# Blind Adaptive Algorithms for Minimum Variance CDMA Receivers

Zhengyuan Xu, *Member, IEEE,* and Michail K. Tsatsanis*, Member, IEEE*

***Abstract*—Constrained optimization methods have received considerable attention as a means to derive blind multiuser receivers with low complexity. The receiver's output variance is minimized subject to appropriate constraints which depend on the multipath structure of the signal of interest. When multipath is present, the constraint equations can be written in parametric form, and the constraint parameters jointly optimized with the linear receiver's parameters. In this paper, we develop adaptive solutions for this joint, constrained optimization problem. Both stochastic gradient and recursive least-square-type algorithms are developed. The performance of the proposed methods is compared with other blind and trained methods and turns out to be close to the trained minimum mean-square-error receiver.**

***Index Terms*—Adaptive interference suppression, blind multiuser detection, code-division multiple access, constrained optimization.**

## I. Introduction

CODE-DIVISION multiple-access (CDMA) systems have gained popularity recently, against rival time-division multiple-access/frequency-division multiple-access (TDMA/FDMA) solutions, despite the increased complexity of the receiver [17]. They are prime candidates for third-generation wide-band wireless systems, due to their improved bandwidth efficiency and flexibility. In the CDMA framework, different users employ distinct spreading codes but transmit at the same time and frequency. Therefore, multiuser interference (MUI) exists at the received signal, reducing the performance and creating "near–far" effects.

Significant effort has focused recently on designing multiuser receivers in order to suppress MUI and deliver the promised capacity gains of CDMA technology [7], [10]–[13], [19], [24]. While optimal solutions have been derived [24], their exponentially increasing complexity renders them inappropriate for systems with a large number of users. For this reason, recent efforts have concentrated on suboptimal linear solutions and have investigated ways to derive the receiver parameters from the data [7], [9], [12], [20], [25], [28].

Multiuser receiver design can be divided into two categories based on whether training sequences are used or not. In the first category, nonblind Minimum mean-square-error (MMSE) receivers can be adaptively implemented if the desired signal is known at the receiver [12], while in the second category, a blind approach has to be employed. Blind subspace methods show good performance [1], [15], [28], but the required singular value decomposition (SVD) of large matrices leads to a heavy computational load. On the contrary, blind adaptive multiuser detectors [4], [7], [14], [16], [18], [20], [26], [27] significantly reduce the computational burden.

A scheme was presented in [7] for the case when multipath interference is absent and showed performance equivalent to that of the MMSE receiver. In [27], an adaptive implementation based on a subspace tracking method was shown to improve performance at the expense of more computational complexity. The method of [7] was later extended by adding more constraints [26]. Unfortunately, in the presence of multipath and multichip interference, minimum variance solutions are known to be sensitive to signature mismatch [7]. A solution for that case was attempted in [20] and later in [18], by forcing the receiver response to delayed copies of the signal of interest to zero. With these additional constraints, minimum variance techniques are applicable, but have inferior performance since they treat part of the useful signal as interference.

Recently however, constrained optimization solutions were developed which combine all multipath components of the signal of interest and jointly minimize MUI while maximizing the signal component at the receiver's output [21]. A related development from a different viewpoint was also reported in [9].

These blind methods exhibit superior performance which is close to that of the trained MMSE receiver [21]. Unfortunately however, adaptive implementations of the solution of [21] are not straightforward, as they correspond to the optimization of a nonlinear cost function.

The goal of the current paper is to derive blind adaptive receivers for a multiuser CDMA communication system with reduced complexity by employing constrained optimization techniques. We recursively minimize the output variance of the received signal subject to certain constraints which are also jointly updated. Three different algorithms are proposed in this paper from different viewpoints. Two variations of constrained least mean-square (LMS) methods are obtained by integrating ideas from array processing (see Frost's adaptive method [3]). Another adaptive algorithm corresponding to the generalized sidelobe canceler (GSC) structure is also presented, which can be thought as an updating rule of the equivalent problem

Paper approved by G. E. Corazza, the Editor for Spread Spectrum of the IEEE Communications Society. Manuscript received September 29, 1998; revised November 10, 1999 and March 20, 2000. This work was supported in part by the National Science Foundation under Grant NSF/CAREER CCR-9733048, Grant NSF/NCR-9706658, and Grant NSF/Wireless CCR-9979288, and the Army Research Office under Grant DAAG55-98-1-0224.

Z. Xu is with the Department of Electrical Engineering, University of California, Riverside, CA 92521 USA (e-mail: dxu@ee.ucr.edu).

M. K. Tsatsanis was with Stevens Institute of Technology, Hoboken, NJ 07030 USA. He is now with Voyan Technology, Santa Clara, CA 95054 USA (e-mail: mtsatsanis@voyan.com).

Publisher Item Identifier S 0090-6778(01)00268-9.

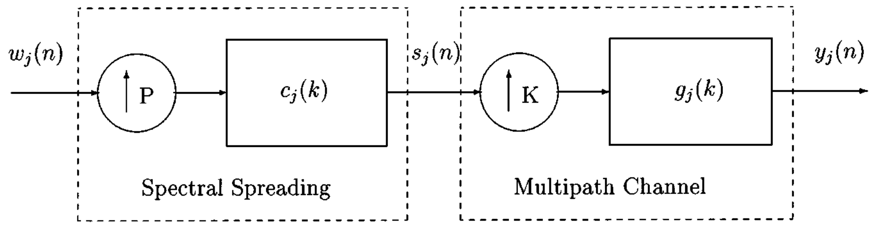


Fig. 1. DS/SS signal in multipath: discrete-time model.

without constraints. These stochastic gradient-based methods suffer from slow convergence as most LMS algorithms do, which offsets their low computational complexity. Therefore the last version of our algorithms with faster convergence is built upon recursive least-square (RLS) ideas.

In contrast with past work in this area [7], the current method can operate in multipath environments by taking into account all multipath components of the signal of interest. Our method also has advantages over the receiver of [9] in the sense that it only updates the coefficients for one filter without using a RAKE receiver structure.

Despite their lower computational complexity however, it turns out that the proposed algorithms exhibit improved performance. Furthermore, it is shown in the sequel that they enjoy global convergence. Simulation results further support these claims according to comparisons with trained MMSE as well as other blind receivers.

The rest of the paper is organized as follows. The discrete-time model of a multiuser CDMA system is described in Section II. Constrained minimum variance techniques are reviewed in Section III in preparation for the development of the proposed adaptive methods. Constrained LMS algorithms are derived in Sections IV and V, while GSC-based and RLS-based methods are obtained in Sections VI and VII, respectively. In Section VIII, the convergence of our algorithms is studied while some simulation results are presented in Section IX. Finally, some conclusions are drawn in the last section.

## II. Problem Statement

In the direct-sequence (DS) CDMA framework, each user transmits digital information after modulating it by a distinct spreading sequence. Let user $j$, $j = 1, \ldots, J$, use a spreading code $c_j(k)$, $k = 0, \ldots, P-1$, of length $P$ to transmit $P$ chips/information symbol. Let the chip sequence be transmitted through a linear channel with a baseband impulse response $g_{c,j}(t)$[1] (including the transmitter and receiver filters), and let the receiver collect $K$ samples/chip. Then, the received discrete-time signal $y_j(n)$ due to user $j$ is (see [21] and Fig. 1)

$$y_j(n) = \sum_{l=-\infty}^{\infty} w_j(l) h_j(n - d_j K - lKP) \quad (1)$$

$$h_j(n) = \sum_{m=-\infty}^{\infty} g_j(m) c_j(n - mK) \quad (2)$$

where $g_j(n) = g_{c,j}(t)|_{t=(nT_c/K)}$ is sampled impulse response of the multipath channel, $h_j(n)$ is the signature of user $j$ [the convolution of the code with the channel—see (2)], $T_c$ is the chip period and $d_j$ is the delay of user $j$ in chip periods.[2] The received signal $y(n)$ at the receiver is a superposition of the signals from all users plus noise $v(n)$

$$y(n) = \sum_{j=1}^{J} y_j(n) + v(n)$$

where $v(n)$ is assumed to be additive white Gaussian noise (AWGN) with zero-mean and variance $\sigma_v^2 = \mathrm{E}\{\|v(n)\|^2\}$. Without loss of generality, we may assume that the delay $0 \le d_j < P$. We will also assume that $g_j(n)$ has finite impulse response of maximum order $q$ (typically $q \ll P$ in many applications). Finally, each user's information bearing sequence $w_j(n)(j = 1, \ldots, J)$ is zeromean, i.i.d., independent of other users with variance $\sigma_{w_j}^2 = \mathrm{E}\{\|w_j(n)\|^2\}$.

Let us collect $L = P + q$ measurements of $y_j(n)$ in a vector $\mathbf{y}_j = [y_j(nP), \ldots, y_j(nP + L - 1)]^T$ $(j = 1, \ldots, J)$. If we consider a single sample per chip ($K = 1$), and if the receiver is synchronized to user 1 ($d_1 = 0$), then signal due to user 1 is[3]

$$\mathbf{y}_1(n) = \mathbf{h}_1 w_1(n) + \tilde{\mathbf{h}}_1 w_1(n-1) + \tilde{\tilde{\mathbf{h}}}_1 w_1(n+1)$$

where $\mathbf{y}_1(n) = [y_1(nP), \ldots, y_1(nP + L - 1)]^T$ is the collection of measurements of $y_1(n)$ and $\mathbf{h}_1 = [h_1(0), \ldots, h_1(P + q - 1)]^T$ is the signature vector of user 1, $\tilde{\mathbf{h}}_1 = [h_1(P), \ldots, h_1(P + q - 1), 0, \ldots, 0]^T$ and $\tilde{\tilde{\mathbf{h}}}_1 = [0, \ldots, 0, h_1(0), \ldots, h_1(q - 1)]^T$ are signatures of the bits $w_1(n-1)$, $w_1(n+1)$. The signal due to other asynchronous users ($j = 2, \ldots, J$, $d_j \ne 0$ and $d_j > q$) is

$$\mathbf{y}_j(n) = \tilde{\mathbf{h}}_{j,a} w_j(n-1) + \tilde{\tilde{\mathbf{h}}}_{j,a} w_j(n)$$

where $\tilde{\mathbf{h}}_{j,a} = [h_j(P - d_j), \ldots, h_j(P + q - 1), 0, \ldots, 0]^T$ and $\tilde{\tilde{\mathbf{h}}}_{j,a} = [0, \ldots 0, h_j(0), \ldots, h_j(P + q - 1 - d_j)]^T$ are signatures from the other $J-1$ interfering users. If we consider the superposition of all users, the received signal will be

$$\mathbf{y}(n) = \sum_{j=1}^{J} \mathbf{y}_j(n) = \mathbf{h}_1 w_1(n) + \mathbf{H}\mathbf{w}(n) + \mathbf{v}(n) \quad (3)$$

where $w_1(n)$ is the signal of interest, while $\mathbf{H} = [\tilde{\mathbf{h}}_1, \tilde{\tilde{\mathbf{h}}}_1, \tilde{\mathbf{h}}_{2,a}, \tilde{\tilde{\mathbf{h}}}_{2,a}, \ldots, \tilde{\mathbf{h}}_{J,a}, \tilde{\tilde{\mathbf{h}}}_{J,a}]$, $\mathbf{w}(n) = [w_1(n-1), w_1(n+1), w_2(n), w_2(n-1), \ldots, w_J(n), w_J(n-1)]^T$, and $\mathbf{v}(n)$ are interference and noise components, respectively.

[1]We use subscript $c$ to denote continuous-time signals.

[2]Fractional delays are absorbed in the channel parameters $g_j(n)$.

[3]In the sequel, we assume without loss of generality that user 1 is the desired user.

Furthermore, notice that the signature of user 1 can be decomposed as [c.f. (2)]

$$\mathbf{h}_1 = \mathbf{C}\mathbf{g}_1 \tag{4}$$

where

$$\mathbf{C} = \begin{bmatrix} c_1(0) & & 0 \\ \vdots & \ddots & c_1(0) \\ c_1(P-1) & & \vdots \\ 0 & \ddots & c_1(P-1) \end{bmatrix}_{(P+q)\times(q+1)}$$
$$\mathbf{g}_1 = \begin{bmatrix} g_1(0) \\ \vdots \\ g_1(q) \end{bmatrix}_{(q+1)\times 1} \tag{5}$$

and $\mathbf{g}_1$ is the unknown multipath parameter vector for user 1. This structure of the user's signature will be exploited in this paper to derive blind adaptive multiuser receivers with the ability to accommodate multipath distortions. In order to construct the code filtering matrix $\mathbf{C}$, we assume that the receiver is synchronized to user 1 and the maximum channel order $q$ (or an upper bound of it) is also available. We will focus on adaptive implementation of the minimum variance batch approach developed in [21]. As will become clear soon, adaptive solutions to this problem are nontrivial due to the nonlinear nature of the cost function. But let us first briefly revisit the method of [21].

## III. Constrained Minimum Variance Receivers

If we focus on linear solutions, the receiver design problem is equivalent to determining a parameter vector $\mathbf{f}$ which provides an estimate of the desired signal

$$\hat{w}_1(n) = \mathbf{f}^H \mathbf{y}(n). \tag{6}$$

Vector $\mathbf{f}$ has the same length as the data vector $\mathbf{y}(n)$. In the case of severe multipath, it is conceivable that the model of (3) can be extended to have $\mathbf{y}(n)$ span multiple bit intervals, resulting in a longer receiver vector. In the interest of clarity, we will maintain the model of (3) in the sequel.

The receiver vector $\mathbf{f}$ may be optimized by minimizing the output variance [7], [20]

$$\mathcal{J} = \mathrm{E}\{\|\hat{w}_1(n)\|^2\} = \mathbf{f}^H \mathbf{R}_y \mathbf{f} \qquad \mathbf{R}_y = \mathrm{E}\{\mathbf{y}(n)\mathbf{y}^H(n)\} \tag{7}$$

subject to the constraint that the response of the user of the interest is a constant

$$\mathbf{f}^H \mathbf{h}_1 = 1 \tag{8}$$

This approach was proposed in [7] in the absence of multipath, in which case the signature $\mathbf{h}_1$ coincides with the user's code $\mathbf{h}_1 = \mathbf{c}_1$ and is *a priori* known. Similar ideas have been extensively studied in array processing in the context of minimum variance distortionless response (MVDR) beamforming [23].

In the multipath case, an extended set of constraints can be used (one constraint for each delayed copy of the signal) [21]

$$\mathbf{C}^H \mathbf{f} = \mathbf{g}. \tag{9}$$

where $\mathbf{g}$ is a parameter vector to be determined. For a given $\mathbf{g}$, the solution to this constrained optimization problem is obtained using Lagrange multipliers [21]

$$\mathbf{f}_{\mathrm{opt}} = \mathbf{R}_y^{-1}\mathbf{C}\left(\mathbf{C}^H \mathbf{R}_y^{-1}\mathbf{C}\right)^{-1}\mathbf{g} \tag{10}$$

while the minimum output variance is

$$\mathcal{J}_{\min} = \mathbf{f}_{\mathrm{opt}}^H \mathbf{R}_y \mathbf{f}_{\mathrm{opt}} = \mathbf{g}^H\left(\mathbf{C}^H \mathbf{R}_y^{-1}\mathbf{C}\right)^{-1}\mathbf{g}. \tag{11}$$

In [21], it was proposed to optimize the unknown vector $\mathbf{g}$ by maximizing $\mathcal{J}_{\min}$, that is, by maximizing the energy of the signal component after the interference has been suppressed.[4] This max/min blind approach exhibits near-optimal performance [21] (see the analysis in Sections IV and V therein) and is related to Capon beamforming ideas in array processing (e.g., [23]).

On the other hand, it involves the inversion of $\mathbf{R}_y$ which may be computationally prohibitive in systems with large spreading factor $P$ [$\mathbf{R}_y$ is a $(P+q)\times(P+q)$ matrix]. In this paper, we derive adaptive algorithms for estimating $\mathbf{f}$ and $\mathbf{g}$ by directly optimizing the Lagrangian cost function and avoiding the inversion of $\mathbf{R}_y$. Both stochastic gradient and RLS solutions are investigated.

## IV. Proposed Constrained LMS Algorithm: Method I

Let us consider the Lagrangian cost function parameterized by $\mathbf{f}$ and $\mathbf{g}$

$$\begin{aligned}\mathcal{J}_1 =& \mathbf{f}^H \mathbf{R}_y \mathbf{f} + \boldsymbol{\lambda}^H\left(\mathbf{C}^H\mathbf{f} - \mathbf{g}\right) \\ &+ \left(\mathbf{f}^H\mathbf{C} - \mathbf{g}^H\right)\boldsymbol{\lambda} + \rho\left(\mathbf{g}^H\mathbf{g} - 1\right)\end{aligned} \tag{12}$$

where $\boldsymbol{\lambda}$ and $\rho$ are the Lagrange multipliers. Notice that we fix $\|\mathbf{g}\| = 1$ in (12). Since we wish to minimize (12) with respect to $\mathbf{f}$ and maximize it with respect to $\mathbf{g}$, we consider the following gradient search procedures:

$$\mathbf{f}_{n+1} = \mathbf{f}_n - \mu_f \nabla_{\mathbf{f}^*}\mathcal{J}_1 \tag{13}$$
$$\mathbf{g}_{n+1} = \mathbf{g}_n + \mu_g \nabla_{\mathbf{g}^*}\mathcal{J}_1 \tag{14}$$

where $\mu_f$, $\mu_g$ are two step sizes. The gradients in (13) and (14) may be obtained from (12) as $\nabla_{\mathbf{f}^*}\mathcal{J}_1 = \mathbf{R}_y\mathbf{f} + \mathbf{C}\boldsymbol{\lambda}$, $\nabla_{\mathbf{g}^*}\mathcal{J}_1 = \rho\mathbf{g} - \boldsymbol{\lambda}$, respectively. Then, (13) and (14) become

$$\mathbf{f}_{n+1} = \mathbf{f}_n - \mu_f(\mathbf{R}_y\mathbf{f}_n + \mathbf{C}\boldsymbol{\lambda}_n) \tag{15}$$
$$\mathbf{g}_{n+1} = \mathbf{g}_n + \mu_g(\rho_n\mathbf{g}_n - \boldsymbol{\lambda}_n) \tag{16}$$

where $\boldsymbol{\lambda}_n$ and $\rho_n$ also need to be updated at each iteration. In order to avoid buildup of roundoff errors in constrained optimization problems, it was suggested in [3] to update the Lagrange multipliers in such a way that constraints are satisfied at

[4]The maximization of (11) is not a well-defined problem unless $\|\mathbf{g}\|$ is fixed to a constant. For this reason a normalized version of (11) is maximized.

TABLE I
PROPOSED CONSTRAINED LMS ALGORITHM: METHOD I

| | |
|---|---|
| Step 1: | Initialize $\mathbf{f}$ and $\mathbf{g}$ at $n=0$, and choose step sizes $\mu_f$, $\mu_g$ |
| Step 2: | Pre-compute $\mathbf{\Pi}_{\mathbf{C}}^{\perp} = \mathbf{I} - \mathbf{C}(\mathbf{C}^H\mathbf{C})^{-1}\mathbf{C}^H$ |
| Step 3: | For $n = 0, 1, 2, \ldots$ |
| | (1) update the receiver $\mathbf{f}_{n+1} = \mathbf{\Pi}_{\mathbf{C}}^{\perp}[\mathbf{f}_n - \mu_f \mathbf{y}(n)\mathbf{y}^H(n)\mathbf{f}_n] + \mathbf{C}(\mathbf{C}^H\mathbf{C})^{-1}\mathbf{g}_n$ |
| | (2) compute $\mathbf{x}_n = \mathbf{g}_n - \frac{\mu_g}{\mu_f}(\mathbf{C}^H\mathbf{C})^{-1}\{\mathbf{C}^H[\mathbf{f}_n - \mu_f\mathbf{y}(n)\mathbf{y}^H(n)\mathbf{f}_n] - \mathbf{g}_n\}$ |
| | (3) compute $a_1 = \mu_g^2\|\mathbf{g}_n\|^2, a_2 = \mu_g(\mathbf{g}_n^H\mathbf{x}_n + \mathbf{x}_n^H\mathbf{g}_n), a_3 = \mathbf{x}_n^H\mathbf{x}_n - 1$ |
| | (4) if $a_2^2 - 4a_1a_3 \geq 0$ then compute $\rho_n = \frac{1}{2a_1}(-a_2 - \sqrt{a_2^2 - 4a_1a_3})$, |
| | and update constraint vector $\mathbf{g}$ as $\mathbf{g}_{n+1} = \mu_g\rho_n\mathbf{g}_n + \mathbf{x}_n$, |
| | otherwise let $\mathbf{g}_{n+1} = \mathbf{g}_n$ (no update for $\mathbf{g}_n$) |

each iteration. By using (15) and enforcing the constraints on $\mathbf{f}$ as $\mathbf{C}^H\mathbf{f}_{n+1} = \mathbf{g}_n$, one can solve for $\boldsymbol{\lambda}_n$

$$\boldsymbol{\lambda}_n = \frac{1}{\mu_f}\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\left(\mathbf{C}^H\mathbf{f}_n - \mu_f\mathbf{C}^H\mathbf{R}_y\mathbf{f}_n - \mathbf{g}_n\right). \quad (17)$$

Substituting (17) in (15), we arrive at the updating rule for $\mathbf{f}$

$$\mathbf{f}_{n+1} = \mathbf{\Pi}_{\mathbf{C}}^{\perp}(\mathbf{f}_n - \mu_f\mathbf{R}_y\mathbf{f}_n) + \mathbf{C}\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{g}_n \quad (18)$$

where

$$\mathbf{\Pi}_{\mathbf{C}}^{\perp} = \mathbf{I} - \mathbf{C}\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{C}^H. \quad (19)$$

It can be seen that $\mathbf{g}_n$ is needed in (18), hence an update equation for $\mathbf{g}$ should be derived as well.

If we substitute (17) in (16), we obtain an update equation for $\mathbf{g}_{n+1}$ as

$$\mathbf{g}_{n+1} = \mu_g\rho_n\mathbf{g}_n + \mathbf{x}_n \quad (20)$$

where

$$\mathbf{x}_n = \mathbf{g}_n - \frac{\mu_g}{\mu_f}\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\left[\mathbf{C}^H(\mathbf{f}_n - \mu_f\mathbf{R}_y\mathbf{f}_n) - \mathbf{g}_n\right]. \quad (21)$$

To get the final update equation for $\mathbf{g}$, we need to determine the unknown $\rho_n$ in (20). Following the ideas of [3] once more, we impose the quadratic constraint on $\mathbf{g}_{n+1}$

$$\mathbf{g}_{n+1}^H\mathbf{g}_{n+1} = \|\mathbf{g}_{n+1}\|^2 = 1 \quad (22)$$

at each iteration. Substituting (20) in (22), a second-order equation for $\rho_n$ is obtained

$$a_1\rho_n^2 + a_2\rho_n + a_3 = 0 \quad (23)$$

where

$$a_1 = \mu_g^2\|\mathbf{g}_n\|^2 \quad a_2 = \mu_g(\mathbf{g}_n^H\mathbf{x}_n + \mathbf{x}_n^H\mathbf{g}_n) \quad a_3 = \mathbf{x}_n^H\mathbf{x}_n - 1. \quad (24)$$

There may be two real solutions if the discriminant of (23) is nonnegative, that is, $\triangle = a_2^2 - 4a_1a_3 \geq 0$. Then, we choose the smaller of the two solutions for $\rho_n$ as

$$\rho_n = \frac{1}{2a_1}\left(-a_2 - \sqrt{a_2^2 - 4a_1a_3}\right) \quad (25)$$

for reasons that will become clear later. The case of $\triangle < 0$ corresponds to incompatible linear and quadratic constraints [23], for certain constrained problems. In our case, however, the optimal solution for $\mathbf{f}$ is guaranteed for every $\mathbf{g}$ [cf. (10)] and therefore the constraints are compatible (at least in the neighborhood of the optimal point). Once $\rho_n$ is obtained, $\mathbf{g}_{n+1}$ can be updated according to (20) and (21). However, no update for $\mathbf{g}$ should be performed in the occasional situation when $\triangle < 0$.

Finally, by using an instantaneous approximation $\hat{\mathbf{R}}_y(n) = \mathbf{y}(n)\mathbf{y}^H(n)$ for $\mathbf{R}_y$, we arrive at a constrained LMS type algorithm. The steps of the algorithm are summarized in Table I.

In this algorithm, all constraints are explicitly considered in one cost function. As will be seen, this will facilitate our theoretical analysis. However, the quadratic constraint on $\mathbf{g}$ results in a second-order equation for $\rho_n$ which has to be solved at each iteration. In order to avoid the extra computations, one might try to abolish the constraint on the norm of $\mathbf{g}$ and instead normalize $\mathbf{g}_n$ at each iteration. This approach results in a different variation of our LMS algorithm, which is discussed next.

## V. PROPOSED CONSTRAINED LMS ALGORITHM: METHOD II

Let us write our Lagrangian cost function now as

$$\mathcal{J}_2 = \mathbf{f}^H\mathbf{R}_y\mathbf{f} + \boldsymbol{\lambda}^H\left(\mathbf{C}^H\mathbf{f} - \mathbf{g}\right) + \left(\mathbf{f}^H\mathbf{C} - \mathbf{g}^H\right)\boldsymbol{\lambda} \quad (26)$$

where $\boldsymbol{\lambda}$ is the Lagrange multiplier corresponding to constraints for our receiver $\mathbf{f}$. Then two update equations for $\mathbf{f}$ and $\mathbf{g}$ can be formed as

$$\mathbf{f}_{n+1} = \mathbf{f}_n - \mu_f\nabla_{\mathbf{f}^*}\mathcal{J}_2 \quad (27)$$

$$\mathbf{g}_{n+1} = \mathbf{g}_n + \mu_g\nabla_{\mathbf{g}^*}\mathcal{J}_2 \quad (28)$$

which will adaptively minimize $\mathcal{J}_2$ with respect to $\mathbf{f}$ and maximize $\mathcal{J}_2$ with respect to $\mathbf{g}$. Since a change in the length of $\mathbf{g}$ does not affects the performance of the receiver $\mathbf{f}$ [see (10), (11)], we may project $\nabla_{\mathbf{g}^*}\mathcal{J}_2$ onto the space orthogonal to $\mathbf{g}$ to obtain the following equation to update $\mathbf{g}$:

$$\mathbf{g}_{n+1} = \mathbf{g}_n + \mu_g\left(\mathbf{I} - \frac{\mathbf{g}_n\mathbf{g}_n^H}{\mathbf{g}_n^H\mathbf{g}_n}\right)\nabla_{\mathbf{g}^*}\mathcal{J}_2. \quad (29)$$

TABLE II
PROPOSED CONSTRAINED LMS ALGORITHM: METHOD II

| | |
|---|---|
| Step 1: | Choose initial conditions for $\mathbf{f}$ and $\mathbf{g}$, and appropriate step sizes $\mu_f$, $\mu_g$ |
| Step 2: | Compute $\mathbf{\Pi}_{\mathbf{C}}^{\perp} = \mathbf{I} - \mathbf{C}(\mathbf{C}^H\mathbf{C})^{-1}\mathbf{C}^H$ |
| Step 3: | For $n = 0, 1, 2, \ldots$ |
| | (1) update receiver $\mathbf{f}_{n+1} = \mathbf{\Pi}_{\mathbf{C}}^{\perp}[\mathbf{f}_n - \mu_f \mathbf{y}(n)\mathbf{y}^H(n)\mathbf{f}_n] + \mathbf{C}(\mathbf{C}^H\mathbf{C})^{-1}\mathbf{g}_n$ |
| | (2) update constraint vector $\mathbf{g}_{n+1}$ |
| | a. compute $\tilde{\mathbf{g}}_{n+1} = \mathbf{g}_n + \frac{\mu_g}{\mu_f}(\mathbf{I} - \frac{\mathbf{g}_n\mathbf{g}_n^H}{\mathbf{g}_n^H\mathbf{g}_n})(\mathbf{C}^H\mathbf{C})^{-1}[\mu_f\mathbf{C}^H\mathbf{y}(n)\mathbf{y}^H(n)\mathbf{f}_n + \mathbf{g}_n - \mathbf{C}^H\mathbf{f}_n]$ |
| | b. normalize $\tilde{\mathbf{g}}_{n+1}$ to obtain $\mathbf{g}_{n+1} = \frac{\tilde{\mathbf{g}}_{n+1}}{\|\tilde{\mathbf{g}}_{n+1}\|}$ |

In order to guarantee the constraint $\|\mathbf{g}\| = 1$ at each iteration, we will normalize $\mathbf{g}_{n+1}$ by

$$\mathbf{g}_{n+1} \leftarrow \frac{\mathbf{g}_{n+1}}{\|\mathbf{g}_{n+1}\|}. \tag{30}$$

By using (26), (27) and (29) become

$$\mathbf{f}_{n+1} = \mathbf{f}_n - \mu_f\left(\mathbf{R}_y\mathbf{f}_n + \mathbf{C}\boldsymbol{\lambda}_n\right) \tag{31}$$

$$\mathbf{g}_{n+1} = \mathbf{g}_n - \mu_g\left(\mathbf{I} - \frac{\mathbf{g}_n\mathbf{g}_n^H}{\mathbf{g}_n^H\mathbf{g}_n}\right)\boldsymbol{\lambda}_n. \tag{32}$$

The Lagrange multiplier $\boldsymbol{\lambda}_n$ is again obtained by enforcing the constraint $\mathbf{C}^H\mathbf{f}_{n+1} = \mathbf{g}_n$ and the result is identical to (17). By substituting (17) in (31) and (32), we obtain the recursions

$$\mathbf{f}_{n+1} = \mathbf{\Pi}_{\mathbf{C}}^{\perp}\left[\mathbf{f}_n - \mu_f\mathbf{R}_y\mathbf{f}_n\right] + \mathbf{C}\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{g}_n \tag{33}$$

$$\begin{aligned}\mathbf{g}_{n+1} = &\mathbf{g}_n + \frac{\mu_g}{\mu_f}\left(\mathbf{I} - \frac{\mathbf{g}_n\mathbf{g}_n^H}{\mathbf{g}_n^H\mathbf{g}_n}\right)\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\\ &\cdot\left[\mu_f\mathbf{C}^H\mathbf{R}_y\mathbf{f}_n + \mathbf{g}_n - \mathbf{C}^H\mathbf{f}_n\right]\end{aligned} \tag{34}$$

where

$$\mathbf{\Pi}_{\mathbf{C}}^{\perp} = \mathbf{I} - \mathbf{C}\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{C}^H. \tag{35}$$

Considering (33)–(35) together with (30) and using the instantaneous approximation $\hat{\mathbf{R}}_y(n) = \mathbf{y}(n)\mathbf{y}^H(n)$ for $\mathbf{R}_y$, we obtain our simpler constrained LMS method, which is summarized in Table II.

## VI. BLIND ADAPTIVE GSC ALGORITHM: METHOD III

In beamforming and array processing literature, linearly constrained optimization problems are often transformed into unconstrained ones. This leads to the separation of the mainlobe (signal part) from the sidelobe (noise part) by using the so-called GSC structure for the receiver [23]. In our problem, we have a linear constraint for the receiver $\mathbf{f}$ as $\mathbf{C}^H\mathbf{f} = \mathbf{g}$ and a quadratic constraint for the constraint vector $\|\mathbf{g}\| = 1$. We can partition this problem into two parts, one part corresponding to the linear constraint which can be transformed to an unconstrained one, the other part corresponding to the normalization step. Following this idea, we begin our derivation by analyzing the matrix $\mathbf{C}$ first.

As can be seen by (5), the columns of $\mathbf{C}$ are shifted versions of the spreading code of the user of interest, so they are linearly independent and make up a signal subspace basis (not necessarily orthogonal). If we consider the matrix $\mathbf{C}_n$ whose columns represent a basis for the orthogonal complement of the space spanned by the columns of $\mathbf{C}$, then the columns of $\mathbf{C}$ and $\mathbf{C}_n$ span the entire space, and we can express any vector $\mathbf{f}$ in this space as

$$\mathbf{f} = \mathbf{C}\mathbf{s} - \mathbf{C}_n\mathbf{u}. \tag{36}$$

Matrix $\mathbf{C}_n$ can be obtained by using SVD on the matrix $\mathbf{C}$ and collecting eigenvectors corresponding to null eigenvalues. Applying the constraint $\mathbf{C}^H\mathbf{f} = \mathbf{g}$ to (36) yields $\mathbf{C}^H\mathbf{C}\mathbf{s} = \mathbf{g}$, and therefore $\mathbf{s} = (\mathbf{C}^H\mathbf{C})^{-1}\mathbf{g}$. Then (36) may be written as

$$\mathbf{f} = \mathbf{C}(\mathbf{C}^H\mathbf{C})^{-1}\mathbf{g} - \mathbf{C}_n\mathbf{u}. \tag{37}$$

In (37), we have managed to parameterize $\mathbf{f}$ by $\mathbf{g}$ and $\mathbf{u}$ in such a way that the constraints are guaranteed to be satisfied. Using this parameterization, we may transform our constrained optimization problem to an unconstrained one for $\mathbf{f}$. We still use the output variance $\mathbf{f}^H\mathbf{R}_y\mathbf{f}$ as our cost function $\mathcal{J}_3$. According to (37), $\mathcal{J}_3$ can be expanded as

$$\begin{aligned}\mathcal{J}_3 = &\mathbf{g}^H\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{C}^H\mathbf{R}_y\mathbf{C}\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{g}\\ &+ \mathbf{u}^H\mathbf{C}_n^H\mathbf{R}_y\mathbf{C}_n\mathbf{u} - \mathbf{g}^H\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{C}^H\mathbf{R}_y\mathbf{C}_n\mathbf{u}\\ &- \mathbf{u}^H\mathbf{C}_n^H\mathbf{R}_y\mathbf{C}\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{g}.\end{aligned} \tag{38}$$

We want to minimize $\mathcal{J}_3$ with respect to $\mathbf{u}$ and maximize it with respect to $\mathbf{g}$. Therefore, we may construct two update equations for $\mathbf{u}$ and $\mathbf{g}$, respectively

$$\mathbf{u}_{n+1} = \mathbf{u}_n - \mu_u\nabla_{\mathbf{u}^*}\mathcal{J}_3 \tag{39}$$

$$\mathbf{g}_{n+1} = \mathbf{g}_n + \mu_g\nabla_{\mathbf{g}^*}\mathcal{J}_3. \tag{40}$$

The derivatives in (39) and (40) are computed by using (38)

$$\nabla_{\mathbf{u}^*}\mathcal{J}_3 = \mathbf{C}_n^H\mathbf{R}_y\mathbf{C}_n\mathbf{u} - \mathbf{C}_n^H\mathbf{R}_y\mathbf{C}\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{g} \tag{41}$$

$$\begin{aligned}\nabla_{\mathbf{g}^*}\mathcal{J}_3 = &\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{C}^H\mathbf{R}_y\mathbf{C}\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{g}\\ &- \left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{C}^H\mathbf{R}_y\mathbf{C}_n\mathbf{u}.\end{aligned} \tag{42}$$

Substituting (41) into (39) and (42) into (40), the recursive algorithm for computing both $\mathbf{u}$ and $\mathbf{g}$ is

$$\begin{aligned}\mathbf{u}_{n+1} = &\mathbf{u}_n - \mu_u\mathbf{C}_n^H\mathbf{R}_y\mathbf{C}_n\mathbf{u}_n\\ &+ \mu_u\mathbf{C}_n^H\mathbf{R}_y\mathbf{C}\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{g}_n\end{aligned} \tag{43}$$

$$\begin{aligned}\mathbf{g}_{n+1} = &\mathbf{g}_n + \mu_g\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{C}^H\mathbf{R}_y\mathbf{C}\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{g}_n\\ &- \mu_g\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{C}^H\mathbf{R}_y\mathbf{C}_n\mathbf{u}_n.\end{aligned} \tag{44}$$

TABLE III
PROPOSED GSC ALGORITHM: METHOD III

| | |
|---|---|
| Step 1: | Choose initial values for $\mathbf{u}$ and $\mathbf{g}$, and appropriate step sizes $\mu_u$, $\mu_g$ |
| Step 2: | Compute $\mathbf{B} = (\mathbf{C}^H\mathbf{C})^{-1}$ and perform SVD on $\mathbf{C}$ to obtain $\mathbf{C}_n$ |
| Step 3: | For $n = 0,\ 1,\ 2,\ \ldots$ |
| | (1) compute $\tilde{\mathbf{y}}_n = \mathbf{C}^H\mathbf{y}_n$, $\bar{\mathbf{y}}_n = \mathbf{C}_n^H\mathbf{y}_n$, $\omega = \tilde{\mathbf{y}}_n^H\mathbf{B}\mathbf{g}_n - \bar{\mathbf{y}}_n^H\mathbf{u}_n$ |
| | (2) update $\mathbf{u}_{n+1} = \mathbf{u}_n + \omega\mu_u\bar{\mathbf{y}}_n$ |
| | (3) update $\mathbf{g}_{n+1} = \mathbf{g}_n + \omega\mu_g\mathbf{B}\tilde{\mathbf{y}}_n$ |
| | (4) normalize $\mathbf{g}_{n+1}$ as $\mathbf{g}_{n+1} \leftarrow \frac{\mathbf{g}_{n+1}}{\|\mathbf{g}_{n+1}\|}$ |
| | (5) construct receiver $\mathbf{f}_{n+1} = \mathbf{C}\mathbf{B}\mathbf{g}_{n+1} - \mathbf{C}_n\mathbf{u}_{n+1}$ |

Each step is followed by normalizing $\mathbf{g}_{n+1}$ as

$$\mathbf{g}_{n+1} \leftarrow \frac{\mathbf{g}_{n+1}}{\|\mathbf{g}_{n+1}\|} \tag{45}$$

in order to guarantee that the quadratic constraint for $\mathbf{g}$ is satisfied. Then, based on (37), our receiver vector at iteration $n+1$ can be constructed as

$$\mathbf{f}_{n+1} = \mathbf{C}\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{g}_{n+1} - \mathbf{C}_n\mathbf{u}_{n+1}. \tag{46}$$

Equations (43)–(46) form our blind adaptive GSC algorithm. Substituting $\mathbf{R}_y$ by $\mathbf{y}(n)\mathbf{y}^H(n)$, we can further simplify the algorithm. If we define

$$\tilde{\mathbf{y}}_n = \mathbf{C}^H\mathbf{y}_n \qquad \bar{\mathbf{y}}_n = \mathbf{C}_n^H\mathbf{y}_n$$
$$\mathbf{B} = \left(\mathbf{C}^H\mathbf{C}\right)^{-1} \qquad \omega = \tilde{\mathbf{y}}_n^H\mathbf{B}\mathbf{g}_n - \bar{\mathbf{y}}_n^H\mathbf{u}_n$$

then (43), (44), and (46) become

$$\mathbf{u}_{n+1} = \mathbf{u}_n + \omega\mu_u\bar{\mathbf{y}}_n \tag{47}$$
$$\mathbf{g}_{n+1} = \mathbf{g}_n + \omega\mu_g\mathbf{B}\tilde{\mathbf{y}}_n \tag{48}$$
$$\mathbf{f}_{n+1} = \mathbf{C}\mathbf{B}\mathbf{g}_{n+1} - \mathbf{C}_n\mathbf{u}_{n+1} \tag{49}$$

respectively. We can thus summarize this LMS-GSC algorithm in Table III.

## VII. PROPOSED RLS ALGORITHM: METHOD IV

All three algorithms we have developed in Sections IV–VI (Methods I–III) are stochastic gradient-based methods where the instantaneous approximation for the data covariance matrix $\mathbf{R}_y$ is used. When communication channels experience fast frequency-selective fading, the structure of the interference may change drastically in short time. Hence, this approximation becomes more inaccurate. Furthermore, various near–far situations and signal-to-noise ratio (SNR) scenarios may affect the eigenvalue spread of $\mathbf{R}_y$. Thus, all previous methods may experience slow convergence rates in some cases. For this reason, we explore RLS-based solutions with faster convergence in the sequel despite their higher computational complexity.

Motivated by the optimal solutions for the receiver $\mathbf{f}$ and our constraint vector $\mathbf{g}$ in (10) and (11), respectively, we need to compute $\mathbf{R}_y^{-1}$ recursively to facilitate the computation of the eigenvector of the matrix $(\mathbf{C}^H\mathbf{R}_y^{-1}\mathbf{C})^{-1}$ corresponding to its maximum eigenvalue (or the eigenvector of the matrix $\mathbf{C}^H\mathbf{R}_y^{-1}\mathbf{C}$ corresponding to its minimum eigenvalue) while reducing the number of computations. As described in [5], we can use Kalman RLS recursions to update $\hat{\mathbf{R}}_y^{-1}$ as follows:

$$\mathbf{k}(n) = \frac{\hat{\mathbf{R}}_y^{-1}(n-1)\mathbf{y}(n)}{\nu + \mathbf{y}^H(n)\hat{\mathbf{R}}_y^{-1}(n-1)\mathbf{y}(n)} \tag{50}$$

$$\hat{\mathbf{R}}_y^{-1}(n) = \frac{1}{\nu}\hat{\mathbf{R}}_y^{-1}(n-1) - \frac{1}{\nu}\mathbf{k}(n)\mathbf{y}^H(n)\hat{\mathbf{R}}_y^{-1}(n-1) \tag{51}$$

where $0 < \nu \leq 1$ is the forgetting factor. The algorithm can be initialized by $\hat{\mathbf{R}}_y^{-1}(0) = \delta\mathbf{I}$ where $\delta$ is a large positive number. In this way, $\hat{\mathbf{R}}_y(n)$ and thus $\mathbf{C}^H\hat{\mathbf{R}}_y^{-1}(n)\mathbf{C}$ are Hermitian. Once $\hat{\mathbf{R}}_y^{-1}(n)$ is updated, we can form the matrix $\mathbf{C}^H\hat{\mathbf{R}}_y^{-1}(n)\mathbf{C}$. Since the latter is a small size matrix with dimension $(q+1)\times(q+1)$ (recall $q \ll P$ under our assumption), its eigenvector $\mathbf{g}_{\min}(n)$ corresponding to its minimum eigenvalue $\alpha_{\min}$ may be obtained by directly applying SVD decomposition. The embedded arbitrary phase, inherent in obtaining $\mathbf{g}_{\min}(n)$ from SVD, may affect the continuity of the receiver vectors $\mathbf{f}(n)$ for successive $n$'s. It might be advisable therefore to normalize the phase of $\mathbf{g}_{\min}(n)$ by setting the phase of one of its elements to zero. Alternatively, subspace tracking procedures may be employed to further reduce the computations [2]. Finally, by utilizing the formula [c.f. (10)]

$$\mathbf{f}_n = \alpha_{\min}\hat{\mathbf{R}}_y^{-1}(n)\mathbf{C}\mathbf{g}_{\min}(n) \tag{52}$$

we can obtain the receiver $\mathbf{f}_n$ at iteration $n$. The steps of this method are summarized in Table IV.

There are more efficient ways to update the matrix $\mathbf{C}^H\hat{\mathbf{R}}_y^{-1}(n)\mathbf{C}$ (e.g., using Cholesky factors [8, p. 408]), but we will not discuss them here due to lack of space. Instead, we will turn our attention to convergence issues.

## VIII. CONVERGENCE ANALYSIS

In this section, we investigate the convergence behavior of our algorithms in terms of global/local minima, trajectory of the mean receiver vector (or tap vector) and the steady-state excess mean square error (MSE). We focus on Method I with all constraints in a single cost function as more convenient for analysis.

TABLE IV
PROPOSED RLS ALGORITHM: METHOD IV

| | |
|---|---|
| Step 1: | Choose $0 < \nu \leq 1$ and $\hat{\mathbf{R}}_y^{-1}(0) = \delta \mathbf{I}$ ($\delta$ is a large positive number) |
| Step 2: | For $n = 1, 2, \ldots$ |
| | (1) compute $\mathbf{k}(n) = \frac{\hat{\mathbf{R}}_y^{-1}(n-1)\mathbf{y}(n)}{\nu + \mathbf{y}^H(n)\hat{\mathbf{R}}_y^{-1}(n-1)\mathbf{y}(n)}$ |
| | (2) update $\hat{\mathbf{R}}_y^{-1}(n) = \frac{1}{\nu}\hat{\mathbf{R}}_y^{-1}(n-1) - \frac{1}{\nu}\mathbf{k}(n)\mathbf{y}^H(n)\hat{\mathbf{R}}_y^{-1}(n-1)$ |
| | (3) compute the matrix $\mathbf{C}^H\hat{\mathbf{R}}_y^{-1}(n)\mathbf{C}$ and apply SVD to get its eigenvector $\mathbf{g}_{min}(n)$ corresponding to its minimum eigenvalue $\alpha_{min}$ |
| | (4) construct the receiver $\mathbf{f}_n = \alpha_{min}\hat{\mathbf{R}}_y^{-1}(n)\mathbf{C}\mathbf{g}_{min}(n)$ |

### A. Global Convergence

The output variance $\mathbf{f}^H\mathbf{R}_y\mathbf{f}$ is a quadratic function of $\mathbf{f}$ which might lead one to think that the issue of global convergence is a trivial one. In the current situation, however, $\mathbf{f}$ has to be jointly optimized with $\mathbf{g}$ [cf. (12)] and the presence of local minima can not be *a priori* dismissed.

By studying the first as well as the second derivatives of the cost function (12), it is shown in Appendix I that our algorithm guarantees a unique convergent point for $(\mathbf{f}, \mathbf{g})$. Any other stationary points are unstable except the desired optimum point. Therefore, the proposed algorithm enjoys global convergence, and thus imposes no restrictions on initialization. Next, we study other properties of the proposed algorithms in terms of the mean tap vector and excess steady-state MSE.

### B. Trajectory of the Mean Tap Vector and Constraint Vector

To proceed, let us first define the tap error vector $\mathbf{e}_f[n]$ and the constraint error vector $\mathbf{e}_g[n]$ at time $n$

$$\mathbf{e}_f[n] = \mathbf{f}_n - \mathbf{f}_{\rm opt},\ \mathbf{e}_g[n] = \mathbf{g}_n - \mathbf{g}_{\rm opt} \tag{53}$$

where $\mathbf{g}_{\rm opt}$ is the optimal constraint vector which is the eigenvector of matrix $(\mathbf{C}^H\mathbf{R}_y^{-1}\mathbf{C})^{-1}$ corresponding to its maximum eigenvalue $\xi_{\max}$, and $\mathbf{f}_{\rm opt}$ is the optimal receiver given by

$$\mathbf{f}_{\rm opt} = \mathbf{R}_y^{-1}\mathbf{C}\left(\mathbf{C}^H\mathbf{R}_y^{-1}\mathbf{C}\right)^{-1}\mathbf{g}_{\rm opt} = \xi_{\max}\mathbf{R}_y^{-1}\mathbf{C}\mathbf{g}_{\rm opt}. \tag{54}$$

These two error vectors have to be considered together because of the joint optimization procedure. It is shown in Appendix II that their trajectory abides by the following equation:

$$\begin{bmatrix} E\{\mathbf{e}_f[n+1]\} \\ E\{\mathbf{e}_g[n+1]\} \end{bmatrix} = \mathbf{A}\begin{bmatrix} E\{\mathbf{e}_f[n]\} \\ E\{\mathbf{e}_g[n]\} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mu_g(\xi_{\max} + \overline{\rho}_n)\mathbf{g}_{\rm opt} \end{bmatrix} \tag{55}$$

where $\overline{\rho}_n = E\{\rho_n\}$, "$E$" represents expectation and matrix A is shown in the equation at the bottom of the page. Equation (55) implies that the stability of the proposed method depends on the stability of matrix $\mathbf{A}$. Therefore, the step sizes $\mu_g$, $\mu_f$ should be chosen such that the eigenvalues of $\mathbf{A}^H\mathbf{A}$ have magnitude less than one. Unfortunately, the problem is further involved due to the presence of $\overline{\rho}_n$ in $\mathbf{A}$; $\overline{\rho}_n$ is not updated through a linear recursion and the study of its trajectory is intractable. Still, some insight on the stability of the algorithm in the neighborhood of the desired solution can be obtained by letting $\overline{\rho}_n \simeq -\xi_{\max}$. Equation (55) also indicates some necessary conditions on the step sizes.

### C. Trajectory of Excess MSE

Our method is a minimum output energy (MOE) approach and its steady-state excess MSE analysis follows the general steps presented in [7], although the latter deals with the nonfrequency-selective channel. In the current context, however, the multipath channel is taken into account.

Let the MSE at time $n+1$ be (see [7])

$$\epsilon[n+1] = {\rm MSE}(\mathbf{f}_{n+1}) = E\{\|w_1(n+1) - \mathbf{f}_{n+1}^H\mathbf{y}_{n+1}\|^2\}. \tag{56}$$

By noticing that $\mathbf{f}_{n+1} = \mathbf{f}_{\rm opt} + \mathbf{e}_f[n+1]$ and using the independence assumption, (56) becomes

$$\begin{aligned}\epsilon[n+1] = &\ \epsilon_{\min} + \sigma_{w_1}^2 E\left\{\mathbf{h}_1^H\mathbf{e}_f[n+1] + \mathbf{e}_f^H[n+1]\mathbf{h}_1\right\} \\ &+ {\rm tr}\, E\left\{\mathbf{R}_y\mathbf{f}_{\rm opt}\mathbf{e}_f^H[n+1]\right\} \\ &+ {\rm tr}\, E\left\{\mathbf{R}_y\mathbf{e}_f[n+1]\mathbf{f}_{\rm opt}^H\right\} + \eta[n+1]\end{aligned} \tag{57}$$

where $\epsilon_{\min}$ is the MSE with optimal receiver $\mathbf{f}_{\rm opt}$

$$\epsilon_{\min} = E\left\{\left\|w_1(n+1) - \mathbf{f}_{\rm opt}^H\mathbf{y}_{n+1}\right\|^2\right\}$$

and

$$\eta[n+1] = {\rm tr}\, E\left\{\mathbf{R}_y\mathbf{e}_f[n+1]\mathbf{e}_f^H[n+1]\right\}. \tag{58}$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{\Pi}_{\mathbf{C}}^{\perp}(\mathbf{I} - \mu_f\mathbf{R}_y) & \mathbf{C}\left(\mathbf{C}^H\mathbf{C}\right)^{-1} \\ \frac{\mu_g}{\mu_f}\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{C}^H(\mu_f\mathbf{R}_y - \mathbf{I}) & (1 + \mu_g\overline{\rho}_n)\mathbf{I} + \frac{\mu_g}{\mu_f}\left(\mathbf{C}^H\mathbf{C}\right)^{-1} \end{bmatrix}$$

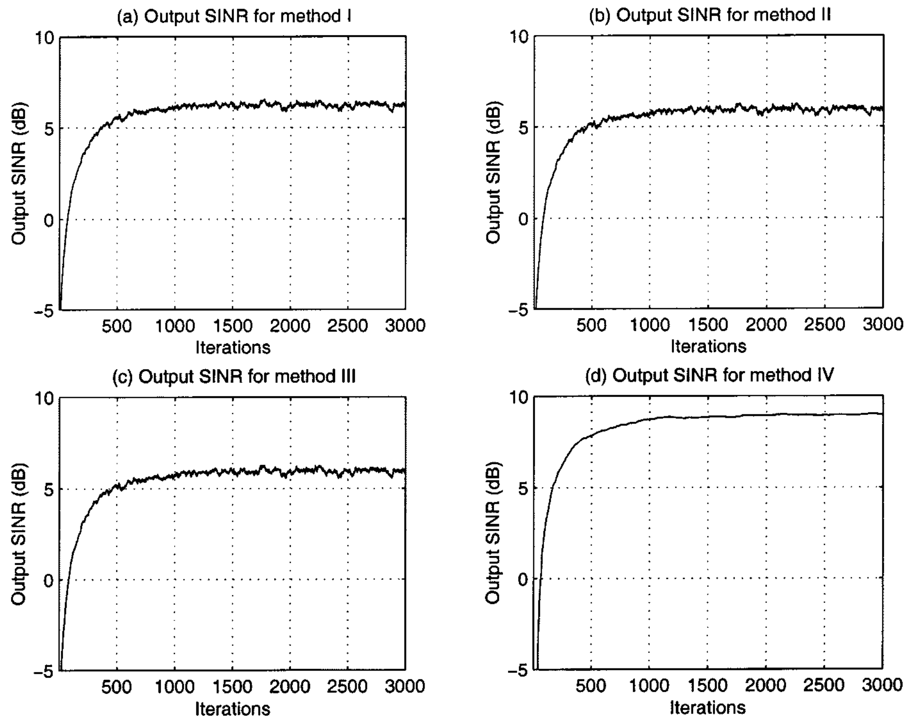


Fig. 2. SINR comparison for different proposed methods.

Since $\lim_{n\to\infty} E\{\mathbf{e}_f[n]\} = \mathbf{0}$, we have

$$\lim_{n\to\infty} \epsilon[n+1] = \epsilon_{\min} + \lim_{n\to\infty} \eta[n+1]. \tag{59}$$

The second term in (59) is the steady-state excess MSE due to adaptation. If we denote it by $\overline{\eta}_{ex}$, then it is related to the tap vector error by

$$\overline{\eta}_{ex} = \lim_{n\to\infty} \operatorname{tr} E\left\{\mathbf{R}_y \mathbf{e}_f[n+1]\mathbf{e}_f^H[n+1]\right\}. \tag{60}$$

It is shown in Appendix III that $\overline{\eta}_{ex}$ is determined by the second- and fourth-order statistics of the received data vector $\mathbf{y}_n$, namely

$$\overline{\eta}_{ex} \approx \mu_f \operatorname{vec}^H(\mathbf{R}_y)\mathbf{\Theta}^{-1}\mathbf{a} \tag{61}$$

where

$$\begin{aligned}\mathbf{\Theta} = &\left(\mathbf{R}_y \mathbf{\Pi}_{\mathbf{C}}^{\perp}\right)^T \otimes \mathbf{I} + \mathbf{I} \otimes \left(\mathbf{\Pi}_{\mathbf{C}}^{\perp}\mathbf{R}_y\right) - \mu_f \left[\left(\mathbf{\Pi}_{\mathbf{C}}^{\perp}\right)^T \otimes \mathbf{\Pi}_{\mathbf{C}}^{\perp}\right]\\ &\cdot E\left\{\left(\mathbf{y}_n\mathbf{y}_n^H\right)^T \otimes \left(\mathbf{y}_n\mathbf{y}_n^H\right)\right\}\\ \mathbf{a} = &\left[\left(\mathbf{\Pi}_{\mathbf{C}}^{\perp}\right)^T \otimes \mathbf{\Pi}_{\mathbf{C}}^{\perp}\right] E\left\{\left(\mathbf{y}_n\mathbf{y}_n^H\right)^T \otimes \left(\mathbf{y}_n\mathbf{y}_n^H\right)\right\}\\ &\cdot \operatorname{vec}(\mathbf{f}_{\text{opt}}\mathbf{f}_{\text{opt}}^H)\end{aligned}$$

and "$\otimes$" represents the Kronecker product, "vec" is the "vec" operation to arrange all elements of a matrix into a vector along column-wise. Without further assumptions, the expression of (61) is quite involved due to the high-order statistics $E\{(\mathbf{y}_n\mathbf{y}_n^H)^T \otimes (\mathbf{y}_n\mathbf{y}_n^H)\}$ of the data $\mathbf{y}_n$. However, it can be seen that $\overline{\eta}_{ex}$ increases almost linearly with the step size $\mu_f$, similar to the result in [7].

It is true that the convergence properties of the adaptive algorithm highly depend on the variations of the communication environment. The algorithm may fail to track the fast changing channel if the adaptation speed is too slow. Thus, the step size should be chosen as a compromise between excess error and convergence rate to make the algorithm more suitable for a particular communication scenario. The choice of step size is also affected by the dimension of the constraint subspace in the current context. When a long channel is experienced which consequently increases the dimension of constraint vector, the receiver vector has less freedom to be adjusted. Then, inferior performance could be expected.

## IX. Simulations

The applicability and performance of the different proposed methods were verified by our simulation results. We used the average output signal-to-interference-and-noise ratio (SINR) at the receiver end as a performance index, and we compared our methods with the trained MMSE receiver and the receivers presented in [7] and [6], which are both MOE-based methods. A CDMA system was simulated with quadrature phase-shift keying modulation and ten users in a 15-dB AWGN environment. Gold sequences of length 31 were used as code sequences. All inputs took values from $\{\pm 1 \quad \pm i\}$ where $i^2 = -1$. User 1 was assumed to be the desired user which was 5 dB weaker than each of the other nine equal-power users. The receiver was synchronized to this user. Signals from other users arriving at the receiver with arbitrary delays $d_j (j = 2, \ldots, 10)$ between 0–31 chip periods. Channel vectors for all users were also randomly generated. Fifty Monte Carlo runs were performed for each experiment with $N = 3000$ bits/realization. The input sequences, the AWGN sequence, and the channel

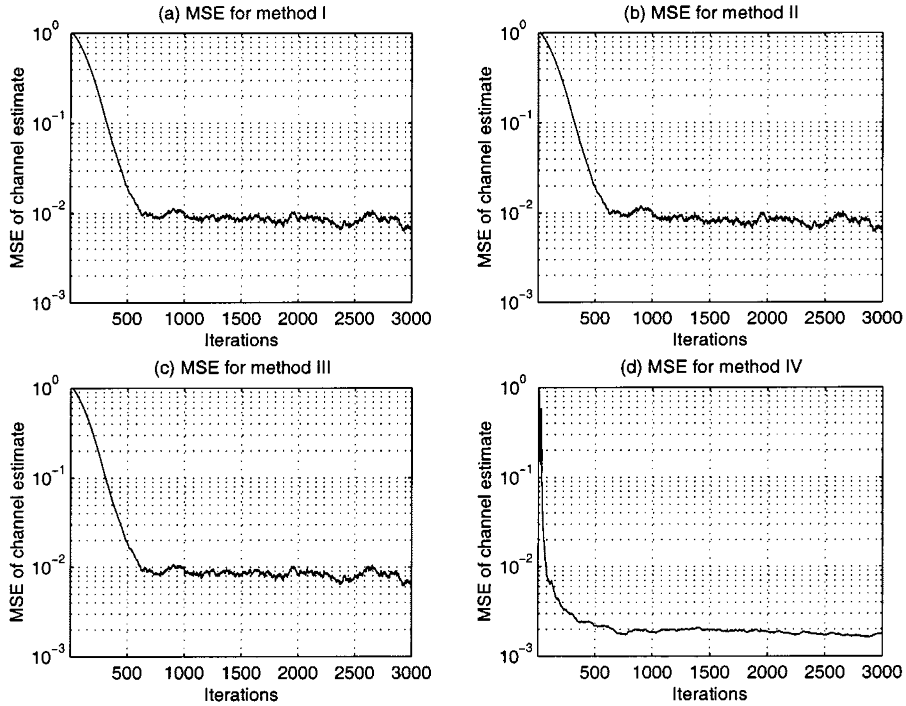


Fig. 3. Comparison of MSE of channel estimates for different proposed methods.

vectors changed from one realization to another, while other parameters were kept unchanged to obtain the average results.

First, we compared four different proposed methods (see Tables I–IV for methods I–IV) and show the corresponding output SINR in Fig. 2(a)–(d), respectively. The initial values were set to be $\mathbf{g} = [1,\ 0,\ 0,\ 0]'$, $\mathbf{u} = \mathbf{0}$, and $\mathbf{f} = \mathbf{0}$. The step sizes for updating these vectors were $\mu_f = \mu_u = 0.0009$, $\mu_g = 0.014$. In implementing Method IV, $\delta = 1000$, $\nu = 0.998$. It can be seen that Methods I–III have similar performance. However, Method IV shows better performance at the expense of increased computational cost. This result is not surprising if we investigate the recursive behavior of our constraint vector $\mathbf{g}$. Motivated by the analysis result in [21] that this vector will asymptotically converge to the normalized channel vector $\mathbf{g}_1/||\mathbf{g}_1||$ within a phase ambiguity, we also compare it with $\mathbf{g}_1/||\mathbf{g}_1||$ by using the phase of $g_1(0)$ (the first element in $\mathbf{g}_1$) as a reference to remove the ambiguity. Thus, the MSEs of channel estimates were obtained and are plotted in Fig. 3(a)–(d). Again similar convergence properties are observed for Methods I–III, while Method IV gives lower error level with the same iteration.

Next, we tested the AWGN effect on the performance of one of our algorithms (Method I) to gain some insight. We obtained the SINR in Fig. 4(a)–(d) by adding Gaussian noise to the system with different levels such as 5, 10, 15, and 20 dB. When $\text{SNR} = 5$ dB, the SINR converges to 0 dB. For other three SNR levels, the output SINR converges to 4, 6, and 8 dB, respectively. These improvement are also reflected by the channel estimation errors for these four different situations, as indicated by Fig. 5(a)–(d). When $\text{SNR} = 5$ dB, the MSE of channel estimate converges to $7 \times 10^{-2}$, while $2 \times 10^{-2}$, $0.9 \times 10^{-2}$, $0.7 \times 10^{-2}$ convergence levels are observed for $\text{SNR} = 10, 15,$ and 20 dB, respectively. It is shown in [21] that as $\text{SNR} \to \infty$, the constraint vector $\mathbf{g}$ will eventually converge to $\mathbf{g}_1/||\mathbf{g}_1||$ within a phase ambiguity. Thus, the vector $\mathbf{g}$ can be used to estimate the multipath channel when the background is almost "noise-free."

Our next experiment was performed to show how our algorithm behaves in a near–far communication environment. All parameters are the same as those for Method I in the first experiment except the relative power for the desired user (user 1). Three cases were tested where the power $\sigma_{w_1}^2$ equals or is 5 dB weaker or 10 dB weaker than each of the other nine equally-powered interfering users. The simulation results are shown in Fig. 6 for SINR and in Fig. 7 for the channel estimation MSE. Solid lines are for the 0 dB weaker (or equal power) case, dash-dotted lines for the 5 dB weaker case, and dashed lines for the 10 dB weaker one. It can be seen from both figures that the transition and convergence rates of the method are different in different cases. When all users are assigned equal power (or under power control), the algorithm converges much faster (about 300 iterations are required), while it requires more than 3000 iterations to converge when user 1 is much weaker (10 dB weaker). However, either SINR or channel estimation MSE converges to approximately the same level despite the different power assignments. Therefore, the algorithm exhibits near–far resistance according to our experimental results.

Next, we compare Method I with the nonblind MMSE receiver, the MOE-based adaptive multiuser detector [7] (simply called MOE) and the adaptive orthogonally anchored MMSE (OAMMSE) detector [6] in terms of the output SINR. All algorithms were implemented with the same step size 0.0009 and initial values to have a fair comparison. User 1 is still 5 dB weaker in power. The equal-gain combining (EGC) technique was used for [6]. All results are depicted in Fig. 8. In the order from top to bottom after 500 iterations, the dashed, solid, dash-dotted, and dotted lines represent the results for MMSE, the

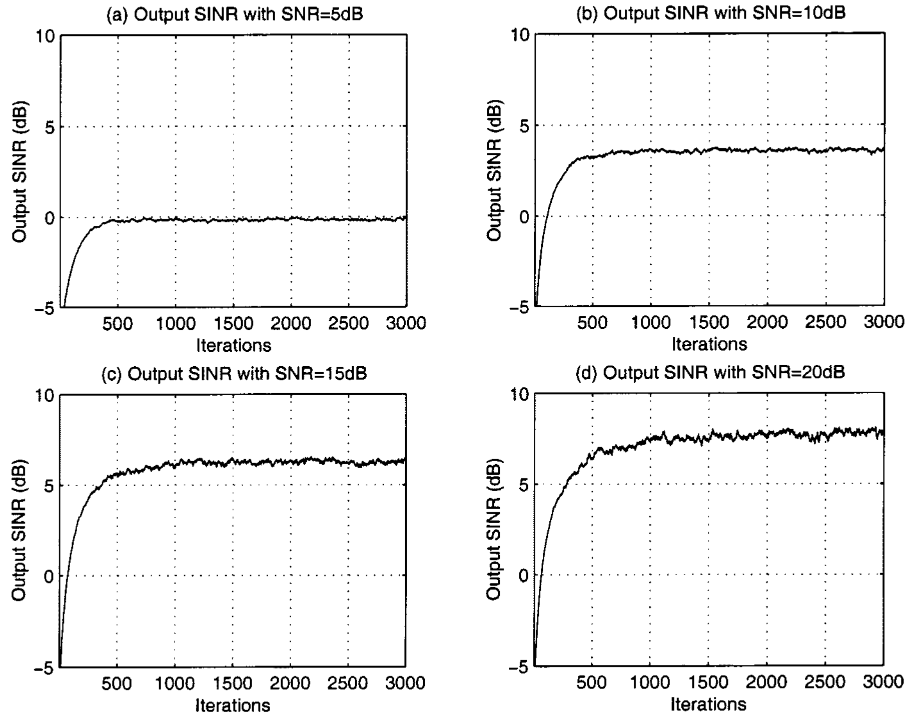


Fig. 4. AWGN effect on the output SINR.

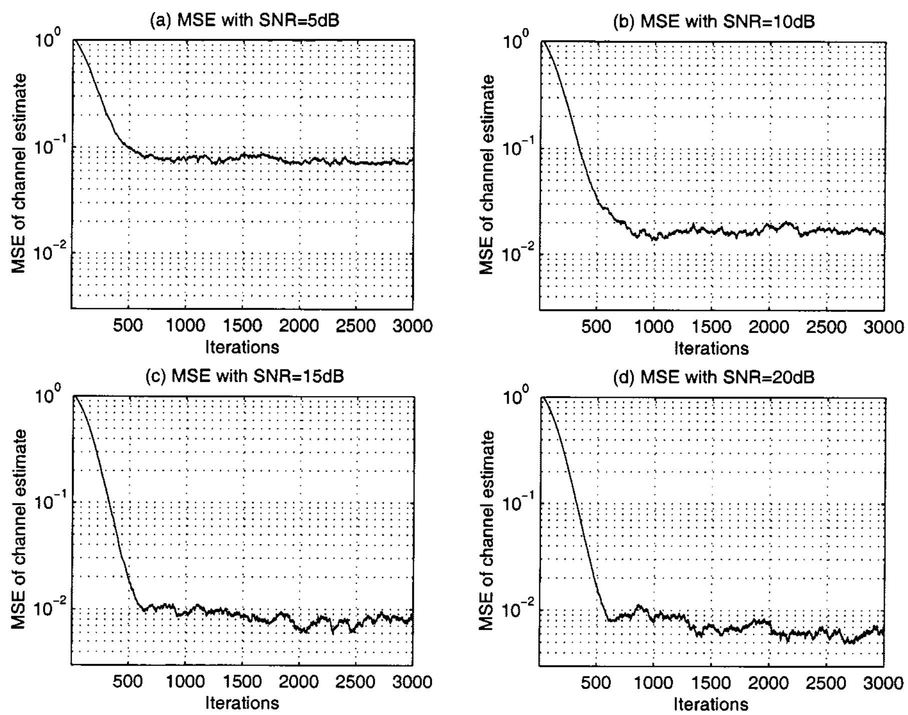


Fig. 5. AWGN effect on channel estimation.

proposed, OAMMSE, and MOE, respectively. Different convergence levels can be seen from this figure. After 3000 iterations, they reach 8, 4.5, 3, and −12 dB, respectively. It is clear that the proposed method can compensate for multipath effect and suppress MUI successfully even in the near–far situation. Our method also performs closer to the trained MMSE receiver, but about 4-dB loss is observed. This performance loss is explained by the greater misadjustment of the blind method due to the fact that the blind cost function is orders of magnitude greater than the MSE one, even at the optimum point. However, our method shows better result than that of OAMMSE (about 1.5-dB difference) for the current setup. One major reason behind this is that our method can somehow estimate the channel parameters adaptively, while without the knowledge for the channel, the EGC technique used in OAMMSE does not always provide satisfactory results for various communication systems with multipath

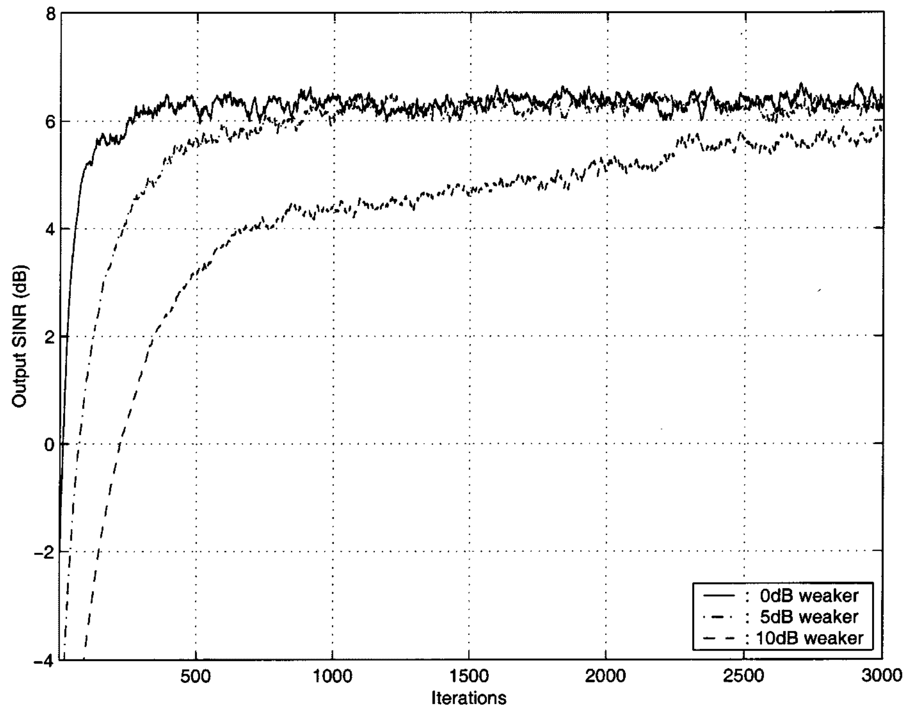


Fig. 6. Near–far effect on the output SINR.

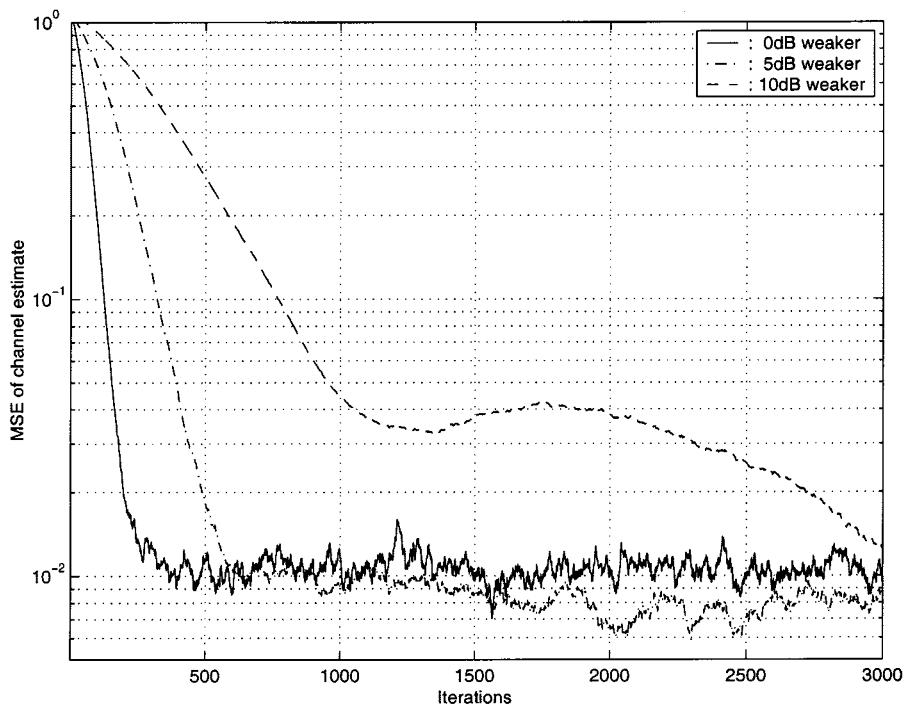


Fig. 7. Near–far effect on channel estimation.

distortions. If we look into the computational complexity for these two methods, it can be found that extra steps have to be taken to update the constraint vector $\mathbf{g}$ in the proposed method, which results in additional complexity of about $O((q+1)^2+(q+1)(P+q))$. This extra cost is not significant if the channel order $q$ is small. But in high speed data communications it becomes not negligible. Since the MOE algorithm [7] was developed for flat channels, it is no surprise that it does not work well in the presence of multipath distortions.

The last experiment was performed to test the performance of our method (Method I) in a Rayleigh fading communication environment. All ten users in the system were assumed to have equal power, and the receiver was synchronized to user 1. The channel parameters for user 1 were obtained by sampling the following continuous function generated from a two-ray Rayleigh fading model (see [17, ch. 4])

$$h(t;\tau) = 0.4\alpha_1(t)\delta(\tau) + 0.7\alpha_2(t)\delta(\tau - \tau_1)$$

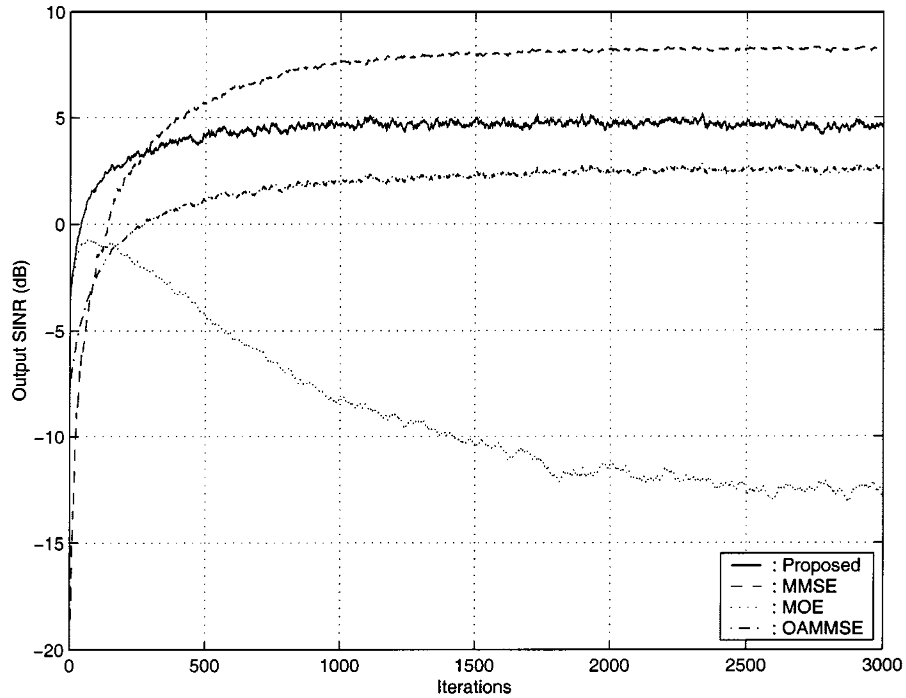


Fig. 8. SINR comparison for the nonblind MMSE receiver, the proposed receiver, the OAMMSE detector, and MOE detector.

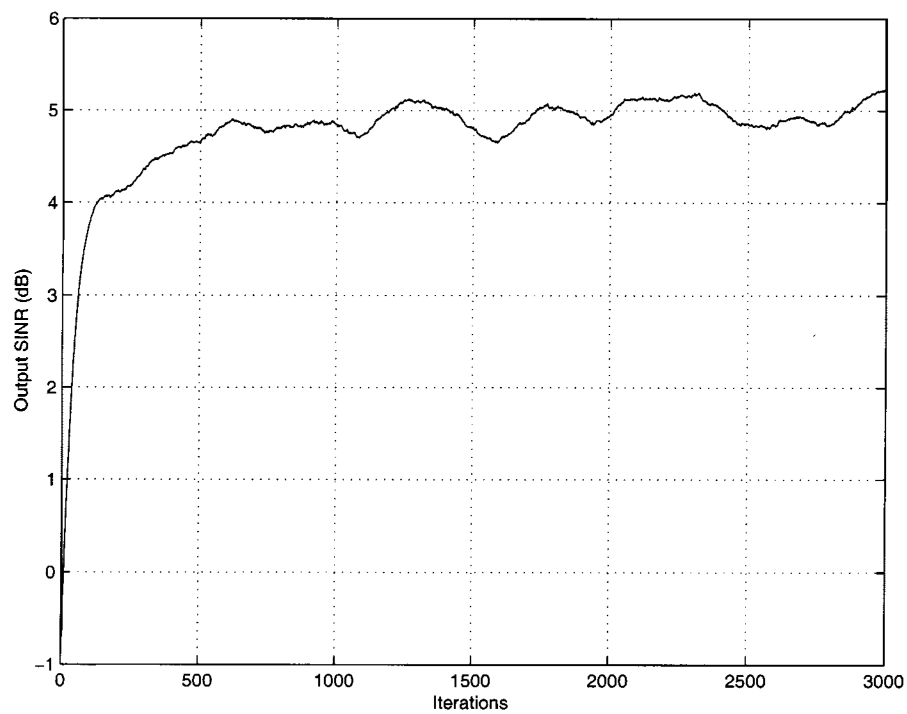


Fig. 9. Output SINR of the proposed receiver over a Rayleigh fading channel.

$\alpha_1(t)$, $\alpha_2(t)$ are two independent complex Gaussian processes with unit power. They were generated by passing i.i.d. white Gaussian inputs through a Doppler filter with cutoff frequency equal to the maximum Doppler shift (for details, see [17, ch. 4]). Delay $\tau_1$ of the second path was equal to 2.8 chip periods. A square pulse shaping function was employed. The symbol rate was set to 48.6 kb/s and the maximum Doppler shift to 100 Hz. Channel parameters for the asynchronous interfering users were also similarly simulated resulting in a time-varying interference structure. Other parameters were same as in the first experiment. The average output SINR of the proposed receiver for 1000 realizations versus number of iterations is shown in Fig. 9. It can be observed that the algorithm is capable of tracking the time-varying channel and suppressing the time-varying interference.

## X. Conclusions

The contribution of this paper lies in the derivation of blind adaptive solutions for interference cancellation in CDMA systems. LMS- and RLS-based constrained optimization methods are developed, which jointly optimize the receiver and constraint parameters. Furthermore, it is shown that the proposed methods enjoy global convergence. Experimental results indicate that their performance is better than existing methods and close to that of the trained MMSE receiver.

## Appendix I
## Proof of Global Convergence

Our cost function $\mathcal{J}_1$ in (12) is a function of $\mathbf{f}$ and $\mathbf{g}$. At the stationary points, its first derivative with respect to these two vectors should be zero

$$\nabla_{\mathbf{f}^*}\mathcal{J}_1 = \mathbf{R}_y\mathbf{f} + \mathbf{C}\boldsymbol{\lambda} = 0 \tag{62}$$

$$\nabla_{\mathbf{g}^*}\mathcal{J}_1 = \rho\mathbf{g} - \boldsymbol{\lambda} = 0. \tag{63}$$

Solving (62) with respect to $\mathbf{f}$ and substituting in the constraint $\mathbf{C}^H\mathbf{f} = \mathbf{g}$, we obtain an expression for $\boldsymbol{\lambda}$

$$\boldsymbol{\lambda} = -\left(\mathbf{C}^H\mathbf{R}_y^{-1}\mathbf{C}\right)^{-1}\mathbf{g}. \tag{64}$$

Substituting (64) back into (62) and (63), we obtain

$$\mathbf{f} = \mathbf{R}_y^{-1}\mathbf{C}\left(\mathbf{C}^H\mathbf{R}_y^{-1}\mathbf{C}\right)^{-1}\mathbf{g} \tag{65}$$

and

$$\left(\mathbf{C}^H\mathbf{R}_y^{-1}\mathbf{C}\right)^{-1}\mathbf{g} = -\rho\mathbf{g} \tag{66}$$

respectively. Equation (66) indicates that $(-\rho, \mathbf{g})$ is an eigenpair of $(\mathbf{C}^H\mathbf{R}_y^{-1}\mathbf{C})^{-1}$. Because $(\mathbf{C}^H\mathbf{R}_y^{-1}\mathbf{C})^{-1}$ is positive definite, its eigenvalue $-\rho$ should be positive. Furthermore, using (65) and (66), the output variance under the constraint $\|\mathbf{g}\| = 1$ becomes $-\rho$. This explains why we choose the smaller solution for $\rho$ in (25) in order to maximize this variance at each iteration. In order to investigate which eigenvectors correspond to stable and which to unstable stationary points, we need to evaluate the Hessian of the cost function. From (63) and (64), we obtain

$$\nabla_{\mathbf{g}^*}\mathcal{J}_1 = \rho\mathbf{g} + \left(\mathbf{C}^H\mathbf{R}_y^{-1}\mathbf{C}\right)^{-1}\mathbf{g}$$

which yields

$$\begin{aligned}\nabla^2_{\mathbf{g}^*}\mathcal{J}_1 &= \rho\mathbf{I} + \left(\mathbf{C}^H\mathbf{R}_y^{-1}\mathbf{C}\right)^{-1}\\ &= \mathbf{V}\begin{bmatrix}\rho+\xi_1 & & 0\\ & \ddots & \\ 0 & & \rho+\xi_{q+1}\end{bmatrix}\mathbf{V}^H\end{aligned} \tag{67}$$

where $\mathbf{V}$ contains the eigenvectors of $(\mathbf{C}^H\mathbf{R}_y^{-1}\mathbf{C})^{-1}$ while $\xi_i (i = 1, \ldots, (q+1))$ represents its eigenvalues. Let us order $\xi_1 \le \xi_2 \le \cdots \le \xi_{q+1}$ without loss of generality. According to (66), at a stationary point we have $\rho = -\xi_i$ for some $i = 1, \ldots, (q+1)$. We may therefore distinguish the following three cases

1) If $\rho = -\xi_1 = -\xi_{\min}$, then $\rho + \xi_i \ge 0\ \forall i$ and therefore $\nabla^2_{\mathbf{g}^*}\mathcal{J}_1 \ge 0$ indicating a minimum point [cf. (67)].
2) If $\rho = -\xi_{q+1} = -\xi_{\max}$, then $\rho + \xi_i \le 0\ \forall i$ and therefore $\nabla^2_{\mathbf{g}^*}\mathcal{J}_1 \le 0$ indicating a maximum point.
3) If $\rho = -\xi_i$, $1 < i < q+1$, then $\rho + \xi_1 \le 0$ while $\rho + \xi_{q+1} \ge 0$ and hence $\nabla^2_{\mathbf{g}^*}\mathcal{J}_1$ is nondefinite, indicating a saddle point.

In conclusion, only the desired solution, corresponding to $\xi_{\max}$ is a stationary point, and therefore the algorithm enjoys global convergence. Once global convergence for $\mathbf{g}$ is guaranteed, $\mathbf{f}$ will also be globally convergent, because for each stationary point $\mathbf{g}$, vector $\mathbf{f}$ is uniquely defined as a function of $\mathbf{g}$ [see (65)].

## Appendix II
## Derivation of (55)

We start with the recursion for the tap vector in (18). Noticing that $\mathbf{C}^H\mathbf{f}_{\text{opt}} = \mathbf{g}_{\text{opt}}$ and replacing $\mathbf{R}_y$ in (18) by $\mathbf{y}_n\mathbf{y}_n^H$, the tap error vector at time $n+1$ becomes

$$\begin{aligned}\mathbf{e}_f[n+1] = &\ \mathbf{\Pi}_{\mathbf{C}}^{\perp}\left(\mathbf{I} - \mu_f\mathbf{y}_n\mathbf{y}_n^H\right)\mathbf{e}_f[n]\\ &+ \mathbf{C}\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{e}_g[n] - \mu_f\mathbf{\Pi}_{\mathbf{C}}^{\perp}\mathbf{y}_n\mathbf{y}_n^H\mathbf{f}_{\text{opt}}.\end{aligned} \tag{68}$$

Taking expectation of both sides of (68) and using the independence assumption (e.g., [5]), we can obtain

$$\begin{aligned}E\{\mathbf{e}_f[n+1]\} = &\ \mathbf{\Pi}_{\mathbf{C}}^{\perp}(\mathbf{I} - \mu_f\mathbf{R}_y)E\{\mathbf{e}_f[n]\}\\ &+ \mathbf{C}\left(\mathbf{C}^H\mathbf{C}\right)^{-1}E\{\mathbf{e}_g[n]\}\end{aligned} \tag{69}$$

where the term $\mathbf{\Pi}_{\mathbf{C}}^{\perp}\mathbf{R}_y\mathbf{f}_{\text{opt}}$ has been zeroed out. Equation (69) involves the mean constraint error vector $\mathbf{e}_g[n]$, hence a recursion for $\mathbf{e}_g[n]$ is also needed. Similarly from (21) and (20), the constraint error vector can be computed as

$$\begin{aligned}\mathbf{e}_g[n+1] = &\left[(1+\mu_g\rho_n)\mathbf{I} + \frac{\mu_g}{\mu_f}\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\right]\mathbf{e}_g[n]\\ &+ \left[\mu_g\rho_n\mathbf{I} + \frac{\mu_g}{\mu_f}\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\right]\mathbf{g}_{\text{opt}}\\ &+ \frac{\mu_g}{\mu_f}\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{C}^H\left(\mu_f\mathbf{y}_n\mathbf{y}_n^H - \mathbf{I}\right)\mathbf{f}_n.\end{aligned} \tag{70}$$

Then, the mean constraint error vector can be easily shown to be

$$\begin{aligned}E\{\mathbf{e}_g[n+1]\} = &\left[(1+\mu_g\bar{\rho}_n)\mathbf{I} + \frac{\mu_g}{\mu_f}\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\right]E\{\mathbf{e}_g[n]\}\\ &+ \mu_g(\xi_{\max} + \bar{\rho}_n)\mathbf{g}_{\text{opt}} + \frac{\mu_g}{\mu_f}\left(\mathbf{C}^H\mathbf{C}\right)^{-1}\mathbf{C}^H\\ &\cdot(\mu_f\mathbf{R}_y - \mathbf{I})E\{\mathbf{e}_f[n]\}\end{aligned} \tag{71}$$

where (54) is employed. By combining (69) and (71) together, we can obtain our matrix recursion (55).

## APPENDIX III
## DERIVATION OF (61)

Let $\mathbf{R}_e[n] = E(\mathbf{e}_f[n]\mathbf{e}_f[n]^H)$ and $\mathbf{R}_e = \lim_{n\to\infty}\mathbf{R}_e[n]$, then from (60) we have

$$\overline{\eta}_{ex} = \mathrm{tr}(\mathbf{R}_y\mathbf{R}_e) = \mathrm{vec}^H(\mathbf{R}_y)\mathrm{vec}(\mathbf{R}_e) \tag{72}$$

where the second equality comes from the property of trace. To evaluate $\overline{\eta}_{ex}$, it is sufficient to study $\mathbf{R}_e$. This quantity depends on the trajectory of the tap error vector as given by (68). For simplicity of analysis, we assume that $\mathbf{e}_g[n] \approx \mathbf{C}^H\mathbf{e}_f[n]$ which is true when adaptation approaches the steady state. Then (68) can be written as

$$\mathbf{e}_f[n+1] \approx \mathbf{P}[n]\mathbf{e}_f[n] - \mu_f\mathbf{\Pi}_{\mathbf{C}}^{\perp}\mathbf{y}_n\mathbf{y}_n^H\mathbf{f}_{\mathrm{opt}} \tag{73}$$

where

$$\mathbf{P}[n] = \mathbf{I} - \mu_f\mathbf{\Pi}_{\mathbf{C}}^{\perp}\mathbf{y}_n\mathbf{y}_n^H. \tag{74}$$

Therefore the following holds:

$$\begin{aligned}\mathbf{e}_f&[n+1]\mathbf{e}_f^H[n+1]\\ &\approx \mathbf{P}[n]\mathbf{e}_f[n]\mathbf{e}_f^H[n]\mathbf{P}^H[n] - \mu_f\mathbf{P}[n]\mathbf{e}_f[n]\mathbf{f}_{\mathrm{opt}}^H\mathbf{y}_n\mathbf{y}_n^H\mathbf{\Pi}_{\mathbf{C}}^{\perp}\\ &\quad - \mu_f\mathbf{\Pi}_{\mathbf{C}}^{\perp}\mathbf{y}_n\mathbf{y}_n^H\mathbf{f}_{\mathrm{opt}}\mathbf{e}_f^H[n]\mathbf{P}^H[n]\\ &\quad + \mu_f^2\mathbf{\Pi}_{\mathbf{C}}^{\perp}\mathbf{y}_n\mathbf{y}_n^H\mathbf{f}_{\mathrm{opt}}\mathbf{f}_{\mathrm{opt}}^H\mathbf{y}_n\mathbf{y}_n^H\mathbf{\Pi}_{\mathbf{C}}^{\perp}.\end{aligned} \tag{75}$$

Substituting (74) in (75) and taking expectation of both sides of it, it is not difficult to show that

$$\begin{aligned}\mathbf{R}_e[n+1] \approx \mathbf{R}_e[n] &- \mu_f\left(\mathbf{R}_e[n]\mathbf{R}_y\mathbf{\Pi}_{\mathbf{C}}^{\perp} + \mathbf{\Pi}_{\mathbf{C}}^{\perp}\mathbf{R}_y\mathbf{R}_e[n]\right)\\ &- \mu_f E\{\mathbf{P}[n]\mathbf{e}_f[n]\mathbf{f}_{\mathrm{opt}}^H\mathbf{y}_n\mathbf{y}_n^H\mathbf{\Pi}_{\mathbf{C}}^{\perp}\\ &\qquad + \mathbf{\Pi}_{\mathbf{C}}^{\perp}\mathbf{y}_n\mathbf{y}_n^H\mathbf{f}_{\mathrm{opt}}\mathbf{e}_f^H[n]\mathbf{P}^H[n]\}\\ &+ \mu_f^2 E\{\mathbf{\Pi}_{\mathbf{C}}^{\perp}\mathbf{y}_n\mathbf{y}_n^H(\mathbf{f}_{\mathrm{opt}}\mathbf{f}_{\mathrm{opt}}^H + \mathbf{R}_e[n])\mathbf{y}_n\mathbf{y}_n^H\mathbf{\Pi}_{\mathbf{C}}^{\perp}\}.\end{aligned} \tag{76}$$

Since as $n \to \infty$, we have

$$\lim_{n\to\infty}\mathbf{R}_e[n+1] = \lim_{n\to\infty}\mathbf{R}_e[n] = \mathbf{R}_e$$

and

$$\lim_{n\to\infty}E\{\mathbf{e}_f[n]\} = \mathbf{0}$$

then taking the limit on both sides of (76), we obtain

$$\begin{aligned}\mathbf{R}_e\mathbf{R}_y\mathbf{\Pi}_{\mathbf{C}}^{\perp} &+ \mathbf{\Pi}_{\mathbf{C}}^{\perp}\mathbf{R}_y\mathbf{R}_e\\ &\approx \mu_f E\left\{\mathbf{\Pi}_{\mathbf{C}}^{\perp}\mathbf{y}_n\mathbf{y}_n^H(\mathbf{f}_{\mathrm{opt}}\mathbf{f}_{\mathrm{opt}}^H + \mathbf{R}_e)\mathbf{y}_n\mathbf{y}_n^H\mathbf{\Pi}_{\mathbf{C}}^{\perp}\right\}\end{aligned} \tag{77}$$

from which $\mathrm{vec}(\mathbf{R}_e)$ can be obtained by the "vec" operation

$$\mathrm{vec}(\mathbf{R}_e) \approx \mu_f\mathbf{\Theta}^{-1}\mathbf{a}. \tag{78}$$

Substituting (78) in (72), (61) follows.

## REFERENCES


[1] S. E. Bensley and B. Aazhang, "Subspace-based channel estimation for code division multiple access communication systems," *IEEE Trans. Commun.*, vol. 44, pp. 1009–1020, Aug. 1996.

[2] P. Comon and G. H. Golub, "Tracking a few extreme singular values and vectors in signal processing," *Proc. IEEE*, vol. 78, pp. 1327–1343, Aug. 1990.

[3] O. L. Frost, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE*, vol. 60, pp. 926–935, Aug. 1972.

[4] R. de Gaudenzi, F. Giannetti, and M. Luise, "Design of a low-complexity adaptive interference-mitigating detector for DS.SS receivers in CDMA radio networks," *IEEE Trans. Inform. Theory*, vol. 46, pp. 125–134, Jan. 1998.

[5] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[6] H. C. Huang and S. Verdu, "Linearly differentially coherent multiuser detector for multipath channels," *Wireless Pers. Commun.*, vol. 6, no. 1–2, pp. 113–136, Jan. 1998.

[7] M. Honig, U. Madhow, and S. Verdu, "Blind adaptive multiuser detection," *IEEE Trans. Inform. Theory*, vol. 41, pp. 944–960, July 1995.

[8] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing: Concepts and Techniques*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[9] H. Liu and K. Li, "Low complexity receivers for CDMA communications over frequency-selective channels," in *Proc. SPAWC'97*, Paris, France, Apr. 16–18, 1997, pp. 265–268.

[10] 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.

[11] ——, "Near–Far resistance of multiuser detectors in asynchronous channels," *IEEE Trans. Commun.*, vol. 38, pp. 496–508, Apr. 1990.

[12] 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.

[13] N. B. Mandayam and S. Verdu, "Analysis of an approximate decorrelating detector," in *Proc. Annu. Allerton Conf. Communication, Control and Computing*, Monticello, IL, Oct. 1995, pp. 1043–1052.

[14] N. R. Mangalvedhe and J. H. Reed, "Blind CDMA interference rejection in multipath channels," in *Proc. 1997 IEEE 47th Vehicular Technology Conf. (VTC'97)*, vol. 1, Phoenix, AZ, May 1997, pp. 21–25.

[15] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, pp. 516–525, Feb. 1995.

[16] S. C. Park and J. F. Doherty, "Generalized projection algorithm for blind interference suppression in DS/CDMA communications," *IEEE Trans. Circuits Syst. II*, vol. 44, pp. 453–460, June 1997.

[17] T. S. Rappaport, *Wireless Communications*. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[18] J. Shen and Z. Ding, "Blind adaptive multiuser CDMA detection based on a linear projection constraint," in *Proc. SPAWC'97*, Paris, France, Apr. 16–18, 1997, pp. 261–264.

[19] M. K. Tsatsanis and G. B. Giannakis, "Optimal decorrelating receivers for DS-CDMA systems: A signal processing framework," *IEEE Trans. Signal Processing*, vol. 44, pp. 3044–3055, Dec. 1996.

[20] M. K. Tsatsanis, "Inverse filtering criteria for CDMA systems," *IEEE Trans. Signal Processing*, vol. 45, pp. 102–112, Jan. 1997.

[21] M. K. Tsatsanis and Z. Xu, "Performance analysis of minimum variance CDMA receivers," *IEEE Trans. Signal Processing*, vol. 46, pp. 3014–3022, Nov. 1998.

[22] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE Acoust., Speech, Signal Processing Mag.*, vol. 5, pp. 4–24, Apr. 1988.

[23] B. D. Van Veen, "Adaptive radar detection and estimation," in *Minimum Variance Beamforming*, S. Haykin and A. Steinhardt, Eds. New York: Wiley, 1992.

[24] S. Verdu, "Minimum probability of error for asynchronous Gaussian multiple-access channels," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 85–96, Jan. 1986.

[25] J. B. Schodorf and D. B. Williams, "Array processing techniques for multiuser detection," *IEEE Trans. Commun.*, vol. 45, pp. 1375–1378, Nov. 1997.

[26] ——, "A constrained optimization approach to multiuser detection," *IEEE Trans. Signal Processing*, vol. 45, pp. 258–262, Jan. 1997.

[27] X. Wang and H. V. Poor, "Blind multiuser detection: A subspace approach," *IEEE Trans. Inform. Theory*, vol. 44, pp. 677–690, Mar. 1998.

[28] ——, "Blind equalization and multiuser detection in dispersive CDMA channels," *IEEE Trans. Commun.*, vol. 46, pp. 91–103, Jan. 1998.

**Zhengyuan Xu** (S'97–M'99) received both the B.E. and M.E. degrees in electronic engineering from Tsinghua University, Beijing, China, in 1989 and 1991, respectively. He received the Ph.D. degree in electrical engineering from Stevens Institute of Technology, Hoboken, NJ, in May 1999.

From 1991 to 1996, he was a System Engineer and Department Manager at the Tsinghua Unisplendour Group Corporation, Tsinghua University. During the summer of 1999, he was a Research Associate at Stevens Institute of Technology, Hoboken, NJ, focusing on signal processing for wireless communications, especially for multiuser CDMA systems. Then, he joined the Department of Electrical Engineering, University of California, Riverside, as an Assistant Professor. His current research interests include digital communication theory, multiuser detection, signal processing for wireless communications, and system identification.

Dr. Xu received the Outstanding Student Award and Motorola Scholarship from Tsinghua University. He also received the Peskin Award from Stevens Institute of Technology for his outstanding performance .

**Michail K. Tsatsanis** (M'93) received his diploma degree in electrical engineering from the National Technical University of Athens, Athens, Greece, in 1987, and the M.Sc. and Ph.D. degrees in electrical engineering from the University of Virginia, Charlottesville, in 1990 and 1993, respectively.

He is currently a Chief Scientist at Voyan Technology, Santa Clara, CA. In the past, he held academic positions in electrical engineering as Research Associate (University of Virginia, 1994–1995) and Assistant and Associate Professor (Stevens Institute of Technology, 1995–1999 and 1999–2000, respectively). His general research interests lie in the areas of statistical signal and array processing, with applications to communications and networking. His current interests focus on signal processing techniques for wireless communications including blind equalization, multiuser detection, fading channel estimation and tracking, and signal processing methods for networking problems.

Dr. Tsatsanis is a member of the IEEE Technical Committee on SPCOM. He has served as a member of the Organizing Committee for the 1996 IEEE Signal Processing Workshop on SSAP and as the Technical Co-Chair of the Organizing Committee for the 1999 IEEE Workshop on Signal Processing Advances in Wireless Communications. He received the 1998 NSF CAREER Award and the 1999 SP Society Young Author Best Paper Award. He is an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and IEEE COMMUNICATIONS LETTERS.