Code-Timing Estimation for CDMA Systems With Bandlimited Chip Waveforms

Rensheng Wang, *Student Member, IEEE*, Hongbin Li, *Member, IEEE*, and Tao Li

***Abstract*—In this paper, we present a novel code-timing estimator for uplink asynchronous direct-sequence code-division multiple-access systems utilizing *bandlimited* chip waveforms. The proposed estimator requires only the spreading code and training of the desired user. We start from a maximum likelihood (ML) approach that models the intersymbol interference and multiple-access interference as a colored Gaussian process with unknown covariance matrix in the frequency domain. The exact ML estimator is highly nonlinear and requires iterative searches over multi-dimensional parameter space that is impractical to implement. To deal with this difficulty, we invoke asymptotic (large-sample) approximations of the ML criterion and reparameterization techniques, which lead to an asymptotic ML estimator that yields code-timing and channel estimates via efficient noniterative quadratic optimizations. To benchmark the proposed estimator, we provide Cramér–Rao bound analysis for the code-timing estimation problem. Numerical simulation results are presented, which show that the proposed scheme is resistant to interference, fading, and modeling errors (e.g., sampling position errors), and compares favorably to several competing schemes in multipath fading channels.**

***Index Terms*—Bandlimited chip waveforms, code-division multiple-access (CDMA), code synchronization, Cramér-Rao bound (CRB), maximum-likelihood (ML), parameter estimation.**

I. Introduction

DIRECT-SEQUENCE (DS) code-division multiple-access (CDMA) is a major air interface for wireless mobile communications [1]. In DS-CDMA systems, all user transmissions overlap in time and frequency. They are differentiated from one another by using a unique spreading code for each user. In order to successfully recover the information of each transmission, the local spreading code generator has to be synchronized to the code-timing of the desired transmission.

Multiuser code-timing estimation, which parallels the well acknowledged research on multiuser detection ([2] and references therein) for CDMA systems, has been receiving increasing interest recently. A variety of code-timing estimation techniques have been proposed so far, including both training-assisted and blind schemes. Examples of the former category include the classical, single-user based correlator [3] and the more recently introduced, multiuser-based minimum mean squared error (MMSE) [4], large-sample maximum likelihood (LSML) [5], [6], exact maximum likelihood (ML) [7], and decoupled multiuser acquisition (DEMA) [8] synchronization schemes. Some recent blind code synchronization algorithms include MUSIC [9], [10] and the variants [11], [12], and the minimum variance-based schemes [13]–[15]. Compared with the single-user based correlator, these multiuser-based code synchronization schemes achieve significantly improved performance in near-far environments, and are able to support more user transmissions without enforcing stringent power control.

Most of these code-timing estimation schemes, however, implicitly assume *rectangular* chip waveforms that are *not bandlimited*. Meanwhile, practical CDMA systems utilize *bandlimited* chip waveforms, such as the square-root raised-cosine pulse [16]. Although extensions of the aforementioned techniques to deal with bandlimited chip waveforms appear conceptually straightforward, implementations of such extensions are often challenging due to the need to solve highly nonlinear cost functions. One such extension was reported in [17], which extends the MUSIC code-timing estimator of [9] by incorporating knowledge of the bandlimited chip waveform in the MUSIC cost function. While the original MUSIC estimator can be efficiently and noniteratively implemented via simple second-order polynomial rooting, the extended MUSIC algorithm involves iterative nonlinear optimization that is computationally intensive and subject to local convergence.

An alternative code-timing estimator that considers bandlimited chip waveforms was recently presented in [18]. It exploits various signal space invariances in the frequency domain to isolate the subspace of interest for the desired user, from which an ESPRIT [19] like procedure is invoked to derive the code-timing estimates. Unlike the extended MUSIC estimator in [17], which is an iterative time-domain based scheme, the shift-invariance-based algorithm in [18] utilizes computationally more efficient, noniterative frequency-domain processing. The shift-invariance-based method is also found statistically more accurate since it is free of the local convergence problem suffered by the former.

While the extended MUSIC [17] and the shift-invariance based [18] estimators are both blind schemes that require no training, we present in this paper a new code-timing estimation scheme for CDMA with bandlimited chip waveforms by exploiting training that exists in most wireless standards anyway. As we shall see, the proposed estimator benefits from training with significantly improved performance over the

Manuscript received January 19, 2003; revised May 1, 2003; accepted May 3, 2003. The editor coordinating the review of this paper and approving it for publication is H. Boelcskei. This work was supported by the Army Research Office under Contract DAAD19-03-1-0184, by the New Jersey Commission on Science and Technology, and by the Center for Wireless Network Security, Stevens Institute of Technology, Hoboken, NJ.

R. Wang and H. Li are with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ 07030 USA (e-mail: rwang1@stevens-tech.edu; hli@stevens-tech.edu).

T. Li is with the Department of Computer, Sichuan University, Chengdu 610065, China (e-mail: litao@scu.edu.cn).

Digital Object Identifier 10.1109/TWC.2004.830858

blind techniques. Like the shift-invariance based method, our proposed estimator is a frequency-domain-based scheme. We first convert the received signal to the frequency domain by fast Fourier transform (FFT). The proposed estimator is then derived by an ML approach that models the overall interference as a *colored* Gaussian random process with an unknown covariance matrix. The exact ML cost function is in general highly nonlinear and difficult to solve. To circumvent this difficulty, we invoke an asymptotic result that renders the ML criterion asymptotically (for large data samples) equivalent to a simpler cost function involving a (nonlinear) weighted least-squares (WLS) fitting. Still, the WLS cost function requires L_k-dimensional searches over the parameter space, where L_k denotes the number of distinct paths for the desired user. To further reduce the complexity, we next reparameterize the WLS cost function by coefficients of an L_kth-order polynomial, by which the code-timing estimates for the desired user are obtained via simple quadratic minimizations.

The rest of the paper is organized as follows. In Section II, we introduce the data model for CDMA with bandlimited chip waveforms, and formulate the problem of interest. In Section III, we present the proposed code-timing estimator, with technical details included in the Appendices. In Section IV, we derive the Cramér–Rao bound (CRB), a lower bound on the variance of any unbiased estimator, for the code-timing estimation problem. Numerical results comparing the proposed and other code-timing estimators for bandlimited CDMA are presented in Section V. Finally, we provide concluding remarks in Section VI.

Note: Vectors (matrices) are denoted by boldface lower (upper) case letters; all vectors are column vectors; superscripts $(\cdot)^T, (\cdot)^*$, and $(\cdot)^H$ denote the transpose, conjugate, and conjugate transpose, respectively; $\mathbf{I}_N$ denotes the $N \times N$ identity matrix; $\mathbf{0}$ denotes a matrix/vector with all zero elements; $\mathrm{diag}\{\cdot\}$ denotes a diagonal matrix; $E\{\cdot\}$ denotes the statistical expectation; $\star$ denotes the linear convolution; $\|\cdot\|$ denotes the matrix/vector Frobenius norm; and finally, $\otimes$ denotes the matrix/vector Kronecker product [20].

II. Problem Formulation

The system under investigation is an asynchronous (uplink) K-user DS-CDMA system with spreading codes of length (processing gain) N. The code waveform for user k is $u_k(t) = \sum_{n=0}^{N-1} c_k(n)p(t-nT_c)$, where $\{c_k(n)\}_{n=0}^{N-1}$ denotes the spreading code for user $k, p(t)$ the chip waveform assumed to be *bandlimited* and identical for all users, and T_c the chip duration. The transmitted signal $s_k(t)$ for user k is formed by multiplying $u_k(t)$ by the mth transmitted data symbol $d_k(m)\ :\ s_k(t) = \sum_{m=0}^{M-1} d_k(m)u_k(t-mT_s)$, where M denotes the number of symbols used for code acquisition, and $T_s = NT_c$ denotes the symbol duration.

Consider a general scenario that the base station is equipped with an array of J receive antennas. It is noted that our scheme works with $J=1$. For mathematical tractability, the derivation of the proposed estimator assumes that the multipath channel remains static during code acquisition. Later, in Section V, we will test the proposed estimator in realistic time-varying multipath channels, using the standard Jakes' model [21]. With that in mind, the received signal from the μth receive antenna at the base station can be expressed by

$$\begin{aligned} r_\mu(t) &= \sum_{k=1}^{K}\sum_{l=1}^{L_k} \alpha_{\mu,k,l} s_k(t-\tau_{k,l}) + n'_\mu(t) \\ \mu &= 1,\ldots,J \end{aligned} \tag{1}$$

where $L_k, \alpha_{\mu,k,l}$, and $\tau_{k,l}$ denote the number of propagation paths, the lth path's (complex-valued) attenuation and delay observed at the μth receive antenna for user k, respectively, and $n'_\mu(t)$ is the additive noise. We assume that the relative delay among different receive antennas is negligible; note that the relative delay manifests itself as a phase shift that makes $\alpha_{\mu,k,l}$ distinct for different μ. The receiver front-end is a chip-matched filter that outputs[1]

$$\begin{aligned} \bar{y}_\mu(t) &= r_\mu(t) \star p(T_c - t) \\ &= \sum_{k=1}^{K}\sum_{m=1}^{M} d_k(m) h_{\mu,k}(t-mT_s) + n_\mu(t) \end{aligned} \tag{2}$$

where $h_{\mu,k}(t)$ denotes the impulse response of the overall channel for user k at receive antenna μ, which includes the transmitter/receiver filters and the physical wireless channel

$$h_{\mu,k}(t) = \sum_{l=1}^{L_k} \alpha_{\mu,k,l}\bar{g}_k(t-\tau_{k,l}) \tag{3}$$

with $\bar{g}_k(t) \triangleq u_k(t) \star p(T_c - t)$ and $n_\mu(t) \triangleq n'_\mu(t) \star p(T_c - t)$.

We assume that the maximum path delay is less than T_s, which may be achieved through a side signaling channel for call set-up [4], [22]. It is customary to truncate $p(t)$ such that it spans only several chips [17], [23]; such a truncation leads to negligible spectral leakage compared to the effect of noise and interference in the system (also see Section III). Under these conditions, we note that $h_{\mu,k}(t)$ has a finite support: $h_{\mu,k}(t) = 0$, for $t \notin [0, 2T_s]$.

Without loss of generality, let user k be the desired user. The problem of interest is to estimate the code-timing $\{\tau_{k,l}\}_{l=1}^{L_k}$ from $\{\bar{y}_\mu(t)\}_{\mu=1}^{J}$, assuming that the training symbols $\{d_k(m)\}_{m=0}^{M-1}$ and spreading waveform $\bar{g}_k(t)$ for user k are known at the base station.

III. Proposed Code-Timing Estimation Scheme

For digital signal processing (DSP), the outputs of the chip-matched filter $\{\bar{y}_\mu(t)\}_{\mu=1}^{J}$ are sampled with a sampling interval $T_i = T_c/Q$. That is, $\bar{y}_\mu(i) = \bar{y}_\mu(t)|_{t=iT_i}, i = 0, 1, \ldots, (M+1)NQ - 1$, where the integer $Q \geq 1$ denotes the *oversampling factor*. Typically, $Q = 2$ is sufficient. Note that because of delay spread, the observation interval that covers M symbols is $(M+1)T_s$. We form M *overlapping* data blocks of $2NQ$ samples, each block consisting of data samples within two adjacent symbol intervals: $\bar{\mathbf{y}}_\mu(m) = [\bar{y}_\mu(mNQ), \ldots, \bar{y}_\mu(mNQ + 2NQ - 1)]^T$ $m = 0, \ldots, M-1; \mu = 1, \ldots, J$. Due to asynchronous transmissions, $\bar{\mathbf{y}}_\mu(m)$ is contributed by three consec-

[1]Throughout this paper, we use notation $(\bar{\cdot})$ to denote a time-domain quantity if its frequency-domain counterpart is also used for estimation.

utive symbols. Let $h_{\mu,k}(i) = h_{\mu,k}(t)|_{t=iT_i}$. The $2NQ \times 1$ *signature* vectors, which take into account the spreading, transmitter/receiver filters, and physical channel, corresponding to the three adjacent symbols $d_k(m-1)$, $d_k(m)$, and $d_k(m+1)$, have the following forms:

$$\mathbf{h}^{-}_{\mu,\boldsymbol{\tau}_k} \triangleq [h_{\mu,k}(NQ), \ldots, h_{\mu,k}(2NQ-1), \mathbf{0}_{1\times NQ}]^T \quad (4)$$

$$\mathbf{h}_{\mu,\boldsymbol{\tau}_k} \triangleq [h_{\mu,k}(0), \ldots, h_{\mu,k}(2NQ-1)]^T \quad (5)$$

$$\mathbf{h}^{+}_{\mu,\boldsymbol{\tau}_k} \triangleq [\mathbf{0}_{1\times NQ}, h_{\mu,k}(0), \ldots, h_{\mu,k}(NQ-1)]^T \quad (6)$$

where the subscript $\boldsymbol{\tau}_k \triangleq [\tau_{k,1}, \ldots, \tau_{k,L_k}]^T$ signifies the dependence of these vectors on the path delays [e.g., (3)]. With these definitions, $\bar{\mathbf{y}}_\mu(m)$ can be expressed as [cf. (2)–(3)]

$$\bar{\mathbf{y}}_\mu(m) = \sum_{k=1}^{K}[d_k(m-1)\mathbf{h}^{-}_{\mu,\boldsymbol{\tau}_k} + d_k(m)\mathbf{h}_{\mu,\boldsymbol{\tau}_k} + d_k(m+1)\mathbf{h}^{+}_{\mu,\boldsymbol{\tau}_k}] + \mathbf{n}_\mu(m) \quad (7)$$

where $\mathbf{n}_\mu(m)$ denotes the $2NQ \times 1$ vectors formed from the noise/interference samples of $n_\mu(t)$. Given that user k is of interest, we rewrite (7) as

$$\bar{\mathbf{y}}_\mu(m) \triangleq d_k(m)\mathbf{h}_{\mu,\boldsymbol{\tau}_k} + \bar{\mathbf{e}}_\mu(m) \quad (8)$$

where $\bar{\mathbf{e}}_\mu(m)$ lumps the channel noise and overall interference, including the multiple access interference (MAI), intersymbol interference (ISI), and additive noise.

To convert the received data to the frequency domain, we take the Fourier transform, or the FFT in particular, of $\bar{\mathbf{y}}_\mu(m)$. It should be noted that the spectrum of $\bar{g}_k(t)$ in (3) usually tapers off at the end frequencies (i.e., frequencies close to $\pm 0.5 f_i$, where $f_i \triangleq 1/T_i$ denotes the sampling frequency, e.g., [23]). In the presence of channel noise, the end frequencies have a lower signal-to-noise ratio (SNR) than elsewhere. Hence, we may discard the end frequencies to avoid noise amplification caused by frequency deconvolution [23]. To do so, let $\eta \in (0,1]$ denote the FFT frequency selection parameter. Typically, one can choose $\eta = 0.5$ when the oversampling factor $Q = 2$ [18]. Define

$$\boldsymbol{\mathcal{F}} = \begin{bmatrix} 1 & \phi^{-N_s} & \cdots & \phi^{-(2NQ-1)N_s} \\ 1 & \phi^{-(N_s-1)} & \cdots & \phi^{-(2NQ-1)(N_s-1)} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \phi^{N_s-1} & \cdots & \phi^{(2NQ-1)(N_s-1)} \end{bmatrix}$$

where $\phi \triangleq e^{-j\pi/(NQ)}$ and $N_s \triangleq \lceil \eta NQ \rceil$, where $\lceil \cdot \rceil$ denotes the smallest integer no less than the argument. One can see that the $2N_s \times 2NQ$ matrix $\boldsymbol{\mathcal{F}}$ is formed by rows of the $2NQ \times 2NQ$ full FFT matrix that correspond to the selected FFT frequencies. Hence, discarding the FFT end frequencies is equivalent to multiplying both sides of (8) by $\boldsymbol{\mathcal{F}}$

$$\mathbf{y}_\mu(m) \triangleq \boldsymbol{\mathcal{F}}\bar{\mathbf{y}}_\mu(m) = d_k(m)\boldsymbol{\mathcal{F}}\mathbf{h}_{\mu,\boldsymbol{\tau}_k} + \mathbf{e}_\mu(m) \quad (9)$$

where $\mathbf{e}_\mu(m) \triangleq \boldsymbol{\mathcal{F}}\bar{\mathbf{e}}_\mu(m)$. To facilitate our derivation, we rewrite $\mathbf{h}_{\mu,\boldsymbol{\tau}_k}$ as [cf. (3)]: $\mathbf{h}_{\mu,\boldsymbol{\tau}_k} = \sum_{l=1}^{L_k} \alpha_{\mu,k,l}\bar{\mathbf{g}}_k(\tau_{k,l})$, where $\bar{\mathbf{g}}_k(\tau_{k,l}) \triangleq [\bar{g}_k(-\tau_{k,l}), \ldots, \bar{g}_k((2NQ-1)T_i - \tau_{k,l})]^T$. Using the time-shifting property of Fourier transform [24], we have

$$\boldsymbol{\mathcal{F}}\mathbf{h}_{\mu,\boldsymbol{\tau}_k} \approx \mathbf{G}_k\boldsymbol{\Phi}_k(\boldsymbol{\tau}_k)\boldsymbol{\alpha}_{\mu,k} \quad (10)$$

where

$$\boldsymbol{\Phi}_k(\boldsymbol{\tau}_k) \triangleq \begin{bmatrix} \phi^{-N_s\tau_{k,1}} & \cdots & \phi^{-N_s\tau_{k,L_k}} \\ \phi^{-(N_s-1)\tau_{k,1}} & \cdots & \phi^{-(N_s-1)\tau_{k,L_k}} \\ \vdots & \vdots & \vdots \\ \phi^{(N_s-1)\tau_{k,1}} & \cdots & \phi^{(N_s-1)\tau_{k,L_k}} \end{bmatrix} \quad (11)$$

$$\boldsymbol{\alpha}_{\mu,k} \triangleq [\alpha_{\mu,k,1}, \ldots, \alpha_{\mu,k,L_k}]^T \quad (12)$$

$$\mathbf{G}_k \triangleq \mathrm{diag}\{\mathbf{g}_k\} \quad (13)$$

$$\mathbf{g}_k \triangleq \boldsymbol{\mathcal{F}}[\bar{g}_k(0), \ldots, \bar{g}_k((2NQ-1)T_i)]^T. \quad (14)$$

Equation (10) holds only approximately because of the aliasing caused by the truncation of $p(t)$ (see Section II). The truncation widens slightly the spectrum of $\bar{g}_k(t)$, and sampling at a rate $1/T_i$ introduces some small aliasing due to spectral folding, which will eventually lead to a small bias in the code-timing estimate. The aliasing, however, can be neglected compared to the noise/interference induced estimation error [23].

Substituting (10) into (9) yields

$$\mathbf{y}_\mu(m) \approx d_k(m)\mathbf{G}_k\boldsymbol{\Phi}_k(\boldsymbol{\tau}_k)\boldsymbol{\alpha}_{\mu,k} + \mathbf{e}_\mu(m)$$
$$m = 0, 1, \ldots, M-1; \quad \mu = 1, \ldots, J. \quad (15)$$

Effectively, $\mathbf{y}_\mu(m)$ can be thought of a sum of L_k complex sinusoids with frequencies

$$f_{k,l} = -\frac{\tau_{k,l}}{2NQ} \quad (16)$$

that are weighted by the diagonal matrix $\mathbf{G}_k$ and corrupted by the interference/noise $\mathbf{e}_\mu(m)$.

In the sequel, we approximate $\{\mathbf{e}_\mu(m)\}$ as complex Gaussian random vectors with zero-mean and an *arbitrary unknown* covariance matrix $E\{\mathbf{e}_{\mu_1}(m_1)\mathbf{e}^H_{\mu_2}(m_2)\} = \mathbf{R}_{\mathbf{e}}\delta(\mu_1-\mu_2)\delta(m_1-m_2)$, where $\delta(\cdot)$ denotes the Kronecker delta. While this approximation may not be observed exactly in practice, it leads to an estimator that works quite well in realistic multiuser environments, as verified in Section V. Briefly stated, the proposed estimator follows an ML approach, starting from the likelihood function of $\mathbf{y}_\mu(m)$ conditioned on the unknown parameters, i.e., the multipath code-timing $\boldsymbol{\tau}_k$, channel fading coefficients $\{\boldsymbol{\alpha}_{\mu,k}\}_{\mu=1}^J$, and the interference/noise covariance matrix $\mathbf{R}_{\mathbf{e}}$. As shown in Appendix A, the exact ML cost function is highly nonlinear and difficult to optimize. To circumvent this difficulty, we invoke an asymptotic result that renders the ML criterion asymptotically (for large data samples) equivalent to a simpler cost function involving a (nonlinear) weighted least-squares (WLS) fitting. Still, the WLS cost function requires L_k-dimensional searches over the parameter space. To further reduce the complexity, we next reparameterize the WLS cost function by coefficients of an L_kth-order polynomial, by which the code-timing estimates for the desired user are obtained via simple quadratic minimizations. Details of the derivation of the proposed code-timing estimator can be found in Appendix A. In the following, we summarize the proposed estimator along with its computational complexity in terms of flops of operation.[2]

Step 1) Compute the ML estimate of the interference/noise covariance matrix

$$\hat{\mathbf{R}}_{\mathbf{e}} = \frac{1}{J}\sum_{\mu=1}^{J}\left(\hat{\mathbf{R}}_{\mathbf{y}_\mu} - \hat{\mathbf{r}}_{\mathbf{y}_\mu d}\hat{\mathbf{r}}_{\mathbf{y}_\mu d}^H\right) \quad \Rightarrow O(J(NQ)^2) \text{ flops} \tag{17}$$

where $\hat{\mathbf{R}}_{\mathbf{y}_\mu}$ and $\hat{\mathbf{r}}_{\mathbf{y}_\mu d}$ denote, respectively,

$$\hat{\mathbf{R}}_{\mathbf{y}_\mu} \triangleq \frac{1}{M}\sum_{m=0}^{M-1}\mathbf{y}_\mu(m)\mathbf{y}_\mu^H(m), \quad \Rightarrow O(JM(NQ)^2) \text{ flops} \tag{18}$$

$$\hat{\mathbf{r}}_{\mathbf{y}_\mu d} \triangleq \frac{1}{M}\sum_{m=0}^{M-1}\mathbf{y}_\mu(m)d_k^*(m), \quad \Rightarrow O(JMNQ) \text{ flops.} \tag{19}$$

Compute the weighting matrix used for WLS fitting (note that $\mathbf{G}_k$ is diagonal)

$$\mathbf{W}^{-1} \triangleq \mathbf{G}_k^{-1}\hat{\mathbf{R}}_{\mathbf{e}}\mathbf{G}_k^{-H}, \quad \Rightarrow O((NQ)^2) \text{ flops} \tag{20}$$

and form the Hankel matrix

$$\mathbf{\Gamma}_\mu \triangleq \begin{bmatrix} \gamma_{\mu,1} & \gamma_{\mu,2} & \cdots & \gamma_{\mu,L_k+1} \\ \gamma_{\mu,2} & \gamma_{\mu,3} & \cdots & \gamma_{\mu,L_k+2} \\ \vdots & \vdots & \vdots & \vdots \\ \gamma_{\mu,2N_s-L_k} & \gamma_{\mu,2N_s-L_k+1} & \cdots & \gamma_{\mu,2N_s} \end{bmatrix} \tag{21}$$

where

$$[\gamma_{\mu,1},\ldots,\gamma_{\mu,2N_s}]^T \triangleq \boldsymbol{\gamma}_\mu = \mathbf{G}_k^{-1}\hat{\mathbf{r}}_{\mathbf{y}_\mu d}, \quad \Rightarrow O(JNQ) \text{ flops.} \tag{22}$$

Step 2) Due to a reparameterization procedure (see Appendix A), the code-timing estimation problem is equivalent to the estimation of the coefficients of an L_kth-order polynomial. Compute an initial estimate, $\hat{\mathbf{b}}^{(0)}$ of the polynomial coefficients $\mathbf{b} = [b_0, b_1, \ldots, b_{L_k}]^T$, by minimizing the following quadratic cost function: [3]

$$C_{\mathrm{qd},0}(\mathbf{b}) = \mathbf{b}^H\left(\sum_{\mu=1}^{J}\mathbf{\Gamma}_\mu^H\mathbf{\Gamma}_\mu\right)\mathbf{b} \quad \Rightarrow O\left(JL_k^2NQ\right) + O\left(L_k^3\right) \text{ flops.} \tag{23}$$

Step 3) Form a $2N_s \times (2N_s - L_k)$ Toeplitz matrix from the initial estimate $\hat{\mathbf{b}}^{(0)}$ as follows:

$$\hat{\mathbf{B}}^{(0)} = \begin{bmatrix} \hat{b}_0^{(0)} & \hat{b}_1^{(0)} & \cdots & \hat{b}_{L_k}^{(0)} & & 0 \\ & \ddots & \ddots & & \ddots & \\ 0 & & \hat{b}_0^{(0)} & \hat{b}_1^{(0)} & \cdots & \hat{b}_{L_k}^{(0)} \end{bmatrix}^H. \tag{24}$$

Refine the initial estimate of the polynomial coefficients and compute a new estimate $\hat{\mathbf{b}}$ of $\mathbf{b}$ by minimizing the following quadratic cost function:

$$C_{\mathrm{qd}}(\mathbf{b}) = \mathbf{b}^H\left[\sum_{\mu=1}^{J}\mathbf{\Gamma}_\mu^H\left(\hat{\mathbf{B}}^{(0)H}\mathbf{W}^{-1}\hat{\mathbf{B}}^{(0)}\right)^{-1}\mathbf{\Gamma}_\mu\right]\mathbf{b} \quad \Rightarrow O((NQ)^3) + O(JL_k(NQ)^2) + O\left(L_k^3\right) \text{ flops.} \tag{25}$$

Step 4) Compute the L_k roots of the L_kth-order polynomial with coefficients $\hat{\mathbf{b}}$. Calculate the phase angles of the L_k roots and denote them by $\{\hat{\zeta}_{k,l}\}_{l=1}^{L_k}$. Compute the code-timing estimates as follows [see (16)]:

$$\hat{\tau}_{k,l} = -\frac{NQ\hat{\zeta}_{k,l}}{\pi}. \tag{26}$$

Remark 1: The minimization of the quadratic functions (23) and (25) follows a similar approach, which is discussed in details in Appendix B.

Remark 2: The *overall complexity* of the proposed scheme, in terms of flop count, is the sum of the number of flops incurred in each step as summarized in the above, plus $O(2JMNQ\log_2(2NQ))$ flops that are incurred in the calculation of $\mathbf{y}_\mu(m), m = 0,1,\ldots,M-1$ and $\mu = 1,\ldots,J$, through JM FFTs of length $2NQ$. For most applications, we typically have J (the number of receive antennas) quite small, and $M \approx NQ \gg L_k$, for which we see that the major complexity comes from the calculation of the data covariance matrices $\hat{\mathbf{R}}_{\mathbf{y}_\mu}$ and the matrix inverse $(\hat{\mathbf{B}}^{(0)H}\mathbf{W}^{-1}\hat{\mathbf{B}}^{(0)})^{-1}$.

Remark 3: Once we have $\hat{\tau}_{k,l}$, an estimate of the channel coefficients can be obtained as

$$\hat{\boldsymbol{\alpha}}_{\mu,k} = \left[\mathbf{\Phi}_k^H(\hat{\boldsymbol{\tau}}_k)\mathbf{W}\mathbf{\Phi}_k(\hat{\boldsymbol{\tau}}_k)\right]^{-1}\mathbf{\Phi}_k^H(\hat{\boldsymbol{\tau}}_k)\mathbf{W}\mathbf{G}_k^{-1}\hat{\mathbf{r}}_{\mathbf{y}_\mu d}. \tag{27}$$

IV. Cramér–Rao Bound

Cramér–Rao bound (CRB) provides a lower bound on the variance of the parameter estimates obtained by any unbiased estimators, and it can be used to assess the accuracy of various code-timing estimation schemes. In this section, we present the CRB for the parameter estimation problem based on the data model in (15), which is repeated below for easy reference:

$$\begin{aligned} \mathbf{y}_\mu(m) &\approx d_k(m)\mathbf{G}_k\mathbf{\Phi}_k(\boldsymbol{\tau}_k)\boldsymbol{\alpha}_{\mu,k} + \mathbf{e}_\mu(m) \\ m &= 0,1,\ldots,M-1 \\ \mu &= 1,\ldots,J. \end{aligned} \tag{28}$$

Let $\boldsymbol{\theta} \triangleq [\boldsymbol{\tau}_k^T, \mathrm{Re}^T(\boldsymbol{\alpha}_{1,k}),\ldots,\mathrm{Re}^T(\boldsymbol{\alpha}_{J,k}), \mathrm{Im}^T(\boldsymbol{\alpha}_{1,k}),\ldots, \mathrm{Im}^T(\boldsymbol{\alpha}_{J,k})]^T$, which collects all unknown parameters of interest. By using the Slepian–Bangs formula [25], we show in the Appendix I that the CRB matrix is given by

$$[\mathrm{CRB}^{-1}(\boldsymbol{\theta})] = 2\mathrm{Re}\left\{\begin{bmatrix} \mathbf{F}_{\boldsymbol{\tau}\boldsymbol{\tau}} & \mathbf{F}_{\boldsymbol{\tau}\boldsymbol{\alpha}} & j\mathbf{F}_{\boldsymbol{\tau}\boldsymbol{\alpha}} \\ \mathbf{F}_{\boldsymbol{\tau}\boldsymbol{\alpha}}^H & \mathbf{F}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} & j\mathbf{F}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} \\ -j\mathbf{F}_{\boldsymbol{\tau}\boldsymbol{\alpha}}^H & -j\mathbf{F}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} & \mathbf{F}_{\boldsymbol{\alpha}\boldsymbol{\alpha}} \end{bmatrix}\right\} \tag{29}$$

[2]In our complexity analysis, the FFT frequency selection factor is set to a standard value of $\eta = 0.5$ [18], [23], which gives $2N_s = NQ$.

[3]The flop count of $O(L_k^3)$ in (23) and (25) comes from a matrix eigendecomposition involved in the quadratic minimization that is detailed in Appendix B.

where

$$\mathbf{F}_{\boldsymbol{\tau\tau}} = M \sum_{\mu=1}^{J} \boldsymbol{\Psi}_{\mu}^{H} \left(\mathbf{G}_k^H \mathbf{R}_{\mathbf{e}}^{-1} \mathbf{G}_k\right) \boldsymbol{\Psi}_{\mu} \tag{30}$$

$$\begin{aligned}\mathbf{F}_{\boldsymbol{\tau\alpha}} = M \big[&\boldsymbol{\Psi}_1^H \left(\mathbf{G}_k^H \mathbf{R}_{\mathbf{e}}^{-1} \mathbf{G}_k\right) \boldsymbol{\Phi}_k(\boldsymbol{\tau}_k), \ldots, \\ &\boldsymbol{\Psi}_J^H \left(\mathbf{G}_k^H \mathbf{R}_{\mathbf{e}}^{-1} \mathbf{G}_k\right) \boldsymbol{\Phi}_k(\boldsymbol{\tau}_k)\big]\end{aligned} \tag{31}$$

$$\mathbf{F}_{\boldsymbol{\alpha\alpha}} = M \left[\mathbf{I}_J \otimes \left(\boldsymbol{\Phi}_k^H(\boldsymbol{\tau}_k) \mathbf{G}_k^H \mathbf{R}_{\mathbf{e}}^{-1} \mathbf{G}_k \boldsymbol{\Phi}_k(\boldsymbol{\tau}_k)\right)\right]. \tag{32}$$

In (30) and (31), we have $\boldsymbol{\Psi}_{\mu} \stackrel{\triangle}{=} [\alpha_{\mu,k,1}\boldsymbol{\psi}_1, \ldots, \alpha_{\mu,k,L_k}\boldsymbol{\psi}_{L_k}] \in \mathbb{C}^{2N_s \times L_k}$, for $\mu = 1, \ldots, J$, where $\boldsymbol{\psi}_l \stackrel{\triangle}{=} (\partial \boldsymbol{\phi}_l)/(\partial \tau_{k,l})$ for $l = 1, \ldots, L_k$, with $\boldsymbol{\phi}_l$ denoting the lth column of $\boldsymbol{\Phi}_k(\boldsymbol{\tau}_k)$.

V. Numerical Results

We consider a K-user asynchronous CDMA system in the uplink using a unit-energy binary phase shift keying (BPSK) constellation and bandlimited chip waveforms. Each user is assigned a randomly generated spreading code of length $N = 16$. We consider an environment with no strict power control. In particular, the transmitted power for all interfering users is assumed P dB higher than that of the desired user. Henceforth, P is referred to as the near-far ratio (NFR). The bandlimited chip waveform is a square-root raised-cosine pulse with roll-off factor 0.8, truncated to a duration of $4T_c$. The oversampling factor is chosen as $Q = 2$, and the FFT frequency selection parameter is $\eta = 0.5$ (see Section III).

While the proposed estimator was derived assuming that the channel remains static during code acquisition, we assess its performance in both time-invariant and time-varying fading channels. We consider two performance measures. One is the *probability of correct acquisition*, which is defined as the probability of the event that the code-timing estimate is within a half chip to the correct code-timing. The other is the *root mean-squared error (RMSE)* normalized by T_c, given correct acquisition. In the multipath case, we evaluate the probability of acquisition for each path regardless the acquisition of the other paths. However, the results reported in the sequel are the *averaged* probability of acquisition for all paths. This implies that if correct acquisition is achieved with only a single path, the overall performance would still be very poor (due to averaging with paths with incorrect acquisition). The RMSE results are reported in a similar fashion.

We compare herein the proposed estimator with the blind shift-invariance based (SIB) scheme [18] and the matched filter (MF) estimator ([3, Sec. 5-5]). The MF estimator is a training based method that uses identical training symbols for the desired user [3], [5]. It treats the overall interference as *white* Gaussian noise. In particular, the MF estimator is implemented by taking the Fourier transform of $\mathbf{G}_k^{-1}\hat{\mathbf{r}}_{\mathbf{y}_{\mu}d}$ [see (13) and (19)], which effectively performs matched filtering/correlation in the frequency domain, follwed by finding the peak of the magnitude spectrum. For multipath code acquisition, the MF estimator estimates the first dominant path by finding the largest correlation peak; then, the dominant path is subtracted from the received signal, and the second dominant path is found by correlating the resulting residual signal with the spreading waveform of the desired user; and so on. The results presented next are averaged over 1000 independent trials, for which the delays, attenuations, information symbols, and channel noise are changed independently from one trial to another.

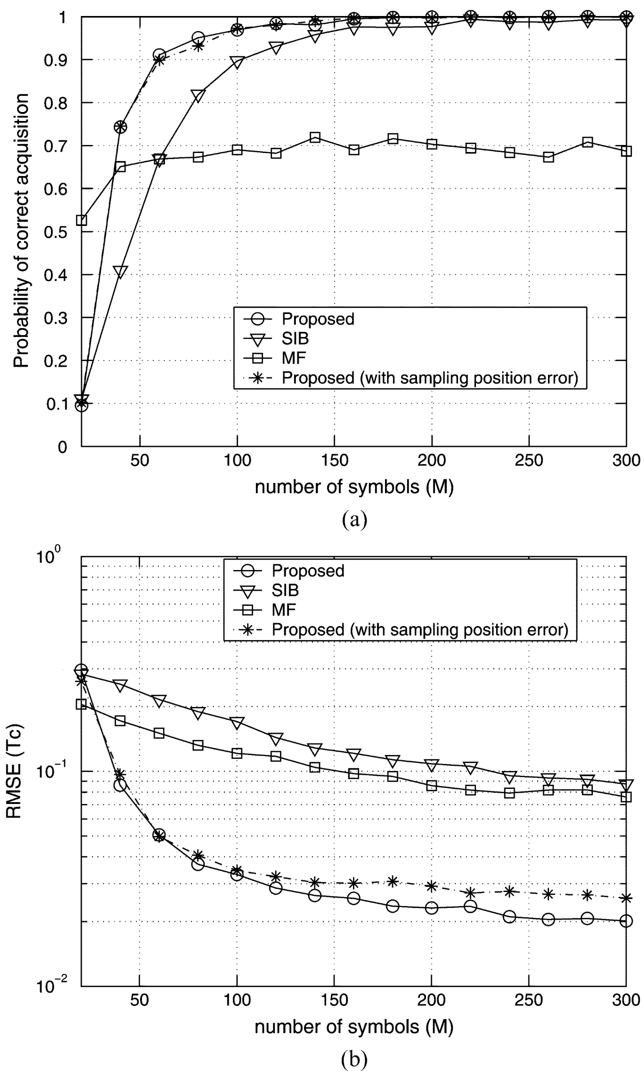


Fig. 1. Performance versus M, the number of symbols used for code acquisition, in flat-fading channels when $J = 1, K = 10, N = 16$, SNR $= 15$ dB, and NFR $= 5$ dB. (a) Probability of correct acquisition. (b) RMSE.

We first examine the performance of the three schemes versus the code acquisition time in flat fading channels. Fig. 1 depicts the performances of the three code acquisition algorithms as a function of M, the number of information symbols when $J = 1$ (one receive antenna), $K = 10$ users, SNR $= 15$ dB, and NFR $= 5$ dB. It is seen that the proposed scheme incurs a faster (i.e., smaller M) acquisition time.

It is of interest to consider the performance of the proposed scheme in the presence of modeling errors, e.g., *sampling position errors*. To this end, we generate the received signal by perturbing the sampling instances with a Gaussian random variable with zero mean and standard deviation $0.1T_i$ (that is, a 10% sampling position error). The result is also shown in Fig. 1 (dash-star line). It is seen that relative small sampling errors lead

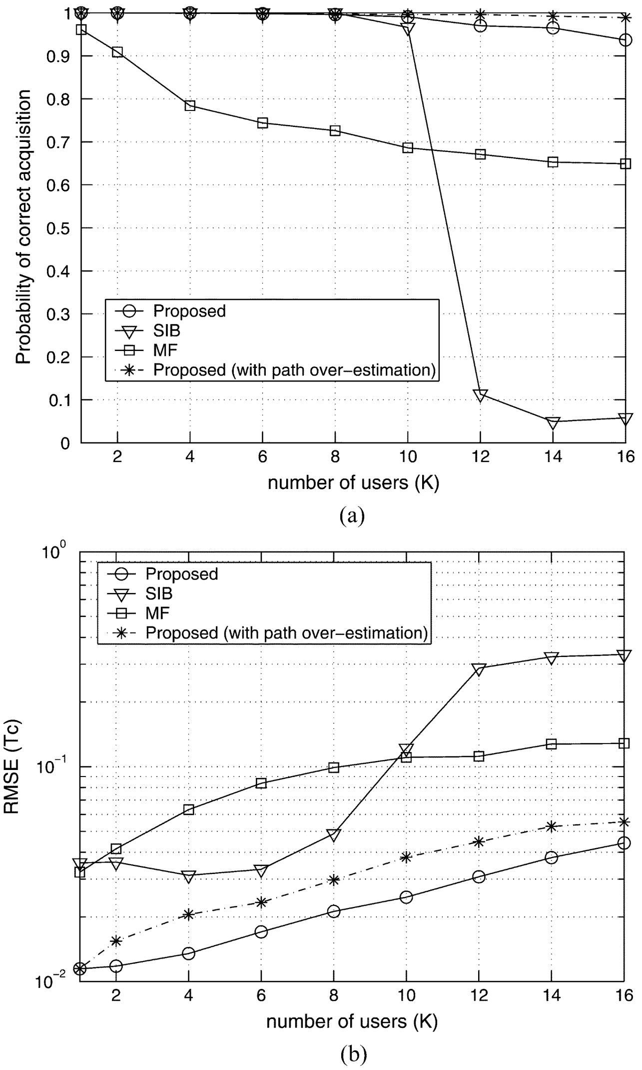


Fig. 2. Performance versus K, the number of users, in flat-fading channels when $J = 1, M = 150, N = 16$, SNR $= 15$ dB, and NFR $= 5$ dB. (a) Probability of correct acquisition. (b) RMSE.

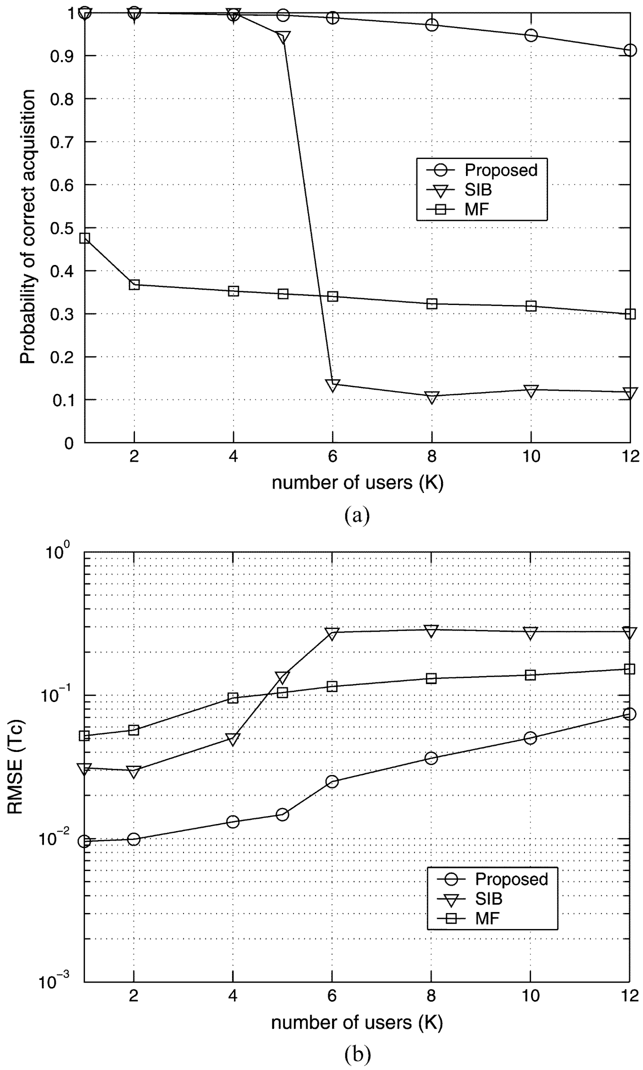


Fig. 3. Performance versus K, the number of users, in multipath fading channels when $J = 1, M = 300, N = 16, L_k = 2\ \forall k$, SNR $= 15$ dB, and NFR $= 10$ dB. (a) Probability of correct acquisition. (b) RMSE.

to negligible performance loss for the proposed scheme, which indicates that our scheme is quite robust to modeling errors. We have also noted that small sampling errors lead to negligible performance loss to the other methods as well; the details are skipped for brevity.

We next examine the user capacity, i.e., the number of users that can be supported by these schemes. The simulation parameters are similar to the previous example except that we fix $M = 150$ and vary K from 1 to 16. The results are depicted in Fig. 2. One can see that the proposed scheme achieves a larger user capacity.

All methods considered herein have implicitly assumed knowledge of L_k, the number of paths for the desired user. To illustrate the performance of the proposed scheme with inaccurate knowledge of L_k, we have included in Fig. 2 the result when the number of paths is *over estimated* (by assuming $L_k = 2$). It is seen that over-estimation leads to negligible degradation in terms of RMSE, at least for the desired path. We have also tried the case when the path number is *under estimated*. In that case, significant performance loss does happen. Hence, we would recommend over-estimating the path number for the proposed scheme if exact knowledge is not available.

Next, consider the user capacity in frequency-selective ($L_k = 2$ for all users) fading channels with NFR $= 10$ dB. The results are shown in Fig. 3. It is seen that the proposed scheme outperforms the others, but there is a performance degradation relative to the flat-fading results in Fig. 2.

We now examine the performance of the schemes versus SNR in multipath fading channels. Fig. 4 depicts the results when $K = 5, L_k = 2$ for all $k, J = 1, M = 150$ and NFR $= 5$ dB. Also shown in Fig. 4(b) is the CRB derived in Section IV. Since the CRB is a function of the propagation delays and attenuations for the desired user, these quantities are fixed from trial to trial in this example. In calculating the CRB, the interference/noise covariance matrix $\mathbf{R_e}$ [cf. (30)–(32)] is replaced by an empirical estimate obtained with 1000 independent realizations. It is seen from Fig. 4(a) that the proposed scheme has a lower SNR threshold. Meanwhile, Fig. 4(b) indicates that the proposed scheme is close to the CRB over a wide range of SNR.

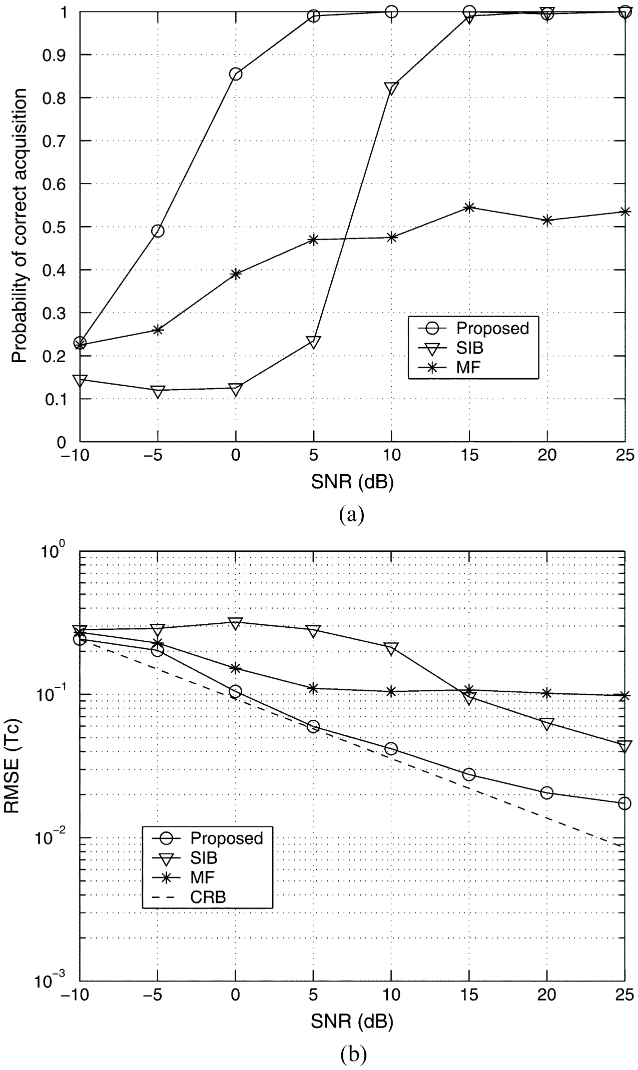


Fig. 4. Performance versus the SNR in multipath fading channels when $J = 1, K = 5, L_k = 2\ \forall k, N = 16, M = 150$, NFR $= 5$ dB. (a) Probability of correct acquisition. (b) RMSE.

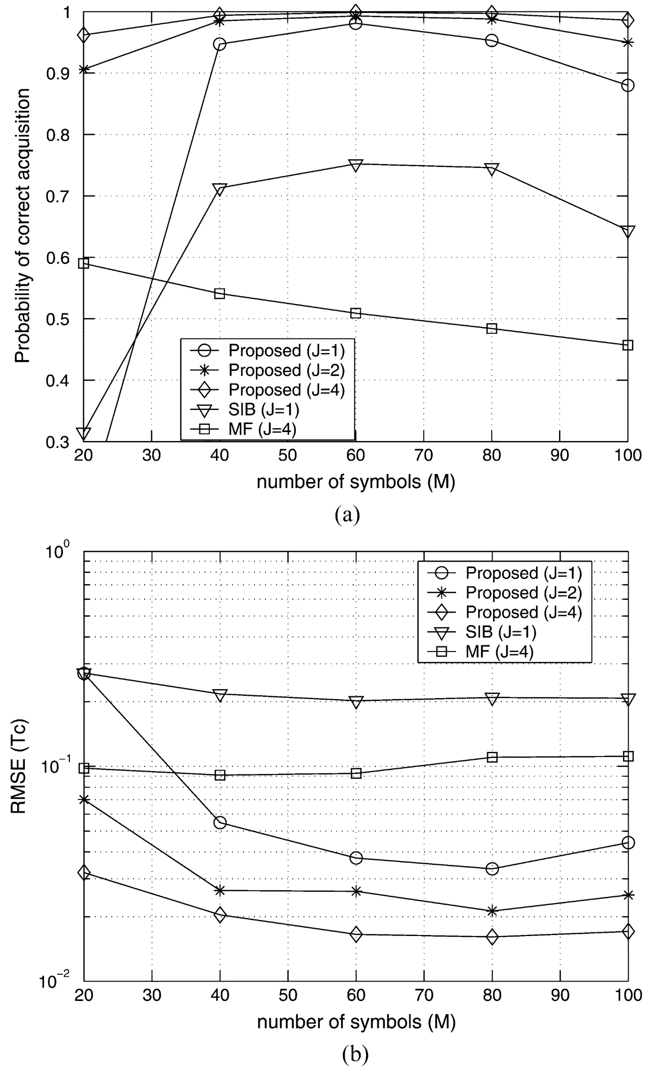


Fig. 5. Performance versus M in time-varying multipath fading channels when $K = 5, L_k = 2\ \forall k, N = 16$, SNR $= 15$ dB and $f_D T_s = 0.0067$. (a) Probability of correct acquisition. (b) RMSE.

From now on, we consider *time- and frequency-selective* channels. The time-varying fading channels are varied sample by sample (i.e., every T_i s, where $T_i \triangleq T_c/Q$ is the sampling interval) according to the Jakes' model [21], which is parameterized by the normalized Doppler rate $f_D T_s$, where f_D denotes the maximum Doppler frequency and T_s the symbol duration. Fig. 5 depicts the performance as a function of M when $K = 5, L_k = 2$ for all $k, f_D T_s = 0.0067$ and SNR $= 15$ dB. Also shown in the figure are the results for the proposed scheme with $J = 2$ and $J = 4$ receive antennas, and the MF estimator with $J = 4$; we do not have the corresponding results for SIB since the estimator is discussed in [18] only for the case of $J = 1$, and the extension to $J > 1$ is nontrivial. Comparing Figs. 1 and 5 for $J = 1$, we note that the proposed scheme is degraded by time-selective channel fading. However, the resistance of the proposed scheme to time-selective channel fading is greatly improved by using multiple receive antennas, as can be seen in Figs. 5(a) and (b) for the case of $J = 2$ and 4.

In the last example, we consider the performance versus the fading rate in time- and frequency-selective channels. The scenario is similar to the previous example except that M is fixed to 40 and the normalized Doppler rate is varying from 0.004 to 0.02. Fig. 6 only shows the results with the proposed scheme and MF for $J = 4$. It is seen that both schemes are affected as the fading rate increases.

Finally, we briefly discuss the relative computational complexity by counting (using Matlab) and comparing the number of flops of each method. We note that the comparison is only to serve the purpose of getting a rough feeling about the relative complexity, since we did not make a special effort to optimize the codes of each method and, furthermore, the comparison may change for different parameters. Consider, for example, the case with $K = 10, L_k = 1$ and $M = 150$, we found that the proposed scheme requires around 10 times more flops than the MF method, while SIB requires about 100 times more. In most scenarios, we found that the SIB is more involved than the proposed scheme.

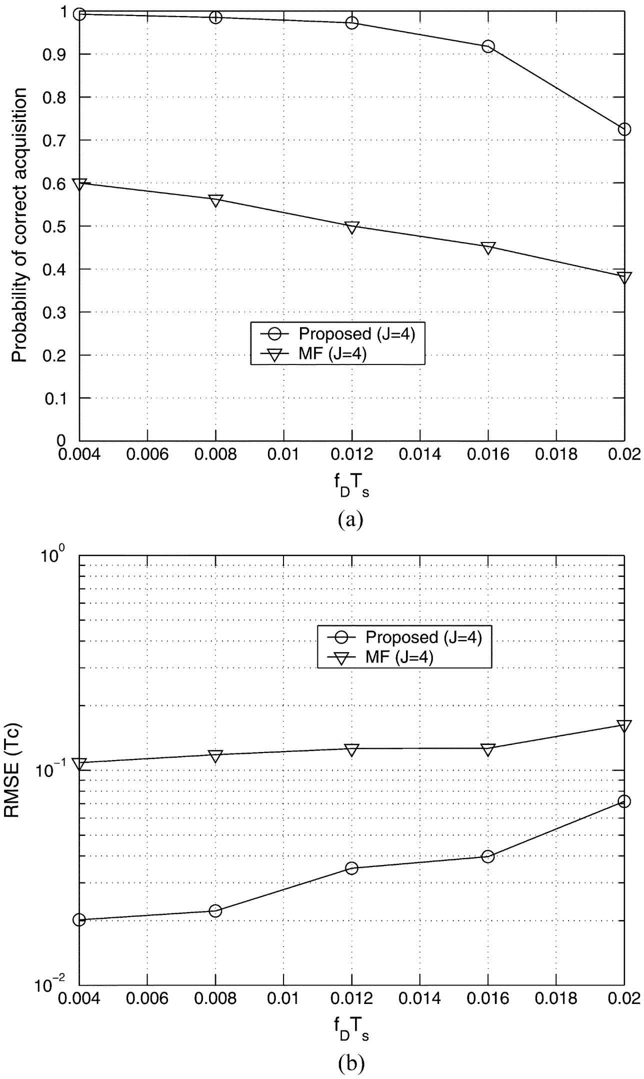


Fig. 6. Performance versus normalized Doppler rate $f_D T_s$ in time-varying multipath fading channels when $N = 16, M = 40, K = 5, L_k = 2\ \forall k$, SNR $= 15$ dB, and NFR $= 5$ dB. (a) Probability of correct acquisition. (b) RMSE.

VI. Conclusion

We have presented a new code-timing estimator for DS-CDMA systems that employ bandlimited chip waveforms. The proposed scheme requires only the training and spreading waveform of the desired user, and can be efficiently implemented by simple noniterative quadratic optimizations. Our scheme can be classified as an asymptotic ML estimator that models the overall interference in the frequency domain as a colored Gaussian process with unknown correlation. If we drop the Gaussian assumption, we may also classify the proposed scheme as a deterministic WLS (estimator that utilizes a weighting matrix for nonlinear least squares fitting in the frequency domain. Irrespective of how it is interpreted, we have shown that the proposed estimator is resistant to MAI and ISI, and can deal with both time- and frequency-selective channel fading, especially when multiple antennas are employed at the base station.

Since the base station has information of all spreading codes, one possible extension of the proposed scheme is to exploit knowledge of codes of all users by following a similar ML derivation in the frequency domain. The resulting scheme, however, will need simultaneous training for all users. Furthermore, the improvement might be limited since we will have to estimate more unknown parameters (for all users) in this case.

The proposed scheme relies on the assumption that the interference in the frequency domain is stationary from symbol to symbol. This is possible only with short spreading codes. Hence, the proposed scheme cannot be directly applied to long-code systems. Efficient code acquisition for long-code CDMA with bandlimited chip waveforms still remains a challenging research problem that needs further attention.

Appendix A
Derivation of the Proposed Code-Timing Estimator

With the assumption that $\{\mathbf{e}_\mu(m)\}$ are complex Gaussian random vectors with zero-mean and an *arbitrary unknown* covariance matrix $E\{\mathbf{e}_{\mu_1}(m_1)\mathbf{e}_{\mu_2}^H(m_2)\} = \mathbf{R_e}\delta(\mu_1-\mu_2)\delta(m_1 - m_2)$, the log-likelihood function conditioned on $\boldsymbol{\tau}_k, \{\boldsymbol{\alpha}_{\mu,k}\}_{\mu=1}^J$ and $\mathbf{R_e}$ is proportional to (in the following, we sometimes use $\boldsymbol{\Phi}_k$ to denote $\boldsymbol{\Phi}_k(\boldsymbol{\tau}_k)$ for notational brevity)

$$\begin{aligned}C_{\mathrm{ml},0}&(\boldsymbol{\tau}_k,\boldsymbol{\alpha}_{\mu,k},\mathbf{R_e})\\ &= -\mathrm{tr}\left\{\mathbf{R_e}^{-1}\frac{1}{MJ}\sum_{\mu=1}^{J}\sum_{m=0}^{M-1}\right.\\ &\quad\times[\mathbf{y}_\mu(m)-\mathbf{G}_k\boldsymbol{\Phi}_k\boldsymbol{\alpha}_{\mu,k}d_k(m)]\\ &\quad\left.\times[\mathbf{y}_\mu(m)-\mathbf{G}_k\boldsymbol{\Phi}_k\boldsymbol{\alpha}_{\mu,k}d_k(m)]^H\right\} - \ln|\mathbf{R_e}|. \quad (33)\end{aligned}$$

Using standard matrix calculus results (e.g., [20]), we take the partial derivative of the likelihood function with respect to $\mathbf{R_e}$ and setting it zero

$$\begin{aligned}\frac{\partial C_{\mathrm{ml},0}}{\partial \mathbf{R_e}} &= -\mathbf{R_e}^{-1} + \mathbf{R_e}^{-1}\frac{1}{MJ}\sum_{\mu=1}^{J}\sum_{m=0}^{M-1}\\ &\quad\times[\mathbf{y}_\mu(m)-\mathbf{G}_k\boldsymbol{\Phi}_k\boldsymbol{\alpha}_{\mu,k}d_k(m)]\\ &\quad\times[\mathbf{y}_\mu(m)-\mathbf{G}_k\boldsymbol{\Phi}_k\boldsymbol{\alpha}_{\mu,k}d_k(m)]^H\mathbf{R_e}^{-1} = \mathbf{0}\end{aligned} \quad (34)$$

from which we obtain

$$\begin{aligned}\mathbf{R_e}&(\boldsymbol{\tau}_k,\boldsymbol{\alpha}_{\mu,k})\\ &= \frac{1}{MJ}\sum_{\mu=1}^{J}\sum_{m=0}^{M-1}[\mathbf{y}_\mu(m)-\mathbf{G}_k\boldsymbol{\Phi}_k\boldsymbol{\alpha}_{\mu,k}d_k(m)]\\ &\quad\times[\mathbf{y}_\mu(m)-\mathbf{G}_k\boldsymbol{\Phi}_k\boldsymbol{\alpha}_{\mu,k}d_k(m)]^H. \quad (35)\end{aligned}$$

Substituting (35) into (33) and discarding quantities irrelevant to the parameters of interest, one can see that maximizing the ML criterion reduces to minimizing

$$C_{\mathrm{ml},1}(\boldsymbol{\tau}_k,\boldsymbol{\alpha}_{\mu,k}) = \ln|\mathbf{R_e}(\boldsymbol{\tau}_k,\boldsymbol{\alpha}_{\mu,k})|. \quad (36)$$

Next, we rewrite (35) as follows:

$$\begin{aligned}\mathbf{R}_{\mathbf{e}}(\boldsymbol{\tau}_k,\boldsymbol{\alpha}_{\mu,k}) &= \frac{1}{J}\sum_{\mu=1}^{J}\Big[\left(\mathbf{G}_k\boldsymbol{\Phi}_k\boldsymbol{\alpha}_{\mu,k}-\hat{\mathbf{r}}_{\mathbf{y}_\mu d}\right)\left(\mathbf{G}_k\boldsymbol{\Phi}_k\boldsymbol{\alpha}_{\mu,k}-\hat{\mathbf{r}}_{\mathbf{y}_\mu d}\right)^H \\ &\quad + \hat{\mathbf{R}}_{\mathbf{y}_\mu} - \hat{\mathbf{r}}_{\mathbf{y}_\mu d}\hat{\mathbf{r}}_{\mathbf{y}_\mu d}^H\Big] \end{aligned} \tag{37}$$

where $\hat{\mathbf{R}}_{\mathbf{y}_\mu}$ and $\hat{\mathbf{r}}_{\mathbf{y}_\mu d}$ are defined in (18)–(19). One can quickly see that the minimum of $\mathbf{R}_{\mathbf{e}}(\boldsymbol{\tau}_k,\boldsymbol{\alpha}_{\mu,k})$ is given by $\hat{\mathbf{R}}_{\mathbf{e}}$ as shown in (17), whereby it is minimized in the sense that $\mathbf{R}_{\mathbf{e}}(\boldsymbol{\tau}_k,\boldsymbol{\alpha}_{\mu,k})-\hat{\mathbf{R}}_{\mathbf{e}}$ is always nonnegative [26]. Once $\mathbf{R}_{\mathbf{e}}(\boldsymbol{\tau}_k,\boldsymbol{\alpha}_{\mu,k})$ is minimized, so is (36), [27], which is a nondecreasing function of $\mathbf{R}_{\mathbf{e}}$. Hence, $\hat{\mathbf{R}}_{\mathbf{e}}$ in (17) is the ML estimate of the interference/noise covariance matrix.

We rearrange the ML criterion (36) as follows:

$$\begin{aligned}&C_{\mathrm{ml},1}(\boldsymbol{\tau}_k,\boldsymbol{\alpha}_{\mu,k}) \\ &\quad= \ln\left|\frac{1}{J}\sum_{\mu=1}^{J}\left(\mathbf{G}_k\boldsymbol{\Phi}_k\boldsymbol{\alpha}_{\mu,k}-\hat{\mathbf{r}}_{\mathbf{y}_\mu d}\right)\times\left(\mathbf{G}_k\boldsymbol{\Phi}_k\boldsymbol{\alpha}_{\mu,k}-\hat{\mathbf{r}}_{\mathbf{y}_\mu d}\right)^H+\hat{\mathbf{R}}_{\mathbf{e}}\right| \\ &\quad= \ln|\hat{\mathbf{R}}_{\mathbf{e}}| + \ln\left|\mathbf{I}_{2N_s}+\hat{\mathbf{R}}_{\mathbf{e}}^{-1}\frac{1}{J}\sum_{\mu=1}^{J}\left(\mathbf{G}_k\boldsymbol{\Phi}_k\boldsymbol{\alpha}_{\mu,k}-\hat{\mathbf{r}}_{\mathbf{y}_\mu d}\right)\times\left(\mathbf{G}_k\boldsymbol{\Phi}_k\boldsymbol{\alpha}_{\mu,k}-\hat{\mathbf{r}}_{\mathbf{y}_\mu d}\right)^H\right|. \end{aligned} \tag{38}$$

Dropping the first term that is independent of the unknown parameters, the ML estimates of the coding-timing $\boldsymbol{\tau}_k$ and channel coefficients $\{\boldsymbol{\alpha}_{\mu,k}\}_{\mu=1}^{J}$ are given by the minimizer of

$$C_{\mathrm{ml},2}(\boldsymbol{\tau}_k,\boldsymbol{\alpha}_{\mu,k}) = \ln\left|\mathbf{I}_{2N_s}+\hat{\mathbf{R}}_{\mathbf{e}}^{-1}\frac{1}{J}\sum_{\mu=1}^{J}\left(\mathbf{G}_k\boldsymbol{\Phi}_k\boldsymbol{\alpha}_{\mu,k}-\hat{\mathbf{r}}_{\mathbf{y}_\mu d}\right)\times\left(\mathbf{G}_k\boldsymbol{\Phi}_k\boldsymbol{\alpha}_{\mu,k}-\hat{\mathbf{r}}_{\mathbf{y}_\mu d}\right)^H\right|. \tag{39}$$

The cost function is highly nonlinear. Exact minimization requires searches over a $(2J+1)L_k$-dimensional (real) parameter space (note that $\boldsymbol{\alpha}_{\mu,k}$ are complex-valued), which is computationally impractical.

To deal with this difficulty, we invoke a result in [28], which shows that minimizing a cost function of the form (39) is asymptotically (for large data samples) equivalent to minimizing the following WLS cost function:

$$\begin{aligned}&C_{\mathrm{wls},0}(\boldsymbol{\tau}_k,\boldsymbol{\alpha}_{\mu,k}) \\ &\quad= \sum_{\mu=1}^{J}\left(\mathbf{G}_k\boldsymbol{\Phi}_k\boldsymbol{\alpha}_{\mu,k}-\hat{\mathbf{r}}_{\mathbf{y}_\mu d}\right)^H\hat{\mathbf{R}}_{\mathbf{e}}^{-1}\left(\mathbf{G}_k\boldsymbol{\Phi}_k\boldsymbol{\alpha}_{\mu,k}-\hat{\mathbf{r}}_{\mathbf{y}_\mu d}\right) \\ &\quad= \sum_{\mu=1}^{J}\left(\boldsymbol{\Phi}_k\boldsymbol{\alpha}_{\mu,k}-\mathbf{G}_k^{-1}\hat{\mathbf{r}}_{\mathbf{y}_\mu d}\right)^H\mathbf{W}\left(\boldsymbol{\Phi}_k\boldsymbol{\alpha}_{\mu,k}-\mathbf{G}_k^{-1}\hat{\mathbf{r}}_{\mathbf{y}_\mu d}\right) \end{aligned} \tag{40}$$

where the weighting matrix $\mathbf{W}$ is defined in (20). This leads immediately to the conditional channel estimate $\boldsymbol{\alpha}_{\mu,k}$, conditioned on any coding-timing estimates, shown in (27). Substituting the conditional estimate back in (40), one can see that minimizing the WLS criterion reduces to minimizing the following function:

$$\begin{aligned}C_{\mathrm{wls},1}(\boldsymbol{\tau}_k) &= \sum_{\mu=1}^{J}\hat{\mathbf{r}}_{\mathbf{y}_\mu d}^H\mathbf{G}_k^{-H}\mathbf{W}^{1/2}\mathbf{P}^{\perp}_{\mathbf{W}^{1/2}\boldsymbol{\Phi}_k}\mathbf{W}^{1/2}\mathbf{G}_k^{-1}\hat{\mathbf{r}}_{\mathbf{y}_\mu d} \\ &= \mathrm{tr}\left\{\mathbf{P}^{\perp}_{\mathbf{W}^{1/2}\boldsymbol{\Phi}_k}\mathbf{W}^{1/2}\mathbf{G}_k^{-1}\times\left(\sum_{\mu=1}^{J}\hat{\mathbf{r}}_{\mathbf{y}_\mu d}\hat{\mathbf{r}}_{\mathbf{y}_\mu d}^H\right)\mathbf{G}_k^{-H}\mathbf{W}^{1/2}\right\} \end{aligned} \tag{41}$$

where $\mathbf{P}^{\perp}_{\mathbf{W}^{1/2}\boldsymbol{\Phi}_k} \triangleq \mathbf{I}_{2N_s}-\mathbf{W}^{1/2}\boldsymbol{\Phi}_k(\boldsymbol{\Phi}_k^H\mathbf{W}\boldsymbol{\Phi}_k)^{-1}\boldsymbol{\Phi}_k^H\mathbf{W}^{1/2}$ denotes the projection matrix that projects onto the orthogonal complement of the range of $\mathbf{W}^{1/2}\boldsymbol{\Phi}_k$.

The WLS estimate of the code-timing, which also coincides asymptotically with the ML estimate, is the minimizer of the WLS criterion (41). Exact minimization requires searches over an L_k-dimensional parameter space. Although the WLS estimator is significantly simpler than the exact ML estimator, it is still quite involved and may suffer from local convergence due to the nonlinearity of the cost function (41).

To further reduce the complexity, we consider reparameterization of the WLS criterion (41). Our reparameterization is in principle iterative quadratic maximum likelihood (IQML) like (e.g., [29]) in that it solves the parameter estimation problem by polynomial rooting. There are also notable differences. In particular, the original IQML [29] assumes that interference is white with identity covariance matrix. In our scheme, however, the interference is colored and we have to invoke whitening at the various stages, which we shall see next.

Let $\mathbf{b} \triangleq [b_0, b_1, \ldots, b_{L_k}]^T$ be coefficients of the following L_kth-order polynomial:

$$b(z) \triangleq b_0 + b_1 z + \cdots + b_{L_k}z^{L_k} = b_{L_k}\prod_{l=1}^{L_k}(z-\phi_l) \tag{42}$$

where $\phi_l \triangleq e^{-j\pi\tau_{k,l}/(NQ)}$. Let $\mathbf{B}$ be a $2N_s\times(2N_s-L_k)$ Toeplitz matrix formed from $\mathbf{b}$, similar to (24). Equation (42) implies that $\mathbf{B}^H\boldsymbol{\Phi}_k=\mathbf{0}$ and, accordingly

$$(\mathbf{W}^{-1/2}\mathbf{B})^H\mathbf{W}^{1/2}\boldsymbol{\Phi}_k = \mathbf{B}^H\mathbf{W}^{-1/2}\mathbf{W}^{1/2}\boldsymbol{\Phi}_k = \mathbf{0}. \tag{43}$$

This, along with the fact that $\mathrm{rank}(\mathbf{W}^{-1/2}\mathbf{B}) = \mathrm{rank}(\mathbf{B}) = 2N_s - L_k$, suggests that $\mathbf{W}^{-1/2}\mathbf{B}$ spans the orthogonal complement of the range of $\boldsymbol{\Phi}_k^H\mathbf{W}^{1/2}$, viz.

$$\begin{aligned}\mathbf{P}^{\perp}_{\mathbf{W}^{1/2}\boldsymbol{\Phi}_k} &= \mathbf{P}_{\mathbf{W}^{-1/2}\mathbf{B}} \\ &\triangleq \mathbf{W}^{-1/2}\mathbf{B}(\mathbf{B}^H\mathbf{W}^{-1}\mathbf{B})^{-1}\mathbf{B}^H\mathbf{W}^{-1/2}. \end{aligned} \tag{44}$$

Therefore, the WLS criterion (41) is equivalent to minimizing the following reparameterized cost function:

$$\begin{aligned}&C_{\mathrm{rp},1}(\mathbf{b}) \\ &\quad= \mathrm{tr}\left\{\mathbf{P}_{\mathbf{W}^{-1/2}\mathbf{B}_k}\mathbf{W}^{1/2}\mathbf{G}_k^{-1}\left(\sum_{\mu=1}^{J}\hat{\mathbf{r}}_{\mathbf{y}_\mu d}\hat{\mathbf{r}}_{\mathbf{y}_\mu d}^H\right)\mathbf{G}_k^{-H}\mathbf{W}^{1/2}\right\} \\ &\quad= \mathrm{tr}\left\{(\mathbf{B}^H\mathbf{W}^{-1}\mathbf{B})^{-1}\left(\sum_{\mu=1}^{J}\mathbf{B}^H\boldsymbol{\gamma}_\mu\boldsymbol{\gamma}_\mu^H\mathbf{B}\right)\right\} \end{aligned} \tag{45}$$

where $\boldsymbol{\gamma}_\mu$ is defined in (22). Note that $\mathbf{B}^H$ is effectively a convolutional matrix, and $\mathbf{B}^H\boldsymbol{\gamma}_\mu$ computes the linear convolution between $\mathbf{b}$ and $\boldsymbol{\gamma}_\mu$. Since convolution is commutative, we can write

$$\mathbf{B}^H\boldsymbol{\gamma}_\mu = \boldsymbol{\Gamma}_\mu\mathbf{b} \tag{46}$$

where $\boldsymbol{\Gamma}_\mu$ is a Hankel matrix formed from $\boldsymbol{\gamma}_\mu$ as shown in (21). Hence, (45) is equivalent to

$$C_{\mathrm{rp},1}(\mathbf{b}) = \mathbf{b}^H\left[\sum_{\mu=1}^{J}\boldsymbol{\Gamma}_\mu^H(\mathbf{B}^H\mathbf{W}^{-1}\mathbf{B})^{-1}\boldsymbol{\Gamma}_\mu\right]\mathbf{b}. \tag{47}$$

It is straightforward to show that $\boldsymbol{\gamma}_\mu$ is a consistent estimate of $\boldsymbol{\Phi}_k\boldsymbol{\alpha}_{\mu,k}$, i.e., $\boldsymbol{\Phi}_k\boldsymbol{\alpha}_{\mu,k} - \boldsymbol{\gamma}_\mu = O(1/\sqrt{M})$. Since $\mathbf{B}^H\boldsymbol{\Phi}_k = 0$, we have $\mathbf{B}^H\boldsymbol{\gamma}_\mu = \boldsymbol{\Gamma}_\mu\mathbf{b} = O(1/\sqrt{M})$. Hence, within a second-order approximation [i.e., by ignoring terms smaller than $O(1/M)$], the $(\mathbf{B}^H\mathbf{W}^{-1}\mathbf{B})^{-1}$ in (47) can be replaced by a consistent estimate without affecting the asymptotic statistical properties of the minimizer of (47) (see, e.g., [30]). A consistent estimate of $(\mathbf{B}^H\mathbf{W}^{-1}\mathbf{B})^{-1}$ can be formed by using an initial consistent estimate of $\mathbf{b}$. One such estimate was described in Step 2 of the proposed estimator in Section III. In particular, that estimate is obtained by a least-squares (LS) fitting to $\boldsymbol{\Gamma}_\mu$, which is consistent due to the consistency of $\boldsymbol{\gamma}_\mu$. Once we have $\mathbf{b}^{(0)}$ as in Step 2, we use it to form $(\hat{\mathbf{B}}^{(0)H}\mathbf{W}^{-1}\hat{\mathbf{B}}^{(0)})^{-1}$, which is a consistent estimate of $(\mathbf{B}^H\mathbf{W}^{-1}\mathbf{B})^{-1}$, and recompute $\mathbf{b}$ according to Step 3. Clearly, the so-obtained estimate of $\mathbf{b}$ approaches asymptotically to the ML estimate, and so are the code-timing estimates.

APPENDIX B
MINIMIZATION OF QUADRATIC FUNCTIONS (23) and (25)

Both cost functions in (23) and (25) can be written as a common form

$$C(\mathbf{b}) = \mathbf{b}^H\left(\sum_{\mu=1}^{J}\boldsymbol{\Gamma}_\mu^H\boldsymbol{\Omega}\boldsymbol{\Gamma}_\mu\right)\mathbf{b} \tag{48}$$

where $\boldsymbol{\Omega} = \mathbf{I}$ for (23) and $\boldsymbol{\Omega} = (\hat{\mathbf{B}}^{(0)H}\mathbf{W}^{-1}\hat{\mathbf{B}}^{(0)})^{-1}$ for (25). Since, theoretically, the polynomial $b(z)$ in (42) has all its roots on the unit circle, a necessary condition on $\mathbf{b}$ is that $\mathbf{b}$ should satisfy the conjugate symmetry property (cf. [30])

$$b_l = b_{L_k-l}, \quad l = 0,1,\ldots,L_k. \tag{49}$$

We would also like to enforce this constraint in our estimation of $\mathbf{b}$. To this end, let $\tilde{\mathbf{b}} = [\mathrm{Re}(b_0),\ldots,\mathrm{Re}(b_{\lfloor L_k/2\rfloor}), \mathrm{Im}(b_0),\ldots,\mathrm{Im}(b_{\lfloor(L_k-1)/2\rfloor})]^T$, where $\lfloor L_k/2\rfloor$ denotes the integer part of $L_k/2$. Then, we have

$$L_k \text{ odd}: \ \mathbf{b} = \begin{bmatrix}\mathbf{I}_{(L_k+1)/2} & j\mathbf{I}_{(L_k+1)/2}\\ \mathbf{J}_{(L_k+1)/2} & -j\mathbf{J}_{(L_k+1)/2}\end{bmatrix}\tilde{\mathbf{b}} \triangleq \mathbf{S}\tilde{\mathbf{b}} \tag{50}$$

$$L_k \text{ even}: \ \mathbf{b} = \begin{bmatrix}\mathbf{I}_{L_k/2} & \mathbf{0}_{(L_k/2)\times 1} & j\mathbf{I}_{L_k/2}\\ \mathbf{0}_{1\times(L_k/2)} & 1 & \mathbf{0}_{1\times(L_k/2)}\\ \mathbf{J}_{L_k/2} & \mathbf{0}_{(L_k/2)\times 1} & -j\mathbf{J}_{L_k/2}\end{bmatrix}\tilde{\mathbf{b}} \triangleq \mathbf{S}\tilde{\mathbf{b}} \tag{51}$$

where $\mathbf{J}_{L_k/2}$ denotes a $(L_k/2)\times(L_k/2)$ exchange matrix with ones on its cross-diagonal and zeros elsewhere. Using (50) or (51) in (48), we have

$$\hat{\tilde{\mathbf{b}}} = \arg\min_{\tilde{\mathbf{b}}\in\mathbb{R}^{(L_k+1)\times 1}}\tilde{\mathbf{b}}^T\left[\mathbf{S}^H\left(\sum_{\mu=1}^{J}\boldsymbol{\Gamma}_\mu^H\boldsymbol{\Omega}\boldsymbol{\Gamma}_\mu\right)\mathbf{S}\right]\tilde{\mathbf{b}}. \tag{52}$$

To avoid the trivial solution $\tilde{\mathbf{b}} = 0$, we also impose the constraint $\|\tilde{\mathbf{b}}\| = 1$. The solution to (52) is the eigenvector of $\mathrm{Re}\{\mathbf{S}^H[\sum_{\mu=1}^{J}\boldsymbol{\Gamma}_\mu^H\boldsymbol{\Omega}\boldsymbol{\Gamma}_\mu]\mathbf{S}\}$ associated with the smallest eigenvalue (e.g., [26]).

APPENDIX C
DERIVATION OF THE CRB

Let $\mathbf{x}_\mu(m) \triangleq d_k(m)\mathbf{G}_k\boldsymbol{\Phi}_k(\boldsymbol{\tau}_k)\boldsymbol{\alpha}_{\mu,k}$. Then, (28) is written as

$$\mathbf{y}_\mu(m) = \mathbf{x}_\mu(m) + \mathbf{e}_\mu(m), \quad m = 0,1,\ldots,M-1, \quad \mu = 1,\ldots,J \tag{53}$$

where $\mathbf{e}_\mu(m)$ is assumed to be a zero-mean colored Gaussian random process with unknown covariance matrix $E\{\mathbf{e}_{\mu_1}(m_1)\mathbf{e}_{\mu_2}^H(m_2)\} = \mathbf{R}_\mathbf{e}\delta(\mu_1-\mu_2)\delta(m_1-m_2)$. Next, we collect all the received vectors and form $\mathbf{y} \triangleq [\mathbf{y}_1^T(0),\ldots,\mathbf{y}_J^T(0),\ldots,\mathbf{y}_1^T(M-1),\ldots,\mathbf{y}_J^T(M-1)]^T$. Moreover, let $\mathbf{x} \triangleq [\mathbf{x}_1^T(0),\ldots,\mathbf{x}_J^T(0),\ldots,\mathbf{x}_1^T(M-1),\ldots,\mathbf{x}_J^T(M-1)]^T$ and $\mathbf{e} \triangleq [\mathbf{e}_1^T(0),\ldots,\mathbf{e}_J^T(0),\ldots,\mathbf{e}_1^T(M-1),\ldots,\mathbf{e}_J^T(M-1)]^T$. Then, (53) can be rewritten compactly as

$$\mathbf{y} = \mathbf{x} + \mathbf{e}. \tag{54}$$

We observe the following structure of $\mathbf{x}$:

$$\mathbf{x} = (\mathbf{I}_M \otimes \{[\mathbf{I}_J \otimes (\mathbf{G}_k\boldsymbol{\Phi}_k(\boldsymbol{\tau}_k))]\boldsymbol{\alpha}_k\})\mathbf{d}_k \tag{55}$$

where $\boldsymbol{\alpha}_k \triangleq [\boldsymbol{\alpha}_{1,k}^T,\ldots,\boldsymbol{\alpha}_{J,k}^T]^T$ and $\mathbf{d}_k \triangleq [d_k(0),d_k(1),\ldots,d_k(M-1)]^T$. The covariance matrix of $\mathbf{e}$ is given by

$$\boldsymbol{\Sigma} \triangleq E\{\mathbf{e}\mathbf{e}^H\} = \mathbf{I}_M \otimes (\mathbf{I}_J \otimes \mathbf{R}_\mathbf{e}). \tag{56}$$

By using the Slepian–Bangs formula [25], the (i,j)th element of the Fisher information matrix is given by

$$[\mathrm{CRB}^{-1}(\boldsymbol{\theta})]_{i,j} = 2\mathrm{Re}\left(\frac{\partial\mathbf{x}^H}{\partial[\boldsymbol{\theta}]_i}\boldsymbol{\Sigma}^{-1}\frac{\partial\mathbf{x}}{\partial[\boldsymbol{\theta}]_j}\right). \tag{57}$$

Let $\mathbf{F}_{\boldsymbol{\tau}} \triangleq [(\partial\mathbf{x})/(\partial\tau_{k,1}),\ldots,(\partial\mathbf{x})/(\partial\tau_{k,L_k})]$ and $\mathbf{F}_{\boldsymbol{\alpha}} \triangleq [(\partial\mathbf{x})/(\partial\boldsymbol{\alpha}_{1,k}^T),\ldots,(\partial\mathbf{x})/(\partial\boldsymbol{\alpha}_{J,k}^T)]$. Then, we can write (57) collectively as

$$\mathrm{CRB}^{-1}(\boldsymbol{\theta}) = 2\mathrm{Re}\left\{[\mathbf{F}_{\boldsymbol{\tau}},\mathbf{F}_{\boldsymbol{\alpha}},j\mathbf{F}_{\boldsymbol{\alpha}}]^H\left(\mathbf{I}_M\otimes\left(\mathbf{I}_J\otimes\mathbf{R}_\mathbf{e}^{-1}\right)\right) \times [\mathbf{F}_{\boldsymbol{\tau}},\mathbf{F}_{\boldsymbol{\alpha}},j\mathbf{F}_{\boldsymbol{\alpha}}]\right\}. \tag{58}$$

By direct calculation, we have

$$\mathbf{F}_{\boldsymbol{\tau}} = [(\mathbf{I}_M\otimes\mathbf{z}_1)\mathbf{d}_k,\ldots,(\mathbf{I}_M\otimes\mathbf{z}_{L_k})\mathbf{d}_k] \tag{59}$$

$$\mathbf{F}_{\boldsymbol{\alpha}} = [(\mathbf{I}_M\otimes\mathbf{a}_{1,1})\mathbf{d}_k,\ldots,(\mathbf{I}_M\otimes\mathbf{a}_{J,L_k})\mathbf{d}_k] \tag{60}$$

where $\mathbf{z}_l \triangleq [\mathbf{I}_J\otimes(\mathbf{G}_k(\partial\boldsymbol{\Phi}_k(\boldsymbol{\tau}_k))/(\partial\tau_{k,l}))]\boldsymbol{\alpha}_k$ and $\mathbf{a}_{\mu,l} \triangleq [\mathbf{I}_J\otimes(\mathbf{G}_k\boldsymbol{\Phi}_k(\boldsymbol{\tau}_k))](\partial\boldsymbol{\alpha}_k)/(\partial\alpha_{\mu,k,l})$, for $l = 1,\ldots,L_k, \mu =$

$1, \ldots, J$. The partial derivatives with respect to $\tau_{k,l}$ and $\alpha_{\mu,k,l}$ are

$$\frac{\partial \boldsymbol{\Phi}_k(\boldsymbol{\tau}_k)}{\partial \tau_{k,l}} = \left[\mathbf{0}_{2N_s \times (l-1)}, \boldsymbol{\psi}_l, \mathbf{0}_{2N_s \times (L_k - l)}\right] \tag{61}$$

$$\frac{\boldsymbol{\alpha}_k}{\partial \alpha_{\mu,k,l}} = \left[\mathbf{0}_{1\times[(\mu-1)L_k+(l-1)]}, 1, \mathbf{0}_{1\times[(J-\mu+1)L_k-l]}\right]^T \tag{62}$$

where $\boldsymbol{\psi}_l = j(\pi/NQ)[N_s\phi^{-N_s\tau_{k,l}}, (N_s - 1)\phi^{-(N_s-1)\tau_{k,l}}, \ldots, 0, -\phi^{\tau_{k,l}}, \ldots, -(N_s-1)\phi^{(N_s-1)\tau_{k,l}}]^T$.

Following (58) and (59), the (i,j)th element of the matrix $\mathbf{F}_{\boldsymbol{\tau}\boldsymbol{\tau}}$ is given by

$$\begin{aligned}
[\mathbf{F}_{\boldsymbol{\tau\tau}}]_{i,j} &= \mathbf{d}_k^H \left(\mathbf{I}_M \otimes \mathbf{z}_i^H\right)\left(\mathbf{I}_M \otimes \left(\mathbf{I}_J \otimes \mathbf{R}_{\mathbf{e}}^{-1}\right)\right)(\mathbf{I}_M \otimes \mathbf{z}_j)\mathbf{d}_k \\
&= \mathbf{d}_k^H \left(\mathbf{I}_M \otimes \left[\mathbf{z}_i^H(\mathbf{I}_J \otimes \mathbf{R}_{\mathbf{e}}^{-1})\mathbf{z}_j\right]\right)\mathbf{d}_k \\
&= M\mathbf{z}_i^H\left(\mathbf{I}_J \otimes \mathbf{R}_{\mathbf{e}}^{-1}\right)\mathbf{z}_j \\
&= M\sum_{\mu=1}^{J} \boldsymbol{\alpha}_{\mu,k}^H \frac{\partial \boldsymbol{\Phi}_k^H(\boldsymbol{\tau}_k)}{\partial \tau_{k,i}} \mathbf{G}_k^H \mathbf{R}_{\mathbf{e}}^{-1} \\
&\quad \times \mathbf{G}_k \frac{\partial \boldsymbol{\Phi}_k(\boldsymbol{\tau}_k)}{\partial \tau_{k,j}} \boldsymbol{\alpha}_{\mu,k} \\
&= M\sum_{\mu=1}^{J} (\boldsymbol{\psi}_i \alpha_{\mu,k,i})^H \mathbf{G}_k^H \mathbf{R}_{\mathbf{e}}^{-1} \mathbf{G}_k (\boldsymbol{\psi}_j \alpha_{\mu,k,j})
\end{aligned} \tag{63}$$

where in the second equality, we used the fact that $(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D}) = (\mathbf{AC})\otimes(\mathbf{BD})$ for arbitrary $\mathbf{A}, \mathbf{B}, \mathbf{C}$, and $\mathbf{D}$ with compatible dimensions [20]. Equation (63) is recognized as the (i,j)th element of the right-hand side (RHS) of (30).

Next, we evaluate the (i,j)th element of the matrix $\mathbf{F}_{\boldsymbol{\tau\alpha}}$. Let $j = (\mu-1)J + l$, for $\mu = 1, \ldots, J$ and $l = 1, \ldots, L_k$. We have

$$\begin{aligned}
[\mathbf{F}_{\boldsymbol{\tau\alpha}}]_{i,j} &= [\mathbf{F}_{\boldsymbol{\tau\alpha}}]_{i,(\mu-1)J+l} \\
&= \mathbf{d}_k^H \left(\mathbf{I}_M \otimes \left[\mathbf{z}_i^H(\mathbf{I}_J \otimes \mathbf{R}_{\mathbf{e}}^{-1})\mathbf{a}_{\mu,l}\right]\right)\mathbf{d}_k \\
&= M\boldsymbol{\alpha}_k^H \left(\mathbf{I}_J \otimes \left[\frac{\partial \boldsymbol{\Phi}_k^H(\boldsymbol{\tau}_k)}{\partial \tau_{k,i}}\right.\right. \\
&\quad \left.\left. \times\ \mathbf{G}_k^H \mathbf{R}_{\mathbf{e}}^{-1} \mathbf{G}_k \mathbf{R}_{\mathbf{e}}^{-1} \mathbf{G}_k \boldsymbol{\Phi}_k(\boldsymbol{\tau}_k)\right]\right) \frac{\partial \boldsymbol{\alpha}_k}{\partial \alpha_{\mu,k,l}} \\
&= M\boldsymbol{\alpha}_{\mu,k}^H \frac{\partial \boldsymbol{\Phi}_k^H(\boldsymbol{\tau}_k)}{\partial \tau_{k,i}} \mathbf{G}_k^H \mathbf{R}_{\mathbf{e}}^{-1} \\
&\quad \times \mathbf{G}_k \boldsymbol{\Phi}_k(\boldsymbol{\tau}_k) \frac{\partial \boldsymbol{\alpha}_{\mu,k}}{\partial \alpha_{\mu,k,l}} \\
&= M(\boldsymbol{\psi}_i \alpha_{\mu,k,i})^H \mathbf{G}_k^H \mathbf{R}_{\mathbf{e}}^{-1} \mathbf{G}_k \boldsymbol{\phi}_l
\end{aligned} \tag{64}$$

where we recall $\boldsymbol{\phi}_l$ is the lth column of $\boldsymbol{\Phi}_k(\boldsymbol{\tau}_k)$. One can see that (64) is the (i,j)th element of the RHS of (31).

Similarly, for $\mathbf{F}_{\boldsymbol{\alpha\alpha}}$, let $i = (\mu_1 - 1)J + l_1$, where $\mu_1 = 1, \ldots, J$ and $l_1 = 1, \ldots, L_k$; $j = (\mu_2 - 1)J + l_2$, where $\mu_2 = 1, \ldots, J$ and $l_2 = 1, \ldots, L_k$. Then

$$\begin{aligned}
[\mathbf{F}_{\boldsymbol{\alpha\alpha}}]_{i,j} &= [\mathbf{F}_{\boldsymbol{\alpha\alpha}}]_{(\mu_1-1)J+l_1,(\mu_2-1)J+l_2} \\
&= \mathbf{d}_k^H \left(\mathbf{I}_M \otimes \left[\mathbf{a}_{\mu_1,l_1}^H(\mathbf{I}_J \otimes \mathbf{R}_{\mathbf{e}}^{-1})\mathbf{a}_{\mu_2,l_2}\right]\right)\mathbf{d}_k \\
&= M\frac{\partial \boldsymbol{\alpha}_k^H}{\partial \alpha_{\mu_1,k,l_1}} \left(\mathbf{I}_J \otimes \left[\boldsymbol{\Phi}_k^H(\boldsymbol{\tau}_k)\right.\right. \\
&\quad \left.\left. \times\ \mathbf{G}_k^H \mathbf{R}_{\mathbf{e}}^{-1} \mathbf{G}_k \boldsymbol{\Phi}_k(\boldsymbol{\tau}_k)\right]\right) \frac{\partial \boldsymbol{\alpha}_k}{\partial \alpha_{\mu_2,k,l_2}} \\
&= M\boldsymbol{\phi}_{l_1}^H \mathbf{G}_k^H \mathbf{R}_{\mathbf{e}}^{-1} \mathbf{G}_k \boldsymbol{\phi}_{l_2} \delta(\mu_1 - \mu_2)
\end{aligned} \tag{65}$$

which is seen to coincide with the (i,j)th element of the RHS of (32).

ACKNOWLEDGMENT

The authors would like to thank N. Petrochilos and A. J. van der Veen for sharing their Matlab code of the shift-invariance based (SIB) code-timing estimator in [18], and the anonymous reviewers for their constructive comments and suggestions.

REFERENCES

[1] R. Prasad and T. Ojanperä, "An overview of CDMA evolution toward wideband CDMA," *IEEE Commun. Surv.*, vol. 1, no. 1, pp. 2–29, 1998.
[2] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.
[3] R. L. Peterson, R. E. Ziemer, and D. E. Borth, *Introduction to Spread Spectrum Communications*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
[4] R. F. Smith and S. L. Miller, "Acquisition performance of an adaptive receiver for DS-CDMA," *IEEE Trans. Commun.*, vol. 47, pp. 1416–1424, Sept. 1999.
[5] D. Zheng, J. Li, S. L. Miller, and E. G. Ström, "An efficient code-timing estimator for DS-CDMA system," *IEEE Trans. Signal Processing*, vol. 45, pp. 82–89, Jan. 1997.
[6] S. E. Bensley and B. Aazhang, "Maximum-likelihood synchronization of a single user for code-division multiple-access communication systems," *IEEE Trans. Commun.*, vol. 46, pp. 392–399, Mar. 1998.
[7] E. G. Ström and F. Malmsten, "A maximum likelihood approach for estimating DS-CDMA multipath fading channels," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 132–140, Jan. 2000.
[8] H. Li, J. Li, and S. L. Miller, "Decoupled multiuser code-timing estimation for code-division multiple-access communication systems," *IEEE Trans. Commun.*, vol. 49, pp. 1425–1436, Aug. 2001.
[9] E. G. Ström, S. Parkvall, S. L. Miller, and B. E. Ottersten, "Propagation delay estimation in asynchronous direct-sequence code-division multiple access systems," *IEEE Trans. Commun.*, vol. 44, pp. 84–93, Jan. 1996.
[10] S. E. Bensley and B. Aazhang, "Subspace-based channel estimation for code division multiple access communications systems," *IEEE Trans. Commun.*, vol. 44, pp. 1009–1020, Aug. 1996.
[11] E. G. Ström, S. Parkvall, S. L. Miller, and B. E. Ottersten, "DS-CDMA synchronization in time-varying fading channels," *IEEE J. Select. Areas Commun.*, vol. 14, pp. 1636–1642, Oct. 1996.
[12] T. Östman, S. Parkvall, and B. Ottersten, "An improved MUSIC algorithm for estimation of time delays in asynchronous DS-CDMA systems," *IEEE Trans. Commun.*, vol. 47, pp. 1628–1631, Nov. 1999.
[13] U. Madhow, "Blind adaptive interference suppression for the near-far resistant acquisition and demodulation of direct-sequence CDMA signals," *IEEE Trans. Signal Processing*, vol. 45, pp. 124–136, Jan. 1997.
[14] M. Latva-aho, J. Lilleberg, J. Iinatti, and M. Juntti, "CDMA downlink code acquisition performance in frequency-selective fading channels," in *Proc. IEEE Int. Symp. Personal, Indoor and Mobile Radio Communication (PIMRC'98)*, Boston, MA, Sept. 1998, pp. 1476–1479.
[15] H. Li and R. Wang, "Filterbank-based blind code synchronization for DS-CDMA systems in multipath fading channels," *IEEE Trans. Signal Processing*, vol. 51, pp. 160–161, Jan. 2003.
[16] J. G. Proakis, *Digital Communications*, 3rd ed. New York: McGraw-Hill, 1995.
[17] T. Östman and B. Ottersten, "Near far robust time delay estimation for asynchronous DS-CDMA systems with bandlimited pulse shapes," in *Proc. IEEE 48th Vehicular Technology Conf.*, Ottawa, ON, Canada, May 1998, pp. 1650–1654.
[18] N. Petrochilos and A. J. van der Veen, "Blind time delay estimation in asynchronous CDMA via subspace intersection and ESPRIT," in *Proc. 2001 IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP'2001)*, Salt Lake City, UT, May 2001.
[19] R. Roy and T. Kailath, "ESPRIT—Estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoustics, Speech, Signal Processing*, vol. 37, pp. 984–995, July 1989.
[20] A. Graham, *Kronecker Products and Matrix Calculus with Applications*. Chichester, U.K.: Ellis Horwood, 1981.
[21] W. C. Jakes Jr., *Microwave Mobile Communications*. New York: Wiley-Interscience, 1974.

[22] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[23] A. J. van der Veen, M. C. Vanderveen, and A. Paulraj, "Joint angle and delay estimation using shift-invariance techniques," *IEEE Trans. Signal Processing*, vol. 46, pp. 405–418, Feb. 1998.

[24] A. V. Oppenheim and R. W. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1989.

[25] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1997.

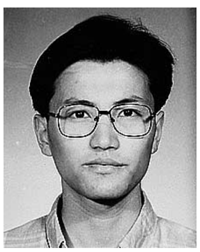
[26] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.

[27] T. Söderström and P. Stoica, *System Identification*. London, U.K.: Prentice-Hall, 1989.

[28] J. Li, B. Halder, P. Stoica, and M. Viberg, "Computationally efficient angle estimation for signals with known waveforms," *IEEE Trans. Signal Processing*, vol. 43, pp. 2154–2163, Sept. 1995.

[29] A. L. Swindlehurst, "Time delay and spatial signature estimation using known asynchronous signals," *IEEE Trans. Signal Processing*, vol. 46, pp. 449–462, Feb. 1998.

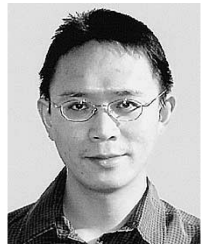
[30] P. Stoica and K. C. Sharman, "Novel eigenanalysis method for direction estimation," *Proc. Inst. Elect. Eng.*, pt. F, vol. 137, no. 1, pp. 19–26, Feb. 1990.



Rensheng Wang (S'03) received the B.E. and M.E. degrees from Harbin Institute of Technology, Harbin, China, in 1995 and 1997, respectively, both in electrical engineering. He is currently working toward the Ph.D. degree in electrical engineering at the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ.

From 1997 to 2000, he was a Researcher at the Institute of Electronics, Chinese Academy of Sciences, Beijing. He is currently working as a Research Assistant with the Department of Electrical and Computer Engineering, Stevens Institute of Technology. His research interests include signal processing for communications with focus on multiuser detection and estimation in CDMA.

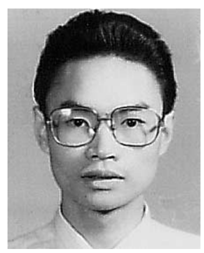
Mr. Wang received the Outstanding Research Assistant Award in 2002 and the Edward Peskin Award in 2004 from Stevens Institute of Technology.



Hongbin Li (M'99) received the B.S. and M.S. degrees from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 1991 and 1994, respectively, and the Ph.D. degree from the University of Florida, Gainesville, in 1999, all in electrical engineering.

From July 1996 to May 1999, he was a Research Assistant with the Department of Electrical and Computer Engineering, University of Florida. Since July 1999, he has been an Assistant Professor with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ. He was a Summer Visiting Faculty Member at the Air Force Research Laboratory, Rome, NY, in the summers of 2003 and 2004. His current research interests include stochastic signal processing, sensor array processing, wireless communications, and radar imaging.

Dr. Li is an Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS. He received the Harvey N. Davis Teaching Award in 2003 and the Jess H. Davis Memorial Award for excellence in research from Stevens Institute of Technology in 2001, and the Sigma Xi Graduate Research Award from the University of Florida in 1999. He is a member of Tau Beta Pi and Phi Kappa Phi.



Tao Li received the B.S. and M.S. degrees in computer science and the Ph.D. degree in electrical engineering from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, in 1986, 1991, and 1995, respectively.

From 1993 to 1994, he was a Visiting Scholar at the University of California, Berkeley. He is currently a Professor with the Department of Computer, Sichuan University, Chengdu, where he founded and directs the Laboratory of Computer Networks and Information Security. His research interests include computer networks, information security, artificial intelligence, neural networks, artificial immune system, computer communications, and wireless communications.