Blind Maximum-Likelihood Carrier-Frequency-Offset Estimation for Interleaved OFDMA Uplink Systems

Hung-Tao Hsieh and Wen Rong Wu, *Member, IEEE*

***Abstract*—Blind maximum-likelihood (ML) carrier-frequency-offset (CFO) estimation is considered to be difficult in interleaved orthogonal frequency-division multiple-access (OFDMA) uplink systems. This is because multiple CFOs have to be simultaneously estimated (each corresponding to a user's carrier), and an exhaustive multidimensional search is required. The computational complexity of the search may be prohibitively high. Methods such as the multiple signal classification and the estimation of signal parameters via the rotational invariance technique have been proposed as alternatives. However, these methods cannot maximize the likelihood function, and the performance is not optimal. In this paper, we propose a new method to solve the problem. With our formulation, the likelihood function can be maximized, and the optimum solution can be obtained by solving a polynomial function. Compared with the exhausted search, the computational complexity can be reduced dramatically. Simulations show that the performance of the proposed method can approach that of the Cramér–Rao lower bound.**

***Index Terms*—Carrier-frequency offset (CFO), multiuser system, orthogonal frequency-division multiple-access (OFDMA) .**

I. Introduction

EMERGING as a promising technology for next-generation broadband wireless network, orthogonal frequency-division multiple access (OFDMA) has received a considerable amount of research interest recently [3]–[10]. An appealing feature of OFDMA is that the transmission signals of different users are orthogonal and that multiple-access interference (MAI) can be avoided. However, if the carrier-frequency offsets (CFOs) between transmitters and the receivers are not properly estimated and compensated, then the orthogonality will be destroyed, and intercarrier interference and MAI will arise. In OFDMA downlink systems, the signals for different users are multiplexed by the same transmitter, and the receiver of each user can estimate and compensate its own CFO easily. In such a scenario, methods for CFO estimation in single-user orthogonal frequency-division-multiplexing (OFDM) systems [1], [2] can directly be applied. However, in OFDMA uplink systems, all of the users' CFOs have to be simultaneously estimated at the base-station (BS) receiver, which is considered to be a more challenging problem. Many methods have been developed for CFO estimation in OFDMA uplink systems [3]–[10]. Those methods can roughly be classified into two categories: methods 1) using and 2) not using training sequences. Methods using training sequences insert a known preamble in front of each data packet, facilitating CFO estimation at the BS receiver [3]–[6]. Methods not using training sequences, which are also referred to as blind methods, manipulate the subcarrier assignment scheme such that the CFO for each user can individually or jointly be estimated [7]–[10] at the BS receiver.

Maximum-likelihood (ML) methods were proposed to estimate the CFOs in [3]–[6] for training-based systems. It turns out that a multidimension (MD) exhaustive search is required to obtain the ML solution, and the computational complexity can be high. As a result, low-complexity ML CFO estimators were then developed. In [3], the CFOs were estimated by maximizing the mean likelihood function, and an important sampling method [13] originally proposed for single-carrier systems was applied. In [4], an iterative scheme, which is referred to as the alternating projection frequency estimator (APFE), is employed to conduct 1-D searches for the ML solution. Although their solutions can approach the Cramér–Rao bound (CRB), the required computational complexity is still high. To solve the problem, a simplified method, i.e., the approximate APFE (AAPFE), was then proposed [4]. However, AAPFE suffers considerable performance degradation. To improve the performance of AAPFE, a divide-and-update frequency estimator [5] was later developed. Another suboptimal approach with complexity lower than APFE was reported in [6]. This scheme achieves complexity reduction by approximating the inverse of a CFO-dependent matrix with that of a predetermined matrix. To achieve better results, the system must have a large number of subcarriers.

When training sequences are not available, the CFO for each user cannot easily be estimated. Fortunately, this problem can be overcome with a proper subcarrier assignment scheme, such as subband-based or interleaved-based estimators. The subband-based CFO estimators [7], [8] require that each user is assigned with some consecutive subcarriers, and the subcarrier sets for different users are well separated in the frequency domain. With the scheme, a filter bank can be used in the BS to extract each user's signal, and then, the conventional CFO estimation methods can be exploited. In interleave-based OFDMA systems, each user's time-domain signal is periodic. With the property, the CFOs for all users can jointly be estimated by the multiple signal classification (MUSIC) [9] or the estimation of

Manuscript received April 2, 2010; revised August 9, 2010 and October 12, 2010; accepted October 14, 2010. Date of publication November 9, 2010; date of current version January 20, 2011. This work was supported in part by the National Chiao Tung University-MediaTek joint research program. The review of this paper was coordinated by Dr. H. Lin.

The authors are with the Department of Communication Engineering, National Chiao Tung University, Hsinchu 300, Taiwan (e-mail: dow.cm93g@nctu.edu.tw; wrwu@faculty.nctu.edu.tw).

Digital Object Identifier 10.1109/TVT.2010.2090179

signal parameters via rotational invariance technique (ESPRIT) methods [10]. Although the computational complexity of these methods is low, the CRB cannot be achieved, and the solutions are not optimal.

In this paper, we investigate the blind CFO estimation problem in the interleaved OFDMA uplink system. Our objective is to develop a low-complexity ML CFO estimation method. The main obstacle in the ML method is that there is an inverted correlation matrix in the likelihood function. The CFOs in the likelihood function become intractable after matrix inversion. Using the matrix inversion lemma, we first transform the correlation matrix into a matrix with smaller size. Then, we express the matrix with a series expansion. By properly truncating the expansion, we can obtain a closed-form expression and solve the optimum CFOs with a root-finding method. Simulations show that the performance of the proposed method can approach the CRB. The computational complexity of the proposed algorithm is as low as that of ESPRIT.

The rest of this paper is organized as follows: Section II briefly describes the system model and derives the proposed CFO estimation method. Section III shows the performance analysis for the proposed method. Section IV provides simulation results and complexity analysis. Finally, Section V draws the conclusions. For convenience, the notations used in this paper are defined as follows: $(\cdot)^H$ and $(\cdot)^*$ denote the Hermitian and complex conjugate operation of a matrix, respectively; $\mathrm{mod}(.,C)$ denotes modulo C; Im(.) and Re(.) denote the image-part-taking and the real-part-taking operation, respectively; $E\{\cdot\}$ denotes the expectation operation; $\delta(\cdot)$ denotes a Dirac delta function; $(\cdot)_p$ denotes the pth element of a vector; $(\cdot)_{p,q}$ denotes the (p,q)th element of a matrix; $\mathbf{0}_l$ denotes an l-by-1 column zero vector, $\mathbf{I}$ denotes an identity matrix; $\mathrm{diag}(\mathbf{x})$ denotes a diagonal matrix with the diagonal entries of $\mathbf{x}$; $\det(\cdot)$ denotes the determinant of a matrix; $\|\cdot\|$ denotes the Frobenius norm of a matrix; and $\mathrm{tr}(\cdot)$ denotes the trace of a matrix.

II. Proposed Carrier-Frequency Offset Estimation Method

A. Signal Model for Interleaved OFDMA Uplink System

In an OFDMA system, let M users share the N_s subcarriers of an OFDM symbol, and let the M users simultaneously transmit their data streams. The subcarriers are divided into Q subchannels, and each subchannel has $N = N_s/Q$ subcarriers. Each user occupies a specific subchannel, and the subcarriers assigned to user m are denoted as s_m^k's, where $k \in \Upsilon_m$. Here, Υ_m denotes a subset of subcarrier indices. For an interleaved OFDMA system [9], [10], the subset for the mth user is defined as $\Upsilon_m = \{q_m, Q + q_m, \ldots, q_m + (N-1)Q\}$, where q_m is the subchannel index, and $q_m \in \{0, 1, \ldots, Q-1\}$. In the system, it is assumed that $\Upsilon_m \bigcap \Upsilon_k = \phi$ for $m \neq k$, where ϕ denotes the empty set. In our system, we assume that the sequence each user transmits is unknown to the BS and $M < Q$.

Consider a specific OFDMA symbol and denote the frequency-domain signal that user m transmits as an $N_s \times 1$ vector $\mathbf{u}_m$. Note that the elements of $\mathbf{u}_m$ are nonzero only in designated subcarriers, i.e., Υ_m. Taking the inverse discrete Fourier transform (IDFT) of $\mathbf{u}_m$, we can obtain the time-domain signal for user m, which is denoted as $\bar{\mathbf{s}}_m = [\bar{s}_m(0), \ldots, \bar{s}_m(N_s-1)]^T$. Inserting a cyclic prefix (CP) of length L at the beginning of the symbol, user m can then serially transmit the resultant signal through a wireless channel. Let the channel response from user m to the BS receiver be denoted as $h_m(l), l = 0, \ldots, L_m - 1$, where L_m is the channel length, and $L_m \leq L$. In addition, let the normalized CFO for user m be denoted as ε_m. Then, the CP-removed received OFDMA symbol at the BS can be expressed as

$$y(k) = \sum_{m=1}^{M} \exp(j2\pi\varepsilon_m k/N_s) \sum_{l=0}^{L_m-1} h_m(l)\bar{s}_m(k-l) + \eta(k) \tag{1}$$

where $k = 0, \ldots, N_s - 1$, and $\eta(k)$ represents the additive white Gaussian noise (AWGN) with a variance of σ_η^2.

As mentioned, subchannel q_m is assigned to user m in the interleaved OFDMA system. It is equivalent to say that user m is assigned to subchannel zero, and a CFO of q_m is introduced. Therefore, the received noiseless symbol from user m can be rewritten as

$$\begin{aligned}\bar{x}_m(k) &= \exp(j2\pi\varepsilon_m k/N_s) \sum_{l=0}^{L_m-1} h_m(l)\bar{s}_m(k-l)\\ &= \exp\left(j2\pi(\varepsilon_m + q_m)k/N_s\right) \sum_{l=0}^{L_m-1} h_m(l)s_m(k-l)\\ &= w^{\varepsilon_{e,m}\cdot k}x_m(k)\end{aligned} \tag{2}$$

where $w = \exp(j2\pi/N_s)$, $\varepsilon_{e,m} = \varepsilon_m + q_m$, $x_m(k) = \sum_{l=0}^{L_m-1} h_m(l)s_m(k-l)$, and $s_m(k)$ is the transmitted signal of user m if subchannel zero is assigned. The term $\varepsilon_{e,m}$ denotes the effective CFO for user m. It includes the virtual CFO caused by the subchannel q_m. Note that the periodicity of the transmitted sequence still remains after it is passed through the channel. Since the time-domain signal has a period of N, we can make an index transformation by letting $k = (p-1)N + n$, where $p = 1, \ldots, Q$, and $n = 0, \ldots, N-1$. With the transformation, we can convert the kth sample of a signal into the nth sample in the pth period. The nth sample in each period, corresponding to a signal, can then be extracted to form a vector. Then, we have

$$\mathbf{y}(n) = \mathbf{U}\mathbf{D}(n)\mathbf{x}(n) + \boldsymbol{\eta}(n) \tag{3}$$

where $\mathbf{y}(n) = [y(n), y(n+N), \ldots, y(n+(Q-1)N)]^T = [y_1(n), y_2(n), \ldots, y_Q(n)]^T$, $\mathbf{U}$ is a Q-by-M matrix and $(\mathbf{U})_{p,q} = w^{\{\varepsilon_{e,q}\cdot(p-1)N\}}$, $\mathbf{D}(n) = \mathrm{diag}([w^{n\cdot\varepsilon_{e,1}}, \ldots, w^{n\cdot\varepsilon_{e,M}}]^T)$, $\mathbf{x}(n) = [x_1(n), \ldots, x_M(n)]^T$, and $\boldsymbol{\eta}(n) = [\eta(n), \eta(n+N), \ldots, \eta(n+(Q-1)N)]^T = [\eta_1(n), \eta_2(n), \ldots, \eta_Q(n)]^T$. We will use (3) as our signal model in the derivation of the ML CFO estimate.

B. Proposed Method

To the best of our knowledge, blind ML CFO estimation has not been studied before in OFDMA uplink systems. Here, we propose a method to solve the problem. For interleaved OFDMA uplink systems, the transmitted time-domain signal is obtained from the IDFT of its frequency-domain signal. From the central limit theorem, we know that if the number of subcarriers is reasonably large, then the corresponding time-domain signal can be approximated as a white Gaussian sequence [2]. Similar to [3], we assume that each user is under perfect power control; therefore, signals arrive at the BS with equal average power. If we further assume that each channel tap independently experiences Rayleigh fading and all users' signals are white and independent of each other, then the received sequence $y(k)$ in (1) can also be approximated as a Gaussian sequence [2] with a variance of $M\sigma_x^2 + \sigma_\eta^2$, where $\sigma_x^2 = E\{|x_m(n)|^2\}$. Let $-0.5 < \varepsilon_m < 0.5$ and $f(.)$ be a probability density function. Then, we can explicitly write out the log-likelihood function, which is shown in [2], as

$$\Lambda(\boldsymbol{\varepsilon}) = \ln\left\{\prod_{n=0}^{N-1} f\left(\mathbf{y}(n)\right)\right\}. \tag{4}$$

Define $\mathbf{R}_y = E\{\mathbf{y}(n)\mathbf{y}^H(n)\}$, and

$$(\mathbf{R}_y)_{p,q} = \sigma_\eta^2\delta(p-q) + \sigma_x^2\Gamma(p,q) \tag{5}$$

where

$$\Gamma(p,q) = \sum_{m=1}^{M} w^{(\varepsilon_{e,m})N(p-q)}. \tag{6}$$

Thus, we can express $f(\mathbf{y}(n))$ as [19], [20]

$$f\left(\mathbf{y}(n)\right) = \left(\pi^Q \det(\mathbf{R}_y)\right)^{-1} \exp\left[-\mathbf{y}(n)^H\mathbf{R}_y^{-1}\mathbf{y}(n)\right]. \tag{7}$$

The log-likelihood function can be expressed as

$$\Lambda(\boldsymbol{\varepsilon}) = \sum_{n=0}^{N}\left\{-Q\cdot\ln(\pi) - \ln\left(\det(\mathbf{R}_y)\right) - \mathbf{y}(n)^H\mathbf{R}_y^{-1}\mathbf{y}(n)\right\}. \tag{8}$$

Let $\mathbf{u}(n) = \mathbf{U}\mathbf{D}(n)\mathbf{x}(n)$. Then, $\mathbf{y}(n) = \mathbf{u}(n) + \mathbf{w}(n)$. As assumed, the transmitted sequences are independent of each other, i.e., $\mathbf{R}_y = \sigma_x^2\mathbf{U}\mathbf{U}^H + \sigma_\eta^2\mathbf{I}$. Note that $\mathbf{U}$ is a Q-by-M matrix. To use (8) and solve the M unknown CFOs, $\mathbf{U}$ must be a full-rank tall matrix. From (3), we see that $\mathbf{U}$ is a Vandermonde matrix ($\varepsilon_{e,m} \neq \varepsilon_{e,n}$ if $m \neq n$) [17]. Since we assume that $M < Q$, the full-rank property then holds. As a result, (8) can be applied.

To find the maximum of the log-likelihood function for the ith user, we take a derivative with respect to $\varepsilon_{e,i}$ [16], i.e.,

$$\begin{aligned}\frac{\partial}{\partial\varepsilon_{e,i}}\Lambda(\boldsymbol{\varepsilon}) = &-N\cdot\mathrm{tr}\left[\mathbf{R}_y^{-1}\frac{\partial}{\partial\varepsilon_{e,i}}\mathbf{R}_y\right]\\ &-\sum_{n=0}^{N}\left\{\mathbf{y}(n)^H\left[\left(\frac{\partial}{\partial\varepsilon_{e,i}}\mathbf{R}_y^{-1}\right)\mathbf{y}(n)\right\}.\right.\end{aligned} \tag{9}$$

We use the matrix inversion lemma [18] to write the inverse of $\mathbf{R}_y$ as

$$\begin{aligned}\mathbf{R}_y^{-1} &= \sigma_\eta^{-2}\mathbf{I} - \sigma_\eta^{-4}\mathbf{U}\left(\sigma_x^{-2}\mathbf{I} + \sigma_\eta^{-2}\mathbf{U}^H\mathbf{U}\right)^{-1}\mathbf{U}^H\\ &= \sigma_\eta^{-2}\mathbf{I} - \sigma_\eta^{-4}\mathbf{U}(\mathbf{R}_s)^{-1}\mathbf{U}^H\end{aligned} \tag{10}$$

where $\mathbf{R}_s = \sigma_x^{-2}\mathbf{I} + \sigma_\eta^{-2}\mathbf{U}^H\mathbf{U}$. With (10), we only need the inverse of an M-by-M matrix $\mathbf{R}_s$ rather than a Q-by-Q matrix $\mathbf{R}_y$.

However, $(\mathbf{R}_s)^{-1}$ is difficult to obtain. Even if it can, the relationship between the likelihood function and the CFOs may not be trackable after the inversion. To solve the problem, we propose using the Neumann series to expand $(\mathbf{R}_s)^{-1}$ [14]. Let $\mathbf{S}$ be a nonsingular matrix and $\beta(\mathbf{S})$ be its maximum absolute eigenvalue. Then, the series $\sum_{k=0}^{\infty}\mathbf{S}^k$ will converge to $(\mathbf{I}-\mathbf{S})^{-1}$ [17] if $\beta(\mathbf{S}) < 1$ [15]. However, the condition of $\beta(\mathbf{S}) < 1$ is not always satisfied for a nonsingular $\mathbf{S}$. This problem can be overcome by dividing $\mathbf{S}$ by a real parameter $\lambda > 0$ and expanding the resultant matrix. It is simple to show that there always exists a λ such that $\beta(\mathbf{R}_s/\lambda) < 1$. Now, we can rewrite $\mathbf{R}_s$ as

$$\begin{aligned}\mathbf{R}_s &= \sigma_x^{-2}\mathbf{I} + \sigma_\eta^{-2}\mathbf{U}^H\mathbf{U}\\ &= \lambda(\mathbf{I}+\mathbf{B})\end{aligned} \tag{11}$$

where $\mathbf{B}$ is obtained as $(1/\lambda)\mathbf{R}_s - \mathbf{I}$, and its (p,q)th element is

$$(\mathbf{B})_{p,q} = \left(\frac{1}{\lambda\sigma_x^2} - 1\right)\delta(p-q) + \frac{1}{\lambda\sigma_\eta^2}\sum_{k=1}^{Q} w^{(-\varepsilon_{e,p}+\varepsilon_{e,q})N(k-1)}. \tag{12}$$

From the Neumann series shown, the inverse of $\mathbf{R}_s$ can be expanded as

$$\begin{aligned}\left(\frac{1}{\lambda}\mathbf{R}_s\right)^{-1} &= (\mathbf{I}+\mathbf{B})^{-1}\\ &= \sum_{k=0}^{\infty}(-1)^k\mathbf{B}^k.\end{aligned} \tag{13}$$

For simplicity, we can retain the first three and truncate high-order terms, i.e.,

$$\mathbf{R}_s^{-1} \approx \frac{1}{\lambda}\sum_{k=0}^{2}(-1)^k\mathbf{B}^k. \tag{14}$$

The determination of the optimum λ and the analysis of the truncation error will be discussed in the next section. From (12), we can find the (p,q)th element of $\mathbf{B}^2$ as

$$\begin{aligned}(\mathbf{B}^2)_{p,q} = &\left(\frac{1}{\lambda\sigma_x^2} - 1\right)^2\delta(p-q)\\ &+ \frac{2}{\lambda\sigma_\eta^2}\left(\frac{1}{\lambda\sigma_x^2} - 1\right)\Gamma_0(p,q)\\ &+ \left(\frac{1}{\lambda\sigma_\eta^2}\right)^2\sum_{k=1}^{M}\Gamma_0(p,k)\Gamma_0(k,q)\end{aligned} \tag{15}$$

where $\Gamma_0(p,q) = \sum_{n=1}^{Q} w^{(-\varepsilon_{e,p}+\varepsilon_{e,q})N(n-1)}$. Substituting (12) and (15) into (14), we then obtain

$$\begin{aligned}\left(\mathbf{R}_s^{-1}\right)_{p,q} = \frac{1}{\lambda}\Bigg[&\left(2 - \frac{1}{\sigma_x^2\lambda} + \left(\frac{1}{\sigma_x^2\lambda} - 1\right)^2\right)\delta(p-q)\\ &+ \frac{1}{\sigma_\eta^2\lambda}\left(\frac{2}{\sigma_x^2\lambda} - 3\right)\Gamma_0(p,q)\\ &+ \left(\frac{1}{\sigma_\eta^2\lambda}\right)^2\sum_{k=1}^{M}\Gamma_0(p,k)\Gamma_0(k,q)\Bigg].\end{aligned} \tag{16}$$

Using (16) in (10), we can approximate the inverse of $\mathbf{R}_y$ as

$$\begin{aligned}\left(\mathbf{R}_y^{-1}\right)_{p,q} =& \sigma_\eta^{-2}\delta(p-q) - C_0\Gamma(p,q)\\ & - C_1\sum_{a=1}^{M} w^{(\varepsilon_{e,a})N(p-1)}\\ & \cdot\sum_{b=1}^{M}\Gamma_0(a,b)w^{(-\varepsilon_{e,b})N(q-1)}\\ & - C_2\sum_{a=1}^{M} w^{-(\varepsilon_{e,a})N(q-1)}\sum_{b=1}^{M} w^{(\varepsilon_{e,b})N(p-1)}\\ & \cdot\sum_{k=1}^{M}\Gamma_0(b,k)\Gamma_0(k,a)\end{aligned} \tag{17}$$

where

$$\begin{aligned}C_0 &= \frac{1}{\sigma_\eta^4\lambda}\left(2-\frac{1}{\sigma_x^2\lambda}+\left(\frac{1}{\sigma_x^2\lambda}-1\right)^2\right)\\ C_1 &= \frac{1}{\sigma_\eta^6\lambda^2}\left(\frac{2}{\sigma_x^2\lambda}-3\right)\\ C_2 &= \frac{1}{\sigma_\eta^8\lambda^3}.\end{aligned} \tag{18}$$

Note that $\Gamma_0(.,.)$ cannot directly be estimated. However, it can be combined with some variables in (17) and converted to $\Gamma(.,.)$, as defined in (6). The value of $\Gamma(p,q)$ can be estimated from that of the $(p-q)$th diagonal term of $\mathbf{R}_y$ as

$$\Gamma(p,q) = \begin{cases}\dfrac{\sum\limits_{p=1}^{Q-m}[\mathbf{R}_y]_{p,p+m}}{(Q-m)\sigma_x^2}, & \text{if } m=q-p\geq 0\\ \dfrac{\sum\limits_{q=1}^{Q-m}[\mathbf{R}_y]_{q+m,q}}{(Q-m)\sigma_x^2}, & \text{if } m=p-q>0.\end{cases} \tag{19}$$

The second and third terms in (17) can be rewritten as

$$\begin{aligned}&\sum_{a=1}^{M}\sum_{b=1}^{M}\Gamma_0(a,b)w^{(\varepsilon_{e,a})N(p-1)}w^{(-\varepsilon_{e,b})N(q-1)}\\ &\quad=\sum_{n=1}^{Q}\sum_{a=1}^{M}\sum_{b=1}^{M}w^{(-\varepsilon_{e,a}+\varepsilon_{e,b})N(n-1)}\\ &\quad\quad\cdot w^{(\varepsilon_{e,a})N(p-1)}w^{(-\varepsilon_{e,b})N(q-1)}\\ &\quad=\sum_{n=1}^{Q}\Gamma(p,n)\Gamma(n,q)\end{aligned} \tag{20}$$

$$\begin{aligned}&\sum_{a=1}^{M}w^{-(\varepsilon_{e,a})N(q-1)}\sum_{b=1}^{M}w^{(\varepsilon_{e,b})N(p-1)}\sum_{k=1}^{M}\Gamma_0(b,k)\Gamma_0(k,a)\\ &\quad=\sum_{m=1}^{Q}\sum_{n=1}^{Q}\Gamma(p,m)\Gamma(m,n)\Gamma(n,q).\end{aligned} \tag{21}$$

Substituting (17)–(21) into (9), we obtain (22), shown at the bottom of the page, where

$$\gamma(p,q) = \sum_{n=0}^{N-1}y^*(n+(p-1)N)y(n+(q-1)N) \tag{23}$$

$$x = \exp\left(j2\pi\frac{N\varepsilon_{e,i}}{N_s}\right). \tag{24}$$

The detailed derivation of (22) is provided in Appendix I. Setting (22) to zero, we can solve all the possible $2(Q-1)$ roots, i.e., $\hat{x}$. The effective CFO can then be obtained by

$$\hat{\varepsilon}_{e,i} = \frac{N_s}{N}\left(\frac{\ln(\hat{x})}{j2\pi}\right). \tag{25}$$

As defined in (2), the true CFO for the ith user is then given by

$$\hat{\varepsilon}_i = -q_i + \frac{N_s}{N}\left(\frac{\ln(\hat{x})}{j2\pi}\right) \tag{26}$$

where q_i is the subchannel index for user i. It is apparent that after adding $-q_i$, there will only be one root falling into the

$$\begin{aligned}\frac{\partial}{\partial\varepsilon_{e,i}}\Lambda(\boldsymbol{\varepsilon})=\frac{j2\pi N}{N_s}\Bigg\{\sum_{p=1}^{Q}\sum_{q=1}^{Q}\Bigg\{&N\sigma_x^2(q-p)x^{q-p}\left[C_0\Gamma(p,q)+C_1\sum_{n=1}^{Q}\Gamma(p,n)\Gamma(n,q)+C_2\sum_{m=1}^{Q}\sum_{n=1}^{Q}\Gamma(p,m)\Gamma(m,n)\Gamma(n,q)\right]\\ &+\gamma(p,q)\Bigg[C_0(p-q)x^{p-q}+C_1\sum_{n=1}^{Q}\big((n-q)x^{n-q}\Gamma(p,n)+(p-n)x^{p-n}\Gamma(n,q)\big)\\ &+C_2\sum_{m=1}^{Q}\Gamma(p,m)\sum_{n=1}^{Q}\Gamma(n,q)\big((m-n)x^{m-n}+(n-q)x^{n-q}\Gamma(p,m)\Gamma(m,n)\\ &+(p-m)x^{p-m}\Gamma(m,n)\Gamma(n,q)\big)\Bigg]\Bigg\}\Bigg\}\end{aligned} \tag{22}$$

range of subchannel 0, and the root is the estimated CFO for user i.

A direct method for solving the roots in (22) is via an exhaustive grid search over the interval spanned by $\varepsilon_{e,i}$. However, the computational complexity is high. Taking a closer look at (22), we find that (22) is a polynomial function of x, i.e.,

$$\frac{\partial}{\partial \varepsilon_{e,i}} \Lambda(\boldsymbol{\varepsilon}) = \sum_{k=1}^{Q-1} \alpha_p(k) x^k + \sum_{k=1}^{Q-1} \alpha_n(k) x^{-k} = 0. \tag{27}$$

The detailed derivation for $\alpha_p(k)$ and $\alpha_n(k)$ is provided in Appendix II. Using (27), we can then use a more efficient root-finding method to obtain the roots.

III. Performance Analysis

A. Truncation Error in (14)

As we can see, the series in (13) is infinite, and truncation has to be conducted. In the previous section, we retain the first three terms in the series. One may be curious about how large the error will be. In this section, we analyze the truncation error in (14).

For a positive-definite Hermitian matrix $\mathbf{R}$ with rank K, we can have its eigen-decomposition as

$$\mathbf{R} = \mathbf{V}\mathbf{G}\mathbf{V}^H \tag{28}$$

where $\mathbf{G} = \mathrm{diag}[g_1, \ldots, g_K]$ is a diagonal matrix, where g_i's being positive are the eigenvalues of $\mathbf{R}$ in descending order, i.e., $g_1 > \cdots > g_K$, and $\mathbf{V}$ is an unitary matrix consisting of the eigenvectors. As shown in Section II, $\mathbf{R}$ can also be expressed as

$$\left(\frac{\mathbf{R}}{\lambda}\right)^{-1} = (\mathbf{I} - \mathbf{A})^{-1} = \sum_{k=0}^{\infty} \mathbf{A}^k \tag{29}$$

where λ is a real number ensuring that the maximum absolute eigenvalue of $\mathbf{R}/\lambda$ is smaller than one, and $\mathbf{A}$ is a matrix to be determined. Substituting (28) and $\mathbf{V}\mathbf{V}^H = \mathbf{I}$ into (29), we can obtain that $\mathbf{A} = \mathbf{V}(\mathbf{I} - \mathbf{G}/\lambda)\mathbf{V}^H$ and that

$$\mathbf{A}^k = \mathbf{V}\left(\mathbf{I} - \frac{\mathbf{G}}{\lambda}\right)^k \mathbf{V}^H. \tag{30}$$

From (30), it is simple to see that for the convergence of (29), $|1 - g_i/\lambda|$, $i = 1, 2, \ldots, K$ has to be smaller than 1. In addition, the smaller the value of $|1 - g_i/\lambda|$, the faster the convergence we can have. Since the values of g_i's may be different, the convergent rate of each $|1 - g_i/\lambda|$ (which is referred to as a mode) may be different. As a result, the overall convergence is dominated by the mode with the maximum $|1 - g_i/\lambda|$. To have the fastest convergence, we then want to find a λ minimizing the maximum $|1 - g_i/\lambda| (1 \leq i \leq K)$. This yields a min–max optimization problem as

$$\min_{\lambda} \max_{i=1,\cdots,K} |1 - g_i/\lambda| \tag{31}$$

subject to the constraints

$$|1 - g_i/\lambda| < 1 \tag{32}$$

where $i = 1, 2, \ldots, K$. The optimum value of λ has been shown to be [22]

$$\lambda = \frac{g_1 + g_K}{2}. \tag{33}$$

Substituting (33) into $|1 - g_i/\lambda|$, we find that there is the same maximum value yielded by g_1 and g_K. Denote the value as the slowest convergence rate (SCR) of $\mathbf{R}$, i.e.,

$$\begin{aligned}\mathcal{M}(\mathbf{R}) &= |1 - g_1/\lambda| = |1 - g_K/\lambda| = \frac{g_1 - g_K}{g_1 + g_K} \\ &= \frac{\mathcal{S}(\mathbf{R}) - 1}{\mathcal{S}(\mathbf{R}) + 1}\end{aligned} \tag{34}$$

where $\mathcal{S}(\mathbf{R}) = g_1/g_K$ is the eigenvalue spread (EVS) of $\mathbf{R}$. It is obvious that a smaller EVS yields a smaller SCR. Furthermore, if the SCR is smaller, then the convergence of the series of (29) will be faster, and the truncation error will be smaller. However, a closed-form expression for the truncation error is difficult to obtain. Instead of the exact value of the error, we will try to derive an upper bound. Let the number of the terms retained in (29) be $\mathcal{L}$ and the power of the truncation error be $\mathcal{E}$. Then, we have

$$\begin{aligned}\mathcal{E} &= \left\| \sum_{k=0}^{\infty} \mathbf{A}^k - \sum_{k=0}^{\mathcal{L}-1} \mathbf{A}^k \right\| = \left\| \sum_{k=\mathcal{L}}^{\infty} \mathbf{A}^k \right\| \\ &\leq \sum_{k=\mathcal{L}}^{\infty} \|\mathbf{A}^k\| \leq \sum_{k=\mathcal{L}}^{\infty} \|\mathbf{A}\|^k \\ &\leq \sum_{k=\mathcal{L}}^{\infty} \mathcal{M}^k(\mathbf{R})\end{aligned} \tag{35}$$

where $\mathcal{M}(\mathbf{R})$ is the maximum diagonal value of $\mathbf{I} - (\mathbf{G}/\lambda)$ in (30). If $\mathcal{M}(\mathbf{R}) \neq 0$, then

$$\mathcal{E} \leq \frac{\mathcal{M}(\mathbf{R})^{\mathcal{L}}}{1 - \mathcal{M}(\mathbf{R})} = \frac{(\mathcal{S}(\mathbf{R}) - 1)^{\mathcal{L}}}{2\,(\mathcal{S}(\mathbf{R}) + 1)^{\mathcal{L}-1}}. \tag{36}$$

It is simple to see that when $\mathcal{S}(\mathbf{R}) = 1$ and $\mathcal{M}(\mathbf{R}) = 0$, $\mathcal{E} = 0$, giving the fastest convergence of (29). In this case, $\mathbf{R}$ is diagonal, and only one term is required in (29).

We now compare the EVSs of $\mathbf{R}_y$ and $\mathbf{R}_s$ in (10) and show that $\mathcal{S}(\mathbf{R}_s) < \mathcal{S}(\mathbf{R}_y)$. As shown in Section II, $\mathbf{R}_y = \sigma_x^2 \mathbf{U}\mathbf{U}^H + \sigma_\eta^2 \mathbf{I}$, and $\mathbf{R}_s = \sigma_x^{-2}\mathbf{I} + \sigma_\eta^{-2}\mathbf{U}^H\mathbf{U} = \sigma_x^{-2}\mathbf{I} + \sigma_\eta^{-2}\mathbf{R}_u$. Let $\{g_{u,1}, \ldots, g_{u,Q}\}$ be the eigenvalues of $\mathbf{U}\mathbf{U}^H$ and $g_{u,1} \geq \cdots \geq g_{u,Q}$. Since the rank of $\mathbf{U}$, which is a Q-by-M matrix, is M, the smallest $Q{-}M$ eigenvalues of $\mathbf{U}\mathbf{U}^H$ are zero, i.e., $g_{u,M+1} = \cdots = g_{u,Q} = 0$. In addition, the nonzero eigenvalues of $\mathbf{U}\mathbf{U}^H$ and $\mathbf{U}^H\mathbf{U}$ are the same. This indicates that we can obtain the eigenvalues of $\mathbf{R}_s$ from $\mathbf{R}_y$ as

$$\mathrm{eig}(\mathbf{R}_s) = \sigma_x^{-2} + \sigma_\eta^{-2}\mathrm{eig}(\mathbf{R}_u) \tag{37}$$

where $\mathrm{eig}(\mathbf{R}_u)$ denotes the first M eigenvalues of $\mathbf{R}_u$. Therefore, the EVSs of $\mathbf{R}_y$ and $\mathbf{R}_s$ can easily be obtained as

$$\begin{aligned}\mathcal{S}(\mathbf{R}_y) &= \left(\sigma_x^2 g_{u,1} + \sigma_\eta^2\right) / \left(\sigma_\eta^2\right) \\ &= \rho \cdot g_{u,1} + 1\end{aligned} \tag{38}$$

$$\begin{aligned}\mathcal{S}(\mathbf{R}_s) &= \left(\sigma_x^{-2} + \sigma_\eta^{-2} g_{u,1}\right) / \left(\sigma_x^{-2} + \sigma_\eta^{-2} g_{u,M}\right) \\ &= (\rho \cdot g_{u,1} + 1)/(\rho \cdot g_{u,M} + 1)\end{aligned} \tag{39}$$

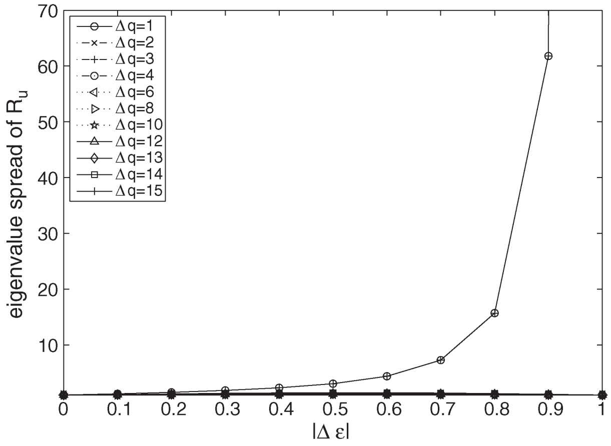

Fig. 1. EVS of $\mathbf{R}_u$ for $M_{\max} = 2$ $(1 \le \Delta q \le 15)$.

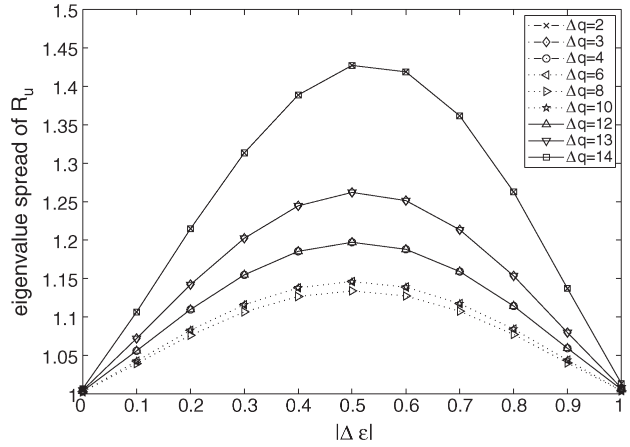

Fig. 2. EVS of $\mathbf{R}_u$ for $M_{\max} = 2$ $(2 \le \Delta q \le 14)$.

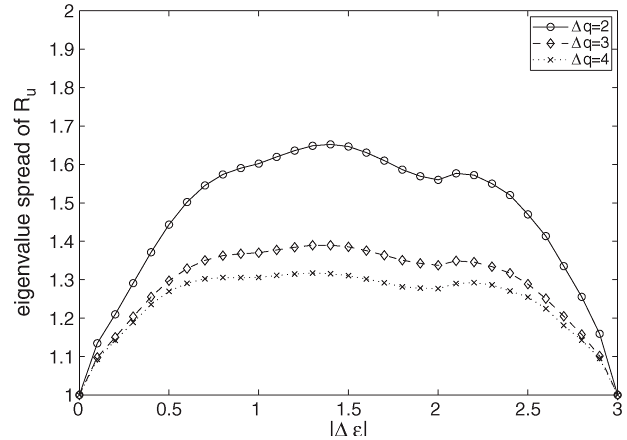

Fig. 3. EVS of $\mathbf{R}_u$ for $M_{\max} = 4$ $(2 \le \Delta q \le 4)$.

where $\rho = \sigma_x^2/\sigma_\eta^2$. Note that the received SNR is defined as $\mathrm{SNR} = M\sigma_x^2/\sigma_\eta^2 = M\rho$. It is easy to see that the EVS of $\mathbf{R}_s$ is smaller than that of $\mathbf{R}_y$. Therefore, the SCR of $\mathbf{R}_s$ is smaller than that of $\mathbf{R}_y$. Thus, the matrix inversion lemma used in (10) not only reduces the computational complexity but also reduces the truncation error in (14).

From (39), it is also simple to see that for low SNR, the EVS of $\mathbf{R}_s$ approaches 1, and therefore, the truncation error in (14) can be ignored. For high SNR, the EVS of $\mathbf{R}_s$ approaches that of $\mathbf{R}_u$. The EVS of $\mathbf{R}_u = \mathbf{U}^H\mathbf{U}$ depends on the subchannel assignment since the (a, b)th entry of $\mathbf{U}$ is $w^{\{\varepsilon_{e,b}\cdot(a-1)N\}}$ and $\varepsilon_{e,b} = \varepsilon_b + q_b$. To analyze the truncation error, we first have to analyze the EVS of $\mathbf{U}^H\mathbf{U}$. Unfortunately, a general closed-form for the EVS is difficult to obtain. Here, we study two special cases to show that the EVS of $\mathbf{U}^H\mathbf{U}$ is low and that the truncation error in (14) can be small in our applications.

Define a $M_{\max}$-user system as a system that can simultaneously handle $M_{\max}$ users at most. The first case we consider is a two-user system. For the system, the EVS of $\mathbf{U}^H\mathbf{U}$ can be solved from (3) in closed form as

$$\mathcal{S}(\mathbf{U}^H\mathbf{U}) = \frac{Q + \vartheta_{1,2}}{Q - \vartheta_{1,2}} \tag{40}$$

where

$$\vartheta_{1,2} = \{[1 - \cos(2\pi\delta_\varepsilon) - \cos(2\pi\delta_\varepsilon/Q) + \cos(2\pi\delta_\varepsilon(Q-1)/Q)]^{0.5} \cdot [1 - \cos(2\pi\delta_\varepsilon/Q)]^{-0.5}.$$

As we can see, the EVS varies with $\delta_\varepsilon = |\varepsilon_{e,1} - \varepsilon_{e,2}| = |q_1 - q_2 + \varepsilon_1 - \varepsilon_2|$. Therefore, $\mathcal{S}(\mathbf{U}^H\mathbf{U})$ is dependent on $\Delta\varepsilon = |\varepsilon_1 - \varepsilon_2|$, and $\Delta q = |q_1 - q_2|$. Note that Δq is the difference of the neighbor subchannel indices. We now use an example to examine the EVS of $(\mathbf{U}^H\mathbf{U})$. Let $N_s = 128$ and $N = 8$. Then, there are $N_s/N = 16$ subchannels, and $1 \le \Delta q \le 15$. Fig. 1 shows the result. From the figure, we can see that the EVSs for $\Delta q = 1$ and $\Delta q = 15$ are much larger than those for $2 \le \Delta q \le 14$. Fig. 2 shows the EVSs for $2 \le \Delta q \le 14$. It is clear that all the values are smaller than 1.5. This indicates that the truncation error will be small as long as the adjacent subchannels are not used simultaneously, i.e., $|\bmod(\Delta q - i \cdot Q, Q)| \ne 1$, where i is an integer.

The second case we consider is a four-user system. For the system, the closed-form solution of the EVS is not obtainable. Simulations are then conducted to obtain numerical results. Note that in this case the EVS is a function of $\{q_1, \ldots, q_4, \varepsilon_1, \ldots, \varepsilon_4\}$. It is difficult to examine the behavior of the EVS in terms of these variables. For convenience, we still define two variables: 1) $\Delta q = |q_1 - q_2| = |q_2 - q_3| = |q_3 - q_4|$ and 2) $\Delta\varepsilon = \sum_{i=1}^{M-1} |\varepsilon_{i+1} - \varepsilon_i|$. Note that the definition and the implication of $\Delta\varepsilon$ are different from those in the two-user case. Using the same simulation setting as that in the two-user system, we obtain the EVS versus $\Delta\varepsilon$ in Fig. 3. Here, we assume that adjacent subchannels are not used. From the figure, we can find that all the EVSs are smaller than 1.7. We can expect that the truncation error in (14) will be small. From Figs. 2 and 3, we can also find that the smallest truncation error can be obtained when $\Delta q = 8$ and $\Delta q = 4$ for the two-user and four-user systems, respectively. However, a large Δq will result in a smaller $M_{\max}$, which is the maximum number of users. Thus, the selection of Δq is a tradeoff between $M_{\max}$ and the SCR of $\mathbf{R}_u$. In this paper, we let the smallest Δq be 2.

Using (33) and (37), we can see that the optimum λ in (11) is equal to $\sigma_x^{-2} + \sigma_\eta^{-2}(g_{u,1} + g_{u,M})/2$. Note that the eigenvalues

TABLE I
NORMALIZED TRUNCATION ERROR VERSUS SNR

SNR	0dB	3dB	6dB	9dB	20dB	30dB
M= 2, $\Delta q = 2$ (dB)	-26.6867	-26.2986	-26.0997	-25.9989	-25.9049	-25.8976
Upper bound in (42)	-25.2057	-24.7714	-24.5481	-24.4348	-24.3291	-24.3209
M= 2, $\Delta q = 3$ (dB)	-32.4019	-32.0139	-31.8150	-31.7141	-31.6202	-31.6128
Upper bound in (42)	-31.5097	-31.0943	-30.8811	-30.7729	-30.6720	-30.6642
M= 2, $\Delta q = 4$ (dB)	-35.8326	-35.4446	-35.2457	-35.1449	-35.0509	-35.0436
Upper bound in (42)	-35.1642	-34.7558	-34.5462	-34.4399	-34.3408	-34.3331
M= 4, $\Delta q = 2$ (dB)	-20.0507	-19.6616	-19.4621	-19.3610	-19.2667	-19.2594
Upper bound in (42)	-18.0176	-17.5638	-17.3303	-17.2117	-17.1010	-17.0924
M= 4, $\Delta q = 3$ (dB)	-25.1858	-24.7970	-24.5978	-24.4967	-24.4026	-24.3952
Upper bound in (42)	-23.8575	-23.4277	-23.2068	-23.0947	-22.9901	-22.9820
M= 4, $\Delta q = 4$ (dB)	-27.2568	-26.8681	-26.6689	-26.5679	-26.4737	-26.4664
Upper bound in (42)	-26.1414	-25.7185	-25.5013	-25.3910	-25.2882	-25.2802

of $\mathbf{R}_u$ can be estimated as $\text{eig}(\sigma_x^{-2}(\mathbf{R}_y - \sigma_\eta^2\mathbf{I}))$, and $\mathbf{R}_y$ can be estimated as $\sum_{n=0}^{N-1}\mathbf{y}(n)\mathbf{y}^H(n)$. Therefore, the optimum λ can then be calculated. To evaluate the performance of the proposed expansion, we define a normalized truncation error as (with the optimum λ)

$$\mathcal{E}_n = E_{\boldsymbol{\varepsilon}}\left\{\left\|\mathbf{R}_s/\lambda \cdot \left[(\mathbf{R}_s/\lambda)^{-1} - \sum_{k=0}^{2}(-1)^k\mathbf{B}^k\right]\right\|\right\} \tag{41}$$

where $\boldsymbol{\varepsilon}$ is the set for all possible $\varepsilon_{e,m}$ $(m = 1, \ldots, M)$. Using the result in (36), we can also define a normalized upper bound as

$$\begin{aligned}\mathcal{E}_n &\le E_{\boldsymbol{\varepsilon}}\left\{\|\mathbf{R}_s/\lambda\| \cdot \left\|(\mathbf{R}_s/\lambda)^{-1} - \sum_{k=0}^{2}(-1)^k\mathbf{B}^k\right\|\right\}\\ &\le E_{\boldsymbol{\varepsilon}}\left\{\|\mathbf{R}_s/\lambda\|\frac{\left(\mathcal{S}(\mathbf{R}_s/\lambda)-1\right)^3}{2\left(\mathcal{S}(\mathbf{R}_s/\lambda)+1\right)^2}\right\}. \end{aligned}\tag{42}$$

We now use some examples to evaluate the normalized truncation error and its upper bound in (42). The result is shown in Table I. It is simple to see that the normalized truncation error increases with the decrease of Δq and with the increase of SNR. In addition, the deviation of the upper bound from the actual error is small; the upper bound overestimates the error by 1–2 dB. From Table I, we can also see that even with $\Delta q = 2$, the truncation error is still quite small, i.e., -20 dB. Section IV gives more results to show the property.

B. CRB Analysis

For the training-based method, only the AWGN is considered as a random variable, and the CRB for CFO estimation can then be derived [4]. In the blind method, the transmit symbol is treated as an additional random variable, and the CRB for blind CFO estimation can also be derived [2], [11]. Here, we generalize the result in [2] and [11] (for single-user OFDM systems) to derive the CRB in OFDMA systems. From (4), the (p,q)th entry of the Fisher information matrix $\mathbf{F}$ is given by

$$(\mathbf{F})_{p,q} = -E\left\{\frac{\partial^2 \ln(\Lambda(\boldsymbol{\varepsilon}))}{\partial\varepsilon_p\partial\varepsilon_q}\right\} = -E\left\{\frac{\partial^2 \ln(\Lambda(\boldsymbol{\varepsilon}))}{\partial\varepsilon_{e,p}\partial\varepsilon_{e,q}}\right\} \tag{43}$$

where $1 \le p, q \le M$. Substituting (8) into (43) yields

$$\begin{aligned}(\mathbf{F})_{p,q} =& N \cdot \text{tr}\left(-\mathbf{R}_y^{-1}\frac{\partial\mathbf{R}_y}{\partial\varepsilon_p}\mathbf{R}_y^{-1}\frac{\partial\mathbf{R}_y}{\partial\varepsilon_q} + \mathbf{R}_y^{-1}\frac{\partial^2\mathbf{R}_y}{\partial\varepsilon_p\partial\varepsilon_q}\right)\\ &- E\left\{\sum_{n=0}^{N-1}\mathbf{y}^H(n)\mathbf{J}\mathbf{y}(n)\right\}\\ =& N \cdot \text{tr}\left(\mathbf{R}_y^{-1}\frac{\partial\mathbf{R}_y}{\partial\varepsilon_p}\mathbf{R}_y^{-1}\frac{\partial\mathbf{R}_y}{\partial\varepsilon_q}\right)\end{aligned}\tag{44}$$

where

$$\begin{aligned}\mathbf{J} =& -\mathbf{R}_y^{-1}\frac{\partial\mathbf{R}_y}{\partial\varepsilon_p}\mathbf{R}_y^{-1}\frac{\partial\mathbf{R}_y}{\partial\varepsilon_q}\mathbf{R}_y^{-1}\\ &+ \mathbf{R}_y^{-1}\left(\frac{\partial^2\mathbf{R}_y}{\partial\varepsilon_p\partial\varepsilon_q}\right)\mathbf{R}_y^{-1}\\ &- \mathbf{R}_y^{-1}\frac{\partial\mathbf{R}_y}{\partial\varepsilon_p}\mathbf{R}_y^{-1}\frac{\partial\mathbf{R}_y}{\partial\varepsilon_q}\mathbf{R}_y^{-1}.\end{aligned}\tag{45}$$

From [16], we see that

$$\begin{aligned}E\left\{\sum_{n=0}^{N-1}\mathbf{y}^H(n)\mathbf{J}\mathbf{y}(n)\right\} &= \sum_{n=0}^{N-1}E\left\{\mathbf{y}^H(n)\mathbf{J}\mathbf{y}(n)\right\}\\ &= \sum_{n=0}^{N-1}E\left\{\text{tr}\left(\mathbf{J}\mathbf{y}(n)\mathbf{y}^H(n)\right)\right\}\\ &= N \cdot \text{tr}(\mathbf{J}\mathbf{R}_y).\end{aligned}\tag{46}$$

Finally, the CRB for the ε_i estimate is obtained as

$$CRB(\varepsilon_i) = (\mathbf{F}^{-1})_{i,i}. \tag{47}$$

We average the diagonal terms of (47) to have a single index for performance comparison.

C. Computational Complexity

Here, the computational complexity of the proposed method is assessed and compared with that of the existing schemes. For the proposed method, there are three operation steps. In the first step, we need to calculate the autocorrelation matrix in (5) and its eigen-decomposition to obtain (37). The computational complexity for this step is $O(Q^3 + Q^2N)$. In the second step, we need to calculate the coefficients in (27). From the received signal and the known autocorrelation matrix, we can obtain $\gamma(p, q)$ in (23) and $\Gamma(p, q)$ in (19). Let $\Gamma(p, q)$ and $\gamma(p, q)$ be the (p, q)th entries of two matrices $\mathbf{\Gamma}$ and $\boldsymbol{\gamma}$, respectively. The computational complexity for constructing these matrices is $O(Q^2N)$. Note that not all the coefficients in (27) (see (57) and (58) in Appendix II) are required for the calculation. Some coefficient pairs in (57) and in (58) are complex conjugates of each other. In addition, some terms in (57) and (58) appear repeatedly. For example, $F_1(p, q)$ in (57), the sixth term of (57), and the seventh term of (57) all include $\sum_m \Gamma(p, m)\Gamma(m, n)$. Without the redundant computations, the complexity for calculating the coefficients in (57) or (58) is found to be $O(6Q^3)$. Then, we evaluated the computational complexity for calculating all the coefficients in the polynomial (27). The last step is the root-searching process in (27) and the CFOs sorting for each user in (26). Since there are $2Q - 1$ terms in (27), the roots can be solved with the complexity of $O(8Q^3)$ [2]. Compared with the root-searching process, the complexity in calculating (26) is small and can be ignored. Adding all together and taking only dominant terms, we can have that the entire complexity for the proposed method is $O(15Q^3 + 2Q^2N)$. For the ESPRIT frequency estimator, the complexity is shown to be $O(5Q^3 + Q^2N)$ [10]. Therefore, the computational complexity of the proposed method is on the same complexity order as that of ESPRIT.

Next, we evaluate the computational complexity of the training-based schemes. For APFE, the total complexity has been shown to be $O(MN_cN_w(L^3 + LN_s^2))$ [3], [4], where N_c denotes the number of iterations, and N_w denotes the number of grid points used for each iteration. For simplified AAPFE, the computational complexity is $O(MN_cN_wN_s^2K)$ [4], [5]. Since the computational complexity of the APFE algorithm is high, suboptimum training-based schemes are then proposed [3], [5], [6]. For the method in [3], the computational complexity is $O(2MTN_s^2 + TN_s(MN)^2)$, where T is the number of the Monte Carlo runs finding a mean likelihood [13]. The computational complexity for the method in [5] is $O(N_c(M^2N^2 + 1/2N_s \log_2(N_s) + 3/2M^3N^3 + 3/8M^2N^2N_s))$, whereas that in [6] is $O(N_c(N_s(MN)^2 + M^3))$ for $\mathcal{L} = 1$. Here, $\mathcal{L}$ is the number of terms retained in an infinite series [6]. Note that for $\mathcal{L} = 1$, the method has the worse performance but the lowest computational complexity. In addition, note that the simulation results in [6] indicate that $\mathcal{L}$ should be at least 3 for acceptable performance. However, as shown in [6], the computational complexity order for $\mathcal{L} = 3$ is difficult to evaluate.

IV. Simulations

In this section, we report simulation results demonstrating the effectiveness of the proposed method. In the first set of simulations, we compare the performance of the proposed method with existing blind methods. Note that ESPRIT is known to be better than the MUSIC algorithm [10]. Thus, we only conduct simulations for ESPRIT [10]. In the second set of simulations, we compare the performance of the proposed method with existing ML methods, such as APFE and AAPFE. Note that existing ML schemes require training sequences. Finally, we compare the computational complexity of all schemes.

A. System Setup

In our simulations, the channel response used for each user is generated according to the HIPERLAN/2 channel model [12]. The channel response, having six taps, follows an exponential power decay profile. Each tap coefficient is modeled as an independent complex Gaussian random variable with zero mean. The CFO of each user is generated with a uniform distribution in the interval (−0.5, 0.5). The symbols used for CFO estimation are modulated with a binary phase-shift-keying scheme, whereas those for data transmission are modulated with a 16-QAM scheme. The interleaved OFDMA system used in our simulations has $N_s = 128$ subcarriers. Since there are multiple CFOs to be estimated, the mean square error (MSE) is used as the performance index, which is defined as

$$MSE = \sum_{m=1}^{M} (\hat{\varepsilon}_m - \varepsilon_m)^2. \tag{48}$$

All the simulation results are obtained by averaging 1000 Monte Carlo runs.

As shown in Section III, Δq, which is the difference of the neighbor subchannel indices, influences the truncation error in (14) and $M_{\max}$, which is the maximum number of users. As shown in Section III, the larger the Δq, the smaller the truncation error and the smaller the $M_{\max}$. We compare the results for $\Delta q = 2$ and $\Delta q = 4$ in the following.

B. Performance Assessment for $\Delta q = 2$

First, we let $\Delta q = 2$, the smallest Δq we use. In this case, the SCR will be maximal, and $M_{\max}$ is also maximal. It corresponds to the worst case in the proposed method. An important design parameter for the proposed method is the number of subcarriers for each user N. The number of subchannels is then $Q = N_s/N$. Observing (27) and (47), we see that a higher Q performs better but requires higher complexity. To see the impact of N, we let $M = 4$ and observe the CRB for different N. The result is shown in Fig. 4. We can see that the CRB is almost the same for $N = 4$ and $N = 8$. To reduce computations, we choose $N = 8$ for the simulations we conducted. Without loss of generality, we assume that the CP length is N, which is larger than the channel length. For the first set of simulations, we compare the performance of the proposed algorithm with that of ESPRIT. Fig. 5 shows the result for $M = 2$, whereas Fig. 6 shows the result for $M = 4$. As expected, the performance of the proposed algorithm is significantly better than that of ESPRIT since the proposed method conducts the ML estimation. We can also see that the proposed method can approach the CRB. At high SNR regions, the performance of the proposed algorithm slightly deviates from the CRB. This is due to our approximation used in (14). When the number of users is larger, the deviation is also larger.

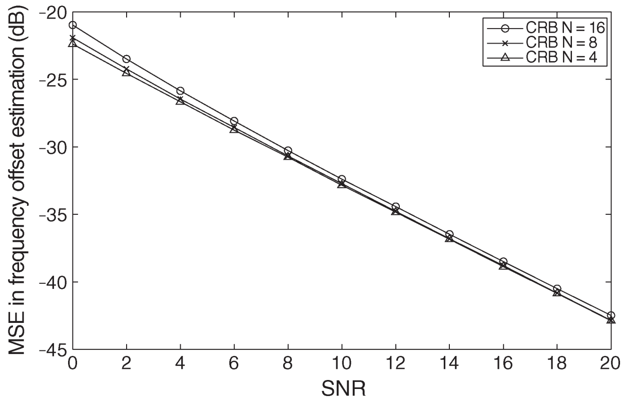

Fig. 4. CRB comparison for various N.

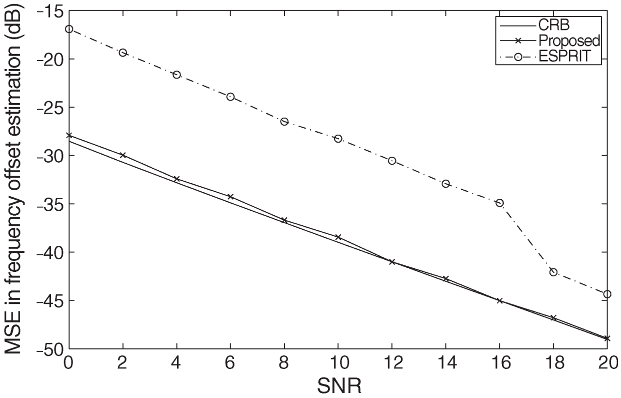

Fig. 5. Performance comparison for ESPRIT and the proposed algorithm ($M = 2$).

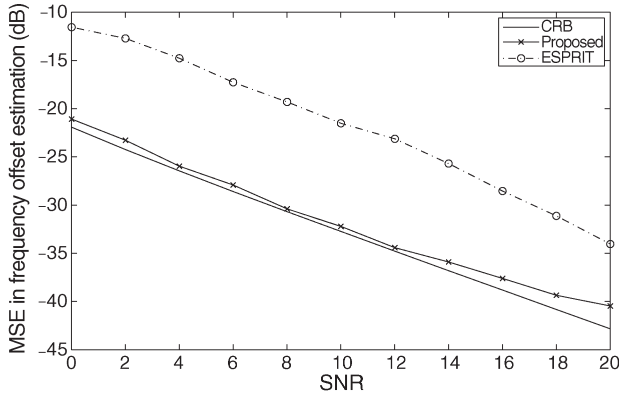

Fig. 6. Performance comparison for ESPRIT and the proposed algorithm ($M = 4$).

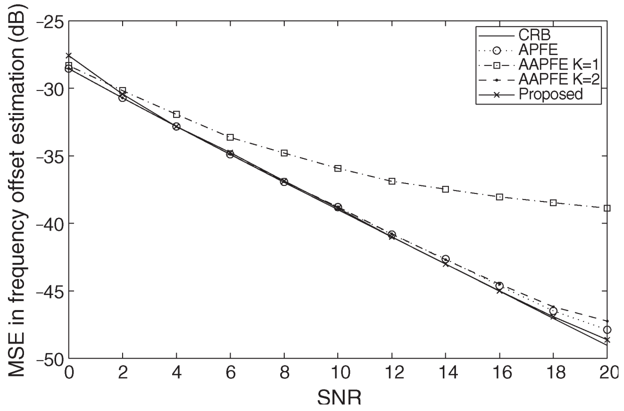

Fig. 7. Performance comparison for training-based and the proposed algorithm ($M = 2$).

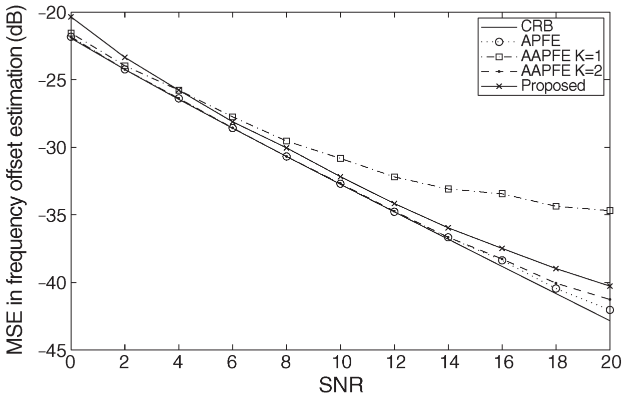

Fig. 8. Performance comparison for training-based and the proposed algorithm ($M = 4$).

In the second set of simulations, we compare the performance of the proposed algorithm with that of other ML algorithms. The simulation setup is the same as that of the first set of simulations. The ML problem can directly be solved by using an exhaustive grid search over the MD space spanning $\{\varepsilon_{e,1}, \ldots, \varepsilon_{e,M}\}$. To reduce the computational complexity, we use the APFE and AAPFE schemes. The AAPFE is a suboptimum solution of APFE, and it also truncates the Neumann series to approximate an inverse matrix (with an order K) [21]. In each iteration, only one user's CFO is updated, whereas the other users' CFOs remain unchanged. The CFO update is conducted by a grid-search method. Note that the purpose of the expansion is different from ours. In addition, AAPFE does not use the optimum λ to achieve the best result. In our simulations, we let $N_c = 2$ and $N_w = 100$. We have tried $K = 1$ and $K = 2$. Fig. 7 shows the result for $M = 2$, and Fig. 8 shows the result for $M = 4$. We can see that the APFE performs the best, and the AAPFE with $K = 1$ performs the worst. Note that the conventional APFE and AAPFE have to use training sequences. From the figure, we see that the performance gaps between the APFE, the AAPFE with $K = 2$, and the proposed blind algorithm are very small. In addition, note that all these algorithms tend to deviate from the CRB when the SNR is high.

C. Performance Assessment for $\Delta q = 4$

Fig. 9 shows the performance comparison between $\Delta q = 2$ and $\Delta q = 4$. The simulation setup is the same as that in the previous section. It is easy to find that the performance of the proposed method with $\Delta q = 4$ is better. In Section III, we have shown the relationship between the SCR and the truncation error (as shown in Table I). From Fig. 9, we can observe a similar result. The smaller the SCR, the better the estimation performance. The performance gap between the APFE and the proposed algorithm becomes smaller when $\Delta q = 4$.

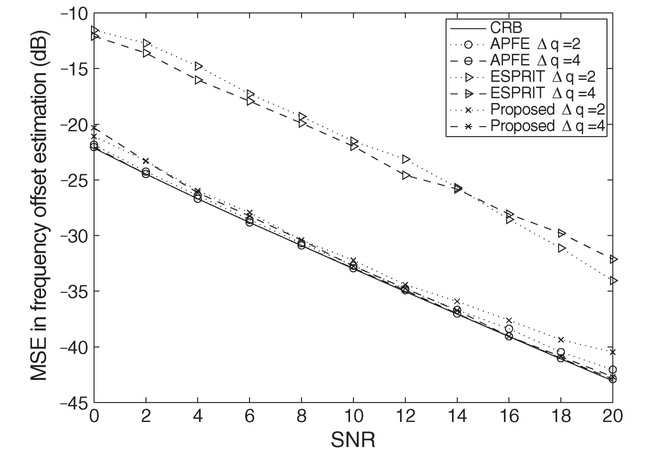

Fig. 9. Performance comparison for all algorithms with $\Delta q = 2$ and $\Delta q = 4$ $(M = 4)$.

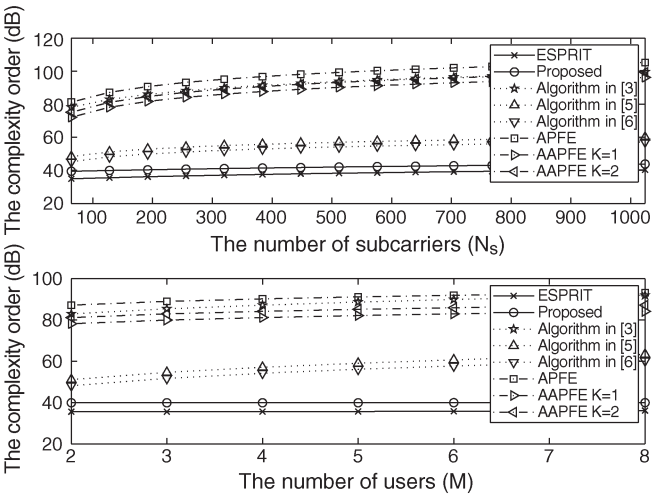

Fig. 10. Computational complexity comparison for the proposed algorithm and algorithms in [3]–[6] and [10].

D. Complexity Comparison

Fig. 10 shows the computational complexity of the schemes we consider. Fig. 10(a) shows the complexity versus the number of subcarriers for the two-user case. Fig. 10(b) shows the complexity versus the number of users for a fixed total number of subcarriers: $N_S = 128$, and $Q = 8$. From the figures, we find that the computational complexity of the proposed method is similar to that of the conventional blind methods. However, the proposed method outperforms the conventional methods by 10 dB (see Figs. 5 and 6). Compared with the training-based methods, the proposed blind method can have similar performance (see Figs. 7 and 8) but much lower computational complexity (see Fig. 10).

V. Conclusion

In this paper, we have developed a new algorithm for blind ML CFO estimation in interleaved OFDMA uplink systems. Conventional methods for this problem require conducting MD exhausted searches, and the computational complexity can be very high. The distinct feature of the proposed algorithm is that it only requires a root-searching procedure. The main idea is to use a series expansion when evaluating the ML function. The performance of the expansion is also analyzed. The operations of the proposed method are simple, and the computational complexity is low. Simulations show that the proposed method can approach the CRB. As shown in Fig. 1, a large EVS will be induced in the full-loaded scenario $(\Delta q = 1)$, and the performance of the proposed method will seriously be affected. The problem can be solved by an expectation-maximization (EM) algorithm referred to as iterative space alternating generalized EM (SAGE) [23], [24]. However, the complexity of the SAGE algorithm can be very high for large N_s. Note that in real-world applications, only a number of users will be activated at a specific time [24]. Thus, only the CFOs of the newly activated users have to be estimated, and the knowledge of the previously estimated CFOs can be exploited in each new estimation. It is interesting to incorporate the SAGE algorithm into the proposed method, which may serve as a topic for further research.

Appendix I
Derivation of (22)

Taking the derivative of $(\mathbf{R}_y)_{p,q}$ with respect to $\varepsilon_{e,i}$, we have the result as

$$\left(\frac{\partial}{\partial \varepsilon_{e,i}}\mathbf{R}_y\right)_{p,q} = \frac{j2\pi N\sigma_x^2}{N_s}(p-q)w^{(\varepsilon_{e,i})N(p-q)}. \tag{49}$$

Then, the first term in the right-hand side of (9) can be derived as

$$\begin{aligned}
&\left(\mathbf{R}_y^{-1}\frac{\partial}{\partial \varepsilon_{e,i}}\mathbf{R}_y\right)_{p,q}\\
&\quad= \sum_{k=1}^{Q}(\mathbf{R}_y^{-1})_{p,k}\frac{j2\pi N\sigma_x^2}{N_s}(k-q)w^{(\varepsilon_{e,i})N(k-q)}\\
&\quad= \frac{j2\pi N\sigma_x^2}{N_s}\Bigg\{\sigma_\eta^{-2}w^{(\varepsilon_{e,i})N(p-q)}\\
&\qquad\qquad - C_0\sum_{k=1}^{Q}(k-q)x^{k-q}\Gamma(p,k)\\
&\qquad\qquad - C_1\sum_{k=1}^{Q}\sum_{b=1}^{Q}(k-q)x^{k-q}\Gamma(p,b)\Gamma(b,k)\\
&\qquad\qquad - C_2\sum_{k=1}^{Q}(k-q)x^{k-q}\\
&\qquad\qquad \cdot\sum_{a=1}^{Q}\sum_{b=1}^{Q}\Gamma(p,a)\Gamma(a,b)\Gamma(b,k)\Bigg\}
\end{aligned} \tag{50}$$

where

$$x = \exp\left(j2\pi\frac{N\varepsilon_{e,i}}{N_s}\right). \tag{51}$$

To obtain the second term in (9), the derivative of the third and fourth terms in the right-hand side of (17) can first be found as

$$\begin{aligned}\frac{\partial}{\partial\varepsilon_{e,i}}\left\{\sum_{n=1}^{Q}\Gamma(n,q)\Gamma(p,n)\right\} &= \frac{j2\pi N}{N_s}\left[\sum_{n=1}^{Q}(n-q)\Gamma(p,n)w^{(\varepsilon_{e,i})N(n-q)}\right.\\ &\qquad\left.+\,(p-n)\Gamma(n,q)w^{(\varepsilon_{e,i})N(p-n)}\right]\end{aligned} \tag{52}$$

$$\begin{aligned}\frac{\partial}{\partial\varepsilon_{e,i}}\left\{\sum_{m=1}^{Q}\sum_{n=1}^{Q}\Gamma(m,n)\Gamma(n,q)\Gamma(p,m)\right\} &= \frac{j2\pi N}{N_s}\left[\sum_{m=1}^{Q}\Gamma(p,m)\sum_{n=1}^{Q}\Gamma(n,q)\cdot(m-n)w^{(\varepsilon_{e,i})N(m-n)}\right.\\ &\qquad+(n-q)\Gamma(p,m)\Gamma(m,n)w^{(\varepsilon_{e,i})N(n-q)}\\ &\qquad\left.+(p-m)\Gamma(m,n)\Gamma(n,q)w^{(\varepsilon_{e,i})N(p-m)}\right].\end{aligned} \tag{53}$$

Thus, we can have the second term in (9) as (54), shown at the bottom of the page.

Substituting (3), (5), (50), and (54) into (9), we can rewrite the log-likelihood function as (55), shown at the bottom of the page.

Appendix II
Derivation of (27)

Rewrite (22) as (56), shown at the top of the next page, where $F_1(p,q) = N\sigma_x^2(q-p)[C_0\Gamma(p,q) + C_1\sum_{n=1}^{Q}\Gamma(p,n)\Gamma(n,q) + C_2\sum_{m=1}^{Q}\sum_{n=1}^{Q}\Gamma(p,m)\Gamma(m,n)\Gamma(n,q)]$, and $F_2(p,1) = C_0(p-q)\gamma(p,q)$. Conducting some variable transformation, we can express the power of x as a single variable of k. We can then collect all the items with positive k into one expression and get (57), shown at the bottom of the next page. Similarly, we can collect all the items with negative k into another expression and get (58), shown at the top of the page after the next page.

Therefore, the derivative of the logarithm likelihood function (22) can then be reexpressed as

$$\frac{\partial}{\partial\varepsilon_{e,i}}\Lambda(\boldsymbol{\varepsilon}) = \sum_{k=1}^{Q-1}\alpha_p(k)x^{k} + \sum_{k=1}^{Q-1}\alpha_n(k)x^{-k}. \tag{59}$$

$$\begin{aligned}\frac{\partial}{\partial\varepsilon_{e,i}}\left(\mathbf{R}_y^{-1}\right)_{p,q} = \frac{j2\pi N}{N_s}\Bigg\{ &-C_0(p-q)w^{(\varepsilon_{e,i})N(p-q)}\\ &-C_1\sum_{n=1}^{Q}(n-q)w^{(\varepsilon_{e,i})N(n-q)}\Gamma(p,n)\\ &-C_1\sum_{n=1}^{Q}(p-n)w^{(\varepsilon_{e,i})N(p-n)}\Gamma(n,q)\\ &-C_2\sum_{m=1}^{Q}\sum_{n=1}^{Q}\Big[(m-n)w^{(\varepsilon_{e,i})N(m-n)}\Gamma(p,m)\Gamma(n,q)\\ &\qquad+(n-q)w^{(\varepsilon_{e,i})N(n-q)}\Gamma(p,m)\Gamma(m,n)\\ &\qquad+(p-m)w^{(\varepsilon_{e,i})N(p-m)}\Gamma(m,n)\Gamma(n,q)\Big]\Bigg\}\end{aligned} \tag{54}$$

$$\begin{aligned}\frac{\partial}{\partial\varepsilon_{e,i}}\Lambda(\boldsymbol{\varepsilon}) = \frac{j2\pi N}{N_s}\Bigg\{&N\sigma_x^2\sum_{p=1}^{Q}\sum_{q=1}^{Q}(q-p)x^{q-p}\left[C_0\Gamma(p,q) + C_1\sum_{n=1}^{Q}\Gamma(p,n)\Gamma(n,q) + C_2\sum_{m=1}^{Q}\sum_{n=1}^{Q}\Gamma(p,m)\Gamma(m,n)\Gamma(n,q)\right]\\ &+\gamma(p,q)\Bigg[C_0(p-q)x^{p-q} + C_1\sum_{n=1}^{Q}\big((n-q)x^{n-q}\Gamma(p,n) + (p-n)x^{p-n}\Gamma(n,q)\big)\\ &\qquad+C_2\sum_{m=1}^{Q}\sum_{n=1}^{Q}\big((m-n)x^{m-n}\Gamma(p,m)\Gamma(n,q) + (n-q)x^{n-q}\Gamma(p,m)\Gamma(m,n)\\ &\qquad\qquad+(p-m)x^{p-m}\Gamma(m,n)\Gamma(n,q)\big)\Bigg]\Bigg\}\end{aligned} \tag{55}$$

$$
\begin{aligned}
\frac{\partial}{\partial \varepsilon_{e,i}}\Lambda(\boldsymbol{\varepsilon}) = \sum_{p=1}^{Q}\sum_{q=1}^{Q}\Bigg\{ & N\sigma_x^2 (q-p)x^{q-p}\Bigg[C_0\Gamma(p,q) + C_1\sum_{n=1}^{Q}\Gamma(p,n)\Gamma(n,q) + C_2\sum_{m=1}^{Q}\sum_{n=1}^{Q}\Gamma(p,m)\Gamma(m,n)\Gamma(n,q)\Bigg] \\
& + \gamma(p,q)\Bigg[C_0(p-q)x^{p-q} + C_1\sum_{n=1}^{Q}\big((n-q)x^{n-q}\Gamma(p,n) + (p-n)x^{p-n}\Gamma(n,q)\big) \\
& + C_2\sum_{m=1}^{Q}\sum_{n=1}^{Q}\big((m-n)x^{m-n}\Gamma(p,m)\Gamma(n,q) + (n-q)x^{n-q}\Gamma(p,m)\Gamma(m,n) \\
& + (p-m)x^{p-m}\Gamma(m,n)\Gamma(n,q)\big)\Bigg]\Bigg\} \\
= & \sum_{p=1}^{Q}\sum_{q=1}^{Q}F_1(p,q)x^{q-p} + \sum_{p=1}^{Q}\sum_{q=1}^{Q}F_2(p,q)x^{p-q} \\
& + \sum_{p=1}^{Q}\sum_{q=1}^{Q}\Bigg\{\gamma(p,q)\Bigg[C_1\sum_{n=1}^{Q}\big((n-q)x^{n-q}\Gamma(p,n) + (p-n)x^{p-n}\Gamma(n,q)\big) \\
& + C_2\sum_{m=1}^{Q}\sum_{n=1}^{Q}((m-n)x^{m-n}\Gamma(p,m)\Gamma(n,q) + (n-q)x^{n-q}\Gamma(p,m)\Gamma(m,n) \\
& + (p-m)x^{p-m}\Gamma(m,n)\Gamma(n,q))\Bigg]\Bigg\} = 0
\end{aligned} \tag{56}
$$

$$
\begin{aligned}
\left\{\frac{\partial}{\partial \varepsilon_{e,i}}\Lambda(\boldsymbol{\varepsilon})\right\}_{+} = \sum_{k=1}^{Q-1}\Bigg\{ & \sum_{p=1,q=p+k}^{Q-k}F_1(p,q)x^k + \sum_{q=1,p=q+k}^{Q-k}F_2(p,q)x^k + C_1\sum_{p=1}^{Q}\sum_{q=1,n=q+k}^{Q-k}\gamma(p,q)\Gamma(p,n)kx^k \\
& + C_1\sum_{q=1}^{Q}\sum_{n=1,p=n+k}^{Q-k}\gamma(p,q)\Gamma(n,q)kx^k + C_2\sum_{p=1}^{Q}\sum_{q=1}^{Q}\sum_{n=1,m=n+k}^{Q-k}\gamma(p,q)\Gamma(p,m)\Gamma(n,q)kx^k \\
& + C_2\sum_{p=1}^{Q}\sum_{m=1}^{Q}\sum_{q=1,n=q+k}^{Q-k}\gamma(p,q)\Gamma(p,m)\Gamma(m,n)kx^k \\
& + C_2\sum_{n=1}^{Q}\sum_{q=1}^{Q}\sum_{m=1,p=m+k}^{Q-k}\gamma(p,q)\Gamma(m,n)\Gamma(n,q)kx^k\Bigg\} \\
= & \sum_{k=1}^{Q-1}\alpha_p(k)x^k
\end{aligned} \tag{57}
$$

$$
\begin{aligned}
\left\{\frac{\partial}{\partial \varepsilon_{e,i}}\Lambda(\boldsymbol{\varepsilon})\right\}_{-} &= \sum_{k=-1}^{1-Q}\Bigg\{ \sum_{p=1-k,q=p+k}^{Q} F_1(p,q)x^k + \sum_{q=1-k,p=q+k}^{Q} F_2(p,q)x^k + C_1\sum_{p=1}^{Q}\sum_{q=1-k,n=q+k}^{Q}\gamma(p,q)\Gamma(p,n)kx^k \\
&\qquad + C_1 \sum_{q=1}^{Q}\sum_{n=1-k,p=n+k}^{Q} \gamma(p,q)\Gamma(n,q)kx^k + C_2\sum_{p=1}^{Q}\sum_{q=1}^{Q}\sum_{n=1-k,m=n+k}^{Q} \gamma(p,q)\Gamma(p,m)\Gamma(n,q)kx^k \\
&\qquad + C_2 \sum_{p=1}^{Q}\sum_{m=1}^{Q}\sum_{q=1-k,n=q+k}^{Q} \gamma(p,q)\Gamma(p,m)\Gamma(m,n)kx^k \\
&\qquad + C_2 \sum_{n=1}^{Q}\sum_{q=1}^{Q}\sum_{m=1-k,p=m+k}^{Q} \gamma(p,q)\Gamma(m,n)\Gamma(n,q)kx^k\Bigg\} \\
&= \sum_{k=1}^{Q-1}\Bigg\{\sum_{q=1,p=q+k}^{Q-k}F_1(p,q)x^{-k} + \sum_{p=1,q=p+k}^{Q-k}F_2(p,q)x^{-k} - C_1\sum_{p=1}^{Q}\sum_{n=1,q=n+k}^{Q-k}\gamma(p,q)\Gamma(p,n)kx^{-k} \\
&\qquad - C_1\sum_{q=1}^{Q}\sum_{p=1,n=p+k}^{Q-k}\gamma(p,q)\Gamma(n,q)kx^{-k} - C_2\sum_{p=1}^{Q}\sum_{q=1}^{Q}\sum_{m=1,n=m+k}^{Q-k}\gamma(p,q)\Gamma(p,m)\Gamma(n,q)kx^{-k} \\
&\qquad - C_2\sum_{p=1}^{Q}\sum_{m=1}^{Q}\sum_{n=1,q=n+k}^{Q-k}\gamma(p,q)\Gamma(p,m)\Gamma(m,n)kx^{-k} \\
&\qquad - C_2\sum_{n=1}^{Q}\sum_{q=1}^{Q}\sum_{p=1,m=p+k}^{Q-k}\gamma(p,q)\Gamma(m,n)\Gamma(n,q)kx^{-k}\Bigg\} \\
&= \sum_{k=1}^{Q-1}\alpha_n(k)x^{-k}
\end{aligned}
\tag{58}
$$

References

[1] J. J. van de Beek, M. Sandell, and P. O. Borjesson, "ML estimation of time and frequency offset in OFDM systems," *IEEE Trans. Signal Process.*, vol. 45, no. 7, pp. 1800–1805, Jul. 1997.

[2] H. T. Hsieh and W. R. Wu, "Maximum likelihood timing and carrier frequency offset estimation for OFDM systems with periodic preamble," *IEEE Trans. Veh. Technol.*, vol. 58, no. 8, pp. 4224–4237, Oct. 2009.

[3] J. Chen, Y. C. Wu, S. C. Chan, and T. S. Ng, "Joint maximum-likelihood CFO and channel estimation for OFDMA uplink using importance sampling," *IEEE Trans. Veh. Technol.*, vol. 57, no. 6, pp. 3462–3470, Nov. 2008.

[4] M. O. Pun, M. Morelli, and C.-C. J. Kuo, "Maximum-likelihood synchronization and channel estimation for OFDMA uplink transmissions," *IEEE Trans. Commun.*, vol. 54, no. 4, pp. 726–736, Apr. 2006.

[5] Z. Wang, Y. Xin, and G. Mathew, "Iterative carrier-frequency offset estimation for generalized OFDMA uplink transmission," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1373–1383, Mar. 2009.

[6] S. Sezginer and P. Bianchi, "Asymptotically efficient reduced complexity frequency offset and channel estimators for uplink MIMO-OFDMA systems," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 964–979, Mar. 2008.

[7] J. J. van de Beek, P. O. Borjesson, M. L. Boucheret, D. Landstrom, J. M. Arenas, O. Odling, M. Wahlqvist, and S. K. Wilson, "A time and frequency synchronization scheme for multiuser OFDM," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 11, pp. 1900–1914, Nov. 1999.

[8] S. Barbarossa, M. Pompili, and G. B. Giannakis, "Channel-independent synchronization of orthogonal frequency division multiple access systems," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 2, pp. 474–486, Feb. 2002.

[9] Z. Cao, U. Tureli, and Y. D. Yao, "Deterministic multiuser carrier frequency offset estimation for interleaved OFDMA uplink," *IEEE Trans. Commun.*, vol. 52, no. 9, pp. 1585–1594, Sep. 2004.

[10] J. Lee, S. Lee, K. J. Bang, S. Cha, and D. Hong, "Carrier frequency offset estimation using ESPRIT for interleaved OFDMA uplink systems," *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 3227–3231, Sep. 2007.

[11] T. J. Lv and J. Chen, "ML estimation of timing and frequency offset using multiple OFDM symbols in OFDM systems," in *Proc. IEEE Global Telecommun. Conf.*, Dec. 2003, vol. 4, pp. 2280–2284.

[12] ETSI, BRAN; HIPERLAN Type 2; Physical (PHY) Layer Specification, Tech. Spec. 101 475, 2001.

[13] S. Kay and S. Saha, "Mean likelihood frequency estimation," *IEEE Trans. Signal Process.*, vol. 48, no. 7, pp. 1937–1946, Jul. 2000.

[14] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*. Philadelphia, PA: SIAM, 2000.

[15] K. Berberidis, S. Rantos, and J. Palicot, "A step-by-step quasi-Newton algorithm in the frequency domain and its application to adaptive channel equalization," *IEEE Trans. Signal Process.*, vol. 52, no. 12, pp. 3335–3344, Dec. 2004.

[16] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part IV, Optimum Array Processing*. Hoboken, NJ: Wiley, 2002.

[17] D. S. Bernstein, *Matrix Mathematics: Theory, Facts, and Formulas With Application to Linear Systems Theory*. Princeton, NJ: Princeton Univ. Press, 2005.

[18] S. Haykin, *Adaptive Filter Theory*, 4th ed. Englewood Cliffs, NJ: Prentice-Hall, 2002.

[19] Y. S. Choi, P. J. Voltz, and F. A. Cassara, "ML estimation of carrier frequency offset for multicarrier signals in Rayleigh fading channels," *IEEE Trans. Veh. Technol.*, vol. 50, no. 2, pp. 644–655, Mar. 2001.

[20] S. M. Kay, *Fundamentals of Statistical Signal Processing—Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[21] D. Kincaid and W. Cheney, *Mathematics of Scientific Computing*, 2nd ed. Pacific Grove, CA: Brooks/Cole, 1996.

[22] A. H. Sayed, *Fundamentals of Adaptive Filtering*. Hoboken, NJ: Wiley, 2003.

[23] M.-O. Pun, M. Morelli, and C.-C. J. Kuo, "Iterative detection and frequency synchronization for OFDMA uplink transmissions," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 629–639, Feb. 2007.

[24] M. Guenach, F. Simoens, H. Wymeersch, and M. Moeneclaey, "Uplink acquisition of a new user accessing a fixed wireless DS-CDMA system," in *Proc. IEEE 6th Workshop SPAWC*, New York, Jun. 2005, pp. 353–357.

Hung-Tao Hsieh received the B.S. degree in physics in 2000 from the National Central University, Jhongli, Taiwan, and the M.S. degree in electrophysics in 2002 from the National Chiao Tung University, Hsinchu, Taiwan, where he is currently working toward the Ph.D. degree.

His research interests include detection/estimation theories and communication signal processing.

Wen-Rong Wu (M'89) received the B.S. degree in mechanical engineering from Tatung Institute of Technology, Taipei, Taiwan, in 1980 and the M.S. degree in mechanical and electrical engineering and the Ph.D. degree in electrical engineering from the State University of New York at Buffalo in 1985, 1986, and 1989, respectively.

Since August 1989, he has been a faculty member with the Department of Communication Engineering, National Chiao Tung University, Hsinchu, Taiwan. His research interests include statistical signal processing and digital communication.