

Multiuser Receiver Design

Huaiyu Dai, Sudharman Jayaweera, H. Vincent Poor, Daryl Reynolds, Xiaodong Wang

1 Introduction

The preceding chapter considered the design of receivers for MIMO systems operating as single-user systems. Increasingly however, as noted in Chapters 2 and 4, wireless communication networks operate as shared-access systems in which multiple transmitters share the same radio resources. This is due largely to the ability of shared access systems to support flexible admission protocols, to take advantage of statistical multiplexing, and to support transmission in unlicensed spectrum. In this chapter we will extend the treatment of Chapter 5 to consider receiver structures for multiuser, and specifically, multiple-access MIMO systems. We will also generalize the channel model considered to include more general situations than the flat-fading channels considered in Chapter 5. To treat these problems, we will first describe a general model for multiuser MIMO signaling, and then discuss the structure of optimal receivers for this signal model. This model will generally include several sources of interference arising in MIMO wireless systems, including multiple-access interference caused by the sharing of radio resources noted above, inter-symbol interference caused by dispersive channels, and inter-antenna interference caused by the use of multiple transmit antennas. Algorithms for the mitigation of all of these types of interference can be derived in this common framework, leading to a general receiver structure for multiuser MIMO communications over frequency-selective channels. As we shall see, these basic algorithms will echo similar algorithms that have been described in Chapters 3 and 5. Since optimal receivers in this situation are often prohibitively complex, the bulk of the chapter will focus on useful lower complexity sub-optimal iterative and adaptive receiver structures that can achieve excellent performance in mitigating interference in such systems. This discussion is organized as follows.

Section 2 of this chapter will introduce a simple, yet useful, model for the signals received by the receiver in a MIMO system. This model is rich enough to capture the important behavior of most wireless communication channels, while being simple enough to allow for the straightforward motivation and understanding of the basic receiver elements arising in practical situations. This section also derives a canonical multiuser MIMO receiver structure, discusses several specific receivers that can be explained within this structure, and provides a digital receiver implementation that will be useful in the discussion of adaptive systems later in the chapter.

As noted above, complexity is a major issue in multiuser receiver design and implementation, and the remainder of this chapter addresses the problem of complexity reduction in multiuser MIMO systems. This complexity takes two forms: computational, or implementational, complexity; and informational complexity.

The first type of complexity refers to the amount of resources needed to implement a given receiver algorithm. Optimal MIMO multiuser receiver algorithms are typically prohibitively complex in this sense, and thus a major issue in this area is complexity reduction. Sections 3 and 4 of this chapter address the principal method for complexity reduction in practical multiuser receivers, namely the use of iterative algorithms in which tentative decisions are made and updated iteratively. There are a number of basic iterative techniques, involving different tradeoffs between complexity and performance, and depending on the type of system under consideration, and these are described in Section 3. In Section 4, we tackle the additional complexity that arises in receiving space-time coded transmissions, such as those described in Chapter 4, in multiuser MIMO systems. Here, iterative algorithms similar to those discussed in Chapter 5 provide the answer to finding algorithms that can exploit the space-time coded structure with only moderate increases in complexity.

The second type of complexity refers to the amount of knowledge that a given receiver needs to have about the structure of received signals in order to effect signal reception. Although, as we will see shortly, optimal MIMO multiuser reception requires knowledge of the waveforms being transmitted by all users sharing the channel and the structure of the physical channel intervening between transmitters and receiver, this type of knowledge is rarely available in practical wireless multiuser systems. Thus, it is

necessary to consider adaptive receiver algorithms that can operate without such knowledge, or with only limited such knowledge. Such algorithms are the topic of Section 5 of the chapter, in which the structure of adaptive MIMO multiuser receivers is reviewed.

The chapter will conclude in Sections 6 and 7 with a summary and pointers to additional reading of interest in this general area.

2 Multiple-Access MIMO Systems

As noted above, this section will provide a general treatment of the multiuser MIMO receiver design problem. Here we will focus on modeling and on the structure of optimal receivers. In doing so, we will expose the principal issues underlying the reception of signals in multiuser MIMO systems, and also will set the stage for more practical algorithms developed in succeeding sections.

2.1 Signal and Channel Models

In order to discuss multiuser MIMO receiver structures, it is useful to first specify a general model for the signal received by a MIMO receiver in a multiuser environment. (See Fig. 1.) In doing so, we will build on the signaling model developed in Chapter 1, and in particular our model is an abstraction of the physical channel described there that is especially useful for the purposes of this chapter. Specifically, a useful received signal model for a multiuser MIMO system having K active users, M_T transmit antennas and M_R receive antennas, and transmitting over a frame of B symbol periods, can be written as follows:

$$r_p(t) = \sum_{k=1}^K \sum_{m=1}^{M_T} \sum_{i=0}^{B-1} b_{k,m}[i] g_{k,m,p}(t - iT_s) + n_p(t), \quad p = 1, \dots, M_R, \quad (1)$$

where the various quantities are as follows:

- $r_p(\cdot)$ = the signal received at the output of the p^{th} receive antenna,
- $b_{k,m}[i]$ = the symbol transmitted by user k from its m^{th} antenna in the i^{th} symbol interval,
- $g_{k,m,p}(\cdot)$ = the waveform on which symbols from the m^{th} antenna of user k arrives at the output of the p^{th} receive antenna,

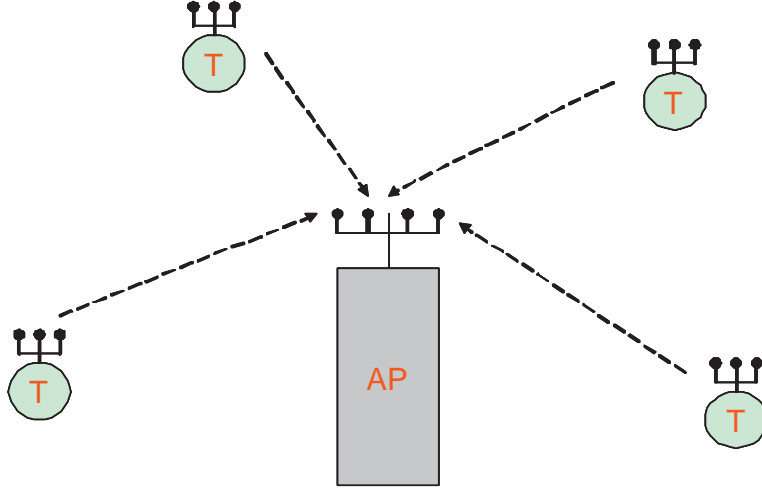


Figure 1: A Multiuser MIMO System.

- T_s = the symbol period, and
- $n_p(\cdot)$ = ambient noise at the p^{th} receive antenna.

Each of the waveforms $g_{k,m,p}(\cdot)$ can be modeled as

$$g_{k,m,p}(t) = \int_{-\infty}^{\infty} s_{k,m}(u) f_{k,m,p}(t-u) du, \quad (2)$$

where

- $s_{k,m}(\cdot)$ = the signaling waveform used by user k on its m^{th} antenna, and
- $f_{k,m,p}(\cdot)$ = the impulse response of the channel between the m^{th} transmit antenna of user k and the p^{th} receive antenna output.

Thus, we are assuming linear modulation and a linear channel model, both of which are reasonable assumptions for wireless systems. Note that, since $g_{k,m,p}(\cdot)$ does not depend on the symbol index i in this model, we are implicitly assuming here that the channel is stable (and time-invariant) over the transmission frame (which is BT_s seconds long) and that the transmitters use the same signaling waveforms in each symbol period. The first of these assumptions is valid for the coherence times and signaling parameters arising in most systems of interest, while the latter is often violated, particularly in cellular systems.

However, with the exception of the adaptive methods of Section 5, this time variation is not difficult to incorporate into any of the results described in this chapter, and is omitted here for the sake of notational simplicity. (See, e.g., [46]).

In order to minimize the number of parameters in this model, we will assume that the signaling waveforms are normalized to have unit total energy, i.e.,

$$\int_{-\infty}^{\infty} [s_{k,m}(t)]^2 dt = 1, \quad k = 1, \dots, K, \quad m = 1, \dots, M_T. \quad (3)$$

In reality, the actual transmitted waveforms will carry differing and non-unit energies, reflecting the transmitted powers of the various users' terminals. However, from the vantage point of *receiver* design, the critical scale parameter is the received power of a user, which will depend on the user's transmitter power and the gain of the intervening channel. Thus, it is convenient to lump all scaling of the signals into the channel impulse response $f_{k,m,p}(\cdot)$, and to simply assume normalized waveforms (3) at the transmitter. Again, from the receiver's point of view, it is impossible anyway to separate the effects of channel gain and transmit power on the received power. Also for convenience, we will assume that the transmitted waveforms have duration of only a single symbol interval; i.e.,

$$s_{k,m}(t) = 0, \quad t \notin [0, T_s]. \quad (4)$$

As with the normalization constraint (3), this assumption does not remove any generality since received waveforms that extend beyond a single symbol interval can be modeled via dispersion in the channel response.

A typical and useful model for the channel response is as a discrete multipath model:

$$f_{k,m,p}(t) = \sum_{\ell=1}^L h_{k,m,p,\ell} \delta(t - \tau_{k,m,p,\ell}), \quad (5)$$

where $\delta(\cdot)$ denotes the Dirac delta function, and where $h_{k,m,p,\ell}$ and $\tau_{k,m,p,\ell} \geq 0$ denote the channel gain and propagation delay, respectively, of the ℓ^{th} path of the channel between the m^{th} transmit antenna of

user k and the output of the p^{th} receive antenna.¹ In this case, the waveforms $g_{k,m,p}(\cdot)$ are of the form

$$g_{k,m,p}(t) = \sum_{\ell=1}^L h_{k,m,p,\ell} s_{k,m}(t - \tau_{k,m,p,\ell}) . \quad (6)$$

That is, in this model, the waveform received at a given receive antenna p from a given transmit antenna m of a particular user k is the superposition of L scaled and delayed copies of the waveform $s_{k,m}(\cdot)$ transmitted from that antenna. Except where noted otherwise, we will assume this particular model for the channel response in the sequel.

The signaling waveforms $s_{k,m}(\cdot)$ can take many forms. Although these waveforms can be thought of as being generic in our discussion, a quintessential example is the case in which the transmitted signals are in direct-sequence code-division multiple-access (DS/CDMA) format. This is a very widely used signaling format in wireless systems (used notably in both major 3G cellular standards), and is the example used in the simulations discussed in succeeding sections of this chapter. In the notation of this section, this format can be described as follows.

DS/CDMA Signaling: In the DS/CDMA format, the signaling waveforms used by all transmitters are in the form of spread-spectrum signals; i.e., the waveforms $\{s_{k,m}(\cdot)\}$ of (1) are of the form

$$s_{k,m}(t) = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} c_{k,m}^{(j)} \psi(t - (j-1)T_c) , \quad 0 \leq t \leq T_s , \quad (7)$$

where N is the spreading gain of the system, $c_{k,m}^{(0)}, c_{k,m}^{(1)}, \dots, c_{k,m}^{(N-1)}$ is the spreading code (or signature sequence) associated with the m^{th} transmit antenna of user k , $T_c = T_s/N$ is the chip interval, and $\psi(\cdot)$ is a chip waveform having unit-energy and approximate duration T_c . (For a general discussion of spread-spectrum signaling, see, e.g., [48].) In studying this format, the chip waveform $\psi(\cdot)$ is often modeled as a unit-energy pulse of duration T_c ; i.e.,

$$\psi(t) = \begin{cases} \frac{1}{\sqrt{T_c}} , & t \in [0, T_c] \\ 0 , & \text{otherwise} \end{cases} . \quad (8)$$

¹For simplicity, we lump the effects of the radio channel itself and the antenna response into the same term $h_{k,m,p,\ell}$. Often these two terms can be separated (see, e.g., [46]). However, no generality is lost in lumping these effects together for the purposes of analysis and exposition.

Again, most of the results of this chapter apply to general signaling waveforms, and it is not necessary to particularize to this specific format except where noted. It should also be mentioned that these signaling waveforms, the symbols, the noise, and the channel responses may be taken to be complex (rather than real as is tacitly assumed here). We will not need this generality here until Section 5, and so we will defer discussions of needed modifications (which are minor) until then. A complex version of the above model can be found in [46], which allows for two-dimensional signaling constellations, such as QPSK and QAM, to fit within this model.

As an additional assumption, we assume that the ambient noise processes $n_p(\cdot)$, $p = 1, \dots, M_R$, are mutually independent white Gaussian processes with common spectral height σ^2 . We also assume that the transmitted symbols take values in a finite alphabet \mathcal{A} containing $|\mathcal{A}|$ elements. Beginning in Section 3, we will specialize this to the binary antipodal case $\mathcal{A} = \{-1, +1\}$. This is primarily for convenience, as most of the results in this chapter hold for more general signaling alphabets.

Finally, we note that M_T , B and L in the above model could vary from user to user, while L could also vary from antenna pair to antenna pair. However, again for simplicity, we will assume them to be constants, as the extensions of the discussions in the chapter to these non-constant cases are quite straightforward.

2.2 Canonical Receiver Structure

A basic MIMO multiuser receiver structure can be usefully decomposed into two parts: a front-end (or hardware) part, and a decision algorithm (or software) part. In practice, these pieces may not be completely distinct, as much of the front-end may be implemented in software; but for the purposes of exposition, it is a useful decomposition.

A canonical front-end for such a system can be derived based on the theory of statistical inference. In particular, it is of interest to examine the so-called *likelihood function* of the observations (1) given the collection of transmitted symbol: $\{b_{k,m}[i]\}_{k=1,\dots,K; m=1,\dots,M_T; i=0,\dots,B-1}$. Due to the assumption of white, Gaussian noise, the logarithm of this likelihood function is given (up to a scalar multiple) by the

Cameron-Martin formula [29] to be:

$$\sum_{k=1}^K \sum_{m=1}^{M_T} \sum_{i=0}^{B-1} b_{k,m}[i] z_{k,m}[i] - \frac{1}{2} \sum_{k,k'=1}^K \sum_{m,m'=1}^{M_T} \sum_{i,i'=0}^{B-1} b_{k,m}[i] b_{k',m'}[i'] C(k, m, i; k', m', i'), \quad (9)$$

where, for $k = 1, \dots, K$, $m = 1, \dots, M_T$, and $i = 0, \dots, B - 1$,

$$z_{k,m}[i] = \sum_{\ell=1}^L \sum_{p=1}^P h_{k,m,p,\ell} \int_{-\infty}^{\infty} r_p(t) s_{k,m}(t - \tau_{k,m,p,\ell} - iT_s) dt, \quad (10)$$

and for $k, k' = 1, \dots, K$, $m, m' = 1, \dots, M_T$, and $i, i' = 0, \dots, B - 1$,

$$C(k, m, i; k', m', i') = \sum_{p=1}^P \sum_{\ell,\ell'=1}^L h_{k,m,p,\ell} h_{k',m',p,\ell'} \int_{-\infty}^{\infty} s_{k,m}(t - \tau_{k,m,p,\ell} - iT_s) s_{k',m'}(t - \tau_{k',m',p,\ell'} - i'T_s) dt. \quad (11)$$

Although the expression (9) may seem somewhat complicated, the key thing to note about it is that the antenna outputs, $r_1(t), r_2(t), \dots, r_P(t)$, enter into the likelihood function only through the collection of “observables” $\{z_{k,m}[i]\}_{k=1,\dots,K; m=1,\dots,M_T; i=0,\dots,B-1}$. This means that this collection of variables is a *sufficient statistic*[29] for making inferences about the corresponding set of transmitted symbols $\{b_{k,m}[i]\}_{k=1,\dots,K; m=1,\dots,M_T; i=0,\dots,B-1}$, which implies in turn that all attention can be restricted to this set of observables when designing and building systems or algorithms for demodulating and detecting the transmitted symbols.

Before turning to some types of algorithms that we might use for this purpose, it is worthwhile to examine the structure of this set of observables a bit more closely. In particular, it can be seen that (10) consists of three basic operations:

1. integration to obtain: $x_{k,m,p,\ell}[i] = \int_{-\infty}^{\infty} r_p(t) s_{k,m}(t - \tau_{k,m,p,\ell} - iT_s) dt$;
2. correlation to obtain: $y_{k,m,\ell}[i] = \sum_{p=1}^P h_{k,m,p,\ell} x_{k,m,p,\ell}[i]$; and
3. summation to obtain: $z_{k,m}[i] = \sum_{\ell=1}^L y_{k,m,\ell}[i]$.

The first operation is a *matched filtering* operation, so that we see that each received antenna output is filtered with a filter that is matched to the waveform received on each path from each transmit antenna in each symbol interval of each user. Thus, there are $K \times M_R \times B \times L \times M_T$ matched filter

outputs, which we can think of as being produced by a bank of linear filters, each of which is sampled at the end of each signaling interval; i.e., samples are taken at times iT_s for $i = 0, \dots, B - 1$.

The second operation, in which the matched-filter outputs $\{x_{k,m,p,\ell}[i]\}$ are correlated across the receive antenna array with the channel/antenna gains $\{h_{k,m,p,\ell}\}$, can be viewed as a form of *beamforming*, through which the spatial dimension afforded by the receive array is exploited. Since the terms $h_{k,m,p,\ell}$ also incorporate channel gains, this is not strictly a simple beamforming operation in general, but it has a similar effect of coherently collapsing the spatial dimension of the array. Note that, after beamforming, there are $K \times B \times L \times M_T$ observables.

Finally, the third operation, in which the beamformer outputs $\{y_{k,m,\ell}[i]\}$ are added, is a *multipath combiner*, or Rake operation through which the spatial dimension introduced by the multipath channel is exploited. Typically a Rake receiver also includes a correlation with the channel multipath coefficients. This is being done here as part of the beamforming operation. So, the combination of the second and third operation is equivalent to beamforming followed by Rake combining, and this combination might be decomposed in other ways in practice. After this third operation, there are $K \times M_T \times B$ observables, one for each symbol in the frame of each user.

These three operations constitute the (hardware) front-end of a canonical multiuser receiver, as illustrated in Fig. 2. This front-end is sometimes known as a *space-time matched filter*. Note that, although this structure may seem complicated, it is essentially composed of standard communication-system components: matched filters, beamformers, Rake receivers.

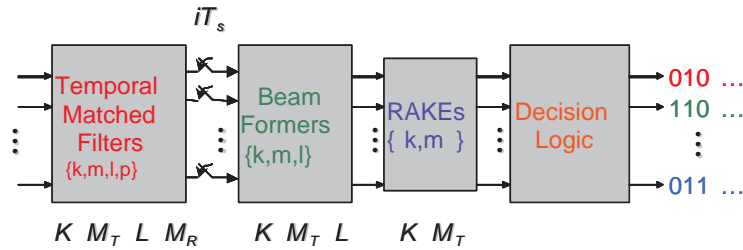


Figure 2: A Canonical MIMO Multiuser Receiver Structure.

It is noteworthy that this formalism and general front-end structure encompasses three standard interference-mitigation problems in communications. To discuss this point, it is useful to define the

parameter

$$\Delta = \left\lceil \frac{\max_{k,m,p,\ell} \{\tau_{k,m,p,\ell}\}}{T_s} \right\rceil, \quad (12)$$

where $\lceil x \rceil$ denotes the smallest integer not less than x . Δ is the maximum delay spread of the wireless channels (5) in units of symbol intervals, and is thus the maximum extent to which symbols of a given user interfere with one another. Returning to the general receiver structure, the case in which $K = M_T = 1$ and $\Delta > 1$ is the channel equalization problem studied notably in the 1970s; the case $M_T = \Delta = 1$ and $K > 1$ is the traditional multiuser detection problem, studied notably in the 1980s; and finally the case in which $K = \Delta = 1$ and $M_T > 1$ is the standard MIMO communications problem, exemplified by the BLAST architecture studied notably in the 1990s. Combinations of these problems and refinements on them have been mainstays of research and development in digital communications throughout the past few decades and continuing to the present day. The applicability of the results in this chapter to these various problems, both individually and jointly, is worth keeping in mind in the subsequent discussions. Thus, the receiver architectures described herein can be applied other than in the multiuser MIMO communications setting, and many of them generalize solutions to the more particular cases noted above.

2.3 Basic MUD Algorithms

As illustrated in Fig. 2, the KM_TB outputs of the canonical multiuser front-end are operated upon by a decision algorithm whose purpose is to infer the values of the KM_TB transmitted symbols $\{b_{k,m}[i]\}$. This decision algorithm can take many forms, ranging through the full toolbox of statistical signal processing: optimal algorithms based on maximum-likelihood or maximum *a posteriori* probability criteria, linear algorithms, iterative algorithms, and adaptive algorithms. Each of these techniques will be discussed briefly in the following paragraphs, and counterparts to these algorithms are discussed in Chapters 3 and 5. However, before discussing these types of algorithms, it is useful to first examine the relationship between the observables $\{z_{k,m}[i]\}$ and the corresponding symbols $\{b_{k,m}[i]\}$ to be inferred. To do so, it is convenient to collect the symbols into a KM_TB -long column vector \mathbf{b} by sorting the symbols $\{b_{k,m}[i]\}$

first by symbol number, then by user number, and finally by antenna number. That is,

$$\mathbf{b} = \begin{pmatrix} \mathbf{b}[0] \\ \mathbf{b}[1] \\ \vdots \\ \mathbf{b}[N-1] \end{pmatrix} \quad (13)$$

where

$$\mathbf{b}[i] = \begin{pmatrix} \mathbf{b}_1[i] \\ \mathbf{b}_2[i] \\ \vdots \\ \mathbf{b}_K[i] \end{pmatrix} \quad (14)$$

with

$$\mathbf{b}_k[i] = \begin{pmatrix} \mathbf{b}_{k,1}[i] \\ \mathbf{b}_{k,2}[i] \\ \vdots \\ \mathbf{b}_{k,M_T}[i] \end{pmatrix}. \quad (15)$$

Similarly, we can denote by \mathbf{z} the set of observations $\{z_{k,m}[i]\}$ collected into a KM_TB -long column vector indexed conformally with \mathbf{b} . We can also define a $KM_TB \times KM_TB$ cross-correlation matrix \mathbf{R} whose (n, n') -th element is given by the cross-correlation $C(k, m, i; k', m', i')$ from (11) where the indices are determined by matching with the corresponding elements of \mathbf{b} (or, equivalently, \mathbf{z}); i.e., $b_n = b_{k,m}[i]$ and $b_{n'} = b_{k',m'}[i']$ with $n = [iK + (k-1)]M_T + m$ and $n' = [i'K + (k'-1)]M_T + m'$.

With these definitions, the observables and transmitted symbols can be related to one another through the relationship

$$\mathbf{z} = \mathbf{R} \mathbf{b} + \mathbf{n}, \quad (16)$$

where \mathbf{n} denotes a KM_TB -long noise vector having the $\mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{R})$ distribution. (Here, $\mathbf{0}$ denotes a KM_TB -long vector having all components equal to zero.)

As a simple example, we can consider the flat-fading, synchronous case, in which all signals arrive at the receive array with the same symbol timing. This corresponds to the discrete multipath model of (5) with $L = 1$ and $\tau_{k,m,p,\ell} \equiv 0$.

$$f_{k,m,p}(t) = h_{k,m,p,1} \delta(t) . \quad (17)$$

In this case, the matrix \mathbf{R} is a block-diagonal matrix having B identical blocks along its diagonal, each of dimension $K M_T \times K M_T$. These square sub-matrices contain the cross-correlations between the signals received from the different antennas of the different users. So, for example, in this case, the first block is given by

$$\mathbf{R}_{n,n'} = \int_{-\infty}^{\infty} s_{k,m}(t) s_{k',m'}(t) dt \times \sum_{p=1}^P h_{k,m,p,1} h_{k',m',p,1} , \quad n, n' = 1, 2, \dots, K M_T , \quad (18)$$

where the indices n and n' correspond, respectively, to antenna m of user k and antenna m' of user k' , both in the zeroth symbol interval. This block is then repeated B times along the diagonal of \mathbf{R} . This example is further illuminated by considering the single-receive-antenna case ($M_R = 1$), in which this first diagonal block simplifies to

$$\mathbf{R}_{n,n'} = \int_{-\infty}^{\infty} s_{k,m}(t) s_{k',m'}(t) dt A_{k,m} A_{k',m'} , \quad n, n' = 1, 2, \dots, K M_T , \quad (19)$$

with $A_{k,m} = h_{k,m,1,1}$ for $k, m = 1, \dots, K, M_T$, $n = (k-1)M_T + m$, and $n' = (k'-1)M_T + m'$. This block is thus of the form:

$$\mathbf{A} \bar{\mathbf{R}} \mathbf{A} \quad (20)$$

where \mathbf{A} is a diagonal matrix having the received amplitudes $A_{1,1}, \dots, A_{1,M_T}, A_{2,1}, \dots, A_{2,M_T}, \dots, A_{K,1}, \dots, A_{K,M_T}$ on its diagonal, and where $\bar{\mathbf{R}}$ is the *normalized* cross-correlation matrix of the signaling multiplex:

$$\bar{\mathbf{R}}_{n,n'} = \int_{-\infty}^{\infty} s_{k,m}(t) s_{k',m'}(t) dt , \quad n, n' = 1, 2, \dots, K M_T . \quad (21)$$

For example, in the DS/CDMA case of (7)-(8), this normalized cross-matrix is given by

$$\bar{\mathbf{R}}_{n,n'} = \frac{1}{N} \sum_{j=0}^{N-1} c_{k,m}^{(j)} c_{k',m'}^{(j)} , \quad n, n' = 1, 2, \dots, K M_T ; \quad (22)$$

that is, the normalized cross-correlation matrix is determined by the cross-correlations of the spreading sequences used by the system. The specific structure of this matrix depends on how the spreading sequences are allocated to the various users' antennas. In some systems, all antennas of the same user

use the same spreading code, while in others, different spreading codes are used for all antennas. As an example, if the spreading codes are so-called m -sequences (see, e.g., [48]), then $\overline{\mathbf{R}}_{n,n'} = 1$ for antennas using identical spreading codes and $\overline{\mathbf{R}}_{n,n'} = -1/N$, for antennas (and users) using different spreading codes.

In the general case in which the channel is not flat or the users are not synchronous, the block diagonal form of this example becomes a block Toeplitz form, as will be discussed in Section 3. From (16) we see that the basic relationship between \mathbf{z} and \mathbf{b} is that of a noisy linear model, and so the basic problem to be solved by the decision algorithm in Fig. 2 is that of fitting such a model. At first glance, this appears to be a rather straightforward problem, as the fitting of linear models is a classical problem in statistics. However, the difficulty in this problem arises because the vector \mathbf{b} to be chosen in this fit has discrete-valued elements (e.g., ± 1), and this significantly increases the complexity of fitting this model (16).

In general, the most powerful techniques for data detection are maximum-likelihood (ML) and maximum *a posteriori* probability (MAP) detection. ML detection makes inferences about the transmitted symbols in (1) by choosing those symbol values that maximize the log-likelihood function of (9). To get a sense of this task, it is useful to use the compact notation of (16) to rewrite the log-likelihood function (9) as

$$\mathbf{b}^T \mathbf{z} - \frac{1}{2} \mathbf{b}^T \mathbf{R} \mathbf{b} . \quad (23)$$

So, the ML symbol decision solve the optimization problem:

$$\max_{\mathbf{b} \in \mathcal{B}} \left[\mathbf{b}^T \mathbf{z} - \frac{1}{2} \mathbf{b}^T \mathbf{R} \mathbf{b} \right] , \quad (24)$$

where $\mathcal{B} = \mathcal{A}^{KM_TB}$. The optimization problem (24) is an integer quadratic program, which is known to be an NP-complete computational problem. Since the size of the search set \mathcal{B} is potentially enormous at $|\mathcal{A}|^{KM_TB}$, solving this problem appears to be impossible². However, for most practical wireless channels, the matrix \mathbf{R} has many zero elements which reduces the complexity of this problem significantly. In particular, assuming that the signaling waveforms $\{s_{k,m}(\cdot)\}$ are limited in duration to a single symbol interval, and given the finite multipath channel model (5), the matrix \mathbf{R} is a banded matrix, meaning that

²Typically, K might be dozens, M_T several and B hundreds.

all of its elements are zero except on a certain number of diagonals; i.e., $\mathbf{R}_{n,n'} = 0$ if $|n - n'| > KM_T\Delta$, where again Δ is the maximum delay spread of the wireless channels (5) in units of symbol intervals (12). This bandedness allows for a complexity reduction from the order of $|\mathcal{A}|^{KM_TB}$ needed to exhaustively search for the ML solution, to the order of $|\mathcal{A}|^{KM_T\Delta}$ (per symbol) to search via dynamic programming (see, e.g., [30]). Although in most wireless channels the maximum delay spread Δ is much less than the frame length B , even this reduced complexity is prohibitive for most applications as the exponent $KM_T\Delta$ could still be fairly large in a typical situation with dozens of users, a few antennas per user, and a few symbols of delay spread. The ML detector is sometimes referred to as the jointly optimal (JO) detector.

MAP detection is applicable to situations in which the receiver knows a prior probability distribution governing the values that the transmitted symbols may assume. In this situation, it is possible to consider the posterior probability distribution of a given symbol, conditioned on the observations, and to infer that value for each symbol that has maximum *a posteriori* probability (APP). That is, a given symbol, say b_n is detected as \hat{b}_n according to the following criterion:

$$\hat{b}_n = \arg \left\{ \max_{a \in \mathcal{A}} P(b_n = a | \mathbf{z}) \right\}. \quad (25)$$

Using Bayes' formula, we can write the APP as

$$P(b_n = a | \mathbf{z}) = \frac{\sum_{\mathbf{b} \in \mathcal{B}_{n,a}} \ell(\mathbf{z} | \mathbf{b}) w(\mathbf{b})}{\sum_{\mathbf{b} \in \mathcal{B}} \ell(\mathbf{z} | \mathbf{b}) w(\mathbf{b})}, \quad (26)$$

where $\mathcal{B}_{n,a}$ denotes the subset of \mathcal{B} in which the n^{th} coordinate is fixed at a , $w(\mathbf{b})$ is the prior probability of \mathbf{b} , and $\ell(\mathbf{y} | \mathbf{b})$ denotes the likelihood function of \mathbf{z} given \mathbf{b} :

$$\ell(\mathbf{z} | \mathbf{b}) = e^{(\mathbf{b}^T \mathbf{z} - \frac{1}{2} \mathbf{b}^T \mathbf{R} \mathbf{b}) / \sigma^2}. \quad (27)$$

Commonly, it is assumed that the symbol vector \mathbf{b} is uniformly distributed in its range \mathcal{B} ; i.e., that

$$w(\mathbf{b}) \equiv |\mathcal{A}|^{-KM_TB}. \quad (28)$$

This assumption is equivalent to assuming that all the symbols are independent and identically distributed (i.i.d.) from time to time, from user to user, and from antenna to antenna, and that each symbol is chosen

equiprobably among the elements of \mathcal{A} . This assumption is not always valid, as we will discuss below. However, when it is valid, the prior distribution drops out of the computation of the APP, and the MAP criterion becomes:

$$\hat{b}_n = \arg \left\{ \max_{a \in \mathcal{A}} \sum_{\mathbf{b} \in \mathcal{B}_{n,a}} \ell(\mathbf{z}|\mathbf{b}) \right\}. \quad (29)$$

(Note that the denominator in the APP (26) is irrelevant to the maximization since it does not depend on the value of any individual symbol.) The MAP detector is sometimes termed the individually optimal (IO) detector since it chooses each symbol decision according to a single-symbol criterion.

Like the ML detector, the computation of symbol decisions using (29) is generally prohibitively complex. In particular, we note that computation of the APP for each individual symbol value involves a summation over $|\mathcal{A}|^{KM_T B-1}$ values of the symbol vector. Also like ML detection, however, this complexity can be reduced via dynamic programming to the order of $|\mathcal{A}|^{KM_T \Delta}$ operations per symbol when the channel has delay spread of Δ symbol intervals [30, 39].

As we see from the above discussion, the basic complexity of ML (JO) or MAP (IO) data detection is quite complex, on the order of $|\mathcal{A}|^{KM_T \Delta}$ operations per detected symbol. So, the complexity grows with the number of users, the number of antennas, and the channel length. It is noteworthy that this issue is present even in the single-user ($K = 1$) case, the single-antenna case ($M_T = 1$), or in the flat-fading case ($\Delta = 1$). Only if all of these conditions is missing do we get a simple detector structure, which reduces in either the ML or MAP case to a simple quantization:

$$\hat{b}_n = Q(z_n), \quad (30)$$

where the quantizer $Q : \mathbb{R} \rightarrow \mathcal{A}$. For example, in the case of binary antipodal symbols ($\mathcal{A} = \{-1, +1\}$), we take Q to be the signum function:

$$Q(z) = \text{sgn}(z) = \begin{cases} -1 & z < 0 \\ +1 & z \geq 0 \end{cases}. \quad (31)$$

Since data detection must be performed on a relatively limited computing platform (i.e, a communications receiver) at essentially the rate of data transmission (i.e., tens to thousands of kilobits per second), it is of interest to consider alternatives to the optimal detectors described above. One family

of such detectors are the *linear* multiuser detectors, which seek to balance the simplicity of the simple detector (30) with the power of IO or JO detection. In linear detection, this is accomplished by first multiplying the sufficient statistic \mathbf{z} by a suitably chosen square matrix, and then quantizing the result:

$$\hat{b}_n = Q(v_n), \quad (32)$$

where

$$\mathbf{v} = \mathbf{M}\mathbf{z} \quad (33)$$

and where \mathbf{M} is a $KM_TB \times KM_TB$ matrix. This type of detector is illustrated in Fig. 3. Various types of detectors can be implemented through different choices of the matrix \mathbf{M} . Three key ones can be described as follows.

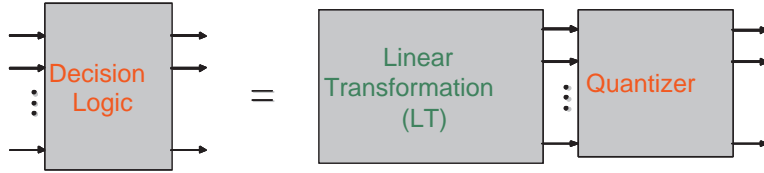


Figure 3: Linear Multiuser Algorithm.

Space-Time Matched Filter/Rake Receiver: The simplest example of a linear detector arises from choosing \mathbf{M} to be the $KM_TB \times KM_TB$ identity matrix \mathbf{I} , in which case the linear detector reduces to the simple detector of (30). This detector is a classical space-time matched filter receiver which is optimal in an additive white Gaussian noise (AWGN) channel. A flaw of this receiver is that it addresses only the ambient noise, while ignoring the cross-correlations between the signals affecting different symbols; i.e., it ignores the off-diagonal elements of \mathbf{R} .

Decorrelating (Zero-Forcing) Receiver: Noting from (16) that the mapping from transmitted symbols \mathbf{b} to the observables \mathbf{z} is in the form of a (square) linear transformation plus noise, a natural detection strategy is a zero-forcing detector that eliminates the interference embodied in the cross-correlation matrix \mathbf{R} . Assuming that \mathbf{R} is non-singular, this can be implemented as a linear detector with $\mathbf{M} = \mathbf{R}^{-1}$. The resulting detector is known as the *decorrelator*. The decorrelator thus quantizes the variables $\mathbf{v} = \mathbf{R}^{-1}\mathbf{z}$,

which are given by

$$\mathbf{v} = \mathbf{b} + \mathbf{R}^{-1}\mathbf{n}. \quad (34)$$

Note that, as expected, these transformed observables are free of (inter-user, inter-antenna, inter-symbol) interference. However, this receiver is the opposite extreme of the matched filter receiver, in that it is tantamount to ignoring the ambient noise to suppress the interference. Using standard properties of the multivariate Gaussian distribution, the noise terms in (34) is distributed according to

$$\mathbf{R}^{-1}\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{R}^{-1}) . \quad (35)$$

Depending on the structure of \mathbf{R} the inverse \mathbf{R}^{-1} can have very large diagonal values, leading to noise enhancement and consequent high error rate. (This problem is well-known in the context of equalization [33].) The assumption that \mathbf{R} be invertible is not overly restrictive in general, as \mathbf{R} is at least non-negative definite. However, there are nontrivial situations in which \mathbf{R} can be singular, in which case the decorrelator is not a viable structure.

MMSE Receiver: While the matched filter addresses the ambient noise and the decorrelator addresses the interference, the minimum-mean-square-error (MMSE) multiuser detector effects a compromise between these two impairments by selecting the transformation \mathbf{M} such as the vector $\mathbf{v} = \mathbf{M}\mathbf{z}$ is an MMSE estimate of the symbol vector \mathbf{b} . For this criterion to make sense, it is necessary to provide a prior model for \mathbf{b} . On making the common assumption that the elements of \mathbf{b} are of zero mean and mutually uncorrelated, the MMSE detector corresponds to the following choice of the matrix \mathbf{M} :

$$\mathbf{M} = (\mathbf{R} + \sigma^2\mathbf{I})^{-1} \quad (36)$$

where, as before, \mathbf{I} denotes the $KM_TB \times KM_TB$ identity matrix. Note that, it is clear from this form that the MMSE detector represents a compromise between the matched filter ($\mathbf{M} = \mathbf{I}$) and the decorrelator ($\mathbf{M} = \mathbf{R}^{-1}$), in which the action of each is tempered with the action of the other. The relative mix of these two is controlled by the noise level (or more properly by the signal-to-noise ratio (SNR), as the signal strength is incorporated into \mathbf{R}). When the interference is dominant (i.e, for high SNR), the MMSE detector mimics the decorrelator, while when the ambient noise is dominant (i.e., for low SNR) it mimics the matched filter. More generally, it balances between these two.

In general, the complexity of linear multiuser detectors is that of matrix inversion, which is on the order of $(KM_TB)^3$. As with the ML and MAP solutions, this complexity can be reduced by exploiting bandedness in the case of short delay spread. In some cases, this complexity may also be amortized over many frames. However, for most wireless systems, such amortization is not possible as the signaling waveforms, the user population, or the channel parameters may change from frame to frame. Thus, although the order of complexity here has been reduced from exponential to polynomial, complexity is still a concern for practical systems. Moreover, in both linear and nonlinear cases, constraints on the transmitted symbols imposed by space-time coding or temporal channel coding can add to this complexity substantially [30].

For these reasons, a number of other techniques for multiuser reception have been developed, with the objective of reducing computational complexity while maintaining good performance in the presence of multiple-access interference. The principal technique for doing this is to make use of iterative algorithms to fit the linear model (16). This can be done either linearly with a final quantization (i.e., iterative linear detection), or nonlinearly with inter-iteration quantization (sometimes known as interference cancellation). Section 3 of this chapter will address this issue in some detail for multiuser MIMO systems. When further complexity is introduced by channel coding, iterative algorithms such as those described in Chapter 5 (in this case “turbo” style algorithms) again allow for excellent performance with moderate complexity. This topic is addressed in Section 4.

As noted above, another form of complexity is *informational* complexity, which arises from the need to know the received waveforms $\{g_{k,m,p}(\cdot)\}$ in the model (1) for the received signal. There are two potential problems with this requirement. One is that the channels intervening the transmitters and receiver are typically dynamic and behave in an apparently random fashion. So, the channel parameters (assuming the channel can be parameterized) are not readily known to the receiver. Another problem is that the signaling waveforms of all users may also not be known to the receiver, because, for example, the receiver may only be intended to receive a subset of the users. In either case, it is thus necessary for the receiver to be able to adapt itself to those properties of the signaling environment that it does not know. Receiver structures for this purpose are described in Section 5 of this chapter. In preparation for this latter

treatment, we turn briefly, in the following sub-section, to a discrete-time model for the received signals considered above that is more suitable for developing and discussing such adaptive receiver algorithms.

2.4 Digital Receiver Implementation

For receiver implementation, and particularly for the adaptive algorithms to be discussed in Section 5 of this chapter, it is useful to consider a digital representation of the signals and observables that we have described in the preceding paragraphs. This type of representation is typically obtained by projecting the received signals (1) onto a finite set of functions arising from a model in which there are finitely many degrees of freedom in the signals of interest. (Most practical signaling methods have this property.) In this sub-section, we will particularize the above structures for this situation, and in particular will consider the common case in which the signaling waveforms are in the DS/CDMA format, described above and in Chapter 1. This model will then be used exclusively in Section 5 of this chapter. It should be noted, however, that similar techniques can be applied in any system allowing for a finite-degree-of-freedom model. A notable alternative example to DS/CDMA is the case of orthogonal frequency-division multiple-access (OFDMA) systems, in which the incoming signal can be decomposed along orthogonal sub-carrier signals using the discrete Fourier transform (DFT).

Recall that, in the DS/CDMA format, the signaling waveforms used by all transmitters are in the form (7). Here, we consider this format in the particular case where the chip waveform is the unit pulse of (8). For this type of system, a natural set of observables can be obtained by projecting the received signals of (1) onto time shifts of the chip waveform $\psi(\cdot)$:

$$r_p[j] = \int_{-\infty}^{\infty} r_p(t)\psi(t - jT_s)dt = \int_{jT_c}^{(j+1)T_c} r_p(t)dt, \quad j = 0, 1, \dots \quad (37)$$

If the system delays are all integer multiples of a chip interval (this is termed the “chip-synchronous” case), then no information is lost in this operation, as the outputs of the matched-filter bank of Fig. 2 and hence the sufficient statistic \mathbf{z} can be extracted from these observables. In the chip-asynchronous case, inferential information may be lost in performing this operation. However, this loss is often minimal and the signal-processing advantages of reducing the observations to a discrete-time sequence outweigh this. (An alternative for the chip-asynchronous case is to integrate over shorter time intervals and thus

effectively to over-sample the signal; however, we will not consider this level of detail here. For further discussion, see [46].)

As noted above, in the chip-synchronous case, the sufficient statistic \mathbf{z} can be written as a function of the observables $\{r_p[j]\}$ and thus the ML and MAP detectors are functions of these observables, as are the linear detectors described in the preceding sub-section. In the latter case it is sometimes convenient to combine all of the linear processing of the receiver front end and the decision algorithm of Fig. 2 into a single linear transformation, in which case symbol detection is of the form:

$$\hat{b}_{k,m}[i] = Q \left(\sum_{p=1}^{M_R} \sum_j w_{k,m,p}^{(j)}[i] r_p[j] \right), \quad (38)$$

where the coefficients $\{w_{k,m,p}^{(j)}[i]\}$ are chosen appropriately. This structure is one that can be adapted using standard adaptive algorithms to adjust the weighting coefficients. Although there are a number of issues surrounding such adaptation, such as the decomposition of spatial and temporal combining, this structure is the essence of many adaptive algorithms for multi-antenna, multiuser receiver design. An extensive treatment of this problem can be found in [46], and we will consider particularly the MIMO case in Section 5.

3 Iterative Space-time Multiuser Detection

Advanced signal processing such as multiuser detection, typically improves system performance at the cost of computational complexity. As noted in Section 1, the optimal maximum likelihood (ML) multiuser detector has prohibitive computational requirements for most current applications, and consequently a variety of linear and nonlinear multiuser detectors have been proposed to ease this computational burden while maintaining satisfactory performance [38, 46]. However, in many situations where the combined system has large dimensions (e.g., large array size, large delay spread, large user population, and combinations of these conditions), direct implementation of these suboptimal techniques still proves to be very complex. In this section, we discuss iterative techniques for efficient space-time multiuser detection in MIMO systems [45, 8, 7]. Iterative methods are among the most practical techniques for multiuser detection. For example, an implementation for 3G cellular systems is described in [19].

As noted in Section 1, linear multiuser detectors in the framework of (39) are of the form

$$\hat{\mathbf{b}} = \text{sgn}(\text{Re}\{\mathbf{M}\mathbf{z}\}), \quad (41)$$

where \mathbf{M} is a linear detection matrix. For the linear decorrelating (zero-forcing) detector, this matrix is given by

$$\mathbf{M}_d = \mathbf{R}^{-1}, \quad (42)$$

while for the linear minimum-mean-square-error (MMSE) detector, we have

$$\mathbf{M}_m = (\mathbf{R} + \sigma^2\mathbf{I})^{-1}. \quad (43)$$

Direct inversion of the matrices in (42) and (43) (after exploiting the block Toeplitz structure) is of complexity $O((KM_T)^2B\Delta)$ per user per symbol.

The linear multiuser detection estimates of (41) can be seen as the solution of a linear equation

$$\mathbf{C}\mathbf{v} = \mathbf{z}, \quad (44)$$

with $\mathbf{C} = \mathbf{R}$ for the decorrelating detector and $\mathbf{C} = \mathbf{R} + \sigma^2\mathbf{I}$ for the MMSE detector. Jacobi and Gauss-Seidel iteration are two common low-complexity iterative schemes for solving linear equations such as (44) [14]. If we decompose the matrix \mathbf{C} as $\mathbf{C} = \mathbf{C}_L + \mathbf{D} + \mathbf{C}_U$ where \mathbf{C}_L denotes the lower triangular part, \mathbf{D} denotes the diagonal part, and \mathbf{C}_U denotes the upper triangular part, then Jacobi iteration can be written as

$$\mathbf{v}_m = -\mathbf{D}^{-1}(\mathbf{C}_L + \mathbf{C}_U)\mathbf{v}_{m-1} + \mathbf{D}^{-1}\mathbf{z}, \quad (45)$$

and Gauss-Seidel iteration is represented as

$$\mathbf{v}_m = -(\mathbf{D} + \mathbf{C}_L)^{-1}\mathbf{C}_U\mathbf{v}_{m-1} + (\mathbf{D} + \mathbf{C}_L)^{-1}\mathbf{z}. \quad (46)$$

From (45), Jacobi iteration can be seen to be a form of linear parallel interference cancellation, the convergence of which is not guaranteed in general. One of the sufficient conditions for the convergence of Jacobi iteration is that $\mathbf{D} - (\mathbf{C}_L + \mathbf{C}_U)$ be positive definite. In contrast, Gauss-Seidel iteration, which (46) reveals to be a form of linear serial interference cancellation, converges to the solution of the linear

equation from any initial value, under the mild conditions that \mathbf{C} be symmetric and positive definite, which is always true for the MMSE detector.

Another general approach to solving the linear equation (44) involves the use of gradient methods, among which are steepest descent and conjugate gradient iteration [14]. Note that solving (44) is equivalent to minimizing the cost function

$$\Phi(\mathbf{v}) = \frac{1}{2} \mathbf{v}^H \mathbf{C} \mathbf{v} - \mathbf{v}^H \mathbf{z}. \quad (47)$$

The idea of gradient methods is to successively minimize this cost function along a set of directions $\{\mathbf{p}_m\}$ via

$$\mathbf{v}_m = \mathbf{v}_{m-1} + \alpha_m \mathbf{p}_m, \quad (48)$$

with

$$\alpha_m = \mathbf{p}_m^H \mathbf{q}_{m-1} / \mathbf{p}_m^H \mathbf{C} \mathbf{p}_m, \quad (49)$$

and

$$\mathbf{q}_m = -\nabla \Phi(\mathbf{v})|_{\mathbf{v}=\mathbf{v}_m} = \mathbf{z} - \mathbf{C} \mathbf{v}_m. \quad (50)$$

Different choices of the set $\{\mathbf{p}_m\}$ give different algorithms. If we choose the search direction \mathbf{p}_m to be the negative gradient of the cost function \mathbf{q}_{m-1} directly, this algorithm is the steepest descent method, global convergence of which is guaranteed. The convergence rate may be prohibitively slow, however, due to the linear dependence of the search directions, resulting in redundant minimization. If instead we choose the search direction to be C-conjugate as follows

$$\mathbf{p}_m = \arg \min_{\mathbf{p} \in \Lambda_m^\perp} \|\mathbf{p} - \mathbf{q}_{m-1}\|, \quad (51)$$

where $\Lambda_m = \text{span}\{\mathbf{C}\mathbf{p}_1, \dots, \mathbf{C}\mathbf{p}_m\}$, then we have the conjugate gradient method, whose convergence is guaranteed and performs well when \mathbf{C} is close to the identity either in the sense of being low rank perturbation or in the sense of a norm. The computational complexity of Gauss-Seidel and conjugate gradient iteration are similar, which is on the order of $O(KM_T \Delta \bar{m})$ per user per symbol, where \bar{m} is the number of iterations. The numbers of iterations required by the Gauss-Seidel and conjugate gradient methods to achieve a stable solution to the associated large system equations have been found to be on the same order in simulations.

3.3 Iterative Nonlinear Space-Time Multiuser Detection

Nonlinear multiuser detectors are often based on bootstrapping techniques, which are iterative in nature. In this section, we will consider the iterative implementation of decision-feedback multiuser detection in the space-time domain. We also discuss briefly the implementation of multistage interference canceling ST MUD, which serves as a reference point for introducing a new expectation-maximization (EM-) based iterative ST MUD, to be discussed in the next subsection. For simplicity, we now restrict the signaling alphabet to the binary antipodal set: $\mathcal{A} = \{-1, +1\}$.

3.3.1 Cholesky Iterative Decorrelating Decision-Feedback ST MUD

Decorrelating decision feedback multiuser detection (DDF MUD) exploits the Cholesky decomposition $\mathbf{R} = \mathbf{F}^H \mathbf{F}$, where \mathbf{F} is a lower triangular matrix, to determine the feedforward and feedback matrix for detection via the algorithm

$$\hat{\mathbf{b}} = \text{sgn}(\mathbf{F}^{-H} \mathbf{z} - (\mathbf{F} - \text{diag}(\mathbf{F})) \hat{\mathbf{b}}). \quad (52)$$

The discussion here applies readily to the implementation of MMSE decision feedback multiuser detection as well.

Suppose we are interested in detecting symbol b_n . The purpose of the feedforward matrix \mathbf{F}^{-H} is to whiten the noise and decorrelate against the "future users" $\{s_{n+1}, \dots, s_{KM_T B}\}$; while the purpose of the feedback matrix $(\mathbf{F} - \text{diag}(\mathbf{F}))$ is to cancel the interference from "previous users" $\{s_1, \dots, s_{n-1}\}$. Note that the performance of DDF MUD is not uniform. While the first "user" is demodulated by its decorrelating detector, the last detected "user" will essentially achieve its single-user lower bound providing the previous decisions are correct. There is another form of Cholesky decomposition, in which the feedforward matrix \mathbf{F} is upper triangular. If we were to use this form instead in (52), then the multiuser detection would operate in the reverse order, as would the performance. The idea of *Cholesky iterative DDF ST MUD* is to employ these two forms of Cholesky decomposition alternatively as follows. For lower triangular Cholesky decomposition \mathbf{F}_1 , first feedforward filtering is applied as

$$\bar{\mathbf{z}}_1 = \mathbf{F}_1^{-H} \mathbf{z}, \quad (53)$$

where it is readily shown that $\bar{z}_{1,i} = \mathbf{F}_{1,ii}b_i + \sum_{j=1}^{i-1} \mathbf{F}_{1,ij}b_j + \bar{n}_{1,i}$, $i = 1, \dots, KM_TB$, with $\bar{n}_{1,i}$, $i = 1, \dots, KB$, being independent and identically distributed (i. i. d.) Gaussian noise components with zero mean and variance σ^2 . We can see that the influence of the "future users" is eliminated and the noise component is whitened. Then we use the feedback filtering to cancel the interference from "previous users" as

$$\mathbf{u}_1 = \bar{\mathbf{z}}_1 - (\mathbf{F}_1 - \text{diag}(\mathbf{F}_1))\hat{\mathbf{b}}, \quad (54)$$

where it is easily seen that $u_{1,i} = \bar{z}_{1,i} - \sum_{j=1}^{i-1} \mathbf{F}_{1,ij}\hat{b}_j \approx \mathbf{F}_{1,ii}b_i + \bar{n}_{1,i}$, $i = 1, \dots, KM_TB$. Similarly, for upper triangular Cholesky decomposition \mathbf{F}_2 , we have

$$\bar{\mathbf{z}}_2 = \mathbf{F}_2^{-H}\mathbf{z}, \quad (55)$$

where $\bar{z}_{2,i} = \mathbf{F}_{2,ii}b_i + \sum_{j=i+1}^{KB} \mathbf{F}_{2,ij}b_j + \bar{n}_{2,i}$, $i = KM_TB, \dots, 1$, and

$$\mathbf{u}_2 = \bar{\mathbf{z}}_2 - (\mathbf{F}_2 - \text{diag}(\mathbf{F}_2))\hat{\mathbf{b}}, \quad (56)$$

where $u_{2,i} = \bar{z}_{2,i} - \sum_{j=i+1}^{KM_TB} \mathbf{F}_{2,ij}\hat{b}_j \approx \mathbf{F}_{2,ii}b_i + \bar{n}_{2,i}$, $i = KM_TB, \dots, 1$. After the above operations are (alternately) executed, the following log-likelihood ratio is calculated,

$$L_i = 2\text{Re}(\mathbf{F}_{1/2,ii}^* u_{1/2,i})/\sigma^2, \quad (57)$$

where $\mathbf{F}_{1/2}$ and $\mathbf{u}_{1/2}$ are used to give a shorthand representation for both alternatives. Then the log-likelihood ratio is compared with the last stored value, which is replaced by the new value if the new one is more reliable, i.e.,

$$L_i^{stored} = \begin{cases} L_i^{stored} & \text{if } |L_i^{stored}| > |L_i^{new}|, \\ L_i^{new} & \text{otherwise.} \end{cases} \quad (58)$$

Finally we make soft decisions $\hat{b}_i = \tanh(L_i/2)$ at an intermediate iteration, which has been shown to offer better performance than making hard intermediate decisions, and make hard decisions $\hat{b}_i = \text{sgn}(L_i)$ at the last iteration. Several iterations are usually enough for the system to achieve an improved steady state without significant oscillation. The structure of Cholesky iterative decorrelating decision-feedback ST MUD is illustrated in Fig. 4.

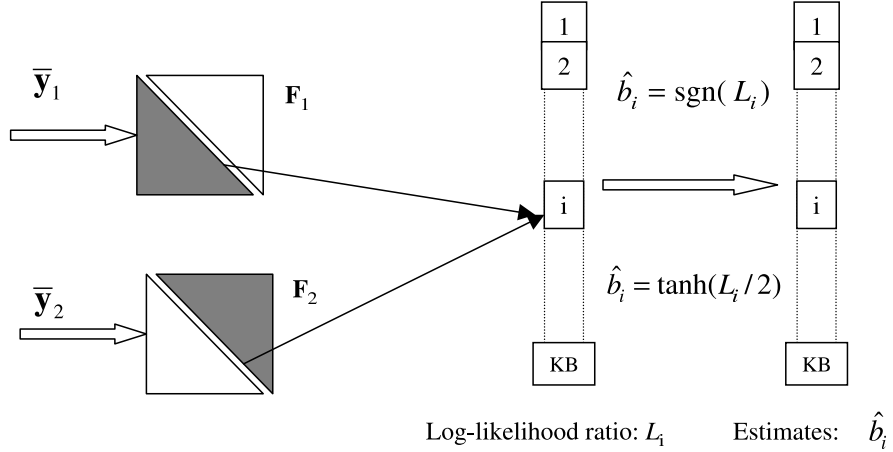


Figure 4: Cholesky iterative decorrelating decision-feedback ST MUD

The Cholesky factorization of the block Toeplitz matrix \mathbf{H} (see (40)) can be performed recursively.

For $\Delta = 1$ we have

$$\mathbf{F} = \begin{bmatrix} \underline{E}_1(0) & 0 & \cdots & 0 & 0 \\ \underline{E}_2(1) & \underline{E}_2(0) & \cdots & 0 & 0 \\ 0 & & & & \\ \vdots & & & & \\ 0 & 0 & \cdots & \underline{E}_M(1) & \underline{E}_M(0) \end{bmatrix}, \quad (59)$$

where the element matrices are obtained recursively as follows.

$$\underline{V}_B = \underline{R}^{[0]}, \quad (60)$$

and, for $i = B, B - 1, \dots, 1$, we perform Cholesky decomposition of the reduced-rank matrix \underline{V}_i to get

$\underline{E}_i(0)$

$$\underline{V}_i = \underline{E}_i^H(0) \underline{E}_i(0), \quad (61)$$

while $\underline{E}_i(1)$ is obtained as

$$\underline{E}_i(1) = (\underline{E}_i^H(0))^{-1} \underline{R}^{[-1]}. \quad (62)$$

Finally we have

$$\underline{V}_{i-1} = \underline{R}^{[0]} - \underline{R}^{[1]} \underline{V}_i^{-1} \underline{R}^{[-1]} \quad (63)$$

for use in the next iteration. The extension of this algorithm to $\Delta > 1$ is straightforward and is omitted here.

3.3.2 Multistage Interference Canceling ST MUD

Multistage interference cancellation (IC) is similar to Jacobi iteration except that hard decisions are made at the end of each stage in place of the linear terms that are fed back in (45). Thus we have

$$\hat{\mathbf{b}}_m = \text{sgn}(\mathbf{z} - (\mathbf{C}_L + \mathbf{C}_U)\hat{\mathbf{v}}_{m-1}) = \text{sgn}(\mathbf{y} - (\mathbf{H} - \mathbf{D})\hat{\mathbf{b}}_{m-1}). \quad (64)$$

The underlying rationale for this method is that the estimator-subtractor structure exploits the discrete-alphabet property of the transmitted data streams. This nonlinear hard-decision operation typically results in more accurate estimates in high SNR situations. Although the optimal decisions are a fixed point of the nonlinear transformation (64), there are problems with the multistage IC such as a possible lack of convergence and oscillatory behavior. In the following section we consider some improvements on space-time multistage IC MUD. Except for the Cholesky factorization, the computational complexity for Cholesky iterative DDF ST MUD is the same as multistage IC ST MUD, which is essentially the same as that of linear interference cancellation, i.e., $O(KM_T\Delta\bar{m})$ per user per symbol.

3.4 EM-based Iterative Space-Time Multiuser Detection

In this section, expectation-maximization-based multiuser detection is introduced to avoid the convergence and stability problem of the multistage IC MUD.

The EM algorithm [10] provides an iterative solution of maximum likelihood estimation problems such as

$$\hat{\boldsymbol{\theta}}(\mathbf{Z}) = \arg \max_{\boldsymbol{\theta} \in \Lambda} \log f(\mathbf{Z}; \boldsymbol{\theta}), \quad (65)$$

where $\boldsymbol{\theta} \in \Lambda$ are the parameters to be estimated, and $f(\cdot)$ is the parameterized probability density function of the observable \mathbf{Z} . The idea of the EM algorithm is to consider a judiciously chosen set of "missing

data" \mathbf{W} to form the complete data $\mathbf{X} = \{\mathbf{Z}, \mathbf{W}\}$ as an aid to parameter estimation, and then to iteratively maximize the following new objective function

$$Q(\boldsymbol{\theta}; \bar{\boldsymbol{\theta}}) = E [\log f(\mathbf{Z}, \mathbf{W}; \boldsymbol{\theta}) | \mathbf{Z} = \mathbf{z}; \bar{\boldsymbol{\theta}}], \quad (66)$$

where it is worth emphasizing again that $\boldsymbol{\theta}$ are the parameters in the likelihood function, which are to be estimated, while $\bar{\boldsymbol{\theta}}$ represent *a priori* estimates of the parameters from the previous iterations. Together with the observations, these previous estimates are used to calculate the expected value of the log-likelihood function with respect to the complete data $\mathbf{X} = \{\mathbf{Z}, \mathbf{W}\}$. To be specific, given an initial estimate $\boldsymbol{\theta}^0$, the EM algorithm alternates between the following two steps:

1. E-step, where the complete-data sufficient statistic $Q(\boldsymbol{\theta}; \boldsymbol{\theta}^i)$ is computed;
2. M-step, where the estimates are refined by $\boldsymbol{\theta}^{i+1} = \arg \max_{\boldsymbol{\theta} \in \Lambda} Q(\boldsymbol{\theta}; \boldsymbol{\theta}^i)$.

It has been shown that EM estimates monotonically increase the likelihood, and converge stably to an ML solution under certain conditions [10].

An issue in using the EM algorithm is the tradeoff between ease of implementation and convergence rate. One would like to add more "missing data" to make the complete data space more informative so that the implementation of the EM algorithm is simpler than the original setting (65). However, the convergence rate of the algorithm is inversely proportional to the Fisher information contained in the complete data space. This tradeoff is essentially due to the simultaneous updating nature of the M-step in the original EM algorithm [11]. Consequently, the space-alternating generalized EM (SAGE) algorithm has been proposed in [11] to improve the convergence rate for multidimensional parameter estimation. The idea is to divide the parameters into several groups (subspaces), with only one group being updated at each iteration. Thus, we can associate multiple less-informative "missing data" sets to improve the convergence rate while maintaining the overall tractability of optimization problems. For each iteration, a subset of parameters $\boldsymbol{\theta}_{S_i}$ and the corresponding missing data \mathbf{Z}^{S_i} are chosen, which is called the definition step. Then similarly to the EM algorithm, in the E-step we calculate

$$Q^{S_i}(\boldsymbol{\theta}_{S_i}; \boldsymbol{\theta}^i) = E [\log f(\mathbf{Z}, \mathbf{W}^{S_i}; \boldsymbol{\theta}_{S_i}, \boldsymbol{\theta}_{\tilde{S}_i}^i | \mathbf{Z} = \mathbf{z}; \boldsymbol{\theta}^i)], \quad (67)$$

where $\boldsymbol{\theta}_{\tilde{S}_i}$ denotes the complement of $\boldsymbol{\theta}_{S_i}$ in the whole parameter set; in the M-step, the chosen parameters are updated while the others remain unchanged as

$$\begin{cases} \boldsymbol{\theta}_{S_i}^{i+1} = \arg \max_{\boldsymbol{\theta}_{S_i} \in \Lambda_{S_i}} Q^{S_i}(\boldsymbol{\theta}_{S_i}; \boldsymbol{\theta}^i), \\ \boldsymbol{\theta}_{\tilde{S}_i}^{i+1} = \boldsymbol{\theta}_{\tilde{S}_i}^i, \end{cases} \quad (68)$$

where Λ_{S_i} denotes the restriction of the entire parameter space to those dimensions indexed by S_i . Like the traditional EM estimates, the SAGE estimates also monotonically increase the likelihood and converge stably to an ML solution under appropriate conditions [11].

The EM algorithm is applied to space-time multiuser detection as follows. Suppose we would like to detect a bit b_n , $n \in \{1, 2, \dots, KM_TB\}$, which can be viewed as the parameter of interest, while the interfering users' bits $\mathbf{b}_{\tilde{k}} = \{b_j\}_{j \neq n}$ are treated as the missing data. The complete-data sufficient statistic is given by (\mathbf{R}_{nm} is the element of matrix \mathbf{R} at the k th row and m th column)

$$Q(b_n; b_n^i) = \frac{1}{2\sigma^2} \left(-R_{nn}b_n^2 + 2b_n(z_n - \sum_{m \neq n} \mathbf{R}_{nm}\tilde{b}_m) \right), \quad (69)$$

with

$$\tilde{b}_m = E[b_m | \mathbf{Z} = \mathbf{z}; b_n = b_n^i] = \tanh \left(\frac{R_{mm}}{\sigma^2} (z_m - \mathbf{R}_{mn}b_n^i) \right), \quad (70)$$

which forms the E-step of the EM algorithm. The M-step is given by

$$b_n^{i+1} = \arg \max_{b_n \in \Lambda} Q(b_n; b_n^i) = \begin{cases} \text{sgn}(z_n - \sum_{m \neq n} \mathbf{R}_{nm}\tilde{b}_m), & \Lambda = \{\pm 1\}, \\ \frac{1}{R_{nn}}(z_n - \sum_{m \neq n} \mathbf{R}_{nm}\tilde{b}_m), & \Lambda = \mathfrak{R}, \end{cases} \quad (71)$$

where $\Lambda = \mathfrak{R}$ (the set of real numbers) means a soft decision is needed, e.g., in an intermediate stage. Note that in the E-step (70), interference from users $j \neq n$ is not taken into account, since these are treated as "missing data". This shortcoming is overcome by the application of the SAGE algorithm, where the symbol vector of all users $\mathbf{b} = \{b_j\}_{j=1}^{KM_TB}$ is treated as the parameter to be estimated and no missing data is needed. The algorithm is described as follows: for $i = 0, 1, \dots$,

1. Definition step: $S_i = 1 + (i \bmod KM_TB)$

2. M-step:

$$\begin{cases} b_n^{i+1} = \text{sgn}(z_n - \sum_{m \neq n} \mathbf{R}_{nm}b_m^i), & n \in S_i, \\ b_m^{i+1} = b_m^i, & m \notin S_i. \end{cases} \quad (72)$$

Note that there is no E-step since there is no missing data, and interference from all other users are recreated from previous estimates and subtracted. The resulting receiver is similar to the multistage interference canceling multiuser receiver (see (64)), except that the symbol estimates are made sequentially rather than in parallel. However, with this simple concept of sequential interference cancellation, the resulting multiuser receiver is convergent, guaranteed by the SAGE algorithm. The multistage interference canceling multiuser receiver discussed in 3.3.2, on the other hand, does not always converge. The computational complexity of this SAGE iterative ST MUD is also $O(K\Delta\bar{m})$ per user per symbol.

3.5 Simulation Results

In this section, the performance of the above described space-time multiuser detectors is examined through simulations on a CDMA example. We assume a $K = 8$ -user CDMA system with spreading gain $N = 16$. Each user, equipped with one single antenna, travels through $L = 3$ paths before it reaches a base station (or access point), equipped with a uniform linear array with $M_R = 3$ elements and half-wavelength spacing. The maximum delay spread is set to be $\Delta = 1$. The complex gains and delays of the multipath and the directions of arrival are randomly generated and kept fixed for the whole data frame. This corresponds to a slow fading situation. The spreading codes of all users are randomly generated and kept fixed for all the simulations.

First we compare the performance of various space-time multiuser receivers and some single-user space-time receivers in Fig. 5. Five receivers are considered: the single-user matched filter (Matched-Filter), the single-user MMSE receiver (SU MMSE), the multiuser MMSE receiver (MU MMSE) implemented using the Gauss-Seidel or conjugate gradient iteration method (the performance is the same for both), the Cholesky iterative decorrelating decision-feedback multiuser receiver (Cholesky Iterative MU DF), and the multistage interference canceling multiuser receiver (MU Multistage IC). An interested reader can refer to [45] for derivations of the single-user based receivers. The performance is evaluated after the iterative algorithms converge. Due to the poor convergence behavior of the multistage IC MUD, we measure its performance after three stages. The single-user lower bound is also depicted for reference. We can see that the multiuser approach greatly outperforms the single-user based methods;

nonlinear MUD offers further gain over the linear MUD; and the multistage IC seems to approach the optimal performance (not always though, as is seen in Fig. 7(c)), when it has good convergence behavior. Note that due to the introduction of spatial (receive antenna) and spectral (RAKE combining) diversity, the SNR for the same BER is substantially lower than that required by normal receivers without these methods.

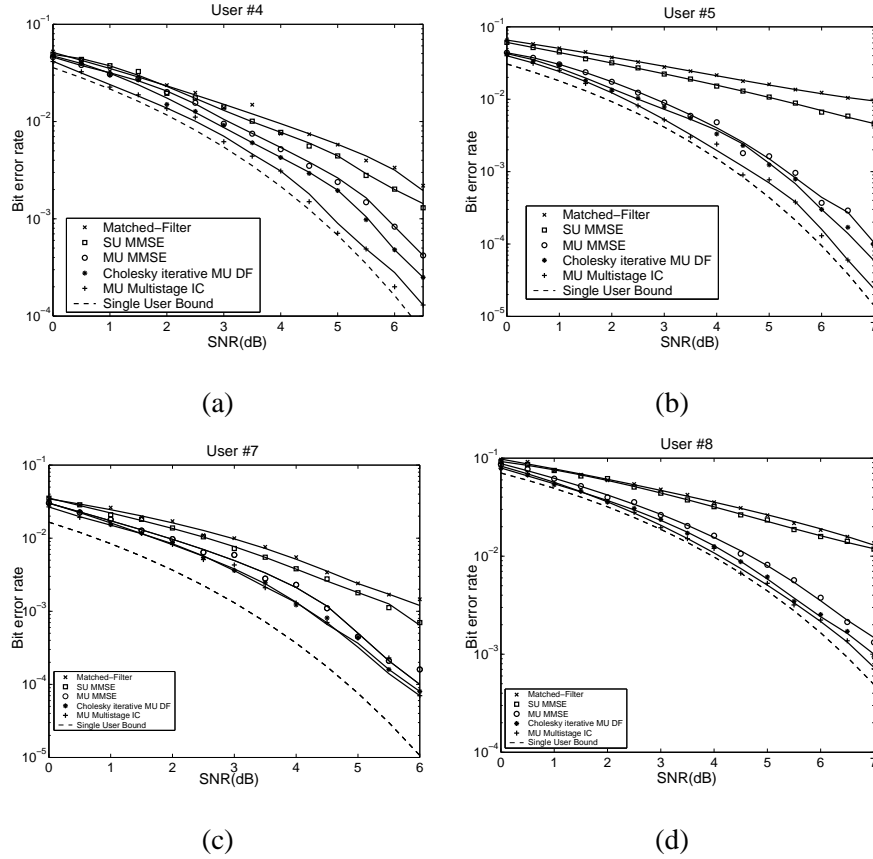


Figure 5: Performance comparison of BER versus SNR for five space-time multiuser receivers

Fig. 6 shows the performance of Cholesky iterative decorrelating decision-feedback ST MUD for two users, which is also typical for other users. We find that the Cholesky iterative method offers uniform gain over its non-iterative counterpart. This gain may be substantial for some users and negligible for others due to the individual characteristics of signals and channels.

Finally, we show the advantage of the EM-based (SAGE) iterative method over the multistage IC method with regard to the convergence of the algorithms. From Fig. 7 we find that, while the multistage

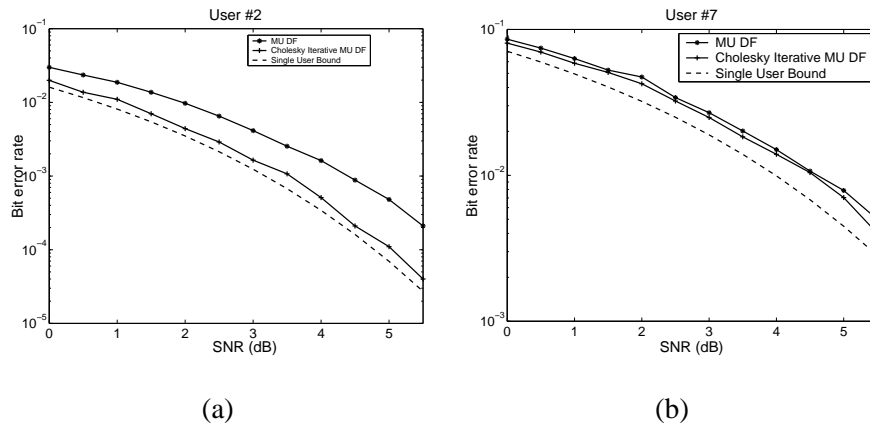


Figure 6: Performance comparison of decision-feedback ST MUD and Cholesky iterative ST MUD

interference canceling ST MUD converges slowly and exhibits oscillatory behavior, the SAGE ST MUD converges quickly and outperforms the multistage IC method. The oscillation of the performance of the multistage IC corresponds to performance degradation as no statistically best iteration number can be chosen.

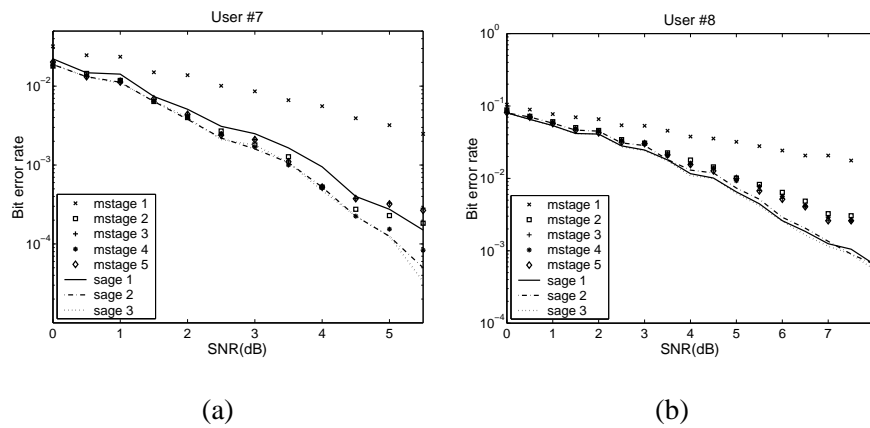


Figure 7: Performance comparison of convergence behavior of multistage interference canceling ST MUD and EM-based iterative ST MUD

3.6 Summary

In this section, we have considered several iterative space-time multiuser detection schemes. It is shown that iterative implementation of these linear and nonlinear multiuser receivers approaches the optimum

performance with reasonable complexity. Among these iterative implementations the SAGE space-time multiuser receiver outperforms the others while requiring similar complexity. While we focus on single-cell communications, all techniques discussed here can be extended to the multicell scenario [6], where the requirement for efficient algorithms only become more stringent.

4 Multiuser Detection in Space-time Coded Systems

With the invention of powerful space-time coding techniques in late 1990s as described in Chapter 4, there has been a growing interest in adapting these to multiple-access communication systems. Although early space-time code construction was concerned with single-user channels [1, 36, 37] (see Chapter 4), subsequently it has been shown that most of the performance criteria developed can still be used effectively in multiuser channels [21]. Space-time block codes have been applied to multiple-access communication systems in [25, 24, 9]. The receivers that explicitly take into account the structure of the space-time block codes have shown to perform well in this context [27, 23].

Here we consider multiuser detection for space-time coded multiple-access systems. As we will see, the joint Maximum Likelihood (ML) decoder for such systems has prohibitively large computational complexity, motivating a search for low-complexity, sub-optimal detector structures. We investigate several partitioned space-time multiuser detectors that separate the multiuser detection and space-time decoding into two stages. Both linear and non-linear schemes are considered for the first stage of the partitioned receiver and the performance versus complexity trade-offs are discussed.

Inspired by the development of turbo codes [3, 4] that were discussed in Chapter 5, various iterative detection and decoding schemes for multiple-access channels have been proposed in recent years. These proposals show that in general iterative receivers can offer significant performance improvements over their non-iterative counterparts. A good example is [44], in which a soft interference cancelling turbo receiver was proposed for convolutionally coded CDMA. The performance results obtained via simulations showed that it is possible to achieve near single-user performance with only a few iterations in an asynchronous, multipath CDMA channel. In this section, among others, we will show a generalization of this idea to a space-time coded CDMA system as in [20, 21, 26].

The development of turbo multiuser receivers for space-time coded systems here closely follows that of [21]. In particular, we assume a multiple-access system based on DS-CDMA signalling as opposed to space-division multiple-access as in [26]. There are two main implementations of CDMA-based multiple transmit antenna systems. One involves assigning a single spreading code to each user so that the signals transmitted from all its antennas are spread by the same code. We will assume a design of this type. In an alternative implementation, each user is assigned multiple spreading codes so that the signals transmitted from different antennas are spread by different codes [18, 17, 9, 28].

Low complexity multiuser receiver structures for space-time coded systems have been described in [26, 47]. For example, a multi-stage receiver suitable for a system employing both turbo and space-time block coding was proposed in [47]. Turbo receiver structures for multiple-access systems with both space-time block and trellis coded systems have been presented in [20, 21, 26]. In general these turbo receivers operate by partitioning the detection and decoding into two separate stages. In the first stage, a multiuser detection technique is employed and a set of soft outputs is generated for each user. The next stage of the receiver is equipped with a bank of decoders (either channel, space-time or both) that decode the individual user channel or space-time codes (or both). These decoders then generate an updated set of soft information about the code symbols which can then be fed back to the first stage to be used as *a priori* information at the next iteration. The process continues by repeating the same steps.

In Section 4.1 we present a simplified signal model for a space-time coded, synchronous multiuser system, while in Section 4.2 we derive the jointly optimal ML detector/decoder. In Section 4.3, we consider low complexity receiver structures for space-time coded multiuser systems by separating the multiuser detection stage from the space-time decoder stage. We consider both linear and non-linear multiuser detection stages. In particular, in this section, we consider partitioned space-time multiuser receivers based on the linear decorrelator and on the linear Minimum-Mean-Square-Error (MMSE) estimator, as well as two partitioned receiver structures based on non-linear interference cancelling multiuser detection stages. Section 4.4 details a soft input soft output (SISO) *maximum a posteriori* (MAP) decoder [2] that can be used as the second stage of the interference cancelling receivers (For more details on MAP decoding refer to Chapter 5).

4.1 Signal Model

Consider a system of K independent users, each employing an independent space-time code with M_T transmitter antennas. The binary information sequence $\{d_k[n]\}_{n=0}^{\infty}$ of user k , for $k = 1, \dots, K$, is first encoded by a space-time encoder, and then the encoded data is divided into M_T streams by passing it through a serial-to-parallel converter. (For simplicity we assume that all the users employ the same number of transmitter antennas, although generalizing to different numbers of transmitter antennas is straightforward.) The code bits in each parallel stream are block interleaved, BPSK symbol-mapped, modulated by an appropriate signature waveform, $s_k(t)$, and are transmitted simultaneously from the M_T transmitter antennas. It should be emphasized that throughout this section we assume that user k employs the same signaling waveform $s_k(t)$ in all its M_T transmitter antennas (i.e. $s_{k,m}(t) = s_k(t)$ for $m = 1, \dots, M_T$).

The k -th user's transmitted signal at time t can thus be written as³

$$x_k(t) = \frac{A_k}{\sqrt{M_T}} \sum_{i=0}^{B-1} \sum_{m=1}^{M_T} b_{k,m}[i] s_k(t - iT), \quad (73)$$

where $\{b_{k,m}[i] \in \{+1, -1\}\}_{i=0}^{B-1}$ is the symbol-mapped space-time encoder output of the k -th user on transmitter antenna m at time i , and B is the number of channel symbols per user in a data frame which is assumed to be the same as the length of a space-time codeword. We assume that the signature waveform of each user is supported only on the interval $0 \leq t \leq T$, and is normalized so that $\int_0^T s_k^2(t) dt = 1$, for $k = 1, \dots, K$. Thus, A_k^2 represents the transmitted energy per bit of user k , independent of the number of transmitter antennas. Note that the model of (73) is otherwise general with regard to the signalling format, and so the following results can be applied to any signalling scheme. However, we are interested here in non-orthogonal signalling schemes such as code-division multiple-access (CDMA).

Assuming that the fading is sufficiently slow to be constant over a received data frame, the corresponding signal received at a single receive antenna can be written as

³Elsewhere in this chapter, we have assumed that the transmitted signals are normalized, and have absorbed the transmitter amplitude into the channel response. In this section, we will decompose the channel response to explicitly show the transmitted amplitude as a separate term, similarly to (20).

$$r(t) = \sum_{i=0}^{B-1} \sum_{k=1}^K \frac{A_k}{\sqrt{M_T}} \sum_{n_T=1}^{M_T} h_{k,m} b_{k,m}[i] s_k(t - iT) + n(t) \quad (74)$$

where $n(t)$ is complex white Gaussian noise with zero mean and variance $N_0/2$ per dimension. The complex fading coefficient, $h_{k,m}$, between the k -th user's m -th transmitter antenna and the receiver, is assumed to be a zero-mean unit variance complex Gaussian random variable with independent real and imaginary parts. Equivalently, $h_{k,m}$ has uniform phase and Rayleigh amplitude; i.e., the so-called Rayleigh fading model. These fading coefficients are assumed to be mutually independent with respect to both k and m . In what follows, we assume that all parameters of the model (74) are known to the receiver. Only the transmitted symbols are unknown.

4.2 Joint ML Multiuser Detection and Decoding for Space-time Coded Multiuser Systems

We start by considering the joint maximum-likelihood detection and decoding of the symbols in the model of Section 4.1. To do so, we first establish some notation.

As before, we denote the k -th user's transmitted symbol vector (on M_T antennas) at time i by the vector $\mathbf{b}_k[i] = [b_{k,1}[i] \dots b_{k,M_T}[i]]^T$. Define, the $BK \times M_T K$ joint codeword matrix \mathbf{D} , of all users, as

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_1 & \mathbf{0}_{B \times M_T} & \dots & \mathbf{0}_{B \times M_T} \\ \mathbf{0}_{B \times M_T} & \mathbf{D}_2 & \dots & \mathbf{0}_{B \times M_T} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{B \times M_T} & \mathbf{0}_{B \times M_T} & \dots & \mathbf{D}_K \end{bmatrix} \quad (75)$$

where we have also introduced the notation, for $k = 1, \dots, K$,

$$\mathbf{D}_k = \begin{bmatrix} \mathbf{b}_k^T[0] \\ \vdots \\ \mathbf{b}_k^T[B-1] \end{bmatrix}. \quad (76)$$

Note that $\mathbf{D}_k \in \{+1, -1\}^{B \times M_T}$, for $k = 1, \dots, K$. We will call the joint codeword, \mathbf{D} , of all users, the *super codeword*. The space-time coded output from all the users at time i is the $K \times KM_T$ matrix

denoted as $\mathbf{D}[i]$, where

$$\mathbf{D}[i] = \begin{bmatrix} \mathbf{b}_1^\top[i] & \mathbf{0}_{1 \times M_T} & \cdots & \mathbf{0}_{1 \times M_T} \\ \mathbf{0}_{1 \times M_T} & \mathbf{b}_2^\top[i] & \cdots & \mathbf{0}_{1 \times M_T} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{1 \times M_T} & \mathbf{0}_{1 \times M_T} & \cdots & \mathbf{b}_K^\top[i] \end{bmatrix}. \quad (77)$$

The fading coefficients of the k -th user can be collected into a vector $\mathbf{h}_k = [h_{k,1}, \dots, h_{k,M_T}]^\top \in \mathbb{C}^{M_T \times 1}$, and we can combine all these fading coefficient vectors into one vector $\mathbf{h} = [\mathbf{h}_1^\top \dots \mathbf{h}_K^\top]^\top \in \mathbb{C}^{KM_T \times 1}$. With this notation, the output, $\mathbf{z}[i] = [z_1[i] \dots z_K[i]]^\top$, of a bank of K matched filters (each matched to a user signature waveform $s_k(t)$) at the i -th symbol interval can be written as,

$$\mathbf{z}[i] = \overline{\mathbf{R}}\mathbf{A}\mathbf{D}[i]\mathbf{h} + \boldsymbol{\eta}[i] \quad (78)$$

where the diagonal matrix \mathbf{A} is defined as $\mathbf{A} = \text{diag}(\frac{A_1}{\sqrt{M_T}}, \dots, \frac{A_K}{\sqrt{M_T}})$, $\overline{\mathbf{R}}$ is the (normalized) cross-correlation matrix of the users' signature waveforms and $\boldsymbol{\eta}(i) \sim \mathcal{N}(\mathbf{0}, N_0\overline{\mathbf{R}})$.

Let us denote the B -vector of the k -th matched filter outputs corresponding to the complete received codeword as $\mathbf{z}_k = [z_k(0) \dots z_k(B-1)]^\top$ and the BK -vector of outputs of all the matched filters corresponding to a complete codeword as $\mathbf{z} = [\mathbf{z}_1 \dots \mathbf{z}_K]^\top$. Then we can write,

$$\mathbf{z} = (\overline{\mathbf{R}}\mathbf{A} \otimes \mathbf{I}_B)\mathbf{D}\mathbf{h} + \boldsymbol{\eta} \quad (79)$$

where $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, N_0\overline{\mathbf{R}} \otimes \mathbf{I}_B)$, \mathbf{I}_B denotes the $B \times B$ identity matrix and \otimes denotes the Kronecker product. The joint ML multiuser decision rule for the space-time coded CDMA system is then given by

$$\begin{aligned} \hat{\mathbf{D}} &= \arg \max_{\mathbf{D}} p(\mathbf{z}|\mathbf{D}, \mathbf{h}) \\ &= 2\text{Re} \left\{ \mathbf{h}^H \mathbf{D}^\top (\mathbf{A} \otimes \mathbf{I}_B) \mathbf{z} \right\} - \mathbf{h}^H \mathbf{D}^\top (\mathbf{A} \otimes \mathbf{I}_B) (\overline{\mathbf{R}} \otimes \mathbf{I}_B) (\mathbf{A} \otimes \mathbf{I}_B) \mathbf{D} \mathbf{h} \end{aligned}$$

where the maximization is over all the valid super codewords and we have used the fact that for general matrices \mathbf{A} , \mathbf{B} , \mathbf{C} and \mathbf{D} we have, $(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = (\mathbf{A}\mathbf{C} \otimes \mathbf{B}\mathbf{D})$ provided the dimensions of the matrices \mathbf{A} , \mathbf{B} , \mathbf{C} and \mathbf{D} are such that the various matrix products are well-defined [22]. Note that this

joint ML detector and decoder searches over a super trellis made up by combining all the users' space-time code trellises.

The asymptotic performance of a space-time code can be quantified by the so-called diversity gain. The diversity gain determines the asymptotic slope of the probability of error curve in log scale. As discussed in Chapter 4, in order to maximize the diversity gain for a Rayleigh fading channel one should design the space-time code such that the minimum rank of codeword difference matrix for any two codewords is as large as possible [15, 36]. When this minimum rank over all pairs of distinct codewords is the largest possible value M_T , then the space-time code is said to achieve full diversity.

In [21], it was shown that the space-time codes designed to achieve full diversity in single user channels will also be able to achieve full diversity asymptotically in CDMA multiuser channel, at least when the SNR is sufficiently large. That is, if the minimum rank of all the valid error codewords $\mathbf{E}_k = \mathbf{D}_k - \hat{\mathbf{D}}_k$ is r_k (where $r_k \leq M_T$), then the asymptotic diversity advantage of the k -th user's space-time code in the multiuser channel is equal to r_k . In particular, if the k -th user's space-time code were to achieve the full diversity M_T in a single-user environment, then it will also achieve the full diversity M_T in the multiuser channel, at least asymptotically in SNR, as long as the signature cross-correlation matrix is non-singular.

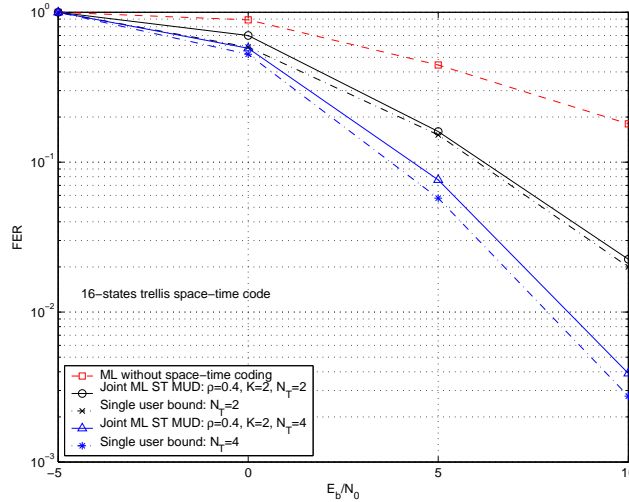


Figure 8: FER Performance Versus E_b/N_0 (in dB) of the Joint Maximum Likelihood Space-time Multiuser Detector. $K = 2$ and $\rho = 0.4$.

Figure 8 shows the performance results for the joint Maximum Likelihood detector in a space-time-coded, synchronous, multiple-access system with two equal-power users having a cross-correlation of 0.4. We set the number of receiver antennas to one, ignoring the possibility of exploiting receiver diversity since our primary concern here is to investigate the transmitter diversity schemes. In Fig. 8 we have shown the joint ML receiver performance results for two systems: one with two transmit antennas and another with four transmit antennas. We make use of full diversity BPSK space-time trellis codes with constraint length $\nu = 5$, given in [16], for both systems. Specifically, we employ space-time codes based on the underlying rate-1/2 convolutional code with octal generators (46, 72), and the underlying rate-1/4 convolutional code with octal generators (52, 56, 66, 76), both given in [16], for the two and four antenna systems, respectively. We use the Frame Error Rate (FER) as the measure of performance. Also shown on this figure is the performance of an equivalent system but without space-time coding. Figure 8 reveals the significant gains that can be achieved with space-time coding in multiuser systems. Moreover, it shows that the joint ML receiver performance is very close to that of the single-user bound as predicted above.

It is easily seen that the above ML path search can be implemented as a maximum-likelihood path search over a super trellis formed by combining all the users' space-time code trellises using the Viterbi algorithm. This is similar to the optimal decoder for convolutionally coded CDMA channels derived in [12]. Assuming (for simplicity) that all the users employ space-time codes based on underlying convolutional codes that have a constraint length ν , this super trellis will have a total of $K(\nu - 1)$ states, resulting in a total complexity per user of about $\mathcal{O}(\frac{2^{K\nu}}{K})$, which is exponential in $K\nu$. Note also that, in order to achieve full diversity gain M_T in an M_T transmitter antenna system we must have $\nu \geq M_T$ [16, 37]. Hence, it is clear that even for a small number of users this could easily become a prohibitively large computational burden at the receiver. This motivates us to look for sub-optimal, low complexity receiver structures for space-time coded multiuser systems.

In order to reduce the computational complexity of joint multiuser detection and space-time decoding while still achieving competitive performance against the joint ML decision rule, one can use partitioned receiver structures. Specifically, the multiuser detection and the space-time decoding can be

separated into two stages, as is done in [13] for the case of (single-antenna) convolutionally coded CDMA channels. At the first stage of the partitioned receiver, multiuser detection is performed. The outputs from the multiuser detection stage are then passed onto a bank of single-user space-time decoders corresponding to the K users in the system. Thus, each user's space-time decoder operates independently from the others. Of course, it is possible to employ either an ML or a maximum *a posteriori* probability (MAP) decoder as the single-user space-time decoder at the second stage of the receiver. Also, it is possible to use any reasonable multiuser detection strategy at the first stage of the receiver. In the following we consider both linear and non-linear multiuser detectors as the first stage of the partitioned space-time multiuser receiver, and compare the performance of these receivers against the best possible performance.

4.3 Partitioned Low Complexity Receivers for Space-time Coded Multiuser Systems

We consider linear multiuser detection based partitioned receivers, followed by the non-linear multiuser detection approaches. For linear multiuser detectors, we investigate both decorrelator and linear MMSE detectors [38]. For non-linear approaches we consider both a simple iterative receiver based on interference cancellation and the turbo principle, and an improved iterative receiver based on instantaneous MMSE filtering after the interference cancellation step.

4.3.1 Decorrelator Based Partitioned Space-time Multiuser Receiver

The decorrelator output at the i -th symbol time is given by [38],

$$\hat{\mathbf{z}}[i] = \overline{\mathbf{R}}^{-1} \mathbf{z}[i] = \mathbf{A}\mathbf{D}[i]\mathbf{h} + \hat{\boldsymbol{\eta}}, \quad (80)$$

where $\hat{\boldsymbol{\eta}} \sim \mathcal{N}(\mathbf{0}, N_0 \overline{\mathbf{R}}^{-1})$. The first stage of the receiver computes soft outputs corresponding to each user's transmitted symbol vectors at time i . The soft outputs are the *a posteriori* probabilities (APPs) of each user's transmitted symbol vectors, defined as below for $l = 1 \dots, 2^{M_T}$, $k = 1 \dots, K$ and $i = 0 \dots, B - 1$ (note that 2^{M_T} is the number of possible transmitted symbol vectors):

$$p_{k,l}[i] = \text{P} [\mathbf{b}_k[i] = \mathbf{s}_l | \hat{\mathbf{z}}[i], \mathbf{h}] \text{ for } \mathbf{s}_l \in \{+1, -1\}^{M_T \times 1}.$$

From (80), we can write this *a posteriori* probability as,

$$p_{k,l}[i] = C_1 \exp \left(-\frac{1}{N_0(\overline{\mathbf{R}}^{-1})_{kk}} \left| \hat{z}_k[i] - \frac{A_k}{\sqrt{M_T}} \mathbf{s}_l^\top \mathbf{h}_k \right|^2 \right),$$

where $(\overline{\mathbf{R}}^{-1})_{kk}$ is the (k, k) -th element of the matrix $\overline{\mathbf{R}}^{-1}$, $\hat{z}_k[i]$ is the k -th component of the vector $\hat{\mathbf{z}}[i]$ and C_1 is a normalizing constant.

The second stage of the partitioned receiver employs a bank of single-user space-time Viterbi decoders that use these *a posteriori* probabilities as inputs. The k -th user's decoder uses only the symbol vector probabilities corresponding to the k -th user. This results in a decentralized implementation of the receiver. Clearly this partitioned receiver is equivalent to a single-user space-time coded system, except for a different noise variance value. This leads to the following upper bound on the pair-wise error probability of the decorrelator-based partitioned space-time multiuser receiver

$$P_e^{k,(d)}[\mathbf{D}_k \rightarrow \hat{\mathbf{D}}_k] \leq \frac{1}{\prod_{n=1}^{r_k} \lambda_{k,n}(\mathbf{E}_k)} \left(\frac{A_k^2/M_T}{4N_0(\overline{\mathbf{R}}^{-1})_{kk}} \right)^{-r_k},$$

where r_k is the rank of the codeword error matrix $\mathbf{E}_k = \mathbf{D}_k - \hat{\mathbf{D}}_k$ and $\lambda_{k,n}(\mathbf{E}_k)$, for $n = 1, \dots, r_k$, are the non-zero eigenvalues of the $M_T \times M_T$ matrix $\mathbf{E}_k^\top \mathbf{E}_k$.

4.3.2 Linear MMSE Based Partitioned Space-time Multiuser Receiver

As is well known, the decorrelator performance degrades when the background noise is dominant, since it completely ignores the presence of background noise [38]. A better compromise between suppressing the multiple-access interference (MAI) and the background noise is obtained by employing a linear MMSE filter at the first stage of the space-time receiver. The linear MMSE multiuser detector output at symbol time i is given by [38]

$$\hat{\mathbf{z}}[i] = \mathbf{A}^{-1}(\overline{\mathbf{R}} + N_0\mathbf{A}^{-2})^{-1}\mathbf{z}[i].$$

The decision statistic corresponding to the k -th user can then be written as,

$$\begin{aligned}
\hat{z}_k[i] &= \frac{A_k}{M_T} \sum_{j=1}^K \mathbf{M}_{kj} A_j \mathbf{b}_j^\top [i] \mathbf{h}_j + \hat{\eta}_k[i] \\
&= \frac{A_k^2}{M_T} \mathbf{M}_{kk} \mathbf{b}_k^\top [i] \mathbf{h}_k + \frac{A_k}{M_T} \sum_{j \neq k} \mathbf{M}_{kj} A_j \mathbf{b}_j^\top [i] \mathbf{h}_j + \hat{\eta}_k[i]
\end{aligned} \tag{81}$$

where we have defined $\mathbf{M} = (\mathbf{A}^2 + N_0 \bar{\mathbf{R}}^{-1})^{-1}$ and $\hat{\eta}_k[i] \sim \mathcal{N}(0, \frac{A_k^2}{M_T} N_0 (\mathbf{M} \bar{\mathbf{R}}^{-1} \mathbf{M})_{kk})$.

In order to compute the soft output *a posteriori* probabilities at the end of the first stage, we make the assumption that the noise at the output of an MMSE multiuser detector (residual MAI plus the background noise) can be modeled as being Gaussian [32]. Therefore, we may model (81) as

$$\hat{z}_k[i] = \frac{A_k^2}{M_T} \mathbf{M}_{kk} \mathbf{b}_k^\top [i] \mathbf{h}_k + \tilde{\eta}_k[i], \tag{82}$$

with $\tilde{\eta}_k[i] \sim \mathcal{N}(0, \nu_k^2[i])$. It can be shown that

$$\nu_k^2[i] = 4 \frac{A_k^2}{M_T} \left[\sum_{j \neq k} \frac{A_j^2}{M_T} \mathbf{M}_{kj}^2 |\mathbf{h}_j[i]|^2 + N_0 (\mathbf{M} \bar{\mathbf{R}}^{-1} \mathbf{M})_{kk} \right]. \tag{83}$$

Using this model, the soft output *a posteriori* probabilities at the output of the linear MMSE multiuser stage can be written as

$$\begin{aligned}
p_{k,l}[i] &= \text{P} [\mathbf{b}_k[i] = \mathbf{s}_l | \hat{\mathbf{z}}[i], \mathbf{h}] \\
&= C_2 \exp \left(-\frac{1}{\nu_k^2[i]} |\hat{z}_k[i] - \frac{A_k^2}{M_T} \mathbf{M}_{kk} \mathbf{s}_l^\top \mathbf{h}_k|^2 \right),
\end{aligned}$$

where C_2 is a normalizing constant.

The second stage of this receiver operates exactly the same way as that in the decorrelator-based partitioned receiver.

Figure 9 shows the FER performance of partitioned space-time multiuser receivers based on linear first stage multiuser detectors and ML single-user decoders, in a four-user system with each having two transmit antennas. As before, we make use of the full diversity BPSK space-time trellis code with

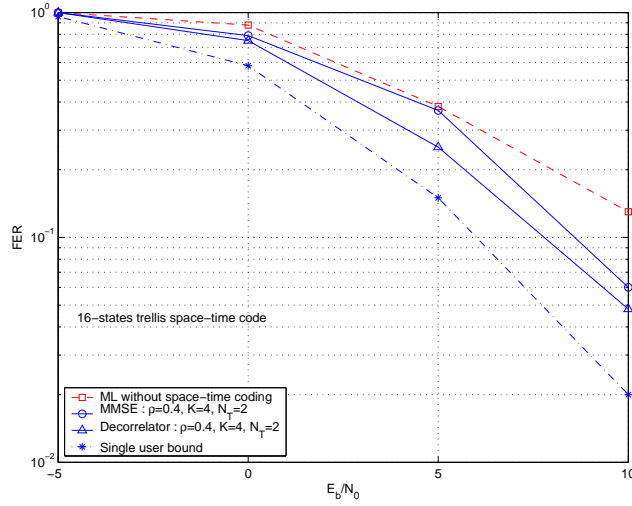


Figure 9: FER Performance Versus E_b/N_0 (in dB) of the Linear First Stage-based Partitioned Space-time Multiuser Detectors. $K = 4$, $\rho = 0.4$ and $M_T = 2$.

constraint length $\nu = 5$ and based on the underlying rate-1/2 convolutional code with octal generators (46, 72) [16]. User cross-correlations are assumed to be $\rho_{jk} = 0.4$ for all $k \neq j$.

From Fig. 9 it can be seen that the linear first stage-based partitioned space-time receivers may offer some diversity gain over single-antenna systems, though they fail to capture the full gains achievable with space-time coding. This is clear from the large performance gap between that of linear first stage-based partitioned receivers and the single-user bound in Fig. 9. This performance degradation becomes severe with increasing user cross-correlations, as one would expect. These results also justify our iterative approach, which is capable of providing near single-user performance even in severe MAI environments (as we will see below). We observe that for the given cross-correlation values, the MMSE first stage performance is no better than that with a decorrelator first stage. Of course in the case of smaller MAI than what we have simulated, the MMSE first stage would out-perform the decorrelator-based receiver, since in this case the background noise would be the dominant noise source. In either case, these linear detectors fail to exploit the large performance gains available with space-time coding.

4.3.3 Iterative MUD with Interference Cancellation for Space-time Coded CDMA

In this section we present a simple iterative receiver structure based on interference cancellation and the turbo principle. Suppose that at the first stage of the receiver, we have available *a priori* probabilities of all users' transmitted symbol vectors, $p_{k,l}[i]_2^p = \text{P}[\mathbf{b}_k[i] = \mathbf{s}_l]$, for $l = 1 \dots, 2^{M_T}$, $k = 1 \dots, K$ and $i = 0 \dots, B - 1$. Note that the subscript 2 and superscript p indicate that these *a priori* probabilities were in fact generated by the second stage of the receiver (i.e. the single-user space-time decoders) at the previous iteration. Using these *a priori* probabilities $p_{k,l}[i]_2^p$, the interference-cancelling multiuser detector at the first stage of the receiver computes soft estimates of the transmitted symbol vectors of all the users as

$$\hat{\mathbf{b}}_k[i] = \sum_{l=1}^{2^{M_T}} \mathbf{s}_l p_{k,l}[i]_2^p. \quad (84)$$

These soft estimates are used to cancel the multiple-access interference at the output of the k -th user's matched filter. The interference cancelled output corresponding to the k -th user is obtained as the k -th component of the vector

$$\hat{\mathbf{z}}_k[i] = \hat{\mathbf{z}}[i] - \bar{\mathbf{R}}\mathbf{A}\hat{\mathbf{D}}_k[i]\mathbf{h}, \quad (85)$$

where $\hat{\mathbf{D}}_k[i] = \text{diag}(\hat{\mathbf{b}}_1[i], \dots, \hat{\mathbf{b}}_{k-1}[i], \mathbf{0}, \hat{\mathbf{b}}_{k+1}[i], \dots, \hat{\mathbf{b}}_K[i])$. From (85), with $\hat{z}_k[i]$ denoting the k -th element of $\hat{\mathbf{z}}_k[i]$, we have that

$$\hat{z}_k[i] = \frac{A_k}{\sqrt{M_T}} \mathbf{b}_k^\top \mathbf{h}_k + \sum_{j \neq k} \rho_{kj} \frac{A_j}{\sqrt{M_T}} (\mathbf{b}_j - \hat{\mathbf{b}}_j)^\top \mathbf{h}_j + \eta_k[i]. \quad (86)$$

Since $\eta_k[i] \sim \mathcal{N}(0, N_0)$, assuming all the previous estimates of the symbol vectors were correct, the iterative interference-cancelling space-time multiuser detector (IC-ST-MUD) computes the soft output *a posteriori* probabilities of the transmitted symbol vectors of user k , for $k = 1, \dots, K$, as

$$\begin{aligned} \text{P} \left[\mathbf{b}_k[i] = \mathbf{s}_l | \mathbf{z}[i], \{\hat{\mathbf{b}}_j\}_{j=1, j \neq k}^K \right] &= C_3 \exp \left[-\frac{1}{N_0} \left| \hat{z}_k[i] - \frac{A_k}{\sqrt{M_T}} \mathbf{s}_l^\top \mathbf{h}_k \right|^2 \right] p_{k,l}[i]_2^p \\ &= p_{k,l}[i]_1 p_{k,l}[i]_2^p, \end{aligned}$$

where C_3 is a normalizing constant.

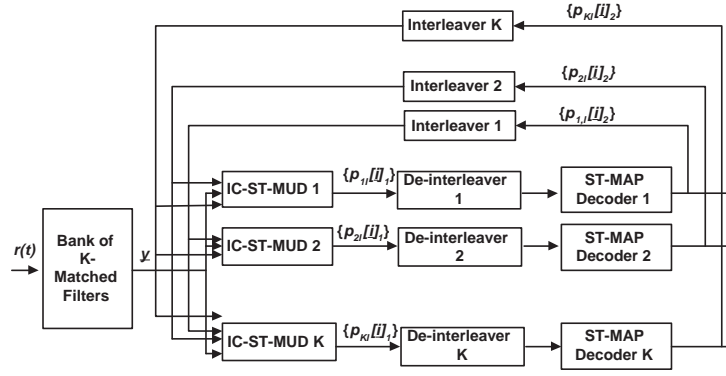


Figure 10: Iterative, Interference-cancelling, Space-time Multiuser Detector.

Following turbo decoding terminology, the term $p_{k,l}[i]_1$ is called the extrinsic *a posteriori* probability as computed by the space-time multiuser detector. These extrinsic *a posteriori* probabilities, $p_{k,l}[i]_1$, are de-interleaved and passed on to a bank of K single-user soft-input/soft-output (SISO) space-time MAP decoders, described in Section 4.4 below (For a more general factor graph interpretation refer to Chapter 5). The k -th user's SISO space-time MAP decoder computes *a posteriori* probabilities of the transmitted symbol vectors for all the symbols in a given frame [44]. The extrinsic component of these symbol vector APPs, $p_{k,l}[i]_2$, are then interleaved and fed back to the first stage of the IC-ST-MUD, to be used as the *a priori* probabilities $p_{k,l}[i]_1^p$, in the next iteration. At the final iteration, the space-time MAP decoders output hard decisions on the information symbols. A block diagram of this iterative, interference-cancelling space-time multiuser detector is shown in Fig. 10.

FER performance of the iterative receiver based on interference cancellation, but without instantaneous linear MMSE filtering, is shown on Fig. 11 for the same four-equal-power-user system in which each user has two antennas considered in Fig. 9. From Fig. 11 we observe that with about four iterations we can achieve most of the gain available from the iterative decoding process. Significantly, we see that for medium values of ρ , this simple interference cancellation scheme can achieve near single-user performance with few iterations, which is not possible with linear first stages as we observed earlier.

However, this simple interference-cancellation-based iterative detector fails when the cross-correlations between users is increased. In this case, the performance becomes almost insensitive to

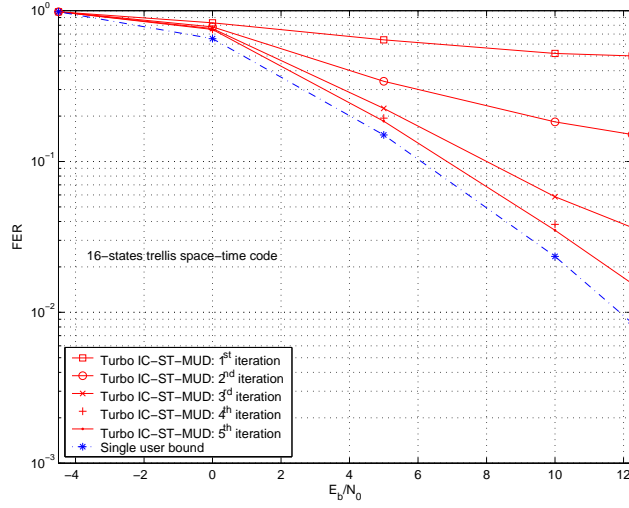


Figure 11: FER Performance Versus E_b/N_0 (in dB) of the Partitioned Iterative Space-time Receiver Based on Interference Cancelling Multiuser Detection. $K = 4$, $\rho = 0.4$ and $M_T = 2$.

the number of iterations since when the user cross-correlations are high our estimates at the end of the initial iteration are very poor (which of course is the same as a system employing a single-user matched filter front end), and thus the subsequent iterations will be based on these poor estimates.

The conventional matched filter complexity is $\mathcal{O}(1)$. At each iteration, the first stage of the receiver needs to compute 2^{M_T} symbol vector *a posteriori* probabilities. Hence, the computational complexity of this partitioned receiver is $\mathcal{O}(2^{M_T} + 2^\nu)$ per user per iteration. Note that even though both MAP and ML decoding have same $\mathcal{O}(2^\nu)$ complexity order, the MAP decoding in general requires more computations compared to the ML decoding. It has been shown that MAP decoding can be done with a complexity roughly four times that of ML decoding [40].

4.3.4 Iterative MUD with Interference Cancellation and Instantaneous MMSE Filtering for Space-time Coded Multiuser Systems

As we mentioned above, the performance of the iterative IC-ST-MUD receiver, proposed in the previous section degrades considerably for medium to large cross-correlation values. Especially when the user cross-correlations are high, the soft estimates at the initial iteration can be very poor and thus the performance does not improve significantly on subsequent iterations. In order to overcome this shortcoming,

in this section we modify the iterative receiver proposed in the previous section by the addition of an instantaneous filter. This becomes similar to the iterative decoder proposed in [44] for a convolutionally coded CDMA channel.

Specifically, we choose a linear MMSE filter that minimize the mean square error between the interference-suppressed output and the k -th user's fading-modulated transmitted symbol vector. Clearly, when the soft estimates of the multiple-access interference are very poor or they are not available at all (as in the case of the first iteration), this filtering helps the receiver to still maintain an acceptable performance level, as we will see by the simulation results.

The k -th user's linear MMSE filter at symbol time i applies weights $\mathbf{w}_k[i]$ to the interference-suppressed output $\hat{\mathbf{z}}_k[i]$ of (85), where $\mathbf{w}_k[i]$ is designed so that,

$$\mathbf{w}_k[i] = \underset{\mathbf{w}}{\operatorname{argmin}} \operatorname{E} \left[\|\mathbf{b}_k^T[i] \mathbf{h}_k - \mathbf{w}^H \hat{\mathbf{z}}_k[i]\|^2 \right]. \quad (87)$$

It can easily be shown that the solution to (87) is given by

$$\mathbf{w}_k[i] = \operatorname{E} \left[\hat{\mathbf{z}}_k[i] \hat{\mathbf{z}}_k^H[i] \right]^{-1} \operatorname{E} \left[\hat{\mathbf{z}}_k[i] \mathbf{b}_k^T[i] \mathbf{h}_k \right], \quad (88)$$

with

$$\operatorname{E} \left[\hat{\mathbf{z}}_k[i] \hat{\mathbf{z}}_k^H[i] \right] = \bar{\mathbf{R}} \mathbf{V}_k[i] \bar{\mathbf{R}} + N_0 \bar{\mathbf{R}},$$

and

$$\operatorname{E} \left[\hat{\mathbf{z}}_k[i] \mathbf{b}_k^T[i] \mathbf{h}_k \right] = \frac{A_k}{\sqrt{M_T}} |\mathbf{h}_k|^2 \bar{\mathbf{R}} \mathbf{e}_k,$$

where we have defined the matrix $\mathbf{V}_k[i]$ as

$$\mathbf{V}_k[i] = \operatorname{diag} \left(\frac{A_1^2}{M_T} \sum_{m=1}^{M_T} (1 - \hat{b}_{1,m}^2) |h_{1,m}|^2, \dots, \frac{A_k^2}{M_T} |\mathbf{h}_k|^2, \dots, \frac{A_K^2}{M_T} \sum_{m=1}^{M_T} (1 - \hat{b}_{K,m}^2) |h_{K,m}|^2 \right),$$

and \mathbf{e}_k is the k -th unit vector. Denoting the matrix $(\bar{\mathbf{R}} \mathbf{V}_k[i] \bar{\mathbf{R}} + N_0 \bar{\mathbf{R}})^{-1}$ by $\mathbf{M}_k[i]$, we can write the instantaneous linear MMSE filter corresponding to the k -th user at symbol time i as

$$\begin{aligned}
\mathbf{w}_k[i] &= \frac{A_k}{\sqrt{M_T}} |\mathbf{h}_k|^2 (\overline{\mathbf{R}} \mathbf{V}_k[i] \overline{\mathbf{R}} + N_0 \overline{\mathbf{R}})^{-1} \overline{\mathbf{R}} \mathbf{e}_k \\
&= \frac{A_k}{\sqrt{M_T}} |\mathbf{h}_k|^2 \mathbf{M}_k[i] \overline{\mathbf{R}} \mathbf{e}_k.
\end{aligned} \tag{89}$$

We again model the residual noise at the linear MMSE filter output as having a Gaussian distribution [32, 44]. Thus, we have the following model for $v_k[i]$, the output of the linear MMSE filter corresponding to the k -th user at symbol time i :

$$v_k[i] = \mathbf{w}_k^H[i] \hat{\mathbf{z}}_k[i] = \mu_k[i] \mathbf{b}_k^T[i] \mathbf{h}_k + u_k[i], \tag{90}$$

where $u_k[i] \sim \mathcal{N}(0, \nu_k^2[i])$. It can be shown that

$$\mu_k[i] = \frac{A_k^2}{M_T} |\mathbf{h}_k|^2 (\mathbf{M}_k[i])_{k,k} \tag{91}$$

and

$$\nu_k^2[i] = |\mathbf{h}_k|^2 (\mu_k[i] - \mu_k^2[i]). \tag{92}$$

The soft-output interference-cancelling multiuser detector with instantaneous MMSE filtering makes use of the model in (90) in order to compute the *a posteriori* probabilities of the transmitted symbol vectors corresponding to the k -th user:

$$\begin{aligned}
\mathbb{P} \left[\mathbf{b}_k[i] = \mathbf{s}_l \mid \mathbf{z}[i], \{\hat{\mathbf{b}}_j\}_{j=1, j \neq k}^K \right] &= C_4 \exp \left[-\frac{|v_k[i] - \mu_k[i] \mathbf{s}_l^T \mathbf{h}_k|^2}{\nu_k^2[i]} \right] p_{k,l}[i]_2^p \\
&= p_{k,l}[i]_1 p_{k,l}[i]_2^p,
\end{aligned}$$

where C_4 is a normalizing constant.

The second stage of this modified iterative receiver is a SISO space-time MAP decoder which operates exactly the same way as the receiver described in the previous section. This decoder is described briefly in the following section.

Figure 12 shows the FER performance of the interference-cancelling space-time multiuser receiver with instantaneous linear MMSE filtering assuming the same four user system but with the cross-correlation between any pair of users being equal to 0.75. We observe that this modified iterative receiver

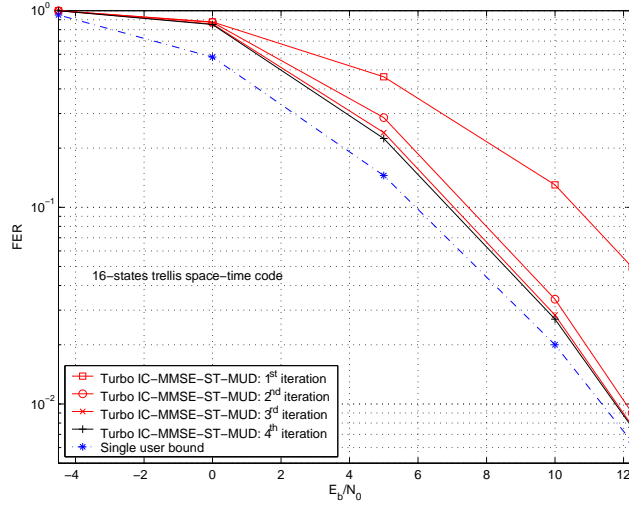


Figure 12: FER Performance Versus E_b/N_0 (in dB) of the Partitioned Iterative Space-time Receiver Based on Interference Cancelling and Linear MMSE Filtering Multiuser Detection Stage. $K = 4$, $\rho = 0.75$ and $M_T = 2$.

provides excellent performance and is able to achieve near single-user performance with only a few iterations (2 – 3 iterations), even in the presence of considerable MAI.

The complexity of this MMSE-based interference cancelling partitioned receiver is roughly about $\mathcal{O}(K^2 + 2^{M_T} + 2^\nu)$ per user per iteration. Note also that this iterative receiver does not rely on spatial diversity for interference suppression but exploits the multiuser signal structure, which is likely to be available at a base station receiver.

4.4 Single-User Soft Input Soft Output Space-time MAP Decoder

In the following we briefly outline the single-user soft-input soft-output space-time MAP decoder assumed in the iterative receivers above. The space-time encoder of each user is assumed to append zero bits to a given information bit block of size B' , so that the trellis is always terminated in the zero state. Thus, the actual space-time code block length is $B = B' + \nu - 1$ (since we assume that the rate of the space-time code is 1), where ν is the constraint length of the underlying convolutional code. In this section, we use the MAP decoding algorithm [2] to compute the *a posteriori* probabilities of all the symbol vectors and the information bits.

Similarly to the notation in [44], we will denote the state of the space-time trellis at time i by a $(\nu - 1)$ -tuple, as $S_i = (s_i^1, \dots, s_i^{\nu-1}) = (d_{i-1}, \dots, d_{i-\nu+1})$, where d_i is the input information bit to the space-time encoder at time i . The corresponding output code symbol vector is denoted by \mathbf{b}_i . (Note that here we are using the subscripts to denote the time index.) Let $d(s', s)$ be the input information bit that causes the state transition from $S_{i-1} = s'$ to $S_i = s$ and $\mathbf{b}(s', s)$ be the corresponding output bit vector, which is of length M_T .

Define the forward and backward recursions [2] as

$$\alpha_i(s) = \sum_{s'} \alpha_{i-1}(s') P[\mathbf{b}_i(s', s)], \quad i = 1, \dots, B, \quad (93)$$

and

$$\beta_i(s) = \sum_{s'} \beta_{i+1}(s') P[\mathbf{b}_{i+1}(s', s)], \quad i = B - 1, \dots, 0, \quad (94)$$

where $P[\mathbf{b}_i(s', s)] = P[\mathbf{b}_i = \mathbf{b}(s', s)]$. Initial conditions for (93) and (94) are given as $\alpha_0(\mathbf{0}) = 1$, $\alpha_0(s \neq \mathbf{0}) = 0$, $\beta_B(\mathbf{0}) = 1$ and $\beta_B(s \neq \mathbf{0}) = 0$. The summations are over all the states s' where the state transition (s', s) is allowed in the code trellis. Normalization of forward and backward variables is done as in [44] to avoid numerical instabilities, though we do not elaborate them here.

Let \mathcal{S}^l denote the set of state pairs (s', s) such that the output symbol vector corresponding to this transition is \mathbf{s}_l . The SISO ST MAP decoder of user k updates the *a posteriori* symbol vector probabilities as

$$\begin{aligned} P[\mathbf{b}_k[i] = \mathbf{s}_l | \{p_{k,l'}[i]_1\}_{i=0}^{B-1}, l' = 1, \dots, L] &= \sum_{(s', s) \in \mathcal{S}^l} \alpha_{i-1}(s') \beta_i(s) P[\mathbf{b}_i(s', s)] \\ &= \left(\sum_{(s', s) \in \mathcal{S}^l} \alpha_{i-1}(s') \beta_i(s) \right) P[\mathbf{b}_k[i] = \mathbf{s}_l] \\ &= p_{k,l}[i]_2 p_{k,l}[i]_1. \end{aligned} \quad (95)$$

The extrinsic part of the above *a posteriori* symbol vector probability, $p_{k,l}[i]_2$, is interleaved and fed back to the interference-cancelling space-time multiuser detector, to be used as the *a priori* probability $p_{k,l}[i]_2^p$, in the next iteration.

In the final iteration the SISO ST MAP decoder also computes the *a posteriori* log-likelihood ratio (LLR) of the information bits. Again, similarly to the notation in [44], let \mathcal{U}^+ denote the set of state pairs (s', s) such that the corresponding input information bit is $+1$. \mathcal{U}^- is defined similarly. Then we have that

$$\begin{aligned}\Lambda[d_k[i]] &= \frac{\text{P}[d_k[i] = +1]}{\text{P}[d_k[i] = -1]} \\ &= \log \frac{\sum_{\mathcal{U}^+} \alpha_{i-1}(s') \beta_i(s) \text{P}[\mathbf{b}_i(s', s)]}{\sum_{\mathcal{U}^-} \alpha_{i-1}(s') \beta_i(s) \text{P}[\mathbf{b}_i(s', s)]}.\end{aligned}$$

Based on these *a posteriori* log-likelihood ratios (LLRs), the decoder outputs a final hard decision on the information bit $d_k[i]$ for $i = 1, \dots, B' - 1$, at the last iteration.

4.5 Summary

In this section, we have considered space-time coding for multiple-access systems in the presence of quasi-static Rayleigh fading. We first obtained the joint ML receiver for a space-time coded CDMA multiuser channel. This joint ML receiver can be shown to achieve full diversity advantage for each user if the individual space-time codes are of full diversity. A better trade-off between performance and computational complexity at the receiver can be obtained by partitioning the multiuser detection and space-time decoding into two stages at the receiver. In particular, a non-linear iterative receiver based on interference cancellation and instantaneous MMSE filtering is capable of capturing most of the gains available with space-time coding in multiple-access channels, with only a few iterations.

5 Adaptive Linear Space-Time Multiuser Detection

We now turn to the situation in which some of the parameters of the model of (1) are not known, and thus the receiver must adapt itself to the environment. To examine this situation, two linear multiuser MIMO reception strategies, diversity and space-time multiuser detection, are presented. Citing advantages of the space-time technique, *linear adaptive* implementations, including batch and sequential-adaptive algorithms for synchronous CDMA in flat fading channels, are then developed. The section concludes with extensions to asynchronous CDMA in multipath fading. Portions of this work first appeared in [35].

5.1 Diversity Multiuser Detection Versus Space-Time Multiuser Detection

We consider a K -user code division multiple access (CDMA) system with processing gain N , operating in flat fading with M_R receiver antennas and M_T transmitter antennas. For simplicity of exposition, we will consider only $M_T = M_R = 2$ and BPSK modulation in this section. Extensions to other antenna configurations and modulation techniques are straightforward. When two antennas are employed at the transmitter, we must first specify how the information symbols are transmitted across the two antennas. Here we adopt the Alamouti space-time block coding scheme [1, 36] discussed in Chapter 1. Specifically, for each user k , two information symbols $b_{k,1}$ and $b_{k,2}$ are transmitted over two symbol intervals. At the first time interval, the symbol pair $(b_{k,1}, b_{k,2})$ is transmitted across the two transmitter antennas; and at the second time interval, the symbol pair $(-b_{k,2}, b_{k,1})$ is transmitted. After chip-matched filtering with respect to $\psi(t)$ and chip-rate sampling, the received signals at antenna 1 during the two symbol intervals are⁴

$$\mathbf{r}_{1,1} = \sum_{k=1}^K [h_{k,1,1}b_{k,1} + h_{k,2,1}b_{k,2}] \mathbf{s}_k + \mathbf{n}_{1,1} \quad (96)$$

and

$$\mathbf{r}_{2,1} = \sum_{k=1}^K [-h_{k,1,1}b_{k,2} + h_{k,2,1}b_{k,1}] \mathbf{s}_k + \mathbf{n}_{2,1}, \quad (97)$$

and the corresponding signals received at antenna 2 are

$$\mathbf{r}_{1,2} = \sum_{k=1}^K [h_{k,1,2}b_{k,1} + h_{k,2,2}b_{k,2}] \mathbf{s}_k + \mathbf{n}_{1,2} \quad (98)$$

and

$$\mathbf{r}_{2,2} = \sum_{k=1}^K [-h_{k,1,2}b_{k,2} + h_{k,2,2}b_{k,1}] \mathbf{s}_k + \mathbf{n}_{2,2}, \quad (99)$$

where $h_{k,i,j}$, $i, j \in \{1, 2\}$ is the complex channel response between transmitter antenna i and receiver antenna j for user k and $\mathbf{s}_k = [c_k^{(0)} \ c_k^{(1)} \ \dots \ c_k^{(N-1)}]^T \in \{\pm 1/\sqrt{N}\}^N$ is the spreading code assigned to user k , as discussed previously in this chapter.

The noise vectors $\mathbf{n}_{1,1}$, $\mathbf{n}_{1,2}$, $\mathbf{n}_{2,1}$, and $\mathbf{n}_{2,2}$ are assumed to be independent and identically distributed (i.i.d) with distribution $\mathcal{N}_c(\mathbf{0}, \sigma^2 \mathbf{I}_N)$.

⁴In this section, we assume complex signaling waveforms and channel coefficients.

5.1.1 Linear Diversity Multiuser Detector

Denote

$$\mathbf{S} \triangleq [\mathbf{s}_1 \cdots \mathbf{s}_K]$$

and

$$\bar{\mathbf{R}} \triangleq \mathbf{S}^T \mathbf{S}.$$

Suppose that user 1 is the user of interest. The combining weights for the linear decorrelating detector [38] for user 1 can be written as

$$\mathbf{w}_1 = \mathbf{S} \bar{\mathbf{R}}^{-1} \mathbf{e}_1, \quad (100)$$

where \mathbf{e}_1 denotes the first unit vector in \mathbb{R}^K . Our first detection strategy, which we call *linear diversity multiuser detection*, applies the linear multiuser detector \mathbf{w}_1 in (100) to each of the four received signals $\mathbf{r}_{1,1}$, $\mathbf{r}_{1,2}$, $\mathbf{r}_{2,1}$, and $\mathbf{r}_{2,2}$ and then performs space-time decoding. Specifically, denote the filter outputs as

$$z_{1,1} \triangleq \mathbf{w}_1^T \mathbf{r}_{1,1} = h_{1,1,1} b_{1,1} + h_{1,2,1} b_{1,2} + u_{1,1}, \quad (101)$$

$$z_{2,1} \triangleq (\mathbf{w}_1^T \mathbf{r}_{2,1})^* = -h_{1,1,1}^* b_{1,2} + h_{1,2,1}^* b_{1,1} + u_{2,1}^*, \quad (102)$$

$$z_{1,2} \triangleq \mathbf{w}_1^T \mathbf{r}_{1,2} = h_{1,1,2} b_{1,1} + h_{1,2,2} b_{1,2} + u_{1,2}, \quad (103)$$

$$z_{2,2} \triangleq (\mathbf{w}_1^T \mathbf{r}_{2,2})^* = -h_{1,1,2}^* b_{1,2} + h_{1,2,2}^* b_{1,1} + u_{2,2}^*, \quad (104)$$

$$\text{with } u_{i,j} \triangleq \mathbf{w}_1^T \mathbf{n}_{i,j} \sim \mathcal{N}_c \left(\mathbf{0}, \frac{\sigma^2}{\eta_1^2} \right), \quad i, j = 1, 2 \quad (105)$$

where $\eta_1^2 \triangleq 1 / [\bar{\mathbf{R}}^{-1}]_{1,1}$.

We define the following quantities:

$$\begin{aligned}\mathbf{z} &\triangleq [z_{1,1} \ z_{2,1} \ z_{1,2} \ z_{2,2}]^T \\ \mathbf{u} &\triangleq [u_{1,1} \ u_{2,1}^* \ u_{1,2} \ u_{2,2}^*]^T \\ \mathbf{h}_{1,1} &\triangleq [h_{1,1,1} \ h_{1,2,1}]^H \\ \bar{\mathbf{h}}_{1,1} &\triangleq [h_{1,2,1} \ -h_{1,1,1}]^T \\ \mathbf{h}_{1,2} &\triangleq [h_{1,1,2} \ h_{1,2,2}]^H \\ \bar{\mathbf{h}}_{1,2} &\triangleq [h_{1,2,2} \ -h_{1,1,2}]^T.\end{aligned}$$

Then (101) - (105) can be written as

$$\mathbf{z} = \underbrace{[\mathbf{h}_{1,1} \ \bar{\mathbf{h}}_{1,1} \ \mathbf{h}_{1,2} \ \bar{\mathbf{h}}_{1,2}]^H}_{\mathbf{H}_1^H} \begin{bmatrix} b_{1,1} \\ b_{2,1} \end{bmatrix} + \mathbf{u}, \quad (106)$$

$$\text{with } \mathbf{u} \sim \mathcal{N}_c\left(\mathbf{0}, \frac{\sigma^2}{\eta_1^2} \cdot \mathbf{I}_4\right). \quad (107)$$

It is readily verified that

$$\mathbf{H}_1 \mathbf{H}_1^H = \begin{bmatrix} E_1 & 0 \\ 0 & E_1 \end{bmatrix}, \quad (108)$$

$$\text{with } E_1 \triangleq |h_{1,1,1}|^2 + |h_{1,1,2}|^2 + |h_{1,2,1}|^2 + |h_{1,2,2}|^2. \quad (109)$$

To form the maximum-likelihood decision statistic, we premultiply \mathbf{z} by \mathbf{H}_1 and obtain

$$\begin{bmatrix} d_{1,1} \\ d_{1,2} \end{bmatrix} \triangleq \mathbf{H}_1 \mathbf{z} = E_1 \begin{bmatrix} b_{1,1} \\ b_{1,2} \end{bmatrix} + \mathbf{v}, \quad (110)$$

$$\text{with } \mathbf{v} \sim \mathcal{N}_c\left(\mathbf{0}, \frac{E_1 \sigma^2}{\eta_1^2} \cdot \mathbf{I}_2\right). \quad (111)$$

The corresponding symbol estimates are given by

$$\begin{bmatrix} \hat{b}_{1,1} \\ \hat{b}_{1,2} \end{bmatrix} = \text{sign} \left(\Re \left\{ \begin{bmatrix} d_{1,1} \\ d_{1,2} \end{bmatrix} \right\} \right). \quad (112)$$

The bit error probability is then given by

$$\begin{aligned} P_1^D(e) &= P\left(\Re\{d_{1,1}\} < 0 \mid b_{1,1} = +1\right) \\ &= P\left[E_1 + \mathcal{N}\left(0, \frac{E_1\sigma^2}{2\eta_1^2}\right) < 0\right] = Q\left(\frac{\sqrt{2E_1}}{\sigma} \cdot \eta_1\right), \end{aligned} \quad (113)$$

which fully exploits the available antenna diversity.

5.1.2 Linear Space-Time Multiuser Detector

Now consider the quantities:

$$\begin{aligned} \tilde{\mathbf{r}} &\triangleq \begin{bmatrix} \mathbf{r}_{1,1} \\ \mathbf{r}_{2,1}^* \\ \mathbf{r}_{1,2} \\ \mathbf{r}_{2,2}^* \end{bmatrix}, \quad \tilde{\mathbf{n}} \triangleq \begin{bmatrix} \mathbf{n}_{1,1} \\ \mathbf{n}_{2,1}^* \\ \mathbf{n}_{1,2} \\ \mathbf{n}_{2,2}^* \end{bmatrix}, \quad \mathbf{h}_k \triangleq \begin{bmatrix} h_{k,1,1} \\ h_{k,2,1}^* \\ h_{k,1,2} \\ h_{k,2,2}^* \end{bmatrix}, \\ \bar{\mathbf{h}}_k &\triangleq \begin{bmatrix} h_{k,2,1} \\ -h_{k,1,1}^* \\ h_{k,2,2} \\ -h_{k,1,2}^* \end{bmatrix}. \end{aligned} \quad (114)$$

Then (96)-(99) may be written as

$$\tilde{\mathbf{r}} = \sum_{k=1}^K \left(b_{k,1} \mathbf{h}_k \otimes \mathbf{s}_k + b_{k,2} \bar{\mathbf{h}}_k \otimes \mathbf{s}_k \right) + \tilde{\mathbf{n}} = \tilde{\mathbf{S}} \mathbf{b} + \tilde{\mathbf{n}}, \quad (115)$$

where

$$\tilde{\mathbf{S}} \triangleq \left[\mathbf{h}_1 \otimes \mathbf{s}_1, \bar{\mathbf{h}}_1 \otimes \mathbf{s}_1, \dots, \mathbf{h}_K \otimes \mathbf{s}_K, \bar{\mathbf{h}}_K \otimes \mathbf{s}_K \right]_{4N \times 2K} \quad (116)$$

$$\mathbf{b} \triangleq \left[b_{1,1} \ b_{1,2} \ b_{2,1} \ b_{2,2} \ \dots \ b_{K,1} \ b_{K,2} \right]^T. \quad (117)$$

Since $\mathbf{h}_k^H \bar{\mathbf{h}}_k = 0$ it is easy to show that the decorrelating detector for detecting the symbol $b_{1,1}$ based on $\tilde{\mathbf{r}}$ is given by

$$\tilde{\mathbf{w}}_{1,1} = \frac{\mathbf{h}_1 \otimes \mathbf{w}_1}{\|\mathbf{h}_1\|^2}, \quad (118)$$

which we call *linear space-time multiuser detection*. Hence the output of the linear space-time detector in this case is given by

$$\tilde{z}_1 = \tilde{\mathbf{w}}_{1,1}^H \tilde{\mathbf{r}} = b_{1,1} + u_1 \quad (119)$$

$$\text{with } u_1 \triangleq \tilde{\mathbf{w}}_{1,1}^H \tilde{\mathbf{n}} \sim \mathcal{N}_c(0, \sigma^2 \|\tilde{\mathbf{w}}_{1,1}\|^2) \quad (120)$$

where

$$\|\tilde{\mathbf{w}}_{1,1}\|^2 = \frac{\|\mathbf{w}_1\|^2}{\|\mathbf{h}_1\|^2} = \frac{1}{E_1 \eta_1^2}. \quad (121)$$

Therefore the probability of error is given by

$$\begin{aligned} P_1^{\text{ST}}(e) &= P\left(\Re\{\tilde{z}_1\} < 0 \mid b_{1,1} = +1\right) \\ &= P\left[1 + \mathcal{N}\left(0, \frac{1}{2E_1 \eta_1^2}\right) < 0\right] = Q\left(\frac{\sqrt{2E_1}}{\sigma} \cdot \eta_1\right). \end{aligned} \quad (122)$$

Comparing (122) with (113) it is seen that when two transmitter antennas and two receiver antennas are employed and the signals are transmitted in the form of space-time block code, then the linear diversity receiver and the linear space-time receiver have identical performance. What, then, are the benefits of the space-time detection technique? They include the following:

1. User capacity for CDMA systems is limited by correlations among composite signature waveforms. This multiple-access interference will tend to decrease as the dimension of the vector space in which the signature waveforms reside increases. The signature waveforms for linear diversity detection are of length N , i.e., they reside in \mathbb{C}^N . Since the received signals are stacked for space-time detection, these signature waveforms reside in \mathbb{C}^{2N} for two transmit and one receive antenna or \mathbb{C}^{4N} for two transmit and two receive antennas. As a result, the space-time structure can support more users than linear diversity detection for a given performance threshold. A specific example of this phenomenon is discussed in Section 5.3.3.
2. For adaptive configurations, linear diversity multiuser detection requires four independent subspace trackers operating simultaneously since the receiver performs detection on each of the four received signals, and each has a different signal subspace. The space-time structure requires only one subspace tracker.

5.2 Adaptive Linear Space-Time Multiuser Detection for Flat Fading CDMA

5.2.1 Signal Model

Motivated by the above discussion, we now discuss *adaptive* space-time multiuser detection algorithms for systems with two transmitter antennas and two receiver antennas. These algorithms are also blind, in the sense that the receiver requires knowledge only of the signature waveform of the user of interest, i.e., neither *a priori* channel knowledge nor the spreading codes of the interfering users are necessary for detection. As before, the Alamouti space-time block code is used for transmission, so that during the first symbol interval of block i , user k transmits $(b_{k,1}[i], b_{k,2}[i])$ from the two transmit antennas. During the second symbol interval, user k transmits $(-b_{k,2}[i], b_{k,1}[i])$. Note that inherent to any blind receiver in multiple transmitter antenna systems is an ambiguity issue. That is, if the same spreading waveform is used for a user at both transmitter antennas, the blind receiver cannot distinguish which symbol is transmitted from which antenna. To resolve this ambiguity, we use two different spreading waveforms for each user, i.e., $\mathbf{s}_{k,j}$, $j \in \{1, 2\}$ is the spreading code for user k for the transmission of symbol $b_{k,j}[i]$. The discrete-time received N -vector at base station antenna 1 during the two symbol periods for block i is

$$\mathbf{r}_{1,1}[i] = \sum_{k=1}^K (h_{k,1,1}b_{k,1}[i]\mathbf{s}_{k,1} + h_{k,2,1}b_{k,2}[i]\mathbf{s}_{k,2}) + \mathbf{n}_{1,1}[i] \quad (123)$$

and

$$\mathbf{r}_{2,1}[i] = \sum_{k=1}^K (-h_{k,1,1}b_{k,2}[i]\mathbf{s}_{k,2} + h_{k,2,1}b_{k,1}[i]\mathbf{s}_{k,1}) + \mathbf{n}_{2,1}[i], \quad (124)$$

and the corresponding signals received at antenna 2 are

$$\mathbf{r}_{1,2}[i] = \sum_{k=1}^K (h_{k,1,2}b_{k,1}[i]\mathbf{s}_{k,1} + h_{k,2,2}b_{k,2}[i]\mathbf{s}_{k,2}) + \mathbf{n}_{1,2}[i] \quad (125)$$

and

$$\mathbf{r}_{2,2}[i] = \sum_{k=1}^K (-h_{k,1,2}b_{k,2}[i]\mathbf{s}_{k,2} + h_{k,2,2}b_{k,1}[i]\mathbf{s}_{k,1}) + \mathbf{n}_{2,2}[i]. \quad (126)$$

We stack the received signal vectors and denote

$$\begin{aligned} \tilde{\mathbf{r}}[i] &\triangleq \begin{bmatrix} \mathbf{r}_{1,1}[i] \\ \mathbf{r}_{2,1}^*[i] \\ \mathbf{r}_{1,2}[i] \\ \mathbf{r}_{2,2}^*[i] \end{bmatrix}, \quad \tilde{\mathbf{n}}[i] \triangleq \begin{bmatrix} \mathbf{n}_{1,1}[i] \\ \mathbf{n}_{2,1}^*[i] \\ \mathbf{n}_{1,2}[i] \\ \mathbf{n}_{2,2}^*[i] \end{bmatrix}, \\ \mathbf{h}_k &\triangleq \begin{bmatrix} h_{k,1,1} \\ h_{k,2,1}^* \\ h_{k,1,2} \\ h_{k,2,2}^* \end{bmatrix}, \quad \bar{\mathbf{h}}_k \triangleq \begin{bmatrix} h_{k,2,1} \\ -h_{k,1,1}^* \\ h_{k,2,2} \\ -h_{k,1,2}^* \end{bmatrix}. \end{aligned} \quad (127)$$

Then we have

$$\tilde{\mathbf{r}}[i] = \sum_{k=1}^K \left(b_{k,1}[i] \mathbf{h}_k \otimes \mathbf{s}_{k,1} + b_{k,2}[i] \bar{\mathbf{h}}_k \otimes \mathbf{s}_{k,2} \right) + \tilde{\mathbf{n}}[i] \quad (128)$$

$$= \tilde{\mathbf{S}} \mathbf{b}[i] + \tilde{\mathbf{n}}[i], \quad (129)$$

where

$$\begin{aligned} \tilde{\mathbf{S}} &\triangleq \left[\mathbf{h}_1 \otimes \mathbf{s}_{1,1}, \bar{\mathbf{h}}_1 \otimes \mathbf{s}_{1,2}, \dots, \mathbf{h}_K \otimes \mathbf{s}_{K,1}, \bar{\mathbf{h}}_K \otimes \mathbf{s}_{K,2} \right]_{4N \times 2K} \\ \mathbf{b}[i] &\triangleq \left[b_{1,1}[i] \ b_{1,2}[i] \ b_{2,1}[i] \ b_{2,2}[i] \ \dots \ b_{K,1}[i] \ b_{K,2}[i] \right]_{2K \times 1}^T \end{aligned}$$

and where \otimes denotes the Kronecker product. The autocorrelation matrix of the stacked signal $\tilde{\mathbf{r}}[i]$, \mathbf{C} , and its eigendecomposition are given by

$$\mathbf{C} = \mathbf{E} [\tilde{\mathbf{r}}[i] \tilde{\mathbf{r}}[i]^H] = \tilde{\mathbf{S}} \tilde{\mathbf{S}}^H + \sigma^2 \mathbf{I}_{4N} \quad (130)$$

$$= \mathbf{U}_s \mathbf{\Lambda}_s \mathbf{U}_s^H + \sigma^2 \mathbf{U}_n \mathbf{U}_n^H, \quad (131)$$

where $\mathbf{\Lambda}_s = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_{2K}\}$ contains the largest $(2K)$ eigenvalues of \mathbf{C} , the columns of \mathbf{U}_s are the corresponding eigenvectors; and the columns of \mathbf{U}_n are the $(4N - 2K)$ eigenvectors corresponding to the smallest eigenvalue σ^2 .

The blind linear space-time MMSE filter for joint suppression of multiple access interference (MAI) and space-time decoding for symbol $\left[\mathbf{b}[i] \right]_1 = b_{1,1}[i]$ is given by the solution to the optimization

problem

$$\mathbf{w}_{1,1} \triangleq \arg \min_{\mathbf{w} \in \mathbb{C}^{4N}} \mathbf{E} \left[|b_{1,1}[i] - \mathbf{w}^H \tilde{\mathbf{r}}[i]|^2 \right]. \quad (132)$$

It has been shown in [43, 46] that a scaled version of the solution can be written in terms of the signal subspace components as

$$\mathbf{w}_{1,1} = \mathbf{U}_s \mathbf{\Lambda}_s^{-1} \mathbf{U}_s^H (\mathbf{h}_1 \otimes \mathbf{s}_{1,1}), \quad (133)$$

and the decision is made according to

$$z_{1,1}[i] = \mathbf{w}_{1,1}^H \tilde{\mathbf{r}}[i], \quad (134)$$

$$\hat{b}_{1,1}[i] = \text{sign} \left[\Re \left(z_{1,1}[i] \right) \right] \quad (\text{coherent detection}), \quad (135)$$

and

$$\hat{\beta}_{1,1}[i] = \text{sign} \left[\Re \left(z_{1,1}[i-1]^* z_{1,1}[i] \right) \right] \quad (\text{differential detection}). \quad (136)$$

Before we address specific batch and sequential adaptive algorithms, we note that these algorithms can also be implemented using linear *group-blind* multiuser detectors [41] which, in contrast to their blind counterparts, are constructed with knowledge of the spreading codes of a subset of the active users. They would be appropriate, for example, in cellular uplink environments in which the receiver has knowledge of the signature waveforms of all of the users in its cell, but not those of interfering users outside the cell. Specifically, we may rewrite (129) as

$$\tilde{\mathbf{r}}[i] = \check{\mathbf{S}} \check{\mathbf{b}}[i] + \bar{\mathbf{S}} \bar{\mathbf{b}}[i] + \tilde{\mathbf{n}}[i], \quad (137)$$

where we have separated the users into two groups. The signature sequences of the known users are the columns of $\check{\mathbf{S}}$. The unknown users' sequences are the columns of $\bar{\mathbf{S}}$. Then the group-blind linear hybrid detector for symbol $b_{1,1}[i]$ is given by [41]

$$\mathbf{w}_{1,1}^{\text{GB}} = \mathbf{U}_s \mathbf{\Lambda}_s^{-1} \mathbf{U}_s^H \check{\mathbf{S}} \left[\check{\mathbf{S}}^H \mathbf{U}_s \mathbf{\Lambda}_s^{-1} \mathbf{U}_s^H \check{\mathbf{S}} \right]^{-1} (\mathbf{h}_1 \otimes \mathbf{s}_{1,1}). \quad (138)$$

This detector offers a significant performance improvement over blind implementations of (133) for environments in which the signature sequences of some of the interfering users are known.

5.2.2 Batch Blind Linear Space-time Multiuser Detection

Implementation of (133) requires knowledge of the signal subspace components and the channel. The subspace components can be estimated blindly from the received signal using the sample autocorrelation matrix of the received signal. In order to obtain an estimate of \mathbf{h}_1 we make use of the orthogonality between the signal and noise subspaces, i.e., the fact that $\mathbf{U}_n^H (\mathbf{h}_1 \otimes \mathbf{s}_{1,1}) = \mathbf{0}$. In particular, we have

$$\begin{aligned}
\hat{\mathbf{h}}_1 &= \arg \min_{\mathbf{h} \in \mathbb{C}^4} \|\mathbf{U}_n^H (\mathbf{h} \otimes \mathbf{s}_{1,1})\|^2 \\
&= \arg \max_{\mathbf{h} \in \mathbb{C}^4} \|\mathbf{U}_s^H (\mathbf{h} \otimes \mathbf{s}_{1,1})\|^2 \\
&= \arg \max_{\mathbf{h} \in \mathbb{C}^4} (\mathbf{h}^H \otimes \mathbf{s}_{1,1}^T) \mathbf{U}_s \mathbf{U}_s^H (\mathbf{h} \otimes \mathbf{s}_{1,1}) \\
&= \arg \max_{\mathbf{h} \in \mathbb{C}^4} \mathbf{h}^H \underbrace{[(\mathbf{I}_4 \otimes \mathbf{s}_{1,1}^T) \mathbf{U}_s \mathbf{U}_s^H (\mathbf{I}_4 \otimes \mathbf{s}_{1,1})]}_{\mathbf{Q}} \mathbf{h} \tag{139} \\
&= \text{principal eigenvector of } \mathbf{Q}. \tag{140}
\end{aligned}$$

In (140), $\hat{\mathbf{h}}_1$ specifies \mathbf{h}_1 up to an arbitrary complex scale factor α , i.e. $\hat{\mathbf{h}}_1 = \alpha \mathbf{h}_1$, but this ambiguity can be circumvented using differential modulation and detection. The following is the summary of a batch blind space-time multiuser detection algorithm for the two transmitter antenna/two receiver antenna configuration. The channel is assumed to be constant for at least the duration of the batch size M .

Algorithm 1 [Batch blind linear space-time multiuser detector – synchronous CDMA, two transmitter antennas and two receiver antennas]

- *Estimate the signal subspace:*

$$\hat{\mathbf{C}} = \frac{1}{M} \sum_{i=0}^{M-1} \tilde{\mathbf{r}}[i] \tilde{\mathbf{r}}[i]^H, \tag{141}$$

$$= \hat{\mathbf{U}}_s \hat{\mathbf{\Lambda}}_s \hat{\mathbf{U}}_s^H + \hat{\mathbf{U}}_n \hat{\mathbf{\Lambda}}_n \hat{\mathbf{U}}_n^H. \tag{142}$$

- *Estimate the channels:*

$$\hat{\mathbf{Q}}_1 = (\mathbf{I}_4 \otimes \mathbf{s}_{1,1}^T) \hat{\mathbf{U}}_s \hat{\mathbf{U}}_s^H (\mathbf{I}_4 \otimes \mathbf{s}_{1,1}), \quad (143)$$

$$\hat{\mathbf{Q}}_2 = (\mathbf{I}_4 \otimes \mathbf{s}_{1,2}^T) \hat{\mathbf{U}}_s \hat{\mathbf{U}}_s^H (\mathbf{I}_4 \otimes \mathbf{s}_{1,2}), \quad (144)$$

$$\hat{\mathbf{h}}_1 = \text{principal eigenvector of } \hat{\mathbf{Q}}_1, \quad (145)$$

$$\hat{\mathbf{h}}_1 = \text{principal eigenvector of } \hat{\mathbf{Q}}_2. \quad (146)$$

- *Form the detectors*

$$\hat{\mathbf{w}}_{1,1} = \hat{\mathbf{U}}_s \hat{\mathbf{\Lambda}}_s^{-1} \hat{\mathbf{U}}_s^H (\hat{\mathbf{h}}_1 \otimes \mathbf{s}_{1,1}), \quad (147)$$

$$\hat{\mathbf{w}}_{1,2} = \hat{\mathbf{U}}_s \hat{\mathbf{\Lambda}}_s^{-1} \hat{\mathbf{U}}_s^H (\hat{\mathbf{h}}_1 \otimes \mathbf{s}_{1,2}). \quad (148)$$

- *Perform differential detection:*

$$z_{1,1}[i] = \hat{\mathbf{w}}_{1,1}^H \tilde{\mathbf{r}}[i], \quad (149)$$

$$z_{1,2}[i] = \hat{\mathbf{w}}_{1,2}^H \tilde{\mathbf{r}}[i], \quad (150)$$

$$\hat{\beta}_{1,1}[i] = \text{sign} \left(\Re \left\{ z_{1,1}[i] z_{1,1}[i-1]^* \right\} \right), \quad (151)$$

$$\hat{\beta}_{1,2}[i] = \text{sign} \left(\Re \left\{ z_{1,2}[i] z_{1,2}[i-1]^* \right\} \right), \quad (152)$$

$$i = 0, \dots, M-1.$$

A batch group-blind space-time multiuser detector algorithm can be implemented with simple modifications to (147) and (148).

5.2.3 Adaptive Blind Linear Space-time Multiuser Detection

To form a sequential blind adaptive receiver, we need adaptive algorithms for sequentially estimating the channel and the signal subspace components \mathbf{U}_s and $\mathbf{\Lambda}_s$. First, we address sequential adaptive channel estimation. Denote by $\mathbf{z}[i]$ the projection of the stacked signal $\tilde{\mathbf{r}}[i]$ onto the noise subspace, i.e.,

$$\mathbf{z}[i] = \tilde{\mathbf{r}}[i] - \mathbf{U}_s \mathbf{U}_s^H \tilde{\mathbf{r}}[i] \quad (153)$$

$$= \mathbf{U}_n \mathbf{U}_n^H \tilde{\mathbf{r}}[i]. \quad (154)$$

Since $\mathbf{z}[i]$ lies in the noise subspace, it is orthogonal to any signal in the signal subspace, and in particular, it is orthogonal to $(\mathbf{h}_1 \otimes \mathbf{s}_{1,1})$. Hence \mathbf{h}_1 is the solution to the following constrained optimization problem:

$$\begin{aligned}
& \min_{\mathbf{h}_1 \in \mathbb{C}^4} \mathbf{E} \left[\|\mathbf{z}[i]^H (\mathbf{h}_1 \otimes \mathbf{s}_{1,1})\|^2 \right] \\
&= \min_{\mathbf{h}_1 \in \mathbb{C}^4} \mathbf{E} \left[\|\mathbf{z}[i]^H (\mathbf{I}_4 \otimes \mathbf{s}_{1,1}) \mathbf{h}_1\|^2 \right] \\
&= \min_{\mathbf{h}_1 \in \mathbb{C}^4} \mathbf{E} \left[\left\| \left[(\mathbf{I}_4 \otimes \mathbf{s}_{1,1}^T) \mathbf{z}[i] \right]^H \mathbf{h}_1 \right\|^2 \right] \quad \text{s.t. } \|\mathbf{h}_1\| = 1. \tag{155}
\end{aligned}$$

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

$$\begin{aligned}
\mathbf{h}_1[i+1] &= \mathbf{h}_1[i], && \text{state equation} \\
0 &= \left[(\mathbf{I}_4 \otimes \mathbf{s}_{1,1}^T) \mathbf{z}[i] \right]^H \mathbf{h}_1[i], && \text{observation equation.}
\end{aligned}$$

The standard Kalman filter can then be applied to the above system as follows. Denote $\mathbf{x}[i] \triangleq (\mathbf{I}_4 \otimes \mathbf{s}_{1,1}^T) \mathbf{z}[i]$.

$$\mathbf{k}[i] = \Sigma[i-1] \mathbf{x}[i] (\mathbf{x}[i]^H \Sigma[i-1] \mathbf{x}[i])^{-1}, \tag{156}$$

$$\begin{aligned}
\mathbf{h}_1[i] &= (\mathbf{h}_1[i-1] - \mathbf{k}[i] (\mathbf{x}[i]^H \mathbf{h}_1[i-1])) / \\
&\quad \|\mathbf{h}_1[i-1] - \mathbf{k}[i] (\mathbf{x}[i]^H \mathbf{h}_1[i-1])\|, \tag{157}
\end{aligned}$$

$$\Sigma[i] = \Sigma[i-1] - \mathbf{k}[i] \mathbf{x}[i]^H \Sigma[i-1]. \tag{158}$$

Once we have obtained channel estimates at block i , we can combine them with estimates of the signal subspace components to form the detector in (133). Subspace tracking algorithms of various complexities exist in the literature. Since we are stacking received signal vectors and subspace tracking complexity increases at least linearly with signal subspace dimension, it is imperative that we choose an algorithm with minimal complexity. The best existing low-complexity algorithm for this purpose appears to be noise-averaged Hermitian-Jacobi fast subspace tracking (NAHJ-FST) [34]. This algorithm has the lowest complexity of any algorithm used for similar purposes and has performed well when used for signal subspace tracking in multipath fading environments. Since the size of \mathbf{U}_s is $4N \times 2K$, the complexity is $40 \cdot 4N \cdot 2K + 3 \cdot 4N + 7.5(2K)^2 + 7 \cdot 2K$ floating point operations per iteration. The algorithm and a multiuser detection application are presented in [34]. The application to the current tracking problem is

straightforward and will not be discussed in detail.

Algorithm 2 [Blind adaptive linear space-time multiuser detector – synchronous CDMA, two transmitter antennas and two receiver antennas]

- Using a suitable signal subspace tracking algorithm, e.g. NAJJ-FST, update the signal subspace components $\mathbf{U}_s[i]$ and $\mathbf{\Lambda}_s[i]$ at each block i .
- Track the channel $\mathbf{h}_1[i]$ and $\bar{\mathbf{h}}_1[i]$ according to the following

$$\mathbf{z}[i] = \bar{\mathbf{r}}[i] - \mathbf{U}_s[i]\mathbf{U}_s[i]^H\bar{\mathbf{r}}[i], \quad (159)$$

$$\mathbf{x}[i] = (\mathbf{I}_4 \otimes \mathbf{s}_{1,1}^T) \mathbf{z}[i], \quad (160)$$

$$\bar{\mathbf{x}}[i] = (\mathbf{I}_4 \otimes \mathbf{s}_{1,2}^T) \mathbf{z}[i], \quad (161)$$

$$\mathbf{k}[i] = \mathbf{\Sigma}[i-1] \mathbf{x}[i] (\mathbf{x}[i]^H \mathbf{\Sigma}[i-1] \mathbf{x}[i])^{-1}, \quad (162)$$

$$\bar{\mathbf{k}}[i] = \bar{\mathbf{\Sigma}}[i-1] \bar{\mathbf{x}}[i] (\bar{\mathbf{x}}[i]^H \bar{\mathbf{\Sigma}}[i-1] \bar{\mathbf{x}}[i])^{-1}, \quad (163)$$

$$\begin{aligned} \mathbf{h}_1[i] &= (\mathbf{h}_1[i-1] - \mathbf{k}[i] (\mathbf{x}[i]^H \mathbf{h}_1[i-1])) / \\ &\quad \|\mathbf{h}_1[i-1] - \mathbf{k}[i] (\mathbf{x}[i]^H \mathbf{h}_1[i-1])\|, \end{aligned} \quad (164)$$

$$\begin{aligned} \bar{\mathbf{h}}_1[i] &= (\bar{\mathbf{h}}_1[i-1] - \bar{\mathbf{k}}[i] (\bar{\mathbf{x}}[i]^H \bar{\mathbf{h}}_1[i-1])) / \\ &\quad \|\bar{\mathbf{h}}_1[i-1] - \bar{\mathbf{k}}[i] (\bar{\mathbf{x}}[i]^H \bar{\mathbf{h}}_1[i-1])\|, \end{aligned} \quad (165)$$

$$\mathbf{\Sigma}[i] = \mathbf{\Sigma}[i-1] - \mathbf{k}[i] \mathbf{x}[i]^H \mathbf{\Sigma}[i-1], \quad (166)$$

$$\bar{\mathbf{\Sigma}}[i] = \bar{\mathbf{\Sigma}}[i-1] - \bar{\mathbf{k}}[i] \bar{\mathbf{x}}[i]^H \bar{\mathbf{\Sigma}}[i-1]. \quad (167)$$

- Form the detectors

$$\hat{\mathbf{w}}_{1,1}[i] = \mathbf{U}_s[i]\mathbf{\Lambda}_s^{-1}[i]\mathbf{U}_s[i]^H (\mathbf{h}_1[i] \otimes \mathbf{s}_{1,1}). \quad (168)$$

$$\hat{\mathbf{w}}_{1,2}[i] = \mathbf{U}_s[i]\mathbf{\Lambda}_s^{-1}[i]\mathbf{U}_s[i]^H (\bar{\mathbf{h}}_1[i] \otimes \mathbf{s}_{1,2}). \quad (169)$$

- Perform differential detection:

$$z_{1,1}[i] = \hat{\mathbf{w}}_{1,1}[i]^H \tilde{\mathbf{r}}[i], \quad (170)$$

$$z_{1,2}[i] = \hat{\mathbf{w}}_{1,2}[i]^H \tilde{\mathbf{r}}[i], \quad (171)$$

$$\hat{\beta}_{1,1}[i] = \text{sign} \left(\Re \left\{ z_{1,1}[i] z_{1,1}[i-1]^* \right\} \right), \quad (172)$$

$$\hat{\beta}_{1,2}[i] = \text{sign} \left(\Re \left\{ z_{1,2}[i] z_{1,2}[i-1]^* \right\} \right). \quad (173)$$

A group-blind sequential adaptive space-time multiuser detector can be implemented similarly. The adaptive receiver structure is illustrated in Figure 13.

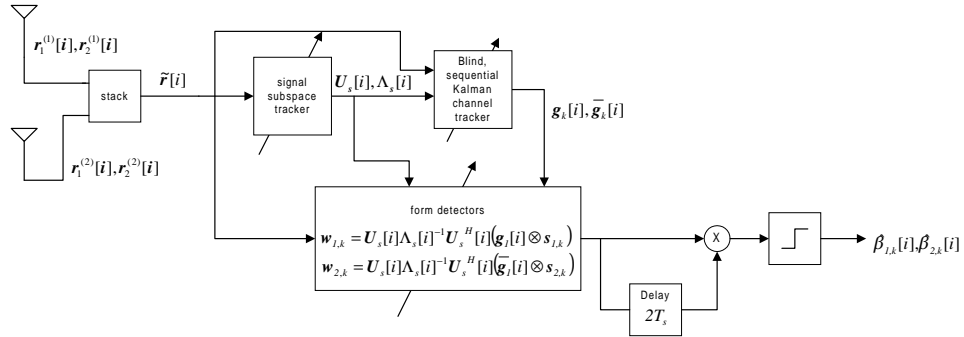


Figure 13: Adaptive receiver structure for linear space-time multiuser detectors.

5.3 Blind Adaptive Space-Time Multiuser Detection for Asynchronous CDMA in Fading Multipath Channels

5.3.1 Signal Model

To extend the previous development to asynchronous multipath channels, we must begin with a continuous-time baseband signal model. The signal transmitted from antennas 1 and 2 due to the k -th user for time interval $i \in \{0, 1, \dots, M-1\}$ is given by

$$x_{k,1}(t) = \sum_{i=0}^{M-1} \left[b_{k,1}[i] s_{k,1}(t - 2iT_s) - b_{k,2}[i] s_{k,2}(t - (2i+1)T_s) \right] \quad (174)$$

$$x_{k,2}(t) = \sum_{i=0}^{M-1} \left[b_{k,2}[i] s_{k,2}(t - 2iT_s) + b_{k,1}[i] s_{k,1}(t - (2i+1)T_s) \right] \quad (175)$$

where M denotes the length of the data frame, T_s denotes the information symbol interval, and $\{b_k[i]\}_i$ is the symbol stream of user k . Although this is an asynchronous system, we have, for notational simplicity, suppressed the delay associated with each user's transmitted signal and incorporated it into the path delays in (3). We assume that for each k , the symbol stream, $\{b_k[i]\}_i$, is a collection of independent random variables that take on values of $+1$ and -1 with equal probability. Furthermore, we assume that the symbol streams of different users are independent. The transmitted signature waveforms $\{s_{k,m}(t)\}$ are described in (26). The k -th user's space-time coded signals, $x_{k,1}(t)$ and $x_{k,2}(t)$, propagate from transmitter to receiver through the multipath fading channel described by (3), where $\tau_{k,m,p,l}$, satisfying $\tau_{k,m,p,1} \leq \tau_{k,m,p,2} \leq \dots \leq \tau_{k,m,p,L}$, is the sum of the corresponding path delay and the initial transmission delay of user k . It is assumed that the channel is slowly varying, so that the path gains and delays remain constant over the duration of one signal frame (MT_s).

The received signal component due to the transmission of $x_{k,1}(t)$ and $x_{k,2}(t)$ through the channel at receiver antennas 1 and 2 is given by

$$y_{k,1}(t) = x_{k,1}(t) \star h_{k,1,1}(t) + x_{k,2}(t) \star h_{k,2,1}(t), \quad (176)$$

$$y_{k,2}(t) = x_{k,1}(t) \star h_{k,1,2}(t) + x_{k,2}(t) \star h_{k,2,2}(t). \quad (177)$$

The total received signal at receiver antenna $b \in \{1, 2\}$ is given by

$$r_b(t) = \sum_{k=1}^K y_{k,b}(t) + n_b(t). \quad (178)$$

At the receiver, the received signal is match filtered to the chip waveform and sampled at the chip rate, i.e., the sampling interval is T_c , N is the total number of samples per symbol interval, and $2N$ is the total number of samples per time slot. The n -th matched filter output during the i -th time slot is given by

$$r_b[i, n] \triangleq \int_{2iT_s+nT_c}^{2iT_s+(n+1)T_c} r_b(t) \psi(t - 2iT_s - nT_c) dt. \quad (179)$$

Denote the maximum delay (in symbol intervals) by

$$\iota_{k,m,p} \triangleq \left\lceil \frac{\tau_{k,m,p,L} + T_c}{T_s} \right\rceil \quad \text{and} \quad \iota \triangleq \max_{k,m,p} \iota_{k,m,p}. \quad (180)$$

Closed form expressions for the matched filter outputs $r_b[i, n]$ are provided in [35].

To fully exploit available diversity, we stack the matched-filter outputs from both receive antennas, forming the vector

$$\underline{r}[i] \triangleq \begin{bmatrix} r_1[i] \\ r_2[i] \end{bmatrix}_{4N \times 1}, \quad (181)$$

where, for $b \in \{1, 2\}$,

$$\underline{r}_b[i] \triangleq \begin{bmatrix} r_b[i, 0] \\ \vdots \\ r_b[i, 2N - 1] \end{bmatrix}_{2N \times 1}. \quad (182)$$

Stacking \bar{m} successive sample vectors, we form

$$\mathbf{r}[i] \triangleq \begin{bmatrix} \underline{r}[i] \\ \vdots \\ \underline{r}[i + m - 1] \end{bmatrix}_{4N\bar{m} \times 1} \quad (183)$$

$$= \mathbf{H}\mathbf{b}[i] + \mathbf{n}[i] \quad (184)$$

where \mathbf{H} is a function of the spreading codes, the channel conditions, and the chip waveform (see [35] for details), $\mathbf{n}[i]$ is additive white Gaussian noise, and where

$$\mathbf{b}[i] \triangleq \begin{bmatrix} \underline{b}[i - \lceil \iota/2 \rceil] \\ \vdots \\ \underline{b}[i + m - 1] \end{bmatrix}_{r \times 1}, \quad \underline{b}[i] \triangleq \begin{bmatrix} b_{1,1}[i] \\ \vdots \\ b_{K,1}[i] \\ b_{1,2}[i] \\ \vdots \\ b_{K,2}[i] \end{bmatrix}_{2K \times 1}, \quad (185)$$

and $r \triangleq 2K(\bar{m} + \lceil \iota/2 \rceil)$.

We will see in Section 5.3.3 that the smoothing factor, \bar{m} , is chosen such that

$$\bar{m} \geq \left\lceil \frac{N(\iota + 1) + K\lceil \iota/2 \rceil + 1}{2N - K} \right\rceil \quad (186)$$

for channel identifiability. Note that the columns of \mathbf{H} (the composite signature vectors) contain information about both the timings and the complex path gains of the multipath channel of each user. Hence an estimate of these waveforms eliminates the need for separate estimates of the timing information $\{\tau_{k,m,p,l}\}$.

5.3.2 Blind MMSE Space-Time Multiuser Detection

Since the ambient noise is white, i.e., $\mathbf{E}[\mathbf{n}[i]\mathbf{n}[i]^H] = \sigma^2\mathbf{I}_{4N\bar{m}}$, the autocorrelation matrix of the received signal in (184) is

$$\mathbf{C}_r \triangleq \mathbf{E}[\mathbf{r}[i]\mathbf{r}[i]^H] = \mathbf{H}\mathbf{H}^H + \sigma^2\mathbf{I}_{4N\bar{m}} \quad (187)$$

$$= \mathbf{U}_s\mathbf{\Lambda}_s\mathbf{U}_s^H + \sigma^2\mathbf{U}_n\mathbf{U}_n^H, \quad (188)$$

where (188) is the eigendecomposition of \mathbf{C}_r . Note that \mathbf{U}_s has size $4N\bar{m} \times r$ and \mathbf{U}_n has size $4N\bar{m} \times (4N\bar{m} - r)$.

The joint MMSE multiuser detector and space-time decoder with corresponding symbol estimate for $b_{k,a}[i]$, $a \in \{1, 2\}$ are given by

$$\mathbf{w}_{k,a}[i] \triangleq \arg \min_{\mathbf{w} \in \mathbb{C}^{4P\bar{m}}} \mathbf{E} \left[|b_{k,a}[i] - \mathbf{w}^H \mathbf{r}[i]|^2 \right], \quad (189)$$

$$\hat{b}_{k,a}[i] = \text{sign} \left[\text{Re} \left\{ \mathbf{w}_{k,a}[i]^H \mathbf{r}[i] \right\} \right]. \quad (190)$$

The solution to (189) can be written in terms of the signal subspace components as [42]

$$\mathbf{w}_{k,a}[i] = \mathbf{U}_s\mathbf{\Lambda}_s^{-1}\mathbf{U}_s^H\mathbf{h}_{k,a}, \quad (191)$$

where $\mathbf{h}_{k,a} \triangleq \mathbf{H}\mathbf{e}_{K(2\lceil \ell/2 \rceil + a - 1) + k}$ is the composite signature waveform of user k for symbol $a \in \{1, 2\}$. As for the synchronous case, this detector can be implemented in blind mode, requiring knowledge only of the signature sequence of the user of interest and a (blind) estimate of the channel.

5.3.3 Blind Sequential Kalman Channel Estimation

The full details of the discrete-time channel model for the asynchronous multipath case appear in [35]. In summary, the composite signature waveform of user k for symbol a can be written as

$$\mathbf{h}_{k,a} = \bar{\mathbf{C}}_{k,a}\mathbf{f}_{k,a} \quad (192)$$

where $\overline{\mathbf{C}}_{k,a}$ is a matrix of size $4N(\lceil \iota/2 \rceil + 1) \times (2N(\iota + 1) + 2)$ that is constructed from the a -th spreading code assigned to user k . The vector $\mathbf{f}_{k,a}$, with size $(2N(\iota + 1) + 2) \times 1$, is a function of the channel state information for user k and is also defined in [35]. The blind channel estimation problem involves the estimation of $\mathbf{f}_{k,a}$ ($1 \leq k \leq K, a = 1, 2$) from the received signal $\mathbf{r}[i]$. As we did for the synchronous case, we will exploit the orthogonality between the signal subspace and noise subspace. Specifically, since \mathbf{U}_n is orthogonal to the column space of \mathbf{H} , we have

$$\mathbf{U}_n^H \mathbf{h}_{k,a} = \mathbf{U}_n^H \overline{\mathbf{C}}_{k,a} \mathbf{f}_{k,a} = \mathbf{0}. \quad (193)$$

Denote by $\mathbf{z}[i]$ the projection of the received signal $\mathbf{r}[i]$ onto the noise subspace, i.e.,

$$\mathbf{z}[i] = \mathbf{r}[i] - \mathbf{U}_s \mathbf{U}_s^H \mathbf{r}[i] \quad (194)$$

$$= \mathbf{U}_n \mathbf{U}_n^H \mathbf{r}[i]. \quad (195)$$

Using (193) we have

$$\mathbf{f}_{k,a}^H \overline{\mathbf{C}}_{k,a}^H \mathbf{z}[i] = 0. \quad (196)$$

Our channel estimation problem, then, involves the solution of the optimization problem

$$\hat{\mathbf{f}}_{k,a} = \arg \min_{\mathbf{f}} \mathbf{E} \left[\left| \mathbf{f}^H \overline{\mathbf{C}}_{k,a}^H \mathbf{z}[i] \right|^2 \right] \quad (197)$$

subject to the constraint $\|\mathbf{f}\| = 1$. If we denote $\mathbf{x}[i] \triangleq \overline{\mathbf{C}}_{k,a}^H \mathbf{z}[i]$ then we can use the Kalman-type algorithm described in (156)-(158) where $\mathbf{h}_1[i]$ is replaced with $\mathbf{f}_{k,a}[i]$.

Note that a necessary condition for the channel estimate to be unique is that the matrix $\mathbf{U}_n^H \overline{\mathbf{C}}_{k,a}$ is tall, i.e. $4N\bar{m} - 2K(\bar{m} + \lceil \iota/2 \rceil) \geq 2N(\iota + 1) + 2$. Therefore we choose the smoothing factor, \bar{m} , such that

$$\bar{m} \geq \left\lceil \frac{N(\iota + 1) + K\lceil \iota/2 \rceil + 1}{2N - K} \right\rceil. \quad (198)$$

Using the same constraint, we find that for a fixed m , the maximum number of users that can be supported is

$$\min \left\{ \left\lfloor \frac{N(2\bar{m} - \iota - 1) - 1}{\bar{m} + \lceil \iota/2 \rceil} \right\rfloor, \left\lfloor \frac{N}{2} \right\rfloor \right\}. \quad (199)$$

Notice that for reasonable choices of \bar{m} and ι , (199) is larger than the maximum number of users for the linear diversity receiver structure, given by

$$\left\lfloor \frac{N(\bar{m} - \iota)}{2(\bar{m} + \iota)} \right\rfloor. \quad (200)$$

This represents a quantitative example of the user capacity benefit of space-time multiuser detection discussed in Section 5.1.2.

Once an estimate of the channel state, $\hat{\mathbf{f}}_{k,a}$, is obtained, the composite signature vector of the k -th user for symbol a is given by (192). Note that there is an arbitrary phase ambiguity in the estimated channel state, which necessitates differential encoding and decoding of the transmitted data.

5.3.4 Algorithm Summary

Algorithm 3 [Blind adaptive linear space-time multiuser detector – asynchronous multipath CDMA, two transmitter antennas and two receiver antennas]

- Stack matched filter outputs in (179) to create $\mathbf{r}[i]$.
- Create $\bar{\mathbf{C}}_{k,a}$.
- Using a suitable signal subspace tracking algorithm, e.g. NAJJ-FST, update the signal subspace components $\mathbf{U}_s[i]$ and $\mathbf{\Lambda}_s[i]$ at each time slot i .
- Track the channel $\mathbf{f}_{k,a}$ ($1 \leq k \leq K, a = 1, 2$) according to the following

$$\mathbf{z}[i] = \mathbf{r}[i] - \mathbf{U}_s[i]\mathbf{U}_s[i]^H\mathbf{r}[i], \quad (201)$$

$$\mathbf{x}[i] = \bar{\mathbf{C}}_{k,a}^H\mathbf{z}[i], \quad (202)$$

$$\mathbf{k}[i] = \mathbf{\Sigma}[i-1]\mathbf{x}[i] (\mathbf{x}[i]^H\mathbf{\Sigma}[i-1]\mathbf{x}[i])^{-1}, \quad (203)$$

$$\mathbf{f}_{k,a}[i] = (\mathbf{f}_{k,a}[i-1] - \mathbf{k}[i] (\mathbf{x}[i]^H\mathbf{f}_{k,a}[i-1])) / \|\mathbf{f}_{k,a}[i-1] - \mathbf{k}[i] (\mathbf{x}[i]^H\mathbf{f}_{k,a}[i-1])\|, \quad (204)$$

$$\mathbf{\Sigma}[i] = \mathbf{\Sigma}[i-1] - \mathbf{k}[i]\mathbf{x}[i]^H\mathbf{\Sigma}[i-1]. \quad (205)$$

- Form the detectors

$$\mathbf{w}_{k,a}[i] = \mathbf{U}_s[i] \mathbf{\Lambda}_s^{-1}[i] \mathbf{U}_s[i]^H \overline{\mathbf{C}}_{k,a} \mathbf{f}_{k,a}[i]. \quad (206)$$

- Perform differential detection:

$$z_{k,a}[i] = \mathbf{w}_{k,a}[i]^H \mathbf{r}[i], \quad (207)$$

$$\hat{\beta}_{k,a}[i] = \text{sign} \left(\Re \left\{ z_{k,a}[i] z_{k,a}[i-1]^* \right\} \right). \quad (208)$$

5.4 Simulation Results

In this section, we present simulation results to illustrate the performance of blind adaptive space-time multiuser detection. We first look at the synchronous flat-fading case; then we consider the asynchronous multipath-fading scenario. For all simulations we use the two transmit/two receive antenna configuration. m -sequences of length 15 and their shifted versions are employed as user spreading sequences. The chip pulse is a raised cosine with roll-off factor 0.5. For the multipath case, each user has $L = 3$ paths. The delay of each path is uniformly distributed on $[0, T_s]$. Hence, the maximum delay spread is one symbol interval, i.e., $\iota = 1$. The fading gain for each user's channel is generated from a complex Gaussian distribution and is fixed for all simulations. The path gains in each users' channel are normalized so that all users' signals arrive at the receiver with the same power. The smoothing factor is $\bar{m} = 2$ and the forgetting factor for the subspace tracking algorithm for all simulations is 0.995. The performance measures are bit-error probability and signal-to-interference-plus-noise ratio, defined by $\text{SINR} \triangleq E^2\{\mathbf{w}^H \mathbf{r}\} / \text{Var}\{\mathbf{w}^H \mathbf{r}\}$, where the expectation is with respect to the data symbols of interfering users and the ambient noise. In the simulations, the expectation operation is replaced by time averaging. SINR is a particularly appropriate figure of merit for MMSE detectors since it has been shown [32] that the output of an MMSE detector is approximately Gaussian distributed. Hence, the SINR values (approximately) translate directly and simply to bit-error probabilities, i.e., $\Pr(e) \approx Q\left(\sqrt{\text{SINR}}\right)$. The labelled horizontal lines on the SINR plot represent bit-error-probability thresholds. For the SINR plots, the number of users for the first 1500 iterations is 4. At iteration 1501, 3 users are added so that the system is fully loaded. At iteration 3001, 5 users are removed.

Figure 14 illustrates the adaptation performance for the synchronous, flat fading case. The SNR is fixed at 8dB. Figure 15 shows the adaptation performance for the asynchronous multipath case. The SNR for this simulation is 11dB. Notice that in both cases the bit-error-probability does not drop below tolerable levels even during transitions, when users enter or leave the system. Convergence of the SINR to its maximum value is almost instantaneous when users leave the system, and requires less than 500 iterations when users are added to the system.

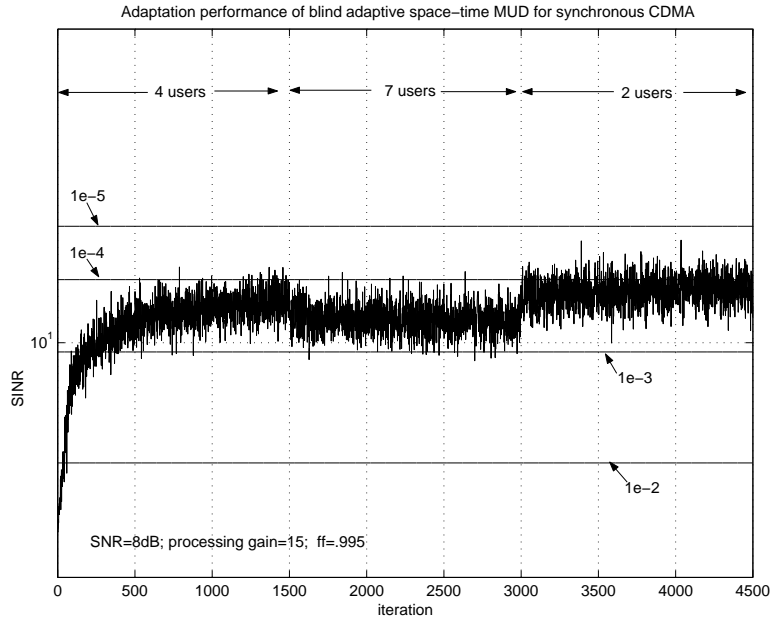


Figure 14: Adaptation performance of space-time multiuser detection for synchronous CDMA. The labelled horizontal lines represent bit-error-probability thresholds.

6 Summary

In this chapter, we have taken the work of the preceding chapters in several directions. In Section 2, we introduced a general model for multiple-access signaling in MIMO channels, and used this model to derive canonical receiver structures for multiuser MIMO systems. This development ties the MIMO multiuser channel models discussed in Chapter 2 together with receiver designs described in Chapters 3 and 5, and then extends the latter to the multiple-access, frequency-selective channels arising in many applications.

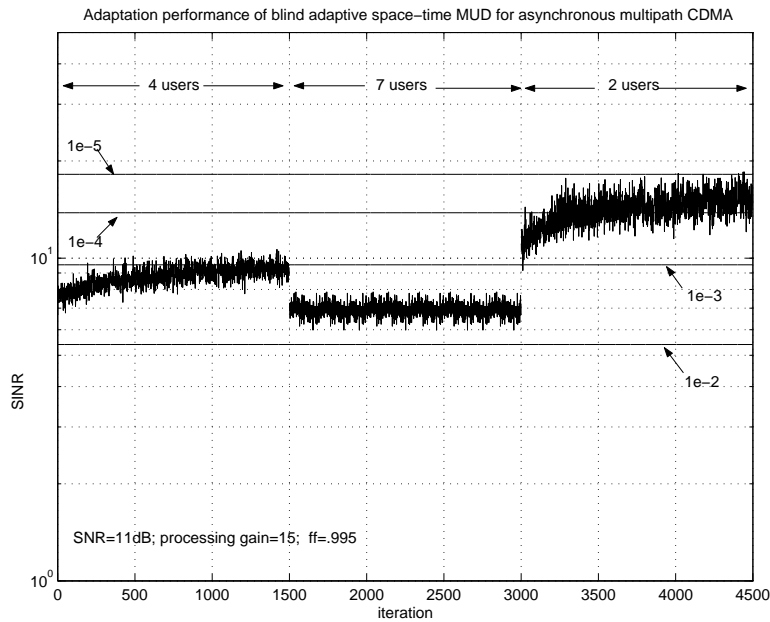


Figure 15: Adaptation performance of space-time multiuser detection for asynchronous multipath CDMA. The labelled horizontal lines represent bit-error-probability thresholds.

Section 3 also echoes the detection problems discussed in Chapters 3 and 5, notably by re-emphasizing the importance of iterative algorithms in complexity reduction for MIMO receivers. Section 4 describes how the structure imposed by space-time coding techniques of Chapter 4 can be exploited, together with the turbo-style iterative methods of Chapter 5, can be used to significantly enhance the overall receiver performance with little attendant increase in complexity. Although most of the work of the techniques in this chapter apply to general interference-type channels (multi-access, inter-symbol, and inter-antenna), the focus has been on the direct-sequence CDMA channels introduced in Chapter 1. Section 5 specifically deals with such channels, which are particularly amenable to adaptive implementation. Moreover, Section 5 exploits the Alamouti space-time coding structure described in Chapters 1 and 4 in its adaptation algorithms.

7 Bibliographic Notes

As noted in Section 2, the methods discussed in this chapter have been developed over a period of several decades. Early work on receiver design for channel-coded systems and intersymbol-interference channels dates from the 1960s and 1970s, respectively, while the techniques for multiple-access and inter-antenna interference channels began largely in the 1980s and 1990s, respectively. A review of these developments is found in [30]. Complexity reduction through iterative algorithms and through adaptation have been major issues throughout this development, with turbo style algorithms gaining significant interest in the 1990s. (An overview of iterative techniques is found in [31].) The current decade has seen a number of developments, particularly in the development of new analytical tools using methods of statistical physics, and in refinement, analysis and understanding of adaptive and iterative methods. However, all of these areas are still areas of active research, and new developments continue today. Perhaps the most critical open issue lies in the transition of these methods into more widespread practice. Although current wireless standards and systems do incorporate some of the ideas exposed in this chapter, there is still considerable opportunity for further practical development. The iterative and adaptive methods are, of course, directed at precisely this goal.

For further additional reading on the subject matter of this chapter, the reader is referred to the books by Verdú [38], Wang and Poor [46], and Comaniciu, et al. [5]. The first of these three books contains an excellent exposition of the fundamentals of multiuser detection, while the second contains further elaboration and additional examples illustrating the model of Section 2, as well as considerable discussion of various methods of adaptive and iterative receiver design. Issues not treated in this chapter, such as fast fading and OFDM systems, are also considered there. Finally, the impact of these methods on higher-layer networking issues, such as resource allocation, quality-of-service provision, and network performance, is discussed in the third of these three books.

References

- [1] S. M. Alamouti, *A simple transmit diversity technique for wireless communications*, IEEE Journ. Select. Areas. Commun. **16** (1998), no. 8, 1451 – 1458.
- [2] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, *Optimum decoding of linear codes for minimizing symbol error rate*, IEEE Trans. Inform. Theory **20** (1974), no. 3, 284 – 287.
- [3] C. Berrou and A. Glavieux, *Near optimum error-correcting coding and decoding: Turbo codes*, IEEE Trans. Commun. **44** (1996), no. 10, 1261 – 1271.
- [4] C. Berrou, A. Glavieux, and P. Thitimajshima, *Near Shannon limit error-correcting coding and decoding: Turbo codes*, Proc. 1993 Int. Conf. Commun. (Geneva, Switzerland), vol. 2, 1993, pp. 1064–1070.
- [5] C. Comaniciu, N. Mandayam, and H. V. Poor, *Wireless networks: Multiuser detection in cross-layer design.*, Springer, New York, NY, 2005.
- [6] H. Dai, A. F. Molisch, and H. V. Poor, *Downlink capacity of interference-limited mimo systems with joint detection*, IEEE Trans. Wireless Comm. **3** (2004), no. 2, 442–453.
- [7] H. Dai and H. V. Poor, *Sample-by-sample adaptive space-time processing for multiuser detection in multipath CDMA systems*, Proc. 2001 Fall IEEE Vehicular Technology Conference (Atlantic City, NJ), Oct. 2001.
- [8] ———, *Iterative space-time processing for multiuser detection in multipath cdma channels*, IEEE Trans. Sig. Proc. **50** (2002), no. 9, 2116–2127.
- [9] M. O. Damen, A. Safavi, and K. Abed-Meriam, *On CDMA with space-time codes over multipath fading channels*, IEEE Trans. Wireless Commun. **2** (2003), 11–19.
- [10] A. P. Dempster, *Maximum likelihood from incomplete data via the EM algorithm*, 1977, pp. 1–38.
- [11] J. A. Fessler and A. O. Hero, *Space-alternating generalized EM algorithm*, IEEE Trans. Signal Processing **42** (1994), no. 10, 2664–2677.

- [12] T. R. Giallorenzi and S. G. Wilson, *Multiuser ML sequence estimator for convolutionally coded asynchronous DS-CDMA systems*, IEEE Trans. Commun. **44** (1996), no. 8, 997–1008.
- [13] ———, *Suboptimum multiuser receivers for convolutionally coded asynchronous DS-CDMA systems*, IEEE Trans. Commun. **44** (1996), no. 9, 1183–1196.
- [14] G. H. Golub and C. F. Van Loan, *Matrix computation*, The John Hopkins University Press, Baltimore, MD, 1996.
- [15] J.-C. Guey, M. P. Fitz, M. R. Bell, and W.-Y. Kuo, *Signal design for transmitter diversity wireless communication systems over Rayleigh fading channels*, Proc. IEEE Veh. Technol. Conf. (Atlanta, GA), 1996, pp. 136–140.
- [16] A. R. Hammons and H. El Gamal, *On the theory of space-time codes for PSK modulation*, IEEE Trans. Inform. Theory **46** (2000), no. 2, 524 – 542.
- [17] B. Hochwald, T. L. Marzetta, and C. B. Papadias, *A novel space-time spreading scheme for wireless CDMA systems*, Proc. 37th Ann. Allerton Conf. Commun., Contr., Comput. (Monticello, IL), Sep. 22-24, 1999.
- [18] ———, *A transmitter diversity scheme for wideband CDMA systems based on space-time spreading*, IEEE Journ. Select. Areas. Commun. **19** (2001), 48–60.
- [19] J. Hou, J. E. Smee, H. D. Pfister, and S. Tomasin, *Implementing interference cancellation to increase the EV-DO REV A link capacity*, IEEE Communications Magazine **44** (2006), no. 2, 96–102.
- [20] S. K. Jayaweera and H. V. Poor, *Iterative multiuser detection for space-time coded synchronous CDMA*, Proc. IEEE Veh. Technol. Conf. (Atlantic City, NJ), vol. 4, Fall 2001, pp. 2736–2739.
- [21] ———, *Low complexity receiver structures for space-time coded multiple-access systems*, EURASIP Journ. Applied Signal Processing (Special Issue on Space-time Coding) **2002** (2002), 275–288.

- [22] P. Lancaster and M. Tismenetsky, *The theory of matrices with applications*, Academic Press, Inc., Orlando, FL, 1985.
- [23] H. Li, X. Lu, and G. B. Giannakis, *Capon multiuser receiver for CDMA systems with space-time coding*, IEEE Trans. Sig. Proc. **50** (2002), 1193 – 1204.
- [24] J. Liu, J. Li, H. Li, and E. G. Larsson, *Differential space-time modulation for interference suppression*, IEEE Trans. Sig. Proc. **49** (2001), 1786–1795.
- [25] Z. Liu, G. B. Giannakis, B. Muquet, and S. Zhou, *Space-time coding for broadband wireless communications*, Wireless Syst. Mobile Comput. **1** (2001), 35–53.
- [26] B. Lu and X. Wang, *Iterative receivers for multiuser space-time coding systems*, IEEE Journ. Select. Areas. Commun. **18** (2000), no. 11, 2322 – 2335.
- [27] A. Naguib and N. Seshadri, *Combined interference cancellation and ML decoding of space-time block codes*, Proc. 7-th Commun. theory mini-conference at Globecom'98 (Sydney, Australia), 1998.
- [28] I. Oppermann, *CDMA space-time coding using an LMMSE receiver*, Proc. Int. Conf. Commun. (ICC'99) (Vancouver, BC, Canada), 1999, pp. 182–186.
- [29] H. V. Poor, *An introduction to signal detection and estimation*, Springer-Verlag, New York, 1994.
- [30] ———, *Dynamic programming in digital communications: Viterbi decoding to turbo multiuser detection*, Journal of Optimization Theory and Applications **115** (2002), no. 3, 629 – 657.
- [31] ———, *Iterative multiuser detection*, IEEE Signal Processing Magazine **21** (2004), no. 1, 81 – 88.
- [32] H. V. Poor and S. Verdú, *Probability of error in MMSE multiuser detection*, IEEE Trans. Inform. Theory (1997), 858–871.
- [33] J. Proakis, *Digital communications*, fourth ed., McGraw-Hill, New York, 2000.
- [34] D. Reynolds and X. Wang, *Adaptive group-blind multiuser detection based on a new subspace tracking algorithm*, IEEE Trans. Commun. **49** (2001), no. 7, 1135–1141.

- [35] D. Reynolds, X. Wang, and H. V. Poor, *Blind adaptive space-time multiuser detection with multiple transmitter and receiver antennas*, IEEE Trans. Signal Processing **50** (2002), no. 6, 1261–1276.
- [36] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, *Space-time block codes from orthogonal designs*, IEEE Trans. Inform. Theory **45** (1999), no. 5, 1456 – 1467.
- [37] V. Tarokh, N. Seshadri, and A. R. Calderbank, *Space-time codes for high rate wireless communication: Performance criterion and code construction*, IEEE Trans. Inform. Theory **44** (1998), no. 2, 744 – 765.
- [38] S. Verdú, *Multiuser detection*, Cambridge Univ. Press, Cambridge, UK, 1998.
- [39] S. Verdú and H. V. Poor, *Abstract dynamic programming models under commutativity conditions*, SIAM Journal on Control and Optimization **25** (1987), no. 4, 990 – 1006.
- [40] A. J. Viterbi, *An intuitive justification and a simplified implementation of the MAP decoder for convolutional codes*, IEEE Journ. Select. Areas. Commun. **16** (1998), no. 2, 260 – 264.
- [41] X. Wang and A. Host-Madsen, *Group-blind multiuser detection for uplink CDMA*, IEEE Journ. Select. Areas Commun. **17** (1999), no. 11, 1971–1984.
- [42] X. Wang and H. V. Poor, *Blind equalization and multiuser detection for CDMA communications in dispersive channels*, IEEE Trans. Commun. **COM-46** (1998), no. 1, 91–103.
- [43] _____, *Blind multiuser detection: A subspace approach*, IEEE Trans. Inform. Theory **44** (1998), no. 2, 677–691.
- [44] _____, *Iterative (turbo) soft interference cancellation and decoding for coded CDMA*, IEEE Trans. Commun. **47** (1999), no. 7, 1046 – 1061.
- [45] _____, *Space-time multiuser detection in multipath CDMA channels*, IEEE Trans. Signal Processing **47** (1999), no. 9, 2356–2374.
- [46] _____, *Wireless communication systems: Advanced techniques for signal reception*, Prentice-Hall, Upper Saddle River, NJ, 2002.

- [47] Y. Zhang and R. S. Blum, *Multistage multiuser detection for CDMA with space-time coding*, Proc. Tenth IEEE Workshop on Statistical Signal and Array Processing (Poconos, PA), Aug. 2000, pp. 1–5.
- [48] R. E. Ziemer, R. L. Peterson, and D. E. Borth, *Introduction to spread spectrum communications*, Prentice-Hall: Upper Saddle River, NJ, 1995.