A Blind Adaptive Decorrelating Detector for CDMA Systems*

Sennur Ulukus **Roy D. Yates**

August 29, 1997

Department of Electrical and Computer Engineering
Wireless Information Networks Laboratory (WINLAB)
Rutgers University, PO Box 909, Piscataway NJ 08855-0909
Tel: (908) 445 5249 Fax: (908) 445 2820
ulukus@winlab.rutgers.edu *ryates@winlab.rutgers.edu*

Abstract

The decorrelating detector is known to eliminate multi access interference when the signature sequences of the users are linearly independent, at the cost of enhancing the Gaussian receiver noise. In this paper, we present a blind adaptive decorrelating detector which is based on the observation of readily available statistics. The algorithm recursively updates the filter coefficients of a desired user by using the output of the current filter. Due to the randomness of the information bits transmitted and the ambient Gaussian channel noise, the filter coefficients evolve stochastically. We prove the convergence of the filter coefficients to a decorrelating detector in the mean squared error (MSE) sense. We develop lower and upper bounds on the MSE of the receiver filter from the convergence point and show that with a fixed step size sequence, the MSE can be made arbitrarily small by choosing a small enough step size; and with a time-varying step size sequence the MSE converges to zero implying an exact convergence. The proposed algorithm is distributed in the sense that no information about the interfering users such as their signature sequences or power levels is needed. The algorithm requires the knowledge of only two parameters for the construction of the receiver filter of a desired user: the desired user's signature sequence and the variance of the additive white Gaussian receiver noise. This detector for an asynchronous CDMA channel converges to the one-shot decorrelating detector.

*Supported by NSF Grant NCR 95-06505

1 Introduction

Code Division Multiple Access (CDMA) systems suffer from the *near-far effect* because of non-orthogonality of the users' signature sequences. Multiuser detection [1] can be used to overcome the near-far problem by exploiting the known structure of the multiple access interference to effectively demodulate the non-orthogonal signals of the users. It was shown in [2] that the optimal multiuser detector has a computational complexity which increases exponentially with the number of active users. Several suboptimum detectors have been proposed to achieve a performance comparable to that of the optimum detector while keeping the complexity low. Examples include the decorrelating detector [3], the decision feedback detector [4], the minimum mean squared error (MMSE) detector [5] and multistage detectors [6].

The decorrelating detector [3, 7] achieves the same asymptotic efficiency as the optimal multiuser detector and has linear (in terms of the number of users) computational complexity. The decorrelating detector was shown to eliminate the multi access interference totally if the signature sequences of the users are linearly independent at the cost of enhancing the additive white Gaussian receiver noise. The decorrelating detector of [3, 7] is centralized and non-iterative. The construction of the decorrelating detector filter for a certain user requires the knowledge of the signature sequences of all the interfering users as well as the signature sequence of the user of interest. In addition, for an N user system, the construction requires inversion of the $N \times N$ cross correlation matrix. Blind adaptive algorithms are desirable to overcome the need for the knowledge about the parameters of the interfering users and iterative algorithms are needed to avoid the matrix inversion which may have a large computational complexity. A blind adaptive algorithm based on the minimization of the output energy was given in [8]. This algorithm was shown to converge to the MMSE multiuser detector [5].

In [9], an adaptive multiuser detector which converges to the decorrelating detector is proposed. This detector still needs the signature sequences of all users. In [10], blind algorithms based on signal subspace tracking are investigated and two algorithms which converge to the decorrelating and MMSE multiuser detectors are proposed. The blind adaptive decorrelating detector proposed in [10] needs information about the variance of the additive white Gaussian channel noise and the number of users, both of which can be estimated again using the subspace tracking techniques. The computational complexity of the algorithm of [10] is $O(GN)$ per iteration, where G is the processing gain and N is the number of active users.

In this paper, we present a blind adaptive multiuser detector which uses observables that are readily available at the receiver and which converges to a decorrelating detector. The detector is constructed via a distributed iterative algorithm which updates the receiver filter coefficients of a desired user by using the previous output of the filter under construction. Since the filter output is random due to the randomness of the multiaccess interference and existence of ambient Gaussian channel noise, the algorithm evolves stochastically. We prove the convergence of the filter coefficients to a decorrelating detector in the mean squared error (MSE) sense. We develop lower and upper bounds for the MSE of the filter coefficients at iteration $n+1$ in terms of the MSE of the filter coefficients at iteration n, and investigate the conditions under which the MSE sequence converges to zero. Two kinds of algorithms using a fixed step size sequence and using a time dependent step size sequence are proposed. For the algorithm using a fixed step size, we obtain the lower and upper bounds for the limiting MSE as n approaches infinity in terms of the step size. We show that the limiting MSE can be made arbitrarily small by choosing a small enough step size. On the other hand we directly prove the convergence of limiting MSE to zero when a time-varying step size sequence of a special structure is used.

The computational complexity of the proposed algorithm is $O(G)$ per iteration. The proposed algorithm requires the knowledge of only two parameters for the construction of the filter coefficients for a desired user: the desired user's signature sequence and the variance of the additive white Gaussian noise (AWGN). The variance of the AWGN is a fixed quantity (not time varying) and can be estimated easily, perhaps before the information transfer starts when only the samples of AWGN can be observed without any interference at the output of an arbitrary (non-zero) receiver filter. In such a case, a reliable estimate of the variance of AWGN can be obtained by time-averaging the squares of the output of the receiver filter.

In the following section we will briefly summarize the well known decorrelating detector and develop the necessary background for the presentation of the blind adaptive decorrelating detector.

2 The Decorrelating Detector

Until Section 5 where the convergence of the proposed algorithm for an asynchronous CDMA channel will be investigated, we will assume a synchronous CDMA system. Throughout this paper a BPSK modulation is assumed to simplify the analysis. We will use G dimensional vector $\boldsymbol{s}_i$ to denote the pre-assigned unique signature sequence of user i. Let us define an $N \times G$ dimensional matrix $\boldsymbol{S}$ with its (i,j)th element being $(\boldsymbol{s}_i)_j$, the jth component of $\boldsymbol{s}_i$. Therefore, the rows of $\boldsymbol{S}$ (equivalently the columns of $\boldsymbol{S}^\top$) are the signature sequences of the users. For future use, we define a subspace $\mathcal{L}$ in G dimensional vector space to be the subspace spanned by the signature sequences of the users, i.e.,

$$\mathcal{L} = \text{span}\{\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N\} = \text{column space of } \boldsymbol{S}^\top \tag{1}$$

We consider the decorrelating detector for the first user without loss of generality and represent it by a G dimensional vector $\boldsymbol{c}$. Then $\boldsymbol{c}$ should satisfy the following condition,

$$\boldsymbol{S}\boldsymbol{c} = \alpha \boldsymbol{e}_1 \tag{2}$$

where $\boldsymbol{e}_1$ is the first unit vector in N dimensional space, i.e., $\boldsymbol{e}_1 = [1\ 0\ 0 \ldots\ 0]^\top$, and α is a non-negative real number. Equation (2) states that the receiver filter for the desired (first) user should be orthogonal to the signature sequences of all the interfering users, and it should have a nonzero correlation with the desired user. Here this nonzero correlation is equal to the scaling factor α. It is not difficult to show that the bit error rate (BER) performance of the decorrelating detector is insensitive to scaling of the filter as long as the filter eliminates the multi access interference totally. The reason for this is that the scalar factor multiplies both the received power level of the desired user and the AWGN. Thus, for any nonnegative value of α we have a decorrelating detector with the same BER performance. The importance of using $\alpha \neq 1$ will be evident in the next section where we will avoid the need for the knowledge about the received power of the desired user by a particular selection of α.

We first note that (2) has more than one solution, because it has N equalities and G unknowns, and typically $G \gg N$. The unique decorrelating detector for the first user, $\tilde{\boldsymbol{c}}$, is given in [3] as

$$\tilde{\boldsymbol{c}} = \boldsymbol{S}^\top (\boldsymbol{S}\boldsymbol{S}^\top)^{-1} \boldsymbol{e}_1 \tag{3}$$

Let us denote any solution of (2) as $\bar{\boldsymbol{c}}$, then inserting (2) into (3) we obtain (assuming $\alpha = 1$ for the time being)

$$\begin{aligned} \tilde{\boldsymbol{c}} &= \boldsymbol{S}^\top (\boldsymbol{S}\boldsymbol{S}^\top)^{-1} \boldsymbol{S}\bar{\boldsymbol{c}} && (4) \\ &= \boldsymbol{P}\bar{\boldsymbol{c}} && (5) \end{aligned}$$

Note that $\boldsymbol{P} = \boldsymbol{S}^\top (\boldsymbol{S}\boldsymbol{S}^\top)^{-1} \boldsymbol{S}$ is the projection matrix which projects any vector onto the column space of $\boldsymbol{S}^\top$. Also note that the column space of $\boldsymbol{S}^\top$ is the subspace spanned by $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N$, i.e., the subspace which was previously denoted as $\mathcal{L}$. Therefore, although (2) has more than one solution, all of the solutions have the same projection onto $\mathcal{L}$; and this projection is equal to the unique decorrelating detector solution of [3].

Let $\boldsymbol{\Lambda}$ be an $N \times N$ dimensional diagonal matrix with received power of user i, p_i, being its ith diagonal element. Multiplying both sides of (2) first with $\boldsymbol{\Lambda}$, then with $\boldsymbol{S}^\top$, we obtain

$$\boldsymbol{S}^\top \boldsymbol{\Lambda} \boldsymbol{S} \boldsymbol{c} = \alpha p_1 \boldsymbol{s}_1 \tag{6}$$

We observe that although (6) has G equations in G unknowns it does not have a unique solution for $\boldsymbol{c}$ since $(G-N)$ eigenvalues of $\boldsymbol{S}^\top \boldsymbol{\Lambda} \boldsymbol{S}$ are equal to zero . The solution spaces of (2) and (6) are related as stated in the following Lemma.

Lemma 1 *If $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N$ are linearly independent, then all solutions of (2) and (6) coincide.*

A proof for Lemma 1 is given in the Appendix.

Let us define $\boldsymbol{A} = \boldsymbol{S}^\top \boldsymbol{\Lambda} \boldsymbol{S}$. At this point we choose $\alpha = 1/p_1$ and devise the following gradient descent algorithm:

$$\boldsymbol{c}(n+1) = \boldsymbol{c}(n) - \mu \left(\boldsymbol{A} \boldsymbol{c}(n) - \boldsymbol{s}_1 \right) \tag{7}$$

Note that the iterative algorithm given in (7) converges to the solution of (6) and equivalently, by Lemma 1, to the solution of (2). The fact that the solution of (2) is not unique is noted earlier. Note that, for any $\boldsymbol{c}$

$$\boldsymbol{P} \boldsymbol{A} \boldsymbol{c} = \boldsymbol{S}^\top (\boldsymbol{S} \boldsymbol{S}^\top)^{-1} \boldsymbol{S} \boldsymbol{S}^\top \boldsymbol{\Lambda} \boldsymbol{S} \boldsymbol{c} = \boldsymbol{S}^\top \boldsymbol{\Lambda} \boldsymbol{S} \boldsymbol{c} = \boldsymbol{A} \boldsymbol{c} \tag{8}$$

This means that for any $\boldsymbol{c}(n)$, $\boldsymbol{A}\boldsymbol{c}(n)$ lies entirely in $\mathcal{L}$. Since $\boldsymbol{s}_1 \in \mathcal{L}$ by definition, the correction term added to $\boldsymbol{c}(n)$ at every iteration in (7) is always in $\mathcal{L}$. Starting from $n = 0$, by induction we obtain

$$\boldsymbol{c}(n) = \boldsymbol{P}\boldsymbol{c}(n) + \boldsymbol{c}(0) - \boldsymbol{P}\boldsymbol{c}(0) \tag{9}$$

Therefore, if the iterative algorithm (7) is started in the subspace $\mathcal{L}$, implying $\boldsymbol{c}(0) = \boldsymbol{P}\boldsymbol{c}(0)$, then from (9) we will have: $\boldsymbol{c}(n) = \boldsymbol{P}\boldsymbol{c}(n)$ for all n. In this case $\boldsymbol{c}(n)$ will always stay in $\mathcal{L}$ and it will converge to the scaled version of the unique decorrelating detector solution $\boldsymbol{c}^* = \frac{1}{p_1}\tilde{\boldsymbol{c}}$ as n goes to infinity. Note that the algorithm converges to the scaled version of $\tilde{\boldsymbol{c}}$, the decorrelating detector of [3, 7], instead of converging to $\tilde{\boldsymbol{c}}$, because α was chosen to be $1/p_1$ instead of 1. Note also that $\tilde{\boldsymbol{c}}$ and $\boldsymbol{c}^* = \frac{1}{p_1}\tilde{\boldsymbol{c}}$ have the same BER performance as discussed earlier.

The restriction that the iterations should be started in $\mathcal{L}$ for algorithm (7) to converge to a decorrelating detector is fairly mild. Selections $\boldsymbol{c}(0) = \boldsymbol{0}$, $\boldsymbol{c}(0) = \boldsymbol{s}_1$ or any linear combination of the signature sequences imply $\boldsymbol{c}(0) \in \mathcal{L}$, satisfying the convergence condition of (7). The signature sequences of all of the users must be known for the algorithm given in

(7). Also, the algorithm of (7) is an off-line algorithm in the sense that it does not utilize any real time measurements or observations. The algorithm of (7) can be run before the real information transmission of mobile to the base station starts. After running the algorithm for some time, the filter coefficients would converge to the decorrelating detector filter and then the communication can be started. In this paper, our aim is to develop a blind adaptive algorithm which would converge to the decorrelating detector solution in a stochastic sense by using real random measurements while the users are active and transmitting bits. To this end, we will propose an algorithm which can be viewed as the stochastic version of the deterministic algorithm given in (7).

3 A Blind Adaptive Decorrelating Detector (BADD)

The received base band signal before the receiver filters can be written as

$$\boldsymbol{r} = \sum_{i=1}^{N} \sqrt{p_i} a_i \boldsymbol{s}_i + \boldsymbol{n} \tag{10}$$

where a_i is the information bit (± 1, equiprobably), p_i is the power and $\boldsymbol{s}_i$ is the signature sequence of user i and $\boldsymbol{n}$ is an a Gaussian random vector with zero mean and $E[\boldsymbol{n}\boldsymbol{n}^\top] = \sigma^2 \boldsymbol{I}$. Note that,

$$\begin{aligned} E\left[\boldsymbol{r}\boldsymbol{r}^\top\right] &= \sum_{i=1}^{N} p_i \boldsymbol{s}_i \boldsymbol{s}_i^\top + \sigma^2 \boldsymbol{I} \\ &= \boldsymbol{A} + \sigma^2 \boldsymbol{I} \end{aligned} \tag{11}$$

Therefore, using (11), the deterministic iteration of (7) can be written in an exact form as

$$\boldsymbol{c}(n+1) = \boldsymbol{c}(n) - \mu \left[\left(E\left[\boldsymbol{r}\boldsymbol{r}^\top\right] - \sigma^2 \boldsymbol{I} \right) \boldsymbol{c}(n) - \boldsymbol{s}_1 \right] \tag{12}$$

Note that $\boldsymbol{r}\boldsymbol{r}^\top - \sigma^2 \boldsymbol{I}$ is an *unbiased* estimate for $\boldsymbol{A}$ matrix. In order to obtain a practical algorithm, we replace $E\left[\boldsymbol{r}\boldsymbol{r}^\top\right]$ in (12) with the estimate $\boldsymbol{r}(n)\boldsymbol{r}^\top(n)$ where $\boldsymbol{r}(n)$ is the received signal vector at the nth iteration.

$$\boldsymbol{c}(n+1) = \boldsymbol{c}(n) - \mu \left[\left(\boldsymbol{r}(n)\boldsymbol{r}^\top(n) - \sigma^2 \boldsymbol{I} \right) \boldsymbol{c}(n) - \boldsymbol{s}_1 \right] \tag{13}$$

Before analyzing the convergence of (13), we state it in terms of available observations, and list the parameters needed at each iteration. We note that the output of the receiver filter of the desired user at time n is $y(n) = \boldsymbol{r}^\top(n)\boldsymbol{c}(n)$. Thus, the implementation oriented version of the algorithm (13) is

$$\boldsymbol{c}(n+1) = \left(1 - \sigma^2\mu\right)\boldsymbol{c}(n) - \mu\left(\boldsymbol{r}(n)y(n) - \boldsymbol{s}_1\right) \tag{14}$$

Since $\boldsymbol{r}(n)$ and $y(n)$ are readily available at the input and output of the receiver filter that is under construction, only two system parameters are needed to be known in order to run the algorithm: the signature sequence of the desired user $\boldsymbol{s}_1$, and the variance of the AWGN, σ^2. The variance of the AWGN is a fixed quantity which can be easily estimated before the communication starts by using the outputs of a nonzero receiver filter as discussed in Section 1.

4 Convergence of the BADD

In this Section we will investigate the convergence of the blind adaptive decorrelating detector (BADD) proposed in the previous section. Let us define the zero mean random vector $\boldsymbol{\eta}(n)$ as

$$\boldsymbol{\eta}(n) = \left(\boldsymbol{r}\boldsymbol{r}^\top - \boldsymbol{A} - \sigma^2\boldsymbol{I}\right)\boldsymbol{c}(n) \tag{15}$$

Noting that $\boldsymbol{A}\boldsymbol{c}^* = \boldsymbol{s}_1$, we can write the stochastic iterations (13) and (14) as

$$\boldsymbol{c}(n+1) = \boldsymbol{c}(n) - \mu\left[\boldsymbol{A}\left(\boldsymbol{c}(n) - \boldsymbol{c}^*\right) + \boldsymbol{\eta}(n)\right] \tag{16}$$

Subtracting $\boldsymbol{c}^*$ from both sides of (16) and defining $\boldsymbol{\epsilon}(n) = \boldsymbol{c}(n) - \boldsymbol{c}^*$ we obtain

$$\boldsymbol{\epsilon}(n+1) = \boldsymbol{\epsilon}(n) - \mu\left(\boldsymbol{A}\boldsymbol{\epsilon}(n) + \boldsymbol{\eta}(n)\right) \tag{17}$$

Note that $\|\boldsymbol{\epsilon}(n)\|^2$, the norm of $\boldsymbol{\epsilon}(n)$, is a measure of the distance of the receiver filter at iteration n from the convergence point. In order to relate the distance measures at iterations n and $n+1$, we square both sides of (17) and obtain

$$\begin{aligned} \|\boldsymbol{\epsilon}(n+1)\|^2 &= \|\boldsymbol{\epsilon}(n)\|^2 - 2\mu\boldsymbol{\epsilon}(n)^\top\boldsymbol{A}\boldsymbol{\epsilon}(n) + 2\mu^2\boldsymbol{\epsilon}(n)^\top\boldsymbol{A}\boldsymbol{\eta}(n) - 2\mu\boldsymbol{\epsilon}(n)^\top\boldsymbol{\eta}(n) \\ &\quad + \mu^2\boldsymbol{\epsilon}(n)^\top\boldsymbol{A}^2\boldsymbol{\epsilon}(n) + \mu^2\|\boldsymbol{\eta}(n)\|^2 \end{aligned} \tag{18}$$

Taking the conditional expectation of both sides of (18), conditioned on $\boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}$, and observing that $E\left[\boldsymbol{\eta}(n) \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}\right] = \boldsymbol{0}$ yields

$$E\left[\|\boldsymbol{\epsilon}(n+1)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}\right] = \|\boldsymbol{\epsilon}\|^2 - 2\mu\boldsymbol{\epsilon}^\top\boldsymbol{A}\boldsymbol{\epsilon} + \mu^2\boldsymbol{\epsilon}^\top\boldsymbol{A}^2\boldsymbol{\epsilon} + \mu^2 E\left[\|\boldsymbol{\eta}(n)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}\right] \tag{19}$$

We will be using the results of the following Lemmas to develop bounds for the right hand side of (19).

Lemma 2 *If* $\boldsymbol{P\epsilon} = \mathbf{0}$ *then* $\boldsymbol{\epsilon}^\top \boldsymbol{A\epsilon} = 0$ *and* $\boldsymbol{\epsilon}^\top \boldsymbol{A}^2\boldsymbol{\epsilon} = 0$*. If* $\boldsymbol{P\epsilon} \neq \mathbf{0}$ *then there exist* $0 < k_0 \leq k_1 < \infty$*, such that*

$$k_0 \|\boldsymbol{\epsilon}\|^2 \leq \boldsymbol{\epsilon}^\top \boldsymbol{A\epsilon} \leq k_1 \|\boldsymbol{\epsilon}\|^2 \qquad \textit{and} \qquad k_0^2 \|\boldsymbol{\epsilon}\|^2 \leq \boldsymbol{\epsilon}^\top \boldsymbol{A}^2\boldsymbol{\epsilon} \leq k_1^2 \|\boldsymbol{\epsilon}\|^2 \tag{20}$$

Lemma 3 *There exist* $0 < c_0 \leq c_1 < \infty$*, such that*

$$0 \leq E\left[\|\boldsymbol{\eta}(n)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}\right] \leq c_0 + c_1 \|\boldsymbol{\epsilon}\|^2 \tag{21}$$

The proofs of Lemmas 2 and 3 can be found in the Appendix.

If $\boldsymbol{P\epsilon} = \mathbf{0}$, we can develop the following lower and upper bounds for the right hand side of (19) using Lemmas 2 and 3,

$$E\left[\|\boldsymbol{\epsilon}(n+1)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}, \boldsymbol{P\epsilon} = \mathbf{0}\right] \leq \left(1 + \mu^2 c_1\right) \|\boldsymbol{\epsilon}\|^2 + c_0\mu^2 \tag{22}$$

$$E\left[\|\boldsymbol{\epsilon}(n+1)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}, \boldsymbol{P\epsilon} = \mathbf{0}\right] \geq \|\boldsymbol{\epsilon}\|^2 \tag{23}$$

If $\boldsymbol{P\epsilon} \neq \mathbf{0}$, the following lower and upper bounds for the right hand side of (19) can be obtained using again Lemmas 2 and 3,

$$E\left[\|\boldsymbol{\epsilon}(n+1)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}, \boldsymbol{P\epsilon} \neq \mathbf{0}\right] \leq \left(1 - 2\mu k_0 + (k_1^2 + c_1)\mu^2\right) \|\boldsymbol{\epsilon}\|^2 + c_0\mu^2 \tag{24}$$

$$E\left[\|\boldsymbol{\epsilon}(n+1)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}, \boldsymbol{P\epsilon} \neq \mathbf{0}\right] \geq \left(1 - 2\mu k_1 + k_0^2\mu^2\right) \|\boldsymbol{\epsilon}\|^2 \tag{25}$$

An important observation toward the convergence proof is that the projection of $\boldsymbol{c}(n)$ (equivalently the projection of $\boldsymbol{\epsilon}(n)$) on $\mathcal{L}$ would be non-zero *almost surely (a.s.)* [11]. This means that for any n the probability of the event $[\boldsymbol{P\epsilon}(n) = \mathbf{0}]$ is zero. Similar to the deterministic result in (9), it can be shown using induction on (17) that (see also [12, Eqn. (16)]),

$$\boldsymbol{\epsilon}(n) = \boldsymbol{P\epsilon}(n) + \boldsymbol{\epsilon}(0) - \boldsymbol{P\epsilon}(0) \qquad \text{a. s.} \tag{26}$$

If $\boldsymbol{\epsilon}(0) \in \mathcal{L}$ then $\boldsymbol{\epsilon}(0) = \boldsymbol{P\epsilon}(0)$ and (26) can be written as

$$\boldsymbol{\epsilon}(n) = \boldsymbol{P\epsilon}(n) \qquad \text{a. s.} \tag{27}$$

Note that $\boldsymbol{\epsilon}(0) = \boldsymbol{c}(0) - \boldsymbol{c}^*$, since $\boldsymbol{c}^* \in \mathcal{L}$, $\boldsymbol{\epsilon}(0) \in \mathcal{L}$ is equivalent to $\boldsymbol{c}(0) \in \mathcal{L}$. Thus, (27) is valid if the iteration (13) is started in $\mathcal{L}$. The upper bound on $E\left[\|\boldsymbol{\epsilon}(n+1)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}\right]$

can be developed as follows.

$$
\begin{aligned}
E\left[\|\boldsymbol{\epsilon}(n+1)\|^2 \mid \boldsymbol{\epsilon}(n)=\boldsymbol{\epsilon}\right] &= E\left[\|\boldsymbol{\epsilon}(n+1)\|^2 \mid \boldsymbol{\epsilon}(n)=\boldsymbol{\epsilon}, \boldsymbol{P\epsilon}=\boldsymbol{0}\right]\mathrm{Prob}\left[\boldsymbol{P\epsilon}=\boldsymbol{0}\right] \\
&\quad + E\left[\|\boldsymbol{\epsilon}(n+1)\|^2 \mid \boldsymbol{\epsilon}(n)=\boldsymbol{\epsilon}, \boldsymbol{P\epsilon}\neq\boldsymbol{0}\right]\mathrm{Prob}\left[\boldsymbol{P\epsilon}\neq\boldsymbol{0}\right] \\
&= E\left[\|\boldsymbol{\epsilon}(n+1)\|^2 \mid \boldsymbol{\epsilon}(n)=\boldsymbol{\epsilon}, \boldsymbol{P\epsilon}\neq\boldsymbol{0}\right] \\
&\leq \left(1-2\mu k_0+(k_1^2+c_1)\mu^2\right)\|\boldsymbol{\epsilon}\|^2+c_0\mu^2 \qquad (28)
\end{aligned}
$$

where we used the facts $\mathrm{Prob}\left[\boldsymbol{P\epsilon}=\boldsymbol{0}\right]=0$ and $\mathrm{Prob}\left[\boldsymbol{P\epsilon}\neq\boldsymbol{0}\right]=1$ and the result in (24). Following similar steps to those in (28), and using the result given in (25) this time, we can obtain the following lower bound for the same term

$$
E\left[\|\boldsymbol{\epsilon}(n+1)\|^2 \mid \boldsymbol{\epsilon}(n)=\boldsymbol{\epsilon}\right] \geq \left(1-2\mu k_1+k_0^2\mu^2\right)\|\boldsymbol{\epsilon}\|^2 \qquad (29)
$$

Taking the expectation of both sides of the inequalities in (28) and (29) with respect to $\boldsymbol{\epsilon}(n)$ and letting $b_n=E\left[\|\boldsymbol{\epsilon}(n)\|^2\right]$, we obtain the following lower and upper bounds for b_{n+1}, the mean squared error (MSE) of the filter coefficients at iteration $n+1$ from $\boldsymbol{c}^*$, in terms of b_n

$$
\left(1-2\mu k_1+k_0^2\mu^2\right)b_n \leq b_{n+1} \leq \left(1-2\mu k_0+(k_1^2+c_1)\mu^2\right)b_n+c_0\mu^2 \qquad (30)
$$

By defining

$$
\alpha_0=1-2\mu k_0+(k_1^2+c_1)\mu^2 \qquad \alpha_1=1-2\mu k_1+k_0^2\mu^2 \qquad (31)
$$

we can rewrite (30) as

$$
\alpha_1 b_n \leq b_{n+1} \leq \alpha_0 b_n+c_0\mu^2 \qquad (32)
$$

We observe from (32) that the nonnegative sequence b_n is sandwiched between the two sequences generated according to $b'_{n+1}=\alpha_0 b'_n+c_0\mu^2$ and $b''_{n+1}=\alpha_1 b''_n$. These two sequences converge to finite numbers if and only if μ is chosen such that $\mid \alpha_0 \mid< 1$ and $\mid \alpha_1 \mid< 1$. Note that both α_0 and α_1 are equal to 1 at $\mu=0$. We also note that both α_0 and α_1 are locally decreasing as μ increases, since

$$
\left.\frac{d\alpha_0}{d\mu}\right|_{\mu=0}=-2k_0<0 \qquad \text{and} \qquad \left.\frac{d\alpha_1}{d\mu}\right|_{\mu=0}=-2k_1<0 \qquad (33)
$$

This means that we can always choose μ small enough so that

$$
|\alpha_0|<1 \qquad \text{and} \qquad |\alpha_1|<1 \qquad (34)
$$

in which case the sequences b'_n and b''_n converge and the limiting MSE, i.e., $\lim_{n\to\infty} b_n$, has finite lower and upper bounds. From the sandwich theorem, we have

$$0 \leq \lim_{n\to\infty} b_n \leq \frac{c_0\mu^2}{1-\alpha_0} \tag{35}$$

We can evaluate the value of the upper bound in the extreme case when $\mu \to 0$ as,

$$\lim_{\mu\to 0} \frac{c_0\mu^2}{1-\alpha_0} = \lim_{\mu\to 0} \frac{c_0\mu}{2k_0 - (k_1^2 + c_1)\mu} = 0 \tag{36}$$

Therefore, if the step size (μ) is chosen extremely small then the MSE of the filter coefficients from the convergence point goes to zero as the number of iterations grows to infinity. But note that as $\mu \to 0$, the numbers α_0 and α_1 go to 1 in which case the convergence rate goes to zero. Thus we observe the trade off between the limiting value of the MSE and the convergence rate: if a large value is chosen as the step size μ then the convergence rate is faster but the limiting MSE is larger; and if a small value is chosen as the step size the limiting MSE is smaller but the convergence rate is slower too. Hence, a time dependent step size sequence which takes large values at the beginning and smaller values at the end may be preferable. An iteration index (n) dependent step size sequence can be used to accomplish this. Replacing the fixed step size μ in (14) with the time varying step size sequence a_n, we obtain the new algorithm to be:

$$\boldsymbol{c}(n+1) = \left(1 - \sigma^2 a_n\right)\boldsymbol{c}(n) - a_n\left(\boldsymbol{r}(n)y(n) - \boldsymbol{s}_1\right) \tag{37}$$

If the step size sequence satisfies two simple conditions shown in (38) then the filter coefficients converge to the decorrelating detector solution in the MSE sense.

$$\sum_n^\infty a_n = \infty \qquad \text{and} \qquad \sum_n^\infty a_n^2 < \infty \tag{38}$$

The convergence is stated in the following Lemma which is proved in the Appendix.

Lemma 4 *Stochastic iteration given in (37) converges to* $\boldsymbol{c}^*$ *in the MSE sense, i.e.,*

$$\lim_{n\to 0} E\left[\|\boldsymbol{c}(n) - \boldsymbol{c}^*\|^2\right] = 0$$

if a_n *satisfies the conditions of (38).*

Note that $a_n = a/(n+n_0)$, for any $a > 0$ and $n_0 > 0$ is a sequence which satisfies (38).

5 The BADD in an Asynchronous CDMA System

Let τ_i denote the delay of the ith user. Without loss of generality, we will assume that τ_1, the desired user's delay, is equal to zero and $\tau_i \geq 0$, for $i > 1$. As in [1], each asynchronous interfering user i, can be replaced with two synchronous users $2i-2$ and $2i-1$ with modified signature sequences; see Figure 1. The modified signature sequences, $\bar{\boldsymbol{s}}_i$, of these synchronous users are given as

$$\begin{aligned}
\bar{\boldsymbol{s}}_{2i-2} &= \begin{cases} s_i(t) & 0 \leq t \leq \tau_i \\ 0 & \tau_i < t \leq T_b \end{cases} \\
\bar{\boldsymbol{s}}_{2i-1} &= \begin{cases} 0 & 0 \leq t \leq \tau_i \\ s_i(t) & \tau_i < t \leq T_b \end{cases}
\end{aligned}$$

where T_b is the bit duration. In this case the chip sampled received signal before the received filter of the desired user at the nth bit interval is given as

$$\boldsymbol{r}(n) = \sum_{i=1}^{N} \sqrt{p_i} \left(a_i(n-1)\bar{\boldsymbol{s}}_{2i-2} + a_i(n)\bar{\boldsymbol{s}}_{2i-1} \right) + \boldsymbol{n} \tag{39}$$

where $a_i(n)$ denotes the information bit transmitted by user i in the nth bit interval. In (39), for convenience, the desired (first) user is also written in terms of the signature of two equivalent users. Note that the signature sequence $\bar{\boldsymbol{s}}_0$ is identically zero and $\bar{\boldsymbol{s}}_1 = \boldsymbol{s}_1$.

From the mutual independence of the transmitted bits and the AWGN we have

$$\begin{aligned}
E\left[\boldsymbol{r}(n)\boldsymbol{r}^\top(n)\right] &= \sum_{i=1}^{N} p_i \left(\bar{\boldsymbol{s}}_{2i-2}\bar{\boldsymbol{s}}_{2i-2}^\top + \bar{\boldsymbol{s}}_{2i-1}\bar{\boldsymbol{s}}_{2i-1}^\top \right) + \sigma^2 \boldsymbol{I} \\
&= \sum_{i=1}^{2N+1} \bar{p}_i \bar{\boldsymbol{s}}_i \bar{\boldsymbol{s}}_i^\top + \sigma^2 \boldsymbol{I} \\
&= \bar{\boldsymbol{A}} + \sigma^2 \boldsymbol{I}
\end{aligned} \tag{40}$$

where $\bar{p}_{2i-2}$ and $\bar{p}_{2i-1}$ are equal to p_i since they denote the received power levels of the two synchronous users representing the asynchronous user i.

If the stochastic iteration (13) is used for the asynchronous system described in this section, the received filter of the desired user will converge to the point $\bar{\boldsymbol{c}}^*$ where

$$\bar{\boldsymbol{A}}\bar{\boldsymbol{c}}^* = \boldsymbol{s}_1 \tag{41}$$

We define a $(2N+1) \times G$ matrix $\bar{\boldsymbol{S}}$ with its ith row being $\bar{\boldsymbol{s}}_i$. Thus, $\bar{\boldsymbol{S}}$ contains the signature

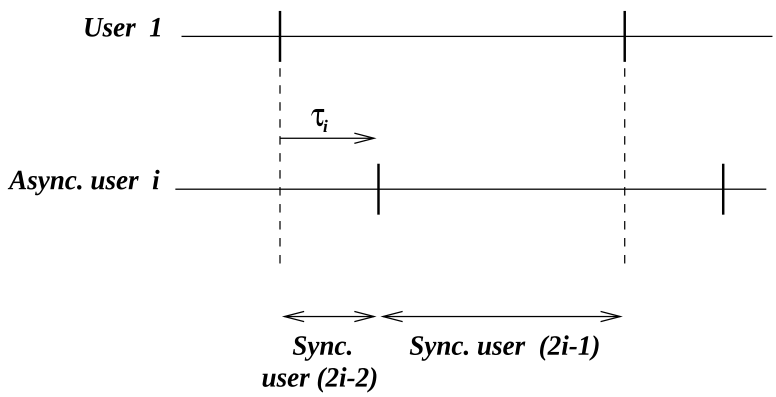

Figure 1: Asynchronous CDMA model.

sequences of the equivalent synchronous system with $(2N+1)$ users. Equation (41) is equivalent to

$$\bar{\boldsymbol{S}}^\top \bar{\boldsymbol{\Lambda}} \bar{\boldsymbol{S}} \bar{\boldsymbol{c}}^* = \boldsymbol{s}_1 \tag{42}$$

where $\bar{\boldsymbol{\Lambda}}$ is a $(2N+1) \times (2N+1)$ dimensional diagonal matrix with $\bar{\boldsymbol{\Lambda}}_{ii} = \bar{p}_i$. By Lemma 1 the solution of (42) is equal to the solution of

$$\bar{\boldsymbol{S}} \bar{\boldsymbol{c}}^* = \boldsymbol{e}_1 \tag{43}$$

if the rows of $\bar{\boldsymbol{S}}$, equivalently the modified signature sequences, are linearly independent. This condition is the *Linear Independence Assumption (LIA)* of [7] which requires $(2N+1) \leq G$. In this paper, for the asynchronous channel, we will assume that LIA is in effect as was done in [7].

The convergence point $\bar{\boldsymbol{c}}^*$ is the *one-shot decorrelator* [2,7] which is the decorrelating receiver filter when only one bit of the desired user is considered. The one-shot decorrelator has lower complexity than asynchronous decorrelating detectors which take into account the dependence of the received signal in different bit intervals. However, the one-shot decorrelating detector does suffer some loss in performance.

6 Simulation Results

In the simulations we consider a synchronous CDMA system with $N = 6$ users using randomly generated signature sequences with a processing gain of $G = 31$. Powers of the users are assigned so that the final signal to interference ratio (SIR) of all users will be 20 dB. Thus, the power of the ith user is found by $p_i = 100\sigma^2[\boldsymbol{\Gamma}^{-1}]_{ii}$ where $\boldsymbol{\Gamma} = \boldsymbol{S}\boldsymbol{S}^\top$ is the cross correlation matrix. The algorithm is run for 100 times and averaged results are plotted in the figures. In Figure 2, we plot the averaged normalized squared error (NSE) of the filter coefficients of the desired (first) user versus iteration index n, where NSE at iteration n is defined as

$$\mathrm{NSE}(n) = \frac{\|\boldsymbol{c}(n) - \boldsymbol{c}^*\|^2}{\|\boldsymbol{c}^*\|^2} \tag{44}$$

The curves in Figure 2 correspond to the blind adaptive decorrelating detector algorithms with fixed step size for different step size values and that with a time dependent step size sequence of the form of $a_n = 1/(n + n_0)$ with $n_0 = 5$. We observe that if the step size μ is too large, then the algorithm does not converge; see increasing NSE curve for $\mu = 0.1$ in Figure 2. We also observe that for large step size values the convergence rate is fast but the limiting NSE is large too; compare the curves corresponding to $\mu = 0.01$ and $\mu = 0.001$ in Figure 2.

Averaged (over 100 runs) SIR of the desired (first) user is plotted in Figure 3 for the same system. At iteration n, the SIR of the desired user is calculated as

$$\mathrm{SIR}(n) = \frac{p_1(\boldsymbol{s}_1^\top \boldsymbol{c}(n))^2}{\sum_{j\neq 1} p_j(\boldsymbol{s}_j^\top \boldsymbol{c}(n))^2 + \sigma^2(\boldsymbol{c}^\top(n)\boldsymbol{c}(n))} \tag{45}$$

We observe that the convergence of the SIRs to the *target* SIR (which is 20 dB in this experiment) is quantitatively similar to the convergence of the NSE to zero.

The blind adaptive decorrelating detector of [10] is also implemented for the same system and the SIR of the desired user is plotted as a function of iteration index in Figure 4 for different values of the forgetting factor, β (see [13]). In comparing Figures 3 and 4, we see that for this instance the blind adaptive decorrelating detector in this work outperforms that of [10].

7 Conclusion

In this paper we presented an iterative and distributed blind adaptive decorrelating detector algorithm which is based on the observation of the available statistics, and studied its convergence. The update equation of the algorithm needs the observation of the chip sampled input signal before the receiver filter and the output of the filter with the current filter coefficients. For the implementation of the algorithm to construct the decorrelating receiver filter of a user only two parameters are needed to be known: the user's signature sequence and the variance of the additive white Gaussian receiver noise.

We studied the convergence of the proposed algorithm both for a fixed step size sequence and for a time varying step size sequence. For the first case we developed the conditions of having lower and upper bounds on the MSE and showed that as the step size goes to zero the algorithm converges in the MSE. For the second case we directly proved the convergence in the MSE.

A Additional Proofs

Proof: Lemma 1 We need to show that if $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N$ are linearly independent, then (i) any solution of (2) is also a solution of (6), and (ii) any solution of (6) is also a solution of (2). Let $\hat{\boldsymbol{c}}$ be a solution of (2). Then $\boldsymbol{S}\hat{\boldsymbol{c}} = \alpha \boldsymbol{e}_1$. Multiplying both sides by first $\boldsymbol{\Lambda}$ and then with $\boldsymbol{S}^\top$ yields, $\boldsymbol{S}^\top \boldsymbol{\Lambda} \boldsymbol{S}\hat{\boldsymbol{c}} = \alpha p_1 \boldsymbol{s}_1$, and the proof of the first part is complete. Now let $\check{\boldsymbol{c}}$ be a solution of (6). Then $\boldsymbol{S}^\top \boldsymbol{\Lambda} \boldsymbol{S}\check{\boldsymbol{c}} = \alpha p_1 \boldsymbol{s}_1$ and equivalently

$$\boldsymbol{S}^\top \left(\boldsymbol{\Lambda} \boldsymbol{S} \check{\boldsymbol{c}} - \alpha p_1 \boldsymbol{e}_1 \right) = \boldsymbol{0} \tag{46}$$

Note that any vector multiplying $\boldsymbol{S}^\top$ yields a linear superposition of the columns of $\boldsymbol{S}^\top$, i.e., $\boldsymbol{S}^\top \boldsymbol{a} = \sum_{i=1}^N a_i \boldsymbol{s}_i$, for any vector $\boldsymbol{a}$. Combined with the independence of $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_N$, (46) implies

$$\boldsymbol{\Lambda} \boldsymbol{S} \check{\boldsymbol{c}} - \alpha p_1 \boldsymbol{e}_1 = \boldsymbol{0} \tag{47}$$

Multiplying both sides of (47) with $\boldsymbol{\Lambda}^{-1}$ yields $\boldsymbol{S}\check{\boldsymbol{c}} = \alpha \boldsymbol{e}_1$, completing the second part of the proof. □

Proof: Lemma 2 First note that $\boldsymbol{A}$ is a positive semidefinite matrix and its eigenvalues are nonnegative. To be specific, since the rank of $\boldsymbol{A}$ is N which is typically much less than G, $(G - N)$ eigenvalues of $\boldsymbol{A}$ are zero and the remaining N eigenvalues are positive. Since $\boldsymbol{A}\boldsymbol{x} \in \mathcal{L}$ by (8), eigenvectors of $\boldsymbol{A}$ are either completely in $\mathcal{L}$ with positive eigenvectors, or completely out of $\mathcal{L}$ (meaning that their projections onto $\mathcal{L}$ are zero) with zero eigenvalues. In this case Lemma 2 is a simple result of Rayleigh quotient [14] which states that for any symmetric matrix $\boldsymbol{M}$ and vector $\boldsymbol{x}$,

$$\lambda_{\min}^M \|\boldsymbol{x}\|^2 \leq \boldsymbol{x}^\top \boldsymbol{M} \boldsymbol{x} \leq \lambda_{\max}^M \|\boldsymbol{x}\|^2 \tag{48}$$

where $\lambda_{\min}^M$ and $\lambda_{\max}^M$ are the smallest and the largest eigenvalues of $\boldsymbol{M}$. Note that in (20) k_1 and k_0 are equal to the largest and the smallest non-zero eigenvalues of $\boldsymbol{A}$. □

Proof: Lemma 3 Using (15) and noting that $\boldsymbol{\epsilon}(n) = \boldsymbol{c}(n) - \boldsymbol{c}^*$, we observe that

$$E\left[\|\boldsymbol{\eta}(n)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}\right] = (\boldsymbol{\epsilon} + \boldsymbol{c}^*)^\top E\left[\left(\boldsymbol{r}\boldsymbol{r}^\top - (\boldsymbol{A} + \sigma^2 \boldsymbol{I})\right)^\top \left(\boldsymbol{r}\boldsymbol{r}^\top - (\boldsymbol{A} + \sigma^2 \boldsymbol{I})\right)\right] (\boldsymbol{\epsilon} + \boldsymbol{c}^*) \tag{49}$$

Note that $E\left[\boldsymbol{r}\boldsymbol{r}^\top\right] = \boldsymbol{A} + \sigma^2\boldsymbol{I}$ from (11). Applying this result to (49) yields

$$E\left[\|\boldsymbol{\eta}(n)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon}\right] = (\boldsymbol{\epsilon} + \boldsymbol{c}^*)^\top \left\{E\left[\boldsymbol{r}\boldsymbol{r}^\top\boldsymbol{r}\boldsymbol{r}^\top\right] - \left(\boldsymbol{A} + \sigma^2\boldsymbol{I}\right)^2\right\}(\boldsymbol{\epsilon} + \boldsymbol{c}^*) \tag{50}$$

First we will evaluate $E\left[\boldsymbol{r}\boldsymbol{r}^\top\boldsymbol{r}\boldsymbol{r}^\top\right]$ below. Using (10), we note that

$$\boldsymbol{r}\boldsymbol{r}^\top = \sum_k\sum_l \sqrt{p_k}\sqrt{p_l}a_k a_l \boldsymbol{s}_k\boldsymbol{s}_l^\top + \left(\sum_k \sqrt{p_k}a_k\boldsymbol{s}_k\right)\boldsymbol{n}^\top + \boldsymbol{n}\left(\sum_k \sqrt{p_k}a_k\boldsymbol{s}_k^\top\right) + \boldsymbol{n}\boldsymbol{n}^\top \tag{51}$$

Therefore,

$$\begin{aligned}
\boldsymbol{r}\boldsymbol{r}^\top\boldsymbol{r}\boldsymbol{r}^\top \;=\;& \sum_i\sum_j\sum_k\sum_l \sqrt{p_i}\sqrt{p_j}\sqrt{p_k}\sqrt{p_l}a_i a_j a_k a_l \boldsymbol{s}_i\boldsymbol{s}_j^\top\boldsymbol{s}_k\boldsymbol{s}_l^\top \\
+& \sum_i\sum_j\sum_k \sqrt{p_i}\sqrt{p_j}\sqrt{p_k}a_i a_j a_k \boldsymbol{s}_i\boldsymbol{s}_j^\top\boldsymbol{s}_k\boldsymbol{n}^\top + \sum_i\sum_j\sum_k \sqrt{p_i}\sqrt{p_j}\sqrt{p_k}a_i a_j a_k \boldsymbol{s}_i\boldsymbol{s}_j^\top\boldsymbol{n}\boldsymbol{s}_k^\top \\
+& \sum_i\sum_j \sqrt{p_i}\sqrt{p_j}a_i a_j \boldsymbol{s}_i\boldsymbol{s}_j^\top\boldsymbol{n}\boldsymbol{n}^\top + \sum_i\sum_j\sum_k \sqrt{p_i}\sqrt{p_j}\sqrt{p_k}a_i a_j a_k \boldsymbol{s}_i\boldsymbol{n}^\top\boldsymbol{s}_j\boldsymbol{s}_k^\top \\
+& \sum_i\sum_j \sqrt{p_i}\sqrt{p_j}a_i a_j \boldsymbol{s}_i\boldsymbol{n}^\top\boldsymbol{s}_j\boldsymbol{n}^\top + \sum_i\sum_j \sqrt{p_i}\sqrt{p_j}a_i a_j \boldsymbol{s}_i\boldsymbol{n}^\top\boldsymbol{n}\boldsymbol{s}_j + \sum_i \sqrt{p_i}a_i\boldsymbol{s}_i\boldsymbol{n}^\top\boldsymbol{n}\boldsymbol{n}^\top \\
+& \sum_i\sum_j\sum_k \sqrt{p_i}\sqrt{p_j}\sqrt{p_k}a_i a_j a_k \boldsymbol{n}\boldsymbol{s}_i^\top\boldsymbol{s}_j\boldsymbol{s}_k^\top + \sum_i\sum_j \sqrt{p_i}\sqrt{p_j}a_i a_j \boldsymbol{n}\boldsymbol{s}_i^\top\boldsymbol{s}_j\boldsymbol{n}^\top \\
+& \sum_i\sum_j \sqrt{p_i}\sqrt{p_j}a_i a_j \boldsymbol{n}\boldsymbol{s}_i^\top\boldsymbol{n}\boldsymbol{s}_j^\top + \sum_i \sqrt{p_i}a_i\boldsymbol{n}\boldsymbol{s}_i^\top\boldsymbol{n}\boldsymbol{n}^\top + \boldsymbol{n}\boldsymbol{n}^\top\sum_i\sum_j \sqrt{p_i}\sqrt{p_j}a_i a_j\boldsymbol{s}_i\boldsymbol{s}_j^\top \\
+& \boldsymbol{n}\boldsymbol{n}^\top\sum_i \sqrt{p_i}a_i\boldsymbol{s}_i\boldsymbol{n}^\top + \boldsymbol{n}\boldsymbol{n}^\top\boldsymbol{n}\sum_i \sqrt{p_i}a_i\boldsymbol{s}_i^\top + \boldsymbol{n}\boldsymbol{n}^\top\boldsymbol{n}\boldsymbol{n}^\top
\end{aligned} \tag{52}$$

Let T_i denote the ith term on the right hand side of (52). Note that

$$\begin{aligned}
E[T_i] &= 0 \qquad i = 2, 3, 5, 8, 9, 12, 14, 15. \\
E[T_i] &= \sigma^2\boldsymbol{A} \qquad i = 4, 6, 11, 13. \\
E[T_7] &= \sigma^2 G\boldsymbol{A} \\
E[T_{10}] &= \sigma^2 p_T\boldsymbol{I} \\
E[T_{16}] &= (G+1)\sigma^4\boldsymbol{I} \\
E[T_1] &= \boldsymbol{A}^2 + \boldsymbol{C}
\end{aligned} \tag{53}$$

where $p_T > 0$ is the total receiver power, i.e., $p_T = \sum_i p_i$, and $\boldsymbol{C}$ is given as

$$\boldsymbol{C} = \sum_i\sum_{j\neq i} p_i p_j \boldsymbol{s}_i\boldsymbol{s}_i^\top + \sum_i\sum_{j\neq i} p_i p_j \boldsymbol{s}_i\boldsymbol{s}_i^\top\boldsymbol{s}_j\boldsymbol{s}_j^\top \tag{54}$$

Note that since

$$\boldsymbol{C} = \frac{1}{2} \sum_{i} \sum_{j \neq i} p_i p_j \left(\boldsymbol{s}_i \boldsymbol{s}_i^\top + \boldsymbol{s}_j \boldsymbol{s}_j^\top \right)^2 \tag{55}$$

$\boldsymbol{C}$ is positive semidefinite and symmetric. Combining the results in (53) yields

$$E \left[\boldsymbol{r}\boldsymbol{r}^\top \boldsymbol{r}\boldsymbol{r}^\top \right] = \boldsymbol{A}^2 + \boldsymbol{C} + (G+4)\sigma^2 \boldsymbol{A} + \sigma^2 p_T \boldsymbol{I} + (G+1)\sigma^4 \boldsymbol{I} \tag{56}$$

Inserting (56) into (50) and defining

$$\boldsymbol{B} = \boldsymbol{C} + (G+2)\sigma^2 \boldsymbol{A} + \sigma^2 p_T \boldsymbol{I} + G\sigma^4 \boldsymbol{I} \tag{57}$$

yields

$$E \left[\|\boldsymbol{\eta}(n)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon} \right] = (\boldsymbol{\epsilon} + \boldsymbol{c}^*)^\top \boldsymbol{B} (\boldsymbol{\epsilon} + \boldsymbol{c}^*) \tag{58}$$

Note that $\boldsymbol{B}$ is a positive definite matrix since $\boldsymbol{A}$ and $\boldsymbol{C}$ are positive semidefinite and $\sigma > 0$. Using Rayleigh quotient [14], and denoting the smallest and largest eigenvalues of $\boldsymbol{B}$ by $\lambda_{\min}$ and $\lambda_{\max}$, we obtain

$$\lambda_{\min} \|\boldsymbol{\epsilon} + \boldsymbol{c}^*\|^2 \leq E \left[\|\boldsymbol{\eta}(n)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon} \right] \leq \lambda_{\max} \|\boldsymbol{\epsilon} + \boldsymbol{c}^*\|^2 \tag{59}$$

Using the fact that $\|\boldsymbol{a} + \boldsymbol{b}\|^2 \leq 2(\|\boldsymbol{a}\|^2 + \|\boldsymbol{b}\|^2)$ for any two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, we obtain the desired result

$$0 \leq E \left[\|\boldsymbol{\eta}(n)\|^2 \mid \boldsymbol{\epsilon}(n) = \boldsymbol{\epsilon} \right] \leq c_0 + c_1 \|\boldsymbol{\epsilon}\|^2 \tag{60}$$

where $c_0 = 2\lambda_{\max} \|\boldsymbol{c}^*\|^2$ and $\boldsymbol{c}_1 = 2\lambda_{\max}$. □

Proof: Lemma 4 We need only the upper bound given in (30). Replacing μ with a_n we get

$$b_{n+1} \leq \left(1 - 2a_n k_0 + (k_1^2 + c_1) a_n^2\right) b_n + c_0 a_n^2 \tag{61}$$

We will follow Sakrison's approach [15, pages 60–61] in the following derivation. Since a_n is a monotonically decreasing sequence, there exists n_0 and δ, $0 < \delta < 2$, such that for $n \geq n_0$,

$$\left(1 - 2k_0 a_n + (k_1^2 + c_1) a_n^2\right) \leq (1 - (2 - \delta) k_0 a_n) \tag{62}$$

Furthermore, we can choose n_0 such that for $n \geq n_0$, we have $(1 - (2-\delta)k_0 a_n) > 0$. For $n \geq n_0$, the inequality (61) can be further bounded as,

$$b_{n+1} \leq (1 - (2 - \delta) k_0 a_n)\, b_n + a_n^2 c_0 \tag{63}$$

Starting at $n = n_0$ and executing the recursion repeatedly yields

$$b_n \leq b_{n_0}\beta_{n_0,n-1} + c_0 \sum_{k=n_0}^{n-1} a_k^2 \beta_{k+1,n-1} \tag{64}$$

where

$$\beta_{k,n} = \begin{cases} \prod_{j=k}^{n} [1 - a_j(2-\delta)k_0] & 0 \leq k \leq n \\ 1 & k > n \end{cases} \tag{65}$$

For $0 < x < 1$, we can use the inequality $\ln(1-x) \leq -x$ to show

$$\beta_{k,n} = \exp\left(\sum_{j=k}^{n} \ln[1 - a_j(2-\delta)k_0]\right) \leq \exp\left\{-(2-\delta)k_0 \sum_{j=k}^{n} a_j\right\} \tag{66}$$

By the first condition in (38) and the fact that $(2-\delta)k_0 > 0$, the exponent in the above equation diverges to negative infinity and we have,

$$\lim_{n\to\infty} \beta_{k,n-1} = 0 \tag{67}$$

This implies that $b_{n_0}\beta_{n_0,n-1}$ on the right hand side of (64) goes to zero as n goes to infinity. Now we will investigate the second term. Let $u(x)$ be the unit step function whose value is 1 for non-negative x and 0 otherwise. Then,

$$\begin{aligned} \lim_{n\to\infty} c_0 \sum_{k=n_0}^{n-1} a_k^2 \beta_{k+1,n-1} &= c_0 \lim_{n\to\infty} \sum_{k=n_0}^{\infty} a_k^2 \beta_{k+1,n-1} u(n-1-k) && (68) \\ &= c_0 \sum_{k=n_0}^{\infty} a_k^2 \lim_{n\to\infty} \beta_{k+1,n-1} u(n-1-k) && (69) \\ &= 0 && (70) \end{aligned}$$

since $\lim_{n\to\infty} \beta_{k,n-1} = 0$ from (67). We could exchange the limit and summation to obtain (69) from (68), because the sum on the right side of (68) is absolutely convergent. Finally, combining the results in (67) and (70) and the fact that b_n is a non-negative sequence, we obtain $\lim_{n\to\infty} b_n = 0$ which concludes the proof. □

References

[1] S. Verdú. Multiuser detection. *Advances in Statistical Signal Processing*, 2:369–409, 1993.

[2] S. Verdú. Minimum probability of error for asynchronous gaussian multiple-access channels. *IEEE Transactions on Information Theory*, 32:85–96, January 1986.

[3] R. Lupas and S. Verdú. Linear multiuser detectors for synchronous code-division multiple-access channels. *IEEE Transactions on Information Theory*, 35(1):123–136, January 1989.

[4] A. Duel-Hallen. Decorrelating decision-feedback multiuser detector for synchronous code-division multiple-access channels. *IEEE Transactions on Communications*, 41(2):285–290, February 1993.

[5] U. Madhow and M. L. Honig. MMSE interference suppression for direct-sequence spread-spectrum CDMA. *IEEE Transactions on Communications*, 42(12):3178–3188, December 1994.

[6] M. K. Varanasi and B. Aazhang. Multistage detection in asynchronous code-division multiple-access communications. *IEEE Transactions on Communications*, 38(4):509–519, April 1990.

[7] R. Lupas and S. Verdú. Near-far resistance of multiuser detectors in asynchronous channels. *IEEE Transactions on Communications*, 38(4):496–508, April 1990.

[8] M. Honig, U. Madhow, and S. Verdú. Blind adaptive multiuser detection. *IEEE Transactions on Information Theory*, 41(4):944–960, July 1995.

[9] D. S. Chen and S. Roy. An adaptive multiuser receiver for CDMA systems. *IEEE Journal on Selected Areas in Communications*, 12(5):808–816, June 1994.

[10] X. Wang and H. V. Poor. Blind adaptive interference suppression for CDMA communications based on eigenspace tracking. In *Conference on Information Sciences and Systems, CISS'97*, pages 468–473, March 1997.

[11] W. F. Stout. *Almost Sure Convergence*. Academic Press, 1974.

[12] L. Gyorfi. Adaptive linear procedures under general conditions. *IEEE Transactions on Information Theory*, IT-30(2):262–267, March 1984.

[13] B. Yang. Projection approximation subspace tracking. *IEEE Transactions on Signal Processing*, 43(1):95–107, January 1995.

[14] G. Strang. *Linear Algebra and Its Applications*. Saunders College Publishing, 1988. 3rd edition.

[15] D. J. Sakrison. Stochastic approximation : A recursive method for solving regression problems. *Advances in Communication Systems 2*, pages 51–106, 1966. A. V. Balakrishnan, Ed.

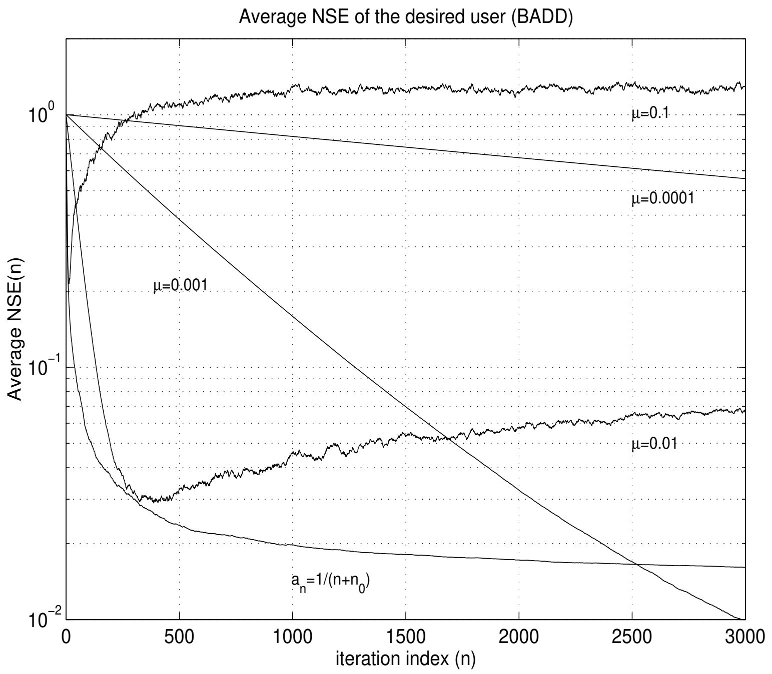

Figure 2: Averaged normalized squared error (NSE) of the desired user.

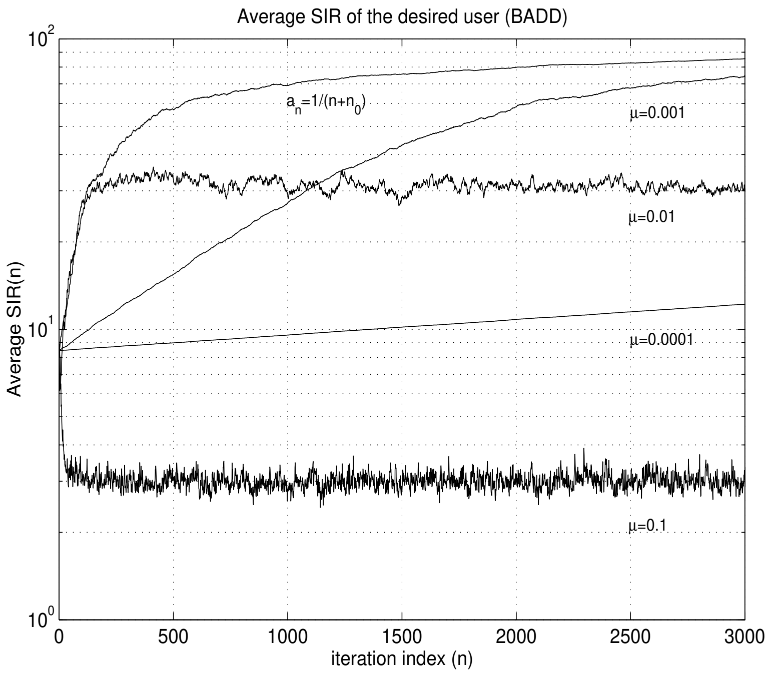

Figure 3: Averaged SIR of the desired user.

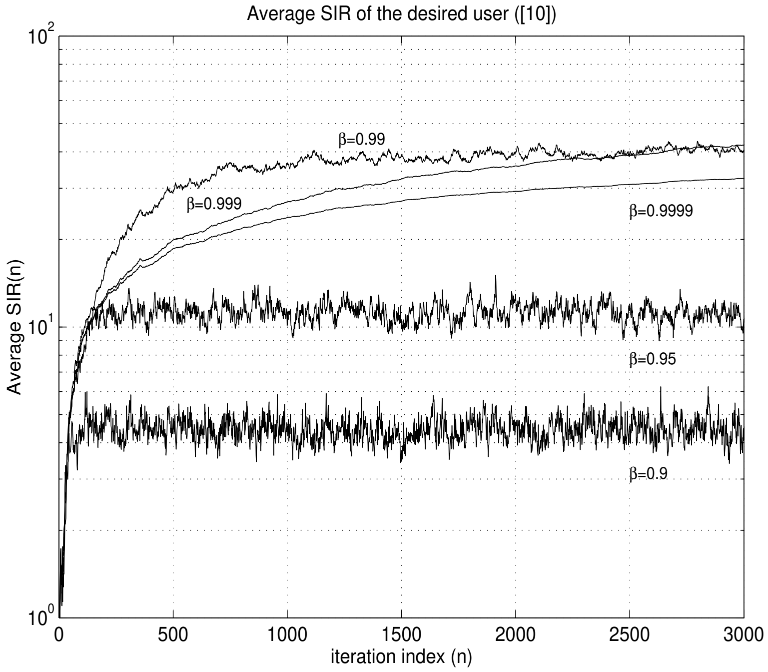

Figure 4: Averaged SIR of the desired user (algorithm of [10]).