Transactions Papers

Blind Recursive Tracking of Carrier Frequency Offset (CFO) Vector in MC-CDMA Systems

Feng-Tsun Chien, *Member, IEEE,* and C.-C. Jay Kuo, *Fellow, IEEE*

***Abstract*— A recursive algorithm for estimating and updating the effective carrier frequency offset (CFO) vector in a multicarrier code-division multiple-access (MC-CDMA) system is proposed in this work. The recursive relation is derived based on the expectation maximization (EM) algorithm with a quadratic constraint. This new approach enables the use of linear estimation theory to tackle the CFO estimation problem with or without training data, which leads to an analytic CFO estimate in closed form. Furthermore, the multiple access interference (MAI) is mitigated using the second order statistics of the interference-plus-noise vector, which is updated in a recursive manner under the EM formulation, too. When reaching a converged estimate, a fixed-norm quadratic constraint is imposed so that the final CFO estimate is robust to an imprecise covariance matrix estimate caused by insufficient data samples. It is demonstrated by computer simulation that the performance of an MC-CDMA system without the CFO information can be restored by the proposed scheme in the sense that its bit error probability (BEP) performance is close to that with perfect CFO knowledge.**

***Index Terms*— Code division multiple access (CDMA), multi-carrier CDMA, carrier frequency offset (CFO), recursive EM algorithm, quadratic constraint.**

I. INTRODUCTION

WITH the success of the code-division multiple-access (CDMA) and the multicarrier modulation systems, the idea to integrate merits from both techniques has stimulated a large amount of research during the past decade. In particular, the multicarrier-CDMA (MC-CDMA) system is a potential candidate for the fourth generation (4G) wireless standard thanks to a number of promising features such as robustness to frequency selective fading channels, immunity to inter-symbol interference (ISI) by adding/removing the cyclic prefix and flexibility in allowing multiple access [1]. However, multicarrier systems, including orthogonal frequency division multiplexing (OFDM) and MC-CDMA, are sensitive to the carrier frequency offset (CFO) effect. A prominent CFO destroys the orthogonality between sub-carriers, causing the undesirable inter-carrier interference (ICI) and dramatically deteriorating the system performance.

The range of CFO, which may come from mismatches between local oscillators and/or from mobility induced Doppler shifts, in practical multicarrier applications is often too large to be acceptable. For example, the frequency accuracy 25 parts per million (ppm) specified in IEEE 802.11g can be translated to $2.4\ \text{GHz} \times 25\ \text{ppm} = 60\ \text{KHz}$ frequency shift. Compared to the sub-carrier spacing 312.5 KHz, the normalized CFO is 0.192, which is well above the 0.01 threshold in maintaining a tolerable bit error probability (BEP) performance [2] and must be compensated by means of signal processing techniques in a later stage. It is expected that, when projected into 4G applications where the MC-CDMA system may come into play, there will be an even larger normalized CFO due to reduced subchannel bandwidth.

Motivated by the above observation, there have been a number of studies dedicated to the development of effective algorithms for CFO estimation in the OFDM system [2]–[8]. However, all these schemes cannot be directly applied to the MC-CDMA system due to the presence of multiple access interference (MAI). In addition to the MAI problem, the underlying nonlinear structure makes the CFO estimation a challenging issue in the receiver design for MC-CDMA. Only a limited amount of research in the literature can be found in this regard. A block-based joint CFO and channel estimation scheme for single carrier CDMA systems was considered in [9] based on the subspace projection and the polynomial root-finding algorithms. A blind channel-independent block-based approach, which is similar to [9], was developed in [10] to find a CFO estimate for MC-CDMA systems, wherein each user's channel impulse response was implicitly assumed to be orthogonal. In [11], an iterative receiver structure for MC-CDMA was developed using the technique of generalized sidelobe cancellation, where an estimate of CFO was obtained relying on exhaustive search.

On the other hand, the application of the expectation maximization (EM) algorithm to parameter estimation for OFDM systems has been extensively studied recently, thanks to its

Manuscript received May 22, 2005; revised June 28, 2006; accepted August 26, 2006. The associate editor coordinating the review of this paper and approving it for publication was M. Uysal.

F.-T. Chien is with the Department of Electronics Engineering, National Chiao Tung University, Hsinchu, Taiwan, R.O.C. (email: ftchien@mail.nctu.edu.tw).

C.-C. Jay Kuo is with the Integrated Media Systems Center and the Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089-2564 (email: cckuo@sipi.usc.edu).

Digital Object Identifier 10.1109/TWC.2007.05380.

power in dealing with a broad range of estimation problems with incomplete observations [12], [13]. Estimation of the symbol arrival time and the carrier phase was considered in [14] with the transmitted symbol being treated as the missing data. A blind CFO estimation scheme in MC-CDMA systems using the EM algorithm was studied in [15], where the gradient decent technique was proposed to deal with the nonlinear optimization problem in the CFO estimate. However, since there exist multiple local optima in the cost function, the solution is sensitive to the initialization of the optimization process, which could be problematic for a hill-climbing adaptive algorithm. Besides, all these schemes were operated in a block-based manner assuming that the target parameters were time invariant inside this time block. However, timely updates of system parameters would be more desirable in real time applications, which motivates the development of online algorithms to track possibly time-varying unknowns.

In this work, we address the blind CFO estimation problem in MC-CDMA systems using the sequential EM algorithm [16], [17] and the constrained optimization technique. The EM algorithm provides an iterative procedure to find the maximum likelihood (ML) estimates of target parameters with appropriate initializations. It allows us to treat the CFO estimation problem blindly, *i.e.* without the aid of training data. The algorithm in the training mode is a direct extension from the blind mode. Specifically, rather than finding CFO directly, we derive a recursive relation of the estimate of a time-varying vector consisting of the exponentials of the desired user's CFO, which is referred to as the CFO vector in the sequel. By doing so, the nonlinear CFO estimation problem can be avoided. That is, the resultant cost function in the E-step is quadratic with respect to this CFO vector, which makes the M-step easier to cope with and leads to an analytically tractable estimate for the CFO vector and, hence, a closed-form solution for the CFO itself. We find that no performance loss in terms of the system bit error probability is observed by examining the CFO vector and then estimating the offset. Furthermore, the joint effect of MAI and AWGN is modeled by a colored Gaussian vector and its effect is whitened using the inverse of its covariance matrix [18]–[20]. This covariance matrix can also be updated using the recursive EM formulation. The computational cost associated with matrix inversion is reduced using the matrix inversion lemma. Finally, the performance of the proposed estimator is demonstrated by computer simulations.

The rest of this paper is organized as follows. The system model of an MC-CDMA system is provided and the problem is formulated in Sec. II. The developed recursive algorithm with a quadratic constraint is discussed in Sec. III. Numerical simulations are presented in Sec. IV. Finally, concluding remarks are given in Sec. V.

II. System Model

A. Continuous-Time Signal Model

Consider an MC-CDMA system with N sub-carriers and bandwidth W. The transmitted signal of user k after performing the inverse discrete Fourier transform (IDFT) and adding the cyclic prefix can be represented as

$$d_k(t) = \sum_{m=-\infty}^{\infty} b_k[m] \sum_{n=-N_G}^{N-1} p_{k,n}\psi(t - mT_s - nT), \quad (1)$$

where $b_k[m]$ is the transmitted symbol of user k at the mth symbol block, N_G is the number of samples employed in the cyclic prefix, $T_s = (N_G + N)\cdot T$ is the time span of the mth transmitted symbol block with $T = 1/W$ being the duration of each time-domain chip $p_{k,n}$, $\psi(t)$ is the pulse-shaping waveform assumed to be rectangular of unity amplitude and duration T in this paper and

$$p_{k,n} = \frac{1}{N}\sum_{i=0}^{N-1} c_{k,i}e^{j\frac{2\pi}{N}in}$$

is the nth IDFT output of the signature sequence $\mathbf{c}_k = [c_{k,0}, c_{k,1}\cdots c_{k,N-1}]^T$ of user k with $c_{k,i} \in \{-1,+1\}$ for $0 \le i \le N-1$. Note that we will use s_m to represent $b_1[m]$ in the sequel for notational convenience.

It is assumed that the system experiences a frequency selective fading channel and the following tapped-delay line model is adopted for the channel impulse response [21]

$$h_{k,m}(\tau) = \sum_{l=0}^{L} h_{k,m}[l]\cdot\delta(\tau - \frac{l}{W}),$$

where $h_{k,m}[l]$ is the fading coefficient of user k on the lth path at the mth symbol block, $\delta(t)$ is the Kronecker delta function and L is the order of the channel depending on the maximum multipath delay spread of the channel. The fading gain $h_{k,m}[l]$ is modeled as a zero mean complex-valued Gaussian random process. Then, the received signal for user k in the absence of ambient noise is given by

$$\begin{aligned} r_k(t) &= d_k(t) \star h_k(\tau) \\ &= \sum_{m=-\infty}^{\infty} b_k[m] \sum_{n=-N_G}^{N-1}\sum_{l=0}^{L} p_{k,n}h_{k,m}[l] \\ &\qquad\qquad \times \psi(t - mT_s - nT - lT), \end{aligned}$$

where $\star$ is the convolution operation. Note that the received signal $r_k(t)$ has been derived above in the absence of CFO. A more realistic model which encompasses the CFO effect will be considered in Section II-C.

B. Discrete-Time Signal Model

At the receiver, after passing through the matched filter with a sampling rate $1/T$, the discrete time signal observed at the mth symbol block and the gth chip interval for user k can be written as

$$r_{k,m}[g] = \frac{1}{T}\int_{mT_s+gT}^{mT_s+(g+1)T} r_k(t)\psi(t - mT_s - gT)dt, \quad (2)$$

for $-N_G \le g \le N-1$. The above matched filtering only involves the following integration

$$\int_{mT_s+gT}^{mT_s+(g+1)T} \psi(t - mT_s - nT - lT)\psi(t - mT_s - gT)dt,$$

which takes nonzero values when $g = n + l$. This yields

$$r_{k,m}[g] = b_k[m] \sum_{n+l=g} p_{k,n} h_{k,m}[l],$$

which is equivalent to the result of performing discrete-time convolution between the channel coefficient $\mathbf{h}_{k,m} = [h_{k,m}[0], h_{k,m}[1] \cdots h_{k,m}[L]]^T$ and the time domain signature sequence $\mathbf{p}_k = [p_{k,0}, p_{k,1} \cdots p_{k,N-1}]^T$. After removing the cyclic prefix, the received signal vector at the mth symbol block is equal to

$$\begin{aligned}\mathbf{r}_{k,m} &= [r_{k,m}[0], r_{k,m}[1] \cdots r_{k,m}[N-1]]^T = \mathbf{H}_{k,m}\mathbf{p}_k \cdot b_k[m] \\ &= \frac{1}{N}\mathbf{W}_N^H \Lambda_{k,m}\mathbf{c}_k \cdot b_k[m]\end{aligned}$$

where $\mathbf{H}_{k,m}$ is the $N \times N$ right circular matrix with $[h_{k,m}[0], h_{k,m}[1] \cdots h_{k,m}[L], 0 \cdots 0]^T$ being its first column and superscript H denotes the Hermitian transposition. It is easy to see that $\mathbf{H}_{k,m} = \frac{1}{N}\mathbf{W}_N^H \Lambda_{k,m}\mathbf{W}_N$, where $\mathbf{W}_N$ is the standard DFT matrix with its (m,n)th element equal to $e^{-j\frac{2\pi}{N}(m-1)(n-1)}$ and $\Lambda_{k,m}$ is the diagonal matrix composed of the N-point DFT of $\mathbf{h}_{k,m}$. The discrete-time model for the received signal with K active users in the system can be represented by

$$\mathbf{r}_m = \frac{1}{N}\sum_{k=1}^{K} \mathbf{W}_N^H \Lambda_{k,m}\mathbf{c}_k \cdot b_k[m] + \mathbf{w}_m,$$

where $\mathbf{w}_m$ is a complex additive white Gaussian noise (AWGN) vector with zero mean and covariance matrix $E[\mathbf{w}_m\mathbf{w}_m^H] = \sigma^2\mathbf{I}$. Note that, for $\mathbf{H}_{k,m}$ to be a circular matrix, the length of the cyclic prefix should be no less than the channel order. More generally, the length of the cyclic prefix is designed to absorb both the channel order as well as the residual timing delay estimation errors to avoid ISI. Here, we take the minimum value for N_G, *i.e.* $N_G = L$. In other words, we do not consider asynchronous transmission delays. Furthermore, we assume perfect channel knowledge in this work.

C. Problem Formulation

In the presence of carrier frequency offset, the received signal for user k needs to be modified by a phase shift as $r_k(t)e^{j2\pi\Delta f_k t}$, where Δf_k is the CFO of user k. The matched filter output for is therefore given by

$$\begin{aligned}&\frac{1}{T}\int_{mT_s+gT}^{mT_s+(g+1)T} \psi(t - mT_s - nT - lT) \\ &\qquad\qquad \times\ \psi(t - mT_s - gT)\cdot e^{j2\pi\Delta f_k t}dt \\ &= \rho_k \cdot e^{j\frac{2\pi}{N}(m(N_G+N)+\frac{1}{2}+g)\varepsilon_k} \quad \text{for} \quad g = n + l, \quad (3)\end{aligned}$$

where $\varepsilon_k = \Delta f_k/(W/N)$ denotes the kth user's normalized CFO, which is a relative measure of the amount of CFO in Hertz compared to the sub-carrier spacing W/N and is assumed to be deterministic as well as time-invariant, and $\rho_k = \mathrm{sinc}_\pi(\frac{\varepsilon_k}{N})$, which is approximately equal to one with a large number of sub-carriers. It is assumed that the absolute value of the normalized CFO is no larger than one half of the sub-carrier spacing, *i.e.* $|\varepsilon_k| < 0.5$. For a large number of sub-carriers and the assumption $|\varepsilon_k| < 0.5$, we have scalar $\rho_k \approx 1$. The approximated result $e^{j\frac{2\pi}{N}\left(m(N_G+N)+\frac{1}{2}+g\right)\varepsilon_k}$ in (3) is a time-varying scalar reflecting the effect caused by the CFO, which not only rotates the symbol constellation by a different amount at different symbol timing block m, but also ruins the orthogonality between sub-carriers.

The received signal vector $\mathbf{r}_m$ of the MC-CDMA system in the presence of CFO observed at the mth symbol block after removing the cyclic prefix of a length equal to L can be expressed by

$$\mathbf{r}_m = \frac{1}{N}\sum_{k=1}^{K} \mathbf{F}_{k,m}\mathbf{W}_N^H \mathbf{\Lambda}_{k,m}\mathbf{c}_k \cdot b_k[m] + \mathbf{w}_m, \quad (4)$$

where the effect of CFO for the kth user is modeled by the diagonal matrix

$$\mathbf{F}_{k,m} = \kappa_{k,m} \cdot \mathrm{d}iag\left\{1, e^{j2\pi\varepsilon_k/N}, \cdots, e^{j2\pi(N-1)\varepsilon_k/N}\right\}, \quad (5)$$

where $\kappa_{k,m} = e^{j\frac{2\pi}{N}\left(m(N_G+N)+\frac{1}{2}\right)\varepsilon_k}$ is the scale factor (which is a function of timing index m). Apparently, the orthogonality between sub-carriers has been ruined due to the presence of $\mathbf{F}_{k,m}$, *i.e.*

$$\mathbf{W}_N\mathbf{F}_{k,m}\mathbf{W}_N^H \neq N \cdot \mathbf{I},$$

where the resulting off-diagonal terms characterize the ICI for all sub-carrier pairs. Therefore, before facilitating the DFT operation in the receiver, this residual CFO needs to be compensated. Without loss of generality, we assume the first user to be the user of interest and its signal component can be written as $\mathbf{d}_m = \mathbf{G}_m \cdot \mathbf{f}_m$, where

$$\mathbf{G}_m = \mathrm{d}iag\{\frac{1}{N}\mathbf{W}_N^H \mathbf{\Lambda}_{1,m}\mathbf{c}_1 \cdot s_m\}$$

is a diagonal matrix describing the system structure with $s_m = b_1[m]$ corresponding to the scalar transmitted symbol at the mth symbol block, and

$$\mathbf{f}_m = \kappa_{k,m} \cdot \left[1, e^{j2\pi\varepsilon_k/N}, \cdots, e^{j2\pi(N-1)\varepsilon_k/N}\right]^T$$

is the CFO vector consisting of the diagonal entries of $\mathbf{F}_{1,m}$. The matrix $\mathbf{G}_m$ collects all the parameters other than CFO, and is generally unknown to the receiver. The signal components from all interfering users and the AWGN are modeled by a colored Gaussian vector $\mathbf{n}_m = \frac{1}{N}\sum_{k=2}^{K} \mathbf{F}_{k,m}\mathbf{W}_N^H \mathbf{\Lambda}_{k,m}\mathbf{c}_k \cdot b_k[m] + \mathbf{w}_m$ with covariance matrix

$$\mathbf{R} = \mathbb{E}\left[\mathbf{n}_m\mathbf{n}_m^H\right].$$

Therefore, we have a generic received signal model for the mth time interval

$$\mathbf{r}_m = \mathbf{G}_m \cdot \mathbf{f}_m + \mathbf{n}_m. \quad (6)$$

It is clear from (5) that the CFO vector $\mathbf{f}_m$ satisfies the following dynamic evolution relation

$$\mathbf{f}_m = \mathbf{E} \cdot \mathbf{f}_{m-1}, \quad (7)$$

where $\mathbf{E} = e^{j2\pi(1+N_G/N)\varepsilon_1}\mathbf{I}$. This dynamic evolution plays an important role in the derivation of the recursive EM algorithm for updating $\mathbf{f}_m$.

In this work, rather than obtaining a direct inference to ε_1, we develop a recursive procedure to find an estimate for the CFO vector $\mathbf{f}_m$. In a practical scenario, system parameters, including CFO's, channel impulse responses and the interference-plus-noise vector correlation matrix, are generally unknown to the receiver. For the sake of clarity and brevity of the presentation, only the estimates for the CFO vector $\mathbf{f}_m$ as well as the interference-plus-noise vector correlation matrix $\mathbf{R}$ are considered. In other words, along with the assumption of synchronous transmission, we also assume perfect knowledge of the channel impulse response here. Our objective is the joint estimation of $\mathbf{f}_m$ and correlation matrix $\mathbf{R}$ in a recursive manner based on all received signals up to time m without the aid of the training sequences. The maximum a posteriori (MAP) symbol detection can also be achieved at each EM iteration.

It is however worthwhile to emphasize that the estimation of possibly time-varying channel impulse responses can be incorporated in the iterative procedure using the expectation conditional maximization (ECM) algorithm [22]. Performance loss is expected due to possible phase ambiguity in the estimation of channel coefficients [23]. Besides, the convergence rate is likely to slow down with more unknown parameters included in the ECM iterations [22]. On the other hand, when considering timing synchronization errors, the received signal with ISI under timing inaccuracy can still be modeled by the generic representation in (6), which implies that the proposed algorithm for the CFO vector update can work as well with timing inaccuracy. Under such circumstances, system matrix $\mathbf{G}_m$ needs be modified to encapsulate the effects of ISI, channel impulse responses and timing errors [24, eq. (10)], and should be updated using the ECM algorithm.

III. Constrained Recursive Algorithm

In this section, we first outline the proposed recursive blind algorithm for $\mathbf{f}_m$ with full knowledge of the covariance matrix $\mathbf{R}$. Then, an approach to determine the Lagrange multiplier for a quadratically constrained optimization problem is detailed. Finally, we relax the assumption of knowing $\mathbf{R}$, and consider joint estimation of $\mathbf{f}_m$ and $\mathbf{R}$ based on the ECM algorithm.

A. Recursive EM Formulation

Let $\mathbf{y}_m = [\mathbf{r}_1^T\mathbf{r}_2^T\cdots\mathbf{r}_m^T]^T$ be all the received signals up to time m and $\theta_m = [\mathbf{f}_1^T\mathbf{f}_2^T\cdots\mathbf{f}_m^T]^T$. Then, assuming that covariance matrix $\mathbf{R}$ is available, the updated inference $\mathbf{f}_{m|m}^{(i+1)}$ of $\mathbf{f}_m$ based on $\mathbf{y}_m$ at the $(i+1)$th iteration can be effectively accomplished using the EM algorithm as stated below [16], [17]:

E-step: Find

$$\mathcal{C}_m(\theta_m,\theta_{m|m}^{(i)}) \triangleq \mathbb{E}_{\mathbf{x}_m}\left[\log p\left(\mathbf{y}_m,\mathbf{x}_m;\theta_m\right)\big|\mathbf{y}_m;\theta_{m|m}^{(i)}\right], \quad (8)$$

M-step: Solve

$$\mathbf{f}_{m|m}^{(i+1)} = arg\max_{\mathbf{f}_m}\mathcal{C}_m(\theta_m,\theta_{m|m}^{(i)}), \quad (9)$$

where $\theta_{m|m}^{(i)} = [\mathbf{f}_{m|m}^{(i)^T},\mathbf{f}_{m-1|m-1}^{(c)^T}\cdots\mathbf{f}_{1|1}^{(c)^T}]^T$ with $\mathbf{f}_{l|l}^{(c)}$ representing the converged estimate of $\mathbf{f}_l$ based on $\mathbf{y}_l$, $p(\cdot;\theta)$ and $\mathbb{E}_{\mathbf{x}_m}[\cdot;\theta]$ are the probability density function (pdf) and the expectation operator averaging over $\mathbf{x}_m$, respectively, parameterized by deterministic θ, $\mathbf{x}_m$ is the vector of missing data which is essential to characterize the incomplete-data likelihood function $p(\mathbf{y}_m;\theta_m)$ and is averaged in the E-step. The iterative procedure of the EM algorithm guarantees convergence. With appropriate initializations, the EM iteration will converge to the ML estimate [22]. The choice of missing data is not unique in general. For the blind scenario considered in this research, $\mathbf{x}_m$ is chosen to be the transmitted signals up to time m; *i.e.* $\mathbf{x}_m = [s_1\cdots s_m]^T$. In contrast, in the training mode, $\mathbf{x}_m$ belongs to the empty set. Consequently, in the training mode, we can remove the expectation in (8).

The complete-data, composed by the received data vector, $\mathbf{y}_m$, and the unknown missing data vector, $\mathbf{x}_m$, the log likelihood function of θ_m in (8) can be computed as

$$\log p\left(\mathbf{y}_m,\mathbf{x}_m;\theta_m\right) = \log p\left(\mathbf{y}_m\big|\mathbf{x}_m;\theta_m\right)+\log p\left(\mathbf{x}_m\right)$$

where $\log p\left(\mathbf{x}_m\right)$ does not depend on θ_m and can be discarded in the EM iteration. Therefore, we can equivalently state the EM iterative procedure in the following two steps:

E-step: Find

$$\mathcal{Q}_m(\theta_m,\theta_{m|m}^{(i)}) \triangleq \mathbb{E}_{\mathbf{x}_m}\left[\log p\left(\mathbf{y}_m\big|\mathbf{x}_m;\theta_m\right)\big|\mathbf{y}_m;\theta_{m|m}^{(i)}\right], \quad (10)$$

M-step: Solve

$$\mathbf{f}_{m|m}^{(i+1)} = arg\max_{\mathbf{f}_m}\mathcal{Q}_m(\theta_m,\theta_{m|m}^{(i)}). \quad (11)$$

Having reached the converged estimate $\mathbf{f}_{m|m}^{(c)}$, the final estimate of the CFO vector at the mth symbol block becomes

$$\hat{\mathbf{f}}_{m|m} = arg\max_{||\mathbf{f}_m||^2=N}\mathcal{Q}_m(\theta_m,\theta_{m|m}^{(c)}), \quad (12)$$

where the constant 2-norm constraint is imposed due to the particular structure of the CFO vector; namely, each element of $\mathbf{f}_m$ lies on the unit circle in the complex domain. This quadratic constraint makes the CFO vector estimate more robust to the error of the estimated covariance matrix of the interference-plus-noise vector [25].

Although it is feasible in finding the newly update of (11) and the final estimate (12) at each symbol block interval m with batch processing, a recursive structure is more desirable for lower computational cost and real time implementations. By applying the Taylor series expansion to $\mathcal{Q}_m(\theta_m,\theta_{m|m}^{(i)})$ at $\mathbf{f}_{m|m}^{(i)}$, we can derive a recursive relation for the newly update $\mathbf{f}_{m|m}^{(i+1)}$ in (11) as [16, eq. (3.20)], [26, theorem 2]

$$\mathbf{f}_{m|m}^{(i+1)} = \mathbf{f}_{m|m}^{(i)} - \left(\left.\frac{\partial^2\mathcal{Q}_m(\theta_m,\theta_{m|m}^{(i)})}{\partial\mathbf{f}_m^2}\right|_{\mathbf{f}_m=\mathbf{f}_{m|m}^{(i)}}\right)^{-1} \times\left(\left.\frac{\partial\mathcal{Q}_m(\theta_m,\theta_{m|m}^{(i)})}{\partial\mathbf{f}_m^*}\right|_{\mathbf{f}_m=\mathbf{f}_{m|m}^{(i)}}\right), \quad (13)$$

where the partial derivative relative to complex variable z is defined as $\frac{\partial}{\partial z} \triangleq \frac{1}{2}\left(\frac{\partial}{\partial\Re(z)} - j\frac{\partial}{\partial\Im(z)}\right)$ with $\Re(z)$ and $\Im(z)$ representing the real and the imaginary parts of z, respectively, and the 2nd order derivative is defined as $\partial^2/\partial\mathbf{f}_m^2 \triangleq \partial^2/\partial\mathbf{f}_m^*\partial\mathbf{f}_m^T$ with the superscript $*$ denoting the complex conjugate [17]. Note that (13) is an equality instead of an approximation since $\mathcal{Q}_m(\theta_m,\theta_{m|m}^{(i)})$ is quadratic with respect to $\mathbf{f}_m$.

B. Blind Recursive EM Algorithm

We describe the procedure to obtain the final estimate in (12) and the recursive update in (13) in this subsection. We focus on the development of the recursive algorithm without the aid of training symbols. The training data based approach can be straightforwardly deduced by removing the expectation operator in (10) since the set of missing data is empty. Let

$$\mathbf{P}_m \triangleq \left.\frac{\partial^2 \mathcal{Q}_m(\theta_m, \theta^{(i)}_{m|m})}{\partial \mathbf{f}_m^2}\right|_{\mathbf{f}_m=\mathbf{f}^{(i)}_{m|m}},$$

where the evaluation at $\mathbf{f}_m = \mathbf{f}^{(i)}_{m|m}$ actually has no effect on the result of the matrix $\mathbf{P}_m$ due to the quadratic structure of $\mathcal{Q}_m(\theta_m, \theta^{(i)}_{m|m})$ with respect to $\mathbf{f}_m$. It is shown in the appendix that $\mathbf{P}_m$ has the following recursive structure

$$\mathbf{P}_m = \mathbf{P}_{m-1} - \mathbf{G}_m^H \mathbf{R}^{-1} \mathbf{G}_m. \tag{14}$$

Then, carrying out the derivatives in (13) yields

$$\mathbf{f}^{(i+1)}_{m|m} = \mathbf{f}^{(i)}_{m|m} - \mathbf{P}_m^{-1} \cdot \mathbb{E}_{s_m}\left[\mathbf{G}_m^H \mathbf{R}^{-1}\big(\mathbf{r}_m - \mathbf{G}_m \mathbf{f}^{(i)}_{m|m}\big)\big|\mathbf{r}_m; \mathbf{f}^{(i)}_{m|m}\right], \tag{15}$$

which is of the standard prediction and correction form. The above recursion is initialized by $\mathbf{f}^{(0)}_{m|m}$, which is the maximum likelihood prediction of $\mathbf{f}_m$ based on $\mathbf{y}_{m-1}$, the received signal vector up to time $m-1$. It can be shown from the dynamic evolution given in (7) that

$$\mathbf{f}^{(0)}_{m|m} = \mathbf{E} \cdot \mathbf{f}^{(c)}_{m-1|m-1}, \tag{16}$$

where transition matrix $\mathbf{E}$ is generally unknown to the receiver and should be estimated separately. The estimate of $\mathbf{E}$, or equivalently the estimate of $e^{j2\pi(1+L/N)\varepsilon_1}$, at the mth time instance can be obtained according to the previously updated estimate $\mathbf{f}^{(c)}_{m-1|m-1}$ through all its elements, as will be detailed later in subsection III-E.

Having reached the converged estimate $\mathbf{f}^{(c)}_{m|m}$ during the mth symbol block time interval, we employ the constrained optimization in (12) to find the final estimate. With a replacement of $\mathcal{Q}_m(\theta_m, \theta^{(i)}_{m|m})$ in (13) by

$$\mathcal{J}_m(\lambda_m, \mathbf{f}_m) = \mathcal{Q}_m(\theta_m, \theta^{(c)}_{m|m}) + \lambda_m(||\mathbf{f}_m||^2 - N), \tag{17}$$

which is the cost function of the constrained optimization problem using the Lagrangian multiplier technique and λ_m is the Lagrange multiplier at the mth symbol block time interval, we have the following prediction-correction form

$$\begin{aligned}\hat{\mathbf{f}}_{m|m} = \mathbf{f}^{(c)}_{m|m} - \Big(\mathbf{P}_m + \lambda_m \mathbf{I}\Big)^{-1} \qquad\qquad (18)\\ \times \Big(\mathbb{E}_{s_m}[\mathbf{G}_m^H \mathbf{R}^{-1}\big(\mathbf{r}_m - \mathbf{G}_m \mathbf{f}^{(c)}_{m|m}\big)\big|\mathbf{r}_m; \mathbf{f}^{(c)}_{m|m}] + \lambda_m \mathbf{f}^{(c)}_{m|m}\Big),\end{aligned}$$

as in (15). After some algebraic manipulations, (18) is reduced to

$$\begin{aligned}\hat{\mathbf{f}}_{m|m} = \Big(\mathbf{P}_m + \lambda_m \mathbf{I}\Big)^{-1}\\ \cdot \Big(\mathbf{P}_{m-1}\mathbf{f}^{(c)}_{m|m} - \mathbb{E}_{s_m}[\mathbf{G}_m^H \mathbf{R}^{-1}\mathbf{r}_m\big|\mathbf{r}_m; \mathbf{f}^{(c)}_{m|m}]\Big).\end{aligned} \tag{19}$$

Note that, through the manipulations from (18) to (19), the Lagrange multiplier λ_m is contained only in the inverse in (19). This rearrangement is particularly useful in determining the Lagrange multiplier, as will be detailed in the next subsection. We relegate the derivation of (19) to the appendix.

C. Lagrange Multiplier

We discuss the determination of the Lagrange multiplier for the constrained optimization in (12) for each symbol block time interval in this subsection. More specifically, we need to determine the Lagrange multiplier λ_m in (19) in order to reach the final estimate $\hat{\mathbf{f}}_{m|m}$ that satisfies the constraint $||\hat{\mathbf{f}}_{m|m}||^2 = N$.

It is clear from (14) that $\mathbf{P}_m$ is Hermitian. Thus, it can be decomposed as

$$\mathbf{P}_m = \mathbf{U}_m \mathbf{\Gamma}_m \mathbf{U}_m^H, \tag{20}$$

where $\mathbf{U}_m$ is an unitary matrix consisting of eigenvectors of $\mathbf{P}_m$ and $\mathbf{\Gamma}_m = \mathrm{d}iag(\gamma_{m,1} \cdots \gamma_{m,N})$ with $\gamma_{m,1} \leq \cdots \leq \gamma_{m,N}$ being the corresponding eigenvalues. The Lagrange multiplier λ_m can be determined as a solution to the equation $||\hat{\mathbf{f}}_{m|m}||^2 = N$. Substituting (20) into (19) and taking the square of its norm gives

$$g(\lambda) \triangleq \Big|\Big| \left(\mathbf{\Gamma}_m + \lambda \mathbf{I}\right)^{-1} \mathbf{q}_m \Big|\Big|^2 = \sum_{l=1}^{N} \frac{|q_{m,l}|^2}{\left(\lambda + \gamma_{m,l}\right)^2}, \tag{21}$$

where $\mathbf{q}_m = \mathbf{U}_m^H\big(\mathbf{P}_{m-1}\mathbf{f}^{(c)}_{m|m} - \mathbb{E}_{s_m}\left[\mathbf{G}_m^H \mathbf{R}^{-1}\mathbf{r}_m\big|\mathbf{r}_m; \mathbf{f}^{(c)}_{m|m}\right]\big)$ with $q_{m,l}$ being its lth element. It was shown in [27] that the solution to the Lagrange multiplier of the constrained optimization problem in least squares formulation is equal to the largest real root of $g(\lambda) = N$, and can be resorted to root-finding algorithms for its solution. Here, we employ the Newton method and derive the upper and lower bounds for this optimal root. The obtained upper and lower bounds provide insights into the determination of a good initial guess for the Newton method.

Observing the relation in (21), it is clear that $g(\lambda)$ is monotonically decreasing in $(-\gamma_{m,1}, \infty)$ since $g'(\lambda) < 0$ in this interval with $g(-\gamma_{m,1}) \to \infty$ and $g(\infty) \to 0$. Therefore, there exists an unique solution $\bar{\lambda}$ in this interval such that $g(\bar{\lambda}) = N$. Apparently, this is the largest real root of $g(\lambda) = N$ and is the solution to the Lagrange multiplier, which may take the form

$$\bar{\lambda} = -\gamma_{m,1} + \triangle,$$

where $\triangle$ is a positive real number. Replacing $\gamma_{m,l}$ in (21) by $\gamma_{m,1}$ and $\gamma_{m,N}$, we can obtain the following upper and lower bounds

$$\max\left\{\frac{||\mathbf{q}_m||}{\sqrt{N}} - (\gamma_{m,N} - \gamma_{m,1}), 0\right\} \leq \triangle \leq \frac{||\mathbf{q}_m||}{\sqrt{N}}.$$

In the above interval lies a good initial guess for $\triangle$ and, hence, $\bar{\lambda}$. The upper bound is not a good candidate of the initial guess λ^1, since we have $\lambda^1 = -\gamma_{m,1} + ||\mathbf{q}_m||/\sqrt{N}$ at this point, where $g(\lambda^1)$ has a mild slope, which is likely to result in a newly update smaller than $-\gamma_{m,1}$ and converge to an undesired root. Therefore, it is advised to initialize the Newton algorithm with the lower bound if it is not zero. Otherwise, the Newton algorithm is initialized by

$$\lambda^1 = -\gamma_{m,1} + \frac{||\mathbf{q}_m||}{2^p\sqrt{N}}$$

with a proper choice of positive integer p to set the initial guess between the desired root and $-\gamma_{m,1}$, where the slope of $g(\lambda)$ is sharper.

D. Interference-plus-Noise Correlation Matrix

In Sec. III-A, it was assumed that correlation matrix $\mathbf{R}$ of the interference-plus-noise vector $\mathbf{n}_m$ is known in the development of the recursive algorithm for updating $\mathbf{f}_m$. However, this is in general not true in practice. Here, we will show that $\mathbf{R}$ can also be jointly estimated in a recursive manner based on the ECM algorithm, which is a generalized EM algorithm aiming at finding the ML estimates of multidimensional parameters that do not have a simultaneous closed form to maximize the obtained cost function in the E-step of the EM algorithm [22], [28].

The ECM algorithm states that, conditioned on the previous estimate of one parameter, it is much easier to obtain a closed form representation of the update for another parameter. Likewise, conditioned on this new update, we can go back to update the original parameter. This procedure yields an analytically tractable way to iteratively update all parameters one by one and guarantees an increase of the likelihood as the iteration goes on.

For the problem under consideration, the set of parameters contains $\mathbf{R}$ and $\mathbf{f}_m$, and the ECM algorithm is applicable for their joint estimation. More specifically, conditioned on the ith update of $\mathbf{R}_{|m}^{(i)}$ and $\mathbf{f}_{m|m}^{(i)}$, the new update $\mathbf{f}_{m|m}^{(i+1)}$ for the CFO vector can be derived using (13) with associated $\mathbf{R}$'s being replaced by $\mathbf{R}_{|m}^{(i)}$. Then, having obtained $\mathbf{f}_{m|m}^{(i+1)}$, the update of the correlation matrix in a sequential form is given by

$$\mathbf{R}_{|m}^{(i+1)} = \left(1-\frac{1}{m}\right)\mathbf{R}_{|m-1}^{(c)} + \frac{1}{m}\mathbb{E}_{s_m}\left[\mathbf{n}_{|m}^{(i)}\mathbf{n}_{|m}^{(i)^H}\Big|\mathbf{r}_m;\mathbf{f}_{m|m}^{(i)},\mathbf{R}_{|m}^{(i)}\right], \quad (22)$$

where $\mathbf{R}_{|m-1}^{(c)}$ is the converged estimate of $\mathbf{R}$ based on all received signals up to time $m-1$ and $\mathbf{n}_{|m}^{(i+1)} = \mathbf{r}_m - \mathbf{G}_m\mathbf{f}_{m|m}^{(i+1)}$. This recursive estimate can be initialized either by a warm-up period of length N_p, or by an identity matrix multiplied with a small scalar. With the recursive and iterative update for the correlation matrix $\mathbf{R}$ in (22), the calculation of matrix $\mathbf{P}_m$ in (14) should be modified as

$$\mathbf{P}_m^{(i+1)} = \mathbf{P}_{m-1}^{(c)} - \mathbf{G}_m^H{\mathbf{R}_{|m}^{(i+1)}}^{-1}\mathbf{G}_m. \quad (23)$$

A final remark can be made that, by using the matrix inversion lemma repeatedly, the inversion of $\mathbf{R}_{|m}^{(i+1)}$ only involves a previously calculated matrix inverse and scalar inversions, which greatly reduces the complexity of the proposed algorithm. The computational cost mostly comes from calculating the inverse of $\mathbf{P}_m^{(i+1)}$ and the eigenvalue decomposition of $\mathbf{P}_m^{(c)}$ required at every time instance.

E. Symbol Detection

After arriving at the convergence stage with a final estimated parameter set $\mathbf{f}_{m|m}^{(c)}$, the maximum *a posteriori* probability (MAP) detection for s_m has already been realized in the E-step via [29]

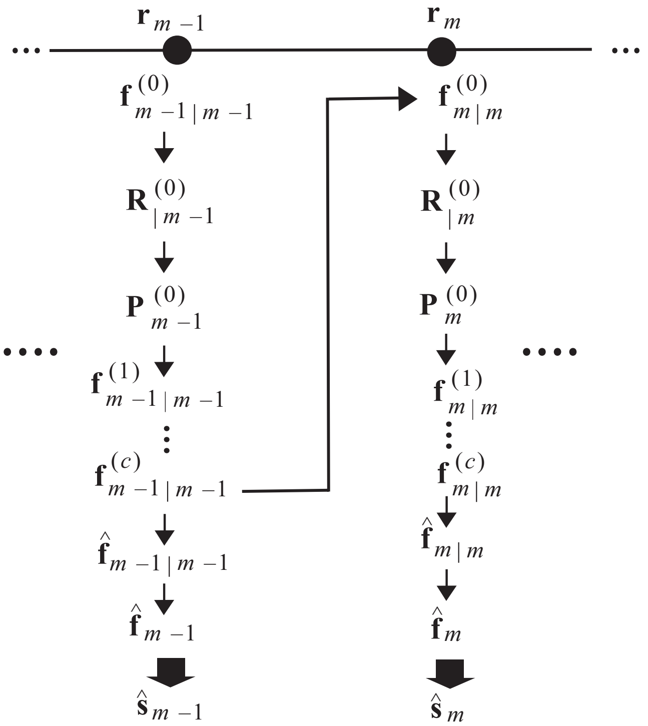

Fig. 1. The update order of estimates in the proposed recursive algorithm at each symbol block m.

$$\hat{s}_m = arg\max_{s_m\in\mathcal{A}} P[s_m|\mathbf{r}_m;\hat{\mathbf{f}}_{m|m}^{(c)},\mathbf{R}_{|m}^{(c)}].$$

However, it should be noted that the EM algorithm does not guarantee convergence to the global optimum. This effect may lead to phase ambiguity in the CFO vector estimate in the complex domain and results in a total erroneous decision for symbol detection.

One way to resolve this problem is to employ the differential encoding/decoding scheme at the cost of 3 dB performance loss. Alternatively, thanks to the particular structure of the CFO vector, the phase ambiguity can be resolved without invoking differential detection, observing that the CFO vector $\mathbf{f}_m$ is actually a geometric sequence with a common factor $e^{j\frac{2\pi}{N}\varepsilon}$ at every symbol block m. For $N \geq 2$, this common factor can be extracted by comparing adjacent entries of the updated CFO vector, regardless of the presence of the phase ambiguity. Therefore, the scalar CFO can be estimated via

$$\hat{\varepsilon}_{|m} = \frac{1}{N-1}\sum_{l=1}^{N-1}\frac{N}{2\pi}\arg\left[(\hat{\mathbf{f}}_{m|m}(l))^*\cdot\hat{\mathbf{f}}_{m|m}(l+1)\right], \quad (24)$$

from which the CFO vector can be reconstructed without the phase ambiguity, and the transition matrix can be inferred, too. Then, the MAP symbol detection can be realized using

$$\hat{s}_m = arg\max_{s_m\in\mathcal{A}} P[s_m|\mathbf{r}_m;\hat{\mathbf{f}}_m,\mathbf{R}_{|m}^{(c)}], \quad (25)$$

where $\hat{\mathbf{f}}_m$ is the reconstructed CFO vector.

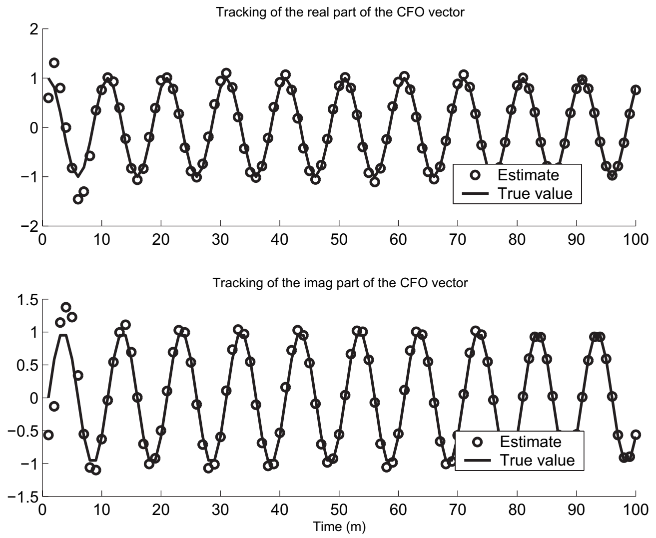

Fig. 2. The tracking behavior of the real and the imaginary parts of the first element of $\mathbf{f}^{(1)}_{m|m}$ in the training mode.

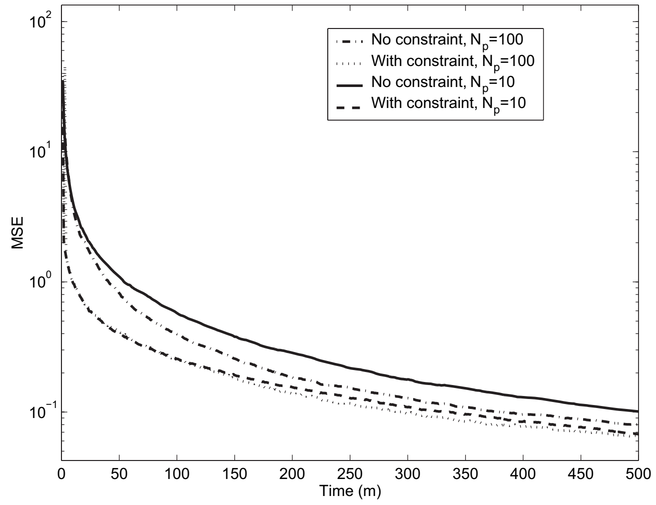

Fig. 4. The MSE of the CFO vector estimate as a function of time using the recursive EM algorithm in the training mode with and without the quadratic constraint.

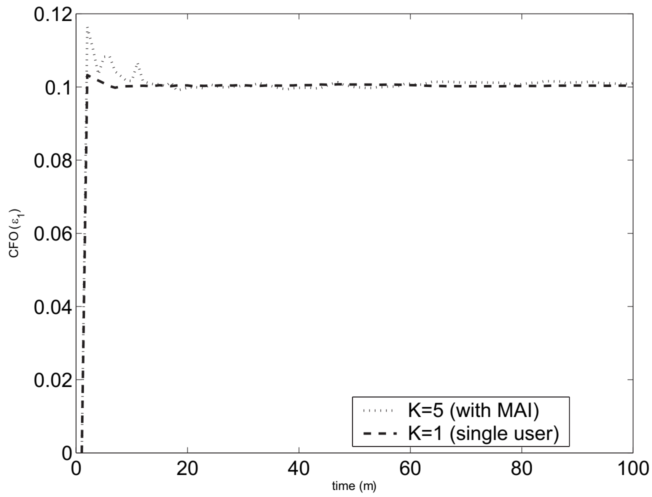

Fig. 3. The normalized CFO vector estimates as a function of tracking time in the training mode.

F. Summary of the Algorithm

The proposed recursive algorithm is summarized below and the order of estimates update is depicted in Fig. 1.

1) Initialize $\mathbf{f}^{(0)}_{m|m}$ using (16) at the beginning of each mth symbol block for $m > 1$.
2) Iteratively update $\mathbf{R}^{(i+1)}_{|m}$, $\mathbf{P}^{(i+1)}_{m}$, and $\mathbf{f}^{(i+1)}_{m|m}$ using (22), (23), and (15), respectively.
3) Impose the quadratic constraint and find $\hat{\mathbf{f}}_{m|m}$ in (19).
4) Calculate $\hat{\varepsilon}_{|m}$ using (24) and reconstruct the CFO vector $\hat{\mathbf{f}}_m$ using $\hat{\varepsilon}_{|m}$.
5) Perform MAP symbol detection in (25).

IV. Simulation Results

In this section, we verify the effectiveness of the proposed recursive algorithm for the CFO vector estimation in the MC-CDMA systems with computer simulations. All numerical experiments were conducted under the environment of the BPSK transmission, the equal power assumption for all users, and a frequency-selective fading channel model with order $L = 5$. Each user's signature sequence was randomly generated, with a total $K = 5$ users in the system. The number of subcarriers and the desired user's CFO were set to $N = 8$ and $\varepsilon_1 = 0.1$, respectively. All simulation tests adopted a common set of parameters, including the channel state information, CFO and thus the covariance matrix of MAI, all assuming time-invariant. The Monte Carlo simulation technique was employed to plot all simulation curves averaged over 250 realizations.

First, we demonstrate the effectiveness of the proposed algorithm in the training mode. Fig. 2 shows the tracking behavior of the first element of the estimated CFO vector. It is observed that the CFO vector is successfully tracked as the number of received signals grows. And, the behavior of the CFO estimate itself is shown in Fig. 3, where the estimate of the CFO itself is obtained via (24). In the presence of MAI, we see that the algorithm requires a larger number of samples for establishing the statistics of the interference-plus-noise correlation matrix until a reliable estimation of the CFO can be reached.

The effect of imposing the quadratic constraint in the training mode with $K = 5$ and $E_b/N_o = 0$ dB is presented in Fig. 4. For the two curves with a warm-up period of $N_p = 10$, the number of data samples is not sufficient in the beginning stage to support a reliable estimate for $\mathbf{R}$ and the recursive algorithm with the quadratic constraint provides a more accurate estimate. This indicates that the proposed algorithm is more robust to imprecise estimation of the correlation matrix. A similar trend is observed in the case when the warm-up period is 100. However, as the number of samples grows, this discrepancy narrows down.

Next, we show the performance of the algorithm in the blind scenario. Fig. 5 demonstrates the mean-squared error (MSE) between the updated and true CFO vectors with respect to the symbol block time index m in the blind case. It is observed that there appears multiple high MSE time instances in the figure. This is due to possible erroneous symbol detection at

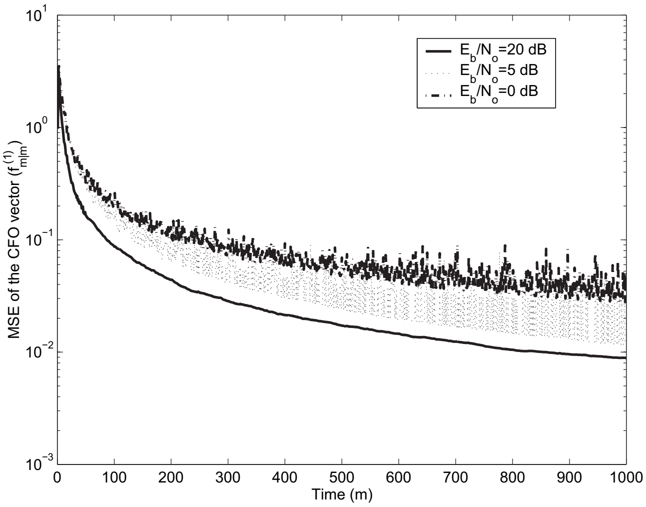

Fig. 5. The MSE of the estimate $\hat{\mathbf{f}}_{m|m}$ as a function of time for different E_b/N_o in the blind scenario.

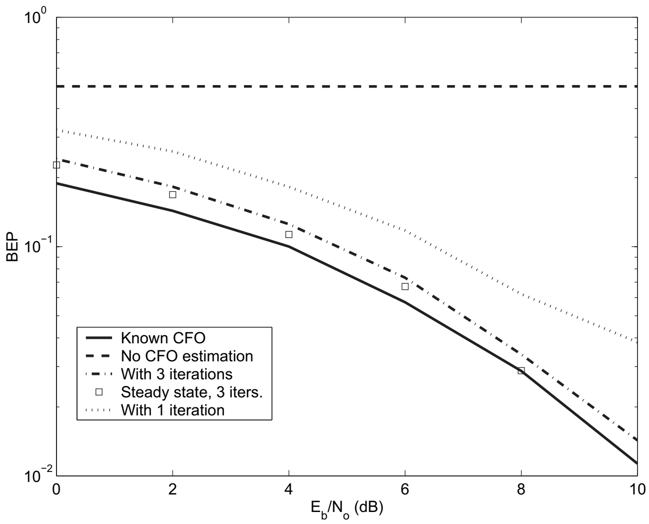

Fig. 6. The evolution of the BEP performance of the MC-CDMA system in the blind scenario with $\varepsilon_1 = 0.1$, $N = 8$, $K = 5$ and $L = 5$.

those time instances occurring in any of the 250 realizations. In other words, a wrong decision of a symbol results in a common sign change to all the components of the CFO vector estimate, which in turn affects the accuracy of the estimate of it. Nevertheless, this effect on the estimate of CFO vector doesn't influence the estimate of the CFO itself, because the CFO estimate is obtained by comparing adjacent components of the updated CFO vector and the phase ambiguity can be resolved in the correlation operation.

The evolution of the BEP performance of the system using the proposed joint blind recursive estimation and detection algorithm is presented in Fig. 6. When no CFO estimation and compensation are employed, the system performance is too bad to be useful. Using the proposed algorithm, we can see from the figure that the BEP performance improves significantly even with only 1 iteration. After 3 iterations, the BEP performance can be recovered within a comparable level to that of an MC-CDMA system with perfect knowledge of CFO.

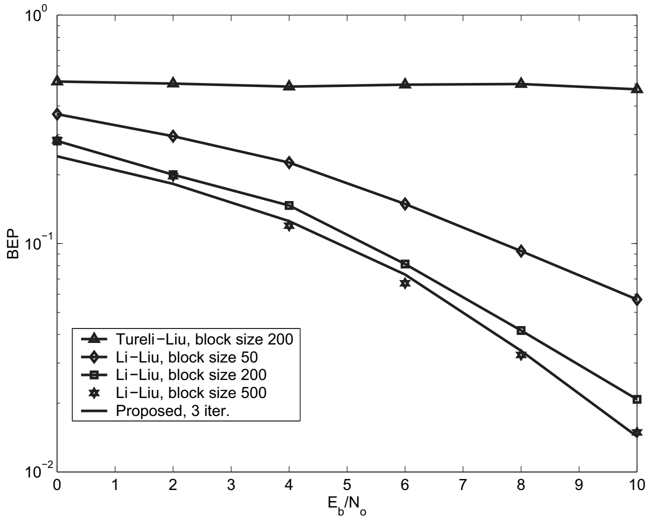

Fig. 7. The BEP performance of the MC-CDMA system in the blind scenario employing the proposed recursive algorithm, Li-Liu's scheme [9] and Toreli-Liu's scheme [10] with $\varepsilon_1 = 0.1$, $N = 8$, $K = 5$, and $L = 5$.

Next, we compare the BEP performance for MC-CDMA systems employing the proposed recursive algorithm, Tureli-Liu's block-based scheme [10] and Li-Liu's block subspace scheme [9] in Fig. 7, where the number of samples M simulated at each point is at least 10^4 for each realization. We assume that the channel is known to the receiver in all schemes for fair comparison. Tureli-Liu's block-based scheme fails to provide an accurate CFO estimate due to its unrealistic assumption on the orthogonality between channel impulse responses of different users. Consequently, we see an unacceptable level of BEP performance in Fig. 7. Li-Liu's block subspace scheme with the minimum mean squared-error (MMSE) detection[1] has performance comparable to that of the proposed recursive algorithm when the number of samples B in each observation block is 500, as seen in Fig. 7. When processing delay and complexity are not of primary concern, increasing the block size in Li-Liu's scheme will improve the system performance. However, there is an implicit assumption in this scheme. That is, the channel impulse responses of all users have to be static for the entire observation block. This requirement can be relaxed in our proposed algorithm. In addition, the computational cost required by the singular value decomposition (SVD) in Li-Liu's scheme with $N = 8$ and block size $B = 500$ is about $4N^2B + 13B^3 \approx 1.625 \times 10^9$ floating point operations (flops) using the R-SVD algorithm [30, p.254]. In contrast, the eigenvalue decomposition in the proposed algorithm using B samples demands about $B(4N^3 + 13N^3) \approx 4.352 \times 10^6$ flops. It is clear that the size of observation samples significantly impedes the efficiency of Li-Liu's scheme in terms of detection delay and computational complexity.

On the other hand, the proposed recursive algorithm does need a sufficient number of samples to get statistics for accurate estimation for both the CFO vector and the correlation

[1]Due to the multiplication factors by the block size in the phase terms in [9, eq. (26)], differential MMSE detection is employed in the simulation to avoid magnifying the compensated residual phase.

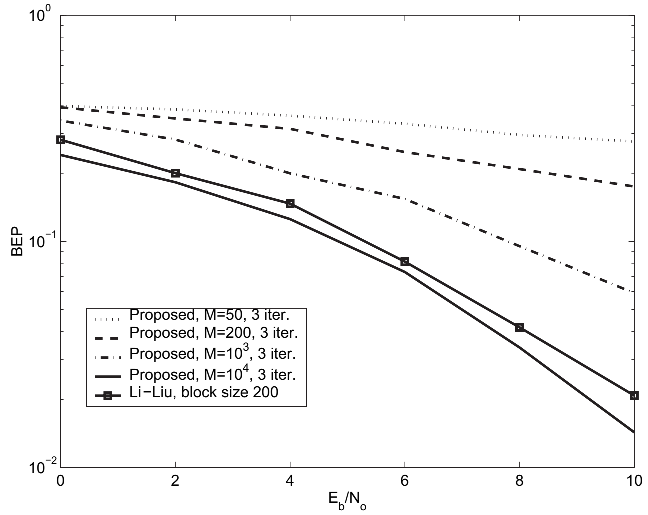

Fig. 8. The BEP performance of the MC-CDMA system in the blind scenario employing the proposed recursive algorithm and Li-Liu's scheme [9] with $\varepsilon_1 = 0.1$, $N = 8$, $K = 5$, and $L = 5$.

matrix. Fig. 8 shows that, when the total number of samples M simulated in each realization is 200 (as opposed to 10^4 samples simulated in Fig. 7), the proposed recursive algorithm does not perform as well as Li-Liu's scheme with a block size of 200 symbols. Nevertheless, we see a trend of performance improvement in the proposed algorithm as the number of the received samples increases. For practical applications such as the IEEE 802.16 standard, the time duration for one symbol block is 100.8 μs. It takes 1.008s to collect 10^4 samples, by which a reasonably acceptable BEP performance can be achieved as shown in Fig. 8.

V. Conclusion

A blind recursive CFO estimation and tracking technique in an MC-CDMA system was proposed in this research. We treated the problem from a new viewpoint so that techniques from linear estimation theory can be used. A recursive relation for the CFO vector was developed based on the EM algorithm with a quadratic constraint. The recursive update for the interference-plus-noise correlation matrix was also derived. It was shown by computer simulation that the BEP of an MC-CDMA system without CFO information can be restored to a level comparable to that with perfect CFO knowledge using the proposed CFO estimation and tracking algorithm.

Appendix

In the appendix, we provide the derivation of (14) and (19). First, we derive the result in (14). From (10) and the fact that each received sample $\mathbf{r}_l$, $1 \le l \le m$, is independent with each other at different time instants, the cost function $\mathcal{Q}_m(\theta_m, \theta^{(i)}_{m|m})$ can be decomposed as

$$\mathcal{Q}_m(\theta_m, \theta^{(i)}_{m|m}) = \sum_{l=1}^{m-1} \mathbb{E}_{s_l}\left[\log p\left(\mathbf{r}_l \mid \mathbf{s}_l; \theta_l\right)\Big|\mathbf{y}_m; \theta^{(i)}_{m|m}\right] + \mathbb{E}_{s_m}\left[\log p(\mathbf{r}_m \mid \mathbf{s}_m; \mathbf{f}_m)\Big|\mathbf{y}_m; \theta^{(i)}_{m|m}\right] \quad (26)$$

The first expectation on the right hand side of (26) can be furthermore carried out as

$$\begin{aligned}
&\mathbb{E}_{s_l}\left[\log p(\mathbf{r}_l \mid s_l; \mathbf{f}_l)\Big|\mathbf{y}_m; \theta^{(i)}_{m|m}\right] \\
&= \sum_{n=1}^{|\mathcal{A}|} p(s_l = \xi_n \mid \mathbf{y}_m; \theta^{(i)}_{m|m}) \log p(\mathbf{r}_l \mid s_l = \xi_n; \mathbf{f}_l) \\
&= \sum_{n=1}^{|\mathcal{A}|} \frac{p(\mathbf{y}_m \mid s_l = \xi_n; \theta^{(i)}_{m|m}) p(s_l = \xi_n)}{p(\mathbf{y}_m; \theta^{(i)}_{m|m})} \log p(\mathbf{r}_l \mid s_l = \xi_n; \mathbf{f}_l) \\
&= \sum_{n=1}^{|\mathcal{A}|} \frac{p(\mathbf{y}_{m-1} \mid \mathbf{s}_l = \xi_n; \theta^{(c)}_{m-1|m-1}) p(\mathbf{r}_m; \mathbf{f}^{(i)}_{m|m}) p(s_l = \xi_n)}{p(\mathbf{y}_{m-1}; \theta^{(c)}_{m-1|m-1}) p(\mathbf{r}_m; \mathbf{f}^{(i)}_{m|m})} \\
&\quad \times \log p(\mathbf{r}_l \mid s_l = \xi_n; \mathbf{f}_l) \quad \Big(\text{since} \quad l \neq m\Big) \\
&= \mathbb{E}_{s_l}\left[\log p(\mathbf{r}_l \mid s_l; \mathbf{f}_l)\Big|\mathbf{y}_{m-1}; \theta^{(c)}_{m-1|m-1}\right]
\end{aligned}$$

For the second expectation in (26), the condition on $\mathbf{y}_m$ can be reduced to $\mathbf{r}_m$. Therefore, we can represent the decomposition in (26) as

$$\begin{aligned}
\mathcal{Q}_m(\theta_m, \theta^{(i)}_{m|m}) &= \mathcal{Q}_{m-1}(\theta_{m-1}, \theta^{(c)}_{m-1|m-1}) \\
&\quad + \mathbb{E}_{s_m}\left[\log p(\mathbf{r}_m \mid \mathbf{s}_m; \mathbf{f}_m)\Big|\mathbf{r}_m; \theta^{(i)}_{m|m}\right].
\end{aligned}$$

With the definition of $\mathbf{P}_m$, we take the 2nd order derivative of the above and obtain

$$\begin{aligned}
\mathbf{P}_m = &\left(\frac{\partial \mathbf{f}^H_{m-1}}{\partial \mathbf{f}^*_m}\right) \mathbf{P}_{m-1} \left(\frac{\partial \mathbf{f}_{m-1}}{\partial \mathbf{f}^T_m}\right) \\
&- \mathbb{E}_{s_m}\left[\mathbf{G}^H_m \mathbf{R}^{-1} \mathbf{G}_m | \mathbf{r}_m; \mathbf{f}^{(i)}_{m|m}\right],
\end{aligned}$$

where the expectation can be further removed since s_m has been cancelled out in $\mathbf{G}_m \mathbf{R}^{-1} \mathbf{G}_m$. Also, from the dynamic evolution in (7), it follows that $\partial \mathbf{f}_{m-1} / \partial \mathbf{f}^T_m = \mathbf{E}^H$ and $\partial \mathbf{f}^H_{m-1} / \partial \mathbf{f}^*_m = \mathbf{E}$, both of which are equivalent to mutually conjugate complex scalars and can also be cleared. Thus, we prove the relationship in (14). Note that $\mathbf{P}_m$ and $\mathbf{P}_{m-1}$ do not depend on any specific realizations of $\mathbf{f}_m$, because the cost function $\mathcal{Q}_m(\theta_m, \theta^{(i)}_{m|m})$ and $\mathcal{Q}_{m-1}(\theta_{m-1}, \theta^{(c)}_{m-1|m-1})$ are quadratic with respect to $\mathbf{f}_m$ and $\mathbf{f}_{m-1}$, respectively.

Next, we show the derivation of (19). It can be easily shown that

$$\begin{aligned}
\hat{\mathbf{f}}_{m|m} = &\mathbf{f}^{(c)}_{m|m} - \Big(\mathbf{P}_m + \lambda_m \mathbf{I}\Big)^{-1} \\
&\times \Big(\mathbb{E}_{s_m}[\mathbf{G}^H_m \mathbf{R}^{-1}(\mathbf{r}_m - \mathbf{G}_m \mathbf{f}^{(c)}_{m|m}) | \mathbf{r}_m; \mathbf{f}^{(c)}_{m|m}] + \lambda_m \mathbf{f}^{(c)}_{m|m}\Big),
\end{aligned} \quad (27)$$

which is the result from replacing $\mathcal{Q}_m(\theta_m, \theta^{(i)}_{m|m})$ in (13) by $\mathcal{J}_m(\theta_m, \lambda_m)$. Gathering all factors associated with $\mathbf{f}^{(c)}_{m|m}$ yields

$$\begin{aligned}
&\Big(\mathbf{I} - \big(\mathbf{P}_m + \lambda_m \mathbf{I}\big)^{-1}\big(-\mathbf{G}_m \mathbf{R}^{-1} \mathbf{G}_m + \lambda_m \mathbf{I}\big)\Big) \mathbf{f}^{(c)}_{m|m} \\
&= \Big(\mathbf{I} - \big(\mathbf{P}_m + \lambda_m \mathbf{I}\big)^{-1}\big(\mathbf{P}_m - \mathbf{P}_{m-1} + \lambda_m \mathbf{I}\big)\Big) \mathbf{f}^{(c)}_{m|m} \\
&= \Big(\mathbf{P}_m + \lambda_m \mathbf{I}\Big)^{-1} \mathbf{P}_{m-1} \mathbf{f}^{(c)}_{m|m},
\end{aligned} \quad (28)$$

where the relation in (14) is used in the derivation. Then, plugging (28) into (27), we obtain the result in (19).

REFERENCES

[1] N. Yee, J. P. Linnartz, and G. Fettweis, "Multi-carrier CDMA in indoor wireless radio networks," in *Proc. IEEE PIMRC 1993*, pp. 109–113.

[2] L. Wei and C. Schlegel, "Synchronization requirements for multi-user OFDM on satellite mobile and two-path Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 43, pp. 887–895, Feb./Mar./Apr. 1995.

[3] P. H. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Trans. Commun.*, vol. 42, pp. 2908–2914, Oct. 1994.

[4] T. Pollet and M. Moeneclaey, "The effect of carrier frequency offset on the performance of band limited single carrier and OFDM signals," in *Proc. IEEE GLOBECOM'96*, vol. 1, pp. 719–723.

[5] M. Luise and R. Reggiannini, "Carrier frequency acquisition and tracking for OFDM systems," *IEEE Trans. Commun.*, vol. 44, pp. 1590–1598, Nov. 1996.

[6] J.-J. van de Beek, M. Sandell, and P.-O. Borjesson, "ML estimation of timing and frequency offset in OFDM Systems," *IEEE Trans. Signal Processing*, vol. 45, pp. 1800–1805, July 1997.

[7] T. M. Schmidl and D. C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Trans. Commun.*, vol. 45, pp. 1613–1621, Dec. 1997.

[8] H. Bolcskei, "Blind estimation of symbol timing and carrier frequency offset in wireless OFDM systems," *IEEE Trans. Commun.*, vol. 49, no. 6, pp. 988–999, 2001.

[9] K. Li and H. Liu, "Joint channel and carrier offset estimation in CDMA communications," *IEEE Trans. Signal Processing*, vol. 47, no. 7, pp. 1811–1822, July 1999.

[10] U. Tureli, D. Kivanc, and H. Liu, "MC-CDMA uplink-blind carrier frequency offset estimation," in *Proc. 34th Asilomar Conference 2000*, vol. 1, pp. 241–245.

[11] J.-H. Deng and T.-S. Lee, "An iterative maximum SINR receiver for multicarrier CDMA systems over a multipath fading channel with frequency offset," *IEEE Trans. Wireless Commun.*, vol. 2, no. 3, pp. 560–569, May 2003.

[12] Y. Xie and C. N. Georghiades, "Two EM-type channel estimation algorithms for OFDM with transmitter diversity," *IEEE Trans. Commun.*, vol. 51, pp. 106–115, Jan. 2003.

[13] W. Mo, Z. Wang, and A. Dogandẑić, "Iterative channel estimation and decoding for coded MIMO system in unknown spatially correlated noise," in *Proc. Allerton Conf. 2003*, pp. 228–232.

[14] E. Panayirci and C. N. Georghiades, "Joint ML timing and phase estimation in OFDM systems using the EM algorithm," in *Proc. IEEE ICASSP 2000*, vol. 1, pp. 2949–2952.

[15] F.-T. Chien and C.-C. J. Kuo, "Joint symbol detection and channel estimation for MC-CDMA systems in the presence of carrier frequency offset," in *Proc. IEEE VTC'04 Fall*.

[16] V. Krishnamurthy and J. B. Moore, "On-line estimation of hidden Markov model parameters based on the Kullback-Leibler information measure," *IEEE Trans. Signal Processing*, vol. 4, no. 8, pp. 2557–2573, Aug. 1993.

[17] H. Zamiri-Jafarian and S. Pasupathy, "EM-based recursive estimation of channel parameters," *IEEE Trans. Commun.*, vol. 47, pp. 1297–1302, Sept. 1999.

[18] 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.

[19] E. Ertin, U. Mitra, and S. Siwamogsatham, "Maximum-likelihood-based multipath channel estimation for code-division multiple-access systems," *IEEE Trans. Commun.*, vol. 49, pp. 290–302, Feb. 2001.

[20] A. A. D'Amico and M. Morelli, "Frequency estimation and timing acquisition in the uplink of a DS-CDMA system," *IEEE Trans. Commun.*, vol. 52, no. 10, pp. 1809–1819, Oct. 2004.

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

[22] X. Meng and D. B. Rubin, "Maximum likelihood estimation via the ECM algorithm: a general framework," *Biometrika*, vol. 80, no. 2, pp. 263–278, June 1993.

[23] S. H. Wu, U. Mitra, and C.-C. J. Kuo, "Graph representation for joint channel estimation and symbol detection," in *Proc. IEEE GLOBECOM'04*, vol. 1, pp. 2329–2333.

[24] M. Morelli, "Timing and frequency synchronization for the uplink of an OFDMA systems," *IEEE Trans. Commun.*, vol. 52, no. 2, pp. 296–306, Feb. 2004.

[25] H. Cox, R. Zeskind, and M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoustic, Speech, Signal Processing*, vol. 35, no. 10, pp. 1365–1376, Oct. 1987.

[26] D. M. Titterington, "Recursive parameter estimation using incomplete data," *J. Roy. Statist. Soc. B*, vol. 46, no. 2, pp. 257–267, 1984.

[27] W. Gandar, "Least squares with a quadratic constraint," *Numer. Math.*, vol. 36, pp. 291–307, 1981.

[28] L. B. Nelson and H. V. Poor, "Iterative multiuser receivers for CDMA channels: an EM based approach," *IEEE Trans. Commun.*, vol. 44, no. 12, pp. 1700–1710, Dec. 1996.

[29] G. K. Kaleh and R. Vallet, "Joint parameter estimation and symbol detection for linear or nonlinear unknown channels," *IEEE Trans. Commun.*, vol. 42, no. 7, pp. 2406–2413, July 1994.

[30] G. H. Golub and C. F. V. Loan, *Matrix Computations, Third Edition* The Johns Hopkins University Press, 1996.

Feng-Tsun Chien (S'02-M'05) received the B.S. degree from the National Tsing Hua University, Hsinchu, in 1995, the M.S. degree from the National Taiwan University, Taipei, in 1997, and the Ph.D. degree from the University of Southern California, Los Angeles, in 2004, respectively, all in Electrical Engineering.

He joined the Department of Electronics Engineering of the National Chiao Tung University, Hsinchu, in July 2005, as an Assistant Professor. His current research interests include signal processing aspects on communications, cross-layer considerations for OFDM and OFDMA systems, multicarrier CDMA, and MIMO-OFDM systems.

C.-C. Jay Kuo (S'83-M'86-SM'92-F'99) received the B.S. degree from the National Taiwan University, Taipei, in 1980 and the M.S. and Ph.D. degrees from the Massachusetts Institute of Technology, Cambridge, in 1985 and 1987, respectively, all in Electrical Engineering. Dr. Kuo was Computational and Applied Mathematics (CAM) Research Assistant Professor in the Department of Mathematics at the University of California, Los Angeles, from October 1987 to December 1988. Since January 1989, he has been with the University of Southern California, where he is currently Professor of Electrical Engineering, Computer Science and Mathematics and Director of the Signal and Image Processing Institute. His research interests are in the areas of digital signal and image processing, multimedia compression, communication and networking technologies. Dr. Kuo has guided about 80 students to their Ph.D. degrees and supervised 15 postdoctoral research fellows. He is co-author of seven books and about 800 technical publications in international conferences and journals.

Dr. Kuo is a Fellow of IEEE and SPIE and a member of ACM. He is Editor-in-Chief for the *Journal of Visual Communication and Image Representation*, and Editor for the *Journal of Information Science and Engineering* and the *EURASIP Journal of Applied Signal Processing*. He was on the Editorial Board of the *IEEE Signal Processing Magazine* in 2003-2004. He served as Associate Editor for *IEEE Transactions on Image Processing* in 1995-98, *IEEE Transactions on Circuits and Systems for Video Technology* in 1995-1997 and *IEEE Transactions on Speech and Audio Processing* in 2001-2003. He received the National Science Foundation Young Investigator Award (NYI) and Presidential Faculty Fellow (PFF) Award in 1992 and 1993, respectively.