VARIABLE STEP-SIZE LMS BLIND CDMA MULTIUSER DETECTOR

Yu Gong

Centre for Signal Processing
Nanyang Technological
University,
Singapore 639798

B. Farhang-Boroujeny

Department of Electrical
Engineering,
University of Utah,
Salt Lake City,
UT 84112-9206, USA.

Teng Joon Lim

Department of Electrical and
Computer Engineering
University of Toronto
Toronto, ON M5S 3G4, Canada

ABSTRACT

Adaptive least mean square (LMS) filters with or without training sequences, which are known as training-based and blind detectors respectively, have been formulated to counter interference in CDMA systems. The convergence characteristics of these two LMS detectors are analyzed and compared in this paper. We show that the blind detector is superior to the training-based detector with respect to convergence rate. On the other hand, the training-based detector performs better in the steady state, giving a lower excess mean-square error (MSE) for a given adaptation step-size. A novel decision-directed LMS detector which achieves the low excess MSE of the training-based detector and the superior convergence performance of the blind detector is proposed.

1. INTRODUCTION

In a multiuser code-division multiple-access (CDMA) system, multiple-access interference (MAI) is the main source of performance degradation. Multiuser detection, which is now an established and popular field of research, has the objective of suppressing MAI. In this paper, we are interested in the LMS implementation of the linear tapped-delay line MMSE detector, both in its training-based and blind (or anchored) forms.

A comprehensive comparison of the convergence properties of the training-based and blind adaptive LMS multiuser detectors is thus far not available in the literature. The training-based detector was closely studied in [1], where it is shown that initialization of the filter tap-weight vector within the signal subspace or to the origin leads to faster convergence compared to initialization in the noise subspace, and that chip-spaced (CS) and fractionally-spaced (FS) detectors with the same time-span exhibit the same convergence behavior. Honig *et al.* [2] gave a simple convergence analysis for the blind (anchored) detector, which unfortunately contains some errors and is incomplete, as we pointed out in [3]. Here, we expand on the precise convergence results for the blind (anchored) detector given in [3], compare its convergence behavior to that of the training-based detector, and suggest an improved algorithm based on an adaptive decision directed implementation.

2. ANALYSIS

2.1. System Model

For simplicity, we consider a synchronous time-invariant K-user CDMA system with processing gain N, where user k transmits over a single-path, time-invariant channel characterized by the attenuation a_k, and all user symbol boundaries are aligned in time. The received signal vector is then written as

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

where the kth column of $\mathbf{S}$ represents the received spreading code of the kth user and is denoted by $\mathbf{s}_k$[1], $\mathbf{d}(i)$ is the data vector, $\mathbf{A} = \text{diag}[a_1, a_2, \cdots, a_K]$, $\mathbf{n}(i)$ is the channel additive white Gaussian noise (AWGN) vector, i is the symbol index, and $\mathbf{y}(i)$ is an $N \times 1$ column vector of the received signal sampled at chip-rate.

The convergene behaviour of the LMS algorithm is determined by the input correlation matrix. For the training-based detector, the correlation matrix is $\mathbf{R}_{yy} = \text{E}[\mathbf{y}(i)\mathbf{y}^H(i)]$. For the blind detector, as shown in [3], the relevant correlation matrix is $\mathbf{R}_{vv} = \text{E}[\mathbf{v}(i)\mathbf{v}^H(i)]$, where $\mathbf{v}(i) = \mathbf{P}_{s1}^{\perp}\mathbf{y}(i)$, $\mathbf{P}_{s1}^{\perp} = \mathbf{I} - \mathbf{s}_1\mathbf{s}_1^H$. Since $\mathbf{R}_{vv}$ is symmetric, the convergence behaviour of the blind detector can be easily obtained, unlike what was implied by the procedure outlined in [2]. Furthermore, when the tap-weights are initialized in the signal

[1]We assume $\|\mathbf{s}_k\| = 1$ for $k = 1, \cdots, K$, without loss of generality

subspace, only the eigenvalues in the signal subspace affect the convergence behaviour [1, 3].

2.2. Convergence Analysis

2.2.1. Eigenvalue Spread

A comparison of the convergence behaviors of the training-based and blind detectors may be made by comparing the eigenvalue spread [4, 5]. First we note that $\mathbf{s}_1$ is an eigenvector of $\mathbf{R}_{vv}$ with an associated eigenvalue of zero. However, this zero eigenvalue plays no role in the convergence of the blind detector, since the tap-weights are never adapted in the direction of $\mathbf{s}_1$. It can be easily verified that σ^2 is the minimum eigenvalue of both $\mathbf{R}_{yy}$ and $\mathbf{R}_{vv}$, which repeats $N-K$ times for $\mathbf{R}_{yy}$ and $N-K-1$ for $\mathbf{R}_{vv}$. Thus,

$$\lambda_{\min}(\mathbf{R}_{yy}) = \lambda_{\min}(\mathbf{R}_{vv}) = \sigma^2 \tag{2}$$

where $\lambda_{\min}(\cdot)$ denotes the minimum non-zero eigenvalue of the indicated matrix.

On the other hand, the maximum eigenvalue of $\mathbf{R}_{yy}$ may be obtained by applying the minimax theorem which states [4, 5]

$$\lambda_{\max}(\mathbf{R}_{yy}) = \max_{\mathbf{q}} \mathbf{q}^H \mathbf{R}_{yy} \mathbf{q}, \quad \text{subject to } \mathbf{q}^H\mathbf{q} = 1. \tag{3}$$

Similarly, the maximum eigenvalue of $\mathbf{R}_{vv}$ can be obtained from

$$\lambda_{\max}(\mathbf{R}_{vv}) = \max_{\mathbf{p}} \mathbf{p}^H \mathbf{R}_{vv} \mathbf{p}, \quad \text{subject to } \mathbf{p}^H\mathbf{p} = 1. \tag{4}$$

Decomposing $\mathbf{p}$ as $\mathbf{p} = \mathbf{p}_1 + \mathbf{p}_2$, where $\mathbf{p}_1$ is the projection of $\mathbf{p}$ onto $\mathbf{s}_1$, and noting that $\mathbf{R}_{vv} = \mathbf{P}_{s1}^{\perp}\mathbf{R}_{yy}\mathbf{P}_{s1}^{\perp}$ and $\mathbf{P}_{s1}^{\perp}\mathbf{p}_1 = 0$, $\mathbf{P}_{s1}^{\perp}\mathbf{p}_2 = \mathbf{p}_2$, we obtain

$$\lambda_{\max}(\mathbf{R}_{vv}) = \max_{\mathbf{p}_2} \mathbf{p}_2^H \mathbf{R}_{yy} \mathbf{p}_2, \tag{5}$$

subject to $\|\mathbf{p}_2\|^2 = 1 - \|\mathbf{p}_1\|^2$. Since $\mathbf{R}_{yy}$ is positive semi-definite, $\mathbf{x}^H\mathbf{R}_{yy}\mathbf{x}$ is a convex function of $\mathbf{x}$. Thus, maximizing (5) with respect to $\mathbf{p}$ requires us to make $\mathbf{p}_2$ as large as possible, while keeping $\|\mathbf{p}\| = 1$. Consequently, $\mathbf{p}_1$ must be set to $\mathbf{0}$ and we arrive at the following result,

$$\lambda_{\max}(\mathbf{R}_{vv}) = \max_{\mathbf{p}} \mathbf{p}^H \mathbf{R}_{yy} \mathbf{p}, \tag{6}$$

subject to $\mathbf{p}^H\mathbf{p} = 1$ and $\mathbf{p}^H\mathbf{s}_1 = 0$.

Comparing (3) and (6), we see that both $\lambda_{\max}(\mathbf{R}_{yy})$ and $\lambda_{\max}(\mathbf{R}_{vv})$ are obtained by maximizing the same quadratic expression, but the latter has an additional constraint. This implies that

$$\lambda_{\max}(\mathbf{R}_{vv}) \leq \lambda_{\max}(\mathbf{R}_{yy}). \tag{7}$$

Furthermore, from (2) and (7), we obtain

$$\frac{\lambda_{\max}(\mathbf{R}_{vv})}{\lambda_{\min}(\mathbf{R}_{vv})} \leq \frac{\lambda_{\max}(\mathbf{R}_{yy})}{\lambda_{\min}(\mathbf{R}_{yy})} \tag{8}$$

which indicates that the blind detector has an eigenvalue spread that is smaller than or equal to its counterpart in the training-based detector.

The same result as (8) has also been developed by Roy [6], using a somewhat different approach. However, using our observation that the Hermitian $\mathbf{R}_{vv}$ can be taken as the convergence-controlling correlation matrix for the blind detector, further results can be derived, as discussed in the next few sections.

2.2.2. Initialization Effects

In some recent works [1], we have studied the phenomenon of different convergence behaviors of MMSE detectors when the adaptive tap-weight vector is initialized in the signal and noise subspaces. The relative performance of the blind and training-based adaptive detectors must also be examined from this angle.

Let $\lambda_{\text{s-max}}(.)$ and $\lambda_{\text{s-min}}(.)$ denote the maximum and minimum eigenvalues in the signal subspace of the indicated matrices, respectively. Obviously $\lambda_{\text{s-max}}(\mathbf{R}_{yy}) = \lambda_{\max}(\mathbf{R}_{yy})$ and $\lambda_{\text{s-max}}(\mathbf{R}_{vv}) = \lambda_{\max}(\mathbf{R}_{vv})$. Thus, in view of (7), $\lambda_{\text{s-max}}(\mathbf{R}_{vv}) \leq \lambda_{\text{s-max}}(\mathbf{R}_{yy})$. Following the same line of derivations to those which led to (7), we also obtain

$$\lambda_{\text{s-min}}(\mathbf{R}_{yy}) = \min_{\mathbf{q}} \mathbf{q}^H \mathbf{R}_{yy} \mathbf{q}, \tag{9}$$

subject to $\mathbf{q}^H\mathbf{q} = 1$ and $\mathbf{q} \in \mathcal{S}$, where $\mathcal{S}$ indicates the signal subspace spanned by $\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_K$. Similarly,

$$\lambda_{\text{s-min}}(\mathbf{R}_{vv}) = \min_{\mathbf{p}} \mathbf{p}^H \mathbf{R}_{yy} \mathbf{p}, \tag{10}$$

subject to $\mathbf{p}^H\mathbf{p} = 1$ and $\mathbf{p} \in \tilde{\mathcal{S}}$, where $\tilde{\mathcal{S}}$ indicates the signal subspace spanned by $\tilde{\mathbf{s}}_2, \tilde{\mathbf{s}}_3, \cdots, \tilde{\mathbf{s}}_K$, where $\tilde{\mathbf{s}}_k = \mathbf{P}_{s1}^{\perp}\mathbf{s}_k$.

Noting that $\tilde{\mathcal{S}}$ is a subspace of $\mathcal{S}$, from (9) and (10), we obtain

$$\lambda_{\text{s-min}}(\mathbf{R}_{vv}) \geq \lambda_{\text{s-min}}(\mathbf{R}_{yy}). \tag{11}$$

Thus,

$$\frac{\lambda_{\text{s-max}}(\mathbf{R}_{vv})}{\lambda_{\text{s-min}}(\mathbf{R}_{vv})} \leq \frac{\lambda_{\text{s-max}}(\mathbf{R}_{yy})}{\lambda_{\text{s-min}}(\mathbf{R}_{yy})}. \tag{12}$$

That is, with signal-subspace initialization of the tap weight vectors, the blind detector also has a lower effective eigenvalue spread than the training-based detector.

2.2.3. Misadjustment and excess MSE

The misadjustment for the training-based and blind LMS detectors is given by [4, 3]

$$\mathcal{M}_{\text{td}} = 0.5\mu\text{tr}[\mathbf{R}_{yy}], \quad \mathcal{M}_{\text{bd}} = 0.5\mu\text{tr}[\mathbf{R}_{vv}], \tag{13}$$

respectively. where tr[.] denotes the trace of the indicated matrix. But $\text{tr}[\mathbf{R}_{vv}] < \text{tr}[\mathbf{R}_{yy}]$. Hence, for equal step-size parameters,

$$\mathcal{M}_{\text{bd}} < \mathcal{M}_{\text{td}}. \tag{14}$$

Even though compared to the training-based detector, the blind detector has smaller eigenvalue spread and lower misadjustment, one should not conclude that, in terms of convergence behavior, the blind detector has more to offer than the training-based detector. In fact, the opposite result may be obtained if we consider the excess MSE.

The error in the blind detector is the desired output plus channel noise. This results in a significantly larger tap-weight perturbation and thus a much higher excess MSE in the blind detector, as compared to the training based detector where error is almost at channel noise level. We assume that the channel noise is significantly lower than the desired signal level. This point has also been mentioned by Honig et al [2]. In the next section we propose a hybrid CDMA detector which works blindly when it is started, but switches to a special decision directed mode later so that, similar to the training-based detector, it achieves a small steady-state excess MSE.

3. HYBRID DETECTOR

In order to develop an algorithm which keeps the advantages of both blind and training-based detectors, we define the performance function

$$\eta = E[|\alpha d_1(i) - (\mathbf{s}_1 + \mathbf{x})^H \mathbf{y}(i)|^2], \tag{15}$$

where the scalar α and vector $\mathbf{x}$ are adjustable parameters which are chosen to minimize η. We show that the optimum choices of α and $\mathbf{x}$ are independent of each other. To this end, by expanding (15), recalling $\mathbf{s}_1^H \mathbf{x} = 0$ and noting that $\text{E}[d_1^*(i)\mathbf{s}_1^H \mathbf{y}(i)] = a_1$, we obtain

$$\eta = |\alpha|^2 - \alpha a_1^* - \alpha^* a_1 + E[|(\mathbf{s}_1 + \mathbf{x})^H \mathbf{y}(i)|^2]. \tag{16}$$

We clearly see that optimization of α involves minimization of the first three terms on the right-hand side of (16) which leads to $\alpha_{\text{opt}} = a_1$, and optimization of $\mathbf{x}$ requires minimization of the last term on the right-hand side of (16). Moreover, the last term is nothing but the cost function of the blind detector. This means the optimum value of $\mathbf{x}$ in (15) is identical to that of the blind detector.

Considering that a good guess of $d_1(i)$ is available upon convergence of $\mathbf{x}$, we may start to adapt $\mathbf{x}$ with $\alpha = 0$, and then begin to adapt α when $\mathbf{x}$ has approached its optimum value. Thus we need a mechanism to automatically identify when $\mathbf{x}$ is close to its optimum value. For this we suggest using the variable step-size LMS (VSLMS) algorithm discussed in [4, 7]. In VSLMS, the step-size is large when the filter tap weights are far from their optimum values, and is small when the filter tap weights have approached their optimum values. A possible adaptation rule for μ is the following [4]:

$$\mu(i) = \left(1 + \rho \left(\mathbf{g}_R^T(i)\mathbf{g}_R(i-1) + \mathbf{g}_I^T(i)\mathbf{g}_I(i-1)\right)\right) \mu(i-1) \tag{17}$$

where ρ is the step-size parameter for adaptation of $\mu(i)$, $\mathbf{g}(i)$ and $\mathbf{g}(i-1)$ are the stochastic gradient at the present iteration and the one before, and the subscripts R and I indicate the real and imaginary parts, respectively. In addition, to make sure that the LMS algorithm remains stable, $\mu(i)$ has to be checked after every iteration not to exceed a predetermined level, μ^+. The step-size parameter $\mu(i)$ should also be checked against a minimum predetermined level, μ^-, and limited. The step-adaptation recursion (17) belongs to the class of VSLMS algorithm which was proposed by Mathews and Xie [7]. A recent work, [8], has shown that the VSLMS algorithm will be improved if the SG vector $\mathbf{g}(i-1)$ (but not $\mathbf{g}(i)$) in (17) is replaced by its short-time averaged values obtained using the recursion

$$\boldsymbol{\psi}(i) = \beta_\psi \cdot \boldsymbol{\psi}(i-1) + (1 - \beta_\psi) \cdot \mathbf{g}(i) \tag{18}$$

where β_ψ is a constant smaller than close to one. Moreover, it is noted that normalization of the parameter ρ to an estimate of the gradient power, $\text{E}[\mathbf{g}^H(i)\mathbf{g}(i)]$, also improves the behavior of the algorithm.

Clearly, the above procedure implies that $\mu(i)$ will remain relatively large when $\mathbf{x}$ has not yet converged, and drops to lower values when $\mathbf{x}$ has converged. Accordingly, the adaptation of α is activated by comparing $\mu(i)$ against a threshold level; say, μ_{th}. The adaptation of α is activated when $\mu(i) < \mu_{\text{th}}$, and stopped otherwise.

We use the LMS algorithm to adapt the parameter α. Here, the error is $\hat{e}(i) = \alpha(i)\hat{d}_1(i) - (\mathbf{s}_1 + \mathbf{x}(i))^H \mathbf{y}(i)$, and the tap input is $d_1(i)$. Since in practice $d_1(i)$ is not available, we use its estimate which for binary data is $\text{sign}[z(i)]$. This lead to the following update equation

$$\alpha(i+1) = \alpha(i) - \mu_\alpha \cdot \hat{e}(i) \cdot \text{sign}[z(i)]. \tag{19}$$

Computer simulations are performed to confirm the reliability of the above hybrid algorithm.

Because of the stochastic nature of the algorithm, the start time of adaptation of the coefficient α varies for different runs, even when all simulation parameters are kept fixed. Fig. 1 presents a histogram of the results of 5,000 independent runs with the following parameters: (i) processing gain $= 16$; (ii) number of active users $= 4$; (iii) desired user and two interfering users have power of 0 dB; (iv) fourth user power $= 10$ dB; (v) spreading codes are randomly generated for each run; (vi) background noise level $= -20$ dB. For the VSLMS algorithm the following parameters are used: (i) $\mu(0) = 0.1$; (ii) $\mu^- = 0.03$; (iii) $\mu^+ = 0.5$; (iv) $\mu_{\text{th}} = 0.06$; (v)$\mu_\alpha = 0.02$; (vi) $\beta_\psi = 0.99$.

From Fig. 1, we see that in most of the cases adaptation of α begins within the first 1000 iterations.

The output signal to interference and noise ratio (SINR) for the blind detector and the hybrid algorithm are shown in Fig. 2, where the VSLMS algorithm is used for both detectors, and time averaging is used to smoothen the curves. As expected, at the beginning of adaptation, both algorithms behave exactly the same. However, upon convergence of the blind detector and the start of adaptation of α, the hybrid algorithm begins to perform better, achieving a higher SINR. It is also observed that the hybrid algorithm converges to a stable SINR, while in the blind detector SINR has significant fluctuations.

4. CONCLUSION

In this paper, we compared the convergence behavior of the training-based detector and the blind detector. It is shown that the blind detector has superior convergence behavior to the training-based detector, while the training-based has smaller excess MSE than the blind detector. A hybrid detector which combines the advantages of the above detectors was then proposed. Computer simulations were done to verify the reliability of the new detector.

5. REFERENCES

[1] T. J. Lim, Y. Gong, and B. Farhang-Boroujeny. Convergence analysis of chip- and fractionally-spaced LMS adaptive multiuser CDMA detectors. *IEEE Trans. on Signal Processing*, 48(8):2219–2228, Aug. 2000.

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

[3] Y. Gong, T. J. Lim, and B. Farhang-Boroujeny. Performance analysis of the LMS blind minimum-output-energy CDMA detection. In *International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages V–2473 – V–2476, Istanbul, Turkey, June 2000.

[4] B. Farhang-Boroujeny. *Adaptive Filters-Theory and Application*. John Wiley & Sons, 1998.

[5] S. Haykin. *Adaptive Filter Theory*. Prentice Hall, Englewood Cliffs, NJ, 1996.

[6] S. Roy. Subspace blind adaptive detection for multiuser CDMA. *IEEE Trans. on Commun.*, 48(1):169–175, Jan. 2000.

[7] V. J. Mathews and Z. Xie. Stochastic gradient adaptive filters with gradient adaptive step sizes. *IEEE Trans. on Signal Processing*, 41(6):2075–2087, June 1993.

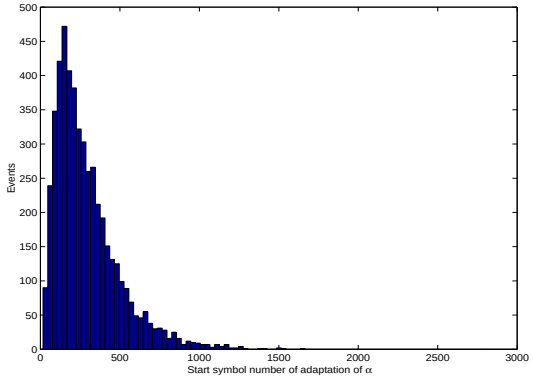

Fig. 1. Histogram for the start time of the adaption of the α, the results are based on 5,000 independent runs.

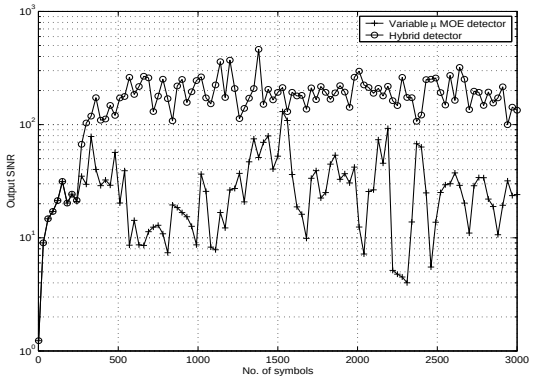

Fig. 2. A typical output SINR for the blind and hybrid detectors. The VSLMS scheme is used to adapt both algorithms.

[8] W. P. Ang and B. Farhang-Boroujeny. Gradient adaptive step-size LMS algorithms: Past results & new developments. In *IEEE Symposium 2000 on adaptive system for Signal Processing, Communications, and Control (ASSPCC)*, Canada, Oct. 2000.