Large System Convergence Analysis of Blind and Training-Based Stochastic Steepest Descent Adaptive CDMA Detectors

Teng Joon Lim† and Yu Gong‡

†Dept. of Electrical & Computer Engineering, University of Toronto,
King's College Road, Toronto, Ontario M5S 3G4.
E-mail: limtj@comm.toronto.edu.

‡Centre for Signal Processing, Nanyang Technological University, Singapore.
E-mail: eygong@ntu.edu.sg

Abstract— **We show that, in the limit as the number of users K and the processing gain N both go to infinity with $\beta = K/N$ constant, there is no difference in the convergence speed of the training-based least mean square (LMS) and blind minimum output energy (MOE) adaptive CDMA detectors. The analysis is based on deriving the asymptotic eigenvalue spread of the two correlation matrices governing convergence of these two algorithms.**

I. Introduction

Adaptive code division multiple access (CDMA) detectors mitigate interference with minimal side information. In the case of the blind MOE detector [1], one only requires the same information as a conventional matched-filter receiver; for the training-based LMS detector [2], [3], one only requires a training sequence (which have to be inserted into the information-bearing data symbol stream at regular intervals) and symbol timing for the desired user. Other forms of these two basic algorithms have also appeared in the literature. They all have the advantage of being applicable on the downlink of a communications system, unlike most other multiuser detectors, but the disadvantages of requiring the use of short codes, and that the channel be slowly and smoothly changing over time at worst.

On balance, it appears that adaptive detectors can be used for low-mobility applications such as wireless local area networks (LANs) or wireless local loops (WLLs) where relatively long setup times[1] are acceptable. In mobile, fading channels, their practicality unfortunately appears to be limited, as recently shown in [4]. From this point of view, adaptive detectors have a place in the lexicon of communications systems, and in this paper, our objective is to apply the theory of large-dimensional matrix eigenstructure [5] to the analysis of convergence speed of both the blind MOE and training-based LMS detectors.

Our main result is that, for the same adaptation step size μ, the speed of convergence from rest is identical for both detectors, in the limit as K and N grow to infinity while keeping K/N constant and finite. Computer simulations indicate that even for values of K and N that are quite small, the behaviour predicted by asymptotic analysis holds to a large degree.

[1]Adaptive algorithms can take many tens or even hundreds of symbols to converge to a steady state.

II. Relevant Past Results

A. Adaptive CDMA Detectors

As is well known from [6] and other multiuser detection publications, the received-signal vector can be expressed as

$$\mathbf{y}(i) = \mathbf{A}(i)\mathbf{d}(i) + \mathbf{n}(i) \tag{1}$$

where $\mathbf{y}(i) \in \mathbb{C}^N$ is the vector of baseband received-signal samples used for the detection of the desired user's ith symbol; $\mathbf{A}(i) \in \mathbb{C}^{N\times K}$ is the channel matrix, incorporating both spreading-sequence and channel information; $\mathbf{d}(i) \in \mathbb{C}^K$ is the vector of transmitted symbols contributing to $\mathbf{y}(i)$; and $\mathbf{n}(i)$ is a vector of samples of an additive white Gaussian noise signal with zero mean and covariance $\frac{N_0}{2}\mathbf{I}$.

A linear adaptive detector forms a decision statistic for user 1 (assumed without loss of generality to be the user of interest) by filtering $\mathbf{y}(i)$:

$$\hat{d}_1(i) = \mathbf{w}^{\mathrm{H}}(i)\mathbf{y}(i) \tag{2}$$

where $\mathbf{w}(i)$ is the length-N filter tap-weight vector in the ith symbol epoch, and is obtained through a time-iterative process. In this paper, we are interested in the following two algorithms:

$$\text{LMS}: \quad \mathbf{w}(i+1) = \mathbf{w}(i) + \mu e^*(i)\mathbf{y}(i)$$

$$\text{MOE}: \quad \left\{ \begin{array}{rcl} \mathbf{x}(i+1) & = & \mathbf{x}(i) - \mu z^*(i)\mathbf{P}\mathbf{y}(i) \\ \mathbf{w}(i+1) & = & \mathbf{x}(i+1) + \mathbf{a}_1 \end{array} \right.$$

The vector $\mathbf{a}_1$ is the column of $\mathbf{A}$ that couples with $d_1(i)$, $z(i) = \mathbf{w}^{\mathrm{H}}(i)\mathbf{y}(i)$ is the filter output, $e(i) = d_1(i) - z(i)$ is the estimation error term needed in the LMS algorithm, and $\mathbf{P}$ performs an orthogonal projection onto the vector space orthogonal to $\mathbf{a}_1$. We note that for adaptive algorithms such as the LMS and MOE to work in the CDMA scenario, the spreading sequences of all users must have periods equal to the symbol rate, and hence $\mathbf{A}(i) = \mathbf{A}$.

B. Eigenvalues of Large-Dimensional Correlation Matrices

From the results presented in Bai and Yin [5], we can deduce the following theorem:

Theorem 1: Given a matrix $\mathbf{S} = \frac{1}{N}\mathbf{X}\mathbf{X}^{\mathrm{H}}$, where the elements of $\mathbf{X}$ are independent and identically distributed (i.i.d.) with zero mean and variance σ^2, and $\mathbf{X}$ has K rows and N columns, the largest and smallest *non-zero* eigenvalues of $\mathbf{S}$ converge in the limit as $K, N \to \infty$ for a fixed $\beta \triangleq K/N$ to

$$\lambda_{max} = (1+\sqrt{\beta})^2\sigma^2 \tag{3}$$
$$\text{and} \quad \lambda_{min} = (1-\sqrt{\beta})^2\sigma^2 \tag{4}$$

respectively.

This theorem (and its corollaries) allows for a deeper analysis of the relative convergence behaviour of the training-based LMS and blind MOE detectors than what is currently known, as we showed in [7]. Although strictly valid only in the limiting case of large K and N, the resulting conclusions appear from simulations to hold even for finite-dimensional systems. In the present paper, we focus only on the speed of convergence, given identical adaptation step sizes.

III. Eigenvalue Spread of Input Correlation Matrices

A. Training-Based LMS Detector

Because the elements of the matrix $\mathbf{X}$ in Theorem 1 are required to be i.i.d., we need to restrict our attention to synchronous, equal-power CDMA systems. Under this constraint we lose no generality in assuming that the spreading-code matrix $\mathbf{A}$ has columns of unit norm. Then clearly $\mathbf{A}^{\mathrm{H}}\mathbf{A} = \frac{1}{N}\mathbf{S}^{\mathrm{H}}\mathbf{S}$ where $\mathbf{S} = \sqrt{N}\mathbf{A} \in \{\pm 1\}^{N\times K}$. Therefore if the spreading codes are randomly drawn from a binary distribution, we have

$$\lim_{K,N\to\infty} \lambda_{max}(\mathbf{A}^{\mathrm{H}}\mathbf{A}) = (1+\sqrt{K/N})^2$$
$$\lim_{K,N\to\infty} \lambda_{min}(\mathbf{A}^{\mathrm{H}}\mathbf{A}) = (1-\sqrt{K/N})^2.$$

This observation has already been made elsewhere (see [8], [9]) and others).

In an LMS adaptive detector, the convergence speed when the weight vector is initialised to zero is determined by the eigenvalue spread (i.e. the ratio of the largest and smallest eigenvalues) of the input correlation matrix $\mathbf{R}_{yy}$ in the signal subspace [10]. Now it is easily shown that

1. $\mathbf{R}_{yy} = \mathbf{A}\mathbf{A}^{\mathrm{H}} + \sigma^2\mathbf{I}$;
2. the eigenvalues of $\mathbf{R}_{yy}$ and $\mathbf{A}\mathbf{A}^{\mathrm{H}}$ differ by σ^2;
3. the non-zero eigenvalues of $\mathbf{A}\mathbf{A}^{\mathrm{H}}$, which are the eigenvalues in the signal subspace, are equal to those of $\mathbf{A}^{\mathrm{H}}\mathbf{A}$.

Putting these three facts together gives us the following corollary.

Corollary 1: In the limit as K and N tend to infinity, and $\beta = K/N$ is constant, the minimum and maximum eigenvalues of $\mathbf{R}_{yy} = \mathbf{A}\mathbf{A}^{\mathrm{H}} + \sigma^2\mathbf{I}$ in the signal subspace converge with probability 1 to

$$\lambda_{min}(\mathbf{R}_{yy}) = (1-\sqrt{K/N})^2 + \sigma^2 \tag{5}$$
$$\lambda_{max}(\mathbf{R}_{yy}) = (1+\sqrt{K/N})^2 + \sigma^2. \tag{6}$$

It is common practice to initialize the filter tap-weight vector to the origin, i.e. $\mathbf{w}(0) = \mathbf{0}$, in which case the eigenvalue spread that interests us is [10]

$$\gamma_{LMS} = \frac{(1+\sqrt{K/N})^2 + \sigma^2}{(1-\sqrt{K/N})^2 + \sigma^2}, \tag{7}$$

an expression which we will compare against the corresponding result for the blind detector.

B. Blind Minimum Output Energy (MOE) Detector

We have shown in [11], [7] that the equivalent correlation matrix that governs convergence of the MOE detector is $\mathbf{R}_{vv} = \mathbf{P}\mathbf{R}_{yy}\mathbf{P}$, where $\mathbf{P} = \mathbf{I} - \mathbf{a}_1\mathbf{a}_1^{\mathrm{H}}$. Furthermore, when we set $\mathbf{x}(0) = \mathbf{0}$, it can be shown [7] that the $K-1$ largest eigenvalues of $\mathbf{R}_{vv}$ determine the convergence speed of the MOE detector. We therefore seek an asymptotic expression for the ratio of the largest to the smallest of these $K-1$ eigenvalues.

Given that $\mathbf{R}_{yy} = \mathbf{A}\mathbf{A}^{\mathrm{H}} + \sigma^2\mathbf{I}$, we can write

$$\mathbf{R}_{vv} = \mathbf{P}\mathbf{A}\mathbf{A}^{\mathrm{H}}\mathbf{P} + \sigma^2\mathbf{P} \tag{8}$$
$$= (\mathbf{P}\mathbf{A})\left(\mathbf{A}^{\mathrm{H}}\mathbf{P}^{\mathrm{H}}\right) + \sigma^2\mathbf{I} - \sigma^2\mathbf{a}_1\mathbf{a}_1^{\mathrm{H}}. \tag{9}$$

Since $\mathbf{P}\mathbf{a}_1 = \mathbf{0}$, it is clear that $\mathbf{R}_{vv}\mathbf{a}_1 = \mathbf{0}$, and hence $\mathbf{a}_1$ is an eigenvector of $\mathbf{R}_{vv}$ with an eigenvalue of zero.

Consider also that $\mathbf{R}_{vv}$ is a symmetric matrix, and therefore has orthogonal eigenvectors. If $\mathbf{e}_i$, $i = 1, \ldots, N-1$ are the eigenvectors of $\mathbf{R}_{vv}$ not equal to $\mathbf{a}_1$, then $\mathbf{a}_1^{\mathrm{H}}\mathbf{e}_i = 0$. Suppose that λ_i is the eigenvalue associated with $\mathbf{e}_i$, then by definition $\mathbf{R}_{vv}\mathbf{e}_i = \lambda_i\mathbf{e}_i$. Substituting (9) into this expression yields

$$\left(\mathbf{A}_P\mathbf{A}_P^{\mathrm{H}} + \sigma^2\mathbf{I}\right)\mathbf{e}_i = \lambda_i\mathbf{e}_i \tag{10}$$
$$\text{or} \quad \mathbf{A}_P\mathbf{A}_P^{\mathrm{H}}\mathbf{e}_i = (\lambda_i - \sigma^2)\mathbf{e}_i \tag{11}$$

where $\mathbf{A}_P = \mathbf{P}\mathbf{A}$. This means that there are $N-1$ eigenvalues of $\mathbf{A}_P\mathbf{A}_P^{\mathrm{H}}$ which are uniformly smaller than those of $\mathbf{R}_{vv}$ by σ^2. (As is readily verified, the one remaining eigenvalue of $\mathbf{A}_P\mathbf{A}_P^{\mathrm{H}}$ is zero, and is paired with the eigenvector $\mathbf{a}_1$.) Finding the maximum and minimum eigenvalues of $\mathbf{A}_P\mathbf{A}_P^{\mathrm{H}}$ thus allows us to deduce the eigenvalue spread of $\mathbf{R}_{vv}$ within its signal subspace.

To proceed, we observe that, since $\mathbf{P}\mathbf{a}_1 = \mathbf{0}$,

$$\mathbf{A}_P\mathbf{A}_P^{\mathrm{H}} = \mathbf{P}\mathbf{A}\mathbf{A}^{\mathrm{H}}\mathbf{P} = \mathbf{P}\mathbf{A}_{\bar{1}}\mathbf{A}_{\bar{1}}^{\mathrm{H}}\mathbf{P}^{\mathrm{H}} \tag{12}$$

where $\mathbf{A}_{\bar{1}}$ denotes the matrix $\mathbf{A}$ with the first column removed. Observation (12) allows us to say that the non-zero eigenvalues of $\mathbf{A}_P\mathbf{A}_P^{\mathrm{H}}$ are identical to those of

$$\left(\mathbf{A}_{\bar{1}}^{\mathrm{H}}\mathbf{P}^{\mathrm{H}}\right)\left(\mathbf{P}\mathbf{A}_{\bar{1}}\right) = \frac{1}{N}\mathbf{S}_P^{\mathrm{H}}\mathbf{S}_P \tag{13}$$

where $\mathbf{S}_P = \sqrt{N}\mathbf{P}\mathbf{A}_{\bar{1}}$.

Assuming that $\mathbf{A}_{\bar{1}}$ is a random matrix with zero-mean i.i.d. elements, $\mathbf{S}_P$ (being a projection of $\mathbf{A}_{\bar{1}}$ onto a random subspace) is also a zero-mean i.i.d. random matrix. Keeping the ratio $(K-1)/N$ constant as K and N grow to infinity will hence admit the use of Theorem 1 to find the asymptotic minimum and maximum non-zero eigenvalues of $\mathbf{A}_P\mathbf{A}_P^{\mathrm{H}}$.

The one remaining unknown is the variance of the i.i.d. elements of $\mathbf{S}_P$, which we find by noting that the kth column ($k = 1, \ldots, K-1$) of $\mathbf{S}_P$ is

$$\mathbf{s}_{p,k} = \sqrt{N}(\mathbf{I} - \mathbf{a}_1\mathbf{a}_1^{\mathrm{H}})\mathbf{a}_{k+1} = \sqrt{N}(\mathbf{a}_{k+1} - \mathbf{a}_1\rho_{1,k+1}), \quad (14)$$

where $\rho_{1,k+1} \triangleq \mathbf{a}_1^{\mathrm{H}}\mathbf{a}_{k+1}$. With some trivial manipulations, we can show that

$$E[\mathbf{s}_{p,k}^{\mathrm{H}}\mathbf{s}_{p,k}] = N \cdot E(1 - |\rho_{1,k+1}|^2). \quad (15)$$

But $\rho_{1,k+1} = \sum_{n=1}^{N} a_{1,n}^* a_{k+1,n}$, and assuming independence between $a_{1,n}$ and $a_{k+1,n}$ for all n, we have

$$E|a_{1,n}^* a_{k+1,n}|^2 = E|a_{1,n}|^2 \cdot E|a_{k+1,n}|^2 = \frac{1}{N^2}. \quad (16)$$

Therefore, $E|\rho_{1,k+1}|^2 = \frac{1}{N^2} \times N = \frac{1}{N}$, and we have

$$E[\mathbf{s}_{p,k}^{\mathrm{H}}\mathbf{s}_{p,k}] = N - 1. \quad (17)$$

The individual elements of $\mathbf{S}_P$ have a variance that is $1/N$ times of that, i.e. $1 - (1/N)$.

Referring to Theorem 1, we can now state that, as K and N tend to infinity but $(K-1)/N$ remains constant, the minimum and maximum non-zero eigenvalues of $\mathbf{R}_{vv}$ converge to

$$\lambda_{min}(\mathbf{R}_{vv}) = \left(1 - \sqrt{\frac{K-1}{N}}\right)^2 \left(1 - \frac{1}{N}\right) + \sigma^2 \quad (18)$$

$$\lambda_{max}(\mathbf{R}_{vv}) = \left(1 + \sqrt{\frac{K-1}{N}}\right)^2 \left(1 - \frac{1}{N}\right) + \sigma^2. \quad (19)$$

Finally, since asymptotically $K - 1 \rightarrow K$ and $1/N \rightarrow 0$, we have

$$\lambda_{min}(\mathbf{R}_{vv}) = \left(1 - \sqrt{\frac{K}{N}}\right)^2 + \sigma^2 \quad (20)$$

$$\lambda_{max}(\mathbf{R}_{vv}) = \left(1 + \sqrt{\frac{K}{N}}\right)^2 + \sigma^2. \quad (21)$$

Comparing these expressions with (5) and (6), we can conclude that in the limit, the blind MOE and training-based LMS CDMA detectors have no differences in convergence behaviour. This conclusion does not contradict the results of [12], which states that the blind MOE detector always converges faster than the training-based one, but does not indicate how much faster it converges. Here, we have shown that for large systems, the difference in convergence performance is negligible.

IV. Simulations

We simulated $K = 8$, $N = 12$ synchronous CDMA systems with spreading sequences randomly selected for each experiment. All users were received with the same power, and $E_b/N_o = 5$ dB for each user. In Figure 1, we let the MOE and LMS algorithms have the same step size μ ($= 0.02$), and averaged the MSE and SIR for user 1 over 1000 independent runs i.e., different noise realizations and transmitted bits were used in each of 1000 separate experiments, but the same spreading sequences were used throughout. Both algorithms reach their steady states after about 100 symbols. However the MOE algorithm had a measured misadjustment of 49.67 % and final SIR of some 5 dB, while the LMS algorithm had a misadjustment of only 9.53 % and a final SIR of over 6 dB.

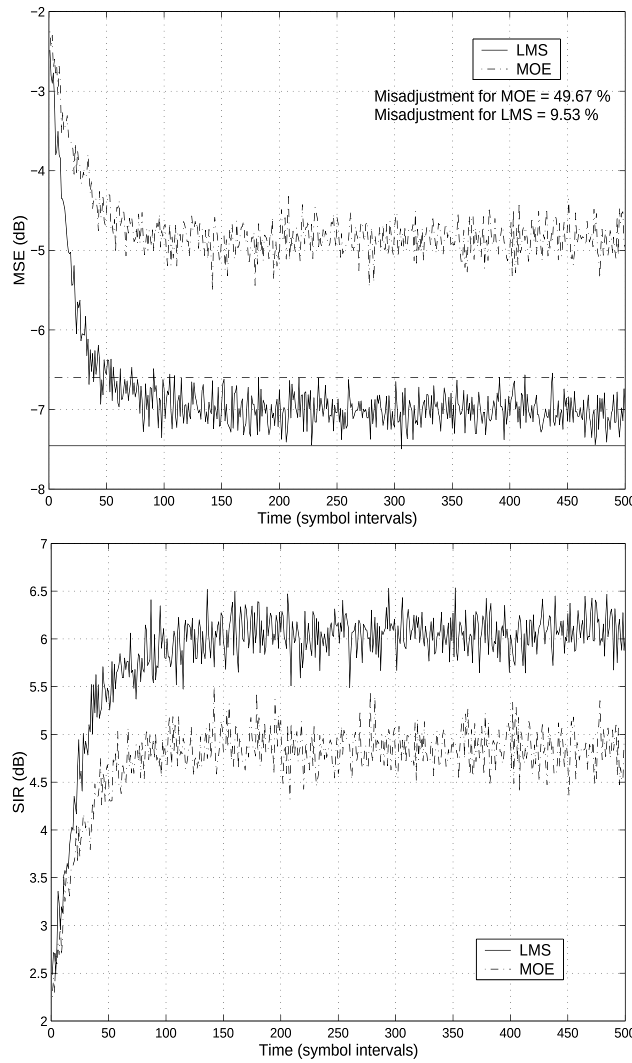

Fig. 1. MSE and SIR curves for the LMS and MOE algorithms when $\mu = 0.02$ for both cases.

Simulations of a more realistic asynchronous, fading CDMA channel are ongoing.

V. Conclusions

We have shown through large-system analysis that asymptotically, as the number of users and processing gain grow to infinity at the same rate, there is no difference in eigenvalue spread of the relevant correlation matrices for the LMS and MOE adaptive CDMA detectors. This implies that for a given step size μ, we should expect no difference in convergence speed, though the final signal-to-interference ratio (SIR) of the two algorithms differ substantially. We should add that more substantive results that show that the MOE algorithm is always expected to have worse performance than the LMS are derived in [7], again using the tool of large-system analysis.

References

[1] M. Honig, U. Madhow, and S. Verdú, "Blind adaptive multiuser detection," *IEEE Trans. Inf. Theory*, vol. 41, no. 4, pp. 944–960, July 1995.

[2] P. B. Rapajic and B. S. Vucetic, "Adaptive receiver structures for asynchronous CDMA systems," *IEEE J. Sel. Areas Comms.*, vol. 12, no. 4, pp. 685–697, May 1994.

[3] S. L. Miller, "An adaptive direct-sequence code-division multiple-access receiver for multiuser interference rejection," *IEEE Trans. Communications*, vol. 43, no. 2/3/4, pp. 1746–1755, Feb/Mar/Apr 1995.

[4] S. L. Miller, M. L. Honig, and L. B. Milstein, "Performance analysis of MMSE receivers for DS-CDMA in frequency-selective fading channels," *IEEE Trans. Communications*, vol. 48, no. 11, pp. 1919–1929, Nov. 2000.

[5] Z. D. Bai and Y. Q. Yin, "Limit of the smallest eigenvalue of a large dimensional sample covariance matrix," *Annals of Probability*, vol. 21, no. 3, pp. 1275–1294, July 1993.

[6] L. K. Rasmussen, P. D. Alexander, and T. J. Lim, "A linear model for CDMA signals received with multiple antennas over multipath fading channels," in *CDMA Techniques for Third Generation Mobile Systems*, P. van Rooyen, I. Oppermann, and M. Lotter, Eds., chapter 2. Kluwer Academic, Norwell, MA, 1999.

[7] T. J. Lim, B. Farhang-Boroujeny, and Y. Gong, "Relative convergence properties of the adaptive minimum output energy and least mean square CDMA detectors," *submitted to the IEEE Trans. Communications*, Feb. 2001.

[8] S. Verdú and S. Shamai, "Spectral efficiency of CDMA with random spreading," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 622–640, Mar. 1999.

[9] A. J. Grant and P. D. Alexander, "Random sequence multisets for synchronous code-division multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 44, no. 7, pp. 2832–2836, Nov. 1998.

[10] T. J. Lim, Y. Gong, and B. Farhang-Boroujeny, "Convergence analysis of chip- and fractionally-spaced LMS adaptive multiuser CDMA detectors," *IEEE Trans. Signal Processing*, vol. 48, no. 8, pp. 2219–2228, Aug. 2000.

[11] Y. Gong, T. J. Lim, and B. Farhang-Boroujeny, "Performance analysis of the LMS blind minimum output energy CDMA detector," in *Proc. of the IEEE Int'l Conf. Acoust., Speech & Signal Proc. (ICASSP)*, Istanbul, Turkey, June 2000, pp. V2473–V2476.

[12] S. Roy, "Subspace blind adaptive detection for multiuser CDMA," *IEEE Trans. Communications*, vol. 48, no. 1, pp. 169–175, Jan. 2000.