ABSTRACT

TRANSIENT BEHAVIOR OF AN ADAPTIVE SYNCHRONOUS CDMA RECEIVER

by
Bin Zhu

A steepest descent algorithm is used to update the adaptive weights of a two-stage synchronous Code-Division Multiple-Access (CDMA) receiver that was proposed recently. An issue of the adaptive CDMA system — the convergence and stability property of the receiver is investigated in this thesis.

This adaptive synchronous CDMA receiver uses a decorrelator at the first stage and adopts a neural network which acts as an interference canceler at the second stage. It can achieve near-optimum performance. Furthermore, its computational complexity is just a square function of the number of users. The only requirement is the knowledge of the users' signature sequences.

The analysis shows that the algorithm for the adaptive weights is convergent and straightforward in implementation. The guaranteed fast convergence of the receiver weights and the tractable theoretical analysis on it, as revealed in this thesis, make this adaptive receiver a promising approach for wireless communications.

TRANSIENT BEHAVIOR OF AN ADAPTIVE SYNCHRONOUS CDMA RECEIVER

by
Bin Zhu

A Thesis
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Master of Science in Electrical Engineering

Department of Electrical and Computer Engineering

January 1994

Blank Page

APPROVAL PAGE

TRANSIENT BEHAVIOR OF AN ADAPTIVE SYNCHRONOUS CDMA RECEIVER

Bin Zhu

Dr. Nirwan Ansari, Thesis Advisor   Date
Associate Professor of Electrical and Computer Engineering
New Jersey Institute of Technology

Dr. Zoran Siveski, Committee Member   Date
Assistant Professor of Electrical and Computer Engineering
New Jersey Institute of Technology

Dr. Alexander Haimovich, Committee Member   Date
Associate Professor of Electrical and Computer Engineering
New Jersey Institute of Technology

BIOGRAPHICAL SKETCH

Author: Bin Zhu

Degree: Master of Science in Electrical Engineering

Date: January 1994

Undergraduate and Graduate Education:

- Master of Science in Electrical Engineering,
  New Jersey Institute of Technology, Newark, NJ, 1994

- Bachelor of Science in Electrical Engineering,
  Shanghai Jiao Tong University, Shanghai, P.R.C., 1988

Major: Electrical Engineering

This thesis is dedicated to my parents
and my husband

ACKNOWLEDGMENT

I would like to thank my advisor, Dr. Nirwan Ansari, for his guidance and support throughout my Master's program.

I would also like to thank my co-advisor, Dr. Zoran Siveski for his valuable advice on my research and thesis.

Special thanks to Dr. Alexander Haimovich for serving as member of the committee.

Furthermore, I would like to mention Lisa Fitton whose timely advice on the writing of the thesis is much appreciated.

Last but not least, I would like to express my appreciation for all the love and support given to me by my husband and friends.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Several techniques are available to transmit a plurality of different source messages over a common channel simultaneously, such as Time-Division Multiplexing (TDM), which uses different time slots; Frequency-Division Multiplexing (FDM), which uses different frequency bands; and Code-Division Multiplexing (CDM), which uses different pre-assigned code waveforms (signature sequence). Code-Division Multiple-Access (CDMA) has recently been adopted as one of the accessing techniques for wireless personal and mobile communications. The reasons lie in its promise of performance, and most of all, its high degree of flexibility — there is no "hard" limit on the number of users that can simultaneously access the system; adding more users just causes a graceful degradation of system performance.

In this thesis, we consider a synchronous CDMA communication system. In CDMA, multiple accessing is actually achieved by spreading the spectrum of trans-mitted signals with pre-assigned signature sequences. Each transmitter-receiver pair is designed to encode and decode the information bit using a specific sequence. Each receiver partially receives other user signals because of the cross-correlation between their signature sequences. Therefore, in a multiple-user communication system, the noise present at the receiver can be modeled as the sum of the channel noise with additive white Gaussian properties and the interfering noise caused by other users. Although the interfering power can be reduced by carefully choosing the signature sequences with low cross-correlation, it results in the reduction of the maximum number of users that can access the system simultaneously, rendering it an impractical solution.

1

There are various approaches in designing CDMA receivers. The conventional single-user receiver that employs a matched filter is optimum in the sense of a maximum output signal-to-noise ratio, where the noise is modeled as a Gaussian noise. In the multiple-user case, since the interfering noise cannot be accurately modeled as a Gaussian noise, the conventional receiver considers the background noise only and, therefore, it can not be considered optimum. As a result of interference "ignorance", the performance of the conventional receiver closely depends on the power of the interfering signals. Severe degradation of the system, known as the "near-far" problem, can happen because of strong interference caused by the large transmitted power of the interfering signals. To cope with this problem, a power control system that adjusts the powers of the transmitted signal and the interfering signals has to be built, which increases the complexity.

An optimum receiver in multiple-user environments that uses a Viterbi algorithm is proposed by Verdu [1]. With the *power* of each user known, a decision can be made by choosing one possible signal vector that minimizes the cost function. The receiver has been proven to be insensitive to the "near-far" problem. Its practicality is hampered by the requirement of the knowledge of each user's power. In addition, its computational complexity increases *exponentially* with the number of users.

Several suboptimum receivers have been proposed to achieve the reduced computational complexity at the cost of lower performance [2], [3]. The suboptimum two-stage decorrelating receiver analyzed in [3] is one of them. The first stage of the receiver is a decorrelator. The original signals are considered as the signals at the input of this stage. Interference-free signals can be achieved at the output of the stage; however, the noise power is enhanced by this stage. To achieve better signal-to-noise ratios, the second stage is implemented. It is an interference canceler

with *fixed* weights which require the knowledge of the users' powers. What it does is subtract the interference from the original signals again based on the results of the first stage and, therefore, avoids the effect of noise enhancement. The receiver has good performance with regard to error probability, and its computational complexity increases linearly with the number of users. However, a shortcoming of the receiver is the required knowledge of the user powers.

Along another path, interference cancelers with an *adaptive* concept were introduced for the multiple-user environment [4]. Later, the same structures and a similar adaptive rule were proposed independently by a group of European researchers [5], [6]. Since then, a lot of work has been done in light of different structures and adaptive algorithms [7], [8], [9], [10], [11]. The adaptive receivers are designed to either track a parameter of the communication environment or adapt to unknown user parameters. In this case, the *convergence* of the algorithm and the *stability* of the system become main issues.

In [10], Aazhang proposed a CDMA receiver which consists of a multi-layer neural network trained by the back-propagation algorithm. The training is considered to be successful with regard to the convergence (unlikely to be trapped in local minima) and stability of the algorithm. The receiver is shown to achieve good performance in terms of error probability after the training. Still, problems arise regarding the number of neurons needed, which increases *exponentially* with the number of users, and the time needed for training cannot be ignored.

In [11], U. Mitra and H.V. Poor proposed a single-layer perceptron scheme with different filtering techniques for single-user demodulation in a multiple-user channel. The weights are shown to converge to the optimal values in a noiseless environment and the convergence is guaranteed. With additive Gaussian channel noise added to

the system, the performance is not satisfactory. Furthermore, the convergence rate is relatively slow.

The above two neural network receivers are trained by *supervised* learning algorithms. They all require reference training sequences to perform the system training task, which is impractical in implementation. Unsupervised training is then desirable in the CDMA environment.

The *"bootstrap"* algorithm was first introduced in [4] for interference cancellation. In [8], [9], three different bootstrap structures (backward-backward, forward-forward, and forward-backward) of the bootstrap blind adaptive algorithm for multi-signal co-channel separation were analyzed. Recently, the same structures were proposed independently in [5] and [6]. In [5], a backward-backward network structure was proposed with an adaptive unsupervised algorithm. This was first applied successfully to some continuous signals. But investigation revealed that this kind of algorithm has severe drawbacks on *convergence* and *stability* [12], [13]. In [6], two other structures (two-stage mixed and forward-forward) were proposed and compared to the backward-backward structure. These structures can not overcome the same drawbacks on convergence and stability.

The synchronous adaptive CDMA receiver investigated in this thesis was recently proposed in [14]. The idea was inspired by the suboptimum two-stage decorrelating receiver mentioned earlier [3]. Instead of using *fixed* weights (which need the knowledge of the users' powers) for the interference canceler, this receiver employs a neural network updated by a steepest descent algorithm. By doing this, the weights are able to adjust their values and converge to the optimal weights. Therefore, the knowledge of the users' powers is *not required* by this receiver. This factor also makes it possible for the receiver to work in a power changing environment, which is a very important feature in mobile communications.

The system has been shown to be "near-far" resistant, achieves very good performance with regard to error probability, and functions without knowing the users' powers. It can even perform at the same level as the optimum receiver when the users' SNRs are high [14]. Moreover, the computational complexity is just a square function of the number of users.

A thorough investigation on the convergence and stability of the system is presented in this thesis. Several conditions for the system to achieve convergence are derived, and their properties are analyzed.

In chapter 2, earlier work on non-adaptive multiple-user receivers is briefly reviewed for comparison. In Chapter 3, we focus on adaptive multiple-user receivers, especially on single-layer neural networks with unsupervised learning algorithm (or "bootstrap" algorithm). The algorithm, as well as their convergence conditions will be revealed. Simulations on this kind of adaptive CDMA receivers were done and the results will also be shown. In Chapter 4, the recently proposed synchronous adaptive CDMA receiver will be introduced and its convergence and stability properties will be analyzed. Details of the two-user case and three-user case, as well as numerical results, are given for the purpose of illustration. The conclusion of the thesis is given in chapter 5.

CHAPTER 2

LITERATURE SURVEY

2.1 System Model

The synchronous code-division multiple-access environment is shown in Figure 2.1.
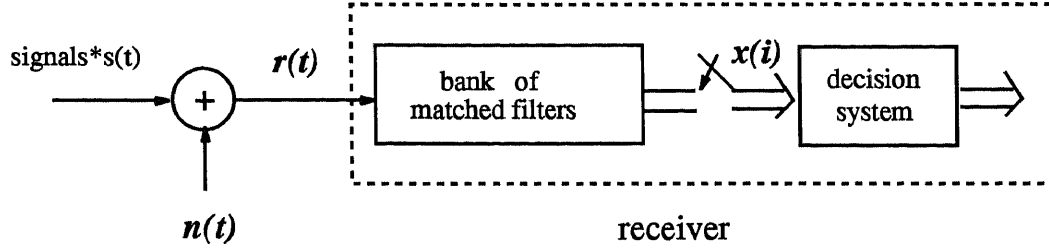
Figure 2.1 The General Receiver Model

Suppose that there are a total of K active users in this environment. Define $b(i) = [b_1(i), b_2(i), \cdots, b_K(i)]^T$ as a column vector corresponding to the K sources' information bits transmitted at the ith time interval, where $b_k \in \{-1, 1\}$ (if assumed antipodal, as in this thesis), $1 \le k \le K$, is the information bit of the kth source signal with duration T_0. Also define $A = diag[\sqrt{A_1}, \sqrt{A_2}, \cdots, \sqrt{A_K}]$ as a diagonal matrix, where A_k is the received power of the kth source signal. Define $s_k(t), 1 \le k \le K$, as the kth user's signature waveform with its time interval limited between $[0, T_0)$. Suppose a code word consists of m chips, each of duration τ, then $T_0 = m\tau$. Without loss of generality, also set that $\int_0^{T_0} s_k^2(t)dt = 1$. $n(t)$ is the background channel noise modeled as an AWGN with a zero mean and a double-sided power spectral density of $N_0/2$.

For the kth source, each information bit b_k is encoded by the signature sequence $s_k(t)$, which means that the spectrum of the transmitted signal is spread and therefore much wider than the bandwidth of the original signal. After being encoded, all the K signals are transmitted through the same channel. Consider the case where all users in the system are bit-synchronous, the signal $r(t)$ received at the output of the channel is the sum of all K signals and the noise $n(t)$, which can be expressed as:

$$r(t) = \sum_{k=1}^{K} \sum_{i} b_k(i)\sqrt{A_k}s_k(t - iT_0) + n(t). \tag{2.1}$$

Consider the ideal case of no intersymbol interference for the signature sequences, then at time interval i, i.e., $iT_0 \le t < (i+1)T_0$, the above equation can be simplified as

$$r(t) = \sum_{k=1}^{K} b_k(i)\sqrt{A_k}s_k(t - iT_0) + n(t). \tag{2.2}$$

The signal $r(t)$ is then passed through a bank of matched filters, which results in

$$x = \rho Ab + n, \tag{2.3}$$

where $x = [x_1, x_2, \ldots, x_K]^T$, $n = [n_1, n_2, \ldots, n_K]^T$. n_k is the output of the channel noise through the kth matched filter. It can be proven that n_k is still a Gaussian noise with zero mean and power spectral density of $N_0/2$, and the covariance matrix of the vector n is

$$E\{nn^T\} = \frac{N_0}{2}\rho,$$

where ρ is the cross-correlation matrix with its (k,j)th element defined as:

$$\rho_{kl} = \int_0^{T_0} s_k(t)s_l(t)dt \qquad k, l \in (1, 2, \ldots, K), \tag{2.4}$$

with $\rho_{kk} = 1$ and $\rho_{kl} = \rho_{lk}$. Therefore, it can be seen that the noise at this point consists of a background noise inherited from the channel and an interference caused by other users. The signal vector x will then be processed by a decision system.

2.2 The Non-adaptive Receiver

2.2.1 The Conventional Receiver

A conventional single-user receiver is a correlation receiver or a matched filter receiver that is optimum in the sense of maximum signal-to-noise ratio. In a K user environment, it refers to a receiver that employs K such single-user receivers. One of the conventional synchronous CDMA receivers is depicted in Figure 2.2.
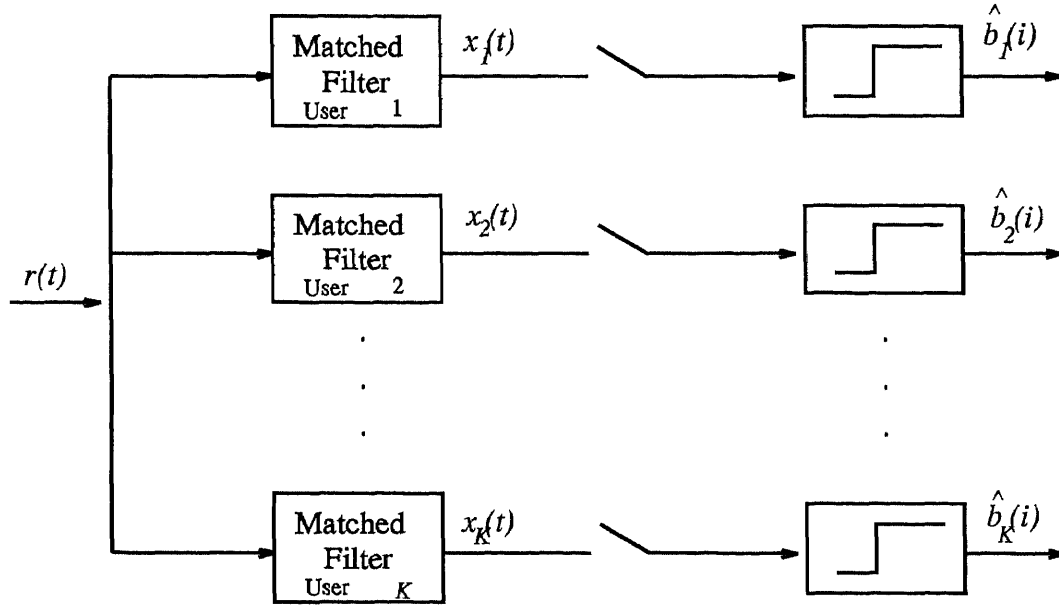
Figure 2.2 The Conventional Receiver

As mentioned earlier, the input to the receiver can be expressed as:

$$r(t) = \sum_{k=1}^{K} \sqrt{A_k} b_k(i) s_k(t - iT_0) + n(t).$$

The signal received at the output of the kth matched filter becomes

$$x_k(t) = \sqrt{A_k} b_k(i) + \sum_{l \neq k} \rho_{kl} \sqrt{A_l} b_l(i) + n_k(t), \tag{2.5}$$

where $b_k(i)$ is the desired signal, $\sum_{l \neq k} \rho_{kl}\sqrt{A_l}b_l(i)$ is the interference from other users, and noise $n_k(t)$ is a Gaussian noise with the following properties:

1. $E\{n_k(t)\} = 0$.

2. $E\{n_k^2(t)\} = N_0/2$.

3. $E\{n_k(t)n_l(t)\} = \rho_{kl}N_0/2, \quad k \neq l$.

These properties can be verified easily by using the fact that $s_k(t)$ is deterministic. A decision is then made on the output of the matched filter:

$$
\begin{aligned}
\hat{b}_k(i) &= sgn[x_k(i)] \\
&= sgn[\sqrt{A_k}b_k(i) + \sum_{l \neq k} \rho_{kl}\sqrt{A_l}b_l(i) + n_k(t)].
\end{aligned}
\tag{2.6}
$$

It is clear that in order to accurately recover the source information bit b_k, the interference from other users, in terms of $\sum_{l \neq k} \rho_{kl}\sqrt{A_l}b_l(i)$, should be small. When the interfering signal power is much larger than the desired signal power (which can happen when a interfering station is much nearer to the receiver than its source station), it is very difficult for the receiver to recover the desired signal. This is the famous "near-far" problem, which is the major drawback of the conventional receiver.

2.2.2 The Optimum Receiver

The optimum receiver in the multiple-user case is then exploited. It can make decisions according to two different decision rules. One is maximum-likelihood sequence detection, and another is minimum-probability-of-error detection. The receiver is depicted in Figure 2.3.

Here, the Viterbi algorithm is used. With all information sequences assumed to be equiprobable, the maximum-likelihood decision or optimum decision on $b(i)$
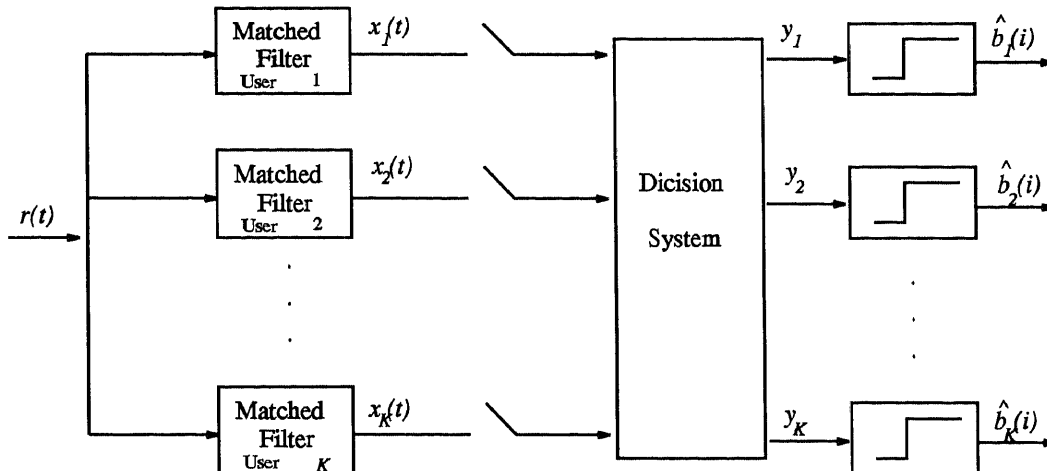
Figure 2.3 The Optimum Receiver

can be made by only observing the signal vector received at ith bit duration. Denote $\hat{\boldsymbol{b}}(i) = [\hat{b}_1(i), \hat{b}_2(i), \cdots, \hat{b}_K(i)]^T$ as the estimation of $\boldsymbol{b}(i)$, the cost function is

$$J = \int_{t_0}^{t_0+T_0} [r(t) - \sum_{k=1}^{K} \hat{b}_k(i) A_k s_k(t - iT_0)]^2 dt, \qquad (2.7)$$

which is the energy of the noise in one bit duration. When $\hat{\boldsymbol{b}}(i) = \boldsymbol{b}(i)$, the cost function reaches its minimum:

$$J_{min} = \int_{t_0}^{t_0+T_0} n^2(t) dt, \qquad (2.8)$$

which is the energy of the channel noise. Therefore, a decision is made according to the following:

$$\hat{\boldsymbol{b}}(i) = \arg\{ \min_{\boldsymbol{b} \in \{-1,1\}^K} [\int_{t_0}^{t_0+T} [r(t) - \sum_{k=1}^{K} \hat{b}_k(i) A_k s_k(t - iT_0)]^2 dt] \}, \qquad (2.9)$$

or equivalently, by maximizing a log likelihood function [3]:

$$\hat{\boldsymbol{b}}(i) = \arg\{ \max_{\boldsymbol{b} \in \{-1,1\}^K} [2\boldsymbol{x}^T \boldsymbol{b} - A\boldsymbol{b}^T \rho \boldsymbol{b}] \}. \qquad (2.10)$$

Although the optimum receiver can efficiently resist the "near-far" problem, it has severe limitations on complexity and it requires the knowledge of all the signal

powers. Since the number of choice on $\hat{b}_k$ is 2^K for antipodal signals in the K-user case, the computational complexity increases exponentially with the number of users, which is impractical in implementation.

2.2.3 The Suboptimum Receiver

Suboptimum receivers are proposed to achieve near-optimum performance with much less complexity. Different interference cancelers are employed in this kind of receiver. The two-stage decorrelating receiver is one that emerged recently. It employs a decorrelator as the first stage to estimate the information bit and performs the interference cancellation at the second stage based on the former estimation [3]. This non-adaptive receiver is depicted in Figure 2.4.
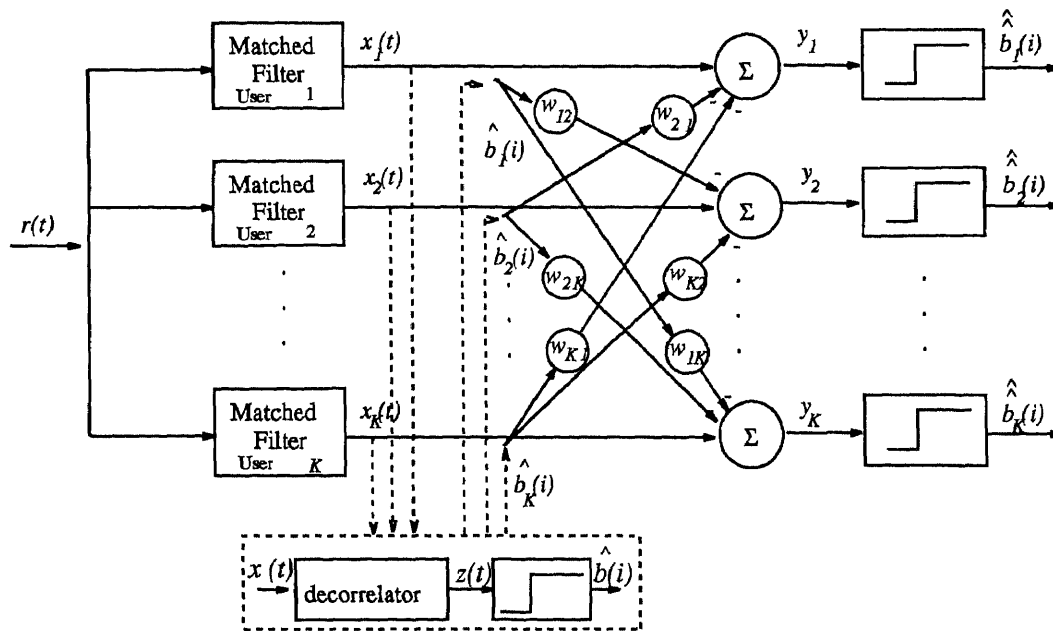

Figure 2.4 The Two-Stage Non-Adaptive Suboptimum Receiver

Assume $\boldsymbol{P}$ is the $K \times K$ decorrelator matrix with its element denoted as p_{kl}, and $\boldsymbol{z} = [z_1, z_2, \cdots, z_K]^T = \boldsymbol{Px}$ is the signal vector obtained at the output of the decorrelator. Since

$$\boldsymbol{x} = \rho \boldsymbol{Ab} + \boldsymbol{n} \quad and \quad \boldsymbol{z} = \boldsymbol{Px},$$

$\boldsymbol{P}$ can be chosen as ρ^{-1} to make the output signals $\boldsymbol{z}$ interference-free. Therefore,

$$\boldsymbol{z} = \boldsymbol{Px} = \rho^{-1}(\rho \boldsymbol{Ab} + \boldsymbol{n}) = \boldsymbol{Ab} + \rho^{-1}\boldsymbol{n}. \tag{2.11}$$

With $\hat{\boldsymbol{b}} = sgn[\boldsymbol{z}]$, the output of the interference canceler can be expressed as

$$\boldsymbol{y} = \boldsymbol{x} - \boldsymbol{W}\hat{\boldsymbol{b}} = \rho \boldsymbol{Ab} + \boldsymbol{n} - \boldsymbol{W}\hat{\boldsymbol{b}}, \tag{2.12}$$

where $\boldsymbol{y} = [y_1, y_2, \cdots, y_K]^T$ is the output signal vector of the second stage, and $\boldsymbol{W}$ is the $K \times K$ weight matrix fixed as $\boldsymbol{W} = (\rho - \boldsymbol{I})\boldsymbol{A}$. When signal SNRs are larger enough, $\hat{\boldsymbol{b}} \approx \boldsymbol{b}$. Then equation (2.12) becomes

$$\boldsymbol{y} \approx \boldsymbol{Ab} + \boldsymbol{n}. \tag{2.13}$$

The output of the receiver is then

$$\hat{\hat{\boldsymbol{b}}} = sgn[\boldsymbol{y}], \tag{2.14}$$

where $\hat{\hat{\boldsymbol{b}}} = [\hat{\hat{b}}_1, \hat{\hat{b}}_2, \cdots, \hat{\hat{b}}_K]^T$ is the final decision of the information vector $\boldsymbol{b}$. The reason for choosing $\hat{\hat{\boldsymbol{b}}}$ instead of $\hat{\boldsymbol{b}}$ is that the SNR at the output of the decorrelator is too vulnerable to the large cross-correlation ρ. This can be seen easily by comparing equation (2.11) and (2.13).

The following is an illustration of the 2×2 case. Denote $\rho = \rho_{12} = \rho_{21}$. The cross-correlation matrix ρ in this case is $\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$. The background noise present in equation (2.11) and (2.13) become

$$\begin{aligned}
\boldsymbol{n_z} &= \rho^{-1}\boldsymbol{n} \\
&= \frac{\begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}}{1 - \rho^2}\boldsymbol{n} \\
&= \frac{1}{1 - \rho^2}\begin{pmatrix} n_1 - \rho n_2 \\ n_2 - \rho n_1 \end{pmatrix},
\end{aligned} \tag{2.15}$$

and

$$\boldsymbol{n_y} = \boldsymbol{n}.$$

Since $E\{n_1\} = E\{n_2\} = 0$, $E\{n_1 n_2\} = \rho\frac{N_0}{2}$, and $E\{n_1^2\} = E\{n_2^2\} = \frac{N_0}{2}$, from equation (2.15), the following hold:

$$\begin{aligned}
E\{\boldsymbol{n_z^2}\} &= \rho^{-2}E\{\boldsymbol{n}^2\} \\
&= \frac{1}{(1 - \rho^2)}\begin{pmatrix} \frac{N_0}{2} \\ \frac{N_0}{2} \end{pmatrix} \\
&= \frac{1}{(1 - \rho^2)}E\{\boldsymbol{n}^2\} \\
&> E\{\boldsymbol{n_y^2}\}.
\end{aligned} \tag{2.16}$$

It is clear that the background noise power present at the output of the decorrelator is larger than the one present at the output of the interference canceler. When the signal SNRs are large, and the weights are fixed ($w_{21} = \rho sqrt(A_2)$ and $w_{12} = \rho sqrt(A_1)$), then $\boldsymbol{y} \approx \boldsymbol{Ab} + \boldsymbol{n}$, the output signals of the second stage are almost interference free. In this case, the error probability of the receiver is better than the one of the first stage. In the case that the signal SNRs are small, the output signals of the receiver is not interference free, then the error probability of the receiver varies depending on the powers of all the users. Evaluations of their probability of error has already been done in [14], [15] .

The receiver achieves good performance on error probability with computational complexity linearly increasing with the number of users. However, the

requirement of a priori knowledge of the transmitted signal powers is the drawback of the receiver.

Since the non-adaptive receivers require the knowledge of the users' signal powers and their signature sequences, neural networks with adaptive concept have been proposed to perform the same separating task with less knowledge of the sources. This kind of receivers as well as their simulations will be discussed in chapter 3.

CHAPTER 3

NEURAL NETWORKS AS SYNCHRONOUS CDMA RECEIVERS

There are two kinds of algorithms used to train neural networks. One is the supervised learning and the other is the unsupervised learning. Since in supervised learning, a training sequence is needed each time to train the system, it is not always desirable in implementation.

3.1   CDMA Receivers with Supervised Learning Algorithms

3.1.1   A Multi-layer Neural Network as a CDMA Receiver

In [10], a multi-layer neural network CDMA receiver is proposed, as shown in Figure 3.1.
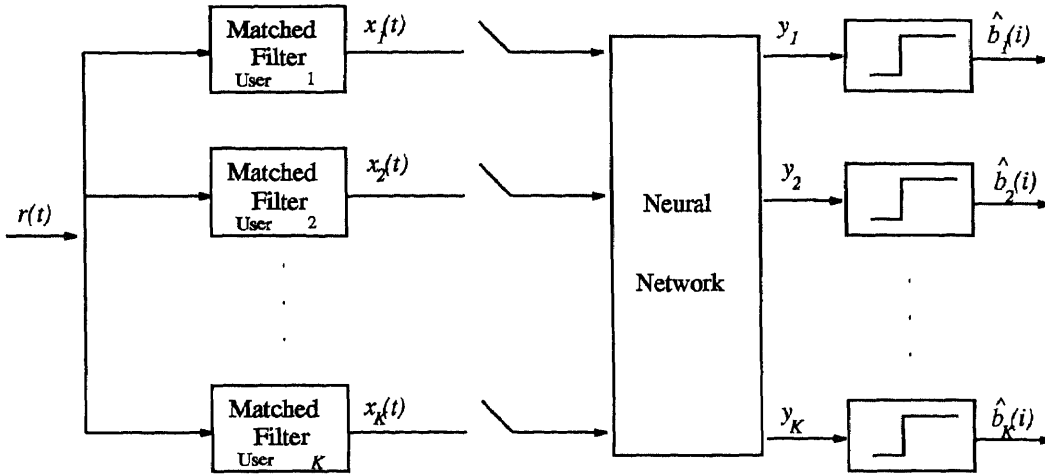

Figure 3.1 A Multi-layer Suboptimum Adaptive CDMA Receiver

In this structure, the neural network is a multi-layer feed-forward network. Employing the gradient descent algorithm to minimize the error function, the back-propagation algorithm is used to train the network. In comparison to the supervisor (the desired output), the derivative of the error function is computed to train each neuron. The training starts at the output layer, then the error function is back-propagated to the former layers. Usually, the algorithm will reach a local minima. In [10], some new techniques were developed to deal with this problem. The receiver is "near-far" resistant and the performance of the receiver has been shown to be suboptimum. Although the convergence problem is solved, the computational complexity is still high. With the number of neurons increasing exponentially with the number of users, the time used to train the network increases exponentially, as well.

3.1.2   A Single-Layer Neural Network as a CDMA Receiver

In [11], a single-layer perceptron with different filter operations as a CDMA receiver is proposed in a noiseless environment. Two nonlinearity functions are used for each neuron. One uses the hard limiter and the other uses the Sigmoid function.

3.1.2.1   The Receiver Model

The scheme is depicted in Figure 3.2.

As shown in the figure, this receiver only considers one source signal as the desired signal. The other user signals are considered interfering signals. Assume the input vector to the receiver is $x = [x_1, x_2, \cdots, x_K]^T$, the output of the receiver at the ith interval is
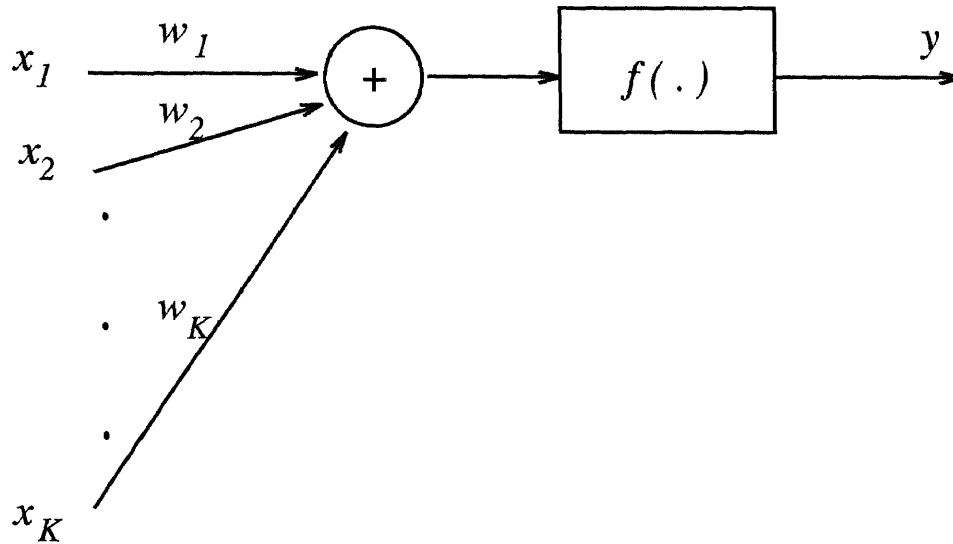
$$y_i = f(w_i^T x_i),$$

Figure 3.2 A Single-layer Suboptimum Adaptive CDMA Receiver

where $w_i = [w_1, w_2, \cdots, w_K]^T$ is the weight vector. The updating algorithm is described as

$$w_{i+1} = w_i + \mu[d_i - y_i]f'(w_i^T x_i)x_i, \tag{3.1}$$

where d_i is the desired output. Using the hard limiter, $f(\cdot) = sgn(\cdot)$, leading to $f'(\cdot) = 1$. Using the Sigmoid function, $f(s) = \frac{1}{1+e^{-s}}$, $f'(s)$ is also a non-linear function. The convergence of the algorithm with the hard limiter is easily established [16]. The convergence of the algorithm using the Sigmoid function is proven [11] with a condition on μ, which is $0 < \mu < \frac{32}{A_{max}^2 M^2 K}$, where M is the length of the code word and A_{max} is the maximum user power. The error probability performance of the receivers using these two algorithms shows that the weights converge to optimal values in a noiseless multiple-user environment. With additive white Gaussian noise, convergence can also be reached, but the receiver is no longer optimal. Comparing the receivers with the hard limiter and the Sigmoid function, the latter is less affected by noise, but it has a slower convergence speed.

Although the above analysis considers only one desired output, it can be easily expanded to the K-user case by using K such receivers together.

3.1.2.2 Simulations

Simulations on the receiver using the hard limiter are shown in two- and three-user cases. Figure 3.3 shows the transient behavior of the weights to users 1 and 2 in the two-user case under the conditions that $SNR_1 = SNR_2 = 8 \ dB$, $\rho = 0.7$, and $\mu = 0.01$. The steady state error probability under the same conditions is depicted in Figure 3.4 with SNR_1 fixed at $8 \ dB$,

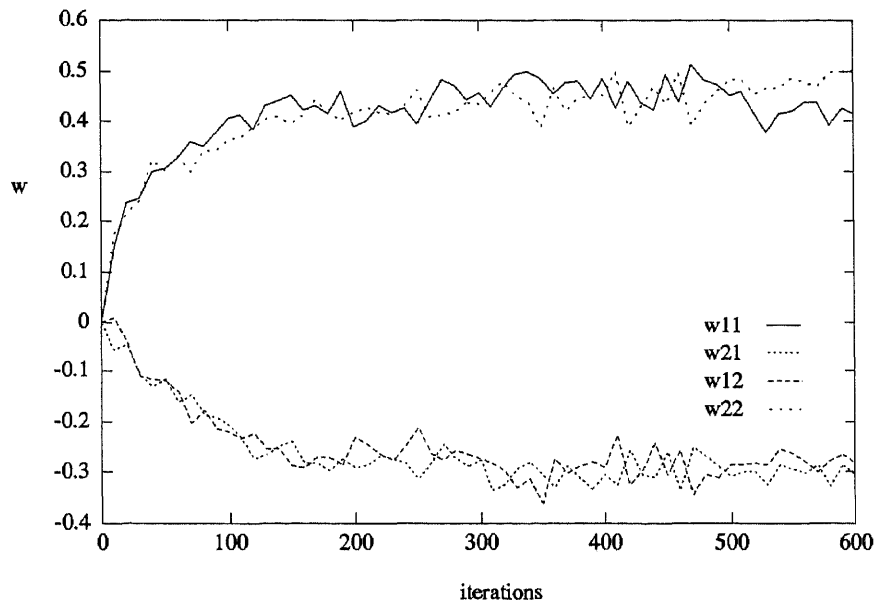**Figure 3.3** The Transient Weights to User 1 and 2 in the Two-User Case

Gold sequences of length seven (frequently used in the literature) are chosen for signature waveforms. The rationale for such a choice is that Gold sequences are regularly used in asynchronous CDMA environments and the study of their proposed synchronous counterparts may provide a useful indication of the performance of the
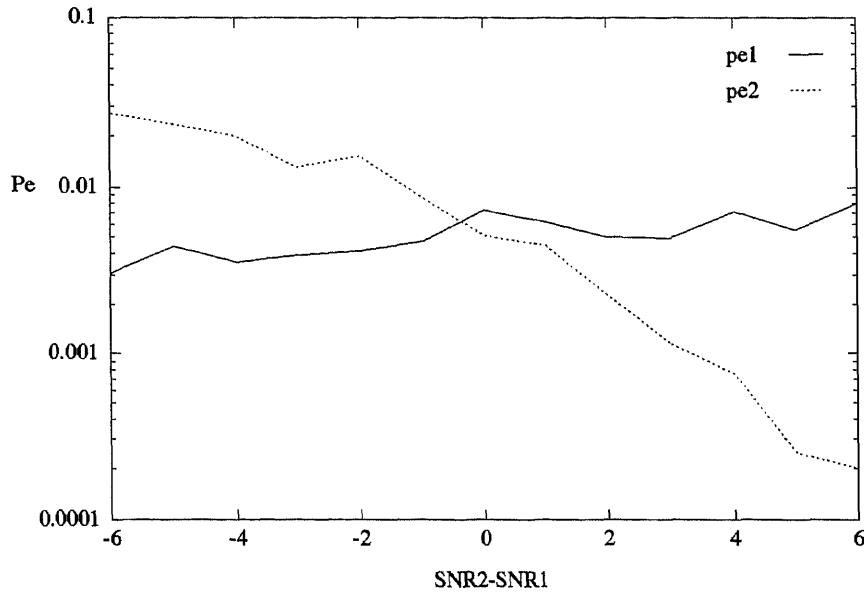
Figure 3.4 The Steady State Error Probability in the Two-User Case

former. The cross-correlation matrix ρ in this case is:

$$\rho = \frac{1}{7} \begin{bmatrix} 7 & -1 & 3 & 3 & 3 \\ -1 & 7 & -1 & 3 & -1 \\ 3 & -1 & 7 & -1 & -1 \\ 3 & 3 & -1 & 7 & -1 \\ 3 & -1 & -1 & -1 & 7 \end{bmatrix}.$$

The simulations for the three-user case are depicted in Figure 3.5 and 3.6 with Gold sequences as their signature waveforms. Figure 3.5 depicts the transient behavior of the weights to user 1 under the conditions that all the signal SNRs are equal 8 dB and $\mu = 0.01$. Figure 3.6 shows the error probability simulated with SNR_1 fixed at 8 dB, SNR_2 and SNR_3 ranging from -6 to 6 dB.
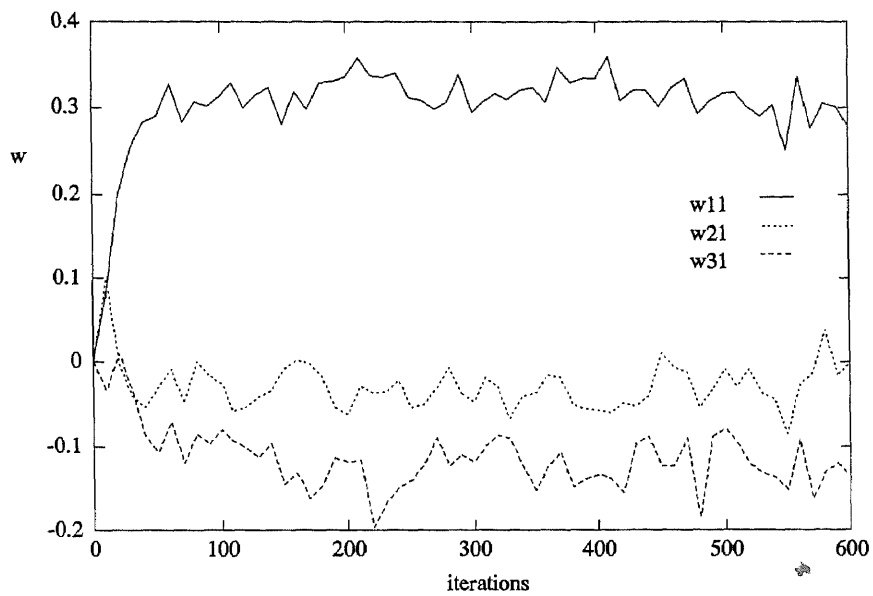
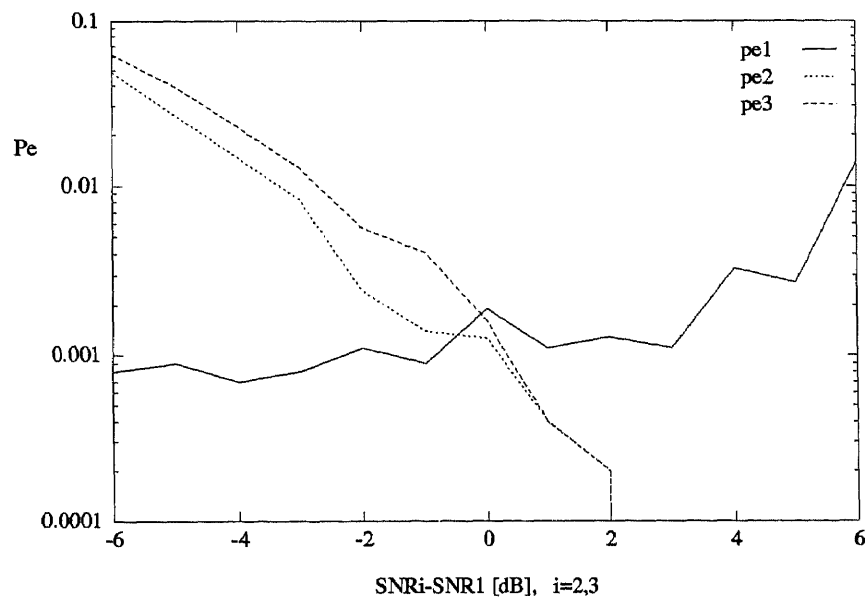**Figure 3.5** The Transient Weights to User 1 in the Three-User Case

Figure 3.6 The Steady State Error Probability in the Three-User Case

3.2 CDMA Receivers with Unsupervised Learning Algorithms

The "bootstrap" algorithm with different structures is dated back to 1981 [4]. These structures function as noise cancelers with or without feedback. Starting with 2×2 case, a lot of work was done to investigate the convergence, stability, and the performance. It was also extended to the multiple-user case. Recently, several scientists in Europe proposed the similar structures independently, to separate superimposed signals in an analog channel with the assumption that sources are statistically independent [5], [6], [12], [13]. Since these structures have similar characteristics, their performance and convergence properties are also similar.

3.2.1 The Structures

3.2.1.1 The Backward Structure

The two-user structure is depicted in Figure 3.7.

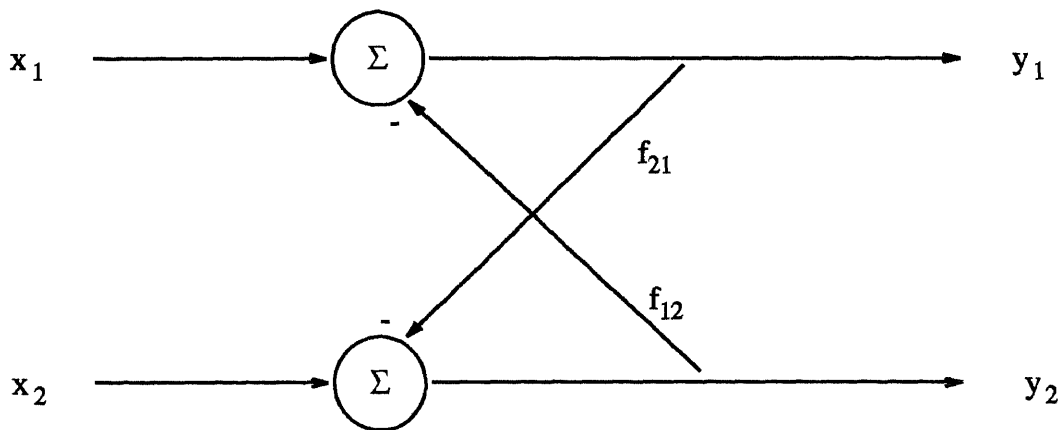**Figure 3.7** The Backward Structure

Denote vectors $\boldsymbol{y} = [y_1, y_2]^T$, $\boldsymbol{x} = [x_1, x_2]^T$, and matrix $\boldsymbol{F} = \begin{bmatrix} 0 & -f_{12} \\ -f_{21} & 0 \end{bmatrix}$, the output vector $\boldsymbol{y}$ can be expressed as:

$$y = x + Fy \Rightarrow y = (I - F)^{-1}x, \tag{3.2}$$

where I is the identity matrix.

Assume that the input vector x is the linear combination of the source vector $b = [b_1, b_2]^T$, i.e., $x = Bb$, where B is the mixing matrix. In order to retrieve the sources at the outputs of the structures, the following must hold:

$$y = (I - F)x \quad and \quad b = B^{-1}x$$

$$\Rightarrow I - F = KB^{-1},$$

where K is a constant matrix with elements k_{ij}, $i, j \in 1, 2$. There are two separation cases:

- $k_{ii} = 0$, which leads to

$$y_1 = k_{12}b_2, \qquad y_2 = k_{21}b_1.$$

The corresponding separating point for (f_{12}, f_{21}) is then $(\frac{b_{12}}{b_{22}}, \frac{b_{21}}{b_{11}})$.

- $k_{ij} = 0$, $\forall i \neq j$. Therefore,

$$y_1 = k_{11}b_1, \qquad y_2 = k_{22}b_2.$$

The corresponding separating point for (f_{12}, f_{21}) is $(\frac{b_{11}}{b_{21}}, \frac{b_{22}}{b_{12}})$.

Since matrix B is unknown, an adaptive rule is proposed to estimate the weights. Because nothing else but the statistic independence of the sources is known, the function to be minimized or maximized must be able to verify this. The updating rule proposed in [5] to realize independence is given as follows:

$$f_{ij}(n + 1) = f_{ij}(n) + \mu g(y_i)l(y_j), \quad i \neq j, \tag{3.3}$$

where μ is the learning step between $(0,1)$, and $g(*)$ and $l(*)$ are two non-linear and different odd functions. Since the source signals usually have even probability densities, verifying the independence of the output signals y_i (proportional to the source signals) simplifies to verifying that $E\{g(y_i)l(y_j)\} = E\{g(y_i)\}E\{l(y_j)\} = 0$. Simple odd functions are chosen, $g(x) = x^3$ and $l(x) = x$. Once the separation of the source signals has been achieved, the weights f_{ij} reach their steady states. This updating rule is also used for the following structures.

3.2.1.2 The Direct Structure

The two-user structure is depicted in Figure 3.8.

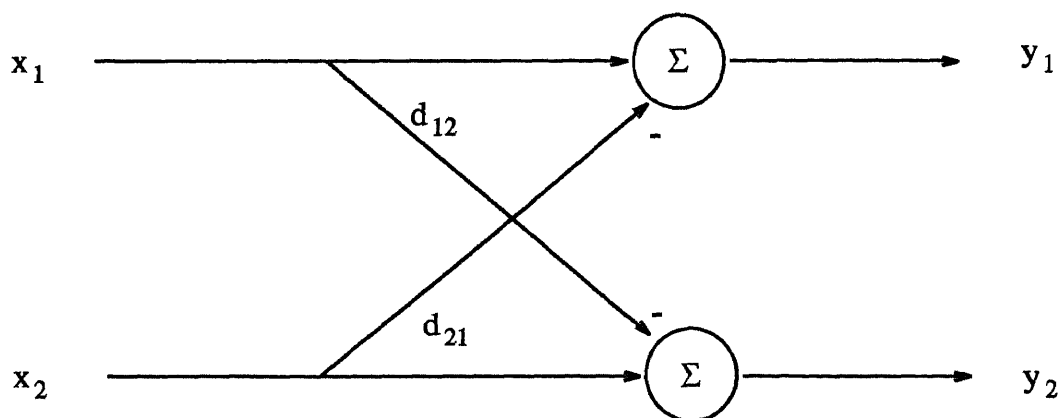**Figure 3.8** The Direct Structure

Defining matrix $D = \begin{bmatrix} 0 & -d_{12} \\ -d_{21} & 0 \end{bmatrix}$, the following holds:

$$y = (I + D)x. \tag{3.4}$$

The same procedure can be used to calculate the separating points for (d_{12}, d_{21}). The results are the same as the feedback structure.

3.2.1.3 The Two-stage Mixed Structure

The two-user two-stage structure is depicted in Figure 3.9.

(M)
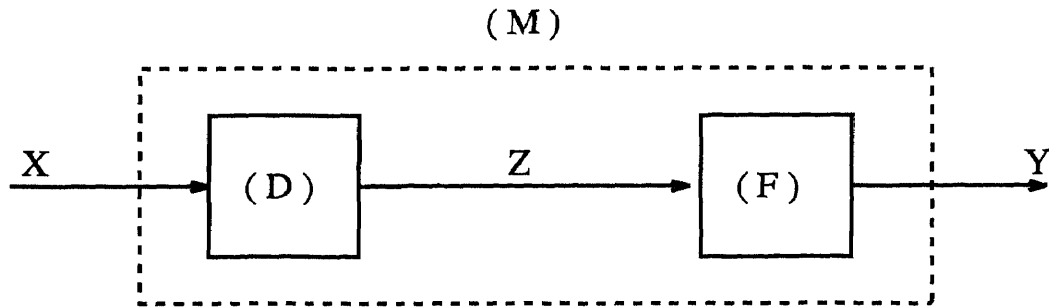
X → (D) → Z → (F) → Y

Figure 3.9 The Two-Stage Mixed Structure

This is obviously a combination of the previous two structures. Using the same updating rule and setting weight $m_{ij} = d_{ij} = f_{ij}$, where $i, j \in 1, 2$, the effect of the receiver is equal to two one-stage receivers, yielding faster convergence. By defining matrix $M = \begin{bmatrix} 0 & -m_{12} \\ -m_{21} & 0 \end{bmatrix}$, the output becomes

$$y = (I - M)^{-1}(I + M)x. \tag{3.5}$$

Solving the separating points [6] for the structure is much more complicated than the previous cases. The structure has two separating points, which correspond to each separating point of the former structures.

3.2.1.4 The Forward-Backward Structure

The structure is depicted in Figure 3.10.

The system outputs are $y_2 = x_2 - w_{21}x_1$ and $y_1 = x_1 - w_{12}y_2$, which can also be expressed as

$$y = \begin{bmatrix} 1 + w_{12}w_{21} & -w_{12} \\ -w_{21} & 1 \end{bmatrix} x, \tag{3.6}$$
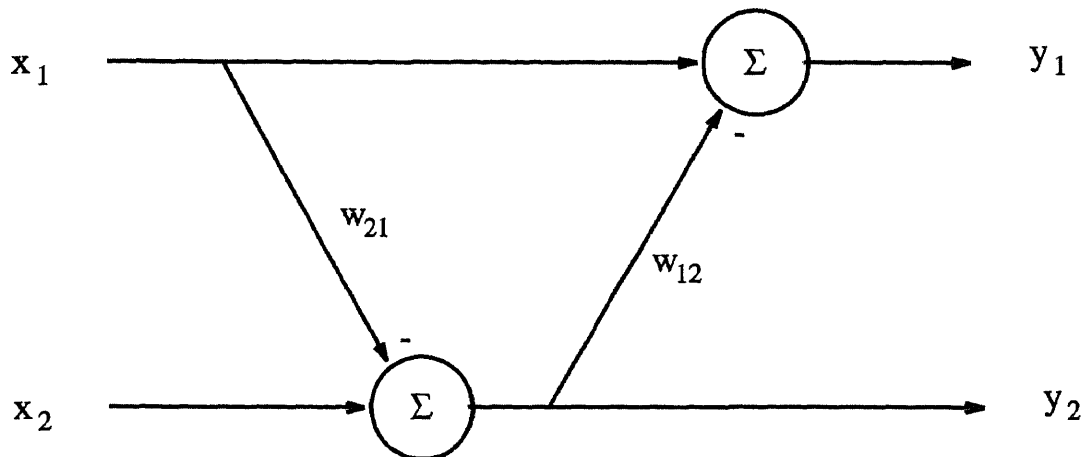
Figure 3.10 The Forward-Backward Structure

and the two separating points for (w_{12}, w_{21}) are $(\frac{1}{b_{22}/b_{12}-b_{12}/b_{11}}, b_{21}/b_{11})$ and $(\frac{1}{b_{12}/b_{11}-b_{22}/b_{12}}, b_{22}/b_{12})$.

3.2.2 Convergence and Stability

It was proved [13] that the feedback structure has four stable equilibrium points. Only two of them (listed previously) are separating points. Furthermore, convergence of the system to any of the four stable points depends on the statistics of the input sources. The necessary condition for this structure to achieve the source separation task is

$$< S_1^4 >< S_2^4 >\leq 9 < S_1^2 >^2 < S_2^2 >^2 ,\qquad (3.7)$$

where $<>$ denotes the time average. Similar situations take place on the other structures mentioned previously. This is the drawback of this kind of neural network receiver.

Once the necessary condition for the convergence is satisfied, the convergence also depends on the value of the learning step μ. On the one hand, the larger the μ is, the faster the weights converge; on the other hand, the weights will start to

diverge once μ is greater than a certain value. This value depends on the structure and is hard to determine.

3.2.3 Simulations

3.2.3.1 The Model

The model proposed to implement neural networks in the synchronous CDMA receiver is depicted in Figure 3.11.
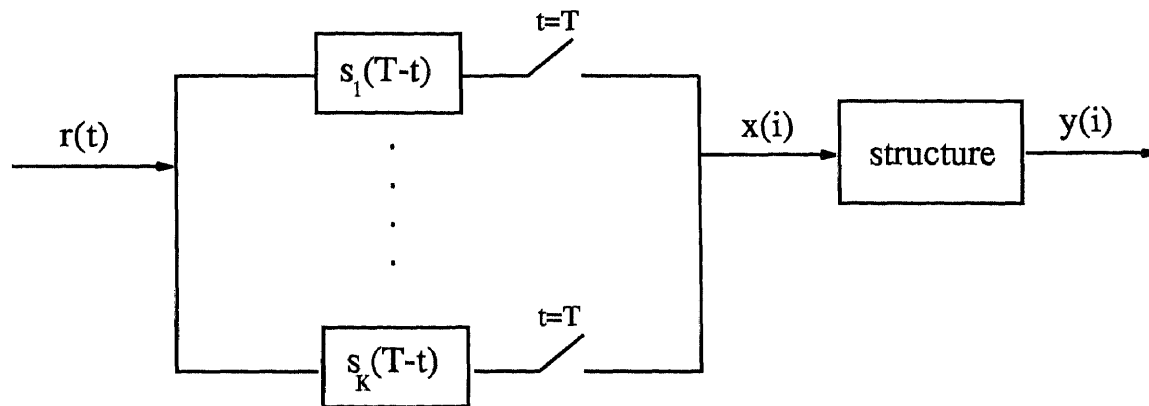
Figure 3.11 The Model Proposed

The structure can be one of the four types stated before. The difference is that the structure used here is for the K-user case.

Considering the two-user case, note that in digital communications, the necessary condition (3.7) for the feedback structure is satisfied. As for the other two structures, due to the difficulty in analyzing them, no conclusion is drawn.

3.2.3.2 Numerical Results

Simulations concerning the weight convergence and system performance, such as bit error probability, are made for the two-user case and three-user case.

Figure 3.12 shows the transient weight behavior of the backward structure receiver with both users' SNRs fixed at 8 dB and a high cross-correlation of $\rho = 0.7$. The learning step, which is a very sensitive parameter in this kind of receiver, is chosen to be 0.00002.
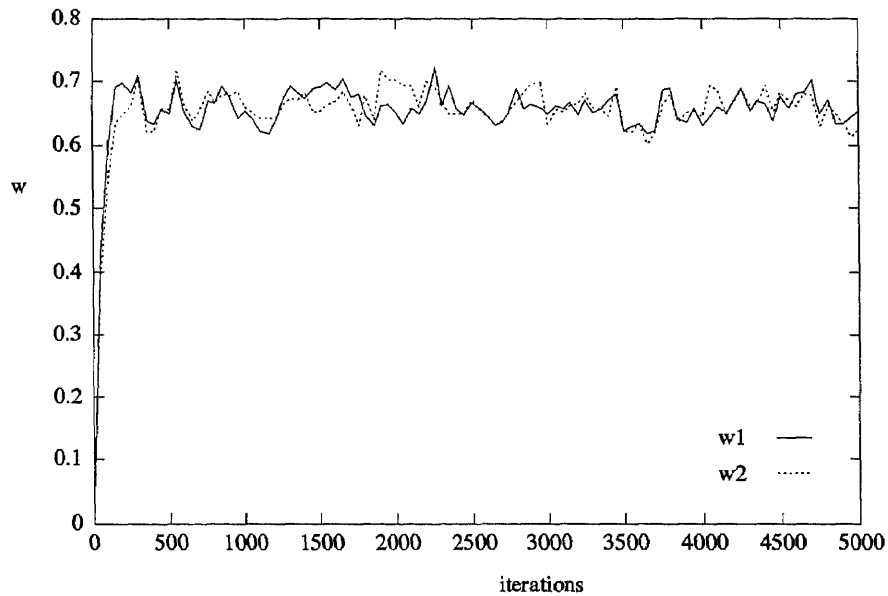

Figure 3.12 The Transient Weights of the Backward Structure Receiver

In Figures 3.13 through 3.15, the transient weight behavior of the other three structures are shown under the same conditions.
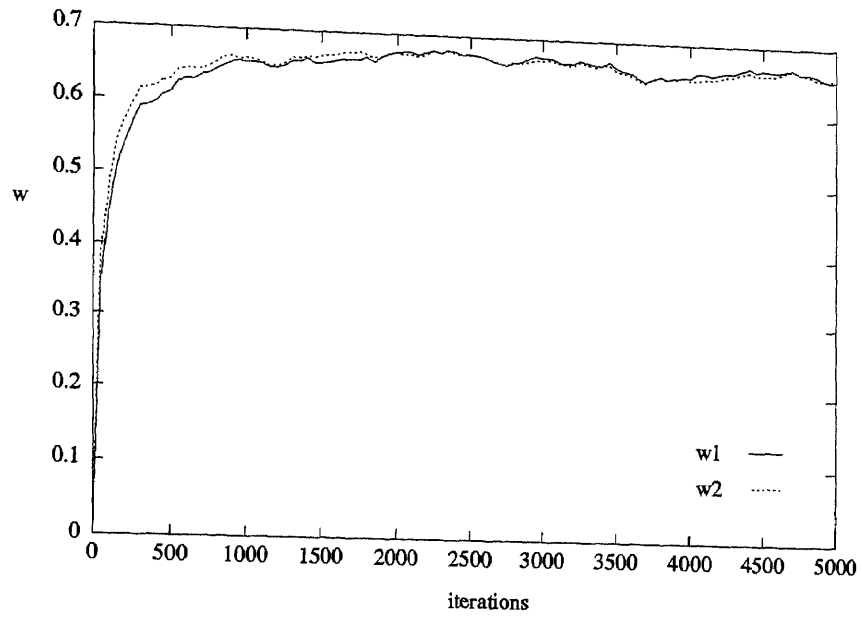
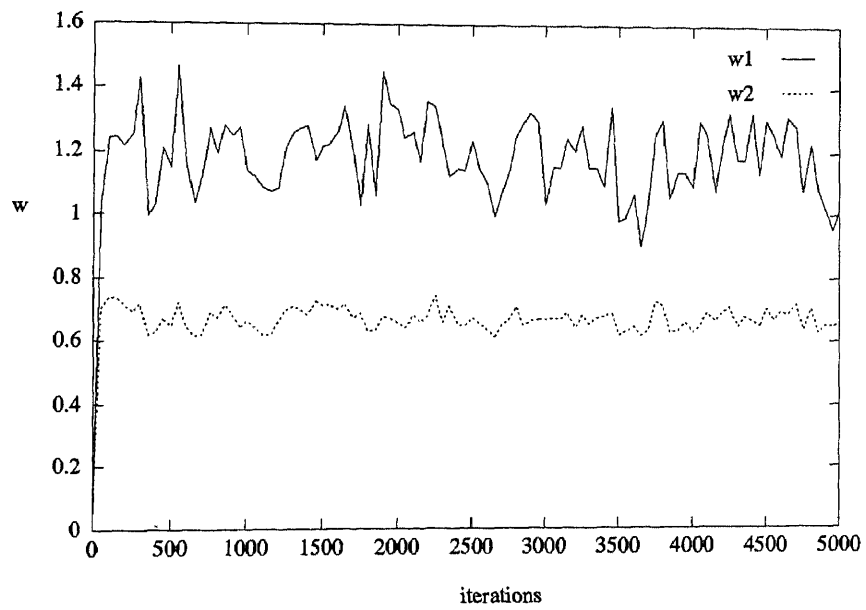**Figure 3.13** The Transient Weights of the Forward Structure Receiver

Figure 3.14 The Transient Weights of the Forward-Backward Structure Receiver
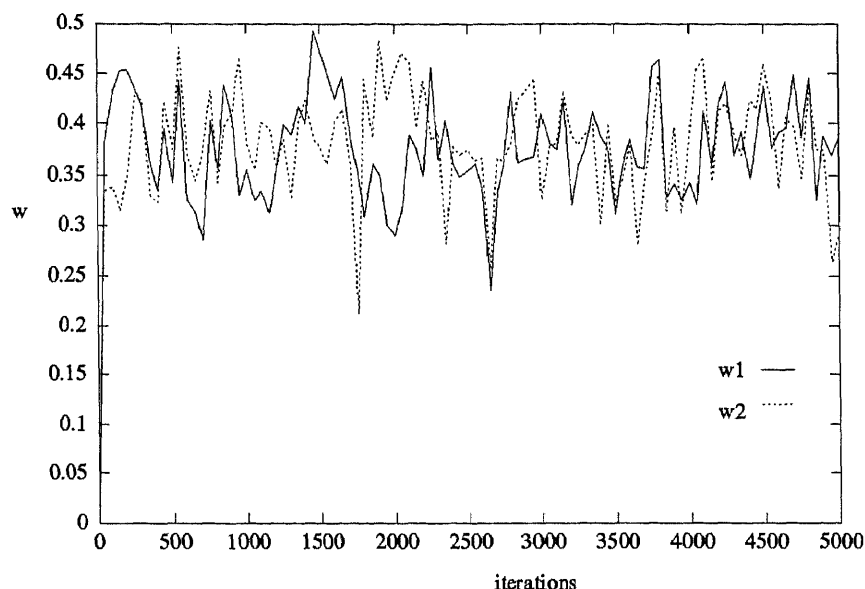
Figure 3.15 The Transient Weights of the Two-Stage Mixed Structure Receiver

It can be seen that under the same conditions, the convergence speed of the weights of the forward structure is the slowest among all structures. The weights of the two-stage mixed structure converge a little faster than those of the forward-backward structure, although they all converge faster than the weights of the other two structures. The figures of the bit error probability of the above four structures under the same conditions are also shown in Figures 3.16 through 3.19. It is clear that their performance with regard to error probability is similar, especially for the forward-backward structure and the two-stage mixed structure. Comparing these two structures, some conclusions can be drawn: these two structures are almost equivalent concerning their performance with regard to error probability, and the weights of the mixed structure converge a little faster than those of the forward-backward structure at the cost of double structural complexity.
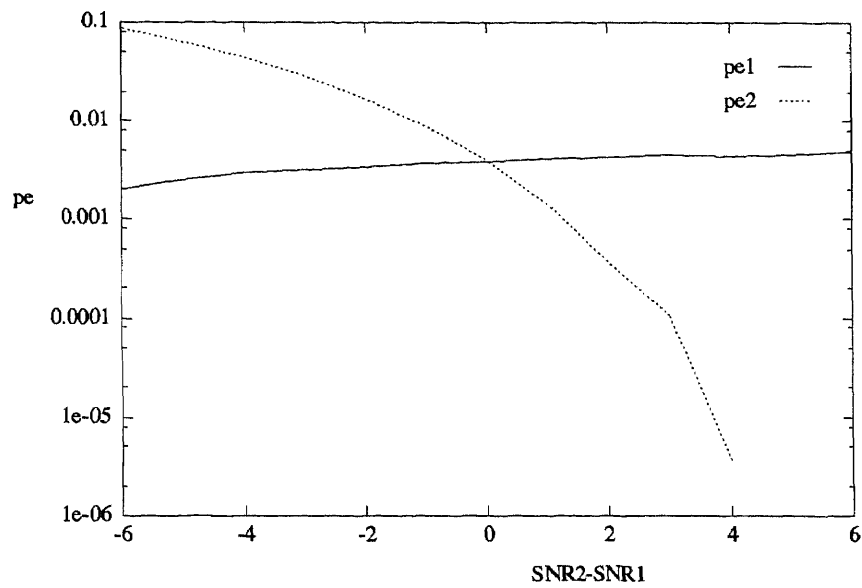
Figure 3.16 The Probability of Error in the Forward-Backward Structure Case
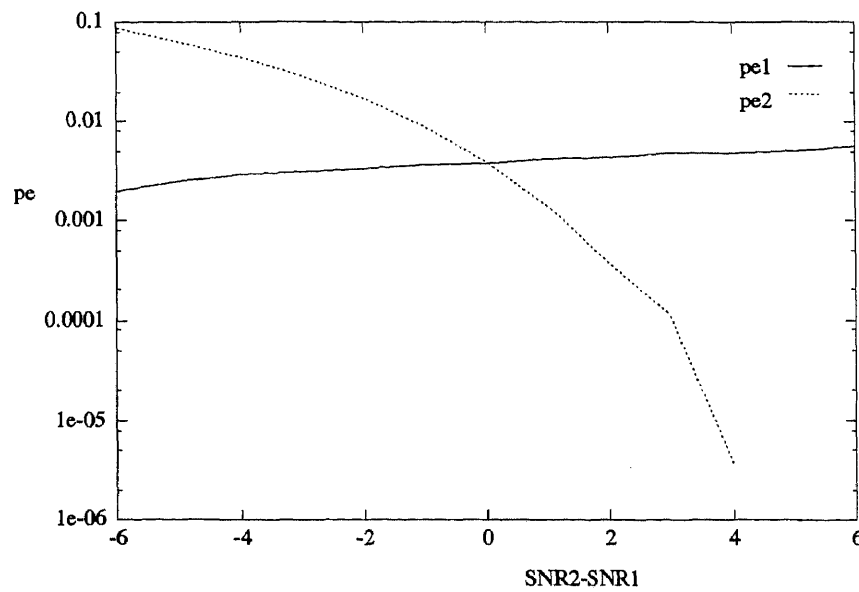
Figure 3.17 The Probability of Error in the Two-Stage Mixed Structure Case
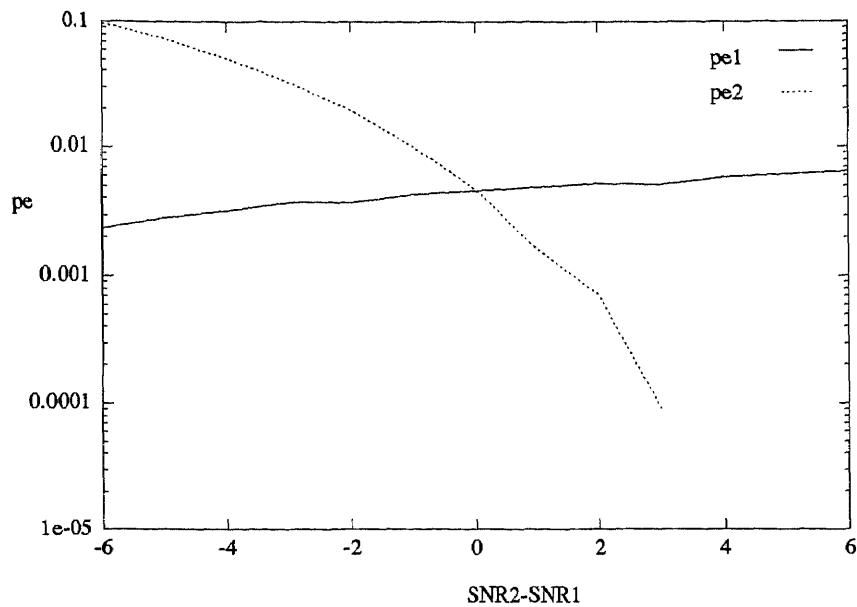
Figure 3.18 The Probability of Error in the Backward Structure Case
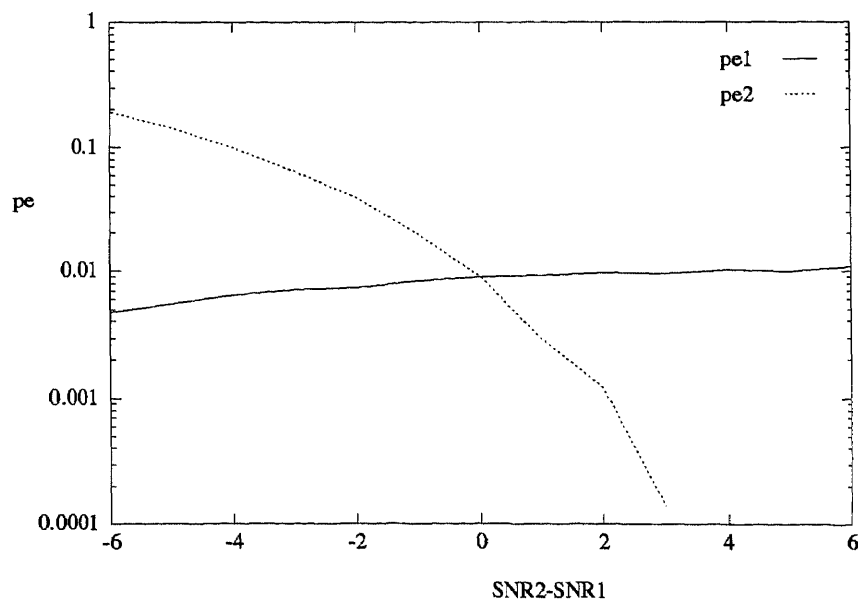
Figure 3.19 The Probability of Error in the Forward Structure Case

Simulations for the three-user case using the forward structure are also done with Gold sequences. Figure 3.20 shows the transient behavior of the weights to user 1 under the conditions that all the signal SNRs are set to 8 dB, and $\mu = 2 \times 10^{-6}$. The steady state error probability is depicted in Figure 3.21 under the same conditions, with SNR_1 fixed at 8 dB, SNR_2 and SNR_3 ranging from -6 to 6 dB.
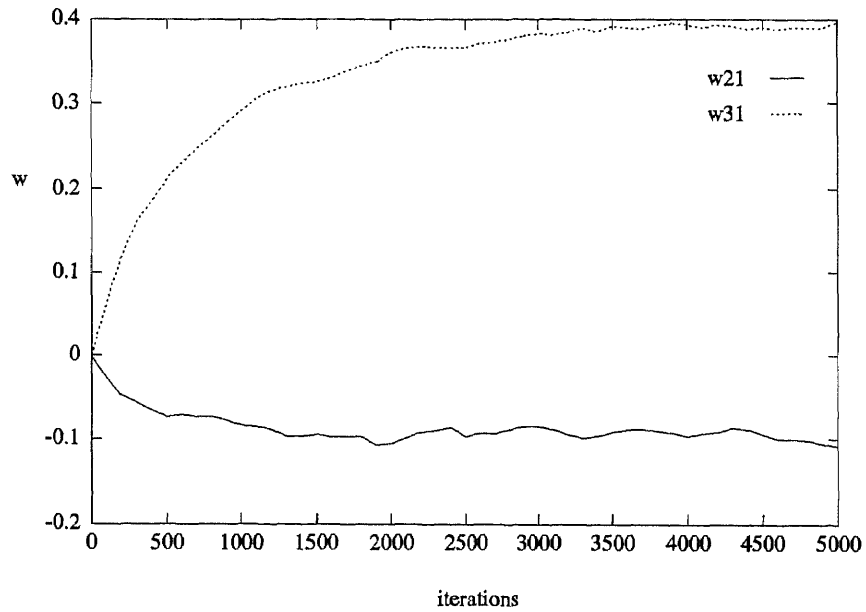
Figure 3.20 The Transient Weights to User 1 in the Three-User Case
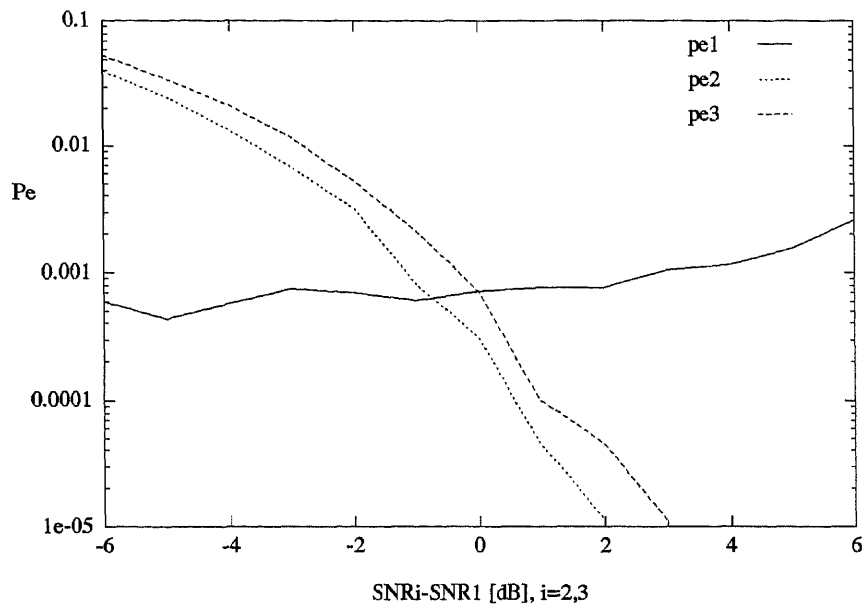
Figure 3.21 The Error Probability in the Three-User Forward Structure Case

CHAPTER 4

CONVERGENCE AND STABILITY ANALYSIS OF A RECENTLY PROPOSED SYNCHRONOUS ADAPTIVE CDMA RECEIVER

4.1 The Synchronous Adaptive CDMA Receiver

4.1.1 The Receiver

An adaptive, synchronous CDMA receiver proposed in [14] is depicted in Figure 4.1.

It consists of a bank of filters (matched to the users' signature sequences), which

comprises the front-end, followed by samplers and the decision system.
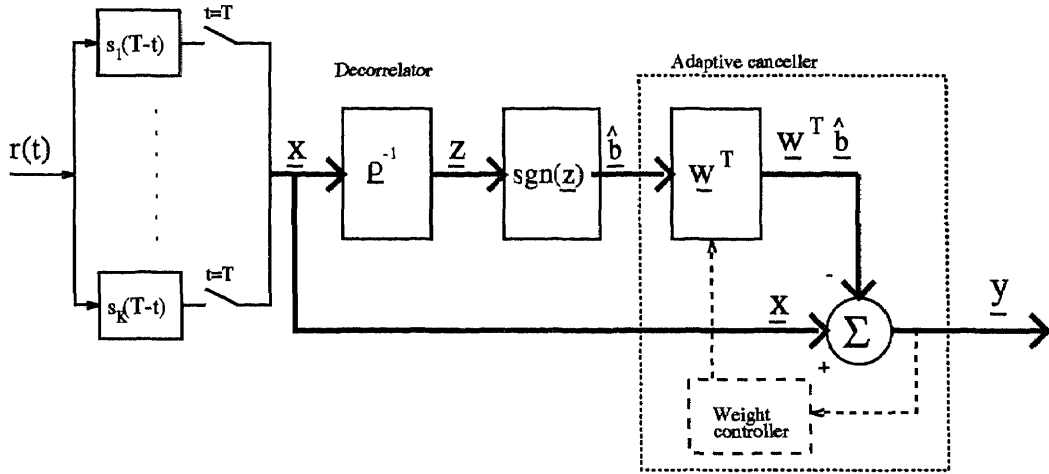
Figure 4.1 The Synchronous Adaptive CDMA Receiver

As shown in equations (2.1) and (2.3), the received signal $r(t)$ is expressed as:

$$r(t) = \sum_{k=1}^{K} \sum_{i} b_k(i)\sqrt{A_k}s_k(t - iT_0) + n(t),$$

where the received energy A_k is unknown to the receiver, the signature sequence

$s_k(t)$ is known to the receiver. The sampled outputs of the bank of matched filters

in the ith bit interval is expressed as:

34

$$x(i) = \rho A b(i) + n(i).$$

For convenience, the index i will be omitted whenever possible in the text.

The output vector of the decorrelator is:

$$z = \rho^{-1} x = A b + \rho^{-1} n.$$

The estimation is:

$$\hat{b} = \text{sgn}(z).$$

The canceler's output is given by:

$$y = x - W^T \hat{b}, \tag{4.1}$$

where

$$W = \begin{bmatrix} 0 & w_{12} & \dots & w_{1K} \\ w_{21} & 0 & \dots & w_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ w_{K1} & w_{K2} & \dots & 0 \end{bmatrix}.$$

The output for the kth user can be expressed as:

$$y_k = x_k - w_k{}^T \hat{b}_k, \tag{4.2}$$

where w_k is the kth column vector of W with the element w_{kk} deleted, i.e., $w_k = [w_{k1}, \cdots, w_{k(k-1)}, w_{k(k+1)}, \cdots, w_{kK}]^T$; and $\hat{b}_k$ is the vector obtained from $\hat{b}$ by deleting the element $\hat{b}_k$, i.e., $\hat{b}_k = [\hat{b}_1, \cdots, \hat{b}_{k-1}, \hat{b}_{k+1}, \cdots, \hat{b}_K]^T$. The output vector of the whole receiver is then evaluated as $sgn(y)$.

4.1.2 Weight Updating

For controlling the weights, the cost function should be set. For the kth signal, the output y_k can be expressed as

$$
\begin{aligned}
y_k &= x_k - \boldsymbol{w}_k^T \hat{\boldsymbol{b}}_k \\
&= x_k - \sum_{i \neq k, i=1}^{K} w_{ik} \hat{b}_i \\
&= \sum_{i=1}^{K} \rho_{ki} \sqrt{A_i} b_i + n_k - \sum_{i \neq k, i=1}^{K} w_{ik} \hat{b}_i \\
&= \sqrt{A_k} b_k + n_k + \sum_{i=1, i \neq k}^{K} (\rho_{ki} \sqrt{A_i} b_i - w_{ik} \hat{b}_i),
\end{aligned}
\tag{4.3}
$$

where b_i, $i = 1, 2, \cdots, K$, are independent. The cost function in this case is $J_k = E\{y_k^2\}$. It is easy to see that if the weights are chosen properly, the cost function will reach its minimum. In the scenario in which the estimations are correct, i.e., $\hat{b}_i = b_i$, $\forall i$, the minimum cost function is

$$
J_{k_{min}} = A_k + N_0/2,
\tag{4.4}
$$

with the optimal weights $w_{ik} = \rho_{ki} \sqrt{A_i}$, $\forall i \neq k$.

The steepest descent algorithm is used to minimize the cost function. That is, for the kth output, the optimum weights are obtained by an iterative search:

$$
\begin{aligned}
\boldsymbol{w}_k(i+1) &= \boldsymbol{w}_k(i) - \frac{\mu}{2} \frac{\partial}{\partial \boldsymbol{w}_k} E\{y_k^2\} \\
&= \boldsymbol{w}_k(i) + \mu E\{y_k \hat{\boldsymbol{b}}_k\} \\
&= \boldsymbol{w}_k(i) + \mu E\{x_k \hat{\boldsymbol{b}}_k - \hat{\boldsymbol{b}}_k \hat{\boldsymbol{b}}_k^T \boldsymbol{w}_k(i)\}.
\end{aligned}
\tag{4.5}
$$

The steady state of the weight vector $\boldsymbol{w}_k$ is readily obtained from the above equation, i.e.,

$$
E\{x_k \hat{\boldsymbol{b}}_k - \hat{\boldsymbol{b}}_k \hat{\boldsymbol{b}}_k^T \boldsymbol{w}_k^0(i)\} = 0,
\tag{4.6}
$$

which leads to

$$
\boldsymbol{w}_k^0 = E\{x_k \hat{\boldsymbol{b}}_k\}[E\{\hat{\boldsymbol{b}}_k \hat{\boldsymbol{b}}_k^T\}]^{-1}.
\tag{4.7}
$$

4.2 Probability of Error

Since the numerical analysis of the probability of error for K-user case, where K is very large, is computationally intensive, most evaluations are made in two- to five-user cases. A detailed analysis of the error probability for the receiver is discussed in [14]. In this section, only figures on the probability of error, based on theoretical calculations [14], are shown. For the purpose of comparison, the performance of the decorrelating detector and the two-stage non-adaptive decorrelating detector are also shown.

The first example, depicted in Figure 4.2, is the probability of error in a simple two-user case, but it nevertheless provides some insight into the steady state behavior of the adaptive detector. The cross-correlation coefficient ρ assumed to be 0.7 (which means strong interference), can certainly be considered to represent a case of high bandwidth efficiency. The SNR_1 is set to 8 dB, while the SNR_2, relative to SNR_1, varies from -10 to 8 dB.
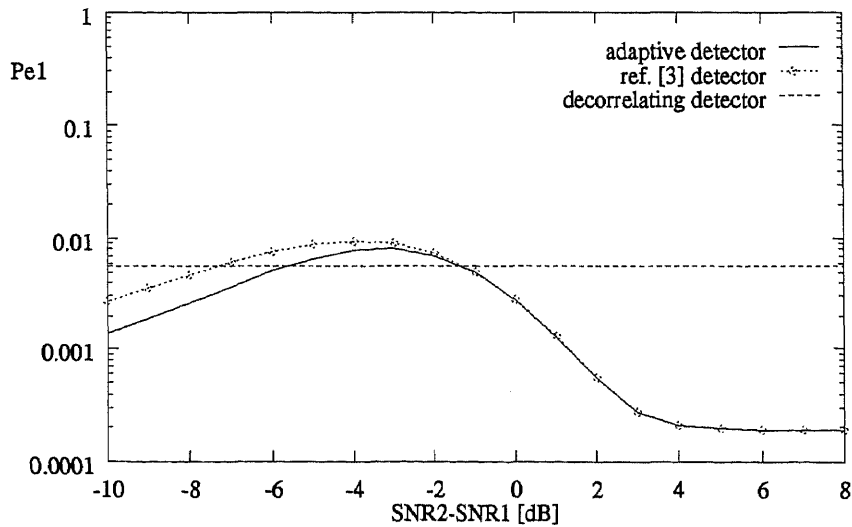
Figure 4.2 The Probability of Error in the Two-User Case with $\rho = 0.7$

Figures 4.3 through 4.5 show the performance of the two- to four-user receivers with Gold sequences. As expected, with K getting large, the decorrelating detector begins to exhibit its inadequacy. The adaptive and the fixed-weights scheme show virtually identical performance, with the former being only slightly better for weak interferers. With the number of simultaneous users increasing, certain trends become more obvious. Due to its unacceptable high error probability, the decorrelating detector clearly does not represent an appropriate choice.
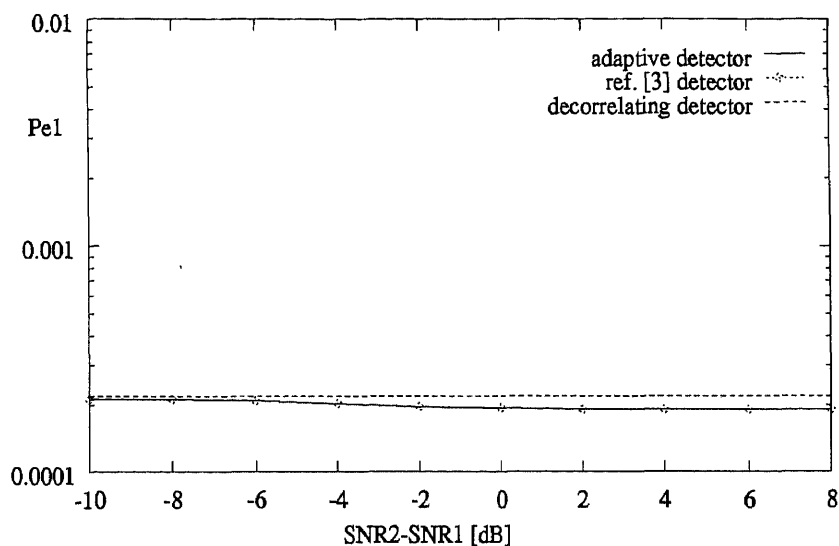
Figure 4.3 The Probability of Error of User 1 in the Two-User Case
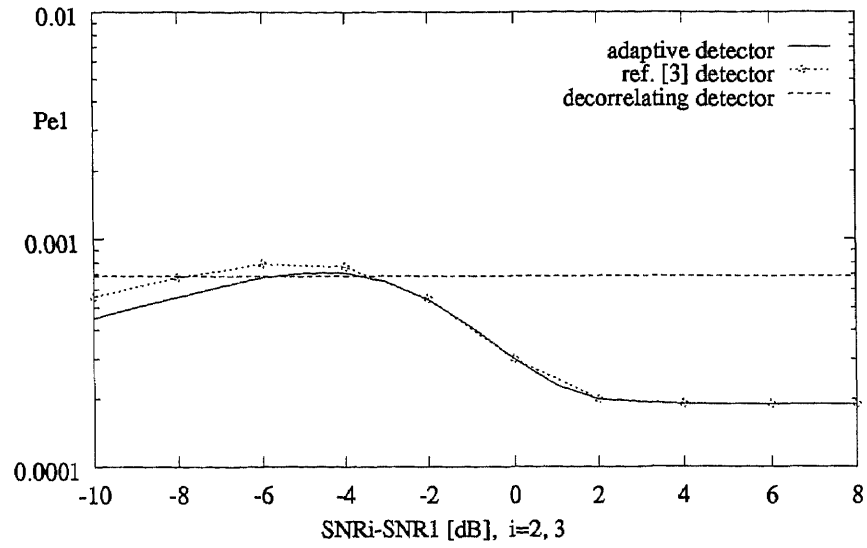
Figure 4.4 The Probability of Error of User 1 in the Three-User Case
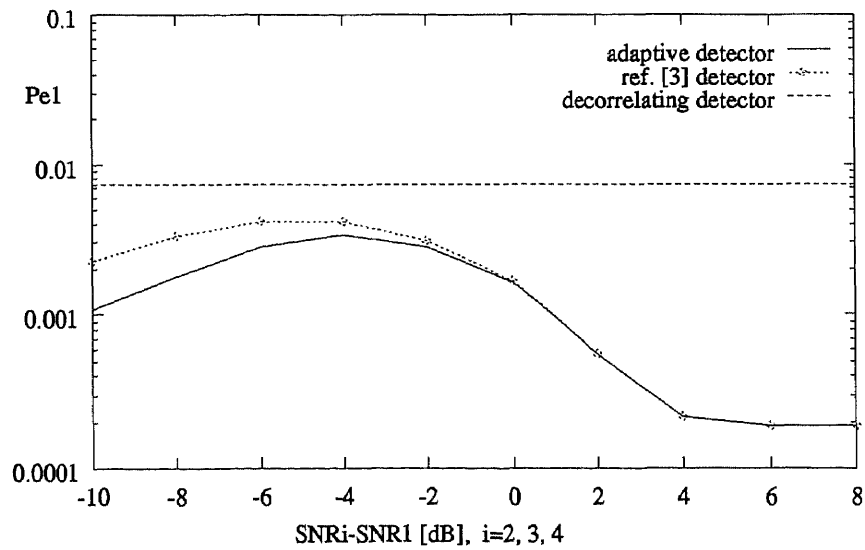
Figure 4.5 The Probability of Error of User 1 in the Four-User Case

Figure 4.6 shows the results from another angle. The error probability of user 1 is measured under two-, three-, and four-user cases with all the user $SNRs$ ranging from 2 to 12 dB. It can be seen that adding more users only cause graceful degradation of the system's performance.
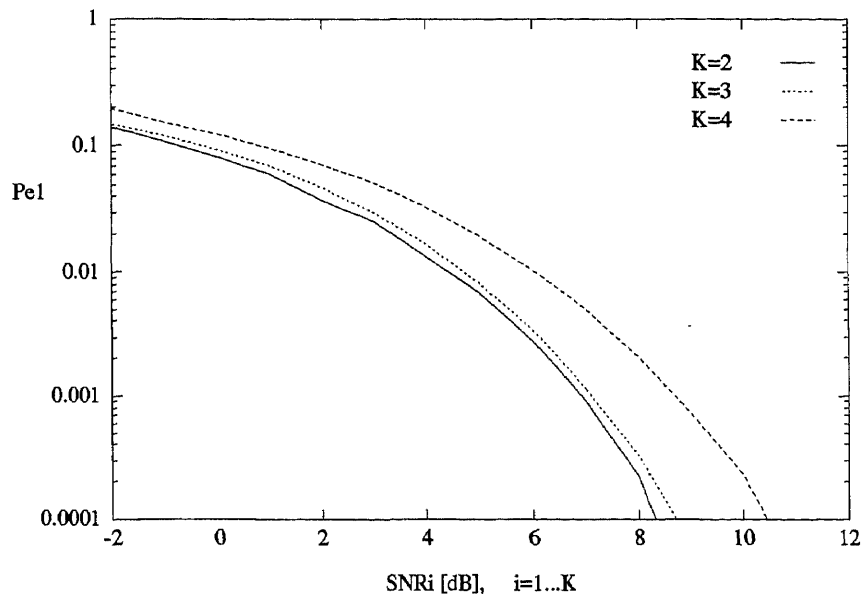
Figure 4.6 The Comparison of Error Probability for the Two- to Four-User Case

From all these figures, some conclusions can be drawn: the performance with regard to error probability of the adaptive receiver is better than the two-stage decorrelating receiver with fixed weights; when the number of users increases, the discrepancy between these two receivers' performance also increases.

4.3 Convergence and Stability Analysis

4.3.1 General Analysis

The updating rule already derived earlier (4.5) is

$$w_k(i+1) = w_k(i) + \mu E\{x_k \hat{b}_k - \hat{b}_k \hat{b}_k^T w_k(i)\}. \tag{4.8}$$

Since w_k is deterministic, which is decided by the expectation function in the weight adaptive, the above equation can also be expressed as

$$
\begin{aligned}
w_k(i+1) &= (I - \mu E\{\hat{b}_k \hat{b}_k^T\}) w_k(i) + \mu E\{x_k \hat{b}_k\} \\
&= (I - \mu E\{\hat{b}_k \hat{b}_k^T\})^{i+1} w_k(0) + \sum_{j=0}^{i} (I - \mu E\{\hat{b}_k \hat{b}_k^T\})^j \mu E\{x_k \hat{b}_k\}. \tag{4.9}
\end{aligned}
$$

For simplicity, the initial weight vector is assumed to be zero, i.e., $w_k(0) = 0$. Thus, equation (4.9) becomes

$$w_k(i+1) = \sum_{j=0}^{i} (I - \mu E\{\hat{b}_k \hat{b}_k^T\})^j \mu E\{x_k \hat{b}_k\}. \tag{4.10}$$

Since $\hat{b}_k \in \{-1, 1\}$, which implies $-1 < E\{\hat{b}_i \hat{b}_j\} < 1 \quad \forall i \neq j$, matrix $E\{\hat{b}_k \hat{b}_k^T\}$ is a $(K-1) \times (K-1)$ symmetric matrix with diagonal elements equal to 1 and non-diagonal elements ranging between $(-1, 1)$. It is the non-linear function $sgn(\cdot)$ that makes the analyzing complicated.

Define $H = I - \mu E\{\hat{b}_k \hat{b}_k^T\}$, a $(K-1) \times (K-1)$ symmetric matrix, equation (4.10) becomes

$$w_k(i+1) = (\sum_{j=0}^{i} H^j) \mu E\{x_k \hat{b}_k\}. \tag{4.11}$$

It is easily seen that the training of the weights converges if and only if

$$\lim_{i \to \infty} |w_k(i+1)| < \infty, \tag{4.12}$$

where $|\cdot|$ means the absolute value.

Equation (4.12) is equivalent to

$$\lim_{i \to \infty} \sum_{j=0}^{i} \boldsymbol{H}^j < \infty, \qquad (4.13)$$

for which one necessary condition is $\lim_{j \to \infty} \boldsymbol{H}^j = \boldsymbol{0}$. Take the determinant of it, $det(\lim_{j \to \infty} \boldsymbol{H}^j) = \lim_{j \to \infty} \{det(\boldsymbol{H})\}^j = 0. \Rightarrow det(\boldsymbol{H}) < 1.$ Therefore, a necessary condition for convergence is $det(\boldsymbol{H}) < 1$.

Since $\boldsymbol{H}$ is a symmetric matrix, it can be diagonalized and all of its $(K-1)$ eigenvalues are real. So, there always exists an orthogonal matrix $\boldsymbol{Q}$ such that

$$\boldsymbol{H} = \boldsymbol{Q}\boldsymbol{D}\boldsymbol{Q}^{-1}$$

where $\boldsymbol{D}$ is a diagonal matrix $\boldsymbol{D} = \text{diag } [\lambda_1, \lambda_2, \cdots, \lambda_{K-1}]$, and $\lambda_i, i = 1, \cdots, K-1$ are the $(K-1)$ eigenvalues of matrix $\boldsymbol{H}$. Thus,

$$
\begin{aligned}
\boldsymbol{w}_k(i+1) &= \sum_{j=0}^{i}(\boldsymbol{Q}\boldsymbol{D}^j\boldsymbol{Q}^{-1})\mu E\{x_k\hat{\boldsymbol{b}}_k\} \\
&= \boldsymbol{Q}(\sum_{j=0}^{i}\boldsymbol{D}^j)\boldsymbol{Q}^{-1}\mu E\{x_k\hat{\boldsymbol{b}}_k\} \\
&= \boldsymbol{Q}\begin{pmatrix} \frac{1-\lambda_1^{i+1}}{1-\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1-\lambda_{K-1}^{i+1}}{1-\lambda_{K-1}} \end{pmatrix}\boldsymbol{Q}^{-1}\mu E\{x_k\hat{\boldsymbol{b}}_k\}.
\end{aligned}
\qquad (4.14)
$$

It is clear that the necessary and sufficient condition for the weights to achieve convergence and stability is

$$| \lambda_i | < 1, \quad \forall i. \qquad (4.15)$$

The following Gershgorin theorem [17] of characteristic roots is used to estimate the eigenvalues of a matrix:

> If $\boldsymbol{M}$ is a square matrix of order n with its element denoted as m_{ij}, every characteristic root of the complex matrix $\boldsymbol{M}$ lies in at least one of the n disks with centres m_{ii}, radii $r_i = \sum_{j=1, j\neq i}^{n} |m_{ij}|$. A sometimes sharper set of bounds can be obtained by applying this theorem to $\boldsymbol{M}$ and $\boldsymbol{M}^T$ simultaneously.

From the theorem, the following holds for matrix H:

$$|\lambda_i - (1 - \mu)| \leq r_i \quad \forall i, \tag{4.16}$$

where r_i is the sum of the absolute value of each element of the ith column vector of matrix H with the diagonal element deleted. Equation (4.16) leads to

$$(1 - \mu) - r_i \leq \lambda_i \leq (1 - \mu) + r_i. \tag{4.17}$$

Assume $e_{max} = max_{i,j}|e_{ij}|$, where e_{ij} is the (i, j)th element of matrix $E\{\hat{b}_k \hat{b}_k^T\}$, which depends on the signal SNRs and cross-correlation matrix ρ, then, $r_i \leq (K - 2)e_{max}$. Combine equations (4.15) and (4.17), the conditions on λ_i ($\forall i$) can be implemented as

$$-1 < 1 - \mu - r_i \leq \lambda_i \leq 1 - \mu + r_i < 1. \tag{4.18}$$

which leads to the following two results.

1.

$$(K - 2)e_{max} \leq 1.$$

This put an balance among three elements of the system: the number of users that can access the system simultaneously, the signal SNRs, and the cross-correlation of their signature sequences.

2.

$$\mu \leq \frac{2}{1 + (K - 2)e_{max}}.$$

This is the condition on learning step μ for the system to achieve convergence and stability. When condition 1 satisfied, this condition implies that μ can be any number between $(0, 1)$.

These two conditions are concerned to be sufficient for the system to converge. The steady state of the weight vector $\boldsymbol{w}_k$ is

$$
\begin{aligned}
\boldsymbol{w}_k^0 &= (\boldsymbol{I} - \boldsymbol{H})^{-1}\mu E\{x_k \hat{\boldsymbol{b}}_k\} \\
&= E\{x_k \hat{\boldsymbol{b}}_k\}[E\{\hat{\boldsymbol{b}}_k \hat{\boldsymbol{b}}_k^T\}]^{-1},
\end{aligned} \tag{4.19}
$$

which is the same as (4.7). It is easy to verify that when the above two conditions satisfied, matrix $E\{\hat{\boldsymbol{b}}_k \hat{\boldsymbol{b}}_k^T\}$ is a non-singular matrix, and its eigenvalues are between (0,2).

Consider the case when the initial weight vector is not set to $\boldsymbol{0}$. According to equation (4.9),

$$
\boldsymbol{w}_k(i+1) = \boldsymbol{H}^{i+1}\boldsymbol{w}_k(0) + \sum_{j=0}^{i} \boldsymbol{H}^j \mu E\{x_k \hat{\boldsymbol{b}}_k\},
$$

it is easy to verify that when the two conditions for convergence are satisfied, $\lim_{i \to \infty} \boldsymbol{H}^{i+1}\boldsymbol{w}_k(0) = \boldsymbol{0}$. Therefore, $\boldsymbol{w}_k$ will converge to the same steady state no matter what its initial value is.

Although the above analysis considers the kth user only, the results can be applied to all the users. Therefore, these two conditions are sufficient for the receiver to achieve convergence.

4.3.2 Transient Behavior Analysis

From equation (4.14), the transient behaviors of the weights is decided by two elements: learning step μ and eigenvalues of matrix $\boldsymbol{H}$. Since the eigenvalues of $\boldsymbol{H}$ is almost untractable in K user case, therefore, no analysis on the effect of the eigenvalue spread is given in this thesis. But when SNR, defined as A_k/N_0 for user k, is high enough, it leads to $E\{b_i b_j\} \approx 0, \forall i \neq j$ and, thus, (4.10) becomes:

$$
\boldsymbol{w}_k(i+1) \approx [1 - (1-\mu)^{i+1}]E\{x_k \hat{\boldsymbol{b}}_k\}, \tag{4.20}
$$

which means that the weight vector is almost free of the effect of the eigenvalue spread, and it changes almost monotonically to the steady state value. Moreover, the larger the μ, the faster the weights converge.

For illustrative purposes, further considerations on the two-user case and three-user case will be given in more detail. The simulations of the updating weights as well as their transient behaviors are performed for those cases.

4.3.3 The Two-User Case

The synchronous CDMA receiver for the two-user case is depicted in Figure 4.7.
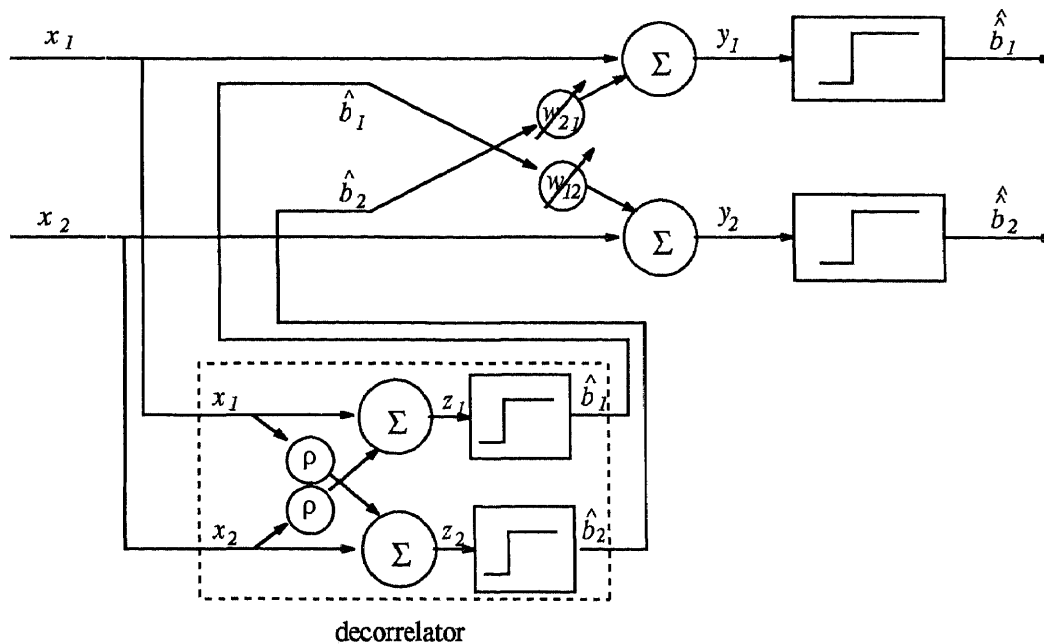

Figure 4.7 The Two-User Synchronous CDMA Receiver

In this case, the outputs are

$$y_1 = x_1 - w_{21}\hat{b}_2$$

and

$$y_2 = x_2 - w_{12}\hat{b}_1.$$

The following weight updating rule is readily obtained from (4.29):

$$w_{21}(i+1) = (1-\mu)w_{21}(i) + \mu E[x_1 \hat{b}_2] \tag{4.21}$$

$$w_{12}(i+1) = (1-\mu)w_{12}(i) + \mu E[x_2 \hat{b}_1]. \tag{4.22}$$

With $\boldsymbol{w}$ initially set to $\boldsymbol{0}$, w_{21} at the ith iteration can be expressed as

$$
\begin{aligned}
w_{21}(i) &= (1-\mu)^i w_{21}(0) + \sum_{j=0}^{i-1}(1-\mu)^j \mu E[x_1 \hat{b}_2] \\
&= [1 - (1-\mu)^i] E[x_1 \hat{b}_2].
\end{aligned}
\tag{4.23}
$$

Similarly,

$$w_{12}(i) = [1 - (1-\mu)^i] E[x_2 \hat{b}_1].$$

H in this case is not a matrix, but a value equal to $(1-\mu)$. In this case, the transient behavior of the weight depends on the learning step μ only.

From the above, the following properties agree with our earlier analysis:

1.

$$\lim_{i \to \infty} w_{21} = E[x_1 \hat{b}_2],$$

and

$$\lim_{i \to \infty} w_{12} = E[x_2 \hat{b}_1],$$

which match equation (4.7).

2. μ can be chosen between $(0,1)$, and convergence is guaranteed provided that $0 < \mu < 1$. The larger μ is, the faster the convergence.

3. w is monotonically increasing with i. If we denote w_{21}^0 as the steady state value of w_{21}, then the convergence speed can be easily established.

Since

$$w_{21}(i) = [1 - (1 - \mu)^i]w_{21}^0,$$

define a parameter δ to measure the distance in percentage between $w_{21}(i)$ and its steady state value w_{21}^0, i.e., $\frac{w_{21}^0 - w_{21}(i)}{w_{21}^0}$. Then, for a given δ, the minimum number of iterations, N, such that

$$1 - \frac{w_{21}(N)}{w_{21}^0} \leq \delta,$$

can be obtained as follows:

$$(1 - \mu)^N \leq \delta$$

$$\Rightarrow N = \lceil \frac{\lg \delta}{\lg(1 - \mu)} \rceil, \tag{4.24}$$

where $\lceil x \rceil$ is the ceiling of x.

Similar analysis can be drawn for w_{12}.

4.3.4   The Three-User Case

The receiver for the K-user case is shown in Figure 4.8. Here, consider the three-user case, i.e., $K = 3$.

Without loss of generality, consider the weight vector to user 1 only, which is $\boldsymbol{w}_1 = [w_{21}, w_{31}]^T$. The output of user 1 is $y_1 = x_1 - w_{21}\hat{b}_2 - w_{31}\hat{b}_3$. The weights are readily obtained from equation (4.29):

$$w_{21}(i + 1) = (1 - \mu)w_{21}(i) - \mu E[\hat{b}_2\hat{b}_3]w_{31}(i) + \mu E[x_1\hat{b}_2]$$

$$w_{31}(i + 1) = (1 - \mu)w_{31}(i) - \mu E[\hat{b}_2\hat{b}_3]w_{21}(i) + \mu E[x_1\hat{b}_3],$$

which in matrix form is

$$\begin{pmatrix} w_{21}(i+1) \\ w_{31}(i+1) \end{pmatrix} = \begin{pmatrix} 1 - \mu & -\mu E[\hat{b}_2\hat{b}_3] \\ -\mu E[\hat{b}_2\hat{b}_3] & 1 - \mu \end{pmatrix} \begin{pmatrix} w_{21}(i) \\ w_{31}(i) \end{pmatrix} + \mu \begin{pmatrix} E[x_1\hat{b}_2] \\ E[x_1\hat{b}_3] \end{pmatrix}. \tag{4.25}$$
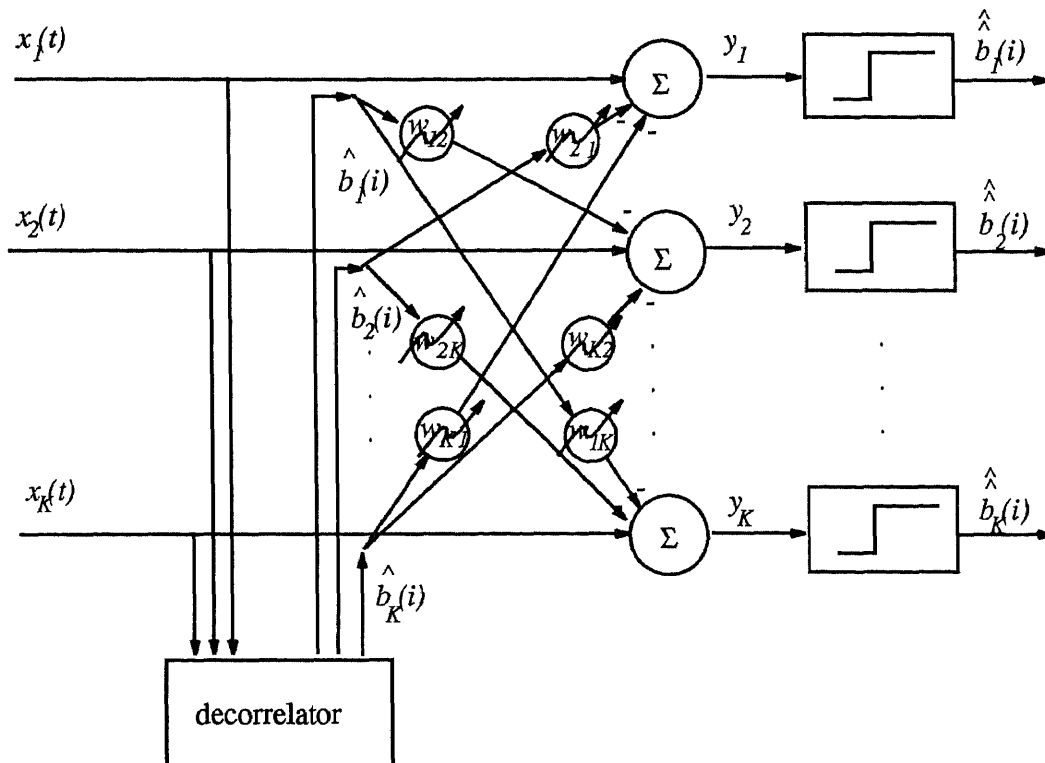
Figure 4.8 The K-User Synchronous CDMA Receiver

In this case, matrix H becomes

$$H = \begin{pmatrix} 1 - \mu & -\mu E\{\hat{b}_2\hat{b}_3\} \\ -\mu E\{\hat{b}_2\hat{b}_3\} & 1 - \mu \end{pmatrix} = QDQ^{-1}, \qquad (4.26)$$

where

$$Q = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

The eigenvalues are

$$\lambda_1 = (1 - \mu) - \mu E\{\hat{b}_2\hat{b}_3\}$$

and

$$\lambda_2 = (1 - \mu) + \mu E\{\hat{b}_2\hat{b}_3\}.$$

Finally, equation (4.25) leads to

$$
\begin{pmatrix} w_{21}(i+1) \\ w_{31}(i+1) \end{pmatrix} = Q \begin{pmatrix} \sum_{j=0}^{i}\lambda_1^j & 0 \\ 0 & \sum_{j=0}^{i}\lambda_2^j \end{pmatrix} Q^{-1}\mu \begin{pmatrix} E\{x_1\hat{b}_2\} \\ E\{x_1\hat{b}_3\} \end{pmatrix}
$$

$$
= \frac{1}{2}\mu \begin{pmatrix} \sum_{j=0}^{i}(\lambda_1^j + \lambda_2^j) & \sum_{j=0}^{i}(\lambda_1^j - \lambda_2^j) \\ \sum_{j=0}^{i}(\lambda_1^j - \lambda_2^j) & \sum_{j=0}^{i}(\lambda_1^j + \lambda_2^j) \end{pmatrix} \begin{pmatrix} E\{x_1\hat{b}_2\} \\ E\{x_1\hat{b}_3\} \end{pmatrix}. \quad (4.27)
$$

From the above, the following properties hold:

1. when $i \to \infty$, the steady state values for w_{21} and w_{31} are w_{21}^0 and w_{31}^0, then

$$
\begin{pmatrix} w_{21}^0 \\ w_{31}^0 \end{pmatrix} = \frac{\begin{pmatrix} 1 & -E[\hat{b}_2\hat{b}_3] \\ -E[\hat{b}_2\hat{b}_3] & 1 \end{pmatrix} \begin{pmatrix} E[x_1\hat{b}_2] \\ E[x_1\hat{b}_3] \end{pmatrix}}{1 - E^2[\hat{b}_2\hat{b}_3]}.
$$

2. for the weights to achieve convergence, λ_i must satisfy $|\lambda_1| \leq 1$ and $|\lambda_2| \leq 1$, which leads to

$$
0 < \mu < \frac{2}{1 + \left|E\{\hat{b}_2\hat{b}_3\}\right|}, \quad (4.28)
$$

i.e., μ can be chosen between $(0,1)$.

3. when the signal $SNRs$ are large enough, $\hat{b}_2 \simeq b_2, \hat{b}_3 \simeq b_3$. so, $E[\hat{b}_2\hat{b}_3] \simeq 0$. $\lambda_1 = \lambda_2 \approx 1 - \mu$.

Then,

$$
\begin{pmatrix} w_{21}(i+1) \\ w_{31}(i+1) \end{pmatrix} \approx \begin{pmatrix} 1 - (1-\mu)^{i+1} & 0 \\ 0 & 1 - (1-\mu)^{i+1} \end{pmatrix} \begin{pmatrix} E[x_1\hat{b}_2] \\ E[x_1\hat{b}_3] \end{pmatrix},
$$

which means that the weights are mostly depend on μ value, and changes almost monotonically with i.

4.3.5 Simulation Results

In practice, the expectation operation is usually implemented by time average. In this case, equation (4.9) becomes

$$w_k(i+1) = (I - \mu < \hat{b}_k \hat{b}_k^T >)w_k(i) + \mu < x_k \hat{b}_k > . \qquad (4.29)$$

where $< \cdot >$ denotes time average. Equation (4.29) revealed that in real implementation, w_k is a random process, which partially accounts for the undulation on weights while updating (the effect of eigenvalue spread may also cause undulation). The subject is similar to the stochastic approximation widely studied in [18]. The only difference is that the learning step in stochastic approximation is a function of time index i.

4.3.5.1 The Two-user Case

In this section, simulations of weight updating in the two-user case is given.

Figure 4.9 depicts the weight updating process under the conditions that $\rho = 0.7$, $SNR_1 = SNR_2 = 8$ dB with learning step $\mu = 0.9$ and 0.002. The window size used for time average here is 2000 steps. It is clear that under the same conditions, the weight updated with the large μ value converges faster than the one with the small μ value. The undulation exists mostly because of the time average used in implementation. Since the weights are random variables, their statistical behaviors in terms of mean values and standard deviations are examined. The statistic behavior of the weights can be seen in Figure 4.10 and Figure 4.11. The results are achieved under the conditions that SNR_1 fixed as 8 dB, and SNR_2 varies as $2, 8, 14$ dB. μ in this case is chosen as 0.002.
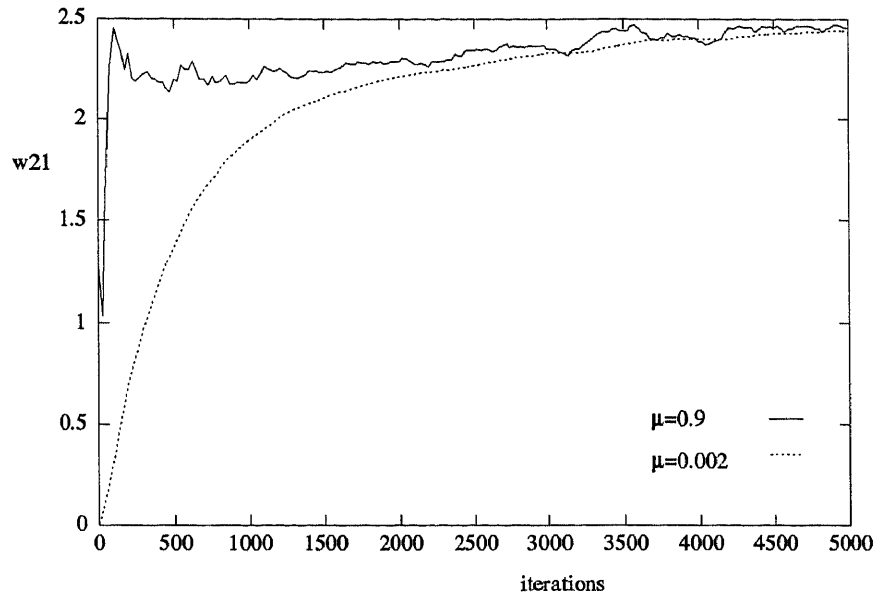
Figure 4.9 The Weight w_{21} with $SNR_1 = SNR_2 = 8$ dB, $\mu = 0.9$ and $\mu = 0.002$
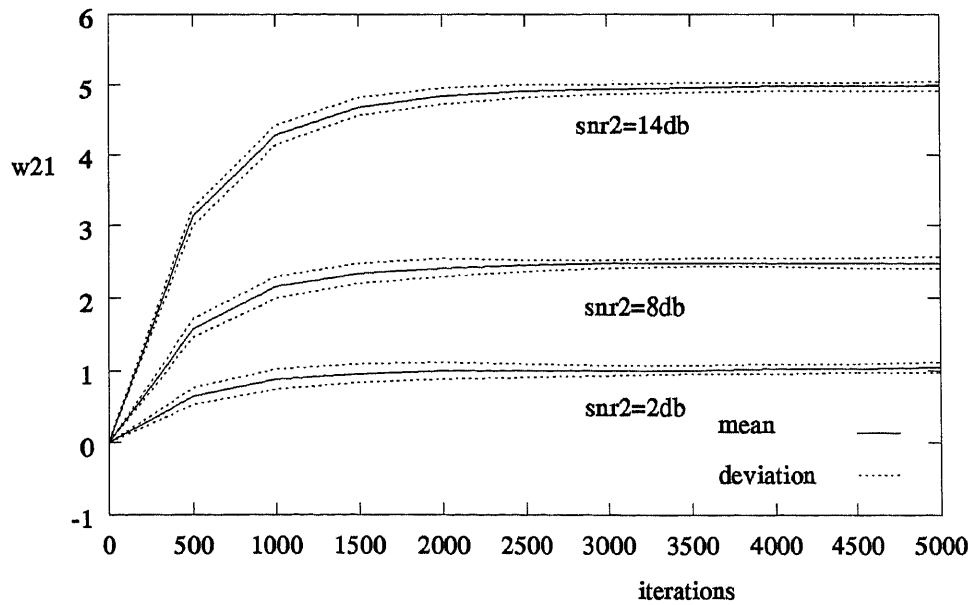

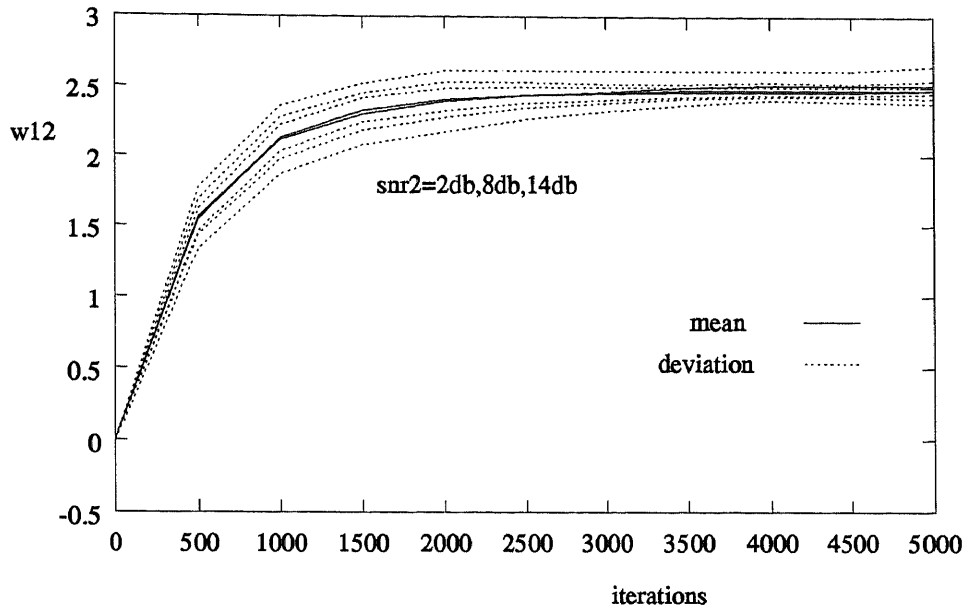**Figure 4.10** The Statistical Behavior of w_{21} with $SNR_1 = 8$ dB, $\mu = 0.002$

Figure 4.11 The Statistical Behavior of w_{12} with $SNR_1 = 8\ dB$, $\mu = 0.002$

Figure 4.12 and Figure 4.13 show the weight behavior with $\mu = 0.2$ under the same user SNRs as in the last case. It can be seen that in both cases, the weights converge to the same values which are their steady state values. Also, they converge with different speed because of different learning steps μ. The number of iterations needed for the weights to converge to their steady states can be calculated using equation (4.24). Assume $\delta = 0.1$, then

$$N = \begin{cases} 1150 & for & \mu = 0.002 \\ 10 & for & \mu = 0.2 \end{cases} \tag{4.30}$$

Assume $\delta = 0.01$, then

$$N = \begin{cases} 2230 & for & \mu = 0.002 \\ 20 & for & \mu = 0.2 \end{cases} \tag{4.31}$$

The results are the same for w_{21} and w_{12}. They agree well with the simulations.
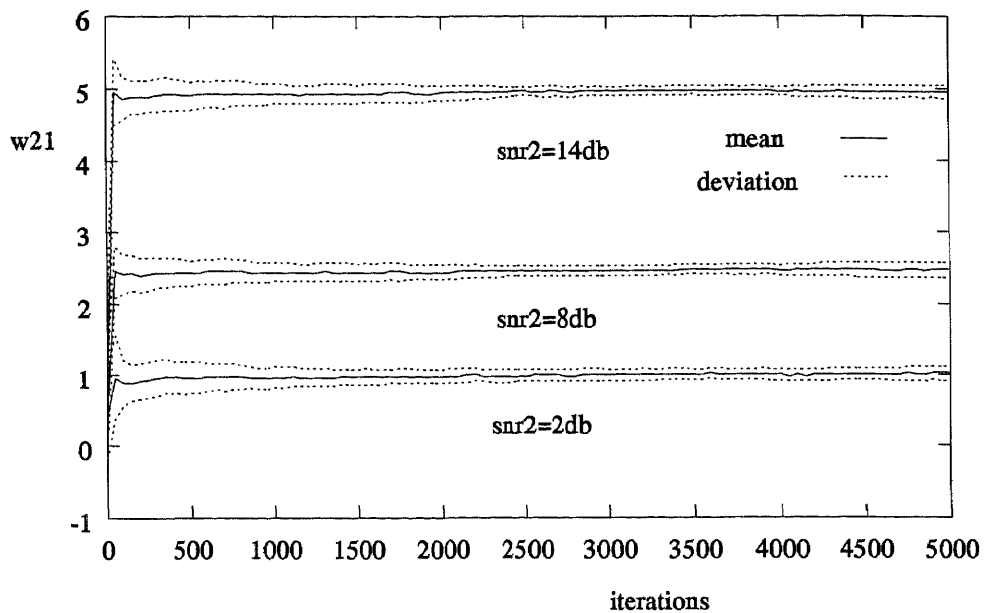
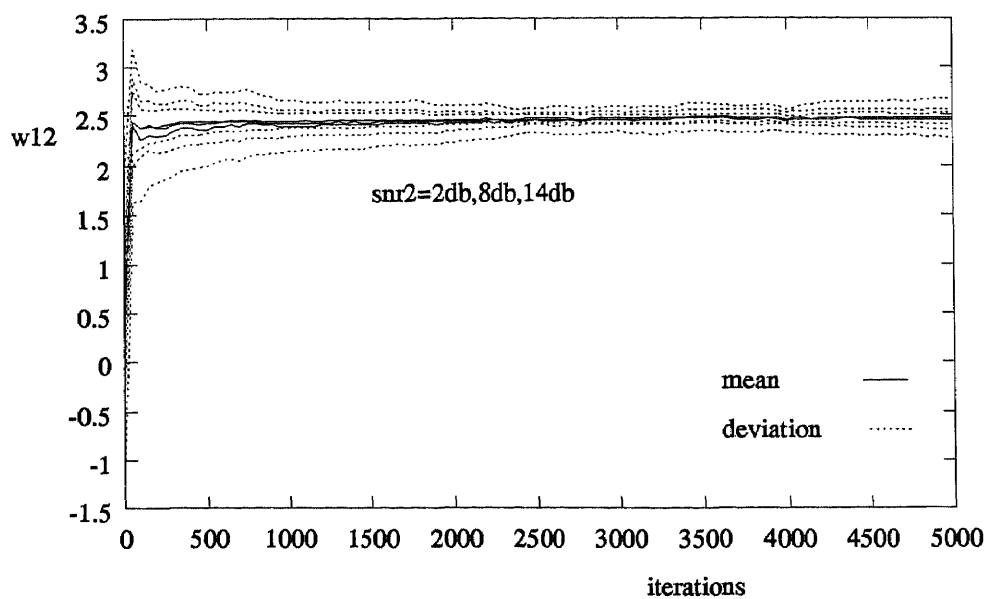**Figure 4.12** The Statistical Behavior of w_{21} with $SNR_1 = 8\ dB$, $\mu = 0.2$

Figure 4.13 The Statistical Behavior of w_{12} with $SNR_1 = 8\ dB$, $\mu = 0.2$

Figure 4.14 depicts the transient behavior of the error probability in the two-user case under the conditions that both signal SNRs are set to 8 dB, $\rho = 0.7$, and $\mu = 0.2$. It is clear that the error probability can reach the steady state in a very fast manner. In this case, with learning step μ set to 0.2, the number of iterations needed for the convergence is about 50.
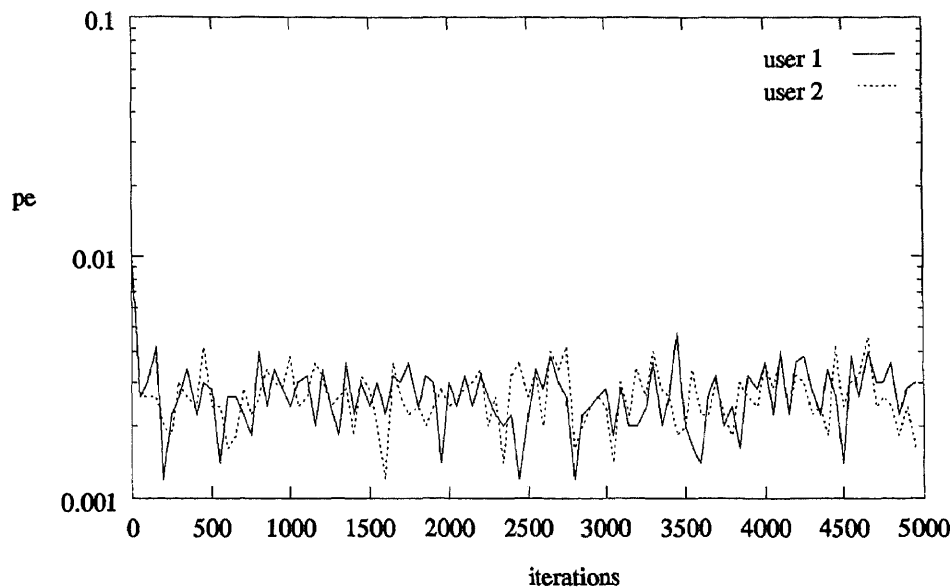

Figure 4.14 The Transient Error Probability in the Two-User Case with $\mu = 0.2$

4.3.5.2 The Three-User Case

In this section, the behavior of the weights in the three-user case, with the Gold sequences introduced earlier as signature sequences, is obtained. The statistical behavior of the weights to user 1 (w_{21} and w_{31}) with $SNR_1 = SNR_2 = SNR_3 = 8\ dB$ and different learning steps $\mu = 0.2$ and $\mu = 0.002$ are shown in Figures 4.15 and 4.16. Their statistic behavior with $\mu = 0.2$ and $SNR_1 = 8\ dB$ but $SNR_2 = SNR_3 = 8$ and 14 dB are shown in Figures 4.17 and 4.18.
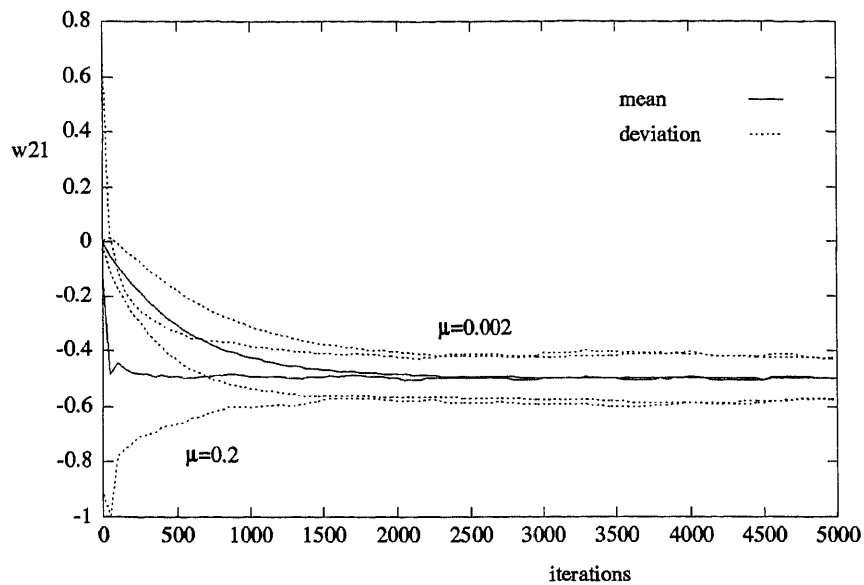
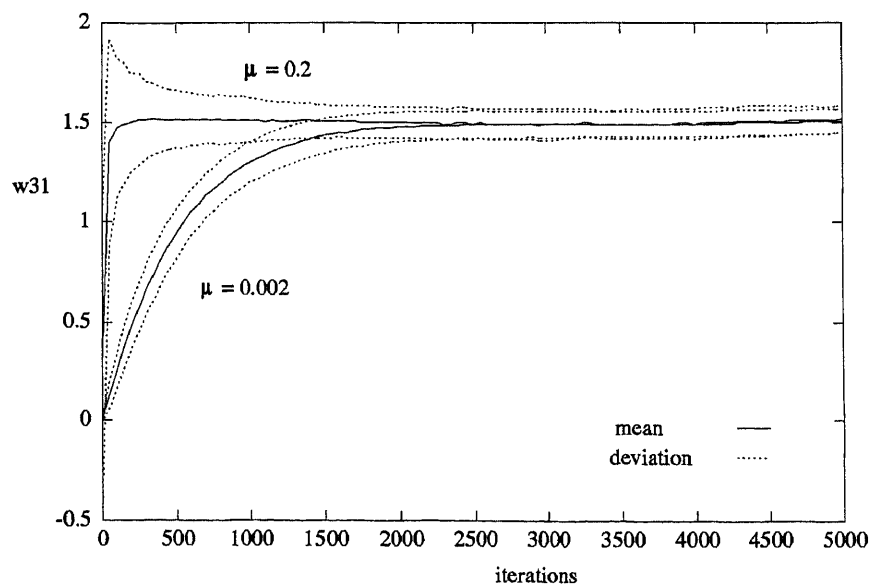**Figure 4.15** The Statistical Behavior of w_{21} with $\mu = 0.2$ and 0.002

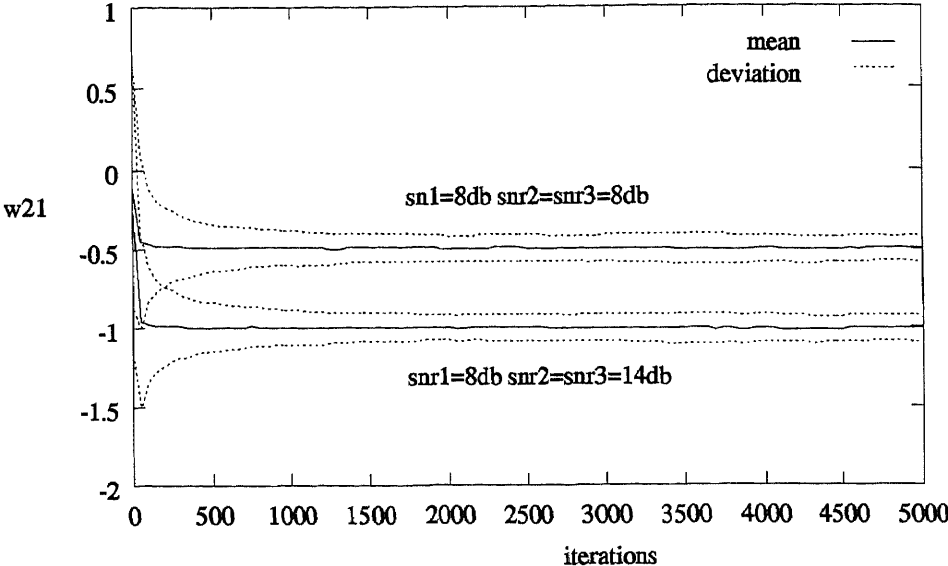**Figure 4.16** The Statistical Behavior of w_{31} with $\mu = 0.2$ and 0.002

Figure 4.17 The Statistical Behavior of w_{21} with $SNR_1 = 8\ dB$, $\mu = 0.2$
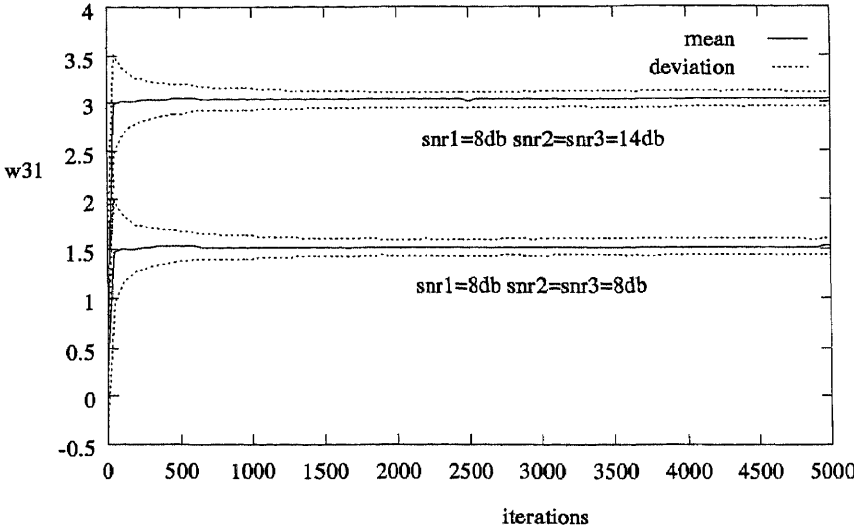

Figure 4.18 The Statistical Behavior of w_{31} with $SNR_1 = 8\ dB$, $\mu = 0.2$

Figure 4.19 shows the result of the transient behavior of the weight vector in the case that the initial value is not set to **0**. The signal SNRs are all set to 8 dB, μ is set to 0.2. The initial weights are set to $w_{21}(0) = 2$ and $w_{31}(0) = 1$. It is clear that the weights still converge to the same steady state, but the number of iterations needed to achieve the steady state is different from the previous case. Therefore, in order to achieve fast convergence, it is suggested to set the initial weight vector to **0**.
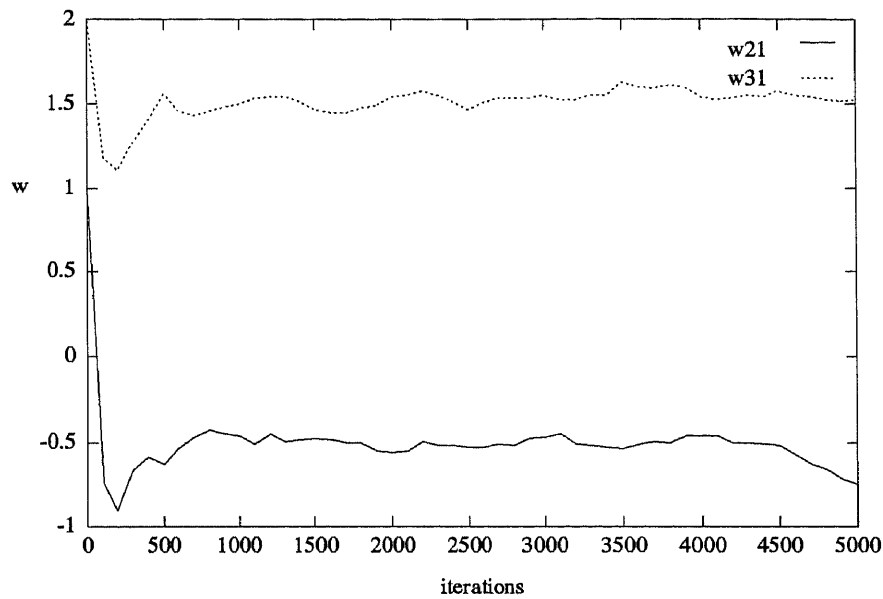

Figure 4.19 The Transient Behavior of Weights with Non-Zero Initial Values

The transient behavior of the probability of error under the conditions $SNR_1 = SNR_2 = SNR_3 = 8$ dB and $\mu = 0.2$ is also given in Figure 4.20. It can be seen that the error probability goes to its steady state very quickly, within 200 steps.
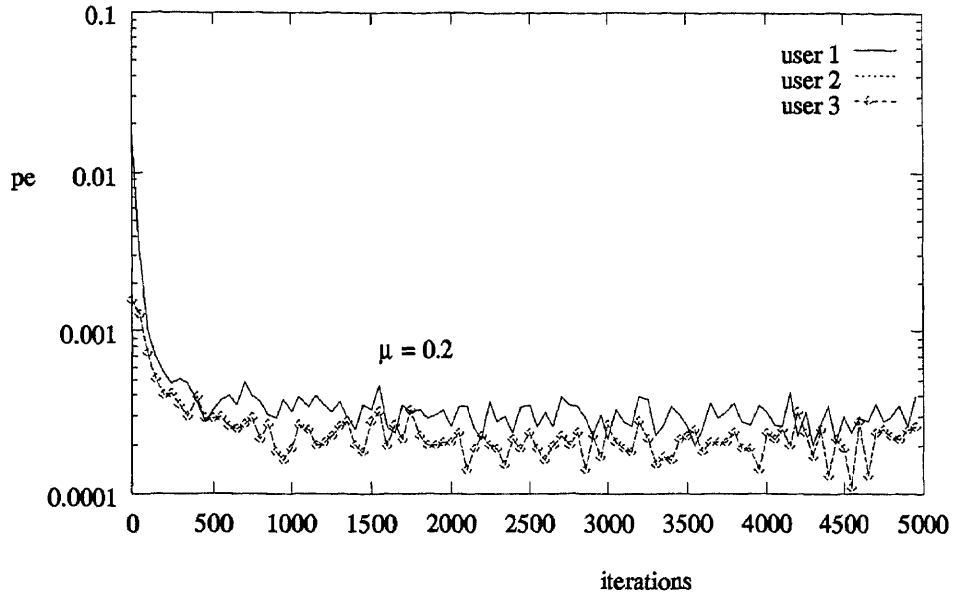
Figure 4.20 The Transient Error Probability in the Three-User Case with $\mu = 0.2$

4.3.5.3 The Four-User Case

Results for the four-user case are also given in Figure 4.21 for the weight vector to user 1 (w_{21}, w_{31}, w_{41}). In this case, all the user $SNRs$ are 8 dB, μ is chosen as 0.2. Figure 4.22 shows their statistical behavior under the same conditions.

In the simulations, it has been shown that the weights achieve their steady state values in a very short period of time, probably within 50 iterations. Their statistic behaviors show that the convergence process is relatively stable, the mean value almost monotonically changes and the deviation range is insensitive to the user $SNRs$ and learning step μ. The transient behavior of the bit-error probability also shows that the receiver can attain steady state performance in a very fast manner. Fast convergence and stability of the adaptive CDMA receiver enhance its practicality and attractiveness as a viable option for future realistic CDMA receivers.
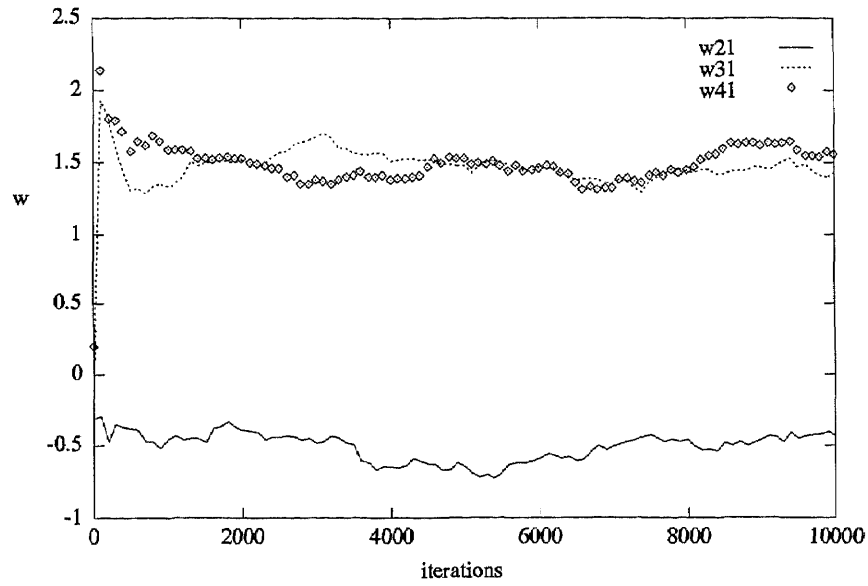
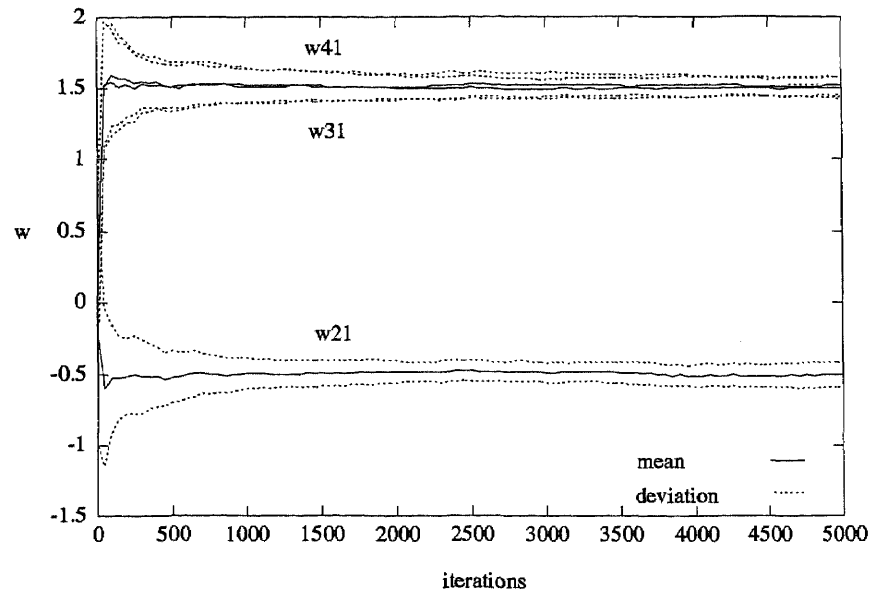**Figure 4.21** The Transient Behavior of the Weights to User 1 with $\mu = 0.2$

Figure 4.22 The Statistical Behavior of the Weights to User 1 with $\mu = 0.2$

CHAPTER 5

CONCLUSIONS

In this thesis, various structures of synchronous CDMA receivers have been reviewed and studied. The key contribution is the investigation of the convergence and stability analysis of a recently proposed adaptive two-stage receiver [14] in a synchronous code-division multiple access communication environment with additive white Gaussian background noise. With the knowledge of all user signature sequences, the receiver has been proven to be "near-far" resistant and near-optimum regarding its error probability, meanwhile its computational complexity is a square function of the number of users.

Analysis on the adaptive rule shows that the convergence of the adaptive weights to the steady states is guaranteed. Two conditions sufficient for convergence are drawn, which put no restriction on the learning step (chosen between (0,1]), though there is a loose constraint among the number of users, the signal $SNRs$, and the cross-correlation of their signature sequences. Simulations are also done to show the transient behaviors of the weights and the user error probabilities. Fast convergence is shown to be achieved by choosing the learning step close to 1. All these properties enhance the practicality of the receiver, making it a viable option for future realistic CDMA receivers.

REFERENCES

[1] S. Verdu, "Minimum Probability of Error for Asynchronous Gaussian Multiple Access Channels," *IEEE Trans. Inform. Theory*, Vol. IT-32, No. 1, pp. 85-96, Jan. 1986.

[2] M. K. Varanasi and B. Aazhang, "Multistage Detector in Asynchronous Code-Division Multiple Access Communications," *IEEE Trans. Commun.*, vol. 38, No. 4, pp. 509-519, Apr. 1990.

[3] M. K. Varanasi and B. Aazhang, "Near-Optimum Detector in Synchronous Code-Division Multiple-Access System," *IEEE Trans. Commun.*, vol. 39, No. 5, pp. 725-736, May 1991.

[4] Y. Bar-Ness and J. Rokah, "Cross-Coupled Bootstrapped Interference Canceler," *Proc. 1981 AP-S Int. Symposium*, pp. 292-295, Los Angeles, CA, June 1981.

[5] C. Herault and J. Jutten, "Blind Separation of Sources, Part I: An Adaptive Algorithm Based on Neuromimetic Architecture," *Signal Processing*, Vol. 24, pp. 1-10, 1991.

[6] E. Moreau and O. Macchi, "Two Novel Architectures for the Self Adaptive Separation of Signals," *IEEE*, Feburary 1993.

[7] A. Dinc and Y. Bar-Ness, "Comparison of Three Defferent Structures of Bootstrap Blind Adaptive Algorithm for Multisignal Co-Channel Separation," *MILCOM "92*, San Diego, CA, October 11-14, 1992.

[8] A. Dinc and Y. Bar-Ness, "Convergence and Performance Comparison of Three Different Structures of Bootstrap Blind Adaptive Algorithm for Multisignal Co-Channel Separation," *MILCOM '92*, San Diego, CA, October 11-14, 1992.

[9] A. Dinc and Y. Bar-Ness, "Error Probability of Bootstrapped Blind Adaptive Cross-Pol Cancelers for M-ary QAM over Non-dispersive Fading Channel," *Proc. ICC '92*, Paper No. 353.5.

[10] B. Aazhang, B. Paris, and G. C. Orsak, "Neural Networks for Multiuser Detection in Code Division Multiple Access Communications," *IEEE Trans. Commun.*, vol. 40, No. 7, July 1992.

[11] U. Mitra and H. V. Poor, "Adaptive Receiver Algorithms for Near-Far Resistant CDMA," *The 3rd IEEE Intl. symposium on Personal, Indoor and Mobile Radio Commun.*, Boston, MA., Oct. 19-21, 1992.

[12] C. Herault, J. Jutten, and P. Comon, "Blind Separation of Sources, Part II: Problems Statement," *Signal Processing*, Vol. 24, pp11-20, 1991.

[13] E. Dorouchyari, "Blind Separation of Sources, Part III: Stability Analysis," *Signal Processing,* Vol. 24, pp. 21-29, 1991.

[14] Z. Siveski, Y. Bar-Ness and D. W. Chen, "Adaptive Signal Separation for Synchronous Code Division Multiple Access Applications" accepted for present at the 1994 *International Zurich Seminar on Digital Communication.*

[15] Y. Jiang, "Adaptive Two-stage Detection Scheme in Synchronous Two-user CDMA Systems," Master's thesis, Department of ECE, NJIT, May, 1993.

[16] B. Widrow, J. M. McDool, M. G. Larimore, and J. C. R. Johnson, "Stationary and Nonstationary Learning Characteristics of the LMS Adaptive Filter," *Proceedings of the IEEE,* vol. 64, no. 8, pp. 1151-1162, August 1976.

[17] M. Marcus and H. Minc, "A Survey of Matrix Theory and Matrix Inequality," *Allyn and Bacon, Inc.,* Boston, 1964.

[18] H. J. Kushner and J. Yang, "Stochastic Approximation with Averaging of the Iterates: Optimal Asymptotic Rate of Convergence of General Processes," *LCDS#91-9,* Brown University, 1991.

[19] R. Lupas and S. Verdu, "Linear Multiuser Detectors for Synchronous Code Division Multiple Access Channels," *IEEE Trans. Inform. Theory,* vol. IT-35, No. 1, pp. 123-136, Jan. 1989.

[20] R. Lupas and S. Verdu, "Near Far Resistance of Multiuser Detectors in Asynchronous Channels," *IEEE Trans. Commun.,* vol. COM-38, No. 4, pp. 496-508, Apr. 1990.

[21] J. S. Lehnert and M. B. Pursley, "Error probabilities for binary direct-sequence spread-spectrum communications with random signature sequences," *IEEE Trans. Commun.,* vol. COM-35, pp. 87-98, Jan. 1987.

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

[23] D. Lu and K. Yao, "Improved Importance Sampling Technique for Efficient Simulation of Digital Communication Systems," *IEEE Selec. Areas Commun.,* vol. 6, pp. 67-75, Jan. 1988.

[24] G. Orsak and B. Aazhang, "Efficient Importance Sampling Techniques for Simulation of Multiuser Communication Systems," *IEEE Trans. Commun.,* vol. 40, pp. 1111-1118, June 1992.

[25] S. Haykin, *Adaptive Filter Theory,* 2nd edition, Prentice-Hall, Inc., 1991.

[26] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, Prentice-Hall, Inc., 1985.

[27] S. Barnett and C. Storey, *Matrix Methods in Stability Theory*, Barnes & Noble, Inc., New York, 1970.

[28] H. J. Kushner, *Approximation and Weak Convergence Methods for Random* Process, with Applications to Stochastic Systems Theory, The MIT Press, 1984.