

Blind Adaptive Noncoherent Multiuser Detection for Nonlinear Modulation

Deepak Das and Mahesh K. Varanasi, *Senior Member, IEEE*

Abstract—Noncoherent multiuser detection for nonlinear modulation was recently studied and the idea of phase-independent noncoherent decorrelation was introduced and three post-decorrelative detectors were obtained and analyzed. However, their implementation requires the knowledge of the signature waveforms of all the users, which may be available only for centralized implementation. In this paper, we obtain a *blind* adaptive noncoherent decorrelative detector for nonlinear modulation that is suitable for distributed implementation with the knowledge of only the normalized signals of the desired user and the additive noise variance. This detector is based on the stochastic approximation method and does not require the overhead of any kind of “training.” Two adaptive algorithms are developed, one guided by every signal in the desired user’s signal set individually, and the other by the user’s entire signal space. While this paper focuses on the particular problem of blind adaptive noncoherent decorrelative detection, it addresses a more general adaptation issue, namely, that of improving convergence properties of an adaptive scheme by effectively using all the information that is known, and adapting only to the part of the desired solution that is truly unknown. Convergence is shown in the mean squared error sense for both the fixed step-size and time-varying step-size versions of the two algorithms.

Index Terms—Blind adaptive receivers, multiuser detection, noncoherent detection, nonlinear modulation, stochastic approximation.

I. INTRODUCTION

NONCOHERENT detection for nonlinear modulation for the multiuser channel was introduced in [1] in the context of the additive white Gaussian noise (AWGN) channel. The need for noncoherent detection arises in applications where factors, such as oscillator phase instability, mobility of the transmitter and/or receiver, etc., cause rapid fluctuations of the carrier phase at the receiver that are difficult to track, and/or when phase-tracking is seen as an expensive proposition. A general nonlinear modulation technique that lends itself to noncoherent detection is nonorthogonal multipulse modulation (NMM) where an M -ary symbol is transmitted by sending one of M possibly nonorthogonal and unequal-energy waveforms [1]. The usual equal-energy orthogonal multipulse modulation (OMM) method is a special case of NMM. The motivation for considering

the more general nonorthogonal signaling framework is that the correlations between the signals in NMM can be regarded as design parameters [to increase bandwidth efficiency of OMM (cf. [2])] or they can arise as a result of a distorting channel in even systems that use OMM. The multiuser channel that we consider has several NMM-based transmitters sending data independently and simultaneously so that a superposition of several NMM waveforms plus AWGN is observed at the receiver. The multiple-access technique employed is correlated waveform multiple access (CWMA) [1], where the signature waveforms employed by the different users can be correlated as well.

Since the complex signal amplitudes are allowed to vary randomly and even independently from symbol to symbol, the blind adaptive noncoherent detection algorithms of this paper are also applicable to slow and even moderately fast flat-fading channels where the only requirement is that the zero-mean complex fading amplitudes (as in Rayleigh fading) remain nearly constant over at least one symbol interval. Moreover, the detection strategies that we discuss need to be modified to combat intersymbol interference in the context of frequency-selective fading channels, an issue we do not address here.

NMM can be viewed as pseudolinear modulation in the expanded signal space of all the signals of all the users. This idea allows the specification of an energy-and-phase-independent decorrelative operation in the expanded signal space to remove the multiple-access interference (MAI) completely [1]. Three post-decorrelative detectors, namely, the optimum, the asymptotically optimum, and the generalized likelihood ratio test (GLRT) based detectors were obtained in [1]. These detection schemes can be extended to *linearly dependent* signaling [3]–[5]. Improvement in terms of spectral efficiency and probability of error performance beyond these initial solutions to the multiuser NMM problem can be realized through signal design as in [2] and superior detection strategies. In the latter context, the reader can refer to the (deterministic) noncoherent decision feedback decorrelative detector for NMM in [4].

The deterministic NMM decorrelator (there are M decorrelators per user), as formulated in this paper, involves a computation of the pseudoinverse of the signal matrix consisting of the MK signal vectors in the expanded signal space. This requires the knowledge of all the signals even to compute the M decorrelators of one user. Receivers at individual users, however, will most likely have knowledge of system parameters pertinent only to themselves and would, therefore, have to *adapt* to their decorrelators in a distributed manner. In this paper, we present two computationally simple blind adaptive noncoherent decorrelators (BANDs) for NMM that are suitable for distributed implementation requiring just the normalized signature waveforms

Paper approved by K. B. Letaief, the Editor for Wireless Communications of the IEEE Communications Society. Manuscript received September 16, 1999; revised February 3, 2000 and April 15, 2000. This work was supported in part by the National Science Foundation under Grant NCR-9725778 and in part by the Army Research Office (ARO) under Grant DADD19-99-1-0291. This paper was presented in part at the 37th Annual Allerton Conference on Communication, Control and Computing, Allerton, IL, September 1999.

The authors are with the Department of Electrical and Computer Engineering, University of Colorado, Boulder, CO 80309-0425 USA (e-mail: dasd@ucsu.colorado.edu; varanasi@schof.colorado.edu).

Publisher Item Identifier S 0090-6778(00)09877-9.

of the user of interest and the estimated additive noise variance. Moreover, the adaptation does not require any kind of "training." The results herein can also be used with the optimum and asymptotically optimum post-decorrelative detectors provided the energies of the M signals of the user of interest are also known. The adaptation algorithms are based on the stochastic approximation technique [6]. Even in a centralized system where the signals are slowly time varying, it may be infeasibly prohibitive to compute the pseudoinverse every time the changing signals are estimated, and one would need to use the adaptation algorithms of this paper.

Past work in blind adaptive multiuser detection deals with coherent detection for linear modulation (where each user employs different levels of a single waveform). Several blind adaptive schemes to estimate the two well-known decorrelating [7] and minimum mean squared error (MMSE) [8] detectors have been obtained that alleviate the requirement of system information to varying extents. For instance, a blind adaptive algorithm based on the minimization of the output energy was given in [9] and shown to be equivalent to adapting to the MMSE detector, requiring only the signal waveform of the user of interest. In [10], an adaptive multiuser detector based on the least-squares approach was proposed that converged to the decorrelating detector and was suitable for centralized implementation in that it needed the signature waveforms of all the users. In [11], blind algorithms based on signal subspace tracking were investigated and two algorithms which converged to the decorrelating and the MMSE multiuser detectors were proposed. A blind adaptive decorrelating detector (BADD), which required the knowledge of the signal of only the user of interest and the channel noise variance, was presented in [12]. The decision rule in each of the above-mentioned schemes, in general, requires the received energy and the carrier phase (coherent detection) of the user of interest. Lack of knowledge of the carrier phase makes them unsuitable for noncoherent reception. Moreover, the decision rules for linear modulation are easily implemented once the detector has been estimated. This is no longer true for noncoherent detection for NMM where the decision rules can be fairly complicated. As we will see later in this paper, blind NMM noncoherent decorrelative detection has to not only estimate the decorrelator blindly, but also has to implement the post-decorrelative decision algorithm in a blind fashion. A generalization of the subspace tracking approach in [11] can be found in [5], which addresses the issue of blind adaptive noncoherent detection for NMM. The subspace tracking method is computationally more intensive than the simpler stochastic approximation based adaptation strategies obtained in this paper.

In order to appreciate the difference between the two BANDs of this paper, we can view the NMM system equivalently as comprising of MK users organized in K M -user groups, so that in every group, only one user selected with uniform randomness transmits its signal in a given signaling interval. With this description, we can devise a blind adaptive decorrelator for each of the MK users, assuming it knows only its own signal. This adaptation strategy, guided by the individual signals of any user, is referred to as the *single-signal BAND* or simply as the SS-BAND. It can be seen as an extension of the BADD algorithm for linear modulation in [12]. In the equivalent MK user

description, we note that in each of the K groups of M pseudosignals each, the knowledge of all the M signals of the corresponding group (or user) is available. Therefore, it is more appropriate to devise a blind adaptive decorrelator that is guided by the group's entire signal space. This is the basis of the second BAND and we shall refer to it as the *group signal-space BAND* or, more succinctly, as the GS-BAND. This is the key contribution in this paper. The GS-BAND attempts in a novel way to use as much as possible the signal space knowledge that is available in the NMM problem, and consequently, we can show both analytically and numerically that it has superior convergence properties at low computational cost.

Beyond the particular multiuser NMM problem considered here, the general issue we deal with is that of trying to improve convergence properties of an adaptation scheme by effectively using all the information that is known, and adapting only to that part of the desired solution that is truly unknown. Without being seen in this light, this strategy has been employed, for instance, in deriving the minimum output energy (MOE)-based adaptation rule in [9], where the authors employ a canonical decomposition of the desired solution, and adapt to only the part that is orthogonal to the known signal. However, there has been no effort to show that the MOE-based rule is better than the simpler blind adaptation rule for the MMSE solution that was derived in [12]. In this paper, we shall explicitly show that the idea of decomposing the NMM decorrelator into two orthogonal parts and adapting only to the unknown part (as in the case of the GS-BAND) ensures better rate of convergence for the GS-BAND than that for the SS-BAND.

The convergence to the deterministic decorrelator [1] in the mean-squared error (MSE) sense, of a special Robbins–Monro [13] iteration-dependent step-size version of both the SS-BAND and the GS-BAND, follows along the lines of the classical proof of convergence of the stochastic approximation algorithm [6]. The proof of quasi-convergence of the fixed step-size version of the SS-BAND is similar to that of the linear BADD convergence proof in [12] (see also [14]). The fixed step-size version of GS-BAND is studied in detail in this paper and shown to converge with increasing iterations to the deterministic decorrelator with a limiting MSE that is proportional to the step-size chosen.

II. SYSTEM MODEL

We consider an NMM-CWMA system, i.e., each user transmits one of M possibly nonorthogonal signals to send $\log_2 M$ bits of information. The symbol synchronous superposition of K signals plus AWGN at the receiver admits the following low-pass representation:

$$r(t) = \sum_{k=1}^K \sqrt{E_{ki_k}} e^{j\phi_{ki_k}} s_{ki_k}(t) + n(t), \quad t \in [0, T] \quad (1)$$

where $n(t)$ is the AWGN with noise power spectral density (one-sided) of σ^2 , and $s_{ki_k}(t)$, $i_k \in \{1, \dots, M\}$ are the M complex, equiprobable, and normalized (to unit-energy) signature signals of user k , time-limited to a period T . i_k is the symbol transmitted by user k and E_{ki_k} and ϕ_{ki_k} denote the energy and phase of the i_k th signal of user k . It is implicitly assumed that the

phases remain constant over each signaling interval. No knowledge of the complex amplitudes $E_{k i_k} e^{j\phi_{k i_k}}$ or their distributions is assumed to be available at the receiver.

The low-pass received signal can be projected onto an N -dimensional basis of orthonormal functions ($N \geq MK$) to yield the following vector representation for the received signal:

$$\mathbf{r} = \sum_{k=1}^K \sqrt{E_{k i_k}} e^{j\phi_{k i_k}} \mathbf{s}_{k i_k} + \mathbf{n} \quad (2)$$

where $\mathbf{s}_{k i_k}$ and \mathbf{n} are the projections of the k th user's i_k th signal and the noise, respectively, on the orthonormal basis. For example, in a system where the signals have a direct-sequence spread-spectrum structure, the orthonormal functions could be the so-called chip waveforms. The noise vector \mathbf{n} is a zero-mean white complex Gaussian random vector with covariance matrix $\sigma^2 \mathbf{I}$, denoted as $\eta(0, \sigma^2 \mathbf{I})$. Representing user k 's information by the random vector $\mathbf{b}_k^T \triangleq [b_{k1}, \dots, b_{kM}]$ which equiprobably takes values in the set of M M -dimensional unit vectors (each unit vector corresponding to one of the M signals that user k could possibly send), the N -length received vector admits the following pseudolinear model:

$$\mathbf{r} = \mathbf{S} \mathbf{A} \mathbf{b} + \mathbf{n} \quad (3)$$

with $\mathbf{S} = [\mathbf{s}_{11} \ \dots \ \mathbf{s}_{1M} \ \dots \ \mathbf{s}_{K1} \ \dots \ \mathbf{s}_{KM}] \in \mathbf{C}^{N \times MK}$, $\mathbf{b}^T = [\mathbf{b}_1^T, \dots, \mathbf{b}_K^T]$, $\mathbf{A} = \text{diag}(\mathbf{A}_1, \dots, \mathbf{A}_K) \in \mathbf{C}^{MK \times MK}$, and $\mathbf{A}_k = \text{diag}\{\sqrt{E_{k1}} e^{j\phi_{k1}}, \dots, \sqrt{E_{kM}} e^{j\phi_{kM}}\}$.

The conventional (single-user) noncoherent multiuser detector forms the MK -length matched-filter outputs $\mathbf{y} = \mathbf{S}^\dagger \mathbf{r}$ (where \dagger denotes conjugate transposition), and then considers as decision statistic for user k the M -length subvector \mathbf{y}_k , i.e., matched filter outputs $M(k-1) + 1$ to Mk to apply the following decision rule:

$$\hat{i}_k = \arg \max_i |y_k(i)|^2, \quad i \in \{1, \dots, M\}. \quad (4)$$

Unfortunately, this detector is susceptible to *near-far* effects since it does not adequately compensate for MAI which is the contribution in \mathbf{y}_k from the information sent by the other users [1].

III. NONCOHERENT DECORRELATIVE DETECTION

Let the decorrelator of interest (in the expanded signal space) be denoted by the $MK \times N$ phase and energy independent complex matrix transformation \mathbf{C} , such that

$$\mathbf{z} = \mathbf{C}^\dagger \mathbf{r} = \mathbf{A} \mathbf{b} + \gamma \quad (5)$$

where γ is $\eta(0, \sigma^2 \mathbf{C}^\dagger \mathbf{C})$. In order to decouple the multiuser channel into K single-user channels, we base the decision as to which signal is transmitted by user k only on the M -length decision statistic \mathbf{z}_k given as

$$\mathbf{z}_k = \mathbf{C}_k^\dagger \mathbf{r} = \mathbf{A}_k \mathbf{b}_k + \gamma_k \quad (6)$$

where \mathbf{C}_k is the k th block of M columns of \mathbf{C} and γ_k is $\eta(0, \sigma^2 \mathbf{C}_k^\dagger \mathbf{C}_k)$. Note that this is an information-lossy operation that enables us to obtain a decision statistic for user k that is free of MAI but ignores the noise correlation with other users.

Letting $\mathbf{P}_k \triangleq (\mathbf{C}_k^\dagger \mathbf{C}_k)^{-1}$, it follows from [1], that the GLRT based post-decorrelative detector for user k determines the decision for user k , as follows:

$$\hat{i}_k = \arg \max_i \left\{ \frac{1}{P_k(i, i)} \left| \sum_{m=1}^M P_k(i, m) z_{km} \right|^2 \right\} \quad (7)$$

where $P_k(i, m)$ denotes the (i, m) th entry of the matrix \mathbf{P}_k . The computation of \mathbf{P}_k can be accomplished by decomposing $\mathbf{C}_k^\dagger \mathbf{C}_k$ into a product of lower and upper triangular matrices involving computational complexity of $\mathbf{O}(NM^2/2)$ when $N \gg M$.

We now need to specify the exact form of the decorrelating transformation \mathbf{C}_k for user k , and note from (6) that \mathbf{c}_{km} , the m th column of \mathbf{C}_k , should satisfy the following condition:

$$\mathbf{S}^\dagger \mathbf{c}_{km} = \mathbf{e}_{M(k-1)+m} \quad (8)$$

where \mathbf{e}_i is the MK -length unit vector with a 1 in the i th position. Since the MK signal vectors are assumed to be linearly independent, there could be more than one solution for \mathbf{c}_{km} . However, it is easy to show that the unique solution lying in the column space of \mathbf{S} is the corresponding column of the pseudoinverse of the signal matrix \mathbf{S} , given as

$$\mathbf{c}_{km}^o = \mathbf{S}(\mathbf{S}^\dagger \mathbf{S})^{-1} \mathbf{e}_{M(k-1)+m} \quad (9)$$

and this is the decorrelator we are interested in. The adaptive algorithms to estimate the matrix \mathbf{C}_k^o that we discuss next have the following desirable properties: they require only the knowledge of the user-of-interest's M normalized signals, thereby facilitating distributed detection, avoid the estimation and inversion of $\mathbf{R}^{-1} = (\mathbf{S}^\dagger \mathbf{S})^{-1}$, and are "blind" in that they do not require a training phase. In any case, it is not clear how the commonly used technique of employing training sequences can be applied to adaptive noncoherent decorrelation for NMM (where the energies and the phases of the received signals are unknown).

IV. BLIND ADAPTIVE NONCOHERENT DECORRELATION

Let us define the $MK \times MK$ matrix $\mathbf{W} \triangleq E(\mathbf{A} \mathbf{b} \mathbf{b}^T \mathbf{A}^\dagger)$ where E is the expectation operator. The (i, j) th $M \times M$ block, \mathbf{W}_{ij} , of the matrix \mathbf{W} is given as $\mathbf{W}_{ij} \triangleq E(\mathbf{A}_i \mathbf{b}_i \mathbf{b}_j^T \mathbf{A}_j^\dagger)$ where $i, j \in \{1, \dots, K\}$. With independent NMM transmitters sending data simultaneously, we can assume that the phases of different users are independent and \mathbf{W}_{ij} is the zero matrix for $i \neq j$. Moreover, equiprobable signaling implies that \mathbf{W}_{ii} is a diagonal matrix whose m th diagonal element is E_{im}/M . We can now write the following equation for the output correlation:

$$E(\mathbf{r} \mathbf{r}^\dagger) = E[(\mathbf{S} \mathbf{A} \mathbf{b} + \mathbf{n})(\mathbf{S} \mathbf{A} \mathbf{b} + \mathbf{n})^\dagger] = \mathbf{S} \mathbf{W} \mathbf{S}^\dagger + \sigma^2 \mathbf{I} \quad (10)$$

where we have assumed that the noise term is independent of the received complex amplitudes and the information vector \mathbf{b} . In a slightly different scenario where the phases of all the users are the same (as in transmission for all users on a single carrier from a base station in a cellular network) $E(\mathbf{r} \mathbf{r}^\dagger)$ has the same form as in (10) if the sign of the transmitted signal of every user is switched randomly and equiprobably from one symbol interval to the next. This does not alter the symbol-error rate

performance of our detection strategy and moreover allows the remaining discussion in this paper to hold even for detection on the downlink in a cellular network.

A. SS-BAND

We will first obtain a blind adaptive algorithm for an arbitrary decorrelating solution \mathbf{c}_{km} , such that the adaptation is guided by \mathbf{s}_{km} individually. Starting with (8), we observe that

$$\mathbf{SWS}^\dagger \mathbf{c}_{km} = \frac{E_{km}}{M} \mathbf{s}_{km}. \quad (11)$$

It can be shown that if the MK signals are linearly independent, then all the solutions of (8) and (11) for \mathbf{c}_{km} coincide. Therefore, any iterative algorithm that converges to a solution of (11) also yields a solution of (8). If we now consider an estimate $\hat{\mathbf{c}}_{km}$ of the vector-valued parameter \mathbf{c}_{km} , and define a random vector $\mathbf{X}(\hat{\mathbf{c}}_{km})$, parameterized by the estimate, whose realization in the l th interval is given by $\mathbf{x}(\hat{\mathbf{c}}_{km}) = (\mathbf{r}(l)\mathbf{r}(l)^\dagger - \sigma^2\mathbf{I})\hat{\mathbf{c}}_{km}$, we can see at once, using (10), that

$$E[\mathbf{X}(\hat{\mathbf{c}}_{km})] = E[(\mathbf{r}(l)\mathbf{r}(l)^\dagger - \sigma^2\mathbf{I})\hat{\mathbf{c}}_{km}] = \mathbf{SWS}^\dagger \hat{\mathbf{c}}_{km}. \quad (12)$$

Using the stochastic approximation method [6], we can devise the following fixed step-size stochastic adaptation algorithm to estimate a scaled version of \mathbf{c}_{km} without the knowledge of the user energies:

$$\begin{aligned} \mathbf{c}_{km}(l+1) &= \mathbf{c}_{km}(l) - \mu((\mathbf{r}(l+1)\mathbf{r}(l+1)^\dagger - \sigma^2\mathbf{I}) \\ &\cdot \mathbf{c}_{km}(l) - \mathbf{s}_{km}), \quad m = 1, \dots, M. \end{aligned} \quad (13)$$

Not using the knowledge of the user energies in (13), however, causes the estimate of the decorrelator to converge at best to a scaled version of the actual decorrelator, scaled by a factor of M/E_{km} . Using the scaled decorrelator, we obtain scaled versions of the decision statistic z_{km} and the terms $P_k(i, i)$ and $P_k(i, m)$ in (7). It is left to the reader to verify that the decision rule in (7) yields the same decisions when the scaled terms are used in place of the unscaled ones. Since only each signal \mathbf{s}_{km} is used individually in the adaptation in (13), we refer to this algorithm as the SS-BAND. If we replace the fixed step-size μ by an iteration index dependent step-size sequence μ_l (satisfying some conditions stated later), we obtain the Robbins–Monro stochastic approximation based algorithm to estimate \mathbf{c}_{km} .

In the rest of this paper, we will denote the projection operators for matrix Ψ as $\mathcal{P}_\Psi = \Psi(\Psi^\dagger\Psi)^{-1}\Psi^\dagger$ and $\mathcal{P}_\Psi^\perp = \mathbf{I} - \mathcal{P}_\Psi$. A key point to note is that since the solution of (11) is not unique, we have to ensure that if the stochastic algorithm in (13) converges, it does so to \mathbf{c}_{km}^o (possibly scaled) as given by (9). This is guaranteed if $\mathbf{c}_{km}(l) = \mathcal{P}_\mathbf{S}\mathbf{c}_{km}(l)$ almost surely (a.s.), i.e., $\Pr[\mathbf{c}_{km}(l) \neq \mathcal{P}_\mathbf{S}\mathbf{c}_{km}(l)] = 0$. It can be shown that a sufficient condition for this is that $\mathbf{c}_{km}(0) = \mathcal{P}_\mathbf{S}\mathbf{c}_{km}(0)$. Note that for user k , any linear combination of its M signals is an acceptable value for $\mathbf{c}_{km}(0)$, including $\mathbf{c}_{km}(0) = \mathbf{0}$.

Computing the decision statistics $\{z_{km}(l+1) = \mathbf{c}_{km}^\dagger(l)\mathbf{r}(l+1)\}$, $m = 1, \dots, M$, the updation of the decorrelator estimate can be done efficiently with the following version of (13):

$$\begin{aligned} \mathbf{c}_{km}(l+1) &= (1 + \mu\sigma^2)\mathbf{c}_{km}(l) - \mu(\mathbf{r}(l+1) \\ &\cdot z_{km}^*(l+1) - \mathbf{s}_{km}), \quad m = 1, \dots, M. \end{aligned} \quad (14)$$

Further, with the stochastic approximation of \mathbf{P}_k , namely, $\mathbf{P}_{k,l+1} \triangleq (\mathbf{C}_k^\dagger(l+1)\mathbf{C}_k(l+1))^{-1}$, which can be computed efficiently as described after (7), user k 's decision in the $(l+1)$ th interval, is given by

$$\hat{i}_k(l+1) = \arg \max_i \left\{ \frac{1}{P_{k,l+1}(i, i)} \cdot \left| \sum_{m=1}^M P_{k,l+1}(i, m) z_{km}(l+1) \right|^2 \right\}. \quad (15)$$

B. Group SS-BAND

The adaptation algorithm in (13) is a single-signal-based adaptation strategy where the correction term at each iteration depends only on the signal \mathbf{s}_{km} . A more appropriate adaptation scheme for NMM would use the knowledge of all M signals. Without loss of generality, we will consider only the decorrelator for user 1. First, for notational simplicity, let us define the signal correlation matrix and its partitions as follows: $\mathbf{R} = \mathbf{S}^\dagger\mathbf{S}$, $\mathbf{S} = [\mathbf{S}_1 | \bar{\mathbf{S}}_1]$, $\mathbf{R}_A = \mathbf{S}_1^\dagger\mathbf{S}_1$, $\mathbf{R}_B = \mathbf{S}_1^\dagger\bar{\mathbf{S}}_1$, and $\mathbf{R}_C = \bar{\mathbf{S}}_1^\dagger\bar{\mathbf{S}}_1$, where clearly $\bar{\mathbf{S}}_1$ is composed of the signals of the remaining $K-1$ users. Let us also define two blocks along the diagonal of \mathbf{W} , namely \mathbf{W}_1 and $\bar{\mathbf{W}}_1$, where \mathbf{W}_1 is the $M \times M$ northwest block and $\bar{\mathbf{W}}_1$ is the $M(K-1) \times M(K-1)$ southeast block.

From the results on inversion of matrices [15], the conjugate transpose of the desired decorrelator matrix \mathbf{C}_1^o (given by the first M columns of the pseudoinverse of \mathbf{S}) can be written as

$$\begin{aligned} (\mathbf{C}_1^o)^\dagger &= \mathbf{R}_A^{-1}\mathbf{S}_1^\dagger - \mathbf{R}_A^{-1}\mathbf{R}_B \left(\mathbf{R}_B^\dagger\mathbf{R}_A^{-1}\mathbf{R}_B - \mathbf{R}_C \right)^{-1} \mathbf{R}_B^\dagger\mathbf{R}_A^{-1}\mathbf{S}_1^\dagger \\ &\quad + \mathbf{R}_A^{-1}\mathbf{R}_B \left(\mathbf{R}_B^\dagger\mathbf{R}_A^{-1}\mathbf{R}_B - \mathbf{R}_C \right)^{-1} \bar{\mathbf{S}}_1^\dagger. \end{aligned} \quad (16)$$

Defining the matrix $\mathbf{D}_1^\dagger \triangleq \mathbf{R}_A^{-1}\mathbf{R}_B \left(\mathbf{R}_B^\dagger\mathbf{R}_A^{-1}\mathbf{R}_B - \mathbf{R}_C \right)^{-1} \bar{\mathbf{S}}_1^\dagger$ and substituting in (16) for \mathbf{R}_B and \mathbf{R}_C in terms of $\mathbf{S}_1, \bar{\mathbf{S}}_1$, we obtain the following:

$$\mathbf{C}_1^o = \mathbf{S}_1\mathbf{R}_A^{-1} + \mathcal{P}_{\bar{\mathbf{S}}_1}^\perp\mathbf{D}_1. \quad (17)$$

Note that the matrix \mathbf{D}_1 captures the interfering signal space information that user 1 has no access to, in a completely distributed system. Consequently, this is the part of the decorrelator that we really need to adapt to. Denoting the m th column of \mathbf{D}_1 as \mathbf{d}_{1m} , and the m th column of \mathbf{R}_A^{-1} as \mathbf{q}_{1m} , we obtain the following representation of the m th decorrelator in terms of user 1's group of signals:

$$\mathbf{c}_{1m}^o = \mathbf{S}_1\mathbf{q}_{1m} + \mathcal{P}_{\bar{\mathbf{S}}_1}^\perp\mathbf{d}_{1m}. \quad (18)$$

We remark here that the orthogonal decomposition of \mathbf{c}_{1m}^o in (18) is a generalization, to the *group* signal space approach, of a single-signal-based orthogonal decomposition, $\mathbf{c}_{1m}^o = \mathbf{s}_{1m} + \pi_{1m}$, $\mathbf{s}_{1m} \perp \pi_{1m}$, as suggested in [9] in the context of the MOE solution for the coherent detection problem for linear modulation. The GS-BAND is not, however, an extension of the MOE rule because our detection strategy is a decorrelative one and not based on minimizing MSE or output energy.

In the case that user 1's signal subspace is orthogonal to the interfering signal subspace, we can exactly construct the corresponding decorrelators, \mathbf{c}_{1m}^o , with just the knowledge of \mathbf{S}_1 , i.e., $\mathbf{C}_1^o = \mathbf{S}_1 \mathbf{R}_A^{-1}$. However, user 1 may not know whether this is true of the signal set of its interferers and will still need to adaptively determine its decorrelator matrix.

Using (10) and (11), we obtain

$$E [\mathcal{P}_{\mathbf{S}_1}^\perp (\mathbf{r}(l+1)\mathbf{r}(l+1)^\dagger - \sigma^2 \mathbf{I}) (\mathbf{S}_1 \mathbf{q}_{1m} + \mathcal{P}_{\mathbf{S}_1}^\perp \mathbf{d}_{1m})] = \mathbf{0} \quad (19)$$

where $\mathbf{0}$ is the N -length zero vector. As in (13), we can now devise the following fixed step-size adaptation algorithm:

$$\mathbf{d}_{1m}(l+1) = \mathbf{d}_{1m}(l) - \mu (\mathcal{P}_{\mathbf{S}_1}^\perp (\mathbf{r}(l+1)\mathbf{r}(l+1)^\dagger - \sigma^2 \mathbf{I}) \times (\mathbf{S}_1 \mathbf{q}_{1m} + \mathcal{P}_{\mathbf{S}_1}^\perp \mathbf{d}_{1m}(l))) \quad (20)$$

to stochastically estimate \mathbf{d}_{1m} . We can replace μ by a step-size sequence μ_l , and with certain conditions on μ_l , obtain the Robbins–Monro formulation of the adaptation algorithm. In either case, at iteration $l+1$, we can reconstruct the estimate $\mathbf{c}_{1m}(l+1)$ from $\mathbf{d}_{1m}(l+1)$ by the relation

$$\mathbf{c}_{1m}(l+1) = \mathbf{S}_1 \mathbf{q}_{1m} + \mathcal{P}_{\mathbf{S}_1}^\perp \mathbf{d}_{1m}(l+1). \quad (21)$$

Note that the adaptation rule does not require the knowledge of the energy or phase of any of the signals received, and that the convergence here is to the actual decorrelator and not its scaled version. The terms $\mathbf{S}_1 \mathbf{q}_{1m}$ and $\mathcal{P}_{\mathbf{S}_1}^\perp$ remain the same for each iteration and, therefore, can be precomputed and stored.

The recursion in (20) followed by (21) can be seen to be equivalent to

$$\mathbf{d}_{1m}(l+1) = \mathbf{d}_{1m}(l) - \mu (\mathbf{r}(l+1)\mathbf{r}(l+1)^\dagger - \sigma^2 \mathbf{I}) \times (\mathbf{S}_1 \mathbf{q}_{1m} + \mathcal{P}_{\mathbf{S}_1}^\perp \mathbf{d}_{1m}(l)) \quad (22)$$

followed by (21). In either adaptation scheme [(20) and (21) or (22) and (21)], we observe that all the M signals of user 1 are used explicitly as a group for each decorrelator and we shall refer to this as the GS-BAND. For the purposes of proving the convergence of the GS-BAND, we will work with the pair (20) and (21). As with the SS-BAND, we need to ensure that the GS-BAND, if it converges, does so to \mathbf{c}_{1m}^o in (9). For this, we require that $\mathbf{c}_{1m}(l) = \mathcal{P}_{\mathbf{S}} \mathbf{c}_{1m}(l)$ a.s. It will be shown in the convergence analysis section that $\mathcal{P}_{\mathbf{S}_1}^\perp \mathbf{d}_{1m}(0) = \mathcal{P}_{\mathbf{S}} \mathcal{P}_{\mathbf{S}_1}^\perp \mathbf{d}_{1m}(0)$ is a sufficient condition to ensure this. This is a mild restriction on the choice of $\mathbf{d}_{1m}(0)$ since any linear combination of user 1's signals is an acceptable one, including $\mathbf{d}_{1m}(0) = \mathbf{0}$.

In terms of the decision statistics $z_{1m}(l+1) = \mathbf{c}_{1m}^\dagger(l+1)\mathbf{r}(l+1)$, for $m = 1, \dots, M$, the adaptation algorithm in (22) can be efficiently implemented as

$$\mathbf{d}_{1m}(l+1) = \mathbf{d}_{1m}(l) + \mu \sigma^2 \mathbf{c}_{1m}(l) - \mu \mathbf{r}(l+1) z_{1m}^*(l+1). \quad (23)$$

With the current decorrelator estimate $\mathbf{C}_1(l)$, the post-decorrelative GLRT-based detector for the GS-BAND is implemented in the same way as for the SS-BAND in (15).

It has been assumed in the derivation of both the SS-BAND and the GS-BAND that the user has perfect knowledge of the noise variance. Details of estimating the noise variance are omitted here and the reader is referred to Section VI [12] for the same. The only additional significant computation for the GS-BAND, compared to the SS-BAND, is the multiplication

$\mathcal{P}_{\mathbf{S}_1}^\perp \mathbf{d}_{1m}(l)$ in (21). More precisely, with the assumption that the signals are linearly independent, the recursion in (23) is $\mathcal{O}(N)$ and that in (21) is $\mathcal{O}(NM)$. The noise variance could be estimated either once, before the filter updates, or iteratively as in [12], the latter being an $\mathcal{O}(N)$ computation. The subspace-based adaptation scheme of [5], on the other hand, involves an $\mathcal{O}(NMK)$ subspace tracking step, followed by an $\mathcal{O}(NMK)$ matrix multiplication step and finally an $\mathcal{O}(NM^2)$ orthonormal basis decomposition step. Note that the noise variance estimation is a by-product of the subspace-tracking step.

The convergence analysis of the two BANDs in the next section provides analytical justification for preferring the GS-BAND over the SS-BAND and offers insight regarding step-size selection. The effects of using an imperfect estimate of the noise variance on the convergence can be studied in a manner very similar to that in [12] and, hence, will not be included in the following discussion.

V. CONVERGENCE ANALYSIS

In this section, we will investigate the convergence of the GS-BAND along the lines of the analysis in [6] for the Robbins–Monro decreasing step-size and in [14] and [12] for the fixed step-size. We remark that the convergence analysis for the SS-BAND is similar to that for the blind adaptive decorrelator for linear modulation in [12], deviating only due to the difference in the model for the received signal, and is omitted here. The interested reader will be able to find the elements that are different, in the following analysis of the GS-BAND. For notational simplicity, we shall omit the iteration index on the received vector \mathbf{r} in the following discussion. We first define the zero-mean vector $\beta(l)$ as

$$\beta(l) \triangleq \mathcal{P}_{\mathbf{S}_1}^\perp (\mathbf{r}\mathbf{r}^\dagger - \mathbf{S}\mathbf{W}\mathbf{S}^\dagger - \sigma^2 \mathbf{I}) \times (\mathbf{S}_1 \mathbf{q}_{1m} + \mathcal{P}_{\mathbf{S}_1}^\perp (\mathbf{d}_{1m} + \varepsilon_{1m}(l))) \quad (24)$$

where $\varepsilon_{1m}(l) = (\mathbf{d}_{1m}(l) - \mathbf{d}_{1m})$ is the error after the l th iteration. Since $\mathbf{S}\mathbf{W}\mathbf{S}^\dagger (\mathbf{S}_1 \mathbf{q}_{1m} + \mathcal{P}_{\mathbf{S}_1}^\perp \mathbf{d}_{1m}) = (E_{1m}/M) \mathbf{S}_{1m}$, we can write the following recursion for the error:

$$\varepsilon_{1m}(l+1) = \varepsilon_{1m}(l) - \mu \left(\mathcal{P}_{\mathbf{S}_1}^\perp \bar{\mathbf{S}}_1 \bar{\mathbf{W}}_1 \bar{\mathbf{S}}_1^\dagger \mathcal{P}_{\mathbf{S}_1}^\perp \varepsilon_{1m}(l) + \beta(l) \right). \quad (25)$$

Recalling that the GS-BAND is computationally more intensive than the SS-BAND, the following result provides justification in terms of the rate of convergence for preferring the former.

Theorem 1: In the fixed step-size case, the optimum rate of convergence in the mean of the GS-BAND is superior to that of the SS-BAND.

The proof of this result can be found in the Appendix.

We now address the mean square convergence for the GS-BAND. Considering the norm squared of both sides of (25), taking the conditional expectation, conditioned on $\varepsilon_{1m}(l) = \varepsilon$, and noting from (24) that $E[\beta(l) | \varepsilon_{1m}(l) = \varepsilon] = \mathbf{0}$, we obtain

$$\begin{aligned} E[\|\varepsilon_{1m}(l+1)\|^2 | \varepsilon_{1m}(l) = \varepsilon] &= \|\varepsilon\|^2 - 2\mu \varepsilon^\dagger \mathcal{P}_{\mathbf{S}_1}^\perp \bar{\mathbf{S}}_1 \bar{\mathbf{W}}_1 \bar{\mathbf{S}}_1^\dagger \mathcal{P}_{\mathbf{S}_1}^\perp \varepsilon \\ &\quad + \mu^2 \varepsilon^\dagger \left(\mathcal{P}_{\mathbf{S}_1}^\perp \bar{\mathbf{S}}_1 \bar{\mathbf{W}}_1 \bar{\mathbf{S}}_1^\dagger \mathcal{P}_{\mathbf{S}_1}^\perp \right)^2 \varepsilon \\ &\quad + \mu^2 E[\|\beta(l)\|^2 | \varepsilon_{1m}(l) = \varepsilon]. \end{aligned} \quad (26)$$

Let us denote the space spanned by the columns of a matrix Ψ by $\langle \Psi \rangle$. In order to develop bounds on the various terms in (26), we note, first, that $\mathbf{d}_{1m}(l)$ [and, therefore, $\varepsilon_{1m}(l)$] will a.s. have a nonzero projection on $\langle \mathcal{P}_{\mathbf{S}_1}^\perp \bar{\mathbf{S}}_1 \rangle$ [16], [17]. This fact will be useful in Lemma 1.

By induction on (25) (as in [16, eq. (16)])

$$\varepsilon_{1m}(l) = \mathcal{P}_{\mathcal{P}_{\mathbf{S}_1}^\perp} \bar{\mathbf{S}}_1 (\varepsilon_{1m}(l)) + \varepsilon_{1m}(0) - \mathcal{P}_{\mathcal{P}_{\mathbf{S}_1}^\perp} \bar{\mathbf{S}}_1 (\varepsilon_{1m}(0)) \text{ a.s.} \quad (27)$$

Clearly, the projection operation $\mathcal{P}_{\mathcal{S}} \mathcal{P}_{\mathbf{S}_1}^\perp$ leaves both the first and the last terms on the right-hand side of (27) unchanged, and so, if $\mathbf{d}_{1m}(0)$ is such that $\mathcal{P}_{\mathcal{S}} \mathcal{P}_{\mathbf{S}_1}^\perp \varepsilon_{1m}(0) = \mathcal{P}_{\mathcal{P}_{\mathbf{S}_1}^\perp} \bar{\mathbf{S}}_1 (\varepsilon_{1m}(0))$, the following will be true: $\mathcal{P}_{\mathbf{S}_1}^\perp \varepsilon_{1m}(l) = \mathcal{P}_{\mathcal{S}} \mathcal{P}_{\mathbf{S}_1}^\perp \varepsilon_{1m}(l)$ a.s. This implies that $\mathbf{c}_{1m}(l) = \mathbf{S}_1 \mathbf{q}_{1m} + \mathbf{P}_{\mathbf{S}_1}^\perp (\mathbf{d}_{1m} + \varepsilon_{1m}(l))$ a.s. stays in the signal subspace, and will converge (as we will show shortly) to the desired decorrelator \mathbf{c}_{1m}^o . The results of the following two lemmas will help us suitably bound $\|\varepsilon_{1m}(l+1)\|^2$ in (26).

Lemma 1: With $\varepsilon_{1m}(l)$ in (26) set equal to ε and $\mathbf{G} = \mathcal{P}_{\mathbf{S}_1}^\perp \bar{\mathbf{S}}_1 \bar{\mathbf{W}} \bar{\mathbf{S}}_1^\dagger \mathcal{P}_{\mathbf{S}_1}^\perp$, there exists $0 < k_0 \leq k_1 < \infty$, such that

$$k_0 \|\varepsilon\|^2 \leq \varepsilon^\dagger \mathbf{G} \varepsilon \leq k_1 \|\varepsilon\|^2$$

and

$$k_0^2 \|\varepsilon\|^2 \leq \varepsilon^\dagger \mathbf{G}^2 \varepsilon \leq k_1^2 \|\varepsilon\|^2. \quad (28)$$

The proof of this lemma can be easily shown using the facts that the error vector ε has a nonzero projection on the subspace $\langle \mathcal{P}_{\mathbf{S}_1}^\perp \bar{\mathbf{S}}_1 \rangle$ a.s. and that the quadratic forms $\varepsilon^\dagger \mathbf{G} \varepsilon$ and $\varepsilon^\dagger \mathbf{G}^2 \varepsilon$ can be bounded using the Rayleigh–Ritz ratio [15].

Lemma 2: There exist constants c_0, c_1 such that

$$0 \leq E[\|\beta(l)\|^2 | \varepsilon_{1m}(l) = \varepsilon] \leq c_0 + c_1 \|\varepsilon\|^2. \quad (29)$$

The proof of this lemma can be found in the Appendix.

The results of the two lemmas above can now be used to prove the mean-square convergence of the GS-BAND along the lines of [12]. From Lemmas 1 and 2, we develop the following bounds on $E[\|\varepsilon_{1m}(l+1)\|^2 | \varepsilon_{1m}(l) = \varepsilon]$:

$$\begin{aligned} & (1 - 2\mu k_1 + k_0^2 \mu^2) \|\varepsilon\|^2 \\ & \leq E[\|\varepsilon_{1m}(l+1)\|^2 | \varepsilon_{1m}(l) = \varepsilon] \\ & \leq \left(1 - 2\mu k_0 + (k_1^2 + c_1) \mu^2\right) \|\varepsilon\|^2 + c_0 \mu^2. \end{aligned} \quad (30)$$

Taking the expectation of the inequality in (30) with respect to ε , and letting $v_l = E[\|\varepsilon_{1m}(l)\|^2]$, we obtain the upper and lower bounds for v_{l+1} , the MSE of the estimated decorrelator $\mathbf{c}_{1m}(l+1)$, in terms of v_l as

$$\begin{aligned} & (1 - 2\mu k_1 + k_0^2 \mu^2) v_l \\ & \leq v_{l+1} \leq (1 - 2\mu k_0 + (k_1^2 + c_1) \mu^2) v_l + c_0 \mu^2. \end{aligned} \quad (31)$$

By defining $\alpha_0 = 1 - 2\mu k_0 + (k_1^2 + c_1) \mu^2$ and $\alpha_1 = 1 - 2\mu k_1 + k_0^2 \mu^2$, we can rewrite (31) as

$$\alpha_1 v_l \leq v_{l+1} \leq \alpha_0 v_l + c_0 \mu^2. \quad (32)$$

We observe from (32) that the nonnegative sequence v_l is sandwiched between the two sequences generated according to $v'_{l+1} = \alpha_0 v'_l + c_0 \mu^2$ and $v''_{l+1} = \alpha_1 v''_l$. These two sequences converge to finite numbers if and only if μ is chosen such that

$|\alpha_0| < 1$ and $|\alpha_1| < 1$. Note that both α_0 and α_1 are equal to 1 at $\mu = 0$. We also note that both α_0 and α_1 are locally decreasing as μ increases from $\mu = 0$, since $(d\alpha_0/d\mu)|_{\mu=0} = -2k_0 < 0$ and $(d\alpha_1/d\mu)|_{\mu=0} = -2k_1 < 0$. This means we can always choose μ small enough so that $|\alpha_0| < 1$ and $|\alpha_1| < 1$, in which case the sequences v'_l and v''_l converge and the limiting MSE, i.e., $\lim_{l \rightarrow \infty} v_l$, has finite lower and upper bounds. Some authors call this a *quasi-convergence* result, since the MSE does not necessarily go to zero. With the normalized squared error (NSE) for user 1 at iteration l for the GS-BAND given by

$$\text{NSE}(l) = \frac{1}{M} \sum_{m=1}^M \frac{\|\mathbf{c}_{1m}^o - \mathbf{c}_{1m}(l)\|^2}{\|\mathbf{c}_{1m}^o\|^2} \quad (33)$$

the convergence result (and the following observations for the MSE) can be easily seen to extend to the expected value of the NSE. From the sandwich theorem for convergence of sequences, we have $0 \leq \lim_{l \rightarrow \infty} v_l \leq (c_0 \mu^2 / (1 - \alpha_0))$. We can evaluate the value of the upper bound in the extreme case, when $\mu \rightarrow 0$, as $\lim_{\mu \rightarrow 0} (c_0 \mu^2 / (1 - \alpha_0)) = \lim_{\mu \rightarrow 0} (c_0 \mu / (2k_0 - (k_1^2 + c_1) \mu)) = 0$. If the step-size μ is chosen arbitrarily small, the MSE is seen to converge to zero as the number of iterations grows to infinity. This comes at a price, though, since $\mu \rightarrow 0$ implies that the rate of convergence goes to zero. Thus, we observe the familiar tradeoff between the limiting value of MSE that we can achieve and the rate of convergence.

A time-dependent step-size μ_l which takes large values at the beginning and smaller values at the end would seem to be a good choice to ensure faster convergence as well as lower limiting MSE. In fact, if the varying step-size sequence satisfies the following two simple conditions, known as the Robbins–Monro conditions [13]

$$\sum_l \mu_l = \infty \quad \text{and} \quad \sum_l \mu_l^2 < \infty \quad (34)$$

then, it can be shown from [6] and [12] that the GS-BAND converges to the true decorrelating detector in the MSE sense, i.e., $\lim_{l \rightarrow \infty} E[\|\mathbf{c}_{1m}(l) - \mathbf{c}_{1m}^o\|^2] = 0$. The corresponding result for the Robbins–Monro version of the SS-BAND can also be shown to be true. An example of a step-size sequence which satisfies the Robbins–Monro conditions is $\mu_l = (a/b + l)$, where a and b are constants.

The decreasing step-size algorithm is better suited for the initial “learning” phase, whereas the fixed step-size algorithm performs better in the “tracking” phase, when the algorithm has sufficiently converged and needs to only track changes in the system. This is because the decreasing step-size sequence is initially relatively larger than a suitable fixed step-size that could be chosen, but eventually becomes too small to effectively track the system.

VI. NUMERICAL EXAMPLES

In this section, we illustrate the comparative performance of the SS-BAND and the GS-BAND in terms of the NSE and the symbol-error rate (SER) for fixed and decreasing step-size cases. We also make a comparison of these algorithms with the “canonical” single-signal decomposition-based algorithm adapted from the MOE approach of [9] to noncoherent decor-

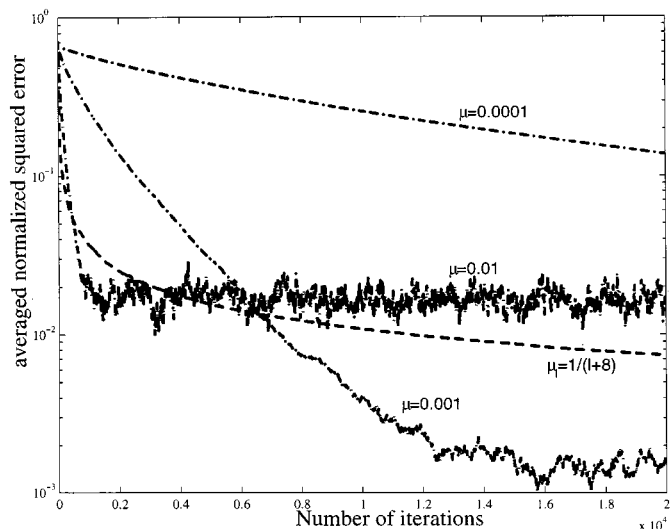


Fig. 1. Averaged NSE of the GS-BAND for different step-sizes.

relative detection for NMM. Also included in our comparison is the more complex subspace based approach of [5].

In all our examples, the SER is computed as an average (over several trajectories of the blind algorithms) of the Monte-Carlo SER estimate for the adaptive decorrelator at every 2000th iteration. User 1 is the user of interest. The NSE for user 1 at iteration l for the GS-BAND is given in (33) and for the SS-BAND [accounting for the scale factors (M/E_{1m})] is given as

$$\text{NSE}(l) = \frac{1}{M} \sum_{m=1}^M \frac{\|(M/E_{1m})\mathbf{c}_{1m}^o - \mathbf{c}_{1m}(l)\|^2}{\|(M/E_{1m})\mathbf{c}_{1m}^o\|^2}. \quad (35)$$

The NSE averaged over several runs provides an estimate of the MSE. The unit-norm signature signals are randomly generated from a uniform distribution for the different users, but once chosen remain the same throughout the adaptation.

Example 1: We consider a three-user, quaternary signaling example with “processing gain” equal to 12—a bandwidth-efficient system under the linearly independent signaling criterion. The SNR is chosen to be 30 dB. In Fig. 1, we compare the averaged NSE for the GS-BAND for different fixed step-sizes, namely, $\mu = 0.0001, 0.001, 0.01$. The tradeoff between the limiting MSE (as indicated by the averaged NSE) and the rate of convergence is evident from the figure. Fast convergence requires a large step-size but comes at the price of a large MSE. If it is required that the blind detector be very close to the deterministic detector (low MSE), then a small step-size must be used, but this has the effect of slowing down the rate of convergence. The requirements of a particular application will determine the number of iterations that is acceptable and this in turn determines the step-size and hence limiting MSE. At the time of system design, it must be ensured that a blind detector with this limiting MSE can deliver acceptable performance. Note that this performance will necessarily be sub-par relative to that of the deterministic solution. This is the price that must be paid for blind detection with limited convergence time. For instance, a step-size slightly larger than 0.01 will allow convergence in only around 100 iterations, but at the cost of high limiting MSE. The

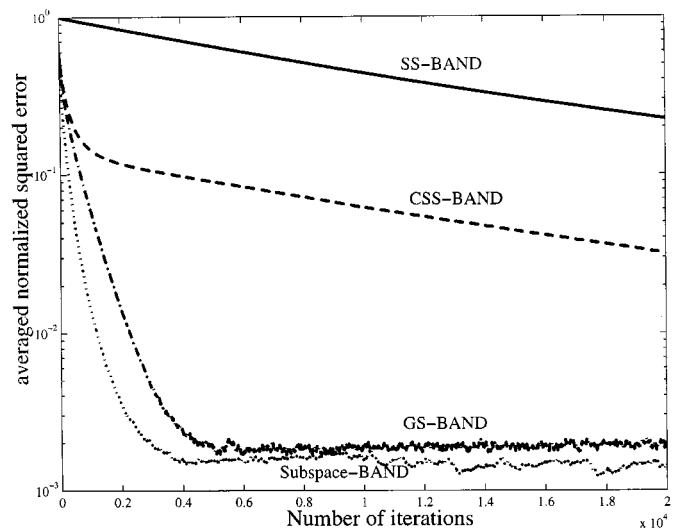


Fig. 2. Averaged NSE of the SS-BAND, GS-BAND, CSS-BAND, and the subspace-based BAND.

decreasing step-size GS-BAND ($\mu_l = 1/l + 8$) shows essentially the following trend: convergence, initially, is fast (when the step-size values are relatively large) and becomes slower with an increasing number of iterations.

Example 2: We next consider a ten-user, quaternary signaling example, with processing gain 64. The SNR is slightly less than 20 dB. Fig. 2 compares the performance of the fixed step-size SS-BAND and the GS-BAND with two other BANDs. The plot labeled “CSS-BAND” corresponds to the stochastic approximation algorithm that uses the “canonical” single-signal decomposition (as in [9]) of the decorrelator. More precisely, the decorrelator is decomposed into the corresponding signal plus a perpendicular part, where only the perpendicular part is adaptively estimated. This is an intermediate decomposition between the SS-BAND and the GS-BAND and the NSE performance confirms this. For the same fixed step-size $\mu = 0.001$, note the significant improvement in convergence as a result of using more of the user’s signal space information. The “subspace-BAND” corresponds to the subspace-based adaptation of [5], where we have chosen the version that uses a sliding window of data in its subspace estimation and, therefore, like the other three fixed step-size adaptations, has tracking ability. All the algorithms, in this comparison, use perfect knowledge of the noise power. Moreover, for the subspace-BAND, an initial estimate of the received signal correlation matrix is computed over 50 symbol intervals. The subspace-BAND shows a *sliding window length versus limiting MSE* tradeoff much like the *fixed step-size versus limiting MSE* tradeoff shown by the other three.

Example 3: Using another ten-user example, we make some observations regarding a comparison of the subspace-BAND and the GS-BAND, and the robustness of the latter to inaccuracies in the knowledge of noise variance.

Consider Figs. 3 and 4. In Fig. 3, we plot the performance of the GS-BAND (with and without errors in noise variance) and that of the subspace-BAND with sliding window length of ≈ 1000 symbol intervals for an SNR of ≈ 20 dB. We notice that although the rate of convergence of the subspace-BAND is better, it shows a higher asymptotic NSE floor than the

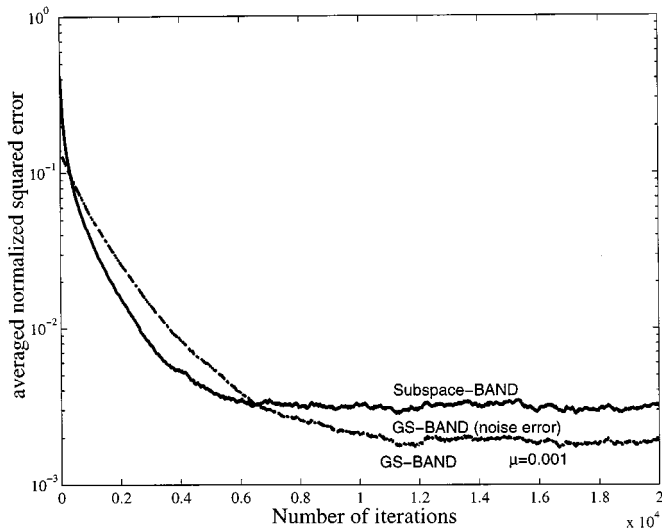


Fig. 3. Averaged NSE for GS-BAND (with and without noise estimate error $\mu = 0.001$) and subspace-BAND (sliding window length ≈ 1000).

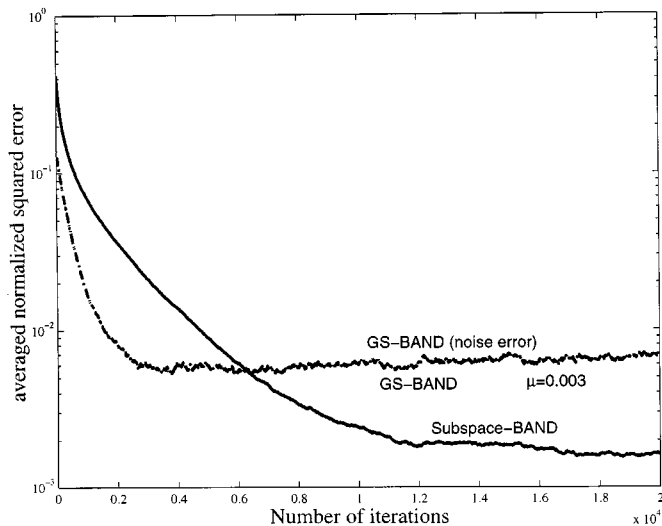


Fig. 4. Averaged NSE for GS-BAND (with and without noise estimate error $\mu = 0.003$) and subspace-BAND (sliding window length ≈ 2000).

GS-BAND. Further, there is almost no difference between the GS-BAND using perfect noise power knowledge and that using noise power estimates that are in error by $\pm 20\%$. In Fig. 4, we increase the sliding window length of the subspace-BAND to ≈ 2000 and increase the step-size for the GS-BAND to 0.003. Now, we see faster convergence for the GS-BAND but with higher asymptotic NSE floor as compared to the subspace-BAND. Again, there is hardly any effect due to error in the noise power estimate in the GS-BAND performance. We have observed this phenomenon over several numerical examples at reasonable signal-to-noise ratios.

We conclude that, because of the dependence on a right choice of adaptation parameters, it is difficult to unequivocally compare the performance of subspace-BAND and GS-BAND.

Example 4: We consider once again the signaling setup in Example 1. The SNR is chosen as 20 dB. The superior convergence of the GS-BAND over the SS-BAND is once again evident from the averaged NSE plots in Fig. 5 (for

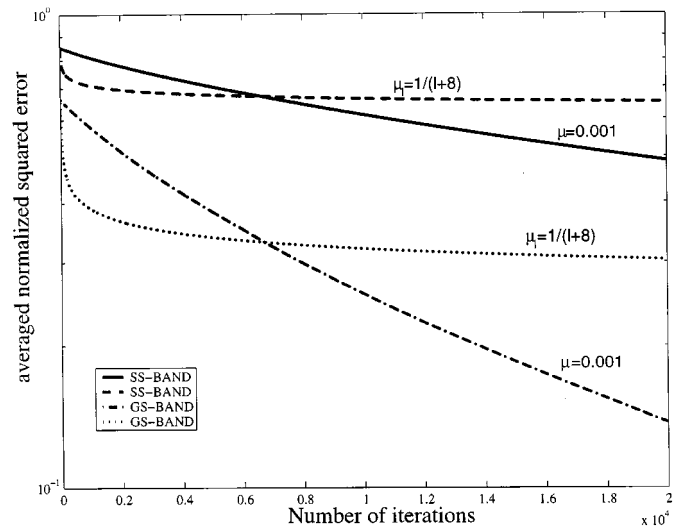


Fig. 5. Averaged NSE of the SS-BAND and GS-BAND.

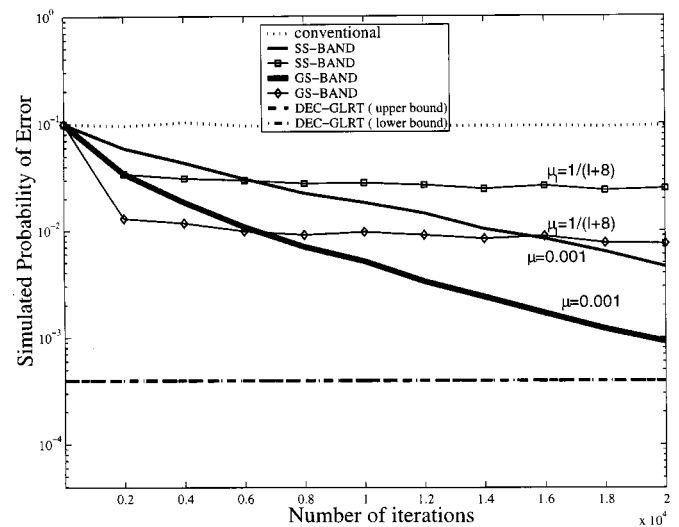


Fig. 6. Simulated SER curves corresponding to Fig. 5 for the conventional detector, SS-BAND, GS-BAND, and DEC-GLRT (DEC-GLRT) detector upper and lower bounds (the last two are indistinguishable).

both fixed and decreasing step-size sequences $\mu = 0.001$ and $\mu_l = 1/l + 8$, respectively). Fig. 6 plots the corresponding simulated SER curves for the conventional detector, the SS-BAND, the GS-BAND, and the deterministic decorrelator (DEC-GLRT). The post-decorrelative detector used in the latter three cases is the GLRT-based detector. Note that the BER upper and lower bounds for the deterministic decorrelator (see [1]) in Fig. 6, which together form the bottom-most line in the plot, are indistinguishable. Clearly the convergence, even for the GS-BAND, is rather poor.

We next try to capture the rapid initial convergence of the decreasing step-size algorithm and the sustained convergence property of the fixed step-size algorithm by using a suitably large step-size in the beginning and switching to smaller step-sizes (every several iterations) as the adaptation proceeds. Adopting such a strategy for this example, where the starting step-size is chosen to be 0.1 and every several iterations decreased by a small amount, differently for the SS-BAND and the GS-BAND,

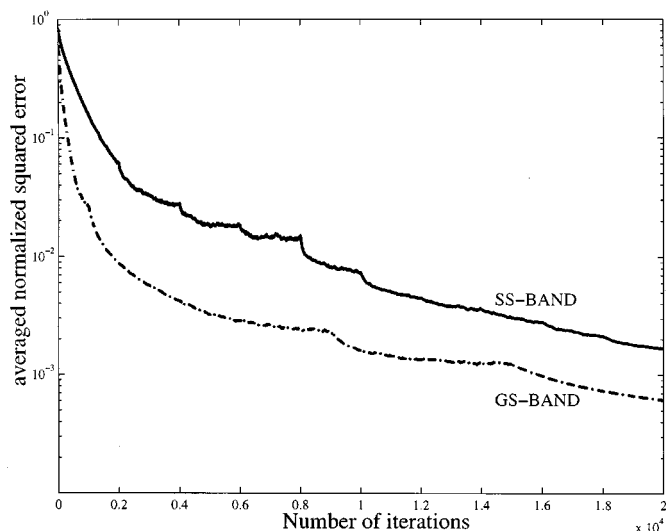


Fig. 7. Averaged NSE of the SS-BAND and GS-BAND with improved step-sizes.

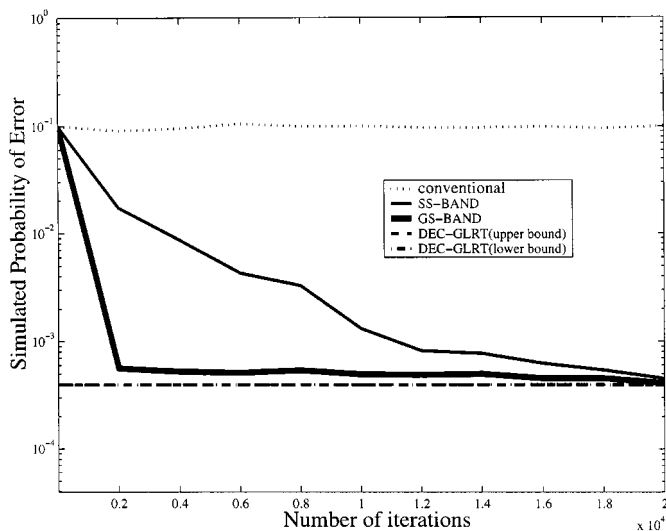


Fig. 8. Simulated SER curves for the conventional detector, SS-BAND, and GS-BAND (with improved step-size sequence) and DEC-GLRT (DEC-GLRT) detector upper and lower bounds (the last two are indistinguishable).

leads to a significant improvement of performance as evidenced by Figs. 7 and 8 (compare with Figs. 5 and 6). The step-size sequences chosen for the SS-BAND and GS-BAND in this example have been arrived at by trial and error. A systematic method to arrive at such a sequence based on known system parameters can, perhaps, be studied along the lines of [18].

Example 5: The SS-BAND and the GS-BAND can be applied in a frequency-nonselctive Rayleigh fading channel such as a channel with several point scatterers [19, Ch. 2] moving slowly enough to cause relatively slow fading (complex amplitude constant over one signal interval), where the amplitude of the received signal is a Rayleigh random variable and the random phase is modeled as uniformly distributed on $[0, 2\pi]$. In general, the blind algorithms are applicable to slow or moderately fast fading channels (where the zero-mean complex amplitudes remain constant over one symbol interval and can change, in the worst case, independently, from symbol to symbol) and

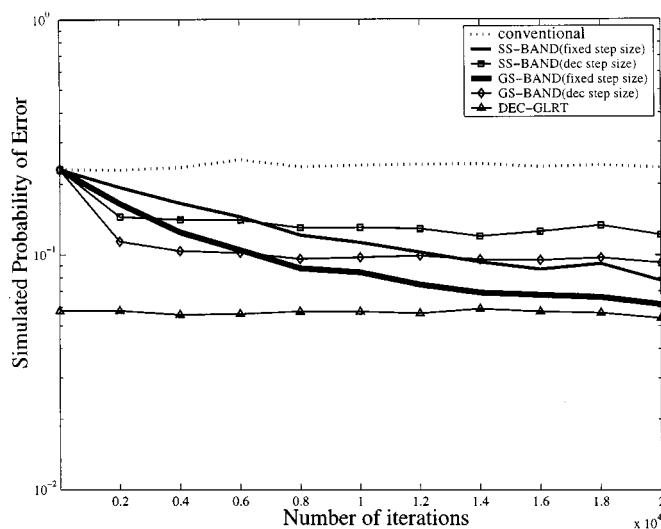


Fig. 9. Simulated SER of the conventional detector, the SS-BAND, GS-BAND, and DEC-GLRT detector for a Rayleigh fading channel. The step-sizes are the same as in Fig. 5.

are, therefore, particularly relevant when phase estimation and tracking over such channels is not dependable.

In this example, we compare the SER performance of the SS-BAND and the GS-BAND with that for the deterministic decorrelator for a Rayleigh fading channel using post-decorrelative GLRT detection. The mean-squared values of the Rayleigh fading amplitudes are the same as the fixed signaling energies in the Gaussian channel case (i.e., the energies of all users are equal to unity). All other system and adaptation parameters are the same as for Fig. 6. Fig. 9 plots the simulated SER curves for the conventional, the SS-BAND, the GS-BAND, and the DEC-GLRT. Comparison of this plot to Fig. 6 shows a deterioration (as is to be expected) of the performance for the fading channel over that for the Gaussian channel. It is clear however that the GS-BAND retains its superiority over the SS-BAND even with fading. Again, it is possible to improve the performance of both the algorithms by proper step-size selection.

VII. CONCLUSION

Two blind adaptive noncoherent decorrelative detectors (BANDs) for NMM are obtained that adapt without “training” and require only the signals of the user of interest and the channel noise variance. They are based on the stochastic approximation technique and are suitable for distributed implementation. The first one, called SS-BAND, obtains each of any user’s decorrelators by employing only the corresponding signal to guide the adaptation algorithm, while the second, GS-BAND, employs the user’s entire signal set in the adaptation design. It has been shown that both the fixed step-size and the decreasing step-size (satisfying the Robbins–Monro conditions) versions of the NMM-BANDs converge, in the MSE sense, to the deterministic decorrelator. The former, capable of tracking system changes, converges with a nonzero-limiting MSE, and the latter with zero-limiting MSE. The superiority of the GS-BAND over the SS-BAND has been shown analytically and numerical examples confirm this in terms of the MSE and SER performance. With suitable step-size selection, the

GS-BAND algorithm in this paper can be competitive with the computationally more intensive subspace-based methods. Although designed for the AWGN channel, they also find applicability in slow or moderately fast frequency-nonselective fading channels.

APPENDIX

Proof of Theorem 1: First, we define the term corresponding to $\beta(l)$ in (24) for the SS-BAND as $\tilde{\beta}(l) \triangleq (\mathbf{r}\mathbf{r}^\dagger - \mathbf{S}\mathbf{W}\mathbf{S}^\dagger - \sigma^2\mathbf{I})\tilde{\mathbf{c}}_{1m}(l)$, where $\tilde{\mathbf{c}}_{1m}(l)$ is the estimate at iteration l given by the SS-BAND rule. The error equation for the SS-BAND can be written in terms of the error at iteration l , $\tilde{\mathbf{e}}_{1m}(l)$, as

$$\tilde{\mathbf{e}}_{1m}(l+1) = \tilde{\mathbf{e}}_{1m}(l) - \mu \left(\mathbf{S}\mathbf{W}\mathbf{S}^\dagger \tilde{\mathbf{e}}_{1m}(l) + \tilde{\beta}(l) \right). \quad (36)$$

Taking the expectation of both sides of (25) and (36), conditioned, first on $\varepsilon_{1m}(l) = \varepsilon$ and $\tilde{\mathbf{e}}_{1m}(l) = \tilde{\mathbf{e}}$, and then over ε and $\tilde{\mathbf{e}}$, respectively, we obtain the following:

$$E[\varepsilon_{1m}(l+1)] = (\mathbf{I} - \mu \mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{S}\mathbf{W}\mathbf{S}^\dagger \mathcal{P}_{\mathbf{S}_1}^\dagger) E[\varepsilon_{1m}(l)] \quad (37)$$

$$E[\tilde{\mathbf{e}}_{1m}(l+1)] = (\mathbf{I} - \mu \mathbf{S}\mathbf{W}\mathbf{S}^\dagger) E[\tilde{\mathbf{e}}_{1m}(l)]. \quad (38)$$

As has been discussed in the convergence analysis section, the error terms in (37) and (38) above can be shown a.s. to lie in the subspaces $\langle \mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{S} \rangle$ and $\langle \mathbf{S} \rangle$, respectively. As is well known in the least mean square adaptive literature [19], the rate of convergence in the mean of the two adaptations GS-BAND and SS-BAND is dictated by the disparity of the nonzero eigenvalues of $\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{S}\mathbf{W}\mathbf{S}^\dagger \mathcal{P}_{\mathbf{S}_1}^\dagger$ and $\mathbf{S}\mathbf{W}\mathbf{S}^\dagger$, respectively. We will show, therefore, that the eigenspread (ratio of the maximum and the minimum nonzero eigenvalues) of $\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{S}\mathbf{W}\mathbf{S}^\dagger \mathcal{P}_{\mathbf{S}_1}^\dagger$ is, indeed, smaller than that for $\mathbf{S}\mathbf{W}\mathbf{S}^\dagger$, i.e.,

$$\frac{\lambda_1(\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{S}\mathbf{W}\mathbf{S}^\dagger \mathcal{P}_{\mathbf{S}_1}^\dagger)}{\lambda_{MK-M}(\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{S}\mathbf{W}\mathbf{S}^\dagger \mathcal{P}_{\mathbf{S}_1}^\dagger)} \leq \frac{\lambda_1(\mathbf{S}\mathbf{W}\mathbf{S}^\dagger)}{\lambda_{MK}(\mathbf{S}\mathbf{W}\mathbf{S}^\dagger)} \quad (39)$$

where the nonzero eigenvalues of the two matrices are in decreasing order: $\lambda_1(\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{S}\mathbf{W}\mathbf{S}^\dagger \mathcal{P}_{\mathbf{S}_1}^\dagger) \geq \lambda_2(\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{S}\mathbf{W}\mathbf{S}^\dagger \mathcal{P}_{\mathbf{S}_1}^\dagger) \geq \dots \geq \lambda_{MK-M}(\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{S}\mathbf{W}\mathbf{S}^\dagger \mathcal{P}_{\mathbf{S}_1}^\dagger) > 0$ and $\lambda_1(\mathbf{S}\mathbf{W}\mathbf{S}^\dagger) \geq \lambda_2(\mathbf{S}\mathbf{W}\mathbf{S}^\dagger) \geq \dots \geq \lambda_{MK}(\mathbf{S}\mathbf{W}\mathbf{S}^\dagger) > 0$. We know that $\lambda_i(\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{S}\mathbf{W}\mathbf{S}^\dagger \mathcal{P}_{\mathbf{S}_1}^\dagger) = \lambda_i(\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{S}\mathbf{W}\mathbf{S}^\dagger)$. Let us also order the singular values as $\sigma_1(\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{S}\mathbf{W}\mathbf{S}^\dagger) \geq \sigma_2(\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{S}\mathbf{W}\mathbf{S}^\dagger) \geq \dots \geq \sigma_{MK-M}(\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{S}\mathbf{W}\mathbf{S}^\dagger) > 0$ and $\sigma_1(\mathbf{S}\mathbf{W}\mathbf{S}^\dagger) \geq \sigma_2(\mathbf{S}\mathbf{W}\mathbf{S}^\dagger) \geq \dots \geq \sigma_{MK}(\mathbf{S}\mathbf{W}\mathbf{S}^\dagger) > 0$. Now, both $\mathcal{P}_{\mathbf{S}_1}^\dagger$ and $\mathbf{S}\mathbf{W}\mathbf{S}^\dagger$, being nonnegative matrices, have unique square roots (denoted by $\sqrt{\mathcal{P}_{\mathbf{S}_1}^\dagger}$ and $\sqrt{\mathbf{S}\mathbf{W}\mathbf{S}^\dagger}$, respectively) that are also nonnegative. Also, $\lambda_i(\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{S}\mathbf{W}\mathbf{S}^\dagger) = \lambda_i(\sqrt{\mathcal{P}_{\mathbf{S}_1}^\dagger} \sqrt{\mathbf{S}\mathbf{W}\mathbf{S}^\dagger} \sqrt{\mathcal{P}_{\mathbf{S}_1}^\dagger}) = \lambda_i^2(\sqrt{\mathcal{P}_{\mathbf{S}_1}^\dagger} \sqrt{\mathbf{S}\mathbf{W}\mathbf{S}^\dagger})$. Therefore

$$\sigma_i \left(\sqrt{\mathcal{P}_{\mathbf{S}_1}^\dagger} \sqrt{\mathbf{S}\mathbf{W}\mathbf{S}^\dagger} \right) = \sqrt{\lambda_i(\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{S}\mathbf{W}\mathbf{S}^\dagger)}. \quad (40)$$

Using [20, Theorem 3.8], we obtain that $\lambda_{MK}(\sqrt{\mathbf{S}\mathbf{W}\mathbf{S}^\dagger}) \lambda_{N-M}(\sqrt{\mathcal{P}_{\mathbf{S}_1}^\dagger}) \leq \sigma_{MK-M}(\sqrt{\mathcal{P}_{\mathbf{S}_1}^\dagger} \sqrt{\mathbf{S}\mathbf{W}\mathbf{S}^\dagger})$, and from [21], the inequality $0 < \sigma_1(\sqrt{\mathcal{P}_{\mathbf{S}_1}^\dagger} \sqrt{\mathbf{S}\mathbf{W}\mathbf{S}^\dagger}) \leq \lambda_1(\sqrt{\mathbf{S}\mathbf{W}\mathbf{S}^\dagger}) \lambda_1(\sqrt{\mathcal{P}_{\mathbf{S}_1}^\dagger})$. Therefore, using (40) and the eigenvalue properties

of $\mathcal{P}_{\mathbf{S}_1}^\dagger$, it is true that $\lambda_{MK}(\mathbf{S}\mathbf{W}\mathbf{S}^\dagger) \leq \lambda_{MK-M}(\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{S}\mathbf{W}\mathbf{S}^\dagger)$ and $\lambda_1(\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{S}\mathbf{W}\mathbf{S}^\dagger) \leq \lambda_1(\mathbf{S}\mathbf{W}\mathbf{S}^\dagger)$. It follows, then, that (39) is true.

Proof of Lemma 2: Using (24) and the fact that $E[\beta(l) | \varepsilon_{1m}(l) = \varepsilon] = \mathbf{0}$, we obtain

$$E[\|\beta(l)\|^2 | \varepsilon_{1m}(l) = \varepsilon] = (\mathbf{S}_1 \mathbf{q}_{1m} + \mathcal{P}_{\mathbf{S}_1}^\dagger (\mathbf{d}_{1m} + \varepsilon))^\dagger \times \mathbf{B} (\mathbf{S}_1 \mathbf{q}_{1m} + \mathcal{P}_{\mathbf{S}_1}^\dagger (\mathbf{d}_{1m} + \varepsilon)) \quad (41)$$

where $\mathbf{B} \triangleq E[\mathbf{r}\mathbf{r}^\dagger \mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{r}\mathbf{r}^\dagger] - (\mathbf{S}\mathbf{W}\mathbf{S}^\dagger + \sigma^2\mathbf{I})^\dagger \mathcal{P}_{\mathbf{S}_1}^\dagger (\mathbf{S}\mathbf{W}\mathbf{S}^\dagger + \sigma^2\mathbf{I})$. Using the model for \mathbf{r} in (3), we obtain the expansion $E[\mathbf{r}\mathbf{r}^\dagger \mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{r}\mathbf{r}^\dagger] = E[(\mathbf{S}\mathbf{A}\mathbf{b} + \mathbf{n})(\mathbf{S}\mathbf{A}\mathbf{b} + \mathbf{n})^\dagger \mathcal{P}_{\mathbf{S}_1}^\dagger (\mathbf{S}\mathbf{A}\mathbf{b} + \mathbf{n})(\mathbf{S}\mathbf{A}\mathbf{b} + \mathbf{n})^\dagger]$. This expansion is made up of the following nonzero terms (the proofs are left to the reader): $E[(\mathbf{S}\mathbf{A}\mathbf{b}\mathbf{b}^\dagger \mathbf{A}^\dagger \mathbf{S}^\dagger) \mathcal{P}_{\mathbf{S}_1}^\dagger (\mathbf{n}\mathbf{n}^\dagger)] = \sigma^2 \mathbf{S}\mathbf{W}\mathbf{S}^\dagger \mathcal{P}_{\mathbf{S}_1}^\dagger$, $E[(\mathbf{S}\mathbf{A}\mathbf{b}\mathbf{b}^\dagger \mathbf{A}^\dagger \mathbf{S}^\dagger) \mathcal{P}_{\mathbf{S}_1}^\dagger (\mathbf{n}\mathbf{b}^\dagger \mathbf{A}^\dagger \mathbf{S}^\dagger)] = \sigma^2 \text{Trace}(\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{S}\mathbf{W}\mathbf{S}^\dagger)$, $E[(\mathbf{n}\mathbf{b}^\dagger \mathbf{A}^\dagger \mathbf{S}^\dagger) \mathcal{P}_{\mathbf{S}_1}^\dagger (\mathbf{S}\mathbf{A}\mathbf{b}\mathbf{n}^\dagger)] = \sigma^2 \text{Trace}(\mathbf{S}\mathbf{W}\mathbf{S}^\dagger \mathcal{P}_{\mathbf{S}_1}^\dagger) \mathbf{I}$, $E[(\mathbf{n}\mathbf{n}^\dagger) \mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{n}\mathbf{n}^\dagger] = \sigma^4 (\mathcal{P}_{\mathbf{S}_1}^\dagger + \text{Trace}(\mathcal{P}_{\mathbf{S}_1}^\dagger) \mathbf{I})$, and $E[(\mathbf{S}\mathbf{A}\mathbf{b}\mathbf{b}^\dagger \mathbf{A}^\dagger \mathbf{S}^\dagger) \mathcal{P}_{\mathbf{S}_1}^\dagger (\mathbf{S}\mathbf{A}\mathbf{b}\mathbf{b}^\dagger \mathbf{A}^\dagger \mathbf{S}^\dagger)] = (\mathbf{S}\mathbf{W}\mathbf{S}^\dagger) \mathcal{P}_{\mathbf{S}_1}^\dagger (\mathbf{S}\mathbf{W}\mathbf{S}^\dagger) + (\mathbf{S}\mathbf{L}\mathbf{S}^\dagger)$, where \mathbf{L} is an $MK \times MK$ matrix with $M \times M$ nonzero blocks along the diagonal. Denoting by $\Psi(a : b, c : d)$, the intersection of rows a to b and columns c to d of any matrix Ψ , we let $\Theta_k \triangleq (\mathbf{S}(1 : N, (k-1)M+1 : kM))^\dagger \mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{S}(1 : N, (k-1)M+1 : kM)$. Further, we define $M \times M$ diagonal matrices \mathbf{W}_i such that $\mathbf{W} = \text{diag}(\mathbf{W}_1, \mathbf{W}_2, \dots, \mathbf{W}_K)$. The j th diagonal element of the k th diagonal block of \mathbf{L} is hence given by $(E_{kj}/M)(\sum_{i \neq k} \text{Trace}(\mathbf{W}_i \Theta_i)) + (E_{kj}^2 (M-1)/M^2) \Theta_k(j, j)$. The off-diagonal terms of the k th diagonal block of \mathbf{L} are given by the negative of the off-diagonal terms of $\mathbf{W}_k \Theta_k \mathbf{W}_k$. It can be shown that \mathbf{L} is positive semidefinite, and therefore, so is $\mathbf{S}\mathbf{L}\mathbf{S}^\dagger$. We can now write $\mathbf{B} = (\mathbf{S}\mathbf{L}\mathbf{S}^\dagger) + \sigma^2 \text{Trace}(\mathcal{P}_{\mathbf{S}_1}^\dagger) (\mathbf{S}\mathbf{W}\mathbf{S}^\dagger) + \sigma^2 \text{Trace}((\mathbf{S}\mathbf{W}\mathbf{S}^\dagger) \mathcal{P}_{\mathbf{S}_1}^\dagger) \mathbf{I} + \sigma^4 \text{Trace}(\mathcal{P}_{\mathbf{S}_1}^\dagger) \mathbf{I}$.

It follows from (41) that

$$E[\|\beta(l)\|^2 | \varepsilon_{1m}(l) = \varepsilon] = (\mathbf{S}_1 \mathbf{q}_{1m})^\dagger \mathbf{B} (\mathbf{S}_1 \mathbf{q}_{1m}) + (\mathbf{d}_{1m} + \varepsilon)^\dagger \mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{B} \mathcal{P}_{\mathbf{S}_1}^\dagger (\mathbf{d}_{1m} + \varepsilon) + 2\text{Re}(\varepsilon^\dagger \mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{B} \mathbf{S}_1 \mathbf{q}_{1m}) + 2\text{Re}(\mathbf{d}_{1m}^\dagger \mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{B} \mathbf{S}_1 \mathbf{q}_{1m}). \quad (42)$$

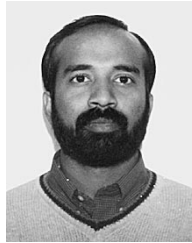
With nonzero σ^2 , the matrix \mathbf{B} is positive definite and using the Rayleigh-Ritz ratio [15], we can upper bound $(\mathbf{d}_{1m} + \varepsilon)^\dagger \mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{B} \mathcal{P}_{\mathbf{S}_1}^\dagger (\mathbf{d}_{1m} + \varepsilon)$ in (42) by $\lambda_{\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{B} \mathcal{P}_{\mathbf{S}_1}^\dagger}^{\max} \|(\mathbf{d}_{1m} + \varepsilon)\|^2$, where $\lambda_{\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{B} \mathcal{P}_{\mathbf{S}_1}^\dagger}^{\max} > 0$ is the maximum eigenvalue of $\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{B} \mathcal{P}_{\mathbf{S}_1}^\dagger$. Also, the term $2\text{Re}(\varepsilon^\dagger \mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{B} \mathbf{S}_1 \mathbf{q}_{1m})$ in (42) can be upper bounded by $\|\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{B} \mathbf{S}_1 \mathbf{q}_{1m}\|^2 + \|\varepsilon\|^2$, and $\|(\mathbf{d}_{1m} + \varepsilon)\|^2$ can be upper bounded by $2(\|\mathbf{d}_{1m}\|^2 + \|\varepsilon\|^2)$. It follows from these upper bounds that

$$0 \leq E[\|\beta(l)\|^2 | \varepsilon_{1m}(l) = \varepsilon] \leq c_0 + c_1 \|\varepsilon\|^2 \quad (43)$$

where $c_0 = (\mathbf{S}_1 \mathbf{q}_{1m})^\dagger \mathbf{B} (\mathbf{S}_1 \mathbf{q}_{1m}) + 2\lambda_{\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{B} \mathcal{P}_{\mathbf{S}_1}^\dagger}^{\max} \|\mathbf{d}_{1m}\|^2 + 2\text{Re}(\mathbf{d}_{1m}^\dagger \mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{B} \mathbf{S}_1 \mathbf{q}_{1m}) + \|\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{B} \mathbf{S}_1 \mathbf{q}_{1m}\|^2$ and $c_1 = 2\lambda_{\mathcal{P}_{\mathbf{S}_1}^\dagger \mathbf{B} \mathcal{P}_{\mathbf{S}_1}^\dagger}^{\max} + 1$.

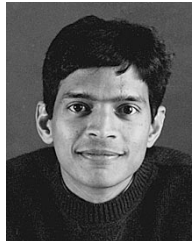
REFERENCES

- [1] M. K. Varanasi and A. Russ, "Noncoherent decorrelative multiuser detection for nonlinear nonorthogonal modulation," *IEEE Trans. Commun.*, vol. 46, pp. 1675–1684, Dec. 1998.
- [2] M. K. Varanasi and M. L. McCloud, "Complex spherical modulation for noncoherent communications," in *Proc. IEEE Int. Symp. Information Theory*, Sorrento, Italy, June 2000, p. 163.
- [3] M. L. McCloud and L. L. Scharf, "Generalized likelihood detection on multiple access channels," in *Proc. 31st Annu. Asilomar Conf. Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 1997, pp. 1033–1037.
- [4] M. K. Varanasi and D. Das, "Noncoherent decision feedback multiuser detection for nonlinear modulation," *IEEE Trans. Commun.*, vol. 48, pp. 259–269, Feb. 2000.
- [5] M. L. McCloud and L. L. Scharf, "Interference estimation with applications to blind multiple access communication," *IEEE Trans. Inform. Theory*, vol. 46, pp. 947–961, May 2000.
- [6] D. J. Sakrison, "Stochastic approximation: A recursive method for solving regression problems," in *Advances in Communication Systems 2*, A. V. Balakrishnan, Ed. New York: Academic, 1966, pp. 51–106.
- [7] R. Lupas and S. Verdu, "Linear multiuser detectors for synchronous code-division multiple-access channels," *IEEE Trans. Inform. Theory*, vol. 35, pp. 123–136, Jan. 1989.
- [8] U. Madhow and M. L. Honig, "MMSE interference suppression for direct-sequence spread-spectrum CDMA," *IEEE Trans. Commun.*, vol. 38, pp. 509–519, Apr. 1990.
- [9] M. L. Honig, U. Madhow, and S. Verdu, "Blind adaptive multiuser detection," *IEEE Trans. Inform. Theory*, vol. 41, pp. 994–960, July 1995.
- [10] D. S. Chen and S. Roy, "An adaptive multiuser receiver for CDMA systems," *IEEE J. Select. Areas Commun.*, vol. 12, pp. 808–816, June 1994.
- [11] X. Wang and V. Poor, "Blind multiuser detection: A subspace approach," *IEEE Trans. Inform. Theory*, vol. 44, pp. 677–689, Mar. 1998.
- [12] S. Ulukus and R. D. Yates, "A blind adaptive decorrelating detector for CDMA systems," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1530–1541, Oct. 1998.
- [13] H. Robbins and S. Monro, "A stochastic approximation method," *Ann. Math. Statist.*, vol. 22, pp. 400–407, 1951.
- [14] Y. Z. Tsytkin, *Adaptation and Learning in Automatic Systems*. New York: Academic, 1971.
- [15] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1993.
- [16] L. Györfi, "Adaptive linear procedures under general conditions," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 262–267, Mar. 1984.
- [17] W. F. Stout, *Almost Sure Convergence*. New York: Academic, 1974.
- [18] G. Ch. Pflug, "Non-Asymptotic Confidence Bounds for Stochastic Approximation Algorithms with Constant Step Size," in *Monatshefte für Mathematik*. Berlin, Germany: Springer-Verlag, 1990, vol. 110, pp. 297–314.
- [19] S. G. Wilson, *Digital Modulation and Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [20] O. Macchi, *Adaptive Processing—The Least Mean Squares Approach with Applications in Transmission*. New York: Wiley, 1995.
- [21] A. R. Amir-Moez, "Extreme properties of eigenvalues of a Hermitian transformation and singular values of the sum and product of linear transformations," *Duke Math. J.*, vol. 23, pp. 463–476, 1956.
- [22] R. Horn and C. Johnson, *Topics in Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1994.



Deepak Das received the B.E. (with honors) degree in electrical and electronics engineering from Birla Institute of Technology and Science, Pilani, India, in 1992, and the M.S. degree in electrical and computer engineering from University of Colorado, Boulder, in 1997.

From 1992 to 1995, he was a Software Engineer/Systems Analyst at Mahindra-British Telecom Ltd., first in Bombay, India, and later in London, U.K. He is currently working toward the Ph.D. degree in electrical engineering at the University of Colorado, Boulder. His research interest has primarily been in the area of multiuser detection and power control.



Mahesh K. Varanasi (S'87–M'89–SM'95) received the B.E. degree in electronics and communication engineering from Osmania University, Hyderabad, India, in 1984, and the M.S. and Ph.D. degrees in electrical engineering from Rice University, Houston, TX, in 1987 and 1989, respectively.

In 1989, he joined the faculty of the University of Colorado at Boulder in the Electrical and Computer Engineering Department where he is now an Associate Professor. His teaching interests include communication theory, information theory, and signal processing. His research interests include multiuser detection, space-time communications, equalization, signal design, diversity communications over fading channels, and power- and bandwidth-efficient multiuser communications.