[23] C. C. Chai, T. T. Tjhung, and L. C. Leck, "Combined power and rate adaptation for wireless cellular systems," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 6–13, Jan. 2005.

[24] P. Ligdas and N. Farvardin, "Optimizing the transmit power for slow fading channels," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 565–576, Mar. 2000.

[25] M. Nakagami, "The m-distribution—A general formula of intensity distribution of rapid fading," in *Statistical Methods in Radio Wave Propagation*. New York: Pergamon, 1960, pp. 3–36.

[26] W. Feller, *An Introduction to Probability Theory and Its Applications,* 2nd ed., vol. II. Hoboken, NJ: Wiley, 1971.

[27] S. W. Kim and A. J. Goldsmith, "Truncated power control in code-division multiple-access communications," *IEEE Trans. Veh. Technol.*, vol. 49, no. 3, pp. 965–972, May 2000.

[28] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flanner, *Numerical Recipes in C*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1997.

Blind MMSE-Constrained Multiuser Detection

Ping Liu, *Member, IEEE*, and Zhengyuan Xu, *Senior Member, IEEE*

***Abstract*—In this paper, blind multiuser detection is studied for a direct-sequence (DS) code-division multiple access (CDMA) system with unknown multipath channels. The receiver is designed to follow a minimum mean-square-error (MMSE) form parameterized by a channel-like constraint vector, which is then optimized by the conventional constant modulus algorithm in a much smaller dimensional space. It is analytically established that the constraint vector asymptotically converges to the channel vector of the user of interest. Correspondingly, the receiver converges to the desired user's MMSE receiver, ensuring detection of the desired signals. When the constraint vector is treated as a channel estimate, the performance of the channel estimator is analyzed using a perturbation technique. Simulation results show a satisfactory performance of the proposed approach, which is also observed to be superior to other approaches in many scenarios.**

***Index Terms*—Constant modulus algorithm (CMA), constrained optimization, multiuser detection.**

I. Introduction

In a direct-sequence (DS) code-division multiple access (CDMA) system, multiuser interference (MUI) is a typical obstacle to be obviated in the detection of input signals. Although the optimal detector is known as the maximum likelihood sequence estimator, linear detectors have received considerable attention due to their low complexity and acceptable performance. Among all linear detectors, blind solutions are particularly suitable for a bandwidth-constrained system. Second-order-statistics-based blind methods have been widely investigated and generally exhibit fast convergence and good channel-tracking capability for either flat fading channels [1], [2] or multipath fading channels [3]–[6].

Recently, there emerges significant interest in studying higher order statistics-based multiuser-detection techniques which may yield better detection performance. The constant modulus algorithm (CMA) [7], which has shown effectiveness in removing intersymbol interference (ISI) in a single-user system, has been revisited for CDMA systems. The CMA-based multiuser detection generally involves constrained optimization, where the detector is forced to satisfy some linear constraints such that signals from the user of interest are detected. Among existing approaches, the approaches in [8] and [9] consider only a flat fading channel. They suffer from a signature mismatch in the presence of multipath fading. Among those capable of multipath mitigation, the approach in [10] exhibits local minima and inability to optimally combine signal components from different paths. The approaches proposed in [11] and [12], although showing good performance, leave the convergence issue open and anticipate comprehensive system performance analysis. Our previously developed CMA-based method [13] requires some initial conditions to be satisfied to ensure convergence. The recently proposed two-stage approach [14] shows a satisfactory performance and allows low-complexity adaptive implementation. However, its performance depends on the front end of the minimum output energy (MOE) technique. The MOE approach has an inherent performance limitation in the presence of strong interference and noise, as shown and improved by [6].

In this paper, the receiver is forced to take a minimum mean-square-error (MMSE) form with an unknown channel-like constraint vector. It is shown that the adopted CMA cost function is indeed parameterized by the constraint vector. Therefore, minimization is performed in a much smaller space. Our analysis demonstrates that the constraint vector converges asymptotically to the desired channel vector and that the corresponding receiver converges to the desired user's MMSE receiver. Channel estimation error and the receiver's signal to interference and noise ratio (SINR) in the presence of small noise are further derived in closed forms. Algorithm implementation based on finite number of noisy observations is discussed to improve its robustness. Extension of the proposed approach to an asynchronous system is also considered. Simulation study shows that the proposed algorithm outperforms other approaches in many scenarios.

A DS/CDMA system model is described in Section II. CMA-based blind multiuser detectors and channel estimators are proposed in Section III and analyzed in Section IV. Implementation of the proposed approach and an extension to asynchronous situations are discussed in Section V. Extensive simulation examples are provided in Section VI, and conclusions are drawn in Section VII.

II. DS/CDMA System Model

Consider a CDMA system with J users, where the jth user's symbols are spread by a code sequence $c_j(k)$ of length P before being transmitted through a multipath channel characterized by coefficients $g_j(n)$. Assume that the maximum order of channels of all users is q. Then, at the receiver, the discrete-time chip-synchronized signal $y_j(n)$ due to user j is the convolution of its information bearing sequence $w_j(n)$ and composite channel $s_j(n)$ [4], [11]

$$y_j(n) = \sum_{l=-\infty}^{\infty} w_j(l) s_j(n - d_j - lP) \tag{1}$$

where d_j is its delay in chip periods and assumed to satisfy $0 \leq d_j < P$, and the signature waveform $s_j(n)$ is the convolution of the

Manuscript received June 21, 2005; revised May 12, 2006, December 1, 2006, May 23, 2007, and May 29, 2007. The work of P. Liu was supported by the Arkansas Center for Energy, Natural Resources, and Environmental Studies. The work of Z. Xu was supported by the National Science Foundation under Grant NSF-CCR 0207931. The review of this paper was coordinated by Prof. M. Juntti.

P. Liu is with the Department of Electrical Engineering, Arkansas Tech University, Russellville, AR 72801 USA (e-mail: pliu@atu.edu).

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

Digital Object Identifier 10.1109/TVT.2007.905430

spreading codes with the propagation channel

$$s_j(n) = \sum_{m=0}^{q} g_j(m)c_j(n-m). \tag{2}$$

Let us define the jth user's code filtering matrix $\boldsymbol{C}_j$ with dimension of $L \times (q+1)$ and its channel vector $\boldsymbol{g}_j$ as

$$\boldsymbol{C}_j = \begin{bmatrix} c_j(0) & & 0 \\ \vdots & \ddots & c_j(0) \\ c_j(P-1) & & \vdots \\ 0 & \ddots & c_j(P-1) \\ \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} \tag{3}$$

and $\boldsymbol{g}_j = [g_j(0), \ldots, g_j(q)]^{\mathrm{T}}$. Let $\boldsymbol{J}$ denote a shift matrix of dimension $L \times L$ with all 1s in the subdiagonal immediately below the main diagonal. If the observation window spans ν symbol intervals, then the received data vector containing $L = \nu P$ chip samples from time nP to $nP + L - 1$ from user j can be concisely expressed as [4]

$$\boldsymbol{y}_j(n) = \sum_{i=-1-i_0}^{\nu} \boldsymbol{s}_j^{(i)} w_j(n+i) \tag{4}$$

where i_0 in the lower limit depends on d_j: $i_0 = 0$ if $0 \le d_j \le P - q$; $i_0 = 1$ if $P - q < d_j < P$. The signature vector $\boldsymbol{s}_j^{(i)}$ of dimension $L \times 1$, corresponding to the symbol $w_j(n+i)$, is defined as

$$\boldsymbol{s}_j^{(i)} = \boldsymbol{J}^{iP-d_j}\boldsymbol{C}_j\boldsymbol{g}_j.$$

If the receiver is assumed to be synchronized to the user of interest (user 1) and $w_1(n+\nu_0)$ is the desired signal at symbol delay ν_0, the received data vector that contains all users' contribution from chip time nP to $nP + L - 1$ can be expressed in the following equation after explicitly separating the desired term from interference and noise:

$$\begin{aligned} \boldsymbol{y}_n &= \boldsymbol{H}\boldsymbol{w}(n) + \boldsymbol{v}_n \\ &= \boldsymbol{h}_1 w_1(n+\nu_0) + \boldsymbol{H}_{\mathrm{int}}\boldsymbol{w}_{\mathrm{int}}(n) + \boldsymbol{v}_n \end{aligned} \tag{5}$$

where

$$\boldsymbol{H} = [\boldsymbol{h}_1, \boldsymbol{H}_{\mathrm{int}}], \quad \boldsymbol{h}_1 \triangleq \boldsymbol{s}_1^{(\nu_0)} = \mathcal{C}_1\boldsymbol{g}_1, \quad \mathcal{C}_1 \triangleq \boldsymbol{J}^{\nu_0 P}\boldsymbol{C}_1$$

$\boldsymbol{w}_{\mathrm{int}}(n)$ is an interference vector including all ISI and MUI, $\boldsymbol{H}_{\mathrm{int}}$ is the corresponding signature matrix, and $\boldsymbol{v}_n$ of dimension $L \times 1$ is the additive white Gaussian noise vector with zero mean and variance $\sigma_v^2\boldsymbol{I}$. If we assume that jth user's channel spans up to ν_j symbol durations, then the user will contribute $\nu + \nu_j + 1$ signatures in $\boldsymbol{H}$ in the ν symbol duration [14], except that the desired user, which is assumed synchronized, has only $\nu + \nu_j$ signatures contributed in $\boldsymbol{H}$. Consequently, the dimension of $\boldsymbol{H}$ can be approximated as $\nu P \times (\sum_{j=1}^{J}(\nu + \nu_j + 1) - 1)$. Correspondingly, $\boldsymbol{H}_{\mathrm{int}}$ will have the same rows as $\boldsymbol{H}$ but one column fewer than $\boldsymbol{H}$. The unique structures of $\boldsymbol{h}_1$ and $\mathcal{C}_1$, with dimensions $L \times 1$ and $L \times (q+1)$, respectively, will be exploited to derive a blind detector which is capable of combating multipath distortion and suppressing MUI.

Through this paper, we make the following assumptions.

Assumption 1: All users' information sequences are mutually independent and temporally independent identically distributed with unit power.

Assumption 2: Channel noise $\boldsymbol{v}_n$ is white Gaussian and independent of input signals.

Assumption 3: The matrix $[\mathcal{C}_1,\ \text{basis of span}\{\boldsymbol{H}_{\mathrm{int}}\}]$ has full column rank.

The first two assumptions are common in most multiuser-detection approaches. Assumption 3) requires only the linear independence between $\mathcal{C}_1$ and the basis of the interfering signatures without further rank requirement on $\boldsymbol{H}_{\mathrm{int}}$ itself. It is similar to the study in [14] and is much less restrictive than that in some existing methods (e.g., in [3]).

In the next section, we will present a CMA-based approach to recover the desired signal sequence $w_1(n+\nu_0)$ from the received noisy data $\boldsymbol{y}_n$.

III. MMSE-Constrained CMA-Based Receivers

If we focus on linear solutions, then the receiver design problem is equivalent to determining a vector $\boldsymbol{f}$ such that the receiver output

$$z_n = \boldsymbol{f}^{\mathrm{H}}\boldsymbol{y}_n$$

contains less interference and is close to the desired signal $w_1(n+\nu_0)$ in some sense. When channel parameters of the desired user are known, the MMSE receiver $\boldsymbol{f} = \boldsymbol{R}^{-1}\mathcal{C}_1\boldsymbol{g}_1$ has been shown to be the best linear receiver in the sense of minimizing the MSE, where $\boldsymbol{R}$ of dimension $L \times L$ is the autocorrelation of $\boldsymbol{y}_n$. Motivated by this fact, we propose to impose an MMSE-form constraint on the receiver and minimize Godard cost function with respect to (w.r.t.) the receiver when a channel is unknown [16]

$$\min_{\boldsymbol{f}} \mathcal{J}_{\mathrm{CMA}} = E\left\{\left(|z_n|^2 - 1\right)^2\right\}$$
$$\text{subject to } \boldsymbol{f} = \boldsymbol{R}^{-1}\mathcal{C}_1\boldsymbol{g}. \tag{6}$$

The CMA cost function is chosen because of its excellent performance in removing ISI [15]. The constraint on $\boldsymbol{f}$ aims at removing user and delay ambiguity and forcing the output of the receiver to converge to the desired user's signal at a particular offset ν_0. $\boldsymbol{g}$ is a parameterized channel-like vector in the constraint. Applying the constraint of $\boldsymbol{f}$ directly to the output yields $z_n = \boldsymbol{g}^{\mathrm{H}}\mathcal{C}_1^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{y}_n$. This way, the constrained optimization problem w.r.t. $\boldsymbol{f}$ described by (6) is transformed to the following unconstrained one w.r.t. $\boldsymbol{g}$:

$$\min_{\boldsymbol{g}} \mathcal{J}(\boldsymbol{g}) = E\left\{\left(\boldsymbol{g}^{\mathrm{H}}\mathcal{C}_1^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{y}_n\boldsymbol{y}_n^{\mathrm{H}}\boldsymbol{R}^{-1}\mathcal{C}_1\boldsymbol{g} - 1\right)^2\right\}. \tag{7}$$

Once the optimal $\boldsymbol{g}$ is obtained from (7), the receiver $\boldsymbol{f}$ can be constructed from the constraint as (6). Next, we will analyze the proposed algorithm.

IV. Performance Analysis

In this section, asymptotic convergence property is analyzed first. The channel estimation error and the receiver's output SINR are then investigated under a small noise assumption. Both real and complex systems are considered, where BPSK and real channels are assumed for the real system and QPSK and complex channels are assumed for the complex system. Without loss of generality, $\|\boldsymbol{g}_1\| = 1$ is assumed throughout the analysis.

Applying the theoretical CMA cost function [15] and the proposed constraint $\boldsymbol{f} = \boldsymbol{R}^{-1}\mathcal{C}_1\boldsymbol{g}$, we first obtain the theoretical unconstrained optimization w.r.t. $\boldsymbol{g}$ for a real system as

$$\min_{\boldsymbol{g}} \mathcal{J}_1(\boldsymbol{g}) = -2\sum_{i=1}^{M}\left(\boldsymbol{g}^{\mathrm{T}}\mathcal{C}_1^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{h}_i\right)^4 + 3(\boldsymbol{g}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{g})^2 - 2(\boldsymbol{g}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{g}) + 1 \tag{8}$$

and that for a complex system as

$$\min_{\boldsymbol{g}} \mathcal{J}_1(\boldsymbol{g}) = -\sum_{i=1}^{M} \left|\boldsymbol{g}^{\mathrm{H}}\mathcal{C}_1^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{h}_i\right|^4 + 2(\boldsymbol{g}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{g})^2 - 2(\boldsymbol{g}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{g}) + 1 \tag{9}$$

where the $(q+1)\times(q+1)$ matrix $\boldsymbol{A} \triangleq \mathcal{C}_1^{\mathrm{H}}\boldsymbol{R}^{-1}\mathcal{C}_1$, M is the total number of signatures (columns $\boldsymbol{h}_i$) contained in $\boldsymbol{H}$, and $\boldsymbol{R} = \boldsymbol{H}\boldsymbol{H}^{\mathrm{H}} + \sigma_v^2\boldsymbol{I}$. If $\boldsymbol{H}$ is decomposed into signal subspace $\boldsymbol{U}_s$ and noise subspace $\boldsymbol{U}_n$, the data covariance matrix $\boldsymbol{R}$ can further be expressed as

$$\boldsymbol{R} = \boldsymbol{U}_s\left(\boldsymbol{\Lambda}_s + \sigma_v^2\boldsymbol{I}\right)\boldsymbol{U}_s^{\mathrm{H}} + \sigma_v^2\boldsymbol{U}_n\boldsymbol{U}_n^{\mathrm{H}} \tag{10}$$

where $\boldsymbol{\Lambda}_s = \mathrm{diag}\{|\lambda_1|^2, \ldots, |\lambda_\zeta|^2\}$, and λ_i is the ith singular value of $\boldsymbol{H}$. Next, let us present and prove a lemma which will be used later.

Lemma: Under Assumptions 1)–3) and the assumption that $\|\boldsymbol{g}_1\| = 1$, matrix $\lim_{\sigma_v^2\to 0}\boldsymbol{A}$ can be decomposed by SVD as $\lim_{\sigma_v^2\to 0}\boldsymbol{A} = \sum_{i=1}^{q+1}\beta_i\boldsymbol{\xi}_i\boldsymbol{\xi}_i^{\mathrm{H}}$ with orthonormal basis $\boldsymbol{\xi}_i$ and singular value β_i, where $\boldsymbol{\xi}_1 = \boldsymbol{g}_1$, $\beta_1 = 1$, and $\beta_i \to +\infty$ for $i > 1$.

Proof: According to (10), $\lim_{\sigma_v^2\to 0}\sigma_v^2\boldsymbol{A} = \mathcal{C}_1^{\mathrm{H}}\boldsymbol{U}_n\boldsymbol{U}_n^{\mathrm{H}}\mathcal{C}_1$. Clearly, $\boldsymbol{g}_1$ is a null vector of $\lim_{\sigma_v^2\to 0}\sigma_v^2\boldsymbol{A}$. Therefore, $\boldsymbol{g}_1$ is also a singular vector of $\lim_{\sigma_v^2\to 0}\boldsymbol{A}$ associated with singular value $\beta_1 = \lim_{\sigma_v^2\to 0}\boldsymbol{g}_1^{\mathrm{H}}\boldsymbol{A}\boldsymbol{g}_1$. Noticing that $\lim_{\sigma_v^2\to 0}\boldsymbol{g}_1^{\mathrm{H}}\mathcal{C}_1^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{h}_i = \delta(i-1)$ with $\delta(\cdot)$ denoting the Kronecker delta function, we have $\beta_1 = 1$. For $i \geq 2$, since $\boldsymbol{\xi}_i$ is linearly independent of $\boldsymbol{g}_1$, it is straightforward to find that $\mathcal{C}_1\boldsymbol{\xi}_i$ is not in the signal space of $\boldsymbol{R}$ by using contradiction. As a result, the corresponding singular value $\beta_i = \lim_{\sigma_v^2\to 0}(\mathcal{C}_1\boldsymbol{\xi}_i)^{\mathrm{H}}\boldsymbol{R}^{-1}\mathcal{C}_1\boldsymbol{\xi}_i$ will be dominated by the projection of $\mathcal{C}_1\boldsymbol{\xi}_i$ on the noise subspace of $\boldsymbol{R}$, i.e., $\lim_{\sigma_v^2\to 0}(\mathcal{C}_1\boldsymbol{\xi}_i)^H((1/\sigma_v^2)\boldsymbol{U}_n\boldsymbol{U}_n^{\mathrm{H}})\mathcal{C}_1\boldsymbol{\xi}_i$ which is $+\infty$. Hence, $\beta_i \to +\infty$ for $i \geq 2$. □

A. Convergence Analysis

A real system is first considered. For notational convenience, $\lim_{\sigma_v^2\to 0}\boldsymbol{R}^{-1}$ is simply denoted by $\boldsymbol{R}^{-1}$ in this section.

1) Real Case: The derivative of (8) is given by

$$\nabla_{\boldsymbol{g}}\mathcal{J}_1 = -8\sum_{i=1}^{M}\left(\boldsymbol{g}^{\mathrm{T}}\mathcal{C}_1^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{h}_i\right)^3\mathcal{C}_1^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{h}_i + (12\boldsymbol{g}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{g} - 4)\boldsymbol{A}\boldsymbol{g}. \tag{11}$$

Premultiplying (11) by $\boldsymbol{g}_1^{\mathrm{T}}$ and forcing the result to be zero yield

$$-8\sum_{i=1}^{M}\left(\boldsymbol{g}^{\mathrm{T}}\mathcal{C}_1^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{h}_i\right)^3\boldsymbol{g}_1^{\mathrm{T}}\mathcal{C}_1^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{h}_i + (12\boldsymbol{g}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{g} - 4)\left(\boldsymbol{g}_1^{\mathrm{T}}\boldsymbol{A}\boldsymbol{g}\right) = 0. \tag{12}$$

Clearly, each equilibrium of the unconstrained approach should satisfy (12). However, (12) may introduce extra stationary points, which need to be removed if they cannot make (11) be zero. Applying Lemma, one can immediately find that $\boldsymbol{g} \in \mathrm{span}\{\boldsymbol{g}_1^{\perp}\}$ is one solution to (12) and that other solutions satisfy

$$-8\left(\boldsymbol{g}_1^{\mathrm{T}}\boldsymbol{g}\right)^2 + 12\boldsymbol{g}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{g} - 4 = 0. \tag{13}$$

Decomposing $\boldsymbol{g}$ into the orthonormal basis as $\boldsymbol{g} = \gamma(\mu_1\boldsymbol{g}_1 + \sum_{i=2}^{q+1}\mu_i\boldsymbol{\xi}_i)$, where $\gamma = \|\boldsymbol{g}\|$ and $\sum_{i=1}^{q+1}|\mu_i|^2 = 1$, other solutions can be derived to be $\boldsymbol{g} = \pm\boldsymbol{g}_1$, which can be verified to be the solutions to (11).

To exclude $\boldsymbol{g} \in \mathrm{span}\{\boldsymbol{g}_1^{\perp}\}$ from the equilibria of (11), we premultiply (11) by $\boldsymbol{g}^{\mathrm{T}}$ and check the corresponding result at the point $\boldsymbol{g} \in \mathrm{span}\{\boldsymbol{g}_1^{\perp}\}$. If $\boldsymbol{g} \in \mathrm{span}\{\boldsymbol{g}_1^{\perp}\}$, then $\boldsymbol{g}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{g} = +\infty$. Noticing that $\lim_{\sigma_v^2\to 0}\boldsymbol{R} = \sum_{i=1}^{M}\boldsymbol{h}_i\boldsymbol{h}_i^{\mathrm{T}}$, one can verify that

$$\sum_{i=1}^{M}\left(\boldsymbol{g}^{\mathrm{T}}\mathcal{C}_1^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{h}_i\right)^4 \leq \left(\sum_{i=1}^{M}\left(\boldsymbol{g}^{\mathrm{T}}\mathcal{C}_1^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{h}_i\right)^2\right)^2 = (\boldsymbol{g}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{g})^2. \tag{14}$$

Then, $\boldsymbol{g}^{\mathrm{T}}\nabla_{\boldsymbol{g}}\mathcal{J}_1 \geq 4\boldsymbol{g}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{g}(\boldsymbol{g}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{g} - 1) > 0$. As a result, $\boldsymbol{g} \in \mathrm{span}\{\boldsymbol{g}_1^{\perp}\}$ could not be an equilibrium, and $\boldsymbol{g} = \pm\boldsymbol{g}_1$ are the only ones. The Hessian matrix is further given by

$$\frac{\partial^2\mathcal{J}_1}{\partial\boldsymbol{g}^2} = -24\sum_{i=1}^{M}(\boldsymbol{g}^{\mathrm{T}}\chi_i)^2\chi_i\chi_i^{\mathrm{T}} + (12\boldsymbol{g}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{g} - 4)\boldsymbol{A} + 24(\boldsymbol{A}\boldsymbol{g})(\boldsymbol{A}\boldsymbol{g})^{\mathrm{T}} \tag{15}$$

where $\chi_i \triangleq \mathcal{C}_1^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{h}_i$. By using the Lemma, one can verify that $(\partial^2\mathcal{J}_1/\partial\boldsymbol{g}^2) = 8\boldsymbol{A}$ at $\boldsymbol{g} = \pm\boldsymbol{g}_1$, which is positive definite. Combining the above analysis, we conclude that $\boldsymbol{g}$ converges to $\boldsymbol{g}_1$ within a phase ambiguity.

2) Complex Case: Consider the theoretical cost function of (9), whose derivative is given by

$$\nabla_{\boldsymbol{g}}\mathcal{J}_1 = -2\sum_{i=1}^{M}\left[\left|\boldsymbol{g}^{\mathrm{H}}\mathcal{C}_1^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{h}_i\right|^2\left(\boldsymbol{h}_i^{\mathrm{H}}\boldsymbol{R}^{-1}\mathcal{C}_1\boldsymbol{g}\right)\mathcal{C}_1^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{h}_i\right] + (4\boldsymbol{g}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{g} - 2)\boldsymbol{A}\boldsymbol{g}. \tag{16}$$

Following similar steps as in a real case, one can verify that stationary points for the complex case satisfy

$$-2\boldsymbol{g}^{\mathrm{H}}\boldsymbol{g}_1\boldsymbol{g}_1^{\mathrm{H}}\boldsymbol{g} + 4\boldsymbol{g}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{g} - 2 = 0. \tag{17}$$

Similarly expressing $\boldsymbol{g}$ in terms of the basis of $\boldsymbol{A}$, the stationary points for the complex case are found to be $\boldsymbol{g} = \boldsymbol{g}_1 e^{j\theta}$, where θ indicates a phase ambiguity. Therefore, $\boldsymbol{g}$ converges to $\boldsymbol{g}_1$ within a phase ambiguity.

To summarize, we conclude the convergence property as follows.

Proposition 1: Under the Assumptions 1)–3) and when $\sigma_v^2 \to 0$, the optimal $\boldsymbol{g}$ obtained by the optimization (7) converges to the multipath channel $\boldsymbol{g}_1$ within a phase ambiguity.

B. Channel Estimation Error

Since noise causes solutions to deviate from the aforementioned desired points, we analyze its effect on channel estimation. Consider only the neighborhood of the desired point $\boldsymbol{g} = \boldsymbol{g}_1$ for small noise power ($\sigma_v^2 \ll 1$). Perturbed solutions around other desired points can similarly be analyzed. Consider the normalized channel estimation error, i.e., $(\boldsymbol{g}/\|\boldsymbol{g}\|) - (\boldsymbol{g}_1/\|\boldsymbol{g}_1\|) = \Delta\boldsymbol{g}_1 + \Delta\boldsymbol{g}_1^{\perp}$, where $\Delta\boldsymbol{g}_1$ represents the in-space estimation error, and $\Delta\boldsymbol{g}_1^{\perp}$ is the error in the orthogonal space. Correspondingly, $\boldsymbol{g} = \|\boldsymbol{g}\|((\boldsymbol{g}_1/\|\boldsymbol{g}_1\|) + \Delta\boldsymbol{g}_1 + \Delta\boldsymbol{g}_1^{\perp})$. Since $\Delta\boldsymbol{g}_1 \ll \Delta\boldsymbol{g}_1^{\perp}$ under a small perturbation (see [17]), only $\Delta\boldsymbol{g}_1^{\perp}$ is considered in the analysis. The result is provided next, and its proof is detailed in the Appendix.

Proposition 2: For small σ_v^2, the channel estimation error is approximated by

$$\Delta\boldsymbol{g}_1^{\perp} \approx -\sigma_v^8\boldsymbol{A}_0^{\dagger}\mathcal{C}_1^{\mathrm{H}}\boldsymbol{U}_s\boldsymbol{\Lambda}_s^{-1}\boldsymbol{U}_s^{\mathrm{H}} \times \sum_{i=2}^{M}\left|\boldsymbol{h}_1^{\mathrm{H}}\boldsymbol{U}_s\boldsymbol{\Lambda}_s^{-2}\boldsymbol{U}_s^{\mathrm{H}}\boldsymbol{h}_i\right|^2\left(\boldsymbol{h}_i^{\mathrm{H}}\boldsymbol{U}_s\boldsymbol{\Lambda}_s^{-2}\boldsymbol{U}_s^{\mathrm{H}}\boldsymbol{h}_1\right)\boldsymbol{h}_i.$$

C. SINR Analysis

The asymptotic SINR ratio will be studied here, with SINR defined as [4]

$$\text{SINR} = \frac{|\boldsymbol{f}^{\rm H}\boldsymbol{h}_1|^2}{\boldsymbol{f}^{\rm H}\boldsymbol{R}\boldsymbol{f} - |\boldsymbol{f}^{\rm H}\boldsymbol{h}_1|^2}. \quad (18)$$

Let us consider the real case. By using $\boldsymbol{f}_{\rm new} = \boldsymbol{R}^{-1}\mathcal{C}_1\boldsymbol{g}$ into (18) and replacing $\boldsymbol{g}$ by $\boldsymbol{g}_1 + \Delta\boldsymbol{g}_1^{\perp}$, we have

$$\begin{aligned}\text{SINR}_{\rm new} &= \frac{\boldsymbol{g}_1^{\rm T}\boldsymbol{A}\boldsymbol{g}_1 + O(\sigma_v^4)}{1 - \boldsymbol{g}_1^{\rm T}\boldsymbol{A}\boldsymbol{g}_1 + O(\sigma_v^4)} \\ &= \frac{1 + O(\sigma_v^2)}{\sigma_v^2\boldsymbol{g}_1^{\rm T}\boldsymbol{A}_2\boldsymbol{g}_1 + O(\sigma_v^4)}. \end{aligned} \quad (19)$$

Since $\text{SINR}_{\rm MMSE} = (1 + O(\sigma_v^2)/\sigma_v^2\boldsymbol{g}_1^{\rm T}\boldsymbol{A}_2\boldsymbol{g}_1 + O(\sigma_v^4))$ by the study in [4], the ratio follows as

$$\frac{\text{SINR}_{\rm new}}{\text{SINR}_{\rm MMSE}} \to 1 \quad \text{as} \quad \sigma_v^2 \to 0. \quad (20)$$

It is trivial to extend the result to the complex case.

V. Implementation and Extension

We briefly discuss some issues in implementation and extend the proposed approach to asynchronous situations.

Following the study in [11], the stochastic gradient-based batch iterative algorithm is applied to find the optimal solution, where both cost function and gradient can be approximated based on a data record of N symbols. Due to the effect of nontrivial noise power and finite data length, the surface of the data-based cost function is generally changed from its ideal one, and thus, the performance of the proposed method might vary with initializations. To obtain the best performance, we propose to implement the proposed algorithm under different initial constraint vector $\boldsymbol{g}$, each having one nonzero element corresponding to a particular path delay. Then, we take the constraint and the corresponding receiver yielding the minimum CMA cost as the optimal solution. The proposed initialization strategy will show a complexity of about $q+1$ times of the complexity in [11] due to the $q+1$ independent optimization processes resulting from $q+1$ initializations. Under this setup, the simulation results show that the proposed scheme outperforms other methods in many situations. On the other hand, similar to the study in [11], the unconstrained CMA approach discussed therein can also be applied here by setting the initial receiver $\boldsymbol{f}$ to be the one obtained from the aforementioned algorithm and then by directly updating $\boldsymbol{f}$ following the CMA cost function without any constraint on $\boldsymbol{f}$. Since the unconstrained approach can closely follow the conventional CMA algorithm, we will skip our discussion on it.

In the presence of channel fading, an adaptive algorithm can be implemented to achieve better tracking performance. The gradient is approximated by $(\partial\mathcal{J}/\partial\boldsymbol{g}) \approx 2\mathcal{C}_1^{\rm H}\boldsymbol{R}^{-1}(|z_n|^2 - 1)z_n^*\boldsymbol{y}_n$, where $\boldsymbol{R}^{-1}$ can be updated by the recursive least squares method based on the received data [18]. To improve robustness, we bound the norm of the constraint vector between a properly selected lower bound and an upper bound [5].

Similar to the study in [11], the proposed approach can be extended to an asynchronous situation by defining a new channel vector to absorb the unknown chip delay d_1 as $\tilde{\boldsymbol{g}}_1 = [0, \ldots, \boldsymbol{g}_1^{\rm H}, 0, \ldots]^{\rm H}$ with $(\nu_1+1)P - d_1 - q - 1$ leading zeros and d_1 trailing zeros, where $\nu_1 = \lceil (q + 1/P) \rceil$ represents the multipath delay in the symbol period. Then, the corresponding code matrix changes to $\tilde{\mathcal{C}}_1 = \boldsymbol{J}^{(\nu_0-1)P}\tilde{\boldsymbol{C}}_1$, and $\tilde{\boldsymbol{C}}_1$ of dimension $L \times (\nu_1+1)P$ is a code filtering matrix similarly constructed as $\boldsymbol{C}_1$. We also modify the MMSE-form constraint as $\boldsymbol{f} = \boldsymbol{R}^{-1}\tilde{\mathcal{C}}_1\tilde{\boldsymbol{g}}$ and the algorithm in (7) by replacing $\mathcal{C}_1$ with $\tilde{\mathcal{C}}_1$. As a result, the solution $\tilde{\boldsymbol{g}}$ to the modified method will be an approximate of $\tilde{\boldsymbol{g}}_1$ instead of $\boldsymbol{g}_1$. The extension will be examined by simulation and shown to achieve a satisfactory performance.

VI. Simulations

We first verify the analytical results concerning asymptotic convergence, channel normalized MSE (NMSE), and output SINR using the theoretical data covariance matrix. Then, we carry out experiments based on finite data samples (termed as data-based) to compare the proposed algorithm with the constrained method [11] (termed as TL), the MOE method [4], and the two-stage CMA approach [14] (termed as TS-CMA) in terms of bit error rate (BER). All approaches are implemented in a batch iterative way. Finally, we compare the tracking performance of the proposed approach with the TS-CMA and with the coherent least square (CLS) receiver [18], which is an MMSE receiver with data covariance matrix estimated from the received data. In the last experiment, all approaches are implemented adaptively.

In our simulation setup, each user's spreading codes, delay, and multipath channel parameters are randomly generated in each realization, except that the desired user is always synchronized in synchronous situations. BPSK signal sources and real Gaussian channels are assumed for a real system, whereas QPSK signal sources and complex Gaussian channels are assumed for a complex system. For a complex channel, both real and imaginary parts are assumed independent Gaussian random variables. Each simulation result is the average over 100 Monte Carlo realizations. The parameters $\nu = 5$ and $\nu_0 = 3$ are set for all simulations. In the data-based implementation, the threshold ε to terminate the iteration process [11] is set to be 0.02 for the proposed TL and TS-CMA approaches. In the adaptive implementation, $\boldsymbol{g}$ is initialized to be $[1, 0]^{\rm T}$, and the smoothing factor for updating $\boldsymbol{R}^{-1}$ is set to 0.998 for all approaches. The variable step size is taken following the idea in [14] and [20].

A. Theoretical Results

We adopt an ideal cost function which is equivalent to setting $N = \infty$. Both real and complex systems are considered for the proposed approach, and only the real system is considered for the MOE method for simplicity. Spreading factor is set to be $P = 8$, and three equal-power users are assumed. Each user's channel has four paths, each of which is Gaussian-distributed of unit variance. The first path has zero delay, and the remaining paths have delays uniformly distributed from 1 to $P-1$. The system is assumed synchronous.

The asymptotic convergence is studied in Fig. 1, where fixed initialization vectors $[1, 0, \ldots, 0]$ and $[1+j, 0, \ldots, 0]$ for the real and the complex systems are implemented, respectively. In Fig. 1(a), one can see that the channel NMSEs converge to an acceptable low level from 25 dB for the real system and from 10 dB for the complex system. The high NMSEs at other SNRs are caused by ill convergence of some realizations, whose frequency of occurrence is further shown in Fig. 1(b). Clearly, the frequency decreases as SNR increases. It drops to zero from 25 dB for the real system and from 10 dB for the complex system, indicating no failure realization among 100 realizations. Fig. 1 verifies the asymptotic convergence property of the proposed approach. It also reveals its sensitivity to large noise, entailing the use of the proposed initialization strategy in a practical situation where only finite noisy observations are available.

Fig. 2 then shows the performance of the approach in the presence of small noise under the proposed initialization scheme. Analytical

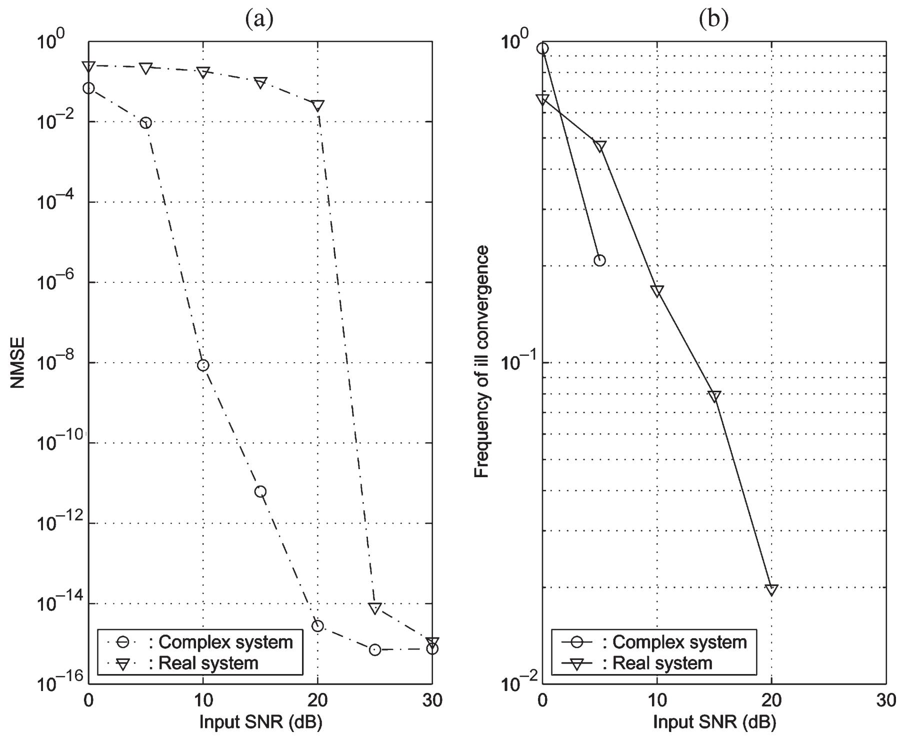

Fig. 1. Asymptotic convergence verification for $N = \infty$. (a) Channel MSE. (b) Frequency of ill convergence.

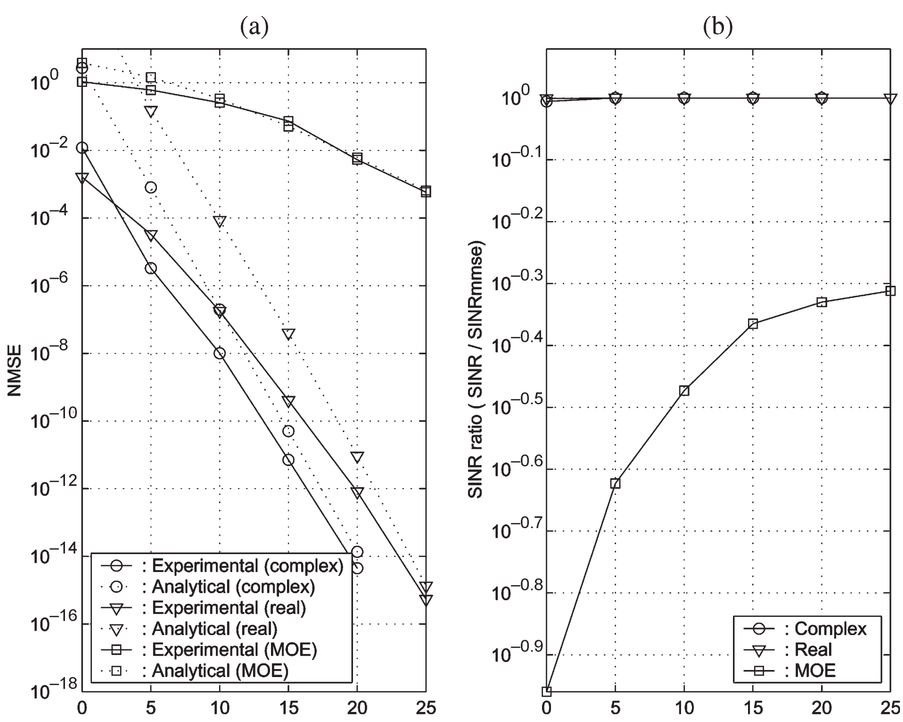

Fig. 2. Performance of the proposed approach for $N = \infty$. (a) Channel estimation error. (b) SINR ratio.

results are also plotted for verification. It can be observed that NMSEs exhibit low levels even at low SNRs, indicating nonexistence of ill convergence. The experimental results are all shown to converge to their analytical counterparts at high SNRs. The gap between analytical and experimental results at low SNR is caused by the approximation error in our analysis. It is also seen that the proposed scheme has much lower NMSEs than the MOE method for all SNRs. On the other hand, the SINR ratio $\mathrm{SINR}/\mathrm{SINR}_{\mathrm{MMSE}}$ converges to one at higher SNRs for the proposed approach. A satisfactory ratio of 0.99 can be achieved even at 0 dB. In contrast, the ratio of the MOE method converges to a constant that is much smaller than one.

From the aforementioned simulation examples, it is also found that the complex system shows better performance than the real system. For this reason and simplicity, we will focus on real systems in the sequel.

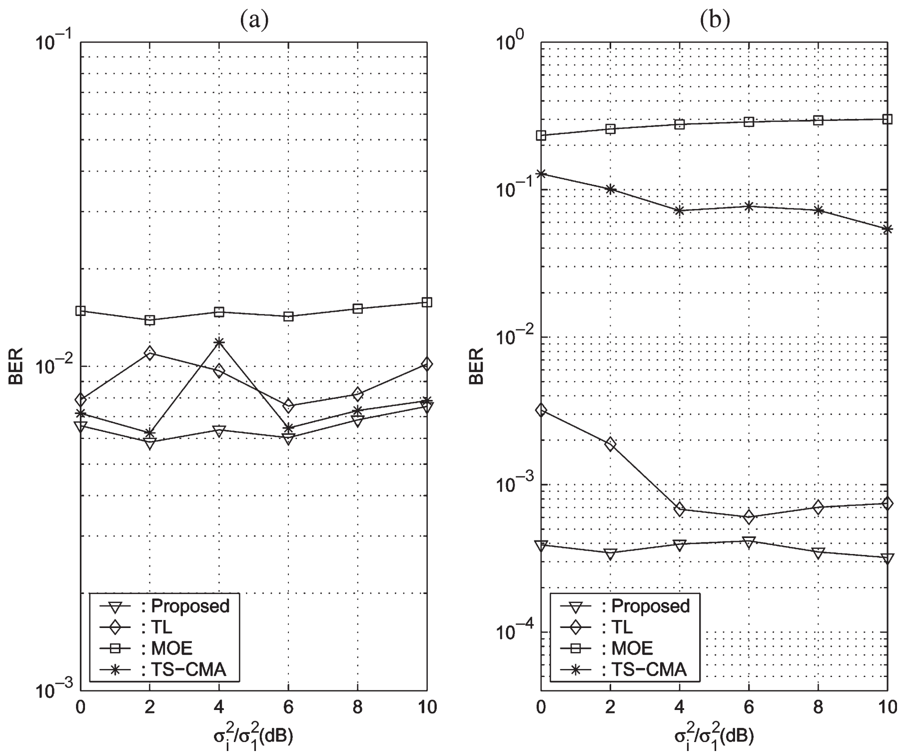

Fig. 3. Near–far effect with $P = 8$, $J = 3$, SNR $= 20$ dB, and $N = 400$. (a) Synchronous. (b) Asynchronous.

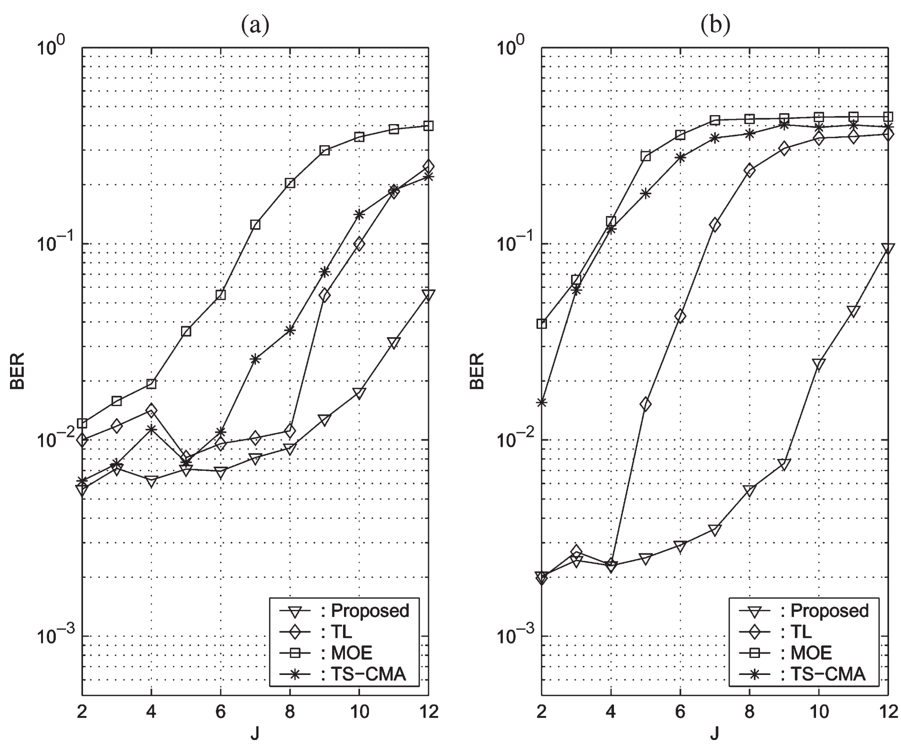

Fig. 4. BER versus number of active users with $P = 12$, SNR $= 20$ dB, and $N = 400$. (a) Synchronous. (b) Asynchronous.

B. Data-Based Results

Case 1—Batch implementation with $P = 8$, $N = 400$, $J = 3$, SNR $= 20$ dB, different interfering powers: The near–far effect is investigated. The power ratio of each equal-power interfering user i over the desired user σ_i^2/σ_1^2 varies from 0 to 10 dB. Fig. 3 shows the BER performance for both synchronous and asynchronous situations. The proposed approach shows higher near–far resistance.

Case 2—Batch implementation with $P = 12$, $N = 400$, SNR $=$ 20 dB, equal-power users, variable J: Fig. 4 shows the BERs of all approaches w.r.t. variable active users from 2 to 12. In the synchronous situation, the proposed approach shows the best performance over all loadings. In contrast, TL and TS-CMA methods show very good performance for a lightly loaded system $(J < 4)$ but degrade drastically otherwise. It is also interesting to

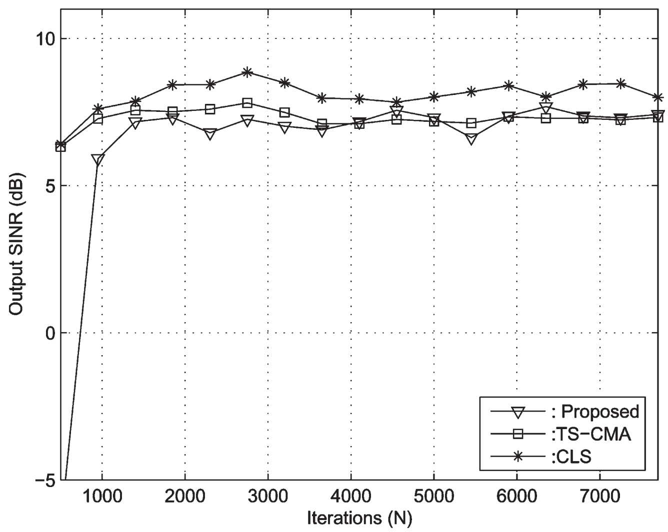

Fig. 5. Output SINR versus N for adaptive implementation with $P = 8$, $q = 3$, $J = 3$, SNR $= 15$ dB, and $f_dT = 2 \times 10^{-3}$.

note that the proposed method slowly degrades as system loading increases.

If we assume that all users' channels span ν_1 symbol durations, a rough analysis, which is based on the full column rank condition of $[\mathcal{C}_1\boldsymbol{H}_{\mathrm{int}}]$ for the synchronous case and $[\tilde{\boldsymbol{C}}_1\boldsymbol{H}_{\mathrm{int}}]$ for the asynchronous case, reveals that the proposed approach can support $(\nu P - q)/(\nu + \nu_1 + 1)$ and $(\nu P - q - 1 - P)/(\nu + \nu_1 + 1)$ users for the synchronous and asynchronous cases, which are seven and five users in our current setting, respectively. The results herein are quite close to the simulation results. On the other hand, the simulation results also show that when the matrix $[\mathcal{C}_1\boldsymbol{H}_{\mathrm{int}}]$ starts to lose rank due to the increased number of users in the system, the proposed approach slowly degrades. This may be credited to the proposed initialization strategy at the cost of higher complexity.

Case 3—Adaptive implementation with $P = 8$, $q = 3$, $J = 3$, SNR $= 15$ dB, equal-power users: We consider two-ray equal-power Raleigh fading channels [19] with $f_dT = 2 \times 10^{-3}$ [1], where f_d represents the maximum Doppler shift, and T denotes the symbol period. The proposed approach is implemented adaptively with relaxed lower and upper bounds, allowing approximately ± 7-dB power variation and yielding a much looser norm-bounded constraint. In implementing TS-CMA [14], the optimal MOE front-end filter is directly computed at each data sample for better performance instead of using low-complexity approximation as suggested therein. The output SINR, which can better demonstrate the adaptation of the approaches in a dynamic environment, is shown in Fig. 5. It is observed that all methods have similar SINR levels with CLS slightly outperforming the other two. The proposed approach shows a little bit slow startup due to its sensitivity to the estimation error for $\boldsymbol{R}^{-1}$, which suggests that the approach is best suitable for tracking very slowly fading channel.

VII. Conclusion

A constrained CMA-based criterion to blindly detect a desired user in a DS/CDMA system with unknown multipath distortion has been proposed. The receiver is forced to take a form of the MMSE receiver parameterized by a constraint vector. The constraint vector is proved to converge asymptotically to the desired channel vector. Accordingly, the receiver converges to the MMSE receiver. Simulation examples demonstrate the effectiveness of the proposed approach in several communication scenarios.

Appendix
Proof of Proposition 2

A. Real Case

Solving the optimal $\boldsymbol{g}$ from (11) and projecting $\boldsymbol{g}$ onto the orthogonal space of $\boldsymbol{g}_1$, one can obtain

$$\Delta\boldsymbol{g}_1^{\perp} = \frac{8}{\|\boldsymbol{g}\|(12\boldsymbol{g}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{g} - 4)}\left(\boldsymbol{I} - \frac{\boldsymbol{g}_1\boldsymbol{g}_1^{\mathrm{T}}}{\boldsymbol{g}_1^{\mathrm{T}}\boldsymbol{g}_1}\right)\boldsymbol{A}^{-1} \times \sum_{i>1}\left(\boldsymbol{g}^{\mathrm{T}}\mathcal{C}_1^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{h}_i\right)^3\mathcal{C}_1^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{h}_i \quad (21)$$

where $\boldsymbol{R}^{-1}$ can be expanded as [4]

$$\boldsymbol{R}^{-1} = \frac{1}{\sigma_v^2}\boldsymbol{U}_n\boldsymbol{U}_n^{\mathrm{T}} + \boldsymbol{U}_s\boldsymbol{\Lambda}_s^{-1}\boldsymbol{U}_s^{\mathrm{T}} - \sigma_v^2\boldsymbol{U}_s\boldsymbol{\Lambda}_s^{-2}\boldsymbol{U}_s^{\mathrm{T}} + \boldsymbol{O}\left(\sigma_v^4\right). \quad (22)$$

It is reasonable to assume that $\|\Delta\boldsymbol{g}_1^{\perp}\|$ is at least at the order of $O(\sigma_v^2)$, which then enables one to conclude from (21) that $\Delta\boldsymbol{g}_1^{\perp} \approx O(\sigma_v^6)$. Consequently, $\boldsymbol{g}^{\mathrm{T}}\mathcal{C}_1^{\mathrm{T}}\boldsymbol{R}^{-1}\boldsymbol{h}_i = -\sigma_v^2\|\boldsymbol{g}\|\boldsymbol{h}_1^{\mathrm{T}}\boldsymbol{U}_s\boldsymbol{\Lambda}_s^{-2}\boldsymbol{U}_s^{\mathrm{T}}\boldsymbol{h}_i + O(\sigma_v^4)$. Replacing it back to (21) and using (22), we have

$$\Delta\boldsymbol{g}_1^{\perp} = -\sigma_v^6\left(\boldsymbol{I} - \frac{\boldsymbol{g}_1\boldsymbol{g}_1^{\mathrm{T}}}{\boldsymbol{g}_1^{\mathrm{T}}\boldsymbol{g}_1}\right)\boldsymbol{A}^{-1}\mathcal{C}_1^{\mathrm{T}}\boldsymbol{U}_s\boldsymbol{\Lambda}_s^{-1}\boldsymbol{U}_s^{\mathrm{T}} \times \sum_{i>1}\left(\boldsymbol{h}_1^{\mathrm{T}}\boldsymbol{U}_s\boldsymbol{\Lambda}_s^{-2}\boldsymbol{U}_s^{\mathrm{T}}\boldsymbol{h}_i\right)^3\boldsymbol{h}_i + O\left(\sigma_v^8\right). \quad (23)$$

By using (22), one can verify that

$$\boldsymbol{A} = \frac{1}{\sigma_v^2}\boldsymbol{B} - \boldsymbol{g}_1\boldsymbol{g}_1^{\mathrm{T}}\boldsymbol{A}_1 + \boldsymbol{A}_1 - \sigma_v^2\boldsymbol{A}_2 + \boldsymbol{O}\left(\sigma_v^4\right) \quad (24)$$

where the four $(q+1) \times (q+1)$ matrices are defined as $\boldsymbol{B} \triangleq \boldsymbol{A}_0 + \sigma_v^2\boldsymbol{g}_1\boldsymbol{g}_1^{\mathrm{T}}\boldsymbol{A}_1\boldsymbol{A}_0 \triangleq \mathcal{C}_1^{\mathrm{T}}\boldsymbol{U}_n\boldsymbol{U}_n^{\mathrm{T}}\mathcal{C}_1$, $\boldsymbol{A}_1 \triangleq \mathcal{C}_1^{\mathrm{T}}\boldsymbol{U}_s\boldsymbol{\Lambda}_s^{-1}\boldsymbol{U}_s^{\mathrm{T}}\mathcal{C}_1$, and $\boldsymbol{A}_2 \triangleq \mathcal{C}_1^{\mathrm{T}}\boldsymbol{U}_s\boldsymbol{\Lambda}_s^{-2}\boldsymbol{U}_s^{\mathrm{T}}\mathcal{C}_1$. Noticing that

$$\boldsymbol{A}_0\boldsymbol{A}_0^{\dagger} + \boldsymbol{g}_1\boldsymbol{g}_1^{\mathrm{T}} = \boldsymbol{I}, \quad \boldsymbol{g}_1^{\mathrm{T}}\boldsymbol{A}_1\boldsymbol{g}_1 = 1, \quad A_0^{\dagger}\boldsymbol{g}_1 = 0 \quad (25)$$

where † denotes pseudoinverse, we have

$$\boldsymbol{B}^{-1} = \left(\boldsymbol{A}_0 + \sigma_v^2\boldsymbol{g}_1\boldsymbol{g}_1^{\mathrm{T}}\boldsymbol{A}_1\right)^{-1} = \frac{1}{\sigma_v^2}\boldsymbol{g}_1\boldsymbol{g}_1^{\mathrm{T}} + \boldsymbol{A}_0^{\dagger} - \boldsymbol{g}_1\boldsymbol{g}_1^{\mathrm{T}}\boldsymbol{A}_1\boldsymbol{A}_0^{\dagger}. \quad (26)$$

By using (25) and (26) repeatedly, we obtain

$$\boldsymbol{A} = \frac{1}{\sigma_v^2}\boldsymbol{B}\left[\boldsymbol{I} + \sigma_v^2\left(\boldsymbol{A}_0^{\dagger}\boldsymbol{A}_1 - \boldsymbol{g}_1\boldsymbol{g}_1^{\mathrm{T}}\boldsymbol{A}_1\boldsymbol{A}_0\dagger\boldsymbol{A}_1 - \boldsymbol{g}_1\boldsymbol{g}_1^{\mathrm{T}}\boldsymbol{A}_2\right) + \boldsymbol{O}\left(\sigma_v^4\right)\right]. \quad (27)$$

By applying Taylor series expansion, we have

$$\boldsymbol{A}^{-1} = \boldsymbol{g}_1\boldsymbol{g}_1^{\mathrm{T}} + \sigma_v^2\left[\boldsymbol{A}_0^{\dagger} + \left(\boldsymbol{g}_1^{\mathrm{T}}\boldsymbol{A}_1\boldsymbol{A}_0^{\dagger}\boldsymbol{A}_1\boldsymbol{g}_1 + \boldsymbol{g}_1^{\mathrm{T}}\boldsymbol{A}_2\boldsymbol{g}_1\right)\boldsymbol{g}_1\boldsymbol{g}_1^{\mathrm{T}} - \boldsymbol{g}_1\boldsymbol{g}_1^{\mathrm{T}}\boldsymbol{A}_1\boldsymbol{A}_0^{\dagger} - \boldsymbol{A}_0^{\dagger}\boldsymbol{A}_1\boldsymbol{g}_1\boldsymbol{g}_1^{\mathrm{T}}\right] + \boldsymbol{O}\left(\sigma_v^4\right) \quad (28)$$

which, together with (25), suggests that

$$\left(\boldsymbol{I} - \frac{\boldsymbol{g}_1\boldsymbol{g}_1^{\mathrm{T}}}{\boldsymbol{g}_1^{\mathrm{T}}\boldsymbol{g}_1}\right)\boldsymbol{A}^{-1} = \sigma_v^2\left(\boldsymbol{A}_0^{\dagger} - \boldsymbol{A}_0^{\dagger}\boldsymbol{A}_1\boldsymbol{g}_1\boldsymbol{g}_1^{\mathrm{T}}\right).$$

Therefore, (23) becomes

$$\Delta \boldsymbol{g}_1^{\perp} \approx -\sigma_v^8 \boldsymbol{A}_0^{\dagger} \mathcal{C}_1^{\mathrm{T}} \boldsymbol{U}_s \boldsymbol{\Lambda}_s^{-1} \boldsymbol{U}_s^{\mathrm{T}} \sum_{i=2}^{M} \left(\boldsymbol{h}_1^{\mathrm{T}} \boldsymbol{U}_s \boldsymbol{\Lambda}_s^{-2} \boldsymbol{U}_s^{\mathrm{T}} \boldsymbol{h}_i \right)^3 \boldsymbol{h}_i. \quad (29)$$

B. Complex Case

Now, (16) is used. Setting it to zero yields

$$\boldsymbol{g} = \frac{2|\boldsymbol{g}^{\mathrm{H}} \boldsymbol{A} \boldsymbol{g}_1|^2 \boldsymbol{g}_1^{\mathrm{H}} \boldsymbol{A} \boldsymbol{g}}{4\boldsymbol{g}^{\mathrm{H}} \boldsymbol{A} \boldsymbol{g} - 2} \boldsymbol{g}_1 + \frac{2}{4\boldsymbol{g}^{\mathrm{H}} \boldsymbol{A} \boldsymbol{g} - 2} \boldsymbol{A}^{-1} \times \sum_{i>1} \left| \boldsymbol{g}^{\mathrm{H}} \mathcal{C}_1^{\mathrm{H}} \boldsymbol{R}^{-1} \boldsymbol{h}_i \right|^2 \left(\boldsymbol{h}_i^{\mathrm{H}} \boldsymbol{R}^{-1} \mathcal{C}_1 \boldsymbol{g} \right) \mathcal{C}_1^{\mathrm{H}} \boldsymbol{R}^{-1} \boldsymbol{h}_i.$$

Following the similar procedures as in the real case, one can verify that channel estimation error in the orthogonal space of $\boldsymbol{g}_1$ is given by

$$\Delta \boldsymbol{g}_1^{\perp} = -\sigma_v^6 \left(\boldsymbol{I} - \frac{\boldsymbol{g}_1 \boldsymbol{g}_1^{\mathrm{H}}}{\boldsymbol{g}_1^{\mathrm{H}} \boldsymbol{g}_1} \right) \boldsymbol{A}^{-1} \mathcal{C}_1^{\mathrm{H}} \boldsymbol{U}_s \boldsymbol{\Lambda}_s^{-1} \boldsymbol{U}_s^{\mathrm{H}} \times \sum_{i>1} \left| \boldsymbol{h}_1^{\mathrm{H}} \boldsymbol{U}_s \boldsymbol{\Lambda}_s^{-2} \boldsymbol{U}_s^{\mathrm{H}} \boldsymbol{h}_i \right|^2 \left(\boldsymbol{h}_i^{\mathrm{H}} \boldsymbol{U}_s \boldsymbol{\Lambda}_s^{-2} \boldsymbol{U}_s^{h} \boldsymbol{h}_1 \right) \boldsymbol{h}_i + O\left(\sigma_v^8 \right). \quad (30)$$

Since (25) and (28) can be shown to hold for a complex system with $^{\mathrm{T}}$ replaced by $^{\mathrm{H}}$, (30) is simplified and approximated as

$$\Delta \boldsymbol{g}_1^{\perp} \approx -\sigma_v^8 \boldsymbol{A}_0^{\dagger} \mathcal{C}_1^{\mathrm{H}} \boldsymbol{U}_s \boldsymbol{\Lambda}_s^{-1} \boldsymbol{U}_s^{\mathrm{H}} \times \sum_{i=2}^{M} \left| \boldsymbol{h}_1^{\mathrm{H}} \boldsymbol{U}_s \boldsymbol{\Lambda}_s^{-2} \boldsymbol{U}_s^{\mathrm{H}} \boldsymbol{h}_i \right|^2 \left(\boldsymbol{h}_i^{\mathrm{H}} \boldsymbol{U}_s \boldsymbol{\Lambda}_s^{-2} \boldsymbol{U}_s^{\mathrm{H}} \boldsymbol{h}_1 \right) \boldsymbol{h}_i. \quad (31)$$

Considering (29) and (31), one can express them in a general form given in the proposition.

References

[1] W. Chen and U. Mitra, "An improved blind adaptive MMSE receiver for fast fading DS-CDMA channels," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 8, pp. 1531–1543, Aug. 2001.

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

[3] X. Wang and H. Poor, "Blind equalization and multiuser detection in dispersive CDMA channels," *IEEE Trans. Commun.*, vol. 46, no. 1, pp. 91–103, Jan. 1998.

[4] M. Tsatsanis and Z. Xu, "Performance analysis of minimum variance CDMA receivers," *IEEE Trans. Signal Process.*, vol. 46, no. 11, pp. 3014–3022, Nov. 1998.

[5] Z. Tian, K. Bell, and H. Van Trees, "Robust constrained linear receivers for CDMA wireless systems," *IEEE Trans. Signal Process.*, vol. 49, no. 7, pp. 1510–1522, Jul. 2001.

[6] Z. Xu, P. Liu, and X. Wang, "Blind multiuser detection: From MOE to subspace methods," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 510–524, Feb. 2004.

[7] D. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. COM-28, no. 11, pp. 1867–1875, Nov. 1980.

[8] J. Miguez and L. Castedo, "A linearly constrained constant modulus approach to blind adaptive multiuser interference suppression," *IEEE Commun. Lett.*, vol. 2, no. 8, pp. 217–219, Aug. 1998.

[9] C. Xu and G. Feng, "Comments on 'A linearly constrained constant modulus approach to blind adaptive multiuser interference suppression,'" *IEEE Commun. Lett.*, vol. 4, no. 9, pp. 280–282, Sep. 2000.

[10] L. Li and H. Fan, "Blind CDMA detection and equalization using linearly constrained CMA," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Jun. 2000, vol. 5, pp. 2905–2908.

[11] J. Tugnait and T. Li, "Blind asynchronous multiuser CDMA receivers for ISI channels using code-aided CMA," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 8, pp. 1520–1530, Aug. 2001.

[12] J. Tugnait and T. Li, "Blind detection of asynchronous CDMA signals in multipath channels using code-constrained inverse filter criterion," *IEEE Trans. Signal Process.*, vol. 49, no. 7, pp. 1300–1309, Jul. 2001.

[13] Z. Xu and P. Liu, "Code-constrained blind detection of CDMA signals in multipath channels," *IEEE Signal Process. Lett.*, vol. 9, no. 12, pp. 389–392, Dec. 2002.

[14] G. Gelli, L. Paura, and F. Verde, "A two-stage CMA-based receiver for blind joint equalization and multiuser detection in high data-rate DS-CDMA systems," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1209–1223, Jul. 2004.

[15] M. Gu and L. Tong, "Geometrical characterizations of constant modulus receivers," *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2745–2756, Oct. 1999.

[16] P. Liu and Z. Xu, "A globally convergent CMA-based approach to blind multiuser detection," in *Proc. Asilomar Conf. Signals, Syst., Comput.*, Nov. 2002, pp. 634–638.

[17] Z. Xu, "Perturbation analysis for subspace decomposition with applications in subspace-based algorithms," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2820–2830, Nov. 2002.

[18] M. Honig, S. Miller, M. Shensa, and L. Milstein, "Performance of adaptive linear interference suppression in the presence of dynamic fading," *IEEE Trans. Commun.*, vol. 49, no. 4, pp. 635–645, Apr. 2001.

[19] T. S. Rappaport, *Wireless Communications*. Upper Saddle River, NJ: Prentice-Hall, 1996.

[20] M. Rupp and S. C. Douglas, "A posteriori analysis of adaptive blind equalizers," in *Proc. 32nd Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, 1998, pp. 369–373.

Antenna Selection for Noncoherent Space–Time–Frequency-Coded OFDM Systems

Qian Ma and Cihan Tepedelenlioğlu, *Member, IEEE*

***Abstract*—This paper studies receive antenna selection (AS) for multiantenna systems that operate over frequency-selective channels, where the channel state information (CSI) is known neither at the transmitter nor at the receiver. We consider AS for noncoherent space–time–frequency (STF)-coded orthogonal frequency-division multiplexing (OFDM) systems that employ unitary signals. An OFDM receiver with a single radio-frequency chain is employed, and the selection is based on the received signal power, where the receive antenna with the largest received signal power averaged over all subcarriers is chosen. Theoretical analysis and simulations show that by using AS, the maximum spatial and multipath diversity can still be obtained without CSI at the receiver, which is the same as the full-complexity multiantenna OFDM systems. We simplify code design and decoding using subcarrier grouping, which enables us to apply existing full-diversity codes for noncoherent and differential STF-OFDM systems to each group, to obtain maximum spatial and multipath diversity when AS is used. We also present a low-complexity suboptimal decoder that can easily be implemented in practice.**

***Index Terms*—Antenna selection (AS), diversity, multiantenna communications, noncoherent space–time–frequency (STF) coding, orthogonal frequency-division multiplexing (OFDM).**

Manuscript received November 16, 2005; revised January 15, 2007 and February 27, 2007. This work was supported by the National Science Foundation CAREER Grant CCR-0133841. The review of this paper was coordinated by Prof. M. Juntti.

Q. Ma was with the Department of Electrical Engineering, Arizona State University, Tempe, AZ 85287 USA. He is now with Qualcomm Incorporated, San Diego, CA 92121 USA (e-mail: Qian.Ma@asu.edu).

C. Tepedelenlioğlu is with the Department of Electrical Engineering, Arizona State University, Tempe, AZ 85287 USA (e-mail: cihan@asu.edu).

Digital Object Identifier 10.1109/TVT.2007.901903