

BLIND CHANNEL ESTIMATION FOR MULTIUSER CDMA SYSTEMS WITH LONG SPREADING CODES¹

Zhengyuan (Daniel) Xu

Michail K. Tsatsanis

Electrical & Computer Engr. Dept.
Stevens Institute of Technology
Castle Point on Hudson
Hoboken, NJ 07030

e-mail: mtsatsan@stevens-tech.edu

ABSTRACT

In this paper, correlation-matching techniques are employed to estimate multipath channel parameters for a multiuser CDMA system with long spreading codes. For given code sequences, the output correlation matrix (parametrized by the unknown channel coefficients) is compared with its instantaneous approximation. By minimizing the Frobenius norm of the resulting error matrix the channel parameters can be estimated up to a scalar ambiguity. Under the assumption of i.i.d. code sequences, identifiability for each channel is guaranteed and the asymptotic convergence of the proposed algorithm is established. Simulation results confirm our claims. Comparisons with other methods are also provided.

1. INTRODUCTION

DS-CDMA technology has several advantages over competing TDMA/FDMA alternatives and is a serious candidate for the next generation wireless networks. Agreement however has not been reached yet on whether future systems will employ short codes (which repeat at every bit) or long codes whose period spans a very large number of bits. Many current systems (e.g. IS-95) employ long codes [4].

From a signal processing viewpoint, the short code case is more tractable since the interference pattern does not change from bit to bit. It is not surprising that most of the past research effort has focused on the short code case (e.g. [1], [5]). Unfortunately, long spreading codes introduce time-varying user signatures rendering classical adaptive multiuser algorithms impractical [5].

Recent research efforts however have focused on the long code case and have contributed to the design of new blind receivers suitable for such systems [8], [10]. In these approaches, the channel parameters are first estimated and receivers can then be constructed. We focus on channel estimation problems in this paper by employing correlation matching techniques which have been extensively analyzed and applied to a multitude of problems in blind identification [7], detection [3] and channel estimation [2]. Extensions to the time-varying systems have been reported in [9].

In the current setup the users' signatures are time-varying due to the changing codes at every bit. However the multipath parameters are kept constant. We therefore can match the output covariance matrix (parameterized by these unknown channel parameters) with instantaneous approxima-

tions based on the received data. By minimizing the resulting error, closed form solutions of channel vectors within a scalar ambiguity are obtained. Moreover, their asymptotic performance is studied.

It is established that our estimates for all corresponding channels strongly converge to their true parameters. Based on simulation results, the proposed method shows better performance compared with subspace based approach [10] for a heavily loaded system.

2. PROBLEM STATEMENT

In a DS-CDMA communication system with M users, let us assume user j ($j = 1, \dots, M$) transmits a zero-mean, i.i.d. bit sequence $w_j(n)$ with variance $\sigma_{w_j}^2 = E\{\|w_j(n)\|^2\}$. Every bit is spread by an independently assigned i.i.d. code sequence. Let us define $c_{j,k}(n)$, $n = 1, \dots, P$, to be the spreading code of user j , bit k , with P chips (n is the chip index). Then at the chip-rate receiver, the output signal contributed by user j is (see Fig. 1)

$$y_j(n) = \sum_{k=-\infty}^{\infty} w_j(k) h_{j,k}(n - kP) \quad (1)$$

$$h_{j,k}(n) = \sum_{m=-\infty}^{\infty} g_j(m) c_{j,k}(n - m - \delta_j) \quad (2)$$

where $h_{j,k}(n)$ is the signature of user j for bit k with chip index n , $0 \leq \delta_j < P$ is the delay in chip periods, $g_j(n)$ is the discrete-time equivalent channel impulse response which includes the transmitter and receiver filters. The overall received signal $y(n)$ is then a superposition of signals from all M users corrupted by AWGN $v(n)$

$$y(n) = \sum_{j=1}^M y_j(n) + v(n) \quad (3)$$

where $v(n)$ has zero-mean and variance $\sigma_v^2 = E\{\|v(n)\|^2\}$. In practice, communication channels are usually modeled as having finite impulse response. In the sequel the maximum order for all multipath channels $g_j(n)$ is assumed to be q .

To obtain a compact form of our model in an observation interval, let us collect $P + q$ samples of $y(n)$ in a vector $\mathbf{y}(n) = [y(nP + 1), \dots, y(nP + P + q)]^T$. Then from (1), (2) and (3), the received signal $\mathbf{y}(n)$ becomes

$$\mathbf{y}(n) = \sum_{j=1}^M [\mathbf{h}_j w_j(n) + \tilde{\mathbf{h}}_j w_j(n-1) + \tilde{\tilde{\mathbf{h}}}_j w_j(n+1)] + \mathbf{v}(n) \quad (4)$$

where $\mathbf{h}_j = [0, \dots, 0, h_{j,n}(1), \dots, h_{j,n}(P + q - \delta_j)]^T$, $\tilde{\mathbf{h}}_j = [h_{j,n-1}(P + 1 - \delta_j), \dots, h_{j,n-1}(P + q), 0, \dots, 0]^T$ and $\tilde{\tilde{\mathbf{h}}}_j = [0, \dots, 0, h_{j,n+1}(1), \dots, h_{j,n+1}(q - \delta_j)]^T$ are signatures of

¹This work was supported in part by the National Science Foundation Grant NSF-NCR 9706658, NSF-CCR 9733048 and the Army Research Office Grant DAAG55-98-1-0224.

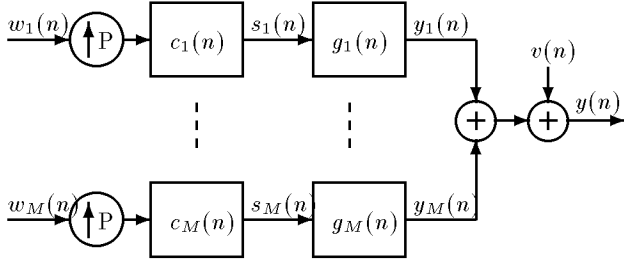


Figure 1. CDMA system with long spreading codes

$w_j(n)$, $w_j(n-1)$ and $w_j(n+1)$ at time n respectively¹. According to (2), these signatures can be expressed as

$$\mathbf{h}_j = \mathbf{C}_{j,1}(n)\mathbf{g}_j, \quad \tilde{\mathbf{h}}_j = \mathbf{C}_{j,2}(n-1)\mathbf{g}_j, \quad \tilde{\tilde{\mathbf{h}}}_j = \mathbf{C}_{j,3}(n+1)\mathbf{g}_j, \quad (5)$$

where $\mathbf{g}_j = [g_j(0), \dots, g_j(q)]^T$ is the channel vector of user j , the code filtering matrices are defined by²

$$\mathbf{C}_j(n) = \begin{bmatrix} c_{j,n}(1) & & \mathbf{0} \\ \vdots & \ddots & c_{j,n}(1) \\ c_{j,n}(P) & & \vdots \\ \mathbf{0} & \ddots & c_{j,n}(P) \end{bmatrix}, \quad \mathbf{C}_{j,1}(n) = \begin{bmatrix} \mathbf{0} \\ \mathbf{C}_j(n)(1 : P+q-\delta_j, :) \end{bmatrix}, \quad \mathbf{C}_{j,2}(n) = \begin{bmatrix} \mathbf{C}_j(n)(P+1-\delta_j : P+q, :) \\ \mathbf{0} \end{bmatrix}, \quad \mathbf{C}_{j,3}(n) = \begin{bmatrix} \mathbf{0} \\ \mathbf{C}_j(n)(1 : q-\delta_j, :) \end{bmatrix}, \quad (6)$$

and notation $\mathbf{X}(l_1 : l_2, :)$ from Matlab is used to take out those rows from l_1 to l_2 of matrix $\mathbf{X}$. By substituting (5) in (4), the output becomes

$$\mathbf{y}(n) = \sum_{j=1}^M [\mathbf{C}_{j,1}\mathbf{g}_j w_j(n) + \mathbf{C}_{j,2}\mathbf{g}_j w_j(n-1) + \mathbf{C}_{j,3}\mathbf{g}_j w_j(n+1)] + \mathbf{v}(n) \quad (7)$$

Let us define the correlation matrix of this received data at time n as $\mathbf{R}_y(n) = \mathbb{E}\{\mathbf{y}(n)\mathbf{y}^H(n) | \mathbf{C}_{j,j=1,\dots,M}\}$ conditioned on the given code sequences for M users, where superscript “ H ” denotes complex conjugate transpose (Hermitian). Then from (7), $\mathbf{R}_y(n)$ can be written as

$$\mathbf{R}_y(n) = \sum_{j=1}^M \tilde{\mathbf{T}}_j + \sigma_v^2 \mathbf{I} \quad (8)$$

$$\tilde{\mathbf{T}}_j = \sigma_{w_j}^2 [\mathbf{C}_{j,1}\mathbf{g}_j\mathbf{g}_j^H \mathbf{C}_{j,1}^H + \mathbf{C}_{j,2}\mathbf{g}_j\mathbf{g}_j^H \mathbf{C}_{j,2}^H + \mathbf{C}_{j,3}\mathbf{g}_j\mathbf{g}_j^H \mathbf{C}_{j,3}^H]$$

It is clear that $\mathbf{R}_y(n)$ is parametrized by constants σ_v^2 and $\sigma_{w_j}\mathbf{g}_j$ under the assumption that $\mathbf{C}_j$ and delays δ_j are available³. Notice that due to the time-varying feature of code sequences, $\mathbf{R}_y(n)$ is also time-varying. Our current problem is how to retrieve the time-invariant parameters $\mathbf{g}_j$ based on the knowledge of $\mathbf{R}_y(n)$. We will focus on estimating all $\mathbf{g}_j$ next based on correlation matching techniques.

¹All signatures $\mathbf{h}_j$, $\tilde{\mathbf{h}}_j$ and $\tilde{\tilde{\mathbf{h}}}_j$ change with time due to the time-varying spreading codes. To simplify the notation, we drop their time indices from now on.

²If $\delta_j = 0$ then $\mathbf{C}_{j,1}(n) = \mathbf{C}_j(n)$.

³The code acquisition problem, i.e., the estimation of the delays δ_j is beyond the scope of this paper.

3. BLIND MULTIUSER CHANNEL ESTIMATION

Let $\hat{\mathbf{R}}_y(n)$ denote some estimator of $\mathbf{R}_y(n)$ and let us build our cost function as

$$J = \frac{1}{N} \sum_{n=1}^N \|\mathbf{R}_y(n) - \hat{\mathbf{R}}_y(n)\|_F^2 \quad (9)$$

where N is the number of bits available. By applying the relationship between $\|\cdot\|_F$ of a matrix and its trace, it can be written as

$$J = \frac{1}{N} \sum_{n=1}^N \text{tr}\{[\mathbf{R}_y(n) - \hat{\mathbf{R}}_y(n)][\mathbf{R}_y^H(n) - \hat{\mathbf{R}}_y^H(n)]\} \quad (10)$$

By minimizing (10) with respect to σ_v^2 and $\sigma_{w_j}\mathbf{g}_j$, these unknowns could in principle be obtained.

However, two types of difficulties will impede our estimation. First, $\hat{\mathbf{R}}_y(n)$ needs to be known in (10). Due to the time-varying property of $\mathbf{R}_y(n)$, the sample average over the data record (which is usually used for a time invariant system) is not applicable here. Instead, we use instantaneous approximations $\hat{\mathbf{R}}_y(n) = \mathbf{y}(n)\mathbf{y}^H(n)$ to substitute in (10) and obtain

$$J = \frac{1}{N} \sum_{n=1}^N \text{tr}\{[\mathbf{R}_y(n) - \mathbf{y}(n)\mathbf{y}^H(n)]^2\} \quad (11)$$

where the Hermitian property of the correlation matrix is used. By minimizing J in (11) with respect to $\mathbf{g}_j$, we can obtain its estimate $\hat{\mathbf{g}}_j$. It may seem that $\hat{\mathbf{R}}_y(n)$ is very inaccurate and hence $\hat{\mathbf{g}}_j$ will not be accurate. But our cost function J employs all data points and results in surprisingly reliable estimates. This is further supported by our consistency results.

Secondly, J is the fourth order function of the unknowns $\mathbf{g}_j$, thus its high nonlinearity may lead to difficulties in estimating the channel vector $\mathbf{g}_j$. With this in mind, if we let $\mathbf{D}_j = \sigma_{w_j}^2 \mathbf{g}_j \mathbf{g}_j^H$, then (8) becomes a linear function of $\mathbf{D}_j$ and σ_v^2

$$\mathbf{R}_y(n) = \sum_{j=1}^M \mathbf{T}_j + \sigma_v^2 \mathbf{I} \quad (12)$$

with $\mathbf{T}_j = \mathbf{C}_{j,1}\mathbf{D}_j\mathbf{C}_{j,1}^H + \mathbf{C}_{j,2}\mathbf{D}_j\mathbf{C}_{j,2}^H + \mathbf{C}_{j,3}\mathbf{D}_j\mathbf{C}_{j,3}^H$. Substituting (12) in (11), our cost function becomes

$$J = \frac{1}{N} \sum_{n=1}^N \text{tr}\{[\sum_{j=1}^M \mathbf{T}_j + \sigma_v^2 \mathbf{I} - \mathbf{y}(n)\mathbf{y}^H(n)]^2\} \quad (13)$$

Thus we arrive at a quadratic function by overparametrizing the problem using $\mathbf{D}_j$ instead of $\mathbf{g}_j$. If we minimize this cost function, a unique closed form solution can be obtained. Let's first define the derivative of J with respect to a matrix $\mathbf{D}_j$ as a matrix, with (k, m) -th element equal to the derivative with respect to the (k, m) -th element of $\mathbf{D}_j$, i.e., $[\nabla_{\mathbf{D}_j} J]_{k,m} = \nabla_{d_j^{(k,m)}} J$. To find the minimum solution of (13), it is sufficient to differentiate it with respect to σ_v^2 and $\mathbf{D}_j$ respectively, and set these derivatives equal to zero. To obtain a closed form solution $\hat{\mathbf{D}}_j$, here we define an unknown vector $\hat{\mathbf{d}}_j$ which is a vector formed by stacking all columns of $\hat{\mathbf{D}}_j$ into one long vector (see [6, Ch. 12]) performed by the *vec* function $\hat{\mathbf{d}}_j = \text{vec}(\hat{\mathbf{D}}_j)$, and further define a vector $\hat{\mathbf{d}} = [(\hat{\mathbf{d}}_1)^T, \dots, (\hat{\mathbf{d}}_M)^T]^T$ which contains all our unknown parameters. Based on properties of the Kronecker product “ $\otimes$ ” (see [6, Ch. 12]), it is shown in Appendix A that the estimate $\hat{\mathbf{d}}$ satisfies the following equation

$$\mathbf{T} \hat{\mathbf{d}} = \mathbf{t} \quad (14)$$

with

$$\mathbf{T} = \frac{1}{N} \sum_{n=1}^N (\mathbf{Q}^T \mathbf{Q}) - \frac{1}{(P+q)N^2} \sum_{n=1}^N \mathbf{b}(n) \sum_{n=1}^N [\mathbf{b}(n)]^H \quad (15)$$

$$\mathbf{t} = \frac{1}{N} \sum_{n=1}^N \{\mathbf{Q}^T \text{vec}[\mathbf{y}(n) \mathbf{y}^H(n)]\} - \frac{\sum_{n=1}^N [\mathbf{y}^H(n) \mathbf{y}(n)]}{(P+q)N^2} \sum_{n=1}^N \mathbf{b}(n) \quad (16)$$

where

$$\mathbf{Q} = [\mathbf{Q}_1, \dots, \mathbf{Q}_M],$$

$$\mathbf{Q}_j = \mathbf{C}_{j,1}^* \otimes \mathbf{C}_{j,1} + \mathbf{C}_{j,2}^* \otimes \mathbf{C}_{j,2} + \mathbf{C}_{j,3}^* \otimes \mathbf{C}_{j,3} \quad (17)$$

$$\mathbf{b}(n) = \text{vec}(\mathbf{H}), \quad \mathbf{H} = [\mathbf{H}_1, \dots, \mathbf{H}_M],$$

$$\mathbf{H}_j = \mathbf{C}_{j,1}^H \mathbf{C}_{j,1} + \mathbf{C}_{j,2}^H \mathbf{C}_{j,2} + \mathbf{C}_{j,3}^H \mathbf{C}_{j,3} \quad (18)$$

“*” denotes complex conjugate, and $\mathbf{C}_{j,1}, \mathbf{C}_{j,2}, \mathbf{C}_{j,3}$ are given by (6). Notice that all code matrices depend on time, but the time index is dropped for the sake of notational convenience, therefore $\mathbf{H}$ and $\mathbf{Q}$ are also time-varying. In (14) there are $M(q+1)^2$ unknown parameters in $\hat{\mathbf{d}}$ and $M(q+1)^2$ equations. $\hat{\mathbf{d}}$ can be uniquely solved as

$$\hat{\mathbf{d}} = \mathbf{T}^{-1} \mathbf{t} \quad (19)$$

as long as the matrix $\mathbf{T}$ is nonsingular as will be discussed in the next section. According to our definitions of $\hat{\mathbf{d}}$, our estimates $\hat{\mathbf{d}}_j$ can be obtained by taking out corresponding elements of $\hat{\mathbf{d}}$. Then the reverse operation of vec function can be performed to obtain $\hat{\mathbf{D}}_j$.

Once $\hat{\mathbf{D}}_j$ is found, SVD on $\hat{\mathbf{D}}_j$ can be performed to obtain its eigenvector corresponding to the unique maximum eigenvalue, which is our estimated normalized channel vector $\frac{\hat{\mathbf{g}}_j}{\|\hat{\mathbf{g}}_j\|}$ for user j . The computational load of this SVD operation is not severe because $\mathbf{D}_j$ is a $(q+1) \times (q+1)$ small size matrix for moderate channel order q .

The proposed algorithm is batch and requires knowledge of all users' codes. Adaptive versions are possible but will be reported elsewhere. Also, modified versions are possible when only knowledge on a single user's code is available (single user receivers). They will not be reported here however, due to lack of space.

4. PERFORMANCE ANALYSIS

As is well-known, the minimization of the quadratic cost function in (13) admits a unique solution, but the question here is if this solution can guarantee consistency as $N \rightarrow \infty$. To establish the identifiability of the problem, we start from eq. (14) and show that our solution in (19) strongly converges to $\mathbf{d}$. The asymptotic result will be presented without proof due to the limited space.

To simplify our analysis, we assume that all M users are synchronous which means $\delta_j = 0$ for $j = 1, \dots, M$. Therefore $\mathbf{C}_{j,1} = \mathbf{C}_j$ according to eq. (6). Moreover, since $q \ll P$ in practice, then $\mathbf{C}_{j,2}, \mathbf{C}_{j,3}$ in (6) are only a small portion of $\mathbf{C}_j(n)$ which will be ignored next. Hence, in our eqs. (17)-(18), $\mathbf{Q}_j = \mathbf{C}_j^* \otimes \mathbf{C}_j$, $\mathbf{H}_j = \mathbf{C}_j^H \mathbf{C}_j$. Furthermore, the long spreading codes are assumed to be real for the same purpose. Then $\mathbf{C}_j^H = \mathbf{C}_j^T$ and $\mathbf{C}_j^* = \mathbf{C}_j$ as well as $\mathbf{b}^H = \mathbf{b}^T$ are valid.

Lemma 1: If all code sequences $\mathbf{c}_j(n)$ ($j = 1, \dots, M$) are assumed i.i.d. taking values from $\{+1, -1\}$ and independent of both the transmitted bits and the AWGN, then it can be shown that as $N \rightarrow \infty$, $\mathbf{T}$ and $\mathbf{t}$ in (15) and (16) converge to $\mathbf{U}$ and $\mathbf{U}\mathbf{d}$ with probability 1 respectively, that is

$$\mathbf{T} \xrightarrow{w.p.1} \mathbf{U}, \quad \mathbf{t} \xrightarrow{w.p.1} \mathbf{U}\mathbf{d}$$

where $\mathbf{U} = \tilde{\mathbf{A}} - \frac{1}{P+q} \tilde{\mathbf{b}} \tilde{\mathbf{b}}^T$, $\tilde{\mathbf{b}} = P \text{vec}([\mathbf{I}_{q+1}, \dots, \mathbf{I}_{q+1}])$ with M blocks, $\mathbf{I}_{q+1}$ is a $(q+1) \times (q+1)$ identity matrix, $\tilde{\mathbf{A}}$ has $M \times M$ blocks and $\mathbf{J}$ is of dimension $(q+1) \times (q+1)$ given by

$$\tilde{\mathbf{A}} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{B}_2 & \cdots & \mathbf{B}_2 \\ \mathbf{B}_2 & \mathbf{B}_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{B}_2 \\ \mathbf{B}_2 & \cdots & \mathbf{B}_2 & \mathbf{B}_1 \end{bmatrix}, \quad \mathbf{J} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ & & & \ddots & 1 \\ 0 & \cdots & & & 0 \end{bmatrix}$$

$$\mathbf{B}_1 = \begin{bmatrix} \mathbf{X}_0 & \mathbf{X}_1 & \cdots & \mathbf{X}_q \\ \mathbf{X}_1 & \mathbf{X}_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{X}_1 \\ \mathbf{X}_q & \cdots & \mathbf{X}_1 & \mathbf{X}_0 \end{bmatrix}, \quad \mathbf{B}_2 = \begin{bmatrix} \mathbf{Y}_0 & \mathbf{Y}_1 & \cdots & \mathbf{Y}_q \\ \mathbf{Y}_1^T & \mathbf{Y}_0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{Y}_1 \\ \mathbf{Y}_q^T & \cdots & \mathbf{Y}_1^T & \mathbf{Y}_0 \end{bmatrix}$$

$$\mathbf{X}_0 = P^2 \mathbf{I}_{q+1}, \quad \mathbf{Y}_0 = P \mathbf{I}_{q+1}$$

$\mathbf{Y}_m = (P-m) \mathbf{J}^m$, $\mathbf{X}_m = \mathbf{Y}_m + \mathbf{Y}_m^T$ for $m = 1, \dots, q$; matrix $\mathbf{X}_m$ only has non-zero elements along its m -th upper and lower diagonal and $\mathbf{Y}_m$ with non-zero elements along its m -th upper diagonal. $\square$

According to Lemma 1, identifiability in this problem is equivalent to the non-singularity of matrix $\mathbf{U}$. Notice that matrix $\mathbf{U}$ depends only on the parameters P , M and q . For a large range of possible P (e.g., up to 256), M (with $M \leq P$) and q (e.g., 1 to 20), it is true that $\mathbf{U}$ is nonsingular. We have been unable however to obtain a general proof of this conjecture.

The nonsingularity of matrix $\mathbf{U}$ indicates that $\hat{\mathbf{D}}_j$ asymptotically converges to $\mathbf{D}_j$ for all possible j ($j = 1, \dots, M$). Since the normalized channel $\frac{\hat{\mathbf{g}}_j}{\|\hat{\mathbf{g}}_j\|}$ is the eigenvector of $\hat{\mathbf{D}}_j$ corresponding to its maximum eigenvalue, all channel parameters can be identified up to a scalar ambiguity.

5. SIMULATIONS

In our experiment, a DS-CDMA system is simulated. All users have equal power. Their transmitted bits and assigned long spreading codes are assumed i.i.d. taking values from $\{+1, -1\}$. Other values are set as $P = 16$ (e.g., [8]), $q = 3$, and the bit $SNR = 15dB$. We will use the mean square error (MSE) of the channel estimates $E\{\|\hat{\mathbf{g}}_j - \mathbf{g}_j\|^2\}$ as the performance measure, where $\mathbf{g}_j$ is the normalized channel vector for user j and $\hat{\mathbf{g}}_j$ is its estimate. This expected value is approximated by the average result from 50 Monte Carlo runs.

The result for the proposed method (batch) is shown in Fig. 2 with $M = 8$. Each user transmitted 500 bits. Since there are no critical differences between estimates for different users, the MSE is presented only for four users, in a), b), c), d) respectively. It can be seen that the MSE for each channel of the proposed method reaches 10^{-2} after 200 bits are transmitted.

Our next experiment is to compare the proposed method with the recently presented subspace method [10] for systems with different loads. We assume all users experience the common multipath channel and implement the subspace approach in [10] by assigning the desired user in group 1 and all other users in group 2. The channel estimation errors for a 2-user system are compared in Fig. 3(a), while for an 8-user system in Fig. 3(b). Solid lines represent the proposed method while dashed lines for the subspace method. It can be observed that the subspace method has better performance than the proposed algorithm when the system

has only a few users according to Fig. 3(a). However it converges to a high error level under heavy load based on Fig. 3(b).

Appendix A: Derivation of Equation (14)

In order to solve for $\mathbf{D}_j$ and σ_v^2 , we follow our definition of the derivative and differentiate (13) with respect to σ_v^2 and $\mathbf{D}_j$ respectively,

$$\nabla_{\sigma_v^2} J = 2(P+q)\sigma_v^2 + \frac{2}{N} \sum_{n=1}^N \left[\text{tr} \left(\sum_{m=1}^M \mathbf{T}_m \right) - \mathbf{y}^H(n) \mathbf{y}(n) \right] \quad (20)$$

$$\begin{aligned} \nabla_{\mathbf{D}_j} J &= \frac{2}{N} \sum_{n=1}^N \left\{ \mathbf{C}_{j,1}^T \left[\sum_{m=1}^M \mathbf{T}_m + \sigma_v^2 \mathbf{I} - \mathbf{y}(n) \mathbf{y}^H(n) \right] \mathbf{C}_{j,1}^* \right. \\ &\quad + \mathbf{C}_{j,2}^T \left[\sum_{m=1}^M \mathbf{T}_m + \sigma_v^2 \mathbf{I} - \mathbf{y}(n) \mathbf{y}^H(n) \right] \mathbf{C}_{j,2}^* \\ &\quad \left. + \mathbf{C}_{j,3}^T \left[\sum_{m=1}^M \mathbf{T}_m + \sigma_v^2 \mathbf{I} - \mathbf{y}(n) \mathbf{y}^H(n) \right] \mathbf{C}_{j,3}^* \right\} \quad (21) \end{aligned}$$

Where “*” represents complex conjugate and $\mathbf{T}_j$ is previously defined. At the equilibrium points $(\hat{\sigma}_v^2, \hat{\mathbf{D}}_j)$ of our cost function J , these derivatives are equal to zero. By setting (20) to zero, $\hat{\sigma}_v^2$ is obtained first as

$$\hat{\sigma}_v^2 = \frac{1}{(P+q)N} \sum_{n=1}^N \left\{ [\mathbf{y}^H(n) \mathbf{y}(n)] - \left[\text{tr} \left(\sum_{m=1}^M \hat{\mathbf{T}}_m \right) \right] \right\} \quad (22)$$

where $\hat{\mathbf{T}}_m$ is obtained via replacing $\mathbf{D}_m$ by $\hat{\mathbf{D}}_m$ in $\mathbf{T}_m$. By using the property of the trace of the product of matrices, (22) becomes

$$\hat{\sigma}_v^2 = \frac{1}{(P+q)N} \sum_{n=1}^N \left\{ [\mathbf{y}^H(n) \mathbf{y}(n)] - [\mathbf{b}(n)]^H \hat{\mathbf{d}} \right\} \quad (23)$$

where $\mathbf{b}(n)$ is given by (18). Substituting (23) in (21), and setting it to zero, we can arrive at an equation for $\hat{\mathbf{d}}$. To obtain a closed form solution, we take *vec* operation on both sides of this new equation. By applying the properties of Kronecker product (see [6, Ch. 12]) and stacking all equations for $j = 1, \dots, M$ together, (14) can be obtained. $\square$

REFERENCES

- [1] S. E. Bensley and B. Aazhang, “Subspace-based channel estimation for code division multiple access communication Systems”, *IEEE Trans. Commun.*, pp. 1009-1020, August 1996.
- [2] G. B. Giannakis and S. D. Halford, “Asymptotically Optimal Blind Fractionally Spaced Channel Estimation and Performance Analysis”, *IEEE Trans. on Signal Processing*, vol. 45, no. 7, pp. 1815-1830, July 1996.
- [3] G. B. Giannakis and M. K. Tsatsanis, “A Unifying Maximum-likelihood View of Cumulant and Polyspectral Measures for Non-Gaussian Signal Classification and Estimation”, *IEEE Trans. on IT*, vol. 38, pp. 386-406, Mar. 1992.
- [4] K. S. Gilhousen, I. M. Jacobs, R. Padovani, A. J. Viterbi, A. Weaver and C. E. Wheatley, “On the capacity of a cellular CDMA system”, *IEEE Transactions on VT*, vol. 40, no. 2, pp. 303-311, May 1991.
- [5] M. Honig, and U. Madhow, and S. Verdu, “Blind Adaptive Multiuser Detection”, *IEEE Transactions on IT*, vol. 41, no. 4, pp. 944-960, July 1995.
- [6] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, Academic Press, San Diego, CA, 1985
- [7] B. Friedlander and B. Porat, “Asymptotically optimal estimation of MA and ARMA parameters of non-Gaussian processes from higher-order moments”, *IEEE Trans. on Auto. Control*, vol. 35, pp. 27-35, Jan. 1990.
- [8] M. Torlak, B. L. Evans and G. Xu, “Blind Estimation of FIR Channels in CDMA Systems with Aperiodic Spreading Sequences”, *Proc. of 31th Asilomar Conf.*, Pacific Grove, CA, Oct. 29-Nov. 1, 1997.
- [9] M. K. Tsatsanis, and G. B. Giannakis, “Equalization of Rapidly Fading Channels: Self-Recovering Methods”, *IEEE Trans. on Commun.*, vol. 44, no. 5, pp. 619-630, May 1996.
- [10] A. J. Weiss and B. Friedlander, “CDMA Downlink Channel Estimation with Aperiodic Spreading”, *Proc. of 31th Asilomar Conf.*, Pacific Grove, CA, Oct. 29-Nov. 1, 1997.

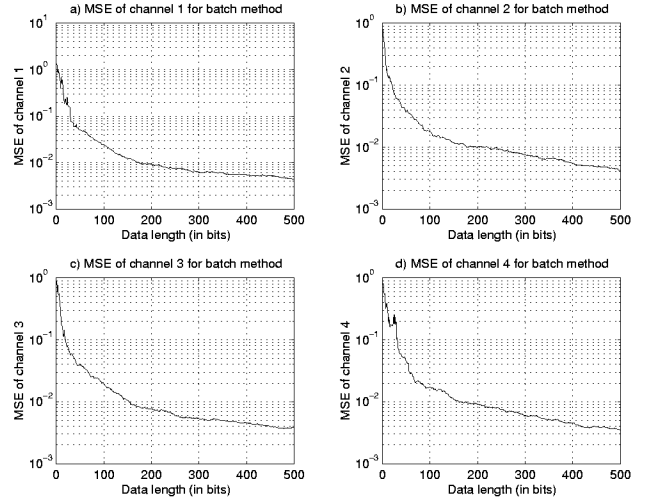


Figure 2. MSE of the proposed batch method for an 8-user system (4 users shown).

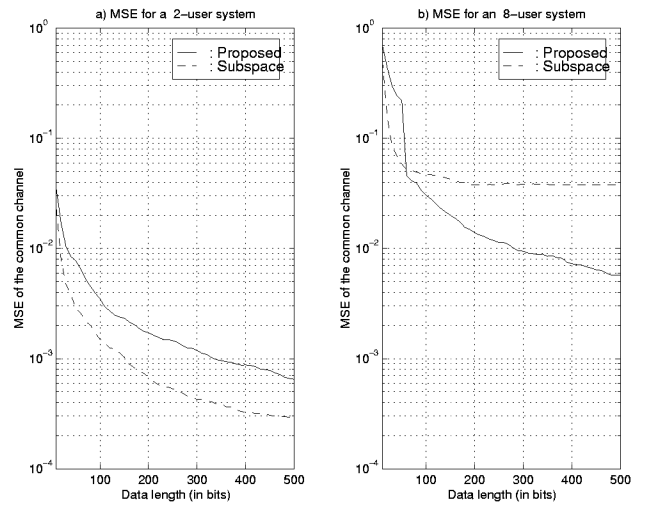


Figure 3. Comparison of the proposed batch method with subspace method for a 2-user and an 8-user system.