A PENALTY FUNCTION APPROACH TO CODE-CONSTRAINED CMA FOR BLIND MULTIUSER CDMA SIGNAL DETECTION

Jinghong Ma   *Jitendra K. Tugnait*

Department of Electrical & Computer Engineering
Auburn University, Auburn, Alabama 36849, USA
tugnait@eng.auburn.edu

ABSTRACT

A code-constrained constant-modulus approach (CMA) was presented recently in Li and Tugnait [4] for blind detection of asynchronous short-code DS-CDMA signals in multipath channels. Only the spreading code of the desired user is assumed to be known; its transmission delay may be unknown. The equalizer was determined by minimizing the Godard/CMA cost function of the equalizer output with respect to the equalizer coefficients subject to the fact that the equalizer lies in a subspace associated with the desired user's code sequence. An iterative projection approach was used in [4] for constrained optimization where at each iteration the equalizer was projected onto the desired subspace. In this paper we investigate an alternative, penalty function-based approach to constrained optimization. Global minima and some of the local minima of the cost function are investigated. A simulation example is presented.

1. INTRODUCTION

Direct sequence code division multiple access (DS-CDMA) systems have been a subject of intense research interest in recent years. In CDMA systems multiple users transmit signals simultaneously leading to multiuser interference (MUI). In addition to MUI, presence of multipath propagation introduces intersymbol interference (ISI) causing distortion of the spreading code sequences. Moreover, in reverse links, unknown transmission delays (user asynchronism) also contribute to performance degradation.

In this paper we consider blind detection (i.e. no training sequence) of the desired user signal, given knowledge of its spreading code, in the presence of MUI, ISI and user asynchronism (lack of knowledge of user transmission delays, including that of the desired user). Past work on blind detection of DS-CDMA signals include [1]-[4], [7],[8], [11] and references therein. In this paper our focus is on extraction of a desired user's signal. Unlike [2],[3], [7], and [8], we do not assume synchronization with the desired user's signal. In [11] we investigated maximization of the normalized fourth cumulant of inverse filtered (equalized) data w.r.t. the equalizer coefficients subject to the equalizer lying in a subspace associated with the desired user's code sequence. Constrained maximization leads to extraction of the desired user's signal whereas unconstrained maximization leads to the extraction of any one of the existing users. In [4] the same ideas were used in conjunction with the Godard cost (CMA) where we minimized the cost. In this paper we investigate an alternative, penalty function-based approach to constrained minimization of the Godard cost.

2. SYSTEM MODEL

Consider an asynchronous short-code DS-CDMA system with M users and N chips per symbol with the j-th user's spreading code denoted by $\mathbf{c}_j = [c_j(0), \cdots, c_j(N-1)]^T$. Consider a baseband discrete-time model representation. Let $s_j(k)$ denote the j-th user's k-th symbol. The sequence $\{s_j(k)\}$ is zero-mean, independently and identically distributed (i.i.d.) either QAM $\forall j$ or binary $\forall j$. For different j's, $\{s_j(k)\}$'s are mutually independent. In the presence of a linear dispersive channe, let $g_j(n)$ denote the j-th user's effective channel impulse response (IR) assuming zero transmission delay, sampled at the chip interval T_c. Let

$$h_j(n) = \sum_{m=0}^{N-1} c_j(m) g_j(n-m), \tag{1}$$

where $h_j(n)$ represents the effective signature sequence of user j (i.e. code $c_j(n)$ "distorted" due to multipath etc.). Define a $[(d+1)N] \times [2N]$ code matrix

$$\mathbf{C}_j^{(d)} := \begin{bmatrix} c_j(0) & 0 & \cdots & 0 \\ c_j(1) & c_j(0) & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ c_j(N-1) & \ddots & \ddots & c_j(0) \\ 0 & c_j(N-1) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & c_j(N-1) \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}. \tag{2}$$

If we collect N chip-rate measurements of received signal (from all users) into N-vector $\mathbf{y}(k)$, then we obtain, at the symbol rate, the MIMO model (additive white Gaussian noise $\mathbf{w}(k)$ is defined in a manner similar to $\mathbf{y}(k)$):

$$\mathbf{y}(k) = \sum_{j=1}^{M} \sum_{l=0}^{L_j} \mathbf{h}_j(l) s_j(k-l) + \mathbf{w}(k) \tag{3}$$

where

$$\mathbf{h}_j(l) = [h_j(lN - d_j), \cdots, h_j(lN - d_j + N - 1)]^T, \tag{4}$$

d_j $(0 \le d_j < N)$ is the (effective) transmission delay (mod N) of user j in chip intervals and $L_j + 1$ is the length of the j-th user's vector IR. It follows that for any $d \ge 0$,

$$\mathbf{h}_j^{(d)} := \begin{bmatrix} \mathbf{h}_j^H(0) & \mathbf{h}_j^H(1) & \cdots & \mathbf{h}_j^H(d) \end{bmatrix}^H = \mathbf{C}_j^{(d)} \mathbf{g}_j \tag{5}$$

where the superscript H denotes the complex conjugate transpose (Hermitian) operation,

$$\mathbf{g}_j := \begin{bmatrix} g_j(-d_j) & g_j(-d_j+1) & \cdots & g_j(2N-d_j-1) \end{bmatrix}^T, \tag{6}$$

$\mathbf{h}_j^{(d)}$ is $(d+1)N$-vector, $\mathbf{g}_j$ is $2N$-vector and we assume that $g_j(l) = 0$ for $l > N$ (in addition to $g_j(l) = 0$ for $l < 0$), i.e. the multipath delays can be of at most one symbol duration (N chips). Not all elements in $\mathbf{g}_j$ are nonzero. It follows that $\mathbf{h}_j(l) = 0$ for $l \ge 3$.

This work was supported by NSF Grant CCR-9803850.

3. CODE-CONSTRAINED CMA (CC-CMA)

3.1. Projection Approach to CC-CMA [4]

Consider an $N \times 1$ vector equalizer $\{\mathbf{f}(i)\}_{i=0}^{L_e-1}$ of length L_e symbols (NL_e chips) operating on the data $\mathbf{y}(n)$ (see (3)) to yield

$$e(n) \;=\; \sum_{i=0}^{L_e-1} \mathbf{f}^H(i)\mathbf{y}(n-i) \tag{7}$$

where $\mathbf{f}(i)$ is $N \times 1$. Define

$$\tilde{\mathbf{f}}^H := \left[\begin{array}{cccc} \mathbf{f}^H(0) & \mathbf{f}^H(1) & \cdots & \mathbf{f}^H(L_e-1) \end{array}\right]. \tag{8}$$

Following [4] consider minimization of the Godard cost

$$J_{cma}(\tilde{\mathbf{f}}) := E\{(|e(n)|^2-1)^2\} \tag{9}$$

for designing the linear equalizer. It is shown in [5],[9] (see also [6]) that under certain mild sufficient conditions, when (9) is minimized w.r.t. $\{\mathbf{f}(i)\}_{i=0}^{L_e-1}$ using a stochastic gradient algorithm, then (11) reduces to

$$e(n) \;=\; \alpha s_{j_0}(n-n_0), \tag{10}$$

where complex $\alpha \neq 0$, $0 \leq n_0 \leq L_e-1+L_j$ is some integer, j_0 indexes some user out of the given M users, i.e., the equalizer output is a possibly scaled and shifted version of one of the users. The problem is that there is no control over which user is extracted.

It has been shown in [4] that in order to extract the desired user ($j_0=1$) with desired delay ($n_0=d$), the linear equalizer should belong to the null space of a matrix $\mathcal{A}$ which is a function of the desired user's code matrix $\mathbf{C}_1^{(d)}$ and the data correlation matrix. It is a $[N(L_e-2)]\times[NL_e]$ matrix given by

$$\mathcal{A} = \mathcal{U}^{(1)H}\mathcal{T}\mathcal{R}_{yy} \tag{11}$$

where $\mathcal{R}_{yy}$ is the $[NL_e]\times[NL_e]$ data correlation matrix with ij-th block element $\mathbf{R}_{yy}(j-i) = E\{\mathbf{y}(k+j-i)\mathbf{y}^H(k)\}$,

$$\mathcal{T} := \left[\begin{array}{cc} \mathcal{T}_d & 0 \\ 0 & I_{N(L_e-1-d)} \end{array}\right] = [NL_e]\times[NL_e] \;\; \text{matrix}, \tag{12}$$

I_K denotes a $K \times K$ identity matrix,

$$\mathcal{T}_d := \left[\begin{array}{cccc} 0 & \cdots & 0 & I_N \\ 0 & \cdots & I_N & 0 \\ \vdots & \ddots & \vdots & \vdots \\ I_N & \cdots & 0 & 0 \end{array}\right] = [N(d+1)]\times[N(d+1)], \tag{13}$$

$$\mathcal{C}_1^{(d)} := \left[\begin{array}{c} \mathbf{C}_1^{(d)} \\ 0 \end{array}\right] \;=\; [NL_e]\times[2N] \;\; \text{matrix} \tag{14}$$

and columns of $\mathcal{U}^{(1)}$ denote an orthonormal basis for the orthogonal complement of $\mathcal{C}_1^{(d)}$. Since $\mathcal{C}_1^{(d)}$ is of full column rank, $\mathcal{U}^{(1)}$ is an $[NL_e]\times[NL_e-2N]$ matrix (it can be obtained via an SVD (singular value decomposition) of $\mathcal{C}_1^{(d)}$). **Thus, the desired solution satisfies (15) in addition to minimizing (9) (in fact, in addition to being a stationary point of (9))** where

$$\mathcal{A}\tilde{\mathbf{f}} = 0. \tag{15}$$

By [4] and [5] there exists an equalizer that minimizes (9) as well satisfies (15).

Let $\Pi_{\mathcal{A}}^{\perp}$ denote the $[NL_e]\times[NL_e]$ projection matrix onto the null space of $\mathcal{A}$. The the following iterative, batch, projection stochastic gradient algorithm was used in [4] to obtain the desired equalizer. Let $\widehat{J}_{cma}(\tilde{\mathbf{f}})$ denote the data-based cost (9) and let $\nabla_{\tilde{\mathbf{f}}^*}\widehat{J}_{cma}(\tilde{\mathbf{f}}')$ denote its gradient (NL_e-column) w.r.t. $\tilde{\mathbf{f}}^*$ evaluated at $\tilde{\mathbf{f}}'$; (the symbol $*$ denotes the complex conjugation operation). Given the equalizer $\tilde{\mathbf{f}}^{(n)}$ at n-th iteration, the equalizer update at $n+1$st iteration is calculated as $\tilde{\mathbf{f}}^{(n+1)} = \tilde{\mathbf{f}}^{(n)} - \rho\Pi_{\mathcal{A}}^{\perp}\nabla_{\tilde{\mathbf{f}}^*}\widehat{J}_{cma}(\tilde{\mathbf{f}}^{(n)})$, where ρ is a suitable step-size (see [4]). It is a projection algorithm since any changes in $\tilde{\mathbf{f}}^{(n)}$ are forced to lie in (projected onto) the null space of $\mathcal{A}$. Of course, we choose the initial guess $\tilde{\mathbf{f}}^{(0)}$ to satisfy (15) [4].

3.2. A Penalty Function Approach to CC-CMA

Now we propose a novel penalty function approach to CC-CMA. Instead of minimizing the cost (9) subject to (15), we propose to consider unconstrained minimization of

$$J_{cma}^{(p)}(\tilde{\mathbf{f}}) := J_{cma}(\tilde{\mathbf{f}}) + \mu\tilde{\mathbf{f}}^H\mathcal{A}^H\mathcal{A}\tilde{\mathbf{f}} \tag{16}$$

where $\mu > 0$ is a design parameter. The cost (16) is unlike the standard penalty function approaches [10] where the solution to the modified cost (such as (16)) approaches the desired (constrained) solution to the unmodified cost (such as (9)) subject to the constraint (such as $\mathcal{A}\tilde{\mathbf{f}}=0$) as $\mu \to \infty$. In our case, given the nature of the underlying system model and cost (9), we do not have to let $\mu \to \infty$. Indeed, for "small" enough μ, we have the unconstrained cost whose stationary points are well understood [5],[9].

3.2.1. GLOBAL MINIMA:

We now consider investigate global minima of (16). Assume no noise: $\mathbf{w}(k) \equiv 0$. When an equalizer is such that (10) is achieved, $J_{cma}(\tilde{\mathbf{f}})$ is minimized [5]. It can be shown that

$$\min_{\tilde{\mathbf{f}}} J_{cma}(\tilde{\mathbf{f}}) = 1 - \frac{(E\{|s_j(n)|^2\})^2}{E\{|s_j(n)|^4\}} = 0. \tag{17}$$

Let $\tilde{\mathbf{f}}_{1o}$ be an equalizer for which $J_{cma}(\tilde{\mathbf{f}}_{1o}) = 0$ with corresponding $e(n) = \alpha_1 s_1(n-d)$ where $|\alpha_1|^2 = E\{|s_j(n)|^2\}/E\{|s_j(n)|^4\}$, i.e. $\tilde{\mathbf{f}}_{1o}$ leads to extraction of user 1 with delay d. Then, by construction, $\mathcal{A}\tilde{\mathbf{f}}_{1o} = 0$ [4]. Hence, $J_{cma}^{(p)}(\tilde{\mathbf{f}}_{1o}) = 0$ and therefore, $\min_{\tilde{\mathbf{f}}} J_{cma}^{(p)}(\tilde{\mathbf{f}}) = 0$. It follows from the results of [5] that if $J_{cma}(\tilde{\mathbf{f}}) \neq 0$, then (10) can not hold true (all stable local minima of $J_{cma}(\tilde{\mathbf{f}})$ lead to (10) for some j_0 and n_0 [5]). Clearly $J_{cma}^{(p)}(\tilde{\mathbf{f}}) = 0$ if and only if $J_{cma}(\tilde{\mathbf{f}}) = 0$ and $\mathcal{A}\tilde{\mathbf{f}} = 0$. Therefore, global minima of $J_{cma}^{(p)}(\tilde{\mathbf{f}})$ are given by those $\tilde{\mathbf{f}}$'s for which $J_{cma}(\tilde{\mathbf{f}}) = 0$ and $\mathcal{A}\tilde{\mathbf{f}} = 0$, equivalently, for which (10) and (15) hold true. The equalizer that yields (10) satisfies [5]

$$\sum_{i=0}^{L_e-1} \mathbf{f}^H(i)\mathbf{h}_j(n-i) = \alpha\delta_{j,j_0}\delta_{n,n_0}, \;\; 1 \leq j \leq M, \;\; n \geq 0, \tag{18}$$

where $\delta_{j,i} = 1$ for $j=i$, 0 otherwise.

We now characterize the equalizer solutions that satisfy both (15) and (18). Define the $[NL_e]$-column vector, for $m = 0, \cdots, L_e-1+L_j$,

$$\tilde{\mathbf{h}}_j^{(m)} := \left[\begin{array}{cccccc} \mathbf{h}_j^H(m) & \cdots & \mathbf{h}_j^H(0) & 0 & \cdots & 0 \end{array}\right]^H. \tag{19}$$

Using (19) and results from [4], [11] and Sec. 3.1, (18) can be rewritten as ($\mathcal{R}_{ss} = \mathcal{R}_{yy}$ under the no noise assumption)

$$\mathcal{T}\mathcal{R}_{yy}\tilde{\mathbf{f}} = \alpha\mathcal{T}\tilde{\mathbf{h}}_{j_0}^{(n_0)}, \; j_0 \in \{1,\cdots,M\}, \; n_0 \in \{0,\cdots,L_e-1+L_{j_0}\}. \tag{20}$$

Therefore, solutions satisfying (15) and (18) are equivalent to solutions satisfying (15) and (20). Consider

(C1) The $[NL_e] \times [2N+1]$ matrix $[\mathcal{C}_1^{(d)} \vdots \mathcal{T}\tilde{\mathbf{h}}_j^{(m)}]$ has full column rank for every $j \in \{2, 3, \cdots, M\}$ and every $m \in \{d-1, d\}$ where $L_e \geq d+1$ and $d \geq 2$.

Claim 1: Under (C1), any solution that satisfies (15) and (20) (i.e. any global minimum of (16)), corresponds to $j_0 = 1$ and $n_0 \in \{d-1, d\}$ in (10), (18) and (20).

Proof: Suppose that for fixed j and m, $\mathcal{T}\tilde{\mathbf{h}}_j^{(m)} \notin \text{sp}\{\mathcal{C}_1^{(d)}\}$, where $\text{sp}\{B\}$ denotes the linear subspace spanned by the columns of B. Then $\mathcal{U}^{(1)H}\mathcal{T}\tilde{\mathbf{h}}_j^{(m)} \neq 0$ (i.e. if $\tilde{\mathbf{f}}$ satisfies (20) with $j_0 = j$ and $n_0 = m$, then it does not satisfy (15)), else $\mathcal{T}\tilde{\mathbf{h}}_j^{(m)} \in \text{sp}\{\mathcal{C}_1^{(d)}\}$. We now establish that $\mathcal{T}\tilde{\mathbf{h}}_j^{(m)} \notin \text{sp}\{\mathcal{C}_1^{(d)}\}$ for any $m > d$ irrespective of the nature of spreading codes and of multipaths. Recall that, by assumption, total delay spread is no more that $N+1$ chips for any user. We have $(m > d)$

$$\mathcal{T}\tilde{\mathbf{h}}_j^{(m)} = \big[\mathbf{h}_j^H(m-d),\ \mathbf{h}_j^H(m-d+1),\ \cdots,\ \mathbf{h}_j^H(m)$$

$$\mathbf{h}_j^H(m-d-1),\ \mathbf{h}_j^H(m-d-2),\ \cdots,\ \mathbf{h}_j^H(0),\ 0,\ \cdots,\ 0\big]^H \tag{21}$$

Since $\mathbf{h}_j(0) \neq 0$, it follows that $\mathcal{T}\tilde{\mathbf{h}}_j^{(m)} \notin \text{sp}\{\mathcal{C}_1^{(d)}\}$ as the rows in $\mathcal{C}_1^{(d)}$ corresponding to the position of $\mathbf{h}_j(0)$ in $\mathcal{T}\tilde{\mathbf{h}}_1^{(m)}$ are all zeros. Therefore, $\mathcal{T}\tilde{\mathbf{h}}_j^{(m)} \notin \text{sp}\{\mathcal{C}_1^{(d)}\}$ for any $m > d$ and $\forall j$. We now turn to the case of $m \leq d$. In this case

$$\mathcal{T}\tilde{\mathbf{h}}_j^{(m)} = \big[\ \mathbf{h}_j^H(m-d) \quad \mathbf{h}_j^H(m-d+1) \quad \cdots \quad \mathbf{h}_j^H(m)$$

$$0 \quad \cdots \quad 0\ \big]^H \tag{22}$$

We have

$$\mathcal{T}\tilde{\mathbf{h}}_j^{(m)} \in \text{sp}\{\mathcal{C}_1^{(d)}\} \ \Leftrightarrow\ [\mathbf{h}_j^H(m-d), \cdots, \mathbf{h}_j^H(m)]^H \in \text{sp}\{\mathbf{C}_1^{(d)}\}. \tag{23}$$

For (23) to be true, there must exist a $2N$-vector $\mathbf{g} \neq \mathbf{0}$ such that

$$[\mathbf{h}_j^H(m-d), \cdots, \mathbf{h}_j^H(m)]^H \stackrel{?}{=} \mathbf{C}_1^{(d)}\mathbf{g}. \tag{24}$$

If $d - m \geq 2$ (assuming that $d \geq 2$), then $\mathbf{h}_j(m-d) = \mathbf{h}_j(m-d+1) = \mathbf{0}$ leading to $\mathbf{g} = \mathbf{0}$ in (24): that is, (23) is never satisfied if $d - m \geq 2$; (since $m \geq 0$, this requires that $d \geq 2$.) Thus, $\mathcal{T}\tilde{\mathbf{h}}_j^{(m)} \notin \text{sp}\{\mathcal{C}_1^{(d)}\}\ \forall j$ and any $m \notin \{d-1, d\}$ for any choice of spreading codes and multipaths. Finally, by construction, $\mathcal{U}^{(1)H}\mathcal{T}\tilde{\mathbf{h}}_1^{(d)} = 0$. This proves the desired result. ⊙

If we pick $d \geq 2$, then the only possible convergence points from among (20) are $\mathcal{T}\tilde{\mathbf{h}}_j^{(m)}$ with $m = d$ or $m = d-1$ and $j = 1, 2, \cdots, M$. If $d = 3$, then both $\mathcal{T}\tilde{\mathbf{h}}_j^{(3)}$ and $\mathcal{T}\tilde{\mathbf{h}}_j^{(2)}$ contain the entire IR of the j-th user (recall that the IR is of maximum length $\overline{L} = 3$ symbols). If $d = 2$, then while $\mathcal{T}\tilde{\mathbf{h}}_j^{(2)}$ contains the entire IR, $\mathcal{T}\tilde{\mathbf{h}}_j^{(1)}$ may not since it does not contain $\mathbf{h}_j(2)$, which may (or may not) be nonzero. In order to better distinguish between two distinct users, it is therefore more prudent to use $d \geq 3$.

3.2.2. LOCAL MINIMA:

Let us allow doubly infinite equalizers $\{\mathbf{f}(i)\}_{i=-\infty}^{\infty}$. Define the scalar composite channel-equalizer impulse response from the j-th user to the equalizer output as

$$r_j(n) := \sum_{l=-\infty}^{\infty} \mathbf{f}^H(l)\mathbf{h}_j(n-l), \tag{25}$$

$$\mathbf{r} := [\cdots, r_1(0), \cdots, r_M(0), r_1(1), \cdots, r_M(1), r_1(2), \cdots]^T. \tag{26}$$

It has been shown in [5],[9] (see also [6, Appendix C]) that the only stable local minima of $\overline{J}_{cma}(\mathbf{r}) (= J_{cma}(\tilde{\mathbf{f}}))$ w.r.t. $\mathbf{r}$ are given by the solutions (18). In particular, let $\mathbf{r}_r$ and $\mathbf{r}_i$ denote the real and the imaginary parts, respectively, of $\mathbf{r}$. Let $\bar{\mathcal{J}}(\mathbf{r}')$ denote the the Hessian (second-order derivative) of $\overline{J}_{cma}(\mathbf{r})$ w.r.t. $[\mathbf{r}_r^T \ \ \mathbf{r}_i^T]^T$ evaluated at $\mathbf{r} = \mathbf{r}'$. Let $\mathbf{r}^{(n_0,j_0)}$ denote the vector $\mathbf{r}$ specified in (26) with all zero entries except for the one corresponding to $r_{j_0}(n_0)$ (see (25)) which equals α (cf. (10) and (18)). Then by [5], [9] and [6, Appendix C], $\bar{\mathcal{J}}(\mathbf{r}^{(n_0,j_0)})$ is positive definite on the set $\mathcal{F}_r = \{\mathbf{u} : \mathbf{u} = \mathbf{r} - \mathbf{r}^{(n_0,j_0)},\ \mathbf{u} \neq (e^{j\theta} - 1)\mathbf{r}^{(n_0,j_0)}\ \forall\theta\}$, i.e. $[\mathbf{u}_r^T \ \ \mathbf{u}_i^T]\bar{\mathcal{J}}(\mathbf{r}^{(n_0,j_0)})[\mathbf{u}_r^T \ \ \mathbf{u}_i^T]^T > 0$ for any $\mathbf{u} \in \mathcal{F}_r$, $\mathbf{u} \neq 0$, and it is positive semidefinite in general. Since any perturbation in the phase of α alone in (18) leaves the cost unchanged (i.e. $\overline{J}_{cma}(e^{j\theta}\mathbf{r}^{(n_0,j_0)}) = \overline{J}_{cma}(\mathbf{r}^{(n_0,j_0)})$, it follows that $[\mathbf{u}_r^T \ \ \mathbf{u}_i^T]\bar{\mathcal{J}}(\mathbf{r}^{(n_0,j_0)})[\mathbf{u}_r^T \ \ \mathbf{u}_i^T]^T = 0$ for $\mathbf{u} = (e^{j\theta} - 1)\mathbf{r}^{(n_0,j_0)}$. Thus, $\bar{\mathcal{J}}(\mathbf{r})$ has a strict local minimum at $\mathbf{r}^{(n_0,j_0)}$ for $\mathbf{r} \in (\{\mathbf{r}^{(n_0,j_0)}\} \cup \mathcal{F}_r')$ where $\mathcal{F}_r' := \{\mathbf{r} : \mathbf{r} = \mathbf{u} + \mathbf{r}^{(n_0,j_0)},\ \mathbf{u} \in \mathcal{F}_r\}$, and $\bar{\mathcal{J}}(\mathbf{r})$ has a local minimum at $\mathbf{r}^{(n_0,j_0)}$. Let $\mathcal{J}(\tilde{\mathbf{f}}')$ denote the Hessian of $J_{cma}(\tilde{\mathbf{f}})$ w.r.t. the real and the imaginary parts $\tilde{\mathbf{f}}_r$ and $\tilde{\mathbf{f}}_i$, respectively, of $\tilde{\mathbf{f}}$, evaluated at $\tilde{\mathbf{f}} = \tilde{\mathbf{f}}'$. By (25), (26) and the definition of $\tilde{\mathbf{f}}$, it follows that there exists a complex-valued matrix $\mathbf{B} = \mathbf{B}_r + j\mathbf{B}_i$, a function only of the MIMO channel IR (3), such that

$$\mathbf{r} = \mathbf{B}\tilde{\mathbf{f}} \ \Rightarrow\ \begin{bmatrix} \mathbf{r}_r \\ \mathbf{r}_i \end{bmatrix} = \begin{bmatrix} \mathbf{B}_r & -\mathbf{B}_i \\ \mathbf{B}_i & \mathbf{B}_r \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{f}}_r \\ \tilde{\mathbf{f}}_i \end{bmatrix} =: \mathcal{B} \begin{bmatrix} \tilde{\mathbf{f}}_r \\ \tilde{\mathbf{f}}_i \end{bmatrix}. \tag{27}$$

Therefore, we have $\mathcal{J}(\tilde{\mathbf{f}}) = \mathcal{B}^T\bar{\mathcal{J}}(\mathbf{r})\mathcal{B}$ for any $\mathbf{r} = \mathbf{B}\tilde{\mathbf{f}}$. Let $\tilde{\mathbf{f}}^{(n_0,j_0)}$ denote an equalizer (not necessarily unique) corresponding to the composite channel-equalizer response $\mathbf{r}^{(n_0,j_0)}$, i.e. $\mathbf{B}\tilde{\mathbf{f}}^{(n_0,j_0)} = \mathbf{r}^{(n_0,j_0)}$. Then it follows that $\mathcal{J}(\tilde{\mathbf{f}}^{(n_0,j_0)})$ is positive definite on the set $\mathcal{F}_f := \{\mathbf{u} : \tilde{\mathbf{f}} - \tilde{\mathbf{f}}^{(n_0,j_0)},\ \mathbf{B}\mathbf{u} \neq (e^{j\theta}-1)\mathbf{r}^{(n_0,j_0)}\ \forall\theta\}$, and positive semidefinite in general. Note that any perturbations in $\tilde{\mathbf{f}}^{(n_0,j_0)}$ that leave $\mathbf{r}^{(n_0,j_0)}$ unperturbed, do not change the CMA cost.

We now turn to the cost (16). Let $\mathcal{J}^{(p)}(\tilde{\mathbf{f}}')$ denote the Hessian of $J_{cma}^{(p)}(\tilde{\mathbf{f}})$ at $\tilde{\mathbf{f}} = \tilde{\mathbf{f}}'$. Then $\mathcal{J}^{(p)}(\tilde{\mathbf{f}}') = \mathcal{J}(\tilde{\mathbf{f}}') + \mu\mathcal{A}^H\mathcal{A}$. Moreover, $\nabla_{\tilde{\mathbf{f}}^*} J_{cma}^{(p)}(\tilde{\mathbf{f}}) = \nabla_{\tilde{\mathbf{f}}^*} J_{cma}(\tilde{\mathbf{f}}) + \mu\mathcal{A}^H\mathcal{A}\tilde{\mathbf{f}}$. If $\tilde{\mathbf{f}}^{(n_0,j_0)}$ is such that $\mathcal{A}\tilde{\mathbf{f}}^{(n_0,j_0)} = 0$, then it is a stationary point of $J_{cma}^{(p)}(\tilde{\mathbf{f}})$ because it is a stationary point of $J_{cma}(\tilde{\mathbf{f}})$. Therefore, the global minima of (16) are stationary points of $J_{cma}^{(p)}(\tilde{\mathbf{f}})$. By the discussion in the previous paragraph, they are also local minima of (16).

Let $\nabla_{\tilde{\mathbf{f}}^*} J_{cma}^{(p)}(\tilde{\mathbf{f}}_\mu') = 0$. Let $\lim_{\mu\to 0} \tilde{\mathbf{f}}_\mu' = \tilde{\mathbf{f}}_0'$. Then $\nabla_{\tilde{\mathbf{f}}^*} J_{cma}(\tilde{\mathbf{f}}_0') = 0$. If $\mathbf{B}\tilde{\mathbf{f}}_0' \neq \mathbf{r}^{(n_0,j_0)}$, then by [5],[9], $\mathcal{J}(\tilde{\mathbf{f}}_0')$ has at least one negative eigenvalue (recall from [5],[9] that unless $\mathbf{B}\tilde{\mathbf{f}}_0' = \mathbf{r}^{(n_0,j_0)}$, all other stationary points of

$J_{cma}(\tilde{\mathbf{f}})$ are either saddle points or local maxima). Since $\lim_{\mu\to 0}\mathcal{J}^{(p)}(\tilde{\mathbf{f}}'_\mu) = \mathcal{J}(\tilde{\mathbf{f}}'_0)$, if $\mathbf{B}\tilde{\mathbf{f}}'_0 \neq \mathbf{r}^{(n_0,j_0)}$, then $\mathcal{J}^{(p)}(\tilde{\mathbf{f}}'_\mu)$ has at least one negative eigenvalue for $0 \leq \mu \leq \bar{\mu}$ for some $\bar{\mu} > 0$. That is, for "small" μ's, all stationary points $\tilde{\mathbf{f}}'_\mu$ of $J^{(p)}_{cma}(\tilde{\mathbf{f}})$ that lead to $\mathbf{B}\tilde{\mathbf{f}}'_0 \neq \mathbf{r}^{(n_0,j_0)}$, are either saddle points or local maxima. The preceding discussion suggests that if we can initialize "close" to $\tilde{\mathbf{f}}'_0$ such that $\mathcal{A}\tilde{\mathbf{f}}'_0 = 0$, then we will achieve convergence to the solution(s) specified in Lemma 1. In Sec. 4 we describe a strategy for such an initialization.

4. SIMULATION EXAMPLE

We consider the case of 3 users, each transmitting 4-QAM signals, and short-codes with 8 chips per symbol. The spreading codes were randomly generated binary (± 1, with equal probability) sequences. The multipath channels for each user have 4 paths with transmission delays uniformly distributed over one symbol interval, and the remaining 3 multipaths having mutually independent delays (w.r.t. the first arrival) uniformly distributed over one symbol interval. All four multipath amplitudes are complex Gaussian with zero-mean and identical variance. The channels for each user were randomly generated in each of the 100 Monte Carlo runs. Complex white zero-mean Gaussian noise was added to the received signal from the 3 users. The SNR refers to the SNR of the desired user, which was user 1, and it equals the energy per symbol divided by N_0 (= one-sided power spectral density of noise $= 2E\{\|\mathbf{w}(k)\|^2\}/N$). In the equal-power case (0dB MUIs), all users have the same power; in the near-far case (10dB MUIs), the desired user power is 10 dB below that of other users.

Equalizer of length (L_e) 5 symbols and desired delay (lag) $d = 3$ was designed using the proposed penalty function algorithm: cost (16) with $\mu = 1$. Initialization was done as in [4] with first ten iterations using the projection approach (so that the initial choice satisfies data-based version of (15) exactly); then we switch to unconstrained minimization of (16) using an iterative, batch, stochastic gradient algorithm. The approach of [3] (equivalent to that of [2]) was also simulated with a "smoothing factor" (m in [3]) of 5 ($=L_e$). The approach of [3] was used to estimate the desired user's channel IR which, in turn, was used in a MMSE equalizer with delay $d = 2$. We also applied the approach of [1] using equalizer of length 5 symbols and desired delay $d = 2$. The projection approach of [4] for CC-CMA ($L_e = 5$, $d = 3$) was also implemented. We also simulate an ideal (clairvoyant) matched filter receiver which is matched to the true effective signature sequence $h_1(n - d_j)$ (or $\mathbf{h}_1(l)$) of user 1. Figs. 1 and 2 show the probability of symbol detection error (P_e) after equalization for the equal power and the near-far scenario, respectively, based on 100 Monte Carlo runs.

5. REFERENCES

[1] D. Gesbert, J. Sorelius, P. Stoica and A. Paulraj, "Blind multiuser MMSE detection of CDMA signals in ISI channels," *IEEE Comm. Lett.*, pp. 233-235, Aug. 1999.

[2] M. Torlak and G. Xu, "Blind multiuser channel estimation in asynchronous CDMA systems," *IEEE Trans. Signal Proc.*, vol. SP-45, pp. 137-147, Jan. 1997.

[3] X. Wang and H.V. Poor, "Blind equalization and multiuser detection in dispersive CDMA channels," *IEEE Trans. Commun.*, vol. COM-46, pp. 91-103, Jan. 1998.

[4] Tong-tong Li and J.K. Tugnait, "A code-constrained constant-modulus approach to blind detection of asynchronous CDMA signals in multipath channels," in *Proc. 2000 CISS*, pp. TAb7-13:17, Princeton U., NJ, March 15-17, 2000.

[5] J.K. Tugnait, "Blind spatio-temporal equalization and impulse response estimation for MIMO channels using a Godard cost function," *IEEE Trans. Signal Proc.*, vol. SP-45, pp. 268-271, Jan. 1997.

[6] J.K. Tugnait, "Identification and deconvolution of multichannel linear non-Gaussian processes using higher-order statistics and inverse filter criteria," *IEEE Trans. Signal Proc.*, vol. SP-45, pp. 658-672, March 1997.

[7] U. Madhow, "Blind adaptive interference suppression for direct-sequence CDMA," *Proc. IEEE*, vol. 86, pp. 2049-269, Oct. 1998.

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

[9] J.K. Tugnait, "Spatio-temporal signal processing for blind separation of multichannel signals," in *Digital Signal Processing Technology*, Joseph Picone, Editor, *Proc. SPIE 2750*, pp. 88-103, 1996.

[10] D.G. Luenberger, *Linear and Nonlinear Programming, Second Ed.*. Addison-Wesley: Reading, MA, 1984.

[11] J.K. Tugnait and Tong-tong Li, "Blind detection of asynchronous CDMA signals in multipath channels using code-constrained inverse filter criteria," in *Proc. IEEE 2000 ICASSP*, pp. V-2461:2464, Istanbul, Turkey, June 2000.

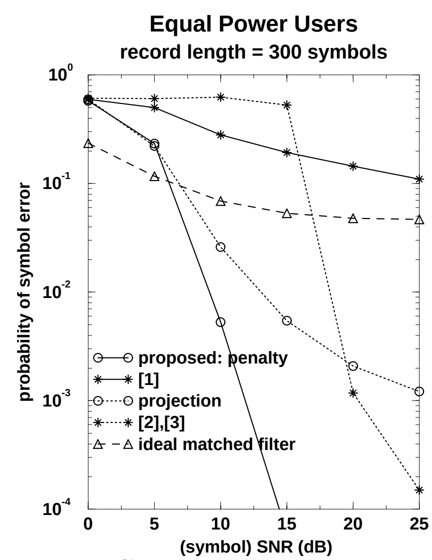

FIG. 1. Based on 100 runs.

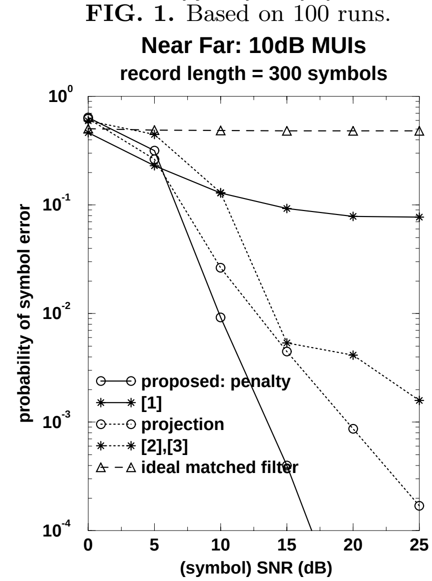

FIG. 2. Based on 100 runs.