A Toeplitz Displacement Method for Blind Multipath Estimation for Long Code DS/CDMA Signals

Carlos J. Escudero, *Member, IEEE*, Urbashi Mitra, *Member, IEEE*, and Dirk T. M. Slock, *Member, IEEE*

***Abstract*—The problem of blind channel identification for direct-sequence/code-division multiple-access (DS/CDMA) multiuser systems is explored. For wideband DS/CDMA signals, multipath distortion is well modeled by a finite-impulse response filter. In this work, a blind channel identification technique based on second-order statistics is investigated. The method exploits knowledge of the spreading code of the user of interest via matched filtering, as well as properties of spreading codes. The current scheme focuses on a method appropriate for randomized long sequence DS/CDMA. This access scheme poses special challenges as the spreading codes are time varying. An analytical approximation of the mean-squared error is derived using perturbation techniques. The performance of the algorithm is studied via simulation and through the mean-squared error approximation, which is observed to be tight.**

***Index Terms*—Blind channel identification, code-division multiple-access systems, perturbation analysis, randomized spreading sequences.**

I. Introduction

DIRECT-SEQUENCE/code-division multiple-access (DS/CDMA) is a promising candidate for next-generation wireless systems [1], [3]. In order to provide high-fidelity detection, it is necessary to either implicitly or explicitly characterize the wireless propagation channels. The propagation effects experienced by wideband signals such as DS/CDMA can be well modeled by a tapped delay line. With knowledge of these multipath coefficients, multiuser equalizers can be constructed (see, e.g., [5] and [21]). In this paper, we explore blind channel identification schemes for multiuser systems employing DS/CDMA based on second-order statistics. The search for blind identification methods is motivated by the desire to have training signal independent schemes to improve the overall data rate of the system.

Manuscript received May 21, 1999; revised August 28, 2000. This work was supported in part by the National Science Foundation under Grants INT-9403734 and NCR-9624375, by CICYT Grant TIC96-0500-C10-02, and by FEDER Grant 1FD97-0082. This paper was presented in part at the Eighth IEEE Digital Signal Processing Workshop, Bryce Canyon, UT, August 1998 and in part at the Bayona Workshop on Emerging Technologies in Telecommunications, Bayona, Spain, September 1999. The associate editor coordinating the review of this paper and approving it for publication was Prof. Dimitrios Hatzinakos.

C. J. Escudero is with the Departamento de Electrónica y Sistemas, La Coruña University, La Coruña, Spain.

U. Mitra was with the Department of Electrical Engineering, Ohio State University, Columbus, OH 43210 USA. She is now with the Department of Electrical Engineering—Systems, University of Southern California, Los Angeles, CA 90007 USA.

D. T. M. Slock is with the Mobile Communications Group, Eurecom Institute, Sophia Antipolis, France.

Publisher Item Identifier S 1053-587X(01)00076-9.

Furthermore, we will consider DS/CDMA systems where the signature waveforms have periods much longer than a symbol duration. Such *randomized* spreading is appropriate for systems based on IS-95 as well as for military systems that require a higher degree of security as afforded by long randomized codes. The contributions of this paper are the development and analysis of a blind channel identification scheme for randomized spreading codes. In particular, an approximation of the mean-squared channel estimation error is derived based on perturbation analysis techniques. Simulation studies verify the tightness of the approximation. The mean-squared error development represents some of the first work on error derivation for channel estimation in systems with randomized codes (in addition to that found in [19]). Due to the spreading operation, DS/CDMA signals lend themselves naturally to a category of blind channel identification techniques that have become recently popular (see e.g., [10]). Although it is possible to estimate the global channel response (see, e.g., [7] and [17]), we will focus on schemes that estimate the pure channel only (e.g., [15]).

We will adopt an approach based on second-order statistics. Such methods typically exploit channel diversity as obtained via oversampling or multiple sensors. However, in DS/CDMA systems, such techniques to create a multichannel system are unnecessary. This is due to the fact that DS/CDMA systems are inherently multichannel as a result of the spreading operation. Thus, while multiple sensor systems are considered in [7] and [15], they are not necessary, as seen in [2]. In essence, each chip of a user's spreading code behaves like a channel.

The basis of our methods will be the direct exploitation of the spreading sequence of the user of interest via matched filtering. For delay spreads that are much smaller than the symbol period, a moderate number of matched filter outputs form a set of statistics that are suitable for performing blind channel identification. This often allows for a significant reduction in computational complexity (see, e.g., [2]) due to the reduction in dimension of the observation vector.

Although there has been a significant amount of research directed toward channel identification for short spreading code systems [2], [15], [17], there has been limited effort with regard to the randomized code scenario. A key challenge in the estimation of channel responses in systems where randomized codes are employed is the time-varying nature of the spreading codes. The two approaches that have been adopted for the long spreading code identification problem can be categorized as *deterministic* or *stochastic*. In deterministic approaches, the long spreading code is considered to be a deterministic sequence. These methods are characterized by the requirement of a significant amount of information about the interfering users. Recently, there have been several contributions to this type of blind identification for long code systems. A key property shared by [6], [16], [18], and [19] is that it is presumed that the receiver has knowledge of all of the users' spreading codes.[1] For randomized code DS/CDMA systems, such a requirement translates to knowing KB different length N spreading codes for K active users and a transmission block of B bits.

In [16], an iterative approach that modifies the channel estimation method of [14] to the multiuser, randomized code case is considered. Essentially, a series of least-squares problems, where, alternatively, the data symbols or the channel is assumed to be known and the other quantity is estimated, are solved. Correlation matching techniques are employed in [18]. In [6], an iterative scheme is derived based on alternating projections. Whereas [6], [16], and [18] consider the uplink estimation problem, [19] focuses on the downlink. In the downlink, one can exploit the fact that all users' signals, including the interferers, are received through the same channel.

The second approach to the problem at hand is to take a stochastic view. Thus, the long spreading sequences are modeled as random sequences, and properties of the resulting statistics are exploited to perform identification. In this category are present the proposed work and the work of [8]. The stochastic approach obviates the need for the complete knowledge of all user signature waveforms. The algorithm presented herein can thus has applicability for both the uplink and the downlink. It will be seen that the matched filtering algorithm developed herein offers a comparable channel estimation error to that of [8]. The new algorithm does, however, offer several tunable design parameters to regulate performance and thus provides additional flexibility. It should be noted that the complexity of the algorithm in [8] is less than the proposed algorithm. The main contributions of the current work are the observation that the matched filtered multiple access interference can be well-modeled as interference that is stationary at the chip rate, a tight approximation of the mean-squared channel estimation error for a randomized spreading code environment, and thorough simulation studies of the proposed algorithm. The simulation studies reveal that neither the proposed algorithm nor the algorithm of [8] are near–far resistant, which is a facet not previously studied. It is noted that the requirement of full knowledge of the spreading codes of all active users for the works in [16], [18], and [19] precludes fair comparison to the technique presented herein.

This paper is organized as follows. Section II presents the signal model and the assumptions employed. Section III describes the technique for blind channel identification for systems with randomized codes. Section IV provides the analysis of the mean-squared error for the new identification scheme. Issues relating to the practical implementation of the proposed algorithm are noted in Section V. Numerical results are presented in Section VI, and concluding remarks are provided in Section VII. The Appendix contains key derivations necessary for the mean-squared error analysis.

[1]A decentralized correlation matching scheme is alluded to in [18] that does not require all users' codes.

II. Signal Model

We will presume a coherent system where the active users transmit DS/CDMA signals. The signal is transmitted over a multipath channel. Coarse synchronization (accuracy to within half a chip) is assumed for the user of interest. Synchronization for DS/CDMA systems is a challenging issue; possible synchronization schemes include [13], [15], and [17]. A synchronization scheme specifically for randomized code DS/CDMA systems was proposed in [20].

The baseband representation of the received signal after coherent reception is given by

$$x(t) = \sum_{n=-\infty}^{\infty} \sum_{p=1}^{P} A_p \tilde{c}_p^n(t - nT - \tau_p) b_p(n) + w(t). \quad (1)$$

In randomized code DS/CDMA, the spreading waveform for each user p, $\tilde{c}_p^n(t)$ changes from bit to bit and is thus a function of n, which is the symbol index. In the equation above, $\tilde{c}_p^n(t)$ denotes the effective spreading waveform for user p and for QAM symbol $b_p(n)$. The effective spreading waveform is constructed through convolution of the original spreading waveform with the channel $\tilde{c}_p^n(t) = c_p^n(t) \star h_p(t)$, where the channel response for user p is $h_p(t)$. The spreading waveform is formed via

$$c_p^n(t) = \sum_{l=1}^{N} c_p^n(l) \psi(t - lT_c)$$

where $\psi(t)$ is the pulse shape of the chip and has duration T_c. For simplicity, we will consider rectangular pulse shapes. The sequence $c_p^n(l)$ is the spreading sequence for user p that changes from symbol to symbol and takes on values equal to $(\pm 1/\sqrt{N})$ with equal probability, where N is the spreading gain (i.e., number of chips/bit). Note that symbol duration is T, and thus, $T = NT_c$. The delay for user p is denoted by τ_p, and the transmitted amplitude for user p is A_p. These amplitudes are modeled as fixed, but unknown quantities. We will assume that the delay values are integer multiples of a chip. The fractional parts of the delays are incorporated into the effective channel impulse response $h_p(t)$. The additive, circularly symmetric Gaussian noise process is $w(t)$ with variance σ_w^2.

We will assume that the data signals and the noise process are mutually independent. We also assume that the channel is of finite length and that the multipath delay spread is less than a symbol interval. For facility of description, we assume that the channel length for each user is M chips ($M < N$). In fact, the techniques explored herein only require knowledge of the desired user's channel length. The received signal is chip-matched filtered and sampled at the chip-rate; then, $aN + M - 1$ samples are concatenated to form the observation vector $\mathbf{x}(n)$. The quantity a represents the number of "whole" data symbols contained in the observation vector. This concept is elucidated further in the sequel. Fig. 1 depicts the scenario where $a = 1$, $N = 5$, and $M = 3$. The filtered and sampled complex channel impulse response is denoted by $\mathbf{h}_p = [h_p(0), \cdots, h_p(M-1)]^T$.

For notational clarity, we begin by providing an expression for the observation vector under the assumption of a synchronized system ($\tau_p = 0 \forall p$). The extension to asynchronous in-

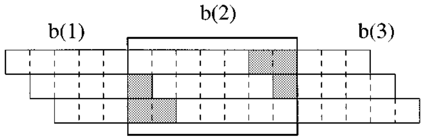

Fig. 1. Abstract depiction of observation vector for a single user system with length 5 (N) spreading sequence; three paths (M) and $a = 1$.

terferers is straightforward once the relevant matrices have been defined. We consider an observation vector containing a whole data symbols (and two partial bits; see the shaded area in Fig. 1). The partially observed bits are due to intersymbol interference. The observation vector of $aN+M-1$ samples (at the chip rate) is expressed as

$$\mathbf{x}(n) = \sum_{p=1}^{P} A_p \mathbf{C}_p(n) \mathbf{H}_p \mathbf{b}_p(n) + \mathbf{w}(n) \tag{2}$$

where the $\mathbf{x}(n) = [x(n), \cdots, x(n + aN + M - 2)]^T$ and $\mathbf{w}(n) = [w(n), \cdots, w(n+aN+M-2)]^T$ are vectors of the received samples and noise samples of size $(aN+M-1)\times 1$, and $\mathbf{b}(n) = [b(\lfloor n/N \rfloor -1), \cdots, b(\lfloor n/N \rfloor + a)]^T$ is a $(a+2)\times 1$ vector of data bits. The operator $\lfloor \cdot \rfloor$ is the floor operator and returns the largest integer smaller than its argument. The channel matrix for user p of dimension $(a+2)M \times (a+2)$ is given by

$$\mathbf{H}_p = \mathbf{I}_{a+2} \otimes \mathbf{h}_p \tag{3}$$

where $\otimes$ denotes the Kronecker product operator, and $\mathbf{I}_{a+2}$ is the identity matrix of dimension $(a+2) \times (a+2)$. We will use the notation $\mathcal{C}(\mathbf{v}, M)$ to denote the Sylvester matrix of width M for vector $\mathbf{v}$. This is equivalent to the convolution matrix for the convolution of vector $\mathbf{v}$ with another vector of length M. $\mathcal{C}(\mathbf{c}_p(n), M)$ is a $(N+M-1)\times M$ matrix since $\mathbf{c}_p(n) = [c_p(n), \cdots, c_p(n+N-1)]^T$, i.e., see (4) at the bottom of the page. We define $\mathbf{C}^1_{p,M}(n)$ to be the first N rows of $\mathcal{C}(\mathbf{c}_p(n), M)$ and $\mathbf{C}^2_{p,M}(n)$ to be the last $M-1$ rows of $\mathcal{C}(\mathbf{c}_p(n), M)$. Then, the spreading code matrix for user p of dimension $(aN+M-1)\times(a+2)M$ is given by the second equation at the bottom of the page, where the matrix $\tilde{\mathbf{C}}^1_{p,M}(n+(a+1)N)$ is composed of the first $M-1$ rows of $\mathbf{C}^1_{p,M}(n+(a+1)N)$.

In order to consider asynchronous systems, we can modify (2) as follows. First, we will presume that user 1 is the user of interest. We define the global channel matrix as $\mathbf{G}_p(n) = \mathbf{C}_p(n)\mathbf{H}_p$. Let $\tau_p = d_p T_c$ be the delay of interfering user p with respect to user 1. We assume that $0 \le \tau_p < T$ and that $d_p \in \{0, 1, \cdots, N-1\}$. The global channel matrices for the interfering users are constructed by forming the matrix $\mathbf{G}^s_p(n)$ as above, but this matrix is of dimension $((a+1)N + M - 1)\times(a+3)$. The channel matrix $\mathbf{G}_p(n)$ is formed by removing the first $N - d_p$ rows and the last d_p rows from $\mathbf{G}^s_p(n)$; finally, the last column is also removed. Note that this global channel matrix can be written as the product of appropriately truncated $\mathbf{C}_p(n)$ and $\mathbf{H}_p$ matrices.

In order to more fully exploit properties of the DS/CDMA signals, the identification scheme presented in this work processes a set of matched filter outputs synchronized to the desired user's signal rather than processing the received signal $\mathbf{x}(n)$ directly. We use M matched filters per received symbol. The $aM \times 1$ observation vector $\mathbf{y}(n)$ is given by

$$\begin{aligned}\mathbf{y}(n) &= \mathbf{S}_1(n)\mathbf{x}(n) \\ &= \mathbf{S}_1(n)\left(\sum_{p=1}^{P} A_p \mathbf{C}_p(n)\mathbf{H}_p\mathbf{b}_p(n)\right) + \mathbf{S}_1(n)\mathbf{w}(n)\end{aligned} \tag{5}$$

$$\mathcal{C}(\mathbf{c}_p(n), M) = \begin{pmatrix} c_p(n) & 0 & \cdots & 0 \\ c_p(n+1) & c_p(n) & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ c_p(n+M-1) & \cdots & \cdots & c_p(n) \\ \vdots & & & \vdots \\ c_p(n+N-1) & \cdots & \cdots & c_p(n+N-M) \\ 0 & c_p(n+N-1) & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & c_p(n+N-1) \end{pmatrix} \tag{4}$$

$$\mathbf{C}_p(n) = \begin{pmatrix} \mathbf{C}^2_{p,M}(n-N) & \mathbf{C}^1_{p,M}(n) & & & & \\ & \mathbf{C}^2_{p,M}(n) & \mathbf{C}^1_{p,M}(n+N) & & \mathbf{0} & \\ & & \mathbf{C}^2_{p,M}(n+N) & \mathbf{C}^1_{p,M}(n+2N) & & \\ & \mathbf{0} & & \ddots & \ddots & \\ & & & & \mathbf{C}^2_{p,M}(n+aN) & \tilde{\mathbf{C}}^1_{p,M}(n+(a+1)N) \end{pmatrix}$$

where the $aM \times (aN + M - 1)$ matched filtering matrix $\mathbf{S}_1(n)$ is given by (6), shown at the bottom of the page, where $\mathbf{C}^1_{1,M}(n)$ and $\mathbf{C}^2_{1,M}(n)$ are the first N and last $M-1$ rows of $\mathcal{C}(\mathbf{c}_1(n), M)$, respectively.

The matrices $\mathbf{S}_1^T(n)$ and $\mathbf{C}_1(n)$ are related in the following manner:

$$\mathbf{C}_1(n) = \begin{pmatrix} \mathbf{C}^2_{1,M}(n-N) & \vdots & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_1^T(n) & \mathbf{0} \\ \mathbf{0} & \vdots & \tilde{\mathbf{C}}^1_{1,M}(n+(a+1)N) \end{pmatrix}. \tag{7}$$

That is, $\mathbf{C}_1(n)$ is formed by augmenting $\mathbf{S}_1^T(n)$ by the appropriate $2M$ columns.

III. Toeplitz Displacement-Based Channel Estimation

The randomized spreading code model implies that the autocorrelation sequence of the desired user's spreading sequence is a delta function. Similarly, the cross-correlation sequence between two distinct users is ideally identically zero. Let us consider the $aM \times 1$ matched filter output vector $\mathbf{y}(n)$ of (5). With our assumptions on the data symbols and the noise, we take expectation with respect to the data symbols and the noise. Thus, the output vector covariance matrix is

$$\mathbf{R}_y(n) = \mathbf{E}_{\mathbf{b},\mathbf{w}}[\mathbf{y}(n)\mathbf{y}^H(n)] \tag{8}$$

$$\tag{9}$$

$$= \sigma_1^2 \mathbf{S}_1(n)\mathbf{C}_1(n)\mathbf{H}_1\mathbf{H}_1^H\mathbf{C}_1^H(n)\mathbf{S}_1^T(n) + \mathbf{R}_I(n) + \mathbf{R}_w(n) \tag{10}$$

where $\sigma_1^2 = A_1^2\mathbf{E}\{b_1^2(n)\}$, and $\mathbf{R}_w(n) = \sigma_w^2\mathbf{S}_1(n)\mathbf{S}_1^T(n)$ is the noise autocorrelation matrix. The contribution of the interfering users at time n is thus

$$\mathbf{R}_I(n) = \sum_{p=2}^{P} \sigma_p^2 \mathbf{S}_1(n)\mathbf{C}_p(n)\mathbf{H}_p\mathbf{H}_p^H\mathbf{C}_p^H(n)\mathbf{S}_1^T(n).$$

We now consider the time-averaged version of $\mathbf{R}_y(n)$. That is

$$\begin{aligned}\overline{\mathbf{R}}_y &= \frac{1}{N_s}\sum_{n=1}^{N_s}\mathbf{R}_y(n) \\ &= \frac{\sigma_1^2}{N_s}\sum_{n=1}^{N_s}\mathbf{S}_1(n)\mathbf{C}_1(n)\mathbf{H}_1\mathbf{H}_1^H\mathbf{C}_1^H(n)\mathbf{S}_1^T(n) \\ &\quad + \overline{\mathbf{R}}_I + \overline{\mathbf{R}}_w \\ &\overset{N,N_s\to\infty}{\approx} \sigma_1^2\overline{\mathbf{SC}}_1\mathbf{H}_1\mathbf{H}_1^H\overline{\mathbf{SC}}_1^H + \sum_{p=2}^{P}\sigma_p^2\mathbf{R}(\mathbf{h}_p) + \sigma_w^2\mathbf{I}\end{aligned} \tag{11}$$

where

$$\overline{\mathbf{SC}}_1 = \frac{1}{N_s}\sum_{n=1}^{N_s}\mathbf{S}_1(n)\mathbf{C}_1(n)$$

where N_s is the number of samples employed for time averaging, and $\mathbf{R}(\mathbf{h}_p)$ is the asymptotic averaged covariance matrix for user p. We note that the asymptotic approximation above follows from key assumptions made about the randomized spreading codes. That is, the components of the code sequences are independently and identically distributed and, furthermore, are stationary at the chip rate. Therefore, we have $\mathbf{S}_1(n)\mathbf{C}_1(n) = \mathbf{SC}_1 + \mathbf{A}(n)$, where $\mathbf{A}(n)$ is a time-varying perturbation matrix, and $\mathbf{SC}_1 = \lim_{N_s\to\infty}\overline{\mathbf{SC}}_1 = [\mathbf{0}\ \mathbf{I}_{aM}\ \mathbf{0}]$. The perturbation $\mathbf{A}(n)$ gets small as N gets large, and hence, terms of the form $\mathbf{A}(n)\mathbf{H}_1\mathbf{H}_1^H\mathbf{A}^H(m)$ become negligible as $N \to \infty$. The effects of the imperfect spreading auto-correlation function are captured in $\overline{\mathbf{SC}}_1$. Asymptotically, as the spreading gain increases ($N \uparrow$) and as the number of samples used to average increases ($N_s \uparrow$), $\overline{\mathbf{R}}_I$ and $\overline{\mathbf{R}}_w$ converge to Toeplitz matrices [$\mathbf{R}(\mathbf{h}_p)$ is the Toeplitz autovariance matrix of the channel $\mathbf{h}_p$].

Channel identification can be performed by noting that the Toeplitz contribution of the channel noise and interference can be removed. To that end, let us focus on the limiting case $\lim_{N_s\to\infty}\overline{\mathbf{R}}_y = \mathbf{R}_y$. Consider now the following Toeplitz displacement operation:

$$\begin{aligned}\mathbf{R}_h &= \mathbf{R}_y(2{:}\,aM, 2{:}\,aM) - \mathbf{R}_y(1{:}\,aM-1, 1{:}\,aM-1) \\ &= \mathbf{R}_y^+ - \mathbf{R}_y^- \\ &= \sigma_1^2\mathbf{SC}_1^+\mathbf{H}_1\mathbf{H}_1^H\mathbf{SC}_1^{+H} - \sigma_1^2\mathbf{SC}_1^-\mathbf{H}_1\mathbf{H}_1^H\mathbf{SC}_1^{-H} && (12)\\ &= \sigma_1^2\mathbf{I}_{aM}^+\tilde{\mathbf{H}}_1\tilde{\mathbf{H}}_1^H\mathbf{I}_{aM}^{+H} - \sigma_1^2\mathbf{I}_{aM}^-\tilde{\mathbf{H}}_1\tilde{\mathbf{H}}_1^H\mathbf{I}_{aM}^{-H} && (13)\\ &= \sigma_1^2\begin{pmatrix}\mathbf{h}_1^+\mathbf{h}_1^{+H} & & & \\ & \mathbf{h}_1\mathbf{h}_1^H & & \\ & & \ddots & \\ & & & \mathbf{h}_1\mathbf{h}_1^H\end{pmatrix} \\ &\quad - \sigma_1^2\begin{pmatrix}\mathbf{h}_1\mathbf{h}_1^H & & & \\ & \mathbf{h}_1\mathbf{h}_1^H & & \\ & & \ddots & \\ & & & \mathbf{h}_1^-\mathbf{h}_1^{-H}\end{pmatrix} && (14)\end{aligned}$$

$$\mathbf{S}_1^T(n) = \begin{pmatrix} \mathbf{C}^1_{1,M}(n) & \mathbf{0} & & & \\ \mathbf{C}^2_{1,M}(n) & \mathbf{C}^1_{1,M}(n+N) & & & \\ \mathbf{0} & \mathbf{C}^2_{1,M}(n+N) & \cdots & & \mathbf{0} \\ & & \ddots & \ddots & \\ & & & & \mathbf{C}^1_{1,M}(n+aN) \\ & & & & \mathbf{C}^2_{1,M}(n+aN) \end{pmatrix}. \tag{6}$$

where the matrix notation $\mathbf{B}(i\colon\ j, i\colon\ j)$ corresponds to the submatrix of $\mathbf{B}$ formed by the appropriately truncated rows i through j and columns i through j, and $\mathbf{SC}_1^+ = \mathbf{SC}_1(2\colon aM, 1\colon aM)$ and $\mathbf{SC}_1^- = \mathbf{SC}_1(1\colon aM-1, 1\colon aM)$. Thus, $\mathbf{SC}_1^-$ and $\mathbf{SC}_1^+$ are formed from $\mathbf{SC}_1$ by removing the last and the first row, respectively. The quantities $\mathbf{I}_{aM}^+$, $\mathbf{I}_{aM}^-$, $\mathbf{h}_1^+$, and $\mathbf{h}_1^-$ are defined similarly to $\mathbf{SC}_1^-$ and $\mathbf{SC}_1^+$. Finally, $\tilde{\mathbf{H}}_1 = \mathbf{I}_a \otimes \mathbf{h}_1$. Note that $\mathbf{R}_h$ is a $(aM-1)\times(aM-1)$ matrix of rank $2a$. The $2a$ nonzero eigenvalues correspond to the contribution of the desired user to $\mathbf{R}_h$. The eigendecomposition of $\mathbf{R}_h$ is given by

$$\mathbf{R}_h = \sum_{i=1}^{aM-1} \lambda_i \mathbf{v}_i \mathbf{v}_i^H. \tag{15}$$

We order the eigenvalues such that $\lambda_1 \geq \lambda_2, \cdots, \geq \lambda_{aM-1}$. With such an ordering, $\lambda_1, \cdots, \lambda_a$ are positive, $\lambda_{aM-a}, \cdots, \lambda_{aM-1}$ are negative, and the remaining eigenvalues are equal to zero so that in fact

$$\mathbf{R}_h = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^H \tag{16}$$

where $\mathbf{V} = [\mathbf{v}_1 \cdots \mathbf{v}_a \mathbf{v}_{aM-a} \cdots \mathbf{v}_{aM-1}]$, and $\lambda = \mathrm{diag}\{\lambda_1, \cdots, \lambda_a, \lambda_{aM-a}, \cdots, \lambda_{aM-1}\}$. Thus, $\mathbf{V}$ is comprised of the eigenvectors corresponding to the nonzero eigenvalues in $\mathbf{\Lambda}$ of $\mathbf{R}_h$. Loosely speaking, $\mathbf{R}_y^+$ contributes a positive eigenvalues and $\mathbf{R}_y^-$ contributes a negative eigenvalues, as seen in (14). For $\mathbf{h}_1$ to be identifiable from $\mathbf{R}_h$, it is clear that a cannot take on arbitrary values given a particular value of M. The valid choices for a are explored in Section VI.

In practice, from the sample average correlation matrix $\hat{\mathbf{R}}_y = (1/N_s)\sum_{n=1}^{N_s} \mathbf{y}(n)\mathbf{y}^H(n)$, we form the estimated $\hat{\mathbf{R}}_h$ and determine $\hat{\mathbf{v}}_1, \cdots, \hat{\mathbf{v}}_a$ and $\hat{\mathbf{v}}_{aM-a}, \cdots, \hat{\mathbf{v}}_{aM-1}$ as the eigenvectors corresponding to the a most positive and a most negative eigenvalues of $\hat{\mathbf{R}}_h$.

We introduce the following subspace fitting problem to estimate the unknown channel:

$$\hat{\mathbf{h}}_1 = \arg \min_{\mathbf{h}_1, \mathcal{W}} \|\mathcal{H} - \hat{\mathbf{V}}\,\mathcal{W}\|_F^2 \tag{17}$$

where $\mathcal{H} = [(\overline{\mathbf{SC}}_1\mathbf{H}_1)^+(\overline{\mathbf{SC}}_1\mathbf{H}_1)^-] = [\overline{\mathbf{SC}}_1^+\overline{\mathbf{SC}}_1^-]\mathbf{H}_1$. We use $\mathcal{H} = [\overline{\mathbf{SC}}_1^+\overline{\mathbf{SC}}_1^-]\mathbf{H}_1$ and not $\mathcal{H} = [\mathbf{I}_{aM}^+\mathbf{I}_{aM}^-]\tilde{\mathbf{H}}_1$ in an attempt to somewhat compensate for the nonideal finite sample correlations of the desired user's spreading sequence [see (12)]. Taking into account that $\|\mathbf{A}\|_F^2 = \mathrm{Trace}\{\mathbf{A}^H\mathbf{A}\}$ is the Frobenious norm, the criterion in (17) can be expressed as

$$\begin{aligned}\hat{\mathbf{h}}_1 &= \arg \min_{\mathbf{h}_1, \mathcal{W}} \mathrm{Trace}\{\mathcal{H}^H\mathcal{H} - \mathcal{H}^H\hat{\mathbf{V}}\mathcal{W} - \mathcal{W}^H\hat{\mathbf{V}}^H\mathcal{H} \\ &\qquad + \mathcal{W}^H\hat{\mathbf{V}}^H\hat{\mathbf{V}}\mathcal{W}\} \\ &= \arg \min_{\mathbf{h}_1, \mathcal{W}} \mathrm{Trace}\{\mathcal{H}^H\mathcal{H} - \mathcal{H}^H\hat{\mathbf{V}}\mathcal{W} - \mathcal{W}^H\hat{\mathbf{V}}^H\mathcal{H} \\ &\qquad + \mathcal{W}^H\mathcal{W}\} \qquad (18) \\ &= \arg \min_{\mathbf{h}_1, \mathcal{W}} \mathcal{J}. \qquad (19)\end{aligned}$$

To solve (19), we must find the matrix where the gradient of $\mathcal{J}$ with respect to $\mathcal{W}$ vanishes,[2] i.e.,

$$\frac{\partial \mathcal{J}}{\partial \mathcal{W}} = \mathbf{0} - \hat{\mathbf{V}}^H\mathcal{H} - \mathbf{0} + \mathcal{W} = \mathbf{0} \Rightarrow \mathcal{W} = \hat{\mathbf{V}}^H\mathcal{H}. \tag{20}$$

Substituting (20) into (19), we obtain

$$\begin{aligned}\hat{\mathbf{h}}_1 &= \arg \min_{\mathbf{h}_1} \mathrm{Trace}\{\mathcal{H}^H\mathcal{H} - \mathcal{H}^H\hat{\mathbf{V}}\hat{\mathbf{V}}^H\mathcal{H} - \mathcal{H}^H\hat{\mathbf{V}}\hat{\mathbf{V}}^H\mathcal{H} \\ &\qquad + \mathcal{H}^H\hat{\mathbf{V}}\hat{\mathbf{V}}^H\mathcal{H}\} \\ &= \arg \min_{\mathbf{h}_1} \mathrm{Trace}\{\mathcal{H}^H(\mathbf{I} - \hat{\mathbf{V}}\hat{\mathbf{V}}^H)\mathcal{H}\}. \qquad (21)\end{aligned}$$

This fitting problem can be expressed as follows:

$$\begin{aligned}\hat{\mathbf{h}}_1 &= \arg \min_{\|\mathbf{h}_1\|=1} \mathrm{Trace}\{\mathcal{H}^H\mathbf{P}\mathcal{H}\} \qquad (22) \\ &= \arg \min_{\|\mathbf{h}_1\|=1} \mathbf{h}^H\mathbf{Q}\mathbf{h} \qquad (23)\end{aligned}$$

where $\mathbf{P} = \mathbf{I} - \hat{\mathbf{V}}\hat{\mathbf{V}}^H$ is the projection matrix onto the orthogonal complement of the column space of $\hat{\mathbf{V}}$. Thus, the desired solution is the eigenvector of $\mathbf{Q}$ corresponding to its minimum eigenvalue, where $\mathbf{Q}$ is defined as

$$\mathbf{Q} = \sum_{i=1}^{a+2} \sum_{\beta=-}^{+} \tilde{\mathbf{D}}_i^H \overline{\mathbf{SC}}_1^{\beta H} \mathbf{P} \overline{\mathbf{SC}}_1^{\beta} \tilde{\mathbf{D}}_i. \tag{24}$$

The $\tilde{\mathbf{D}}_i$ are $(a+2)M \times M$ permutation matrices defined as follows:

$$\tilde{\mathbf{D}}_i(l, m) = \begin{cases} 1, & m = l - (i-1)M \text{ for } i = 1, \cdots, a+2 \\ 0, & \text{otherwise.} \end{cases} \tag{25}$$

Using these permutation matrices, $\mathbf{H}_p = \mathbf{I}_{a+2} \otimes \mathbf{h}_p$ can be alternatively represented as $\mathbf{H}_p = [\tilde{\mathbf{D}}_1\mathbf{h}_p\tilde{\mathbf{D}}_2\mathbf{h}_p \cdots \tilde{\mathbf{D}}_{a+2}\mathbf{h}_p]$.

The recognition of the Toeplitz structure of the interference matrix was also made in [8]. However, instead of working with the Toeplitz displacement of $\mathbf{R}_y$, the difference of the covariance matrices before and after code-matched filtering was manipulated. Under ideal spreading code assumptions, it can be shown that this difference is, in principle, $\mathbf{h}_1\mathbf{h}_1^H$. Thus, the principal eigenvector of the difference between the matched filtered covariance matrix and the nonmatched filtered covariance matrix will yield the desired channel estimate. The work in [8] focuses on equalization but alludes to a corresponding identification scheme as just noted. We will compare the proposed channel identification scheme to that alluded to in [8] in Section VI.

IV. Approximation of Mean-Squared Estimaton Error

In this section, we provide an analytical approximation to the mean-squared channel estimation error. We show, via simulation studies in Section VI, that the approximation derived herein matches quite closely to simulation data. A perturbation technique (see, e.g., [11]) is employed to determine the approxima-

[2]Note that $(\partial\, \mathrm{Trace}\{\mathbf{A}\mathcal{W}\}/\partial\mathcal{W}) = \mathbf{A}^H$ and $(\partial\, \mathrm{Trace}\{\mathcal{W}^H\mathbf{A}\}/\partial\mathcal{W}) = \mathbf{0}$, where $\mathbf{A}$ is an arbitrary matrix independent of $\mathcal{W}$ [22].

tion of the mean-squared error. Such a technique exploits the following identities:

$$\left.\begin{aligned}\mathbf{Q}\mathbf{h} &= \mathbf{0}\\ \hat{\mathbf{Q}}\hat{\mathbf{h}} &\approx \mathbf{0}\\ \hat{\mathbf{h}} &= \mathbf{h}+\Delta\mathbf{h}\\ \hat{\mathbf{Q}} &= \mathbf{Q}+\Delta\mathbf{Q}\end{aligned}\right\} \Rightarrow (\mathbf{Q}+\Delta\mathbf{Q})(\mathbf{h}+\Delta\mathbf{h}) \approx \mathbf{0} \tag{26}$$

$$\text{where } \hat{\mathbf{h}} = \mathbf{v}_{\min}(\hat{\mathbf{Q}}).$$

The minimum eigenvalue of $\hat{\mathbf{Q}}$ is denoted $\lambda_{\min}(\hat{\mathbf{Q}})$, and the corresponding eigenvector is denoted $\mathbf{v}_{\min}(\hat{\mathbf{Q}})$. We note that in the set of expressions above, $\hat{\mathbf{Q}}\hat{\mathbf{h}} \approx \mathbf{0}$. This approximation is justified as follows. Technically, $\hat{\mathbf{Q}}\hat{\mathbf{h}} = \lambda_{\min}(\hat{\mathbf{Q}})\mathbf{v}_{\min}(\hat{\mathbf{Q}})$. As $N_s \to \infty$, $\hat{\mathbf{Q}} \to \mathbf{Q}$, we can thus assume that $\lim_{N_s\to\infty} \lambda_{\min}(\hat{\mathbf{Q}}) = \lambda_{\min}(\mathbf{Q}) = 0$ [12]. For a reasonable number of samples, N_s, $\hat{\mathbf{Q}}\hat{\mathbf{h}}$ will closely approximate the zero vector. This is borne out via empirical evidence collected through simulation studies.

Note that in the sequel, we are going to remove the subindex $_1$ of the desired user vector channel $\mathbf{h}_1$. Since $\mathbf{Q}\mathbf{h} = \mathbf{0}$ and second-order effects are presumed negligible, $(\Delta\mathbf{Q})(\Delta\mathbf{h}) \approx \mathbf{0}$, and we can modify the expression in (26) to yield the following approximation:

$$\mathbf{Q}(\Delta\mathbf{h}) \simeq -(\Delta\mathbf{Q})\mathbf{h} \tag{27}$$

and therefore

$$\begin{aligned}\Delta\mathbf{h} &\simeq -\mathbf{Q}^{\dagger}(\Delta\mathbf{Q})\mathbf{h}\\ &= -\mathbf{Q}^{\dagger}(\hat{\mathbf{Q}}-\mathbf{Q})\mathbf{h}\\ &= -\mathbf{Q}^{\dagger}\hat{\mathbf{Q}}\mathbf{h}\end{aligned} \tag{28}$$

where $\mathbf{Q}^{\dagger}$ denotes the Moore–Penrose pseudo-inverse of $\mathbf{Q}$. The kth component of $\Delta\mathbf{h}$ is given by

$$\begin{aligned}\Delta h(k) &\simeq -\mathbf{q}_k^H\hat{\mathbf{Q}}\mathbf{h}\\ &= -\mathbf{q}_k^H \sum_{i=1}^{a+2}\sum_{\beta=-,+} \tilde{\mathbf{D}}_i^H \overline{\mathbf{SC}}_1^{\beta H}\hat{\mathbf{P}}\overline{\mathbf{SC}}_1^{\beta}\tilde{\mathbf{D}}_i\mathbf{h}\\ &= -\sum_{i=1}^{2(a+2)} \mathbf{q}_k^H\mathbf{D}_i^H\hat{\mathbf{P}}\mathbf{D}_i\mathbf{h}\\ &= -\text{Trace}\{\mathbf{Q}_k^H\hat{\mathbf{P}}\mathcal{H}\}\end{aligned} \tag{29}$$

where $\hat{\mathbf{P}}$ is the estimated noise subspace projection matrix, and $\mathbf{q}_k$ is the kth column of $(\mathbf{Q}^{\dagger})^H$. The matrix $\mathbf{Q}_k^H$ is formed via $\mathbf{q}_k$ and the definition of the following matrices:

$$\mathbf{D}_i = \overline{\mathbf{SC}}_1^{-}\tilde{\mathbf{D}}_i,\ \mathbf{D}_{i+a+2} = \overline{\mathbf{SC}}_1^{+}\tilde{\mathbf{D}}_i \qquad \text{for } i = 1, \cdots, a+2. \tag{30}$$

Then, $\mathbf{Q}_k = [\mathbf{D}_1\mathbf{q}_k \cdots \mathbf{D}_{2(a+2)}\mathbf{q}_k]$. The channel super matrix $\mathcal{H}$ is also constructed using the matrices $\mathbf{D}_i$: $\mathcal{H} = [\mathbf{D}_1\mathbf{h} \cdots \mathbf{D}_{2(a+2)}\mathbf{h}]$. This is an alternative, but equivalent, definition to that provided in (17).

Based on the results of [12, eq. (4.11)], we obtain the following identity:

$$\hat{\mathbf{P}}\mathcal{H} = -\mathbf{P}\hat{\mathbf{V}}\mathbf{V}^H\mathcal{H}. \tag{31}$$

From [12, eq. (A.2)] and the fact that the eigenvalues associated with the noise subspace are ideally zero, we obtain

$$\mathbf{P}\hat{\mathbf{V}} = \mathbf{P}\hat{\mathbf{R}}_h\mathbf{V}\mathbf{\Lambda}^{-1}. \tag{32}$$

Note that the Toeplitz displacement shown in (12) ideally removes any contribution due to the noise and the interferers. Substituting (31) and (32) into (29), we obtain the following expression:

$$\Delta h(k) = \text{Trace}\{\mathbf{Q}_k^H\mathbf{P}\hat{\mathbf{R}}_h\mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{V}^H\mathcal{H}\}. \tag{33}$$

We seek to compute $\mathbf{E}[\Delta h(k)\Delta h^*(k)]$. Deriving this expectation relies on exploiting the fourth-order statistics of binary and Gaussian random variables. As these derivations are somewhat lengthy, they are relegated to the Appendix.

V. Practical Considerations

There are three practical considerations in the implementation and evaluation of the proposed algorithm that require discussion. These three considerations are

1) the need for normalization due to the scalar ambiguity present in the channel estimate;
2) a cleaning operation for the estimated difference covariance matrix to improve performance;
3) overestimation of the channel order.

A. Channel Estimate Normalization

It is well known that a scalar ambiguity exists in the resulting channel estimate for identification employing second-order statistics since the channel can only be estimated up to a scale factor. To remove the effect of the unknown scalar, we consider a normalization of the vector channel estimate. The chosen normalization is

$$\hat{\hat{\mathbf{h}}} = \alpha\hat{\mathbf{h}} \tag{34}$$

where α is a complex scalar selected such that $\mathbf{h}^H\hat{\hat{\mathbf{h}}} = \mathbf{h}^H\mathbf{h}$. The Cramér–Rao bound (CRB) for this type of normalization corresponds to the Moore–Penrose pseudo-inverse of the Fisher information matrix (FIM), which is singular due to the scalar ambiguity [4]. This particular normalization leads to the lowest possible CRB and corresponds to the smallest introduction of additional information since information is introduced only in the nullspace of the FIM.

The desired normalization is $\alpha = (\mathbf{h}^H\mathbf{h}/\mathbf{h}^H\hat{\mathbf{h}})$. However, note that this normalization is a function of the channel estimate. In order to facilitate the analysis of the approximate mean-squared error in Section IV, we seek to approximate the normalization such that the resulting channel estimation error is adjusted by the appropriate linear transformation.

We observe that the estimated channel can be written in terms of the true channel and a perturbation: $\hat{\mathbf{h}} = \mathbf{h} + \Delta\mathbf{h}$. We will assume that the scaling factor in $\hat{\mathbf{h}}$ has already been adjusted so that $\Delta\mathbf{h}$ is small. In that case, we can approximate the normalization scalar α up to first order in $\Delta\mathbf{h}$ as $\alpha = (1/(1 + (\mathbf{h}^H\Delta\mathbf{h}/\mathbf{h}^H\mathbf{h}))) \approx 1 - (\mathbf{h}^H\Delta\mathbf{h}/\mathbf{h}^H\mathbf{h})$. Then, continuing to

TABLE I
CHANNEL IMPULSE RESPONSES FOR DIFFERENT USERS USED FOR THE SIMULATIONS IN THE NUMERICAL RESULTS SECTION

User	$h_p(0)$	$h_p(1)$	$h_p(2)$	$h_p(3)$	$h_p(4)$
$p=1$	0.5044 + 0.2940i	0.0695 + 0.0820i	0.0630 - 0.2483i	0.3847 + 0.3439i	0.0173 + 0.5617i
$p=2$	-0.2986 + 0.0897i	0.4610 + 0.4635i	-0.1588 - 0.5624i	-0.3462 + 0.0203i	0.1081 + 0.0452i
$p=3$	-0.5310 - 0.0072i	-0.1070 - 0.4971i	0.5343 + 0.0988i	0.2285 - 0.0317i	0.1743 - 0.2838i
$p=4$	-0.0832 - 0.1942i	-0.4008 - 0.3086i	-0.3532 - 0.4961i	0.0506 - 0.1902i	0.0412 - 0.5368i
$p=5$	-0.0732 + 0.0346i	0.0062 - 0.1027i	0.2873 + 0.4010i	-0.5975 + 0.4138i	0.2379 - 0.3932i
$p=6$	-0.0705 + 0.1222i	0.2832 + 0.2327i	0.0081 + 0.4078i	0.3544 + 0.2719i	0.5233 + 0.4539i
$p=7$	0.5055 + 0.4016i	0.5151 - 0.2220i	-0.0212 + 0.2852i	0.0647 - 0.2764i	-0.2194 - 0.2409i
$p=8$	0.1239 + 0.1393i	-0.5517 - 0.0410i	0.1532 + 0.2978i	0.6352 - 0.3194i	-0.1999 + 0.0400i
$p=9$	0.5588 + 0.0070i	0.2968 - 0.1685i	-0.2135 + 0.1854i	0.0748 - 0.4746i	0.4501 - 0.2403i
$p=10$	0.1136 - 0.1623i	-0.0123 + 0.2168i	0.6700 + 0.1196i	0.0089 + 0.3254i	0.5111 - 0.2884i

approximate up to first order in $\Delta\mathbf{h}$, we get, for the normalized channel estimation error

$$\begin{aligned}\Delta\mathbf{h}_{norm} &= \alpha\hat{\mathbf{h}} - \mathbf{h} \approx \hat{\mathbf{h}} - \frac{\mathbf{h}^H\Delta\mathbf{h}}{\mathbf{h}^H\mathbf{h}}(\mathbf{h}+\Delta\mathbf{h}) - \mathbf{h}\\ &\approx \Delta\mathbf{h} - \frac{\mathbf{h}^H\Delta\mathbf{h}}{\mathbf{h}^H\mathbf{h}}\mathbf{h}\\ &= \Delta\mathbf{h} - \frac{\mathbf{h}\mathbf{h}^H}{\mathbf{h}^H\mathbf{h}}\Delta\mathbf{h} = \mathbf{P}_h^{\perp}\Delta\mathbf{h} \end{aligned} \tag{35}$$

where $\mathbf{P}_h^{\perp} = \mathbf{I} - \mathbf{h}(\mathbf{h}^H\mathbf{h})^{-1}\mathbf{h}^H$ is the projection onto the nullspace of $\mathbf{h}^H$. Thus, we have neglected second-order terms of $\Delta\mathbf{h}$.

To incorporate the normalization of (34), (28) is modified as

$$\Delta\mathbf{h}_{norm} = -\mathbf{P}_h^{\perp}\mathbf{Q}^{\dagger}\hat{\mathbf{Q}}\mathbf{h} \tag{36}$$

and therefore, the $\mathbf{q}_k$ vector in (29) can be considered to be the kth column of $(\mathbf{P}_h^{\perp}\mathbf{Q}^{\dagger})^H$. Note that (28) is an expression for $\Delta\mathbf{h}$ to which (35) can be applied appropriately because the derivation of (28) assumed that $\Delta\mathbf{h}$ was small. Note, however, also that $\mathbf{P}_h^{\perp}\mathbf{Q}^{\dagger} = \mathbf{Q}^{\dagger}$ since $\mathbf{Q}$ has a one-dimensional (1-D) null space in the direction of $\mathbf{h}$.

B. Cleaning Operation

To further improve the channel estimate, we apply a "cleaning" operation to the sample covariance matrix. Under the assumption that the Toeplitz displacement results in a matrix free of the contribution of any interference, $\mathbf{R}_h$ would have the form exhibited in (14). In practice, $\hat{\mathbf{R}}_h$ will not be block diagonal. Thus, to impose a block diagonal structure and remove the effects of nonideal correlation functions, we propose to modify the construction of $\hat{\mathbf{R}}_h$ via

$$\hat{\mathbf{R}}_h = (\hat{\mathbf{R}}_y \odot \mathbf{K})^+ - (\hat{\mathbf{R}}_y \odot \mathbf{K})^- \tag{37}$$

where

$$\mathbf{K}(i, j) = \begin{cases} 1, & \text{if } (m-1)M+1 \le i \le mM \\ & \text{and } (m-1)M+1 \le j \le mM \\ & \text{for } m \in [1, \cdots, a] \\ 0, & \text{else.} \end{cases}$$

Componentwise matrix multiplication (the Schur product) is indicated by $\odot$. Essentially, the cleaning operation replaces $\hat{\mathbf{R}}_y$ by $\hat{\mathbf{R}}_y \odot \mathbf{K}$ before applying the Toeplitz displacement to obtain $\hat{\mathbf{R}}_h$.

C. Channel Order Overestimation

In practice, the exact channel order may not be known *a priori*. If the maximum value is known, the algorithm can be designed to employ this maximum number of matched filters. Thus, the channel to be estimated would be a vector with several of its final components as zero. Empirically, we have observed that the algorithm performance appears to be insensitive to overestimating the channel order by one or two.

VI. NUMERICAL RESULTS

The proposed algorithm's performance is studied through the calculated and simulated mean-squared channel estimation error (MSE). The MSE found via simulation for the user of interest (user 1) is determined as follows:

$$\text{MSE}_1 = \frac{1}{N_r}\sum_{i=1}^{N_r}\left\|\frac{\mathbf{h}_1^T\mathbf{h}_1}{\mathbf{h}_1^T\hat{\mathbf{h}}_1^i}\hat{\mathbf{h}}_1^i - \mathbf{h}_1\right\|^2 \tag{38}$$

where N_r is the number of Monte Carlo runs in the simulation; this quantity is distinct from N_s, which is the number of symbols employed to determine the channel estimate. For each simulation, $N_r = 50$. The true channel is denoted by $\mathbf{h}_1$, and the channel estimate for Monte Carlo run i is given by $\hat{\mathbf{h}}_1^i$. Thus, the channel estimate is normalized, as discussed previously. For the asynchronous scenario considered, each of the relative delays are generated randomly from a uniform distribution. The channel coefficients are drawn from independent complex Gaussian random variables; the channel impulse responses used for the simulations are shown in Table I. The delays are fixed throughout a simulation, as are the amplitudes. The spreading codes for each user and for each symbol are selected from independent equiprobable binary random variables, and they are changed in every Monte Carlo run.

Fig. 2 investigates the accuracy of the MSE analysis for the new algorithm. The MSE is plotted as a function of the number of symbols used to estimate the channel. An environment with $K = 6$ asynchronous DS/CDMA users (i.e., only the first six channel impulse responses of Table I were used) employing spreading gains of length $N = 30$ is considered. The common channel length is $M = 5$, whereas the SNR was 15 dB. Three curves are exhibited.

1) simulation data [for which the MSE is evaluated as in (38)];

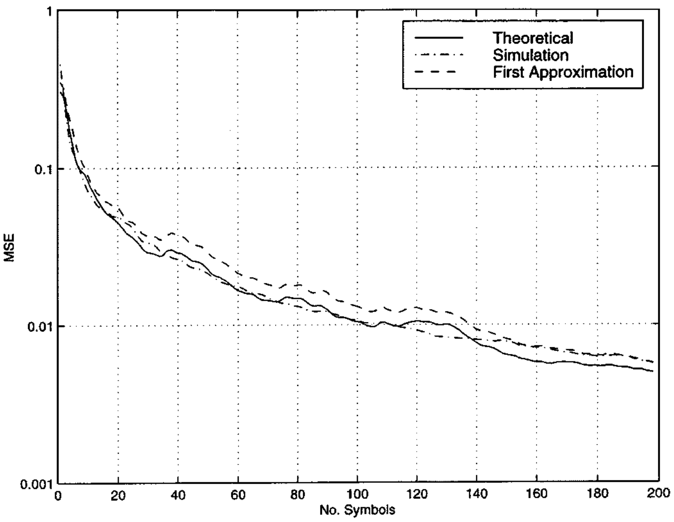

Fig. 2. MSE evolution versus N_s. Empirical MSE, theoretical MSE given by (58), and a "first approximation" of the MSE obtained via (33), as explained in Section VI.

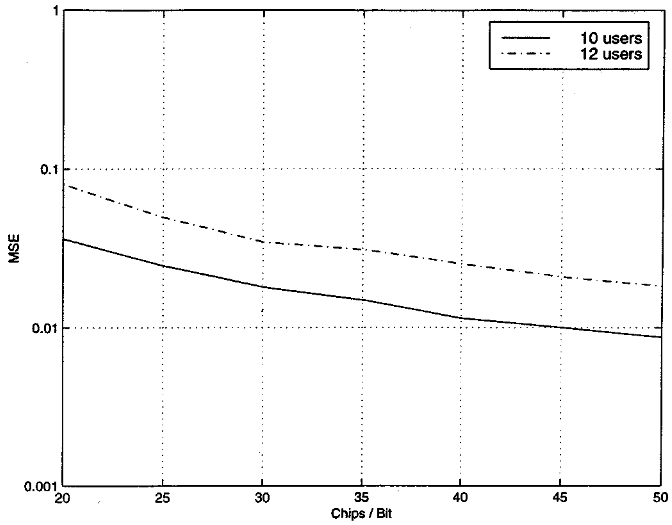

Fig. 3. MSE versus spreading gain.

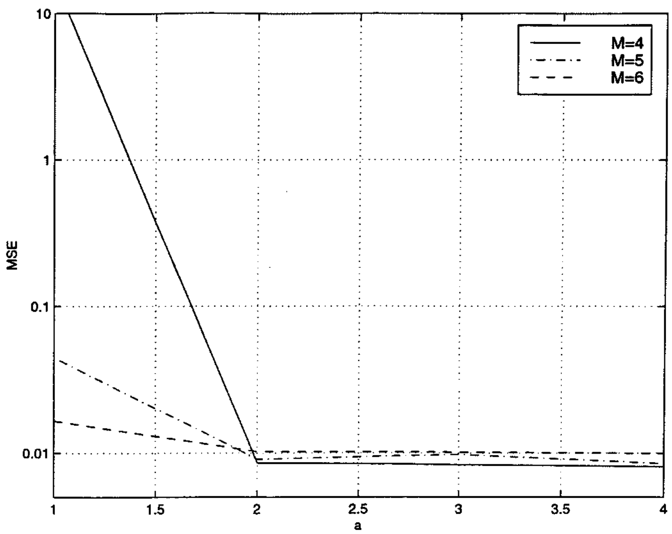

Fig. 4. MSE versus number a.

2) the "first approximation" expression [for which the MSE is computed as $(1/N_r)\sum_{i=1}^{N_r}\sum_{k=0}^{M-1}|\Delta h_1^i(k)|^2$ with $\Delta h_1^i(k)$ as obtained in (33)];
3) the theoretical MSE expression of (58) for the value $a = 3$.

The "first approximation" expression is considered as a very tight approximation to the theoretical MSE with lower computational complexity than that of (58). For the simulations, $N_r = 50$ runs were conducted. It is clear that (58) yields an expression for the MSE that is very close to the one obtained by means of simulation.

We next consider performance as a function of the spreading gain. Fig. 3 shows the theoretical MSE versus the spreading gain for two environments with 10 and 12 users. Different random channels with length 5 were considered in $N_r = 100$ Monte Carlo runs, $a = 2$, and SNR $= 15$ dB for all the users. The MSE value was computed for 300 transmitted symbols. As was predicted, performance improves as the code length increases since the spreading codes become increasingly more orthogonal to each other. In addition, we see that there is a constant difference in performance between the $K = 10$ and $K = 12$ user cases.

We next explore performance as a function of the number of whole symbols in the observation vector a. Fig. 4 depicts the theoretical MSE versus a for three different channel lengths (different random channel impulse responses for the $N_r = 100$ Monte Carlo run). The environment is for $K = 8$ users, SNR $=$ 20 dB, length $N = 30$ spreading codes, and $N_s = 200$. Note that for $M = 4$, a value of $a = 1$ leads to exceptionally poor performance. This can be explained on the basis of identifiability considerations. The matrix $\mathbf{R}_h$ is of dimension $aM - 1$. As $2a$ eigenvectors are the major contributors to the signal subspace, the dimension of the noise subspace is thus $a(M-2)-1$. The existence of a noise subspace leads to the requirement $M \geq 3$. The set of equations that allows us to compute the channel corresponds to the inner products of the noise subspace vectors with the columns of $\tilde{\mathbf{H}}_1^+$ and $\tilde{\mathbf{H}}_1^-$ (these inner products being zero theoretically). These matrices have a columns each. Therefore, we have $M-1$ unknowns (for the channel of length M minus the unidentifiable scale factor) and $2a(a(M-2)-1)$ equations. In order to obtain a solution, we need for the number of equations to be at least equal to the number of unknowns, i.e.,

$$2a^2(M-2) - 2a - M + 1 \geq 0. \tag{39}$$

Given this expression, it is clear that there is no valid value of a for $M = 1, 2$. For the Toeplitz displacement method, it is clear that channels must have length $M \geq 3$ in order for $\mathbf{R}_h$ to have a noise subspace. For $M = 3, 4$, the minimum value for a is 2. For $M \geq 5$, we find $a \geq 1$ satisfies (39). Therefore, in particular for $M = 4$, $a = 1$, the channel is unidentifiable as confirmed by the simulation.

Next, we compare the performance of the proposed algorithm with the one described in [8]. We denote the algorithm of [8] as LZ and the proposed algorithm as TOEP. Fig. 5 shows the results of three different simulations for $N_r = 50$; the first one uses the

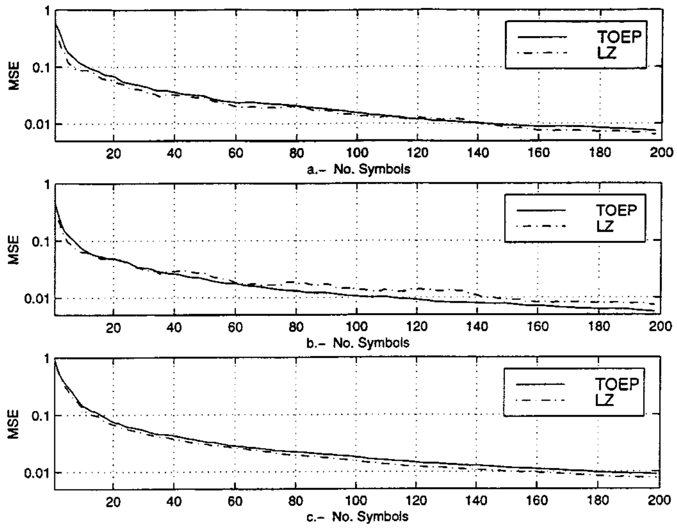

Fig. 5. MSE evolution for a power-controlled environment.

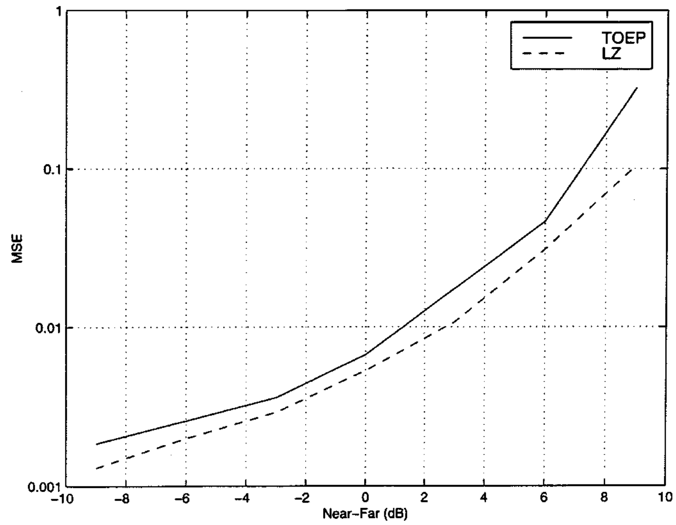

Fig. 6. MSE evolution for a near–far environment.

channel impulse responses shown in Table I, and the second and third use different random channels for each Monte Carlo run. Fig. 5(a) considers an environment with $K = 8$ users, $N = 30$, SNR $= 20$ dB, $a = 3$, and $M = 5$. Fig. 5(b) examines an environment with $K = 6$, $N = 30$, SNR $= 15$ dB, $a = 3$, and $M = 5$. Finally, Fig. 5(c) considers $K = 10$, $N = 50$, SNR $=$ 15 dB, $a = 2$, and $M = 7$. It is clear that both algorithms provide comparable performance.

For the limited number of channels we have simulated, we see slightly different asymptotic MSE for the two methods. Both approaches can provide the superior performance, depending on the channel realization. The classes of channels for which one method outperforms the other one are not yet known. The theoretical MSE approximations in (33) and (58) for the TOEP method facilitate investigating the performance of this method without resorting to Monte Carlo simulations, and a similar MSE analysis of the LZ method would facilitate a more detailed performance comparison of the two methods.

We next consider a near–far scenario. The near–far ratio (NFR) is defined to be $10\log(A_1/A_k)$, where A_1 is the received amplitude of the desired user, and A_k is the received amplitude for all interfering users. In the previous simulations, all users transmitted at the same power. Now, we consider the situation where the desired user is at fixed power and the interfering users all share the same power, which is varied. We evaluate the system in an environment with $K = 8$, $N = 30$, SNR $= 15$ dB, $a = 3$, $M = 5$ with the channel shown in Table I, $N_r = 50$ Monte Carlo runs, and $N_s = 200$ transmitted symbols. In Fig. 6, we see previously unreported results. That is, both algorithms are not near–far resistant. In theory, both algorithms should be near–far resistant. It is conjectured that as the near–far ratio increases, increasingly more samples are necessary to provide high-fidelity estimates of the relevant covariance matrices, and hence, for fixed N_s, performance degrades as the near–far ratio increases.

It should be noted that it was anticipated that the TOEP algorithm would provide superior performance to that of the LZ algorithm. The reasoning was that the matched filtering preprocessing would provide sufficient multiple-access interference suppression and that the Toeplitz displacement would significantly remove any remaining interference. However, it appears that due to the method of computation of the covariance matrix prior to matched filtering in [8], comparable performance is achievable. The maximum amount of averaging is possible for the covariance matrix prior to matched filtering, which is able to truly take advantage of the statistics of the randomized codes. Performance improvements for the TOEP algorithm are possible if improved methods of calculating the matched filtered data covariance matrix can be found. Any new averaging method would also benefit the LZ algorithm.

Finally, we construct a RAKE receiver using the estimated and true channels. Fig. 7 plots the bit error rate (BER) versus the number of users for a scenario, where $a = 5$, SNR $= 10$ dB, $N = 30$, $M = 5$, and the channel impulse responses employed are those noted in Table I. We can see that the curves for the TOEP and LZ algorithms are almost coincident. Furthermore, they are very close to the curve for perfect knowledge of the channel. It can be seen that for a large number of users, the RAKE receivers exhibit poor performance. This is due to the fact that the multiple access interference is not explicitly suppressed.

VII. Conclusions

In this paper, we have proposed and analyzed a blind channel identification scheme appropriate for use with randomized DS/CDMA systems. In such systems, the spreading sequences assigned to each user have periods that are much longer than the symbol duration. As a result, cross-correlation functions are time varying. To overcome this challenge, the proposed identification algorithm exploits the asymptotic statistics of such spreading sequences. An approximation to the mean-squared channel estimation error is derived for the new identification scheme. The analytical approximation is observed to be quite tight via simulation study. In addition, the proposed algorithm is compared with an existing algorithm proposed in [8]. The simulated performance of the two algorithms are comparable. More complete comparison of the two algorithms will be

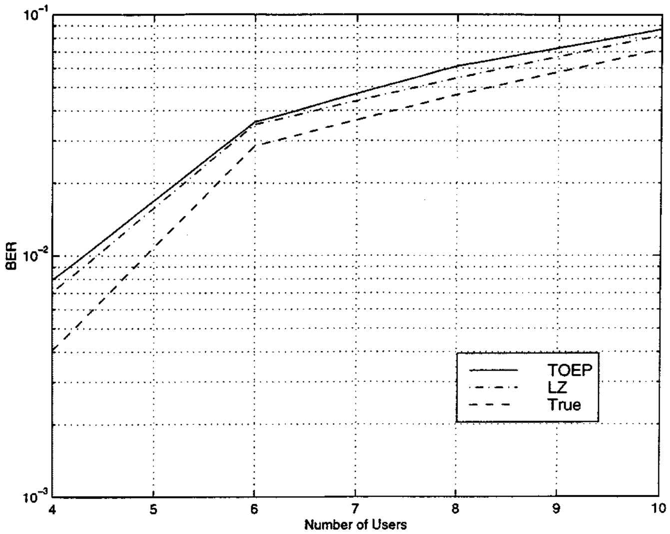

Fig. 7. Bit error rate versus the number of users.

possible when an analytical expression for the mean-squared error for the algorithm in [8] is derived. Future work will focus on designing improved algorithms to estimate the data covariance matrix.

APPENDIX
MEAN-SQUARED ERROR ANALYSIS

In this Appendix, we complete the derivation of the mean-squared error that was begun in Section IV. We calculate $\mathbf{E}[\Delta h(k)\Delta h^*(k)]$, given the following expression for $\Delta h(k)$:

$$\Delta h(k) = \text{Trace}\{\mathbf{Q}_k^H \mathbf{P}\hat{\mathbf{R}}_h \mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{V}^H \mathcal{H}\}. \tag{40}$$

Taking into account that[3] $\hat{\mathbf{R}}_h = \hat{\mathbf{R}}_y^+ - \hat{\mathbf{R}}_y^-$, (33) can be expressed as follows:

$$\begin{aligned}\Delta h(k) &= \text{Trace}\{\mathbf{Q}_k^H \mathbf{P}(\hat{\mathbf{R}}_y^+ - \hat{\mathbf{R}}_y^-)\mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{V}^H \mathcal{H}\} \\ &= \sum_{i=1}^{2(a+2)} \sum_{\beta=-,+} \beta \mathbf{q}_k^H \mathbf{D}_i^H \mathbf{P}\hat{\mathbf{R}}_y^\beta \mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{V}^H \mathbf{D}_i \mathbf{h} \\ &= \sum_{i=1}^{2(a+2)} \sum_{\beta=-,+} \beta \mathbf{u}_{ki}^H \hat{\mathbf{R}}_y^\beta \mathbf{g}_i \end{aligned} \tag{41}$$

where the super-index β, with $\beta = -, +$, denotes the type of displacement in the matrix $\hat{\mathbf{R}}_y$; $\mathbf{u}_{ki} = (\mathbf{q}_k^H \mathbf{D}_i^H \mathbf{P})^H$ and $\mathbf{g}_i = \mathbf{V}\mathbf{\Lambda}^{-1}\mathbf{V}^H \mathbf{D}_i \mathbf{h}$. The mean-squared error (MSE) for the kth element is

$$\mathbf{E}[\Delta h(k)\Delta h^*(k)] = \sum_{i=1}^{2(a+2)} \sum_{\beta=-,+} \sum_{j=1}^{2(a+2)} \sum_{\gamma=-,+} \beta\gamma \times \mathbf{E}[\mathbf{u}_{ki}^H \hat{\mathbf{R}}_y^\beta \mathbf{g}_i \mathbf{g}_j^H \hat{\mathbf{R}}_y^\gamma \mathbf{u}_{kj}] \tag{42}$$

where $\beta\gamma$ equals $+1$ when $\beta = \gamma$ and equals -1 when $\beta \neq \gamma$. Expectation heretoforth is with respect to the random data and the additive noise. The autocorrelation matrix given β is defined as follows:

$$\hat{\mathbf{R}}_y^\beta = \frac{1}{N_s} \sum_{l=1}^{N_s} \mathbf{y}^\beta(l)\mathbf{y}^{\beta^H}(l) \tag{43}$$

where $\mathbf{y}^+(l) = \mathbf{y}(l+2{:}\ l+aM)$, $\mathbf{y}^-(l) = \mathbf{y}(l+1{:}\ l+aM-1)$, and therefore

$$\begin{aligned}&\mathbf{E}[\Delta h(k)\Delta h^*(k)] \\ &\quad = \frac{1}{N_s^2} \sum_{i=1}^{2(a+2)} \sum_{\beta=-,+} \sum_{j=1}^{2(a+2)} \sum_{\gamma=-,+} \sum_{l=1}^{N_s} \sum_{m=1}^{N_s} \beta\gamma \\ &\qquad \times \mathbf{E}[\mathbf{u}_{ki}^H \mathbf{y}^\beta(l)\mathbf{y}^{\beta^H}(l)\mathbf{g}_i \mathbf{g}_j^H \mathbf{y}^\gamma(m)\mathbf{y}^{\gamma^H}(m)\mathbf{u}_{kj}]\end{aligned} \tag{44}$$

where

$$\mathbf{y}^\beta(l) = \mathbf{S}_l^\beta \sum_{p=1}^{P+1} \mathbf{x}_p(l) = \mathbf{S}_l^\beta \sum_{p=1}^{P+1} \mathbf{x}_p(l) \tag{45}$$

and

$\mathbf{S}_l = \mathbf{S}_1(l) \Rightarrow$ Matched filter matrix for the desired user.

$\mathbf{C}_p(l) \Rightarrow$ Code matrix for pth user ($p = 1$ desired user).

$$\mathbf{x}_p(l) = \begin{cases} \mathbf{C}_p(l)\mathbf{H}_p\mathbf{b}_p(l) & p = 1, \cdots, P \\ \mathbf{w}(l) & p = P+1 \text{ (i.e., noise)} \end{cases} \tag{46}$$

where $[\mathbf{A}]^\beta$ represents the β displacement of the matrix $\mathbf{A}$. Taking into account (45), the expectation in (44) can be expressed as

$$\begin{aligned}&\mathbf{E}[\mathbf{u}_{ki}^H \mathbf{y}^\beta(l)\mathbf{y}^{\beta^H}(l)\mathbf{g}_i \mathbf{g}_j^H \mathbf{y}^\gamma(m)\mathbf{y}^{\gamma^H}(m)\mathbf{u}_{kj}] \\ &\quad = \sum_{p1=1}^{P+1} \sum_{p2=1}^{P+1} \sum_{q1=1}^{P+1} \sum_{q2=1}^{P+1} \mathbf{E}[\mathbf{u}_{ki}^H \mathbf{S}_l^\beta \mathbf{x}_{p1}(l)\mathbf{x}_{p2}^H(l)\mathbf{S}_l^{\beta^H} \mathbf{g}_i \mathbf{g}_j^H \mathbf{S}_m^\gamma \\ &\qquad \times \mathbf{x}_{q1}(m)\mathbf{x}_{q2}^H(m)\mathbf{S}_m^{\gamma^H}\mathbf{u}_{kj}].\end{aligned} \tag{47}$$

Due to the independence between users and noise, we consider the following scenarios to compute the expectation in (47):

$$\begin{cases} \text{I} & p_1 = p_2 = q_1 = q_2, \quad \text{i.e., same user (or noise)} \\ \text{II} & p_1 = p_2 \neq q_1 = q_2 \\ \text{III} & p_1 = q_2 \neq q_1 = p_2. \end{cases} \tag{48}$$

Note that for the case $p_1 = q_1 \neq q_2 = p_2$, the expectation is zero since we are using QAM modulation for the user symbols and complex Gaussian noise. For the three cases, (47) is simplified as follows:

$$\begin{aligned}&\text{I} \quad \sum_{p=1}^{P+1} \mathbf{E}[\mathbf{u}_{ki}^H \mathbf{S}_l^\beta \mathbf{x}_p(l)\mathbf{x}_p^H(l)\mathbf{S}_l^{\beta^H} \mathbf{g}_i \mathbf{g}_j^H \mathbf{S}_m^\gamma \mathbf{x}_p(m)\mathbf{x}_p^H(m) \\ &\qquad \times \mathbf{S}_m^{\gamma^H}\mathbf{u}_{kj}] \\ &\text{II} \quad \sum_{p=1}^{P+1} \sum_{\substack{q=1 \\ p\neq q}}^{P+1} \mathbf{E}[\mathbf{u}_{ki}^H \mathbf{S}_l^\beta \mathbf{x}_p(l)\mathbf{x}_p^H(l)\mathbf{S}_l^{\beta^H} \mathbf{g}_i \mathbf{g}_j^H \mathbf{S}_m^\gamma \mathbf{x}_q(m) \\ &\qquad \times \mathbf{x}_q^H(m)\mathbf{S}_m^{\gamma^H}\mathbf{u}_{kj}] \\ &\text{III} \quad \sum_{p=1}^{P+1} \sum_{\substack{q=1 \\ p\neq q}}^{P+1} \mathbf{E}[\mathbf{u}_{ki}^H \mathbf{S}_l^\beta \mathbf{x}_p(l)\mathbf{x}_p^H(m)\mathbf{S}_m^{\gamma^H}\mathbf{u}_{kj}\mathbf{g}_j^H \mathbf{S}_m^\gamma \mathbf{x}_q(m) \\ &\qquad \times \mathbf{x}_q^H(l)\mathbf{S}_l^{\beta^H}\mathbf{g}_i].\end{aligned} \tag{49}$$

[3]Recall that $\hat{\mathbf{R}}_y^+ = \hat{\mathbf{R}}_y(2{:}\ aM, 2{:}\ aM)$ and $\hat{\mathbf{R}}_y^- = \hat{\mathbf{R}}_y(1{:}\ aM-1, 1{:}\ aM-1)$.

In (49), the analysis of cases II and III is obvious, but to develop case I, we consider the general expectation $\mathbf{E}[\mathbf{c}_l^H \mathbf{z}(l)\mathbf{z}^H(l)\mathbf{d}_l\mathbf{d}_m^H\mathbf{z}(m)\mathbf{z}^H(m)\mathbf{c}_m]$, where $\mathbf{c}_n$ and $\mathbf{d}_n$ represent deterministic vectors, and $\mathbf{z}(n)$ is a random process. Equation (49-I) is considered for two different situations.

1) For $p = P+1$ (i.e., Gaussian noise), $\mathbf{c}_n$ and $\mathbf{d}_n$ are $(aN + M - 1 \times 1)$ vectors, and $\mathbf{z}(n)$ is a vector corresponding to a zero mean Gaussian process. It can be shown[4] that

$$\begin{aligned}\mathbf{E}[\mathbf{c}_l^H \mathbf{z}(l)\mathbf{z}^H(l)\mathbf{d}_l\mathbf{d}_m^H\mathbf{z}(m)\mathbf{z}^H(m)\mathbf{c}_m] &= \mathbf{E}[\mathbf{c}_l^H \mathbf{z}(l)\mathbf{z}^H(l)\mathbf{d}_l]\mathbf{E}[\mathbf{d}_m^H\mathbf{z}(m)\mathbf{z}^H(m)\mathbf{c}_m] \\ &\quad + \mathbf{E}[\mathbf{c}_l^H \mathbf{z}(l)\mathbf{z}^H(m)\mathbf{c}_m]\mathbf{E}[\mathbf{d}_m^H\mathbf{z}(m)\mathbf{z}^H(l)\mathbf{d}_l] \\ &= \sigma_w^4(\mathbf{c}_l^H\mathbf{d}_l\mathbf{d}_m^H\mathbf{c}_m + \mathbf{c}_l^H\mathbf{J}_{N(m-l)}^{aN+M-1}\mathbf{c}_m\mathbf{d}_m^H\mathbf{J}_{N(l-m)}^{aN+M-1}\mathbf{d}_l)\end{aligned} \tag{50}$$

where σ_w^2 is the noise variance, and

$$\begin{aligned}\mathbf{J}_k^M &= M \times M \text{ matrix whose element } (i,\, j) \\ &= \begin{cases} 1, & \text{if } (i-j) = k \text{ and } |k| \le M-1 \\ 0, & \text{otherwise.}\end{cases}\end{aligned} \tag{51}$$

2) For $p = 1, \cdots, P$ (i.e., users), $\mathbf{c}_n$ and $\mathbf{d}_n$ are $((a+2)M \times 1)$ vectors, and $\mathbf{z}(n)$ is a vector whose elements are the QAM transmitted symbols for a single user. It can be shown that

$$\begin{aligned}\mathbf{E}[\mathbf{c}_l^H \mathbf{z}(l)\mathbf{z}^H(l)\mathbf{d}_l\mathbf{d}_m^H\mathbf{z}(m)\mathbf{z}^H(m)\mathbf{c}_m] &= \mathbf{E}[\mathbf{c}_l^H \mathbf{z}(l)\mathbf{z}^H(l)\mathbf{d}_l]\mathbf{E}[\mathbf{d}_m^H\mathbf{z}(m)\mathbf{z}^H(m)\mathbf{c}_m] \\ &\quad + \mathbf{E}[\mathbf{c}_l^H \mathbf{z}(l)\mathbf{z}^H(m)\mathbf{c}_m]\mathbf{E}[\mathbf{d}_m^H\mathbf{z}(m)\mathbf{z}^H(l)\mathbf{d}_l] \\ &\quad - \sum_{i=0}^{(a+2)M-1}\sum_{j=0}^{(a+2)M-1} c_l^*(i)d_l(i)d_m^*(j)c_m(j) \\ &\quad \times \text{`}\mathbf{E}[z(l+i)z^*(m+j)] \\ &= \sigma_b^4(\mathbf{c}_l^H\mathbf{d}_l\mathbf{d}_m^H\mathbf{c}_m + \mathbf{c}_l^H\mathbf{J}_{m-l}^{(a+2)M}\mathbf{c}_m\mathbf{d}_m^H\mathbf{J}_{l-m}^{(a+2)M}\mathbf{d}_l \\ &\qquad - \rho_l^H\mathbf{J}_{m-l}^{(a+2)M}\rho_m)\end{aligned} \tag{52}$$

where σ_b^2 is the user power, $c_l(i),\, d_l(i),\, z(l+i)$ are the ith elements of $\mathbf{c}_l,\, \mathbf{d}_l,\, \mathbf{z}(l)$, respectively, and

$$\begin{aligned}\boldsymbol{\rho}_n &= [c_n(0)d_n^*(0),\, c_n(1)d_n^*(1),\, \cdots \\ &\qquad c_n((a+2)M-1)d_n^*((a+2)M-1)]^T \\ &= \operatorname{diag}(\mathbf{c}_n\mathbf{d}_n^H)\end{aligned} \tag{53}$$

where the operator diag$(\mathbf{A})$ results in a vector of the diagonal elements of the matrix $\mathbf{A}$.

Taking into account (49), (50), and (52), (47) can be expressed as follows:

$$\begin{aligned}&\mathbf{E}[\mathbf{u}_{ki}^H\mathbf{y}^\beta(l)\mathbf{y}^{\beta^H}(l)\mathbf{g}_i\mathbf{g}_j^H\mathbf{y}^\gamma(m)\mathbf{y}^{\gamma^H}(m)\mathbf{u}_{kj}] \\ &= \sum_{p=1}^{P+1}\sum_{q=1}^{P+1}(\mathbf{u}_{ki}^H\mathbf{S}_l^\beta\boldsymbol{\Gamma}_p(l,\, m)\mathbf{S}_m^{\gamma^H}\mathbf{u}_{kj}\mathbf{g}_j^H\mathbf{S}_m^\gamma\boldsymbol{\Gamma}_q(m,\, l)\mathbf{S}_l^{\beta^H}\mathbf{g}_i \\ &\quad + \mathbf{u}_{ki}^H\mathbf{S}_l^\beta\boldsymbol{\Gamma}_p(l,\, l)\mathbf{S}_l^{\beta^H}\mathbf{g}_i\mathbf{g}_j^H\mathbf{S}_m^\gamma\boldsymbol{\Gamma}_q(m,\, m)\mathbf{S}_m^{\gamma^H}\mathbf{u}_{kj}) \\ &\quad - \sum_{p=1}^{P}\boldsymbol{\Phi}_p^i(l,\, l)\mathbf{J}_{m-l}^{(a+2)M}\boldsymbol{\Phi}_p^j(m,\, m)\end{aligned} \tag{54}$$

where

$$\boldsymbol{\Gamma}_p(l,\, m) = \begin{cases}\sigma_p^2\mathbf{C}_p(l)\mathbf{H}_p\mathbf{J}_{m-l}^{(a+2)M}\mathbf{H}_p^H\mathbf{C}_p^H(m) \\ \quad p = 1, \cdots, P \\ \sigma_w^2\mathbf{J}_{N(m-l)}^{(aN+M-1} \\ \quad p = P+1\end{cases}$$

$$\boldsymbol{\Phi}_p^i(l,\, l) = \operatorname{diag}(\mathbf{H}_p^H\mathbf{C}_p^H(l)\mathbf{S}_l^{\beta^H}\mathbf{u}_{ki}\mathbf{g}_i^H\mathbf{S}_l^\beta\mathbf{C}_p(l)\mathbf{H}_p). \tag{55}$$

Substituting (54) into (44), we obtain

$$\begin{aligned}&\mathbf{E}[\Delta h(k)\Delta h^*(k)] \\ &= \frac{1}{N_s^2}\sum_{\beta=-,+}\sum_{\gamma=-,+}\sum_{l=1}^{N_s}\sum_{m=1}^{N_s}\beta\gamma \\ &\quad \times \text{Trace}\left\{\mathbf{U}_k^H\mathbf{S}_l^\beta\left(\sum_{p=1}^{P+1}\boldsymbol{\Gamma}_p(l,\, m)\right)\mathbf{S}_m^{\gamma^H}\mathbf{U}_k\mathbf{G}^H\mathbf{S}_m^\gamma\right. \\ &\quad \times \left.\left(\sum_{q=1}^{P+1}\boldsymbol{\Gamma}_q(m,\, l)\right)\mathbf{S}_l^{\beta^H}\mathbf{G}\right\} \\ &\quad + \text{Trace}\left\{\mathbf{U}_k^H\mathbf{S}_l^\beta\left(\sum_{p=1}^{P+1}\boldsymbol{\Gamma}_p(l,\, l)\right)\mathbf{S}_l^{\beta^H}\mathbf{G}\right\} \\ &\quad \times \text{Trace}\left\{\mathbf{G}^H\mathbf{S}_m^\gamma\left(\sum_{q=1}^{P+1}\boldsymbol{\Gamma}_q(m,\, m)\right)\mathbf{S}_m^{\gamma^H}\mathbf{U}_k\right\} \\ &\quad - \sum_{p=1}^{P}\boldsymbol{\Phi}_p(l,\, l)\mathbf{J}_{m-l}^{(a+2)M}\boldsymbol{\Phi}_p(m,\, m) \\ &= \frac{1}{N_s^2}\sum_{\beta=-,+}\sum_{\gamma=-,+}\beta\gamma\left\{\boldsymbol{\Psi}_T^\beta\boldsymbol{\Psi}_T^\gamma + \sum_{l=1}^{N_s}\sum_{\substack{m=l-(a+1)\\1\le m\le N_s}}^{l+(a+1)}\right. \\ &\quad \times \text{Trace}\{\mathbf{U}_k^H\mathbf{S}_l^\beta\mathbf{X}_{lm}\mathbf{S}_m^{\gamma^H}\mathbf{U}_k\mathbf{G}^H\mathbf{S}_m^\gamma\mathbf{X}_{lm}^H\mathbf{S}_l^{\beta^H}\mathbf{G}\} \\ &\quad \left. - \sum_{p=1}^{P}\boldsymbol{\Phi}_p(l,\, l)\mathbf{J}_{m-l}^{(a+2)M}\boldsymbol{\Phi}_p(m,\, m)\right\}\end{aligned} \tag{56}$$

where

$$\begin{aligned}\boldsymbol{\Phi}_p(l,\, l) &= \sum_{i=1}^{2(a+2)}\boldsymbol{\Phi}_p^i(l,\, l) \\ &= \operatorname{diag}(\mathbf{H}_p^H\mathbf{C}_p^H(l)\mathbf{S}_l^{\beta^H}\mathbf{U}_k\mathbf{G}^H\mathbf{S}_l^\beta\mathbf{C}_p(l)\mathbf{H}_p) \\ \mathbf{U}_k &= [\mathbf{u}_{k,\,1},\, \cdots,\, \mathbf{u}_{k,\,2(a+2)}] \\ \mathbf{G} &= [\mathbf{g}_1,\, \cdots,\, \mathbf{g}_{2(a+2)}] \\ \boldsymbol{\Psi}_T^\beta &= \sum_{l=1}^{N_s}\text{Trace}\left\{\mathbf{U}_k^H\mathbf{S}_l^\beta\right. \\ &\qquad \times \left.\left(\sum_{p=1}^{P}\sigma_p^2\mathbf{C}_p(l)\mathbf{H}_p\mathbf{H}_p^H\mathbf{C}_p^H(l) + \sigma_w^2\mathbf{I}\right)\mathbf{S}_l^{\beta^H}\mathbf{G}\right\} \\ \mathbf{X}_{lm} &= \left(\sum_{p=1}^{P}\sigma_p^2\mathbf{C}_p(l)\mathbf{H}_p\mathbf{J}_{m-l}^{(a+2)M}\mathbf{H}_p^H\mathbf{C}_p^H(m)\right) \\ &\quad + \sigma_w^2\mathbf{J}_{N(m-l)}^{aN+M-1}.\end{aligned} \tag{57}$$

We have used the fact that $\mathbf{J}_{m-l}^{(a+2)M}$ and $\mathbf{J}_{N(m-l)}^{aN+M-1}$ are zero matrices for $|m-l| > a+1$ and $N|m-l| > aN+M$ (i.e.,

[4]If $z_1,\, z_2,\, z_3,\, z_4$ are four zero mean Gaussian random variables, then $\mathbf{E}[z_1z_2^*z_3z_4^*] = \mathbf{E}[z_1z_2^*]\mathbf{E}[z_3z_4^*] + \mathbf{E}[z_1z_4^*]\mathbf{E}[z_3z_2^*]$.

$|m - l| > a$), respectively. Therefore, only the values $m = l - (a+1), \cdots, l + (a+1)$ are considered.

Finally, to obtain the MSE of the channel estimation algorithm, we have to consider the contribution of all of the components, that is

$$\mathbf{E}[\Delta \mathbf{h}^H \Delta \mathbf{h}] = \sum_{k=1}^{M} \mathbf{E}[\Delta h(k) \Delta h^*(k)]. \tag{58}$$

In Section VI, the accuracy of this approximation is explored via comparison to simulation results.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewer whose comments improved the clarity of the paper.

REFERENCES

[1] F. Adachi, M. Sawahashi, and H. Suda, "Wideband DS-CDMA for next-generation mobile communication systems," *IEEE Commun. Mag.*, vol. 36, pp. 56–69, Sept. 1998.

[2] E. Aktas and U. Mitra, "Reduced complexity subspace based blind identification for DS/CDMA systems," *IEEE Trans. Commun.*, to be published.

[3] E. Dahlman, B. Gudmundson, M. Nilsson, and J. Sköld , "UMTS/IMT-2000 based on wideband CDMA," *IEEE Commun. Mag.*, vol. 36, pp. 70–81, Sept. 1998.

[4] E. De Carvalho and D. T. M. Slock, "Cramer–Rao bounds for semi-blind, blind and training sequence based channel estimation," in *Proc. IEEE Signal Process. Adv. Wireless Commun. Workshop*, Paris, Apr. 16-18, 1997, pp. 129–132.

[5] A. Klein, G. K. Kaleh, and P. W. Baier, "Zero forcing and minimum mean-square-error equalization for multiuser detection in code-division multiple-access channels," *IEEE Trans. Veh. Technol.*, vol. 45, pp. 21–42, May 1996.

[6] K. Li and H. Liu, "Channel estimation for DS-CDMA with aperiodic spreading codes," in *Proc. ICASSP*, vol. 5, Mar. 1999, pp. 2535–2538.

[7] H. Liu and G. Xu, "Closed-Form blind symbol estimation in digital communications," *IEEE Trans. Signal Processing*, vol. 43, pp. 2714–2723, Nov. 1995.

[8] H. Liu and M. Zoltowski, "Blind equalization in antenna array CDMA systems," *IEEE Trans. Signal Processing*, vol. 45, pp. 161–172, Jan. 1997.

[9] U. Mitra, D. Slock, and C. Escudero, "Blind equalization techniques for CDMA communications," in *Proc. 31st Annu. Conf. Inform. Sci. Syst.*. Baltimore, MD, Mar. 1997, pp. 696–701.

[10] 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.

[11] W. Qiu and Y. Hua, "Performance analysis of the subspace method for blind channel identification," *Signal Process.*, no. 50, pp. 71–81, 1996.

[12] P. Stoica and T. Söderström, "Statistical analysis and subspace rotation estimates of sinusoidal frequencies," *IEEE Trans. Signal Processing*, vol. 39, pp. 1836–1847, Aug. 1991.

[13] E. Ström , S. Parkvall, S. Miller, and B. Ottersten, "Propagation delay estimation in asynchronous direct-sequence code-division multiple access systems," *IEEE Trans. Commun.*, vol. 44, pp. 84–93, Jan. 1996.

[14] S. Talwar, M. Viberg, and A. Paulraj, "Blind estimation of multiple co-channel digital signals using an antenna array," *IEEE Signal Processing Lett.*, vol. 1, pp. 29–41, Feb. 1994.

[15] M. Torlak and G. Xu, "Blind multi-user channel estimation in asynchronous CDMA systems," *IEEE Trans. Signal Processing*, vol. 45, pp. 137–147, Jan. 1997.

[16] M. Torlak, B. Evans, and G. Xu, "Blind estimation of FIR channels in CDMA systems with aperiodic spreading sequences," in *Proc. 31st Asilomar Conf.*, Oct. 1997, pp. 495–499.

[17] M. K. Tsatsanis and G. B. Giannakis, "Blind estimation of direct sequence spread spectrum signals in multipath," *IEEE Trans. Signal Processing*, vol. 45, pp. 1241–1252, May 1997.

[18] Z. Xu and M. Tsatsanis, "Blind channel estimation for multiuser CDMA systems with long spreading codes," in *Proc. ICASSP*, vol. 5, Mar. 1999, pp. 2531–2534.

[19] A. Weiss and B. Friedlander, "Channel estimation for DS-CDMA downlink with aperiodic spreading codes," in *Proc. 31st Asilomar Conf.*, Oct. 1997.

[20] A. Mantravadi and V. Veeravalli, "Multiple-access interference resistant acquisition for CDMA systems with long spreading codes," in *Proc. 32nd Annu. Conf. Inform. Sci. Syst.*. Princeton, NJ, Mar. 1998, pp. 141–146.

[21] Z. Zvonar and D. Brady, "Linear multipath-decorrelating receivers for CDMA frequency-selective fading channels," *IEEE Trans. Commun.*, vol. 44, pp. 650–653, June 1996.

[22] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Reading, MA: Addison-Wesley, 1991.

Carlos J. Escudero (M'99) was born in La Coruña, Spain, in 1968. He received the B.S. degree in telecommunications from the University of Vigo, Pontevedra, Spain, in 1991 and the M.S. and Ph.D. degrees in electrical engineering from the University of La Coruña in 1994 and 1998, respectively.

He joined the faculty of the University of La Coruña in 1992, where he is now an Associate Professor in the Department of Electrónica y Sistemas. He made two research visits to the Information Processing Systems Laboratory, Department of Electrical Engineering, The Ohio State University, Columbus, in 1996 and 1998, respectively. His main research interests include multiuser communications, channel identification, and equalization.

Urbashi Mitra (S'90–M'94) received the B.S. degree (with high honors) and the M.S. degree from the University of California, Berkeley in 1987 and 1989, respectively, both in electrical engineering and computer science. She received the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, in 1994.

Prior to beginning work on the M.S. degree, she was a Research Assistant with the Technical University of Tampere, Tampere, Finland, during the Summer of 1987. From 1989 to 1990, she worked as a Member of Technical Staff at Bellcore, Red Bank, NJ, where she wrote SS7 switch generic requirements and actively participated in T1S1 standards. She is currently an Assistant Professor with the Department of Electrical Engineering, The Ohio State University, Columbus. On several occasions from 1994 to 1997, she was a Visiting Scholar with the Mobile Communications Group of the Institut EURÉCOM, Sophia Antipolis, France, where she considered blind equalization schemes for multiuser detectors. Her current research interests include multiuser detection theory, code-division multiple-access communications for personal wireless and mobile applications, nonparametric and robust detection, equalization techniques, and wireless networking.

Dr. Mitra was a recipient of an NSF CAREER award in 1996 and a MacQuigg teaching award for the College of Engineering, The Ohio State University, in 1997.

Dirk T. M. Slock (M'89) received the B.E. degree in electrical engineering from the University of Gent, Gent, Belgium in 1982. In 1984, he was awarded a Fulbright scholarship and went to Stanford University, Stanford, CA, where he received the M.S. degree in electrical engineering, the M.S. degree in statistics, and the Ph.D. degree in electrical engineering in 1986, 1989, and 1989, respectively.

While at Stanford, he developed new fast recursive least-squares (RLS) algorithms for adaptive filtering. In 1989, he joined the Philips Research Laboratory, Heverlee, Belgium. He investigated the relationship between fast lattice and fast QR RLS algorithms, introduced fractionally spaced subband adaptive filtering, proposed an optimal choice for the wavelet in wavelet transform-based image coding, and introduced the notion of backward consistency in the round-off error analysis of recursive least-squares algorithms. He is now with the Mobile Communications Group, Eurecom Institute, Sophia Antipolis, France, where he teaches statistical signal processing, cellular system design, and speech coding for mobile communications. His research interests include digital signal processing for mobile communications [antenna arrays for (blind) equalization and spatial division multiple access] and adaptation techniques for acoustic echo cancellation in teleconferencing and handsfree telephone applications.

Dr. Slock received one best paper award from the IEEE and one from the EURASIP in 1992. He is the Co-Author of two IEEE-Globecom'98 best student paper awards. He was an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1994 to 1996.