16) is not feasible, especially when parameter dimension d is large.

4.3.2 Monte Carlo methods

In many Bayesian analysis, the computation of the marginal distribution is as difficult as the above description, therefore some analytical or numerical approximations have been resorted to solve the problem. In late 1980s and early 1990s, statisticians discovered that a wide variety of Monte Carlo strategies can be applied to overcome the computational difficulties encountered in almost all likelihood-based interference procedures. Soon afterwards, Monte Carlo method is demonstrated as a powerful computational tool for Bayesian inference in many application fields. The basic idea of the MCMC method is to generate the random samples $\{\theta\}$ from the joint distribution and then approximate the marginal distribution by using these samples.

Bayesian MC techniques fall into one of the two categories - Markov chain Monte Carlo (MCMC) methods for batch signal processing and sequential Monte Carlo (SMC) methods for adaptive signal processing. In this chapter, we will give the general introduction about MCMC methods, and the SMC methods will be described in next chapter.

4.3.3 Markov chain Monte Carlo

Markov Chain Monte Carlo (MCMC) techniques [105] are well developed and especially useful for computing Bayesian solutions based on the Markov Chain theory. MCMC is a class of algorithms that allow one to draw random samples from an arbitrary target probability distribution, $p(\theta)$, known up to a normalized constant. The basic idea behind these algorithms is that one can achieve the sampling from $p(\theta)$ by running a Markov chain whose equilibrium distribution is exactly $p(\theta)$. The derivation of MCMC methods can be traced back to the well-known Metropolis algorithm which was firstly used in a statistical context in [106].

4.3.3.1 Gibbs sampler

As one of the most popular MCMC methods, the Gibbs sampler is a special case of the Metropolis algorithm. Given the initial values $\theta^{(0)} = [\theta_1^{(0)}, \ldots, \theta_d^{(0)}]^T$, the Gibbs sampler procedure iterates the following loop:

- Draw sample $\theta_1^{(n+1)}$ from $p(\theta_1|\theta_2^{(n)}, \ldots, \theta_d^{(n)}, \mathbf{Y})$.
- Draw sample $\theta_2^{(n+1)}$ from $p(\theta_2|\theta_1^{(n+1)}, \theta_3^{(n)}, \ldots, \theta_d^{(n)}, \mathbf{Y})$.

$$\vdots$$

- Draw sample $\theta_d^{(n+1)}$ from $p(\theta_d|\theta_1^{(n+1)}, \ldots, \theta_{d-1}^{(n+1)}, \mathbf{Y})$.

Then the sequence of sample vectors $\ldots, \theta^{(n-1)}, \theta^{(n)}, \theta^{(n+1)}, \ldots$ is a realization of a homogeneous Markov chain.

The convergence behavior of the Gibbs sampler is analyzed in [107][108][109]. The Gibbs sampler requires an initial transient period to converge to equilibrium. The initial period of length n_0 is known as the "burn-in" period, and the first n_0 samples should always be discarded.

4.3.3.2 Adaptive sampling algorithm

The Gibbs sampler must have a burn-in period to reach the convergence, which inevitably slows the convergence speed. The recently developed adaptive sampling method is shown to be able to avoid this problem through the feedback of the available observations [85].

The optimized adaptive sampling algorithm is developed from the *evidence weighting sampling* method. The difference lies in the adaptive sampling algorithm relates the sample allotment to the maximum *a posteriori* probabilities through an optimization formulation. The solution to this formulation gives indication where the samples should be distributed to be more effective. The allocation process biases the sample towards making the most significant contribution to MAP solution. Therefore, samples are properly weighted, do not introduce estimation bias, and do not need an initial transient "burn-in" period for convergence as in Gibbs sampler.

The Bayesian inference with adaptive sampling method is performed by the following steps.

- Draw data samples $\{\theta^{(m)}\}_{m=1}^{M}$ from the instantaneous prior sampling probabilities of unknown parameter

$$\pi(\theta) = \{\pi(i) = p(\theta_i)\} \tag{4.17}$$

where M is the total number of iterations to be performed for the probability inference. The instantaneous sampling priors $\{\pi(i)\}$ are periodically updated with every $\triangle M$ increment of the iteration index, i.e.,

$$\pi(i) \propto N_i \tag{4.18}$$

where N_i is defined as the current sample allocation among all available iterations.

- For each sample set $\{\theta^{(m)}\}$, compute the likelihood function of the observation $\mathbf{Y}$.

$$\lambda(\mathbf{Y}|\theta^{(m)}) = \{\lambda^{(m)}(i) = p(\mathbf{Y}|\theta_i^{(m)})\} \tag{4.19}$$

- The posterior probability of θ is computed by the ensemble expectation

$$P(\theta|\mathbf{Y}) = \mathbf{F}\mathbb{E}_m\{\lambda(\mathbf{Y}|\theta^{(m)})\} \tag{4.20}$$

where $\mathbf{F}$ is the normalization factor matrix for θ.

The basic idea of the adaptive sampling method is to bias the samples towards the MAP solution of parameter estimation through the adaptive sample allotment optimization.

4.4 Bayesian multiuser detection

Bayesian detector is based on the Bayesian inference of all unknown quantities to exploit the structure of the signals.

For the system over flat fading channel described in (4.4), we denote $\mathbf{R}=\{\mathbf{r}(0), \mathbf{r}(1), \cdots, \mathbf{r}(N-1)\}$, and $\mathbf{H} = \{g_k\}_{k=1}^{K}$. With the Bayesian inference, let us consider the problem of estimating the *a posteriori* symbol probabilities

$$P(b_k(n)|\mathbf{R}), \quad k = 1, \cdots, K; \quad n = 0, 1, \cdots, N-1 \tag{4.21}$$

based on the received signals $\mathbf{R}$ without knowing the channel response $\mathbf{H}$ and noise level σ^2. Then the probabilities are used to estimate the data symbols $\mathbf{B} = \{b_k(n)\}_{k=1:K}^{n=0:N-1}$.

Similarly, for the system over frequency-selective fading channels described in (4.12), our problem is to estimate the *a posteriori* probabilities of the symbols based on the receiver signals $\mathbf{R}=\{\mathbf{r}(0), \mathbf{r}(1), \cdots, \mathbf{r}(N-1)\}$, without knowing the channel response $\mathbf{H} = \{\mathbf{h}_k\}_{k=1}^{K}$ and noise level σ^2.

4.4.1 Application of adaptive sampling

We incorporate the adaptive sampling method with the above detection problem. The inference procedure needs to be updated and expanded to deal with the noise and the fading problem. Therefore, for both the flat fading channel system (4.4) and frequency-selective fading channel system (4.12), the Bayesian MC detection based on the adaptive sampling implements the procedures as:

- Draw data samples $\{\mathbf{B}^{(m)}\}_{m=1}^{M}$ from the instantaneous prior sampling probabilities of symbol

$$\pi(\mathbf{B}) = \{\pi_{k,j}(n) = P(b_k(n) = \beta_j)\}. \tag{4.22}$$

The instantaneous sampling priors $\{\pi_{k,j}(n)\}$ are periodically updated with every ΔM increment of the iteration index, i.e., $\pi_{k,j}(n) \propto N_{k,j,n}$, where $N_{k,j,n}$ is defined as the total number of instantiations in which the samples $b_k(n)$ is equal to β_j among all available M iterations.

- For each sample set $\{\mathbf{B}^{(m)}\}$, the likelihood function of the observation $\mathbf{R}$ is computed by

$$\lambda(\mathbf{R}|\mathbf{B}^{(m)}, \mathbf{H}, \sigma^2) = \{\lambda_{k,j}^{(m)}(n) = p(\mathbf{R}|b_k^{(m)}(n) = \beta_j, \mathbf{H}, \sigma^2)\} \tag{4.23}$$

- the *a posteriori* probability of $\mathbf{B}$ is obtained by the ensemble expectation

$$P(\mathbf{B}|\mathbf{R}) = \mathbf{F}\mathbb{E}_m\{\lambda(\mathbf{R}|\mathbf{B}^{(m)}, \mathbf{H}, \sigma^2)\}. \tag{4.24}$$

4.4.2 Channel estimation

As seen from (4.23), in order to compute the conditional distribution of $\mathbf{R}$, the estimation of channels $\mathbf{H}$ and noise variance σ^2 should be available. Thus, the estimation of the channel and noise need to be combined with the symbol inference procedure.

For the estimation of the channel and noise, we present the derivation for the system (4.4) and the system (4.12) respectively.

A. Flat fading channels

These estimates can be obtained from the conditional posterior distributions of channels and noise, that is $p(g_k|\mathbf{H}_k, \sigma^2, \mathbf{B}, \mathbf{R})$ and $p(\sigma^2|\mathbf{H}, \mathbf{B}, \mathbf{R})$, where $\mathbf{H}_k = \mathbf{H}\backslash g_k$ denotes excluding g_k from $\mathbf{H}$.

In principle, prior distributions are used to incorporate the prior knowledge about the unknown parameters, and less restrictive (i.e., non-informative) priors should be employed when such knowledge is limited. The property that the posterior distribution belongs to the same distribution family as the prior distribution is called conjugacy. The prior distribution which is chosen to satisfy the conjugacy is known to be conjugate prior. Conjugate priors are usually used to obtain simple analytical forms for the resulting posterior distributions, such that the conditional posterior distributions are easy to compute and simulate. Following the general guideline in Bayesian analysis [80], we choose a complex Gaussian prior distribution for the unknown channel, i.e., $p(g_k) \sim \mathcal{N}_c(g_{k0}, \Sigma_{k0})$, and an inverse chi-square prior distribution for the noise variance, i.e., $p(\sigma^2) \sim \chi^{-2}(2\nu_0, \lambda_0)$. According to the Bayesian theory, the conditional posterior distributions for the channels are derived as follows.

$$\begin{aligned} p(g_k|\mathbf{H}_k, \sigma^2, \mathbf{B}, \mathbf{R}) &\propto p(\mathbf{R}|\mathbf{H}, \sigma^2, \mathbf{B})p(g_k) \\ &\propto exp\{-\frac{1}{\sigma^2}\sum_{n=0}^{N-1}\|\mathbf{r}(n) - \sum_{k=1}^{K} A_k\mathbf{S}_{k,n}g_k\|^2\}exp\{-(g_k - g_{k0})^H\Sigma_{k0}^{-1}(g_k - g_{k0})\} \\ &\propto exp\{-(g_k - g_{k*})^H\Sigma_{k*}^{-1}(g_k - g_{k*})\} \propto \mathcal{N}_c(g_{k*}, \Sigma_{k*}) \end{aligned} \quad (4.25)$$

with

$$\Sigma_{k*}^{-1} = \Sigma_{k0}^{-1} + \frac{1}{\sigma^2}\sum_{n=0}^{N-1}\mathbf{S}_{k,n}^H\mathbf{S}_{k,n} \quad (4.26)$$

$$g_{k*} = \Sigma_{k*}\left[\Sigma_{k0}^{-1}g_{k0} + \frac{1}{\sigma^2}\sum_{n=0}^{N-1}\mathbf{S}_{k,n}^H(\mathbf{r}(n) - \sum_{j\neq k}\mathbf{S}_{j,n}g_j)\right] \quad (4.27)$$

where $\mathbf{S}_{k,n} = b_k(n)\mathbf{C}_{k,0}(n) + b_k(n-1)\mathbf{C}_{k,1}(n-1)$. The conditional posterior distributions for the noise variance is derived by

$$\begin{aligned} p(\sigma^2|\mathbf{H}, \mathbf{B}, \mathbf{R}) &\propto p(\mathbf{R}|\mathbf{H}, \sigma^2, \mathbf{B})p(\sigma^2) \\ &\propto (\frac{1}{\sigma^2})^{NP}exp\{-\frac{1}{\sigma^2}\sum_{n=0}^{N-1}\|\mathbf{r}(n) - \sum_{k=1}^{K}A_k\mathbf{S}_{k,n}g_k\|^2\}(\frac{1}{\sigma^2})^{\nu_0+1}exp(-\frac{\nu_0\lambda_0}{\sigma^2}) \\ &= (\frac{1}{\sigma^2})^{\nu_0+NP+1}exp(-\frac{\nu_0\lambda_0 + s^2}{\sigma^2}) \\ &\sim \chi^{-2}\left(2[\nu_0 + NP], \frac{\nu_0\lambda_0 + s^2}{\nu_0 + NP}\right) \end{aligned} \quad (4.28)$$

where $s^2 = \sum_{n=0}^{N-1}\|\mathbf{r}(n) - \sum_{k=1}^{K}A_k\mathbf{S}_{k,n}g_k\|^2$.

Using the conditional posterior distributions, the estimates of the channel response $\tilde{\mathbf{H}}$ and noise variance $\tilde{\sigma}^2$ can be obtained, and then the conditional probability of $\mathbf{R}$ is given as

$$p(\mathbf{R}|b_k(n) = \beta_j, \tilde{\mathbf{H}}, \tilde{\sigma}^2) \propto exp\{-\frac{1}{2\tilde{\sigma}^2}\sum_{n=0}^{N-1}\|\mathbf{r}(n) - \sum_{k=1}^{K} A_k \mathbf{S}_{k,n}\tilde{g}_k\|^2\} \tag{4.29}$$

B. Frequency-selective fading channels

Similarly, the estimates of the channel and noise level are obtained from their conditional posterior distributions, i.e., $p(\mathbf{h}_k|\mathbf{H}_k, \sigma^2, \mathbf{B}, \mathbf{R})$ and $p(\sigma^2|\mathbf{H}, \mathbf{B}, \mathbf{R})$. $\mathbf{H}_k = \mathbf{H}\backslash\mathbf{h}_k$ denotes excluding $\mathbf{h}_k$ from $\mathbf{H}$.

Again, we choose a complex Gaussian prior distribution for the unknown channel, i.e., $p(\mathbf{h}_k) \sim \mathcal{N}_c(\mathbf{h}_{k0}, \Sigma_{k0})$, and an inverse chi-square prior distribution for the noise variance, i.e., $p(\sigma^2) \sim \chi^{-2}(2\nu_0, \lambda_0)$. According to the Bayesian theory, the conditional posterior distributions for the channels are derived as follows.

$$\begin{aligned}
&p(\mathbf{h}_k|\mathbf{H}_k, \sigma^2, \mathbf{B}, \mathbf{R}) \propto p(\mathbf{R}|\mathbf{H}, \sigma^2, \mathbf{B})p(\mathbf{h}_k) \\
&\propto exp\{-\frac{1}{\sigma^2}\sum_{n=0}^{N-1}\|\mathbf{r}(n) - \sum_{k=1}^{K} A_k\mathbf{S}_{k,n}\mathbf{h}_k\|^2\}exp\{-(\mathbf{h}_k - \mathbf{h}_{k0})^H\Sigma_{k0}^{-1}(\mathbf{h}_k - \mathbf{h}_{k0})\} \\
&\propto exp\{-(\mathbf{h}_k - \mathbf{h}_{k*})^H\Sigma_{k*}^{-1}(\mathbf{h}_k - \mathbf{h}_{k*})\} \propto \mathcal{N}_c(\mathbf{h}_{k*}, \Sigma_{k*})
\end{aligned} \tag{4.30}$$

with

$$\Sigma_{k*}^{-1} = \Sigma_{k0}^{-1} + \frac{1}{\sigma^2}\sum_{n=0}^{N-1}\mathbf{S}_{k,n}^H\mathbf{S}_{k,n} \tag{4.31}$$

$$\mathbf{h}_{k*} = \Sigma_{k*}\left[\Sigma_{k0}^{-1}\mathbf{h}_{k0} + \frac{1}{\sigma^2}\sum_{n=0}^{N-1}\mathbf{S}_{k,n}^H(\mathbf{r}(n) - \sum_{j\neq k}\mathbf{S}_{j,n}\mathbf{h}_j)\right] \tag{4.32}$$

where $\mathbf{S}_{k,n} = b_k(n)\mathbf{C}_{k,0}(n) + b_k(n-1)\mathbf{C}_{k,1}(n-1) + b_k(n-2)\mathbf{C}_{k,2}(n-2)$. The conditional posterior distributions for the noise variance is derived by

$$\begin{aligned}
p(\sigma^2|\mathbf{H},\mathbf{B},\mathbf{R}) &\propto p(\mathbf{R}|\mathbf{H},\sigma^2,\mathbf{B})p(\sigma^2) \\
&\propto (\frac{1}{\sigma^2})^{NP} exp\{-\frac{1}{\sigma^2}\sum_{n=0}^{N-1}\|\mathbf{r}(n) - \sum_{k=1}^{K} A_k\mathbf{S}_{k,n}\mathbf{h}_k\|^2\}(\frac{1}{\sigma^2})^{\nu_0+1} exp(-\frac{\nu_0\lambda_0}{\sigma^2}) \\
&= (\frac{1}{\sigma^2})^{\nu_0+NP+1} exp(-\frac{\nu_0\lambda_0 + s^2}{\sigma^2}) \\
&\sim \chi^{-2}\left(2[\nu_0 + NP], \frac{\nu_0\lambda_0 + s^2}{\nu_0 + NP}\right)
\end{aligned} \tag{4.33}$$

where $s^2 = \sum_{n=0}^{N-1}\|\mathbf{r}(n) - \sum_{k=1}^{K} A_k\mathbf{S}_{k,n}\mathbf{h}_k\|^2$.

After obtaining the estimates of the channel response $\tilde{\mathbf{H}}$ and noise variance $\tilde{\sigma}^2$, then the conditional probability of $\mathbf{R}$ is given as

$$p(\mathbf{R}|b_k(n) = \beta_j, \tilde{\mathbf{H}}, \tilde{\sigma}^2) \propto exp\{-\frac{1}{2\tilde{\sigma}^2}\sum_{n=0}^{N-1}\|\mathbf{r}(n) - \sum_{k=1}^{K} A_k\mathbf{S}_{k,n}\tilde{\mathbf{h}}_k\|^2\} \tag{4.34}$$

4.4.3 Multiuser detector

With above considerations, a blind Bayesian procedure is devised for the joint process of the symbol detection and channel estimation based on the adaptive sampling algorithm. For convenience, we use the unified description in the procedure for both the flat fading channel system and the frequency-selective fading channel systems.

The complete procedure is described as follows. In the procedure, $B^{(c)}$ denotes the symbol estimates available at the current iteration, while $B^{(m)}$ means the samples drawn at the mth iteration.

FOR $m = 1 : M$,

1. For $k = 1, 2, \cdots, K$

   Draw the samples $g_k^{(m)}$ or $\mathbf{h}_k^{(m)}$ for channel response g_k or $\mathbf{h}_k$ from $p(g_k|\mathbf{H}_k^{(m-1)}, \sigma^{2(m-1)}, \mathbf{B}^{(c)}, \mathbf{R})$ or $p(\mathbf{h}_k|\mathbf{H}_k^{(m-1)}, \sigma^{2(m-1)}, \mathbf{B}^{(c)}, \mathbf{R})$ given by (4.25) or (4.30).

   End

2. Draw the sample $\sigma^{2(m)}$ for noise σ^2 from $p(\sigma^2|\mathbf{H}^{(m)}, \mathbf{B}^{(c)}, \mathbf{R})$ according to (4.28) or (4.33).

3. Draw the samples $\mathbf{B}^{(m)}$ for data symbols $\mathbf{B}$ from instantaneous priors $\pi_{k,j}(n)$.

4. If iteration increment reaches ΔM

   - Compute a new set of sample allocation indices $\{N_{k,j,n}\}$;
   - Update the instantaneous symbol priors $\pi_{k,j}(n)$ according to $\pi_{k,j}(n) \propto N_{k,j,n}$.

   End

5. Compute likelihood weights according to (4.23) and (4.29) or (4.34), i.e.,

$$\lambda(\mathbf{R}|b_k^{(m)}(n), \sigma^{2(m)}, \mathbf{H}^{(m)}) = exp(-\frac{1}{2\sigma^{2(m)}}\|\mathbf{r}(n) - \{\sum_{k=1}^{K} A_k[b_k^{(c)}(n)\mathbf{C}_{k,0}(n) + b_k^{(c)}(n-1)\mathbf{C}_{k,1}(n-1)]g_k^{(m)}\}\|^2) \tag{4.35}$$

or

$$\lambda(\mathbf{R}|b_k^{(m)}(n), \sigma^{2(m)}, \mathbf{H}^{(m)}) = exp(-\frac{1}{2\sigma^{2(m)}}\|\mathbf{r}(n) - \{\sum_{k=1}^{K} A_k[b_k^{(c)}(n)\mathbf{C}_{k,0}(n) + b_k^{(c)}(n-1)\mathbf{C}_{k,1}(n-1) + b_k^{(c)}(n-2)\mathbf{C}_{k,2}(n-2)]\mathbf{h}_k^{(m)}\}\|^2). \tag{4.36}$$

6. Perform symbol detection at the lth iteration by

$$P(b_k(n) = \beta_j|\mathbf{R}) \propto \mathbb{E}_l\{\lambda(\mathbf{R}|b_{k,j}^{(l)}(n), \sigma^2, \mathbf{H})\}, \tag{4.37}$$

7. Estimate channel response $\{g_k\}$ or $\{\mathbf{h}_k\}$ from sample means

$$\mathbb{E}_l\{g_k^{(l)}|\mathbf{R}\} \cong \frac{1}{l}\sum_{m=1}^{l} g_k^{(m)} \quad or \quad E_l\{\mathbf{h}_k^{(l)}|\mathbf{R}\} \cong \frac{1}{l}\sum_{m=1}^{l} \mathbf{h}_k^{(m)} \tag{4.38}$$

8. Estimate noise level σ^2 from sample means

$$\mathbb{E}_l\{\sigma^{2(l)}|\mathbf{R}\} \cong \frac{1}{l}\sum_{m=1}^{l} \sigma^{2(m)}. \tag{4.39}$$

END

As an MCMC method, the adaptive sampling method generates independent random samples and the current samples are biased by the information taken from previous samples.

4.4.4 Comparisons to Gibbs sampler detector

Compared to the blind multiuser receivers based on Gibbs sampler [84], the proposed receiver takes an adaptive procedure to find the most effective distribution for the generation of samples. For iterations of the procedure, an optimized dynamic sample allocation scheme is adopted to give indication where the samples should be distributed. Thus faster convergence of the proposed receiver is achieved. For the proposed inference procedure, since the "burn-in" period is not necessary, all samples from all iterations can be used to calculate the Bayesian estimates of the unknown quantities.

As a result, the computation cost of the proposed receiver is lower than the Gibbs receiver. For the Gibbs sampler, the computation of the conditional posterior distributions of data $\mathbf{B}$ is necessary for every iteration to update the sampling distribution. While for the adaptive sampling, only after every $\Delta M > 1$ increment of the iteration index, the sampling distributions are updated according to the previous samples. Thus the computational complexity for the inference procedure of the proposed receiver decreases obviously.

4.5 Simulation results

This section presents simulation examples that are used to test the performance of the proposed Bayesian multiuser detection for the long code CDMA systems over fading channels. All users' spreading sequences are generated randomly from equal probability binary code $\{\pm 1\}$, and the processing gain is chosen as $P = 10$. The fading coefficients of the channels are generated according to uncorrelated circular complex Gaussian distribution. The initial delay ζ_k for user k is generated randomly from 0 to $P - 1$.

The simulation is setup with the following non-informative conjugate prior distributions. For the case of flat fading channel

$$p(g_k^{(0)}) \sim \mathcal{N}(g_{k0}, \Sigma_{k0}) : g_{k0} = 0, \Sigma_{k0} = 10,$$

$$p(\sigma^{2(0)}) \sim \chi^{-2}(\nu_0, \lambda_0) : \nu_0 = 1, \lambda_0 = 0.1;$$

and for the case of frequency-selective fading channel

$$p(\mathbf{h}_k^{(0)}) \sim \mathcal{N}(\mathbf{h}_{k0}, \Sigma_{k0}) : \mathbf{h}_{k0} = \mathbf{0}, \Sigma_{k0} = 1000\boldsymbol{I},$$

$$p(\sigma^{2(0)}) \sim \chi^{-2}(\nu_0, \lambda_0) : \nu_0 = 1, \lambda_0 = 0.3.$$

For both cases, the initial values for data symbols are generated randomly. Note that the performance of the detector is insensitive to the values of the parameters in these priors, since the priors are non-informative.

Both the uncoded system and the coded system are tested by the simulations. For uncoded system, the data block size is chosen as $N = 128$. For coded system, one half of the constraint length-5 convolutional code [6] (with generator 23 and 35 in octal notation) is employed. We choose 128 information bits, i.e., $N = 256$. The adaptive sampling is performed for 100 iterations for both systems.

4.5.1 Convergence of the data detection

The convergence behaviors of the Bayesian multiuser detector are studied firstly. Let us consider five active users, and the amplitudes of the users and the noise variance are given as

$$A_1^2 = -4dB, A_2^2 = -2dB, A_3^2 = 0dB,$$

$$A_4^2 = 2dB, A_5^2 = 4dB, \sigma^2 = -1.5dB.$$

For the case of the flat fading channel, we consider only the coded system. In Fig 4.1, Fig 4.2 and Fig 4.3, we plot the first 100 samples of the symbols $b_2(50)$, $b_3(80)$, $b_5(120)$ drawn by the Gibbs sampler and the proposed method respectively for the coded system. Same as the definition in the Section 4.2, $b_k(n)$ means the nth symbol transmitted by user k.

For the case of the frequency-selective fading channel, both the uncoded system and the coded system are simulated. Fig 4.4, Fig 4.5 and Fig 4.6 present 100 samples of $b_1(90)$, $b_4(45)$, $b_5(60)$ obtained from the Gibbs sampler and the proposed detection procedure for uncoded system, respectively. Fig 4.7, Fig 4.8 and Fig 4.9 present 100 samples of $b_1(10)$, $b_2(70)$, $b_3(70)$ obtained from the Gibbs sampler and the proposed detection procedure for coded system, respectively.

Fig 4.1 - Fig 4.9 show that the Gibbs sampler method needs a burn-in period to begin the inference procedure, while the proposed method based on adaptive sampling enters the procedure directly and reaches the convergence quickly. Therefore, as seen from the simulation results, the proposed receiver improves the convergence speed substantially. Since "burn-in" period is not needed, all the samples can be used to calculate the Bayesian estimates.

4.5.2 Channel estimation

The channel estimation results of the proposed receiver are illustrated for coded system. Fig 4.10 and Fig 4.11 show the first 100 samples of the channel coefficients g_2 and g_4 drawn by the proposed method for the flat-fading channels. Fig 4.12 - Fig 4.15 present the first 100 samples of channel responses $h_1(1)$, $h_3(1)$, $h_3(3)$ and $h_5(2)$ drawn by the proposed method for the frequency-selective fading channels. It is shown that the inference procedure converges quickly because the samples of all the quantities converge to their real values in a few iterations.

4.5.3 Noise level estimation

For simplicity, the estimation of noise level is tested only for system over the coded frequence-selective fading channel. Fig 4.16 plots the first 100 samples of noise variance σ^2 drawn by the proposed method. Since the samples converge to the real value within only several iterations, it is again demonstrated the fast convergence speed of the inference procedure.

4.5.4 Detection performance

Let us now consider the BER performance of the proposed Bayesian multiuser detector for the long code systems. With the assumption that all users have the same amplitudes, we test the BER versus signal-to-noise ratio (SNR) for the coded systems. Based on the results of the previous simulation examples, it is enough to choose the number of iterations to be 50 because all parameter samples reach convergence after 50 iterations. The simulation results for the systems over flat fading channel and frequency-selective fading channel are shown in Fig 4.17 and Fig 4.18, respectively.

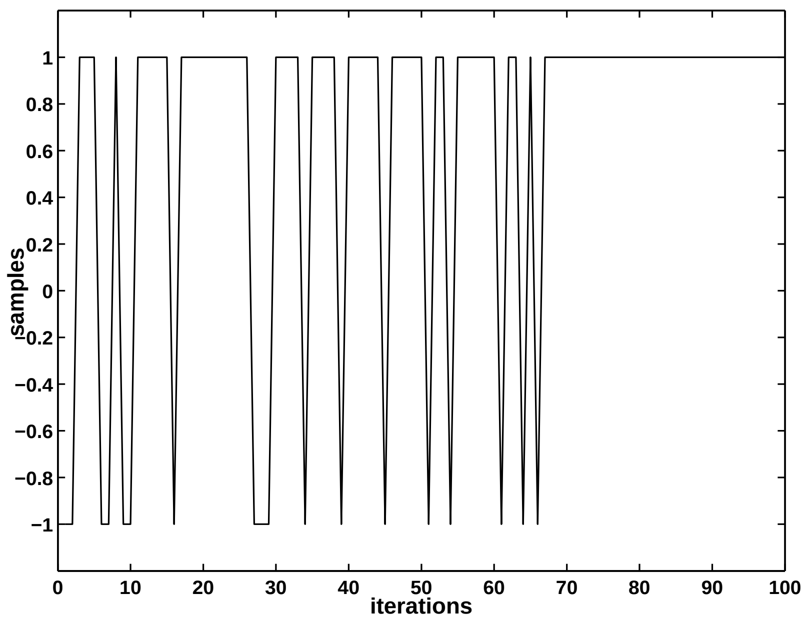
Comparisons are made on the BER performances achieved by the proposed receiver and the Gibbs sampler receiver. The results are presented in Fig 4.19 and Fig 4.20 for the systems over flat fading channel and frequency-selective fading channel respectively. It is seen from these figures, that the proposed receiver achieves the comparable performance to the Gibbs sampler receiver.

4.5.5 Near-far resistance

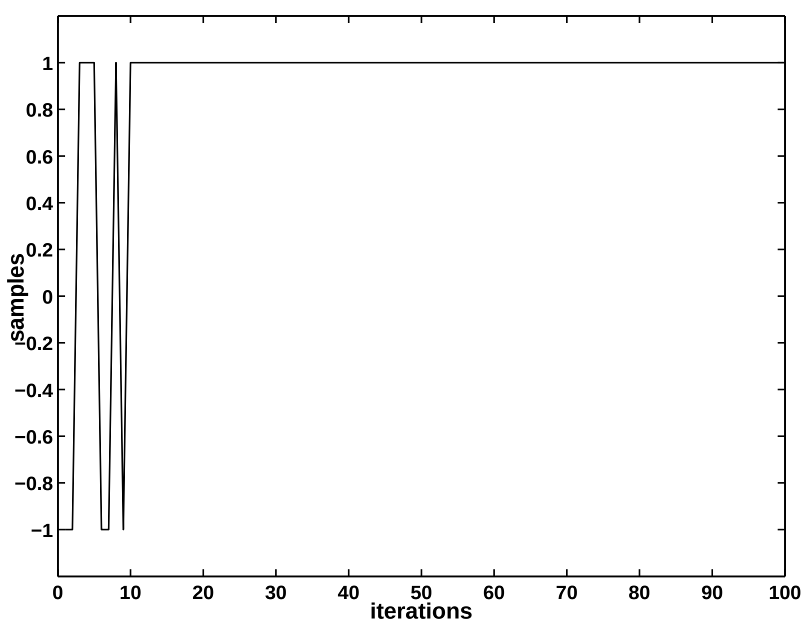
The capability of near-far resistance is illustrated in Fig 4.21 for the system over flat fading channel, and in Fig 4.22 for the system over frequency-selective fading channel. The near-far ratio is defined as the ratio between the power of interfering users and the power of the desired user. Let us fix the power of the desired user and change the power of interfering users, and assume that all interfering users have the same power. The BER performance is tested as the function of near-far ratio for the systems with $K = 5$, 8 and 10. The results in Fig 4.21 and Fig 4.22 demonstrate that the propose technique performs well in near-far situations.

4.6 Conclusions

This chapter proposes a blind Bayesian receiver for the DS-CDMA systems employing long spreading codes in the presence of unknown fading channels. An efficient Bayesian MCMC inference procedure is derived based on the adaptive sampling algorithm for the joint process of multiuser detection and channel estimation. The implementation of the adaptive sampling is effective to improve the convergence speed with a reduced computational complexity and comparable good performance. Simulation results are provided to demonstrate the effectiveness of the proposed technique.

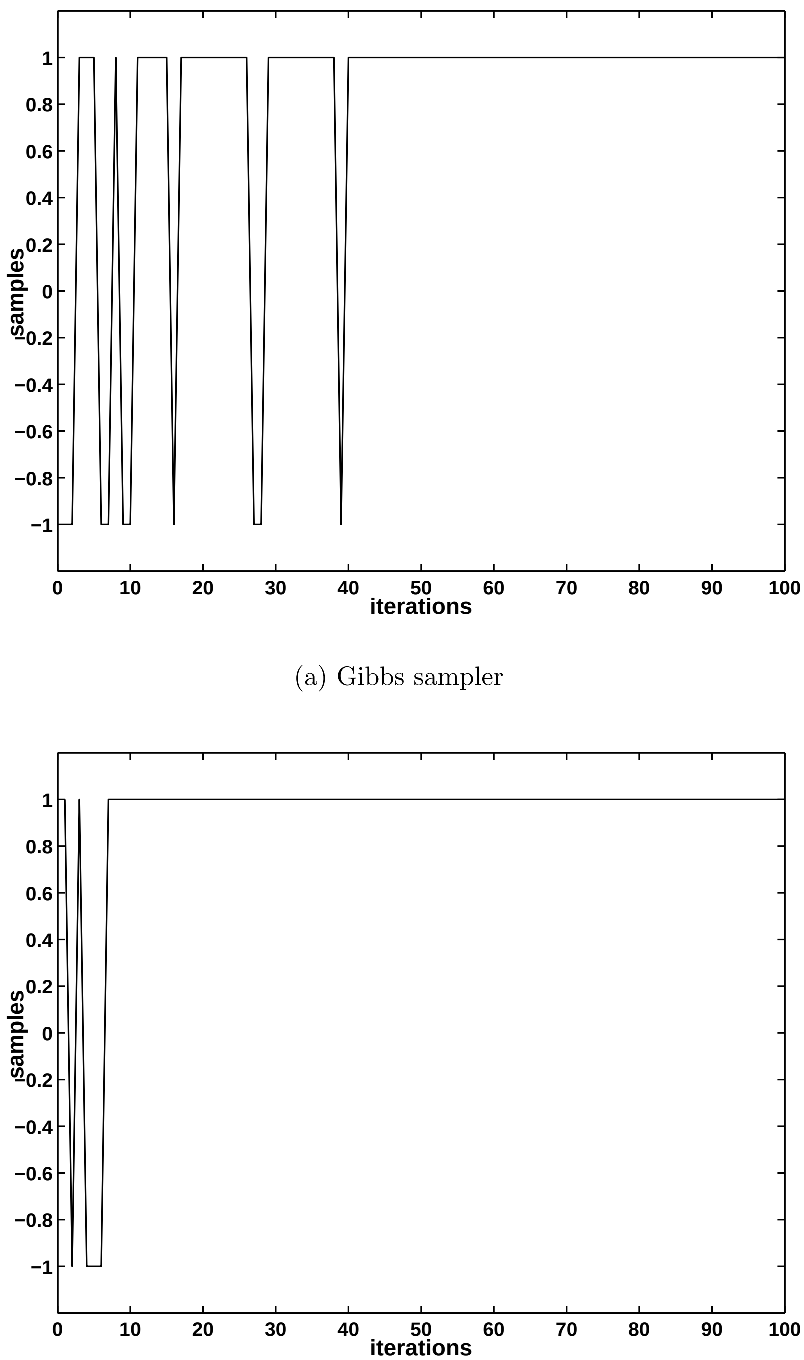


(a) Gibbs sampler



(b) Adaptive sampling

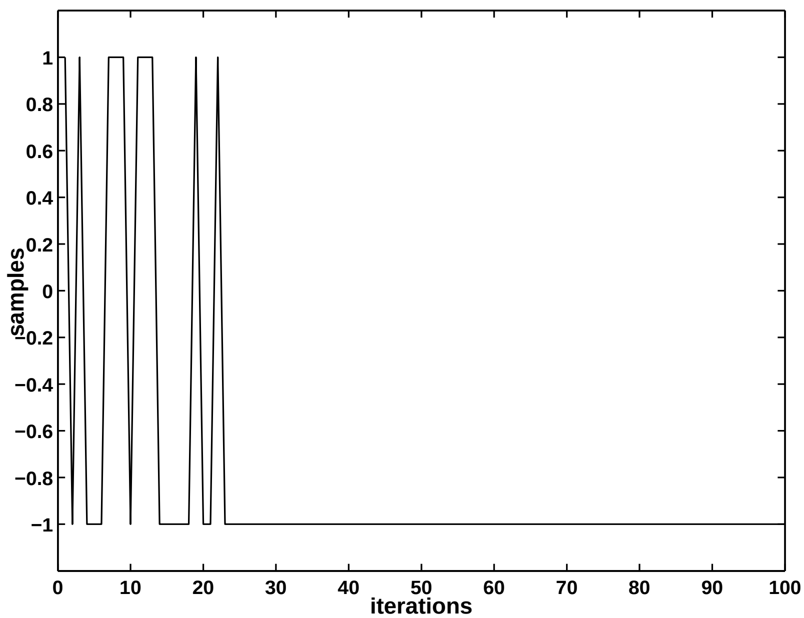
Figure 4.1: Samples of data bit $b_2(50) = 1$ in the coded system (flat fading channel)



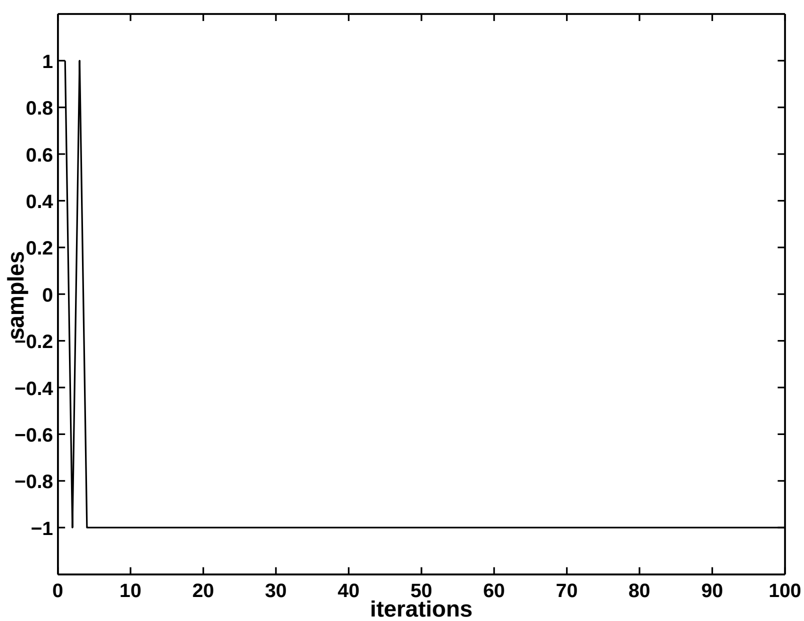
(a) Gibbs sampler

(b) Adaptive sampling

Figure 4.2: Samples of data bit $b_3(80) = 1$ in the coded system (flat fading channel)

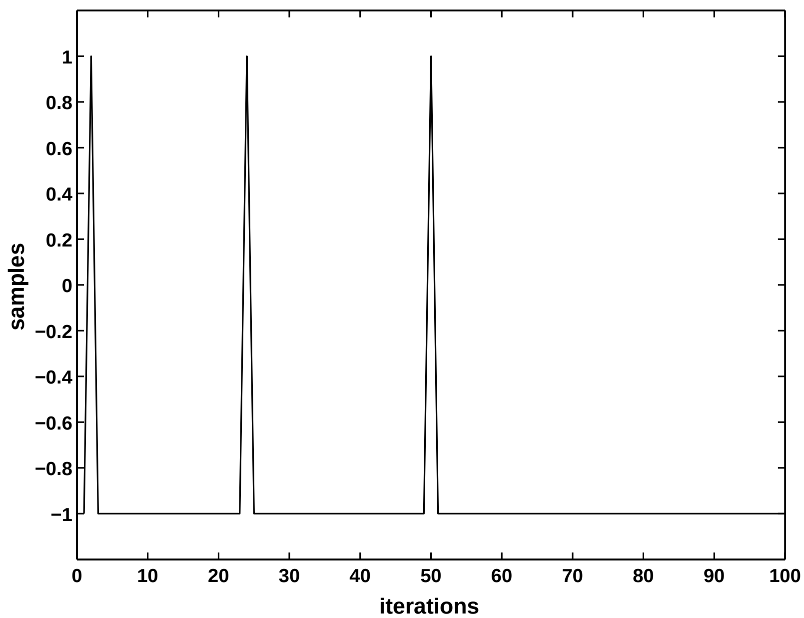


(a) Gibbs sampler

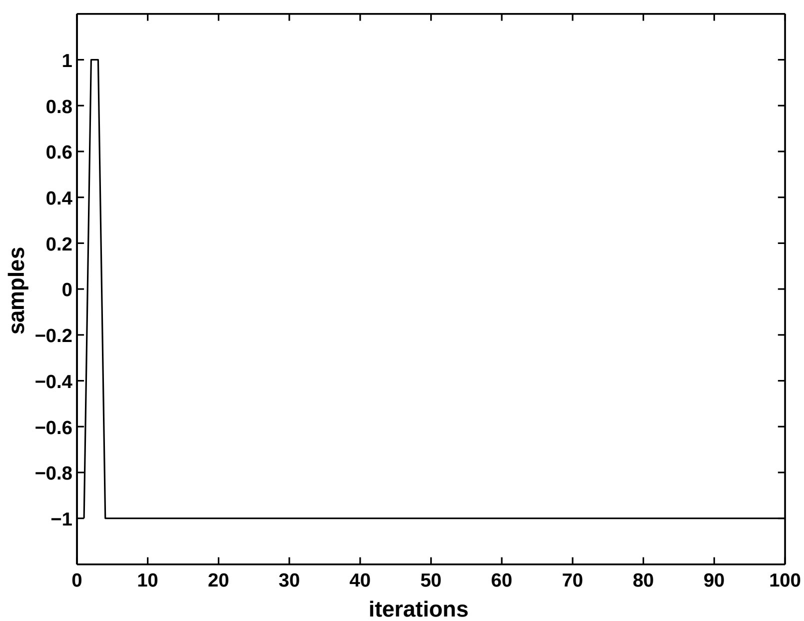


(b) Adaptive sampling

Figure 4.3: Samples of data bit $b_5(120) = -1$ in the coded system (flat fading channel)

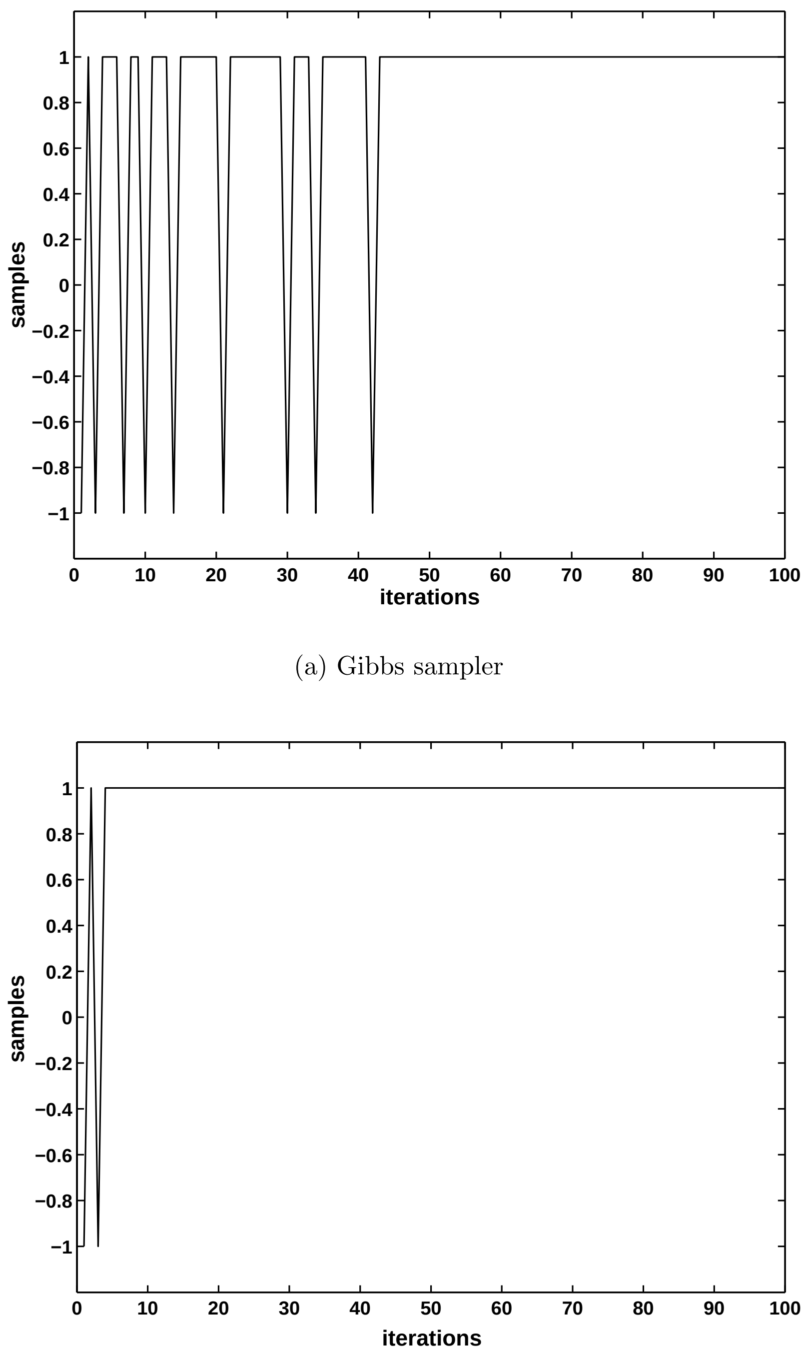


(a) Gibbs sampler



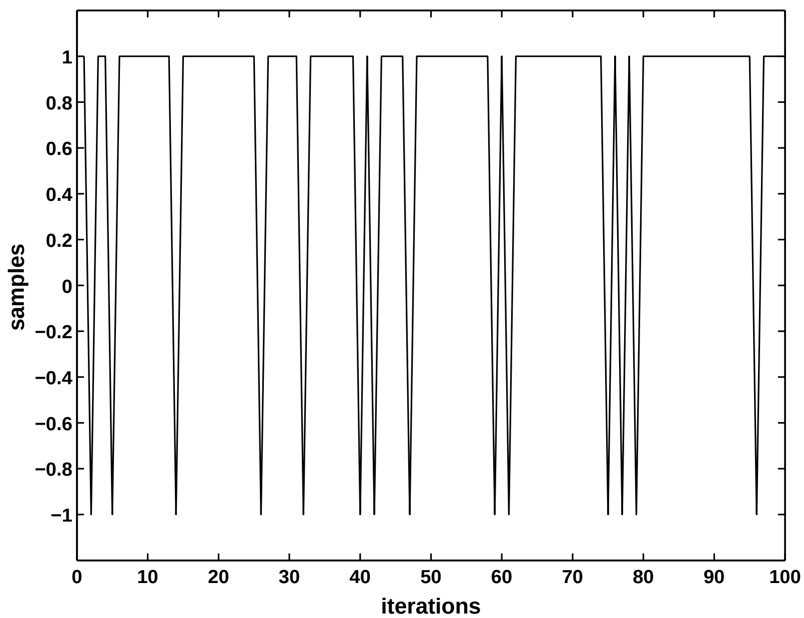
(b) Adaptive sampling

Figure 4.4: Samples of data bit $b_1(90) = -1$ in the uncoded system (frequency-selective fading channel)

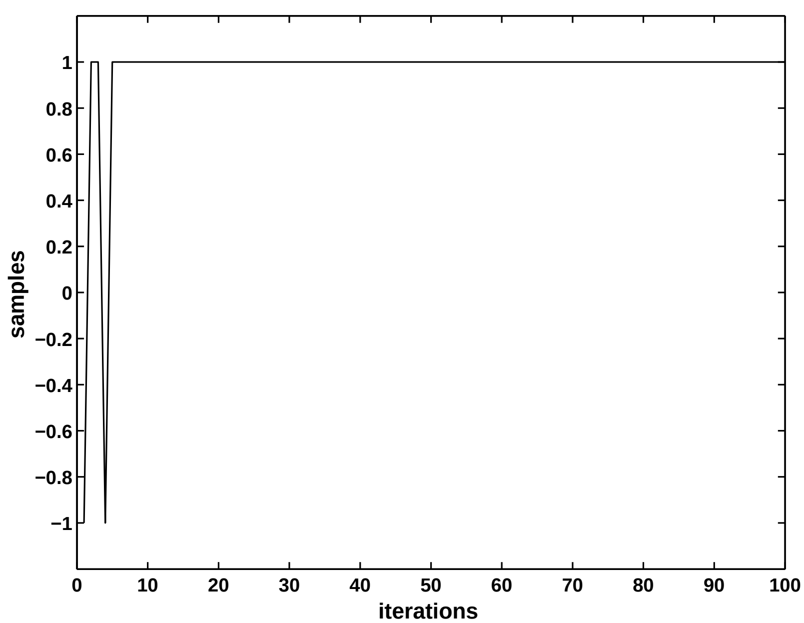


(b) Adaptive sampling

Figure 4.5: Samples of data bit $b_4(45) = 1$ in the uncoded system (frequency-selective fading channel)

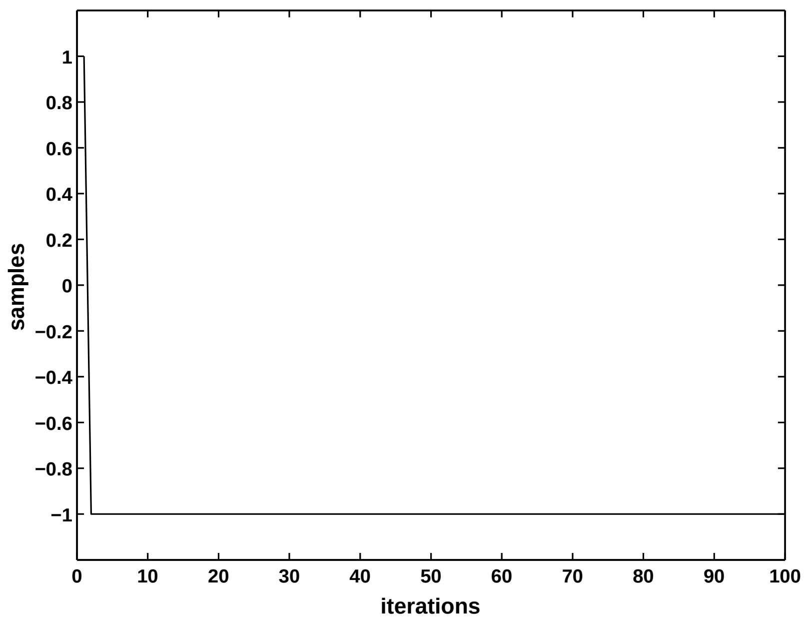


(a) Gibbs sampler

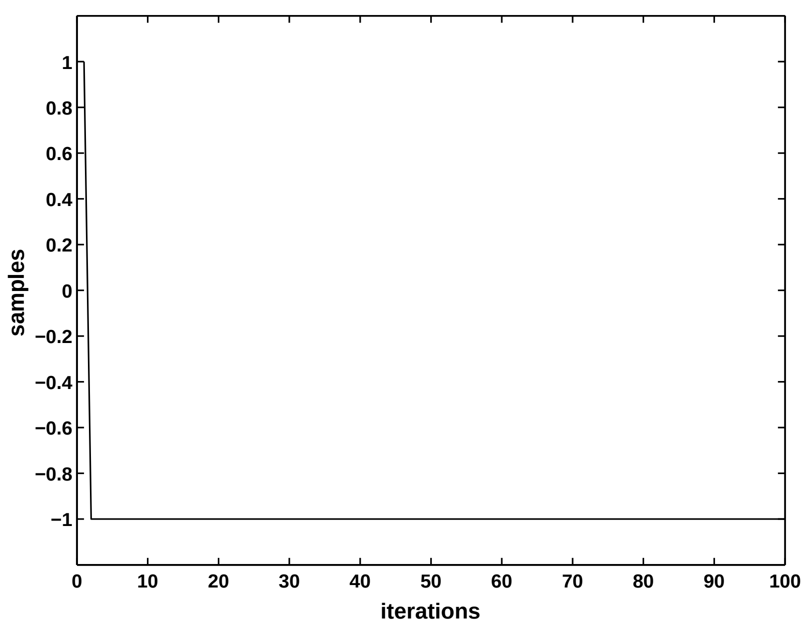


(b) Adaptive sampling

Figure 4.6: Samples of data bit $b_5(60) = 1$ in the uncoded system (frequency-selective fading channel)

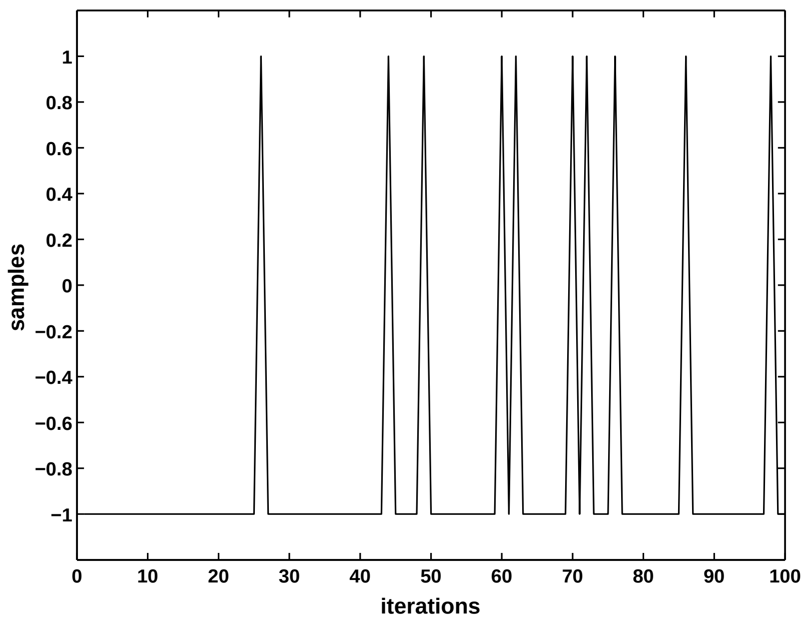


(a) Gibbs sampler

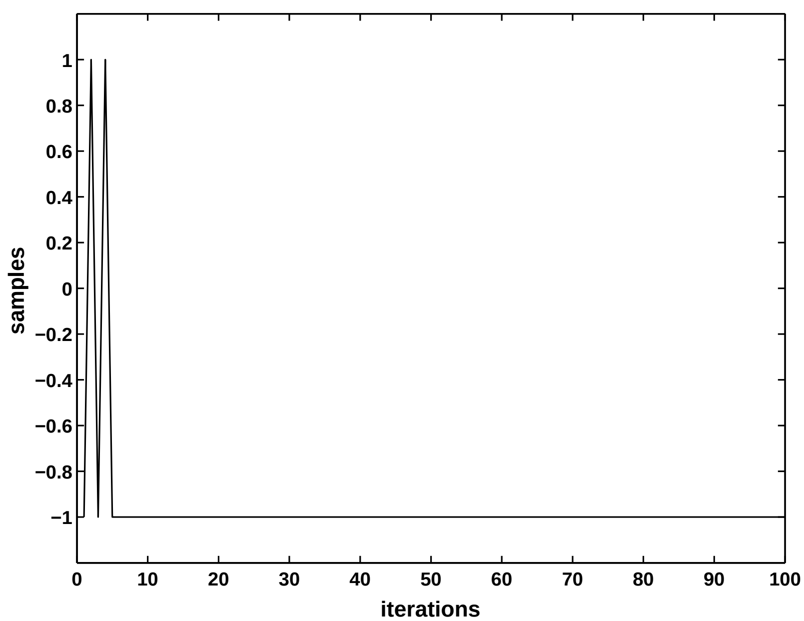


(b) Adaptive sampling

Figure 4.7: Samples of data bit $b_1(10) = -1$ in the coded system (frequency-selective fading channel)

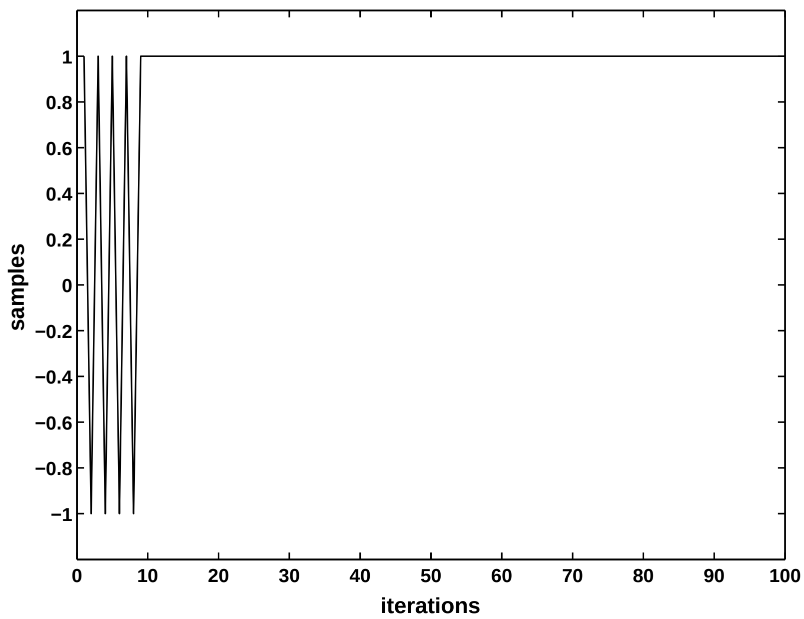


(a) Gibbs sampler

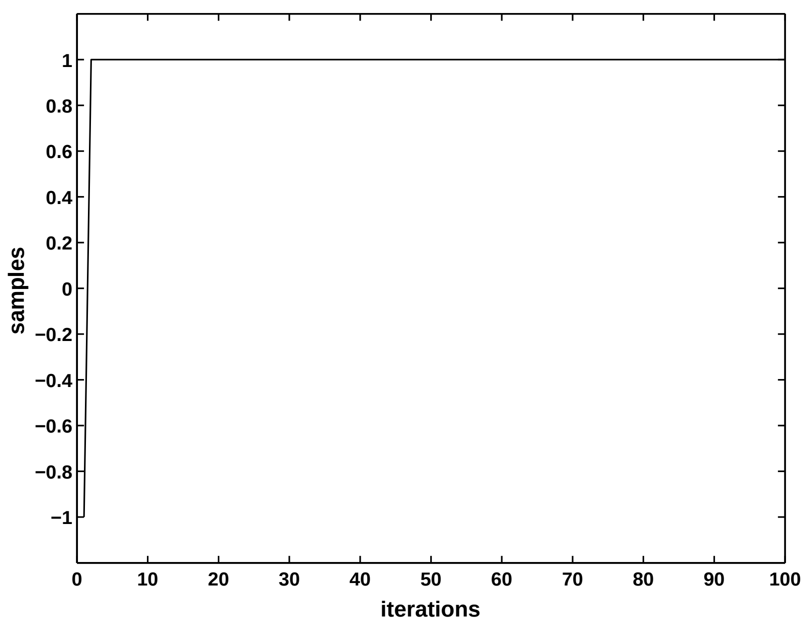


(b) Adaptive sampling

Figure 4.8: Samples of data bit $b_2(70) = -1$ in the coded system (frequency-selective fading channel)



(a) Gibbs sampler



(b) Adaptive sampling

Figure 4.9: Samples of data bit $b_3(70) = 1$ in the coded system (frequency-selective fading channel)

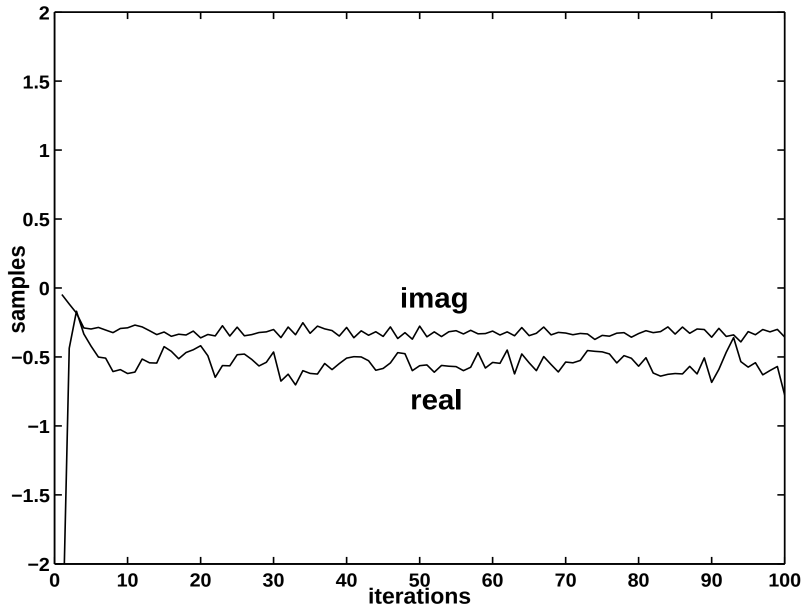


Figure 4.10: Samples of flat fading channel coefficient ($g_2 = -0.5082 - 0.2715i$)

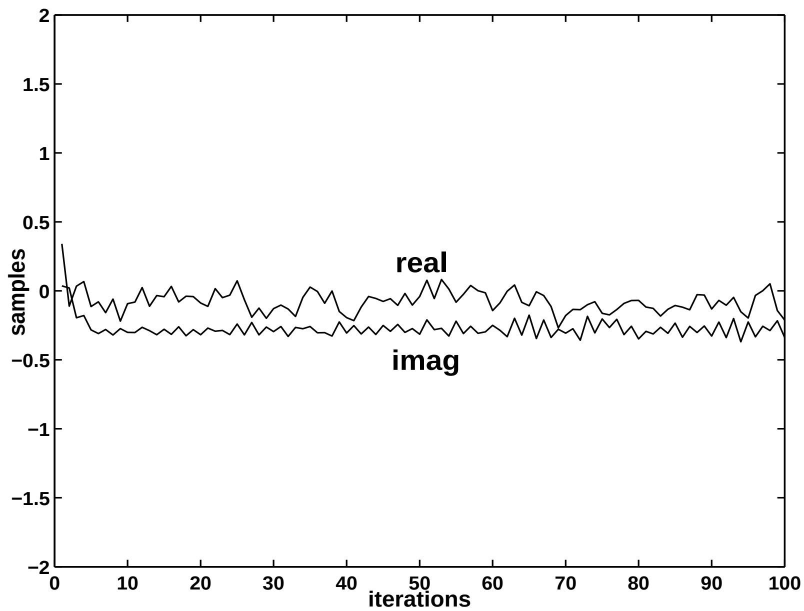


Figure 4.11: Samples of flat fading channel coefficient ($g_4 = -0.1865 - 0.4120i$)

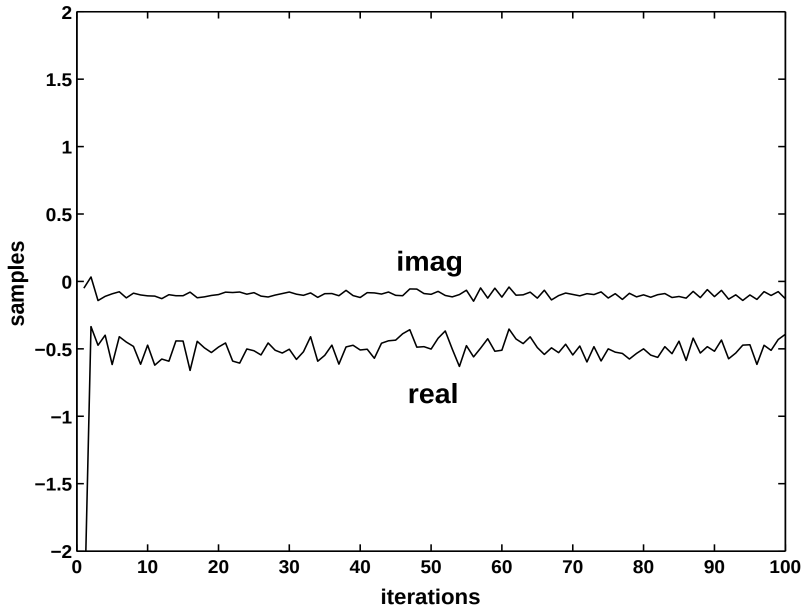


Figure 4.12: Samples of multipath parameters ($h_1(1) = -0.5044 - 0.1665i$)

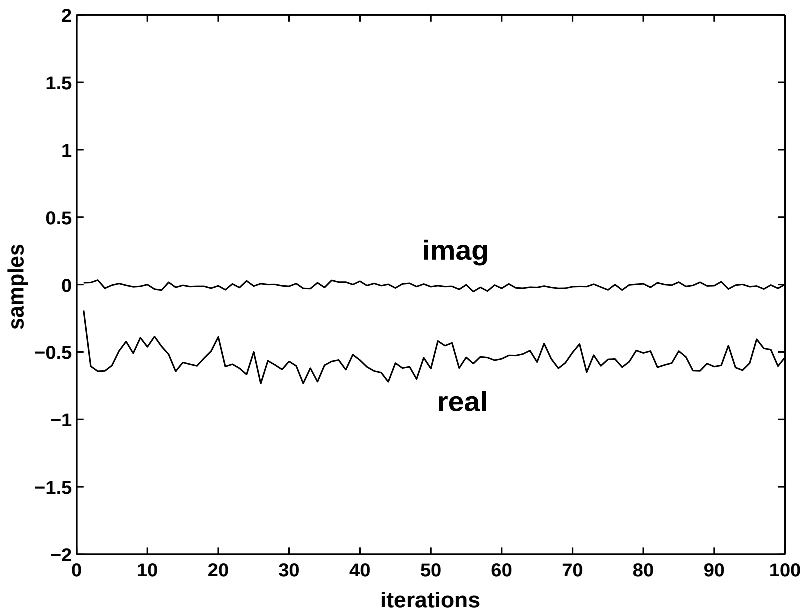


Figure 4.13: Samples of multipath parameters ($h_3(1) = -0.5110 + 0.0072i$)

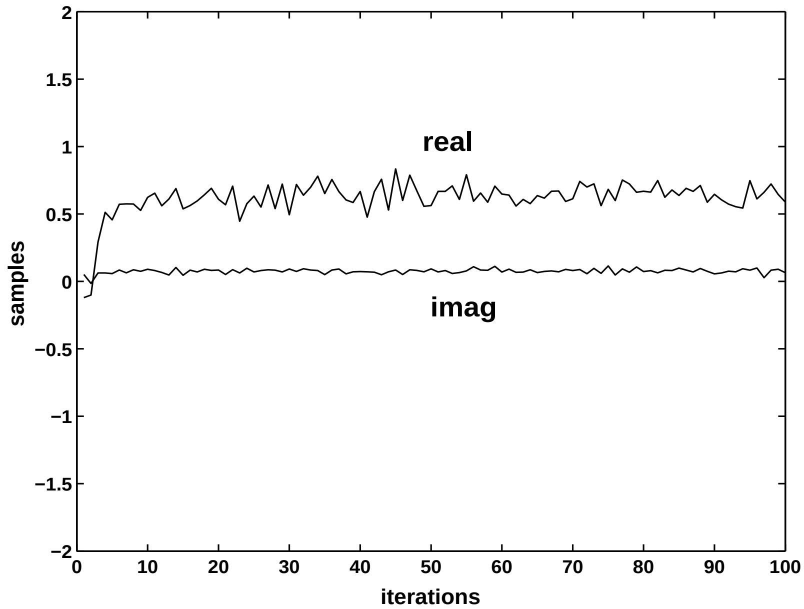


Figure 4.14: Samples of multipath parameters ($h_3(3) = 0.5205 + 0.0888i$)

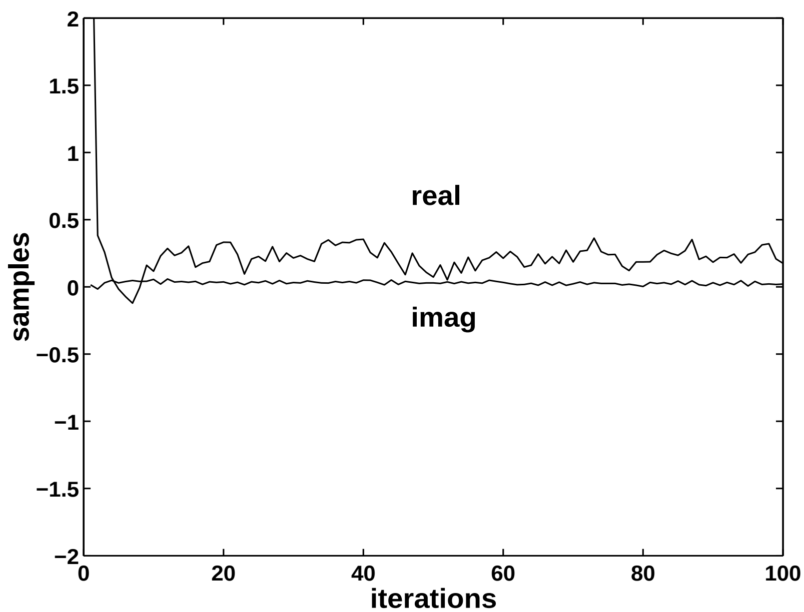


Figure 4.15: Samples of multipath parameters ($h_5(2) = 0.2685 + 0.0565i$)

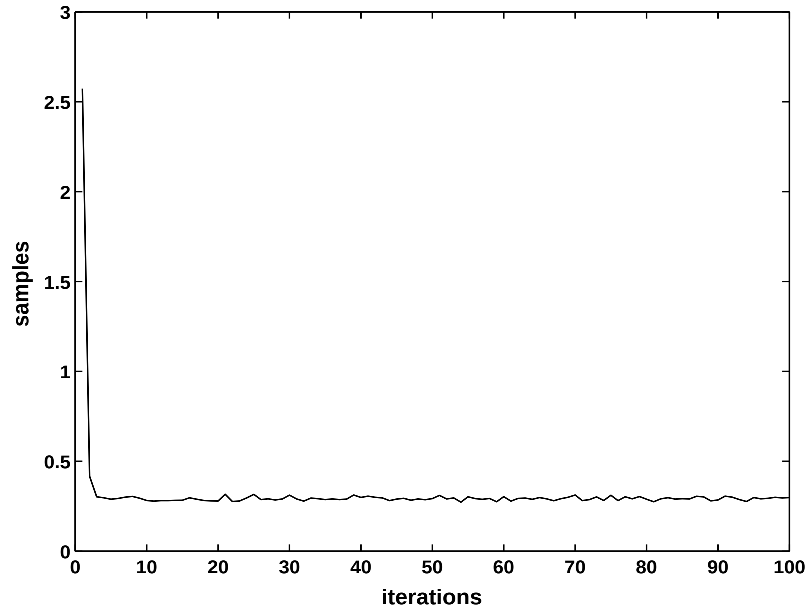


Figure 4.16: Samples of noise σ^2 ($\sigma^2 = -1.5dB$) (frequency-selective fading channel)

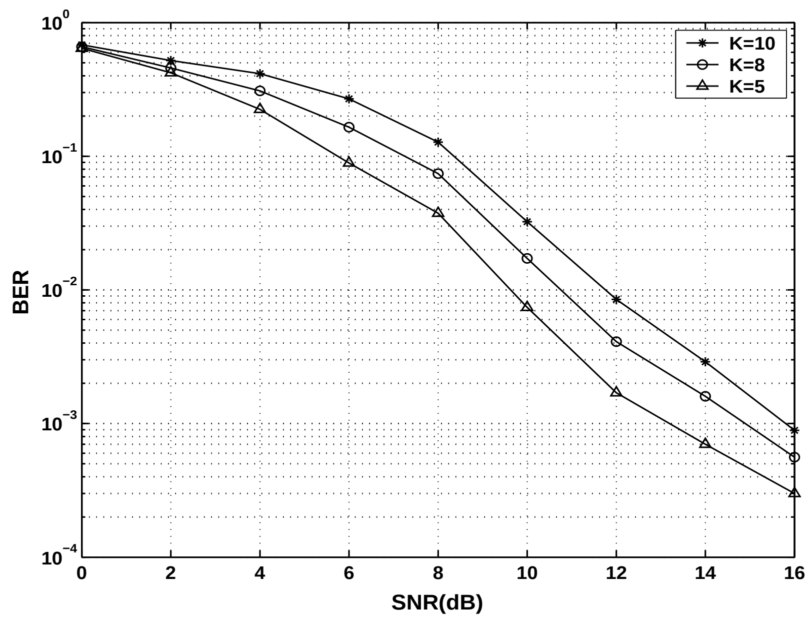


Figure 4.17: BER versus SNR (flat fading channel)

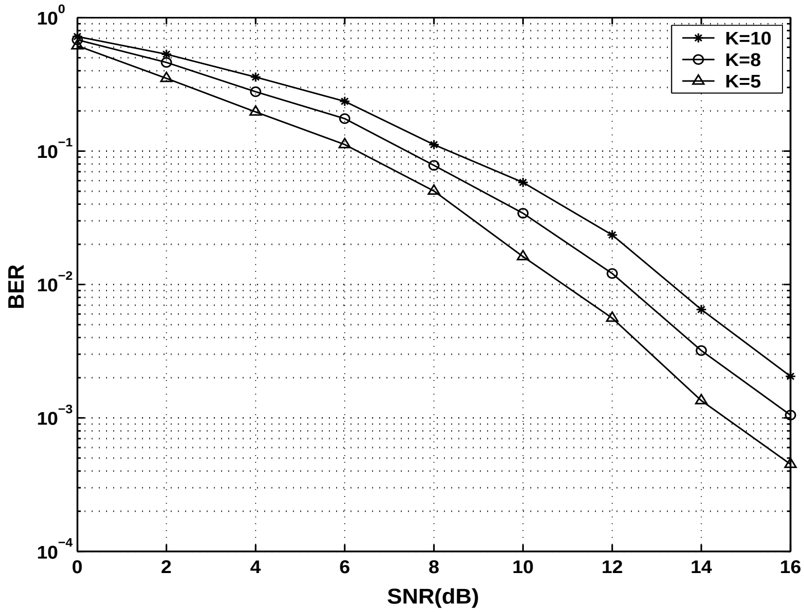


Figure 4.18: BER versus SNR (frequency-selective fading channel)

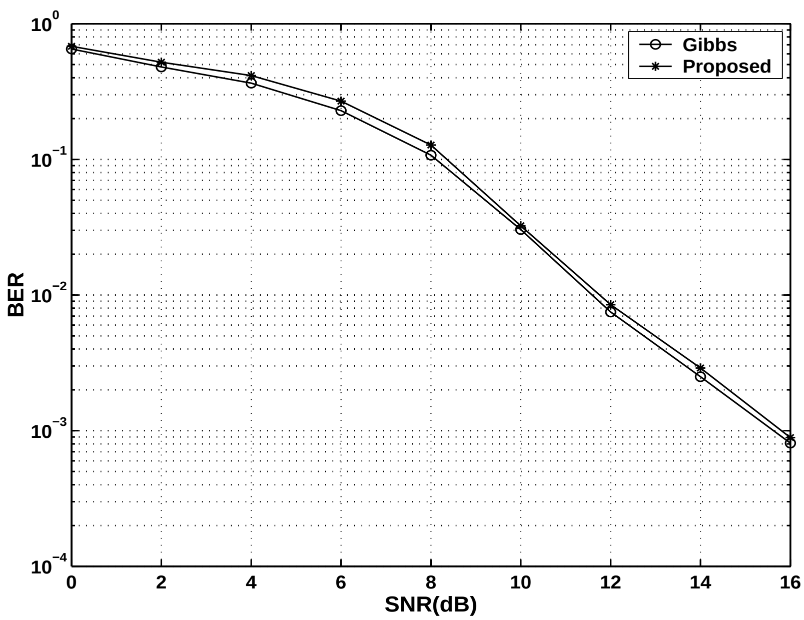


Figure 4.19: Comparison on BER performance, $K = 10$ (flat fading channel)

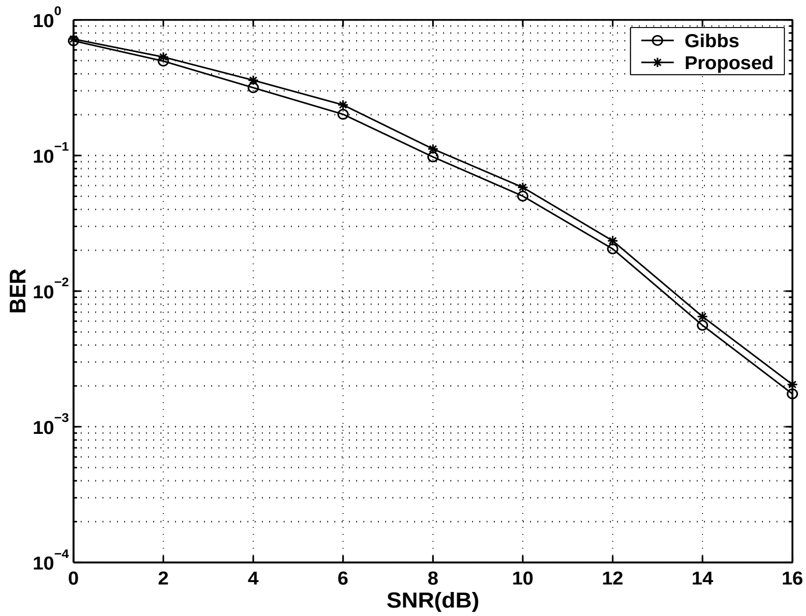


Figure 4.20: Comparison on BER performance, $K = 10$ (frequency-selective fading channel)

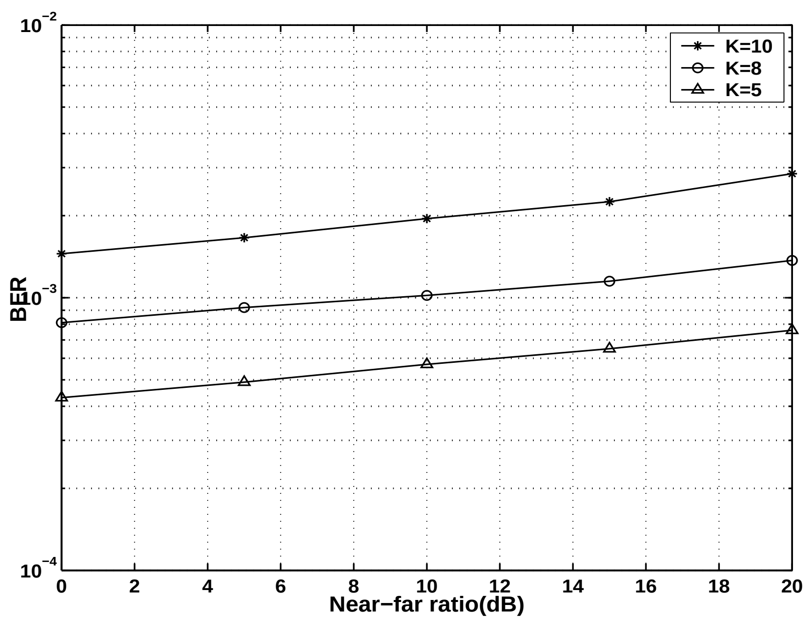


Figure 4.21: BER versus near-far ratio, $SNR = 15\ dB$ (flat fading channel)

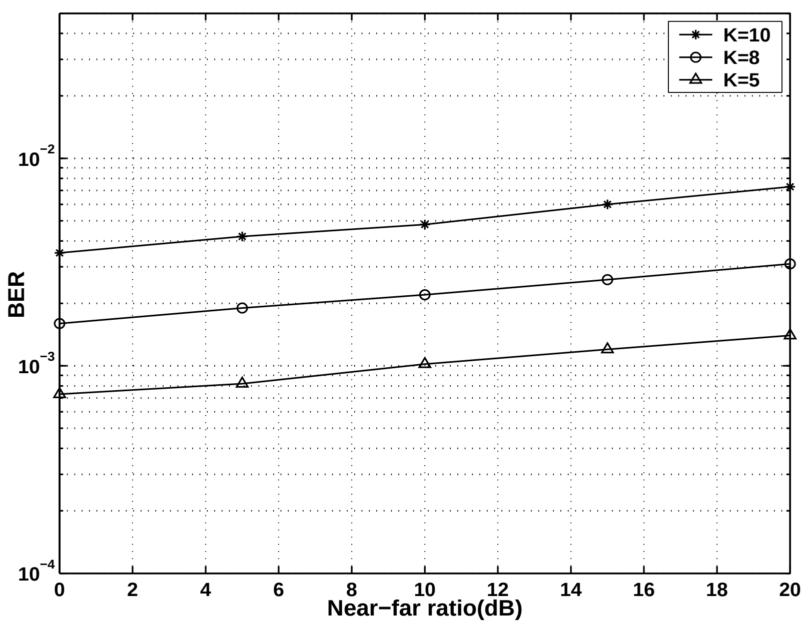


Figure 4.22: BER versus near-far ratio, $SNR = 15\ dB$ (frequency-selective fading channel)

Chapter 5

Blind SMC multiuser detection based on EM framework

5.1 Introduction

The Sequential Monte Carlo (SMC) methods, which are the other category of Monte Carlo signal processing, have also begun to show a great potential for solutions to a wide range of statistical inference problems [86]. Comparing to the MCMC methods, the SMC provides a better performance achieved by parallel processing and is suited to practical applications. By iteratively generating Monte Carlo samples of the state variables or other latent variables, the posterior distribution of any system parameter of interest can be approximated. The complete theoretical framework using SMC methods is described in [87]. The SMC methods have been successfully applied to a few problems in communications, such as blind equalization and detection in fading channels [79] [88] [89] [90] [91] [110]. The solution to the detection problem for general MIMO systems was presented in [90]. The demodulation algorithm based on SMC and QR decomposition was developed in [111] for the MIMO systems over flat-fading channel. For applications to CDMA systems, the SMC-KF reported in [91] combines the conventional Kalman filter and importance sampling technique to approximate the multi-access interference (MAI) as circular Gaussian for the problem of single-user detection. The particle filtering methods are developed for the multiuser framework of CDMA system [92] [93], the required computational complexity of all these methods grows exponentially with the number of users. Therefore the conventional SMC method is inefficiently used for multiuser detection of DS-CDMA system with a large number of users.

A computationally efficient decomposition algorithm was reported in [112] for parameter estimation of superimposed signals. This algorithm is based on the expectation-maximization (EM) framework to decompose the observed data into a number of signal components and then estimate the parameters of each signal component separately. This generalized EM algorithm has been used in some literatures [113] [114] to demonstrate the effectiveness of avoiding exponential complexity in multiuser detection. Therefore, it is possible to adopt this decomposition algorithm to divide a complicated SMC estimation problem into smaller ones so that the total computational complexity is minimized and the error propagation problem is avoided at the same time.

This chapter presents an SMC-based formulation of multiuser detection for DS-CDMA system with unknown fading channels. Firstly, the multiuser system is decomposed into separate single user systems by the EM decomposition algorithm. Then the SIS and the Kalman filter are combined to perform the data detection and channel estimation for every single user system. With the decomposition of the superimposed observation signals, the total computational complexity of the SMC based method can be reduced from $O(|A|^K)$ to $O(|A|K)$, where $|A|$ is the size of transmitted symbol set $\{a_1, \cdots, a_{|A|}\}$ and K is the number of the active users. Based on these concepts, a novel iterative receiver EM-SMC is developed. Simulation results are presented to show that the receiver performs well over fading channel with a significantly reduced computational complexity.

5.2 Signal model

5.2.1 Flat fading channels

Let us consider a DS-CDMA system that has K active users whose signals are transmitted over flat fading channel with additive white Gaussian noise. The representation of the received signals is given by

$$r(t) = \sum_{k=1}^{K} \sum_{n=0}^{N-1} g_k s_k(n) \bar{c}_k(t - nT - \tau_k) + \omega(t) \tag{5.1}$$

where T is the symbol duration, τ_k and $s_k(n)$ denote the delay of the signal and the transmitted symbol for user k, respectively. It is assumed that the transmitted symbols are independent and taken from a finite alphabet set $\mathbf{A} = \{a_1, \cdots, a_i, \cdots, a_{|A|}\}$, and for each user $k = 1, \cdots, K$, the transmitted symbols $\{s_k(n)\}_{n=0}^{N-1}$ are differentially modulated with the binary information symbols $\{d_k(n)\}_{n=1}^{N-1}$. The spreading waveform, $\bar{c}_k(t)$, for user k is formed by the pulse shape $\phi(t)$ and spreading sequence $\bar{c}_k(p)$ for user k is defined as

$$\bar{c}_k(t) = \sum_{p=0}^{P-1} \bar{c}_k(p)\phi(t - pT_c) \tag{5.2}$$

where P is the spreading gain, $T_c = T/P$ is the chip duration, and g_k is the complex fading channel gain between the transmitter and the receiver for user k.

The received signals are processed by a chip-matched filter and sampled at the chip-rate to generate the $P \times 1$ observation vector expressed as

$$\mathbf{r}(n) = \sum_{k=1}^{K} s_k(n)\mathbf{C}_k g_k + \omega(n). \tag{5.3}$$

Here we assume that the delay spread is small compared with the symbol interval, so that the inter-symbol interference is negligible. Therefore, $\mathbf{C}_k$ denotes as

$$\mathbf{C}_k = [\bar{c}_k(0), \bar{c}_k(1), \ldots, \bar{c}_k(P-1)]^T. \tag{5.4}$$

The noise term $\{\omega(n)\}_{n=0}^{N-1}$ in (5.3) is assumed to be an i.i.d complex white Gaussian vector with zero-mean and variance σ^2, i.e., $\omega(\mathbf{n}) \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I}_P)$.

Let $\mathbf{R} = \{\mathbf{r}(n)\}_{n=0}^{N-1}$ and $\mathbf{d}(n) = \{d_k(n)\}_{k=1}^{K}$. Our objective is to estimate the *a posteriori* probabilities of the information symbols

$$P(\mathbf{d}(n) = \mathbf{a}_i|\mathbf{R}), \mathbf{a}_i \in \mathbf{A}^K, n = 1, \cdots, N-1 \tag{5.5}$$

based on the received signals $\mathbf{R}$ without knowing the channel information $\{g_k\}_{k=1}^{K}$.

5.2.2 Frequency-selective fading channels

Now consider a DS-CDMA system with the same signals which transmit through the frequency-selective fading channel. The multipath channel for user k is modeled by the impulse response $h_k(t)$ which is defined as

$$h_k(t) = \sum_{l=1}^{L} \alpha_{k,l}\delta(t - \tau_{k,l}) \tag{5.6}$$

where L is the total number of resolvable paths in the channel, $\alpha_{k,l}$ and $\tau_{k,l}$ are, respectively, the complex path gain and delay of the lth path for user k.

Then the representation of the received signals is given by

$$r(t) = \sum_{k=1}^{K}\sum_{n=0}^{N-1} s_k(n)c_k(t - nT - \tau_k) + \omega(t) \tag{5.7}$$

where the effective spreading waveform, $c_k(t)$, for user k is constructed from convolution of the original spreading waveform and the channel response, i.e., $c_k(t) = \bar{c}_k(t) * h_k(t)$.

It is assumed that the maximum channel delay in terms of number of chips is smaller than P, and the initial delay is much smaller than the symbol interval. The received signals are processed by a chip-matched filter and sampled at the chip-rate to generate the $P \times 1$ observation vector expressed as

$$\mathbf{r}(n) = \sum_{k=1}^{K} s_k(n)\mathbf{C}_k\mathbf{h}_k + \omega(n) \tag{5.8}$$

where $\mathbf{C}_k$ denotes a $P \times L$ matrix which is expressed as

$$\mathbf{C}_k = \begin{bmatrix} \bar{c}_k(0) & 0 & \dots & 0 \\ \bar{c}_k(1) & \bar{c}_k(0) & & \vdots \\ & \bar{c}_k(1) & \ddots & 0 \\ \vdots & & \ddots & \bar{c}_k(0) \\ \vdots & \vdots & & \bar{c}_k(1) \\ & & & \vdots \\ \bar{c}_k(P-1) & \bar{c}_k(P-2) & \cdots & \bar{c}_k(P-L) \end{bmatrix} \tag{5.9}$$

and

$$\mathbf{h}_k = [h_{k,1}, h_{k,2}, \cdots, h_{k,L}]^T. \tag{5.10}$$

The noise term $\{\omega(n)\}_{n=0}^{N-1}$ in (5.8) is also assumed to be an i.i.d complex white Gaussian vector with a zero-mean and a variance σ^2, i.e., $\omega(\mathbf{n}) \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_P)$.

It is seen that the system over flat-fading channel (5.3) and the system over the frequency-selective fading channel (5.8) are described by similar equations, only with different forms of the variable involved.

Again, without knowing the channel response $\{\mathbf{h}_k\}_{k=1}^{K}$, the *a posteriori* probabilities of the information symbols $\{d_k(n)\}_{n=1:N-1}^{k=1:K}$ are to be estimated based on the received signals $\{\mathbf{r}(n)\}_{n=0}^{N-1}$ and the *a priori* information symbol probabilities.

5.3 Sequential Monte Carlo method

Sequential Monte Carlo (SMC) is a family of probability approximation methods which use Monte Carlo samples to efficiently estimate the posterior distribution of the unknown variables in dynamic systems [86]. SMC methods are very flexible, easy to implement, parallelisable and applicable in very general settings.

5.3.1 Problem statement

Let us considering a dynamic system modeled in the following state-space form

$$\begin{aligned} state\ \ equation: \mathbf{z}_t &= f_t(\mathbf{z}_{t-1}, \mathbf{u}_t) \\ observation\ \ equation: \mathbf{y}_t &= g_t(\mathbf{z}_t, \mathbf{v}_t) \end{aligned} \tag{5.11}$$

where $\mathbf{z}_t$, $\mathbf{y}_t$, $\mathbf{u}_t$ and $\mathbf{v}_t$ are, respectively, the state variable, the observation, the state noise, and the observation noise at time t, which can be either scalars or vectors.

Let $\mathbf{Z}_t = (\mathbf{z}_0, \mathbf{z}_1, \cdots, \mathbf{z}_t)$ and $\mathbf{Y}_t = (\mathbf{y}_0, \mathbf{y}_1, \cdots, \mathbf{y}_t)$, an online inference of $\mathbf{Z}_t$ is of interest. That is, at current time instant t, a timely estimate is to be made for the function, say $h(\mathbf{Z}_t)$, of the state variable $\mathbf{Z}_t$ based on the currently available observation, $\mathbf{Y}_t$. With the Bayes theorem, the optimal solution to this problem is found to be

$$E\{h(\mathbf{Z}_t)|\mathbf{Y}_t\} = \int h(\mathbf{Z}_t)p(\mathbf{Z}_t|\mathbf{Y}_t)d\mathbf{Z}_t. \tag{5.12}$$

In most cases, an exact evaluation of this expectation is analytically intractable because of the prohibitive computational complexity of such dynamic systems. Monte Carlo methods provide us a viable alternative to the direct integration. Specifically, if m random samples $\{\mathbf{Z}_t^{(j)}\}_{j=1}^m$ are drawn from the distribution $p(\mathbf{Z}_t|\mathbf{Y}_t)$, $E\{h(\mathbf{Z}_t)|\mathbf{Y}_t\}$ can be estimated with the functional mean of the samples, i.e.,

$$\hat{E}\{h(\mathbf{Z}_t)|\mathbf{Y}_t\} = \frac{1}{m}\sum_{j=1}^{m} h(\mathbf{Z}_t^{(j)}) \tag{5.13}$$

It is often not feasible to directly obtain samples from $p(\mathbf{Z}_t|\mathbf{Y}_t)$ at any time t. In applied statistics, Markov chain Monte Carlo (MCMC) methods are a popular approach to sampling from such complex probability distributions. However, MCMC methods are iterative algorithm unsuited to such recursive estimation problems.

5.3.2 Importance sampling

An alterative classical solution for this case is the *importance sampling* method [115]. It is easy to obtain samples from some trial distributions based on the concept of *importance sampling*.

Let us assume a trial distribution $q(\mathbf{Z}_t|\mathbf{Y}_t)$, which is the so-called *importance sampling distribution*, or *importance function*, and denote the *importance weights* as

$$w_t = \frac{p(\mathbf{Z}_t|\mathbf{Y}_t)}{q(\mathbf{Z}_t|\mathbf{Y}_t)}. \tag{5.14}$$

Consequently, if we generate a set of random samples $\{\mathbf{Z}_t^{(j)}\}_{j=1}^{m}$ according to the importance function $q(\mathbf{Z}_t|\mathbf{Y}_t)$ and associate the weights

$$w_t^{(j)} = \frac{p(\mathbf{Z}_t^{(j)}|\mathbf{Y}_t)}{q(\mathbf{Z}_t^{(j)}|\mathbf{Y}_t)} \tag{5.15}$$

to the sample $\mathbf{Z}_t^{(j)}$, the quantity of interest, $E\{h(\mathbf{Z}_t)|\mathbf{Y}_t\}$, can be estimated by

$$\hat{E}\{h(\mathbf{Z}_t)|\mathbf{Y}_t\} = \frac{1}{W_t}\sum_{j=1}^{m} h(\mathbf{Z}_t^{(j)})w_t^{(j)} \tag{5.16}$$

where $W_t = \sum_{j=1}^{m} w_t^{(j)}$. The pair $(\mathbf{Z}_t^{(j)}, w_t^{(j)})$, $j = 1, \cdots, m$, is called a properly weighted sample with respect to distribution $p(\mathbf{Z}_t|\mathbf{Y}_t)$.

Importance sampling is a general Monte carlo integration method. However, in its simplest form, it is not adequate for recursive estimation. That is because each time new data $\mathbf{Y}_{t+1}$ become available, the importance weights need to be recomputed over the entire state sequence. The computational complexity of this operation increases with time.

5.3.3 Sequential importance sampling

Because the state equation of the system possesses a Markovian structure, a recursive importance sampling strategy, *sequential importance sampling* (SIS), can be implemented. The importance sampling can be modified so that it becomes possible to compute the estimation of $p(\mathbf{Z}_t|\mathbf{Y}_t)$ without modifying the past generated samples and weights. This means that the importance function at time t, $q(\mathbf{Z}_t|\mathbf{Y}_t)$, admits as marginal distribution of the importance function at time $t-1$, $q(\mathbf{Z}_{t-1}|\mathbf{Y}_{t-1})$.

$$q(\mathbf{Z}_t|\mathbf{Y}_t) = q(\mathbf{Z}_{t-1}|\mathbf{Y}_t)q(\mathbf{z}_t|\mathbf{Z}_{t-1}, \mathbf{Y}_t) \tag{5.17}$$

By iterating we can obtain

$$q(\mathbf{Z}_t|\mathbf{Y}_t) = q(\mathbf{z}_0)\prod_{s=1}^{t} q(\mathbf{z}_s|\mathbf{Z}_{s-1}, \mathbf{Y}_s). \tag{5.18}$$

It is easy to see that this importance function allows us to update recursively the importance weights in time, that is

$$w_t^{(j)} = w_{t-1}^{(j)} \cdot \frac{p(\mathbf{Z}_t^{(j)}|\mathbf{Y}_t)}{p(\mathbf{Z}_{t-1}^{(j)}|\mathbf{Y}_{t-1})q(\mathbf{z}_t^{(j)}|\mathbf{Z}_{t-1}^{(j)}, \mathbf{Y}_t)}. \tag{5.19}$$

The SIS, which is a constrained version of importance sampling, is the basis of all the SMC techniques.

5.3.4 SMC framework

Based on the SIS method, we describe a general framework of the SMC methods as follows.

Suppose a set of properly weighted samples $\{(\mathbf{Z}_{t-1}^{(j)}, w_{t-1}^{(j)})\}_{j=1}^m$ (with respect to $p(\mathbf{Z}_{t-1}|\mathbf{Y}_{t-1})$) is given at time $(t-1)$. Based on these previous samples, an SMC algorithm generates a set of properly weighted samples, $\{(\mathbf{Z}_t^{(j)}, w_t^{(j)})\}_{j=1}^m$, at time t with respect to $p(\mathbf{Z}_t|\mathbf{Y}_t)$. It is noted that for most applications we are only able to evaluate $p(\mathbf{Z}_t|\mathbf{Y}_t)$ up to a normalizing constant, which is sufficient for using (5.16) in Monte Carlo estimation. For $j = 1, \cdots, m$, the algorithm is described as follows.

- Draw a sample $\mathbf{z}_t^{(j)}$ from a trial distribution $q(\mathbf{z}_t|\mathbf{Z}_{t-1}^{(j)}, \mathbf{Y}_t)$ and let $\mathbf{Z}_t^{(j)} = (\mathbf{Z}_{t-1}^{(j)}, \mathbf{z}_t^{(j)})$;

- Compute the importance weight according to (5.19).

The algorithm is initialized by drawing a set of i.i.d. samples $\mathbf{z}_0^{(1)}, \ldots, \mathbf{z}_0^{(m)}$ from $p(\mathbf{z}_0|\mathbf{y}_0)$. When $\mathbf{y}_0$ represents the "null" information, $p(\mathbf{z}_0|\mathbf{y}_0)$ corresponds to the prior distribution of $\mathbf{z}_0$. It is proven in [86] that the above algorithm indeed generates properly weighted samples with respect to the distribution $p(\mathbf{Z}_t|\mathbf{Y}_t)$.

5.3.5 Resampling procedure

The importance sampling weight $w_t^{(j)}$ measures the "quality" of the corresponding imputed sequence $\mathbf{Z}_t^{(j)}$. A relatively small weight implies that the sample is drawn far from the main body of the posterior distribution and has a small contribution in the final estimation. Such a sample is said to be ineffective. If there are too many ineffective samples, the Monte Carlo procedure becomes inefficient. To avoid the degeneracy, a useful *resampling* method for reducing ineffective samples and enhancing effective ones was suggested in [116]. Roughly speaking, resampling is to multiply the streams with the larger importance weights, while eliminate the ones with small importance weights. A simple but efficient resampling procedure consists of the following steps.

- Sample a new set of streams $\{\hat{\mathbf{Z}}_t^{(j)}\}_{j=1}^m$ from $\{\mathbf{Z}_t^{(j)}\}_{j=1}^m$ with a probability proportional to the importance weights $\{w_t^{(j)}\}_{j=1}^m$.

- Assign equal weight to each stream in the new samples, i.e., $\hat{w}_t^{(j)} = 1/m$ for $j = 1, \ldots, m$.

Resampling can be implemented at every fixed-length time interval, e.g., every five steps, or it can be conducted dynamically. It is shown in [88] that the samples drawn by the above resampling procedure are also indeed properly weighted with respect to $p(\mathbf{Z}_t|\mathbf{Y}_t)$, provided that m is sufficiently large.

Heuristically, resampling procedure can provide chances for effective sample streams to amplify themselves, hence rejuvenate the sampler to produce a better result for future states as system evolves.

5.4 EM-SMC receiver

In this section, we consider the SMC formulation for multiuser detection and channel estimation of DS-CDMA systems. The applications of SMC to multiuser systems were considered in [90]-[93]. The main problem of these reported work is that the computational complexity of such SMC receivers grows exponentially with the number of inputs. The reason for the high complexity of the conventional SMC methods is that the algorithms make decision for all users at a time, thus the prediction and the update at each step involve $m \times |A|^K$ complexity for computing the trial sample distributions, and for some particle filtering methods, the additional selection step is implemented in $O(m \times |A|^K)$ operations. Therefore, it is difficult to directly use the conventional SMC methods to support a large number of users in DS-CDMA systems.

In this section, a novel receiver framework is proposed to combine the EM algorithm and SMC estimation to avoid the exponential complexity of the conventional SMC methods.

5.4.1 Decomposition of the superimposed signals based on EM algorithm

For developing the new SMC detection scheme with a reasonable complexity, we consider to decompose the multiuser system into K single user systems firstly. Therefore, the EM algorithm is applied here for the decomposition of the superimposed signals.

The expectation maximization algorithm, developed in [117], is a popular numerical method for locating modes of likelihood functions. The basic idea is, rather than directly maximizing the likelihood function of the observed data which are complicated and intractable, specifying the augmented data to simplify the calculation and then performing a series of maximization. Each iteration of the EM consists of two steps: the E-step (expectation) which approximates the augmented data with conditional expectation, and the M-step (maximization) which maximizes the augmented data likelihood.

In [112], the EM method suggests a specific way of decomposing superimposed signals. Based on this method, the observation data is decomposed into their signal components, and then the parameters of each signal component can be estimated separately.

A. Flat fading channels

According to the algorithm, the observation vector given in (5.3) at the output of filter can be decomposed into K components such that

$$\mathbf{r}(n) = \sum_{k=1}^{K} \mathbf{x}_k(n) \tag{5.20}$$

where

$$\mathbf{x}_k(n) = s_k(n)\mathbf{C}_k g_k + \omega_k(n) \tag{5.21}$$

and $\omega_k(n)$ are obtained by arbitrarily decomposing the total noise $\omega(n)$ into K components, that is,

$$\omega(n) = \sum_{k=1}^{K} \omega_k(n). \tag{5.22}$$

One convenient way is to make $\omega_k(n)$ be statistically independent, zero-mean, and Gaussian with a covariance $N_k = \beta_k \sigma^2$, where the weight coefficients $\{\beta_k\}_{k=1}^K$ are real-valued scalars satisfying

$$\sum_{k=1}^{K} \beta_k = 1, \quad \beta_k \geq 0 \tag{5.23}$$

To achieve the maximum convergence rate, the weight coefficients $\{\beta_k\}_{k=1}^K$ are chosen to be equal for different values of k, i.e., $\beta_1 = \cdots = \beta_K = 1/K$ [118].

Then our EM framework assumes the following form for the qth iteration:

E-Step: For $k = 1, 2, \cdots, K$, compute

$$\hat{\mathbf{x}}_k^{(q)}(n) = \hat{s}_k^{(q)}(n)\mathbf{C}_k \hat{g}_k^{(q)} + \beta_k \left[\mathbf{r}(n) - \sum_{l=1}^{K} \hat{s}_l^{(q)}(n)\mathbf{C}_l \hat{g}_l^{(q)} \right] \tag{5.24}$$

M-Step: For $k = 1, 2, \cdots, K$, obtain the ML estimate of $\{s_k^{(q+1)}(n), g_k^{(q+1)}\}$ based on the $\hat{\mathbf{x}}_k^{(q)}(n)$.

B. Frequency-selective fading channels

Similarly, the observation vector given in (5.8) at the output of the filter is decomposed into K components such that

$$\mathbf{r}(n) = \sum_{k=1}^{K} \mathbf{x}_k(n) \tag{5.25}$$

where

$$\mathbf{x}_k(n) = s_k(n)\mathbf{C}_k\mathbf{h}_k + \omega_k(n), \tag{5.26}$$

and $\omega_k(n)$ are obtained by decomposing the total noise $\omega(n)$ into K components, i.e.,

$$\omega(n) = \sum_{k=1}^{K} \omega_k(n). \tag{5.27}$$

Similarly, $\omega_k(n)$ is assumed as statistically independent, zero-mean, and Gaussian with a covariance $N_k = \beta_k\sigma^2$, where the weight coefficients $\{\beta_k\}_{k=1}^{K}$ are chosen to be equal for different values of k, i.e., $\beta_1 = \cdots = \beta_K = 1/K$.

Then the corresponding EM framework has the following form for the qth iteration:

E-Step: For $k = 1, 2, \cdots, K$, compute

$$\hat{\mathbf{x}}_k^{(q)}(n) = \hat{s}_k^{(q)}(n)\mathbf{C}_k\hat{\mathbf{h}}_k^{(q)} + \beta_k \left[\mathbf{r}(n) - \sum_{l=1}^{K} \hat{s}_l^{(q)}(n)\mathbf{C}_l\hat{\mathbf{h}}_l^{(q)}\right] \tag{5.28}$$

M-Step: For $k = 1, 2, \cdots, K$, obtain the ML estimate of $\{s_k^{(q)}(n), \mathbf{h}_k^{(q)}\}$ based on the $\hat{\mathbf{x}}_k^{(q)}(n)$.

As seen from above two frameworks, for every E-step and M-step, there are only one unknown user's parameters to be estimated at a time. Therefore, the complicated multi-parameter optimization is decoupled into K separate optimizations.

5.4.2 Framework of EM-SMC receiver

After the above EM decomposition, the SMC method is used to approach the optimum ML estimations $\{\hat{s}_k^{(q)}(n), \hat{g}_k^{(q)}\}$ or $\{\hat{s}_k^{(q)}(n), \hat{\mathbf{h}}_k^{(q)}\}$ for every user separately. Note that in the conventional SMC approach, K unknown users need to be detected all together to result in a computation complexity exponentially increasing with K. In contrast, the proposed method implements SMC estimation for every unknown user separately and in parallel. This difference is the key leading to the efficient implementation of the proposed SMC approach in terms of both complexity and performance.

The entire framework of this approach is presented in Fig 5.1. The algorithm achieves improvement in the subsequent parameter estimates by using the feedback of the current parameter estimates to more effectively decompose the observed data. $\hat{\theta}_k^{(q)}$ in the feedback path represents the estimates of unknown parameters for user k. For flat fading system, $\hat{\theta}_k^{(q)} = \{\hat{s}_k^{(q)}(n), \hat{q}_k^{(q)}\}$, and for frequency-selective fading system, $\hat{\theta}_k^{(q)} = \{\hat{s}_k^{(q)}(n), \hat{\mathbf{h}}_k^{(q)}\}$. One important property of the computational structure is that $\hat{\theta}_k^{(q)}$ for different values of k can be estimated simultaneously by parallel computation.

5.4.3 SMC detection

Now let us consider the estimation of the unknown parameters $\hat{\theta}_k^{(q)}$ based on each signal component $\hat{\mathbf{x}}_k^{(q)}(n)$. According to the description of SMC method in section 4, we develop a blind SMC estimation algorithm to solve this problem, which combines the SIS for the data detection and the Kalman filter (KF) for the channel estimation.

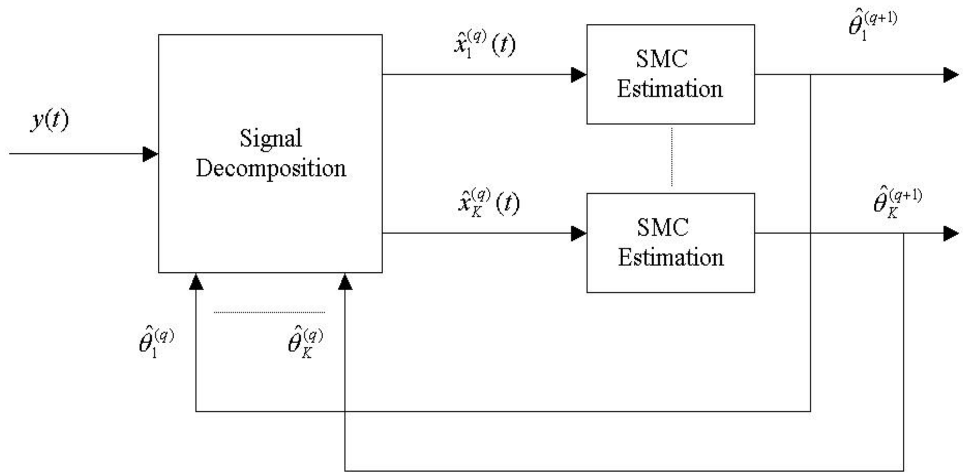


Figure 5.1: The EM-SMC receiver framework

For simplicity of presentation, the hat and superscript of the expressions thereafter are omitted when the systems described in (5.21) and (5.26) are considered. To be convenient, only the model (5.26) is taken in developing the following SMC inference procedure. The results for the model (5.21) are similar and simpler.

Denote $\mathbf{X}_{k,n} \triangleq \{\mathbf{x}_k(0), \mathbf{x}_k(1), \cdots, \mathbf{x}_k(n)\}$ and $\mathbf{D}_{k,n} \triangleq \{d_k(1), \cdots, d_k(n)\}$. The objective is to use SMC method to perform a blind estimate of the *a posteriori* symbol probability

$$P(d_k(n) = a_i | \mathbf{X}_{k,n}), \quad a_i \in A; \quad n = 1, \cdots, N-1 \tag{5.29}$$

based on the observation component $\mathbf{X}_{k,n}$ up to time n and *a priori* symbol probability of $\mathbf{D}_{k,n}$, without knowing the channel response $\mathbf{h}_k$.

Let $\{s_k^{(j)}(n)\}_{j=1}^m$ be samples drawn by the SMC at time n and denote $\mathbf{S}_{k,n}^{(j)} =$

$\{s_k^{(j)}(0), \cdots, s_k^{(j)}(n)\}$. For each value of n, a set of Monte Carlo samples of transmitted symbols, $\{\mathbf{S}_{k,n}^{(j)}, w_{k,n}^{(j)}\}_{j=1}^{m}$, which are properly weighted with respect to the distribution $p(\mathbf{S}_{k,n}|\mathbf{X}_{k,n})$, are to be obtained. For every symbol $a_i \in A$, the *a posteriori* probability of the information symbol $d_k(n)$ can be estimated as

$$\begin{aligned} P(d_k(n) = a_i|\mathbf{X}_{k,n}) &= P(s_k(n)s_k^*(n-1) = a_i|\mathbf{X}_{k,n}) \\ &= E\{1(s_k(n)s_k^*(n-1) = a_i)|\mathbf{X}_{k,n}\} \\ &\cong \frac{1}{W_{k,n}} \sum_{j=1}^{m} 1(s_k^{(j)}(n)s_k^{(j)*}(n-1) = a_i) w_{k,n}^{(j)} \end{aligned} \quad (5.30)$$

where $W_{k,n} \triangleq \sum_{j=1}^{m} w_{k,n}^{(j)}$ and $1(\cdot)$ is an indicator function defined as

$$1(x = a) = \begin{cases} 1, & if\ x = a \\ 0, & if\ x \neq a. \end{cases} \quad (5.31)$$

Based on the general principle of the SMC, the samples $\{s_k^{(j)}(n)\}_{j=1}^{m}$ are drawn from the trial sampling density

$$q(s_k^{(j)}(n)|\mathbf{S}_{k,n-1}^{(j)}, \mathbf{X}_{k,n}) \triangleq p(s_k^{(j)}(n)|\mathbf{S}_{k,n-1}^{(j)}, \mathbf{X}_{k,n}) \quad (5.32)$$

and the importance weight can be updated according to

$$w_{k,n}^{(j)} \propto w_{k,n-1}^{(j)} p(\mathbf{X}_{k,n}|\mathbf{S}_{k,n-1}^{(j)}, \mathbf{X}_{k,n-1}) \quad (5.33)$$

$$= w_{k,n-1}^{(j)} \sum_{a_i \in A} p(\mathbf{x}_k(n)|\mathbf{S}_{k,n-1}^{(j)}, s_k(n) = a_i, \mathbf{X}_{k,n-1}) P(s_k(n) = a_i|\mathbf{S}_{k,n-1}^{(j)}, \mathbf{X}_{k,n-1})$$

$$= w_{k,n-1}^{(j)} \sum_{a_i \in A} \alpha_{k,n,i}^{(j)}. \quad (5.34)$$

The derivation of (5.33) is found in Appendix. To compute the predictive density $p(\mathbf{x}_k(n)|\mathbf{S}^{(j)}_{k,n-1}, s_k(n) = a_i, \mathbf{X}_{k,n-1})$, a Gaussian distribution is assigned to the channel $\mathbf{h}_k$, i.e.,

$$\mathbf{h}_k \sim \mathcal{N}_c(\bar{\mathbf{h}}_k, \bar{\mathbf{\Sigma}}_k). \tag{5.35}$$

Then, the distribution of $\mathbf{h}_k$, conditioned on $\mathbf{S}^{(j)}_{k,n}$ and $\mathbf{X}_{k,n}$, can be computed as

$$\begin{aligned} p(\mathbf{h}_k|\mathbf{S}^{(j)}_{k,n}, \mathbf{X}_{k,n}) &\propto p(\mathbf{X}_{k,n}|\mathbf{S}^{(j)}_{k,n}, \mathbf{h}_k)p(\mathbf{h}_k) \\ &\sim \mathcal{N}_c(\mathbf{h}^{(j)}_{k,n}, \mathbf{\Sigma}^{(j)}_{k,n}) \end{aligned} \tag{5.36}$$

where

$$\mathbf{h}^{(j)}_{k,n} \triangleq \mathbf{\Sigma}^{(j)}_{k,n}\left[\bar{\mathbf{\Sigma}}^{-1}_k\bar{\mathbf{h}}_k + \frac{1}{N_k}\sum_{i=0}^{n}\mathbf{\Psi}^{(j)H}_{k,i}\mathbf{x}_k(i)\right] \tag{5.37}$$

$$\mathbf{\Sigma}^{(j)}_{k,n} \triangleq \left[\bar{\mathbf{\Sigma}}^{-1}_k + \frac{1}{N_k}\sum_{i=0}^{n}\mathbf{\Psi}^{(j)H}_{k,i}\mathbf{\Psi}^{(j)}_{k,i}\right]^{-1} \tag{5.38}$$

and

$$\mathbf{\Psi}^{(j)}_{k,i} = s^{(j)}_k(i)\mathbf{C}_k. \tag{5.39}$$

Hence, the conditional density $p(\mathbf{x}_k(n)|\mathbf{S}^{(j)}_{k,n-1}, s_k(n) = a_i, \mathbf{X}_{k,n-1})$ is given by

$$\begin{aligned} &p(\mathbf{x}_k(n)|\mathbf{S}^{(j)}_{k,n-1}, s_k(n) = a_i, \mathbf{X}_{k,n-1}) \\ &= \int p(\mathbf{x}_k(n)|\mathbf{S}^{(j)}_{k,n-1}, s_k(n) = a_i, \mathbf{X}_{k,n-1}, \mathbf{h}_k)p(\mathbf{h}_k|\mathbf{S}^{(j)}_{k,n-1}, \mathbf{X}_{k,n-1})d\mathbf{h}_k. \end{aligned} \tag{5.40}$$

Because it is an integral of a Gaussian probability density function (pdf) with respect to another Gaussian pdf, the resulting pdf is still Gaussian, i.e.,

$$p(\mathbf{x}_k(n)|\mathbf{S}^{(j)}_{k,n-1}, s_k(n) = a_i, \mathbf{X}_{k,n-1}) \sim \mathcal{N}_c(\mu^{(j)}_{k,n,i}, \mathbf{\Theta}^{(j)}_{k,n,i}) \tag{5.41}$$

with a mean

$$\begin{aligned}\mu_{k,n,i}^{(j)} &\triangleq E\left\{\mathbf{x}_k(n)|\mathbf{S}_{k,n-1}^{(j)}, s_k(n) = a_i, \mathbf{X}_{k,n-1}\right\} \\ &= \mathbf{\Phi}_{k,i}\mathbf{h}_{k,n-1}^{(j)} \end{aligned} \tag{5.42}$$

and a covariance

$$\begin{aligned}\mathbf{\Theta}_{k,n,i}^{(j)} &\triangleq Cov\left\{\mathbf{x}_k(n)|\mathbf{S}_{k,n-1}^{(j)}, s_k(n) = a_i, \mathbf{X}_{k,n-1}\right\} \\ &= N_k\mathbf{I}_p + \mathbf{\Phi}_{k,i}\mathbf{\Sigma}_{k,n-1}^{(j)}\mathbf{\Phi}_{k,i}^H \end{aligned} \tag{5.43}$$

where

$$\mathbf{\Phi}_{k,i} = a_i\mathbf{C}_k. \tag{5.44}$$

Then, $\alpha_{k,n,i}^{(j)}$ in (5.34) can be computed by

$$\begin{aligned}\alpha_{k,n,i}^{(j)} = &|\mathbf{\Theta}_{k,n,i}^{(j)}|^{-1}exp\left\{-(\mathbf{x}_k(n) - \mu_{k,n,i}^{(j)})^H(\mathbf{\Theta}_{k,n,i}^{(j)})^{-1}(\mathbf{x}_k(n) - \mu_{k,n,i}^{(j)})\right\} \\ &\times P(s_k(n) = a_i|\mathbf{S}_{k,n-1}^{(j)}, \mathbf{X}_{k,n-1}) \\ = &|\mathbf{\Theta}_{k,n,i}^{(j)}|^{-1}exp\left\{-(\mathbf{x}_k(n) - \mu_{k,n,i}^{(j)})^H(\mathbf{\Theta}_{k,n,i}^{(j)})^{-1}(\mathbf{x}_k(n) - \mu_{k,n,i}^{(j)})\right\} \\ &\times P(d_k(n) = a_i s_k^{(j)*}(n-1)) \end{aligned} \tag{5.45}$$

which holds because $s_k(n)$ is independent of $\mathbf{X}_{k,n-1}$ given $\mathbf{S}_{k,n-1}^{(j)}$, and $\{s_k(n)\}$ is a first-order Markov chain due to the differential encoding rule. $P(d_k(n) = a_i s_k^{(j)*}(n-1))$ is the *a priori* probability of the unknown symbol.

The trial distribution in (5.32) can be computed as

$$\begin{aligned}
&p(s_k^{(j)}(n)|\mathbf{S}_{k,n-1}^{(j)}, \mathbf{X}_{k,n}) \\
&\quad \propto p(\mathbf{x}_k(n)|\mathbf{S}_{k,n-1}^{(j)}, s_k^{(j)}(n) = a_i, \mathbf{X}_{k,n-1}) P(s_k^{(j)}(n) = a_i|\mathbf{S}_{k,n-1}^{(j)}, \mathbf{X}_{k,n-1}) \\
&\quad = \alpha_{k,n,i}^{(j)}. \qquad (5.46)
\end{aligned}$$

It is noted that the *a posteriori* mean and covariance of the channel in (5.37) and (5.38) can be updated recursively by Kalman filter as follows. At the nth step, the new sample of $s_k^{(j)}(n)$ and the past samples $\mathbf{S}_{k,n-1}^{(j)}$ are combined to form $\mathbf{S}_{k,n}^{(j)}$. Let $\mu_{k,n}^{(j)}$ and $\mathbf{\Theta}_{k,n}^{(j)}$ be the quantities computed by (5.42) and (5.43) for the imputed $s_k^{(j)}(n)$. Based on a matrix inversion lemma, (5.37) and (5.38) become

$$\mathbf{h}_{k,n}^{(j)} = \mathbf{h}_{k,n-1}^{(j)} + \mathbf{\Omega}_{k,n}^{(j)}(\mathbf{\Theta}_{k,n}^{(j)})^{-1}(\mathbf{x}_k(n) - \mu_{k,n}^{(j)}) \qquad (5.47)$$

$$\mathbf{\Sigma}_{k,n}^{(j)} = \mathbf{\Sigma}_{k,n-1}^{(j)} - \mathbf{\Omega}_{k,n}^{(j)}(\mathbf{\Theta}_{k,n}^{(j)})^{-1}\mathbf{\Omega}_{k,n}^{(j)H} \qquad (5.48)$$

with

$$\mathbf{\Omega}_{k,n}^{(j)} = \mathbf{\Sigma}_{k,n-1}^{(j)}\mathbf{\Psi}_{k,n}^{(j)H}. \qquad (5.49)$$

Finally, the SMC blind detector for each decomposed signal component is given as follows.

FOR $k = 1, \cdots, K$,

Initialization:

For $j = 1, \cdots, m$

- Set the initial values of channel vector as $\mathbf{h}_k^{(j)} \sim \mathcal{N}_c(\mathbf{0}, 100\mathbf{I}_L)$;
- Initial all importance weights as $w_{k,-1}^{(j)} = 1$.

End

Estimation:

For $n = 0, \cdots, N-1$

I. Update weighted samples

For $j = 1, \cdots, m$

1. For each $a_i \in A$

   - Compute $\mu_{k,n,i}^{(j)}$, $\mathbf{\Theta}_{k,n,i}^{(j)}$ according to (5.42) and (5.43);
   - Compute trial sampling distribution $\alpha_{k,n,i}^{(j)}$ according to (5.45).

   End

2. Draw a sample $s_k^{(j)}(n)$ according to the trial sampling distribution (5.46).

3. Compute the importance weight

$$\hat{w}_{k,n}^{(j)} = w_{k,n-1}^{(j)} \sum_{a_i \in A} \alpha_{k,n,i}^{(j)}; \tag{5.50}$$

   Normalize as

$$w_{k,n}^{(j)} = \frac{\hat{w}_{k,n}^{(j)}}{\sum_{j=1}^{m} \hat{w}_{k,n}^{(j)}}. \tag{5.51}$$

4. If the imputed samples $s_k^{(j)}(n) = a_i$

   - Set $\mu_{k,n}^{(j)} = \mu_{k,n,i}^{(j)}, \mathbf{\Theta}_{k,n}^{(j)} = \mathbf{\Theta}_{k,n,i}^{(j)}$;
   - Update the *a posteriori* mean $\mathbf{h}_{k,n}^{(j)}$ and covariance $\mathbf{\Sigma}_{k,n}^{(j)}$ of the

channel according to (5.47) and (5.48).

End

5. Compute the *a posteriori* probability of the information symbol $d_k(n)$ according to (5.30).

End (for **I**)

II. Resampling

If n is a multiple of the resampling interval

For $j = 1, \cdots, m$

- Draw a new set of $\{\mathbf{S}_{k,n}^{(j)}, \mathbf{h}_{k,n}^{(j)}, \mathbf{\Sigma}_{k,n}^{(j)}\}_{j=1}^{m}$ from the original set with a probability proportional to the importance weights $\{w_{k,n}^{(j)}\}_{j=1}^{m}$;
- Assign equal weight for each new samples, i.e., $\hat{w}_{k,n}^{(j)} = 1/m$.

End

End (for **II**)

End

END

It is observed that, for each signal component, i.e., each user in our system, the dominant computation for the SMC receiver needs $N \times m \times |A|$ one-step predictions for computing $\alpha_{k,n,i}^{(j)}$ and $N \times m$ one-step Kalman filter updates for $\{(\mathbf{h}_{k,n}^{(j)}, \mathbf{\Sigma}_{k,n}^{(j)})\}$. Since the m samples operate independently, the proposed SMC estimation is well suited for parallel implementation.

5.4.4 Computational complexity

For the entire detection algorithm, it is known that the computational complexity of the conventional SMC method is in the order of $O(|A|^K)$, because the dominant computation for computing trial distribution is in the order of $N \times m \times |A|^K$. However, the computational complexity required by the proposed one is $Q \times O(|A| \times K)$, where Q is the number of the EM iterations. When the number of users is large, the total computational complexity of the proposed method is obviously much lower than that of the conventional SMC method. The simulations in the next section will demonstrate that the value of required Q is a small number. Thus, the proposed EM-SMC algorithm can be applied to the DS-CDMA system with a manageable computational complexity.

5.5 Extension to system with long codes

The above EM-SMC receiver is developed for the system with short spreading codes. As we see, the bit-interval cyclostationarity properties of the short codes are not necessary for this SMC-based method and the intersymbol interference (ISI) is negligible for the system model. Therefore, this method can be extended to the long code systems by changing the spreading waveform (5.2) to

$$\bar{c}_k^n(t) = \sum_{p=0}^{P-1} \bar{c}_k^n(p)\phi(t - pT_c). \tag{5.52}$$

Correspondingly, the spreading vector/matrix $\mathbf{C}_k$ should be changed to $\mathbf{C}_k(n)$, whose elements are substituted from $\{\bar{c}_k(p), p = 0, 1, \cdots, P-1\}$ to $\{\bar{c}_k^n(p), p = 0, 1, \cdots, P-1\}$. Then, the proposed algorithm can be accommodated to the

CDMA system employing long spreading codes.

5.6 Simulation results

This section provides simulation results to illustrate the performance of the blind EM-SMC multiuser receiver in both flat fading channels and frequency-selective fading channels for DS-CDMA system. The channels are assumed to be block fading, that is the fading coefficients remain constant over the entire block of N symbols. And the fading coefficients of the channels are generated according to uncorrelated circular complex Gaussian distribution. All the users' spreading sequences are chosen as short sequence with a processing gain $P = 10$, and are generated randomly from equal probability binary code and then normalized to $\{\pm 1/\sqrt{P}\}$. A rate 1/2 constraint length-5 convolutional code (with generator 23 and 35 in octal notation) is employed. We choose 128 information bits, i.e., the coded bit block size is $N = 256$. Two types of channels are considered, i.e, flat fading channel and frequency-selective fading channel with $L = 3$.

5.6.1 Convergence

Let us first study the convergence of the proposed algorithm with the iterations of the EM signal decomposition. The number of Monte Carlo samples is chosen as $m = 50$, and the number of the users is $K = 8$. The performance of the proposed receiver in terms of the BERs versus signal-to-noise ratios (SNR) for different numbers of iterations is shown in Fig 5.2 and Fig 5.3. Fig 5.2 shows the performance of the system over the flat fading channel, while Fig 5.3 shows the performance of the system over the frequency-selective fading channel. It is observed from both the figures that the BER performance can be improved when the number of iterations increases. However, little gain can be obtained after 4 iterations, i.e., the receiver reaches the convergence with only several iterations. The good capability of convergence ensures the superiority of the proposed EM-SMC to the conventional SMC method in terms of computational complexity. In the following simulations, the number of EM steps is taken as 6.

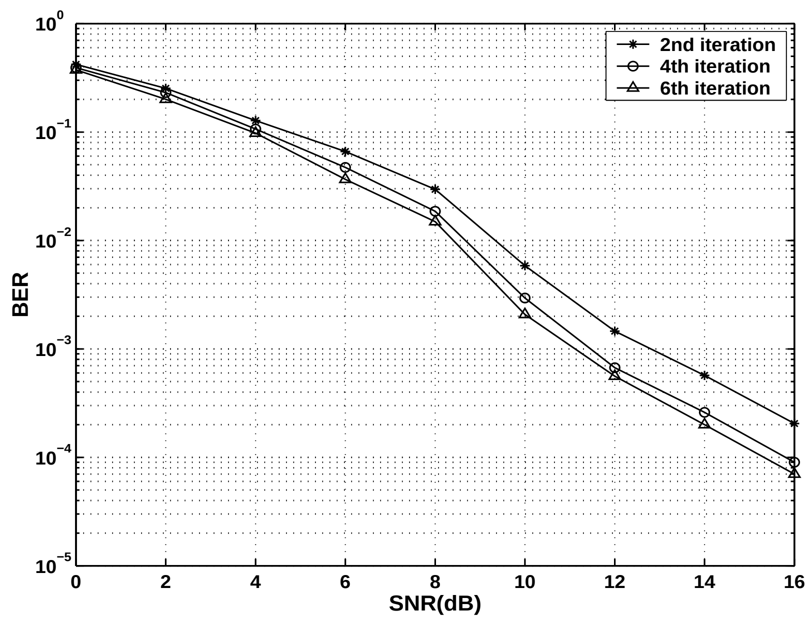


Figure 5.2: BER versus SNR for various number of iterations (flat fading channel)

5.6.2 Selection of the number of Monte Carlo samples

Next we consider the selection for the number of Monte Carlo samples. Fig 5.4 and Fig 5.5 present the BER performance as a function of the number of Monte Carlo samples for the system with $SNR = 10$, 12 and 15 dB, respectively. As illustrated in the figures, for both flat fading and frequency-selective fading

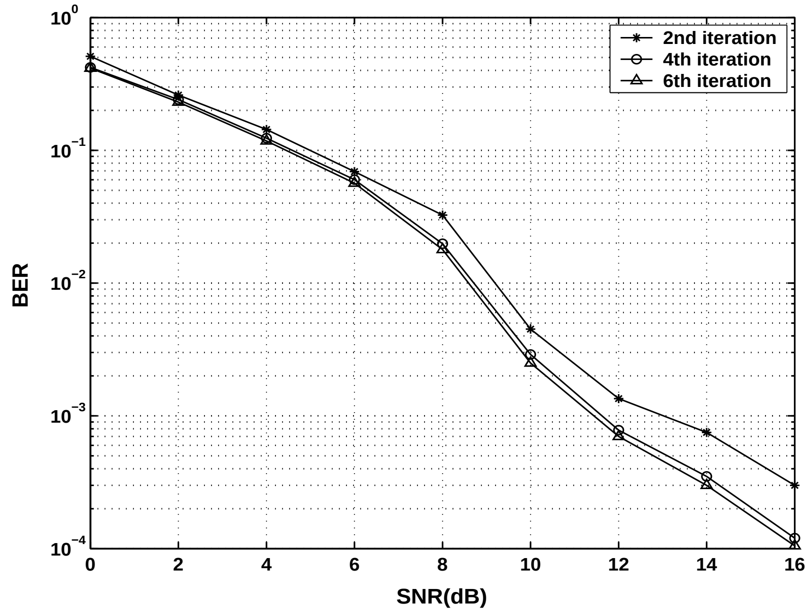


Figure 5.3: BER versus SNR for various number of iterations (frequency-selective fading channel)

systems, the BERs decrease as the number of samples increases and achieve convergence when m reaches about 50. In the following simulations, therefore, the number of Monte Carlo samples is taken as $m = 50$.

5.6.3 BER performance

Let us now test the BER performance of the proposed method with different number of users. The BER performances as the function of SNR are tested for the proposed detector with $K = 8$, 10 and 12. Fig 5.6 and Fig 5.7 plot the results of systems in the flat fading channel and frequency-selective fading channel, respectively.

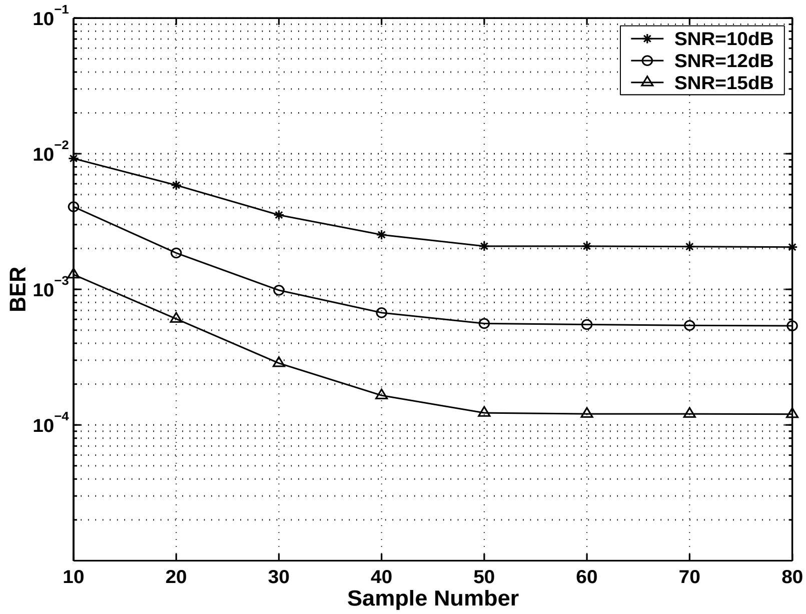


Figure 5.4: BER versus the number of Monte Carlo samples, $K = 8$ (flat fading channel)

5.6.4 Near-far resistance

The capability of near-far resistance is illustrated in Fig 5.8 for the system over flat fading channel, and in Fig 5.9 for the system over frequency-selective fading channel. The near-far ratio is defined as the ratio between the power of interfering users and the power of the desired user. Let us fix the power of the desired user and change the power of interfering users. It is assumed that all interfering users have the same power. The BER performance is tested as the function of near-far ratio for the systems with $K = 8$, 10 and 12. The results in Fig 5.8 and Fig 5.9 demonstrate that the proposed receiver achieves a very good near-far resistance.

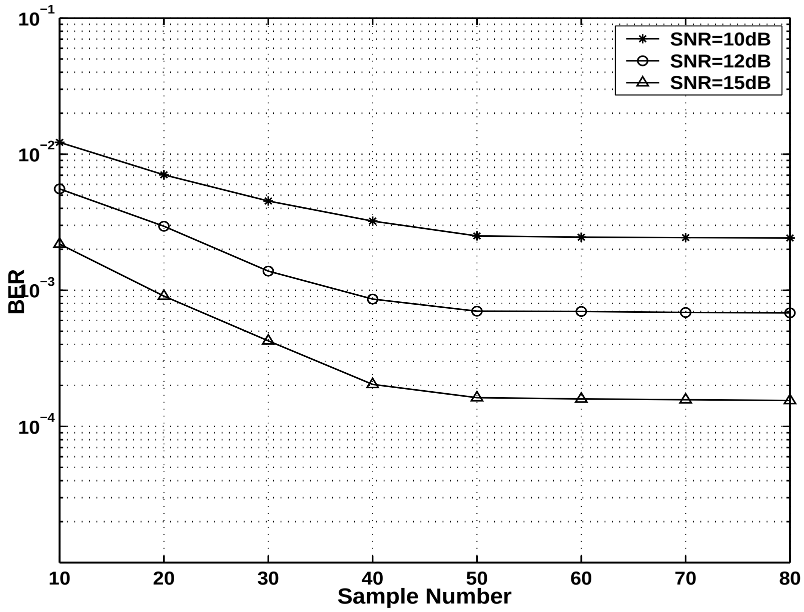


Figure 5.5: BER versus the number of Monte Carlo samples, $K = 8$ (frequency-selective fading channel)

5.6.5 Comparisons of BER performance

Finally, comparisons are made on the performances achieved by the proposed receiver and other reported receivers. Fig 5.10 and Fig 5.11 present the performances achieved by the Gibbs sampler [84], QRD-M-EKF [119], particle filtering (PF) [95], the proposed EM-SMC receiver and the conventional SMC receiver [90]. The Gibbs sampler is a kind of Bayesian MCMC methodologies described in [84]. QRD-M-EKF is a deterministic method developed in [119], which is based on EKF and QRD-M algorithm.

For a fair comparison, the Gibbs sample of MCMC receiver is performed for 100 iterations with the first 50 iterations as the burning-in period. The number of Monte Carlo samples used for both the PF method [95] and the conventional SMC method [90] is 50, which is the same as that used for the proposed receiver. The number of paths M for the tree-search detection QRD-M-EKF is selected

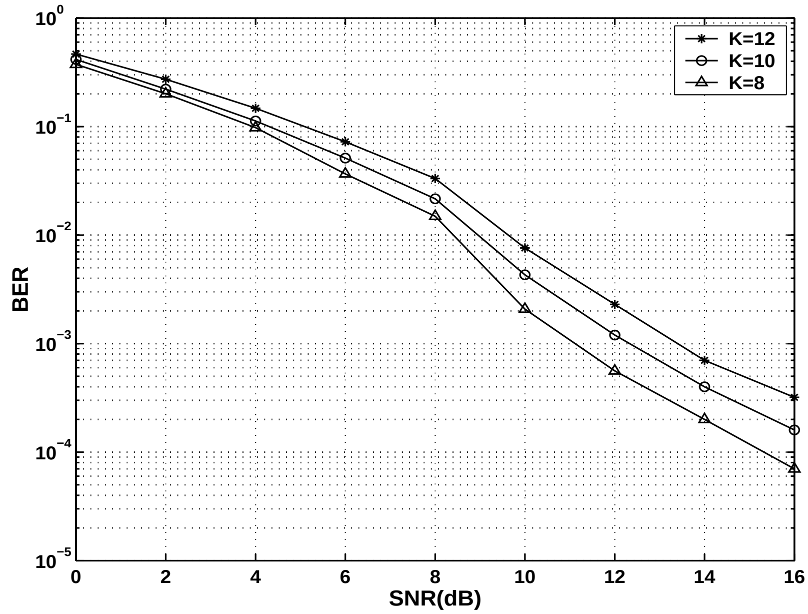


Figure 5.6: BER versus the SNR (flat fading channel)

to be 32. The results for the flat fading system are shown in Fig 5.10, while the results for the frenquency-selective fading system are shown in Fig 5.11.

Fig 5.10 and Fig 5.11 show that the proposed receiver achieves the best performance among the methods which have the similar complexity. Fig 5.10 and Fig 5.11 also show that the performance of proposed receiver is close to conventional SMC receiver with a little inferiority. The simulation results demonstrate that the proposed EM-SMC detector achieves nearly the same performance as that achieved by the conventional SMC detector with a much lower computational complexity, and also outperforms the other deterministic and MC sampling based detectors with the similar computational complexity.

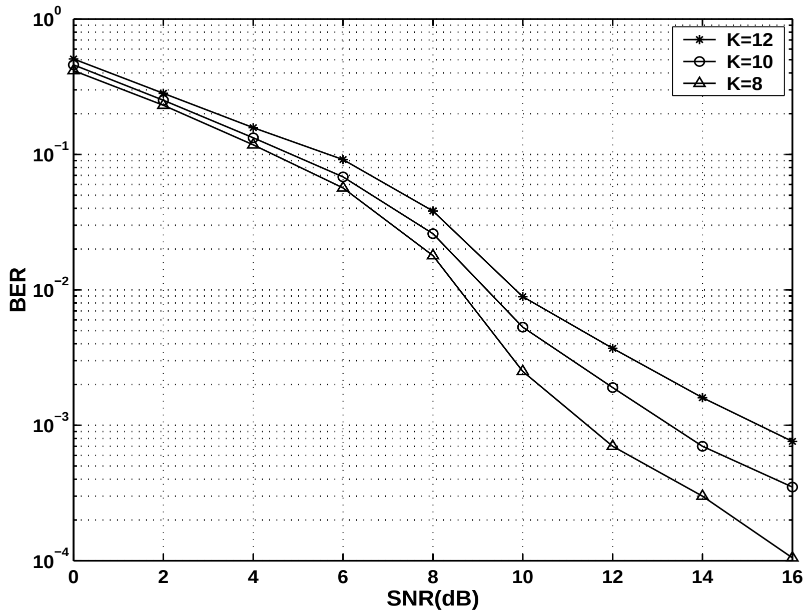


Figure 5.7: BER versus the SNR (frequency-selective fading channel)

5.7 Conclusions

In this chapter, a blind multiuser receiver is developed based on the EM framework to decompose the observation signals. The methods of Bayesian SMC estimation are applied to detect the unknown transmitted symbols according to the decomposed signal components. With a substantially reduced computational complexity, the proposed EM-SMC receiver outperforms other methods with similar complexity and achieves the comparable performance as that obtained by the conventional SMC detector. With the parallel computational property, the proposed receiver is proved to be well suited to DS-CDMA multiuser systems.

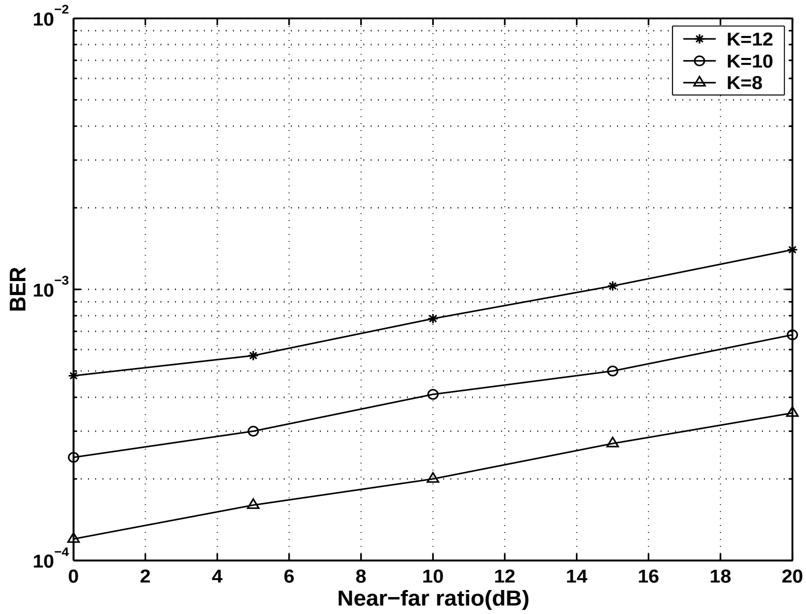


Figure 5.8: BER versus the near-far ratio, $SNR = 15\ dB$ (flat fading channel)

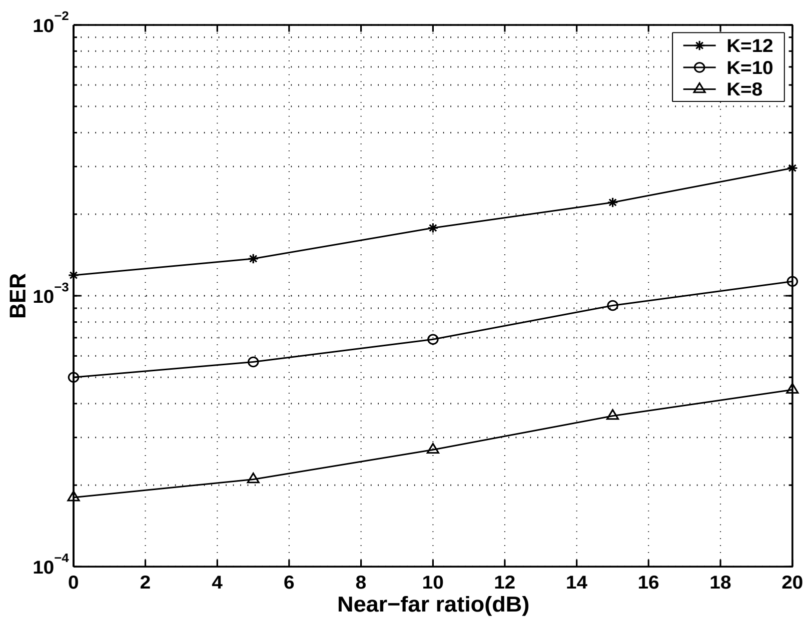


Figure 5.9: BER versus the near-far ratio, $SNR = 15\ dB$ (frequency-selective fading channel)

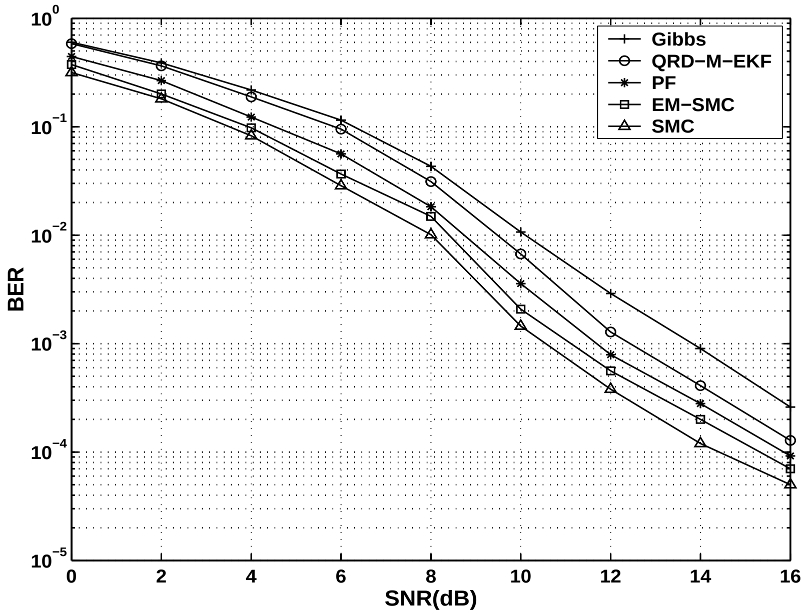


Figure 5.10: Comparisons on BER performance, $K = 8$ (flat fading channel)

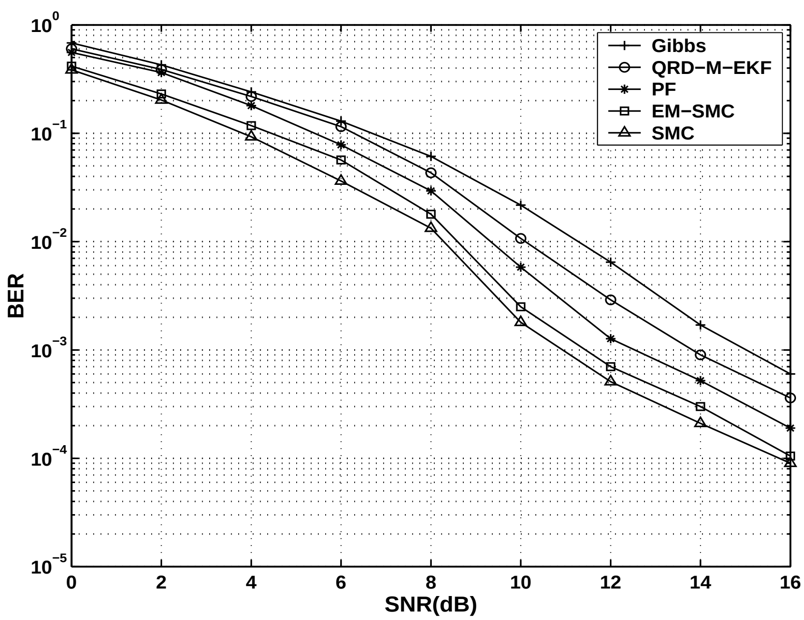


Figure 5.11: Comparisons on BER performance, $K = 8$ (frequency-selective fading channel)

Chapter 6

Blind SMC multiuser detection based on decision feedback

6.1 Introduction

As discussed in the previous chapter, conventional SMC method is inefficient for multiuser detection of DS-CDMA system with a large number of users. The reason for the high computational complexity of conventional SMC method is that the binary data of all users in a symbol interval are considered as a super symbol by using the conventional representation of the system model. Thus we can reduce the computational complexity by finding the decomposed representation of the system model.

Chapter 5 proposed a low complexity SMC-based receiver, EM-SMC, to the multiuser detection of DS-CDMA system, which uses the EM algorithm to decompose the system into separate single user systems, before the SMC method is applied to perform the parameter estimation in parallel. With the application of the proposed EM-SMC receiver, the exponential computational complexity is reduced to be linear with the number of users. However, since the algorithm is based on the EM framework which decomposes the system approximately, the detection of EM-SMC needs to be iterated until the convergence is accomplished. Thus, the whole computational complexity is the product of the complexity of one user detection multiplied by not only the number of users but also the number of the EM iterations.

In order to reduce the computational complexity further while keeping the competitive performance, system decomposition without approximation is necessary. Based on the similar consideration, an SMC demodulation algorithm is proposed in [111] by utilizing the existing BLAST detection scheme which is based on the QR decomposition. However, only flat fading channels are considered in this algorithm, and the channel parameters need to be estimated before detection through the training sequences, thus it's not a blind method.

In this chapter, we derive another new SMC-based formulation to blind multiuser detection of DS-CDMA systems in both flat fading and frequency selective fading channels. A novel transformation of system models is implemented before the SMC estimation, which is based on the Cholesky factorization of the cross-correlation. We make use of the Cholesky factorization algorithm to decompose the observed data into the separate signals according to the number of users. Compared to the QR decomposition used in [111], there is no need to compute the inverse of the cross-correlation matrix for Cholesky decomposition, thus the computation complexity is effectively reduced. Then under the decision-feedback framework, the parameters of each user are estimated by SMC method sequentially. The new detection algorithm samples one user at a time and therefore permits efficient implementation that reduces the computational complexity associated with the SMC inference. Because there is no need for iterations, much lower complexity can be achieved compared to the EM-SMC. The computational complexity of the new SMC detection is in the order of $O(|A|K)$, which is the product of the complexity of one user detection multiplied by only the number of users.

Based on these concepts, an efficient blind DF-SMC receiver is developed. Simulation results are presented to show that the receiver performs well for both flat fading channels and frequency selective fading channels with a further reduced computational complexity.

6.2 Signal model

6.2.1 Flat fading channels

Let us consider a DS-CDMA system that has K active users whose signals are transmitted over flat fading channel with additive Gaussian noise. Let T denote the symbol duration and $s_k(t)$ the normalized spreading waveform assigned to the kth user. Then the received signal $r(t)$ at the nth symbol interval is given by

$$r(t) = \sum_{k=1}^{K} g_k b_k(n) s_k(t) + \omega(t). \tag{6.1}$$

It is assumed that the transmitted symbols are independent, and taken from a finite alphabet set $\mathbf{A} = \{a_1, \cdots, a_i, \cdots, a_{|A|}\}$. For each user $k = 1, \cdots, K$, the transmitted symbols $\{b_k(n)\}_{n=0}^{N-1}$ are differentially modulated from the source information symbols $\{d_k(n)\}_{n=1}^{N-1}$, g_k is the fading coefficient of the kth user's channel, and $\omega(t)$ is the received zero mean additive complex white Gaussian noise with variance σ^2.

The cross-correlation between the signature waveforms of the users is given by the cross-correlation matrix $\mathbf{R}$, where the element $\mathbf{R}_{i,j}$ represents the cross-correlation between the signature waveforms of the ith and jth users. The $\mathbf{R}_{i,j}$ is defined according to

$$\mathbf{R}_{i,j} =< s_i, s_j >= \int_{(n-1)T}^{nT} s_i(t)s_j(t)dt. \tag{6.2}$$

We process the received signal with a bank of matched filters, then the set of matched filter outputs $\mathbf{y}(n) = [y_1(n), \cdots, y_K(n)]^T$ is obtained, where

$$y_k(n) =< r(t), s_k(t) >= \int_{(n-1)T}^{nT} r(t)s_k(t)dt \tag{6.3}$$

It is convenient to express the $K \times 1$ vector $\mathbf{y}(n)$ in the form as:

$$\mathbf{y}(n) = \mathbf{RHb}(n) + \mathbf{w}(n) \tag{6.4}$$

where $\mathbf{H} = diag\{g_1, \cdots, g_K\}$ is the $K \times K$ diagonal matrix of the channel state information, $\mathbf{b}(n) = [b_1(n), \cdots, b_K(n)]^T$ is the data vector, and $\mathbf{w}(n)$ is the $K \times 1$ complex Gaussian noise vector with covariance matrix equal to $\sigma^2\mathbf{R}$.

Let $\mathbf{Y} = \{\mathbf{y}(n)\}_{n=0}^{N-1}$, and $\mathbf{d}(n) = \{d_k(n)\}_{k=1}^{K}$. Our objective is to estimate the *a posteriori* probabilities of the information symbols

$$P(\mathbf{d}(n) = \mathbf{a}_i|\mathbf{Y}), \mathbf{a}_i \in \mathbf{A}^K, n = 1, \cdots, N-1 \tag{6.5}$$

based on the received signals $\mathbf{Y}$ without knowing the channel information $\mathbf{H}$.

6.2.2 Frequency-selective fading channels

Now we consider the DS-CDMA system with the same signals which transmit through the frequency-selective fading channels. That is, the transmitted signal of kth user at nth symbol interval is

$$x_k(t) = b_k(n)s_k(t), \tag{6.6}$$

and the multipath channel is modeled by

$$h_k(t) = \sum_{l=1}^{L} g_{k,l}\delta(t - \tau_{k,l}) \tag{6.7}$$

where L is the number of paths in each user's channel, $g_{k,l}$ and $\tau_{k,l}$ are, respectively, the complex gain and delay of lth path of the kth user's signal. Then the total received signal at the receiver is the superposition of the signals from the K users plus the additive ambient noise given by

$$\begin{aligned} r(t) &= \sum_{k=1}^{K} x_k(t) * h_k(t) + \omega(t) \\ &= \sum_{k=1}^{K} b_k(n) \sum_{l=1}^{L} g_{k,l} s_k(t - \tau_{k,l}) + \omega(t) \end{aligned} \tag{6.8}$$

where $*$ denotes the convolution, and $\omega(t)$ is zero mean additive complex white Gaussian noise with variance σ^2.

The received signals are processed by a bank of matched filters for each path of each user to generate the observation vector expressed as

$$\begin{aligned} y_{k,l}(n) &= < r(t), s_k(t-\tau_{k,l}) >= \int_{(n-1)T}^{nT} r(t) s_k(t-\tau_{k,l}) \\ &= \sum_{k'=1}^{K} b_{k'}(n) \sum_{l'=1}^{L} g_{k',l'} \rho_{(k,l)(k',l')} + w_{k,l}(n) \end{aligned} \tag{6.9}$$

where $\rho_{(k,l)(k',l')}$ is defined as the correlation between the spreading waveforms of the kth user's lth path and the k'th user's l'th path.

$$\rho_{(k,l)(k',l')} = \int_{(n-1)T}^{nT} s_k(t-\tau_{k,l}) s_{k'}(t-\tau_{k',l'}) dt \tag{6.10}$$

We denote the set of the matched filter outputs as the KL vector $\mathbf{y}(n)$, that is,

$$\mathbf{y}(n) = [y_{1,1}(n), \cdots, y_{1,L}(n), \cdots, y_{K,1}(n), \cdots, y_{K,L}(n)]^T \tag{6.11}$$

and the correlation matrix as $KL \times KL$ matrix $\mathbf{R}$, i.e.,

$$\mathbf{R} = \begin{bmatrix} \rho_{(1,1)(1,1)} & \cdots & \rho_{(1,1)(1,L)} & \cdots & \rho_{(1,1)(K,1)} & \cdots & \rho_{(1,1)(K,L)} \\ \rho_{(2,1)(1,1)} & \cdots & \rho_{(2,1)(1,L)} & \cdots & \rho_{(2,1)(K,1)} & \cdots & \rho_{(2,1)(K,L)} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{(K,L)(1,1)} & \cdots & \rho_{(K,L)(1,L)} & \cdots & \rho_{(K,L)(K,1)} & \cdots & \rho_{(K,L)(K,L)} \end{bmatrix}. \tag{6.12}$$

Then we obtain the expression for the observation vector as:

$$\mathbf{y}(n) = \mathbf{RHb}(n) + \mathbf{w}(n) \tag{6.13}$$

where $\mathbf{b}(n) = [b_1(n), \cdots, b_K(n)]^T$ is the data vector, and the $KL \times K$ matrix $\mathbf{H}$ is the channel response information, expressed by $\mathbf{H} = diag\{\mathbf{g}_1, \cdots, \mathbf{g}_K\}$ with $\mathbf{g}_k = [g_{k,1}, \cdots, g_{k,L}]^T$.

The noise term $\mathbf{w}(n) = [w_{1,1}(n), \cdots, w_{1,L}(n), \cdots, w_{K,1}(n), \cdots, w_{K,L}(n)]^T$ is the KL complex Gaussian vector with a zero-mean and a covariance matrix $\sigma^2\mathbf{R}$. Again, without knowing the channel response $\mathbf{H}$, the *a posteriori* probabilities of the information symbols $\{d_k(n)\}_{n=1:N-1}^{k=1:K}$ are to be estimated based on the received signals $\{\mathbf{y}(n)\}_{n=0}^{N-1}$ and the *a priori* information symbol probabilities.

6.3 Cholesky factorization

Now let us consider the problem of blind multiuser detection for the systems described by (6.4) and (6.13). Before the derivation of the SMC-based formulation for the detection problem, we first introduce the theory of the Cholesky factorization.

Cholesky factorization is one of the most efficient techniques for the solution of linear system equations. We describe the corresponding proposition here.

Proposition: *For every positive definite matrix* $\mathbf{B}$, *there exists a unique lower triangular matrix* $\mathbf{L}$ (i.e., $L_{i,k} = 0$ for $i < k$) *with positive diagonal elements such that*

$$\mathbf{B} = \mathbf{L}^H\mathbf{L}. \tag{6.14}$$

It means that Cholesky factorization decomposes the positive-definite Matrix $\mathbf{B}$ into a lower triangular matrix $\mathbf{L}$ and the conjugate transpose of the lower triangular matrix $\mathbf{L}^H$.

In developing the SMC-based detection method, we start with the Cholesky factorization of the cross-correlation matrix $\mathbf{R}$ as

$$\mathbf{R} = \mathbf{F}^H \mathbf{F} \tag{6.15}$$

Next, the matched filter outputs $\mathbf{y}(n)$ are processed by multiplying the matrix $(\mathbf{F}^H)^{-1}$, we obtain

$$\begin{aligned}\bar{\mathbf{y}}(n) &= (\mathbf{F}^H)^{-1}\mathbf{y}(n) \\ &= (\mathbf{F}^H)^{-1}\mathbf{RHb}(n) + (\mathbf{F}^H)^{-1}\mathbf{w}(n) \\ &= \mathbf{FHb}(n) + \bar{\mathbf{w}}(n)\end{aligned} \tag{6.16}$$

which is called the output of the *whitened matched filter*. Because there is a one-to-one correspondence between $\bar{\mathbf{y}}(n)$ and $\mathbf{y}(n)$, both models contain the same information about the data. The covariance matrix of $\bar{\mathbf{w}}(n)$ is

$$E[\bar{\mathbf{w}}(n)\bar{\mathbf{w}}(n)^H] = \sigma^2 (\mathbf{F}^H)^{-1}\mathbf{R}\mathbf{F}^{-1} = \sigma^2 \mathbf{I} \tag{6.17}$$

For the flat fading channel system (6.4), $\mathbf{I}$ is a $K \times K$ identity matrix, and for the frequency-selective fading channel system (6.13), the $\mathbf{I}$ is a $KL \times KL$ identity matrix.

6.4 Decomposition of the models

As seen from the model (6.16), the matrix $\mathbf{F}$ is lower triangular matrix, therefore the system can be decomposed into the components for each user as follows:

A. Flat fading channels

$$x_1(n) = \bar{y}_1(n) = \mathbf{F}_{1,1} g_1 b_1(n) + \bar{w}_1(n) \tag{6.18}$$

$$for \;\; k = 2 : K$$

$$x_k(n) = \bar{y}_k(n) - \sum_{i=1}^{k-1} \mathbf{F}_{k,i} g_i b_i(n) = \mathbf{F}_{k,k} g_k b_k(n) + \bar{w}_k(n) \tag{6.19}$$

$$end$$

where $\mathbf{F}_{i,j}$ is the element of the ith row and jth column for the matrix $\mathbf{F}$, $\bar{y}_i(n)$ and $\bar{w}_i(n)$ are the ith elements of the vector $\bar{\mathbf{y}}(n)$ and the vector $\bar{\mathbf{w}}(n)$ respectively. Now $\bar{y}_k(n)$ contains contributions from users $1, \ldots, k$ but not from users $k+1, \ldots, K$.

B. For frequency-selective fading channels

$$\mathbf{x}_1(n) = \bar{\mathbf{y}}_1(n) = \mathbf{F}_{1,1} \mathbf{g}_1 b_1(n) + \bar{\mathbf{w}}_1(n) \tag{6.20}$$

$$for \;\; k = 2 : K$$

$$\mathbf{x}_k(n) = \bar{\mathbf{y}}_k(n) - \sum_{i=1}^{k-1} \mathbf{F}_{k,i} \mathbf{g}_i b_i(n) = \mathbf{F}_{k,k} \mathbf{g}_k b_k(n) + \bar{\mathbf{w}}_k(n) \tag{6.21}$$

$$end$$

where $\mathbf{F}_{i,j}$ is the $L \times L$ submatrix of the matrix $\mathbf{F}$, i.e., $\mathbf{F}_{i,j} = \mathbf{F}((i-1)L+1 : iL, (j-1)L+1 : jL)$, $\bar{\mathbf{y}}_i(n)$ and $\bar{\mathbf{w}}_i(n)$ are the $L \times 1$ subvectors of the vector $\bar{\mathbf{y}}(n)$ and the vector $\bar{\mathbf{w}}(n)$, respectively. That is, $\bar{\mathbf{y}}_i(n) = \bar{\mathbf{y}}(n)((i-1)L+1 : iL)$ and $\bar{\mathbf{w}}_i(n) = \bar{\mathbf{w}}(n)((i-1)L+1 : iL)$. Here we denote $\mathbf{U}(i : j, k : l)$ as the submatrix of the matrix $\mathbf{U}$ by taking the rows from i to j and columns from k to l of $\mathbf{U}$, and $\mathbf{u}(i:j)$ is denoted as the subvector of the vector $\mathbf{u}$ by taking the elements from i to j of $\mathbf{u}$. Similarly, $\bar{\mathbf{y}}_k(n)$ contains contributions from users $1, \ldots, k$ but not from users $k+1, \ldots, K$.

6.5 DF-SMC receiver

According to the decomposed signal model, we derive the SMC-based detection with the decision-feedback framework. It should be understood that under this framework, the symbols are detected sequentially from the first user to the last user. An SMC inference for the first user is made, then since the decision for the first user is available, the SMC inference for the second user can be made by using the feedback of the first user's inference results. Similarly, for the kth user, the inference can be made depend on all the previous users' feedback. The complete framework is illustrated in Fig 6.1, where the parameters have the different expressions for different fading channel systems as shown in the corresponding models.

Here, the SMC inference methods are used to achieve the optimum estimations for $\{b_k(n), g_k\}_{k=1}^K$ or $\{b_k(n), \mathbf{g}_k\}_{k=1}^K$ according to the models (6.19) or (6.21). For convenience, only the model (6.21) is taken in developing the following SMC inference procedure. The results for the model (6.19) are similar and simpler. We now consider the estimation of the unknown parameters $\{b_k(n), \mathbf{g}_k\}_{k=1}^K$ based on each signal component $\{\mathbf{x}_k(n)\}_{k=1}^K$.

Denote $\mathbf{X}_{k,n} \triangleq \{\mathbf{x}_k(0), \mathbf{x}_k(1), \cdots, \mathbf{x}_k(n)\}$ and $\mathbf{D}_{k,n} \triangleq \{d_k(1), \cdots, d_k(n)\}$. The objective is to use SMC method to perform a blind estimation to compute

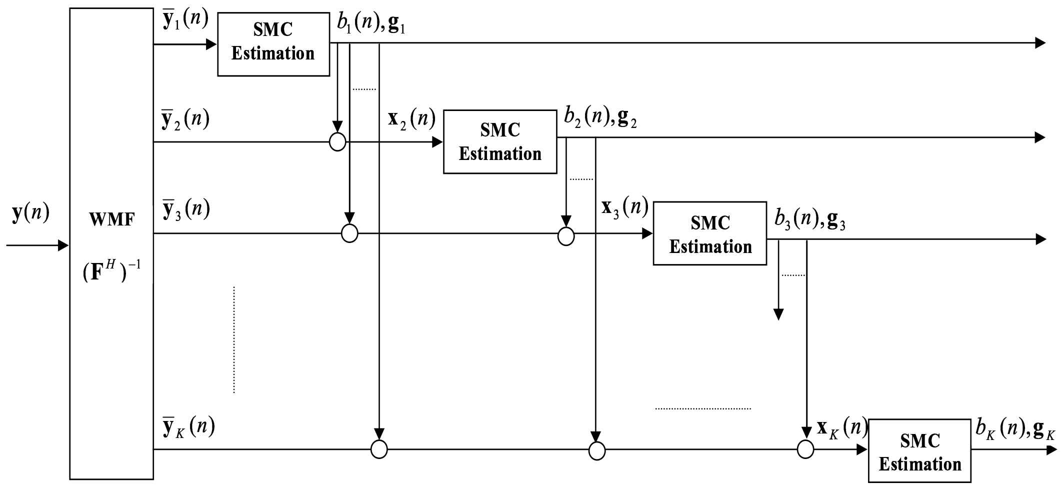


Figure 6.1: The DF-SMC receiver framework

the *a posteriori* symbol probability

$$P(d_k(n) = a_i|\mathbf{X}_{k,n}), \quad a_i \in A; \quad n = 1, \cdots, N-1 \tag{6.22}$$

based on the observation component $\mathbf{X}_{k,n}$ up to time n and *a priori* symbol probability of $\mathbf{D}_{k,n}$, without knowing the channel response $\mathbf{g}_k$.

Let $\{b_k^{(j)}(n)\}_{j=1}^m$ be samples drawn by the SMC at time n and denote $\mathbf{B}_{k,n}^{(j)} = \{b_k^{(j)}(0), \cdots, b_k^{(j)}(n)\}$. For each value of n, a set of Monte Carlo samples of transmitted symbols, $\{\mathbf{B}_{k,n}^{(j)}, w_{k,n}^{(j)}\}_{j=1}^m$, which are properly weighted with respect to the distribution $p(\mathbf{B}_{k,n}|\mathbf{X}_{k,n})$, are to be obtained. For every symbol $a_i \in A$, the *a posteriori* probability of the information symbol $d_k(n)$ can be estimated as

$$\begin{aligned} P(d_k(n) = a_i|\mathbf{X}_{k,n}) &= P(b_k(n)b_k^*(n-1) = a_i|\mathbf{X}_{k,n}) \\ &= E\{1(b_k(n)b_k^*(n-1) = a_i)|\mathbf{X}_{k,n}\} \\ &\cong \frac{1}{W_{k,n}} \sum_{j=1}^m 1(b_k^{(j)}(n)b_k^{(j)*}(n-1) = a_i)w_{k,n}^{(j)} \end{aligned} \tag{6.23}$$

where $W_{k,n} \triangleq \sum_{j=1}^{m} w_{k,n}^{(j)}$ and $1(\cdot)$ is an indicator function.

The samples $\{b_k^{(j)}(n)\}_{j=1}^{m}$ are drawn from the trial sampling density

$$q(b_k^{(j)}(n)|\mathbf{B}_{k,n-1}^{(j)}, \mathbf{X}_{k,n}) \triangleq p(b_k^{(j)}(n)|\mathbf{B}_{k,n-1}^{(j)}, \mathbf{X}_{k,n}) \tag{6.24}$$

and the importance weight can be updated according to

$$\begin{aligned} w_{k,n}^{(j)} &\propto w_{k,n-1}^{(j)} p(\mathbf{X}_{k,n}|\mathbf{B}_{k,n-1}^{(j)}, \mathbf{X}_{k,n-1}) \\ &= w_{k,n-1}^{(j)} \sum_{a_i \in A} p(\mathbf{x}_k(n)|\mathbf{B}_{k,n-1}^{(j)}, b_k(n) = a_i, \mathbf{X}_{k,n-1}) P(b_k(n) = a_i|\mathbf{B}_{k,n-1}^{(j)}, \mathbf{X}_{k,n-1}) \\ &= w_{k,n-1}^{(j)} \sum_{a_i \in A} \alpha_{k,n,i}^{(j)} . \end{aligned} \tag{6.25}$$

To compute the predictive density $p(\mathbf{x}_k(n)|\mathbf{B}_{k,n-1}^{(j)}, b_{k,n}(n) = a_i, \mathbf{X}_{k,n-1})$, a Gaussian distribution is assigned to the channel $\mathbf{g}_k$, i.e.,

$$\mathbf{g}_k \sim \mathcal{N}_c(\bar{\mathbf{g}}_k, \bar{\boldsymbol{\Sigma}}_k). \tag{6.26}$$

Then, the distribution of $\mathbf{g}_k$, conditioned on $B_{k,n}^{(j)}$ and $X_{k,n}$, can be computed as

$$\begin{aligned} p(\mathbf{g}_k|\mathbf{B}_{k,n}^{(j)}, \mathbf{X}_{k,n}) &\propto p(\mathbf{X}_{k,n}|\mathbf{B}_{k,n}^{(j)}, \mathbf{g}_k) p(\mathbf{g}_k) \\ &\sim \mathcal{N}_c(\mathbf{g}_{k,n}^{(j)}, \boldsymbol{\Sigma}_{k,n}^{(j)}) \end{aligned} \tag{6.27}$$

where

$$\mathbf{g}_{k,n}^{(j)} \triangleq \boldsymbol{\Sigma}_{k,n}^{(j)} \left[\bar{\boldsymbol{\Sigma}}_k^{-1} \bar{\mathbf{g}}_k + \frac{1}{\sigma^2} \sum_{i=0}^{n} \boldsymbol{\Psi}_{k,i}^{(j)H} \mathbf{x}_k(i) \right] \tag{6.28}$$

$$\boldsymbol{\Sigma}_{k,n}^{(j)} \triangleq \left[\bar{\boldsymbol{\Sigma}}_k^{-1} + \frac{1}{\sigma^2} \sum_{i=0}^{n} \boldsymbol{\Psi}_{k,i}^{(j)H} \boldsymbol{\Psi}_{k,i}^{(j)} \right]^{-1} \tag{6.29}$$

and

$$\boldsymbol{\Psi}_{k,i}^{(j)} = b_k^{(j)}(i)\mathbf{F}_{k,k}. \tag{6.30}$$

Hence, the conditional density $p(\mathbf{x}_k(n)|\mathbf{B}_{k,n-1}^{(j)}, b_k(n) = a_i, \mathbf{X}_{k,n-1})$ is given by

$$\begin{aligned}&p(\mathbf{x}_k(n)|\mathbf{B}_{k,n-1}^{(j)}, b_k(n) = a_i, \mathbf{X}_{k,n-1})\\ =&\int p(\mathbf{x}_k(n)|\mathbf{B}_{k,n-1}^{(j)}, b_k(n){=}a_i, \mathbf{X}_{k,n-1}, \mathbf{g}_k)p(\mathbf{g}_k|\mathbf{B}_{k,n-1}^{(j)}, \mathbf{X}_{k,n-1})d\mathbf{g}_k.\end{aligned} \tag{6.31}$$

Because it is an integral of a Gaussian probability density function (pdf) with respect to another Gaussian pdf, the resulting pdf is still Gaussian, i.e.,

$$p(\mathbf{x}_k(n)|\mathbf{B}_{k,n-1}^{(j)}, b_k(n) = a_i, \mathbf{X}_{k,n-1}) \sim \mathcal{N}_c(\mu_{k,n,i}^{(j)}, \boldsymbol{\Theta}_{k,n,i}^{(j)}) \tag{6.32}$$

with a mean

$$\begin{aligned}\mu_{k,n,i}^{(j)} &\triangleq E\left\{\mathbf{x}_k(n)|\mathbf{B}_{k,n-1}^{(j)}, b_k(n) = a_i, \mathbf{X}_{k,n-1}\right\}\\ &= \boldsymbol{\Phi}_{k,i}\mathbf{g}_{k,n-1}^{(j)}\end{aligned} \tag{6.33}$$

and a covariance

$$\begin{aligned}\boldsymbol{\Theta}_{k,n,i}^{(j)} &\triangleq Cov\left\{\mathbf{x}_k(n)|\mathbf{B}_{k,n-1}^{(j)}, b_k(n) = a_i, \mathbf{X}_{k,n-1}\right\}\\ &= \sigma^2\mathbf{I}_L + \boldsymbol{\Phi}_{k,i}\boldsymbol{\Sigma}_{k,n-1}^{(j)}\boldsymbol{\Phi}_{k,i}^H\end{aligned} \tag{6.34}$$

where

$$\boldsymbol{\Phi}_{k,i} = a_i\mathbf{F}_{k,k}. \tag{6.35}$$

Then, $\alpha_{k,n,i}^{(j)}$ in (6.25) can be computed by

$$\alpha_{k,n,i}^{(j)} = |\mathbf{\Theta}_{k,n,i}^{(j)}|^{-1} exp\left\{-(\mathbf{x}_k(n) - \mu_{k,n,i}^{(j)})^H(\mathbf{\Theta}_{k,n,i}^{(j)})^{-1}(\mathbf{x}_k(n) - \mu_{k,n,i}^{(j)})\right\} P(d_k(n) = a_i b_k^{(j)*}(n-1)) \tag{6.36}$$

It is noted that the *a posteriori* mean and covariance of the channel in (6.28) and (6.29) can be updated recursively as follows. At the nth step, the new sample of $b_k^{(j)}(n)$ and the past samples $\mathbf{B}_{k,n-1}^{(j)}$ are combined to form $\mathbf{B}_{k,n}^{(j)}$. Let $\mu_{k,n}^{(j)}$ and $\mathbf{\Theta}_{k,n}^{(j)}$ be the quantities computed by (6.33) and (6.34) for the imputed $b_k^{(j)}(n)$. Based on a matrix inversion lemma, (6.28) and (6.29) become

$$\mathbf{g}_{k,n}^{(j)} = \mathbf{g}_{k,n-1}^{(j)} + \mathbf{\Omega}_{k,n}^{(j)}(\mathbf{\Theta}_{k,n}^{(j)})^{-1}(\mathbf{x}_k(n) - \mu_{k,n}^{(j)}) \tag{6.37}$$

$$\mathbf{\Sigma}_{k,n}^{(j)} = \mathbf{\Sigma}_{k,n-1}^{(j)} - \mathbf{\Omega}_{k,n}^{(j)}(\mathbf{\Theta}_{k,n}^{(j)})^{-1}\mathbf{\Omega}_{k,n}^{(j)H} \tag{6.38}$$

with

$$\mathbf{\Omega}_{k,n}^{(j)} = \mathbf{\Sigma}_{k,n-1}^{(j)}\mathbf{\Psi}_{k,n}^{(j)H}. \tag{6.39}$$

Finally, the SMC blind detector for each decomposed signal component is summarized as follows.

Initialization:

For $j = 1, \cdots, m$

- Set the initial values of channel vector as $\mathbf{g}_k^{(j)} \sim \mathcal{N}_c(\mathbf{0}, 100\mathbf{I}_L)$;
- Initial all importance weights as $w_{k,-1}^{(j)} = 1$.

End

Estimation:

For $n = 0, \cdots, N-1$

I. Update weighted samples

For $j = 1, \cdots, m$

1. For each $a_i \in A$

   - Compute $\mu_{k,n,i}^{(j)}$, $\mathbf{\Theta}_{k,n,i}^{(j)}$ according to (6.33) and (6.34);

   - Compute trial sampling distribution $\alpha_{k,n,i}^{(j)}$ according to (6.36).

   End

2. Draw a sample $b_k^{(j)}(n)$ from the set A with the probability

$$P(b_k(n) = a_i | \mathbf{B}_{k,n-1}^{(j)}, \mathbf{X}_{k,n}) \propto \alpha_{k,n,i}^{(j)}, \quad a_i \in A. \tag{6.40}$$

3. Compute the importance weight

$$\hat{w}_{k,n}^{(j)} = w_{k,n-1}^{(j)} \sum_{a_i \in A} \alpha_{k,n,i}^{(j)}, \tag{6.41}$$

   Normalize as

$$w_{k,n}^{(j)} = \frac{\hat{w}_{k,n}^{(j)}}{\sum_{j=1}^{m} \hat{w}_{k,n}^{(j)}}. \tag{6.42}$$

4. If the imputed samples $b_k^{(j)}(n) = a_i$

   - Set $\mu_{k,n}^{(j)} = \mu_{k,n,i}^{(j)}, \mathbf{\Theta}_{k,n}^{(j)} = \mathbf{\Theta}_{k,n,i}^{(j)}$;

   - Update the *a posteriori* mean $\mathbf{g}_{k,n}^{(j)}$ and covariance $\mathbf{\Sigma}_{k,n}^{(j)}$ of the channel according to (6.37) and (6.38).

End

5. Compute the *a posteriori* probability of the information symbol $d_k(n)$ according to (6.23).

6. Compute $b_k(n)$ according to $b_k(n) = d_k(n)b_k^*(n-1)$

End (for **I**)

II. Resampling

If n is a multiple of the resampling interval

For $j = 1, \cdots, m$

- Draw a new set of $\{\mathbf{B}_{k,n}^{(j)}, \mathbf{g}_{k,n}^{(j)}, \mathbf{\Sigma}_{k,n}^{(j)}\}_{j=1}^{m}$ from the original set with a probability proportional to the importance weights $\{w_{k,n}^{(j)}\}_{j=1}^{m}$;
- Assign equal weight for each new samples, i.e., $\hat{w}_{k,n}^{(j)} = 1/m$.

End

End (for **II**)

End

It is observed that, for each signal component, the dominant computation for the SMC receiver needs $N \times m \times |A|$ one-step predictions for computing $\alpha_{k,n,i}^{(j)}$ and $N \times m$ one-step updates for $\{(\mathbf{g}_{k,n}^{(j)}, \mathbf{\Sigma}_{k,n}^{(j)})\}$. Since the m samples operate independently, the sampling is well suited for parallel implementation. For the entire detection algorithm, the proposed new method samples one user at a time and therefore permits efficient implementation. The required computational complexity is reduced to be in the order of $O(|A| \times K)$. Compared to the

EM-SMC, the proposed method achieves a lower complexity, since there is no iteration for convergence. Thus, the proposed DF-SMC detection algorithm can be applied to the DS-CDMA system with a manageable computational complexity.

6.6 Extension to system with long codes

The above proposed detection algorithm is developed for the system with short spreading codes. It is seen that, the bit-interval cyclostationarity properties of the short codes are not necessary for this SMC-based method, and the inter-symbol interference (ISI) is not considered for the system model. Therefore, we can extend this method to the long code CDMA systems by making necessary changes.

For long spreading codes, the spreading waveform generally changes from bit to bit, thus the cross-correlation matrix is a function of the symbol index n, that is, $\mathbf{R}(n)$. For the case of the flat fading channel, the elements of the cross-correlation matrix are changed from (6.2) to

$$\mathbf{R}_{i,j}(n) = \int_{(n-1)T}^{nT} s_i(t)s_j(t)dt. \tag{6.43}$$

For the case of frequency-selective fading channel, the corresponding elements of the cross-correlation matrix are changed from (6.10) to

$$\rho_{(k,l)(k',l')}(n) = \int_{(n-1)T}^{nT} s_k(t-\tau_{k,l})s_{k'}(t-\tau_{k',l'})dt. \tag{6.44}$$

Correspondingly, the whitened matched filter is also the function of the symbol index n, that is, $\mathbf{F}(n)$. With these substitutions, the proposed detection algorithm can be applied to the CDMA system employing long spreading codes directly. Of course, the whitened matched filter should be update every symbol, so that the implementation costs are inevitably increased.

6.7 Simulation results

This section provides simulation results to illustrate the performance of the blind DF-SMC multiuser receiver in both flat fading channels and frequency-selective fading channels for DS-CDMA system. The channels are assumed to be block fading, that is the fading coefficients remain constant over the entire block of N symbols. The fading coefficients of the channels are generated according to uncorrelated circular complex Gaussian distribution. All the users' spreading sequences are chosen as short sequence with a processing gain $P = 10$, and they are generated randomly from equal probability binary code and then normalized to $\{\pm 1/\sqrt{P}\}$. A rate 1/2 constraint length-5 convolutional code (with generator 23 and 35 in octal notation) is employed. We choose 128 information bits, i.e., the coded bit block size is $N = 256$. Two types of channels are considered, i.e, flat fading channel and frequency-selective fading channel with $L = 3$.

6.7.1 Selection of the number of Monte Carlo samples

Let us first consider the selection of the number of Monte Carlo samples. Fig 6.2 and Fig 6.3 present the bit-error-rate (BER) performance as the function of the number of the Monte Carlo samples for flat-fading system and frequency selective fading system, respectively. As illustrated in the figures, for both flat fading system and frequency-selective fading system, the performance improves as the number of samples increases, and the performance achieves convergence when m is about 50. Therefore, in the following simulations, the number of Monte Carlo samples is taken as $m = 50$.

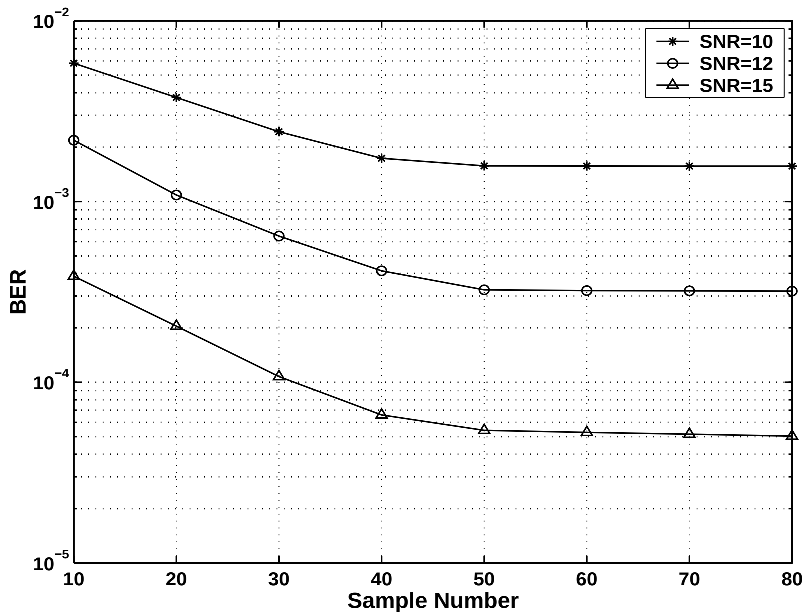


Figure 6.2: BER versus the number of Monte Carlo samples, $K = 8$ (flat fading channel)

6.7.2 BER performance

Let us now test the BER performance of the proposed method for different number of users. Fig 6.4 and Fig 6.5 illustrate the BER performance of the proposed blind detector in the flat fading channel and frequency-selective fading channel, respectively. The BER performances as the function of SNR are plotted in the figures for different number of users, i.e., $K = 8$, 10 and 12.

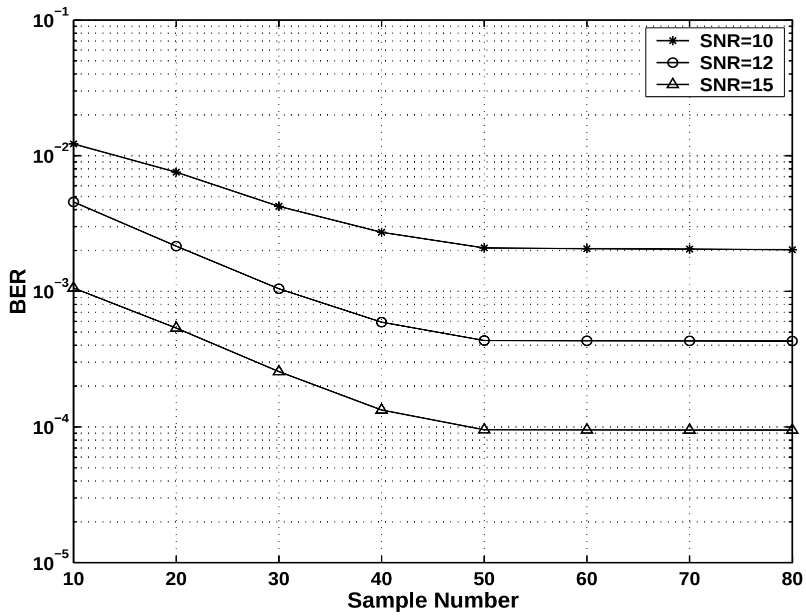


Figure 6.3: BER versus the number of Monte Carlo samples, $K = 8$ (frequency-selective fading channel)

6.7.3 Near-far resistance

The capability of near-far resistance is illustrated in Fig 6.6 for the system over flat fading channel, and in Fig 6.7 for the system over frequency-selective fading channel. The near-far ratio is defined as the ratio between the power of interfering users and the power of the desired user. Let us fix the power of the desired user and change the power of interfering users. It is assumed that all interfering users have the same power. The BER performance is tested as the function of near-far ratio for the systems with $K = 8$, 10 and 12. The results in Fig 6.6 and Fig 6.7 demonstrate the good near-far resistance of the proposed receiver.

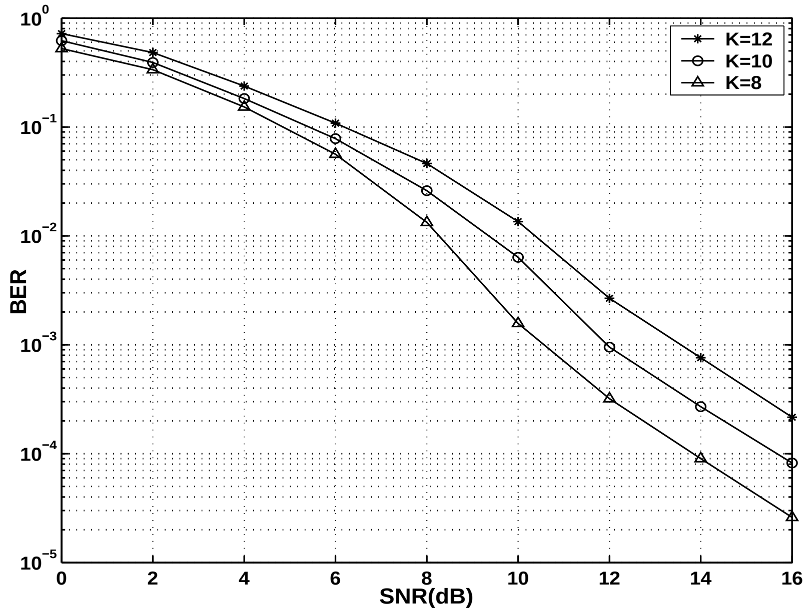


Figure 6.4: BER versus the SNR (flat fading channel)

6.7.4 Comparisons of performance

Finally, comparisons are made on the performances achieved by the new proposed receiver, the receiver derived in previous chapter and other reported receivers. Fig 6.8 and Fig 6.9 present the performances achieved by the Gibbs sampler [84], QRD-M-EKF [119], the EM-SMC receiver, the conventional SMC receiver [90] and the DF-SMC receiver.

For a fair comparison, the Gibbs sampler is performed for 100 iterations with the first 50 iterations as the burning-in period. The number of Monte Carlo samples used for both the EM-SMC receiver and the conventional SMC receiver [90] is 50, which is the same as that used for the proposed receiver. The number of paths M for the tree-search detection QRD-M-EKF is selected to be 32. The results for the flat fading system are shown in Fig 6.8, while the results for the frequency-selective fading system are shown in Fig 6.9.

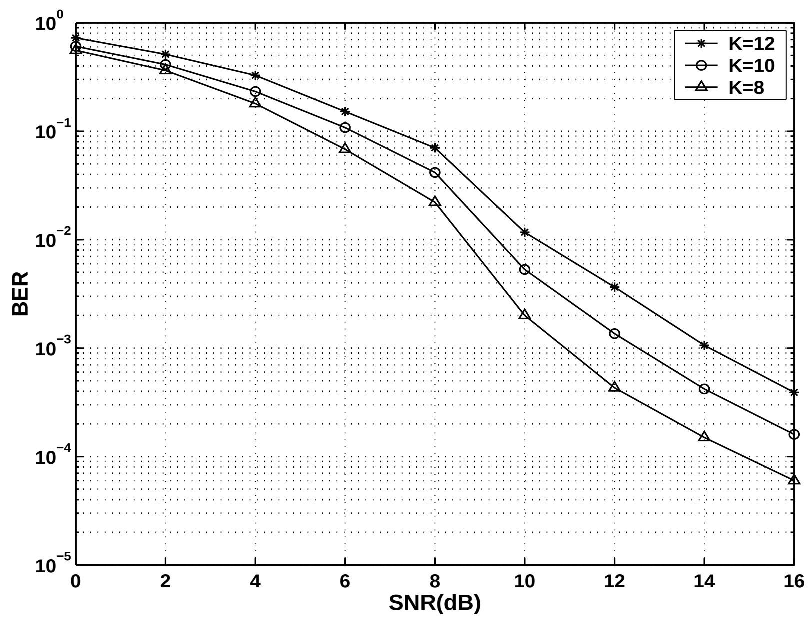


Figure 6.5: BER versus the SNR (frequency-selective fading channel)

Fig 6.8 and Fig 6.9 show that the proposed receiver also outperforms the other deterministic and MC sampling based detectors, and achieves the comparable performance with the EM-SMC receiver and the conventional SMC receiver. When the SNR is in small values, the proposed DF-SMC receiver is a little inferior to the other two SMC-based receivers. When the value of the SNR becomes larger, the performance of the DF-SMC receiver becomes superior to that of the other two receivers. It is also seen from the figures, with the increase of the SNR, the improvement made by the proposed receiver is more obvious. This is because the decision-feedback framework attempts to cancel all multiuser interference provided that the feedback data are correct. With the MAI cancellation, the detection of each user is similar to that for single user system. For systems (6.19) and (6.21), without the MAI effects, the detection performance can improve more significantly when the SNR increases. When the background noise is hypothetically absent, i.e. $\sigma = 0$, the detector should

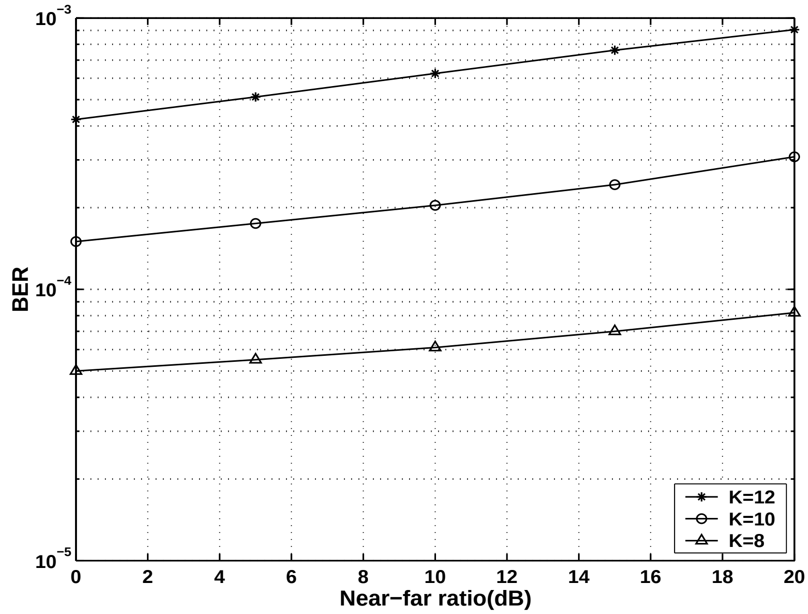


Figure 6.6: BER versus the near-far ratio, $SNR = 15\ dB$ (flat fading channel)

guarantee error-free demodulation approximately.

6.8 Conclusions

In this chapter, a blind multiuser receiver is developed for the DS-CDMA system over both the flat fading channels and the frequency-selective fading channels based on the Bayesian SMC inference method. The Cholesky factorization is utilized before the implementation of the SMC detection to achieve significant reduction in computational complexity. As seen in the simulation results, the proposed receiver obtains a performance that is better than that achieved by the conventional SMC receiver and the EM-SMC receiver when SNR is large enough. And with much smaller computational complexity, the proposed receiver is proved to be well suited to DS-CDMA multiuser systems.

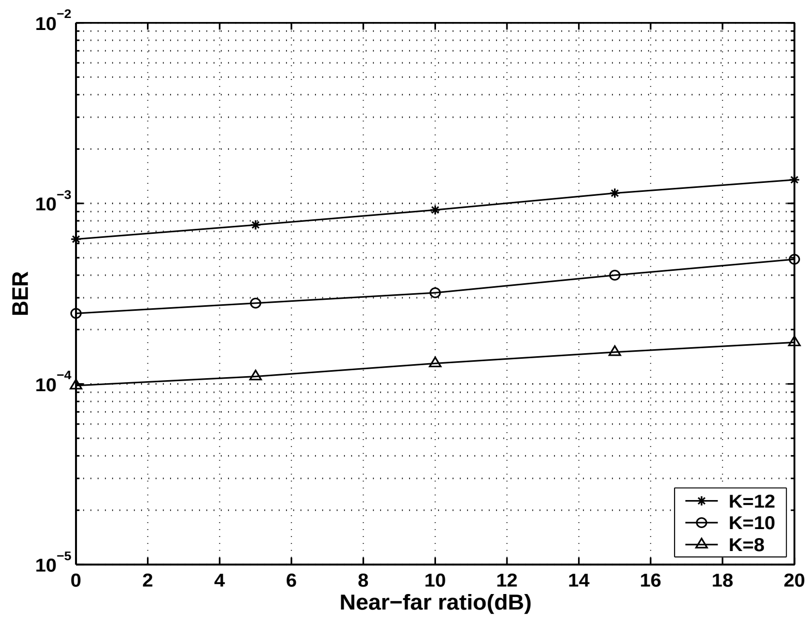


Figure 6.7: BER versus the near-far ratio, $SNR = 15\ dB$ (frequency-selective fading channel)

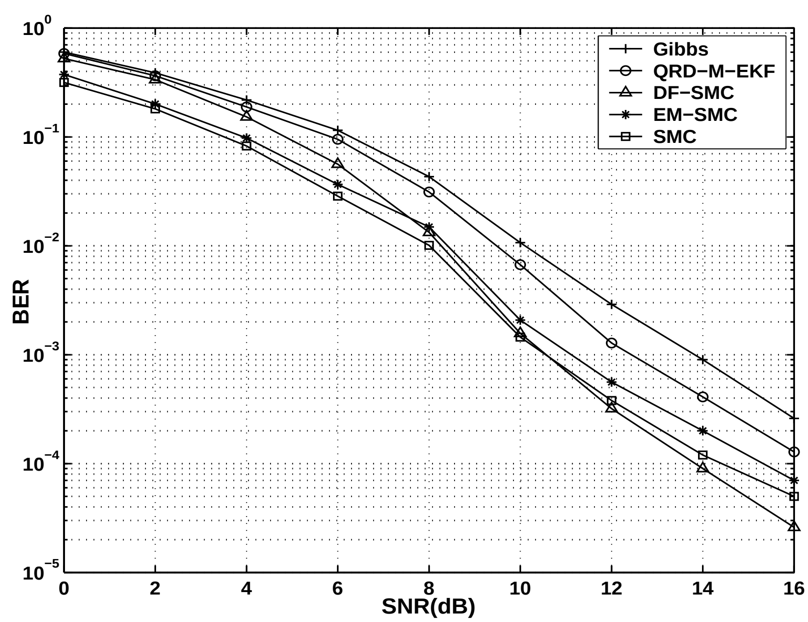


Figure 6.8: Comparisons on BER performance, $K = 8$ (flat fading channel)

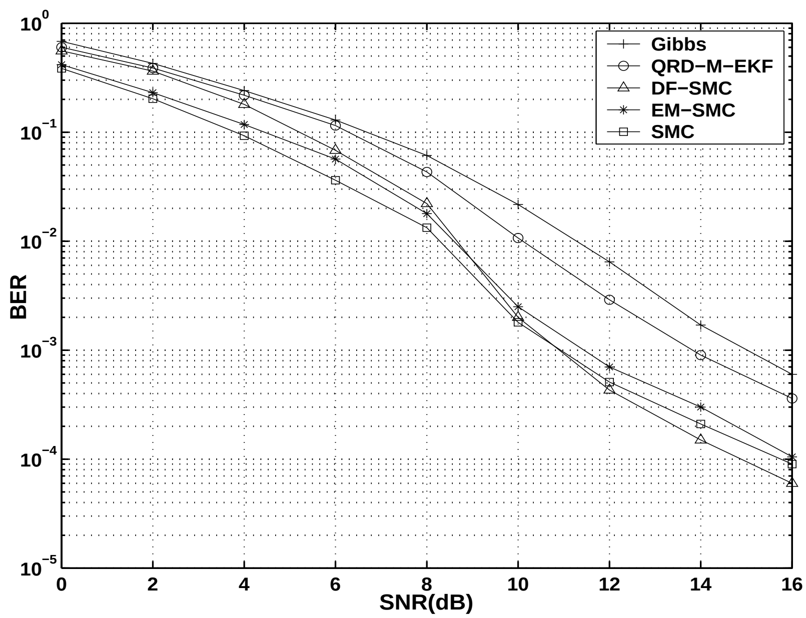


Figure 6.9: Comparisons on BER performance, $K = 8$ (frequency-selective fading channel)

Chapter 7

Comparisons of multiuser receivers

In this chapter, some general conclusions and comparisons are made among the multiuser receivers which are proposed in previous chapters. The comparisons include the requirements of coefficient knowledge, the implementation complexity, and the performance of the receiver. The corresponding simulation results are provided to illustrate the conclusions.

7.1 Requirements

Firstly, we discuss the required knowledge of the coefficients for the various proposed methods. Since our methods are all blind, the training sequence is not needed for any receiver. For all these methods, we have the assumptions that the number of users, the length of the channels and the spreading gain of the signature waveforms are the available knowledge for the systems. Therefore, the main considerations in the following discussion involve the signature waveform, timing, received amplitudes and noise level.

For the algorithm developed in Chapter 3, which is denoted as TD-CM, the contribution focuses on the channel estimation rather than multiuser detection, therefore we consider only the requirements for the channel estimation. The prior knowledge which should be required includes the received amplitudes of the desired user, the timing of desired user and the signature waveform of desired user, since the effects of the interfering users and noise are removed by the Toeplitz displacement.

The Bayesian MCMC detector developed in Chapter 4, which is denoted as MCMC-AS here, is based on the adaptive sampling method. In order to update the *a posteriori* conditional distributions of the channel and noise, the received amplitudes, the timing and the signature waveforms of all users are required.

As for both the Bayesian SMC detectors proposed in Chapter 5 and Chapter 6, the received amplitudes, the timing and the signature waveforms of all users are necessary for the requirements of the signal decomposition before the SMC detection. For the SMC detection procedure, the knowledge of noise level need to be available.

For clarity, the comparison of the requirements for various methods is shown in Table 7.1.

Table 7.1: Comparison of requirements for the proposed multiuser detectors

	TD-CM	MCMC-AS	EM-SMC	DF-SMC
Signature waveform of desired user	$\surd$	$\surd$	$\surd$	$\surd$
Signature waveform of interfering users		$\surd$	$\surd$	$\surd$
Timing of desired user	$\surd$	$\surd$	$\surd$	$\surd$
Timing of interfering users		$\surd$	$\surd$	$\surd$
Received amplitude of desired user	$\surd$	$\surd$	$\surd$	$\surd$
Received amplitude of interfering users		$\surd$	$\surd$	$\surd$
Noise level			$\surd$	$\surd$

7.2 Complexity

The computation complexity for the TD-CM channel estimation has been discussed in section 3.4.3. Since TD-CM method focuses on only the channel estimation and the detector need to be designed separately based on the channel estimates, the complexity of the entire receiver not only depends on the complexity of TD-CM, but also the chosen detection algorithm. However Bayesian Monte Carlo method is applied in this thesis for developing the detection algorithm directly. Therefore, it's not very meaningful to compare just the TD-CM method with the proposed Monte Carlo receivers. As for the three Monte Carlo receivers proposed in this thesis, the computational complexity of the DF-SMC receiver is lower than that of the EM-SMC receiver which has been discussed in the Chapter 6; while the complexity of the MCMC-AS receiver is lower than those of both two SMC based receivers. The reason is that the computation of the sampling distributions is necessary in every iteration for the SMC based methods, while for the MCMC-AS receiver, the sampling distributions are updated only after every several iterations.

The similar conclusion can be made for comparing the implementation complexity of the proposed receivers. As for the implementation framework of the receiver, the simplest one is the MCMC detector based on the adaptive sampling method, which is implemented only through the proposed detection procedure. For the receiver employing TD-CM channel estimator, a set of matched filters are necessary to exploit the properties of the signals, and the estimator should be combined with other separate detector. EM-SMC receiver needs the iterations of the EM decomposition, in each iteration, the detection procedures of all users are performed in parallel. DF-SMC requires the whitened matched filter (WMF) to decompose the superimposed signals, while the decomposed signals for each user are detected sequentially.

7.3 Performance

The performances of various proposed multiuser receivers have been discussed in previous chapters, respectively. Here we make some comparisons for these multiuser receivers to obtain some conclusions about their performances.

In the simulations, we experiment with a DS-CDMA system which employs the long spreading codes. The number of users is 10. The spreading codes are generated randomly from equal probability binary code and then normalized to $\{\pm 1/\sqrt{P}\}$. Both the cases of the flat fading channels and the frequency-selective fading channels are considered, and for the frequency-selective fading channels, the length of channel is taken as $L = 3$. The fading coefficients of the channels are generated according to uncorrelated circular complex Gaussian distribution. For convenience, the initial delay for each user is assumed as 0. The initial values for the setup of the detection procedures are the same as in the previous chapters.

7.3.1 Channel estimation

Let us firstly consider the performance of channel estimation obtained by the proposed methods. In order to achieve the good performance for the TD-CM channel estimation, we choose the spreading gain to be 35 for all the methods.

7.3.1.1 MSE performance

The MSE results for the flat fading channel and for the frequency-selective fading channel are plotted in Fig 7.1 and Fig 7.2 respectively. It is shown that, the TD-CM obtains the best MSE performance among all the receivers, the EM-SMC receiver performs better channel estimation than the MCMC-AS receiver. The DF-SMC receiver has relative poor performance when the SNR is small, and with the increase of SNR, it achieves better performance than the MCMC-AS receiver and EM-SMC receiver.

7.3.1.2 Near-far resistance

The robustness of near-far resistance for channel estimation is compared in Fig 7.3 for the system over flat fading channel, and in Fig 7.4 for the system over frequency-selective fading channel. We fix the power of the desired user and change the power of interfering users. It is assumed that all interfering users

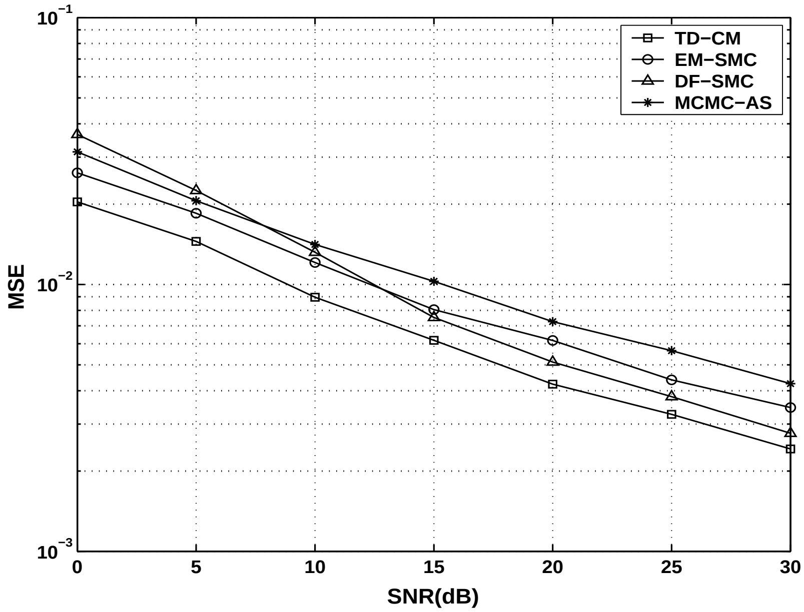


Figure 7.1: MSE versus SNR (flat fading channel)

have the same power. The MSE performance is tested as the function of near-far ratio for the systems with $SNR = 10\ dB$. It is shown in Figure 7.3 and Figure 7.4, the TD-CM and MCMC-AS methods have the comparable capability of the near-far resistance which is more robust than the two SMC-based methods. At the same time, the EM-SMC and DF-SMC methods obtain the comparable robustness of near-far resistance.

7.3.2 Signal detection

Since the method proposed in Chapter 3 is about channel estimation rather than multiuser detection, we consider only the Bayesian methods in the comparisons for the detection performance. As well as the method proposed in [65] which is denoted as LZ here, is implemented for the purpose of comparison. LZ method performs identification by exploiting the statistics of the covariance matrix of the outputs and is known to be a conventional (non-sampling) algorithm with

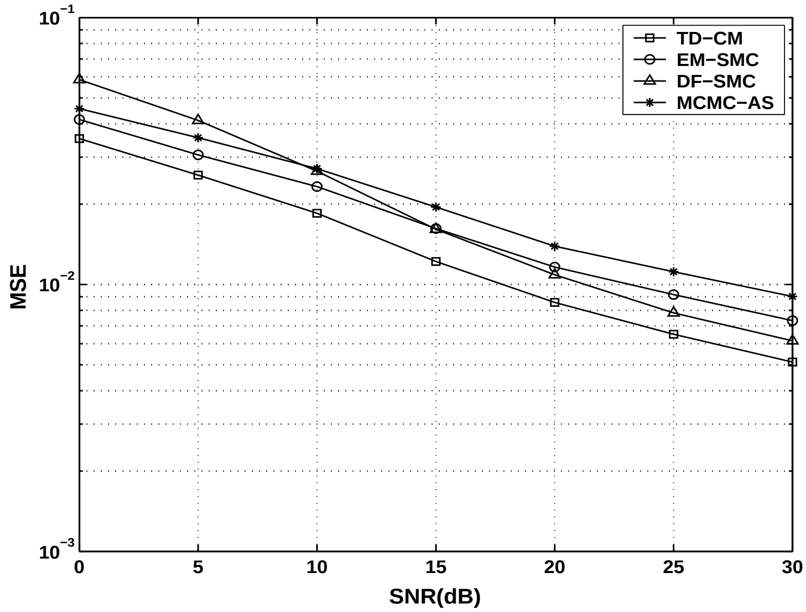


Figure 7.2: MSE versus SNR (frequency-selective fading channel)

the simplicity essential to practical implementation. The performance of LZ method is shown together with the proposed Monte Carlo Receivers in the following comparisons to demonstrated the advantages in performance of the proposed Bayesian Monte Carlo based receivers. For all these detectors, the spreading gain is assumed as $P = 10$.

7.3.2.1 BER performance

Now let us consider the BER performance of the proposed multiuser detectors for the system.

With the assumption that all users have the same amplitudes, we test the BER versus signal-to-noise ratio (SNR) for the coded system. For the MCMC-AS detector, we choose the number of iterations to be 50, while for both the EM-SMC detector and the DF-SMC, the numbers of the Monte Carlo samples are chosen as 50.

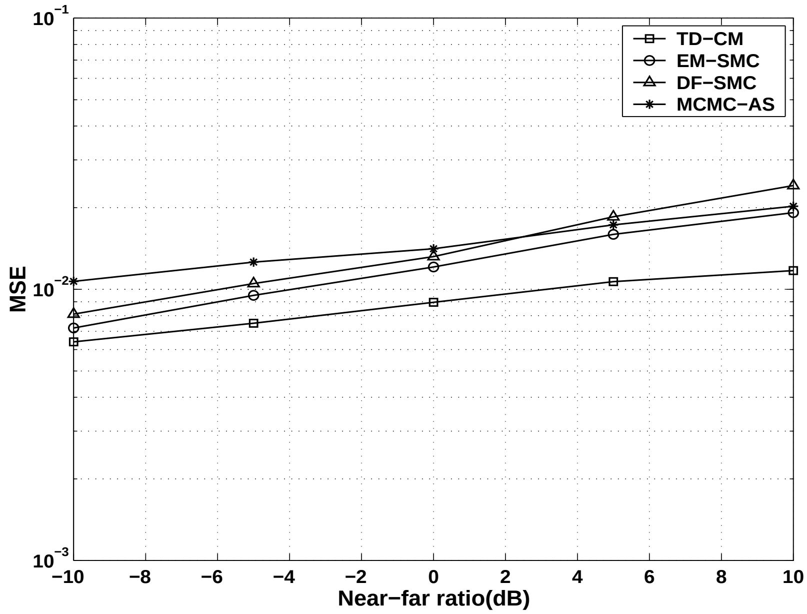


Figure 7.3: MSE versus near-far ratio, $SNR = 10\ dB$ (flat fading channel)

Fig 7.5 shows the results for the flat fading system, while Fig 7.6 shows the results for the frequency-selective fading system. As seen from these two figures, all the proposed Bayesian receivers outperform the LZ, and the two SMC-based detectors outperform the MCMC-AS detector. It is also seen that the EM-SMC detector achieves better performance than the DF-SMC detector when the SNR is relative small, and the DF-SMC detector is superior to the EM-SMC detector when the SNR becomes larger.

7.3.2.2 Near-far resistance

The capabilities of near-far resistance are compared in Fig 7.7 and Fig 7.9 for the system over flat fading channel, and in Fig 7.8 and Fig 7.10 for the system over frequency-selective fading channel. The desired user is at fixed power and all interfering users have the same power. The BER performance is tested as the function of near-far ratio for the systems with $SNR = 8\ dB$ and

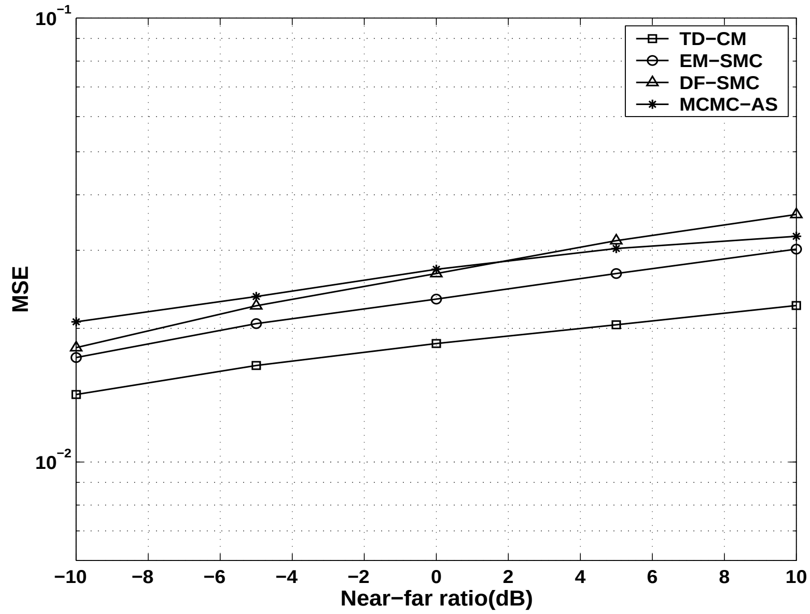


Figure 7.4: MSE versus near-far ratio, $SNR = 10\ dB$ (frequency-selective fading channel)

$SNR = 15\ dB$ respectively. Obviously, the figures show that the LZ has the weakest robustness to the near-far resistance among all receivers. It is seen in Figure 7.7 and Figure 7.8, when $SNR = 8\ dB$, the MCMC-AS receiver is the most robust for the near-far resistance among the receivers. At the same time, the capability of near-far resistance of EM-SMC receiver is superior to that of the DF-SMC receiver. Figure 7.9 and Figure 7.10 demonstrate that, with the larger SNR, i.e., 15 dB, the MCMC-AS receiver is a little better than the other two receivers for the near-far resistance, and the two SMC-based detectors have the comparable capability of the near-far resistance.

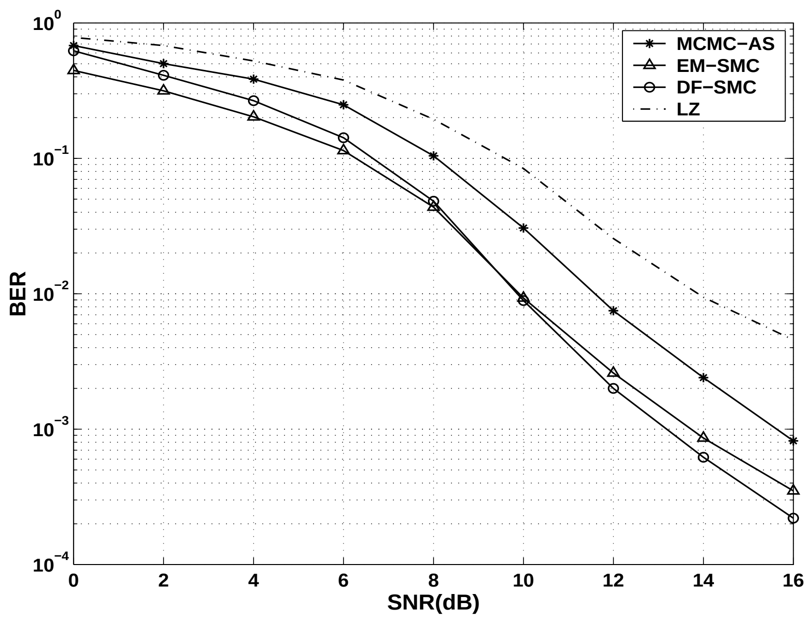


Figure 7.5: BER versus SNR (flat fading channel)

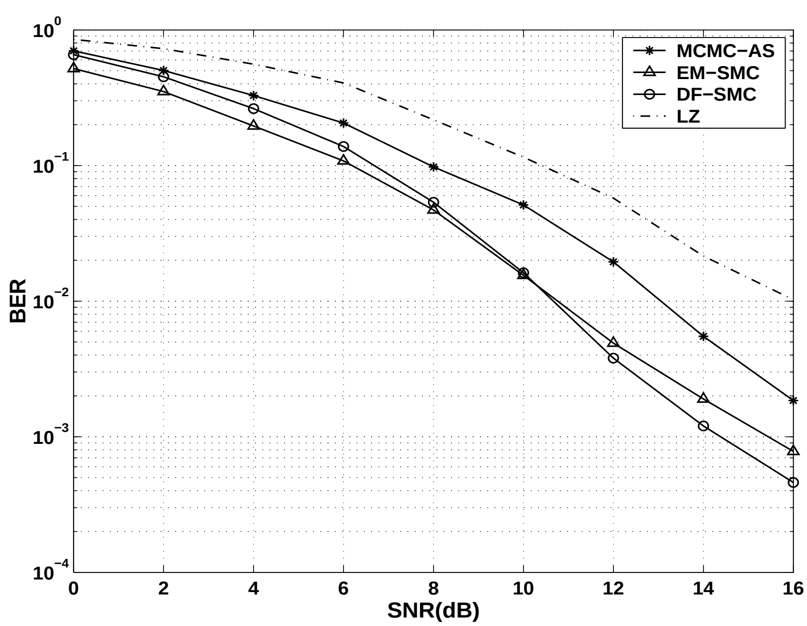


Figure 7.6: BER versus SNR (frequency-selective fading channel)

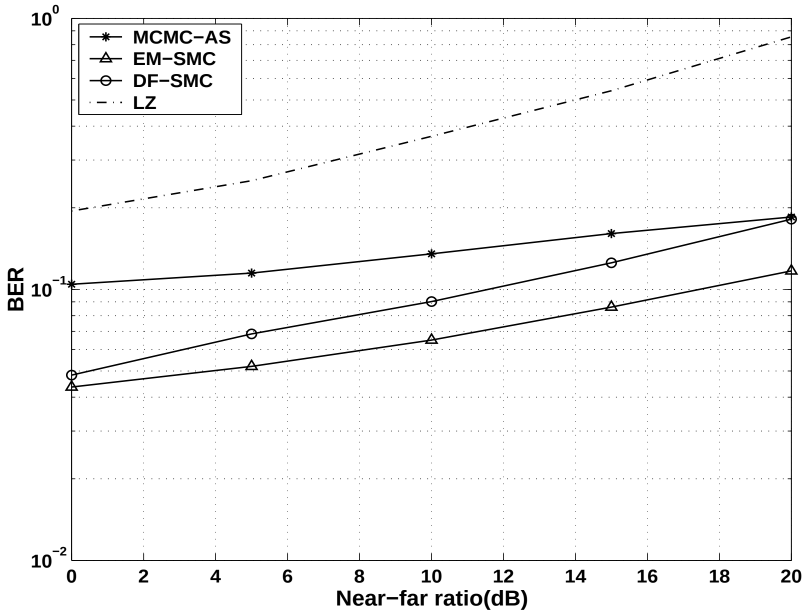


Figure 7.7: BER versus near-far ratio, $SNR = 8\ dB$ (flat fading channel)

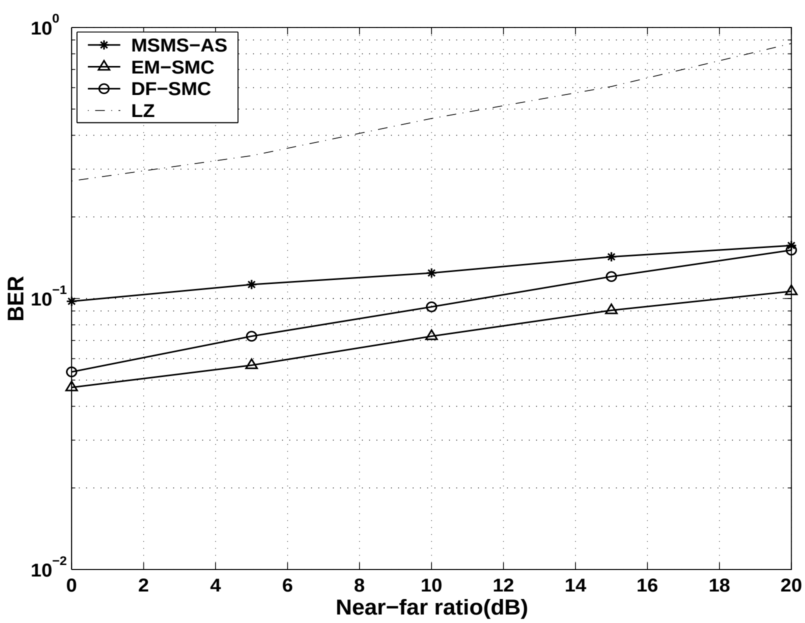


Figure 7.8: BER versus near-far ratio, $SNR = 8\ dB$ (frequency-selective fading channel)

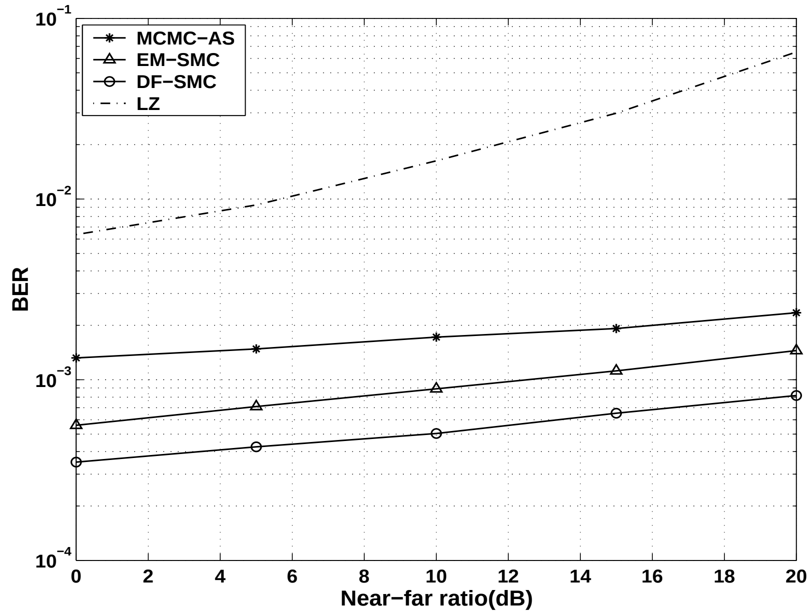


Figure 7.9: BER versus near-far ratio, $SNR = 15\ dB$ (flat fading channel)

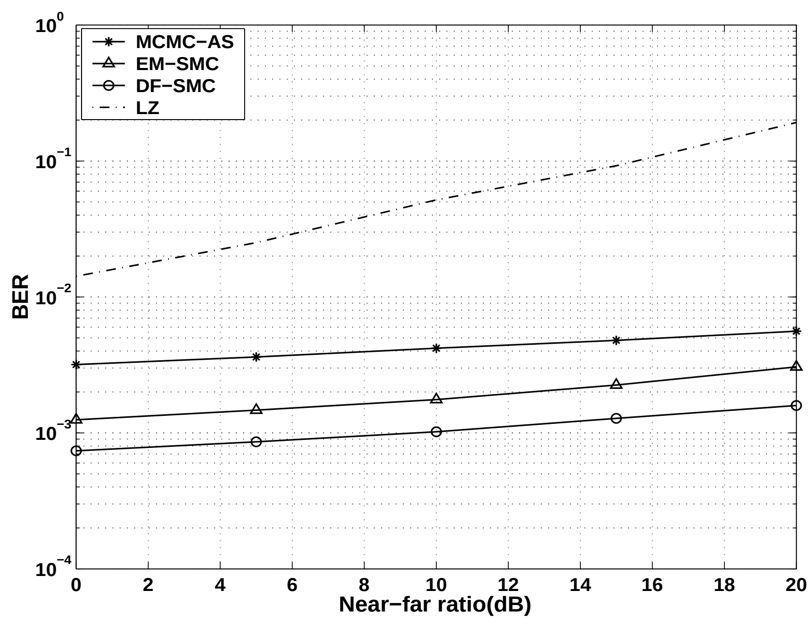


Figure 7.10: BER versus near-far ratio, $SNR = 15\ dB$ (frequency-selective fading channel)

Chapter 8

Conclusions and future work

In this chapter, we draw the conclusions of the thesis. The contents of the previous chapters are reviewed. Some suggestions for future work are provided.

8.1 Conclusions

This thesis deals with the problem of the channel estimation and multiuser detection for the DS-CDMA system over the fading channels. Several methods are developed for both cases of the flat fading channels and frequency-selective fading channels in order to improve the performance and/or reduce the computational complexity of the multiuser detection.

An efficient blind adaptive channel estimation method has been presented for the DS-CDMA system employing long spreading codes. The Toeplitz displacement is operated to the second order statistics, i.e., correlation matrix of the output vector after matched filter, so that the effects of the channel noise and other interfering users can be removed from the identification scheme. Then the correlation matching technique is explored to blindly estimate the channel parameters, that is, the channel parameters are estimated by minimizing the norm of the error matrix between the output correlation matrix (parameterized by the unknown channel coefficients) and its instantaneous approximation. Based on this approach, one LMS type recursion is derived to formulate the adaptive estimation algorithm. The performance of the proposed method is examined in both flat fading and frequency-selective fading channels by simulations. The comparisons are made with the conventional correlation matching method and the subspace Toeplitz method. The simulation results show the proposed method achieves substantial improvements on the MSE performance and near-far resistance.

A blind Bayesian MCMC multiuser receiver has been proposed for the DS-CDMA systems employing long spreading codes. The receiver deals with the joint of channel estimation and the symbol detection. The detection is based on the Bayesian MCMC methodology and uses the recently developed adaptive sampling procedure. Incorporated with the Bayesian channel estimation, an efficient blind multiuser receiver is derived. Compared to the Bayesian MCMC receivers based on current popular MCMC procedure, Gibbs sampler, the proposed receiver utilizes an adaptive procedure to find the most effective distribution for the generation of samples, thus achieves faster convergence and higher efficiency of sampling. On the other hand, the computation cost of updating the sampling distribution for the proposed receiver is lower than that for Gibbs receivers. Simulation results are provided to demonstrate that the proposed receiver is highly efficient in the convergence and performance.

A blind Bayesian SMC multiuser receiver has been developed for the joint of channel estimation and symbol detection. The conventional SMC estimation is not efficient in application to multiuser detection of CDMA system, because the required computational complexity grows exponentially with the number of users. In order to deal with the problem, the EM algorithm is adopted to decompose the superimposed observation signal into separate signals which contain only the information of one user. Then a Bayesian SMC procedure is developed to detect the unknown transmitted symbols according to the decomposed signal components. Combined with the Kalman filter for channel estimation, the symbol detection and channel estimation for every user are performed in parallel. Based on these concepts, a novel iterative SMC-based receiver EM-SMC is developed with the substantially reduced computational complexity which is linear to the number of users. The performance of the proposed EM-SMC receiver is examined and compared with other receivers. The simulation results demonstrated that EM-SMC receiver outperforms other methods with similar complexity and achieves about the comparable performance as that obtained by the conventional SMC detector.

Again in order to deal with the complexity problem of the conventional SMC method, another scheme is provided to develop the efficient SMC-based receiver. A different solution is presented to decompose the superimposed observation signal, which utilizes the Cholesky factorization to decouple the signal model into separate components according to the number of users. Then under the decision-feedback framework, the parameters of each user are estimated by SMC method and Kalman filter sequentially. According to these considerations, a new blind SMC-based multiuser receiver is proposed with complexity linear to the number of users. Because of no need for iterations like EM-SMC receiver, this receiver achieves more reduction in computational complexity. The performance of this receiver is examined and compared with some other receivers again. As seen in the simulation results, the proposed receiver is a little inferior to the conventional SMC receiver and the EM-SMC receiver when the SNR of the system is small. However the proposed receiver obtains the performance that is better than that achieved by the conventional SMC receiver and the EM-SMC receiver with the large SNRs. With much smaller computational complexity, the proposed receiver is proved to be better suited to DS-CDMA multiuser systems.

- Comparisons are made for the multiuser receivers which are proposed in the thesis. The requirements of coefficient knowledge, the implementation complexity, and the performance of channel estimation and symbol detection are considered together for all the receivers discussed in this thesis. All the requirements for various receivers are listed for comparison. The implementation complexities of the receiver frameworks are discussed generally. The performances of all the receivers are compared by the simulation results, which include the MSE performance of the channel estimation, BER performance of the symbol detection and the near-far resistance capability.

8.2 Future research suggestions

This thesis has solved some of the problems in the blind multiuser detection and channel estimation for DS-CDMA systems. The research and development in this area have been and will be one of the most active and vibrant branches of digital communications for a long period. And so far there are still many unsolved problems in these topics which need further investigation, and much more work should be done to develop better algorithms to deal with these problems. Based on the present work, future research can be recommended as follows.

- In this thesis, the communication channels are assumed to be slowly varying for both flat-fading channels and frequency-selective fading channels. That is, they remain constant for the entire symbol block. Hence, the proposed methods accommodate only the delay-insensitive DS-CDMA communication application in slow fading channels. Further research should be conducted on the effective methods that are suitable for a fast fading environment.

- All the work in the thesis considers that the channel ambient noise is Gaussian. However, in many physical channels where multiuser detection may be applied, such as urban and indoor radio channels, the ambient noise is known to be non-Gaussian, due to the impulsive nature of the man-made electromagnetic interference and a great deal of natural noises as well. Therefore, the extension of the proposed methods to the non-Gaussian CDMA channels is an interesting further research direction.

- Since the use of the spatial processing with the antenna arrays can substantially enhances the capacity of DS-CDMA systems, it is also an interesting and important future research direction to combine the proposed methods with array signal processing techniques.

- The techniques of delayed-weight estimation and delayed-sample estimation can be used to enhance the SIS which is the basis of all SMC methods. Thus, in order to improve the SMC-based methods developed in Chapter 5 and Chapter 6, the combination of the proposed SMC-based methods with these techniques deserves to be considered.

• In order to reduce the high complexity of the SMC application to multiuser CDMA system, it may be a feasible solution to develop a new novel dynamic model which describes the variation of each user's signal separately, so that only one unknown user's symbol is detected one time. The possible method is to make some transformations of the common dynamic state-space model. This viewpoint to multiuser systems is an interesting and challenging future research topic.

Author's Publications

1. Qian Yu, Guoan Bi and Chunru Wan, "SMC-Based blind detection for DS-CDMA systems over multipath fading channels", accepted by IEEE Transaction on Communications, 2005.

2. Qian Yu, Guoan Bi and Gaonan Zhang, "Improved blind multipath estimation for long code DS-CDMA", accepted by Journal of Communications and Networks, 2005.

3. Gaonan Zhang, Guoan Bi and Qian Yu, "Intersymbol decorrelating detector for asynchronous CDMA networks with multipath", EURASIP Journal on Wireless Communications and Networking, vol. 2005, no. 3, pp. 419-425, 2005.

4. Gaonan Zhang, Guoan Bi and Qian Yu, "Blind intersymbol decorrelating detector for asynchronous multicarrier CDMA System", Signal Processing, vol. 85, no. 8, pp.1511-1522, 2005.

5. Qian Yu, Guoan Bi and Liren Zhang, "New SMC Method of Blind Multiuser Detection Based on Cholesky Factorization for DS-CDMA Systems", revised and submitted to IEEE Transaction on Vehicle Technology, 2005.

6. Qian Yu, Guoan Bi and Liren Zhang, "Blind multiuser detection for long code multipath DS-CDMA systems with Bayesian MC techniques", submitted to Personal Wireless Communication, 2005.

7. Qian Yu, Guoan Bi, Liren Zhang and Liping Sun, "Application of Sequential Monte Carlo for Multiuser Detection of DS-CDMA Systems in Fading Channels", *Proc. 40th IEEE International Conference on Communications*, Seoul, Korea, May 2005.

8. Qian Yu, Guoan Bi and Liren Zhang, "Bayesian blind multiuser detection for long code multipath DS-CDMA systems", *Proc. 2nd IEEE International Conference on Communications, Circuits and Systems*, Chengdu, China, Jun. 2004.

9. Qian Yu and Guoan Bi, "Blind multiuser receiver for DS-CDMA systems based on sequential Monte Carlo estimation", *Proc. Asia-Pacific Conference on Circuits and Systems*, Tainan, Taiwan, Dec. 2004.

10. Qian Yu, Guoan Bi and Gaonan Zhang, "Improved blind channel estimation method with Toeplitz displacement for long code DS-CDMA", *Proc. 4th IEEE International Workshop on Signal Processing Advances for Wireless Communications*, Rome, Italy, Jun. 2003.

11. Qian Yu, Guoan Bi and Gaonan Zhang, "Blind multipath estimation with Toeplitz displacement for long code DS-CDMA", *Proc. 38th IEEE International Conference on Communications*, Anchorage, Alaska, USA, May 2003.

Bibliography

[1] W. Lee, "Overview of cellular CDMA," *IEEE Trans. Vehic. Tech.*, vol. 40, pp. 291–302, May. 1991.

[2] R. Pickholtz, L. Milstein, and D. Schilling, "Spread spectrum for mobile communications," *IEEE Trans. Vehic. Tech.*, vol. 40, pp. 313–322, May. 1991.

[3] R. Kohno, R. Meidan, and L. Milstein, "Spread spectrum access methods in wireless communications," *IEEE Communications Magazine*, vol. 33, pp. 58–67, Jan. 1995.

[4] F. Adachi, M. Sawahashi, and H. Suda, "Wideband DS-CDMA for next-generation mobile communication systems," *IEEE Pers. Commun.Mag.*, vol. 36, pp. 56–69, Sept. 1998.

[5] E. D. et al., "WCDMA-the radio interface for future mobile multimedia communications," *IEEE Trans. Veh. Technol.*, pp. 1105–1118, Nov. 1998.

[6] B. Sklar, *Digital Communication*, pp. 382–384. Prentice Hall, 2002.

[7] E. H. Dinan, B. Jabbari, and G. Mason, "Spreading codes for direct sequence CDMA and wideband CDMA cellular networks," *IEEE Commun. Magazine*, vol. 36, pp. 48–54, Sep. 1998.

[8] K. Tang, P. H. Siegel, and L. B. Milstein, "A comparison of long code versus short spreading sequences in coded asynchronous DS-CDMA systems," *IEEE Journal On Selected Areas In Communications*, vol. 19, pp. 1614–1624, Aug. 2001.

[9] J. Kaiser, J. W. Schwarts, and J. M. Aein, "Multiple access to a communication satellite with a hard-limiting repeater," *Tech. Rep. R-108, Institute for Defense Analyses*, vol. I: Modulation techniques and their applications, Jan. 1965.

[10] M. Pursley, "Performance evaluation for phase-coded spread-spectrum multiple-access communication - Part I: System analysis," *IEEE Trans. Commun.*, vol. 25, pp. 795–799, Aug. 1977.

[11] K. Yao, "Error probability of asynchronous spread spetrum multiple access communication systems," *IEEE Transactions on Communications*, vol. 25, pp. 803–809, Aug. 1977.

[12] J. Holtzman, "CDMA power control for wireless networks," *Third Generation Wireless Information Networks*, vol. 32, pp. 299–311, 1992.

[13] A. D. Hallen, J. Holtzman, and Z.Zvonar, "Multiuser detection for CDMA systems," *IEEE Pers. Commun.Mag.*, pp. 45–58, April 1995.

[14] S. Verdu, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Information Theory*, vol. 32, pp. 85–96, Jan. 1986.

[15] D. North, "Analysis of the factors which determine signal/noise discrimination in radar," in *Report PPR-6C, RCA*, (Princeton, NJ), Jun. 1943.

[16] M. Simon, J. Omura, R. Scholtz, and B. Levitt, *Spread Spectrum Communications Handbook*. New York: McGraw Hill, 1994.

[17] J. Savage, "Signal detection in the presence of multiple-access noise," *IEEE Transactions on Information Theory*, vol. 20, pp. 42–49, Jan. 1974.

[18] R. P. B. Vojcic and L. Milstein, "Performance of DS-CDMA with imperfect power control operating over a low earth orbit satellite link," *IEEE Journal On Selected Areas In Communications*, vol. 12, pp. 560 –567, May. 1994.

[19] S. Verdu, "Minimum probability of error for asynchronous multiple access communication systems," in *Proc. 1983 IEEE Conf. Military Communications*, pp. 213–219, Nov. 1983.

[20] Z. Zvonar and D. Brady, "Multiuser detection in single-path fading channels," *IEEE Trans. Communications*, vol. 42, pp. 1729–1739, Feb. 1994.

[21] S. Vasudevan and M. Varanasi, "Optimum diversity combiner based multiuser detection for time-dispersive Rician fading CDMA channels," *IEEE Journal On Selected Areas In Communications*, vol. 12, pp. 580–592, May 1994.

[22] R. Lupas and S. Verdu, "Linear multiuser detectors for synchronous code-division multiple access channels," *IEEE Trans. Information Theory*, vol. 35, pp. 123–136, Jan. 1989.

[23] S. Verdu, "Computational complexity of optimum multiuser detection," *Algorithmica*, vol. 4, pp. 303–312, Mar. 1989.

[24] S. Verdu, *Multiuser detection.* Cambridge University Press, 1998.

[25] R. Lupas and S. Verdu, "Optimum near-far resistance of linear detectors for code-division multiple-access channels," *Abstr. IEEE Int. Symp. Information Theory*, p. 14, Jun. 1988.

[26] R. Lupas and S. Verdu, "Near-far resistance of multiuser detectors in asynchronous channels," *IEEE Trans. Communications*, vol. 38, pp. 496–508, Apr. 1990.

[27] A. Klein and P. Baier, "Linear unbiased data estimation in mobile radio systems applying CDMA," *IEEE Journal On Selected Areas In Communications*, vol. 11, pp. 1058–1066, Sept. 1993.

[28] A. Klein, G. K. Kaleh, and P. Baier, "Zero forcing and minimum mean-square-error equalization for multiuser detection in code-division multiple-access channels," *IEEE Trans. Vehicular Technology*, vol. 45, pp. 276–287, May 1996.

[29] X. Wang and H. Poor, "Blind equalization and multiuser detection in dispersive CDMA channels," *IEEE Trans. Communications*, vol. 46, pp. 91–103, Jan. 1998.

[30] M. Varanasi and B. Aazhang, "Optimally near-far resistant multiuser detection in differentially coherent synchronous channels," *IEEE Trans. Information Theory*, vol. 37, pp. 1006–1018, July 1991.

[31] M. Varanasi, "Noncoherent detection in asynchronous multiuser channels," *IEEE Trans. Information Theory*, vol. 39, pp. 157–176, Jan. 1993.

[32] Z. Zvonar and D. Brady, "Differentially coherent multiuser detection in asynchronous CDMA flat Rayleigh fading channels," *IEEE Trans. Communications*, vol. 43, pp. 1251–1255, Apr. 1995.

[33] H. Liu and Z. Siveski, "Differentially coherent decorrelating detector for CDMA single-path time-varying rayleigh fading channels," in *Proc. 1996 Conf. on Information Sciences and Systems*, pp. 86–89, Mar. 1996.

[34] D. Chen and S. Roy, "An adaptive multiuser receiver for CDMA systems," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 808–816, June 1994.

[35] U. Mitra and H. V. Poor, "Adaptive receiver algorithms for near-far resistant CDMA," *IEEE Trans. Communications*, vol. 43, pp. 1251–1255, Apr. 1995.

[36] U. Mitra and H. V. Poor, "Analysis of an adaptive decorrelating detector for synchronous CDMA channels," *IEEE Trans. Communications*, pp. 257–268, Feb. 1996.

[37] T. Lim and L. Rasmussen, "Adaptive symbol and parameter estimation in asynchronous multiuser CDMA detectors," *IEEE Trans. Communications*, vol. 45, pp. 213–220, Feb. 1997.

[38] S. Ulukus and R. Yates, "A blind adaptive decorrelating detector for CDMA systems," in *Proc. 1997 IEEE Conf. Global Telecommunications*, pp. 664–668, Nov. 1997.

[39] D. Geockel and W. Stark, "Throughput optimization in multiple-access communication systems with decorrelator reception," in *Proc. 1996 Int. Symp. on Information Theory and Its applications*, pp. 653–656, Sept. 1996.

[40] X. Wang and H. Poor, "Adaptive multiuser detection in non-gaussian channels," in *Proc. 35th Allerton Conf. on Communications, Control and Computing*, pp. Monticello, IL, Sept. 1997.

[41] U. Madhow, "Blind adaptive interference suppression for near-far resistant acquisition and demodulation of direct-sequence CDMA signals," *IEEE Trans. Signal Processing*, vol. 45, pp. 124–136, Jan. 1997.

[42] H. V. Poor and X. Wang, "Code-aided interference suppression for DS/CDMA communications - Part II: Parallel blind adaptive implementations," *IEEE Trans. Commun.*, vol. 45, pp. 1112–1122, Sep. 1997.

[43] D. Z. et al., "An efficient code-timing estimator for DS-CDMA signals," *IEEE Trans. Signal Processing*, vol. 45, pp. 82–89, Jan. 1997.

[44] E.-R. Jeong, G. Choi, and Y. H. Lee, "Data-aided frequency estimation for PSK signaling in frequency-selective fading," *IEEE Journal On Selected Areas In Communications*, vol. 19, pp. 1408–1419, Jul. 2001.

[45] Z. Xie, R. Short, and C. Rushforth, "A family of suboptimum detectors for coherent multiuser communications," *IEEE J. Select. Areas Commun.*, vol. 8, pp. 683–690, May 1990.

[46] U. Madhow and M. L. Honig, "MMSE interference suppression for DS/SS CDMA," *IEEE Trans. Commun.*, vol. 42, pp. 3178–3188, Dec. 1994.

[47] M. Rupf, F. Tarkoy, and J. L. Massey, "User-separating demodulation for code-division multiple-access systems," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 786–795, June 1994.

[48] T. K. Kashihara, "Adaptive cancellation of mutual interference in spread spectrum multiple access," in *Proc. 1980 IEEE Conf. Int. Communications*, pp. 44.4.1–5, 1980.

[49] S. L. Miller, "An adaptive direct-sequence code-division multiple access receiver for multiuser interference rejection," *IEEE Trans. Commun.*, vol. 43, pp. 1746–1755, Feb. 1995.

[50] P. Rapajic and B. Vucetic, "Adaptive receiver structures for asynchronous CDMA systems," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 685–697, May 1994.

[51] M. L. Honig, U. Madhow, and S. Verdu, "Blind adaptive multiuser detection," *IEEE Trans. Inform. Theory*, vol. 41, pp. 944–960, 1995.

[52] M. L. Honig, M.Sensha, S. Miller, and L. Milstein, "Performance of adaptive linear interference suppression for DS-CDMA in the presence of flat Rayleigh fading," in *Proc. 47th IEEE Conf. on Vehicular Technology*, pp. 2191–2195, May 1997.

[53] H. Huang and S. Verdu, "Linear differentially coherent multiuser detection for multipath channels," *Wireless Personal Communication*, vol. 6, pp. 113–136, Jan. 1998.

[54] S. Miller, M. L. Honig, M.Sensha, and L. Milstein, "MMSE reception of DS-CDMA for frequency-selective fading channels," in *Proc. 1997 IEEE Int. Symp. on Information Theory*, p. 51, July 1997.

[55] R. Kohno, P. Rapajic, and B. Vucetic, "An overview of adaptive techniques for interference minimization in CDMA systems," *Wireless Personal Communications*, vol. 1, no. 1, pp. 3–21, 1994.

[56] Z. Xie, C. K. Rushforth, R. Short, and T. K. Moon, "Joint signal detection and parameter estimation in multiuser communications," *IEEE Trans. Commun.*, vol. 41, pp. 1208–1216, Aug. 1993.

[57] E.-W. Bai and Y. Ye, "The least squares: output error sensitivity and the constrained logarithmic algorithm," in *American Control Conference, 1998. Proceedings of the 1998*, vol. 6, pp. 3570 –3574, 1998.

[58] S. Haykin, *Adaptive Filter Theory.* Prentice Hall, 1991.

[59] A. Y. Wu and K. J. R. Liu, "Split recursive least-squares: algorithms, architectures, and applications; Circuits and Systems II: Analog and digital signal processing," *IEEE Trans. Commun.*, vol. 43, pp. 645 –658, Sep. 1996.

[60] A. M. Haimovich and Y. Bar-Ness, "An eigenanalysis interference canceller," *IEEE Trans. Signal Processing*, vol. 39, pp. 76–84, Jan. 1991.

[61] E. Strom and S. Miller, "Optimum complexity reduction of minimum mean square error DS-CDMA receivers," in *Proc. IEEE Conf. Vehicular Technology*, pp. 568–572, June 1994.

[62] M. L. Honig, "Performance of adaptive interference suppression for DS-CDMA with a time-varying user population," in *Proc. 41th Int. Symp. on Spread Spetrum Techniques and applications*, pp. 267–271, Sept. 1996.

[63] Y. Bar-Ness, W. Chen, and E. Panayiric, "Eigenanalysis for interference cancellation with minimum redundancy array structure," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, pp. 977–988, July 1997.

[64] A. Veen, S. Talwar, and A. Paulraj, "A subspace approach to blind space-time signal processing for wireless communication systems," *IEEE Trans. Signal Processing*, vol. 45, pp. 173–190, Jan. 1997.

[65] H. Liu and M. D. Zoltowski, "Blind equalization in antenna array CDMA systems," *IEEE Transactions on Signal Processing*, vol. 45, no. 1, pp. 161–172, Jan. 1997.

[66] X. Wang and H. Poor, "Blind multiuser detection : A subspace approach," *IEEE Trans. Information Theory*, vol. 44, pp. 677–690, Mar. 1998.

[67] P. Dent, B. Gudmundson, and M. Ewerbring, "CDMA-IC: A novel code division multiple access scheme based on inference interference cancellation," in *Proc. 3rd IEEE Int. Symp. Personal, Indoor and Mobile Radio Communications*, pp. 98–102, Sept. 1992.

[68] Y. Yoon, R. Kohno, and H. Imai, "A spread-spectrum multi-access system with co-channel interference cancellation over multipath fading channels," *IEEE J. Select. Areas Commun.*, vol. 11, pp. 1067–1075, Sept. 1993.

[69] P. Patel and J. Holtzman, "Aanalysis of a DS/CDMA successive interference cancellation scheme using corelations," in *Proc. 1997 IEEE Conf. Global Telecommunications*, pp. 76–80, Nov. 1993.

[70] J. Holtzman, *Code Division Multiple Access Communications*, ch. DS/CDMA successive interference cancellation, pp. 161–182. Dordrecht, Netherlands: Kluwer Academic, 1995.

[71] D. Divsalar, M. Simon, and D. Raphaeli, "Improved parallel interference cancellation for CDMA," *IEEE Trans. Communications*, vol. 46, pp. 258–268, Feb. 1998.

[72] M. Varanasi and B. Aazhang, "Near-optimally detection in synchronous code-division multiple-access systems," *IEEE Trans. Communications*, vol. 39, pp. 725–735, May 1991.

[73] M. Varanasi and B. Aazhang, "Multistage detections in asynchronous code-division multiple access communications," *IEEE Trans. Communications*, vol. 38, pp. 509–519, Apr. 1990.

[74] R. Kohno, H. Imai, M. Hatori, and S. Pasupathy, "An adaptive canceller of co-channel interference for spread-spectrum multi-access communication networks in a power line," *IEEE Journal On Selected Areas In Communications*, vol. 8, pp. 691–699, Apr. 1990.

[75] A. D. Hallen, "Decorrelating decision-feedback multiuser detectors for synchronous CDMA," *IEEE Trans. Communications*, vol. 41, pp. 285–290, Feb. 1993.

[76] A. D. Hallen, "A family of multiuser decision-feedback detectors for asynchronous code-division multiple access channels," *IEEE Trans. Communications*, vol. 43, pp. 421–434, Feb. 1995.

[77] J. Yang and S. Roy, "Joint transmitter/reveiver optimization for multi-input multi-output systems with decision feedback," *IEEE Trans. Information Theory*, vol. 40, pp. 1334–1347, Sept. 1994.

[78] C. Tidestav, A. Ahlen, and M. Sternad, "Narrowband and broad-band multiuser detection using a multivariable DFE," in *Proc. 3rd IEEE Int. Symp. Personal, Indoor and Mobile Radio Communications*, vol. 2, pp. 732–736, Sept. 1995.

[79] X. Wang, R. Chen, and J. S. Liu, "Monte Carlo Bayesian signal processing for wireless communication," *Journal of VLSI Signal Processing*, vol. 30, pp. 89–105, 2002.

[80] G. E. Box and G. C. Tiao, *Bayesian Inference in Statistical Analysis.* Addison-Wesley, 1973.

[81] D. J. C. Mackay, *Learning in Graphical Models, M. I. Jordan*, ch. Introduction to Monte Carlo methods. MIT Press, 1999.

[82] X. Wang and R. Chen, "Adaptive Bayesian multiuser detection for synchronous CDMA with Gaussian and impulsive noise," *IEEE Transaction on Signal Processing*, vol. 47, no. 7, pp. 2013–2028, 2000.

[83] Z. Yang, B. Lu, and X. Wang, "Bayesian Monte Carlo multiuser receiver for space-time coded multicarrier CDMA systems," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 8, pp. 1625–1637, 2001.

[84] Z. Yang and X. Wang, "Blind turbo multiuser detection for long-code multipath CDMA," *IEEE Transactions on Communications*, vol. 50, no. 1, pp. 112–125, Jan. 2002.

[85] Z. Tian, "A high-efficient Monte Carlo receiver for digital communications," in *Proc. IEEE International Conference on Communications*, vol. 3, pp. 1471–1475, 2002.

[86] J. S. Liu and R. Chen, "Sequential Monte Carlo methods for dynamic systems," *J. Amer. Statist. Assoc.*, vol. 93, pp. 1032–1044, 1998.

[87] A. Doucet, N. de Freitas, and N. Gordon, *Sequential Monte Carlo methods in practice.* Springer Press, 2000.

[88] R. Chen, X. Wang, and J. S. Liu, "Adaptive joint detection and decoding in flat-fading channels via mixture Kalman filtering," *IEEE Transaction on Information Theory*, vol. 46, no. 6, pp. 2079–2094, Sept. 2000.

[89] Z. Yang and X. Wang, "A sequential Monte Carlo blind receiver for OFDM systems in frequency-selective fading channels," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 271–280, Feb. 2002.

[90] D. Guo and X. Wang, "Blind detection in MIMO systems via sequential Monte Carlo," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 3, pp. 464–473, 2003.

[91] R. A. Iltis, "A sequential Monte Carlo filter for joint linear/nonlinear state estimation with application to DS-CDMA," *IEEE Transactions on Signal Processing*, vol. 51, no. 2, pp. 417–426, 2003.

[92] C. Andrieu, A. Doucet, and A. Touzni, "Adaptive MAP multiuser detection for fading CDMA channels," in *Proceedings of the 10th IEEE workshop on Statistical Signal and Array Processing*, pp. 6–9, Aug. 2000.

[93] E. Punskaya, C. Andrieu, A. Doucet, and W. J. Fitzgerald, "Particle filtering for multiuser detection in fading CDMA channels," in *Proceedings of the 11th IEEE Signal Processing workshop on Statistical Signal Processing*, pp. 38–41, Aug. 2001.

[94] J. Zhang, Y. Huang, and P. M. Djuric, "Multiuser detection with particle filtering," in *Proceedings of the 11th European Signal Processing*, Sept. 2002.

[95] P. M. Djuric, J. H. Kotecha, J. Zhang, Y. Huang, T. Ghirmai, M. F. Bugallo, and J. Miguez, "Particle filtering," *IEEE Signal Processing Magazine*, vol. 20, no. 5, pp. 19–38, Sept. 2003.

[96] W. Chen and U. Mitra, "An improved blind adaptive MMSE receiver for fast fading DS-CDMA channels," *IEEE Journal On Selected Areas In Communications*, vol. 19, pp. 758–762, Aug. 2001.

[97] A. J. Weiss and B. Friedlander, "Channel estimation for DS-CDMA downlink with aperiodic spreading codes," *IEEE Trans. Commun.*, vol. 47, pp. 1561–1569, Oct. 1999.

[98] M. Torlak, B. L. Evans, and G. Xu, "Blind estimation of FIR channels in CDMA systems with aperiodic spreading sequences," in *Proc. 31st Asilomar Conf. Signals, Syst., Comput.*, vol. 1, pp. 495–499, Nov. 1997.

[99] Z. Xu and M. K. Tsatsanis, "Blind channel estimation for long code multiuser CDMA systems," *IEEE Trans. Signal Processing*, vol. 48, pp. 988–1001, Apr. 2000.

[100] C. Escudero, U. Mitra, and D. Slock, "A Toeplitz displacement method for blind multipath estimation for long code DS/CDMA signals," *IEEE Trans. Signal Processing*, vol. 49, pp. 654–665, Mar. 2001.

[101] V. Tripathi, A. Mantravadi, and V. Veeravalli, "Channel acquisition for wideband CDMA signals," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 1483–1495, Aug. 2000.

[102] S. Buzzi and H. Poor, "Channel estimation and multiuser detection in long-code DS/CDMA systems," *IEEE Journal On Selected Areas In Communications*, vol. 19, pp. 1476–1487, Aug. 2001.

[103] S. Buzzi and H. Poor, "On parameter estimation in long-code DS/CDMA systems: Cramer-Rao bounds and least-squares algorithms," *IEEE Trans. Signal Processing*, vol. 51, pp. 545–559, Feb. 2003.

[104] A. J. Viterbi, *CDMA: principles of spread spectrum communication*. Addison-Wesley, 1995.

[105] W. Gilks, S. Richardson, and D. Spiegelhalter, *Markov Chain Monte Carlo in Practice*. Chapman and Hall, 1996.

[106] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, pp. 97–109, 1970.

[107] K. S. Chan, "Asymptotic behavior of the Gibbs sampler," *J. Amer. Stat. Assoc.*, vol. 88, pp. 320–326, 1993.

[108] J. S. Liu, A. Kong, and W. H. Wong, "Covariance structure of the Gibbs sampler with applications to the comparisons of estimators and augmentation schemes," *Biometrika*, vol. 81, pp. 27–40, 1994.

[109] R. Chen, J. S. Liu, and X. Wang, "Convergence analyses and comparisons of Markov chain Monte Carlo algorithms in digital communications," *IEEE Transaction on Signal Processing*, vol. 50, no. 2, pp. 255–270, Feb. 2002.

[110] E. Punskaya, C. Andrieu, and A. Doucet, "Particle filtering for demodulation in fading channels with non-Gaussian additive noise," *IEEE Trans. Commun.*, vol. 49, no. 4, pp. 579–582, Apr. 2001.

[111] B. Dong, X. Wang, and A. Doucet, "A new class of soft MIMO demodulation algorithms," *IEEE Transaction on Signal Processing*, vol. 51, no. 11, pp. 2752–2763, Nov. 2002.

[112] M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," *IEEE Transaction on Acoustics, Speech, and Signal Processing*, vol. 36, no. 4, pp. 477–489, 1988.

[113] L. B. Nelson and V. Poor, "Iterative multiuser receivers for CDMA channels: An EM-based approach," *IEEE Transactions on Communications*, vol. 44, no. 12, pp. 1700–1710, 1996.

[114] A. Kocian and B. H. Fleury, "EM-based joint data detection and channel estimation of DS-CDMA signals," *IEEE Transactions on Communications*, vol. 51, no. 10, pp. 1709–1720, 2003.

[115] J. F. Geweke, "Bayesian inference in econometric models using Monte Carlo integration," *Econometrica*, vol. 24, pp. 1317–1399, 1989.

[116] J. S. Liu and R. Chen, "Blind deconvolution via sequential imputations," *J. Amer. Statist. Assoc.*, vol. 90, pp. 567–576, 1995.

[117] N. M. Laird, A. P. Dempster, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Ann. Roy. Stat. Soc.*, pp. 1–38, Dec. 1977.

[118] J. A. Fessler and A. O. Hero, "Complete-data spaces and generalized EM algorithm," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, pp. IV1–IV4 2535–2538, Apr. 1993.

[119] K. J. Kim and R. A. Iltis, "Joint detection and channel estimation algorithms for QS-CDMA signals over time-varying channels," *IEEE Transactions on Communications*, vol. 50, no. 5, pp. 845–855, 2002.

Appendix

The derivation of the importance weight

According to the (5.18), the importance weight should be updated by

$$
\begin{aligned}
w_{k,n}^{(j)} &= w_{k,n-1}^{(j)} \frac{p(\mathbf{S}_{k,n}^{(j)}|\mathbf{X}_{k,n})}{p(\mathbf{S}_{k,n-1}^{(j)}|\mathbf{X}_{k,n-1})p(s_k^{(j)}(n)|\mathbf{S}_{k,n-1}^{(j)},\mathbf{X}_{k,n})} \\
&= w_{k,n-1}^{(j)} \frac{p(\mathbf{S}_{k,n}^{(j)}|\mathbf{X}_{k,n})p(\mathbf{X}_{k,n})}{p(\mathbf{S}_{k,n-1}^{(j)}|\mathbf{X}_{k,n-1})p(s_k^{(j)}(n)|\mathbf{S}_{k,n-1}^{(j)},\mathbf{X}_{k,n})p(\mathbf{X}_{k,n})} \\
&= w_{k,n-1}^{(j)} \frac{p(s_k^{(j)}(n),\mathbf{S}_{k,n-1}^{(j)},\mathbf{X}_{k,n})}{p(\mathbf{S}_{k,n-1}^{(j)}|\mathbf{X}_{k,n-1})p(s_k^{(j)}(n)|\mathbf{S}_{k,n-1}^{(j)},\mathbf{X}_{k,n})p(\mathbf{X}_{k,n})} \\
&= w_{k,n-1}^{(j)} \frac{p(\mathbf{S}_{k,n-1}^{(j)},\mathbf{X}_{k,n})}{p(\mathbf{S}_{k,n-1}^{(j)}|\mathbf{X}_{k,n-1})p(\mathbf{X}_{k,n})} \\
&= w_{k,n-1}^{(j)} \frac{p(\mathbf{S}_{k,n-1}^{(j)}|\mathbf{X}_{k,n})}{p(\mathbf{S}_{k,n-1}^{(j)}|\mathbf{X}_{k,n-1})} \\
&= w_{k,n-1}^{(j)} \frac{p(\mathbf{S}_{k,n-1}^{(j)}|\mathbf{X}_{k,n})p(\mathbf{X}_{k,n})p(\mathbf{X}_{k,n-1})}{p(\mathbf{S}_{k,n-1}^{(j)}|\mathbf{X}_{k,n-1})p(\mathbf{X}_{k,n-1})p(\mathbf{X}_{k,n})} \\
&= w_{k,n-1}^{(j)} \frac{p(\mathbf{S}_{k,n-1}^{(j)},\mathbf{X}_{k,n-1},\mathbf{X}_{k,n})p(\mathbf{X}_{k,n-1})}{p(\mathbf{S}_{k,n-1}^{(j)},\mathbf{X}_{k,n-1})p(\mathbf{X}_{k,n})} \\
&\propto w_{k,n-1}^{(j)} p(\mathbf{X}_{k,n}|\mathbf{S}_{k,n-1}^{(j)},\mathbf{X}_{k,n-1})
\end{aligned}
$$