

# Subspace Blind Adaptive Detection for Multiuser CDMA

Sumit Roy, *Member, IEEE*

**Abstract**—Direct adaptive realizations of the linear minimum mean-square error (MMSE) receiver for direct-sequence code-division multiple access possess the attractive feature of not requiring any explicit information of interference parameters such as timing, amplitudes, or spreading sequences; however, they need a training sequence for the desired user. Recently, a new *blind* adaptive receiver was proposed based on an anchored least mean-squared (LMS) algorithm that requires only the spreading code and symbol timing of the desired user but obviates the need for a training sequence. In this work, it is analytically demonstrated that the blind LMS algorithm *always* provides (nominally) faster convergence than the training driven LMS-MMSE receiver of but at the cost of increased tap-weight fluctuations or misadjustment. Second, the property that the optimal MMSE or minimum-output energy filter coefficients lies in the signal subspace is exploited to propose a new *efficient* blind adaptive receiver requiring fewer adaptive coefficients. Improved detector characteristics (superior convergence rates and steady-state signal-to-interference-plus-noise ratios) is indicated by analysis and supported by simulation.

**Index Terms**—Blind adaptive filtering, LMS algorithm, multiuser CDMA.

## I. INTRODUCTION

THE MITIGATION of multiple-access interference (MAI) in code-division multiple-access (CDMA) systems is a problem of continuing interest since MAI (and not additive noise) imposes a fundamental limitation on the capacity of CDMA multiuser (MU) systems. Several suboptimal MU detectors—notably the class of *linear* detectors—have been shown to significantly outperform the conventional receiver (bank of matched filters), particularly at high levels of MAI for which they are specifically designed. Such improved performance is incurred at the cost of increased *complexity* as well as presumed *side information* regarding the structure of the desired and/or interfering users' signals. Of particular current interest for next-generation digital wireless receivers are *adaptive* low-complexity MU detectors, since they possess desired features of adaptability to changing user and channel conditions.

Adaptive minimum mean-square error (MMSE) receivers based on training sequences were developed in [2]–[4]. The drawback of training in a distributed multipoint environment

Paper approved by S. L. Miller, the Editor for Spread Spectrum of the IEEE Communications Society. Manuscript received August 4, 1997; revised July 17, 1998 and December 14, 1998. This work was supported by the Army Research Office under Grant DAAH 04-96-10088 and AFOSR Grant F49620-1-0472

The author was with the University of Texas, San Antonio, TX USA. He is now with the Department of Electrical Engineering, University of Washington, Seattle, WA 98195 USA (e-mail: roy@ee.washington.edu).

Publisher Item Identifier S 0090-6778(00)00488-8.

has led to interest in blind adaptation schemes. The novel blind minimum-output energy (MOE) receiver of [2] based on an “anchored least mean-squared” (LMS) algorithm for the detection of a single desired user (and hence particularly suited for downlink applications) forms the immediate motivation for our work. The key innovation in architectural terms is a subspace (equivalently, reduced-order) receiver based on the fact that the optimal MMSE and the MOE detectors are constrained to lie in the space generated by vectors corresponding to the active users (hence, the signal subspace), as recently recognized by [5]–[8] independently. The primary contribution of this work is thus to quantify the performance of the subspace (reduced-order) receiver vis-a-vis the full-order MOE adaptive receiver [2].

## II. SIGNAL MODEL

We assume a synchronous CDMA system where the users employ binary phase-shift keying modulation of independent bit streams. Individual (short) spreading codes comprising of  $N$  chips that are periodically repeated within a symbol duration  $T = NT_c$  are used by each of  $K$  users to transmit  $b_k(i)$ —the  $i$ th bit for the  $k$ th user. The equivalent received pulse shape corresponding to the  $k$ th user,  $s_k(t)$ , is assumed known at the receiver and are energy-normalized, i.e.,  $\int_0^T s_k^2(t) dt = 1$ . The receiver consists of a chip matched filter front-end whose output is synchronously sampled at the instants  $lT_c$  to produce the sufficient statistics. The sampled chip matched filter output is given by

$$y(l) \triangleq y(lT_c) = \sum_{i=0}^M \sum_{k=1}^K \sqrt{w_k} b_k(i) s_k(lT_c - iT - \tau_k) + n(l) \quad (1)$$

where  $n(l)$  is a zero-mean, white Gaussian noise sequence with variance  $\sigma_n^2$ , and  $w_k, \tau_k$  are the received signal energies and user delays, respectively, for the  $k$ th user. The  $\{b_k(i)\}$  are assumed to be mutually independent sequences consisting of equiprobable,  $\pm 1$  independently, identically distributed symbols with  $E[b_k(i)] = 0$ ,  $E[b_k^2(i)] = 1$ . The propagation delays are assumed to be arranged in nondecreasing order, i.e.,  $\tau_1 \leq \tau_2 \leq \dots \leq \tau_K$ , without loss of generality.

### A. Blind Adaptive Detection: Synchronous Case

For the synchronous case, setting  $\tau_i = 0$ ,  $\forall i$ , yields the following well-known vector model

$$\mathbf{y}_i = \mathbf{A} \mathbf{W} \mathbf{b}_i + \mathbf{n}_i \quad (2)$$

where  $\mathbf{y}_i = [y(iT), \dots, y(iT + (N-1)T_c)]$  is an  $N$ -vector of chip-rate samples collected during the  $i$ th symbol duration,  $\mathbf{b}_i$  is

the  $K$ -vector of (unknown) binary symbols transmitted during this interval,  $\mathbf{W} = \text{diag}[\sqrt{w_1}, \sqrt{w_2}, \dots, \sqrt{w_K}]$  is the  $K \times K$  matrix of received user amplitudes and

$$\mathbf{A} = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_K] \quad (3)$$

where  $\mathbf{s}_k^T = [s_k(0), s_k(T_c), \dots, s_k((N-1)T_c)]$  is the *code sequence* assigned to the  $k$ th user. Without loss of generality in the sequel, the desired user's amplitude  $w_1 = 1$ .

In [2], the MOE detector described by

$$\min_{\mathbf{c}_{\text{moe}}} J_1(\mathbf{c}_{\text{moe}}) = \mathcal{E} \left\| \mathbf{c}_{\text{moe}}^T \mathbf{y}(i) \right\|^2 \quad \text{s.t.} \quad \mathbf{c}_{\text{moe}}^T \mathbf{s}_1 = 1 \quad (4)$$

was proposed that led to an attractive *blind* adaptation rule, where  $\mathcal{E}$  denotes statistical expectation. The solution to (4), given by  $\mathbf{c}_{\text{moe},o} = 1/(\mathbf{s}_1^T \mathbf{R}^{-1} \mathbf{s}_1) \mathbf{R}^{-1} \mathbf{s}_1$ , is well known to be related to that for the *MMSE* problem, i.e.,  $\min_{\mathbf{c}_{\text{mmse}}} J_2(\mathbf{c}_{\text{mmse}}) = \mathcal{E} \left\| b_1(i) - \mathbf{c}_{\text{mmse}}^T \mathbf{y}(i) \right\|^2$ , in the sense that the two solutions are identical up to a constant, i.e.,  $\mathbf{c}_{\text{moe},o} \propto \mathbf{c}_{\text{mmse},o}$ , and yield identical output signal-to-interference-plus-noise ratio (SINR), where for any coefficient vector  $\mathbf{c}$

$$\text{SINR} \triangleq \frac{(\mathbf{c}^T \mathbf{s}_1)^2}{\sum_{i \neq 1} w_i (\mathbf{c}^T \mathbf{s}_i)^2 + \sigma_n^2 \mathbf{c}^T \mathbf{c}}. \quad (5)$$

A gradient search-based algorithm that iteratively searches for the optimum of (4) requires the unconstrained gradient given by

$$\nabla_{\mathbf{c}_{\text{moe}}} E(\mathbf{c}_{\text{moe}}^T \mathbf{y}(i))^2 = 2E \left\{ [\mathbf{c}_{\text{moe}}^T \mathbf{y}(i)] \mathbf{y}(i) \right\} \quad (6)$$

which is approximated by the instantaneous value  $2Z(i)\mathbf{y}(i)$ , where  $Z(i) = \hat{\mathbf{c}}_{\text{moe}}^T(i)\mathbf{y}(i)$  is the decision variable for the  $i$ th bit of the desired user. Noting that only the component of the above gradient in the space orthogonal to  $\mathbf{s}_1$  is relevant for the update leads to the final form

$$\hat{\mathbf{c}}_{\text{moe}}(i) = \hat{\mathbf{c}}_{\text{moe}}(i-1) - \mu Z(i) (\mathbf{y}(i) - \mathbf{y}^T(i) \mathbf{s}_1 \mathbf{s}_1) \quad (7)$$

obtained by replacing  $\mathbf{y}(i)$  with  $\mathbf{y}(i) - [\mathbf{y}^T(i) \mathbf{s}_1] \mathbf{s}_1$ .

1) *Blind Adaptive Detection: Synchronous Case:* A new parameterization for the blind LMS algorithm (7) is introduced that facilitates a performance analysis and also suggests the new (and more efficient) blind subspace adaptive detector.

The adaptive coefficient vector  $\mathbf{c}_{\text{moe}}(i)$  may be rewritten as

$$\mathbf{c}_{\text{moe}}(i) = \mathbf{s}_1 + \mathbf{x}_1(i) \quad (8)$$

$$= \mathbf{s}_1 + \mathbf{U} \alpha(i) \quad (9)$$

where  $\alpha(i) \in \mathcal{R}^{N-1}$  and the  $N-1$  columns of the matrix  $\mathbf{U} \in \mathcal{R}^{N \times N-1}$  span the space orthogonal to  $\mathbf{s}_1$ , i.e.,  $\mathbf{s}_1^T \mathbf{U} = \mathbf{0}$ . While any spanning set (not necessarily orthogonal) of  $N-1$  vectors will suffice, we assume for simplicity and without loss of

generality that these are also orthonormal, i.e.,  $\mathbf{U}^T \mathbf{U} = \mathbf{I}_{N-1}$ . Using (8) into (7), we arrive at the update in terms of  $\alpha(i)$  as follows:

$$\mathbf{U} \alpha(i+1) = \mathbf{U} \alpha(i) - \mu (\mathbf{s}_1 + \mathbf{U} \alpha(i))^T \mathbf{y}(i) (\mathbf{y}(i) - \mathbf{y}^T(i) \mathbf{s}_1 \mathbf{s}_1) \quad (10)$$

$$= \mathbf{U} \alpha(i) - \mu [\mathbf{I} - \mathbf{s}_1 \mathbf{s}_1^T] \mathbf{y}(i) \mathbf{y}^T(i) (\mathbf{s}_1 + \mathbf{U} \alpha(i)). \quad (11)$$

Multiplying both sides by  $\mathbf{U}^T$  yields

$$\alpha(i+1) = \alpha(i) - \mu \mathbf{U}^T \mathbf{y}(i) \mathbf{y}^T(i) (\mathbf{s}_1 + \mathbf{U} \alpha(i)) \quad (12)$$

upon using  $\mathbf{U}^T \mathbf{s}_1 = \mathbf{0}$ . Taking expectations of both sides and invoking the *independence assumption* that  $\alpha(i)$  and  $\mathbf{y}(i)$  are independent leads to

$$E[\alpha(i+1)] = (\mathbf{I} - \mu \mathbf{U}^T \mathbf{R}_y \mathbf{U}) E[\alpha(i)] - \mu \mathbf{U}^T \mathbf{R}_y \mathbf{s}_1. \quad (13)$$

Recall that the analysis of the LMS adapted transversal (FIR) filter where the adaptation rate equals the input sampling rate invokes the “independence” assumption [9] of the current input  $\mathbf{y}(i)$  and filter coefficients  $\alpha(i)$ , which is only approximate (and is usually justified in the limit of  $\mu \rightarrow 0$ ). However, in our present case, this assumption is *exact* since the adaptation rate (equal to the symbol rate) is  $1/N$  of the input rate (equal to chip rate), whereby the adaptive filter inputs at instants  $i$  and  $i+1$  are truly independent.

From (13), it follows that the stability of the coefficient iteration in the mean is guaranteed iff  $|\mu \lambda_{\max}[\mathbf{U}^T \mathbf{R}_y \mathbf{U}]| \leq 1$ . Further, the rates of convergence are governed by the eigenvalues of the matrix  $\mathbf{U}^T \mathbf{R}_y \mathbf{U}$ —specifically, the ratio

$$\eta(\mathbf{U}^T \mathbf{R}_y \mathbf{U}) = \frac{\lambda_{\max}(\mathbf{U}^T \mathbf{R}_y \mathbf{U})}{\lambda_{\min}(\mathbf{U}^T \mathbf{R}_y \mathbf{U})}. \quad (14)$$

We now show that  $\eta(\mathbf{U}^T \mathbf{R}_y \mathbf{U}) \leq \eta(\mathbf{R}_y)$  implying the important result that *convergence rate* of the blind adaptive LMS algorithm is always superior (and depending on the scenario, considerably better) to that of the training-driven LMS algorithm. To do so, one invokes the *Poincare separation theorem* that states the following.

*Theorem [10]:* For a symmetric matrix,  $\mathbf{A} \in \mathcal{R}^{N \times N}$ , and  $\mathbf{U} \in \mathcal{R}^{N \times r}$  ( $r < N$ ), where  $\mathbf{U}^T \mathbf{U} = \mathbf{I}_r$ , let  $\mathbf{B} = \mathbf{U}^T \mathbf{A} \mathbf{U}$ . Then, if the eigenvalues of  $\mathbf{A}$  and  $\mathbf{B}$  are arranged in increasing order ( $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_N$ ), we have

$$\lambda_k(\mathbf{A}) \leq \lambda_k(\mathbf{B}) \leq \lambda_{k+N-r}(\mathbf{A}), \quad k = 1, 2, \dots, r. \quad (15)$$

It then follows from the above that  $\lambda_1(\mathbf{A}) \leq \lambda_1(\mathbf{B})$  and  $\lambda_r(\mathbf{B}) \leq \lambda_N(\mathbf{A})$ , implying that

$$\frac{\eta(\mathbf{A})}{\eta(\mathbf{B})} \geq 1.$$

### III. NEW EFFICIENT BLIND MMSE ADAPTIVE DETECTOR

Despite the above encouraging result, the blind MOE-LMS algorithm of [2] suffers from the characteristic slow convergence

of the reference driven MSE-LMS algorithm. This can be qualitatively appreciated from the following: the blind adaptive filter has  $N - 1$  coefficients (which for practical multiuser systems could be a large number, e.g., order of hundreds), of which only a subset  $(K - 1)$  is effectively used for interference suppression. In high signal-to-noise ratio (SNR) and low signal-to-interference (SIR) situations, the remaining  $N - (K - 1)$  degrees of freedom (that correspond to the “noise” dimensions) contribute little to improving the output SINR, while being the primary reason for the slow convergence. Accordingly, we seek a new and improved algorithm that requires only  $K - 1$  coefficients, in anticipation of corresponding improvements in transient convergence, particularly for lightly loaded systems  $[(K/N) \leq 0.5]$ .

We next demonstrate that the optimum MOE vector lies in the “signal” subspace—and hence an adaptive algorithm that constrains the adaptation coefficient to lie in this subspace achieves the desired improvement in convergence rates.<sup>1</sup>

We rewrite (4) as

$$\min_{\mathbf{c}_1} \text{MOE} = E(\mathbf{c}_1^T \mathbf{R}_y \mathbf{c}_1) \text{ subject to } \mathbf{c}_1^T \mathbf{s}_1 = 1 \quad (16)$$

and utilize the eigen-decomposition (EVD)

$$\mathbf{R}_y = \mathbf{E}_s \Lambda_s \mathbf{E}_s^T + \mathbf{E}_n \Lambda_n \mathbf{E}_n^T \quad (17)$$

$$= \mathbf{R}_s + \mathbf{R}_n = \mathbf{A} \mathbf{S} \mathbf{A}^T + \sigma_n^2 \mathbf{I}_N \quad (18)$$

where  $\mathbf{S} \in \mathcal{R}^{K \times K}$  is the covariance matrix of the users normalized spreading codes, and  $\mathbf{E}_s \in \mathcal{R}^{N \times K}$ ,  $\mathbf{E}_n \in \mathcal{R}^{N \times N-K}$  are orthogonal matrices whose columns span two orthogonal subspaces, denoted as the signal and noise subspaces, respectively, i.e.,  $\mathbf{E}_s^T \mathbf{E}_n = \mathbf{0}$ . Further, the matrix  $\mathbf{E} = [\mathbf{E}_s, \mathbf{E}_n]$  is orthogonal, i.e.,  $\mathbf{E} \mathbf{E}^T = \mathbf{E}^T \mathbf{E} = \mathbf{I}$ , implying that  $\mathbf{E}_s \mathbf{E}_s^T + \mathbf{E}_n \mathbf{E}_n^T = \mathbf{I}$ . The corresponding signal and noise eigenvalues, denoted by  $\lambda_{s,i}$ ,  $i = 1, \dots, K$ ;  $\lambda_{n,i}$ ,  $i = 1, \dots, N - K$ , are the elements of  $\Lambda_s$ ,  $\Lambda_n$ , respectively. Note from (17), it is clear that  $\lambda_{y,i} = \lambda_i(\mathbf{R}_s) + \sigma_n^2$ ,  $1 \leq i \leq K$  where  $\lambda_i(\mathbf{R}_s)$  are the  $K$  nonzero eigenvalues of signal-only component  $\mathbf{R}_s$ , and  $\lambda_{n,i} = \sigma_n^2$ . Finally, the column spaces spanned by  $\mathbf{A}$  and  $\mathbf{E}_s$  are identical. Since,  $\mathbf{s}_1 \in \text{Sp}\{\mathbf{A}\}$  (i.e., lies in the signal subspace), it follows that  $\mathbf{s}_1^T \mathbf{E}_n = \mathbf{0}$ .

Utilizing the EVD of  $\mathbf{R}_y$ , we reparametrize the unknown coefficient vector as  $\mathbf{c}_{\text{moe}} = \mathbf{E}_s \mathbf{c}_s + \mathbf{E}_n \mathbf{c}_n$ , where  $\mathbf{c}_s \in \mathcal{R}^K$ ,  $\mathbf{c}_n \in \mathcal{R}^{N-K}$ , respectively. Substituting the above in (16) yields

$$\min_{\mathbf{c}_s, \mathbf{c}_n} \text{MOE} = (\mathbf{c}_s^T \mathbf{E}_s^T + \mathbf{c}_n^T \mathbf{E}_n^T) (\mathbf{E}_s \Lambda_s \mathbf{E}_s^T + \mathbf{E}_n \Lambda_n \mathbf{E}_n^T) \cdot (\mathbf{E}_s \mathbf{c}_s + \mathbf{E}_n \mathbf{c}_n) \quad (19)$$

$$= \mathbf{c}_s^T \Lambda_s \mathbf{c}_s + \mathbf{c}_n^T \Lambda_n \mathbf{c}_n \quad (20)$$

$$\text{s.t. } \mathbf{s}_1^T \mathbf{E}_s \mathbf{c}_s + \mathbf{s}_1^T \mathbf{E}_n \mathbf{c}_n = \mathbf{s}_1^T \mathbf{E}_s \mathbf{c}_s = 1.$$

<sup>1</sup>Note that the MMSE detector also lies in the signal subspace, as shown in [8].

Since  $\mathbf{c}_n$  does not appear in the constraint, it follows that the optimum value of  $\mathbf{c}_{n, \text{opt}} = \mathbf{0}$ , yielding the equivalent simpler minimization problem

$$\min_{\mathbf{c}_s} \text{MOE} = E(\mathbf{c}_s^T \mathbf{E}_s^T \mathbf{y}(i))^2 = E(\mathbf{c}_s^T \mathbf{E}_s^T \mathbf{R}_y \mathbf{E}_s \mathbf{c}_s) \quad (21)$$

s.t.  $\tilde{\mathbf{s}}_1^T \mathbf{c}_s = 1$ , where  $\tilde{\mathbf{s}}_1^T = \mathbf{s}_1^T \mathbf{E}_s$ . The optimum coefficient vector  $\mathbf{c}_{\text{moe}, o} = \mathbf{c}_{s, o}$  lies in the signal subspace and motivates an adaptive algorithm that adapts only a  $K - 1$ -dimensional vector constrained to the signal subspace. We initially assume that an *exact* basis for this subspace is available in the form of the columns of the matrix  $[\mathbf{s}_1, \mathbf{U}]$  where  $\mathbf{U} \in \mathcal{R}^{N \times K-1}$ , where  $\mathbf{s}_1^T \mathbf{U} = \mathbf{0}$ , and the columns of  $\mathbf{U}$  are orthonormal, i.e.,  $\mathbf{U}^T \mathbf{U} = \mathbf{I}$ . Hence, the coefficient vector may be efficiently represented as  $\mathbf{c}_1 = \mathbf{s}_1 + \mathbf{U} \alpha$ , where the optimal  $\alpha \in \mathcal{R}^{K-1}$  is to be determined iteratively. Note that this model assumes that the desired user’s code vector  $\mathbf{s}_1$  is exactly known at the receiver.

Accordingly, we propose the following modified optimization problem:

$$\min_{\alpha} \text{MOE} = E(\mathbf{c}_1^T \mathbf{y}(i))^2 \quad \text{subject to } \mathbf{c}_1 = \mathbf{s}_1 + \mathbf{U} \alpha \quad (22)$$

where the assumption  $\mathbf{s}_1^T \mathbf{U} = \mathbf{0}$  implies that  $\mathbf{c}_1^T \mathbf{s}_1 = 1$ .

We seek a gradient descent procedure based on (22). Taking the derivative of the objective function yields the update

$$\alpha(i+1) = \alpha(i) - \mu \nabla_{\alpha(i)} (\mathbf{c}_1^T(i) \mathbf{y}(i))^2 \quad (23)$$

$$= \alpha(i) - 2\mu \mathbf{z}(i) \mathbf{U}^T \mathbf{y}(i) \quad (24)$$

where  $\mathbf{z}(i) = (\mathbf{s}_1(i) + \mathbf{U} \alpha(i))^T \mathbf{y}(i)$  is the current output of the adaptive filter.

Clearly, an estimate of the interference subspace  $\hat{\mathbf{U}}$  is required in practice. We assume that a one-shot estimate  $\hat{\mathbf{U}}$  of the “signal” subspace is extracted from the sample covariance matrix based on  $N_{in}$  samples of the received vector  $\mathbf{y}$  with the desired signal component  $\mathbf{s}_1$  removed, i.e.,

$$\tilde{\mathbf{y}}_k = \mathbf{y}_k - (\mathbf{y}_k^T \mathbf{s}_1) \mathbf{s}_1 \quad (25)$$

$$\hat{\mathbf{R}}_y = \frac{1}{N_{in}} \sum_{k=1}^{N_{in}} \tilde{\mathbf{y}}_k \tilde{\mathbf{y}}_k^T. \quad (26)$$

The matrix  $\hat{\mathbf{R}}_y$  is of rank  $N - 1$ , and its interference subspace is spanned by the  $K - 1$  eigenvectors corresponding to the  $K - 1$  largest eigenvalues.

#### A. Analysis of Subspace Algorithm

The reduced dimension update (23) can be rewritten as

$$\alpha(i+1) = \alpha(i) - 2\mu \mathbf{U}^T \mathbf{y}(i) \mathbf{y}^T(i) \mathbf{U} \alpha(i) - 2\mu \mathbf{U}^T \mathbf{y}(i) \mathbf{y}^T(i) \mathbf{s}_1. \quad (27)$$

The above iteration is reminiscent of the standard LMS update but for a modified input  $\mathbf{U}^T \mathbf{y}(i)$ . To see this clearly, we take expectations of both sides, yielding the time-evolution of the average coefficient vector

$$E[\alpha(i+1)] = [\mathbf{I} - 2\mu \mathbf{U}^T \mathbf{R}_y \mathbf{U}] E[\alpha(i)] - 2\mu \mathbf{U}^T \mathbf{R}_y \mathbf{s}_1. \quad (28)$$

Comparing (28) with the standard LMS update [9] reveals an identical form of the update with a modified input covariance matrix  $\tilde{\mathbf{R}}_y = \mathbf{U}^T \mathbf{R}_y \mathbf{U}$ . As a result, we may use the well-known results relating to the two important characteristics of the LMS algorithm as follows: 1) convergence of the transient response; 2) steady-state excess MSE (or misadjustment) with minor modifications. Both these figures of merit are governed by the eigenvalues  $\{\tilde{\lambda}_i(\tilde{\mathbf{R}}_y)\}$ ,  $i = 1, \dots, K$ . We again invoke the Poincare separation theorem [10] that relates the eigenvalues  $\tilde{\lambda}_i$  with  $\lambda_i(\mathbf{R}_y)$  as follows.

*Theorem:* For a symmetric matrix,  $\mathbf{A} \in \mathbf{R}^{N \times N}$ , and  $\mathbf{U} \in \mathbf{R}^{N \times K}$  ( $K < N$ ), where  $\mathbf{U}^T \mathbf{U} = \mathbf{I}_K$ , let  $\mathbf{B} = \mathbf{U}^T \mathbf{A} \mathbf{U}$ . Then, if the eigenvalues of  $\mathbf{A}$  ( $\mathbf{B}$ ) are arranged in increasing order ( $\lambda_1 \leq \lambda_2 \leq \dots$ )

$$\lambda_l(\mathbf{A}) \leq \lambda_l(\mathbf{B}) \leq \lambda_{l+N-K}(\mathbf{A}), \quad l = 1, 2, \dots, K. \quad (29)$$

Denote the eigenvalue spread of  $\tilde{\mathbf{R}}_y$  as  $\rho(\tilde{\mathbf{R}}_y) = (\tilde{\lambda}_K/\tilde{\lambda}_1)$ , and that of  $\mathbf{R}_y$  as  $\rho(\mathbf{R}_y) = (\lambda_N/\lambda_1)$ . Direct application of the above (with minor manipulations) yields that

$$\rho(\tilde{\mathbf{R}}_y) \leq \rho(\mathbf{R}_y) \quad (30)$$

indicating that the convergence of the new algorithm is always superior (and depending on the eigenvalue spread of the original input covariance matrix  $\mathbf{R}_y$ , possibly substantially so) to that of the algorithm in [2].

Further, it is well known that with stochastic gradient-type algorithms, the MSE at steady state  $\xi(\infty)$  exceeds  $\xi_{\min}$  (the minimum MSE corresponding to the Wiener solution) due to the tap weight fluctuations at convergence. The relative excess MSE  $M$  at steady state for the LMS algorithm is well known to be given by [9]

$$M = \frac{\xi(\infty) - \xi_{\min}}{\xi_{\min}} = \frac{K}{1 - K} \quad (31)$$

where  $K = \sum_{i=1}^N (\mu \lambda_i / (1 - 2\mu \lambda_i)) < 1$ , and  $0 < \mu < 1/(2\lambda_i)$ , such that each term in the above summation is positive and less than one. It straightforwardly follows from (31) that  $M$  is monotonic increasing with  $K$ . Accordingly, the relative excess MSE  $\tilde{M}$  for the subspace LMS algorithm is given by

$$\tilde{M} = \frac{\tilde{\xi}(\infty) - \xi_{\min}}{\xi_{\min}} = \frac{\tilde{K}}{1 - \tilde{K}} \quad (32)$$

where  $\tilde{K} = \sum_{i=1}^K (\mu \tilde{\lambda}_i / (1 - 2\mu \tilde{\lambda}_i))$ . Using the second inequality in the Poincare separation theorem yields

$$\frac{\mu \tilde{\lambda}_i}{1 - 2\mu \tilde{\lambda}_i} \leq \frac{\mu \lambda_{N+i-K}}{1 - 2\mu \lambda_{N+i-K}}. \quad (33)$$

Thus

$$\tilde{K} = \sum_{i=1}^K \frac{\mu \tilde{\lambda}_i}{1 - 2\mu \tilde{\lambda}_i} \leq \sum_{i=1}^K \frac{\mu \lambda_{N+i-K}}{1 - 2\mu \lambda_{N+i-K}} \quad (34)$$

$$\leq \sum_{i=1}^N \frac{\mu \lambda_i}{1 - 2\mu \lambda_i} = K. \quad (35)$$

Since  $\tilde{K} \leq K$ , it follows from the monotonic dependence of  $M$  on  $K$  that  $\tilde{M} \leq M$ . Since both approaches—that of [2] and that presented here—achieve the same  $\xi_{\min}$  ideally, it follows that the MSE achieved by the new subspace algorithm at convergence is always less than that of [2]. We have thus demonstrated analytically that the new algorithm outperforms that of [2] with respect to both initial convergence speed and the final MSE (respectively, final SINR, which is known to be monotonic decreasing with MSE).

While both algorithms achieve identical SINR, the choice of the (blind) adaptive algorithm results in different amounts of additional excess MSE (respectively, lowered SINR). The result above demonstrates that the subspace-based blind algorithm offers potential improvements over the [2] detector in this regard. The extent of such improvements are explored in the numerical/simulation examples presented in the next section.

## B. Performance Evaluation

In all examples, the power of the desired signal is normalized to unity (0 dB). All  $K - 1$  interfering users have equal power ( $P_i$ ), while the additive Gaussian noise vector have individual variance  $\sigma_n^2$  components. The users' spreading codes were randomly generated but kept fixed over different simulation runs. The signal-to-interference ratio SIR (decibels) is defined by  $\text{SIR} = -10 \log P_i$ , while the SNR is defined by  $\text{SNR}(\text{dB}) = -10 \log \sigma_n^2$ . The scenarios of greatest interest are those where the desired user is near-far dominated (i.e.,  $\text{SIR} \ll 0$  dB) over a range of SNR's. The interference subspace is assumed to be unknown but constant in Figs. 1–4; hence, an (orthonormal)  $K - 1$  basis set  $\mathbf{U}_I$  for the interference subspace is obtained once initially via eigendecomposition of the *modified* sample covariance matrix from (25) and is not updated subsequently during the detector iteration.<sup>2</sup> It was found that typically  $N_{in} = 30$ –50 samples sufficed for the sample covariance estimator used in the eigendecomposition for the SNR scenarios studied.

*Example 1:* Fig. 1 shows simulation results for a lightly loaded system with  $K = 3$ ,  $N = 15$ , using  $\text{SIR} = -10$  dB,  $\text{SNR} = 10$  dB for the two algorithms, with identical step-size  $\mu = 0.01$ . The maximum (ideal) output SINR is shown (7.1 dB) for comparison, whereas that achieved by the two algorithms are 4.6 and 4.85 dB, respectively. Note the significant losses (of 2.5 and 2.25 dB) vis-a-vis the optimum for both algorithms,

<sup>2</sup>Note that the removal of the desired user ( $\mathbf{s}_1$ ) from the total received signal reveals a  $K - 1$ -dimension interfering signal subspace (corresponding to the dominant eigenvalues) upon eigendecomposition.

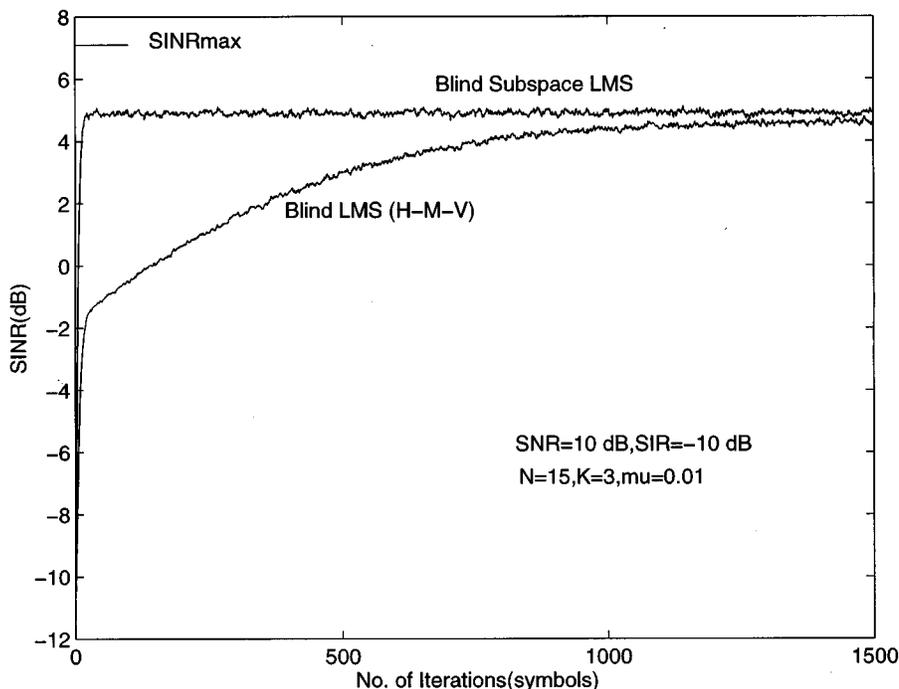


Fig. 1. Output SINR for blind LMS and subspace LMS algorithm ( $K = 3, N = 15$ ).

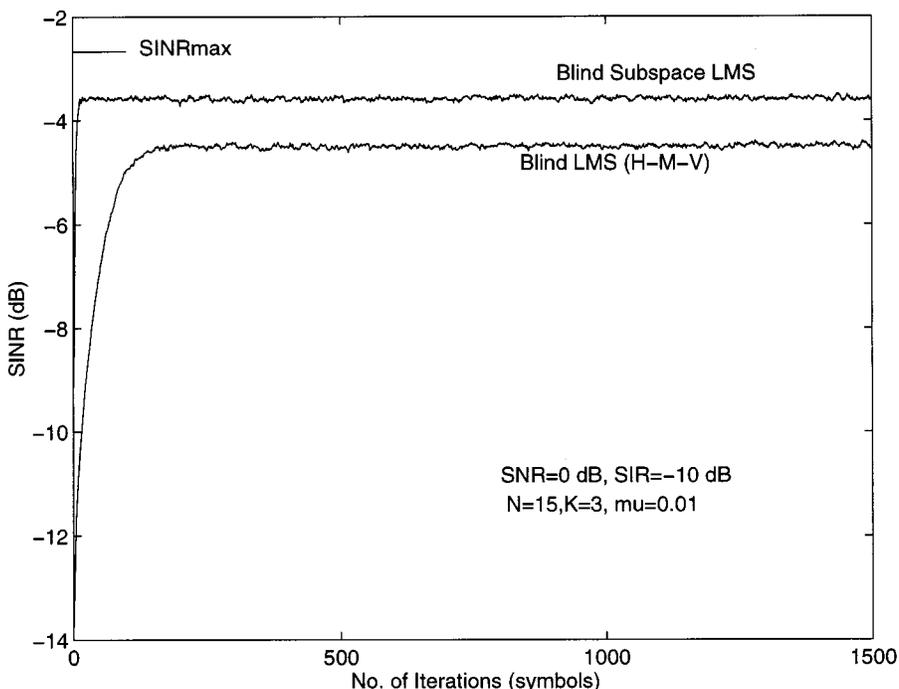


Fig. 2. Output SINR for LMS and subspace LMS algorithm ( $K = 3, N = 15$ ).

and that the blind subspace algorithm offers a minimal gain of 0.25 dB in this case; however, the improvement in convergence speed for the subspace algorithm is quite significant. We note that the blind LMS algorithm of [2] suffers from the same slow transient convergence as the reference driven LMS algorithm. The time constant for the smallest eigenvalue (noise eigenvalues = 0.1) is given by  $\tau_{\max} = 1/(0.01 \times 0.1) = 1000$  for step-size  $\mu = 0.01$ , implying 1000 iterations prior to convergence—as supported by Fig. 1. On the other hand, the subspace LMS

converges in approximately 30 symbols. This corroborates with the time-constant for the smallest signal space eigenvalue for this example (computed to be  $\lambda_s(2) = 2.7$ ).

*Example 2:* In this example (results in Fig. 2), we consider a scenario with the same parameters as in Example 1 but with (lower)  $\text{SNR} = 0$  dB. For this case, the maximum output  $\text{SINR} = -2.67$  dB, whereas that achieved by the two detectors are  $-4.54$  and  $-3.6$  dB. While the losses relative to the optimum SINR are lower, the subspace LMS algorithm now

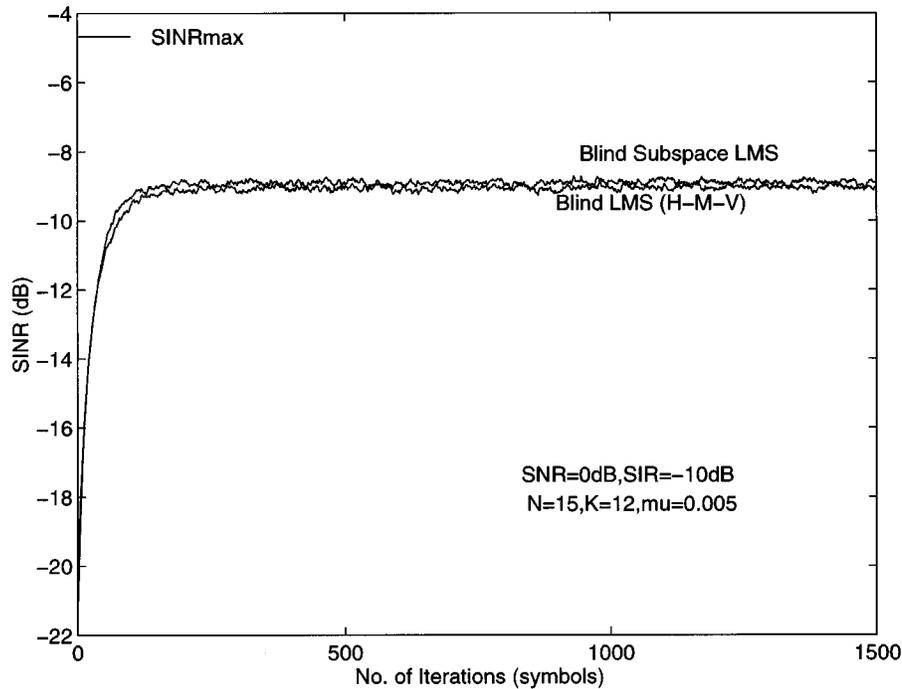


Fig. 3. Output SINR for LMS and subspace LMS Algorithm ( $K = 12$ ,  $N = 15$ ).

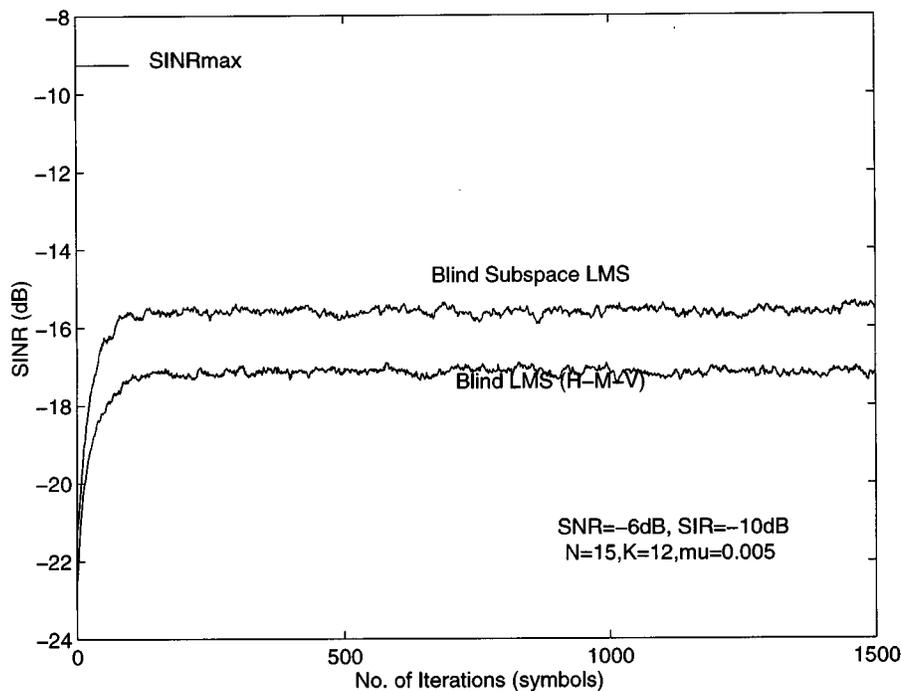


Fig. 4. Output SINR for LMS and subspace LMS Algorithm ( $K = 12$ ,  $N = 15$ ).

has a (nearly) 1 dB advantage in SINR. The subspace LMS algorithm still converges in approximately 30 symbols (since the convergence rate for this algorithm is independent of the additive noise variance), while the blind LMS now converges in approximately 200 symbols. This example suggests that as SNR decreases, the relative improvement (in terms of asymptotic SINR) of the subspace algorithm vis-a-vis [2] is more significant in terms of final SINR, and less by way of convergence rates.

The next two examples in Figs. 3 and 4 examine a heavily loaded system ( $N = 15$ ,  $K = 12$ ) for two low SNR's and compare with the lightly loaded case.

*Examples 3 and 4:* We first consider (as in Example 1)  $\text{SNR} = 0$  dB,  $\text{SIR} = -10$  dB (Fig. 3). For this, the maximum  $\text{SINR} = -4.68$  dB, and both algorithms achieve the same approx.  $\text{SINR} = -9.0$  dB at convergence at approximately the same rate (200 symbols). Note the significant loss of 4.3 dB with respect to the maximum output SINR, which exceeds that for the lightly

loaded system scenarios in Figs. 1 and 2. Further lowering the SNR (−6 dB) as in Fig. 4 leads to 1.8-dB performance gain of the subspace algorithm; however, the convergence rate is identical for both algorithms and is unchanged from that in Example 3. This is understood by examining the signal space eigenvalues for  $(K/N) \rightarrow 1$ . For larger number of users with randomly generated codes, the normalized code cross correlations are high, and hence, the smallest signal eigenvalue approaches the noise eigenvalue (from above, since the former is always greater). As a result, the convergence times of the two algorithms are approximately equal and largely insensitive to changes in input SNR.

In summary, the following insights as to performance characteristics of the subspace LMS algorithm can be stated as follows. 1) It yields significantly faster convergence times vis-a-vis the blind LMS implementation of [2] for *low loading* and modest-to-high SNR; this advantage is surrendered at higher loads. 2) The asymptotic SINR loss (vis-a-vis optimum SINR) increases monotonically with input SNR and is significant at low SNR's (<0 dB); in this range, the subspace LMS algorithm provides noticeable gains in SINR at convergence over the H-M-V algorithm. Note that for highly loaded systems, there appears to be little performance advantage of the subspace LMS algorithm vis-a-vis H-M-V for “pragmatic” ranges of input SNR and SIR values, such as > 0 dB and > −3 dB, respectively, as can be extrapolated from Fig. 3.

#### IV. CONCLUDING REMARKS

A new adaptive detector for interference suppression constrained to the signal subspace was proposed. Analysis of this detector's performance suggests gains in convergence speed or steady-state SINR depending on system load and input SNR, vis-a-vis the blind adaptive detector of [2]. The new detector is based on perfect knowledge of the desired user's code  $\mathbf{s}_1$  just as [2]; however, some decreased sensitivity to perturbations is anticipated. Quantifying this effect and development of newer robust algorithms are promising avenues for future consideration. Further, for fast fading scenarios (where the channel must be tracked), the subspace estimate  $\hat{U}$  must be performed within the loop of the detector update (for online implementation). The performance analysis of such adaptive detectors remains an important issue for future investigation.

#### APPENDIX EQUIVALENCE OF (7) AND (12)

We show that (7) and (12) are equivalent in the sense that the mean coefficient adaptation possesses identical stationary points. To do so, premultiply (12) by  $U$  and add  $\mathbf{s}_1$  to both sides, yielding

$$\mathbf{c}_1(i+1) = \mathbf{c}_1(i) - 2\mu U U^T \mathbf{y}(i) \mathbf{y}^T(i) \mathbf{c}_1(i).$$

Taking expectation of both sides yields

$$E[\mathbf{c}_1(i+1)] = [\mathbf{I} - 2\mu U U^T \mathcal{R}_y] E[\mathbf{c}_1(i)] \quad (36)$$

using the independence of  $\mathbf{y}(i)$  and  $\mathbf{c}_1(i)$ . It follows that asymptotically, the stationary points of (36) are given by letting  $\lim_{i \rightarrow \infty} E[\mathbf{c}_1(i+1)] = \lim_{i \rightarrow \infty} E[\mathbf{c}_1(i)] = E[\mathbf{c}_1(\infty)]$

$$U U^T \mathcal{R}_y E[\mathbf{c}_1(\infty)] = \mathbf{0}. \quad (37)$$

Note that  $rk.[U U^T] = N - 1$ , implying a null space of dimension equal to one. Further by construction, we know that  $U U^T \mathbf{s}_1 = \mathbf{0}$ . Hence, the stationary point of (37) must be given by

$$\mathcal{R}_y E[\mathbf{c}_1(\infty)] = \gamma \mathbf{s}_1$$

or equivalently,  $E\mathbf{c}_1(\infty) = \gamma \mathcal{R}_y^{-1} \mathbf{s}_1$ . Invoking the constraint  $\mathbf{s}_1^T \mathbf{c}_1(\infty) = 1$  yields  $\gamma = 1/(\mathbf{s}_1^T \mathcal{R}_y^{-1} \mathbf{s}_1)$ , thereby establishing the desired equivalence in the mean of the two updates.

#### REFERENCES

- [1] U. Madhow and M. L. Honig, “MMSE interference suppression for direct-sequence spread-spectrum CDMA,” *IEEE Trans. Commun.*, vol. 42, pp. 3178–3188, Dec. 1994.
- [2] M. Honig, U. Madhow, and S. Verdú, “Blind adaptive multiuser detection,” *IEEE Trans. Inform. Theory*, vol. 41, pp. 944–960, July 1995.
- [3] P. B. Rapajic and B. S. Vucetic, “Adaptive receiver structures for asynchronous CDMA systems,” *IEEE J. Select. Areas Commun.*, vol. 12, pp. 685–697, May 1994.
- [4] S. L. Miller, “An adaptive direct-sequence code-division multiple-access receiver for multiuser interference rejection,” *IEEE Trans. Commun.*, vol. 43, pp. 1746–1755, Feb./Mar./Apr. 1995.
- [5] M. Honig, “Adaptive linear interference suppression for packet DS-SS-CDMA,” *Eur. Trans. Telecommun.*, vol. 9, pp. 173–182, Mar./Apr. 1998.
- [6] —, “A comparison of subspace adaptive filtering techniques for ds-ss-ss interference suppression,” in *Proc. Military Communications Conf.*, Monterey, CA, 1997, pp. 836–840.
- [7] X. Wang and H. V. Poor, “Blind multiuser detection: A subspace approach,” *IEEE Trans. Inform. Theory*, vol. 44, pp. 677–690, Mar. 1998, to be published.
- [8] —, “Blind adaptive interference suppression for CDMA communications based on eigenspace tracking,” in *Proc. Conf. Information Sciences and System*, Baltimore, MD, Mar. 1997.
- [9] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [10] R. Horn and C. A. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985, ch. 4.



**Sumit Roy** (S'84–M'88) received the B. Tech. degree from the Indian Institute of Technology, Kanpur, India, in 1983, and the M. S. and Ph. D. degrees from the University of California, Santa Barbara, in 1985 and 1988, respectively, all in electrical engineering. He also received the M.A. degree in statistics and applied probability in 1988.

His previous academic appointments were at the Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, and at the University of Texas, San Antonio. Currently, he is an Associate

Professor of Electrical Engineering, University of Washington, Seattle, where his research interests include analysis/design of communication systems/networks, with a topical emphasis on next-generation mobile/wireless networks.

Dr. Roy's activities for the IEEE Communications Society include membership of several technical committees and serving as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS and for the Wireless Series, the IEEE JOURNAL OF SELECTED AREAS IN COMMUNICATIONS.