Multiuser Detection in Fading Channels

H.V. Poor   Daryl Reynolds   Xiaodong Wang
## Contents

1 Introduction ........................................... 1

2 Blind Adaptive Subspace Multiuser Detection in Slow Fading Channels 2
   2.1 System Descriptions and Blind Multiuser Detectors .................. 2
   2.1.1 Synchronous Multipath CDMA System ................................ 2
   2.1.2 Synchronous Multi-antenna CDMA System ........................... 4
   2.1.3 Asynchronous Multi-antenna Multipath CDMA ....................... 5
   2.2 Performance of Blind Multiuser Detectors ............................ 9
   2.2.1 Complex Gaussian Distribution ...................................... 9
   2.2.2 Performance of Blind Multiuser Detectors with Known Channels 10
   2.2.3 Performance of Blind Multiuser Detector with Blind Channel Estimation ......................................................... 12
   2.2.4 Numerical Results ..................................................... 13
   2.2.5 Adaptive Implementation .............................................. 14
   2.2.6 Algorithm Summary .................................................... 16

3 Bayesian Multiuser Detection for Long-code CDMA 17
   3.1 System Descriptions ................................................... 17
   3.1.1 Channel Model ....................................................... 17
   3.1.2 Noise Model ......................................................... 18
   3.1.3 Blind Bayesian Multiuser Detection ............................... 19
   3.1.4 The Gibbs Sampler ................................................... 20
   3.2 Bayesian MCMC Multiuser Detectors ................................... 21
   3.2.1 White Gaussian Noise ................................................ 21
   3.2.2 Colored Gaussian Noise ............................................. 23
   3.3 Simulation Examples ................................................... 25

4 Multiuser Detection for Long-code CDMA in Fast-Fading Channels 29
   4.1 Channel Model and Sequential EM Algorithm .......................... 29
   4.2 Sequential Blind Multiuser Detector .................................. 31
   4.3 Simulation Results ...................................................... 31

5 Transmitter-Based Multiuser Precoding for Fading Channels 32
   5.1 Basic Approach and Adaptation ........................................ 34
   5.1.1 Uplink Signal Model and Blind Channel Estimation ............... 34
   5.1.2 Downlink Signal Model and Matched Filter Detection ............. 35
   5.1.3 Transmitter Precoding for a Synchronous Multipath Downlink .... 36
   5.1.4 Adaptive Implementation ............................................ 38
   5.1.5 Algorithm Summary ................................................ 39
   5.2 Precoding with Multiple Transmit Antennas ............................ 40
   5.2.1 Downlink Signal Model .............................................. 40
   5.2.2 Precoder Design for Orthogonal Spreading Codes .................. 40
   5.2.3 Precoder Design for Non-Orthogonal Spreading Codes ............. 41
   5.3 Precoding for Multipath ISI Channels ................................ 42
   5.3.1 Prerake-Diversity Combining ....................................... 43
   5.3.2 Precoder Design .................................................... 43
5.4 Performance Analyses ................................................. 46
  5.4.1 Performance of Transmitter Precoding with Blind Channel Estimation 46
  5.4.2 Performance and Achievable Diversity for Multi-antenna Precoding 47

6 Conclusion .......................................................... 50
1 Introduction

The three primary characteristics that distinguish wireless communications from wireline communications are dynamism, fading and interference. In this chapter, we address all of these three characteristics by considering adaptive methods for multiuser detection in fading channels.

Fading refers to variations in the gain of a communications channel. The primary source of fading is multipath, which leads to constructive and destructive self-interference of signals at a communications receiver. Depending on various properties of the communication link (bandwidth, mobility, etc.) fading can vary with time (i.e., time-selective fading) and frequency (frequency-selective fading). Of course, given the fact that the multipath profile depends on geometry, fading is also dependent on position and angle of arrival.

A key issue in the mitigation of fading is that the fading is caused by the channel, which is not known a priori to the receiver. There are two approaches to dealing with this issue. One is to treat the channel gains as random quantities with known distributions (e.g., Rayleigh, Rician, etc.) and to devise optimal receiver algorithms based on this statistical model. Examples of this approach can be found in [32, 59, 45]. An alternative approach, which is considered here, is for the receiver to use the outputs of the channel to adapt to the fading. I.e., adaptive systems essentially estimate the channel and then use the channel estimates in a multiuser receiver corresponding to a known channel. This approach has the advantage of also allowing adaptation to other unknown quantities, such as interference, and also has performance advantages over systems that simply treat the fading as being random and unknowable. A general overview of adaptive receivers for multiuser systems can be found in [52].

In this chapter, we address specifically the problem of blind adaptation in which the system does not make use of training symbols, but rather learns the optimal receiver directly from data-modulated signals. This approach has the obvious advantage of not requiring additional overhead to transmit training data. We consider several aspects of this problem, including the quasi-static case, in which the fading can be considered to be constant over a processing window, and the fast-fading (or time-selective) case, in which the fading can change within the processing window. We also consider both frequency-flat and frequency-selective cases, and we consider the use of multiple receive antennas. Performance issues are also addressed. We begin in Section 2 with a general approach to this problem based on subspace methods. Here, we see that such techniques provide a powerful framework for addressing a variety of issues arising in fading multiuser channels. One limitation of many adaptive techniques, including these subspace methods, is that are applicable primarily to systems in which the signaling multiplex (e.g., the spreading codes in a code-division multiple-access (CDMA) system) is time-invariant. In many systems, however, this condition is violated (e.g., in so-called ”long code” CDMA systems). Thus, alternative methods are often needed, and in Section 3, we consider methods for such problems based on Markov chain Monte Carlo (MCMC) techniques. This provides a further powerful technique for the treatment of fading in multiuser settings. In Section 4 we turn to the problem of fast fading, to which a sequential expectation-maximization (EM) approach can be successfully applied. Finally, in Section 5, we consider situations in which it is advantageous to move computational complexity from the receiver to the transmitter, e.g., a cellular downlink. In particular, we develop transmitter precoding strategies that are designed to jointly combat multiple access interference and exploit any available transmit
antenna or multipath diversity. We include analyses of achievable diversity and SINR/BER performance for the adaptive case.

2 Blind Adaptive Subspace Multiuser Detection in Slow Fading Channels

2.1 System Descriptions and Blind Multiuser Detectors

In what follows, we present the signal models as well as the corresponding blind multiuser detection algorithms for several systems. We first describe two simple systems, namely, the synchronous multipath CDMA system with no intersymbol interference (ISI), and the synchronous multi-antenna CDMA system. We then discuss the general asynchronous multipath multi-antenna CDMA system. As will be seen, all the three systems considered here share essentially the same signal model and the same blind multiuser detection algorithm.

2.1.1 Synchronous Multipath CDMA System

We start by considering a $K$-user discrete-time synchronous multipath CDMA system with no ISI. Such a system is realized by either neglecting the ISI when the multipath delay spread is small compared with the symbol interval, or by inserting guard intervals between symbols when the delay spread is large. The received $N$-dimensional signal during the $i$-th symbol interval in such a system can be written as [7, 48]

$$
r[i] = \sum_{k=1}^{K} b_k[i] \sum_{l=1}^{L} s_{l,k} h_{l,k} + n[i], 
$$

where $L$ is the number of resolvable paths, $s_{l,k}$ and $h_{l,k}$ are respectively the delayed version of the spreading waveform (with zero-padding when a guard interval is inserted) and the complex channel fading gain corresponding to the $l$-th path of the $k$-th user; $n[i] \sim \mathcal{N}(0, \eta I_N)$ is the circularly symmetric complex white Gaussian noise vector. Denote

$$
S_k \triangleq [s_{1,k}, s_{2,k}, \ldots, s_{L,k}], \\
h_k \triangleq [h_{1,k}, h_{2,k}, \ldots, h_{L,k}]^T.
$$

Then (2) can be rewritten as

$$
r[i] = \sum_{k=1}^{K} \tilde{S} h_k b_k[i] + n[i] 
$$

$$
= \tilde{S} b[i] + n[i],
$$

where

$$
\tilde{S} \triangleq [\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_K], \\
b[i] \triangleq [b_1[i], b_2[i], \ldots, b_K[i]]^T.
$$
Let the autocorrelation matrix of the received signal \( r[i] \) be

\[
C_r \triangleq E[r[i]r[i]^H] = \tilde{S}\tilde{S}^H + \eta I_N
\]

\[
= U_s\Lambda_sU_s^H + \eta U_nU_n^H, \tag{4}
\]

where in (5), \( \Lambda_s = \text{diag}(\lambda_1, \cdots, \lambda_K) \) contains the largest \( K \) eigenvalues of \( C_r; \) \( U_s = [u_1, \cdots, u_K] \) contains the eigenvectors corresponding to the eigenvalues in \( \Lambda_s; \) and \( U_n = [u_{K+1}, \cdots, u_N] \) contains the \((N - K)\) eigenvectors corresponding to the smallest eigenvalue \( \eta \) of \( C_r. \) Suppose that user 1 is the user of interest. Since \( \tilde{s}_1 \triangleq S_1h_1, \) and \( U_n^H\tilde{s}_1 = 0, \) it then follows that \( h_1 \) can be obtained from the following relationship

\[
h_1 = \arg \min_{\|h\|=1} \|U_n^H(S_1h)\|^2
\]

\[
= \arg \min_{\|h\|=1} h^H(S_n^H U_n U_n^H S_1) h
\]

\[
= \text{minimum eigenvector of } Q. \tag{6}
\]

Note that, however, (6) specifies \( h_1 \) only up to a phase ambiguity, i.e., if \( h_1 \) is the solution to (6), so is \( e^{j\phi}h_1 \) for any \( \phi. \) It is clear that \( h_1 \) is uniquely determined by (6) up to a phase ambiguity if and only if \( \text{rank}(Q) = L - 1. \) Moreover, the constraint \( \|h\| = 1 \) also implies a scale ambiguity.

The (exact) linear MMSE detector for user 1 given by [48, 49, 50]

\[
w_1 = C_r^{-1}\tilde{s}_1
\]

\[
= U_s\Lambda_s^{-1}U_s^H\tilde{s}_1. \tag{7}
\]

In blind multiuser detection, it is assumed that the receiver has only the knowledge of the spreading sequence of the desired user, \( S_1, \) and the detector is estimated from the received signals as follows. First the sample autocorrelation of the received signals is formed, and its eigendecomposition is computed:

\[
\hat{C}_r \triangleq \frac{1}{M} \sum_{i=0}^{M-1} r[i]r[i]^H
\]

\[
= \hat{U}_s\hat{\Lambda}_s\hat{U}_s^H + \hat{U}_n\hat{\Lambda}_n\hat{U}_n^H. \tag{9}
\]

Then a channel estimation is performed by replacing the quantities in (6) by their estimates, i.e.,

\[
\hat{Q} = S_1^H\hat{U}_n\hat{U}_n^H S_1, \tag{11}
\]

\[
\hat{h}_1 = \text{minimum eigenvector of } \hat{Q}. \tag{12}
\]

Finally, the estimated blind linear MMSE detector is given by

\[
\hat{w}_1 = \hat{U}_s\hat{\Lambda}_s^{-1}\hat{U}_s^H S_1\hat{h}_1. \tag{13}
\]
2.1.2 Synchronous Multi-antenna CDMA System

We next consider a $K$-user synchronous CDMA system employing $P$ receive antennas. Let $h_{p,k}$ be the complex fading gain between the transmit antenna and the $p$-th receive antenna for the $k$-th user. At the $p$-th receive antenna, the received discrete-time signal during the $i$-th symbol interval is given by

$$r^{(p)}[i] = \sum_{k=1}^{K} h_{p,k} b_k[i] s_k + n^{(p)}[i], \quad p = 1, 2, \ldots, P,$$

(14)

where $s_k$ is the spreading waveform of the $k$-th user; $n^{(p)}[i] \sim \mathcal{C}(0, \eta I_N)$ is the circularly symmetric complex white Gaussian noise vector at antenna $p$. It is assumed that the noise vectors at different antennas are independent.

Denote

$$h_k \triangleq [h_{1,k} \ h_{2,k} \ \cdots \ h_{P,k}]^T,$$

$$\tilde{s}_k \triangleq h_k \otimes s_k,$$

$$r[i] \triangleq \begin{bmatrix} r^{(1)}[i]^T & r^{(2)}[i]^T & \cdots & r^{(P)}[i]^T \end{bmatrix}^T,$$

$$n[i] \triangleq \begin{bmatrix} n^{(1)}[i]^T & n^{(2)}[i]^T & \cdots & n^{(P)}[i]^T \end{bmatrix}^T,$$

$$b[i] \triangleq \begin{bmatrix} b_1[i] & b_2[i] & \cdots & b_K[i] \end{bmatrix}^T,$$

$$\tilde{S} \triangleq \begin{bmatrix} \tilde{s}_1 & \tilde{s}_2 & \cdots & \tilde{s}_K \end{bmatrix}.$$ 

Then (14) can be written as

$$r[i] = \sum_{k=1}^{K} b_k[i] \tilde{s}_k + n[i]$$

(15)

$$= \tilde{S} b[i] + n[i],$$

(16)

with $n \sim \mathcal{C}(0, \eta I_{PN})$.

Let the autocorrelation matrix of the received signal and its eigendecomposition be

$$C_r \triangleq E[r[i]r[i]^H] = \tilde{S}\tilde{S}^H + \eta I_{PN}$$

(17)

$$= U_s \Lambda_s U_s^H + \eta U_n U_n^H,$$

(18)

where $\Lambda_s = \text{diag}(\lambda_1, \cdots, \lambda_K)$ contains the largest $K$ eigenvalues of $C_r$; $U_s = [u_1, \cdots, u_K]$ contains the eigenvectors corresponding to the eigenvalues in $\Lambda_s$; and $U_n = [u_{K+1}, \cdots, u_{PN}]$ contains the $(PN - K)$ eigenvectors corresponding to the smallest eigenvalue $\eta$ of $C_r$. The linear MMSE detector for user 1 is then given by

$$w_1 = C_r^{-1} \tilde{s}_1$$

(19)

$$= U_s \Lambda_s^{-1} U_s^H \tilde{s}_1.$$ 

(20)

Since $\tilde{s}_1 \triangleq h_1 \otimes s_1$, and $U_n^H \tilde{s}_1 = 0$, it then follows that $h_1$ satisfies

$$h_1 = \arg \min_{\|h\|=1} \|U_n^H (h \otimes s_1)\|^2$$

4
\[
\text{arg min}_{\|h\|=1} h^H Q h = \text{minimum eigenvector of } Q,
\] (21)

with \[ Q \triangleq (I_P \otimes s_1^H) U_n U_n^H (I_P \otimes s_1). \] (22)

Note that \( h_1 \) is determined uniquely by (21) up to a scale and phase ambiguity if and only if \( \text{rank}(Q) = P - 1 \).

A blind estimate of the linear MMSE detector in this case is then given by the following procedure.

\[
\hat{C}_r \triangleq \frac{1}{M} \sum_{i=0}^{M-1} r[i] r[i]^H
\] (23)

\[
= \hat{U}_s \hat{A}_s^H \hat{U}_s^H + \hat{U}_n \hat{A}_n \hat{U}_n^H,
\] (24)

\[
\hat{Q} = (I_P \otimes s_1^H) \hat{U}_n \hat{U}_n^H (I_P \otimes s_1),
\] (25)

\[
\hat{h}_1 = \text{minimum eigenvector of } \hat{Q},
\] (26)

\[
\hat{w}_1 = \hat{U}_s \hat{A}_s^{-1} \hat{U}_s^H (\hat{h}_1 \otimes s_1).
\] (27)

2.1.3 Asynchronous Multi-antenna Multipath CDMA

We now consider a general asynchronous CDMA system with \( K \) users signaling through their respective multipath channels and employing \( P \) receive antennas. We start with the continuous-time signal model. Let the channel impulse response between the \( k \)-th user’s transmitter and the \( p \)-th receive antenna be

\[
g_k^{(p)}(t) = \sum_{l=1}^{L} \alpha_{l,k}^{(p)} \delta \left( t - \tau_{l,k}^{(p)} \right),
\] (28)

where \( L \) is the total number of paths in the channel; \( \alpha_{l,k}^{(p)} \) and \( \tau_{l,k}^{(p)} \) are respectively the complex path gain and the delay of the \( k \)-th user’s \( l \)-th path corresponding to the \( p \)-th receive antenna, \( \tau_{1,k}^{(p)} < \tau_{2,k}^{(p)} < \cdots < \tau_{L,k}^{(p)} \). The received continuous-time signal at the \( p \)-th receive antenna is given by

\[
r^{(p)}(t) = \sum_{k=1}^{K} \sum_{i=0}^{M-1} b_k[i] \left\{ s_k(t - iT) \ast g_k^{(p)}(t) \right\} + n^{(p)}(t)
\]

\[
= \sum_{k=1}^{K} \sum_{i=0}^{M-1} b_k[i] \sum_{l=1}^{L} \alpha_{l,k}^{(p)} s_k \left( t - iT - \tau_{l,k}^{(p)} \right) + n^{(p)}(t),
\] (29)

where \( \ast \) denotes convolution and \( s_k(t) \) is the spreading waveform of the \( k \)-user given by

\[
s_k(t) = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} s_{j,k} \psi(t - j T_c), \quad 0 \leq t < T,
\] (30)

where \( N \) is the processing gain; \( \{s_{j,k}\}_{j=0}^{N-1} \) is a signature sequence of \( \pm 1 \)’s assigned to the \( k \)-th user; and \( \psi(\cdot) \) is a chip waveform of duration \( T_c = T/N \) and with unit energy, i.e.,

\[
\int_0^{T_c} \psi(t)^2 dt = 1.
\]
At the receiver, the received signal \( r^{(p)}(t) \) is filtered by a chip-matched filter and sampled at the chip-rate. Let

\[
\tau = \max_{1 \leq k \leq K, 1 \leq p \leq P} \left\{ \frac{\tau_{L,k} + T_c}{T} \right\},
\]

be the maximum delay spread in terms of symbol intervals. Substituting (30) into (29), the \( q \)-th signal sample during the \( i \)-th symbol is given by

\[
r^{(p)}_q[i] = \int_{iT+qT_c}^{iT+(q+1)T_c} r^{(p)}(t) \psi(t - iT - qT_c) dt
\]

\[
= \int_{iT+qT_c}^{iT+(q+1)T_c} \sum_{k=1}^{K} \sum_{m=0}^{M-1} b_k[m] \sum_{l=1}^{L} \alpha_{l,k}^{(p)} \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} s_{j,k} \left( t - mT - \tau_{l,k}^{(p)} - jT_c \right) dt + n^{(p)}_q[i]
\]

\[
= \sum_{k=1}^{K} \sum_{m=0}^{i} b_k[i - m] \sum_{j=0}^{N-1} s_{j,k} \frac{1}{\sqrt{N}} \sum_{l=1}^{L} \alpha_{l,k}^{(p)} \int_{iT+qT_c}^{iT+(q+1)T_c} \psi(t) \psi\left( t - \tau_{l,k}^{(p)} + mT - jT_c + qT_c \right) dt + n^{(p)}_q[i], \quad (31)
\]

where \( n^{(p)}_q[i] = \int_{iT+qT_c}^{iT+(q+1)T_c} n^{(p)}(t) \psi(t - iT - qT_c) dt \). Denote

\[
r^{(p)}[i] \triangleq \begin{bmatrix} r^{(p)}_0[i] \\ \vdots \\ r^{(p)}_{N-1}[i] \end{bmatrix}, \quad b[i] \triangleq \begin{bmatrix} b_1[i] \\ \vdots \\ b_K[i] \end{bmatrix}, \quad n^{(p)}[i] \triangleq \begin{bmatrix} n^{(p)}_0[i] \\ \vdots \\ n^{(p)}_{N-1}[i] \end{bmatrix},
\]

\[
\tilde{s}^{(p)}[j] \triangleq \begin{bmatrix} \tilde{s}_1^{(p)}[jN] \\ \vdots \\ \tilde{s}_K^{(p)}[jN + N - 1] \end{bmatrix}, \quad j = 0, \cdots, \iota.
\]

Then (31) can be written in terms of vector convolution as

\[
\tilde{r}^{(p)}[i] = \tilde{S}^{(p)}[i] \ast b[i] + \tilde{n}^{(p)}[i], \quad p = 1, 2, \cdots, P.
\] (32)

By stacking \( m \) successive sample vectors, we define the following quantities

\[
\tilde{r}^{(p)}[i] \triangleq \begin{bmatrix} r^{(p)}[i] \\ \vdots \\ r^{(p)}[i + m - 1] \end{bmatrix}, \quad \tilde{n}^{(p)}[i] \triangleq \begin{bmatrix} n^{(p)}[i] \\ \vdots \\ n^{(p)}[i + m - 1] \end{bmatrix}, \quad b[i] \triangleq \begin{bmatrix} b[i - \iota] \\ \vdots \\ b[i + m - \iota] \end{bmatrix},
\]

\[
\tilde{s}^{(p)} \triangleq \begin{bmatrix} \tilde{s}^{(p)}[i] \\ \vdots \\ \tilde{s}^{(p)}[0] \end{bmatrix}, \quad \tilde{S}^{(p)} \triangleq \begin{bmatrix} \tilde{s}^{(p)}[i] \cdots \tilde{s}^{(p)}[0] \cdots 0 \\ \vdots \\ 0 \cdots \tilde{s}^{(p)}[i] \cdots \tilde{s}^{(p)}[0] \end{bmatrix}.
\]

6
We can then write (32) in a matrix forms as
\[ r^{(p)}[i] = \tilde{S}^{(p)} b[i] + n^{(p)}[i], \quad p = 1, \ldots, P. \] (33)

Finally denote
\[ \begin{bmatrix} r[1][i] \\ \vdots \\ r[Pm][i] \end{bmatrix}_{P \times Nm} \triangleq \begin{bmatrix} r^{(1)}[i] \\ \vdots \\ r^{(P)}[i] \end{bmatrix}, \quad \begin{bmatrix} S[1][i] \end{bmatrix}_{P \times K} \triangleq \begin{bmatrix} \tilde{S}^{(1)}[i] \\ \vdots \\ \tilde{S}^{(P)}[i] \end{bmatrix}, \quad \begin{bmatrix} n[1][i] \end{bmatrix}_{P \times Nm} \triangleq \begin{bmatrix} n^{(1)}[i] \\ \vdots \\ n^{(P)}[i] \end{bmatrix}. \]

Then (33) can be written as
\[ r[i] = \tilde{S} b[i] + n[i]. \] (34)

The smoothing factor \( m \) is chosen according to \( m = \left\lceil \frac{NP+K\eta}{NP-K} \right\rceil \). Note that for such \( m \), the matrix \( \tilde{S} \) is a “tall” matrix, i.e., \( PNm \geq K(m+\eta) \).

Let the autocorrelation matrix of the augmented received signal and its eigendecomposition be
\[ C_r \triangleq E \left[ r[i] r[i]^H \right] = \tilde{S} \tilde{S}^H + \eta I_{NPm} \]
\[ = U_s \Lambda_s U_s^H + \eta U_n U_n^H, \] (35) (36)

where we assume that the matrix \( \tilde{S} \) has full column rank, and hence the signal subspace has a dimension \( K(m+\eta) \). Denote \( \tilde{s}_1 \) as the \((K\eta+1)\)-th column of \( \tilde{S} \). The linear MMSE detector for detecting the \( i \)-th bit of user 1, \( b_1[i] \), based on \( r[i] \) in (34) is then given by
\[ w_1 = C_r^{-1} \tilde{s}_1 \]
\[ = U_s \Lambda_s^{-1} U_s^H \tilde{s}_1. \] (37) (38)

We next address the method for estimating \( \tilde{s}_1 \). From (31),
\[ \tilde{s}_k^{(p)}[n] = \sum_{j=0}^{N-1} s_{j,k}^{(p)}[n-j], \] (39)
\[ n = 0, 1, \ldots, (\mu+1)N-1, \] (40)
with
\[ h_k^{(p)}[n] \triangleq \frac{1}{\sqrt{N}} \sum_{l=1}^{L} a_{l,k}^{(p)} \int_0^{T_c} \psi(t) \psi(t - \tau^{(p)}_{l,k} + nT_c) \]
\[ n = 0, 1, \ldots, \mu - 1, \] (41)

where the length of the channel response \( \{h_k^{(p)}[n]\}_{n=0}^{\mu-1} \) satisfies
\[ \mu \triangleq \left[ \frac{T_c \tau_{l,k}}{T} \right] = \left[ \frac{\tau_{l,k}}{T_c} \right] \leq \iota N. \] (42)

Denote
\[ s_k^{(p)} \triangleq \begin{bmatrix} s_k^{(p)}[0] \\ \vdots \\ s_k^{(p)}[(\mu+1)N-1] \end{bmatrix}_{(\mu+1)N \times 1}, \quad h_k^{(p)} \triangleq \begin{bmatrix} h_k^{(p)}[0] \\ \vdots \\ h_k^{(p)}[\mu N-1] \end{bmatrix}_{\mu \times 1}. \]
and \( \Xi_k \triangleq \begin{bmatrix} s_{0,k} & s_{0,k} & \cdots & s_{0,k} \\ s_{1,k} & s_{0,k} & \cdots & s_{0,k} \\ \vdots & \vdots & \ddots & \vdots \\ s_{N-1,k} & \cdots & \cdots & s_{0,k} \\ s_{N-1,k} & \cdots & \cdots & \cdots & s_{0,k} \end{bmatrix} \). 

Then (40) can be written in matrix form as

\[ \tilde{s}_k^{(p)} = \Xi_k h_k^{(p)}. \]  

Finally, let \( \tilde{s}_k \) be the \((K_1 + k)\)-th column of the matrix \( \tilde{S} \). Then we have

\[
\tilde{s}_k \triangleq \begin{bmatrix} \tilde{s}_{k}^{(1)} \\ \vdots \\ \tilde{s}_{k}^{(P)} \\ 0_{(m-1)PN \times 1} \end{bmatrix} = \begin{bmatrix} \Xi_k \\ \vdots \\ \Xi_k \end{bmatrix} \begin{bmatrix} \tilde{s}_{k}^{(1)} \\ \vdots \\ \tilde{s}_{k}^{(P)} \\ 0_{(m-1)PN \times P_\mu} \end{bmatrix} = \begin{bmatrix} \Xi_k \\ \vdots \\ \Xi_k \end{bmatrix} \begin{bmatrix} h_{k}^{(1)} \\ \vdots \\ h_{k}^{(P)} \end{bmatrix} \text{.} \tag{44}
\]

In order to estimate the desired user’s channel \( h_1 \), we again resort to the orthogonality condition

\[
h_1 = \arg\min_{\|h\|=1} \left\| U_n^H \Xi_1 h \right\|^2 = \arg\min_{\|h\|=1} h^H \left( \Xi_1^H U_n U_n^H \Xi_1 \right) h = \text{minimum eigenvector of } Q. \tag{45}
\]

Note that \( h_1 \) is determined uniquely by (45) up to a scale and phase ambiguity if and only if \( \text{rank}(Q) = P_\mu - 1 \).

A blind estimate of the linear MMSE detector for user 1 is then given by the following procedure.

\[
\hat{C}_r \triangleq \frac{1}{M} \sum_{i=1}^M r[i] r[i]^H = \hat{U}_s \hat{A}_s \hat{U}_s^H + \hat{U}_n \hat{A}_n \hat{U}_n^H, \tag{47}
\]

\[
\hat{Q} = \Xi_1^T \hat{U}_n \hat{U}_n^H \Xi_1, \tag{48}
\]

\[
\hat{h}_1 = \text{minimum eigenvector of } \hat{Q}, \tag{49}
\]

\[
\hat{w}_1 = \hat{U}_s \hat{A}_s^{-1} \hat{U}_s^H \Xi_1 \hat{h}_1. \tag{50}
\]
Remark: From the above discussion, it is seen that the three systems considered here share the same signal model of the form
\[ r[i] = \tilde{S}b[i] + n[i]. \]  
(51)
Furthermore, for these three systems, a linear MMSE decision on the \( i \)-th symbol of user 1 is made based on the output of the linear MMSE detector \( w_1^H r[i] \), where \( w_1 \) has the form
\[ w_1 = C^{-1}_r \tilde{s}_1 = U_s A_s^{-1} U_s^H \tilde{s}_1. \]  
(52)
Moreover, the composite signature waveform \( \tilde{s}_1 \) of the desired user is determined by the original signature waveform of this user, and a channel vector \( h_1 \), which is given by the minimum eigenvector of a matrix \( Q \). The matrix \( Q \) is in turn determined by the noise subspace \( U_n \) and the original signature waveform of the desired user.

2.2 Performance of Blind Multiuser Detectors

In this subsection, when analyzing the performance of the blind receivers, we focus on the synchronous multipath CDMA system described in Section 2.1.1. The results, however, are directly applicable to the synchronous multi-antenna CDMA system described in Section 2.1.2, and the asynchronous multipath multi-antenna CDMA system described in Section 2.1.3, with appropriate interpretation of the quantities such as \( A_s, U_s, U_n, \) and \( Q \) in the corresponding system. The mathematical proofs of the results presented below can be found in [23, 24].

2.2.1 Complex Gaussian Distribution

Let \( x \in \mathbb{C}^n \) be a complex random vector. We say that \( x \) is complex Gaussian distributed if the vector \( \begin{bmatrix} \Re x \\ \Im x \end{bmatrix} \in \mathbb{R}^{2n} \) has a (real-valued) Gaussian distribution. Hence a complex Gaussian vector \( x \) is completely specified by its mean \( \mu \triangleq E[x] \), and the following covariance matrix
\[
\text{Cov} \left\{ \begin{bmatrix} \Re x \\ \Im x \end{bmatrix} \right\} = \begin{bmatrix} \text{Cov} \{ \Re x, \Re x \} & \text{Cov} \{ \Re x, \Im x \} \\ \text{Cov} \{ \Im x, \Re x \} & \text{Cov} \{ \Im x, \Im x \} \end{bmatrix}.
\]  
(53)
An equivalent characterization of the complex Gaussian vector \( x \) is through the following two complex-valued covariance matrices
\[
C \triangleq E \left[ (x - \mu)(x - \mu)^H \right], \\
\bar{C} \triangleq E \left[ (x - \mu)(x - \mu)^T \right].
\]
We call \( C \) the Hermitian covariance matrix and \( \bar{C} \) the symmetric covariance matrix. The real-valued covariance matrix (53) can be expressed by \( C \) and \( \bar{C} \), i.e.,
\[
\text{Cov} \left\{ \begin{bmatrix} \Re x \\ \Im x \end{bmatrix} \right\} = \frac{1}{2} \begin{bmatrix} C + \bar{C} & \Im C - \Im \bar{C} \\ \Im C + \Im \bar{C} & C - \bar{C} \end{bmatrix}.
\]  
(54)
Hence in what follows, we use the following notation to represent a complex Gaussian vector
\[
x \sim \mathcal{N}_c(\mu, C, \bar{C}).
\]  
(55)
When \( \bar{C} = 0 \), \( x \) is said to have a circularly symmetric complex Gaussian distribution. In this case, \( \Re x \) and \( \Im x \) are independent and have the same (real-valued) Gaussian distribution.
### 2.2.2 Performance of Blind Multiuser Detectors with Known Channels

First we assume that the receiver has the knowledge of the original signature waveform and the channel of the desired user, i.e., $\mathbf{s}_1$ and $\mathbf{h}_1$ in the synchronous multipath CDMA case, $\mathbf{s}_1$ and $\mathbf{h}_1$ in the asynchronous multipath multi-antenna CDMA case, and $\mathbf{\bar{s}}_1$ and $\mathbf{h}_1$ in the asynchronous multipath multi-antenna CDMA case. Equivalently, the composite signature waveform $\hat{\mathbf{s}}_1$ is assumed known. This corresponds to systems where the desired user’s channel is obtained through non-blind methods, such as pilot channels or pilot symbols.

As mentioned earlier, we focus on the synchronous multipath CDMA system described in Section 2.1.1. Based on the sample autocorrelation matrix $\mathbf{C}_w$ in (9) and its eigendecomposition (10), we can obtain two forms of the estimated linear MMSE detector, namely, the direct-matrix-inversion (DMI) detector, and the subspace detector, given respectively by

$$\hat{\mathbf{w}}_1 = \hat{\mathbf{C}}_r^{-1}\hat{\mathbf{s}}_1, \quad \text{[DMI blind detector]}$$  \hspace{1cm} (56)

and

$$\hat{\mathbf{w}}_1 = \hat{\mathbf{U}}_s\hat{\mathbf{A}}_s^{-1}\hat{\mathbf{U}}_s^H\hat{\mathbf{s}}_1, \quad \text{[subspace blind detector]}$$  \hspace{1cm} (57)

In what follows, we first present results on the asymptotic distributions of the DMI blind detector (56) and the subspace blind detector (57), when the number of received signals $M$ is large. We then give expressions of the output signal-to-interference-plus-noise ratio (SINR) as well as the bit error rate (BER) for the two blind detectors.

The following result gives the asymptotic distribution of the DMI blind detector and that of the subspace blind detector.

**Theorem 1** Let

$$\mathbf{w}_1 = \mathbf{C}_r^{-1}\mathbf{\bar{s}}_1 = \mathbf{U}_s\mathbf{A}_s^{-1}\mathbf{U}_s^H\mathbf{\bar{s}}_1$$  \hspace{1cm} (58)

be the exact linear MMSE detector, and let $\hat{\mathbf{w}}_1$ be the DMI blind detector given by (56) or the subspace blind detector given by (57). Then

$$\sqrt{M}(\hat{\mathbf{w}}_1 - \mathbf{w}_1) \rightarrow \mathcal{N}_c(\mathbf{0}, \mathbf{C}_w^0, \hat{\mathbf{C}}_w^0), \quad \text{in distribution, as } M \rightarrow \infty,$$

with

$$\mathbf{C}_w^0 = (\mathbf{w}_1^H\mathbf{\bar{s}}_1)\mathbf{U}_s\mathbf{A}_s^{-1}\mathbf{U}_s^H + \tau \mathbf{I}_n\mathbf{U}_n^H$$

$$+ \mathbf{U}_s\mathbf{A}_s^{-1}\mathbf{U}_s^H\mathbf{\hat{S}}\left[\mu \left(\mathbf{S}^T\mathbf{w}_1\mathbf{w}_1^T\mathbf{S}\right) + \nu \mathbf{D}\right]\mathbf{\hat{S}}^H\mathbf{U}_s\mathbf{A}_s^{-1}\mathbf{U}_s^H,$$  \hspace{1cm} (61)

and

$$\hat{\mathbf{C}}_w^0 = \mathbf{w}_1\mathbf{w}_1^T$$

$$+ \mathbf{U}_s\mathbf{A}_s^{-1}\mathbf{U}_s^H\mathbf{\hat{S}}\left[\mu \left(\mathbf{w}_1^T\mathbf{\hat{S}}^T\mathbf{w}_1\mathbf{\hat{S}}\right)\mathbf{I}_K + \nu \mathbf{D}\right]\mathbf{\hat{S}}^T\mathbf{U}_s^*\mathbf{A}_s^{-1}\mathbf{U}_1^T,$$  \hspace{1cm} (62)

where

$$\mathbf{D} \triangleq \text{diag}\left\{|\mathbf{s}_1^H\mathbf{w}_1|^2, |\mathbf{s}_2^H\mathbf{w}_1|^2, \cdots, |\mathbf{s}_K^H\mathbf{w}_1|^2\right\},$$  \hspace{1cm} (63)

$$\bar{\mathbf{D}} \triangleq \text{diag}\left\{(\mathbf{s}_1^H\mathbf{w}_1)^2, (\mathbf{s}_2^H\mathbf{w}_1)^2, \cdots, (\mathbf{s}_K^H\mathbf{w}_1)^2\right\},$$  \hspace{1cm} (64)

$$\mu \triangleq \mathbb{E}[|b|^2], \quad \nu \triangleq \mathbb{E}\left[|b|^4\right] - 2\mathbb{E}[|b|^2]^2 - \mathbb{E}[|b|^2]^2,$$  \hspace{1cm} (65)

and $b$ denotes the transmitted modulation symbols.
Denote $\Delta w_1 \triangleq \hat{w}_1 - w_1$ as the error of the estimated detector. The blind detector $\hat{w}_1$ given by (56) or (57) is correlated with the received signal $r[i]$ in (1) to yield the decision statistic

$$\hat{w}_1^H r[i] = w_1^H \left( \sum_{k=1}^{K} s_k b_k[i] + n[i] \right) + \Delta w_1^H r[i]. \quad (66)$$

The signal-to-interference-plus-noise ratio (SINR) at the output of this estimated detector is then given by

$$\text{SINR} \triangleq \frac{|E[\hat{w}_1^H r[i] | b_1[i]]|^2}{E[\text{Var}(\hat{w}_1^H r[i] | b_1[i])]} = \frac{(w_1^H \hat{s}_1)^2 E[|b|^2]}{E[|b|^2] \sum_{k=2}^{K} |w_1^H \hat{s}_k|^2 + \eta \|w_1\|^2 + \frac{1}{M} \text{tr}(C_w^0 C_r)}. \quad (67)$$

Denote

$$\tilde{R} \triangleq \tilde{S}^H \hat{S}. \quad (68)$$

Then after some manipulations, the quantities in (67) are given in terms of $\tilde{R}$ and $(\eta, \mu, \nu, N, K, M)$ as follows:

$$w_1^H \hat{s}_j = \tilde{s}_i^H w_j = \left[ (I_K + \eta \tilde{R})^{-1} \right]_{i,j}, \quad (69)$$

$$\|w_1\|^2 = \left[ (\tilde{R} + \eta I_K)^{-1} \tilde{R} (\tilde{R} + \eta I_K)^{-1} \right]_{1,1}, \quad (70)$$

$$\tau_\eta = \left\{ \begin{array}{cl}
\tilde{s}_i^H w_1, & \text{DMI} \\
\eta^2 \left[ (\tilde{R} + \eta I_K)^{-1} \tilde{R}^{-1} \right]_{1,1}, & \text{subspace}
\end{array} \right. \quad (71)$$

and

$$\text{tr}(C_w^0 C_r) = K w_1^H \hat{s}_1 + (N - K) \tau_\eta + \mu \sum_{i=1}^{K} \sum_{j=1}^{K} (\tilde{s}_i^H w_i) (\tilde{s}_i^H w_j) (\tilde{s}_j^H w_j) + \nu \sum_{k=1}^{K} |\tilde{s}_k^H w_1|^2 (\tilde{s}_k^H w_k). \quad (72)$$

We next give the expressions of $\tilde{R}$ in terms of the spreading waveforms and the channels of all users, for the following four systems:

1. **Synchronous flat-fading CDMA:** This corresponds to $L = 1$ in (1). Let $H \triangleq \text{diag}(h_{1,1}, h_{1,2}, \cdots, h_{1,K})$ and $S \triangleq [s_1, s_2, \cdots, s_K]$. Then we have $\tilde{S} = H \tilde{s}$. Denote $R \triangleq S^H S$. Hence we have

$$\tilde{R} = H^H R H. \quad (73)$$
2. **Synchronous multipath CDMA:** Consider the signal model (2). Let \( S \triangleq [S_1, S_2, \ldots, S_K] \) and \( H \triangleq \text{diag}(h_1, h_2, \ldots, h_K) \). Then we have \( \tilde{S} = S H \). Hence we have
\[
\tilde{R} = H^H S H. \tag{74}
\]

3. **Synchronous multi-antenna CDMA:** Consider the signal model (16). Note that
\[
[\tilde{R}]_{ij} = \tilde{s}_i^H \tilde{s}_j = (h_i \otimes \tilde{s}_i)^H (h_j \otimes \tilde{s}_j) = (h_i^H h_j) (\tilde{s}_i^H \tilde{s}_j). \tag{75}
\]
Denote \( H \triangleq [h_1, h_2, \ldots, h_K] \). It then follows that
\[
\tilde{R} = R \circ (H^H H), \tag{76}
\]
where \( \circ \) denotes element-wise matrix product.

4. **Asynchronous multipath multi-antenna CDMA:** Consider the general signal model (44). Let \( \tilde{\Xi} \triangleq [\tilde{\Xi}_1, \tilde{\Xi}_2, \ldots, \tilde{\Xi}_K] \), and \( H \triangleq \text{diag}(h_1, h_2, \ldots, h_K) \). Then we have \( \tilde{S} = \tilde{\Xi} H \). Hence
\[
\tilde{R} = H^H \tilde{\Xi}^H \tilde{\Xi}. \tag{77}
\]

### 2.2.3 Performance of Blind Multiuser Detector with Blind Channel Estimation

Next we present results on the asymptotic distributions of the blind detector with blind channel estimation described in Section 2.1 and the corresponding detector output SINR. The following result gives the asymptotic distribution of the blind detector with blind channel estimation, given by (9)-(13).

**Theorem 2** Let
\[
w_1 = U_s \Lambda_s^{-1} U_n^H \tilde{s}_1 \tag{78}
\]
be the exact linear MMSE detector, \( h_1 \) be the true channel of user 1, and \( \hat{w}_1 \) be the estimated blind detector given by (9)-(13). Let \( C_w^0 \) and \( \bar{C}_w \) be the quantities given by (59) and (60) respectively. Then there exists a phase factor \( e^{j\phi} \) such that
\[
\sqrt{M} (\hat{w}_1 - \| h_1 \|^{-1} e^{j\phi} w_1) \to \mathcal{N}_c (0, \| h_1 \|^{-2} C_w, \| h_1 \|^{-2} \bar{C}_w), \quad \text{in distribution, as } M \to \infty,
\]
with
\[
C_w = C_w^0 + \beta_1 \Psi Q^\dagger \Psi^H + \beta_2 (\Psi Q^\dagger S_1 H U_n U_n^H + U_n U_n^H S_1 Q^\dagger \Psi), \tag{79}
\]
\[
\bar{C}_w = \bar{C}_w^0, \tag{80}
\]
where
\[
\Psi \triangleq U_s \Lambda_s^{-1} U_n^H S_1, \tag{81}
\]
\[
\beta_1 \triangleq \eta \tilde{s}_1^H U_s \Lambda_s (\Lambda_s - \eta I_K)^{-2} U_n^H \tilde{s}_1, \tag{82}
\]
\[
\beta_2 \triangleq \eta \tilde{s}_1^H U_s (\Lambda_s - \eta I_K)^{-2} U_n^H \tilde{s}_1, \tag{83}
\]
and \( Q^\dagger \) is the pseudo-inverse of the matrix \( Q \triangleq S_1^H U_n U_n^H S_1 \).
We can also obtain the asymptotic distribution of the channel estimate, as given by the following result.

**Corollary 1** Let $h_1$ be the true channel of user 1, and let $\hat{h}_1$ be the channel estimate given by (12). Then there exists a phase factor $e^{j\phi}$, such that

$$\sqrt{M} \left( h_1 - \|h_1\|^{-1}e^{j\phi} h_1 \right) \rightarrow N_c \left( 0, \beta_1 \|h_1\|^{-2} Q^\dagger, 0 \right),$$

in distribution, as $M \rightarrow \infty$, where $\beta_1$ is given by (82). Thus asymptotically $\hat{h}_1$ is circularly symmetric complex Gaussian.

Hence the blind detector with blind channel estimation can be written as

$$\hat{w}_1 = \|h_1\|^{-1}e^{j\phi} w_1 + \Delta w_1.$$  \hspace{1cm} (84)

It is straightforward to show that the detector output SINR is given by

$$\text{SINR} = \frac{(\hat{w}_1^H \tilde{s}_1)^2}{\sum_{k=2}^{K} |w_1^H \tilde{s}_k|^2 + \eta \|w_1\|^2 + \frac{1}{M} \text{tr}(C_w C_r)},$$

where $w_1^H \tilde{s}_k$ and $\|w_1\|^2$ are given respectively by (69) and (70), and using (79) after some manipulations, we obtain

$$\text{tr}(C_\alpha C_r) = \text{tr}(C_w^0 C_r) + \beta_1 \text{tr}\left[ (S_1^T U_s A_s^{-1} U_s^H S_1)^\dagger \right],$$

where $\text{tr}(C_w^0 C_r)$ is given by (72) and

$$\beta_1 = \eta \left[ I_K + \eta \bar{R}^{-1} \right]_{1,1},$$

where $\bar{R}$ is given by (68). Hence it is clear that the effect of channel estimation on the output SINR of the blind detector is to add the second term in (86) to the denominator of (85).

### 2.2.4 Numerical Results

The simulated system is a synchronous multipath CDMA system. The number of users is $K = 18$. Each user’s original spreading sequence is randomly generated and has length 21. The channel of each user has length 11 and is randomly generated. Both the spreading sequences and the channels of all users are fixed throughout the simulations. Note that the length of the composite signature waveform of each user (i.e., $\tilde{s}_k$) is 31. Here we insert a guard interval of length 10 chips between two consecutive symbols to avoid intersymbol interference. Note such a setup is merely for the purpose of verifying the theoretical results in this paper.

As noted earlier, when the channel is unknown and is blindly estimated, the blind channel estimator introduces a phase ambiguity $\alpha = e^{j\phi}$ which must be estimated for detection purpose. In the simulations, we employ the following simple phase estimator:

$$\hat{\alpha} = \left[ \frac{1}{M} \sum_{i=1}^{M} (\hat{w}_1^H r[i])^2 \right]^{\frac{1}{2}} = \left( \hat{w}_1^H \hat{C}_r \hat{w}_1^* \right)^{\frac{1}{2}}.$$  \hspace{1cm} (88)
Note that the above phase estimator still contains a phase ambiguity of $\pi$ for BPSK and $\frac{\pi}{2}$ for QPSK, which is inherent to any blind detector.

In Fig. 1, assuming BPSK modulation the simulated output SINR of three blind detectors, namely, the DMI detector with known channel, the subspace detector with known channel, and the subspace detector with unknown channel, are plotted as a function of input $E_b/N_0$ of each user, for two block sizes, i.e., $M = 300$ and $M = 2000$. The corresponding theoretical values, as well as the theoretical and simulated SINR of the exact MMSE detector, are also plotted in the same figure. Moreover, the SINR results for QPSK modulation for the same simulated system are plotted in Fig. 2. It is seen from these figures that the analytical expressions match very well with the simulation results. Furthermore, in unknown channels, the simple phase estimator (88) incurs little performance loss compared with the case where the phase ambiguity is perfectly known. Finally, we note that by invoking a Gaussian assumption on the detector output, the SINR can be translated into the bit error rate, as shown in [24].

### 2.2.5 Adaptive Implementation

We next discuss adaptive algorithms for sequentially estimating the blind subspace detector. First, we address adaptive implementation of the blind channel estimator discussed above. Suppose the signal subspace $U_s$ is known. Denote by $z[i]$ the projection of the received signal $r[i]$ onto the noise subspace, i.e.,

$$z[i] = r[i] - U_s U_s^H r[i]$$

$$= U_n U_n^H r[i].$$

Since $z[i]$ lies in the noise subspace, it is orthogonal to any signal in the signal subspace. In particular, it is orthogonal to $\hat{s}_1 = \Xi_1 h_1$. Hence $h_1$ is the solution to the following constrained optimization problem

$$\min_{h_1_i \in \mathbb{C}^{p_n}} E \left[ \left\| (\Xi_1 h_1)^H z[i] \right\|^2 \right], \quad \text{s.t.} \quad \|h_1\| = 1. \quad (91)$$

In order to obtain a sequential algorithm to solve the above optimization problem, we write it in the following (trivial) state space form

$$h_1[i] = h_1[i], \quad \text{(state equation)}$$

$$0 = (\Xi_1 h_1[i])^H h_1[i], \quad \text{(observation equation)} \quad (90)$$

The standard Kalman filter can then be applied to the above system, as follows. (We define $x[i] = \Xi_1^H z[i]$.)

$$k[i] = \Sigma[i-1] x[i] (x[i]^H \Sigma[i-1] x[i])^{-1}, \quad (92)$$

$$h_1[i] = h_1[i-1] - k[i] (x[i]^H h_1[i]) / \|h_1[i-1] - k[i] (x[i]^H h_1[i])\|, \quad (93)$$

$$\Sigma[i] = \Sigma[i-1] - k[i] x[i]^H \Sigma[i-1]. \quad (94)$$

Note that (93) contains a normalization step to satisfy the constraint $\|h_1[i]\| = 1$.

Since the subspace blind detector may be written in closed-form as a function of the signal subspace components, one may use a suitable subspace tracking algorithm in conjunction with this detector and a channel estimator to form an adaptive detector that is able
Figure 1: Output SINR versus input $E_b/N_0$ for BPSK.

Figure 2: Output SINR versus input $E_b/N_0$ for QPSK.
to track changes in the number of users and their composite signature waveforms. Fig. 3 contains a block diagram of such a receiver. The received signal $r[i]$ is fed into a subspace tracker that sequentially estimates the signal subspace components $(U_s, \Lambda_s)$. The signal $r[i]$ is then projected onto the noise subspace to obtain $z[i]$, which is in turn passed through a linear filter that is determined by the signature sequence of the desired user. The output of this filter is fed into a channel tracker that estimates the channel state of the desired user. Finally, the linear MMSE detector is constructed in closed-form based on the estimated signal subspace components and the channel state.

2.2.6 Algorithm Summary

The adaptive receiver algorithm is summarized as follows.

**Algorithm 1** [Blind Adaptive Subspace Multiuser Detection]

1. Using a suitable signal subspace tracking algorithm, update the signal subspace components $U_s[i], \Lambda_s[i], \sigma_s^2[i]$, and the subspace rank $r[i]$ at each time slot $i$.
2. Update the channel estimate $h[i]$ using (92)-(94).
3. Form detector and perform differential detection:

   $w[i] = U_s[i] \Lambda_s[i]^{-1} U_s[i]^H s_1[i], \quad (95)$
   $z_1[i] = w[i]^H r[i], \quad (96)$
   $\hat{\beta}_1[i] = \text{sign} \{ \Re(z_1[i] z_1[i-1]^*) \}. \quad (97)$

---

![Figure 3: Blind adaptive subspace multiuser detector.](image_url)

We next give a simulation example illustrating the performance of the blind adaptive receiver in an asynchronous CDMA system with multipath channels. The processing gain $N = 15$ and the spreading codes are Gold codes of length 15. Each user’s channel has $L = 3$ paths. The delay of each path $\tau_{l,k}$ is uniformly distributed on $[0, 10T_c]$. Hence, as in the preceding example, the maximum delay spread is one symbol interval. The fading gain of each path in each user’s channel is generated from a complex Gaussian distribution and is fixed for all simulations. The path gains in each user’s channel are normalized so that all users’ signals arrive at the receiver with the same power. The received signal is sampled at
twice the chip-rate. Fig. 4 shows the performance of subspace blind adaptive receiver using the NAHJ subspace tracking algorithm [36], in terms of output SINR. During the first 1000 iterations there are 8 total users. At iteration 1000, 4 new users are added to the system. At iteration 2000, one additional known user is added and three existing users vanish. We see that this blind adaptive receiver can closely track the dynamics of the channel.

Figure 4: Performance of the blind adaptive subspace multiuser detector in an asynchronous CDMA system with multipath.

3 Bayesian Multiuser Detection for Long-code CDMA

In the previous section, we discussed subspace blind multiuser detection techniques. The key underlying assumption for this approach is that the users employ periodic spreading codes, i.e., the sequence $\{s_{j,k}\}_{j=0}^{N-1}$, $\forall k$, in (29)-(30) is independent of the symbol time index $i$. On the other hand, existing CDMA standards (such as IS-95, WCDMA, CDMA2000) employ long spreading codes on the reverse link, i.e., PN sequences with very long periods. The theme of this section is on the design of blind multiuser receiver for an uplink asynchronous coded CDMA system employing long spreading sequences. It is assumed that the blind receiver has only the knowledge of the spreading sequences and the initial delays of the desired users within the cell.

3.1 System Descriptions

3.1.1 Channel Model

Consider a $K$-user uplink CDMA system, employing normalized long pseudo-random spreading sequences, and signaling through multipath channels with additive white Gaussian noise and other unknown interference. The $k$-th user’s spreading waveform in (30) corresponding
to the $i$-th symbol interval is now written as

$$s_k^{[i]}(t) = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} s_{j,k}[i] \psi(t - jT_c), \quad 0 \leq t < T,$$

(98)

where $\{s_{j,k}[i], j = 0, \ldots, N - 1\}, \forall k, \forall i$ is the segment of the $k$-th user’s signature sequence corresponding to the $i$-th symbol. For simplicity we consider the single receive antenna case. Then the received signal in (29) is now modified as

$$r(t) = \sum_{k=1}^{K} \sum_{i=0}^{M-1} b_k[i] \sum_{l=1}^{L} \alpha_{l,k} s_k^{[i]}(t - iT - \tau_{l,k}) + v(t),$$

(99)

where $v(t)$ is the ambient noise plus the unknown interfering signals, as will be explained below. At the receiver, the received signal $r(t)$ is filtered by a chip-matched filter and sampled at the chip-rate. Define $\tau_k \triangleq \lfloor \frac{\tau_{k,1}}{T_c} \rfloor - 1$ as the initial delay in terms of number of chips for the $k$-th user’s signal; Define $\mu \triangleq \max_k \lfloor \frac{\tau_{k,1}}{T_c} \rfloor$ as the maximum channel delay among all users; and define $h_k \triangleq [h_k[(k+1)]_1, \ldots, h_k[(k+\mu)]^T]$ as the channel response for the $k$-th user. We assume that both the maximum initial delay $\max_k \{\tau_k\}$ and the maximum channel delay $\mu$ are less than $N$. Hence, the maximum symbol delay satisfies $\tau \leq 2$.

The received discrete-time signal vector corresponding to the $i$-th symbol interval is given by

$$r[i] = \sum_{k=1}^{K} \left( b_k[i] C_k^{[0]} + b_k[i - 1] C_k^{[1]} + b_k[i - 2] C_k^{[2]} \right) h_k + v[i], \quad i = 0, 1, \ldots, M - 1,$$

(100)

where $r[i] \triangleq [r_0[i], r_1[i], \ldots, r_N[i]]^T$; $v[i] \triangleq [v_0[i], \ldots, v_{P-1}[i]]^T$. It is easy to verify that $C_k^{(n)}$ can be expressed as

$$\begin{bmatrix}
C_k^{[0]} \\
C_k^{[1]} \\
C_k^{[2]}
\end{bmatrix}_{N \times \mu} =
\begin{bmatrix}
0_{k \times 1} & 0 & \ldots & 0 \\
s_0,k[i] & s_1,k[i] & \ldots & s_0,k[i] \\
\vdots & \vdots & \ddots & \vdots \\
s_{N-1,k}[i] & \ldots & s_{N-1,k}[i] & s_1,k[i] \\
0 & 0 & \ldots & 0
\end{bmatrix}_{3N \times \mu}.$$

(101)

### 3.1.2 Noise Model

In the simplest case, it is assumed that $v(t)$ in (99) contains the channel ambient noise only, which is a white complex Gaussian process. It is further assumed that the chip waveform $\psi(t)$ is a rectangle pulse with duration $T_c$. Hence all the noise samples $\{v_q[i]\}_{q,i}$ are i.i.d. zero mean complex Gaussian random variables with variance $\sigma^2$. Moreover, $\{v[i]\}_i$ is a sequence of zero-mean i.i.d. complex Gaussian vectors, i.e.,

$$v[i] \sim \mathcal{N}(0, \sigma^2 I).$$

(102)
In cellular DS-CDMA, the same uplink/downlink pair of frequency bands are reused for each cell. Therefore, a signal transmitted in one cell may cause interference in neighboring cells, resulting in out-cell multiple-access interference (OMAI). In addition, narrowband communication systems sometimes can overlay with CDMA systems, and thus cause narrowband interference (NBI) to the latter [29, 31]. Hence, in general, the noise component \( v[i] \) in (100) consists of white Gaussian noise (WGN), OMAI and NBI, i.e.,
\[
v = v_{\text{WGN}} + v_{\text{OMAI}} + v_{\text{NBI}}.
\]
The WGN has zero mean and covariance \( \Sigma_{\text{WGN}} = \sigma^2 I \). The OMAI has the same structure as the in-cell CDMA signals, i.e.,
\[
v_{\text{OMAI}}[i] = \sum_{k=K+1}^{K+K'} \left[ b_k[i] C_{k,i}^{(0)} + b_k[i-1] C_{k,i-1}^{(1)} + b_k[i-2] C_{k,i-2}^{(2)} \right] h_k,
\]
where \( K' \) denote the total number of out-cell users. When \( K' \) is large, by the central limit theorem, \( v_{\text{OMAI}}[i] \) approaches a Gaussian vector with zero mean and a covariance matrix, denoted by \( \Sigma_{\text{OMAI}} \). Note that both the encoded bits \( \{b_k[i]\}_{k,i} \) and the elements of the spreading sequences \( \{s_{j,k}[i]\}_{k,i,j} \) are zero-mean independent random variables. After some manipulations, the element of \( \Sigma_{\text{OMAI}} \) can be written as
\[
\Sigma_{\text{OMAI}}[u, v] = \sum_{k=1}^{K} \sum_{j,l=u-v}^{j+l=M} h_k[j] h_k^*[l], \quad u, v = 0, 1, \ldots, N - 1.
\]
The NBI signal is typically modeled as a correlated Gaussian process. For example, it can be represented as an \( m_0 \)th-order autoregressive (AR) signal [33], where \( m_0 \ll N \), i.e.,
\[
v_{\text{NBI}}[n] = - \sum_{j=1}^{m_0} a_j v_{\text{NBI}}[n-j] + c[n],
\]
where \( v_{\text{NBI}}[n] \) denote the noise sample component due to the NBI signal; \( c[n] \) is a white Gaussian process with variance \( \sigma^2 \). Hence \( v_{\text{NBI}}[i] \triangleq [v_{\text{NBI}}[iN], v_{\text{NBI}}[iN+1], \ldots, v_{\text{NBI}}[iN+N-1]]^T \) is Gaussian with zero mean and a covariance matrix, denoted by \( \Sigma_{\text{NBI}} \).

Combining these three components, the noise vectors \( \{v[i]\} \) can be modeled as colored Gaussian vectors with zero mean and a covariance matrix \( \Sigma = \Sigma_{\text{WGN}} + \Sigma_{\text{OMAI}} + \Sigma_{\text{NBI}} \), i.e.,
\[
v[i] \sim \mathcal{N}(0, \Sigma).
\]
Note that when the OMAI and the NBI are present, the noise vectors \( \{v[i]\} \) are correlated. Nevertheless we ignore such temporal correlations to simplify the algorithms.

### 3.1.3 Blind Bayesian Multiuser Detection

The binary information bits \( \{d_k[n]\} \) for user \( k \) are encoded using some channel code (e.g., block code, convolutional code or turbo code). A code-bit interleaver is used to reduce the influence of the error bursts at the input of the channel decoder. The interleaved code bits are then mapped to BPSK symbols \( \{a_k[i]\}_{i=0}^{M-1} \). These BPSK symbols are differentially encoded to yield the symbol stream \( \{b_k[i]\}_{i=0}^{M-1} \). Differential encoding is used to resolve the phase ambiguity inherent in any blind receiver, and is given by
\[
\begin{align*}
&\begin{cases}
  b_k[0] = 1, \\
  b_k[i] = b_k[i-1] a_k[i], \quad i = 1, 2, \ldots, M - 1.
\end{cases}
\end{align*}
\]
Each symbol $b_k[i]$ is then modulated by a spreading waveform and transmitted through a multipath channel. The received signal is given by (99).

The Bayesian multiuser detector assumes the knowledge of the spreading sequence and the initial delay information for each in-cell user, i.e., \(\{C_{k,i}^{(0)}, C_{k,i}^{(1)}, C_{k,i}^{(2)}\}_{k=1;i=0}^{K;M-1}\) in (99) are known to the receiver. Note that the initial delay is the first non-zero channel coefficient, which may be obtained through CDMA timing acquisition techniques.

For convenience, define the a priori log-likelihood ratios (LLR) of the interleaved code bits as

\[
\rho_k[i] \triangleq \log \frac{P(a_k[i] = +1)}{P(a_k[i] = -1)}. \tag{108}
\]

Further define \(Y \triangleq \{r[0], r[1], \ldots, r[M-1]\}\). The Bayesian multiuser detector estimates a posteriori probabilities of the code bits

\[
P(a_k[i] = +1 | Y), \quad i = 1, \ldots, M-1; \quad k = 1, \ldots, K, \tag{109}
\]

based on the received signals \(Y\), the signal model (99) and the prior information \(\{\rho_k[i]\}_{k=1;i=1}^{K;M-1}\), without knowing the channel response \(\{h_k\}_{k=1}^{K}\) and the noise parameters (i.e., \(\sigma^2\) for white Gaussian noise; \(\Sigma\) for colored Gaussian noise). Note that although \(Y\) is directly determined by the differentially encoded symbols \(\{b_k[i]\}_{k,i}\) as seen in the signal model (99), the channel decoders require the posterior distributions of the code bits \(\{a_k[i]\}_{k,i}\).

Note that such a Bayesian multiuser detector can be employed as the front-end soft-input soft-output demodulator in a turbo receiver for coded CDMA systems [51, 55, 56].

### 3.1.4 The Gibbs Sampler

The blind Bayesian multiuser detectors discussed below are based on the Gibbs sampler [11], a Markov chain Monte Carlo (MCMC) procedure for numerical Bayesian computation. Let \(\theta = [\theta_1, \theta_2, \ldots, \theta_d]^T\) be a vector of unknown parameters. Let \(Y\) be the observed data. Suppose that we are interested in finding the a posteriori marginal distribution of some parameter, say \(\theta_j\), conditioned on the observation \(Y\), i.e., \(p(\theta_j | Y)\). Direct evaluation involves integrating out the rest of the parameters from the joint posterior density \(p(\theta | Y)\), which in most cases is computationally infeasible. The basic idea behind the Gibbs sampler is to generate random samples from the joint posterior distribution \(p(\theta | Y)\), and then to estimate any marginal distribution using these samples. Given the samples at time \((n-1)\), \(\theta^{(n-1)} = [\theta_1^{(n-1)} \ldots \theta_d^{(n-1)}]^T\), at the nth iteration, this algorithm performs the following operation to obtain samples at time \(n\), \(\theta^{(n)} = [\theta_1^{(n)} \ldots \theta_d^{(n)}]^T\):

- For \(i = 1, \ldots, d\), draw \(\theta_i^{(n)}\) from the conditional distribution

  \[
p\left(\theta_i | \theta_1^{(n)}, \ldots, \theta_{i-1}^{(n)}, \theta_{i+1}^{(n-1)}, \ldots, \theta_d^{(n-1)}, Y\right).
  \]

It is known that under regularity conditions [12, 3, 13],

- The distribution of \(\theta^n\) converges geometrically to \(p(\theta | Y)\), as \(n \to \infty\).

- \(\frac{1}{N} \sum_{n=1}^{N} f(\theta^{(n)}) \overset{a.s.,}{\to} \int f(\theta)p(\theta | Y)d\theta\), as \(N \to \infty\), for any integrable function \(f\).
3.2 Bayesian MCMC Multiuser Detectors

3.2.1 White Gaussian Noise

We first consider the problem of computing the \textit{a posteriori} bit probabilities in (109) under the assumption that the ambient noise distribution is white and Gaussian. That is, the pdf of $v[i]$ in (100) is given by

$$p(v[i]) = \left(\frac{1}{\pi \sigma^2}\right)^P \exp\left(-\frac{||v[i]||^2}{\sigma^2}\right). \quad (110)$$

Denote

$$b_k \triangleq \{b_k[i]\}_{i=0}^{M-1}, \quad B \triangleq \{b_k\}_{k=1}^{K},
A \triangleq \{a_k[i]\}_{k=1,i=1}^{K,M-1}, \quad H \triangleq \{h_k\}_{k=1}^{K},
S_{k,i}(b_k) \triangleq b_k[i]C_{k,i}^{(0)} + b_k[i-1]C_{k,i-1}^{(1)} + b_k[i-2]C_{k,i-2}^{(2)}.$$ 

Then (100) can be written as

$$r[i] = \sum_{k=1}^{K} S_{k,i}(b_k)h_k + v[i], \quad i = 0, 1, \ldots, M - 1. \quad (111)$$

The problem is solved under a Bayesian framework, by treating the unknown quantities $H$, $\sigma^2$, and $B$ as realizations of random variables with some prior distributions. The Gibbs sampler is then employed to calculate the marginal distribution of those unknown parameters. Note that although the code bits $A$ are of interest, it is more convenient to sample the differentially encoded bits $B$ in the Gibbs sampler.

\textbf{Prior Distributions:} 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 priors should also be chosen such that the conditional posterior distributions are easy to compute and simulate. To that end, we choose the following prior distributions $p(H), p(\sigma^2)$ and $p(B)$, for the unknown parameters:

1. For the unknown channel $h_k$, a complex Gaussian prior distribution is assumed,

$$p(h_k) \sim \mathcal{N}(h_{k0}, \Sigma_{k0}). \quad (112)$$

Note that large value of $\Sigma_{k0}$ corresponds to less informative prior.

2. For the noise variance $\sigma^2$, an inverse chi-square prior distribution is assumed,

$$p(\sigma^2) = \frac{(\nu_0\lambda_0)^{\nu_0}}{\Gamma(\nu_0)} \left(\frac{1}{\sigma^2}\right)^{\nu_0+1} \exp\left(-\frac{\nu_0\lambda_0}{\sigma^2}\right) \sim \chi^{-2}(2\nu_0, \lambda_0). \quad (113)$$

Small value of $\nu_0\lambda_0$ corresponds to the less informative priors.
3. The data bit sequence $b_k$ is a Markov chain, encoded from $\{a_k[i]\}_{i=1}^{M-1}$. Its prior distribution can be expressed as

$$p(b_k) = p(b_k[0])p(b_k[1] | b_k[0]) \cdots p(b_k[M-1] | b_k[M-2]) = p(b_k[0])p(a_k[1] = b_k[1]b_k[0]) \cdots p(a_k[M-1] = b_k[M-1]b_k[M-2]) = \frac{1}{2} \prod_{i=1}^{M-1} \frac{\exp(\rho_k[i] b_k[i-1] b_k[i])}{1 + \exp(\rho_k[i] b_k[i-1] b_k[i])}. \quad (114)$$

Notice that we set $p(b_k[0]) = \frac{1}{2}$ to count for the phase ambiguity in $b_k[0]$.

**Conditional Posterior Distributions:** The following conditional posterior distributions are required by the Bayesian MCMC multiuser detector.

1. The conditional distribution of the $k$-th user’s channel response $h_k$ given $\sigma^2$, $B$, $H_k$ and $Y$ is (where $H_k \triangleq H \setminus h_{k\cdot}$)

$$p(h_k | B, \sigma^2, H_k, Y) \sim \mathcal{N}(h_{k^*}, \Sigma_{k^*}), \quad (115)$$

with $\Sigma_{k^*} = \Sigma_k^{-1} + \frac{1}{\sigma^2} \sum_{i=0}^{M-1} S_{k,i}^H(b_k) S_{k,i}(b_k)$,

and $h_{k^*} \triangleq \Sigma_k \left[ \Sigma_k^{-1} h_{k0} + \frac{1}{\sigma^2} \sum_{i=0}^{M-1} S_{k,i}^H(b_k) \left( r[i] - \sum_{j \neq k} S_{j,i}(b_j) h_j \right) \right]. \quad (117)$

2. The conditional distribution of the noise variance $\sigma^2$ given $H$, $B$, and $Y$ is given by

$$p(\sigma^2 | H, B, Y) \sim \chi^{-2} \left( 2[\nu_0 + MP], \frac{\nu_0 \lambda_0 + s^2}{\nu_0 + MP} \right), \quad (118)$$

with $s^2 \triangleq \sum_{i=0}^{K} \| r[i] - \sum_{k=1}^{K} S_{k,i}(b_k) h_k \|^2. \quad (119)$

3. The conditional distribution of the data bit $b_k[i]$ given $H$, $\sigma^2$, $B_{ki}$, and $Y$ can be obtained from (where $B_{ki} \triangleq B \setminus b_{ki}[i]$)

$$P(b_k[i] = +1 | H, \sigma^2, B_{ki}, Y) = \exp \left( b_k[i+1] \rho_k[i+1] + b_k[i-1] \rho_k[i] - \frac{\Delta s^2}{\sigma^2} \right), \quad (120)$$

with $\Delta s^2 \triangleq -4\Re \left\{ h_k^H \left( C_{k,i}^{(0)T} r_k[i] + C_{k,i}^{(1)T} r_k[i+1] + C_{k,i}^{(2)T} r_k[i+2] \right) \right\}, \quad (121)$

$$r_k[i] \triangleq r[i] - \sum_{j \neq k} S_{j,i}(b_j) h_j - S_{k,i}(b_{kl}) h_k, \quad (122)$$

where $b_{kl} \triangleq \{b_k[0], \ldots, b_k[l-1], 0, b_k[l+1], \ldots, b_k[M-1]\}$.

Using the above conditional posterior distributions, the Gibbs sampling implementation of the blind Bayesian multiuser detector in white Gaussian noise proceeds iteratively as follows. Note that the samples of code bits $A$ are computed based on the samples of differentially encoded bits $B$ in $(\star)$. 

22
Algorithm 2 [Blind Bayesian Multiuser Detection in White Gaussian Noise]

1. Draw the initial values $\theta^{(0)} = \{H^{(0)}, \sigma^{2(0)}, B^{(0)}\}$ from their prior distributions.

2. For $n = 1, 2, \ldots$
   (a) For $k = 1, 2, \ldots, K$
      i. Draw $h_k^{(n)}$ from $p(h_k \mid H_{k}^{(n-1)}, \sigma^{2(n-1)}, B_{ki}^{(n-1)}, Y)$ given by (115).
   (b) Draw $\sigma^{2(n)}$ from $p(\sigma^2 \mid H^{(n)}, B^{(n-1)}, Y)$ given by (118).
   (c) For $i = 0, 1, \ldots, M - 1$
      i. For $k = 1, 2, \ldots, K$
         A. Draw $b_k^{(n)}[i]$ from $P(b_k[i] \mid h_k^{(n)}, \sigma^{2(n)}, B_{ki}^{(n-1)}, Y)$ given by (120).
         B. Compute $a_k^{(n)}[i] = b_k^{(n)}[i]b_k^{(n)}[i-1]$ (\textsuperscript{\textasteriskcentered}).

where $H_{k}^{(n-1)} \triangleq \{h_1^{(n)}, \ldots, h_k^{(n)}, h_{k+1}^{(n-1)}, \ldots, h_{K}^{(n-1)}\}$ and $B_{ki}^{(n-1)} \triangleq \{b_1^{(n)}[0], \ldots, b_i^{(n)}[M-1], \ldots, b_k^{(n)}[i-1], b_{k+1}^{(n-1)}[i+1], \ldots, b_{K}^{(n-1)}[M-1]\}$.

To ensure convergence, the above procedure is usually carried out for $(n_0 + P)$ iterations and samples from the last $P$ iterations are used to calculate the Bayesian estimates of the unknown quantities. The posterior symbol probabilities (109) can be approximated as

$$P(a_k[i] = +1 \mid Y) \approx \frac{1}{P} \sum_{n=n_0+1}^{n_0+P} \delta_{ki}^{(n)}, \quad k = 1, \ldots, K; i = 1, \ldots, M - 1, \quad (123)$$

where $\delta_{ki}^{(n)}$ is an indicator such that $\delta_{ki}^{(n)} = 1$, if $a_k^{(n)}[i] = +1$ and $\delta_{ki}^{(n)} = 0$, if $a_k^{(n)}[i] = -1$.

3.2.2 Colored Gaussian Noise

We next discuss the blind Bayesian multiuser detector for colored Gaussian noise. It is assumed that $\nu[i]$ in (100) have a complex joint Gaussian distribution, i.e.,

$$p(\nu[i]) = (\pi|\Sigma|)^{-N} \exp(-\nu^H[i] \Sigma^{-1} \nu[i]). \quad (124)$$

As mentioned before, the noise vectors $\{\nu[i]\}_i$ are temporally correlated. However, in what follows we ignore this correlation to reduce the receiver complexity.

Prior Distributions: The unknown quantities in this case are $(H, \Sigma^{-1}, B)$, which are assumed to be independent with each other. As in the case of Gaussian noise, the prior distribution of $H$ and $B$ are given respectively by (112) and (114). For the noise covariance matrix $\Sigma$, an inverse complex Wishart prior is assumed, i.e.,

$$p(\Sigma) = \frac{|\Psi|^m \cdot |\Sigma|^{-(m+P+1)} \cdot \exp[-\text{tr}(\Psi \Sigma^{-1})]}{2^{mP} \Gamma_P(m)} \sim \mathcal{W}_c^{-1}(\Psi, m), \quad (125)$$

where $\text{tr}(\cdot)$ denotes the trace of a matrix; $\Gamma_P(m) \triangleq \prod_{i=1}^{P} \Gamma(m + 1 - i)$. Small values of $m$ and $\Psi$ correspond to less informative prior. The inverse of the covariance matrix $\Sigma$ has a complex Wishart distribution, i.e.,

$$p(\Sigma^{-1}) \sim \mathcal{W}_c(\Psi^{-1}, m). \quad (126)$$
A random matrix with a Wishart distribution with \( m \) degrees of freedom (126) can be generated by \( \sum_{i=1}^{m+1} u_i u_i^H \), where \( \{u_i\} \) are i.i.d. Gaussian random vectors with zero mean and covariance \( \Psi^{-1} \).

**Conditional Posterior Distributions:** The following conditional posterior distributions are required by the blind Bayesian multiuser detector.

1. The conditional distribution of the \( k \)-th user's channel response \( h_k \) given \( \Sigma^{-1}, B, H_k, \) and \( Y \) is (where \( H_k \overset{\Delta}{=} H \setminus h_k \))

   \[
   p(h_k \mid B, \Sigma^{-1}, H_k, Y) \sim \mathcal{N}_c(h_{ks}, \Sigma_{ks}),
   \]  
   (127)

   with \( \Sigma_{ks} \overset{\Delta}{=} \Sigma_{k0}^{-1} + \sum_{i=0}^{M-1} S_{k,i}^H(b_k) \Sigma^{-1} S_{k,i}(b_k), \) 

   and \( h_{ks} \overset{\Delta}{=} \Sigma_{ks} \left[ \Sigma_{k0}^{-1} h_{k0} + \sum_{i=0}^{M-1} S_{k,i}(b_k) \Sigma^{-1} \left( r[i] - \sum_{j \neq k} S_{j,i}(b_j) h_j \right) \right] \) (129)

2. The conditional distribution of the noise covariance matrix \( \Sigma \) given \( H, B, \) and \( Y \) is

   \[
   p(\Sigma \mid H, B, Y) \sim \mathcal{W}_c^{-1}(\Psi + Q, m + M),
   \]  
   (130)

   with \( Q \overset{\Delta}{=} \sum_{i=0}^{M-1} \left( r[i] - \sum_{k=1}^{K} S_{k,i}(b_k) h_k \right) \left( r[i] - \sum_{k=1}^{K} S_{k,i}(b_k) h_k \right)^H \).  
   (131)

   Therefore, the conditional distribution of the inverse covariance matrix \( \Sigma^{-1} \) given \( H, B, \) and \( Y \) is

   \[
   p(\Sigma^{-1} \mid H, B, Y) \sim \mathcal{W}_c \left( (\Psi + Q)^{-1}, m + M \right).
   \]  
   (132)

3. The conditional distribution of the data bit \( b_k[i] \) given \( H, \Sigma^{-1}, B_{ki}, \) and \( Y \) can be obtained from (where \( B_{ki} \overset{\Delta}{=} B \setminus b_k[i] \))

   \[
   \frac{P(b_k[i] = +1 \mid H, \Sigma^{-1}, B_{ki}, Y)}{P(b_k[i] = -1 \mid H, \Sigma^{-1}, B_{ki}, Y)} = \exp \left( b_k[i + 1] \rho_k[i + 1] + b_k[i - 1] \rho_k[i] - \text{tr} \left( \Delta Q \Sigma^{-1} \right) \right),
   \]  
   (133)

   with \( \Delta Q \overset{\Delta}{=} -4 \mathcal{R} \left\{ r_k[i] h_k^H C_{k,i}^{(0)T} + r_k[i + 1] h_k^H C_{k,i+1}^{(1)T} + r_k[i + 2] h_k^H C_{k,i+2}^{(2)T} \right\} \) (134)

   \[
   r_k[l] \overset{\Delta}{=} r[l] - \sum_{j \neq k} S_{j,i}(b_j) h_j - S_{k,i}(b_{kl}^0) h_k,
   \]  
   (135)

   where \( b_{kl}^0 \overset{\Delta}{=} \{b_k[0], \ldots, b_k[l - 1], 0, b_k[l + 1], \ldots, b_k[M - 1] \} \).

Using the above conditional posterior distribution, the Gibbs sampling implementation of the blind Bayesian multiuser detector in colored Gaussian noise proceeds iteratively as follows. As in the case of white Gaussian noise, the a posteriori symbol probability \( P[a_k[i] = +1 \mid Y] \) can also be computed by (123).

**Algorithm 3** [Blind Bayesian Multiuser Detection in Colored Gaussian Noise]
1. Draw the initial values $\theta^{(0)} = \{H^{(0)}, \Sigma^{-1(0)}, B^{(0)}\}$ from their prior distributions.
2. For $n = 1, 2, \ldots$
   (a) For $k = 1, 2, \ldots, K$
   i. Draw $h_k^{(n)}$ from $p(h_k | H_k^{(n-1)}, \Sigma^{-1(n-1)}, B^{(n-1)}, Y)$ given by (127).
   (b) Draw $\Sigma^{-1(n)}$ from $p(\Sigma^{-1(n)} | H^{(n)}, B^{(n-1)}, Y)$ given by (132).
   (c) For $i = 0, 1, \ldots, M - 1$
      i. For $k = 1, 2, \ldots, K$
         A. Draw $b_k^{(n)}[i]$ from $P(b_k[i] | h_k^{(n)}, \Sigma^{-1(n)}, B_k^{(n-1)}, Y)$ given by (133).
         B. Compute $a_k^{(n)}[i] = b_k^{(n)}[i] b_k^{(n)}[i - 1]$ (\star).

where $H_k^{(n-1)} \triangleq \{h_1^{(n)}, \ldots, h_{k-1}^{(n)}, h_{k+1}^{(n-1)}, \ldots, h_K^{(n-1)}\}$ and $B_k^{(n-1)} \triangleq \{b_1^{(n)}[0], \ldots, b_1^{(n)}[M - 1], \ldots, b_K^{(n)}[i - 1], b_K^{(n-1)}[i + 1], \ldots, b_K^{(n-1)}[M - 1]\}$.

3.3 Simulation Examples

In this section, we provide a number of simulation examples to illustrate the performance of the blind Bayesian multiuser detectors. We consider a CDMA system with processing gain $N = 10$. The long spreading sequences of all users are generated randomly. In all the simulations described in this section, the following non-informative conjugate prior distributions are used in the Gibbs sampler. For the case of white Gaussian noise,

\[
p(h_k^{(0)}) \sim \mathcal{N}(h_{k0}, \Sigma_{k0}) : h_{k0} = 0, \quad \Sigma_{k0} = 1000 I;
\]

\[
p(\sigma^{2(0)}) \sim \chi^{-2}(\nu_0, \lambda_0) : \nu_0 = 2, \quad \lambda_0 = 0.3;
\]

and for the case of colored Gaussian noise,

\[
p(h_k^{(0)}) \sim \mathcal{N}(h_{k0}, \Sigma_{k0}) : h_{k0} = 0, \quad \Sigma_{k0} = 1000 I;
\]

\[
p(\Sigma^{-1(0)}) \sim \mathcal{W}_c(\Psi^{-1}, m) : \Psi = 0.01 I, \quad m = 2.
\]

In all the simulations related to colored Gaussian noise, SNR is used to denote the in-cell user signal to WGN ratio, SIR is used to denote the in-cell user signal to NBI ratio. The narrowband interference (NBI) is modeled as a 2nd order AR model with coefficients $a_1 = 1.8, a_2 = -0.81$ in (105). The out-cell multiple-access interference (OMAI) is generated according to (103) with energy $12$dB below the in-cell user and the number of out-cell users is set as $K' = 6K$. The number of path for each user is $L = 3$; the transmitter delay $t_k$ is generated randomly with the restriction $t_k < P$. For each data block, the Gibbs sampling is performed for 100 iterations, with the first 50 iterations as the “burning-in” period, i.e., $n_0 = P = 50$ in (123).

**Convergence behavior of Bayesian multiuser detectors:** We first illustrate the convergence behavior of proposed blind Bayesian multiuser detector in white Gaussian noise. In Fig. 5, we plot the first 100 samples drawn by the Gibbs sampler of $h_1$ and $\sigma^2$. The corresponding true values of these quantities are also shown in the same figure with dashed lines. It is seen that the Gibbs sampler reaches convergence within the first several iterations.
Next, we illustrate the convergence behavior of the blind Bayesian multiuser detector in colored Gaussian noise. The channel responses of in-cell users are generated randomly with normalized energy, and the channel response of the out-cell users are generated randomly with energy 12dB below. In Fig. 6, we plot the first 100 samples drawn by the Gibbs sampler of \( h_2 \) and \( \Sigma^{-1}(1, 2) \). The corresponding true values of \( -h_2 \) and \( \Sigma^{-1}(1, 2) \) are also shown in the same figure with dashed lines. Again, it is seen that the Gibbs sampler reaches convergence within the first several iterations. The channel response samples converges to \( -h_k \) or \( h_k \) randomly due to the phase ambiguity. It is seen that \( \Sigma^{-1}(1, 2) \) is far from 0, which indicates that the noise covariance matrix is not diagonal any more with the existence of OMAI and NBI.

![Figure 5: Samples drawn by the Gibbs sampler for the case of white Gaussian noise with \( K = 3, E_b/N_0 = 8dB \) for all the users.](image)

Performance of Bayesian multiuser detectors:

Fig. 7 illustrates the performance of the Gibbs multiuser detector in white Gaussian noise with different number of in-cell users. The bit error rate for \( \{a_k[i]\} \) is averaged among all the users, and then plotted. The RAKE receiver and the nonlinear parallel interference cancellation (PIC) receiver [44] are also implemented assuming perfect channel knowledge. The performance of the RAKE receiver and the performance of PIC after five iterations are also shown in the same figure for the purpose of comparison. It is seen that at reasonable SNR, the performance of the Gibbs multiuser detector is better than that of the other two methods where perfect channel knowledge is assumed. The performance gain over the other two methods increases as the number of users increases.

In order to demonstrate the performance of the Gibbs multiuser detector in colored Gaussian noise, in Fig. 8, we compare its performance with that of the following receiver schemes: (1) Linear MMSE multiuser detector, where we assume that the multipath channels for all in-cell users are known to the receiver. (2) Genie-aided PIC detector, where we assume that a genie provides the receiver with an observation of signal-free NBI corrupted by additive ambient noise and OMAI with the same statistics. We can then use a Kalman
filter to obtain an estimate of the NBI signal [46]. After subtracting the estimated NBI signal from the observation \( \{r_i\}_i \), a PIC receiver is implemented assuming perfect channel knowledge. (3) Single-user bound without NBI, where we assume that there is no NBI. Rake receiver is implemented for single user CDMA system with the same component of white ambient noise and OMAI. It is clear that this detector provides a lower bound to the system we discussed here.

Note that the three approaches given above assume perfect channel knowledge as well as other side information about the channel. For example, in the linear MMSE detector, the covariance matrix of the combined NBI, OMAI and noise is assumed known; in genie-aided PIC detector, a genie observation is assumed to be available for estimating NBI signal; in single user bound, both NBI and other in-cell users are assumed perfectly known. Hence such performance comparisons are unfavorable to the Gibbs multiuser detector. Nevertheless, it is seen in Fig. 8 that at reasonable SIR, the Gibbs detector outperforms the other two receivers (linear MMSE detector and genie-aided PIC detector) and approaches a near single-user bound performance, which demonstrates that the blind Bayesian multiuser detector under the colored Gaussian noise assumption is effective for combating unknown NBI and OMAI.

Finally, as mentioned earlier, when employed in the context of iterative demodulation and decoding, the Bayesian multiuser detector becomes the key component of a blind turbo multiuser detector, which performs joint channel estimation, demodulation, and decoding. See [56] for details.
Figure 7: Performance of blind Bayesian multiuser detector in white Gaussian noise, assuming all the in-cell users have same energy, (a) $K = 3$; (b) $K = 5$; (c) $K = 7$.

Figure 8: Performance of blind Bayesian multiuser detector in colored Gaussian noise with $K = 3$, $K' = 18$ and fixed SNR= 15dB for all in-cell users.
4 Multiuser Detection for Long-code CDMA in Fast-Fading Channels

In the preceding two sections we have assumed that that fading channels remain fixed for the duration of an entire data burst, i.e., the slow-fading case. In this section, we treat the fast-fading scenario where the fading channels vary from symbol to symbol. A sequential multiuser detection algorithms is outlined for asynchronous long-code CDMA uplink over unknown multipath fast fading channels. With the prior knowledge of only the signature waveforms, the delays and the second-order statistics of the fading channel, the receivers sequentially estimate the channel using the sequential EM algorithm. The snapshot estimates for each path are tracked by linear MMSE filters and the user data are detected by maximum likelihood sequence estimation conditioned on the channel estimates.

4.1 Channel Model and Sequential EM Algorithm

The received continuous-time signal model is of similar form as (99), except now the channel responses depend on the symbol index $i$, i.e.,

$$ r(t) = \sum_{k=1}^{K} \sum_{i=0}^{M-1} b_k[i] \sum_{l=1}^{L} \alpha_{l,k}[i] s_k[i] (t - iT - \tau_{l,k}) + v(t). \quad (136) $$

Accordingly the received discrete-time signal vector corresponding to the $i$-th symbol interval is given by

$$ r[i] = \sum_{k=1}^{K} \left( b_k[i] C_{k,i}^{(0)} + b_k[i-1] C_{k,i-1}^{(1)} + b_k[i-2] C_{k,i-2}^{(2)} \right) h_k[i] + v[i], \quad i = 0, 1, \ldots \quad (137) $$

Our goal is to sequentially track the fading channel responses $\{h_k[i], i = 0, 1, \ldots\}, \forall k$, and at the same time to estimate the transmitted data symbols $\{b_k[i], i = 0, 1, \ldots\}, \forall k$. To that end, we make use of the sequential EM algorithm [42, 53]. Next we briefly introduce this technique.

Suppose $y_1, y_2, \ldots$ are a sequence of observations with probability density function (pdf) $f(y|t)$, where $t \in \mathbb{C}^m$ is a static parameter vector, for some $m$. A class of sequential estimators derived from the maximum-likelihood principle is given by

$$ t[i+1] = t[i] + \Pi(y_{i+1}, t[i]) s(y_{i+1}, t[i]), \quad (138) $$

where $t[i]$ is the estimate of $t$ at the $i$-th step; $\Pi(y_{i+1}, t[i])$ is an $m \times m$ matrix defined below; and

$$ s(y_{i+1}, t[i]) \triangleq \left[ \frac{\partial}{\partial \theta_1^*} \log f(y_{i+1}|t), \ldots, \frac{\partial}{\partial \theta_m^*} \log f(y_{i+1}|t) \right]^T \bigg|_{t=t[i]} \quad (139) $$

is the score (i.e., the gradient of the log-likelihood function). Let $H(y_i, t[i])$ denote the Hessian matrix of $\log f(y_i|t[i])$, where

$$ H_{j,k}(y_i, t[i]) = \left. \frac{\partial^2}{\partial \theta_j^* \partial \theta_k} \log f(y_i|t) \right|_{t=t[i]}, \quad j = 1, \ldots, m, \quad k = 1, \ldots, m. \quad (140) $$
Let $x_i$ denote a “complete” data related to $y_i$, for $i = 1, 2, \ldots$. The complete data $x_i$ is selected in the (sequential) EM algorithms such that $y_i$ can be obtained through a many-to-one mapping $x_i \rightarrow y_i$, and their knowledge makes the estimation problem easy (for example, the conditional density $f(x_i|t)$ can be easily obtained.) Denote the Fisher information matrix of the data $y_i$ and $x_i$, respectively, as

$$I(t^{[i]}) = -E[H(y_i, t^{[i]})]$$

and

$$I_c(t^{[i]}) = -E[H(x_i, t^{[i]})].$$

Different algorithms are characterized by different choices of the function $\Pi(y_{i+1}, t^{[i]})$ in (138).

- The sequential EM algorithm:

$$\Pi(y_{i+1}, t^{[i]}) = \frac{1}{i} I_c^{-1}(t^{[i]}) = \frac{1}{i} I^{-1}(t^{[i]}).$$

(141)

The consistency and asymptotic normality of the algorithm is reported in [42].

- The Newton-Raphson algorithm:

$$\Pi(y_{i+1}, t^{[i]}) = -H^{-1}(y_{i+1}, t^{[i]}).$$

(142)

- A stochastic approximation procedure:

$$\Pi(y_{i+1}, t^{[i]}) = \frac{1}{i} I^{-1}(t^{[i]}).$$

(143)

Note that, for independent and identically distributed (i.i.d.) observations $\{y_i\}$, if $i$ in (143) is substituted by $[i + 1]$, we obtain the maximum-likelihood estimator (MLE) of $t$ for exponential families [42]. The asymptotic distribution of this procedure can be found in [39, 8].

- If $\Pi(y_{i+1}, t^{[i]})$ is a constant diagonal matrix with small elements, then (138) is the conventional steepest-descent algorithm. Some other choices of $\Pi(y_{i+1}, t^{[i]})$ are suggested in [42].

- For time-variant parameters $\{t[i]\}$, a conventional approach suggested in [10, 28] is to substitute the converging series $1/i$ in (141) with a small positive constant $\lambda_0$. The new estimator is given by

$$\hat{\theta}[i + 1] = \hat{\theta}[i] + \lambda_0 I_c^{-1}\left(\hat{\theta}[i]\right) s(y_{i+1}, \hat{\theta}[i]),$$

(144)

where $\hat{\theta}[i]$ is the estimate of $t[i]$. 

30
4.2 Sequential Blind Multiuser Detector

In the sequential EM-based multiuser detector, at time \(t\), we have the following steps:

1. Localization:
   \[
   \hat{h}[i] = \hat{h}[i] + \lambda_0 I_c^{-1}(\tilde{h}[i]) s(r[i], \tilde{h}[i]),
   \]  
   (145)

   where \(\tilde{h}[i]\) is a one-step prediction of \(g[i]\) provided by Step 3.

2. Multiuser detection:
   \[
   \hat{b}[i] = \arg \max_{b[i]} f(r[i], \hat{h}[i], b[i]).
   \]  
   (146)

3. Tracking:
   \[
   \tilde{h}_{kl}[i+1] = \alpha_{kl}[i] \hat{h}_{kl}(i-q) \cdots \hat{h}_{kl}[i]^T, 
   \]  
   (147)

   where \(\alpha_{kl}[i]\) is a \((q+1)\)-order linear MMSE filter for \(\hat{h}_{kl}[i]\).

Based on the signal model (137), the the quantities in the algorithm above [i.e., \(s(r(i), \tilde{h}(i)), I_c(\tilde{h}(i))\), \(\hat{b}(i)\), and \(\alpha_{kl}(i)\)] In particular, a direct evaluation of the score and the information matrix of the incomplete data is prohibitive. By introducing appropriate complete data, an approximation to \(s(z(i), \tilde{h}(i))\) based on the EM algorithm can be obtained. As a byproduct, a low-complexity detector of the data \(b(i)\) can also be obtained. The instantaneous MMSE filter \(\alpha_{kl}(i)\) can also be derived. Due to space limitations, we do not provide the detailed derivations here. They can be found in [26].

4.3 Simulation Results

Next we illustrate the performance of sequential EM multiuser detector in an unknown multipath fading CDMA channel by a simulation example. The delays of users’ paths were randomly generated and then fixed for the simulation. Each user has two \((L = 2)\) equal-energy paths. The time-variant fading coefficients were randomly generated from Clarke’s model to simulate a Rayleigh fading channel with carrier frequency 1850 MHz, data rate 144kb/s, vehicle speed = 52 miles/hour, i.e., normalized bandwidth-time product \(BT = 0.001\) for all paths. We consider a reverse link of an asynchronous CDMA system with five users \((K = 5)\). The spreading sequences of each user and the data bits of each user are independently and randomly generated. The processing gain is \(N = 12\). All users have equal signal amplitudes. To remove the sign ambiguity of the tracked fading process, the transmitted bits are differentially encoded.

From simulation results, after 200 symbols, the estimator enters a steady state. A transition state is plotted in Fig. 9. The average bit error rates (BER) of five users versus \(\frac{E_b}{N_0}\) are plotted in Fig. 10, where \(E_b\) is the energy per bit. The figure shows the performance of the pilot-assisted multiuser receiver with pilot insertion rates 1/500 and 1/100 reaches an error floor due to ISI and MAI. The figure also shows the performance of the pilot-assisted single-user RAKE receiver in a single-user environment with pilot insertion rate
1/500 and the performance of the proposed sequential-EM and approximate sequential-EM receivers, including analytical bounds for each. It is seen that significant performance gains are achieved by the proposed receivers compared to pilot-assisted receivers. Both the performance of the approximate sequential-EM receiver with three iterations and the performance of the sequential-EM receiver are very close to that for known channels. This can be attributed to the quality of the channel estimation for both receivers, as seen in Fig. 9 and Fig. 10. Moreover, the performance difference between the sequential-EM receiver and the approximate sequential-EM receiver is small in the low and medium SNR regions, and it is about 1dB at high SNRs. Furthermore, it is seen from the simulation results that the proposed multiuser receivers in a multiuser channel even outperform the RAKE receiver in a single-user channel. This is because the RAKE receiver makes the assumption that the delayed signals from different paths for each user are orthogonal, which effectively neglects the intersymbol interference.

5 Transmitter-Based Multiuser Precoding for Fading Channels

The receiver-based multiuser detection techniques presented in this chapter allow system designers to trade receiver complexity for improved multiuser system performance. In many applications, however, it is useful to have the option of moving complexity away from the receiver to the transmitter. Cellular service providers, for example, would prefer to keep mobile unit costs to a minimum so they can continue to entice customers with free phones. Similarly, heterogeneous ad hoc or wireless sensor networks may be composed of nodes
Figure 10: BER Performance comparison between the proposed multiuser receivers and the pilot-assisted multiuser receivers in a multipath fading channel with the number of users $K = 5$ and $1$, the number of paths $L = 2$, processing gain $N = 12$, $BT = 0.001$.

with widely varying power constraints and computational capabilities, making the option of moving complexity where it can be managed most efficiently an attractive option.

In this section, we present transmitter-based techniques for enhancing multiuser system performance, i.e., transmitter-based multiuser detection. More specifically, we develop linear transmitter precoding strategies for multiple-access interference suppression and diversity exploitation when no receiver channel state information (CSI) is available and receivers are restricted to low-complexity matched-filter detection. In the context of cellular systems, transmitter precoding would be appropriate for the downlink, where the complexity of mobile units should be kept to a minimum. There has been a significant volume of work in the area of joint transmitter/receiver design when receiver CSI or feedback is available [41, 40, 54, 9] and on diversity transmission without receiver CSI, but with maximum-likelihood reception [15, 25, 16]. Our focus, however, is on precoding jointly for diversity and interference suppression in systems that require ultra-low complexity receivers, that is, matched filtering without receiver-based channel estimation. Precoding for multiple access systems, as previously developed, focuses on transmitter-based MAI suppression. The authors in [47], for example, developed minimum mean square error precoders for synchronous code division multiple access (CDMA) in additive white Gaussian noise channels. They also presented an extension to multipath channels, but a RAKE receiver is required and the channel is assumed perfectly known at the receiver. These initial results were promising, showing that precoding outperformed decorrelating receiver-based multiuser detection in some cases. In [2], the authors considered transmitter precoding for multipath fading channels but, in contrast to the present work, their prefilter is applied to the output of the spread spectrum encoder, rather than applying the filter first, followed by spreading. It was shown that this approach has inferior average performance unless the spreading codes
themselves are allowed to be adaptive. In [5], the authors developed a simple but remarkable precoding technique for exploiting multipath diversity that requires no receiver CSI. This technique, called pre-rake diversity combining, will be used in the present work.

Our approach differs from that of existing information-theoretic precoding work [4] in that we are precoding to minimize mean square error instead of maximizing capacity, mutual information or some other information-theoretic criterion. That is, we are optimizing performance, in terms of interference suppression and diversity gain, while keeping the rate fixed. In contrast to these works, we are also interested in very low complexity joint decoding and detection (via the matched filter). The QR-decomposition, “writing-on-dirty-paper” based pre-subtraction approach [57] and related non-linear TH precoding schemes are also not immediately applicable because it can require more receiver complexity than we are willing to tolerate here. Portions of this material were first published in [35, 38].

5.1 Basic Approach and Adaptation

5.1.1 Uplink Signal Model and Blind Channel Estimation

We consider a $K$-user discrete-time synchronous multipath CDMA system with no inter-symbol interference (ISI)$^1$. Such a system is realized either by neglecting the ISI when the multipath delay spread is small compared with the symbol interval, or by inserting guard intervals between symbols when the delay spread is large. The path delays are also assumed to be an integral number of chip periods and are known. We first consider the chip-match filtered uplink signal received at the base station which, during the $i$-th symbol interval, can be written as

$$
r[i] = \sum_{k=1}^{K} b_k[i] \sum_{l=1}^{L} s_{l,k} f_{l,k} + n[i] \quad (148)
$$

where $L$ is the number of resolvable paths, $b_k[i]$ is the $i$-th symbol for the $k$-th user, $s_{l,k}$ and $f_{l,k}$ are, respectively, the delayed versions of the spreading waveform (with zero-padding when a guard interval is inserted) and the complex channel fading gain corresponding to the $l$-th path of the $k$-th user; $n[i] \sim \mathcal{N}_c(0, \sigma^2 I_N)$ is a complex white Gaussian noise vector. Note that $r[i], n[i] \in \mathbb{C}^N$ where $N$ is the processing gain. Denote

$$
S_k \triangleq [s_{1,k} \ s_{2,k} \cdots \ s_{L,k}]^T, \quad (149)
$$

$$
f_k \triangleq [f_{1,k} \ f_{2,k} \cdots \ f_{L,k}]^T. \quad (150)
$$

Then (148) can be written as

$$
r[i] = \sum_{k=1}^{K} S_k f_k b_k[i] + n[i] \quad (151)
$$

$$
= Hb[i] + n[i] \quad (152)
$$

where

$$
H \triangleq [h_1 \ h_2 \cdots \ h_K], \quad (153)
$$

$$
b[i] \triangleq [b_1[i] \ b_2[i] \cdots \ b_K[i]]^T. \quad (154)
$$

$^1$We will consider multipath ISI channels later in this chapter.
A block diagram of the uplink system appears in Fig. 11.

As in Section 2.1, we can define the autocorrelation matrix of the received signal \( r[i] \) as

\[
C_r \triangleq E \left[ r[i] r[i]^H \right] = HH^H + \sigma^2 I_N \tag{155}
\]

\[
= U_s \Lambda_s U_s^H + \sigma^2 U_n U_n^H \tag{156}
\]

where (156) is the eigendecomposition of \( C_r \). Since the matrix \( H \) has full column rank \( K \), the matrix \( HH^H \) in (155) has rank \( K \). Therefore, in (156) \( \Lambda_s \) contains the \( K \) largest eigenvalues of \( C_r \); \( U_s \) contains the corresponding orthonormal eigenvectors; and \( U_n \) contains the \((N - K)\) orthonormal eigenvectors that correspond to the smallest eigenvalue, \( \sigma^2 \).

Suppose User 1 is the user of interest. Then since \( U_n^H h_1 = U_n^H S_1 f_1 = 0 \), we can estimate \( f_1 \) at the base station in the following way \([1, 27, 43, 49]\)

\[
\hat{f}_1 = \arg \min_{\|f\|=1} \|U_n^H S_1 f\|^2 \tag{157}
\]

\[
= \arg \min_{\|f\|=1} f^H S_1^H U_n U_n^H S_1 f \tag{158}
\]

\[
= \text{minimum eigenvector of } Q. \tag{159}
\]

Note that (159) specifies \( f_1 \) up to a scale and phase ambiguity and that, in practice, (159) can be implemented blindly in a batch or sequential adaptive manner. In batch mode, we simply replace the noise subspace parameters in (158) with parameters obtained from the eigendecomposition of the sample autocorrelation matrix of the received signal, as discussed in Section 2.1. In sequential adaptive mode, where we update the channel estimates at each time slot, we can employ a suitable subspace tracking algorithm and use the sequential Kalman filtering technique described in Section 2.2.5\(^2\). Note that a necessary condition for our channel estimate to be unique is that \( H \) have rank \( K \), which necessitates that this matrix be tall, i.e., \( K \leq N \).

5.1.2 Downlink Signal Model and Matched Filter Detection

The (downlink) signal transmitted from the base station during the \( i \)-th symbol interval can be written

\[
x[i] = SMb[i] \tag{160}
\]

where

\[
S \triangleq [s_1 \ s_2 \ \cdots \ s_K] \tag{161}
\]

is the matrix of spreading waveforms and \( M \in \mathbb{C}^{K \times K} \) is a complex precoding filter which we will optimize in the following section. Throughout this section, we assume that the CDMA system is operating in the time division duplex mode, so that the downlink and uplink operate using the same carrier frequency in different time slots. We also assume that the time elapsing between uplink and downlink transmissions is sufficiently small compared to the coherence time of the channel that the channel impulse response is the same for the

\(^2\)See Section 5.1.4 for more specifics.
uplink and downlink. Then from (149) and (150), the received signal at User 1’s mobile unit can be written as

\[ r_1[i] = \begin{bmatrix} S_1 f_1 & S_2 f_1 & \cdots & S_K f_1 \end{bmatrix} H_1 \begin{bmatrix} Mb[i] + n_1[i] \end{bmatrix} \] (162)

where \( S_1, S_2, \ldots, S_K \) contain shifted versions of their respective signature waveforms as in (149) except that the \( L \) shifts are the same for each user’s waveform since all spreading codes have been transmitted over User 1’s downlink channel. Detection of the downlink information bits is accomplished via matched filtering of the received signal \( r_1[i] \) with User 1’s signature waveform, \( s_1 \). Fig. 12 contains a block diagram of the signal processing that takes place at the base station when an adaptive implementation is employed. We will see more details in Section 5.1.4.

5.1.3 Transmitter Precoding for a Synchronous Multipath Downlink

We seek to choose the precoding matrix \( M \) so as to provide the best downlink performance possible when the mobile units are constrained to the use of a matched filter receiver. We choose the minimum mean-square error criterion, so \( M \) is chosen to minimize

\[ J = E \left[ \left\| b - H Mb - v \right\|^2 \right] \] (163)

where we have dropped the time index for clarity. It is easy to see that

\[ \begin{bmatrix} s_1^H r_1 \\ s_2^H r_2 \\ \vdots \\ s_K^H r_K \end{bmatrix} = \begin{bmatrix} s_1^H H_1 \\ s_2^H H_2 \\ \vdots \\ s_K^H H_K \end{bmatrix} \mathcal{H} + \begin{bmatrix} s_1^H n_1 \\ s_2^H n_2 \\ \vdots \\ s_K^H n_K \end{bmatrix} \] (164)

Then

\[ J = E \left[ \left\| b - \mathcal{H}Mb - v \right\|^2 \right] \] (165)

The following proposition gives the optimal precoding matrix.
Figure 12: Adaptive precoding transmitter structure at the base station for the downlink signal.
Proposition 1 The choice of $M$ that minimizes $J$ is $M = \mathcal{H}^{-1}$.

The proof appears in [35].

Denote by $\hat{H}_i$ ($1 \leq i \leq K$) the matrix $H_i$ where the channel $f_i$ has been replaced with the blind estimate $\hat{f}_i$ obtained from (159). Then we may form an initial blind estimate of $M$ at the base station as

$$\hat{M} = \begin{bmatrix} s_H^H \hat{H}_1 \\ s_H^H \hat{H}_2 \\ \vdots \\ s_H^H \hat{H}_K \end{bmatrix}^{-1}.$$  \hspace{1cm} (166)

There remain amplitude and phase ambiguities in $\hat{M}$ that are addressed in the following sections.

Notice that choosing the optimal precoding matrix by minimizing $J$ places no explicit constraint on average transmit power. In fact, it was found in a related work on MMSE precoding [47] that unconstrained optimization with simple power scaling provides superior performance at high SNR to constrained optimization. As a result, we shall focus on the former. We will also suppress the power scale factor for simplicity.

5.1.4 Adaptive Implementation

In this section we present an adaptive implementation of the transmitter precoding strategy discussed in Section 5.1.3 that updates the precoding matrix $M$ at each time slot. This sequential updating allows the implementation to adapt as the channel changes and as users enter and leave the system. A block diagram of the signal processing at the base station appears in Fig. 12. Note that we have suppressed the signal processing necessary for detection of the uplink bits. The uplink signal received at the base station is used in a signal subspace tracker, along with the known spreading codes of all users, to construct channel estimates as discussed in the following section. Recall that since we are assuming TDD mode, the uplink channel estimates also serve as downlink channel estimates that can be used to construct $M$. As previously mentioned, these channel estimates have amplitude and phase ambiguities. Since nearly all cellular CDMA systems employ power control, it is likely that the base station has some knowledge of each users’ transmit power. This information, coupled with estimates of the received power, can be used to estimate the channel amplitude for each user. More specifically, let the diagonal matrices $A = \text{diag}(\alpha_1, \alpha_2, \ldots, \alpha_K)$ and $P = \text{diag}(\sqrt{p_1}, \sqrt{p_2}, \ldots, \sqrt{p_K})$ contain the unknown channel amplitudes and the known uplink transmit powers, respectively. Also define $\tilde{H}$ such that $H = \tilde{H}AP$, so that the columns of $\tilde{H}$ have unit norm.

We propose an estimator based on the following fact.

Proposition 2

$$A = \left[ \tilde{H}^H U_s \left( \Lambda_s - \sigma^2 I_K \right)^{-1} U_s^H \tilde{H}P^2 \right]^{-1/2} \hspace{1cm} (167)$$

where $U_s, \Lambda_s$ are signal subspace components derived from an eigendecomposition of the autocorrelation matrix, $C_r$, of the received signal, as in (155) and (156).
The proof appears in [35].

We may obtain an estimate, \( \hat{A} \), of \( A \) by replacing \( \bar{H}, U, \Lambda \) and \( \sigma^2 \) of (167) with their respective estimates obtained from subspace tracking and the Kalman channel estimator as discussed in the next section. The subspace tracker we have chosen to use for this adaptive implementation is NAHJ-FST (noise averaged Hermitian-Jacobi fast subspace tracking), which has complexity \( O(NK) \) floating operations per user per bit and which performs close to the lower bound for all SVD-based subspace trackers [37]. The application of NAHJ-FST to the current tracking problem is a straightforward modification of [37] and will not be discussed in detail. The channel phase ambiguity can be circumvented by the use of differential encoding and decoding of the data. After channel and amplitude estimation, the amplitude corrected channel information is then used, along with the known spreading codes, to construct the precoding matrix \( \hat{M} \). Finally, the downlink information bits are differentially encoded and filtered with \( \hat{M} \) before spreading and transmission. At the mobile unit, matched filtering and differential detection are performed to obtain estimates of the downlink information bits.

5.1.5 Algorithm Summary

Making use of the sequential adaptive Kalman channel estimator discussed in Section 2.2.5\(^3\), we can produce an adaptive implementation of transmitter precoding at the base station as follows.

Algorithm 4 [Sequential adaptive transmitter precoding for synchronous multipath CDMA]

1. Using a suitable signal subspace tracking algorithm, e.g. NAHJ-FST, update the signal subspace components \( U_s[i], \Lambda_s[i], \) and \( \sigma^2[i] \) at each time slot \( i \) using the uplink signals.

2. Track the channels \( \{f_k\}_{k=1}^K \) as follows:

\[
\begin{align*}
\mathbf{z}[i] &= \mathbf{r}[i] - U_s[i]U_s[i]^H \mathbf{r}[i], \\
\mathbf{x}[i] &= S_k^H \mathbf{z}[i], \\
f_k[i] &= \Sigma[i-1] \mathbf{x}[i] (\mathbf{x}[i]^H \Sigma[i-1] \mathbf{x}[i])^{-1}, \\
f_k[i] &= \|f_k[i-1] - \mathbf{k}[i] (\mathbf{x}[i]^H f_k[i-1])\|, \\
\Sigma[i] &= \Sigma[i-1] - \mathbf{k}[i] \mathbf{x}[i]^H \Sigma[i-1].
\end{align*}
\]

3. Calculate the channel amplitudes via (167) using the channel estimates, the signal subspace parameters, the known spreading codes, and the known transmit powers.

4. Using (164) and the information from steps 1-3, calculate \( \mathcal{H} \) and set \( \mathbf{M} = \mathcal{H}^{-1} \).

5. Differentially encode the downlink bit streams for each user to form \( \mathbf{b}[i] \).

6. Transmit the precoded downlink signal \( \mathbf{x}[i] = \mathbf{S}\mathbf{b}[i] \).

7. Perform matched filtering and differential detection at the mobile units.

\(^3\)The parameters \( \mathbf{h}_1 \) and \( \Xi_1 \) from Section 2.2.5 are replaced with \( f_1 \) and \( S_1 \) in the current section.
5.2 Precoding with Multiple Transmit Antennas

5.2.1 Downlink Signal Model

Without receiver channel state information (CSI), it is difficult to fully exploit receive antenna diversity because the only diversity combining available at the receiver is (non-coherent) addition of the antenna outputs. This provides no diversity in fading environments [14]. In a block fading environment with transmitter CSI, however, we can employ selection diversity with multiple receive antennas simply by adding a few bits to each frame to instruct the receiver to use the “best” antenna. This possibility notwithstanding, we consider a $K$-user downlink CDMA system in flat block fading with two transmit antennas and a single receive antenna for each user. Extensions to more than two transmit antennas are straightforward. The discrete-time BPSK modulated signal transmitted from antenna $a \in \{1, 2\}$ is

$$x^{(a)} = \alpha S M^{(a)}b$$ (168)

where, similar to (160), the columns of $S \in \mathbb{C}^{N \times K}$ are the normalized spreading codes of the $K$ users, $b \in \{\pm 1\}^K$ contains the downlink bits corresponding to the $K$ users, $N$ is the processing gain, and $M^{(a)} \in \mathbb{C}^{K \times K}$ is a complex precoding matrix used for multiple-access interference (MAI) suppression and transmitter antenna diversity exploitation and is optimized in later sections. The scalar $\alpha$ is a transmit power factor that will be addressed in a later section. For now, we assume $\alpha = 1$.

The goal is to choose $M^{(1)}$ and $M^{(2)}$ to optimize downlink performance when no receiver CSI is available and the receiver is constrained to matched filter detection. The precoders must not only suppress interference, but they must also exploit available diversity. We are interested in situations in which we have either perfect or partial CSI available at the transmitter.

5.2.2 Precoder Design for Orthogonal Spreading Codes

After chip-matched filtering, the noise free received signal at user 1’s mobile unit is

$$r_1 = h_1^{(1)}x^{(1)} + h_1^{(2)}x^{(2)}$$

$$= h_1^{(1)}SM^{(1)}b + h_1^{(2)}SM^{(2)}b$$ (170)

where $h_b^{(a)}$ is the complex channel gain between transmit antenna $a$ and user $b$’s mobile unit and where we have set $\alpha = 1$. The channel gains are assumed mutually independent. The mobile units are restricted to (spreading-code) matched filter detection. If $s_1 \triangleq [S]_{:,1}$ and we have orthogonal spreading codes, the decision statistic for user 1 is

$$d_1 = s_1^H r_1 + \sigma n_1$$ (171)

$$= h_1^{(1)} [M^{(1)}b]_1 + h_1^{(2)} [M^{(2)}b]_1 + \sigma n_1$$ (172)

where $n_1 \sim \mathcal{N}_c(0, 1)$ and is independent of $b$ and the channel and $\sigma^2$ is the noise power. For now, define $M^{(a)}$, $a \in \{1, 2\}$ to be diagonal matrices whose elements are given by

$$[M^{(a)}]_{i,i} = \frac{h_i^{(a)*}}{\sqrt{|h_i^{(1)}|^2 + |h_i^{(2)}|^2}}.$$ (173)
Then we have

\[ d_1 = \sqrt{|h_1^{(1)}|^2 + |h_1^{(2)}|^2 b_1 + \sigma n_1} \]  

and the corresponding bit estimate is

\[ \hat{b}_1 = \text{sign}\{\text{Re}[d_1]\}. \]

This achieves an instantaneous SNR of

\[ \text{SNR}_1 = \frac{|h_1^{(1)}|^2 + |h_1^{(2)}|^2}{\sigma^2} \]  

which provides full two branch diversity for every user and has a \( X_4^2 \) distribution when the channel gains are independent complex Gaussian random variables. Precoding for this scenario reduces to maximal ratio weighting [14], which has the same performance as beamforming to a single receive antenna. Note that \( M^{(1)}, M^{(2)} \) in (173) are normalized in the sense that the sum of the average (with respect to \( b \)) transmit power from both antennas is \( K \). That is,

\[ E_b \left\{ \|\text{SM}^{(1)}b\|^2 \right\} + E_b \left\{ \|\text{SM}^{(2)}b\|^2 \right\} = \text{tr} \left( \text{M}^{(1)H}\text{M}^{(1)} \right) + \text{tr} \left( \text{M}^{(2)H}\text{M}^{(2)} \right) = K \]  

for every channel realization. Therefore, we can set \( \alpha = 1 \) in (168).

### 5.2.3 Precoder Design for Non-Orthogonal Spreading Codes

Let \( \mathbf{\rho}_1^T \triangleq s_1^H \mathbf{S} \). The decision statistic for user 1, assuming for the moment that \( \alpha = 1 \), is

\[ d_1 = \frac{h_1^{(1)} \mathbf{\rho}_1^T \mathbf{M}^{(1)} b + h_1^{(2)} \mathbf{\rho}_1^T \mathbf{M}^{(2)} b + \sigma n_1}{d_1^{(1)} + d_1^{(2)}}. \]  

Our goal is to choose \( \mathbf{M}^{(1)}, \mathbf{M}^{(2)} \) to maximize the collective performance of all users in some sense, assuming no receiver CSI and matched filter detection. We form minimum mean-square error (MMSE) cost functions for the optimization of \( \mathbf{M}^{(1)}, \mathbf{M}^{(2)} \) by stacking \( d_k^{(1)} (1 \leq k \leq K) \) and \( d_k^{(2)} (1 \leq k \leq K) \), respectively. The result for transmit antenna \( a \in \{1, 2\} \) is

\[ J^{(a)} = E \left[ \left\| \begin{bmatrix} |h_1^{(a)}|^2 (|h_1^{(1)}|^2 + |h_1^{(2)}|^2)^{-\frac{1}{2}} b_1 \\ |h_2^{(a)}|^2 (|h_2^{(1)}|^2 + |h_2^{(2)}|^2)^{-\frac{1}{2}} b_2 \\ \vdots \\ |h_K^{(a)}|^2 (|h_K^{(1)}|^2 + |h_K^{(2)}|^2)^{-\frac{1}{2}} b_K \end{bmatrix} \right\|_F^2 \right] \]

\[ \left\| \mathbf{M}^{(a)} b - \begin{bmatrix} \sigma n_1 \\ \sigma n_2 \\ \vdots \\ \sigma n_K \end{bmatrix} \right\|_2^2 \]  

\[ = E \left[ \left\| \mathbf{D}^{(a)} b - \mathbf{H}^{(a)} \mathbf{R} \mathbf{M}^{(a)} b - \mathbf{n} \right\|^2 \right] \]
where

\[
D^{(a)} = \text{diag} \left( \left[ h_1^{(a)} \right]^2 \left[ \left| h_1^{(1)} \right|^2 + \left| h_1^{(2)} \right|^2 \right]^{-\frac{1}{2}}, \left[ h_2^{(a)} \right]^2 \left[ \left| h_2^{(1)} \right|^2 + \left| h_2^{(2)} \right|^2 \right]^{-\frac{1}{2}}, \ldots, \right)
\]

\[
H^{(a)} = \text{diag} \left( h_1^{(a)}, h_2^{(a)}, \ldots, h_K^{(a)} \right) \quad a \in \{1, 2\}, \quad R \triangleq S^H S,
\]

\[
n = \begin{bmatrix}
\sigma n_1 \\
\sigma n_2 \\
\vdots \\
\sigma n_K 
\end{bmatrix}
\]

and where the expectations are with respect to n and b. At this stage, the cost functions implicitly assume that the channel gains are deterministic and known at the transmitter.

The motivation behind the construction of the cost functions is self evident except, perhaps, for the presence of \(D^{(1)}\) and \(D^{(2)}\). This is related to the transmit power constraint and power loading. If we allow for an infinite peak-to-average power ratio at the transmitter, we can replace \(D^{(1)}\) and \(D^{(2)}\) in (180) with \(I_K\) and the resulting optimal precoding matrix will completely eliminate the detrimental effects of fading\(^4\). Because real transmitters cannot operate with an infinite dynamic range, this is not a reasonable assumption. Therefore, we will insist that the sum of the average (with respect to b) transmit power from all antennas be equal to the number of users. For diversity transmission (instead of multiplexing \([58]\)) with this power constraint and orthogonal codes, the best precoding scheme is maximum ratio weighting, as in 5.2.2. It is therefore important that the precoding matrices that minimize the cost functions \(J^{(1)}, J^{(2)}\) reduce to (173) when spreading codes are orthogonal. It is easy to verify that this is true when \(D^{(a)}\) satisfies (181).

**Proposition 3** The choice of \(M^{(1)}\) that minimizes \(J^{(1)}\) and the choice of \(M^{(2)}\) that minimizes \(J^{(2)}\) are given by

\[
M^{(1)} = R^{-1} \left[ H^{(1)} \right]^{-1} D^{(1)}
\]

\[
M^{(2)} = R^{-1} \left[ H^{(2)} \right]^{-1} D^{(2)}.
\]

The proof appears in [38].

These results show that optimal precoding with perfect transmitter CSI and non-orthogonal codes is maximum ratio weighting followed by transmitter-based decorrelation. Precoder design when partial channel knowledge (defined as quantities statistically dependent upon the true channel) is available at the transmitter is discussed in [38].

### 5.3 Precoding for Multipath ISI Channels

The conventional technique for diversity exploitation in multipath is RAKE reception, i.e., maximum ratio combining of each path at the receiver. Because we are considering

\(^4\)The precoding matrix for this situation can be found using (184)-(185) and by solving for \(\alpha\) using the procedure in Section 5.4.2. Essentially, the transmitter will increase power (perhaps without bound) during fades and decrease power during channel peaks, resulting in infinite peak-to-average transmit power.
applications that do not allow for receiver channel information, we will, instead, use a form of pre-rake diversity combining. We will see that Propositions 1 and 2 can be applied with minor modifications to fully exploit multipath and transmit antenna diversity while completely eliminating multiple-access interference.

5.3.1 Prerake-Diversity Combining

We will assume a synchronous, block fading, $L$-path multipath channel [6] with no intersymbol interference$^5$, where the impulse response between transmit antenna $a$ and user $k$’s receive antenna can be modelled as

$$ h_k^{(a)}(t) = \sum_{l=0}^{L-1} h_k^{(a)}[l] \delta(t - lT_c), \quad (186) $$

where $\{h_k^{(a)}[l]\}_{l=0}^{L-1}$ is a set of i.i.d complex Gaussian random variables and $T_c$ is the chip duration, i.e., the symbol duration divided by the processing gain. Synchronism is a reasonable assumption for downlink transmissions (where precoding is most practical) and intersymbol interference can be eliminated using guardbands or it can simply be neglected if the channel delay spread is small relative to the symbol interval.

The general idea behind pre-rake diversity combining [5] is to transmit precoded versions of the chip stream $\mathbf{Sb} \in \mathbb{C}^{N\times1}$ during $L$ consecutive chip intervals so that after the $L$-th transmission, all paths add up coherently at the receiver. Fig. 13 illustrates the approach for the single antenna, single user case in a 3-path channel. The discrete-time transmitted signal for this scenario is

$$ [\tilde{x}]_i = \left( \sum_{l=0}^{L-1} |h_1[l]|^2 \right)^{-\frac{1}{2}} \sum_{l=0}^{L-1} h_1^*[L - 1 - l] \cdot [s_1 b_1]_{i-l}, \quad i = 1, 2, \ldots N + L - 1. \quad (187) $$

The desired portion of the noise-free received signal $\mathbf{r}_1$ is available between relative chip intervals $L$ and $N + L - 1$ and is given by

$$ \mathbf{r}_1 = \left( \sum_{l=0}^{L-1} |h_1[l]|^2 \right)^{\frac{1}{2}} \mathbf{s}_1 b_1 + \text{multipath/interchip interference}. \quad (188) $$

In the next section we will show that MMSE precoding for a $K$-user system can fully exploit multipath and transmit antenna diversity for every user while completely eliminating multipath and multiuser interference.

5.3.2 Precoder Design

For a $K$-user multi-antenna CDMA system using pre-rake diversity combining, the discrete-time signal transmitted from antenna $a$ is

$$ \mathbf{x}^{(a)} = \sum_{l=0}^{L-1} \hat{\mathbf{S}}[l] \mathbf{M}^{(a)}[l] \mathbf{b} \quad (189) $$

$^5$Inter-chip interference constitutes the multipath interference in this model and in Fig. 13.
Figure 13: Pre-rake diversity combining (precoding) for a single-antenna, single user CDMA signal in a 3-path multipath channel with processing gain $N$. The $N$ elements of $\mathbf{x}$ are scaled and transmitted over $N + 2$ chip intervals. The desired portion of the received signal $r_1$ indicates full diversity is achievable if the multipath/interchip interference can be suppressed via additional precoding.

where $\mathbf{x} \in \mathbb{C}^{(N+L-1) \times 1}$, $\mathbf{M}^{(a)}[l]$ is the precoding matrix for transmit antenna $a$ and path $l$, and $\mathbf{S}[l]$ is defined by

$$
\mathbf{S}[l] \triangleq \begin{bmatrix}
\mathbf{O}_{L-1-l,K} \\
\mathbf{S} \\
\mathbf{0}_{l,K}
\end{bmatrix}.
$$

Then we have

$$
\mathbf{x}^{(a)} = \mathbf{\tilde{S}} \mathbf{M}^{(a)} \mathbf{b}
$$

where

$$
\mathbf{M}^{(a)} \triangleq \begin{bmatrix}
\mathbf{M}^{(a)}[0] \\
\mathbf{M}^{(a)}[1] \\
\vdots \\
\mathbf{M}^{(a)}[L-1]
\end{bmatrix}_{K \times KL}
$$

so that the total average transmit power from antenna $a$ required to send a single symbol vector is

$$
P^{(a)} = E_b \left[ ||\mathbf{x}^{(a)}||^2 \right]
$$

$$
= \text{tr} \left( \mathbf{M}^{(a)H} \mathbf{\tilde{S}} \mathbf{\tilde{S}} \mathbf{M}^{(a)} \right).
$$

The noise free received signal at user 1’s mobile unit due to the signal transmitted from antenna $a$ is given by

$$
r_1^{(a)} = \sum_{l=0}^{L-1} \sum_{i=0}^{L-1} h_1^{(a)}[i] \mathbf{S}[i-l] \mathbf{M}^{(a)}[l] \mathbf{b}
$$

$$
= \sum_{l=0}^{L-1} h_1^{(a)}[l] \mathbf{SM}^{(a)}[l] \mathbf{b} + \sum_{l=0}^{L-1} \sum_{i=0}^{L-1} h_1^{(a)}[i] \mathbf{S}[i-l] \mathbf{M}^{(a)}[l] \mathbf{b}
$$

$$
\text{inter-chip interference}
$$
where $S[p] \in \mathbb{C}^{N \times K}$ is a matrix of $p$-shifted spreading codes with zero padding. If $p = 1$, for example, the $k$-th column of $S[p]$ is $[0 [s_k]_1 [s_k]_2 \cdots [s_k]_{N-1}]^T$. For negative $p$, the spreading codes are shifted up and zeros are inserted at the bottom of the matrix.

As before, we assume matched filter detection so that the decision statistic for user 1 due to the signal transmitted from antenna $a$ is

$$d_1^{(a)} = s_1^H r_1^{(a)}.$$  \hfill (197)

Stacking decision statistics from each user, we have

$$d^{(a)} \triangleq \left[ d_1^{(a)} d_2^{(a)} \ldots d_K^{(a)} \right]^T$$ \hfill (198)

$$= \sum_{i=0}^{L-1} \sum_{l=0}^{L-1} H^{(a)}[i] R[i - l] M^{(a)}[l] b$$ \hfill (199)

$$= \mathcal{H}^{(a)} \mathcal{R} \mathcal{M}^{(a)} b$$ \hfill (200)

where

$$H^{(a)}[i] \triangleq \text{diag}(h_1^{(a)}[i], h_2^{(a)}[i], \ldots, h_K^{(a)}[i])$$ \hfill (201)

$$R[i - l] \triangleq S^H S[i - l]$$ \hfill (202)

$$\mathcal{H}^{(a)} \triangleq \begin{bmatrix} H^{(a)}[0] & H^{(a)}[1] & \cdots & H^{(a)}[L - 1] \end{bmatrix}_{K \times KL}$$ \hfill (203)

and

$$\mathcal{R} \triangleq \begin{bmatrix} R[0] & R[-1] & \cdots & R[-(L - 1)] \\ R[1] & \cdots & \cdots & R[-1] \\ \vdots & \cdots & \cdots & \vdots \\ R[L - 1] & R[1] & \cdots & R[0] \end{bmatrix}_{KL \times KL}$$ \hfill (204)

The cost function for determining the optimal precoder supermatrix $\mathcal{M}^{(a)}$ is formed as

$$J_{mp}^{(a)} = E \left[ \left\| \mathcal{D}^{(a)} b - \mathcal{H}^{(a)} \mathcal{R} \mathcal{M}^{(a)} b - n \right\|^2 \right]$$ \hfill (205)

where $\mathcal{D}^{(a)}$ is a diagonal power loading matrix whose elements are given by

$$\left[ \mathcal{D}^{(a)} \right]_{i,i} = \frac{\sum_{l=0}^{L-1} \left| h_i^{(a)}[l] \right|^2}{\left[ \sum_{a=1}^{2} \sum_{j=0}^{L-1} \left| h_i^{(a)}[j] \right|^2 \right]^{\frac{1}{2}}}.$$ \hfill (206)

The optimal precoding supermatrix for antenna $a$, assuming perfect or partial channel information, can be found using straightforward modifications of Propositions 1 and 2.
Proposition 4 The precoding supermatrix $\mathcal{M}^{(a)}$ that minimizes $J_{\text{mp}}^{(a)}$ satisfies

$$\mathcal{M}^{(a)} = \left[ \mathcal{H}^{(a)} \mathcal{R} \right]^\dagger \mathcal{D}^{(a)}$$ (207)

for $a \in \{1, 2\}$.

The proof appears in [38]. The individual precoding matrices $\{\mathcal{M}^{(a)[l]}\}_{l=0}^{L-1}$ can be found from the optimal precoding supermatrix via

$$\mathcal{M}^{(a)[l]} = \mathcal{M}_K^{(a)} \triangleq \mathcal{M}^{(a)}_{Kl+1:Kl+K:c}$$ (208)

5.4 Performance Analyses

5.4.1 Performance of Transmitter Precoding with Blind Channel Estimation

Previously in this chapter, we developed developed analytical tools to investigate the performance of blind and group-blind linear MMSE multiuser detection. In this section, we adapt these tools to the analysis of single-antenna transmitter precoding with blind channel estimation. In particular, we will derive signal-to-interference-plus-noise (SINR) and BER expressions that take residual multiple-access interference and channel estimation error into account.

Notice that the estimate $\hat{M}$ given by (166) is not a consistent estimate of $M$ because of the unknown phase and scaling factors. However, there is a diagonal matrix $\Phi$ so that $\hat{M}\Phi^{-1}$ is a consistent estimate. The matrix $\Phi$ is of the form

$$\Phi \triangleq \text{diag}(\|f_1\|e^{j\phi_1}, \|f_2\|e^{j\phi_2}, \ldots, \|f_K\|e^{j\phi_K})$$ (209)

where $\phi_k, k = 1, \ldots, K$ are phase factors that depend on how the estimation is implemented. With this in mind, we state the following result, which is proved in the appendix.

Theorem 3 Let $\hat{M}$ be given by (166), and let $b$ be i.i.d. QPSK symbols independent of $M$. Then

$$\sqrt{M} \left( [\hat{M}\Phi^{-1} - M]b \right) \rightarrow \mathcal{N}_c(0, C_m) \text{ in distribution as } M \rightarrow \infty$$

with

$$C_m = \mathcal{H}^{-1} \mathcal{D} \mathcal{H}^{-H}$$ (210)

where the diagonal elements of $\mathcal{D}$ are given by

$$[\mathcal{D}]_{i,i} = \beta_i \sum_{k=1}^{K} \sum_{l=1}^{K} [\mathcal{H}^{-1} \mathcal{H}^{-H}]_{k,l} s_i^H s_k Q_i^H S_i s_i$$ (211)

and

$$\beta_i \triangleq \sigma^2 h_i^H U_s A_s (A_s - \eta I_K)^{-2} U_s^H h_i,$$ (212)

while the off-diagonal elements can be ignored with good accuracy. Here $Q_i^H$ denotes the Moore-Penrose generalized inverse [17] of the matrix $Q_i \triangleq S_i^H U_n U_n^H S_i$. 

46
The proof appears in [35].

The SINR at the output of the matched filter for User 1 is given by [19, 18]

\[
\text{SINR} \triangleq \frac{|E\{ s^H r_1[i] | b_1[i] \}|^2}{E\{ \text{Var}( s^H r_1[i] | b_1[i] ) \}}.
\] (213)

Now suppose that the phase and amplitude factors in \( \Phi \) have been determined. Write the estimated matrix, \( \hat{M} \), as \( \hat{M} \Phi^{-1} = M + \Delta M \), where \( \Delta M \) is the estimation error. Dropping the time index for clarity, the received signal can then be written as

\[
r_1 = s_1^H r_1
\] (214)

\[
= (s_1^H) \hat{M} \Phi^{-1} b + s_1^H n_1
\] (215)

\[
= (s_1^H H_1) Mb + (s_1^H H_1) \Delta Mb + s_1^H n_1
\] (216)

\[
= (s_1^H H_1) [\hat{M}]_{:,1} b_1 + (s_1^H H_1) [\hat{M}]_{:,2:K} [b]_{2:K} + (s_1^H H_1) \Delta Mb + s_1^H n_1
\] (217)

where the notation \([\hat{M}]_{:,2:K}\) indicates the matrix composed of columns 2 through \( K \) of the matrix \( \hat{M} \). According to Theorem 3, for large \( M \) the third term in (217) is also Gaussian distributed (independent of the other terms) with variance

\[
v_i^2 = \frac{1}{M} (s_1^H H_1) C_m H_1^H s_1.
\] (218)

Since \( M \) represents an MMSE detector, we can also make the approximate assumption that the multiple-access interference is Gaussian distributed [30]. We can therefore calculate the BER via a single \( Q \)-function as

\[
P_b(e) \approx Q(\sqrt{\text{SINR}})
\] (219)

with

\[
\text{SINR} = \frac{\sum_{k=2}^{K} |(s_1^H H_1)[\hat{M}]_{:,k}|^2 + \sigma^2 \|s_1\|^2 + \frac{1}{M} (s_1^H H_1) C_m H_1^H s_1}{(s_1^H H_1)[\hat{M}]_{:,1}^2}.
\] (220)

Notice that the first term in the denominator of the SINR expression is due to residual multiple-access interference. The second term is the ambient noise, and the third term is due to the channel estimation error.

A comparison of these results with simulation appears in [35], where good agreement is found.

5.4.2 Performance and Achievable Diversity for Multi-antenna Precoding

We have seen that with perfect channel knowledge at the transmitter and orthogonal spreading codes, we can achieve full transmit diversity with precoding. We will see here that non-zero spreading code crosscorrelations lead to an SNR loss, but full diversity is still achievable.

The General Case

Stacking decision statistics from all users obtained using the optimal \( M^{(1)}, M^{(2)} \), we define
the composite received signal as

\[
\mathbf{r} \triangleq \begin{bmatrix}
d_1 \\
d_2 \\
\vdots \\
d_K
\end{bmatrix}
\]

\[
= \alpha [\mathbf{H}^{(1)} \mathbf{R} \mathbf{M}^{(1)} + \mathbf{H}^{(2)} \mathbf{R} \mathbf{M}^{(2)}] \mathbf{b} + \sigma \mathbf{n}
\]

(222)

\[
= \alpha \mathbf{E} \mathbf{b} + \sigma \mathbf{n}
\]

(223)

where

\[
\mathbf{E} \triangleq \text{diag} \left( \left[ \mathbf{h}^{(1)}_1, \mathbf{h}^{(1)}_2, \ldots, \mathbf{h}^{(1)}_K, \mathbf{h}^{(2)}_1, \mathbf{h}^{(2)}_2, \ldots, \mathbf{h}^{(2)}_K \right] \right).
\]

(224)

Because \( \mathbf{E} \) is diagonal, multiple access interference is completely eliminated. Furthermore, we have seen in Section 5.2.2 that with orthogonal codes, we can set \( \alpha = 1 \) (i.e., no transmit power adjustment is necessary) and achieve full transmit diversity. In general, however, we must set \( \alpha \leq 1 \) to constrain average transmit power. For our purposes, average transmit power normalization requires

\[
\alpha^2 \mathbf{E} \mathbf{b} \left[ \| \mathbf{S} \mathbf{M}^{(1)} \mathbf{b} \|^2 \right] + \alpha^2 \mathbf{E} \mathbf{b} \left[ \| \mathbf{S} \mathbf{M}^{(2)} \mathbf{b} \|^2 \right] = K
\]

(225)

for every channel realization. That is, the sum of the average transmit power from the two antennas is equal to the number of users. Dropping the antenna superscripts for notational convenience, we have

\[
\mathbf{E} \left[ \| \mathbf{S} \mathbf{M} \mathbf{b} \|^2 \right] = \text{tr} \left( \mathbf{M}^H \mathbf{R} \mathbf{M} \right) = \text{tr} \left( \left( \mathbf{H}^{-1} \mathbf{D} \right)^H \mathbf{R}^{-1} \left( \mathbf{H}^{-1} \mathbf{D} \right) \right).
\]

(226)

(227)

For transmit antenna \( a \in \{1, 2\} \), the diagonal structures of \( \mathbf{D} \) and \( \mathbf{H} \) yield

\[
\text{tr} \left( \mathbf{M}^a \mathbf{H} \mathbf{R} \mathbf{M}^a \right) = \sum_{i=1}^{K} \left[ \mathbf{R}^{-1} \right]_{i,i} \frac{\left| \mathbf{h}^{(a)}_i \right|^2}{\left| \mathbf{h}^{(1)}_i \right|^2 + \left| \mathbf{h}^{(2)}_i \right|^2}.
\]

(228)

Summing the average transmit power contributions from each antenna, we have

\[
\text{tr} \left( \mathbf{M}^{(1)} \mathbf{H} \mathbf{R} \mathbf{M}^{(1)} \right) + \text{tr} \left( \mathbf{M}^{(2)} \mathbf{H} \mathbf{R} \mathbf{M}^{(2)} \right) = \sum_{i=1}^{K} \left[ \mathbf{R}^{-1} \right]_{i,i}
\]

(229)

\[
= \sum_{i=1}^{K} \frac{1}{\lambda_i}
\]

(230)

where \( \{\lambda_i\}_{i=1}^{K} \) are the eigenvalues of \( \mathbf{R} \). Therefore, by (225), the power scaling factor \( \alpha \) must satisfy

\[
\alpha = \sqrt{\frac{1}{\frac{1}{K} \sum_{i=1}^{K} \frac{1}{\lambda_i}}}.\]

(231)
Notice that $\alpha^2$ is simply the inverse of the average of the diagonal elements of $R^{-1}$. It is interesting to relate this result to the performance of receiver-based decorrelating multiuser detection (MUD), in which the performance of user $k$ is dependent upon the inverse of $[R^{-1}]_{k,k}$, but is not dependent upon the other diagonal elements of $R^{-1}$. In this sense, we can think of the performance of precoding, which is the same for every user, as the performance of decorrelating MUD “averaged” over every user. This interpretation is supported by the simulation results reported in [47].

Assuming all channel gains are independent and have the same statistics, the average bit-error-probability (BEP) of every user will be the same and is given by [34, p. 825]

$$
\overline{Pr}(\epsilon) = E \left[ Q \left( \frac{\alpha}{\sigma} \sqrt{ |h_1^{(1)}|^2 + |h_1^{(2)}|^2} \right) \right] = \frac{1}{4} (\mu^3 - 3\mu + 2)
$$

where

$$
\mu \triangleq \sqrt{\frac{\gamma}{1 + \gamma}}, \quad \gamma \triangleq \frac{\xi_h \alpha^2}{2\sigma^2}, \quad \xi_h \triangleq E \left[ |h_i^{(a)}|^2 \right], \quad a = 1, 2; i = 1, 2, \ldots, K.
$$

This performance is the same as two-branch maximum ratio combining with an SNR penalty of $10 \log_{10} \alpha^2$ dB. Hence diversity, defined here as the slope of the BEP curve, is unaffected by signature waveform crosscorrelations, but we do suffer SNR loss.

**Equicorrelated Spreading Codes**

As an important special case, we consider the scenario in which the normalized spreading code crosscorrelations satisfy

$$
s_k^H s_l = \begin{cases} 1 & k = l \\ \rho & k \neq l \end{cases}
$$

for some $\rho \in [0, 1)$. It is easy to show using the matrix inversion lemma [17, p. 19] that

$$
R^{-1} = \frac{1}{1 - \rho} I_K - \frac{\rho}{(1 - K) \rho^2 + (K - 2) \rho + 1} 1_{K,K},
$$

which yields

$$
\text{tr} \left( M^{(1)}^H R M^{(1)} \right) + \text{tr} \left( M^{(2)}^H R M^{(2)} \right) = K \left( \frac{1}{1 - \rho} - \tilde{\rho} \right)
$$

and

$$
\alpha(\rho, K) = \left[ \frac{1}{1 - \rho} - \frac{\rho}{(1 - K) \rho^2 + (K - 2) \rho + 1} \right]^{-\frac{1}{2}}.
$$

Clearly,

$$
\lim_{K \to \infty} \alpha(\rho, K) = \sqrt{1 - \rho}.
$$
In fact, $\alpha(\rho, K)$ tends to its limit rather quickly, as we see from Fig. 14, which plots $\alpha(\rho, K)$ as a function of $\rho$, for several values of $K$. Note that $\rho$ is, implicitly, a function of the number of users, $K$, and the processing gain, $N$. In order to increase $K$ while maintaining a constant $\rho$, the processing gain (and, hence, the bandwidth) will, in general, have to be increased as well.

For a moderate or large number of users, the performance is nearly equivalent to two-branch maximum ratio combining with a SNR penalty of $10 \log_{10}(1 - \rho)$ dB. Fig. 15 illustrates the BEP performance, calculated using (233), for 20 users and for various values of the crosscorrelation parameter $\rho$. The $10 \log_{10}(1 - \rho)$ dB SNR penalty is clearly visible.

6 Conclusion

In this chapter, we have treated the three main problems arising in wireless systems - dynamism, fading and interference - in a unified setting. We have examined several power approaches to these problems, including subspace methods, Markov chain Monte Carlo, the sequential EM algorithm, and transmitter-based techniques. We have seem that these methods can perform can perform quite well, thereby allowing for effective signal reception in multiuser fading environments.

References

Figure 15: The bit-error-probability for the equicorrelated spreading code case, averaged over all 20 users and their channel gains, versus $E_h/\sigma^2$ for precoding with two transmit antennas and one receiver antenna and for crosscorrelation values of $\rho = 0, 0.3, 0.7, 0.9$. The transmit energy per user per bit is 1.


