A Stochastic Gradient-Based Decorrelation Algorithm with Applications to Multicarrier CDMA

A. Haimovich, Y. Bar-Ness and R. Manzo
Center for Communications and Signal Processing Research
Department of Electrical and Computer Engineering
New Jersey Institute of Technology
University Heights, Newark, NJ 07102
tel:(201)596-3534 fax:(201)596-8473

Abstract— A stochastic gradient-based decorrelation algorithm is suggested for separation of an unknown linear mixture of signals. It is shown that while the decorrelation algorithm is similar in cost to the LMS algorithm, its rate of convergence is significantly faster, making it more attractive for signal separation. Analysis of the decorrelator algorithm shows that the faster speed of convergence is a consequence of the eigenevalue spread associated with the decorrelation problem, which is smaller than the spread associated with the corresponding mean square problem. Operation of the algorithm is illustrated as an adaptive multiuser detector in a multiple carrier CDMA system.

INTRODUCTION

In this paper we address the problem of the recovery of unknown independent sources from observations of a linear mixture of the sources. We are concerned with a multiple input- multiple output system where each output is an unknown linear combination of the inputs. The following linear statistical model is assumed:

$$\mathbf{x} = \mathbf{M}\mathbf{b} + \mathbf{v}, \tag{1}$$

where $\mathbf{x}$, $\mathbf{b}$, and $\mathbf{v}$ are random vectors with values in $\mathcal{R}^N$. The mixture matrix $\mathbf{M}$ is unknown but is invertible. The problem is to estimate realizations of the vector $\mathbf{b}$ given observations of $\mathbf{x}$. This type of problem arises in numerous applications. For example, if the vector $\mathbf{b}$ consists of samples of a time series and $\mathbf{M}$ is Toeplitz triangular, the model in Eq. 1 represents a convolution. The recovery of $\mathbf{b}$ when only the model's output observations $\mathbf{x}$ are available is referred to as *blind signal separation*. In addition to the output observations, to find a solution to this problem it is necessary to have some information about the statistical properties of the components of $\mathbf{b}$. Examples of blind signal separation applications are the cancellation of cross-pol interference in dually polarized systems and multiuser detection in multiple access communications systems. The latter application will be used to illustrate the techniques considered in this paper.

One approach to signal separation is to view it as an interference cancellation problem. When a reference signal is available, the mean square error (MSE) between the output and the reference signal is minimized by the classical Wiener filter, which can be implemented using steepest descent algorithms such as the LMS and RLS.

A different approach is to treat the separation problem in its own right. Beginning with the early eighties a class of adaptive signal separators have been proposed that in effect estimate the inverse of the mixture matrix $\mathbf{M}^{-1}$, [1], [2]. Their research was mainly aimed at communications applications, such as cross-pol separation in dually polarized links. It is assumed that the signals to be separated are uncorrelated. This condition however, resulted in an indeterminate solution. Separation was shown, provided that "some knowledge" is available by which the signals can be discriminated (for example, slightly different spectral characteristics). Later, similar separator structures were suggested for implementation using neural networks [3]. They solved the indetermination problem by assuming that the signals to be recovered are independent. An independence test was approximated by way of decorrelation using nonlinear functions [4].

In this paper we analyze the performance of a stochastic gradient-based decorrelation algorithm for blind signal separation. The algorithm has a computational complexity comparable with the LMS algorithm, but is shown to be faster than the latter. The main contribution of the paper is to prove that the differences in the speed of convergence are a con-

[1] This work is partially supported by a grant from ROME Air Development Center (AFSC, Griffiss Air Force Base, NY, under contract F30602-93-C-0118.

sequence of the different eigenvalue spreads associated with the implementation of the two algorithms. While the problem statement and proofs are quite general, the specific application considered is a synchronous multiuser multicarrier CDMA communications system in a fading environment. In this application the decorrelator is used as a multiuser detector. The decorrelator has been previously suggested as a low complexity near-far resistant multiuser detector [5]. While some authors proposed non-adaptive and adaptive multiuser detectors incorporating the decorrelator [6], [7], in this paper we focus on the transient response of the adaptive decorrelator.

SYSTEM MODEL

The general system model is given in Eq. 1 where the received signal consists of a linear transformation of the transmitted data and additive noise. To be more specific and to lend credibility to the model, we consider a synchronous multiuser multicarrier CDMA communication system such as depicted in [8]. In multicarrier (MC)-CDMA a bit is transmitted on multiple subcarriers. The subcarrier waveforms are chosen as mutually orthogonal. Each of N users is assigned a unique code expressed by a set of $0/\pi$ offsets of the subcarriers phase. The main advantage of this type of modulation over conventional direct sequence spread spectrum is that each subcarrier is subject to frequency-nonselective fading, thus obviating the complex RAKE receiver. The signal transmitted by the n-th user for the i-th bit is given by:

$$s_n(t) = \sqrt{\xi_n} b_n(i) \sum_{m=1}^{M} c_n(m) p(t) \cos\left(\omega_c t + \frac{m-1}{T}\Omega t\right), \quad (2)$$

where ξ_n is the waveform's energy. The information symbols are assumed binary $b_n(i) \in \{-1, 1\}$, independent and equiprobable. The signature code sequence for the n-th user is given by $c_n(m) = \pm 1/\sqrt{M}$, $m = 1, \ldots, M$, where M is the number of subcarriers. $p(t) = 2/\sqrt{T}$, $0 \leq t < T$ is the pulse shape and T is the symbol duration. To ensure orthogonality between waveforms, the frequency offset Ω is chosen to be an integer multiple of $2\pi T$.

The n-th user signal after coherent amplitude detection at the receiver is given by:

$$r_n(t) = \sqrt{\xi_n} b_n(i) \sum_{m=1}^{M} c_n(m) h_{nm}, \quad (3)$$

where h_{nm} is the amplitude channel response at the m-th subcarrier. In vector notation,

$$\mathbf{r}_n(i) = \sqrt{\xi_n} b_n(i) \mathbf{H}_n \mathbf{c}_n, \quad (4)$$

where $\mathbf{r}_n(i)$ is the $M \times 1$ vector $\mathbf{H}_n$ is a diagonal matrix of the channel coefficients, and $\mathbf{c}_n^T = [c_n(1), \ldots, c_n(M)]$ is the n-th user code sequence vector. For brevity of notation explicit time dependency is dropped. Collecting the contributions of all the users and the additive noise, we have at the output of the coherent detectors:

$$\mathbf{q} = \overline{\mathbf{H}} \mathbf{E} \mathbf{b} + \mathbf{z}_M, \quad (5)$$

where $\mathbf{E} = \text{diag}\left(\sqrt{\xi_n}\right)$, $n = 1, \ldots, N$, is a diagonal matrix and $\overline{\mathbf{H}} = [\mathbf{H}_1 \mathbf{c}_1, \ldots, \mathbf{H}_N \mathbf{c}_N]$. $\mathbf{b}$ is the vector of information bits, and the components of the noise vector $\mathbf{z}_M$ are assumed i.i.d zero-mean gaussian with variance σ^2. The signal after cross-correlation with the code of the j-th user can be written,

$$x_j = \mathbf{c}_j^T \overline{\mathbf{H}} \mathbf{E} \mathbf{b} + \mathbf{c}_j^T \mathbf{z}_M, \quad (6)$$

and for all the users in vector form,

$$\mathbf{x} = \mathbf{A} \mathbf{E} \mathbf{b} + \mathbf{v}, \quad (7)$$

where $\mathbf{A} = \mathbf{C}^T \overline{\mathbf{H}}$ is the mixture matrix, and $\mathbf{v}(i)$ is the noise vector with the j-th component equal to $\mathbf{c}_j^T \mathbf{z}_M$. Since the channel coefficients are unknown, the matrix $\mathbf{A}$ is unknown. It can be readily shown that the noise covariance matrix is given by $E\left[\mathbf{v}\mathbf{v}^T\right] = \sigma^2 \mathbf{C}^T \mathbf{C}$. Notice that the MC-CDMA model in Eq. 7 is the same as the general model assumed in Eq. 1. Direct detection of the elements of vector $\mathbf{x}$ is impeded by the presence of the cochannel interference. Therefore, the goal is to add a receiver stage that provides signal separation through cochannel interference cancellation.

SIGNAL SEPARATION CRITERIA

Two signal separation criteria are considered: separation through minimization of the Mean Square Error (MSE) and separation through signal decorrelation. The criteria are used to develop control algorithms for the network weights. The network weights are represented by a matrix $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_N]$, where $\mathbf{w}_n$ is the $N \times 1$ weight vector applied to the network input vector $\mathbf{x}$ to obtain the output for the n-th user, $y_n = \mathbf{w}_n^T \mathbf{x}$. An estimate of the data bits is then obtained from $\widehat{\mathbf{b}} = \text{sgn}\, \mathbf{y}$, where $\mathbf{y} = \mathbf{W}^T \mathbf{x}$.

A. MSE Separation

The MSE signal separator minimizes the mean squared error between its output and a reference signal. Typically, the reference is initially supplied by a training signal. When the adaptive weights converge and the errors with respect to the training signal are small, the detector is switched to operate in decision directed mode, and the reference signal is supplied by the estimated symbol. For the MSE signal separator, the network weights are given by,

$$\begin{aligned} \mathbf{W}_m &= \arg\min_w E\left[\left\|\widehat{\mathbf{b}} - \mathbf{W}^T\mathbf{x}\right\|^2\right] \\ &= \mathbf{R}_x^{-1}\mathbf{R}_{\widehat{b}x}, \end{aligned} \tag{8}$$

where $\mathbf{R}_x = E\left[\mathbf{x}\mathbf{x}^T\right]$ is the correlation matrix of the inputs, and $\mathbf{R}_{\widehat{b}x} = E\left[\mathbf{x}\widehat{\mathbf{b}}^T\right]$ is the cross-correlation matrix between the estimated bits and the input. We will show that the application of this weight matrix leads to signal separation in the high SNR case. Neglecting the noise contribution we have:

$$\mathbf{R}_x \simeq \mathbf{A}\mathbf{E}^2\mathbf{A}^T, \tag{9}$$

where we used $E\left[\mathbf{b}\mathbf{b}^T\right] = \mathbf{I}$, since the signals are assumed independent between users. The cross-correlation matrix is:

$$\mathbf{R}_{\widehat{b}x} = p\mathbf{A}\mathbf{E}, \tag{10}$$

where one can easily prove that $p = 1 - 2P\left[\widehat{b}_i \neq b_i\right] = 1 - 2P_{ei}$, and

$$E\left[\widehat{b}_i b_j\right] = \begin{cases} 1 - 2P_{ei} & i = j \\ 0 & i \neq j. \end{cases} \tag{11}$$

Then the weights are given by $\mathbf{W}_m = p\mathbf{A}^{-T}\mathbf{E}^{-1}$. When the effect of the noise can be neglected, this linear transformation applied to the input vector recovers the transmitted signals vis. $\mathbf{y} = \mathbf{W}_m^T\mathbf{x} = p\mathbf{b} + \mathbf{W}_m^T\mathbf{v} \simeq p\mathbf{b}$.

B. Signal Separation by Decorrelation

The decorrelator seeks to decorrelate each output y_n from all the other users' bit estimates. This criterion can be expressed as:

$$\mathbf{W}_d = \arg\left\{E\left[\widehat{\mathbf{b}}\mathbf{y}^T\right] = \mathbf{\Lambda}\right\}, \tag{12}$$

where $\mathbf{\Lambda}$ is a diagonal matrix with positive entries. Signal separation by decorrelation is different from the MSE separator. The MSE weight vector for any user is derived using previous symbol estimates of that user as a reference signal. In contrast, the decorrelator can be interpreted as using reference signals derived from all other users bit estimates. Suggesting an operation where processing for each user is helped by the other users, this structure has been referred to as *bootstrapped,* [9].

We now develop an expression for the weight matrix solution to the decorrelation criterion of Eq. 12 and then show that this weight vector indeed performs signal separation. Again assuming negligible noise and using the results of Eq. 11, we have:

$$\begin{aligned} E\left[\widehat{\mathbf{b}}\mathbf{y}^T\right] &\simeq E\left[\widehat{\mathbf{b}}\mathbf{b}^T\mathbf{E}\mathbf{A}^T\mathbf{W}\right] \\ &= p\mathbf{E}\mathbf{A}^T\mathbf{W}, \end{aligned} \tag{13}$$

and from Eq. 12 the decorrelation criterion is equivalent to the condition $p\mathbf{E}\mathbf{A}^T\mathbf{W} = \mathbf{\Lambda}$. It results in the decorrelation weight matrix solution given by:

$$\mathbf{W}_d = \mathbf{A}^{-T}\mathbf{E}^{-1}\mathbf{\Lambda}, \tag{14}$$

where the constant p^{-1} has been embedded in the matrix $\mathbf{\Lambda}$. Indeed when this weight matrix is applied to the noiseless input vector $\mathbf{x}$ it results in signal separation vis. $\mathbf{y} = \mathbf{W}_d^T\mathbf{x} \simeq \mathbf{W}_d^T\mathbf{A}\mathbf{E}\mathbf{b} = \mathbf{\Lambda}\mathbf{b}$. Note that while the amplitude of the separated signals cannot be determined by this procedure, the bit detection is not affected since $\text{sign}(\mathbf{\Lambda}\mathbf{b}) = \text{sign}(\mathbf{b})$.

Adaptive Decorrelation

Adaptive signal separation algorithms can be directly derived from the MSE and decorrelation criteria formulated in Eqs. 8 and 12, respectively. The MSE criterion can be implemented using the well known LMS adaptation. In this section we formulate, analyze and compare the decorrelation algorithm with the LMS algorithm. The LMS update of the signal separation network weights is expressed by the relations:

$$\mathbf{W}(k+1) = [\mathbf{I} - \mu\mathbf{x}(k)\mathbf{x}^T(k)]\mathbf{W}(k) + \mu\mathbf{x}(k)\hat{\mathbf{b}}(k). \tag{15}$$

The convergence analysis of the LMS algorithm is well known. Under a set of assumptions often referred to as independence assumptions, a necessary and sufficient condition for convergence is that the step size parameter meets the condition $0 < \mu < 2/\lambda_{\max}(\mathbf{R}_x)$. It is also well known that the convergence speed of the LMS algorithm is determined by eigenvalue spread $\chi = \lambda_{\max}(\mathbf{R}_x)/\lambda_{\min}(\mathbf{R}_x)$.

A stochastic gradient-based decorrelation algorithm can be formulated directly from Eq. 12. It seeks to null the instantaneous cross-correlation between the outputs of the signal separation networks and the detected symbols:

$$\begin{aligned} \mathbf{W}(k+1) &= \mathbf{W}(k) - \mu[\widehat{\mathbf{b}}(k)\mathbf{y}^T(k) \\ &\quad - \text{diag}(\widehat{\mathbf{b}}(k)\mathbf{y}^T(k))] \\ \text{diag}(\mathbf{W}(k+1)) &= \mathbf{I}, \end{aligned} \tag{16}$$

where the operator $\text{diag}(\cdot)$ generates a diagonal matrix of the elements on the main diagonal of the argument. Note that the adaptive decorrelator has a computational complexity similar to that of the LMS.

A. Analysis of the Adaptive Decorrelator

The adaptive decorrelator has been observed to converge faster than the conventional LMS error algorithm [9]. We prove that this is due to the smaller eigenvalue spread of the former. For the following analysis we assume that the mixture matrix $\mathbf{A}$ is symmetric. This assumption is adequate for the downlink of a synchronous MC-CDMA system. Each of the mobiles receives all the signals transmitted by the base station, and all signals experience the same channel, i.e., $\mathbf{H}_n = \mathbf{H}$. With this condition the mixture matrix is given by $\mathbf{A} = \mathbf{C}^T\mathbf{H}\mathbf{C}$, which is clearly symmetric.

To simplify the notation and without loss of generality, consider the weight vector $\mathbf{w}_1$ which decorrelates the output for the first user y_1 from the other users detected symbols $\widehat{b}_2, \ldots, \widehat{b}_N$. Define the vector $\widehat{\mathbf{B}} = \left[0, \widehat{b}_2, \ldots, \widehat{b}_N\right]^T = \mathbf{U}_1\widehat{\mathbf{b}}$, where $\mathbf{U}_1$ is a unity matrix with its first element zero'ed. We can rewrite Eq. 16 for $\mathbf{w}_1$:

$$\begin{aligned} \mathbf{w}_1(k+1) &= \mathbf{w}_1(k) - \mu y_1(k)\widehat{\mathbf{B}}(k) \\ &= \left[\mathbf{I} - \mu\widehat{\mathbf{B}}(k)\mathbf{x}^T(k)\right]\mathbf{w}_1(k). \end{aligned} \quad (17)$$

We assume that the independence assumptions hold as they do for the LMS algorithm. We can write:

$$\begin{aligned} \mathrm{E}[\mathbf{w}_1(k+1)] &= \left[\mathbf{I} - \mu\mathbf{R}_{\widehat{B}x}\right]\mathrm{E}[\mathbf{w}_1(k)] \\ &= [\mathbf{I} - \mu\mathbf{U}_1\mathbf{A}\mathbf{E}]\,\mathrm{E}[\mathbf{w}_1(k)], \end{aligned} \quad (18)$$

where we used $\mathbf{R}_{\widehat{B}x} = E\left[\widehat{\mathbf{B}}(k)\mathbf{x}^T(k)\right] = \mathbf{U}_1\mathbf{R}_{\widehat{b}x}$ and we used the relation,

$$\begin{aligned} \mathbf{U}_1\mathbf{R}_{\widehat{b}x} &= \mathbf{U}_1 E\left[(\mathbf{A}\mathbf{E}\mathbf{b} + \mathbf{v})\,\mathbf{b}^T\right] \\ &= \mathbf{U}_1\mathbf{A}\mathbf{E}. \end{aligned} \quad (19)$$

Define the following matrix and vector partitions: $\mathbf{w}_1^T = [w_1 \vdots \overline{\mathbf{w}}^T]$, $\widehat{\mathbf{B}}^T = (0 \vdots \overline{\mathbf{B}}^T)$, $\mathbf{x}^T = [x_1 \vdots \overline{\mathbf{x}}^T]$ and $I = \mathrm{diag}(1, \overline{\mathbf{I}})$. Then the adaptive decorrelator in Eq. 17 can be rewritten as follows:

$$\begin{aligned} E[\overline{\mathbf{w}}(k+1)] &= \left[\overline{\mathbf{I}} - \mu\mathbf{R}_{\overline{B}\overline{x}}\right]E[\overline{\mathbf{w}}(k)] \\ &\quad -\mu E\left[\overline{\mathbf{B}}(k)\,x_1(k)\,w_1(k)\right] \end{aligned} \quad (20)$$

where $\mathbf{R}_{\overline{B}\overline{x}} = E\left[\overline{\mathbf{B}}\overline{\mathbf{x}}^T\right]$. Form this relation we deduce that the convergence properties of the adaptive decorrelator are controlled by the eigenvalues of $\mathbf{R}_{\overline{B}\overline{x}}$.

Therefore similar to the LMS algorithm for convergence of the decorrelation algorithm the condition $0 < \mu < 2 \,/\, \lambda_{\max}\left(\mathbf{R}_{\overline{B}\overline{x}}\right)$ has to be satisfied:

Proposition 1:

$$\lambda_{\max}\left(\mathbf{R}_{\widehat{b}x}\right) \geq \lambda_{\max}\left(\mathbf{R}_{\overline{B}\overline{x}}\right). \quad (21)$$

Proof: Let,

$$\mathbf{R}_{\widehat{b}x} = \begin{pmatrix} r_1 & \mathbf{r}_1^T \\ \mathbf{r}_1 & \mathbf{R}_{\overline{B}\overline{x}} \end{pmatrix} \quad (22)$$

Applying the Courant-Fisher Minimax theorem [10], $\lambda_{\max}\left(\mathbf{R}_{\widehat{b}x}\right) = \max_z \;\rho(z) = \frac{\mathbf{z}^T\mathbf{R}_{\widehat{b}x}\mathbf{z}}{\mathbf{z}^T\mathbf{z}}$. Using the matrix partition in Eq. 22 and maximizing $\rho(z)$ over the restriction $z_1 = 0$, where z_1 is the first component of the vector $\mathbf{z}$, we get

$$\lambda_{\max}\left(\mathbf{R}_{\widehat{b}x}\right) \geq \max_{\substack{z \\ z_1=0}} \;\rho(z) = \lambda_{\max}\left(\mathbf{R}_{\overline{B}\overline{x}}\right). \quad (23)$$

Proposition 2:

$$\lambda_{\min}\left(\mathbf{R}_{\widehat{b}x}\right) \leq \lambda_{\min}\left(\mathbf{R}_{\overline{B}\overline{x}}\right). \quad (24)$$

Proof: According to the Interlacing property which follows from the Courant-Fisher Minimax theorem [10], and noting that $\mathbf{R}_{\overline{B}\overline{x}}$ is the $(N-1)\times(N-1)$ leading principal submatrix of $\mathbf{R}_{\widehat{b}x}$, we have:

$$\lambda_{\min}\left(\mathbf{R}_{\widehat{b}x}\right) \leq \lambda_{N-1}\left(\mathbf{R}_{\overline{B}\overline{x}}\right) = \lambda_{\min}\left(\mathbf{R}_{\overline{B}\overline{x}}\right). \quad (25)$$

If the noise is negligible then $\mathbf{R}_x \simeq \mathbf{R}_{\widehat{b}x}^T\mathbf{R}_{\widehat{b}x}$. Since $\mathbf{R}_{\widehat{b}x}$ is symmetric, it follows that $\lambda_i(\mathbf{R}_x) = \lambda_i^2\left(\mathbf{R}_{\widehat{b}x}\right)$ for $1 \leq i \leq N$. In particular, we have for the eigenvalue spread,

$$\chi(\mathbf{R}_x) = \frac{\lambda_{\max}(\mathbf{R}_x)}{\lambda_{\min}(\mathbf{R}_x)} = \left(\frac{\lambda_{\max}\left(\mathbf{R}_{\widehat{b}x}\right)}{\lambda_{\min}\left(\mathbf{R}_{\widehat{b}x}\right)}\right)^2 = \chi^2\left(\mathbf{R}_{\widehat{b}x}\right). \quad (26)$$

Finally, using relation Eqs. 23 and 24,

$$\chi(\mathbf{R}_x) = \chi^2\left(\mathbf{R}_{\widehat{b}x}\right) \geq \chi\left(\mathbf{R}_{\overline{B}\overline{x}}\right) \quad (27)$$

which provides the explanation to why the decorrelator algorithm is faster than the LMS algorithm.

Simulation Results

The LMS and the decorrelator algorithms were compared in terms of the probability of error for detecting the transmitted data. At the output of the separator we can compute the probability of error as a function of the weight vector $\mathbf{w}_i^T$:

$$\begin{aligned} P_{ei} &= P\left(\widehat{b}_i \neq b_i \mid b_i\right) \\ &= 2^{1-N}\sum_{\substack{b\in\{-1,1\}^N \\ b_i=-1}} Q\left(\frac{t_{ii} - \sum_{j\neq i} t_{ij}b_j}{\sigma\sqrt{\mathbf{w}_i^T\mathbf{w}_i}}\right), \end{aligned}$$

where $Q(z) = \frac{1}{\sqrt{2\pi}}\int_z^{\infty} e^{-\frac{y^2}{2}}dy$, t_{nj} is the j-th element of the vector $\mathbf{t}_i = \mathbf{w}_i^T\mathbf{A}\mathbf{E}$.

The convergence of the algorithms was studied by simulations. The simulations consisted of an $N = 4$-user system, each using an $M = 15$ chip Gold code spreading sequence. The multipath fading channel was modeled as Rayleigh random variables. In the following figures the SNR is defined with respect to the first user, and the signal-to-signal ratio (SSR) is defined as the ratio of the bit energy of the first user to the bit energy of any of the other users. The probability of error was evaluated with respect to the first user. Each curve is the result of 50 Monte Carlo runs over different channels. Figure 1 shows the average learning curves of the probability of error for SNR = 8 dB and SSR = -5 dB. The higher rate of convergence of the decorrelator is evident.

To estimate the convergence region of each algorithm we took the approach suggested in [11]. A figure of merit γ is defined that relates the initial and final probabilities of error, P_{ei}, and P_{ef}, as follows:

$$\gamma = 1 - \frac{P_{ef}}{P_{ei}} \tag{28}$$

when $P_{ef} \ll P_{ei}$, $\gamma \approx 1$. Note that $\gamma = 0$ corresponds to no convergence, is indicated by $P_{ef} = P_{ei}$. Figure 2 shows the convergence curves for SSR = -5 dB. The SNR is varied from -10 dB to +10 dB to result in the initial probabilities of error indicated on the abscissa. The figure shows that not only is the decorrelator faster, but it also has a wider region of convergence than the LMS.

Conclusions

In this paper we analyzed the transient response of a stochastic gradient-based decorrelation algorithm and showed that it has a faster convergence rate than due to a lesser eigenvalue spread. Furthermore, the algorithm was shown to have wider regions of convergence than the LMS. The decorrelation and the LMS algorithms are of the same complexity. The application of the adaptive decorrelator was suggested as a multiuser detector for a synchronous multiple carrier CDMA system in a fading environment.

References

[1] J.W. Carlin Y. Bar-Ness and M.L. Steinberger. Bootstrapping adaptive cross-pol cancelers for satellite communications. *ICC*, 1982.

[2] M.L. Steinberger J.W.Carlin Y. Bar-Ness, S. Gross and W.E. Studdiford. An if cross-pol canceler for microwave radio. *JSAC-Advances in Digital Comm. by Radio*, 3:502–514, Apr. 1987.

[3] C. Jutten and J. Herault. Blind separation of sources, Part I: An adaptive algorithm based on neuromimetric architecture. *Signal Processing, Elsevier*, 24:1–10, July 1991.

[4] C. Jutten and J. Herault. Blind separation of sources, Part II: Problems statement. *Signal Processing, Elsevier*, 24:11–20, July 1991.

[5] R. Lupas and S. Verdu. Linear multiuser detectors for synchronous code-division multiple-access channels. *IEEE Transactions on Information Theory*, IT-35:123–136, Jan. 1989.

[6] M. K. Varanasi and B. Aazhang. Near-optimum detection in synchronous code-division multiple-access systems. *IEEE Transactions on Communications*, 39:725–736, May 1991.

[7] Z. Siveski, Y. Bar-Ness, and D. W. Chen. Error performance of synchronous multiuser Code Division Multiple Access detector with multidimensional adaptive canceler. *European Transactions on Telecommunications*, Dec. 1994.

[8] Y. Bar-Ness, J. P. Linnartz, and X. Liu. Synchronous multi-user multi-carrier cdma communication system with decorrelating interference canceler. In *Proceedings of the 1994 IEEE Personal Indoor and Mobile Radio Communications (PIMRC)*, pages 184–188, 1994.

[9] A. Dinc and Y. Bar-Ness. Bootstrap: A fast adaptive signal separator. In *ICASSP 92*, pages II.325–II.328, 1992.

[10] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, MD, second edition, 1983.

[11] S. J. Nowlan and G. E. Hinton. A soft decision-directed LMS algorithm for blind equalization. *IEEE Transactions on Communications*, 41:275–279, Feb. 1993.

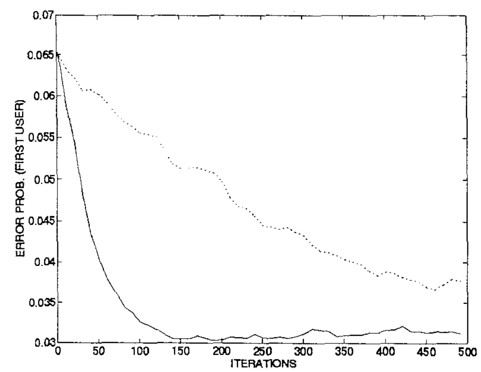

Fig. 1. Learning curve of the probability of error of the first user for N=4 users, SNR=8dB, SSR=-5dB.

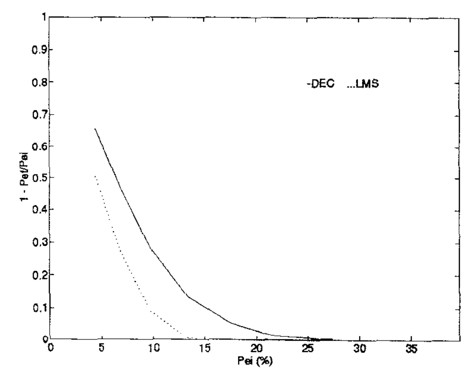

Fig. 2. Convergence regions after 100 iterations for the probability of error of the first user (N=4 users, SNR=8dB, SSR=-5 dB.)