Decorrelating Multiuser Code-Timing Estimation for Long-Code CDMA With Bandlimited Chip Waveforms

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

***Abstract*—This paper addresses the problem of *multiuser* code-timing estimation for asynchronous uplink code-division multiple-access (CDMA) systems with *aperiodic* spreading codes and *bandlimited* chip waveforms. Two decorrelating code-timing estimation schemes, namely the frequency-domain least-squares (FLS) and frequency-domain weighted least-squares (FWLS) estimators, are developed. The two proposed estimators offer different tradeoffs between complexity and estimation accuracy. A critical step for decorrelating-based estimation is to decompose the received signal into subsignals of shorter duration. We discuss how to perform the decomposition to ensure improved identifiability and statistical stability of the proposed schemes. Due to a unique signal structure in the frequency domain, both the FLS and FWLS estimators admit efficient implementations that result in significant complexity reductions. The Cramér–Rao bound for the estimation problem under study is derived and used as an assessment tool for the proposed estimators. Numerical results show that both of the proposed estimators can support overloaded systems (with more users than the processing gain) in multipath fading environments and significantly outperform a conventional technique based on matched-filter processing.**

***Index Terms*—Aperiodic/long spreading codes, bandlimited chip waveforms, code-division multiple access (CDMA), parameter estimation, synchronization.**

I. Introduction

CODE-TIMING estimation or acquisition is a challenging problem in wireless mobile CDMA systems, due to channel impediments such as fading, multipath propagation, and multiple-access interference (MAI). Traditional acquisition techniques are typically based on matched-filter (MF) processing [1, ch. 5], which models the MAI as an additive white noise without taking into account any inherent structure. They were found inadequate in multiuser environments, especially when there exist significant power variations among users, which cause the so-called near-far problem [2].

A. Prior Work

A wealth of recent studies on multiuser code acquisition aimed to enhance the performance by exploiting the structure of the MAI. They use explicit training (e.g., [3]–[7]), or work in a blind (self-recovering) fashion (e.g., [8]–[13]). While these schemes offer improved performance over the MF-based techniques in near-far environments, they also rely on the crucial assumption that the system employs *short* (or *symbol-periodic*) spreading codes and *rectangular* chip waveforms that are not bandlimited. In contrast, most practical CDMA systems, including the IS-95 standard and the majority of 3G CDMA-based wireless networks (e.g., [14]), use *long* (or *aperiodic*) spreading codes, to randomize the interference, and *bandlimited* chip waveforms, such as the square-root raised-cosine pulses.

In principle, the above techniques can be extended to bandlimited systems with short codes (see, e.g., [15]). Most such extensions require intensive iterative nonlinear searches over the parameter space. Alternatively, bandlimited pulses can be dealt with in the frequency domain, usually at a lower complexity (e.g., [16] and [17]). However, these frequency-domain based schemes, like the time-domain based techniques in [3]–[13], rely on a symbol-level cyclostationary signal structure imposed by short-code spreading and, therefore, cannot be applied to long-code systems.

A major challenge in dealing with long codes is that the aforementioned cyclostationary reliance has to be dropped. Along this line, there have been a number of recent studies on channel estimation for long-code CDMA (e.g., [18]–[21] and references therein). The complementary research on multiuser code-timing estimation or acquisition, however, appears more scarce, partly due to the nonlinear nature of the problem. Among limited studies, Mantravadi and Veeravalli [22] developed an acquisition scheme for a scenario in which the code-timing of only one user is to be estimated, while the code-timings and spreading codes for all other users are assumed known. A similar scheme for frequency-selective channels was considered in [23]. More recently, Buzzi and Poor [24] proposed both centralized and decentralized acquisition schemes in flat-fading channels. Their schemes, however, assume rectangular chip waveforms. As noted by the authors, it is nontrivial to extend their schemes to the bandlimited case.

B. Current Contributions

In this paper, we consider multiuser code-timing estimation for asynchronous (i.e., uplink) CDMA with long spreading codes and bandlimited chip waveforms in frequency-selective channels. Like the centralized scheme of [24], we assume that the data symbols and spreading codes for all users are

Manuscript received January 22, 2004; revised August 15, 2004. This work was supported in part by the Army Research Office under Grant DAAD19-03-1-0184 and by the New Jersey Commission on Science and Technology. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Helmut Bölcskei.

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

Digital Object Identifier 10.1109/TSP.2005.849177

known, and seek joint estimation of their code-timings and channel attenuations. We consider decorrelating based estimation techniques which entail relatively low complexity. Two decorrelating code-timing estimators are developed, one based on least-squares (LS) while the other on weighted LS (WLS) processing in the frequency domain. For brevity, they will be henceforth referred to as the FLS and FWLS estimators. A critical step for decorrelating based estimation is to decompose the received signal into subsignals of shorter duration. We discuss how to form such subsignals, in a way leading to improved identifiability and numerical stability of the proposed estimators.

While decorrelating based code acquisition is known for short-code CDMA (e.g., [7]), direct application to long-code CDMA is impractical due to excessive computational complexity and storage requirement (see [21] for discussions on decorrelating based channel estimation for long-code CDMA). In this paper, we develop efficient implementations of the proposed decorrelating FLS and FWLS estimators, which is made possible by exploiting a unique signal structure in the frequency domain. This results in reduction of complexity by *orders of magnitude* compared with direct implementations. Another contribution of this work is an expression of the CRB for the problem under study (note that the CRB in [24] is for rectangular chip pulses). Numerical simulations show that the proposed FLS and FWLS estimators can support *overloaded* systems (systems with more users than the processing gain) in frequency-selective channels and achieve significantly better acquisition performance than a MF-based technique.

The rest of the paper is organized as follows. In Section II, we introduce the data model for bandlimited long-code CDMA, and formulate the problem of interest. In Section III, we develop the decorrelating FLS and FWLS estimators. Their efficient implementations are presented in Section IV. The CRB expression is derived in Section V, followed by numerical results in Section VI. Finally, we summarize this work in Section VII.

Notation: Vectors (matrices) are denoted by boldface lower (upper) case letters; all vectors are column vectors; superscripts $(\cdot)^T$, $(\cdot)^*$, and $(\cdot)^H$ denote transpose, conjugate, and conjugate transpose, respectively; $\star$ denotes linear convolution; $\lceil\cdot\rceil$ denotes the smallest integer no less than the argument; $\mathbf{I}_N$ denotes the $N\times N$ identity matrix; diag$\{\cdot\}$ denotes a diagonal or block diagonal matrix, which should be clear in context; $\otimes$ denotes the Kronecker product; and finally, $\odot$ denotes the elementwise Hadamard product.

II. Problem Formulation

Consider an asynchronous K-user DS-CDMA system in the uplink (mobile-to-base) with *long (aperiodic)* spreading codes. Let $p(t)$ denote the *bandlimited* chip waveform assumed identical for all users. The transmitted signal for user k is

$$x'_k(t)=\sum_{m=0}^{M-1} d_k(m)s'_{k,m}(t-mT_s),$$

where M is the number of symbols used for code acquisition, T_s the symbol period, $d_k(m)$ the mth symbol of user k, and $s'_{k,m}(t)$ the *time-varying* spreading waveform that modulates the mth symbol of user k:

$$s'_{k,m}(t)=\sum_{n=0}^{N-1} c_{k,m}(n)p(t-nT_c).$$

Here and henceforth, N is the number of chips per symbol (i.e., *processing gain*), $T_c = T_s/N$ the chip interval, and $\{c_{k,m}(n)\}_{n=0}^{N-1}$ the spreading code of the mth symbol for user k.

The spread spectrum signal $x'_k(t)$ passes through a frequency-selective channel with L_k distinct propagation paths. The baseband signal received at the base station, after chip-matched filtering, is (e.g., [25])

$$y(t)=\sum_{k=1}^{K}\sum_{l=1}^{L_k}\alpha_{k,l}x_k(t-\tau_{k,l})+w(t) \tag{1}$$

where $\alpha_{k,l}$ and $\tau_{k,l}$ are, respectively, the channel attenuation and *code-timing* for the lth path of user k, $w(t)$ the channel noise, and $x_k(t)$ the output of the chip-matched filter with impulse response $p(T_c-t)$ given input $x'_k(t)$:

$$x_k(t)=\sum_{m=0}^{M-1} d_k(m)s_{k,m}(t-mT_s) \tag{2}$$

and where $s_{k,m}(t)\triangleq s'_{k,m}(t)\star p(T_c-t)$. It is noted that $x_k(t)$ lumps together the time-varying spreading waveform as well as data symbols for user k.

The problem of interest is to estimate the multiuser multipath code-timings $\tau_{k,l}$, $\forall k$, $\forall l$, from the received signal $y(t)$. As a side product, we will also provide estimates for the channel attenuations $\alpha_{k,l}$. Like the centralized scheme of [24], we assume that all users are transmitting training symbols. We also assume that the base station knows the number of propagation paths L_k, determined by prior channel measurements and characterization [14] along with some statistical model order detection scheme (e.g., [26] and references therein).

We provide herein a brief discussion why bandlimited chip waveforms complicate the above estimation problem. Specifically, if the chip waveform is rectangular, then the chip-matched filter (CMF) at the receiver front reduces to an integrate-and-dump filter (IDF) [8]. In that case, the IDF outputs are *linearly dependent* on the *fractional delays* along with a finite number of uncertainties on the *integer delays* [8] (see equations (2)–(4) therein).[1] Many earlier proposed acquisition schemes for short-code systems, e.g., [5], [8], and [12], employ a cost function that is quadratic in the IDF outputs and, in turn, quadratic in the fractional delays. Hence, those schemes can be conveniently implemented via a finite number of *quadratic* optimizations.[2] With bandlimited chip waveforms, the above property is lost. The cost functions of the above schemes become *highly nonlinear*. The difficulty is further complicated by long spreading codes because of the loss of cyclostationarity.

[1]The delay parameter $\tau_{k,l}$ can be written as $\tau_{k,l}=(p_{k,l}+\delta_{k,l})T$, where T is the sampling interval, $0\le p_{k,l}<N-1$ is an integer called the *integer delay*, and $0\le\delta_{k,l}<1$ is the *fractional delay*.

[2]Some schemes incur a cost function that is a ratio of quadratic terms, which makes them slightly more involved.

Despite the challenges, two standard approaches can be used to solve the above nonlinear parameter estimation problem. One is the matched filter (MF) approach [1, ch. 5], by which one correlates the received signal $y(t)$ with the desired signature waveform $x_k(t)$ (supposing user k is of interest) and identifies the peaks of the correlator output. The MF estimator ignores the MAI and is optimum only in a single-user environment (see, e.g., [5]). The other approach is based on ML estimation, which is statistically optimum but computationally highly involved. In particular, it solves a $3L$-dimensional nonlinear least squares (NLS) problem that minimizes the energy of the residual signal $w(t) \triangleq y(t) - \sum_{k=1}^{K}\sum_{l=1}^{L_k} \alpha_{k,l} x_k(t-\tau_{k,l})$ with respect to the unknowns $\alpha_{k,l}$ and $\tau_{k,l}$, where $L \triangleq \sum_{k=1}^{K} L_k$, and note that each complex-valued $\alpha_{k,l}$ is counted as two real unknowns. Exact ML estimates are difficult to find due to the exponential complexity involved in the multidimensional NLS problem. Approximate implementations based on iterative one-dimensional searches are possible; however, accurate initialization for such implementations is critical (e.g., [27] and references therein).

In view of the above discussions, the objective of this work is to develop alternative suboptimum techniques that are not as computationally demanding as the ML estimator and, meanwhile, are immune to the MAI problem that plagues the MF method. The proposed schemes may also be used to initialize the ML based schemes if further improvement in estimation accuracy is of interest.

III. Decorrelating Code-Timing Estimation

We present herein two code-timing estimators that rely on decorrelating operation for interference cancellation and parameter estimation. Unlike the MF and ML estimators which process the received signal $y(t)$ over the whole observation time, our proposed estimators decompose $y(t)$ into subsignal blocks of shorter duration. As will be shown in Section III-B, such a decomposition is necessary for decorrelating operation. The proposed estimators also rely on frequency-domain processing, which is found more convenient to deal with bandlimited chip waveforms than time-domain based processing (e.g., [16] and [17]). The two proposed estimators—one employing frequency-domain least squares (FLS) for decorrelating while the other utilizing frequency-domain weighted LS (FWLS)—offer different tradeoffs between performance and complexity.

In the following, we first discuss how to decompose $y(t)$ into subsignals and their structure in the frequency domain. We then introduce the FLS and FWLS based decorrelating schemes. Following decorrelation, a multitude of methods can be used to derive the code-timings from the decorrelated signals. We discuss the one based on the well-known MUSIC algorithm [28].

A. Subsignal Formation and Frequency-Domain Structure

Several approaches can be utilized to decompose $y(t)$ into subsignal blocks. One is to split $y(t)$ into nonoverlapping blocks of duration M_0T_s seconds, where $M_0 \leq M$. An alternative way is to form overlapping blocks. We adopt the latter, which has at least two advantages, namely, improved identifiability and improved statistical stability, respectively, over the former. These will be discussed in more details in Sections III-B and IV-B, respectively.

To elaborate, we split $y(t)$ into overlapping subsignal blocks of duration M_0T_s seconds as follows: $y_\mu(t) \triangleq y(t)$, $\mu T_s \leq t \leq (\mu + M_0)T_s$ and zero elsewhere, $\mu = 0, 1, \ldots$. The decomposition process is also depicted graphically in Fig. 1(a). Note that any two adjacent blocks $y_{\mu-1}(t)$ and $y_\mu(t)$ are offset by T_s sec, which is appropriate when $\max_{k,l} \tau_{k,l} \leq T_s$. For larger delays, we can increase the offset between adjacent blocks. For example, if it is known *a priori* that $\max_{k,l} \tau_{k,l} \leq 2T_s$, we can define $y_\mu(t) = y(t)$, $2\mu T_s \leq t \leq (2\mu + M_0)T_s$, and zero elsewhere, and $\mu = 0, 1, \ldots$ so that two adjacent blocks are offset by $2T_s$. Note that the maximum delay depends on the cell size and transmission rate and can be determined *a priori*.

For simplicity, we assume in the sequel that $\max_{k,l} \tau_{k,l} \leq T_s$. This implies that the transmission of M symbols leads to an observation $y(t)$ for at most $(M+1)T_s$ sec. In turn, we have a total of $J \triangleq M - M_0 + 2$ subsignal blocks $y_\mu(t)$ formed as above. Let $x_k(t)$ be decomposed into overlapping blocks of duration $(M_0 - 1)T_s$ sec: $x_{k,\mu}(t) \triangleq x_k(t)$, $\mu T_s \leq t \leq (\mu + M_0 - 1)T_s$ and zero elsewhere, which is also illustrated graphically in Fig. 1(b). Note that $x_{k,\mu}(t)$ is formed by the μth, $(\mu+1)$th, ..., and $(\mu + M_0 - 2)$th symbols of the stream. Due to the constraint on the maximum delay, $y_\mu(t)$ is contributed by $x_{k,\mu}(t)$ from all users, plus the *residual* intersymbol interference (ISI) caused by the $(\mu - 1)$th symbol preceding $x_{k,\mu}(t)$ and the $(\mu + M_0 - 1)$th symbol that follows. The residual ISI is not modeled due to the unknown delay. The above decomposition is illustrated in Fig. 1(c) for the lth path of user k. Hence, we can write [cf. (1)]

$$y_\mu(t) = \sum_{k=1}^{K}\sum_{l=1}^{L_k} \alpha_{k,l} x_{k,\mu}(t-\tau_{k,l}) + e_\mu(t), \quad \mu = 0, \cdots, J-1 \tag{3}$$

where $e_\mu(t)$ lumps together the channel noise $w(t)$ and residual ISI from all users and all propagation paths.

For digital processing, $y_\mu(t)$ is sampled with a sampling interval $T \triangleq T_c/Q$: $y_\mu(i) = y_\mu(t)|_{t=\mu T_s + iT}$, $i = 0, 1, \ldots, M_0NQ - 1$, where Q is an integer called oversampling factor. Typically, we choose $Q = 1$ or $Q = 2$. Let $\bar{\mathbf{y}}_\mu \triangleq [y_\mu(0), \cdots, y_\mu(M_0NQ-1)]^T$. It follows from (3) that[3]

$$\bar{\mathbf{y}}_\mu = \sum_{k=1}^{K}\sum_{l=1}^{L_k} \alpha_{k,l} \bar{\mathbf{x}}_{k,\mu}(\tau_{k,l}) + \bar{\mathbf{e}}_\mu, \quad \mu = 0, \cdots, J-1 \tag{4}$$

where $\bar{\mathbf{x}}_{k,\mu}(\tau_{k,l})$ and $\bar{\mathbf{e}}_\mu$ are $M_0NQ \times 1$ vectors formed from $x_{k,\mu}(t - \tau_{k,l})$ and $e_\mu(t)$, respectively. Note that $\bar{\mathbf{x}}_{k,\mu}(\tau_{k,l})$ may have a few zeros at the beginning and/or in the end determined by $\tau_{k,l}$.

Next, we take the discrete Fourier transform (DFT) of both sides of (4). Note that $x_{k,\mu}(t)$ consists of $M_0 - 1$ data symbols modulating a *bandlimited* spreading waveform [cf. (2)], whose spectrum typically tapers off at the end frequencies (i.e., frequencies close to $\pm 0.5 f_s$, where $f_s \triangleq 1/T$ denotes the sampling frequency) [29]. In the presence of channel

[3]We henceforth use notation $\bar{(\cdot)}$ to denote a time-domain quantity if its frequency-domain counterpart is also used for estimation.

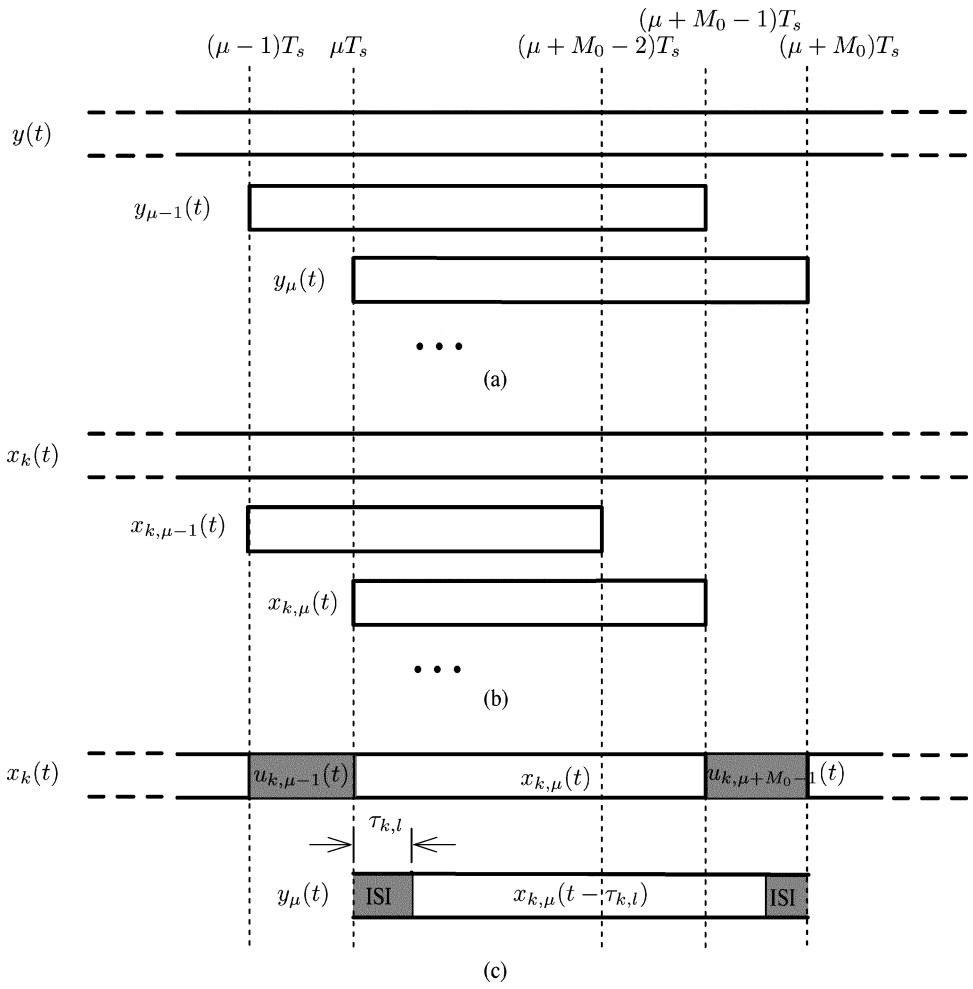

Fig. 1. (a) Splitting received signal $y(t)$ into overlapping subsignal blocks $y_\mu(t)$ of duration M_0T_s : $y_\mu(t)$ starts at $t=\mu T_s$ and ends at $t=(\mu+M_0)T_s$, and two adjacent blocks $y_{\mu-1}(t)$ and $y_\mu(t)$ overlap by $(M_0-1)T_s$. (b) Splitting transmitted signal $x_k(t)$ for user k into overlapping subsignal blocks $x_{k,\mu}(t)$ of duration $(M_0-1)T_s$: $x_{k,\mu}(t)$ starts at $t=\mu T_s$, similar to $y_\mu(t)$, but ends at $t=(\mu+M_0-1)T_s$; two adjacent blocks $x_{k,\mu-1}(t)$ and $x_{k,\mu}(t)$ overlap by $(M_0-2)T_s$. (c) Graphical illustration of contribution to the μth received subsignal block $y_\mu(t)$ from the lth path of user k : $y_\mu(t)=x_{k,\mu}(t-\tau_{k,l})+\mathrm{ISI}$, where contributions from other users and other paths of user k are not shown, and the residual inter-symbol interference (ISI) is due to $u_{k,\mu-1}(t)\triangleq d_k(\mu-1)s_{k,\mu-1}(t-(\mu-1)T_s)$ and $u_{k,\mu+M_0-1}(t)\triangleq d_k(\mu+M_0-1)s_{k,\mu+M_0-1}(t-(\mu+M_0-1)T_s)$.

noise, the end frequencies have a lower signal-to-noise ratio (SNR) than at the middle (i.e., low) frequencies. Hence, we will discard the end frequencies to avoid the noise amplification caused by frequency decorrelation to be applied later. To this end, let $\boldsymbol{\mathcal{F}}$ be the $M_0NQ\times M_0NQ$ matrix with the mnth element given by (assuming M_0NQ is even): $\exp[-j(2\pi/M_0NQ)(m-(M_0NQ/2)-1)(n-1)]$. Note that $\boldsymbol{\mathcal{F}}$ computes the M_0NQ-point *shifted* DFT so that the DC frequency is shifted to the middle. Let $\eta\in(0,1]$ denote the *DFT frequency selection parameter*, and $N_s\triangleq 2\lceil\eta NQ/2\rceil$. Let $\boldsymbol{\mathcal{F}}_s\in\mathbb{C}^{M_0N_s\times M_0NQ}$ be formed from the middle M_0N_s rows of $\boldsymbol{\mathcal{F}}$ (equivalently, by deleting the first and last $(M_0NQ-M_0N_s)/2$ rows). Then, discarding the end-frequency samples of the DFT of $\bar{\mathbf{y}}_\mu$ is equivalent to computing $\mathbf{y}_\mu\triangleq\boldsymbol{\mathcal{F}}_s\bar{\mathbf{y}}_\mu$. By the time-shifting property of Fourier transform, we have

$$\boldsymbol{\mathcal{F}}_s\bar{\mathbf{x}}_{k,\mu}(\tau_{k,l})\approx\mathrm{diag}(\mathbf{x}_{k,\mu})\boldsymbol{\phi}_0(\tau_{k,l})\triangleq\mathbf{X}_{k,\mu}\boldsymbol{\phi}_0(\tau_{k,l}) \tag{5}$$

where $\mathbf{x}_{k,\mu}\triangleq\boldsymbol{\mathcal{F}}_s\bar{\mathbf{x}}_{k,\mu}$, and $\bar{\mathbf{x}}_{k,\mu}$ denotes the $M_0NQ\times 1$ vector formed by samples of $x_{k,\mu}(t)$ with NQ zeros padded at the tail [note that zero-padding is needed since $x_{k,\mu}(t)$ has duration of $(M_0-1)T_s$], and

$$\boldsymbol{\phi}_0(\tau_{k,l})\triangleq\left[e^{\frac{-M_0N_s}{2}\frac{-j2\pi\tau_{k,l}}{M_0NQ}},\cdots,e^{(\frac{M_0N_s}{2}-1)\frac{-j2\pi\tau_{k,l}}{M_0NQ}}\right]^T. \tag{6}$$

Equation (5) holds only approximately because of the aliasing caused by the blocking/windowing operation that produces $x_{k,\mu}(t)$. Such windowing widens slightly the spectrum, and sampling at a rate $1/T$ introduces some small aliasing due to spectral folding, which will eventually lead to a small bias in the parameter estimate. The aliasing, however, can be neglected compared to the noise/interference induced estimation error [29].

Let $\boldsymbol{\tau}_k\triangleq[\tau_{k,1},\cdots,\tau_{k,L_k}]^T$, $\boldsymbol{\alpha}_k\triangleq[\alpha_{k,1},\cdots,\alpha_{k,L_k}]^T$, $\boldsymbol{\Phi}_0(\boldsymbol{\tau}_k)\triangleq[\boldsymbol{\phi}_0(\tau_{k,1}),\cdots,\boldsymbol{\phi}_0(\tau_{k,L_k})]$, and $\boldsymbol{\beta}_k\triangleq\boldsymbol{\Phi}_0(\boldsymbol{\tau}_k)\boldsymbol{\alpha}_k$. Following the above discussions, we have

$$\mathbf{y}_\mu\triangleq\sum_{k=1}^{K}\mathbf{X}_{k,\mu}\boldsymbol{\beta}_k+\mathbf{e}_\mu,\quad\mu=0,\cdots,J-1 \tag{7}$$

where $\mathbf{e}_\mu \in \mathbb{C}^{M_0 N_s \times 1}$ includes both the spectrum of $\bar{\mathbf{e}}_\mu$ as well as any aliasing-induced error. The above equations are fundamental to the proposed estimators, which follow a two-step procedure. Specifically, we first use a decorrelating procedure to eliminate MAI and obtain an unstructured estimate $\hat{\boldsymbol{\beta}}_k$ of $\boldsymbol{\beta}_k$; then, the structure $\boldsymbol{\beta}_k \triangleq \boldsymbol{\Phi}_0(\boldsymbol{\tau}_k)\boldsymbol{\alpha}_k$ is imposed on $\hat{\boldsymbol{\beta}}_k$ to yield estimates of $\boldsymbol{\tau}_k$ and $\boldsymbol{\alpha}_k$.

B. Decorrelating

Let $\mathbf{y} \triangleq [\mathbf{y}_0^T, \cdots, \mathbf{y}_{J-1}^T]^T_{JM_0N_s \times 1}$, $\mathbf{X}_\mu \triangleq [\mathbf{X}_{1,\mu}, \cdots, \mathbf{X}_{K,\mu}]$, and $\boldsymbol{\beta} \triangleq [\boldsymbol{\beta}_1^T, \cdots, \boldsymbol{\beta}_K^T]^T$. Then, (7) can be compactly written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} \tag{8}$$

where $\mathbf{e} \triangleq [\mathbf{e}_0^T, \cdots, \mathbf{e}_{J-1}^T]^T$, and

$$\mathbf{X} \triangleq \begin{bmatrix} \mathbf{X}_0 \\ \vdots \\ \mathbf{X}_{J-1} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_{1,0} & \cdots & \mathbf{X}_{K,0} \\ \vdots & \vdots & \vdots \\ \mathbf{X}_{1,J-1} & \cdots & \mathbf{X}_{K,J-1} \end{bmatrix}_{JM_0N_s \times KM_0N_s}.$$

Applying least-squares (LS) on the frequency-domain (8), the decorrelating FLS estimator is given by

$$\hat{\boldsymbol{\beta}}_{\mathrm{FLS}} = (\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H\mathbf{y}. \tag{9}$$

Substituting (8) into (9), we have $\hat{\boldsymbol{\beta}}_{\mathrm{FLS}} = \boldsymbol{\beta} + (\mathbf{X}^H\mathbf{X})^{-1}\mathbf{X}^H\mathbf{e}$. Hence, the MAI is fully eliminated. It should be noted, however, that the correlation of the residual interference, due to ISI and aliasing, in $\mathbf{e}$ is ignored. The FLS estimate can be improved by appropriate weighting that takes into account the correlation of the residual interference. Doing so leads to the FWLS estimator:

$$\hat{\boldsymbol{\beta}}_{\mathrm{FWLS}} = (\mathbf{X}^H\mathbf{W}^{-1}\mathbf{X})^{-1}(\mathbf{X}^H\mathbf{W}^{-1}\mathbf{y}) \tag{10}$$

where $\mathbf{W}$ denotes a weighting matrix which, ideally, should be chosen as the covariance matrix of $\mathbf{e}$ in order to minimize the variance of the estimate [30]. Before we discuss how to obtain $\mathbf{W}$, we highlight two important issues that need be properly addressed for such decorrelating estimators:

1) *Rank*: To ensure identifiability, both (9) and (10) require that $\mathbf{X}$ is tall and has full column rank.
2) *Complexity*: Although the idea of decorrelating is straightforward, direct application of (9) or (10) is in general not feasible due to the high complexity and large memory requirement. Consider, for example, a modest scenario with $N = 16$, $K = 10$, $M = 100$, $M_0 = 10$ (hence, $J = M - M_0 + 2 = 92$), $Q = 2$, and $N_s = NQ = 32$ (i.e., no frequencies discarded). In this case, $\mathbf{X}$ is of size $29\,440 \times 3200$. Clearly, it is impractical to invert a matrix of this size.

The complexity issue will be addressed in Section IV, where efficient implementations of (9) and (10) are presented by exploiting the structure of $\mathbf{X}$ and $\mathbf{W}$. We discuss in the following the rank of $\mathbf{X}$ and the relevant issue of how to choose the block length M_0 or, equivalently, J.

Note that $\mathbf{X}$ is formed by $J \times K$ blocks of diagonal matrices. To ensure $\mathbf{X}$ is tall, we have to choose M_0 such that $J \geq K$. This corroborates our earlier statement that it is necessary to split $y(t)$ into shorter blocks for decorrelating. The condition $J \geq K$ is necessary but not sufficient for $\mathbf{X}$ to be full rank. However, increasing J, which makes $\mathbf{X}$ taller, increases the likelihood of having a full-rank matrix. This explains why we chose overlapping, rather than nonoverlapping, in splitting $y(t)$ into blocks, which results in larger J for fixed M and M_0. Increasing J has another advantage of providing sufficient averaging for better statistical stability, which will be detailed in Section IV, when we discuss the implementation issue. In our simulation, we never experienced a rank-deficient $\mathbf{X}$, most likely due to the fact that the spreading codes are independent for different users, whereas for the same user, the spreading code changes independently from symbol to symbol (aperiodic long code). Discarding frequency samples at the end frequencies, the places where spectral nulls are usually located also helps to make $\mathbf{X}$ full rank. Finally, it is important to note that since $\mathbf{X}$ is known at the base station, we can always check its rank condition. Should it be rank deficient or ill conditioned, we can discard certain rows of $\mathbf{X}$ (accordingly, the corresponding elements of $\mathbf{y}$) so that the resulting matrix is full rank and, thus, ensure identifiability.

The above discussion seems to suggest that M_0 should be chosen small so that J is as large as possible. There is a tradeoff, however, due to the residual ISI contained in $\mathbf{e}$. In particular, if M_0 is too small, the residual interference may be significant compared to the desired signal in each block, and ignoring the interference can degrade the performance of the FLS estimator quite a bit. While the FWLS estimator can handle some additional interference, the residual ISI, in general, cannot be fully eliminated since the optimum weighting matrix $\mathbf{W}$ is unknown (see discussions later). Hence, for both FLS and FWLS, it is important to keep the residual interference small by choosing M_0 large enough. We found that a choice of $M_0 \geq 6$ is sufficient, provided that the power variations among different users is less than 20 dB (also see numerical examples in Section VI). Note that power variation in real systems is in general much smaller due to the use of power control [14].

To summarize, we recommend choosing $M_0 \geq 6$ while keeping $J \geq K$. In the event that $\mathbf{X}$ is still rank deficient or ill conditioned, we remove some rows of $\mathbf{X}$ (especially those containing small entries, which correspond to spectral nulls) to ensure identifiability.

Finally, we discuss how to obtain the weighting matrix $\mathbf{W}$ for the FWLS estimator (10). The optimum weighting matrix for linear estimator is given by the covariance matrix $\mathbf{R}_e \triangleq E\{\mathbf{e}\mathbf{e}^H\} \in \mathbb{C}^{JM_0N_s \times JM_0N_s}$ [30], which is unknown and has to be estimated. Finding an unstructured estimate of the $JM_0N_s \times JM_0N_s$ matrix is an ill-defined problem, since we have more unknowns than data samples. Hence, we have to impose certain structure on $\mathbf{R}_e$ to make its estimation feasible. In the following, we assume $\mathbf{R}_e = \mathbf{I}_{JM_0} \otimes \boldsymbol{\Gamma}$, where $\boldsymbol{\Gamma} \in \mathbb{C}^{N_s \times N_s}$ captures the correlation within N_s adjacent samples in the frequency domain. While this assumption is *ad hoc*, it leads to quite good

estimation results, as we will see in Section VI. With this assumption, we can obtain an estimate of $\bar{\mathbf{\Gamma}}$, using either (7) or (8), as follows:

$$\hat{\mathbf{\Gamma}} = \frac{1}{JM_0}\sum_{\mu=0}^{J-1}\sum_{m=1}^{M_0}\left(\mathbf{y}_\mu(m) - \mathbf{X}_\mu(m)\hat{\boldsymbol{\beta}}_{\mathrm{FLS}}\right) \times \left(\mathbf{y}_\mu(m) - \mathbf{X}_\mu(m)\hat{\boldsymbol{\beta}}_{\mathrm{FLS}}\right)^H \quad (11)$$

where $\hat{\boldsymbol{\beta}}_{\mathrm{FLS}}$ denotes the initial FLS estimate given by (9), $\mathbf{y}_\mu(m) \in \mathbb{C}^{N_s\times 1}$ is the mth subvector formed from $\mathbf{y}_\mu$, i.e., $\mathbf{y}_\mu = [\mathbf{y}_\mu^T(1),\cdots,\mathbf{y}_\mu^T(M_0)]^T$, and $\mathbf{X}_\mu = [\mathbf{X}_\mu^T(1),\ldots,\mathbf{X}_\mu^T(M_0)]^T$. Using this estimate in (10), we have the following FWLS estimate:

$$\hat{\boldsymbol{\beta}}_{\mathrm{FWLS}} = \left[\mathbf{X}^H\left(\mathbf{I}_{JM_0}\otimes\hat{\mathbf{\Gamma}}^{-1}\right)\mathbf{X}\right]^{-1}\left[\mathbf{X}^H\left(\mathbf{I}_{JM_0}\otimes\hat{\mathbf{\Gamma}}^{-1}\right)\mathbf{y}\right]. \quad (12)$$

C. Code-Timing and Attenuation Estimation

Let $\hat{\boldsymbol{\beta}}$ be either $\hat{\boldsymbol{\beta}}_{\mathrm{FLS}}$ or $\hat{\boldsymbol{\beta}}_{\mathrm{FWLS}}$, and let $\hat{\boldsymbol{\beta}}_k$ be the kth segment corresponding to user k. We then have

$$\hat{\boldsymbol{\beta}}_k = \mathbf{\Phi}_0(\boldsymbol{\tau}_k)\boldsymbol{\alpha}_k + \mathbf{n}_k \quad (13)$$

where $\mathbf{n}_k$ denotes the estimation error. Note that $\mathbf{\Phi}_0(\boldsymbol{\tau}_k)$ is an $M_0N_s \times L_k$ Vandermonde matrix with the lth column given by (6). We can see that, effectively, $\hat{\boldsymbol{\beta}}_k$ consists of a group of L_k complex sinusoids with frequency $\omega_{k,l} \triangleq -(2\pi\tau_{k,l}/M_0NQ)$ and amplitude $\alpha_{k,l}$. Hence, the problem reduces to a sinusoidal parameter estimation problem, and a multitude of methods can be used (see, e.g., [28]) to estimate the sinusoidal parameters. In what follows, we briefly discuss the one based on the well-known MUSIC [31] algorithm. Specifically, let $N_v < M_0N_s$ and form $N_v \times 1$ overlapping subvectors $\hat{\boldsymbol{\beta}}_{k,n} \triangleq [\hat{\beta}_k(n),\cdots,\hat{\beta}_k(n+N_v-1)]^T$, $n = 1,\ldots,M_0N_s - N_v + 1$. In general, N_v should be large enough to provide good estimation accuracy but not close to M_0N_s [28]. Next, compute

$$\hat{\mathbf{R}}_{\beta_k} = \frac{1}{M_0N_s - N_v + 1}\sum_{n=1}^{M_0N_s-N_v+1}\hat{\boldsymbol{\beta}}_{k,n}\hat{\boldsymbol{\beta}}_{k,n}^H. \quad (14)$$

Let its eigendecomposition be $\hat{\mathbf{R}}_{\beta_k} = \hat{\mathbf{E}}_s\hat{\mathbf{\Lambda}}_s\hat{\mathbf{E}}_s^H + \hat{\mathbf{E}}_n\hat{\mathbf{\Lambda}}_n\hat{\mathbf{E}}_n^H$, where $\hat{\mathbf{\Lambda}}_s$ is a diagonal matrix formed by the L_k largest eigenvalues, $\hat{\mathbf{E}}_s \in \mathbb{C}^{N_v\times L_k}$ spans the signal subspace, and $\hat{\mathbf{E}}_n \in \mathbb{C}^{N_v\times(N_v-L_k)}$ spans the noise subspace that is orthogonal to $\hat{\mathbf{E}}_s$. Let $z \triangleq e^{-j2\omega_{k,l}}$, and $\boldsymbol{\varphi}(z) \triangleq [1,z,\cdots,z^{N_v-1}]^T$. The L_k frequency estimates $\hat{\omega}_{k,l}$ can be obtained as the phases of the L_k roots of the following polynomial that are closest to the unit circle [28, Sec. 4.6]:

$$V(z) = \boldsymbol{\varphi}^T(z^{-1})\hat{\mathbf{E}}_s\hat{\mathbf{E}}_s^H\boldsymbol{\varphi}(z). \quad (15)$$

Then, the code-timing estimates are given by $\hat{\tau}_{k,l} = -(M_0NQ\hat{\omega}_{k,l}/2\pi)$. Following code-timing estimation, the attenuations can be estimated using (13):

$$\hat{\boldsymbol{\alpha}}_k = \left[\mathbf{\Phi}_0^H(\hat{\boldsymbol{\tau}}_k)\mathbf{\Phi}_0(\hat{\boldsymbol{\tau}}_k)\right]^{-1}\mathbf{\Phi}_0^H(\hat{\boldsymbol{\tau}}_k)\hat{\boldsymbol{\beta}}_k. \quad (16)$$

IV. Efficient Implementations

As mentioned before, direct applications of (9) and (12) are in general impractical. In this section, we discuss efficient implementations of the proposed FLS and FWLS estimators by exploiting the sparse structure of the large matrices in (9) and (12), which can reduce the complexity significantly. We first present an efficient implementation for the FLS estimator (9). We then extend it to the FWLS estimator (12). Finally, we summarize the complexity of the proposed efficient implementations.

A. Implementation for FLS

The complexity of (9) is dominated by the calculation of $\mathbf{G} \triangleq \mathbf{X}^H\mathbf{X}$ and its inverse $\mathbf{G}^{-1}$. Let

$$\mathbf{G} = \begin{bmatrix}\mathbf{G}_{1,1} & \cdots & \mathbf{G}_{1,K}\\ \vdots & \ddots & \vdots\\ \mathbf{G}_{K,1} & \cdots & \mathbf{G}_{K,K}\end{bmatrix}_{KM_0N_s\times KM_0N_s} \quad (17)$$

where the klth submatrix $\mathbf{G}_{k,l} \triangleq \mathrm{diag}\{\mathbf{g}_{k,l}\}$ is diagonal, with

$$\mathbf{g}_{k,l} = \sum_{\mu=0}^{J-1}\mathbf{x}_{k,\mu}^*\odot\mathbf{x}_{l,\mu}, \quad k,l = 1,\ldots,K. \quad (18)$$

For notational brevity, define $\boldsymbol{\mathcal{G}} \triangleq \mathbf{G}^{-1}$ and $B \triangleq M_0N_s$. We show in Appendix A that due to the block structure of $\mathbf{G}$, $\boldsymbol{\mathcal{G}}$ can be efficiently computed as follows:

```
for  i = 1:B
  G(i:B:KB, i:B:KB) = [G(i:B:KB, i:B:KB)]^{-1}
end
```
(19)

where the Matlab colon operator is used to simplify the notation. Specifically, $\mathbf{G}(i:B:KB, i:B:KB)$ denotes a $K\times K$ submatrix formed by rows $i, i+B,\ldots,i+(K-1)B$, and columns $i, i+B,\ldots,i+(K-1)B$ of $\mathbf{G}$. As shown in Appendix A, $\boldsymbol{\mathcal{G}}$ has the same structure of $\mathbf{G}$ in (17). For later use, let $\boldsymbol{\mathcal{G}}_{kl} \triangleq \mathrm{diag}(\boldsymbol{g}_{k,l}) \in \mathbb{C}^{M_0N_s\times M_0N_s}$ denote the klth subblock of $\boldsymbol{\mathcal{G}}$.

We next summarize the implementation of the FLS estimator (9) along with its computational complexity.

- **Step 1**: Compute $\mathbf{G}$ using (18).
  $\Rightarrow$ Total $O(M_0N_sJK^2)$ flops.
- **Step 2**: Compute $\boldsymbol{\mathcal{G}}$ (i.e., $\mathbf{G}^{-1}$) using (19).
  $\Rightarrow$ Total $O(M_0N_sK^3)$ flops.
- **Step 3**: Compute $\boldsymbol{\nu} \triangleq \mathbf{X}^H\mathbf{y} = [\boldsymbol{\nu}_1^T,\ldots,\boldsymbol{\nu}_K^T]^T$, where $\boldsymbol{\nu}_k \in \mathbb{C}^{M_0N_s\times 1}$ is determined by

$$\boldsymbol{\nu}_k = \sum_{\mu=0}^{J-1}\mathbf{x}_{k,\mu}^*\odot\mathbf{y}_\mu, \quad k = 1,\ldots,K$$
$$\Rightarrow \text{Total } O(JKM_0N_s) \text{ flops.} \quad (20)$$

- **Step 4**: Compute $\hat{\boldsymbol{\beta}}_{\mathrm{FLS}} = \boldsymbol{\mathcal{G}}\boldsymbol{\nu} = [\hat{\boldsymbol{\beta}}_1^T,\ldots,\hat{\boldsymbol{\beta}}_K^T]^T$ by

$$\hat{\boldsymbol{\beta}}_k = \sum_{l=1}^{K}\boldsymbol{g}_{k,l}\odot\boldsymbol{\nu}_l, \quad k = 1,\ldots,K.$$
$$\Rightarrow \text{Total } O(K^2M_0N_s) \text{ flops.} \quad (21)$$

B. Implementation of FWLS

For the FWLS estimator (12), let $\mathbf{G} \triangleq \mathbf{X}^H(\mathbf{I}_{JM_0} \otimes \hat{\mathbf{\Gamma}}^{-1})\mathbf{X}$. Let $\mathbf{G}_{k,l} \in \mathbb{C}^{M_0N_s \times M_0N_s}$ denote the klth subblock of $\mathbf{G}$. Note that $\mathbf{G}_{k,l}$ is block diagonal:

$$\begin{aligned}\mathbf{G}_{k,l} &= \sum_{\mu=0}^{J-1} \mathbf{X}_{k,\mu}^H \left(\mathbf{I}_{M_0} \otimes \hat{\mathbf{\Gamma}}^{-1}\right) \mathbf{X}_{l,\mu} \\ &\triangleq \text{diag}\left\{\mathbf{G}_{k,l}(1), \ldots, \mathbf{G}_{k,l}(M_0)\right\}. \end{aligned} \tag{22}$$

The $N_s \times N_s$ diagonal blocks $\mathbf{G}_{k,l}(m)$ can be computed as follows. Let $\mathbf{x}_{k,\mu}$ (recall $\mathbf{X}_{k,\mu} = \text{diag}(\mathbf{x}_{k,\mu})$) be sliced into M_0 equal-length subvectors: $\mathbf{x}_{k,\mu} = [\mathbf{x}_{k,\mu}^T(1), \cdots, \mathbf{x}_{k,\mu}^T(M_0)]^T$. Then, it is ready to verify

$$\mathbf{G}_{k,l}(m) = \sum_{\mu=0}^{J-1} \left[\mathbf{x}_{k,\mu}^*(m)\mathbf{x}_{l,\mu}^T(m)\right] \odot \hat{\mathbf{\Gamma}}^{-1}. \tag{23}$$

Although $\mathbf{G}_{k,l}$ is no longer diagonal, its block diagonal structure can be exploited for efficient calculation of $\boldsymbol{\mathcal{G}} \triangleq \mathbf{G}^{-1}$. In particular, we show in Appendix B that $\boldsymbol{\mathcal{G}}$ has the same structure as $\mathbf{G}$. Let $\boldsymbol{\mathcal{G}}_{k,l} = \text{diag}\{\boldsymbol{\mathcal{G}}_{k,l}(1), \cdots, \boldsymbol{\mathcal{G}}_{k,l}(M_0)\} \in \mathbb{C}^{M_0N_s \times M_0N_s}$ be the klth subblock of $\boldsymbol{\mathcal{G}}$, where $\boldsymbol{\mathcal{G}}_{k,l}(m) \in \mathbb{C}^{N_s \times N_s}$, $m = 1, \cdots, M_0$. Then, $\boldsymbol{\mathcal{G}}$ can be efficiently computed as

$$\begin{aligned}&\texttt{for} \quad m = 1 : M_0 \\ &\qquad \begin{bmatrix} \boldsymbol{\mathcal{G}}_{1,1}(m) & \cdots & \boldsymbol{\mathcal{G}}_{1,K}(m) \\ \vdots & \ddots & \vdots \\ \boldsymbol{\mathcal{G}}_{K,1}(m) & \cdots & \boldsymbol{\mathcal{G}}_{K,K}(m) \end{bmatrix} \\ &\qquad = \begin{bmatrix} \mathbf{G}_{1,1}(m) & \cdots & \mathbf{G}_{1,K}(m) \\ \vdots & \ddots & \vdots \\ \mathbf{G}_{K,1}(m) & \cdots & \mathbf{G}_{K,K}(m) \end{bmatrix}^{-1}_{KN_s \times KN_s}. \\ &\texttt{end} \end{aligned} \tag{24}$$

We next summarize the implementation of the FWLS estimator (12) and its complexity as follows:

- **Step 1**: Compute $\mathbf{G}$ using (22) and (23).
  $\Rightarrow$ Total $O(JK^2M_0N_s^2)$ flops.
- **Step 2**: Compute $\boldsymbol{\mathcal{G}}$ (i.e., $\mathbf{G}^{-1}$) using (24).
  $\Rightarrow$ Total $O(M_0K^3N_s^3)$ flops.
- **Step 3**: Compute $\boldsymbol{\nu} \triangleq \mathbf{X}^H(\mathbf{I}_{JM_0} \otimes \hat{\mathbf{\Gamma}}^{-1})\mathbf{y} = [\boldsymbol{\nu}_1^T, \ldots, \boldsymbol{\nu}_K^T]^T \in \mathbb{C}^{KM_0N_s \times 1}$. Let $\boldsymbol{\nu}_k(m) \in \mathbb{C}^{N_s \times 1}$ be the mth segment of $\boldsymbol{\nu}_k$, which can be computed as

$$\begin{aligned}\boldsymbol{\nu}_k(m) &= \sum_{\mu=0}^{J-1} \text{diag}\left\{\mathbf{x}_{k,\mu}^*(m)\right\} \hat{\mathbf{\Gamma}}^{-1} \mathbf{y}_\mu(m) \\ &\qquad k = 1, \ldots, K; \quad m = 1, \ldots, M_0 \\ &\Rightarrow \text{Total } O\left(JKM_0N_s^2\right) \text{ flops} \end{aligned} \tag{25}$$

where $\mathbf{y}_\mu(m)$ and $\mathbf{x}_{k,\mu}(m)$ denote the mth $N_s \times 1$ segment of $\mathbf{y}_\mu$ and $\mathbf{x}_{k,\mu}$, respectively.

- **Step 4**: Compute $\hat{\boldsymbol{\beta}}_{\text{FWLS}} = \boldsymbol{\mathcal{G}}\boldsymbol{\nu} = [\hat{\boldsymbol{\beta}}_1^T, \ldots, \hat{\boldsymbol{\beta}}_K^T]^T$. Let $\hat{\boldsymbol{\beta}}_k(m) \in \mathbb{C}^{N_s \times 1}$ be the mth segment of $\hat{\boldsymbol{\beta}}_k$, which can be computed as

$$\begin{aligned}\hat{\boldsymbol{\beta}}_k(m) &= \sum_{l=1}^{K} \boldsymbol{\mathcal{G}}_{k,l}(m)\boldsymbol{\nu}_l(m) \\ &\qquad k = 1, \ldots, K; \quad m = 1, \ldots, M_0 \\ &\Rightarrow \text{Total } O\left(K^2M_0N_s^2\right) \text{ flops.} \end{aligned} \tag{26}$$

C. Complexity

We now compare the complexity of the above efficient implementations with direct implementations of (9) and (12). First, let us examine the FWLS estimator. From Section IV-B, it is straightforward to show that the complexity of the efficient implementation of $\hat{\boldsymbol{\beta}}_{\text{FWLS}}$ is $O(JK^2M_0N_s^2) + O(K^3M_0N_s^3)$ flops. In general, we have $J < KN_s$. Hence, the complexity can be further approximated to be $O(K^3M_0N_s^3)$. On the other hand, direct implementation has a complexity of about $O(K^3M_0^3N_s^3)$. Hence, the efficient implementation for the FWLS estimator reduces the complexity approximately by a factor of M_0^2.

Computational saving for the FLS estimator is even more pronounced. In particular, the complexity associated with the efficient implementation in Section IV-A is $O(JK^2M_0N_s) + O(K^3M_0N_s)$. Supposing J is comparable to K, the complexity is approximately $O(K^3M_0N_s)$. Meanwhile, direct implementation of (9) still has a complexity of about $O(K^3M_0^3N_s^3)$. Thus, the efficient implementation for the FLS estimator reduces the complexity approximately by a factor of $M_0^2N_s^2$.

The above analysis reveals that the efficient FLS estimator is about N_s^2 times faster than the efficient FWLS estimator. In general, the former is slightly inferior to the latter in terms of acquisition performance (see Section VI). Hence, the efficient FLS estimator is the recommended one in most cases from the computational perspective. Their difference in estimation accuracy could be significant though, especially when the MAI is very strong (i.e., in a near–far environment with interfering powers 30 dB higher than that of the desired signal; see Section VI). The FLS estimator suffers from some performance loss in such cases due to the residual interference that is ignored by the estimator. See Section VI for numerical examples.

V. Cramér–Rao Bound

As a benchmark for the proposed estimators, we derive in this section the Cramér–Rao bound (CRB) for the estimation problem under study. The CRB provides a lower variance bound for all unbiased estimators and has been routinely used as an assessment tool. It is useful to present our derivation here since the CRB for code-timing estimation with long code and band-limited waveform appears not available in the literature.

Recall the received signal $y(t)$ as defined in (1). Denote the sampled signal by $y(i) \triangleq y(t)|_{t=iT}$, $i = 0, 1, \cdots, (M+1)NQ - 1$. We collect all samples in a vector: $\bar{\boldsymbol{y}} = [y(0), \cdots, y((M+1)NQ - 1)]^T$. Likewise, let $\bar{\boldsymbol{x}}_k(\tau_{k,l})$ and $\bar{\boldsymbol{w}}$ be $(M+1)NQ \times 1$ vectors formed by samples of $x_k(t - \tau_{k,l})$ and $w(t)$, respectively. Then, we have

$\bar{\boldsymbol{y}} = \sum_{k=1}^{K}\sum_{l=1}^{L_k}\alpha_{k,l}\bar{\boldsymbol{x}}_k(\tau_{k,l}) + \bar{\boldsymbol{w}}$, which is similar to (4) except that we now include all received samples, rather than samples within a block there. Hence, there is no residual ISI, and $\bar{\boldsymbol{w}}$ is only due to the channel noise. Multiplying the above equation with $\boldsymbol{\mathcal{F}}_M$, where $\boldsymbol{\mathcal{F}}_M$ denotes the $(M+1)NQ \times (M+1)NQ$ DFT matrix, yields

$$\boldsymbol{y} = \sum_{k=1}^{K} \boldsymbol{X}_k \boldsymbol{\Phi}(\boldsymbol{\tau}_k)\boldsymbol{\alpha}_k + \boldsymbol{w} \triangleq \boldsymbol{m} + \boldsymbol{w} \tag{27}$$

where $\boldsymbol{y} \triangleq \boldsymbol{\mathcal{F}}_M\bar{\boldsymbol{y}}$, $\boldsymbol{w} \triangleq \boldsymbol{\mathcal{F}}_M\bar{\boldsymbol{w}}$, and $\boldsymbol{X}_k \triangleq \mathrm{diag}(\boldsymbol{x}_k) = \mathrm{diag}(\boldsymbol{\mathcal{F}}_M\bar{\boldsymbol{x}}_k)$, with $\bar{\boldsymbol{x}}_k$ denoting an $(M+1)NQ \times 1$ vector formed from MNQ samples of $x_k(t)$ (with no delay) padded with NQ zeros at the tail, and

$$\boldsymbol{\Phi}(\boldsymbol{\tau}_k) \triangleq [\boldsymbol{\phi}(\tau_{k,1}), \ldots, \boldsymbol{\phi}(\tau_{k,L_k})]$$
$$\boldsymbol{\phi}(\tau_{k,l}) \triangleq \left[1, e^{\frac{-j2\pi}{(M+1)NQ}\tau_{k,l}}, \ldots, e^{\frac{-j2\pi((M+1)NQ-1)}{(M+1)NQ}\tau_{k,l}}\right]^T.$$

Let $\boldsymbol{\tau} \triangleq [\boldsymbol{\tau}_1^T, \cdots, \boldsymbol{\tau}_K^T]^T$, $\boldsymbol{\alpha} \triangleq [\boldsymbol{\alpha}_1^T, \cdots, \boldsymbol{\alpha}_K^T]^T$, and $\boldsymbol{\theta} \triangleq [\boldsymbol{\tau}^T, \mathrm{Re}\{\boldsymbol{\alpha}\}^T, \mathrm{Im}\{\boldsymbol{\alpha}\}^T]^T$, the latter collecting all $3L$ unknown parameters of interest, where $L \triangleq \sum_{k=1}^{K} L_k$.

By the Slepian–Bang formula, the CRB for the unknown vector $\boldsymbol{\theta}$ is given by (e.g., [28, App. B])

$$\begin{aligned}\mathrm{CRB}^{-1}(\boldsymbol{\theta}) &= 2\mathrm{Re}\left[\frac{\partial \boldsymbol{m}^H(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\boldsymbol{\Sigma}^{-1}\left(\sigma_e^2\right)\frac{\partial \boldsymbol{m}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T}\right] \\ &= \frac{2}{(M+1)NQ\sigma_e^2}\mathrm{Re}\left[\frac{\partial \boldsymbol{m}^H(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\frac{\partial \boldsymbol{m}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T}\right]\end{aligned} \tag{28}$$

where we used the fact that $\boldsymbol{\Sigma} \triangleq E\{\boldsymbol{w}\boldsymbol{w}^H\} = E\{\boldsymbol{\mathcal{F}}_M\bar{\boldsymbol{\omega}}\bar{\boldsymbol{\omega}}^H\boldsymbol{\mathcal{F}}_M^H\} = (M+1)NQ\sigma_e^2\mathbf{I}_{(M+1)NQ}$, assuming the channel noise is white, and where σ_e^2 denotes the variance of the time-domain samples of the noise.

Let $\boldsymbol{\psi}_{k,l} \triangleq \partial\boldsymbol{\phi}(\tau_{k,l})/\partial\tau_{k,l} = -j2\pi/((M+1)NQ)[0, e^{-j2\pi\tau_{k,l}/(M+1)NQ}, \ldots, ((M+1)NQ-1)e^{-j2\pi((M+1)NQ-1)\tau_{k,l}/(M+1)NQ}]^T$, and $\boldsymbol{\Psi}_k \triangleq [\alpha_{k,1}\boldsymbol{\psi}_{k,1}, \cdots, \alpha_{k,L_k}\boldsymbol{\psi}_{k,L_k}]$. It can be shown that the partial differentiations in (28) can be evaluated as

$$\frac{\partial \boldsymbol{m}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T} = [\mathbf{J}_{\boldsymbol{\tau}},\ \mathbf{J}_{\boldsymbol{\alpha}},\ j\mathbf{J}_{\boldsymbol{\alpha}}] \tag{29}$$

where

$$\mathbf{J}_{\boldsymbol{\tau}} = [\boldsymbol{X}_1\boldsymbol{\Psi}_1, \cdots, \boldsymbol{X}_K\boldsymbol{\Psi}_K], \tag{30}$$
$$\mathbf{J}_{\boldsymbol{\alpha}} = [\boldsymbol{X}_1\boldsymbol{\Phi}(\boldsymbol{\tau}_1), \cdots, \boldsymbol{X}_K\boldsymbol{\Phi}(\boldsymbol{\tau}_K)]. \tag{31}$$

Substituting (29) into (28) yields

$$\mathrm{CRB}(\boldsymbol{\theta}) = \frac{(M+1)NQ\sigma_e^2}{2}\left(\mathrm{Re}\begin{bmatrix}\mathbf{J}_{\boldsymbol{\tau\tau}} & \mathbf{J}_{\boldsymbol{\tau\alpha}} & j\mathbf{J}_{\boldsymbol{\tau\alpha}} \\ \mathbf{J}_{\boldsymbol{\tau\alpha}}^H & \mathbf{J}_{\boldsymbol{\alpha\alpha}} & j\mathbf{J}_{\boldsymbol{\alpha\alpha}} \\ -j\mathbf{J}_{\boldsymbol{\tau\alpha}}^H & -j\mathbf{J}_{\boldsymbol{\alpha\alpha}} & \mathbf{J}_{\boldsymbol{\alpha\alpha}}\end{bmatrix}\right)^{-1}$$

where

$$\begin{aligned}\mathbf{J}_{\boldsymbol{\tau\tau}} &= \mathbf{J}_{\boldsymbol{\tau}}^H\mathbf{J}_{\boldsymbol{\tau}} \\ &= \begin{bmatrix}\boldsymbol{\Psi}_1^H\boldsymbol{X}_1^H\boldsymbol{X}_1\boldsymbol{\Psi}_1 & \cdots & \boldsymbol{\Psi}_1^H\boldsymbol{X}_1^H\boldsymbol{X}_K\boldsymbol{\Psi}_K \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Psi}_K^H\boldsymbol{X}_K^H\boldsymbol{X}_1\boldsymbol{\Psi}_1 & \cdots & \boldsymbol{\Psi}_K^H\boldsymbol{X}_K^H\boldsymbol{X}_K\boldsymbol{\Psi}_K\end{bmatrix} \\ \mathbf{J}_{\boldsymbol{\tau\alpha}} &= \mathbf{J}_{\boldsymbol{\tau}}^H\mathbf{J}_{\boldsymbol{\alpha}} \\ &= \begin{bmatrix}\boldsymbol{\Psi}_1^H\boldsymbol{X}_1^H\boldsymbol{X}_1\boldsymbol{\Phi}(\boldsymbol{\tau}_1) & \cdots & \boldsymbol{\Psi}_1^H\boldsymbol{X}_1^H\boldsymbol{X}_K\boldsymbol{\Phi}(\boldsymbol{\tau}_K) \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Psi}_K^H\boldsymbol{X}_K^H\boldsymbol{X}_1\boldsymbol{\Phi}(\boldsymbol{\tau}_1) & \cdots & \boldsymbol{\Psi}_K^H\boldsymbol{X}_K^H\boldsymbol{X}_K\boldsymbol{\Phi}(\boldsymbol{\tau}_K)\end{bmatrix} \\ \mathbf{J}_{\boldsymbol{\alpha\alpha}} &= \mathbf{J}_{\boldsymbol{\alpha}}^H\mathbf{J}_{\boldsymbol{\alpha}} \\ &= \begin{bmatrix}\boldsymbol{\Phi}^H(\boldsymbol{\tau}_1)\boldsymbol{X}_1^H\boldsymbol{X}_1\boldsymbol{\Phi}(\boldsymbol{\tau}_1) & \cdots & \boldsymbol{\Phi}^H(\boldsymbol{\tau}_1)\boldsymbol{X}_1^H\boldsymbol{X}_K\boldsymbol{\Phi}(\boldsymbol{\tau}_K) \\ \vdots & \ddots & \vdots \\ \boldsymbol{\Phi}^H(\boldsymbol{\tau}_K)\boldsymbol{X}_K^H\boldsymbol{X}_1\boldsymbol{\Phi}(\boldsymbol{\tau}_1) & \cdots & \boldsymbol{\Phi}^H(\boldsymbol{\tau}_K)\boldsymbol{X}_K^H\boldsymbol{X}_K\boldsymbol{\Phi}(\boldsymbol{\tau}_K)\end{bmatrix}.\end{aligned}$$

Finally, we note that the CRB derived here is approximate and may not be achieved since (27) ignores the aliasing caused by windowing [cf. (5)]. Our numerical results, however, show that the performance of the proposed estimators is quite close to this lower bound.

VI. Numerical Results

We consider a K-user asynchronous CDMA system with BPSK modulation, aperiodic long spreading codes, and bandlimited chip waveforms. The spreading codes are randomly generated with processing gain $N = 16$, and normalized to unit energy per symbol interval (i.e., scaled by $1/\sqrt{N}$). The bandlimited chip waveform is a square-root raised-cosine pulse oversampled with $Q = 2$. We simulate a frequency-selective Rayleigh fading channel with $L_k = 2$ propagation paths for each user, and the path gain $\alpha_{k,l}$ is a complex Gaussian random variable with zero-mean and identical variance P_k/L_k, where P_k denotes the average power of user k. We consider a near-far scenario whereby the average power of the desired user, say user one, is scaled to unity: $P_1 = 1$, while all interfering users transmit at an identical mean power $P_k = P \geq 1$, $k \geq 2$. The *near–far ratio* (NFR) is defined as $10\log_{10} P$ in decibels.

Two performance measures are considered. 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 of the true code-timing. The other measure, which characterizes the acquisition accuracy at a *finer scaler*, is the *root mean-squared error* (RMSE) of the code-timing estimate normalized by the chip duration T_c, given correct acquisition. For the multipath case, we evaluate the performance measures for each path and present the average results averaged over all path estimates. In the sequel, we compare the proposed decorrelating FLS and FWLS estimators with the matched filter (MF) scheme briefly discussed in Section II (also see [1, ch. 5]). All results presented next are based on 400 independent Monte Carlo trials, for which the code-timings, attenuations, data symbols, and channel noise are changed independently from one trial to another.

We first examine the *user capacity*, i.e., the number of users that can be supported within the system, of the three schemes in multipath fading channels. Fig. 2 depicts their performance

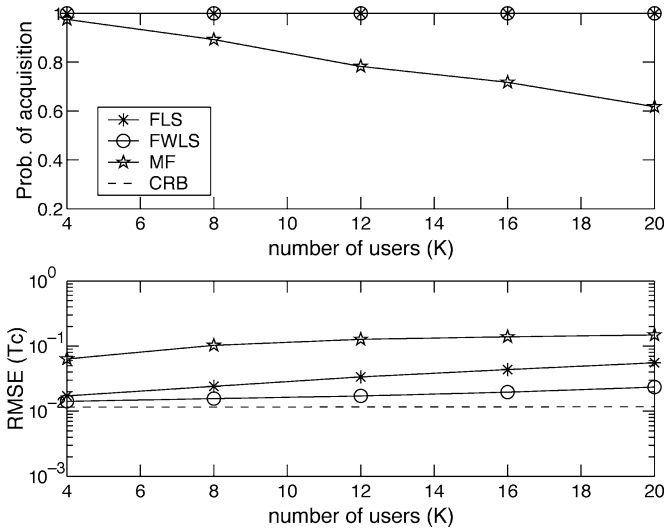

Fig. 2. Performance versus K (user capacity) in multipath fading channels with $N = 16$, $M = 100$, $M_0 = 6$, SNR $= 15$ dB, and NFR $= 10$ dB.

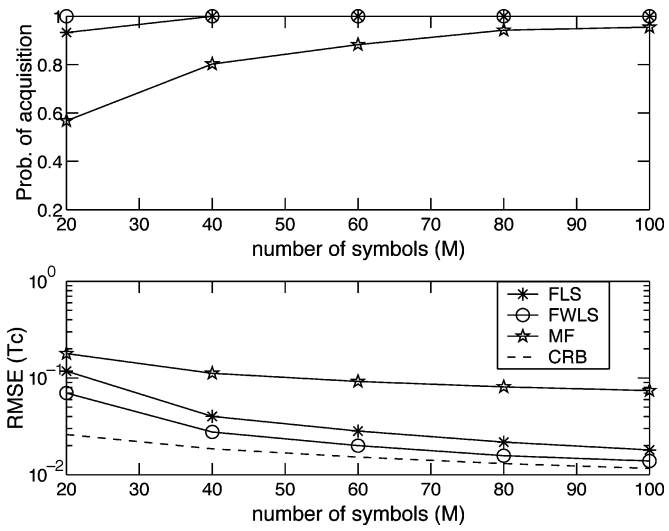

Fig. 3. Performance versus M (acquisition time) in multipath fading channels with $N = 16$, $K = 5$, $M_0 = 6$, SNR $= 15$ dB, and NFR $= 10$ dB.

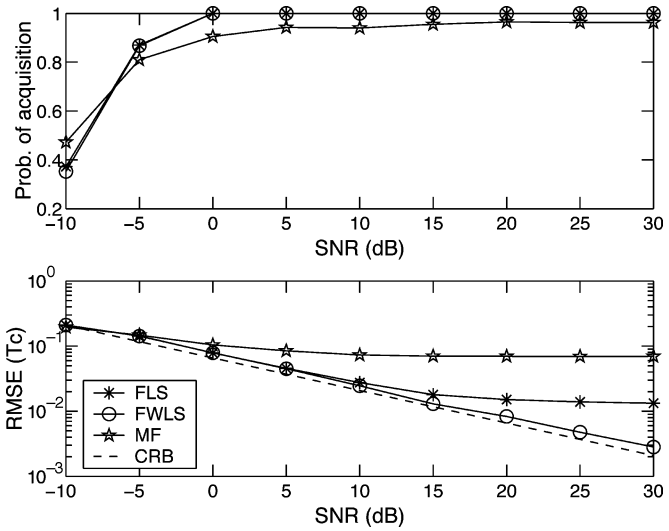

Fig. 4. Performance versus the SNR in multipath fading channels with $N = 16$, $K = 5$, $M = 100$, $M_0 = 6$, and NFR $= 10$ dB.

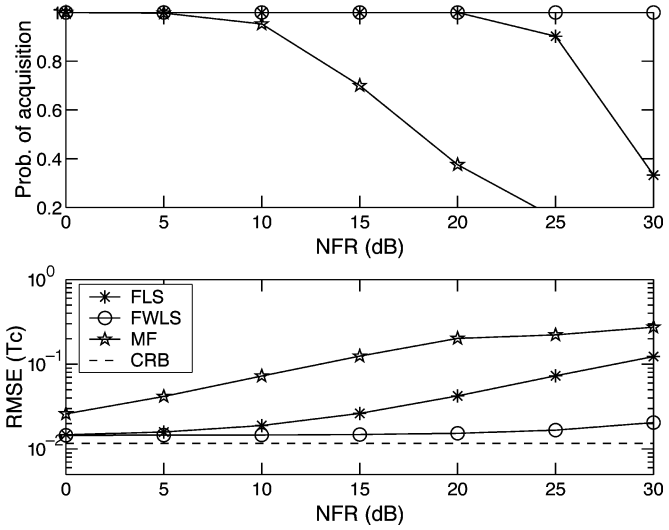

Fig. 5. Performance versus near–far ratio (NFR) in multipath fading channels with $N = 16$, $K = 5$, $M = 100$, $M_0 = 6$, and SNR $= 15$ dB.

when $M = 100$ (number of symbols used for acquisition), $M_0 = 6$ (block size), the DFT frequency selection parameter $\eta = 0.7$ (see Section III-A), SNR $= 15$ dB, and NFR $= 10$ dB. It is seen from Fig. 2 that the proposed FLS and FWLS estimators can support *overloaded* systems (i.e., systems with $K > N$) with little performance degradation and, hence, are MAI-resistant. Meanwhile, the MF estimator degrades quickly as K increases. Fig. 2 shows that the FWLS estimates are statistically more accurate than the FLS estimates and closer to the CRB derived in Section V. Note that the CRB is invariant to K. Although not shown here due to space limitations, we found that the empirical bias of the proposed estimators is at least one order of magnitude smaller than their RMSE. As such, they can be considered as (approximately) unbiased, and the CRB is a suitable lower bound.

We next consider the impact of the *observation time* on acquisition. The simulation parameters are identical to those in the previous example, except that we now fix $K = 5$ users and vary M from 20 to 100. The results are depicted in Fig. 3. We see that the proposed FLS and FWLS estimators require much shorter observation time for acquisition (i.e., faster acquisition) than the MF estimator. For example, the FWLS estimator achieves correct acquisition even with $M = 20$ symbols. In terms of RMSE, the FWLS yields again the most accurate code-timing estimates.

The next example shown in Fig. 4 illustrates the performance of the three methods at different SNR, when $K = 5$, $M = 100$, and the other parameters are identical to those in the previous examples. The advantage of the FLS and FWLS estimators over the MF estimator is again evident. The RMSE results indicate that the FWLS estimator appears the only one that is SNR consistent, i.e., whose RMSE continues to decrease as the SNR increases. The inconsistency of the MF estimator is not surprising, due to the MAI that is not accounted for. For the FLS estimator, it is caused by the residual interference discussed in Section III-B.

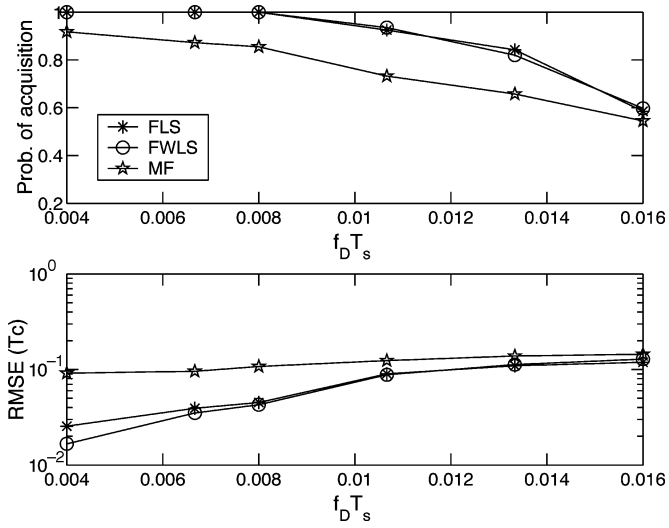

Fig. 6. Performance versus the normalized Doppler rate $f_D T_s$ in time-varying multipath fading channels with $N = 16$, $K = 5$, $M = 60$, $M_0 = 6$, SNR $= 15$ dB, and NFR $= 10$ dB.

We proceed to examine the *near–far resistance* of the three estimators. Fig. 5 depicts their performance as a function of the NFR when $K = 5$, $M = 100$, SNR $= 15$ dB, and the other parameters are identical to those in the previous examples. It is seen that the MF estimator, whose performance degrades quickly when NFR > 5 dB, is not near–far resistant in general. The FLS estimator is able to hold correct acquisition up to NFR $= 20$ dB, which is quite impressive given the relatively small processing gain $N = 16$ employed here. The FWLS estimator is least sensitive to the near–far problem, achieving a good acquisition performance even with an NFR as high as 30 dB. Note that a lower SNR yields similar probability of acquisition in this case, but the RMSE is higher.

For mathematical tractability, the proposed FLS and FWLS estimators were derived based on the assumption that the channel remains unchanged within an observation time of MT_s sec. We now relax this assumption, and test their performance in both time- and frequency-selective channels. The channel attenuations $\alpha_{k,l}$ are varied every $T = T_c/Q$ sec, according to the Jakes' model [32], and the fading rate is characterized by the normalized Doppler frequency $f_D T_s$, where f_D is the maximum Doppler frequency and T_s the symbol interval. Fig. 7 depicts the performance of the three estimators as a function of $f_D T_s$, when $K = 5$, $M = 60$, SNR $= 15$ dB, and NFR $= 10$ dB. It is seen that the proposed estimators perform reasonably well in slow fading channels (i.e., $f_D T_s \leq 0.008$; see Fig. 6). As the fading rate further increases, their performance falls off quickly. In general, code acquisition for long-code bandlimited CDMA in fast fading channels is a much more challenging problem, and good multiuser acquisition schemes for such a scenario have yet to be discovered.

Finally, we briefly comment on the need for L_k, which is the number of path for user k, for the proposed estimators. It is noted that L_k does not need to be known during the *decorrelating* stage, as in Section III-B. The decorrelating stage, which is the major focus of our paper, produces an estimate $\hat{\boldsymbol{\beta}}_k$ that lumps all path contributions in the frequency domain into a single vector. The knowledge of L_k is needed only during the second parameter estimation stage as in Section III-C, where we apply a standard sinusoidal parameter estimation algorithm, namely MUSIC, to fit a complex sinusoidal model to the estimated vector $\hat{\boldsymbol{\beta}}_k$. Model order (i.e., L_k) selection and its impact on sinusoidal parametric model fitting are well studied ([28] and references therein). For the MUSIC estimator, which is a subspace based technique, overestimating L_k will lead to delay estimates for paths that do not exist. Nevertheless, the delay estimates for existing paths are generally not very affected, provided that the estimated signal subspace dimension is much smaller than the estimated noise subspace dimension, i.e., $\hat{L}_k \ll N_v - \hat{L}_k$, where $\hat{L}_k$ is the estimated value of L_k [see (14)]. The condition is easily satisfied since N_v can be chosen to be much larger than L_k. On the other hand, underestimating L_k will result in acquisition of the strongest paths, missing the weaker ones, provided that the signal eigenvectors are not mistaken as noise eigenvectors and used for estimation. The latter can be prevented by using only the eigenvectors associated with relatively small eigenvalues for delay estimation.

VII. Conclusions

In this paper, we have investigated the problem of decorrelating-based multiuser code-timing estimation for CDMA systems with long spreading codes and bandlimited chip waveforms. We have developed two decorrelating FLS and FWLS code-timing estimators by utilizing LS and WLS processing in the frequency domain. We have shown that direct implementations of these decorrelating estimators are, in general, not realistic due to high complexity. To address this difficulty, we have developed efficient implementations of the proposed estimators by exploiting the sparse structure of the large matrices involved. We have shown that such efficient implementations lead to significant complexity reductions compared with direct implementations for the FWLS and, especially, FLS estimators. We have found that the identifiability and statistical stability of the proposed decorrelating estimators are closely related to a signal decomposition process that is critical for decorrelating estimation and provided guidelines how to perform the decomposition to improve the identifiability and stability of the proposed estimators. We have derived the CRB for the estimation problem under study, which is useful to benchmark the proposed and any future estimators for the problem. We have presented extensive simulation results, which show that the FWLS estimator is extremely robust to the MAI and near–far problem, and leads to very accurate estimation in general; meanwhile, the FLS estimator is computationally more appealing than the FWLS estimator and offers good acquisition in most interference environments of practical interest.

Appendix A
Efficient Calculation of $\mathbf{G}^{-1}$ for FLS

In this Appendix, we provide a constructive proof of the efficient algorithm (19) for inverting $\mathbf{G}$ in the FLS estimator. Our proof also leads to the block extension (24) for the FWLS estimator.

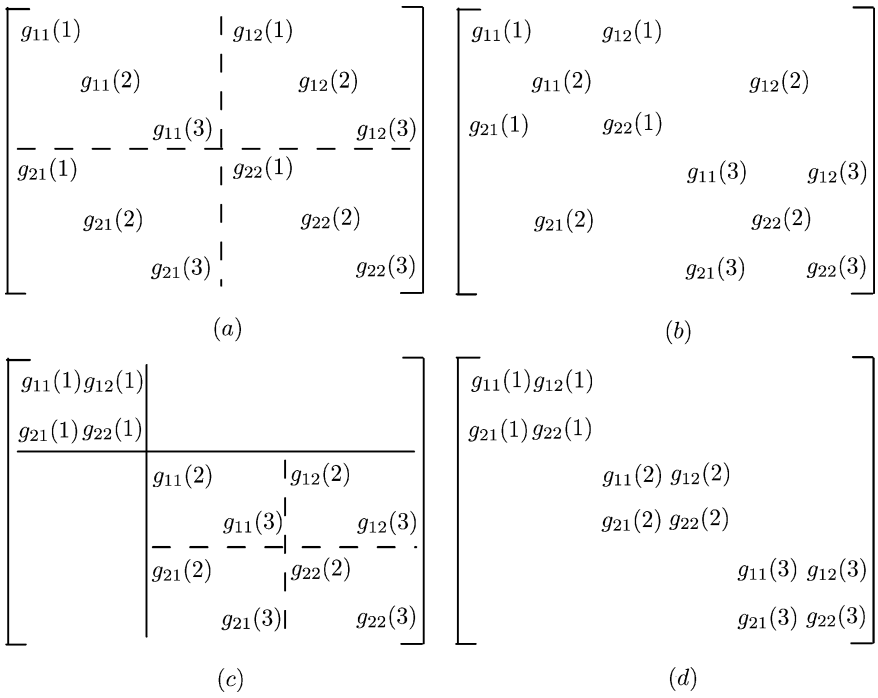

Fig. 7. Graphical illustration of obtaining $\mathbf{C}$ from $\mathbf{G}$ by adjacent row/column permutations, where $M_0N_s = 3$ and $K = 2$, in Example 1. (a) Original matrix $\mathbf{G}$ or $\mathbf{D}^{(3)}$. (b) Swap rows 3 and 4, and columns 3 and 4, by $\mathbf{P}_{3,4}\mathbf{D}^{(3)}\mathbf{P}_{3,4}$. (c) Swap rows 2 and 3 and columns 2 and 3, by $\mathbf{P}_{2,3}\mathbf{P}_{3,4}\mathbf{D}^{(3)}\mathbf{P}_{3,4}\mathbf{P}_{2,3} = \text{diag}\{\mathbf{C}_1, \mathbf{D}^{(2)}\}$, which gives $\boldsymbol{\mathcal{P}}^{(1)} = \mathbf{P}_{3,4}\mathbf{P}_{2,3}$. (d) Swap rows 4 and 5 and columns 4 and 5 by $\mathbf{P}_{4,5}\mathbf{P}_{2,3}\mathbf{P}_{3,4}\mathbf{D}^{(3)}\mathbf{P}_{3,4}\mathbf{P}_{2,3}\mathbf{P}_{4,5} = \text{diag}\{\mathbf{C}_1, \mathbf{C}_2, \mathbf{C}_3\}$, which gives $\boldsymbol{\mathcal{P}}^{(2)} = \mathbf{P}_{4,5}$.

Let the ith element of $\mathbf{g}_{k,l}$ be $g_{k,l}(i)$, $i = 1, \ldots, M_0N_s$. Let $\mathbf{C}_i \triangleq \mathbf{G}(i : M_0N_s : KM_0N_s, i : M_0N_s : KM_0N_s)$, $i = 1, \ldots, M_0N_s$. Note that $\mathbf{C}_i$ is given by

$$\mathbf{C}_i = \begin{bmatrix} g_{1,1}(i) & \cdots & g_{1,K}(i) \\ \vdots & \ddots & \vdots \\ g_{K,1}(i) & \cdots & g_{K,K}(i) \end{bmatrix}_{K \times K}. \tag{32}$$

Let $\mathbf{C} \triangleq \text{diag}\{\mathbf{C}_1, \ldots, \mathbf{C}_{M_0N_s}\} \in \mathbb{C}^{KM_0N_s \times KM_0N_s}$, which can be obtained from $\mathbf{G}$ by row and column permutations. Although there are numerous alternatives to get $\mathbf{C}$ from $\mathbf{G}$, we discuss next a solution that involves only *successive permutations of adjacent rows and columns*. Doing so, we are able to present the solution conveniently in a recursive fashion.

Let $\mathbf{P}_{n,n+1}$ be the permutation matrix obtained by permuting the nth and $(n+1)$st rows of $\mathbf{I}_{KM_0N_s}$. Then, $\mathbf{P}_{n,n+1}\mathbf{G}$ permutes the nth and $(n+1)$st rows of $\mathbf{G}$, $\mathbf{G}\mathbf{P}_{n,n+1}$ permutes the nth and $(n+1)$st columns of $\mathbf{G}$, and $\mathbf{P}_{n,n+1}$ is a symmetric orthogonal matrix [33]. One can see that $(\prod_{n=2}^{M_0N_s} \mathbf{P}_{n,n+1})\mathbf{G}(\prod_{n=2}^{M_0N_s} \mathbf{P}_{n,n+1})^T$ moves up the (M_0N_s+1)st row and, respectively, the (M_0N_s+1)st column of $\mathbf{G}$ to the second row and second column via the $(M_0N_s - 1)$ adjacent row and column permutations (also see the example below). To facilitate our presentation, let $\mathbf{D}^{(\rho)} \in \mathbb{C}^{K\rho \times K\rho}$ be such that its klth subblock $\mathbf{D}^{(\rho)}_{k,l} \triangleq \text{diag}(\mathbf{g}_{k,l}(M_0N_s - \rho + 1 : M_0N_s)) \in \mathbb{C}^{\rho \times \rho}$, $k, l = 1, \ldots, K$ and $\rho = 1, \ldots, M_0N_s$. One can verify that $\mathbf{D}^{(M_0N_s)} = \mathbf{G}$ and $\mathbf{D}^{(1)} = \mathbf{C}_{M_0N_s}$.

Our solution consists of finding $M_0N_s - 1$ *batch* permutation matrices $\boldsymbol{\mathcal{P}}^{(\rho)}$ (formed by product of multiple matrices like $\mathbf{P}_{n,n+1}$) $\rho = 1, \ldots, M_0N_s - 1$ and applying them in succession on $\mathbf{D}^{(M_0N_s)}$ (i.e.,$\mathbf{G}$), which ultimately leads to $\mathbf{C}$. For $\rho = 1$, it can be verified that

$$\boldsymbol{\mathcal{P}}^{(1)^T}\mathbf{D}^{(M_0N_s)}\boldsymbol{\mathcal{P}}^{(1)} = \begin{bmatrix} \mathbf{C}_1 & \\ & \mathbf{D}^{(M_0N_s-1)} \end{bmatrix} \tag{33}$$

where $\boldsymbol{\mathcal{P}}^{(1)} \triangleq \prod_{k=2}^{K}\prod_{n=(k-1)M_0N_s}^{k} \mathbf{P}_{n,n+1}$. It is important to note that $\mathbf{D}^{(M_0N_s-1)}$ has the same structure as that of $\mathbf{D}^{(M_0N_s)}$, due to the use of adjacent row and column permutations (also see the example below). Likewise, we can form $\boldsymbol{\mathcal{P}}^{(2)}, \ldots, \boldsymbol{\mathcal{P}}^{(M_0N_s-1)}$ in a similar fashion, which gives

$$\begin{aligned} &\boldsymbol{\mathcal{P}}^{(M_0N_s-1)^T} \cdots \boldsymbol{\mathcal{P}}^{(2)^T}\boldsymbol{\mathcal{P}}^{(1)^T}\mathbf{D}^{(M_0N_s)} \\ &\quad \times \boldsymbol{\mathcal{P}}^{(1)}\boldsymbol{\mathcal{P}}^{(2)} \cdots \boldsymbol{\mathcal{P}}^{(M_0N_s-1)} \\ &\quad = \text{diag}\{\mathbf{C}_1, \ldots, \mathbf{C}_{M_0N_s}\} = \mathbf{C} \end{aligned} \tag{34}$$

where we recall that $\mathbf{D}^{(M_0N_s)} = \mathbf{G}$.

Example 1: For illustration purposes, consider the simple case where $M_0N_s = 3$ and $K = 2$. The process of adjacent row and column permutation is depicted in Fig. 7. According to the previous discussion, we need two batch permutation matrices to complete the permutation. Fig. 7(a) shows the original $\mathbf{G}$ or $\mathbf{D}^{(3)}$. Fig. 7(b) shows $\mathbf{P}_{3,4}\mathbf{G}\mathbf{P}_{3,4}$, which permutes rows 3 and 4 and columns 3 and 4, respectively, of $\mathbf{G}$. Next, we swap rows 2 and 3 and columns 2 and 3 of $\mathbf{P}_{3,4}\mathbf{G}\mathbf{P}_{3,4}$, as shown in Fig. 7(c), which shows the same structure as in (33). Note that like $\mathbf{D}^{(3)}$, $\mathbf{D}^{(2)}$ is formed by diagonal blocks. In addition, note that the first batch permutation matrix is $\boldsymbol{\mathcal{P}}^{(1)} = \mathbf{P}_{3,4}\mathbf{P}_{2,3}$. Finally, we swap rows 4 and 5 and columns 4 and 5, as shown in Fig. 7(d), which gives $\mathbf{C}$, which is block diagonal, and the second batch permutation matrix is $\boldsymbol{\mathcal{P}}^{(2)} = \mathbf{P}_{4,5}$.

Let $\boldsymbol{\mathcal{P}} \triangleq \prod_{\rho=1}^{M_0N_s-1} \boldsymbol{\mathcal{P}}^{(\rho)}$. Then, (34) can be simplified to $\boldsymbol{\mathcal{P}}^T\mathbf{G}\boldsymbol{\mathcal{P}} = \mathbf{C}$. Since $\boldsymbol{\mathcal{P}}$ is orthogonal, we have

$$\mathbf{G} = \boldsymbol{\mathcal{P}}\mathbf{C}\boldsymbol{\mathcal{P}}^T \tag{35}$$

$$\mathbf{G}^{-1} = \boldsymbol{\mathcal{P}}\mathbf{C}^{-1}\boldsymbol{\mathcal{P}}^T \tag{36}$$

where $\mathbf{C}^{-1} = \text{diag}\{\mathbf{C}_1^{-1}, \ldots, \mathbf{C}_{M_0N_s}^{-1}\}$. Since $\mathbf{C}$ and $\mathbf{C}^{-1}$ are both block diagonal, (36) and (35) indicate that $\mathbf{G}$ and $\mathbf{G}^{-1}$ have the same structure, and [also see (32) and recall $\boldsymbol{\mathcal{G}} = \mathbf{G}^{-1}$]

$$\mathbf{C}_i^{-1} = \boldsymbol{\mathcal{G}}(i : M_0N_s : KM_0N_s, i : M_0N_s : KM_0N_s) \qquad i = 1, \ldots, M_0N_s$$

which is identical to (19).

APPENDIX B
EFFICIENT CALCULATION OF $\mathbf{G}^{-1}$ FOR FWLS

Let $\mathbf{C} \triangleq \text{diag}\{\mathbf{C}_1, \ldots, \mathbf{C}_{M_0}\}$, where

$$\mathbf{C}_m \triangleq \begin{bmatrix} \mathbf{G}_{1,1}(m) & \cdots & \mathbf{G}_{1,K}(m) \\ \vdots & \ddots & \vdots \\ \mathbf{G}_{K,1}(m) & \cdots & \mathbf{G}_{K,K}(m) \end{bmatrix}_{KN_s \times KN_s} \qquad m = 1, \ldots, M_0. \tag{37}$$

Similar to the FLS estimator, $\mathbf{C}$ can be obtained from $\mathbf{G}$ by row and column permutations. However, it can be made more efficient in the current case by *permuting adjacent groups of rows or columns*. In particular, we slice $\mathbf{G}$ into KM_0 equal-sized *groups of rows* (or *groups of columns*), where the first group of rows (columns) are formed by rows (columns) $1, 2, \ldots, N_s$, the second group of rows (columns) are formed by rows (columns) $N_s+1, N_s+2, \ldots, 2N_s$, and so on and so forth. Let $\mathbf{P}_{n,n+1} \triangleq \mathbf{P}'_{n,n+1} \otimes \mathbf{I}_{N_s}$, where $\mathbf{P}'_{n,n+1}$ is the permutation matrix obtained by permuting the nth and $(n+1)$st rows of $\mathbf{I}_{KM_0}$. Then, we can verify that $\mathbf{P}_{n,n+1}\mathbf{G}$ permutes the nth and $(n+1)$st groups of rows of $\mathbf{G}$, whereas $\mathbf{G}\mathbf{P}_{n,n+1}$ permutes the nth and $(n+1)$st groups of columns.

Let $i_{\rho,k} = (\rho-1)K + (k-1)(M_0-\rho+1)$ and $j_{\rho,k} = (\rho-1)K+k$ be the starting and ending indices for matrix permutations. Let $\boldsymbol{\mathcal{P}}^{(\rho)} \triangleq \prod_{k=2}^{K}\prod_{n=i_{\rho,k}}^{j_{\rho,k}} \mathbf{P}_{n,n+1}$, $\rho = 1, \ldots, M_0-1$, which denotes the ρth batch permutation matrix. Then, by direct verification, we have

$$\boldsymbol{\mathcal{P}}^{(M_0-1)^T} \cdots \boldsymbol{\mathcal{P}}^{(2)^T}\boldsymbol{\mathcal{P}}^{(1)^T}\mathbf{G}\boldsymbol{\mathcal{P}}^{(1)}\boldsymbol{\mathcal{P}}^{(2)} \cdots \boldsymbol{\mathcal{P}}^{(M_0-1)} = \text{diag}\{\mathbf{C}_1, \ldots, \mathbf{C}_{M_0}\} = \mathbf{C}. \tag{38}$$

Let $\boldsymbol{\mathcal{P}} \triangleq \prod_{\rho=1}^{M_0-1} \boldsymbol{\mathcal{P}}^{(\rho)}$. Then, (38) reduces to $\boldsymbol{\mathcal{P}}^T\mathbf{G}\mathbf{P} = \mathbf{C}$. It follows that ($\boldsymbol{\mathcal{P}}$ is orthogonal)

$$\mathbf{G} = \boldsymbol{\mathcal{P}}\mathbf{C}\boldsymbol{\mathcal{P}}^T \tag{39}$$

$$\mathbf{G}^{-1} = \boldsymbol{\mathcal{P}}\mathbf{C}^{-1}\boldsymbol{\mathcal{P}}^T \tag{40}$$

where $\mathbf{C}^{-1} = \text{diag}\{\mathbf{C}_1^{-1}, \ldots, \mathbf{C}_{M_0}^{-1}\}$. Since both $\mathbf{C}$ and $\mathbf{C}^{-1}$ are block diagonal, (39) and (40) show that $\mathbf{G}$ and $\mathbf{G}^{-1}$ have the same structure. This, along with (37), indicates (24), where we recall that $\boldsymbol{\mathcal{G}} = \mathbf{G}^{-1}$.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their constructive comments, which led to a significant improvement of the manuscript. One reviewer has our special thanks for suggesting (19) for the FLS estimator. Our original implementations for both FLS and FWLS were based on recursive applications of the matrix inversion lemma. Although the complexities of both implementations in terms of flops are approximately identical, (19) is conceptually simpler. We have thus adopted (19) and extended it for the FWLS estimator.

REFERENCES

[1] R. L. Peterson, R. E. Ziemer, and D. E. Borth, *Introduction to Spread Spectrum Communications*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[2] S. Verdú, *Multiuser Detection*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[3] R. F. Smith and S. L. Miller, "Acquisition performance of an adaptive receiver for DS-CDMA," *IEEE Trans. Commun.*, vol. 47, no. 9, pp. 1416–1424, Sep. 1999.

[4] E. G. Ström and F. Malmsten, "A maximum likelihood approach for estimating DS-CDMA multipath fading channels," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 1, pp. 132–140, Jan. 2000.

[5] D. Zheng, J. Li, S. L. Miller, and E. G. Ström, "An efficient code-timing estimator for DS-CDMA system," *IEEE Trans. Signal Process.*, vol. 45, no. 1, 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, no. 3, pp. 392–399, Mar. 1998.

[7] 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, no. 8, pp. 1425–1436, Aug. 2001.

[8] 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, no. 1, pp. 84–93, Jan. 1996.

[9] S. E. Bensley and B. Aazhang, "Subspace-based channel estimation for code division multiple access communications systems," *IEEE Trans.Commun.*, vol. 44, no. 8, pp. 1009–1020, Aug. 1996.

[10] 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, no. 11, pp. 1628–1631, Nov. 1999.

[11] U. Madhow, "Blind adaptive interference suppression for the near-far resistant acquisition and demodulation of direct-sequence CDMA signals," *IEEE Trans. Signal Process.*, vol. 45, no. 1, pp. 124–136, Jan. 1997.

[12] H. Li and R. Wang, "Filterbank-based blind code synchronization for DS-CDMA systems in multipath fading channels," *IEEE Trans. Signal Process.*, vol. 51, no. 1, pp. 160–171, Jan. 2003.

[13] I. N. Psaromiligkos, S. N. Batalama, and M. J. Medley, "Rapid combined synchronization/demodulation structures for DS-CDMA systems—part I: algorithmic development," *IEEE Trans. Commun.*, vol. 51, no. 6, pp. 983–994, Jun. 2003.

[14] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Upper Saddle River, NJ: Prentice-Hall, 1996.

[15] T. Östman and B. Ottersten, "Near far robust time delay estimation for asynchronous DS-CDMA systems with bandlimited pulse shapes," in *Proc. IEEE 48th Veh. Technol. Conf.*, Ottawa, ON, Canada, May 1998, pp. 1650–1654.

[16] N. Petrochilos and A. J. van der Veen, "Blind time delay estimation in asynchronous CDMA via subspace intersection and ESPRIT," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Salt Lake City, UT, May 2001.

[17] R. Wang, H. Li, and T. Li, "Code-timing estimation for CDMA systems with bandlimited chip waveforms," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1338–1349, Jul. 2004.

[18] M. Torlak, B. L. Evans, and G. Xu, "Blind estimation of FIR channels in CDMA systems with aperiodic spreading sequences," in *Proc. 31th Asilomar Conf. Signals, Syst., Comput.*, vol. 1, Pacific Grove, CA, Nov. 1997, pp. 495–499.

[19] A. J. Weiss and B. Friedlander, "Channel estimation for DS-CDMA downlink with aperiodic spreading codes," *IEEE Trans. Commun.*, vol. 47, no. 10, pp. 1561–1569, Oct. 1999.

[20] Z. Xu, "Low-complexity multiuser channel estimation with aperiodic spreading codes," *IEEE Trans. Signal Process.*, vol. 49, no. 11, pp. 2813–2822, Nov. 2001.

[21] L. Tong, A. van der Veen, P. Dewilde, and Y. Sung, "Blind decorrelating RAKE receiver for long code WCDMA," *IEEE Trans. Signal Process.*, vol. 51, no. 6, pp. 1642–1655, Jun. 2003.

[22] A. Mantravadi and V. V. Veeravalli, "Multiple-access interference-resistant acquisition for band-limited CDMA systems with random sequences," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 7, pp. 1203–1213, Jul. 2000.

[23] V. Tripathi, A. Mantravadi, and V. V. Veeravalli, "Channel acquisition for wideband cdma signals," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 8, pp. 1483–1494, Aug. 2000.

[24] S. Buzzi and H. V. Poor, "On parameter estimation in long-code DS/CDMA systems: Cramér-Rao bounds and least-squares algorithms," *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 545–559, Feb. 2003.

[25] J. G. Proakis, *Digital Communications*, Fourth ed. New York: McGraw-Hill, 2000.

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

[27] B. Fleury, M. Tschudin, R. Heddergott, D. Dahlhaus, and K. Pedersen, "Channel parameter estimation in mobile radio environments using the SAGE algorithm," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 434–450, Mar. 1999.

[28] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*. Upper Saddle River, NJ: Prentice-Hall, 1997.

[29] A. J. van der Veen, M. C. Vanderveen, and A. Paulraj, "Joint angle and delay estimation using shift-invariance techniques," *IEEE Trans. Signal Process.*, vol. 46, no. 2, pp. 405–418, Feb. 1998.

[30] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ: Prentice-Hall, 1993.

[31] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propagat.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[32] W. C. Jakes, Jr., *Microwave Mobile Communications*. New York: Wiley-Interscience, 1974.

[33] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Third ed. Baltimore, MD: Johns Hopkins Univ. Press, 1996.

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 pursuing the Ph.D. degree in electrical engineering with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ.

From 1997 to 2000, he was a Researcher with the Institute of Electronics, Chinese Academy of Sciences, Beijing. He is currently 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, 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. He was a Summer Visiting Faculty Member at the Air Force Research Laboratory, Rome, NY, in the summers of 2003 and 2004. Since July 1999, he has been an Assistant Professor with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ. His current research interests include wireless communications, statistical signal processing, and radars.

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