# Adaptive Interference Suppression in Wireless Communication Systems*

Michael L. Honig† and H. Vincent Poor ‡

December 4, 1997

## 1 Introduction

One of the major features that distinguish modern wireless communication channels from wireline channels is the significant amount of structured interference that must be contended with in wireless channels. This interference is inherent in many wireless systems due to their operation as *multiple-access* systems, in which multiple transmitter/receiver pairs communicate through the same physical channel using non-orthogonal multiplexing. Structured interference also arises because of other non-systemic features of wireless systems, such as the desire to share bandwidth with other, dissimilar, communication services.

Signal processing plays a central role in the suppression of the structured interference arising in wireless communication systems. In particular, the use of appropriate signal processing methods can make a significant difference in the performance of such systems. Moreover, since many wireless systems operate under highly dynamic conditions due to the mobility of the transceivers and to the random nature of the channel access, *adaptive* signal processing is paramount in this context.

The study of adaptive processing techniques for interference suppression in wireless systems has been a very active area of research in recent years. The purpose of this chapter is to introduce the reader both to the basic problems arising in this area, and to the key methods that have been developed for dealing with these problems. This presentation will focus primarily on the problem of suppressing multiple-access interference (MAI), which is the limiting source of interference for the wireless systems being proposed for many emerging applications areas such as third-generation mobile telephony [94, 125] and wireless personal communications [19, 70]. However, we will also touch briefly on the related and important problems of multipath mitigation and narrowband interference suppression.

*This work was prepared in part under the support of the U. S. Army Research Office under Grant DAAH04-96-1-0378, and in part under the support of the U. S. Office of Naval Research under Grant N00014-94-1-0115.

†M. L. Honig is with the Department of Electrical and Computer Engineering, Northwestern University, Evanston, IL 60208 USA.

‡H. V. Poor is with the Department of Electrical Engineering, Princeton University, Princeton, NJ 08544 USA.

In treating the problem of MAI suppression, it is useful to consider a general multiple-access signal model that arises in the context of a wireless digital communications network operating with a coherent modulation format. The waveform received by a given terminal in such a network can be modeled as consisting of a set of superimposed modulated data signals observed in additive noise:

$$r(t) = S_t(\mathbf{b}) + n(t), \quad -\infty < t < \infty, \tag{1.1}$$

where $S_t(\mathbf{b})$ and $n(t)$ represent the useful signal and the ambient channel noise, respectively.

The useful signal $S_t(\mathbf{b})$ in this model is comprised of the data signals of $K$ active users in the channel, and can be written as

$$S_t(\mathbf{b}) = \sum_{k=1}^{K} A_k \sum_{i=-B}^{B} b_{i,k} s_k(t - iT - \tau_k), \tag{1.2}$$

where $2B+1$ is the number of symbols per user in the data frame of interest, $T$ is the symbol interval, and where $A_k$, $\tau_k$, $\{b_{i,k}\}$, and $\{s_k(t); 0 \le t \le T\}$, denote, respectively, the received amplitude, delay, symbol stream, and normalized modulation waveform (or pulse shape) of the $k^{th}$ user. The matrix $\mathbf{b}$ denotes the $K \times (2B+1)$ matrix whose $k,i-$th element is $b_{i,k}$. The data signals of the individual users may be asynchronous, in which case the relative delays with which the various data signals arrive at the receiver are distinct. However, when considering analytical properties, it is often sufficient to examine the synchronous case (i.e., $\tau_1 = \tau_2 = \cdots = \tau_K$), since asynchronous problems can be viewed as large synchronous problems. It should be noted further that, although this model does not explicitly include effects such as fading, multipath, intersymbol interference, or narrowband interference, such effects can be included without loss of tractability, as will also be discussed further in the sequel. A further generalization of this model allows for diversity at the receiver, in which multiple waveforms are observed, each of which contains information about the data sequences. Such a model arises when considering the use of antenna arrays for reception, as will be described below.

The principal feature that distinguishes multiuser formats of the type described in (1.1) - (1.2) from one another is the choice of the set of signaling waveforms (i.e., the signal constellation, $s_1, s_2, \ldots, s_K$). We are interested here in problems in which these waveforms are not orthogonal. The use of non-orthogonal waveforms has several advantages in certain wireless channels, including a greater bandwidth utilization under the common conditions of channel fading and bursty traffic. One of the most important formats of this type is the direct-sequence spread-spectrum multiple-access format, which corresponds to a set of signaling waveforms of the form

$$s_k(t) = \begin{cases} 2^{\frac{1}{2}} a_k(t) \sin(\omega_c t + \phi_k) & , \quad t \in [0, T] \\ 0 & , \quad t \notin [0, T] \end{cases} \tag{1.3}$$

where $\omega_c$ is a common carrier frequency, $\phi_k$ is the phase of the $k^{th}$ user relative to some reference, and where the spreading waveforms $a_k(t)$ are of the form:

$$a_k(t) = \sum_{j=0}^{N-1} a_{k,j} \psi(t - jT_c). \tag{1.4}$$

Here, $a_{k,0}, a_{k,1}, \ldots, a_{k,N-1}$ is a signature sequence of +1's and -1's assigned to the $k^{th}$ user, and $\psi$ is a normalized chip waveform of duration $T_c$ (where $NT_c = T$). The signature sequences (or *spreading codes*) and chip waveform are typically chosen to have autocorrelation and cross-correlation properties that reduce multipath, multiple-access interference, and unintended detectability, criteria that generally lead to signaling waveforms with nearly flat spectral characteristics. Note that, in this type of signaling, the bandwidth of the underlying data signal is spread by a factor of $N$. This particular model, which is sometimes termed direct-sequence code-division multiple-access (DS-CDMA) signaling, will be discussed further below.

It is the non-orthogonality of the signaling waveforms $s_1, s_2, \ldots, s_K$, that gives rise to the multiple-access interference with which the receiver must contend. In particular, if the receiver wishes to infer the data stream of a given user, say user # 1, then the fact that the other users' signaling waveforms are not orthogonal to $s_1$ makes it impossible to isolate user 1's signal without diminishing the detectability of user 1's data. However, through proper signal processing, the effects of the interfering signals can be minimized so that little is lost to this source of error. The area of study that deals with such problems is *multiuser detection,* and this chapter is primarily concerned with performing this task efficiently and adaptively.

This treatment is organized as follows. In Section 2, we provide a brief review of the elements of multiuser detection, which provides a framework for the development of most of the remainder of the chapter. In Section 3, we consider in more detail a specific class of multiuser detectors - linear multiuser detectors - that contains many of the most promising structures for introducing adaptivity into this problem. This treatment generalizes the model (1.1) - (1.2) to include diversity and other effects. Section 4 discusses the particularization of linear multiuser detection to the DS-CDMA format described above. Next, in Section 5 we consider the adaptation of linear multiuser detectors. In particular, several basic adaptive algorithms are discussed in the context of their complexity, convergence and performance characteristics. Section 6 considers the implications of some non-ideal effects arising in wireless channels, and also discusses some systems issues impacting the application of adaptive interference suppression in wireless systems. An extensive, but not exhaustive, bibliography of key sources in this area is also included.

# 2 Elements of Multiuser Detection

Almost by definition, the performance characteristics of multiple-access channels featuring traditional demodulation techniques are limited by multiple-access interference. It can be shown, however, that such limitations are due largely to the use of non-optimal signal processing in the demodulator, and are not due to fundamental characteristics of the channel. Multiuser detection seeks to remove this MAI limitation by the use of appropriate signal processing. Essentially, through the use of multiuser detection (or derivative signal processing techniques), performance in multiple-access channels can be returned to that of corresponding single-access channels, or at least to a situation in which performance is no longer MAI limited. This property is obviously very desirable, even in radio networks using power control or other protocols that seek to limit the effects of MAI.

The basic problem of multiuser detection is that of inferring the data contained in one

or more signals embedded in a non-orthogonal multiplex, the entire multiplex of which is received in ambient noise. Equations (1.1) - (1.2) describe a model of such a received signal. Within this context, multiuser detection refers to the problem of detecting all or part of the symbol matrix $\mathbf{b}$ from the multiplex (1.1) - (1.2) with non-orthogonal signalling waveforms, such as those arising in DS-CDMA. In the demodulation of any given user in such a multiplex, it is necessary to process the received signal in such a way as to minimize two types of detrimental effects - the multiple-access interference caused by the remaining $K-1$ users in the channel, and the ambient channel noise. In order to focus on the multiple-access interference, the great majority of research on this problem has ascribed the simplest possible model to the ambient channel noise; namely, that the only ambient channel noise is additive white Gaussian noise (AWGN) with fixed spectral height, say $\sigma^2$, and that this noise is independent of the data signals. In the following paragraphs, we also assume this model for the ambient noise.

Within the multiple-access signaling model there are two general scenarios of interest: an *uplink* (or *reverse link*) scenario, in which all signaling waveforms and data timing are known; and a *downlink* (or *forward link*) scenario, in which only a single user's waveform and data timing are known. In the uplink scenario, the general characteristics of optimal demodulation schemes for (1.1) - (1.2) under an AWGN model can be inferred by examining the likelihood function of the observed waveform (1.1), conditioned on the knowledge of all data symbols (i.e., conditioned on $\mathbf{b}$). On assuming that the received amplitudes are known, this likelihood function can be written via the Cameron-Martin formula [76] as

$$\ell(\{r(t); -\infty < t < \infty\}|\mathbf{b}) = C \exp\{\Omega(\mathbf{b})/2\sigma^2\} \tag{2.5}$$

where

$$\Omega(\mathbf{b}) = 2\int_{-\infty}^{\infty} S_t(\mathbf{b}) r(t) dt - \int_{-\infty}^{\infty} S_t^2(\mathbf{b}) dt \tag{2.6}$$

and where $C$ is a constant. The part of (2.6) that depends on the received waveform can be written as

$$\int_{-\infty}^{\infty} S_t(\mathbf{b}) r(t) dt = \sum_{k=1}^{K} A_k \sum_{i=-B}^{B} b_{i,k} y_{i,k} \tag{2.7}$$

where

$$y_{i,k} \equiv \int_{-\infty}^{\infty} s_k(t - iT - \tau_k) r(t) dt \tag{2.8}$$

is the output of a filter matched to the $k^{th}$ user's signaling waveform shifted to the $i^{th}$ symbol interval of the $k^{th}$ user.

It follows from (2.5) - (2.8) that the matrix of matched filter outputs,

$$\mathbf{y} = \{y_{i,k}; k = 1, \ldots, K; i = -B, \ldots, B\}, \tag{2.9}$$

forms a sufficient statistic for the matrix $\mathbf{b}$ of data symbols; that is, all information in the received waveform that is relevant to making inferences about $\mathbf{b}$ is contained in $\mathbf{y}$. So, the main job of optimal multiuser detection is to map the matrix $\mathbf{y}$ of observables to a matrix $\hat{\mathbf{b}}$ of symbol decisions. Thus, the general structure of optimal systems for determining the data symbols from the received waveform consists of an analog front-end that extracts the matched filter outputs, followed by a decision algorithm that infers optimal decisions from

the collection of these outputs. The nature of the decision algorithm in this process depends on the optimality criterion that one wishes to apply to the decision. If one adopts either a *maximum-likelihood* criterion,

$$\max_{\hat{\mathbf{b}}} \ell(\{r(t); -\infty < t < \infty\}|\hat{\mathbf{b}}) \tag{2.10}$$

or a *minimum-error-probability* criterion,

$$\min_{\hat{b}_{i,k}} P(\hat{b}_{i,k} \neq b_{i,k}|\{r(t); -\infty < t < \infty\}), \tag{2.11}$$

then, assuming the signaling waveforms $s_k$ satisfy $s_k(t) = 0$ for $t \notin [0, T]$, this optimal decision algorithm can be implemented as a dynamic program (i.e., a *sequence detector*) having $O(|A|^K)$ time complexity per binary decision (see [119]), where $|A|$ is the size of the symbol alphabet. In the case of synchronous signals, dynamic programming is unnecessary and the optimal detectors essentially involve either exhaustive search over the $|A|^K$ symbol choices in each symbol interval in the case of maximum-likelihood detection, or $O(|A|^K)$ computation of posterior probabilities in the case of minimum-error-probability detection.

Using these techniques it has been shown by Verdú (cf., [118]) that, for reasonably high symbol-energy-to-ambient-noise ratios, performance very near that of single-user communications is possible with the optimal multiuser detector. This is a considerable performance gain over conventional matched-filter detection (which demodulates $b_{i,k}$ by simple scalar quantization of $y_{i,k}$), which suffers from substantial performance losses in some multiple-access situations. (Such situations include the "near-far" situation, in which interfering users are received with much larger power than users of interest.) However, the much improved performance afforded by optimal multiuser algorithms comes at the expense of both *computational* complexity (i.e., the $O(|A|^K)$ computational cost per binary decision); and *informational* complexity due to the need for knowing all delays, amplitudes and modulation waveforms to extract the matrix sufficient statistic $\mathbf{y}$.

During the late 1980's and early 1990's, a significant amount of research addressed the problem of reducing the computational complexity of multiuser detection. A key approach to this problem is to restrict the optimal detector to be of the form of a *linear multiuser detector,* in which the data is demodulated by scalar quantization of a linear mapping on the matrix $\mathbf{y}$. In view of the definition of $y_{i,k}$, this type of detector is effectively comprised of a linear filter applied to the received waveform, followed by a scalar quantizer. (Of course the filter may depend on both $k$ and $i$.) Two types of linear detectors of interest are the *decorrelating detector* (or *decorrelator*), which chooses the linear filter to have zero output multiple-access interference [49]; and the *MMSE detector*, which chooses the linear filter to have minimum output energy within the constraint that the response of the filter to $s_k(t - iT - \tau_k)$ is fixed [31, 44, 51, 58, 82, 83, 93, 96, 133]. Such detectors can be shown to also satisfy other optimality criteria. Although such detectors fall short of optimal (maximum-likelihood) detection in terms of error probability, they are still far superior to conventional detection in terms of their error-probability performance in interference-limited environments. Linear detectors form the basis for the results described in the remainder of this chapter, and a detailed description of their properties is found in Sections 3 and 4.

Several useful nonlinear lower-complexity multiuser detectors also have been developed. These are based primarily on various techniques for successive cancellation of interference.

Also, methods for combating fading, multipath, etc. have been combined with multiuser detection as well. Some works in these areas include [17, 45, 68, 72, 111, 112, 113, 114, 115, 116, 135, 139, 140], although this list is hardly exhaustive. A survey of basic multiuser detection methods, and a more complete bibliography up to 1993, can be found in [119]. (See, also the forthcoming textbook [121].)

The issue of informational complexity in multiuser detection has been addressed through the use of adaptivity. This issue is particularly critical in the context of downlink demodulation, in which the direct implementation of non-adaptive versions of the above-noted detectors is neither practical nor desirable. However, uplink adaptivity is also of interest in practice, due to the dynamic nature of practical multiple-access channels. (Some discussion of this issue is found in Section 6.) Recent progress on adaptive multiuser detection is the subject of the remainder of this chapter.

# 3 Linear Interference Suppression

In this section we look more carefully at linear multiuser detection, in which linear filtering at the receiver can be used to suppress wideband multiple-access interference. This general approach to interference suppression has its origins in the related problem of of linear equalization in the presence of synchronous interfering data signals[1], and seems to have been first studied nearly thirty years ago [24], [42]. There are several reasons why linear interference suppression is attractive for wireless applications:

1. it can suppress both narrowband and wideband multiple-access interference;
2. it has modest complexity relative to other interference suppression and multiuser detection techniques, as noted in the preceding section; and
3. the linear filter can be implemented as a digital filter, or tapped-delay line, making it convenient for adaptation using conventional adaptive algorithms [23],[32].

To discuss this approach, it is convenient to view the multiple-access channel considered in Section 1 as a many-input linear system, as shown in Figure 3.1. In this depiction, the (linear) operations of modulation and transmission of the data sequences of the various users are lumped into the linear time-invariant transfer functions $\hat{H}_1(f), \hat{H}_2(f), \ldots, \hat{H}_K(f)$. In particular, (1.1) - (1.2) is realized by this model if we choose $\hat{H}_k(f)$ to be the system with impulse response

$$H_k(t) = A_k s_k(t - \tau_k). \tag{3.1}$$

In this context, we will first derive the optimal linear receiver for this type of channel, where the optimality criterion is minimum Mean Squared Error (MSE). We will then discuss some properties of this detector, such as the number of users that can be suppressed completely versus available bandwidth, and the implementation of the detector as a fractionally-spaced tapped-delay line. Application to the DS-CDMA model of (1.3) - (1.4) is discussed in Section 4.

[1]The multiuser signaling model of (1.1) - (1.2) includes as a special case the single-user intersymbol-interference channel, which corresponds to the case $K = 1$ with $s_1$ nonzero over more than one symbol interval.

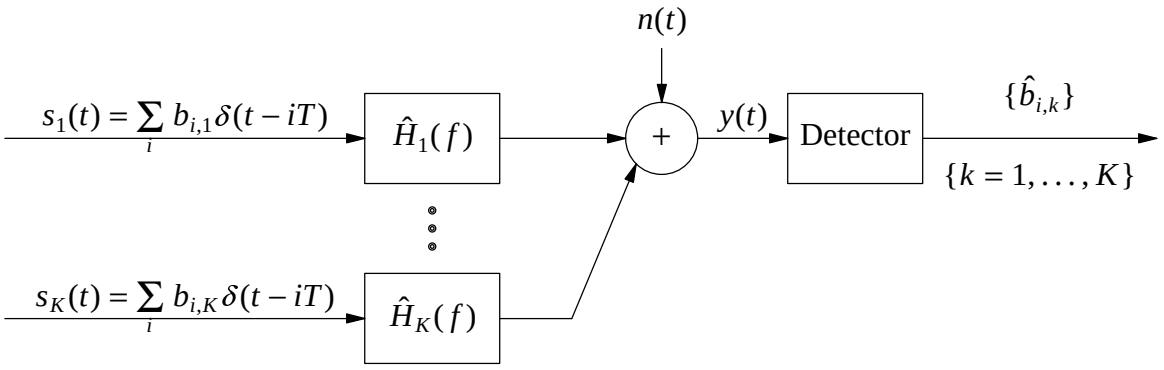


Figure 3.1: Multiple-access channel model.

## 3.1 Multi-input/Multi-output (MIMO) Minimum Mean Squared Error (MMSE) Linear Detector

To examine the MMSE linear multiuser detector, we will generalize the above signaling model slightly. First, in order to treat some nonideal effects later, we will allow the signaling waveforms and received amplitudes to be complex. And, secondly, we will allow for $M^{th}$-order reception diversity, in which we have $M$ observation channels similar to that depicted in Figure 3.1. Usually, these $M$ channels correspond to the outputs of $M$ elements in an antenna array. In this situation, it is convenient to represent the multiple-access channel as a special case of a $K$-input/$M$-output linear channel with $M \times K$ transfer function $\hat{\mathbf{H}}(f)$.

Refering to Figure 3.2, the input to the channel $\hat{\mathbf{H}}(f)$ is

$$\mathbf{s}(t) = \sum_i \mathbf{b}_i \delta(t - iT) \tag{3.2}$$

where $\mathbf{b}_i$ is the $K$-vector of transmitted symbols at time $i$ (that is, $\mathbf{b}_i$ is the $i^{th}$ column of the matrix $\mathbf{b}$ introduced in Section 1) and $1/T$ is the symbol rate, assumed to be the same for all users. The $k$th component of $\mathbf{b}_i$, denoted as $b_{i,k}$, is the $i$th transmitted symbol from user $k$. We assume that the signaling waveforms corresponding to each user, as well as relative amplitudes and delays for each observation channel, are included in the channel transfer function $\hat{\mathbf{H}}(f)$. (We remark that this model is general enough to account for both transmitter and receiver diversity.)

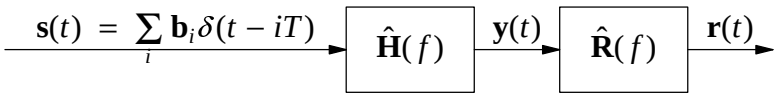


Figure 3.2: Multi-input/Multi-output (MIMO) channel and receiver model.

Both channel and receiver filters are illustrated in Figure 3.2. Let $\mathbf{H}(t)$ and $\mathbf{R}(t)$ be the impulse responses associated with the filters $\hat{\mathbf{H}}(f)$ and $\hat{\mathbf{R}}(f)$, respectively. The output of the receiver filter $\hat{\mathbf{R}}$ at time $kT$ is:

$$\mathbf{r}(kT) = \sum_i \{\mathbf{R} * \mathbf{H}[(k-i)T]\}\mathbf{b}_i + \mathbf{R} * \mathbf{n}(kT) \tag{3.3}$$

where "$*$" denotes convolution and $\mathbf{n}(t)$ is a noise vector with $M$ components. We wish to find the receiver filter $\hat{\mathbf{R}}$ that minimizes the MSE, $E\{\|\mathbf{r}(0) - \mathbf{b}_0\|^2\}$. This filter has been derived in [30] and [100], and is illustrated in Figure 3.3. It consists of a front-end matched filter with $K \times M$ matrix transfer function $\hat{\mathbf{H}}^*(f)\mathbf{S}_n^{-1}(f)$, where $\mathbf{S}_n(f)$ is the noise spectral density matrix, followed by a multi-input/multi-output discrete-time filter with $K \times K$ matrix coefficients. The transfer function (i.e., $z$-transform) of this discrete-time filter is

$$\mathbf{C}(z) = \mathbf{S}_d(z)[\mathbf{S}_H(z)\mathbf{S}_d(z) + \mathbf{I}]^{-1} \tag{3.4}$$

where $\mathbf{S}_H(z)$ is the equivalent discrete-time transfer function that maps the sequence of input symbol vectors $\{\mathbf{b}_i\}$ to the sequence of matched filter outputs $\{\mathbf{r}(iT)\}$, and $\mathbf{S}_d(z)$ is the spectrum of the data sequence. We can therefore write $\mathbf{S}_H(z)$ for $z$ on the unit circle as the aliased version of $\hat{\mathbf{H}}^*\mathbf{S}_n^{-1}\hat{\mathbf{H}}$, i.e., for $z = e^{j2\pi fT}$,

$$\mathbf{S}_H(e^{j2\pi fT}) = \frac{1}{T}\sum_k \hat{\mathbf{H}}^*\left(f - \frac{k}{T}\right)\mathbf{S}_n^{-1}\left(f - \frac{k}{T}\right)\hat{\mathbf{H}}\left(f - \frac{k}{T}\right) \tag{3.5}$$

and we have

$$\mathbf{S}_d(z) = \sum_i z^{-1}\rho_i \tag{3.6}$$

where $\rho_i = E\{\mathbf{b}_m\mathbf{b}_{m+i}^*\}$. In Figure 3.3, $\mathbf{d}_i$ denotes the output of the MIMO discrete-time filter $\mathbf{C}$ at symbol time $i$.

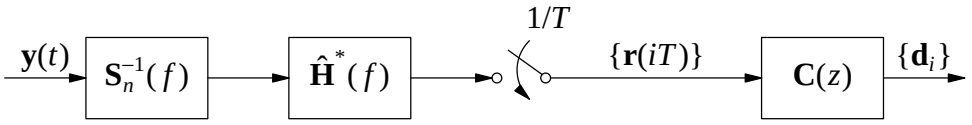


Figure 3.3: MIMO MMSE linear equalizer.

If the noise and data sequences are jointly timewise and componentwise uncorrelated, then the matched filter becomes $\hat{\mathbf{H}}^*(f)/\sigma_n^2$, where $\sigma_n^2$ is the noise variance per channel, and the discrete-time filter becomes

$$\mathbf{C}(z) = \sigma_n^2[\tilde{\mathbf{S}}_H(z) + \xi\mathbf{I}]^{-1} \tag{3.7}$$

with

$$\tilde{\mathbf{S}}_H(e^{j2\pi fT}) = \frac{1}{T}\sum_k \hat{\mathbf{H}}^*\left(f - \frac{k}{T}\right)\hat{\mathbf{H}}\left(f - \frac{k}{T}\right), \tag{3.8}$$

$\xi = \sigma_n^2/\sigma_b^2$, and $\sigma_b^2$ is the data variance per channel. (Note that the only difference between $\tilde{\mathbf{S}}_H$ and $\mathbf{S}_H$ is that $\tilde{\mathbf{S}}_H$ does not include the noise variance $\sigma_n^2$. This representation is convenient for what follows.) The minimum value of the MSE associated with the MMSE filter $\hat{\mathbf{R}}$ is then given by

$$\text{MMSE} = \text{trace}\left\{T\int_{-1/(2T)}^{1/(2T)} \mathbf{C}(e^{j2\pi fT})df\right\} \tag{3.9}$$

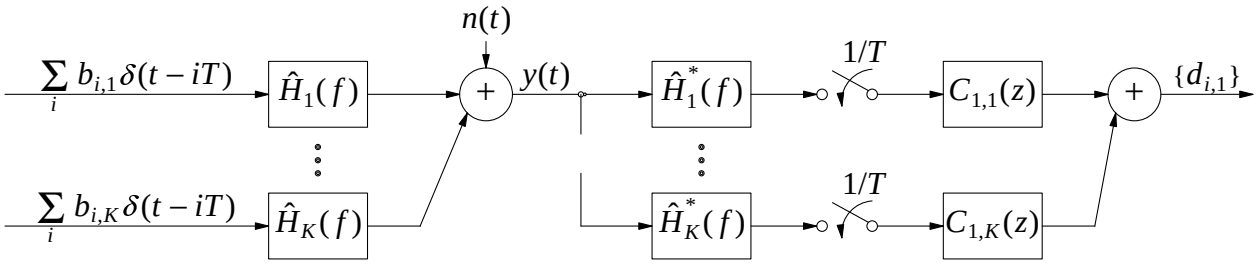


Figure 3.4: MMSE linear detector for the multiple-access channel. Only the detector for user 1 is shown.

This form of the MMSE detector for the multiple-access channel is shown in Figure 3.4. Only the detector for user 1 is shown, which consists of a bank of matched-filters, symbol-rate samplers, and discrete-time filters. The MMSE transfer function $C_{1,k}(z)$ in Figure 3.4 is the $(1,k)$th component of the matrix $\mathbf{C}(z)$ given by (3.7).

In the remainder of this chapter we will assume that the sequences of noise and symbol vectors are timewise and componentwise uncorrelated, in which case Figure 3.4 gives the canonical linear MMSE detector structure.

## 3.2 Zero-Forcing ( Decorrelating) Detector

As the noise variance diminishes to zero, the MMSE linear filter approaches the matched filter $\hat{\mathbf{H}}^*(f)$, followed by a symbol-rate sampler and a discrete-time matrix filter with transfer function

$$\mathbf{C}_{zf}(z) = [\tilde{\mathbf{S}}_H(z)]^{-1}. \tag{3.10}$$

The resulting MSE is given by

$$\mathrm{MSE}_{zf} = \sigma_n^2 \mathrm{trace}\left\{ T \int_{-1/(2T)}^{1/(2T)} \mathbf{C}_{zf}(e^{j2\pi fT}) df \right\} \tag{3.11}$$

Because the transfer function $\mathbf{C}_{zf}(z)$ inverts the equivalent discrete-time transfer function $\tilde{\mathbf{S}}_H$, which maps the source symbols to the matched filter outputs, it eliminates all inter-symbol and multiple-access interference (at the expense of enhancing the background noise). For this reason the matched-filter $\hat{\mathbf{H}}^*$ followed by the transfer function $\mathbf{C}_{zf}(z)$ is known as the *zero-forcing* detector for the multiple-input/multiple-output channel $\hat{\mathbf{H}}(f)$. When applied to the multiple-access channel in Figure 3.1, this detector is also known as the *decorrelating detector*, or *decorrelator*, since it removes the correlation among users due to nonorthogonal pulse shapes [49],[50].

It is apparent from (3.10) that the zero-forcing solution exists provided that the matrix $\tilde{\mathbf{S}}_H(z)$ is nonsingular for $z$ on the unit circle.[2] This condition has a special interpretation for the multiple-access channel. Specifically, in this case $\hat{\mathbf{H}}(f)$ is a $1 \times K$ row vector, so that

[2] Of course, it is possible that $\mathbf{S}_H(z)$ is singular for some set of $z$ on the unit circle, and that the MSE that results from substituting (3.10) into (3.11) is finite. However, finite MSE requires that $\mathbf{S}_H(e^{j2\pi fT})$ cannot be singular for $f$ in some interval with positive length.

$\hat{\mathbf{H}}^*(f)\hat{\mathbf{H}}(f)$ is an outer product matrix, which has rank one. $\tilde{\mathbf{S}}_H(e^{j2\pi fT})$ in (3.8) is therefore the sum of $L$ rank-one matrices, where for each $f \in [-1/(2T), 1/(2T)]$, $L$ is the number of Nyquist zones[3] where $\hat{\mathbf{H}}(f) \neq 0$. Since $\tilde{\mathbf{S}}_H(e^{j2\pi fT})$ is a $K \times K$ matrix, a necessary condition for $\tilde{\mathbf{S}}_H(e^{j2\pi fT})$ to be nonsingular for all $f \in [-1/(2T), 1/(2T)]$ is that $K \leq L$ for each $f$. This implies that for the zero-forcing solution to exist, there must be at least $K$ Nyquist zones available to the users. This property was first observed by Petersen and Falconer [73] in the context of wire (twisted-pair) channels with crosstalk. (See also [4].) Note that these Nyquist zones can be spread among the users so that either (1) the users do not overlap in frequency (namely, Frequency-Division Multiple-Access (FDMA)), (2) all of the users overlap at all frequencies (CDMA or TDMA), or (3) some users overlap at some frequencies, but not at other frequencies (combined FDMA/TDMA/CDMA).

For the multiple-access channel, the availability of $K$ Nyquist zones for $K$ users is necessary but not sufficient to ensure the existence of the zero-forcing solution. That is, it may happen that even with more than $K$ Nyquist zones available, the zero-forcing solution does not exist for $f$ in some positive interval contained in $[-1/(2T), 1/(2T)]$. For sufficiency, there must be at least $K$ vectors $\hat{\mathbf{H}}(f - k/T)$ appearing in the sum (3.8), that are linearly independent at each $f$. (This set of $K$ vectors may depend on $f$.)

Consider the case where the channels for each user shown in Figure 3.1 are the same, i.e., $H_k(f) = H(f)$ for each $k$. The preceding discussion suggests that each additional Nyquist zone in $H(f)$ can be viewed as an additional "dimension", or "degree of freedom", which can support an additional user without causing interference to existing users. This "dimensionality" interpretation will be useful when the tapped-delay line implementation of the linear MMSE detector is discussed in the next section. Note that for "orthogonal" multiple-access systems such as FDMA and TDMA, this observation is equivalent to stating that each user requires at least one Nyquist zone to ensure the existence of the zero-forcing equalizer [46, Ch. 10].

We now examine the effect of receiver diversity on the preceding results. In this case each channel $H_k(f)$ in Figure 3.1 becomes a $M \times 1$ column vector, where $M$ is the order of the receiver diversity, so that $\hat{\mathbf{H}}(f)$ is an $M \times K$ matrix ($K$ inputs, $M$ outputs). Consequently, the rank of $\hat{\mathbf{H}}^*(f)\hat{\mathbf{H}}(f)$ is at most $M$, and we conclude that a necessary condition for $\tilde{\mathbf{S}}_H(e^{j2\pi fT})$ to be nonsingular for all $f \in [-1/(2T), 1/(2T)]$ is that the number of users $K \leq LM$, where $L$ is again the number of Nyquist zones available to the users. This upper bound can be achieved if the matrices $\hat{\mathbf{H}}(f - k/T)$ in the sum (3.8) contain $LM$ linearly independent columns at each $f$. The number of dimensions or degrees of freedom available to suppress users is therefore given by the number of Nyquist zones times the number of antenna elements. (A similar treatment of dimensionality in the frequency and spatial domains is given in [18].)

A final remark about the zero-forcing detector is that even when it does not exist (i.e., $\mathbf{S}_H(e^{j2\pi fT})$ is singular), the MMSE detector is still well defined. Namely, it is always possible to select a filter to minimize output MSE. Consequently, the zero-forcing detector can be viewed more generally as the limit of the MMSE detector as the level of background noise tends to zero. This limit always exists even though the zero-forcing solution may not exist. This representation for the zero-forcing detector is useful in situations where the number of

[3] In this context, a Nyquist zone is a translate of the basic Nyquist interval, $[-1/(2T), 1/(2T)]$ by an integral multiple of $1/T$.

interferers exceeds the available dimensions that the detector has to supress multiple-access interference. Although the zero-forcing solution technically does not exist in this situation, the more general representation may still offer a substantial performance improvement relative to a simpler (e.g., matched-filter) detector.

## 3.3 Implementation as a Tapped-Delay Line (TDL)

The preceding formulation of the MMSE and zero-forcing detectors assumes knowledge of the user signaling waveforms along with relative timing and phase, the channel characteristics for each user, and the noise spectral density. In Section 5 we show that the MMSE detector can be implemented without this knowledge. This depends on an alternative representation of the MIMO MMSE linear filter as a bank of fractionally-spaced tapped-delay lines or discrete-time filters, which we now develop.

A classical result for single-user channels is that the optimal (MMSE) linear equalizer can be implemented as a fractionally-spaced tapped-delay line (TDL) [41, Ch. 10],[91]. To see this, let $W$ denote the two-sided bandwidth of the received data signal. Referring to Figure 3.4, the combination of the matched-filter $H_k^*(f)$, $1/T$ sampler, and discrete-time filter $C_k(z)$ in each branch of the MMSE detector can be replaced by a lowpass filter $B(f)$ with two-sided bandwidth $W$, a sampler at rate $W$, and a discrete-time filter $\tilde{C}_k(z)$ with frequency response that has period $W$. Stated another way, each front-end continuous-time matched-filter in the MMSE linear equalizer can be moved to the associated discrete-time filter, provided that the sampling rate is increased from $1/T$ to at least $W$.

The preceding discussion implies that each branch of the MMSE detector shown in Figure 3.4 can be replaced by a low-pass filter $B(f)$ followed a rate $W$ sampler and fractionally-spaced TDL, where the tap spacing is $1/W$. Choosing $B(f)$ to be the same for each branch allows the $K$ branches shown in Figure 3.4 to be collapsed into a *single* branch consisting of $B(f)$ a rate $W$ sampler, and discrete-time filter with transfer function given by the sum of the transfer functions for each branch.

To summarize, the MMSE multiuser linear detector for the multiple-access channel can be replaced by the bank of fractionally-spaced TDLs as shown in Figure 3.5. The filter $C_k(z)$, $1 \leq k \leq K$, is selected to minimize MSE for user $k$. The sampling rate, or tap spacing in the discrete-time filters thus depends on the bandwidth of the received signal. If the zero-forcing solution exists, then the bandwidth $W$ must be at least $K/T$ where $1/T$ is the symbol rate for the users. To interpret this result another way, observe that to distinguish $K$ symbols transmitted by $K$ (noncooperative) users, the receiver must sample at least $K$ times per symbol. Furthermore, these $K$ samples must be linearly independent. From Nyquist sampling theory, this implies that the bandwidth of the received signal must be greater than or equal to the sampling rate $K/T$.

With spatial diversity at the receiver, it is straightforward to show that the MMSE receiver can be implemented by summing the outputs of fractionally-spaced TDLs associated with each antenna. This is illustrated in Figure 3.6 for the case of two antennas. Figure 3.6 shows the MMSE detector for user 1, which contains two TDLs. Given $M$ antennas, the MMSE detector has $M$ TDLs for each user (indicating an $M$-fold increase in computational complexity associated with computing the filter outputs).

In general, the discrete-time impulse response associated with the MMSE filter $C_k(z)$

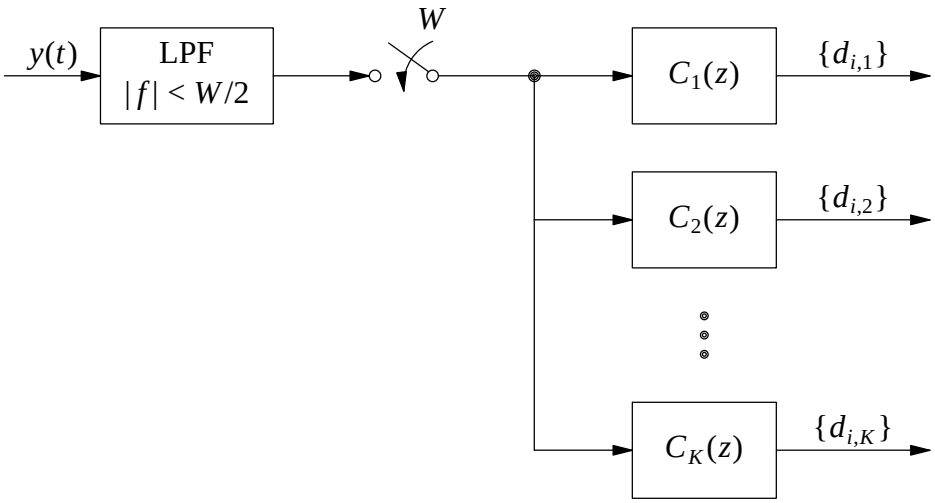


Figure 3.5: Implementation of the linear multiuser detector as a bank of fractionally-spaced TDLs.

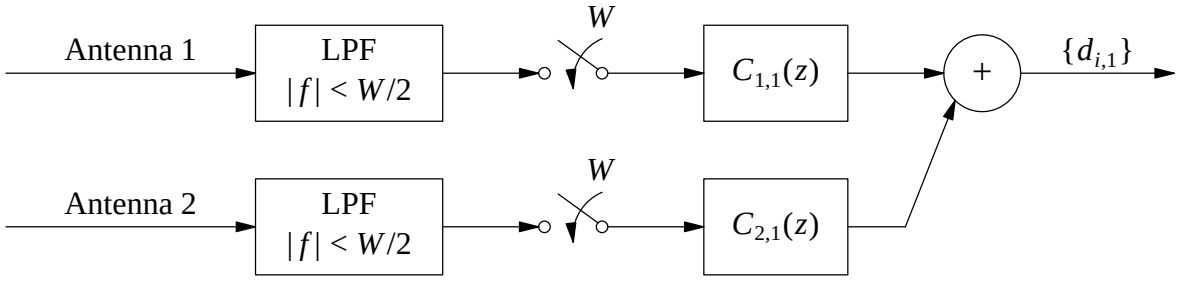


Figure 3.6: MMSE linear filter for user 1 with two-branch diversity.

in Figure 3.5 can be of infinite length. This is a problem since an infinite-length impulse response (IIR) filter cannot be implemented as a TDL, and is difficult to optimize when channel and interference parameters are changing. However, it is always possible to approximate each $C_k(z)$ with a finite-length impulse response (FIR) filter. Of course, there is some performance degradation associated with this truncation, which will depend on how fast the filter impulse response associated with $C_k(z)$ decays to zero.

Finally, we remark that an important benefit of the fractionally-spaced TDL implementation is that it eases timing recovery. That is, it is well known that for single-user channels a fractionally-spaced adaptive equalizer (with taps spaced at $T/k$, $k > 1$) is more robust with respect to timing offset than an adaptive equalizer with $T$-spaced taps. For DS-CDMA signals, timing recovery can be combined with interference suppression by using the adaptive algorithms discussed in Section 5 [52], [102], [129].

# 4 Application to DS-CDMA

We now apply the developments of Section 3 to the DS-CDMA model of (1.3) - (1.4). It is convenient to write this signal using a complex baseband model, in which case the received

signal can be written as

$$y(t) = \sum_{k=1}^{K} A_k \left[ \sum_i b_{i,k} p_k(t - iT - \tau_k) \right] + n(t) \tag{4.1}$$

where the pulse shape $p_k$ for user $k$ is given by

$$p_k(t) = \sum_{n=0}^{N-1} a_{k,n} \psi(t - nT_c); \tag{4.2}$$

that is $p_k$ is $s_k$ from (1.3) with $\omega_c = \phi_k \equiv 0$. Here, as before, $T_c$ is the chip duration and $\psi(t)$ is the normalized chip waveform, which are assumed to be the same for all users, and $A_k, \tau_k, \{b_{i,k}\}$, and $\{a_{k,n}\}$ are the received amplitude, delay, bit stream, and spreading sequence of user $k$. In order to represent the phase differences between signals, we will allow the amplitudes $A_k$ to be complex. The noise $n(t)$, representing noise in the complex baseband, is also assumed to be complex. For generality, we also allow $\psi$ to be complex. For the purposes of exposition, in what follows we will assume that the desired user to be demodulated is user # 1, and that $A_1 = 1$ and $\tau_1 = 0$.

Roughly speaking, the bandwidth spreading factor for DS-CDMA is the processing gain $N$ (assuming non-rectangular bandwidth-efficient chip waveforms). That is, each user spreads the transmitted bandwidth across $N$ Nyquist bands, so that the receiver has $N$ dimensions available with which to suppress interferers. We therefore conclude that the zero-forcing solution exists provided that $K \leq N$, and that the received pulse shapes are linearly independent.

For the TDL implementation of the MMSE detector, the front-end analog filter must cover the signal bandwidth, which is approximately $N/T$ where $T$ is the symbol duration. (This assumes a bandwidth-efficient chip waveform. If rectangular chips are used, then the bandwidth is approximately $2N/T$.) The sampling rate is then $N/T$, and the TDL has taps spaced at $T/N$. If the TDL is of infinite length, then in principle, it can effectively suppress $N - 1$ strong interferers. In what follows we will assume that the front-end analog filter is a chip-matched filter with impulse response $\psi^*(-t)$, which maximizes the Signal-to-Noise Ratio at the output of this filter in the absence of interference.

## 4.1 Discrete-Time Representation

We first specify the TDL coefficients in term of the received samples at the output of the chip matched-filter. Define the vector of received samples at the output of the chip matched-filter during the $i^{th}$ symbol as

$$\mathbf{r}'_i = \{r[iT], r[iT + T_c], \ldots, r[iT + (N-1)T_c]\} \tag{4.3}$$

where

$$r(t) = \int_{-\infty}^{\infty} y(t - s)\psi^*(-s)ds, \tag{4.4}$$

and $y(t)$ is the channel output given by (4.1). If $\psi(t)$ is confined to $[0, T_c]$, then the integral is from $t$ to $t + T_c$. For the time being we assume that all users are both chip- and symbol-synchronous. That is, referring to (4.3), $\tau_k = 0$, $1 \leq k \leq K$. Combining (4.1), (4.2), and

(4.4), we can write the vector $\mathbf{r}_i$ as a linear combination of vectors contributed by each of the users plus noise

$$\mathbf{r}_i = \sum_{k=1}^{L} b_{i,k} A_k \mathbf{p}_k + \mathbf{n}_i \tag{4.5}$$

where the upper index $L$ depends on whether the multiuser data signal is synchronous or asynchronous. For synchronous CDMA, we have $L = K$ and $\mathbf{p}_k$ is the vector of samples at the output of the chip matched-filter in response to the $k$th user's input waveform. The $m$th component of $\mathbf{p}_k$ in this case is therefore

$$\mathbf{p}_{k,m} = \int_{-\infty}^{\infty} p_k(mT_c - s)\psi^*(-s)ds. \tag{4.6}$$

Assuming zero inter-chip interference (i.e., $\psi_k(t - iT_c)$ are orthogonal waveforms for different $i$), then this integral becomes

$$\mathbf{p}_{k,m} = \int_{-\infty}^{\infty} a_{k,m}|\psi(s - mT_c)|^2 ds = a_{k,m} \tag{4.7}$$

since $\psi(t)$ is a unit energy pulse, and where $a_{k,m}$ is the $m$th spreading coefficient for user $k$. Consequently, for the case of synchronous DS-CDMA, we have

$$\mathbf{r}_i = \sum_{k=1}^{K} b_{i,k} A_k \mathbf{a}_k + \mathbf{n}_i \tag{4.8}$$

where $\mathbf{a}_k$ is the vector of spreading coefficients assigned to user $k$.

To specify the received samples for asynchronous DS-CDMA, the delay associated with user $k$ is expressed as

$$\tau_k = (\iota_k + \delta_k)T_c \tag{4.9}$$

where $\iota_k$ is an integer between 0 and $N - 1$, and $\delta_k = \tau_k/T_c - \iota_k$ lies in the interval $[0, 1)$. The delay $\iota_k$ specifies the number of whole chips by which user $k$ is shifted relative to user 1, and $\delta_k$ represents the additional partial chip delay. (In *chip-synchronous* DS-CDMA, $\delta_k = 0$, although $\iota_k \neq 0$ in general.) The computation of $\mathbf{p}_{k,m}$ for asynchronous DS-CDMA is illustrated in Figure 4.1. First note that user $k$ transmits *two* symbols, $b_{i-1,k}$ and $b_{i,k}$, within the time window $(i-1)T$ to $iT$ associated with $b_{i,1}$. This implies that user $k$ contributes *two* vectors to the sum (4.5), associated with the left and right parts of $p_k(t - iT - \tau_k)$ within $(i-1)T$ to $iT$. We therefore rewrite (4.8) as

$$\mathbf{r}_i = b_{i,1} A_1 \mathbf{p}_1 + \sum_{k=2}^{K} A_k (b_{i-1,k}\mathbf{p}_k^- + b_{i,k}\mathbf{p}_k^+) + \mathbf{n}_i \tag{4.10}$$

where $\mathbf{p}_k^-$ and $\mathbf{p}_k^+$ are the sampled outputs of the matched filter during symbol $i$ in response to $p_k[t + (T - \tau_k)]$ and $p_k(t - \tau_k)$, respectively. (This assumes that intersymbol interference is negligible.) The $m^{th}$ components of $\mathbf{p}_k^+$ and $\mathbf{p}_k^-$ are then

$$\begin{aligned} \mathbf{p}_{k,m}^+ &= \int_{-\infty}^{\infty} p_k(x)\psi^*[x - (mT_c - \iota_k) + \delta_k T_c]dx \\ &= \begin{cases} a_{k,m-\iota_k-1}\phi_2 + a_{k,m-\iota_k}\phi_1 & \text{for } m > \iota_k \\ a_{k,0}\phi_1 & \text{for } m = \iota_k \\ 0 & \text{for } m < \iota_k \end{cases} \end{aligned} \tag{4.11a}$$

$$\begin{aligned}\mathbf{p}_{k,m}^{-} &= \int_{-\infty}^{\infty} p_k(x)\psi^*[x-(m+N-\iota_k)T_c+\delta_k T_c)]dx \\ &= \begin{cases} a_{k,m+N-\iota_k}\phi_1 + a_{k,m+N-\iota_k-1}\phi_2 & \text{for } m < \iota_k \\ a_{k,N-1}\phi_2 & \text{for } m = \iota_k \\ 0 & \text{for } m > \iota_k \end{cases} \end{aligned} \tag{4.11b}$$

where

$$\phi_1 = \int_{-\infty}^{\infty} \psi(s)\psi^*(s+\delta_k T_c)ds, \qquad \phi_2 = \int_{-\infty}^{\infty} \psi(s)\psi^*[s-(1-\delta_k)T_c]ds \tag{4.12}$$

and where we are accounting only for the contribution of the two chips from $p_k(t-\tau_k)$ (or $p_k[t+(T-\tau_k)]$) centered next to the $m$th chip of $p_1(t)$.

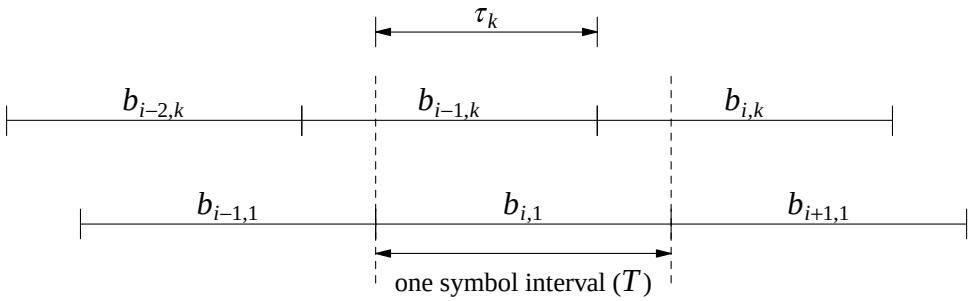


Figure 4.1: Illustration of interference from asynchronous user $k$. The dashed lines designate the time window spanned by the detector $\mathbf{c}$.

As an example, suppose that $\psi(t) = 1/\sqrt{T_c}$ for $0 \leq t < T_c$, and is zero elsewhere (rectangular chips). Then we have that

$$\mathbf{p}_{k,m}^{+} = \begin{cases} a_{k,m-\iota_k-1}\delta_k + a_{k,m-\iota_k}(1-\delta_k) & \text{for } m > \iota_k \\ a_{k,0}(1-\delta_k) & \text{for } m = \iota_k \\ 0 & \text{for } m < \iota_k \end{cases} \tag{4.13a}$$

and

$$\mathbf{p}_{k,m}^{+} = \begin{cases} a_{k,N-(\iota_k-m+1)}\delta_k + a_{k,N-(\iota_k-m)}(1-\delta_k) & \text{for } m < \iota_k \\ a_{k,N-1}\delta_k & \text{for } m = \iota_k \\ 0 & \text{for } m > \iota_k \end{cases} \tag{4.13b}$$

Note that except for $m = \iota_k$, if $\mathbf{p}_{k,m}^{+} \neq 0$ then $\mathbf{p}_{k,m}^{-} = 0$, and vice versa.

We therefore conclude that for both synchronous and asynchronous DS-CDMA, the received vector of chip matched-filter outputs during time $i$ can be written as (4.5). For synchronous DS-CDMA, $L = K$, and the vectors in the sum (4.5) are the spreading sequences assigned to the users. For asynchronous DS-CDMA, $L \leq 2K-1$, and the vectors in the sum (4.5) are given by $\mathbf{p}_k^{+}$ and $\mathbf{p}_k^{-}$.

## 4.2 Computation of MMSE Coefficients

As noted in Section 3, the optimal discrete-time filter for MMSE detection is not necessarily an FIR filter. Thus, in order to limit the complexity of the MMSE detector in this setting, it is desirable to truncate the number of taps in the TDL. To consider this issue, let us define an "extended" received vector

$$\bar{\mathbf{r}}_i' = [\mathbf{r}_{i-M}', \mathbf{r}_{i-M+1}', \ldots, \mathbf{r}_i', \ldots, \mathbf{r}_{i+M}'] \tag{4.14}$$

where $2M+1$ is the width of the truncated processing window to be considered. That is, $\bar{\mathbf{r}}_i$ consists of the vectors $\mathbf{r}_{i-M}, \ldots, \mathbf{r}_{i+M}$ stacked on top of each other, and has dimension $N(2M+1)$. Letting $\mathbf{c}$ denote the vector of TDL coefficients, the output of the TDL at time $iT$ can written as

$$d_i = \mathbf{c}^\dagger \bar{\mathbf{r}}_i \tag{4.15}$$

where "$\dagger$" denotes complex conjugate transpose. The estimate of the transmitted symbol $b_{i,1}$ can then be obtained by quantizing this output. In the case of binary transmissions ($b_{i,k} \in \{\pm 1\}$), the detected (uncoded) symbol is $\hat{b}_{i,1} = \mathrm{sgn}(d_i)$.

For the MMSE detector, the TDL coefficient vector is selected to minimize

$$\mathrm{MSE} = E\{|\mathbf{c}^\dagger \bar{\mathbf{r}}_i - b_{i,1}|^2\} = 1 + \mathbf{c}^\dagger \mathbf{R}\mathbf{c} - 2\mathrm{Re}\{\mathbf{c}^\dagger \bar{\mathbf{p}}_1\} \tag{4.16}$$

where

$$\bar{\mathbf{p}}_1' = [0 \cdots 0\, \mathbf{p}_1'\, 0 \cdots 0] \tag{4.17}$$

with the number of zeros that precede/succeed $\mathbf{p}_1'$ equal to $NM$,

$$\mathbf{R} = E\{\bar{\mathbf{r}}_i \bar{\mathbf{r}}_i^\dagger\}, \tag{4.18}$$

the noise samples at the output of the chip matched filter are white with variance $\sigma_n^2$, we normalize $E\{|b_{i,1}|^2\} = 1$, and the transmitted symbols are assumed to be uncorrelated.

Selecting $\mathbf{c}$ to minimize the MSE gives

$$\mathbf{c}_{mmse} = \mathbf{R}^{-1} \bar{\mathbf{p}}_1 \tag{4.19}$$

and

$$\mathrm{MMSE} = 1 - \mathbf{c}_{mmse}^\dagger \bar{\mathbf{p}}_1 = 1 - \bar{\mathbf{p}}_1^\dagger \mathbf{R}^{-1} \bar{\mathbf{p}}_1 \tag{4.20}$$

For synchronous DS-CDMA, we note that $\mathbf{R}$ is a block-diagonal matrix, where each $N \times N$ diagonal block is given by

$$\mathbf{R}_N = \sum_{k=1}^{K} A_k^2 \mathbf{p}_k \mathbf{p}_k^\dagger + \sigma_n^2 \mathbf{I} \tag{4.21}$$

Consequently, the MMSE TDL coefficient vector, specified by (4.19), has the form

$$\mathbf{c}_{mmse} = [0 \cdots 0\, \tilde{\mathbf{c}}'\, 0 \cdots 0] \tag{4.22}$$

where $NM$ zeros precede/succeed the $N$-vector $\tilde{\mathbf{c}}' = \mathbf{R}_N^{-1} \mathbf{p}_1$. We therefore conclude that for synchronous DS-CDMA, the MMSE detector consists of a chip-matched filter followed by a *finite-length* TDL that spans only one symbol (i.e., we can set $M = 0$ in (4.14)).

For asynchronous DS-CDMA, the MMSE TDL is no longer finite in general; however, we can still consider a truncated version that spans one symbol. This detector, assuming $T_c$-spaced taps and a front-end chip matched-filter, is shown in Figure 4.2. The only difference between this "$N$-tap MMSE detector" and the conventional matched-filter detector is the way in which the TDL coefficients are selected. For the matched-filter detector $\mathbf{c} = \mathbf{a}_1$ (the spreading coefficients for user 1), whereas for the MMSE detector,

$$\mathbf{c}_{mmse} = \mathbf{R}^{-1}\mathbf{p}_1 \tag{4.23}$$

where $\mathbf{R} = \mathbf{R}_N$ in (4.21). Note that in the absence of background noise, $\mathbf{R}$ is singular if $L < N$. However, it is easily shown that any $\mathbf{c}$ that satisfies $\mathbf{Rc} = \bar{\mathbf{p}}_1$ minimizes MSE even in this singular case.

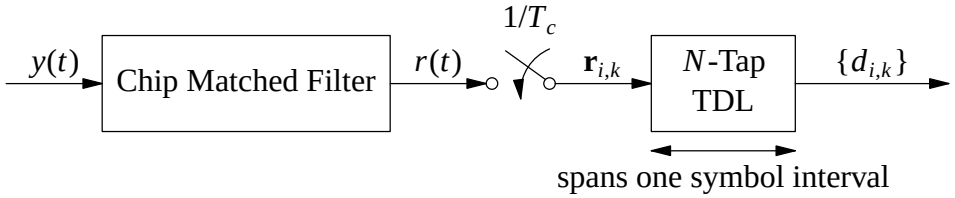


Figure 4.2: $N$-tap MMSE Detector for user $k$.

## 4.3 Geometric Interpretation

Throughout the rest of this section we focus on the $N$-tap detector shown in Figure 4.2. This is for simplicity. The following discussion is easily generalized to account for a TDL that spans multiple symbol intervals. The vectors $\mathbf{p}_1, \ldots, \mathbf{p}_L$ that appear in the sum (4.21) are illustrated in Figure 4.3. The space spanned by these vectors is the *signal subspace*, denoted as $\mathbf{S}$. The *interference subspace*, denoted as $\mathbf{S}_I$, is the space spanned by $\mathbf{p}_2, \ldots, \mathbf{p}_L$. (We continue to assume that user # 1 is the user of interest.) If $\mathbf{p}_1, \ldots, \mathbf{p}_L$ are linearly independent, then $\mathbf{S}$ has dimension $L$, and $\mathbf{S}_I$ has dimension $L-1$.

We first observe that the MMSE solution $\mathbf{c}$ must lie in $\mathbf{S}$. Otherwise, we can write

$$\mathbf{c} = \mathbf{c}_s + \mathbf{c}_s^{\perp} \tag{4.24}$$

where $\mathbf{c}_s \in \mathbf{S}$ and $\mathbf{c}_s^{\perp}$ is orthogonal to $\mathbf{S}$. We then have that $\mathbf{p}_k^{\dagger}\mathbf{c}_s^{\perp} = 0$ for each $k = 1, \ldots, K$, and $(\mathbf{c}_s^{\perp})^{\dagger}\mathbf{r}_i = (\mathbf{c}_s^{\perp})^{\dagger}\mathbf{n}_i$. We therefore conclude that the component $\mathbf{c}_s^{\perp}$ in (4.24) adds a noise term to the filter output $d_i$, which increases the MSE. To minimize MSE, we must take $\mathbf{c}_s^{\perp} = \mathbf{0}$.

Because the MMSE solution $\mathbf{c}_{mmse} \in \mathbf{S}$, we can express $\mathbf{c}_{mmse}$ as a linear combination of the signal vectors. Let $\mathbf{P}$ denote the $N \times L$ matrix with columns $\mathbf{p}_1, \ldots, \mathbf{p}_L$. Then from (4.5) we can write the received vector as

$$\mathbf{r}_i = \mathbf{PAb}_i + \mathbf{n}_i \tag{4.25}$$

where $\mathbf{A} = \mathrm{diag}[A_1\, A_2 \cdots A_K]$ and $\mathbf{b}_i' = [b_{i,1}\, b_{i,2} \cdots b_{i,K}]$. Now define the $N \times K$ matrix $\mathbf{C}$, where the $k$th column of $\mathbf{C}$ is the vector of TDL coefficients used to demodulate user $k$.

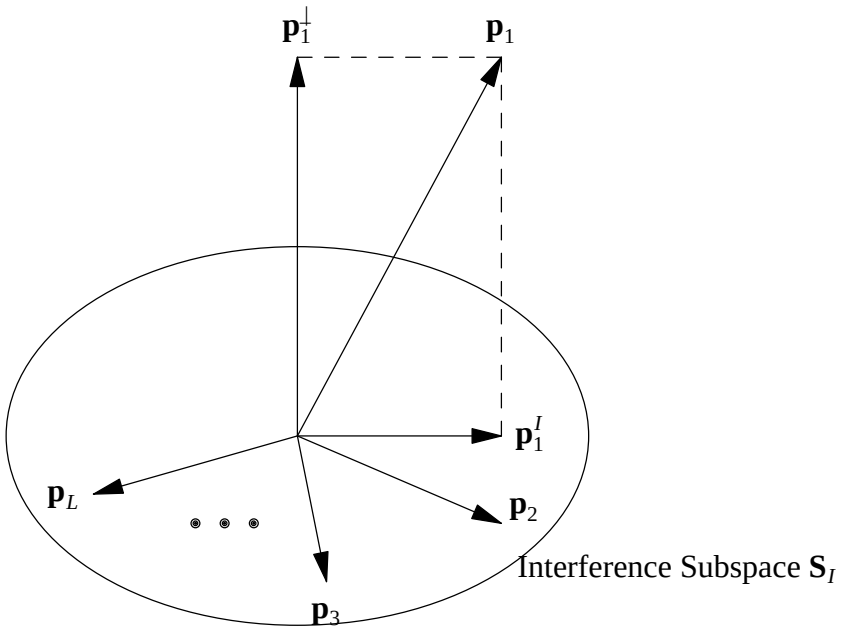


Figure 4.3: Geometric representation of desired signal and interference vectors. $\mathbf{p}_1^I$ is the projection of $\mathbf{p}_1$ onto the interference subspace $\mathbf{S}_I$.

The first column of $\mathbf{C}$ is therefore the vector $\mathbf{c}$ used in (4.15) to demodulate user 1. From the previous discussion it follows that each column of $\mathbf{C}_{mmse}$ can be expressed as a linear combination of the signal vectors. We therefore write $\mathbf{C} = \mathbf{P}\Gamma$, where $\Gamma$ is a $K \times K$ matrix, and note that $\mathbf{C}_k = \mathbf{P}\Gamma_k$, where the subscript $k$ denotes the $k$th column of the matrix. The total MMSE summed over all $K$ users is

$$\text{MSE} = E\{\|\mathbf{C}\mathbf{r}_i - \mathbf{b}_i\|^2\} = \text{trace}\{(\Gamma^\dagger\mathbf{P}^\dagger\mathbf{P}\mathbf{A} - \mathbf{I})(\mathbf{A}^\dagger\mathbf{P}^\dagger\mathbf{P}\Gamma - \mathbf{I}) + \sigma_n^2\Gamma^\dagger\mathbf{P}^\dagger\mathbf{P}\Gamma\} \tag{4.26}$$

and selecting $\Gamma$ to minimize this expression gives (cf., [129])

$$\Gamma = \mathbf{A}[\mathbf{A}\mathbf{P}^\dagger\mathbf{P}\mathbf{A} + \sigma_n^2\mathbf{I}]^{-1} \tag{4.27}$$

The matrix $\mathbf{P}^\dagger\mathbf{P}$ is the *cross-correlation* matrix for the set of signal vectors $\mathbf{p}_1, \ldots, \mathbf{p}_K$. Since $\mathbf{c} = \mathbf{P}\Gamma_1$, this relation gives an alternative method to (4.23) for computing the MMSE solution $\mathbf{c}_{mmse}$. When $K < N$, and the background noise is small, it is better to compute $\mathbf{c}_{mmse}$ via (4.27) since the matrix $\mathbf{R}_N$ defined by (4.21) is likely to be ill-conditioned. Finally, we note that this expression is analogous to the expression (3.7) for the matrix (multiuser) discrete-time transfer function with a bank of front-end matched filters.

## 4.4 Zero-Forcing (Decorrelating) Solution

In analogy with the zero-forcing solution for the MIMO MMSE detector discussed in Section 3.2, it may be possible to choose the $N$-vector $\mathbf{c}$ to completely remove multiple-access interference. From Figure 4.3 it is apparent that this zero-forcing, or decorrelating solution is

proportional to the orthogonal projection of the desired user vector $\mathbf{p}_1$ onto the interference subspace $\mathbf{S}_I$. Denoting the zero-forcing solution for $\mathbf{c}$ as $\mathbf{c}_{zf}$, the filter output is given by

$$d_i = \mathbf{c}_{zf}^\dagger \mathbf{r}_i = \mathbf{c}_{zf}^\dagger (b_{i,1}\mathbf{p}_1 + \mathbf{n}_i) \tag{4.28}$$

Namely, the output of the zero-forcing filter has only two components, one due to the desired signal and one due to background noise. Let $\mathbf{P}_I$ denote the $N \times (K-1)$ matrix with columns given by $\mathbf{p}_2, \ldots, \mathbf{p}_K$. The orthogonal projection of $\mathbf{p}_1$ onto $\mathbf{S}_I$ is denoted as

$$\mathbf{p}_1^\perp = \mathbf{p}_1 - \mathbf{P}_I (\mathbf{P}_I^\dagger \mathbf{P}_I)^{-1} (\mathbf{P}_I^\dagger \mathbf{p}_1), \tag{4.29}$$

and the zero-forcing solution is

$$\mathbf{c}_{zf} = \mathbf{p}_1^\perp / \eta \tag{4.30}$$

where the scale factor $1/\eta$ is selected so that $|\mathbf{c}_{zf}^\dagger \mathbf{p}_1| = |b_1| = 1$. (The quantity $\eta$ is known as the "near-far resistance", and has special significance, which will be explained later.) It is easily shown that $(\mathbf{p}_1^\perp)^\dagger \mathbf{p}_1 = \|\mathbf{p}_1^\perp\|^2$, so that $\eta = \|\mathbf{p}_1^\perp\|^2$.

It is apparent from Figure 4.3 that the zero-forcing solution for $\mathbf{c}$ exists provided that $\mathbf{p}_1$ is not contained in $\mathbf{S}_I$. In that case the dimension of the signal subspace $\mathbf{S}$ must be no greater than $N$. If the vectors $\mathbf{p}_1, \ldots, \mathbf{p}_L$ are linearly independent, then we must have $L \leq N$. For synchronous DS-CDMA this implies that the number of users $K \leq N$, and for asynchronous DS-CDMA, $2K - 1 \leq N$. Of course, even if this latter condition does not hold (as in a heavily loaded cellular system), the MMSE solution is still well defined. (Also, the addition of receiver diversity allows one to increase $K$ beyong this bound, as is discussed below.) As $K$ increases, the performance of the MMSE detector improves relative to the zero-forcing solution.

## 4.5 Asymptotic Behavior of the MMSE Solution

Here we examine the behavior of the MMSE solution as (i) the noise level diminishes to zero, and (ii) the interferers increase in energy. If the noise variance $\sigma_n^2 = 0$, then we observe that the zero-forcing solution, assuming it exists, gives zero MSE. We therefore conclude that the MMSE solution converges to the zero-forcing solution as $\sigma_n^2 \to 0$. It can be shown by matrix manipulations that the zero-forcing solution (4.30) is equivalent to the preceding expressions (4.23) and (4.27) where the noise variance $\sigma_n^2 = 0$.

Now consider what happens as user $k$'s amplitude $A_k \to \infty$. It is easily seen that $\mathbf{c}_{mmse}^\dagger \mathbf{p}_k \to 0$. Otherwise, we would have $\left(\mathbf{c}_{mmse}^\dagger \mathbf{p}_k\right)^2 > \varepsilon > 0$, which implies $\left(A_k \mathbf{c}_{mmse}^\dagger \mathbf{p}_k\right)^2 > \varepsilon A_k^2$. As $A_k \to \infty$, (4.16) implies that MSE $\to \infty$, which contradicts the fact that MMSE $\leq 1$ (i.e., $\mathbf{c} = 0$ gives MSE $= 1$). In fact it can be shown that

$$\lim_{A_k \to \infty} \left( A_k \mathbf{c}_{mmse}^\dagger \mathbf{p}_k \right) = 0, \tag{4.31}$$

which implies that as $A_k \to \infty$, the contribution to the MSE from user $k$ diminishes to zero (see [51],[82]).

To generalize (4.31), if $A_k \to \infty$ for $k$ in some subset $K_s$, then (4.31) applies for each $k \in K_s$ (assuming $\mathbf{c}$ has enough degrees of freedom to suppress these interferers). If the

set $K_s$ contains all $K-1$ interferers, then clearly $\mathbf{c}_{mmse} \to \kappa \mathbf{p}_1^{\perp}$, where $\kappa$ is a constant. Substituting for $\mathbf{c}_{mmse}$ in (4.28), and selecting $\kappa$ to minimize MSE gives $\kappa = 1/(\eta + \sigma_n^2)$. We therefore conclude that as the interfering amplitudes $A_k \to \infty$, $k \neq 1$,

$$\mathbf{c}_{mmse} \to \frac{\mathbf{p}_1^{\perp}}{\eta + \sigma_n^2}, \qquad \text{MMSE} \to \frac{\sigma_n^2}{\eta + \sigma_n^2} \tag{4.32}$$

Note that the filter $\mathbf{c}_{mmse}$ gives a biased estimate of $b_{i,1}$.

## 4.6 Performance Measures

In addition to MMSE, two other performance measures of interest are Signal-to-Interference-Plus-Noise Ratio (SINR) and error probability. The SINR is defined to be the ratio of the desired signal power to the sum of the powers due to noise and multiple-access interference at the output of the filter $\mathbf{c}$. That is,

$$\text{SINR} = \frac{(\mathbf{c}^{\dagger}\mathbf{p}_1)^2}{\sum_{k=2}^{L} A_k^2 (\mathbf{c}^{\dagger}\mathbf{p}_k)^2 + \sigma_n^2 \|\mathbf{c}\|^2} \tag{4.33}$$

It can be shown that the MMSE solution $\mathbf{c}$ also maximizes the SINR, and that this maximum SINR is

$$\text{MSIR} = \frac{\mathbf{c}_{mmse}^{\dagger}\mathbf{p}_1}{1 - \mathbf{c}_{mmse}^{\dagger}\mathbf{p}_1} = \frac{1}{\text{MMSE}} - 1 \tag{4.34}$$

To study the error probability, we restrict attention to the case in which all users transmit binary, equiprobable symbols. In this case, we have $\Pr\{\hat{b}_1 \neq b_1\} = \Pr\{\hat{b}_1 \neq b_1 | b_1 = 1\}$. Conditioning on all users' symbols, and assuming white Gaussian noise gives

$$P_{e|\mathbf{b}}(\mathbf{b}) = P(\hat{b}_1 \neq b_1 | \mathbf{b}, b_1 = 1) = Q\left(\frac{\mathbf{c}^{\dagger}\mathbf{p}_1 + \sum_{k=2}^{K} b_k A_k (\mathbf{c}^{\dagger}\mathbf{p}_k)}{\sigma_n \|\mathbf{c}\|}\right) \tag{4.35}$$

where $Q(x) = (2\pi)^{-\frac{1}{2}} \int_x^{\infty} e^{-t^2/2} dt$. The *average* error probability is then obtained by averaging (4.35) over the distribution for the bit vectors $\mathbf{b}$, i.e., $P_e = E\{P_{e|\mathbf{b}}\}$.

Two additional performance measures that are related to the asymptotic performance discussed in the preceding section are *asymptotic efficiency* and *near-far resistance*. Let $P_e(\sigma_n)$ denote the average error probability for a specific detector as a function of the noise variance $\sigma_n^2$. The asymptotic efficiency of the detector is then defined in [119], [120] as

$$\gamma = \sup\left\{\kappa : \lim_{\sigma_n \to 0} P_e(\sigma_n)/Q(\sqrt{\kappa A_1}/\sigma_n) > 0\right\} \tag{4.36}$$

and is a limiting measure, as the noise tends to zero, of how well the detector performs in the presence of multiple-access interference relative to optimal performance in the absence of multiple-access interference. The near-far resistance of the detector is defined in [119] as

$$\eta = \inf_{A_2, \ldots, A_L} \gamma. \tag{4.37}$$

That is, the near-far resistance is the asymptotic efficiency evaluated for *worst case* interference energies, and is a measure of the robustness of the detector with respect to variations in the received interference energies.

As $\sigma_n \to 0$, the error probability for the MMSE TDL detector satisfies

$$\lim_{\sigma_n \to 0} \frac{P_e(\sigma_n)}{Q(\|\mathbf{p}_1^{\perp}\|/\sigma_n)} = \lim_{\sigma_n \to 0} \frac{\min_{\mathbf{b}} P_{e|\mathbf{b}}(\sigma_n)}{Q(\|\mathbf{p}_1^{\perp}\|/\sigma_n)} = 1 \tag{4.38}$$

The asymptotic efficiency of the MMSE detector is therefore $\|\mathbf{p}_1^{\perp}\|^2$. Since this quantity is independent of the energies of the interference vectors, we also have that

$$\eta = \|\mathbf{p}_1^{\perp}\|^2 \tag{4.39}$$

That is, the near-far resistance of the MMSE detector considered is the squared norm of the component of the desired signal vector that is orthogonal to the space spanned by the interference vectors. From (4.39), it is clear that if $\eta > 0$, then the desired vector is not contained in the interference subspace $\mathbf{S}_I$, which in turn implies that the number of vectors contributed by the users $L \leq N$.

From the discussion in Section 4.4, we observe that the near-far resistance is closely related to the zero-forcing solution. Specifically, the zero-forcing solution is given by (4.30), which includes $\eta$ has a scale factor. Also, it is easily shown that the MSE corresponding to the zero-forcing solution is

$$\mathrm{MSE}_{zf} = \frac{\sigma_n^2}{\eta}, \tag{4.40}$$

so that the noise enhancement associated with the zero-forcing detector is $1/\eta$. Note that $0 \leq \eta \leq 1$ implies that $\sigma_n^2 \leq \mathrm{MSE}_{zf} \leq \infty$. In particular, if $\mathbf{p}_1$ lies in the space spanned by the interferers (i.e., if $\mathrm{rank}(\mathbf{S}_I) > N$), then $\eta = 0$ and $\mathrm{MSE}_{zf} = \infty$.

## 4.7 Space-Time Filtering

It is conceptually straightforward to extend the preceding discussion to combined space-time filtering. Given multiple antennas, the MMSE linear filter for user $k$ is shown in Figure 3.6 (for two antennas), and consists of a chip matched-filter and TDL for each antenna. The TDL outputs are simply added together to form the symbol estimate. To compute the TDL coefficients in terms of the received samples on each branch, we define $\mathbf{r}_i^{(m)}$ as the $N \times 1$ received vector of chip matched filter outputs for symbol $i$ on branch $m$. Then $\bar{\mathbf{r}}_i$ is the $(MN) \times 1$ vector consisting of $\mathbf{r}_i^{(1)}, \ldots, \mathbf{r}_i^{(M)}$ stacked on top of each other. (This assumes $1/T_c$ sampling, and that each TDL spans a single symbol. The generalizations to other sampling rates and to multi-symbol TDLs are straightforward.)

Let $\mathbf{c}^{(k)}$ be the vector of TDL coefficients associated with the $k$th branch. As before, the filter output can be expressed as

$$d_i = \bar{\mathbf{c}}^{\dagger} \bar{\mathbf{r}}_i \tag{4.41}$$

where $\bar{\mathbf{c}}$ is the $(MN) \times 1$ vector of TDL coefficient vectors $\mathbf{c}^{(1)}, \ldots, \mathbf{c}^{(M)}$ stacked on top of each other. The preceding expressions for the MMSE coefficients, zero-forcing solution, and performance measures can therefore be directly applied to this situation. Note, in particular, that with $M$ antennas, a necessary condition for the existence of the zero-forcing solution is that the number of "effective" users $L < MN$. For asynchronous DS-CDMA ($L = 2K - 1$)

adding an additional antenna therefore increases the number of strong interferers that can be (completely) suppressed by approximately $N/2$.

Increasing the amount of spatial diversity leads to a substantial increase in system capacity, but at the expense of additional complexity. Specifically, analog front-end filtering, as well as conversion to baseband (if necessary), is needed for each antenna element. The number of TDL coefficients also increases from $N$ to $MN$, which can adversely affect the performance of the adaptive algorithms discussed in the next section.

## 4.8 Effect of Multipath

Multipath will be discussed in more detail in Section 6. For now we note that reflections of the transmitted signal off of surrounding objects cause the received signal to consist of the sum of weighted and delayed versions of the transmitted signals:

$$r(t) = \sum_{k=1}^{K} \sum_{m=1}^{M_k} \alpha_{k,m} A_k b_k p_k(t - iT - \tau_{k,m}) + n(t) \tag{4.42}$$

where $M_k$ is the number of paths associated with user $k$, and $\tau_{k,m}$ and $\alpha_{k,m}$ are respectively the delay and the (complex) coefficient associated with path $m$ for user $k$. (Without loss of generality, we assume that $\tau_{k,m} \geq 0$, $m = 1, \ldots, M$; $k = 1, 2, \ldots, K$.) We can write the sampled received vector $\mathbf{r}_i$ defined earlier as

$$\mathbf{r}_i = \sum_{k=1}^{K} A_k \left[ \sum_{m=1}^{M_k} \alpha_{k,m} \left( b_{i,k} \mathbf{p}_k^+(m) + b_{i-1,k} \mathbf{p}_k^-(m) \right) \right] + \mathbf{n}_i \tag{4.43}$$

where $\mathbf{p}_{k,m}^+$ and $\mathbf{p}_{k,m}^-$ contain the chip matched-filter output samples within the time window spanned by $\mathbf{r}_i$ in response to the inputs $p_k(t - \tau_{k,m})$ and $p_k(t + T - \tau_{k,m})$, respectively. According to the discussion in Section 4.1, the vectors $\mathbf{p}_{k,m}^+$ and $\mathbf{p}_{k,m}^-$ can be computed according to (4.11) where $\tau_{k,m}$ replaces $\tau_k$.

Note that we can rewrite (4.43) as

$$\mathbf{r}_i = \sum_{k=1}^{K} A_k \left( b_{i,k} \mathbf{p}_k^+ + b_{i-1,k} \mathbf{p}_k^- \right) + \mathbf{n}_i \tag{4.44}$$

where

$$\mathbf{p}_k^{\pm} = \sum_{m=1}^{M_k} \alpha_{k,m} \mathbf{p}_k^{\pm}(m). \tag{4.45}$$

Consequently, the received vector can once again be expressed as (4.5), where the vectors in the sum are computed according to (4.11) and (4.45). The MMSE and zero-forcing solutions for $\mathbf{c}$, and performance measures previously discussed can then be directly applied. For DS-CDMA applications, it is typically assumed that the path delays for user 1, $\tau_{1,m}$, $m = 1, \ldots, M_1$, span at most a few chips. The intersymbol interference due to the multipath vectors $\mathbf{p}_1^-(m)$ is then quite small, and is typically ignored.[4]

[4] In high-data-rate systems, such as arise in some indoor wireless applications, ISI can be significant. Linear detection techniques for dealing jointly with ISI and MAI in such situations are developed in [89] and [128].

To summarize the preceding discussion, when multipath is present the geometric interpretation represented by Figure 4.3 applies, where the signal vectors are the *received* vectors, including the effect of multipath. We therefore conclude that the MMSE solution *coherently combines all multipath within the window spanned by the filter* $\mathbf{c}$. Of course, this interpretation applies in practice only when the MMSE solution can be accurately estimated. That is, the estimation algorithm must be able to compensate for the changing multipath amplitudes and phases of all strong users. Techniques for performing combined linear multiuser detection and channel tracking are developed in [85], [126] and [127]. These methods will be discussed briefly in the sequel.

# 5 Adaptive Algorithms

The expressions for the MMSE vector $\mathbf{c}$ given in the preceding section, (4.19) and (4.27), seem to indicate that the MMSE receiver requires explicit knowledge of all user and channel parameters (i.e., spreading sequences, relative timing, phase, amplitudes, and multipath parameters). In this section we show that the MMSE solution for $\mathbf{c}$ can be accurately estimated without this knowledge. In fact, if the user and channel parameters are time-invariant, then the algorithms in this section can estimate $\mathbf{c}_{mmse}$ to arbitrary accuracy (given a sufficient number of received vectors $\mathbf{r}_i$).

The adaptive algorithms in this section require either (i) a training sequence of transmitted symbols, which are known to the receiver for initial adaptation, or (ii) accurate knowledge of the received vector corresponding to the *desired* user ($\mathbf{p}_1$) and associated timing. In the absence of multipath, the latter knowledge, which is simply the spreading code of the desired user and associated timing, is also required by the matched filter receiver. When multipath is present, the received vector can be estimated with a RAKE receiver [90, Ch. 7].

Three categories of adaptive algorithms are presented herein. The first category consists of the conventional stochastic gradient and least squares algorithms well known in adaptive filtering [23],[32]. These have been applied to obtain MMSE symbol estimates for DS-CDMA in [5], [58] and [92]. (Prior to that, MMSE estimation applied to DS-CDMA was considered in [133].) Application of these techniques to narrowband TDMA systems with co- and adjacent-channel intereference is reported in [47],[48]. The algorithms in the second category are "blind" in the sense that a training sequence is not required. Instead, knowledge of the received vector $\mathbf{p}_1$ and associated timing is assumed. Finally, the algorithms in the third category are "subspace" algorithms, in which each received vector $\mathbf{r}_i$ is projected onto a lower dimensional subspace. These techniques are potentially useful when the dimension of the received vectors is much greater than the dimension of the signal subspace. This may be the case when (i) the processing gain is very large relative to the number of users, (ii) an adaptive antenna array is available with TDLs on each branch, or (iii) the filter $\mathbf{c}$ spans multiple symbol intervals.

## 5.1 Stochastic Gradient (LMS) Algorithm

The stochastic gradient or LMS (Least Mean Squares) algorithm has been successfully applied to many signal processing applications such as noise cancellation, equalization, echo

cancellation, and adaptive beamforming [23],[32]. The approach to adaptive interference suppression presented here is in fact analogous to adaptive equalization for single-user channels. The main difference between the two applications is that for adaptive equalization, the TDL must span multiple symbols to suppress intersymbol interference (ISI), but can have as few as one tap per symbol. In contrast, for interference suppression, the TDL must have multiple taps per symbol, but can span a single symbol interval. Of course, a TDL that spans multiple symbols with multiple taps per symbol can suppress both ISI and multiple-access interference.

Let $\mathbf{c}_i$ denote the TDL vector at symbol time $i$. The LMS algorithm for updating $\mathbf{c}_i$ is given by

$$\mathbf{c}_i = \mathbf{c}_{i-1} + \mu e_i^* \mathbf{r}_i \tag{5.1}$$

where

$$e_i = b_{i,1} - \mathbf{c}_{i-1}^\dagger \mathbf{r}_i \tag{5.2}$$

is the estimation error at time $i$, and $\mu$ is a constant step size, which controls the tradeoff between convergence speed and excess MSE due to random coefficient fluctuations about the mean. Because the LMS algorithm (5.1) assumes knowledge of the symbols $b_{i,1}$, it must be implemented as a *supervised* or *decision-directed* algorithm. In practice, $b_{i,1}$ must be generated via a training sequence for initial adaptation (supervision), after which the symbol estimates $\hat{b}_{i,1}$ are used (decision direction).

There have been numerous analyses of the convergence properties of the LMS algorithm (e.g., see [23],[32]). A detailed analysis of this algorithm for the DS-CDMA interference suppression application considered here is given in [59] (see also [31]). A summary of the main results, given a stationary set of interferers and channels, are as follows.

1. Assuming that the received vector $\mathbf{r}_i$ is statistically independent[5] from past vectors $\mathbf{r}_m$, $m < i$, for each $i$, it can be shown that

$$\delta \mathbf{c}_i = (\mathbf{I} - \mu \mathbf{R}) \delta \mathbf{c}_{i-1} \tag{5.3}$$

   where $\delta \mathbf{c}_i = E\{\mathbf{c}_i\} - \mathbf{c}_{mmse}$.The mean coefficient vector therefore converges exponentially to $\mathbf{c}_{mmse}$ according to $N$ normal modes. The time constant associated with the $n$th mode is $1 - \mu\lambda_n$, where $\lambda_n$ is the $n$th eigenvalue of $\mathbf{R}$.

2. An approximate analysis shows that the MSE remains bounded provided that the step-size

$$\mu < \frac{2}{\text{trace } (\mathbf{R})} \tag{5.4}$$

   If $\mathbf{R}$ is given by (4.21), then

$$\text{trace } (\mathbf{R}) = \sum_{k=1}^{K} A_k^2 + N\sigma_n^2. \tag{5.5}$$

[5]This independence assumption holds for synchronous DS-CDMA,but not for asynchronous DS-CDMA. Nevertheless, even for asynchronous DS-CDMA it gives substantial insight.

3. The asymptotic MSE achieved with the LMS algorithm is greater than the MMSE due to random coefficient fluctuations about the mean. Denoting the MMSE as $\varepsilon_{min}$, the *excess* MSE due to these fluctuations can be approximated as

$$\xi_{ex} = \varepsilon_{min} \frac{\frac{\mu}{2} \text{ trace } \mathbf{R}}{1 - \frac{\mu}{2} \text{ trace } \mathbf{R}} \tag{5.6}$$

where trace $\mathbf{R}$ is given by (5.5).

To interpret the preceding results, suppose that the vectors $\mathbf{p}_1, \ldots, \mathbf{p}_K$ in the sum (4.5) are orthonormal. In that case each of these vectors $\mathbf{p}_k$ is an eigenvector of $\mathbf{R}$ with associated eigenvalue $A_k^2 + \sigma_n^2$. The remaining eigenvectors of $\mathbf{R}$ form a basis for the $N - K$ dimensional subspace that is orthogonal to the signal space. Each of these eigenvectors is associated with eigenvalue $\sigma_n^2$. If $L < K$, then there are $N - K$ modes of convergence for $E\{\mathbf{c}_i\}$ associated with exponential decay factor $1 - \mu\sigma_n^2$. Typically, $\sigma_n^2$ is very small, so that convergence associated with these modes is very slow. If $\mathbf{c}_i$ is in (or close to) the signal space, then this is no problem, since the dominant modes of convergence lie in the signal space. However, if $\mathbf{c}_i$ lies outside the signal space (such as when the filter has converged to a set of users, and a user subsequently departs), then the excess MSE due to the component of $\mathbf{c}_i$ outside the signal space can take a very long time to disappear.

We also note from the preceding discussion that the LMS algorithm is adversely affected by a near-far situation in which $A_k$ is very large. Namely, there will be a slow mode corresponding to an exponential decay factor approximately equal to $1 - \mu/A_k^2$. Also, note that according to (5.4) and (5.5), the larger the interfering amplitudes, the smaller $\mu$ must be for stability. To ensure that $\mu$ satisfies the stability condition (5.4) in the presence of a changing interference environment, it is useful to *normalize* the step size by an estimate of the input power. Specifically, $\mu$ in (5.1) can be replaced by $\bar{\mu} = \mu/\hat{\xi}(i)$, where

$$\hat{\xi}(i) = w\hat{\xi}(i-1) + (1-w)\|\mathbf{r}_i\|^2 \tag{5.7}$$

is a moving average estimate of the input energy, and $w$ is the averaging constant.

## 5.2 Least Squares (LS) Algorithm

An alternative to the stochastic gradient method is to choose the vector $\mathbf{c}_i$ to minimize the least squares (LS) cost function

$$\varepsilon_{ls}(i) = \sum_{n=0}^{i} w^{i-n} |d_n - \hat{b}_{n,1}|^2 \tag{5.8}$$

where $d_n = \mathbf{c}^\dagger \mathbf{r}_n$, and $0 < w < 1$ is an exponential weighting factor that discounts past data. This weighting is important in nonstationary environments where the vector $\mathbf{c}_i$ is computed at each iteration $i$. The LS solution for $\mathbf{c}_i$ is

$$\mathbf{c}_i = \hat{\mathbf{R}}_i^{-1} \hat{\mathbf{p}}_{i,1} \tag{5.9}$$

where

$$\hat{\mathbf{R}}_i = \sum_{n=0}^{i} w^{i-n} \mathbf{r}_n \mathbf{r}_n^\dagger \tag{5.10}$$

and

$$\hat{\mathbf{p}}_{i,1} = \sum_{n=0}^{i} w^{i-n} b_{i,1}^* \mathbf{r}_n. \tag{5.11}$$

Note that $\hat{\mathbf{R}}_i$ and $\hat{\mathbf{p}}_{i,1}$ are estimates of $\mathbf{R}$ and $\mathbf{p}_1$, respectively. In the absence of noise, it is possible for $\hat{\mathbf{R}}_i$ to be singular. In that case any solution to the set of linear equations $\hat{\mathbf{R}}_i \mathbf{c}_i = \hat{\mathbf{p}}_{i,1}$ minimizes the LS cost function $\varepsilon_{ls}(i)$.

The LS criterion is deterministic, as opposed to the stochastic gradient cost criterion (MSE), which is defined in terms of a statistical expectation. In general, LS algorithms converge much faster than stochastic gradient algorithms, but are more complex to implement. Specifically, if the signal vectors $\mathbf{p}_1, \ldots, \mathbf{p}_L$ are linearly independent, and $L < N$, then in the absence of noise, the LS solution for the vector $\mathbf{c}_i$ gives $\varepsilon_{ls}(i) = 0$ for $i > L$, provided that the received vectors $\mathbf{r}_0, \ldots, \mathbf{r}_{L-1}$ used to compute $\mathbf{c}$ from (5.9) are linearly independent. (This implies that the MSE is zero as well.)

A precise analysis of the convergence properties of the LS algorithm in the presence of noise is quite difficult; however, a useful rule of thumb is that it typically takes approximately $2N$ iterations for the LS algorithm to converge (again assuming stationary noise and interference), where $N$ is the filter order. In contrast, the stochastic gradient algorithm typically takes between 5 to 10 times longer to reach steady-state performance, assuming the spread in eigenvalues of the matrix $\mathbf{R}$ is relatively small. Because the convergence rate of the LS algorithm is insensitive to this eigenvalue spread, a large eigenvalue spread, corresponding to a near-far situation, will lead to a more dramatic difference in performance. Some numerical examples that compare the performance of LS and LMS adaptive algorithms will be presented in Section 5.5.

A "Recursive" LS (RLS) algorithm computes the LS solution $\mathbf{c}_i$ for each $i$. In this case, the matrix inverse $\hat{\mathbf{R}}_i^{-1}$ can be propagated in time by using the matrix inversion lemma:

$$\hat{\mathbf{R}}_i^{-1} = \hat{\mathbf{R}}_{i-1}^{-1} - \frac{\mathbf{g}_i \mathbf{g}_i^\dagger}{1 + (\mathbf{r}_i^\dagger \mathbf{g}_i)} \tag{5.12}$$

where $w = 1$ and

$$\mathbf{g}_i = \mathbf{R}_{i-1}^{-1} \mathbf{r}_i. \tag{5.13}$$

The vector $\hat{\mathbf{p}}_{i,1}$ can also be updated recursively as

$$\hat{\mathbf{p}}_{i,1} = w \hat{\mathbf{p}}_{i-1,1} + b_{i,1}^* \mathbf{r}_i \tag{5.14}$$

Although the matrix inversion lemma substantially reduces the complexity of a recursive LS (RLS) algorithm, the update (5.12) is sensitive to numerical roundoff errors, and must therefore be closely monitored or stabilized in some manner. Also, the basic complexity per update is $\mathcal{O}(N^2)$ for RLS, as compared with $\mathcal{O}(N)$ for LMS. (The complexity of RLS can be mitigated through parallelization, as is discussed below.) As with the LMS algorithm, the RLS algorithm defined by (5.9), (5.12)-(5.14) requires estimates of the symbols $\{b_{i,1}\}$. This can be accomplished initially through a training sequence, and subsequently by switching to decision-directed mode.

Rather than compute $\mathbf{c}_i$ for each $i$, it is also possible to update $\mathbf{c}_i$ periodically using the most recent received data vectors. Specifically, a *block* LS algorithm computes $\mathbf{c}_i$ every

$B$ iterations using the data vectors $\mathbf{r}_{i-B+1}, \ldots, \mathbf{r}_i$. An iterative method for obtaining the estimates $b_{i,1}$ in decision-directed mode is described in [29]. The results in [29] indicate that this type of block decision-directed algorithm can sometimes perform significantly better than the RLS algorithm with exponential weighting.

## 5.3 Orthogonally-Anchored (Blind) Algorithms

The decision-directed algorithms presented in the preceding section generally require reliable symbol estimates. Numerical results indicate that the performance of these algorithms begins to degrade when the error rate exceeds 10% [29], [33]. Much higher error rates sustained over many symbols can potentially cause the algorithm to lose track of the desired user. For a mobile wireless channel this situation can occur when the desired user experiences a deep fade, or when a strong interferer suddenly appears. It is therefore desirable to have an adaptive algorithm that does not require symbol estimates. We refer to such an algorithm as a "blind" adaptive algorithm.

An approach to blind adaptation, which was presented in [31], is illustrated in Figure 5.1. The vector $\mathbf{c}$ at time $i$ is expressed as

$$\mathbf{c}_i = \mathbf{p}_1 + \mathbf{w}_i \tag{5.15}$$

where $\mathbf{w}_i$ is constrained to be *orthogonal* to $\mathbf{p}_1$ for all $i$. The filter output is then

$$d_i = \mathbf{c}_i^\dagger \mathbf{r}_i = b_{i,1} + \sum_{k=2}^{L} A_k [\mathbf{p}_1 + \mathbf{w}_i]^\dagger \mathbf{p}_k + (\mathbf{c}^\dagger \mathbf{n}_i) \tag{5.16}$$

where it is assumed that $A_1 \|\mathbf{p}_1\|^2 = 1$. Note that $\mathbf{w}_i$ affects only the interference and noise at the output. Selecting $\mathbf{w}_i$ to minimize the output *variance* $E\{|d_i|^2\}$ therefore minimizes the output interference plus noise energy. In fact, the output MSE is (again, we normalize $E\{|b_{i,1}|^2\} = 1$)

$$\begin{aligned} E\{|b_{i,1} - d_i|^2\} &= 1 + E\{|d_i|^2\} - 2\ \mathrm{Re}\{(\mathbf{p}_1 + \mathbf{w}_i)^\dagger \mathbf{p}_1\} \\ &= E\{|d_i|^2\} - 1 \end{aligned} \tag{5.17}$$

since

$$(\mathbf{p}_1 + \mathbf{w}_i)^\dagger \mathbf{p}_1 = \mathbf{c}^\dagger \mathbf{p}_1 = 1 \tag{5.18}$$

We therefore conclude that selecting $\mathbf{w}_i$ to minimize the output variance $E\{|d_i|^2\}$ also minimizes output MSE. Minimizing the output variance does not require knowledge of the symbol estimates $b_{i,1}$, although it does require knowledge of the desired user's vector $\mathbf{p}_1$ (and associated timing). We remark that this minimum variance technique is analogous to the minimum variance technique in adaptive beamforming where the direction of arrival of the desired signal is known [41].

The minimum variance vector $\mathbf{c}_{mv}$ can be derived by defining the Lagrangian

$$L(\mathbf{c}) = E\{|d_i|^2\} - \xi\ \mathbf{c}^\dagger \mathbf{p}_1, \tag{5.19}$$

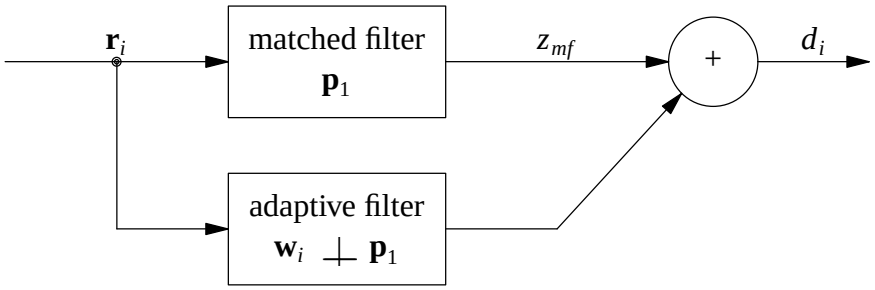


Figure 5.1: Orthogonally-anchored adaptive filter.

where $\xi$ is the Lagrange multiplier, and setting the gradient with respect to $\mathbf{c}$ equal to zero. This gives

$$\mathbf{c}_{mv} = \xi \ \mathbf{R}^{-1}\mathbf{p}_1 = \xi \ \mathbf{c}_{mmse} \tag{5.20}$$

where

$$\xi = E|\mathbf{c}_{mv}^{\dagger}\mathbf{r}_i|^2 = \frac{1}{\mathbf{p}_1^{\dagger}\mathbf{R}^{-1}\mathbf{p}_1} \tag{5.21}$$

is the constrained minimum output variance. If the signal vectors $\mathbf{p}_1, \ldots, \mathbf{p}_L$ are orthogonal, and $L < N$, then the mean output energy is $\xi = 1 + \sigma_n^2$.

Both stochastic gradient and LS adaptive algorithms can be derived based on the preceding minimum variance approach. Before deriving the stochastic gradient algorithm, we note that the constrained minimum variance cost function is the intersection of the quadratic form $|\mathbf{c}^{\dagger}\mathbf{p}_1|^2$ with the hyperplane defined by (5.18). This cost function has a unique global minimum, which can be found by gradient search.

Taking the gradient of the output energy with respect to $\mathbf{w}_i$ gives

$$\nabla_{\mathbf{w}_i}\left(E\{|d_i|^2\}\right) = 2 \ \mathrm{Re}\{E\{d_i^*\mathbf{r}_i\}\}. \tag{5.22}$$

To obtain the stochastic gradient algorithm, we drop the expectation, and take the orthogonal projection with respect to $\mathbf{p}_1$, which gives

$$d_i^*\left(\mathbf{r}_i - z_{mf}^*(i)\mathbf{p}_1\right) \tag{5.23}$$

where

$$z_{mf}(i) = \mathbf{p}_1^{\dagger}\mathbf{r}_i \tag{5.24}$$

is the matched-filter output. The orthogonally-anchored stochastic gradient algorithm is therefore

$$\mathbf{w}_i = \mathbf{w}_{i-1} - \mu d_i^*[\mathbf{r}_i - z_{mf}^*(i)\mathbf{p}_1] \tag{5.25}$$

The convergence properties of the algorithm (5.25) are analyzed in [31]. The main results, which parallel the results for the LMS algorithm in Section 3.1, are summarized as follows.

1. Defining

$$\delta\mathbf{c}_I = \mathbf{c}_i - \mathbf{c}_{mv} \tag{5.26}$$

and assuming that the received vector $\mathbf{r}_i$ is statistically independent from past vectors $\mathbf{r}_m$, $m < i$, for each $i$, it can be shown that

$$\delta \mathbf{c}_i = (\mathbf{I} - \mu \mathbf{R}_{vr}) \delta \mathbf{c}_{i-1} \tag{5.27}$$

where

$$\mathbf{v}_i = (\mathbf{I} - \mathbf{p}_1 \mathbf{p}_1^\dagger) \mathbf{r}_i \tag{5.28}$$

and

$$\mathbf{R}_{vr} = E\{\mathbf{v}_i \mathbf{r}_i^\dagger\} = (\mathbf{I} - \mathbf{p}_1 \mathbf{p}_1^\dagger) \mathbf{R} \tag{5.29}$$

The mean coefficient vector therefore converges exponentially to $\mathbf{c}_{mv}$ according to $N$ normal modes, associated with the eigenvalues of $\mathbf{R}_{vr}$.

2. An approximate analysis shows that the MSE remains bounded provided that the step-size satisfies (5.4).

3. The excess MSE, defined as the asymptotic MSE minus the MSE associated with $\mathbf{c}_{mv}$, can be approximated as

$$\xi_{ex} = \xi_{min} \frac{\frac{\mu}{2} \text{ trace } \mathbf{R}_{vr}}{1 - \frac{\mu}{2} \text{ trace } \mathbf{R}_{vr}} \tag{5.30}$$

where

$$\text{trace } \mathbf{R}_{vr} = \sum_{k=1}^{K} A_k^2 (1 - |\rho_{1k}|^2) + (N-1)\sigma_n^2 \tag{5.31}$$

and $\rho_{1k} = \mathbf{p}_1^\dagger \mathbf{p}_k$.

If the vectors $\mathbf{p}_1, \ldots, \mathbf{p}_L$ in the sum (4.5) are orthonormal, then each of these vectors is an eigenvector of $\mathbf{R}_{vr}$. The eigenvalue associated with $\mathbf{p}_1$ is zero, whereas the eigenvalue associated with $\mathbf{p}_k$, $k > 1$, is $A_k^2 + \sigma_n^2$. The remaining eigenvectors of $\mathbf{R}$ form a basis for the $N - K$ dimensional subspace that is orthogonal to the signal space. Each of these eigenvectors is associated with eigenvalue $\sigma_n^2$.

Given orthogonal signal vectors, the eigenvalues of $\mathbf{R}_{vr}$ are nearly the same as those for $\mathbf{R}$. The convergence of the mean coefficient vector should therefore be similar for both the minimum variance and standard LMS algorithms, given the same step-size $\mu$. However, the excess MSE given by (5.30) is substantially larger than the corresponding MSE for the LMS algorithm (5.6). Specifically,

$$\frac{\xi_{ex}^{(mv)}}{\xi_{ex}^{(lms)}} = \frac{\xi_{min}}{\varepsilon_{min}} \tag{5.32}$$

When the signal vectors are approximately orthogonal, it is easily shown that

$$\frac{\xi_{min}}{\varepsilon_{min}} \approx (1 + \sigma_n^2) \cdot \frac{1 + \sigma_n^2}{\sigma_n^2} \tag{5.33}$$

which can be quite large. Consequently, the blind algorithm (5.25) is quite "noisy", and it is best to switch to a decision-directed algorithm once reliable symbol estimates are available.

An LS minimum variance adaptive algorithm is obtained by selecting $\mathbf{c}_i$ to minimize the cost function

$$V(i) = \sum_{n=0}^{i} w^{i-n} |\mathbf{c}_i^\dagger \mathbf{r}_i|^2 \tag{5.34}$$

subject to the constraint $\mathbf{c}_i^\dagger \mathbf{p}_1 = 1$. The solution is given by

$$\mathbf{c}_i = \hat{\xi}_i \hat{\mathbf{R}}_i^{-1} \mathbf{p}_1 \tag{5.35}$$

where $\hat{\mathbf{R}}_i$ is given by (5.10), and

$$\hat{\xi}_i = \left( \mathbf{p}_1^\dagger \hat{\mathbf{R}}_i^{-1} \mathbf{p}_1 \right)^{-1} \tag{5.36}$$

The LS solution for $\mathbf{c}_i$ therefore has the same form as (5.20), where expectations are replaced by time averages.

It is interesting to compare the minimum variance LS solution (5.35) with the decision-directed LS solution (5.9). The only differences are: (i) The scale factor $\hat{\xi}_i$ appears in (5.35), and (ii) $\hat{\mathbf{p}}_1$ in (5.9) is replaced by $\mathbf{p}_1$ in (5.35). If the magnitude of $b_{i,1}$ is constant for each $i$, which corresponds to phase modulation, then the scale factor $\hat{\xi}_i$ in (5.35) is irrelevant. That is, omitting $\hat{\xi}_i$ does not affect the error rate. In that case, replacing $\hat{\mathbf{p}}_1$ by $\mathbf{p}_1$ is the only real difference between the minimum variance and decision-directed LS algorithms. As a first-order approximation, the minimum variance LS algorithm performs the same as the decision-directed LS algorithm.

Both block and RLS versions of the minimum variance LS algorithm are possible, depending on how often the matrix $\hat{\mathbf{R}}_i^{-1}$ is updated. The matrix inversion lemma can again be applied to the RLS version to reduce the amount of computation.

A potential problem with the minimum variance approach is that the vector $\mathbf{p}_1$ that appears in the adaptive algorithm may not be exactly equal to the received vector contributed by user 1. This may be due to unknown multipath or other types of distortion. This type of receiver *mismatch* can cause a substantial degradation in performance due to suppression of the desired signal.

To illustrate the mismatch problem, suppose that the actual received vector contributed by user 1 is $\mathbf{p}_1$, but that the receiver uses the mismatched estimate $\tilde{\mathbf{p}}_1$. These two vectors are shown in Figure 5.2. According to Figure 5.2, it is possible to choose a vector $\mathbf{w}$, which is orthogonal to $\tilde{\mathbf{p}}_1$, and such that $\tilde{\mathbf{p}}_1 + \mathbf{w}$ is orthogonal to $\mathbf{p}_1$. Consequently, with mismatch, the minimum variance approach attempts to suppress the desired signal down to the level of the interference.

Figure 5.2 indicates that the closer $\tilde{\mathbf{p}}_1$ is to $\mathbf{p}_1$, the longer $\mathbf{w}$ must be to suppress the desired signal. Consequently, one way to mitigate the effect of mismatch is to constrain the length of the vector $\mathbf{w}$. Referring to Figure 5.2, let $\theta_k$ denote the angle between $\mathbf{p}_k$ and $\mathbf{p}_1$, and let $\tilde{\theta}$ denote the angle between $\tilde{\mathbf{p}}_1$ and $\mathbf{p}_1$. Typically, we expect that $\theta_k \gg \tilde{\theta}$. As shown in Figure 5.2, $\|\mathbf{w}\|$ corresponding to the $\mathbf{w}$ needed to suppress $\mathbf{p}_k$ is much less than $\|\mathbf{w}\|$ corresponding to the $\mathbf{w}$ needed to suppress the desired signal. We can therefore mitigate the effect of mismatch by incorporating the constraint

$$\|\mathbf{w}_i\|^2 < \chi \tag{5.37}$$

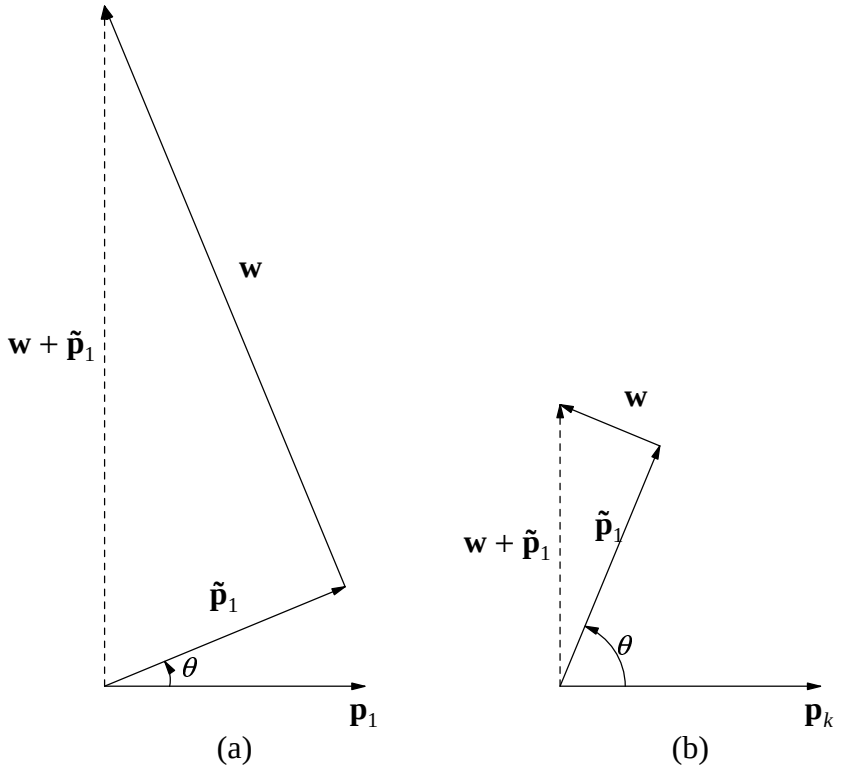


Figure 5.2: Illustration of desired signal suppression with mismatch. (a) shows the vector $\mathbf{w}$ needed to suppress the desired signal, and (b) shows the vector $\mathbf{w}$ needed to suppress an interfering signal $\mathbf{p}_k$.

where $\chi$ is a constant, into the adaptive algorithm. From the preceding discussion, a reasonable choice for $\chi$ is the length of $\mathbf{w}$ needed to suppress the user $k$ corresponding to the smallest value of $\theta_k$. Further discussion and numerical results illustrating how the choice of $\chi$ affects the performance of the minimum variance approach is given in [31].

The constraint (5.37) is easily incorporated into the minimum variance approach, and results in a vector $\mathbf{c}_{mv}$ that again has the form (5.20). The only difference is that the noise variance $\sigma_n^2$, which appears in the definition of $\mathbf{R}$ (4.21), is replaced by $\sigma_n^2 + \nu$, where $\nu$ is a Lagrange multiplier selected to satisfy (5.37). The constraint (5.37) therefore has the same effect on $\mathbf{c}_{mv}$ as increasing the background noise variance. Similarly, incorporating the constraint (5.37) into the LS optimization results in the solution (5.35) where

$$\hat{\mathbf{R}}_i = \sum_{n=0}^{i} w^{i-n} \mathbf{r}_i \mathbf{r}_i^\dagger + \nu \mathbf{I} \tag{5.38}$$

Finally, incorporating this constraint into the stochastic gradient algorithm (5.25) results in the algorithm

$$\mathbf{w}_i = (1 - \mu\nu)\mathbf{w}_{i-1} - \mu d_i^* [\mathbf{r}_i - z_{mf}^*(i)\mathbf{p}_1] \tag{5.39}$$

which is analogous to the tap-leakage algorithm introduced in [21].

Another possible solution to the mismatch problem, presented in [28], is to combine the orthogonally-anchored approach with the decision-directed cost function $E\{|d_i - \mathrm{sgn}(d_i)|^2\}$

(assuming $b_{i,1} \in \{\pm 1\}$). This cost function eliminates the problem of desired signal suppression; however, it can introduce local optima. The stochastic gradient algorithm based on this approach has been observed to perform somewhat better than the analogous minimum variance algorithm in the presence of mismatch.

## 5.4 Projection-Based Approaches

The discussion in preceding sections indicates that the performance (convergence speed) of the adaptive algorithms discussed degrades as the number of filter coefficients increases. Furthermore, increasing the number of filter coefficients generally increases the complexity of the adaptive algorithms. In some situations, it may be desirable to have a TDL $\mathbf{c}$ with high dimensionality. For example, $\mathbf{c}$ may include TDLs on multiple antennas and/or may span many symbols. Also, some military applications require a very large processing gain $N$ for covertness. In these situations it is desirable to reduce the number of *adaptive* coefficients.

One way to reduce the number of adaptive coefficients is to project the received vectors onto a lower dimensional subspace. Specifically, let $\mathbf{S}_D$ be the $N \times D$ matrix with columns vectors that are the basis vectors for a $D$-dimensional subspace, where $D < N$. We wish to restrict $\mathbf{c}$ to lie in this subspace, so we can write

$$\mathbf{c} = \mathbf{S}_D \boldsymbol{\alpha} \tag{5.40}$$

where $\boldsymbol{\alpha}$ is a $D \times 1$ vector of coefficients that must be estimated. Given $\mathbf{S}_D$, it is straightforward to derive stochastic gradient and LS algorithms for estimating $\boldsymbol{\alpha}$. (Note that $\mathbf{c}^\dagger \mathbf{r}_i = \boldsymbol{\alpha}^\dagger \tilde{\mathbf{r}}_i$, where $\tilde{\mathbf{r}}_i = \mathbf{S}_D^\dagger \mathbf{r}_i$ is the projected received vector.)

A few different suggestions for the lower dimensional subspace represented by $\mathbf{S}_D$ have been proposed [51], [101], [104] and [129]. For example, in [101] the columns of $\mathbf{S}_D$ are taken to be nonoverlapping segments of the desired spreading sequence, where each segment is of length $N/D$. (The interpretation is that partial despreading is performed before the adaptive filtering.) Specifically,

$$[\mathbf{S}_D]'_m = [0 \cdots 0 \, \tilde{\mathbf{p}}'_1(m) \, 0 \cdots 0] \tag{5.41}$$

where $1 \le m \le D$,

$$\tilde{\mathbf{p}}'_1(m) = [\mathbf{p}_{1,(m-1)N/D+1}, \cdots, \mathbf{p}_{1,mN/D}], \tag{5.42}$$

$(m-1)N/D$ zeros precede $\tilde{\mathbf{p}}'_1$, $(D-m)N/D$ zeros succeed $\tilde{\mathbf{p}}'_1$, and $N/D$ is assumed to be an integer. Note that $D = N$ corresponds to the MMSE detector previously discussed ($N$ adaptive coefficients), and $D = 1$ corresponds to the matched-filter detector. Choosing $D$ between 1 and $N$ therefore trades off complexity ($D$ adaptive coefficients) with performance (which is between that of the matched-filter and that of MMSE detectors).

If the dimension of the signal space $\mathbf{S}$ is less than the dimension of $\mathbf{c}$, then projecting the received vectors onto the signal space reduces the number of adaptive coefficients without sacrificing optimality (cf., [129]). Generally, this reduction in the number of adaptive components will improve convergence and tracking. Signal subspace methods have received considerable attention in the array processing literature (see [41] and the references within). If the dimension of the signal space is known to be $L$, then an orthogonal basis for the signal space is given by the $L$ eigenvectors of $\mathbf{R}$ that correspond to the $L$ largest eigenvalues. In practice, a basis for the signal space can be estimated by forming an eigen-decomposition

of the matrix $\hat{\mathbf{R}}_i$ given by (5.10). The columns of $\mathbf{S}_D$ in (5.40) are then the eigenvectors corresponding to the $D$ largest eigenvalues.

The dimension of the signal space is typically unknown *a priori*, so that $D$ can either be fixed in advance, or be selected as a consequence of the threshold rule

$$\begin{array}{ll} \lambda_k(\hat{\mathbf{R}}_i) > \Lambda, & \longrightarrow \text{ include } \mathbf{v}_k \\ \lambda_k(\hat{\mathbf{R}}_i) < \Lambda, & \longrightarrow \text{ discard } \mathbf{v}_k \end{array} \tag{5.43}$$

where $\mathbf{v}_k$ is the eigenvector associated with $\lambda_k$, and $\Lambda$ is a constant. More sophisticated alternative dimension estimation techniques can also be used [129].

Interference suppression based on a subspace decomposition is discussed in several works, including [22], [29], [35], [109], and [129].Timing estimation for DS-CDMA based on an analogous type of subspace decomposition is presented in [8], [105] and [129]. From the viewpoint of adaptivity the eigen-decomposition needed to estimate the signal space nominally defeats any reduction in complexity achieved by reducing the number of adaptive coefficients. However, very recent work in [129] has shown that a low-complexity *subspace tracking* algorithms of $\mathcal{O}(KN)$ complexity per update can be used to provide subspace-based adaptivity with practical levels of complexity.

## 5.5 Numerical Examples

In this section we present simulation results that illustrate the performance of some of the adaptive interference suppression algorithms discussed in the preceding subsections. We first present some convergence results assuming a stationary environment and synchronous users. Figure 5.3 shows averaged Signal-to-Interference Plus Noise Ratio (SINR) as a function of time for each of the following algorithms: (i) decision-directed LMS, (ii) decision-directed Recursive Least Squares (RLS), and (iii) orthogonally-anchored (blind) stochastic gradient. (The performance of the orthogonally-anchored RLS algorithm is nearly the same as the decision-directed algorithm.) The curve for the blind algorithm assumes perfect knowledge of the received pulse shape from the desired user (no mismatch). The convergence curves are obtained by averaging 400 simulation runs, assuming that the spreading codes assigned to all users are fixed. The received amplitudes corresponding to the interferers are twice that of the desired signal. The processing gain is $N = 10$, there are 7 users, and the Signal-to-Noise Ratio is 12 dB. In each case the filter is initialized as the matched filter.

This example shows that the LS algorithm converges much faster than the stochastic gradient algorithms. Specifically, the LS algorithm converges in approximately 50 iterations, whereas the stochastic gradient algorithms require approximately 700 iterations to converge within 1 dB of the steady-state SINR. The step-sizes for both the blind stochastic gradient and LMS algorithms are the same, so that the steady-state SINR for the blind algorithm is somewhat lower than the steady-state SINR for the decision-directed LMS algorithm. This difference in steady-state SINR becomes more significant as the power of the interferers increases [31]. It is therefore desirable to switch to decision-directed mode whenever decisions are reliable.

Figure 5.4 shows convergence plots for the blind stochastic gradient and LS algorithms in the presence of mismatch. The mismatch was created by adding a single multipath

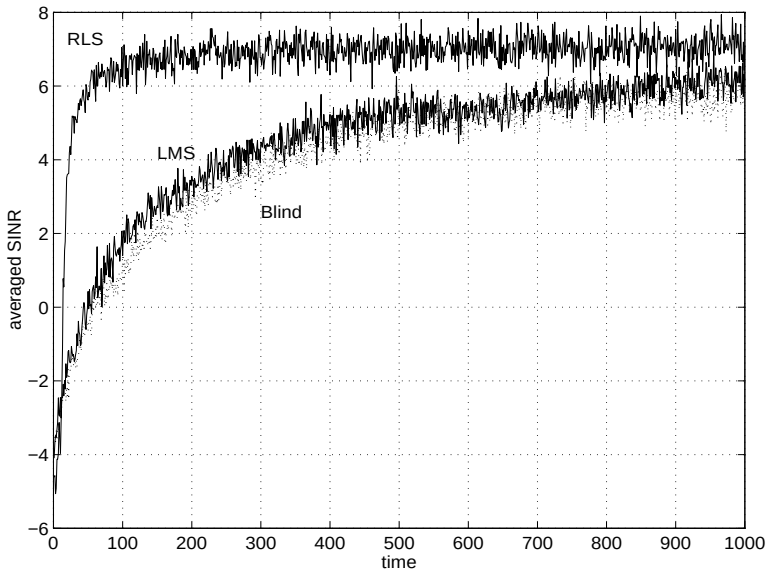


Figure 5.3: Averaged SINR vs. time for the decision-directed LMS, decision-directed RLS, and orthogonally-anchored (blind) stochastic gradient algorithms.

component offset by one chip, and attenuated by 3 dB relative to the main component. The blind algorithms are anchored to the strongest path. In each case, the algorithm switches to decision-directed mode after a fixed number of iterations (100 for the LS algorithm and 300 for the stochastic gradient algorithm). The blind algorithms are able to improve the SINR initially, but then subsequently suppress the signal. (Figure 5.4 does not show this, since the algorithms switch to decision-directed mode before the SINR starts to decrease.) The relatively slow convergence of the LS algorithm is due to the large term added to the diagonal of the matrix $\hat{\mathbf{R}}_i$, given by (5.38), which constrains the length of the adaptive filter vector. (Referring to (5.38), for this example $\nu = 50$.) Note that the additional multipath component in this example improves the asymptotic SINR relative to that shown in Figure 5.3.

Figure 5.5 shows the how the arrival of a new strong interferer affects the performance of the decision-directed RLS algorithm. The scenario used to generate Figure 5.5 is the same as that used to generate Figure 5.3, except that there are only 6 users initially, and the new (7th) user appears at iteration 200. The power of the new user is 18 dB above the desired user, representing a severe near-far situation.

This example shows that the performance (average SINR) of the LS algorithm is temporarily degraded by the appearance of the new user. This degradation in performance is largely due to the sudden transient created by allowing the new user to begin transmitting with full power. Also shown in Figure 5.5 is the performance curve corresponding to the situation where the new user gradually increases the transmitted power (linearly in dB) to the power limit within 50 iterations. Although this mitigates the transient performance degradation, it does add overhead in the form of additional training. (In a packet data system this additional overhead must be included in each packet.) Adaptive techniques that detect the appearance of a new user are discussed in [34] and [64]. In principle, this additional information can be used to mitigate the degradation in performance caused by the associated

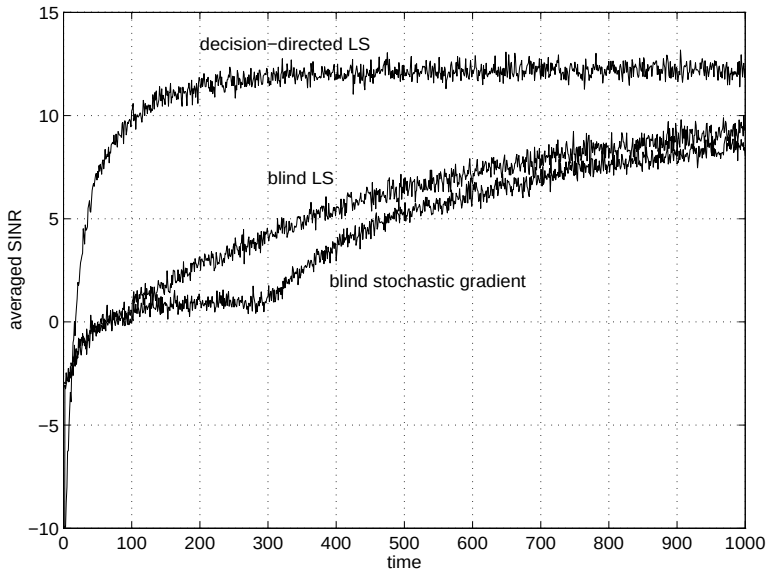


Figure 5.4: Averaged SINR vs. time for the orthogonally-anchored (blind) LS and stochastic gradient algorithms with a mismatched anchor. The blind LS and stochastic gradient algorithms switch to decision-directed mode at times 100 and 300, respectively. Also shown is SINR vs. time for the decision-directed RLS algorithm.

transient without additional overhead.

Finally, Figure 5.6 shows a comparison of the subspace tracking algorithm of [129] with RLS. This algorithms makes use of the Projection Approximation Subspace Tracking - dilation (PASTd) of [134]. Note that this simulation illustrates the potential gain in SINR that can be obtained by reducing the number of dimensions to be adapted from $N$ to $K$. (In this example, we have increased the dimension of the received signal to $N = 31$, while keeping the number of users small to illustrate the dimension-reduction advantages of subspace methods.)

.

# 6 Further Issues and Refinements

In the preceding three sections, we have considered basic elements of adaptive linear multiuser detection. The actual application of these methods requires consideration of a number of further issues, on which we touch in this section. In particular, we discuss briefly some salient features of the mobile wireless communications environment and some additional issues arising with adaptive interference suppression and multiuser detection in this context. Many of these issues apply to multiuser detection in general, although the focus of our discussion is on adaptive linear interference suppression.

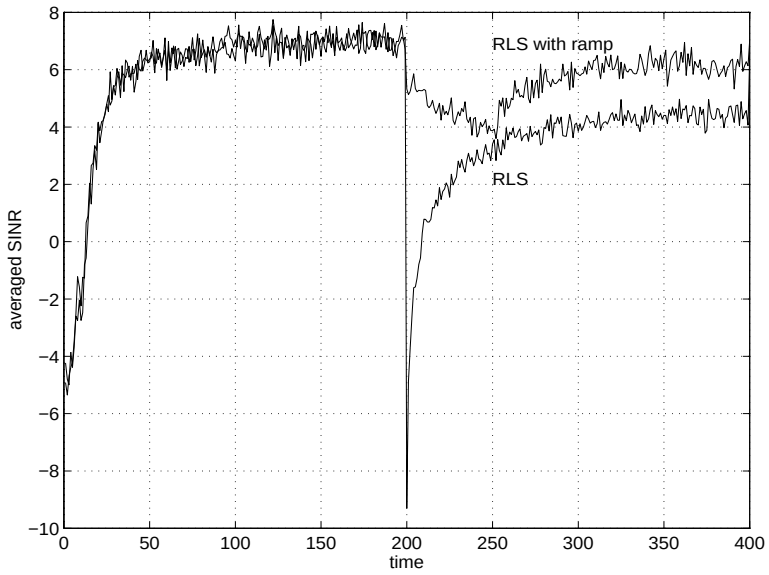


Figure 5.5: Averaged SINR vs. time for the RLS algorithm. A new user appears at time 200. The curve marked "ramp" corresponds to the situation where the new user increases the transmitted power in equal increments (in dB) for a period of 50 iterations.

## 6.1 The Mobile Wireless Environment

In the preceding sections, we have treated the adaptive multiuser detection problem primarily for the situation in which the parameters of the environment are essentially stable. (An exception is the discussion concerning Figure 5.5, in which the user population changes with time.) One of the primary challenges to multiuser detection offered by the mobile wireless environment is that essentially all user parameters, such as received pulse shapes, amplitudes, relative carrier phase, timing, and whether or not a particular user is active, are time-varying. Consequently, these parameters must be estimated, either directly or indirectly, and the detector must be robust with respect to inaccurate estimates. Also, the ideal model (1.1) - (1.2) of a multiple-access data signal observed in white Gaussian noise is not necessarily accurate for many situations. Thus, for the practical use of adaptive interference suppression methods, many aspects of channel behavior beyond those described in the preceding sections must be considered. In this section we briefly describe some of these impairments present in mobile wireless channels. The purpose of this discussion is just to present the models of channel impairments that are commonly used in the literature. More detailed treatments of propagation along with justification for the channel models can be found, for example, in [40] or [94].

### 6.1.1 Distance-Related Attenuation and Shadowing

For terrestrial wireless communications, the received signal strength associated with a particular user in general depends on (1) the transmitted power, (2) the distance between the transmitter and receiver, (3) the presence of large objects, such as buildings, foliage, or vehicles, that lie between the transmitter and receiver line-of-sight, and (4) the relative am-

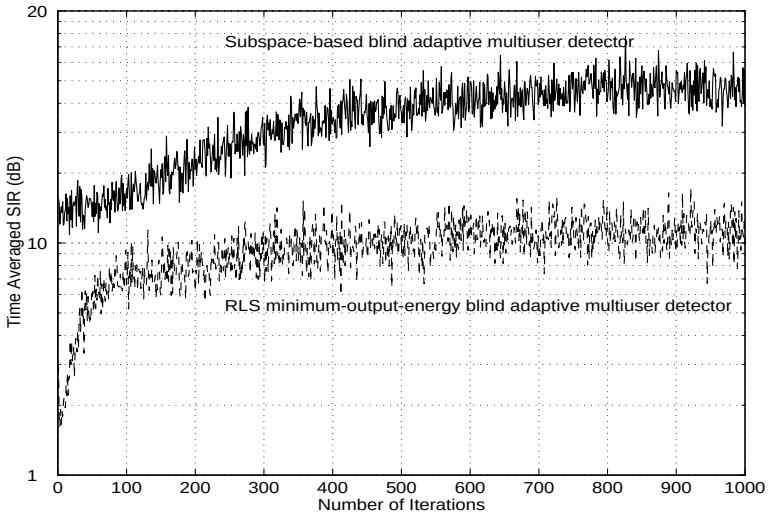


Figure 5.6: Comparison of SINR vs. time for the subspace-tracking and RLS algorithms ($N = 31$ and $K = 6$). Here there are four MAI's 10dB above the intended user, and one MAI 20dB above the intended user. The post-despreading SNR of the intended user is 20dB.

plitudes and phases of received paths associated with scattering off of surrounding objects. The dynamic variation in signal strength due to the motion of the transmitter relative to the receiver (or vice versa) is called *fading*. The fade *rate* is the rate at which the signal experiences fades, and depends on the speed of the mobile.

The second and third items listed above are considered *large-scale* effects, whereas item (4) is a *small-scale* effect [94]. Large-scale effects determine the mean signal strength averaged over a region spanning a few wavelengths in each direction. These cause relatively slow variations in the (mean) signal strength as a mobile moves through space. Small-scale effects cause large swings in signal strength over just a fraction of a wavelength, and are superimposed on top of the large-scale effects. As a transmitter moves relative to the receiver, the *mean* received signal strength therefore varies relatively slowly, but the actual signal strength may experience large rapid variations (i.e., 20 to 40 dB) around the mean.

Given an isolated transmitter and receiver in free space separated by distance $d$, the received signal power is inversely proportional to $d^2$. Although the presence of buildings and other objects greatly complicates the modes of radio wave propagation, both analysis and measurements indicate that the loss in signal strength, or *path loss*, is proportional to $d^n$, where $n$ is an integer. The value of $n$ is typically chosen between 2 and 5, depending on the environment considered. In general, the denser the urban environment, the greater the path loss exponent. The value $n = 4$ is typically assumed for modeling urban cellular systems.

In addition to the deterministic path loss due to distance, there is a random component due to the location-dependent spatial distribution of objects relative to the mobile. That is, the path loss experienced by the mobile depends on both the separation from the transmitter and on the particular placement of surrounding objects that may prevent line-of-sight communications. This latter effect is called *shadowing*. Measurements have shown that the

random variations in path loss around the distance-dependent mean can be modeled as a log-normal random variable. That is, the received strength, measured in dB, has a Gaussian distribution with mean specified by the distance-dependent path loss, and standard deviation $\sigma$ also given in dB. A typical value of $\sigma$ for urban cellular environments is 8 dB.

Based on the preceding discussion, we can write the received power, taking into account distance-based attenuation and shadow fading, as

$$\bar{P}(d) = \bar{P}_0 \xi \left(\frac{d_0}{d}\right)^n \tag{6.1}$$

where $P_0$ is the benchmark received power at distance $d_0$, and $\xi$ is a log-normal Gaussian variable with probability density

$$p_\xi(x) = 10^{-x/10}. \tag{6.2}$$

### 6.1.2 Multipath

As noted in Section 4.8, multipath is caused by scattering and/or reflections of the transmitted signal off of surrounding objects. Given a complex baseband transmitted signal $s(t)$, the effect of multipath is to produce the sum of many delayed and weighted versions of the transmitted signal. Specifically, the received signal (in the absence of noise) is given by

$$y(t) = \sum_{m=1}^{M} a_m s(t - \nu_m) \tag{6.3}$$

where each term in the sum corresponds to a different path, $M$ is the total number of paths, and $a_m$ and $\nu_m$ are the the path weight and delay associated with path $m$. Given a complex baseband transmitted signal $s(t)$, the path weights $a_m$ are also complex in general. (Note that the multiuser multipath signal of (4.42) consists of the superposition of $K$ such signals, in which for the $k^{th}$ signal we have $M = M_k$, $a_m = \alpha_{k,m} A_k$, and $\nu_m = \tau_{k,m}$.)

If the delays $\nu_m$ in (6.3) are sufficiently large, then the paths represented by the terms in the sum in (6.3) are said to be *resolvable*. That is, the receiver is able to distinguish the different paths, and possibly combine them. For two paths to be resolvable, the relative time delay between them, $\nu_m$, must be greater than $1/W$, where $W$ is the signal bandwidth. In urban environments where significant scattering occurs, each path in (6.3) generally represents the sum of many "micro-paths", which arrive within the resolution time. The relative phases of these scattered paths cause the path weight $a_m$ to fluctuate randomly. If there is no line-of-sight path, as is often the case in an urban environment, then the real and imaginary parts of $a_m$ are typically modeled as Gaussian random variables. In this case, the magnitude of $a_m$ has the Rayleigh probability density

$$p_R(r) = \begin{cases} \frac{r}{\sigma_R^2} e^{-r^2/(2\sigma_R^2)} & r \geq 0 \\ 0 & r < 0 \end{cases} \tag{6.4}$$

where $\sigma_R^2$ determines the mean and variance. The phase of $a_m$ is uniformly distributed. Adding a line-of-sight component to the received path results in an envelope that has a Ricean distribution.

From (6.3), the transfer function of a single-user multipath channel can be written as

$$H(f) = \sum_{m=1}^{M} a_m e^{-2\pi f \nu_m} \tag{6.5}$$

If there is only a single resolvable path, then the magnitude of $H(f)$ is the magnitude of $a_1$, which is independent of frequency. This type of channel is called a "flat fading" channel, since the fading occurs uniformly across the entire signal bandwidth. If there is more than one resolvable path, then the magnitude of $H(f)$ depends on $f$, so that this type of channel is called a "frequency-selective" fading channel.

The average received power corresponding to each path specifies the *multipath power delay profile* of the multipath channel. This power delay profile can vary considerably, depending on the mobile environment (e.g., rural, urban, hilly, etc.). An assumption that is sometimes made for transmission of DS-CDMA signals is that the power delay profile is continuous and decays exponentially. If the received signal is sampled at the chip rate, then the channel is modeled by (6.5) where the delays are integer multiples of the chip duration.

It remains to describe how the multipath channel varies with time. The resolvable paths, associated with the delays $\nu_m$, tend to change slowly in comparison with the coefficients $a_m$. Consequently, it is reasonable to assume that the paths are fixed, but that the coefficients $a_m$ are time-varying. The time-variation in the coefficient $a_m$ is due to Doppler shift, which causes the phases associated with all of the unresolvable paths contributing to path $i$ to vary with time. If the mobile is receiving a carrier with wavelength $\lambda$, and is traveling with velocity $v$ at an angle $\theta$ relative to the transmitter, the change in frequency, or Doppler shift, is $f_d = (v/\lambda)\cos\theta$.

The effect of the Doppler shift is to cause each coefficient $a_m$ to rotate. This assumes, however, that the mobile receives a single path coming from a specific direction. If the mobile receives many such paths, arriving at different angles, then the time variations of the multipath coefficients becomes more complicated. In that case, $a_m(t)$ is modeled as a random process. For urban mobile cellular systems, uniform scattering is often assumed, which means that the received spatial power density is a constant function of angle. In that case it can be shown that the power spectral density associated with $a_m(t)$ is given by

$$S_D(f) = \frac{K}{\sqrt{f_d^2 - f^2}} \tag{6.6}$$

for $|f| < f_d$, where $K$ is a constant that determines the power of the random process, and $f_d = v/\lambda$ is the maximum Doppler shift.

For the case of uniform scattering, the coefficients $a_m(t)$ can therefore be modeled as the output of a filter with frequency response $\sqrt{S_D(f)}$ in response to a complex white Gaussian noise input [94]. An alternative, known as Jake's model [40], is to generate $a_m(t)$ by summing complex sinusoids at different frequencies between $-f_d$ and $f_d$, each weighted by the associated value of $\sqrt{S_D}$. An example of a Rayleigh fading process, which was generated according to the first method, is shown in Figure 6.1.

Consider a wireless multiple access channel in which each user is subject to flat fading. In principle, the MMSE solution for the time-varying linear filter (for a desired user) must take

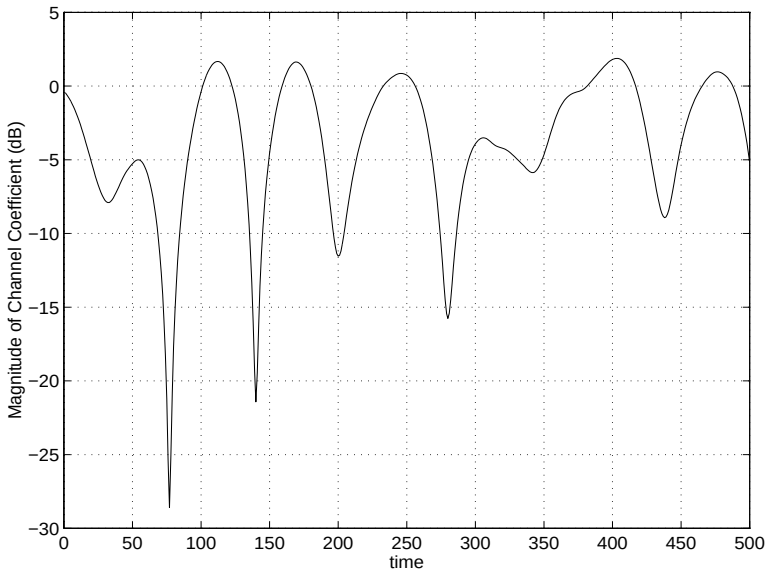


Figure 6.1: Sample path of the magnitude of a Rayleigh fading process.

into account the channels associated with all users. However, if the number of users is much less than the processing gain, and the background Signal-to-Noise Ratio is very high, then the MMSE solution can be approximated by the zero-forcing solution, which does not depend on the channel coefficients. That is, the space spanned by the interferers does not depend on the complex channel coefficients. The adaptive algorithm is therefore relieved from the task of tracking the channels associated with the interferers. Furthermore, the adaptive algorithm does not need to track the flat fading channel associated with the desired user when either differential detection or a pilot signal is used. Consequently, we conclude that for flat fading channels, when the number of users is small relative to the processing gain, the performance of adaptive algorithms should be insensitive to the fade rate (provided that the desired user's channel can be tracked). Simulation results that support this observation are presented in [33].

Now consider the case where each user experiences frequency selective fading. Recall that the received signal vector contributed by user $k$ after chip-matched filtering and sampling is given by (4.45), and depends on the time-varying complex coefficients associated with each path. If each path fades independently, then the interference space is *time-varying*, so that the adaptive algorithm attempts to track the time-varying multipath coefficients associated with *all* users. If the fade rate is sufficiently fast, then the adaptive algorithm is unable to track the combined set of paths for each user, and attempts to suppress each path individually. This, however, degrades the performance of the adaptive filter, since it effectively treats each path as a separate interferer. In the worst case, the multipath contributed by the interferers exceeds the number of available dimensions (e.g., the processing gain) that the filter has to suppress interferers, and the performance becomes equivalent to the matched filter. Tracking is therefore a critical issue for frequency-selective fast fading channels.

Results showing the performance of adaptive interference suppression algorithms in the

context of DS-CDMA with Rayleigh fading channels are presented in [33], [60], [85], [126], and [127]. In [60] a phase predictor is combine with differential coding and detection, whereas in [85], [126] and [127] phase prediction is combined with coherent detection by using a training sequence that must be transmitted periodically. A differential LS algorithm that does not rely on phase prediction is described in [33], which also shows performance results for a cellular type of model with flat Rayleigh fading channels.

### 6.1.3 Delay

Another time variation associated with mobile wireless channels is caused by propagation delay. As the mobiles move, the arrival times of the transmitted signals change, which changes the crosscorrelations between received signals. However, this change in delay occurs very slowly relative to the chip duration for chip rates and mobile speeds of interest. (For example, assuming a chip rate of $10^7$ chips/s, and that the mobile is approaching the base station at a speed of 65 mph, the propagation delay changes by less than one chip/s.)

### 6.1.4 Power Control

Power control is a technique used in currently implemented DS-CDMA mobile telephony systems to alleviate the near-far problem. The basic idea of power control is to provide feedback to mobile transmitters to control their transmitted power levels to yield equal power at the receiver from all mobile transmitters. Since interference suppression techniques can potentially alleviate the near-far problem in DS-CDMA, their use can loosen the requirements on power control. However, power control can still be beneficial performance, with or without interference suppression, and can also reduce the power dissipated by mobile handsets, thereby extending battery life. For the matched-filter detector, the objective of power control for the reverse link is to ensure that all users detected at the base station are received with equal power. For mobile cellular systems, effective power control requires a feedback channel with which the receiver informs the transmitter to raise or lower the transmitted power in small increments (e.g., 1 dB). The effectiveness of the power control depends, of course, on how frequent the power updates are transmitted over the feedback channel, and the probability of power control errors (i.e., a "raise" command is received as a "lower" command, or vice versa).

For mobile cellular, it is generally assumed that a practical power control algorithm can respond quickly enough to compensate for shadowing and distance-related attenuation, but that it cannot compensate for fast Rayleigh fading due to multipath. Consequently, the received power of a signal that experiences flat Rayleigh fading will experience large short-term variations in power, but the power averaged over these short-term fades can be set at some target value. It is observed in [123] that for the type of closed-loop adaptive power control previously described, the distribution of the average received power due to variations caused by updates and power control errors is log-normal. The variance (in dB) of the received power reflects how "tight" the power control is. For the matched filter receiver very tight power control is required for adequate performance, which means that the standard deviation of the received signal power must be approximately 1 to 1.5 dB.

To optimize performance, the power control algorithm should ensure that the error rate

for each user is at the maximum acceptable value. For the matched filter receiver, this is the same as equalizing received powers; however, this is no longer true for receivers with interference suppression.

### 6.1.5 Time-Varying User Population

In addition to time-varying channels, in a mobile cellular environment the set of interferers is also time-varying. The interference suppression algorithm must be able to compensate for the appearance of new users, as well as for the disappearance of existing users. "Users" may be associated with calls, in the case of circuit-switched traffic, or with individual packets, in the case of packet-switched traffic. Note that in the latter case, packets may arrive and depart frequently, causing frequent transients in the interference environment that an adaptive algorithm must track. Even in the case of circuit-switched voice traffic, an adaptive algorithm must adapt to the set of users *currently speaking* in order to obtain the potential gains in capacity due to voice inactivity. (In practice, when a user is silent, the power of the transmitted signal is not set to zero, but is significantly reduced so that synchronization and channel tracking can be maintained.)

The rate at which users arrive and depart determines the average *traffic load*, measured in *Erlangs per cell* [124]. Specifically, assuming Poisson arrivals at rate $\lambda$ per cell (assumed to be the same for all cells), and an average service rate (per call or packet) given by $\mu$, the average number of users in the system is $C(\lambda/\mu)$, where $C$ is the number of cells and $\lambda/\mu$ is measured in Erlangs per cell. It is typically the case that in a DS-CDMA system the average number of users present in the system greatly exceeds the processing gain. This implies that the zero-forcing solutions previously discussed do not exist. However, if the number of *strong* interferers is significantly less than the processing gain, then the linear MMSE detector can effectively suppress these interferers, while treating weak users (e.g., in other cells) as background noise.

We saw from Figure 5.5 that the appearance of a new strong interferer can cause a transient performance degradation in adaptive algorithms. As the traffic load increases, these transients become more frequent. Since, in packet data systems, the appearance of a new user does not necessarily refer to a new call but rather a new data packet, rapid convergence in response to the appearance of new users is therefore a requirement for adaptive interference suppression in a packet data cellular system. Simulation results for a cellular type of model with stochastic arrivals and departures indicate that an adaptive interference suppression filter using the stochastic gradient algorithm is inadequate for this application, even under moderate traffic loads [29]. Thus, more rapid adaptation techniques are needed for this application.

### 6.1.6 Narrowband Interference

The fact that DS-CDMA systems spread transmitter power over a wide bandwidth allows the possibility that such systems can be overlaid on existing narrowband communication services, without undue degradation of either the narrowband or the spread-spectrum service. (The same property allows antijamming capability in military spread spectrum systems.) Although spread spectrum communications is inherently resistant to the narrowband inter-

ference (NBI) caused by such co-existence with conventional communications, it has been demonstrated that the performance of spread-spectrum systems in the presence of narrowband signals can be enhanced significantly through the use of active NBI suppression prior to despreading. In particular, not only does active suppression improve error-rate performance [9], but it also can lead to increased CDMA cellular system capacity [75] and improved acquisition capability [62].

Over the past two decades, a significant body of research has been concerned with the development of techniques for active NBI suppression in spread-spectrum systems. All of these techniques essentially seek to form a replica of the narrowband signals that can be subtracted from the received signal before data demodulation takes place. The formation of the replica may use predictors or interpolators to explicitly exploit the narrowband nature of the NBI against the wideband nature of the DS-CDMA signal (cf., [61], [98], [122], [131]), or it may use more detailed structural information in the case of digital NBI (which also arises in multirate CDMA systems) [83], [99]. In the latter case, a form of linear multiuser detection is essentially being used. Surveys of advances in this area are found in [61] and [80]. More recently, the adaptive MMSE detection techniques described in Sections 3 through 5 have been shown to work quite well against combined MAI and NBI of all types [87], [88].

### 6.1.7 Non-Gaussian Ambient Noise

Much of the development and analysis of interference suppression techniques for wireless systems has focused on situations in which the ambient noise is Gaussian. As noted in Section 2, this model has allowed the research in this area to focus on the main interference sources, namely structured interference (MAI and NBI). However, for many of the physical channels arising in wireless applications, the ambient noise is known through experimental measurements to be decidedly non-Gaussian. This is particularly true of urban and indoor radio channels [54, 55, 57] and underwater acoustic modem channels [11, 12, 56]. For these channels, the ambient noise is likely to have an impulsive component that gives rise to larger tail probabilities than is predicted by the Gaussian model. When the structured interference dominates, the lack of realism of the ambient noise model is perhaps not crucial. However, with multiuser detection, the MAI-limited nature of multiple-access channels is mitigated and the nature of the ambient noise is more important.

It is widely known in the single-user context that non-Gaussian noise can be quite detrimental to the performance of conventional systems based on the Gaussian assumption. On the other hand, the performance of signaling through non-Gaussian channels can be much better than that for corresponding Gaussian channels if the non-Gaussian nature of the channel is appropriately modeled and ameliorated. (The latter typically involves the use of nonlinear signal processing.) Neither of these properties is surprising. The first is a result of the lack of robustness of linear and quadratic type signal processing procedures to many types of non-Gaussian statistical behavior [43]. The second is a manifestation of the well-known least-favorability of Gaussian channels.

In view of the lack of realism of an AWGN model for ambient noise arising in many practical channels in which multiuser detection techniques may be applied, natural questions arise concerning the applicability, optimization, and performance of multiuser detection in non-Gaussian channels. Although performance indices such as MSE and SINR for linear

detectors are not affected by the distribution of the noise (only the spectrum matters), the more crucial bit-error rate can depend heavily on the shape of the noise distribution. The results of an early study of error rates in non-Gaussian DS-CDMA channels are found in [1], [2] and [3], in which the performance of conventional and modified conventional detectors is shown to depend significantly on the shape of the ambient noise distribution. In particular, impulsive noise can seriously degrade the error probability for a given level of ambient noise variance. In the context of multiple-access capability, this implies that fewer users can be supported with conventional detection in an impulsive channel than in a Gaussian channel. However, since non-Gaussian noise can, in fact, be beneficial to system performance if properly treated, the problem of joint mitigation of structured interference and non-Gaussian ambient noise is of interest [79]. An approach to this problem for NBI in spread-spectrum systems is described in [20]. Some very recent results along these lines for the case of MAI are reported in [81] and [130], the latter of which describes nonlinear adaptive methods that generalize the MMSE approach described in Sections 3 through 5.

## 6.2 System Issues

In addition to algorithmic issues such as performance and complexity, it is important to determine how adaptive interference suppression will affect other communication system requirements. These system issues are currently not well understood, so the following discussion is necessarily very brief (relative to importance).

### 6.2.1 Coding

Coding and interleaving are necessary to achieve reliable communications over fading channels. For example, in the commercial IS-95 standard DS-CDMA air interface, a rate 1/3 binary convolutional code is used on the uplink, and simultaneously serves to spread the bandwidth. Consequently, to achieve the same degree of spreading with coding, as compared to without coding, the length of the pseudo-random (PN) sequence associated with each bit must be reduced by a factor of three. When used with linear interference suppression, this reduction in PN-sequence length reduces the "degrees of freedom" available to suppress interference. (The tradeoff is that the coding may be able to compensate for the residual interference.) In other words, a low-rate code robs "dimensions" from the interference suppression filter. We add that this tradeoff does not apply to interference *cancellation* techniques, in which the interference is regenerated and substracted from the received signal [72].

The preceding observation implies that it is best to use high-rate codes with linear interference suppression, rather than low-rate codes. This has also been noted in [97], and in [69]. However, existing high-rate codes, such as Ungerboeck codes [110], rely on relatively dense constellations (such as 8-PSK), which may pose additional problems (e.g., with phase tracking) in a fading environment.

Another issue is that if *uncoded* symbols are used for decision-directed adaptation, then a more powerful coding scheme implies a higher decision error rate in the adaptive algorithm (assuming a fixed target error rate). There is, therefore, a tradeoff between the performance gain due to coding, and the performance loss due to uncoded decision errors in the adaptive

algorithm. (This assumes that the delay associated with using decoded symbols to direct the adaptation is unacceptable in a fast-fading environment.)

### 6.2.2 Power Control

For DS-CDMA systems that use the conventional matched-filter receiver, power control is crucial for mitigating the near-far problem. Also, as noted previously, power control has the advantage of lowering the transmitted power for each user, and thereby extends battery life. For the conventional matched filter receiver, power control is typically used to equalize the received powers [123]. (A further advantage of this is that it helps avoid saturation of nonlinear receiver elements, such as fixed-point processors.) However, in general, the goal of optimum power control is to adjust the transmitted powers so that the received SINRs corresponding to all users being detected are equal [6]. Although multiuser detection and interference suppression techniques can alleviate power control requirements, power control can enhance the performance of interference suppression techniques. The interaction of adaptive power control with adaptive interference suppression has so far received little attention.

### 6.2.3 Timing Recovery

The discussion in preceding sections has assumed that the receiver is perfectly synchronized to the desired user. It has been observed that timing offsets can significantly degrade the performance of some multiuser detectors [13], [137]. This observation, however, is implementation dependent. An advantage of the adaptive tapped-delay line implementation for linear interference suppression is that it is analogous to an adaptive fractionally-spaced equalizer (for a single-user channel), which is known to be insensitive to timing offset [91]. Furthermore, timing recovery for the multiple-access channel with an adaptive fractionally-spaced tapped-delay line (TDL) can be accomplished in an analogous fashion as for a single-user channel. The combination of timing estimation with adaptive linear interference suppression is studied in [52], [102] and [129].

### 6.2.4 Nonuniform Quality of Service

An important difference between linear interference suppression techniques and the matched-filter receiver is that the former relies on the use of short spreading sequences, whereas the latter can use either or short or long spreading sequences. A very long spreading sequence, such as used in IS-95, is equivalent to selecting a different random spreading sequence for each bit ("code hopping"). An advantage of code hopping is that each user sees approximately the same performance, assuming perfect power control, since averaging the performance over the sequence of transmitted bits is equivalent to averaging over spreading sequences.

In contrast, the solution for MMSE, given by (4.20), implies that the MMSE depends on the particular assignment of spreading codes to users (as well as relative amplitudes and phases). Furthermore, the MMSE will generally be different for different users. Consequently, even though *average* performance (i.e., error probability averaged over all spreading sequences) may be very good, some users may have relatively poor performance. This is be-

cause the spreading codes assigned to two users may have relatively high cross-correlation. This observation is made in [117], and is analyzed in [36].

To illustrate the preceding observation, Figure 6.2 shows the distribution (computed via simulation) for SINR assuming that the user signature sequences are selected randomly (i.e., each element of the vectors $\mathbf{p}_k$, $k = 1, \ldots, K$, is determined by a fair coin toss). In this example the processing gain $N = 30$, there are 10 "strong" (i.e., intra-cell) users, 50 "weak" (i.e., other-cell) users, and the signal-to-background noise ratio is 25 dB. The received powers (for members of each set of strong and weak users) were selected from a log-normal distribution. The SIR distribution for the matched-filter receiver is also shown.

The results in Figure 6.2 indicate that there is a significant spread in performance ( $> 10$ dB) over the user population. Note that by using a very long spreading sequence (as in the IS-95 standard) the distribution for the matched-filter becomes a point mass at the average of the distribution shown in Figure 6.2 (-10 dB).

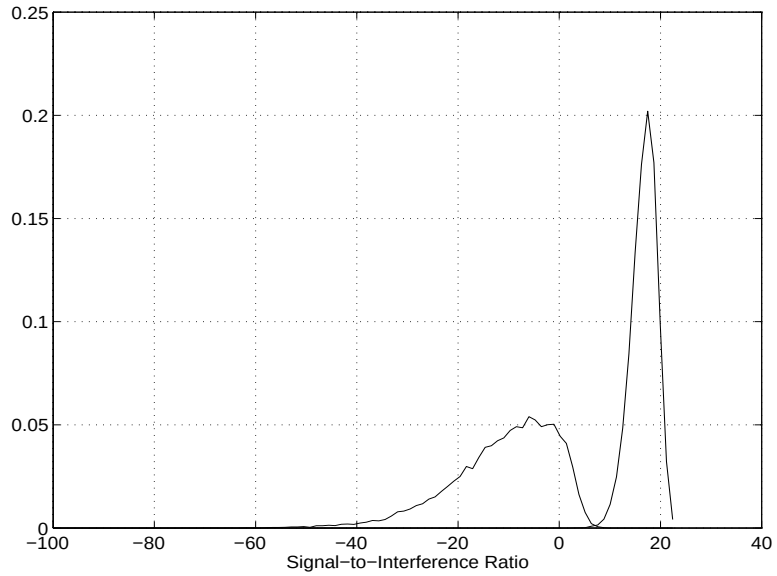


Figure 6.2: Distributions for Signal-to-Interference Ratio assuming random signature sequences. Results for both the linear MMSE detector and the matched filter are shown.

Power control may help to improve the performance of users experiencing poor performance only in some situations. For example, two adjacent users may be assigned "nearby" codes, meaning that their cross-correlation is large. If the adjacent users are transmitting to the same receiver, then power control cannot significantly improve performance for both users. However, if the users are transmitting to different receivers (i.e., if they are in different cells), then it may be possible for one of the users to reduce power, thereby reducing interference to the other user.

Finally, we remark that the significant spread in performance shown in Figure 6.2 assumes a static situation in which the set of users, and relative amplitudes and phases are fixed. In a mobile wireless network these parameters are time-varying, which should alleviate this problem to some degree. In addition, path and/or space diversity may also reduce the likelihood of having relatively high cross-correlations with neighboring users.

### 6.2.5 Very Long Spreading Sequences

The model that we have proposed in this chapter has addressed primarily the situation in which the received signaling waveform of each user is the same in each symbol interval (aside from fading and other channel effects). In some current DS-CDMA systems (such as the IS-95 digital cellular standard), this model is not accurate because the period of the spreading waveform spans many bits. From a theoretical point of view, this distinction is not overly significant. However, from a practical point of view, it is quite significant. Since the key parameter determining performance of DS-CDMA systems is the number of chips per symbol, not the number of chips per period of the spreading code, this use of long spreading codes is primarily of value in providing uniform quality of service over the user population, as discussed in the preceding section, and in avoiding the need to assign (or reassign) codes to each new call (or to an existing call which is handed off to an adjacent cell). In deciding whether or not to use short or long codes in future DS-CDMA standards, these advantages should be weighed against the performance advantages offered by practical multiuser detection.

### 6.2.6 Power Consumption

Since mobility is one of the main motivations for using wireless communications, the practicality of many of the techniques described in this chapter depends heavily on the ability to implement them in portable, battery-operated handsets. Thus, the issue of energy consumption is of considerable importance in the development of interference suppression algorithms for wireless systems. In cellular systems, there is an asymmetry with respect to this issue, in that the base station (i.e., the uplink transceiver) is relatively unconstrained by energy consumption, whereas the mobiles (i.e., the downlink transceivers) are severely constrained. So, the use of sophisticated signal processing, such as multiuser detection, at the base station does not pose a serious energy-consumption problem. Since these techniques allow better performance for a given level of received signal energy than do conventional methods, the use of such methods in the base station can reduce required transmitter power at the mobiles, thereby reducing overall battery requirements for portable transceivers. However, the advantages of adaptive linear multiuser detection, multipath mitigation, etc., can significantly enhance the downlink performance as well. (Also, point-to-point systems do not necessarily feature fixed, non-portable transceivers.) Thus, energy-efficient techniques for the linear adaptive algorithms discussed in this paper are of considerable interest.

One technique for achieving energy efficiency in the MMSE detector is described in [88]. In particular, energy consumption in integrated circuits is an increasing function of gate speed. Since the algorithms of interest here should be implemented at the signaling rate (i.e., at the symbol rate), the algorithms cannot be slowed down to reduce energy consumption. However, as shown in [88], the blind RLS MMSE detection algorithm can be implemented with a slower gate speed by mapping it to a systolic array. This mapping allows the individual gate speed to be reduced significantly without a corresponding reduction in the speed at which the algorithm updates its coefficients.

# 7 Summary and Conclusions

In this chapter, we have discussed the use of adaptive signal processing techniques to suppress structured interference in wireless systems. We have focussed on the suppression of multiple-access interference, and have primarily considered techniques that are based on the MMSE method of linear multiuser detection. As we have seen, MMSE detection provides many of the performance advantages of optimal multiuser detection, without its attendant complexity. Moreover, the MMSE detector lends itself to a great variety of adaptive methods, and it is relatively robust to other types of interference (such as narrowband interference).

The results presented in this chapter have largely been of a research nature, as this is the primarily intention of the present volume. However, we have also mentioned, in Section 6, a variety of other issues that are of concern in bringing these methods to practice. These practical issues present a wealth of other research questions, many of which are currently being addressed by the community.

# References


[1] B. Aazhang and H. V. Poor,"Performance of DS/SSMA Communications in Impulsive Channels - Part I: Linear Correlation Receivers," *IEEE Trans. Commun.*, Vol. 35, No. 11, pp. 1179-1188, November 1987.

[2] B. Aazhang and H. V. Poor, "Performance of DS/SSMA Communications in Impulsive Channels - Part II: Hard-limiting Correlation Receivers," *IEEE Trans. Commun.*, Vol. 36, No. 1, pp. 88-97, January 1988.

[3] B. Aazhang and H. V. Poor, "An Analysis of Nonlinear Direct-sequence Correlators," *IEEE Trans. Commun.*, Vol. 37, No. 7, pp. 723-731, July 1989.

[4] M. Abdulrahman and D. D. Falconer, "Cyclostationary Crosstalk Suppression by Decision Feedback Equalization on Digital Subscriber Loops," *IEEE J. Select. Areas Commun.*, Vol. 10, No. 3, pp. 640-649, April 1992.

[5] M. Abdulrahman, A. U. H. Sheikh, and D. D. Falconer, "Decision Feedback Equalization for CDMA in Indoor Wireless Communications," *IEEE J. Select. Areas Commun.*, Vol. 12, No. 4, pp. 698-704, May 1994.

[6] S. Ariyavisitakul, "Signal and Interference Statistics of a CDMA System with Feedback Power Control Part II," *IEEE Trans. Commun.* Vol. 42, No. 2/3/4, Feb/March/April 1994.

[7] J. Barry, J. Kahn, E. Lee and D. Messerschmidt, "High-speed Nondirective Optical Communication for Wireless Networks," *IEEE Network Mag.*, Vol. 29, pp. 44 - 54, Nov. 1991.

[8] S. E. Bensley and B. Aazhang, "Subspace-Based Estimation of Multipath Channel Parameters for CDMA Communication Systems," *Proc. 1994 IEEE GLOBECOM/Comm. Theory Mini-Conf.*, pp. 154-157, San Francisco, CA, Dec. 1994.

[9] N. Bershad, "Error Probabilities of DS Spread-spectrum Systems Using and ALE for Narrow-band Interference Rejection," *IEEE Trans. Commun.*, Vol. COM-36, No. 5, p. 587 - 595, 1988.

[10] D. Brady and J. Catipovic, "Adaptive Multiuser Detection for Underwater Acoustical Channels," *IEEE J. Oceanic Engineering,* Vol. 19, No. 2, pp. 158 - 165, April 1994.

[11] P. L. Brockett, M. Hinich, and G. R. Wilson, "Nonlinear and Non-Gaussian Ocean Noise," *J. Acoust. Soc. Am.*, vol. 82, pp. 1286-1399, 1987.

[12] P. L. Brockett, M. Hinich, and G. R. Wilson, "Bispectral Characterization of Ocean Acoustic Time Series: Nonlinearity and Non-Gaussianity," in E. J. Wegman, S. C. Schwartz, and J. B. Thomas (Eds.), *Topics in Non-Gaussian Signal Processing*, Berlin: Springer-Verlag, 1989.

[13] R. M. Buehrer, N. S. Correal, B. D. Woerner, "A Comparison of Multiuser Receivers for Cellular CDMA," *Proc. IEEE GLOBECOM'96*, London, UK, November 1996.

[14] J. Catipovic, D. Brady, and A. S. Etchemendy, "Development of Underwater Acoustic Modems and Networks," *Oceanography Magazine,* Vol. 6, pp. 112 - 119, 1993.

[15] K.-C. Chen, L. J. Cimini, Jr., B. D. Woerner and S. Yoshida, Eds., *Wireless Local Communications.* Special Issue of *IEEE J. Select. Areas Commun.,* Vol. 14, No. 3, April 1996.

[16] A. Duel-Hallen, "A Family of Multiuser Decision-Feedback Detectors for Asynchronous Code-Division Multiple-Access Channels," *IEEE Trans. Commun.*, Vol. 43, No. 2/3/4, pp. 421-434, Feb./March/April 1995.

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

[18] D. D. Falconer, M. Abdulrahman, N. W. K. Lo, B. R. Petersen, and A. U. H. Sheikh, "Advances in Equalization and Diversity for Portable Wireless Systems," *Digital Signal Processing 3*, pp. 148-162, 1993.

[19] V. K. Garg and J. E. Wilkes, *Wireless and Personal Communications Systems.* (Prentice-Hall: Upper Saddle River, NJ, 1996)

[20] L. M. Garth and H. V. Poor, "Narrowband Interference Suppression in Impulsive Channels," *IEEE Trans. Aerosp. Electron. Syst.,* Vol. 28, No. 1, pp. 15-34, January 1992.

[21] R. D. Gitlin, H. C. Meadows, and S. B. Weinstein, "The Tap-Leakage Algorithm: An Algorithm for the Stable Operation of a Digitally Implemented, Fractionally-Spaced, Adaptive Equalizer," *Bell Syst. Tech. J.*, Vol. 61, pp. 1817-1839, Oct. 1982.

[22] A. M. Haimovich and Y. Bar-Ness, "An Eigenanalysis Interference Canceler," *IEEE Trans. Signal Process.*, Vol. 39, No. 1, pp. 76-84, Jan. 1991.

[23] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, 2nd edition, 1994.

[24] D. J. Harrison, "Adaptive Equalization for Channels with Crosstalk," M. Eng. thesis, Carleton University, Ottawa, Ontario, Canada, December 1969.

[25] Z. J. Hass, R. Alonso, D. Duchamp and B. Gopinath, *Mobile and Wireless Computing Newtorks.* Special Issue of *IEEE J. Select. Areas Commun.,* Vol. 13, June 1995.

[26] T. Helleseth and P. V. Kumar, "Pseudonoise Sequences." In *Mobile Communications Handbook*, J. Gibson, Ed. (CRC Press: Boca Raton, FL, 1996).

[27] J. Holtzman, "A Simple Accurate Method to Calculate Spread-spectrum Multiple Access Error Probabilities," *IEEE Trans. Commun.*, Vol. 40, pp. 461 - 464, March 1992.

[28] M. L. Honig, "Orthogonally Anchored Blind Interference Suppression Using the Sato Cost Criterion," *Proc. 1995 IEEE Int. Symp. on Inf. Theory*, p. 314, Whistler, British Columbia, Ca., Sept. 1995.

[29] M. L. Honig, "Performance of Adaptive Interference Suppression for DS-CDMA With a Time-Varying User Population," *Proc. IEEE Fourth Int. Symp. on Spread Spectrum Techniques & Applications*, Mainz, Germany, pp. 267-271, Sept. 1996.

[30] M. L. Honig, P. Crespo, and K. Steiglitz, "Suppression of Near- and Far-end Crosstalk by Linear Pre- and Post-filtering," *IEEE J. Select. Areas Commun.*, Vol. 10, No. 3, pp. 614-629, April 1992.

[31] M. L. Honig, U. Madhow, and S. Verdú, "Adaptive Blind Multi-User Detection," *IEEE Trans. Inform. Theory*, Vol. 41, No. 4, pp. 944-960, July 1995.

[32] M. L. Honig and D. G. Messerschmitt, *Adaptive Filters: Structures, Algorithms, and Applications*, Kluwer Academic Publishers, Boston, MA, 1985.

[33] M. L. Honig, S. L. Miller, L. B. Milstein, and M. J. Shensa, "Performance of Adaptive Linear Interference Suppression for DS-CDMA in the Presence of Flat Rayleigh Fading," *Proceedings of the 46th IEEE Vehicular Technology Conference (VTC'96),* Atlanta, GA, May 1996.

[34] M. L. Honig, "Rapid Detection and Suppression of Interference in DS-CDMA," *IEEE Int. Conference on Acoustics, Speech, and Signal Processing*, Detroit, MI, May 1995.

[35] M. L. Honig, "A Comparison of Subspace Adaptive Filtering Techniques for DS-CDMA Interference Suppression," *Proc. MILCOM'97*, Monterey, CA, November 1997.

[36] M. Honig and W. Veerakachen, "Performance Variability of Linear Multiuser Detection for DS-CDMA," *Proceedings of the 46th IEEE Vehicular Technology Conference (VTC'96),* Vol. 1, pp. 372-376, Atlanta, Georgia, May 1996.

[37] D. Horwood and R. Gagliardi, "Signal Design for Digital Multiple Access Communications," *IEEE Trans. Commun.*, Vol. COM-23, pp. 378-383, March 1975.

[38] H. C. Huang, *Combined Multipath Processing, Array Processing, and Multiuser Detection for DS-CDMA Channels*, Ph.D. Thesis, Electrical Engineering Dept., Princeton University, Princeton, NJ, 1995.

[39] R. A. Iltis, "An Adaptive Multiuser Detector with Joint Amplitude and Delay Estimation," *IEEE J. Select. Areas Commun.,* Vol. 12, No. 5, pp. 774 - 785, June 1994.

[40] W. C. Jakes (Ed.), *Microwave Mobile Communications.* IEEE Press, 1974.

[41] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing – Concepts and Techniques*, Prentice Hall, Englewood Cliffs, NJ, 1993.

[42] A. R. Kaye and D. A. George, "Transmission of Multiplexed PAM Signals over Multiple Channel and Diversity Systems," *IEEE Trans. Commun. Technol.*, Vol. COM-18, No. 5, pp. 520-526, October 1970.

[43] S. A. Kassam and H. V. Poor, "Robust Techniques for Signal Processing: A Survey," *Proceedings of the IEEE,* Vol. 73, No. 3, pp. 433-481, March 1985.

[44] A. Klein, G. K. Kaleh and P. W. Baier, "Zero Forcing and Minimum-mean-square-error Equalization for Multiuser Detection in Code-division Multiple-access Channels," *IEEE Trans. Vehic. Technol.,* Vol. 45, No. 2, pp. 276 - 287, May 1996.

[45] R. Kohno, "Psuedo-noise Sequences and Interference Cancellation Techniques for Spread-spectrum Systems - Spread Spectrum Theory and Techniques in Japan," *IEICE Trans.*, Vol. E.74, pp. 1083 - 1092, May 1991.

[46] E. A. Lee and D. G. Messerschmitt, *Digital Communication*, Kluwer Academic Publishers, Boston, 1994.

[47] N. W. K. Lo, D. D. Falconer, and A. U. H. Sheikh, "Adaptive Equalizer MSE Performance in the Presence of Multipath Fading, Interference, and Noise," *Proceedings of the 45th IEEE Vehicular Technology Conference (VTC'95),* pp. 409-413, Chicago, IL, July 1995.

[48] N. W. K. Lo, D. D. Falconer, and A. U. H. Sheikh, "Adaptive Equalization for Co-Channel Interference in a Multipath Fading Environment," *IEEE Trans. Commun.*, Vol. 43, No. 2/3/4, pp. 1441-1453, Feb/March/April 1995.

[49] R. Lupas and S. Verdú, "Linear Multi-user Detectors for Synchronous Code-division Multiple-access Channels," *IEEE Trans. Inform. Theory*, Vol. IT-35, No. 1, pp. 123-136, January 1989.

[50] R. Lupas and S. Verdú, "Near-far Resistance of Multi-user Detectors in Asynchronous Channels," *IEEE Trans. Commun.*, Vol. COM-38, No. 4, pp. 496-508, April 1990.

[51] U. Madhow and M. Honig, "MMSE Interference Suppression for Direct-sequence Spread-spectrum CDMA," *IEEE Trans. Commun.*, Vol. 42, pp. 3178 - 3188, Dec. 1994.

[52] U. Madhow, "MMSE Interference Suppression for Acquisition and Demodulation of Direct-Sequence CDMA Signals," *IEEE Trans. Commun.,* to appear.

[53] N. Mandayam and S. Verdú, "Analysis of an Approximate Decorrelating Detector," *Wireless Personal Communications - Special Issue on Interference in Mobile Wireless Systems*, 1997.

[54] D. Middleton, "Man-made Noise in Urban Environments and Transportation Systems: Models and Measurements," *IEEE Trans. Commun.*, Vol. COM-21, pp. 1232 - 1241, 1973.

[55] D. Middleton, "Statistical-physical Models of Electromagnetic Interference," *IEEE Trans. Electromag. Compat.*, Vol. EMC-19, 106-127, Aug. 1977.

[56] D. Middleton, "Channel Modeling and Threshold Signal Processing in Underwater Acoustics: An Analytical Overview," *IEEE J. Oceanic Eng.* Vol. OE-12, no. 1, pp. 4-28, January 1987.

[57] D. Middleton and A. D. Spaulding, "Elements of Weak-signal Detection in Non-Gaussian Noise," in *Advances in Statistical Signal Processing - Vol. 2: Signal Detection*, H. V. Poor and J. B. Thomas, Eds. (JAI Press: Greenwich, CT, 1993).

[58] S. L. Miller, "An Adaptive Direct-Sequence Code-Division Multiple-Access Receiver for Multiuser Interference Rejection," *IEEE Trans. Commun.*, Vol. 43, No. 2/3/4, pp. 1746-1755, Feb./March/April 1995.

[59] S. L. Miller, "Training Analysis of Adaptive Interference Suppression for Direct-Sequence Code-Division Multiple-Access Systems," *IEEE Trans. Commun.*, Vol. 44, No. 4, pp. 488-495.

[60] S. L. Miller and A. N. Barbosa, "A Modified MMSE Receiver for Detection of DS-CDMA Signals in Fading Channels," *Proc. MILCOM'96*, Washington, D.C.

[61] L. B. Milstein, "Interference Rejection Techniques in Spread Spectrum Communications," *Proceedings of the IEEE,* Vol. 76, no. 6, pp. 657 - 671, 1988.

[62] L. B. Milstein, "Interference Suppression to Aid Acquisition in Direct-sequence Spread-spectrum Communications," *IEEE Trans. Commun.*, Vol. 36, no. 11, pp. 1200 - 1202, 1988.

[63] U. Mitra and H. V. Poor, "Activity Detection in a Multi-user Environment," *Wireless Personal Communications,* Vol. 3, pp. 149 - 174, 1996.

[64] U. Mitra and H. V. Poor, "Adaptive Decorrelating Detectors for CDMA Systems," *Wireless Personal Commun.*, Vol. 2, No. 4, pp. 415-550, 1996.

[65] U. Mitra and H. V. Poor, "Neural Network Techniques for Adaptive Multi-user Demodulation," *IEEE J. Select. Areas Commun.*, Vol. 12, no. 9, pp. 1460 - 1470, December 1994.

[66] U. Mitra and H. V. Poor, "Adaptive Receiver Algorithms for Near-far Resistant CDMA," *IEEE Trans. Commun.*, Vol. 43, no. 2/3/4, Part III pp. 1713 - 1724, February/March/April 1995.

[67] U. Mitra and H. V. Poor, "An Adaptive Decorrelating Detector for Synchronous CDMA Channels," *IEEE Trans. Commun.*, Vol. 44, No.2, pp. 257 - 268, February 1996.

[68] L. B. Nelson and H. V. Poor, "Iterative Multiuser Receivers for CDMA Channels: An EM-based Approach," *IEEE Trans. Commun.*, Vol. 44, No. 12, December 1996.

[69] I. Oppermann, P. Rapajic, and B. S. Vucetic, "Capacity of a Band-Limited CDMA MMSE Receiver Based System When Combined With Trellis or Convolutional Coding," preprint, August 1996.

[70] K. Pahlavan and A. H. Levesque, *Wireless Information Networks,* (Wiley-Interscience: New York, 1995)

[71] K. Pahlavan, T. Probert and M. Chase, "Trends in Local Wireless Networks," *IEEE Commun. Mag.*, Vol. 33, pp. 88 - 95, March 1995.

[72] P. Patel and J. Holtzman, "Analysis of a Simple Successive Interference Cancellation Scheme in a DS/CDMA System," *IEEE J. Select. Areas Commun.*, Vol. 12, No. 5, pp. 796-807, June 1994.

[73] B. R. Petersen and D. D. Falconer, "Minimum Mean Square Equalization in Cyclostationary and Stationary Inteference – Analysis and Subscriber Line Calculations," *IEEE J. Select. Areas Commun.*, Vol. 9, No. 6, pp. 931-940, August 1991.

[74] D. Parsavand and M. H. Varanasi, "RMS Bandwidth Constrained Signature Waveforms that Maximize the Total Capacity of PAM-synchronous CDMA Channels," *IEEE Trans. Commun.*, Vol. 44, No. 1, January 1996.

[75] R. L. Pickholz, L. B. Milstein, and D. L .Schilling, "Spread Spectrum for Mobile Communications," *IEEE Trans. Vehic. Technol.*, Vol. 40, no. 2, pp. 313 - 322, 1991.

[76] H. V. Poor, *An Introduction to Signal Detection and Estimation - Second Edition*, New York: Springer-Verlag, 1994.

[77] H. V. Poor, "Adaptivity in Multiple-access Communications," *Proc. 34th IEEE Conference on Decision and Control*, New Orleans, LA, December 13-15, 1995.

[78] H. V. Poor, "Adaptive Interference Suppression in CDMA Systems," *Proceedings of the Symposium on Interference Rejection and Signal Separation in Wireless Communications,* New Jersey Institute of Technology, Newark, NJ, March 19, 1996.

[79] H. V. Poor, "Non-Gaussian Signal Processing Problems in Multiple-access Communications," in *Proceedings of the 1996 USC/CRASP Workshop on Non-Gaussian Signal Processing*, Ft. George Meade, MD, May 24, 1996.

[80] H. V. Poor and L. A. Rusch, "Narrowband Interference Suppression in Spread-spectrum Systems," *IEEE Personal Communications*, Vol. 1, No. 3, pp. 14 - 27, August 1994.

[81] H. V. Poor and M. Tanda, "An Analysis of Some Multiuser Detectors in Impulsive Noise," *Proceedings of the 16th GRETSI Symposium on Signal and Image Processing,* Grenoble, France, September 15 - 17, 1997.

[82] H. V. Poor and S. Verdú, "Probability of Error in MMSE Multiuser Detection," *IEEE Trans. Inform. Theory*, Vol. 43, No. 3, pp. 858 - 871, May 1997.

[83] H.V. Poor and X. Wang, "Adaptive Suppression of Narrowband Digital Interferers from Spread Spectrum Signals", *Proceedings of the 1996 International Conference on Acoustics, Speech and Signal Processing (ICASSP'96)*, Vol.II, pp.1061-1064, Atlanta, GA, May 7-10, 1996.

[84] H. V. Poor and X. Wang,"Signal Processing for Adaptive Interference Suppression in CDMA Systems," *Proceedings of the 5th European Workshop on DSP Applied to Space Communications,* Barcelona, Spain, September 25 - 27, 1996, pp. 11.1 - 11.10.

[85] H. V. Poor and X. Wang,"Adaptive Multiuser Detection in Fading Channels," *Proceedings of the 34th Annual Allerton Conference on Communications, Control, and Computing,* University of Illinois, Urbana, IL, October 2 - 4, 1996, pp.603-612.

[86] H. V. Poor and X. Wang, "Blind Adaptive Suppression of Multiuser Narrowband Digital Interferers from Spread-spectrum Signals," *Wireless Personal Communications - Special Issue on Interference in Mobile Wireless Systems*, 1997.

[87] H. V. Poor and X. Wang, "Code-aided Interference Suppression in DS/CDMA Spread Spectrum Communications - Part I: Interference Suppression Capability," *IEEE Trans. Commun.,* Vol. 45, No. 9, pp. 1101 - 1111, September 1997.

[88] H. V. Poor and X. Wang, "Code-aided Interference Suppression in DS/CDMA Spread Spectrum Communications - Part II: Parallel Blind Adaptive Implementation," *IEEE Trans. Commun.*, Vol. 45, No. 9, pp. 1112 - 1122, September 1997.

[89] H. V. Poor and X. Wang, "Blind Adaptive Joint Suppression of MAI and ISI in CDMA Channels," *Proceedings of the Thirty-second Asilomar Conference on Signals, Systems and Computers,* Pacific Grove, CA, November 2 - 5, 1997.

[90] J. G. Proakis, *Digital Communications - Third Edition*, McGraw-Hill, Boston, 1995.

[91] S. Qureshi, "Adaptive Equalization," *Proceedings of the IEEE*, Vol. 73, No. 9, pp. 1349-1387, September 1985.

[92] P. B. Rapajic and B. S. Vucetic, "Adaptive Receiver Structures for Asynchronous CDMA Systems," *IEEE J. Select. Areas Commun.*, Vol. 12, No. 4, pp. 685-697, May 1994.

[93] P. Rapajiv and B. S. Vucetic, "Linear Adaptive Transmitter-receiver Structures for Asynchronous CDMA Systems," *European Trans. on Commun.,* pp. 21 - 27, Jan./Feb. 1995.

[94] T. S. Rappaport, *Wireless Communications: Principles and Practice,* (Prentice-Hall: Upper Saddle River, NJ, 1996)

[95] M. Rupf, F. Tarkoy and J. L. Massey, "User-separating Demodulation for Code-division Multiple-access Systems," *IEEE J. Select. Areas Commun.,* Vol. 12, No. 5, pp. 786 - 795, June 1994.

[96] M. Rupf and J. L. Massey, "Optimum Sequence Multisets for Synchronous Code-division Multiple-access Channels," *IEEE Trans. Inform. Theory,* Vo. IT-40, pp. 1261 - 1266, July 1994.

[97] M. Rupf, F. Tarköy, and J. L. Massey, "User-Separating Demodulation for Code-Division Multiple-Access Systems," *IEEE J. Select. Areas Commun.*, Vol. 12, No. 5, pp. 786-795, June 1994.

[98] L. A. Rusch and H. V. Poor, "Narrowband Interference Suppression in CDMA Spread-spectrum Communications," *IEEE Trans. Commun.,* Vol. 42, No. 4, pp. 1969-1979, April 1994.

[99] L. A. Rusch and H. V. Poor, "Multiuser Detection Techniques for Narrowband Interference Suppression in Spread-spectrum Communications," *IEEE Trans. Commun.,* Vol. 43, No. 2/3/4, Part III, pp. 1725 - 1737, February/March/April 1995.

[100] J. Salz, "Digital Transmission over Cross-Coupled Linear Channels," *Bell Syst. Tech. J.*, Vol. 64, No. 6, pp. 1147-59, July-Aug. 1985.

[101] R. Singh and L. B. Milstein, "Adaptive Interference Suppression in Direct-Sequence CDMA," submitted to *IEEE Trans. Commun.*

[102] R. F. Smith and S. L. Miller, "Code Timing Estimation in a Near-Far Environment for Direct-Sequence Code-Division Multiple-Access," *Proc. MILCOM'94.*

[103] J. M. Straus, *et al., Universal Communications: Proceedings of MILCOM'95.* San Diego, CA, November 5 - 8, 1995.

[104] E. G. Strom and S. L. Miller, "A Reduced Complexity Adaptive Near-Far Resistant Receiver for DS-CDMA," *Proc. IEEE GLOBECOM'93*, pp. 1734-1738.

[105] E. G. Strom, S. Parkvall, S. L. Miller, and B. E. Ottersten, "Propagation Delay Estimation of DS-CDMA Signals in a Fading Environment," *Proc. 1994 IEEE GLOBECOM/Comm. Theory Mini-Conf.*, pp. 85-89, San Francisco, CA, Dec. 1994.

[106] Y. Steinberg and H. V. Poor, "Multiuser Delay Estimation," *Proc. 1993 Conf. Inform. Sci. Syst.*, Baltimore, MD, March 17-19, 1993.

[107] Y. Steinberg and H. V. Poor, "Sequential Amplitude Estimation in Multiuser Communications," *IEEE Trans. Inform. Theory*, Vol. 38, pp. 11 - 20, 1994.

[108] J. Travis, "Dialing up Undersea Data - Long Distance," *Science,* Vol. 263, pp. 1223 - 1224, March 4, 1994.

[109] D. W. Tufts and A. A. Shah, "Rapid Interference Suppression and Channel Identification for Digital, Multipath Wireless Channels," *Proceedings of the 44th IEEE Vehicular Technology Conference (VTC'94),* Stockholm, Sweden, 1994.

[110] G. Ungerboeck, "Channel Coding with Multilevel/Phase Signalsm," *IEEE Trans. Inform. Theory*, Vol. IT-28, No. 1, pp. 55-67, Jan. 1982.

[111] M. K. Varanasi, "Noncoherent Detection in Asynchronous Multiuser Channels, *IEEE Trans. Inform. Theory,* Vol. 39, no. 1, pp. 157 - 176, January 1993.

[112] M. K. Varanasi, "Group Detection for Synchronous Gaussian Code-division Multiple-access Channels, *IEEE Trans. Inform. Theory,* Vol. 41, no. 4, July 1995.

[113] M. K. Varanasi, "Parallel Group Detection for Synchronous CDMA Communication over Frequency Selective Raleigh Fading Channels," *IEEE Trans. Inform. Theory,* Vol. 42, no. 1, pp. 116 - 123, January 1996.

[114] M. K. Varanasi and B. Aazhang, "Multistage Detection for Asynchronous Code-Division Multiple-Access Communications," *IEEE Trans. Commun.*, Vol. COM-38:4, pp. 509-519, April 1990.

[115] M. K. Varanasi and B. Aazhang, "Near-optimum Detection in Synchronous Code-division Multiple-access Systems," *IEEE Trans. Commun.*, Vol. 39, No. 5, pp. 725 - 736, May 1991.

[116] S. Vasudevan and M. K. Varanasi, "A Near-optimum Receiver for CDMA Communication over the Time-varying Rayleigh Fading Channel, *IEEE Trans. Commun.*, Vol. 44, 1996.

[117] S. Vembu and A. J. Viterbi, "Two Different Philosophies in CDMA – A Comparison," *Proceedings of the 46th IEEE Vehicular Technology Conference (VTC'96),* Vol. 2, pp. 869-873, Atlanta, Georgia, May 1996.

[118] S. Verdú, "Minimum Probability of Error for Asynchronous Gaussian Multiple-access Channels," *IEEE Trans. Inform. Theory,* Vol. IT-32, pp. 85 - 96, Jan. 1986.

[119] S. Verdú, "Multiuser Detection," in *Advances in Statistical Signal Processing - Vol. 2: Signal Detection*, H. V. Poor and J. B. Thomas, Eds. (JAI Press: Greenwich, CT, 1993).

[120] S. Verdú, "Optimum Multi-user Asymptotic Efficiency," *IEEE Trans. Commun.*, Vol. COM-38, No. 4, pp. 496-508, April 1990.

[121] S. Verdú, *Multiuser Detection.* (Cambridge University Press: Cambridge, UK, 1998, in press)

[122] R. Vijayan and H. V. Poor, "Nonlinear Techniques for Interference Suppression in Spread-Spectrum Systems," *IEEE Trans. Commun.,* Vol. 38, No. 7, pp. 1060-1065, July 1990.

[123] A. J. Viterbi, A. M. Viterbi, and E. Zehavi, "Performance of Power-Controlled Wideband Terrestrial Digital Communication," *IEEE Trans. Commun.*, Vol. 41, No. 4, April 1993.

[124] A. M. Viterbi and A. J. Viterbi, "Erlang Capacity of a Power Controlled CDMA System," *IEEE IEEE J. Select. Areas Commun.*, Vol. 11, No. 6, pp. 892-900, Aug. 1993.

[125] A. J. Viterbi, *CDMA: Principles of Spread Spectrum Communications.* (Addison-Wesley: Reading, MA, 1995)

[126] X. Wang and H. V. Poor, "Adaptive Multiuser Diversity Receivers for Frequency-Selective Rayleigh Fading CDMA Channels," *Proceedings of the 47th IEEE Vehicular Technology Conference (VTC'97),* Phoenix, AZ, May 5 - 7, 1997, Vol. I, pp. 198 - 202.

[127] X. Wang and H. V. Poor, "Subspace-based Blind Adaptive Joint Interference Suppression and Channel Estimation in Multipath CDMA Channels," *Proceedings of the 1997 International Conference on Universal Personal Communications (ICUCP'97),* San Diego, CA, October 13 - 17, 1997. [Also to appear in *IEEE Trans. Signal Process.*]

[128] X. Wang and H. V. Poor, "Blind Equalization and Multiuser Detection for CDMA Communications in Dispersive Channels," *IEEE Trans. Commun.,* Vol. 46, No. 1, January 1998.

[129] X. Wang and H. V. Poor, "Blind Multiuser Detection: A Subpace Approach," *IEEE Trans. Inform. Theory,* Vol. 44, No. 2, March 1998.

[130] X. Wang and H. V. Poor, "Adaptive Multiuser Detection in Non-Gaussian Channels," *Proceedings of the 35th Annual Allerton Conference on Communications, Control, and Computing,* University of Illinois, Urbana, IL, October 1 - 3, 1997.

[131] P. Wei, J. R. Zeidler, and W. H. Ku, "Adaptive Interference Suppression for CDMA Overlay Systems," *IEEE J. Selected Areas in Communications*, Vol. 12, No. 9, pp. 1510 - 1523, Dec. 1994.

[132] Z. Xie, C. K. Rushforth, R. T. Short and T. Moon, "Joint Signal Detection and Parameter Estimation in Multi-user Communications," *IEEE Trans. Commun.*, Vol. 41, pp. 1208 - 1216, August 1993.

[133] Z. Xie, R. T. Short, and C. K. Rushforth, "A Family of Suboptimum Detectors for Coherent Multi-user Communications," *IEEE J. Select. Areas Commun.*, Vol. 8, No. 4, pp. 683-690, May 1990.

[134] B. Yang, "Projection Approximation Subspace Tracking," *IEEE Trans. Signal Processing,* Vol. 44, No. 1, pp. 95 - 107, 1995.

[135] X. Zhang and D. Brady, "Soft-decision Multistage Detection for Asynchronous AWGN Channels," *Proc. 31st Annual Allerton Conf. on Communication, Control and Computing,* Monticello, IL, Sept. 1993.

[136] X. Zhang and D. Brady, "Narrow Bandwidth Waveform Design for Near-far Resistant Multiuser Detection," *Proc. MILCOM'94*, Ft. Monmouth, NJ, October 1994.

[137] F-C Zheng and S. K. Barton, "On the Performance of Near-Far Resistant CDMA Detectors in the Presence of Synchronization Errors," *IEEE Trans. Commun.*, Vol. 43, No. 12, pp. 3037-3045, Dec. 1995.

[138] Z. Zvonar, D. Brady, and J. Catipovic, "Adaptive Receivers for Cochannel Interference Suppression in Shallow-water Acoustic Telemetry Channels," *Proceedings of the 1994 International Conference on Acoustics, Speech and Signal Processing (ICASSP'94)*, Adelaide, Australia.

[139] Z. Zvonar and D. Brady, "Multiuser Detection in Single-path Fading Channels," *IEEE Trans. Commun.*, Vol. 42, No. 2/3/4, pp. 1729 - 1739, 1994.

[140] Z. Zvonar and D. Brady, "Differentially Coherent Multiuser Detection in Asynchronous CDMA Flat Rayleigh Fading Channels," *IEEE Trans. Commun.*, Vol. 43, pp. 1252 - 1255, 1995.