Iterative Detection and Frequency Synchronization for OFDMA Uplink Transmissions

Man-On Pun, *Student Member, IEEE,* Michele Morelli, *Member, IEEE,* and C.-C. Jay Kuo, *Fellow, IEEE*

Abstract— The problem of frequency synchronization, channel estimation and data detection for all active users in the up-link of an OFDMA system is investigated in this work. Since the exact maximum likelihood (ML) solution to this problem turns out to be too complex for practical purposes, we derive an alternative scheme that operates in an iterative fashion. At each step, the superimposed signals arriving at the base station (BS) are separated by means of the space-alternating generalized expectation-maximization (SAGE) algorithm. Each separated signal is then passed to an expectation-conditional maximization (ECM)-based processor that updates frequency estimates and performs channel estimation and data detection for each user. The resulting architecture is reminiscent of the parallel interference cancelation (PIC) receiver, where interference is generated and removed from the received signal to improve the system performance. Simulations indicate that the proposed scheme outperforms other benchmark solutions at the price of increased computational complexity.

Index Terms— Channel estimation, expectation-conditional maximization (ECM), expectation-maximization (EM), frequency synchronization, orthogonal frequency-division multiple-access (OFDMA).

I. INTRODUCTION

THE increasing demand for multimedia communications with variable data rates and different Quality of Service (QoS) requirements has recently led to a strong interest in orthogonal frequency-division multiple-access (OFDMA), which is recognized as a promising technology for wireless metropolitan area networks (MANs) [1], satellite communications [2] and cable TV (CATV) transmissions [3]. In an OFDMA system, several users simultaneously transmit their own data by modulating an exclusive set of orthogonal subcarriers. This approach offers increased robustness against intra-cell interference while allowing straightforward dynamic channel assignment. Furthermore, in contrast with single-carrier systems, channel equalization in OFDMA can be easily accomplished in the frequency-domain by means of a bank of one-tap complex-valued multipliers.

M.-O. Pun and C.-C. Jay Kuo are with the Department of Electrical Engineering and Signal and Image Processing Institute, University of Southern California, Los Angeles, CA 90089-2564 USA (e-mail: mpun@sipi.usc.edu; cckuo@sipi.usc.edu).

M. Morelli is with the Department of Information Engineering, University of Pisa, 56126 Pisa, Italy (e-mail: michele.morelli@iet.unipi.it).

Two critical issues in the design of an OFDMA uplink system are frequency synchronization and channel estimation. Similar to other multicarrier-based techniques, OFDMA is particularly sensitive to carrier frequency offsets (CFOs) caused by oscillator mismatches and/or Doppler shifts. Inaccurate CFO correction results in loss of orthogonality among subcarriers and produces intercarrier interference (ICI) as well as multiple-access interference (MAI) with ensuing limitations of system performance. In addition, knowledge of the channel response of each user is indispensable for coherent detection of transmitted data. Frequency synchronization and channel estimation are challenging tasks in uplink communications due to the existence of multiple CFOs and transmission channels.

A well established scheme to achieve frequency synchronization in OFDMA transmissions is based on a three-step procedure. In the first step, each mobile terminal (MT) performs CFO estimation by exploiting a reference signal transmitted from the base station (BS) at the beginning of each downlink time slot. This operation reduces frequency errors at the MT within a tolerable range and can be easily accomplished using the same techniques employed in OFDM systems [4]–[6]. The frequency estimates are then used by MTs not only to detect the downlink data stream, but also as synchronization references in the subsequent uplink transmission [7].

Due to estimation errors and/or Doppler effects, uplink signals arriving at the BS may be affected by some residual CFOs (typically smaller than the subcarrier spacing). The second step of the frequency synchronization procedure is thus represented by CFO estimation in the uplink. This problem has received some attention in the last few years, and several solutions have been devised for systems employing either a sub-band [8], [9] or interleaved [10] subcarrier assignment scheme (CAS). However, the current trend in OFDMA favors a more flexible generalized CAS (GCAS) in which users can select the best subcarriers (namely, those with the highest channel gains) that are currently available. Since there is no rigid association between subcarriers and users, GCAS allows dynamic resource allocation and provides more flexibility than sub-band or interleaved schemes. Two methods for jointly estimating CFOs and channel responses of all active users in the uplink of an OFDMA system with GCAS have been proposed in [11], [12]. These schemes are based on the maximum likelihood (ML) criterion and they demand that all users send one pilot block (carrying known symbols) at the beginning of each uplink frame.

Once residual CFOs that affect uplink signals have been estimated, they are employed to restore orthogonality among subcarriers. This operation is known as "frequency correction"

and represents the third stage of the synchronization process. A solution to this problem has been discussed in [7] and [9], where CFO estimates are returned to MTs via a downlink control channel and exploited by users to adjust their transmit carrier frequencies. Only after all users have corrected their oscillator's frequency does the BS start to detect the uplink data stream.

In a time-varying scenario, users must be periodically provided with updated CFO estimates, which may result in excessive overhead and outdated information due to feedback delay. A promising alternative is to use advanced signal processing techniques to compensate for residual CFOs *directly* at the BS, *i.e.*, without the need of returning frequency estimates back to subscribers. Solutions derived along this line of reasoning are largely inherited from multi-user detection research, including linear schemes as well as interference-cancelation (IC) architectures. In particular, the method discussed in [13] aims at restoring orthogonality among subcarriers by means of a linear transformation applied to frequency-domain samples. For this purpose, both the least squares (LS) and minimum mean square error (MMSE) criteria are employed. Since this scheme can be regarded as a linear multiuser receiver, we expect that ICI and MAI mitigation is achieved at the price of some noise enhancement, especially in the presence of relatively large CFOs. In [14], uplink signals are tentatively restored at the BS by means of circular convolutions and windowing functions applied at the output of the discrete Fourier transform (DFT) unit. This scheme is simple to implement, but it cannot reduce ICI and MAI effectively. Better results are obtained in [15], where interference is reconstructed and removed from original frequency-domain samples. A possible drawback of this method is the error propagation phenomenon, which may occur in the presence of unreliable data decisions. A solution to this problem is provided in [16] where interference cancelation is accomplished without employing any data decision and the signal of a given user is reconstructed by first windowing the DFT output so as to remove all the energy present in subcarriers allocated to other subscribers. As shown later, however, this operation results in performance degradation in the presence of non-negligible CFOs since the un-designated subcarriers may contain a significant portion of the user's energy which is discarded by the scheme proposed in [16].

In this work, we examine the problem discussed in [13]–[16] and propose a new receiver architecture that jointly performs frequency synchronization, channel estimation and data detection in OFDMA uplink transmissions. The proposed scheme is suitable for any CAS and operates in an iterative fashion. The combined signals from all active users are first separated through the space-alternating generalized expectation-maximization (SAGE) algorithm [17], and then passed to an expectation-conditional maximization (ECM)-based processor [18], where the frequency and channel estimation tasks are coupled with the decision-making process. Compared with previous work in [13]–[16], the proposed scheme can correct larger CFOs without incurring severe performance degradation with respect to a perfectly synchronized system where neither ICI nor MAI is present. It can also be effectively used in a high-mobility environment such as air traffic control and management [19], where CFO and channel variations

must be tracked continuously for reliable data detection. It is nevertheless fair to say that these advantages come at the price of a higher computational complexity with respect to methods presented in [13]–[16].

The rest of the paper is organized as follows. We introduce the signal model for OFDMA uplink transmissions in Section II. The problem of joint estimation of CFOs, channel responses and data symbols of all active users is addressed in Section III. Simulation results are shown in Section IV while concluding remarks are given in Section V.

<u>Notation</u>: Vectors and matrices are denoted by boldface letters. $\|\cdot\|$ represents the Euclidean norm of the enclosed vector. $\boldsymbol{I}_N$ is the $N \times N$ identity matrix, $\boldsymbol{0}_N$ denotes an N-dimensional vector with all zero entries, and $\boldsymbol{A} = \text{diag}\{a_0, a_1, \cdots, a_{N-1}\}$ is a diagonal matrix with $\{a_j; j = 0, 1, \cdots, N - 1\}$ along its main diagonal. $\boldsymbol{M}^{-1}$ and $\text{tr}\{\boldsymbol{M}\}$ indicate the inverse and trace of a square matrix $\boldsymbol{M}$, respectively. We use $E\{\cdot\}$, $(\cdot)^*$, $(\cdot)^T$ and $(\cdot)^H$ for expectation, complex conjugation, transposition and Hermitian transposition. Finally, $\Re\{\cdot\}$ and $\Im\{\cdot\}$ indicate the real and imaginary components of a complex-valued quantity, respectively, and $|\cdot|$ is used for the corresponding amplitude.

II. SIGNAL MODEL FOR OFDMA UPLINK TRANSMISSION

We consider the uplink of an OFDMA system, in which K users are simultaneously active and communicate with the BS as depicted in Fig. 1. We denote N the total number of subcarriers and call $\boldsymbol{s}_k(n)$ the nth block of frequency-domain symbols sent by the kth user, where $k \in \{1, 2, \cdots, K\}$. The jth entry of $\boldsymbol{s}_k(n)$, say $s_{k,j}(n)$, is non-zero if and only if the jth subcarrier is modulated by the kth user, with $j \in \{0, 1, \cdots, N - 1\}$. This means that $\boldsymbol{s}_k(n)$ has only N_k non-zero elements, where N_k is the number of subcarriers assigned to the kth user. The corresponding time-domain vector is given by

$$\boldsymbol{x}_k(n) = \boldsymbol{F}^H \boldsymbol{s}_k(n), \tag{1}$$

where $\boldsymbol{F}$ is the N-point DFT matrix with entries $[\boldsymbol{F}]_{p,q} = \frac{1}{\sqrt{N}} \exp\left(\frac{-j2\pi pq}{N}\right)$ for $0 \leq p, q \leq N - 1$. A cyclic prefix (CP) of length N_g is appended in front of $\boldsymbol{x}_k(n)$ to eliminate the interblock interference (IBI). The resulting vector $\boldsymbol{u}_k(n)$ (of length $N_B = N + N_g$) is then transmitted over the channel.

For simplicity, the channel impulse response (CIR) is assumed static over an OFDMA block, even though it may vary from block to block (slowly-variant channel). Then, we call $\boldsymbol{\xi}_k(n) = [\xi_k(n, 0), \xi_k(n, 1), \cdots, \xi_k(n, L_k - 1)]^T$ the discrete-time baseband CIR of the kth user during the nth block and assume that the channel length L_k remains constant over all blocks. Since L_k is usually unknown, we replace $\boldsymbol{\xi}_k(n)$ by the following L_ξ-dimensional vector

$$\boldsymbol{\xi}'_k(n) \stackrel{\text{def}}{=} \begin{bmatrix} \boldsymbol{\xi}_k^T(n) & \boldsymbol{0}_{L_\xi - L_k}^T \end{bmatrix}^T, \tag{2}$$

where $L_\xi \geq \max_k \{L_k\}$ is a design parameter that depends on the maximum channel delay spread.

The waveform arriving at the BS is the superposition of signals from all active users. As shown in Fig. 1, the discrete-time output of the BS receive filter is divided into adjacent segments of length N_B, each corresponding to a received OFDMA
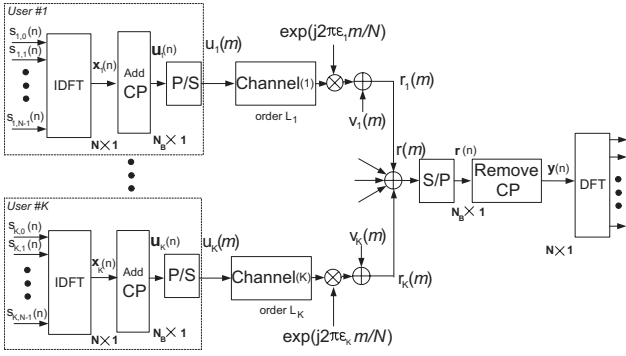
Fig. 1. The discrete-time baseband model of the investigated OFDMA system.

block (in the BS time reference). The samples belonging to the nth block are serial-to-parallel (S/P) converted to form $\boldsymbol{r}(n)$. Next, the CP is removed and the remaining samples are collected into an N-dimensional vector $\boldsymbol{y}(n)$. We consider a quasi-synchronous system where each user performs timing and frequency acquisition through a downlink synchronization channel before initiating the uplink transmission [7]. In this case frequency and timing offsets in the uplink are only dictated by estimation errors occurring at the MTs as well as by Doppler shifts and (two-way) propagation delays from the users' positions to the BS location. The latter are limited to $\tau_{\max} = 2R/c$, where R is the cell radius and $c = 3 \times 10^8$ m/s is the speed of light. In the following, we denote τ_k the timing error of the kth user with respect to the BS time reference and call $\epsilon_k(n)$ the kth CFO (normalized to the subcarrier spacing) during the nth OFDMA block. For convenience, we decompose τ_k into an integer part plus a fractional part with respect to the sampling period T_s, *i.e.*, we let $\tau_k = (\mu_k + \delta_k)T_s$ with $\mu_k = \text{int}\{\tau_k/T_s\}$ and $0 \le \delta_k < 1$. As explained in [7], the fractional part can be incorporated into the CIR and it is not considered in the sequel. Without loss of generality, we concentrate on the nth received block and omit temporal index n for notational simplicity. Then, letting $\mu_{\max} = \max\limits_{k}\{\mu_k\}$ and assuming $N_g \ge L_\xi + \mu_{\max}$ to avoid IBI, we have

$$\boldsymbol{y} = \sum_{k=1}^{K} \boldsymbol{\Gamma}(\epsilon_k)\boldsymbol{F}^H\boldsymbol{D}(\boldsymbol{s}_k)\boldsymbol{W}\boldsymbol{h}_k + \boldsymbol{v}, \qquad (3)$$

where

- $\boldsymbol{\Gamma}(\epsilon_k) = \text{diag}\left\{1, e^{j2\pi\epsilon_k/N}, \cdots, e^{j2\pi(N-1)\epsilon_k/N}\right\}$;
- $\boldsymbol{D}(\boldsymbol{s}_k) = \text{diag}\{s_{k,0}, s_{k,1}, \cdots, s_{k,N-1}\}$ is a diagonal matrix with $\boldsymbol{s}_k$ on its main diagonal;
- $\boldsymbol{h}_k = \begin{bmatrix} \boldsymbol{0}_{\mu_k}^T & \boldsymbol{\xi}_k'^T & \boldsymbol{0}_{L_h - L_\xi - \mu_k}^T \end{bmatrix}^T$ is a vector of dimension $L_h = L_\xi + \mu_{\max}$ that encapsulates both the timing error and channel response of the kth user;
- $\boldsymbol{W}$ is an $N \times L_h$ matrix with elements $[\boldsymbol{W}]_{p,q} = e^{-j2\pi pq/N}$ for $0 \le p \le N-1$ and $0 \le q \le L_h - 1$. In practice, the columns of $\boldsymbol{W}$ are scaled versions of the first L_h columns of $\boldsymbol{F}$;
- $\boldsymbol{v}$ is circularly symmetric white Gaussian noise with zero-mean and covariance matrix $\sigma_v^2 \boldsymbol{I}_N$.

III. ITERATIVE DETECTION AND FREQUENCY SYNCHRONIZATION

Since timing errors μ_k do not appear explicitly in the signal model in (3), timing estimation is not strictly necessary in the proposed system. Thus, in the following we concentrate on the joint estimation of $\boldsymbol{\epsilon} = [\epsilon_1, \epsilon_2, \cdots, \epsilon_K]^T$, $\boldsymbol{h} = \begin{bmatrix} \boldsymbol{h}_1^T, \boldsymbol{h}_2^T, \cdots, \boldsymbol{h}_K^T \end{bmatrix}^T$ and $\boldsymbol{s} = \begin{bmatrix} \boldsymbol{s}_1^T, \boldsymbol{s}_2^T, \cdots, \boldsymbol{s}_K^T \end{bmatrix}^T$ based on received vector $\boldsymbol{y}$. In doing so, we follow an ML approach. Recalling that entries of $\boldsymbol{v}$ are independent Gaussian random variables with zero-mean and variance σ_v^2, the log-likelihood function for unknown parameters $\boldsymbol{\epsilon}$, $\boldsymbol{h}$ and $\boldsymbol{s}$ takes the following form

$$\Lambda\left(\tilde{\boldsymbol{\epsilon}}, \tilde{\boldsymbol{h}}, \tilde{\boldsymbol{s}}\right) = -N\ln\left(\pi\sigma_v^2\right)$$
$$- \frac{1}{\sigma_v^2}\left\| \boldsymbol{y} - \sum_{k=1}^{K}\boldsymbol{\Gamma}(\tilde{\epsilon}_k)\boldsymbol{F}^H\boldsymbol{D}(\tilde{\boldsymbol{s}}_k)\boldsymbol{W}\tilde{\boldsymbol{h}}_k \right\|^2, \qquad (4)$$

where $\tilde{\lambda}$ indicates a trial value of λ. The joint ML estimates of $\boldsymbol{\epsilon}$, $\boldsymbol{h}$ and $\boldsymbol{s}$ are found by searching for the maximum of $\Lambda\left(\tilde{\boldsymbol{\epsilon}}, \tilde{\boldsymbol{h}}, \tilde{\boldsymbol{s}}\right)$ with respect to $\tilde{\boldsymbol{\epsilon}}$, $\tilde{\boldsymbol{h}}$ and $\tilde{\boldsymbol{s}}$. Unfortunately, this operation would require an exhaustive search over the multi-dimensional space spanned by $\tilde{\boldsymbol{\epsilon}}$, $\tilde{\boldsymbol{h}}$ and $\tilde{\boldsymbol{s}}$, which is prohibitively complex for practical implementation. To circumvent this obstacle, we propose an iterative scheme, in which a SAGE-based processor [17] is first used to extract the contribution of each user, say $\hat{\boldsymbol{y}}_k$ ($k = 1, 2, \cdots, K$), from received vector $\boldsymbol{y}$. Each $\hat{\boldsymbol{y}}_k$ is then exploited to jointly estimate ϵ_k, $\boldsymbol{h}_k$ and $\boldsymbol{s}_k$ by following an ECM approach [18].

A. SAGE-Based Signal Decomposition

In a variety of ML problems, direct maximization of the likelihood function is often analytically challenging. The EM algorithm proves to be effective as it tends to achieve the same final result with a comparatively simpler iterative procedure. In the EM formulation, the observed measurements are replaced with some *complete data* from which original measurements could be obtained through a many-to-one mapping [18]. The EM algorithm iteratively alternates between an E-step, calculating the log-likelihood function of the complete data, and an M-step, maximizing the expectation with respect to unknown parameters. The parameter estimates are updated at each iteration, and the process continues until no significant changes in the updates are observed.

The SAGE algorithm improves upon EM in that it has a faster convergence rate. The maximization in the EM algorithm is performed with respect to all unknown parameters *simultaneously*, which results in a slow process since it requires searches over a space with many dimensions. In contrast, the maximization in the SAGE algorithm is performed by updating a smaller group of parameters at a time. The SAGE algorithm was first proposed in [17] and provides a practical solution to parameter estimation from superimposed signals [20]. In particular, it is now exploited to decompose the maximization of $\Lambda\left(\tilde{\boldsymbol{\epsilon}}, \tilde{\boldsymbol{h}}, \tilde{\boldsymbol{s}}\right)$ in (4) into K simpler maximization problems. For this purpose, we view received vector $\boldsymbol{y}$ as the observed data and $\boldsymbol{y}_k$ ($k = 1, 2, \cdots, K$) as the complete data, where

$\boldsymbol{y}_k$ is the contribution of the kth user to $\boldsymbol{y}$ in form of

$$\boldsymbol{y}_k = \boldsymbol{\Gamma}(\epsilon_k)\boldsymbol{F}^H \boldsymbol{D}(\boldsymbol{s}_k)\boldsymbol{W}\boldsymbol{h}_k + \boldsymbol{v}_k, \quad k = 1, 2, \cdots, K \quad (5)$$

and $\boldsymbol{v}_k$ $(k = 1, 2, \cdots, K)$ are circularly symmetric and statistically independent Gaussian vectors satisfying $\boldsymbol{v} = \sum_{k=1}^K \boldsymbol{v}_k$ [20].

The SAGE algorithm is applied in such a way that the parameters of a single user are updated at a time. This leads to a procedure consisting of *iterations* and *cycles*, where K cycles make an iteration and each cycle updates the parameters of a given user. To proceed further, we call $\hat{\epsilon}_k^{(i)}$, $\hat{\boldsymbol{h}}_k^{(i)}$ and $\hat{\boldsymbol{s}}_k^{(i)}$ estimates of ϵ_k, $\boldsymbol{h}_k$ and $\boldsymbol{s}_k$ after the ith iteration, respectively. Given initial estimates $\hat{\epsilon}_k^{(0)}$, $\hat{\boldsymbol{h}}_k^{(0)}$ and $\hat{\boldsymbol{s}}_k^{(0)}$, we compute

$$\hat{\boldsymbol{z}}_k^{(0)} = \boldsymbol{\Gamma}(\hat{\epsilon}_k^{(0)})\boldsymbol{F}^H \boldsymbol{D}(\hat{\boldsymbol{s}}_k^{(0)})\boldsymbol{W}\hat{\boldsymbol{h}}_k^{(0)}, \quad k = 1, 2, \cdots, K. \quad (6)$$

Then, during the mth cycle of the ith iteration (with $m = 1, 2, \cdots, K$), the SAGE procedure proceeds as follows [17].

E-Step:

Compute

$$\hat{\boldsymbol{y}}_m^{(i)} = \boldsymbol{y} - \sum_{k=1}^{m-1} \hat{\boldsymbol{z}}_k^{(i)} - \sum_{k=m+1}^{K} \hat{\boldsymbol{z}}_k^{(i-1)} \quad (7)$$

where $\sum_l^u$ is zero if $u < l$.

M-Step:

Compute

$$\begin{aligned} &\left[\hat{\epsilon}_m^{(i)}, \hat{\boldsymbol{h}}_m^{(i)}, \hat{\boldsymbol{s}}_m^{(i)}\right] \\ &= \underset{\tilde{\epsilon}_m, \tilde{\boldsymbol{h}}_m, \tilde{\boldsymbol{s}}_m}{\arg\min} \left\{ \left\| \hat{\boldsymbol{y}}_m^{(i)} - \boldsymbol{\Gamma}(\tilde{\epsilon}_m)\boldsymbol{F}^H \boldsymbol{D}(\tilde{\boldsymbol{s}}_m)\boldsymbol{W}\tilde{\boldsymbol{h}}_m \right\|^2 \right\}, \end{aligned} \quad (8)$$

and then use updated parameters to obtain the following vector

$$\hat{\boldsymbol{z}}_m^{(i)} = \boldsymbol{\Gamma}(\hat{\epsilon}_m^{(i)})\boldsymbol{F}^H \boldsymbol{D}(\hat{\boldsymbol{s}}_m^{(i)})\boldsymbol{W}\hat{\boldsymbol{h}}_m^{(i)}. \quad (9)$$

We see from (8) that the SAGE algorithm splits the maximization of $\Lambda\left(\tilde{\epsilon}, \tilde{\boldsymbol{h}}, \tilde{\boldsymbol{s}}\right)$ in (4) into a series of K simpler optimization problems. However, the multi-dimensional minimization in (8) still remains a formidable task. An iterative solution to this problem is presented in the next subsection by resorting to the ECM algorithm.

B. ECM-Based Iterative Estimator

Substituting (3) into (7) yields

$$\hat{\boldsymbol{y}}_m^{(i)} = \boldsymbol{\Gamma}(\epsilon_m)\boldsymbol{F}^H \boldsymbol{D}(\boldsymbol{s}_m)\boldsymbol{W}\boldsymbol{h}_m + \boldsymbol{\eta}_m^{(i)}, \quad (10)$$

where

$$\boldsymbol{\eta}_m^{(i)} = \boldsymbol{v} + \sum_{k=1}^{m-1} \left[\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k^{(i)}\right] + \sum_{k=m+1}^{K} \left[\boldsymbol{z}_k - \hat{\boldsymbol{z}}_k^{(i-1)}\right], \quad (11)$$

and $\boldsymbol{z}_k = \boldsymbol{\Gamma}(\epsilon_k)\boldsymbol{F}^H \boldsymbol{D}(\boldsymbol{s}_k)\boldsymbol{W}\boldsymbol{h}_k$ is the signal received from the kth user. Note that $\boldsymbol{\eta}_m^{(i)}$ is a disturbance term that accounts for thermal noise and residual MAI after the ith SAGE iteration, where the MAI is linearly related to the data symbols

of all interfering users. Then, assuming that the latter are independent and identically distributed with zero-mean, it follows from the central limit theorem that the entries of $\boldsymbol{\eta}_m^{(i)}$ are nearly Gaussian distributed with zero-mean and some variance $\sigma_\eta^2(i)$. Under the above assumption, it turns out that the minimization problem in (8) is equivalent to the ML estimation of ϵ_m, $\boldsymbol{h}_m$ and $\boldsymbol{s}_m$ starting from the observations $\hat{\boldsymbol{y}}_m^{(i)}$.

The ECM algorithm offers a practical solution to this problem. The only difference between this technique and the conventional EM algorithm is that the maximization step in the ECM algorithm is divided into several stages. At each stage, only one parameter is updated while the others are kept constant at their most updated values. This makes the ECM algorithm suitable for *multi-dimensional* ML estimation problems, where the likelihood function must be optimized over several parameters [18]. In the following, we apply the ECM algorithm to solve the optimization problem in (8).

In doing so, we view $\hat{\boldsymbol{y}}_m^{(i)}$ as the observed data and $\left[\begin{array}{cc} \hat{\boldsymbol{y}}_m^{(i)T} & \boldsymbol{h}_m^T \end{array}\right]^T$ as the complete set of data. Also, we denote $\boldsymbol{\theta}_m \stackrel{\text{def}}{=} \left[\begin{array}{cc} \boldsymbol{s}_m^T & \epsilon_m \end{array}\right]^T$ the parameters to be estimated and $\hat{\boldsymbol{\theta}}_m^{(i,c)} = \left[\begin{array}{cc} \hat{\boldsymbol{s}}_m^{(i,c)T} & \hat{\epsilon}_m^{(i,c)} \end{array}\right]^T$ the estimate of $\boldsymbol{\theta}_m$ at the cth ECM and ith SAGE iterations. After initializing $\hat{\boldsymbol{s}}_m^{(i,0)} = \hat{\boldsymbol{s}}_m^{(i-1)}$ and $\hat{\epsilon}_m^{(i,0)} = \hat{\epsilon}_m^{(i-1)}$, the ECM algorithm alternates between an E-step and an M-step as follows.

E-Step:

We define

$$\begin{aligned} Q\left(\tilde{\boldsymbol{\theta}}_m \middle| \hat{\boldsymbol{\theta}}_m^{(i,c)}\right) = \\ E_{\boldsymbol{h}_m}\left\{\ln\left[p\left(\hat{\boldsymbol{y}}_m^{(i)} \middle| \boldsymbol{h}_m, \tilde{\boldsymbol{\theta}}_m\right)\right] p\left(\hat{\boldsymbol{y}}_m^{(i)} \middle| \boldsymbol{h}_m, \hat{\boldsymbol{\theta}}_m^{(i,c)}\right)\right\}, \end{aligned} \quad (12)$$

where $p\left(\hat{\boldsymbol{y}}_m^{(i)} \middle| \boldsymbol{h}_m, \tilde{\boldsymbol{\theta}}_m\right)$ and $p\left(\hat{\boldsymbol{y}}_m^{(i)} \middle| \boldsymbol{h}_m, \hat{\boldsymbol{\theta}}_m^{(i,c)}\right)$ are conditional probability density functions (pdf), $E_{\boldsymbol{h}_m}\{\cdot\}$ denotes the statistical expectation over the pdf of $\boldsymbol{h}_m$ and $\tilde{\boldsymbol{\theta}}_m = \left[\begin{array}{cc} \tilde{\boldsymbol{s}}_m^T & \tilde{\epsilon}_m \end{array}\right]^T$ is a trial value of $\boldsymbol{\theta}_m$. Function $Q\left(\tilde{\boldsymbol{\theta}}_m \middle| \hat{\boldsymbol{\theta}}_m^{(i,c)}\right)$ is computed in Appendix for PSK/QAM constellations after modeling $\boldsymbol{h}_m$ as a Gaussian vector with zero-mean and covariance matrix $\boldsymbol{C}_m = E\left\{\boldsymbol{h}_m \boldsymbol{h}_m^H\right\}$. Skipping the additive terms and multiplicative factors independent of $\tilde{\boldsymbol{\theta}}_m$, $Q\left(\tilde{\boldsymbol{\theta}}_m \middle| \hat{\boldsymbol{\theta}}_m^{(i,c)}\right)$ can be equivalently replaced by $\Phi\left(\tilde{\boldsymbol{\theta}}_m \middle| \hat{\boldsymbol{\theta}}_m^{(i,c)}\right)$ shown in (13), where

$$\begin{aligned} \hat{\boldsymbol{h}}_{m,\text{MMSE}}\left(\hat{\boldsymbol{\theta}}_m^{(i,c)}\right) = \\ \boldsymbol{P}^{-1}\left(\hat{\boldsymbol{s}}_m^{(i,c)}\right) \boldsymbol{W}^H \boldsymbol{D}^H\left(\hat{\boldsymbol{s}}_m^{(i,c)}\right) \boldsymbol{F}\boldsymbol{\Gamma}^H\left(\hat{\epsilon}_m^{(i,c)}\right) \hat{\boldsymbol{y}}_m^{(i)} \end{aligned} \quad (14)$$

is the MMSE estimate of $\boldsymbol{h}_m$ obtained with $\boldsymbol{\theta}_m = \hat{\boldsymbol{\theta}}_m^{(i,c)}$, while

$$\boldsymbol{P}\left(\hat{\boldsymbol{s}}_m^{(i,c)}\right) = \boldsymbol{W}^H \boldsymbol{E}_m\left(\hat{\boldsymbol{s}}_m^{(i,c)}\right) \boldsymbol{W} + \sigma_\eta^2(i)\boldsymbol{C}_m^{-1} \quad (15)$$

with $\boldsymbol{E}_m\left(\hat{\boldsymbol{s}}_m^{(i,c)}\right) = \text{diag}\left\{\left|\hat{s}_{m,j}^{(i,c)}\right|^2; j = 0, 1, \cdots, N-1\right\}$.

$$\Phi\left(\tilde{\boldsymbol{\theta}}_m \left| \hat{\boldsymbol{\theta}}_m^{(i,c)}\right.\right) = -\left\| \hat{\boldsymbol{y}}_m^{(i)} - \boldsymbol{\Gamma}(\tilde{\epsilon}_m)\boldsymbol{F}^H \boldsymbol{D}(\tilde{\boldsymbol{s}}_m)\boldsymbol{W}\hat{\boldsymbol{h}}_{m,\text{MMSE}}\left(\hat{\boldsymbol{\theta}}_m^{(i,c)}\right)\right\|^2 - \sigma_\eta^2(i)\cdot\text{tr}\left\{ \boldsymbol{D}\left(\tilde{\boldsymbol{s}}_m\right)\boldsymbol{W}\boldsymbol{P}^{-1}\left(\hat{\boldsymbol{s}}_m^{(i,c)}\right)\boldsymbol{W}^H\boldsymbol{D}^H\left(\tilde{\boldsymbol{s}}_m\right)\right\}, \tag{13}$$

$$\hat{\boldsymbol{h}}_{m,\text{LS}}\left(\hat{\boldsymbol{\theta}}_m^{(i,c)}\right) = \left[\boldsymbol{W}^H\boldsymbol{E}_m\left(\hat{\boldsymbol{s}}_m^{(i,c)}\right)\boldsymbol{W}\right]^{-1}\boldsymbol{W}^H\boldsymbol{D}^H\left(\hat{\boldsymbol{s}}_m^{(i,c)}\right)\boldsymbol{F}\boldsymbol{\Gamma}^H\left(\hat{\epsilon}_m^{(i,c)}\right)\hat{\boldsymbol{y}}_m^{(i)}. \tag{17}$$

We see from (13)-(15) that evaluating $\Phi\left(\tilde{\boldsymbol{\theta}}_m \left| \hat{\boldsymbol{\theta}}_m^{(i,c)}\right.\right)$ requires knowledge of $\boldsymbol{C}_m$ and $\sigma_\eta^2(i)$. Thus, suitable schemes must be found to estimate these parameters. A practical solution to this problem can be derived by considering high SNR values. In this case, we expect that $\sigma_\eta^2(i)$ becomes vanishingly small and $\Phi\left(\tilde{\boldsymbol{\theta}}_m \left| \hat{\boldsymbol{\theta}}_m^{(i,c)}\right.\right)$ can be reasonably approximated by

$$\bar{\Phi}\left(\tilde{\boldsymbol{\theta}}_m \left| \hat{\boldsymbol{\theta}}_m^{(i,c)}\right.\right) = \\ -\left\| \hat{\boldsymbol{y}}_m^{(i)} - \boldsymbol{\Gamma}(\tilde{\epsilon}_m)\boldsymbol{F}^H \boldsymbol{D}(\tilde{\boldsymbol{s}}_m)\boldsymbol{W}\hat{\boldsymbol{h}}_{m,\text{LS}}\left(\hat{\boldsymbol{\theta}}_m^{(i,c)}\right)\right\|^2, \tag{16}$$

where $\hat{\boldsymbol{h}}_{m,\text{LS}}\left(\hat{\boldsymbol{\theta}}_m^{(i,c)}\right)$ is the LS estimate of $\boldsymbol{h}_m$ and takes the form shown in (17).

Since this approach leads to a significant reduction of system complexity, $\bar{\Phi}\left(\tilde{\boldsymbol{\theta}}_m \left| \hat{\boldsymbol{\theta}}_m^{(i,c)}\right.\right)$ is used in place of $\Phi\left(\tilde{\boldsymbol{\theta}}_m \left| \hat{\boldsymbol{\theta}}_m^{(i,c)}\right.\right)$ in the sequel. Although this procedure may entail some performance degradation at low-medium SNRs, it has the advantage of being practically implementable, while computing $\Phi\left(\tilde{\boldsymbol{\theta}}_m \left| \hat{\boldsymbol{\theta}}_m^{(i,c)}\right.\right)$ seems hardly viable in practice.

M-Step:

The M-step aims at maximizing the RHS of (16) with respect to $\tilde{\boldsymbol{s}}_m$ and $\tilde{\epsilon}_m$. This goal is achieved using a two-stage procedure. Following the notation of [18], we denote $\hat{\boldsymbol{\theta}}_m^{(i,c+g/2)}$ the estimate of $\boldsymbol{\theta}_m$ at the gth stage of the cth ECM iteration, where $g = 1, 2$. Then, the maximum of $\bar{\Phi}\left(\tilde{\boldsymbol{\theta}}_m \left| \hat{\boldsymbol{\theta}}_m^{(i,c)}\right.\right)$ can be found as follows.

Step 1:

$$\hat{\boldsymbol{\theta}}_m^{(i,c+1/2)} = \left[\left(\hat{\boldsymbol{s}}_m^{(i,c)}\right)^T \quad \hat{\epsilon}_m^{(i,c+1)}\right]^T, \tag{18}$$

where

$$\hat{\epsilon}_m^{(i,c+1)} = \arg\max_{\tilde{\epsilon}_m}\{ \\ -\left\| \hat{\boldsymbol{y}}_m^{(i)} - \boldsymbol{\Gamma}(\tilde{\epsilon}_m)\boldsymbol{F}^H \boldsymbol{D}\left(\hat{\boldsymbol{s}}_m^{(i,c)}\right)\boldsymbol{W}\hat{\boldsymbol{h}}_{m,\text{LS}}\left(\hat{\boldsymbol{\theta}}_m^{(i,c)}\right)\right\|^2 \}. \tag{19}$$

Note that the quantity

$$\left\| \boldsymbol{\Gamma}(\tilde{\epsilon}_m)\boldsymbol{F}^H \boldsymbol{D}\left(\hat{\boldsymbol{s}}_m^{(i,c)}\right)\boldsymbol{W}\hat{\boldsymbol{h}}_{m,\text{LS}}\left(\hat{\boldsymbol{\theta}}_m^{(i,c)}\right)\right\|^2$$

in (19) is independent of $\tilde{\epsilon}_m$ since $\boldsymbol{\Gamma}^H(\tilde{\epsilon}_m)\boldsymbol{\Gamma}(\tilde{\epsilon}_m) = \boldsymbol{I}_N$. Thus, (19) can be equivalently replaced by (20).

Step 2:

$$\hat{\boldsymbol{\theta}}_m^{(i,c+1)} = \left[\left(\hat{\boldsymbol{s}}_m^{(i,c+1)}\right)^T \quad \hat{\epsilon}_m^{(i,c+1)}\right]^T, \tag{21}$$

where

$$\hat{\boldsymbol{s}}_m^{(i,c+1)} = \arg\min_{\tilde{\boldsymbol{s}}_m} \\ \left\{ \sum_{j=0}^{N-1}\left| \hat{Y}_{m,j}^{(i)}\left(\hat{\epsilon}_m^{(i,c+1)}\right) - \tilde{s}_{m,j}\hat{H}_{m,j,\text{LS}}\left(\hat{\boldsymbol{\theta}}_m^{(i,c)}\right)\right|^2\right\}, \tag{22}$$

with $\left\{ \hat{Y}_{m,j}^{(i)}\left(\hat{\epsilon}_m^{(i,c+1)}\right); j = 0, 1, \cdots, N-1\right\}$ and $\left\{ \hat{H}_{m,j,\text{LS}}\left(\hat{\boldsymbol{\theta}}_m^{(i,c)}\right); j = 0, 1, \cdots, N-1\right\}$ being the N-point DFTs of $\boldsymbol{\Gamma}^H(\hat{\epsilon}_m^{(i,c+1)})\hat{\boldsymbol{y}}_m^{(i)}$ and $\hat{\boldsymbol{h}}_{m,\text{LS}}\left(\hat{\boldsymbol{\theta}}_m^{(i,c)}\right)$, respectively. It is worth mentioning that we only need to make decisions on the N_m non-zero entries of $\tilde{\boldsymbol{s}}_m$.

An approximation of the CFO estimate in (20) can be obtained in closed-form after replacing $\boldsymbol{\Gamma}(\tilde{\epsilon}_m)$ with its Taylor series expansion truncated to the second order term and using $\hat{\epsilon}_m^{(i,c)}$ as the starting point, *i.e.*,

$$\boldsymbol{\Gamma}\left(\tilde{\epsilon}_m\right) \approx \boldsymbol{\Gamma}\left(\hat{\epsilon}_m^{(i,c)}\right) + j\left(\tilde{\epsilon}_m - \hat{\epsilon}_m^{(i,c)}\right)\boldsymbol{\Gamma}'\left(\hat{\epsilon}_m^{(i,c)}\right) \\ -\frac{1}{2}\left(\tilde{\epsilon}_m - \hat{\epsilon}_m^{(i,c)}\right)^2\boldsymbol{\Gamma}''\left(\hat{\epsilon}_m^{(i,c)}\right), \tag{23}$$

where $\boldsymbol{\Gamma}'\left(\hat{\epsilon}_m^{(i,c)}\right) = \boldsymbol{\Psi}\boldsymbol{\Gamma}\left(\hat{\epsilon}_m^{(i,c)}\right)$, $\boldsymbol{\Gamma}''\left(\hat{\epsilon}_m^{(i,c)}\right) = \boldsymbol{\Psi}^2\boldsymbol{\Gamma}\left(\hat{\epsilon}_m^{(i,c)}\right)$ and $\boldsymbol{\Psi} = \frac{2\pi}{N}\cdot\text{diag}\{0, 1, \cdots, N-1\}$. Substituting (23) into (20) and setting the derivative with respect to $\tilde{\epsilon}_m$ to zero yields (24).

After C iterations, where C is a design parameter, we terminate the ECM process and replace (8) with

$$\left[\hat{\epsilon}_m^{(i)}, \hat{\boldsymbol{h}}_m^{(i)}, \hat{\boldsymbol{s}}_m^{(i)}\right] = \left[\hat{\epsilon}_m^{(i,C)}, \hat{\boldsymbol{h}}_{m,\text{LS}}\left(\hat{\boldsymbol{\theta}}_m^{(i,C)}\right), \hat{\boldsymbol{s}}_m^{(i,C)}\right]. \tag{25}$$

The iterative scheme relying on (17), (22) and (24) is referred to as the EM-based receiver (EMBR).

C. Initialization

It is well known that a good initialization is essential to EM-type algorithms. Hence, the problem arises of how to obtain the initial estimates $\hat{\epsilon}_k^{(0)}$, $\hat{\boldsymbol{h}}_k^{(0)}$ and $\hat{\boldsymbol{s}}_k^{(0)}$ before initiating the SAGE procedure. If $\boldsymbol{\epsilon}$ and $\boldsymbol{h}$ vary slowly in time, the frequency and channel estimates obtained in a given block can be used to initialize the iterative process in the next block. As to the first data block, the estimates may be obtained in a data-aided fashion provided that a training sequence is placed at the beginning of the uplink frame [11], [12].

$$\hat{\epsilon}_m^{(i,c+1)} = \arg\max_{\tilde{\epsilon}_m} \left\{ \Re \left[\hat{\boldsymbol{y}}_m^{(i)H} \boldsymbol{\Gamma}(\tilde{\epsilon}_m) \boldsymbol{F}^H \boldsymbol{D}(\hat{\boldsymbol{s}}_m^{(i,c)}) \boldsymbol{W} \hat{\boldsymbol{h}}_{m,\text{LS}} \left(\hat{\boldsymbol{\theta}}_m^{(i,c)} \right) \right] \right\}. \tag{20}$$

$$\hat{\epsilon}_m^{(i,c+1)} = \hat{\epsilon}_m^{(i,c)} + \frac{\Im \left\{ \hat{\boldsymbol{y}}_m^{(i)H} \boldsymbol{\Gamma}' \left(\hat{\epsilon}_m^{(i,c)} \right) \boldsymbol{F}^H \boldsymbol{D} \left(\hat{\boldsymbol{s}}_m^{(i,c)} \right) \boldsymbol{W} \hat{\boldsymbol{h}}_{m,\text{LS}} \left(\hat{\boldsymbol{\theta}}_m^{(i,c)} \right) \right\}}{\Re \left\{ \hat{\boldsymbol{y}}_m^{(i)H} \boldsymbol{\Gamma}'' \left(\hat{\epsilon}_m^{(i,c)} \right) \boldsymbol{F}^H \boldsymbol{D} \left(\hat{\boldsymbol{s}}_m^{(i,c)} \right) \boldsymbol{W} \hat{\boldsymbol{h}}_{m,\text{LS}} \left(\hat{\boldsymbol{\theta}}_m^{(i,c)} \right) \right\}}. \tag{24}$$

CFO estimates $\hat{\boldsymbol{\epsilon}}^{(0)} = \left[\hat{\epsilon}_1^{(0)}, \hat{\epsilon}_2^{(0)}, \cdots, \hat{\epsilon}_K^{(0)} \right]^T$ are then exploited to accomplish frequency correction by resorting to one of the methods discussed in [13]–[16]. This operation aims at restoring orthogonality among subcarriers and produces the following N-dimensional vectors (one for each user)

$$\boldsymbol{\psi}_k = \boldsymbol{D}\left(\boldsymbol{s}_k \right) \boldsymbol{W} \boldsymbol{h}_k + \boldsymbol{\gamma}_k, \quad k = 1, 2, \cdots, K \tag{26}$$

where $\boldsymbol{\gamma}_k$ is a disturbance term and we have skipped the dependence of $\boldsymbol{\psi}_k$ on $\hat{\boldsymbol{\epsilon}}^{(0)}$ for notational simplicity. Finally, initial data decisions can be obtained as in conventional OFDM transmission, i.e.,

$$\hat{\boldsymbol{s}}_k^{(0)} = \arg\min_{\tilde{\boldsymbol{s}}_k} \left\{ \sum_{j=0}^{N-1} \left| \psi_{k,j} - \tilde{s}_{k,j} \hat{H}_{k,j}^{(0)} \right|^2 \right\}, \tag{27}$$

where $\psi_{k,j}$ is the jth entry of $\boldsymbol{\psi}_k$ and $\left\{ \hat{H}_{k,j}^{(0)}; j = 0, 1, \cdots, N-1 \right\}$ is the N-point DFT of $\hat{\boldsymbol{h}}_k^{(0)}$.

In applications characterized by high user mobility, initializing the SAGE iterations with channel estimates from the previous block may result in poor performance due to fast fading. In these circumstances, a possible solution is to insert scattered pilots in each OFDMA block and compute $\hat{\boldsymbol{h}}_k^{(0)}$ through conventional pilot-aided estimation techniques [21]. Albeit being robust against rapidly varying channels, this approach inevitably results in a reduction of overall data throughput due to the increased overhead.

D. Remarks

It is worthwhile to point out several interesting properties related to the proposed iterative detection and frequency synchronization scheme.

1) For PSK transmissions, matrix $\boldsymbol{E}_m \left(\tilde{\boldsymbol{s}}_m \right) = \text{diag} \left\{ |\tilde{s}_{m,j}|^2; j = 0, 1, \cdots, N-1 \right\}$ becomes independent of $\tilde{\boldsymbol{s}}_m$ since $|\tilde{s}_{m,j}|^2$ is either unitary or zero depending on whether the jth subcarrier is assigned to the mth user or not. In this case, evaluating $\hat{\boldsymbol{h}}_{m,\text{LS}} \left(\hat{\boldsymbol{\theta}}_m^{(i,c)} \right)$ in (17) does not require any on-line matrix inversion since $\left[\boldsymbol{W}^H \boldsymbol{E}_m \left(\tilde{\boldsymbol{s}}_m \right) \boldsymbol{W} \right]^{-1}$ can be pre-computed and stored in the receiver. A further simplification is possible if the subcarriers of the mth user are uniformly distributed over the signal bandwidth with separation interval N/N_m. Then, $\boldsymbol{W}^H \boldsymbol{E}_m \left(\tilde{\boldsymbol{s}}_m \right) \boldsymbol{W}$ reduces to $N_m \cdot \boldsymbol{I}_{L_h}$ and (17) becomes

$$\hat{\boldsymbol{h}}_{m,\text{LS}} \left(\hat{\boldsymbol{\theta}}_m^{(i,c)} \right) =$$
$$\frac{1}{N_m} \boldsymbol{W}^H \boldsymbol{D}^H \left(\hat{\boldsymbol{s}}_m^{(i,c)} \right) \boldsymbol{F} \boldsymbol{\Gamma}^H \left(\hat{\epsilon}_m^{(i,c)} \right) \hat{\boldsymbol{y}}_m^{(i)}. \tag{28}$$

2) Intuitively speaking, the SAGE procedure should be stopped when no significant variations are observed in the log-likelihood function, i.e.,

$$\Lambda \left(\hat{\boldsymbol{\epsilon}}^{(i)}, \hat{\boldsymbol{h}}^{(i)}, \hat{\boldsymbol{s}}^{(i)} \right) - \Lambda \left(\hat{\boldsymbol{\epsilon}}^{(i-1)}, \hat{\boldsymbol{h}}^{(i-1)}, \hat{\boldsymbol{s}}^{(i-1)} \right) < \lambda_{th}$$

for some design parameter λ_{th}. A simpler stopping criterion is to terminate the SAGE procedure after a pre-assigned number of iterations. This strategy is adopted in our scheme to reduce system complexity.

3) The computational load of EMBR can be assessed as follows. Multiplications by matrices $\boldsymbol{W}$ and $\boldsymbol{F}^H$ in (9) can be efficiently performed by means of DFT units and need $3N \log_2 N$ operations for each vector $\hat{\boldsymbol{z}}_m^{(i)}$. Since multiplication by either $\boldsymbol{\Gamma}(\hat{\epsilon}_m^{(i)})$ or $\boldsymbol{D}(\hat{\boldsymbol{s}}_m^{(i)})$ is equivalent to N products, the overall complexity involved in computing $\hat{\boldsymbol{z}}_m^{(i)}$ for all K active users is in the order of $KN(2 + 3\log_2 N)$. Using similar arguments and assuming for simplicity an interleaved CAS, it can be shown that the same number of operations is required to evaluate $\hat{\boldsymbol{h}}_{m,\text{LS}} \left(\hat{\boldsymbol{\theta}}_m^{(i,c)} \right)$ in (17) for $m = 1, 2, \cdots, K$. Computing $\hat{\boldsymbol{\epsilon}}^{(i,c+1)} = \left[\hat{\epsilon}_1^{(i,c+1)}, \hat{\epsilon}_2^{(i,c+1)}, \cdots, \hat{\epsilon}_K^{(i,c+1)} \right]^T$ in (24) needs $3KN(1 + \log_2 N)$ operations, while $NK(1 + 1.5\log_2 N)$ operations are approximately required in (22) for data detection. From the above figures, it follows that the overall complexity involved in each SAGE iteration is $\mathcal{O}(10.5KCN \log_2 N)$. Clearly, this result does not take into account the operations required to compute $\hat{\epsilon}_k^{(0)}$, $\hat{\boldsymbol{h}}_k^{(0)}$ and $\hat{\boldsymbol{s}}_k^{(0)}$ during the initialization stage. A certain reduction of complexity is possible if the CFOs are nearly constant over the uplink frame. In this case the initial frequency estimates obtained during the training period need not be updated over the data blocks, and the number of operations per SAGE iteration reduces to $\mathcal{O}(7.5KCN \log_2 N)$. It is worth noting that the complexity of EMBR is mainly due to DFT and IDFT operations. This may not represent a serious problem in practice since DFT and IDFT units are anyhow present at the BS and, therefore, no extra circuitry is required.

IV. SIMULATION RESULTS

The performance of EMBR has been assessed by computer simulation in an OFDMA scenario inspired by the IEEE 802.16 standard for WirelessMAN [1]. Without loss of generality, we only provide results for user #1.

A. System Parameters

The simulated system has $N = 128$ subcarriers and a signal bandwidth of 1.429 MHz corresponding to a sampling period

of $T_s = 0.7 \ \mu s$. The useful part of each OFDMA block has length $T = NT_s = 89.6 \ \mu s$ while the inter-carrier spacing is $1/T = 11.16$ kHz. We consider an interleaved CAS where each user is provided with a set of 32 subcarriers (called a subchannel) uniformly spaced over the signal bandwidth. In this way, the maximum number of active users in each OFDMA block is $K_{\max} = 4$. We assume a fully-loaded system where $K = K_{\max}$. With the exception of Fig. 10, users' CFOs are $\boldsymbol{\epsilon} = \rho \cdot [1, -1, 1, -1]^T$, where ρ is modeled as a deterministic parameter belonging to interval $[0, 0.5]$ and it is referred to as the CFO *attenuation factor* [16]. Unless otherwise specified, information bits are mapped onto uncoded QPSK symbols through a Gray map. We assume a cell radius of $R = 0.3$ km so that the maximum two-way propagation delay (normalized to T_s) is $\text{int}\{2R/cT_s\} = 3$. Timing errors μ_k are independently generated at the beginning of each uplink frame. They are taken from the set $\{0, 1, 2, 3\}$ with equal probability and kept constant over the frame. The channel responses $\boldsymbol{\xi}_k(n)$ have length $L_\xi = 5$, corresponding to a CIR duration of $3.5 \ \mu s$. This means that each $\boldsymbol{h}_k(n)$ has dimension $L_h = L_\xi + \mu_{\max} = 8$. A cyclic prefix of length $N_g = 8$ is used to avoid IBI so that the duration of the extended OFDMA block (including the cyclic prefix) is $T_B = (N + N_g)T_s = 95.2 \ \mu s$.

The channel taps $\xi_k(n, l)$ are modeled as statistically independent narrow-band Gaussian processes with zero-mean and autocorrelation function

$$E\{\xi_k(n, l)\xi_k^*(n + m, l)\} = \sigma_{\xi_k}^2(l)J_0\left(2\pi m B_D T_B\right), \quad (29)$$

for $l = 0, 1, 2, 3, 4$, where B_D is the Doppler bandwidth, $J_0(x)$ is the zeroth-order Bessel function of the first kind and

$$\sigma_{\xi_k}^2(l) = E\left\{|\xi_k(n, l)|^2\right\} = \beta_k \cdot \exp(-l), \quad l = 0, 1, 2, 3, 4. \quad (30)$$

In (30), β_1 is chosen such that the signal power of user #1 is normalized to unity, *i.e.*, $E\left\{\|\boldsymbol{\xi}_1\|^2\right\} = 1$, while parameters β_k $(k \geq 2)$ affect the signal-to-interference ratio. The Doppler bandwidth is related to carrier frequency f_0 and mobile velocity v through $B_D = f_0 v/c$. Letting $f_0 = 2$ GHz and $v = 60$ km/h, we obtain $B_d \approx 110$ Hz, corresponding to 1% of subcarrier spacing.

The uplink frame is composed of 10 OFDMA blocks. As mentioned previously, the frequency and channel estimates obtained in a given block are used to initialize the iterative process in the next block while initialization for the first data block is achieved as discussed in [11] using a training sequence placed at the beginning of the frame. For each block, the initial CFO estimates $\hat{\boldsymbol{\epsilon}}^{(0)}$ are employed to restore orthogonality among subcarriers by resorting to the scheme proposed by Cao, Tureli, Yao and Honan (CTYH) in [13], where a linear transformation is applied to the DFT output to obtain vectors $\boldsymbol{\psi}_k$ $(k = 1, 2, \cdots, K)$ in (26). The latter are exploited to get initial channel estimates $\hat{\boldsymbol{h}}_k^{(0)}$. For this task we employ the pilot-aided estimator described in [21] and assume that 8 pilots are uniformly placed within each subchannel at a distance $1/(8T_s)$ from each other. Initial data decisions are eventually obtained according to (27). Unless otherwise specified, the number C of ECM iterations is set to
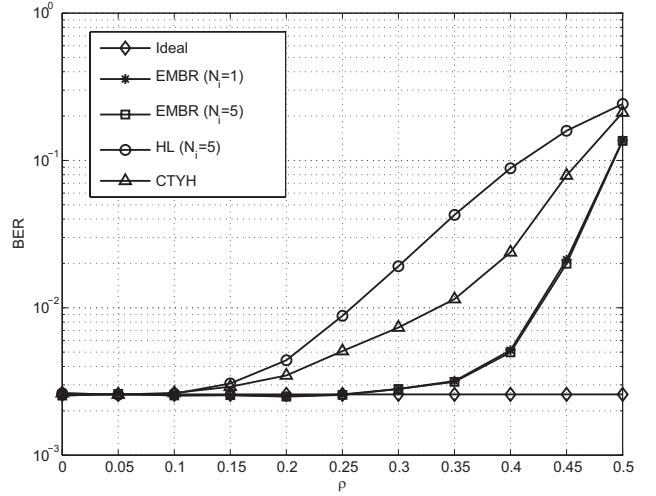

Fig. 2. BER performance v.s. ρ for uncoded QPSK with $E_b/N_0 = 20$ dB and perfect knowledge of the CFOs and channel responses.

1 while the number N_i of SAGE iterations is varied throughout simulations to assess its impact on the system performance.

B. Performance Assessment

Case 1: Performance with perfect channel and frequency estimates

Fig. 2 shows the BER performance as a function of ρ obtained with perfect knowledge of CFOs and channel responses, *i.e.*, we let $\hat{\epsilon}_k = \epsilon_k$ and $\hat{\boldsymbol{h}}_{k,\text{LS}} = \boldsymbol{h}_k$ for $k = 1, 2, 3, 4$. This scenario was also considered in [13]–[16] and it is used here to assess the ability of the proposed scheme to mitigate ICI and MAI produced by frequency offsets. Users have equal power with $E_b/N_0 = 20$dB. Comparisons are made with both CTHY [13] and the iterative scheme proposed by Huang and Letaief (HL) in [16], where frequency correction is accomplished at the output of the receive DFT by means of interference cancelation techniques and windowing functions. Five iterations are employed in HL while the number of SAGE iterations is either $N_i = 1$ or 5. The curve labeled "ideal" is obtained by assuming that all CFOs have been perfectly corrected at MTs, *i.e.*, $\epsilon_k = 0$ for $k = 1, 2, 3, 4$. This provides a benchmark for the BER performance since in this case users' signals at the DFT output are perfectly orthogonal and no interference is present. As expected, the BER of the considered schemes degrades with ρ due to the increased amount of ICI and MAI. Interestingly, the proposed receiver provides similar results with either $N_i = 1$ or $N_i = 5$, meaning that it achieves convergence in a single iteration. Also, it largely outperforms the other methods. A possible explanation is that the CTHY scheme operates similarly to a linear multi-user detector where interference is mitigated at the price of non-negligible noise enhancement. As to the HL method, the windowing functions applied at the DFT output may lead to a significant loss of the signal energy in the presence of relatively large CFOs.

Case 2: Performance with different CFO attenuation factors

Next, we assess the performance of EMBR when the frequency and channel estimation tasks are coupled with the
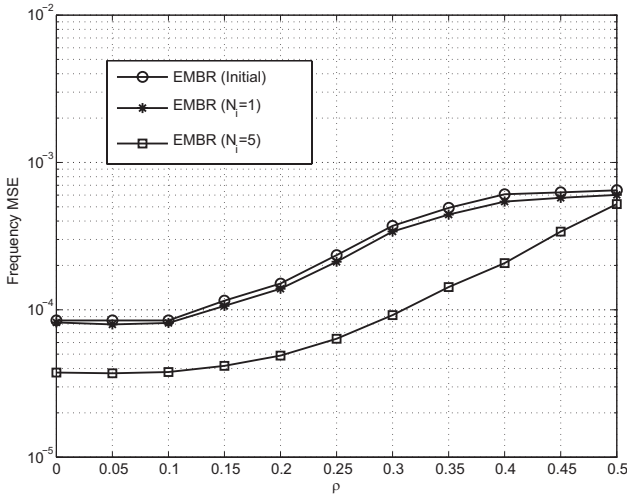
Fig. 3.   Accuracy of the frequency estimates vs. ρ for $E_b/N_0 = 20$ dB.
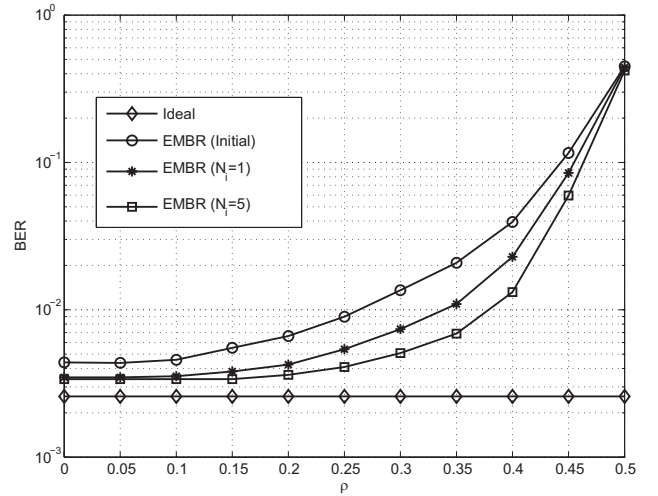
Fig. 5.   BER performance vs. ρ for uncoded QPSK and $E_b/N_0 = 20$ dB.
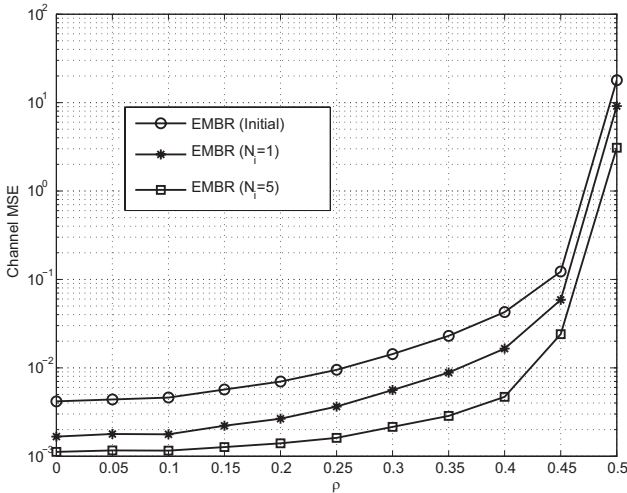
Fig. 4.   Accuracy of channel estimates vs. ρ for $E_b/N_0 = 20$ dB.
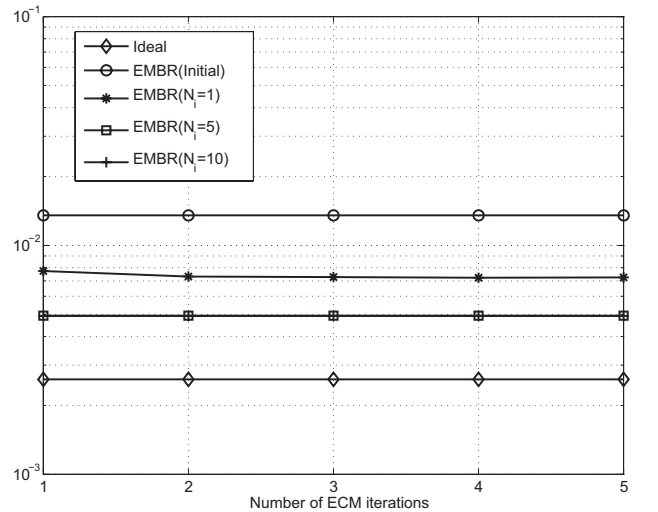
Fig. 6.   BER performance vs. the number C of ECM iterations for $E_b/N_0 = 20$ dB and $\rho = 0.3$.

decision making process as explained in Section III. In this case, the estimation accuracy at the end of the ith SAGE iteration is measured in terms of the following MSE indicators

$$\text{MSE}_\epsilon^{(i)} = E\left\{\left|\epsilon_1^{(i,C)} - \epsilon_1\right|^2\right\}, \qquad (31)$$

$$\text{MSE}_h^{(i)} = E\left\{\left\|\boldsymbol{h}_{1,\text{LS}}\left(\hat{\boldsymbol{\theta}}_1^{(i,C)}\right) - \boldsymbol{h}_1\right\|^2\right\}, \qquad (32)$$

with $C = 1$. Figs. 3 and 4 illustrate $\text{MSE}_\epsilon^{(i)}$ and $\text{MSE}_h^{(i)}$, respectively, as a function of ρ for $i = 1$ and 5, where users have equal power with $E_b/N_0 = 20$dB. For comparison, we show the accuracy of initial frequency and channel estimates provided by the data-aided schemes in [11] and [21]. Again, we see that the performance deteriorates as ρ increases. A considerable improvement is observed by passing from $i = 1$ to 5, especially in the frequency estimation accuracy. The corresponding BER results are illustrated in Fig. 5. Compared to initial data decisions, a non-negligible reduction of the error rate is achieved after the first SAGE iteration, while marginal gains are obtained with more iterations if $\rho \leq 0.2$. Recalling that CTYH is used for the initialization purpose, the BER

measured over the initial data decisions corresponds to the error rate of CTYH.

Case 3: Performance with different numbers of ECM iterations

Fig. 6 illustrates the impact of the number C of ECM iterations on the BER performance of EMBR when $\rho = 0.3$ and $E_b/N_0 = 20$dB. For comparison, we also show the reliability of the initial data decisions as well as the error rate of the ideal system that operates with frequency-synchronized users and perfect knowledge of channel responses. It turns out that the system performance is only marginally affected by parameter C. Hence, we can set $C = 1$ without incurring in significant BER degradations.

Case 4: BER performance for uncoded QPSK

Fig. 7 shows the BER of the considered schemes as a function of E_b/N_0, where users have equal power with $\rho = 0.3$ and the number of iterations is $N_i = 5$ with both EMBR and HL. Again, the best performance is achieved by the
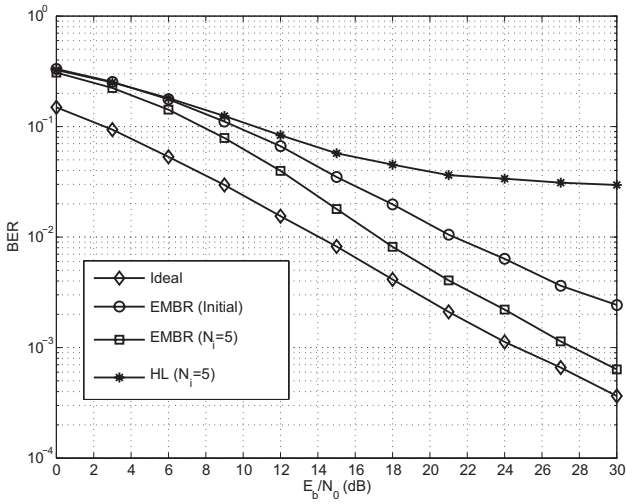
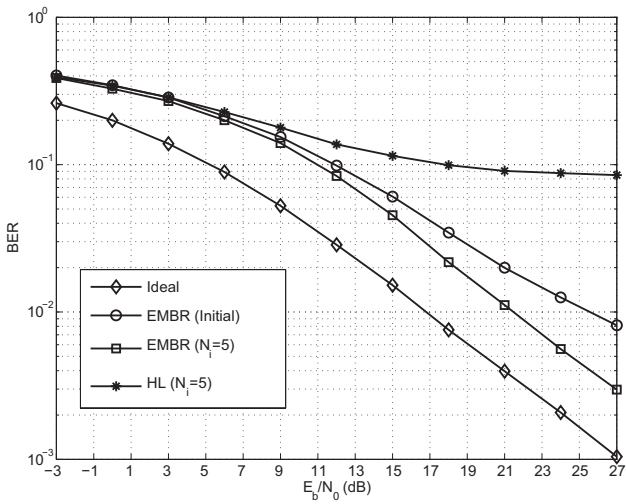Fig. 7. BER performance vs. E_b/N_0 for uncoded QPSK and $\rho = 0.3$.

Fig. 9. BER performance in the presence of a strong interferer for uncoded QPSK with $E_b/N_0 = 20$ dB and $\rho = 0.3$.

Fig. 8. BER performance vs. E_b/N_0 for uncoded 16-QAM and $\rho = 0.3$.

Fig. 10. BER performance vs. ϵ_2 for uncoded QPSK with $E_b/N_0 = 20$ dB.

proposed scheme. In particular, at an error rate of 10^{-2}, the gain over CTYH (initial data decisions) is approximately 4dB while a loss of 3dB is incurred with respect to the ideal system. As for HL, it performs poorly and exhibits an error floor at high SNRs.

Case 5: BER performance for uncoded 16-QAM

Fig. 8 illustrates BER results obtained in the same operating conditions as those for Fig. 7, except that symbols are now taken from a 16-QAM constellation. We see that the performance degradation of the considered schemes with respect to the ideal system is larger than that in Fig. 7. The reason is that the sensitivity to synchronization and/or channel estimation errors increases with the constellation size.

Case 6: Resistance to the near-far effect

In all practical systems, power control is employed to mitigate the near-far problem. However, power control cannot be assumed when a new user is entering the system as its
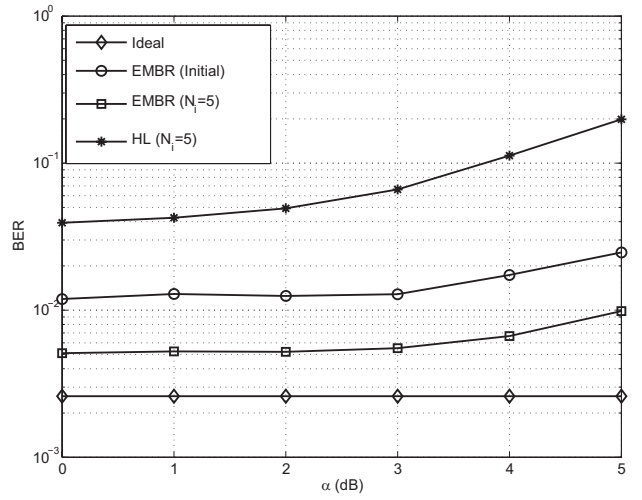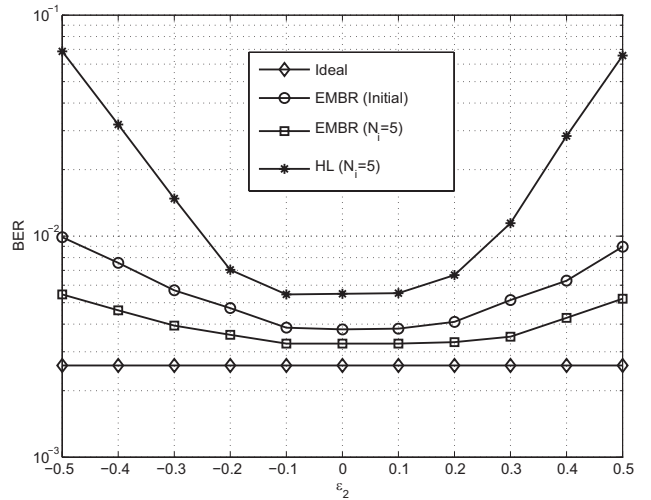
power level is still to be measured. Therefore, it is of interest to assess the performance of the considered schemes in the presence of a strong interferer. For this purpose, we consider a scenario in which the power of user #2 is larger than that of the others by a factor $\alpha \geq 1$. This can be easily achieved setting $\beta_2 = \sqrt{\alpha} \cdot \beta_1$ in (30), while keeping $\beta_k = \beta_1$ for $k = 3, 4$. Simulation results illustrating the BER of user #1 are shown in Fig. 9 as a function of α (expressed in dB) for $\rho = 0.3$ and $E_b/N_0 = 20$dB. As expected, the system performance degrades with α. In particular, the BER of EMBR and CTYH increases by a factor of two when α passes from 0 to 5dB, while larger degradations occur with HL.

Case 7: Performance in the presence of an interferer with a large CFO

We assess the performance of the considered schemes when the CFO of a given user is significantly larger than others. In particular, we assume $\epsilon = [0.1, \epsilon_2, 0.1, -0.1]^T$, with ϵ_2 varying in the interval $[-0.5, 0.5]$. Fig. 10 shows the BER of

$$
\begin{aligned}
Q_1\left(\tilde{\boldsymbol{\theta}}_m \,\Big|\, \hat{\boldsymbol{\theta}}_m^{(i,c)}\right) \;=\;& 2\int_\Omega \Re\left\{\hat{\boldsymbol{y}}_m^{(i)H}\boldsymbol{\Gamma}(\tilde{\epsilon}_m)\boldsymbol{F}^H\boldsymbol{D}(\tilde{\boldsymbol{s}}_m)\boldsymbol{W}\boldsymbol{h}_m\right\}e^{-\frac{1}{\sigma_\eta^2(i)}\left\|\hat{\boldsymbol{y}}_m^{(i)}-\tilde{\boldsymbol{z}}_m\right\|^2-\boldsymbol{h}_m^H\boldsymbol{C}_m^{-1}\boldsymbol{h}_m}\,d\boldsymbol{h}_m \\
&-\int_\Omega \operatorname{tr}\left\{\boldsymbol{D}(\tilde{\boldsymbol{s}}_m)\boldsymbol{W}\boldsymbol{h}_m\boldsymbol{h}_m^H\boldsymbol{W}^H\boldsymbol{D}^H(\tilde{\boldsymbol{s}}_m)\right\}e^{-\frac{1}{\sigma_\eta^2(i)}\left\|\hat{\boldsymbol{y}}_m^{(i)}-\tilde{\boldsymbol{z}}_m\right\|^2-\boldsymbol{h}_m^H\boldsymbol{C}_m^{-1}\boldsymbol{h}_m}\,d\boldsymbol{h}_m,
\end{aligned}
\tag{A5}
$$

$$
\begin{aligned}
Q_2\left(\tilde{\boldsymbol{\theta}}_m \,\Big|\, \hat{\boldsymbol{\theta}}_m^{(i,c)}\right) \;=\;& 2\Re\left\{\hat{\boldsymbol{y}}_m^{(i)H}\boldsymbol{\Gamma}(\tilde{\epsilon}_m)\boldsymbol{F}^H\boldsymbol{D}(\tilde{\boldsymbol{s}}_m)\boldsymbol{W}\hat{\boldsymbol{h}}_{m,\mathrm{MMSE}}\right\} \\
&-\operatorname{tr}\left\{\boldsymbol{D}\left(\tilde{\boldsymbol{s}}_m\right)\boldsymbol{W}\left[\sigma_\eta^2(i)\boldsymbol{P}^{-1}\left(\hat{\boldsymbol{s}}_m^{(i,c)}\right)+\hat{\boldsymbol{h}}_{m,\mathrm{MMSE}}\hat{\boldsymbol{h}}_{m,\mathrm{MMSE}}^H\right]\boldsymbol{W}^H\boldsymbol{D}^H\left(\tilde{\boldsymbol{s}}_m\right)\right\}.
\end{aligned}
\tag{A7}
$$

user #1 as a function of ϵ_2 for $E_b/N_0 = 20\mathrm{dB}$ and equal-powered users. It is intuitively clear that the MAI produced by user #2 increases with ϵ_2 and this inevitably leads to some performance degradation. However, the results in Fig. 10 indicate that both EMBR and CTYH can tolerate relatively large values of ϵ_2, while the BER of HL rapidly deteriorates with ϵ_2.

V. CONCLUSIONS

An iterative receiver architecture for OFDMA uplink transmissions was proposed in this work, where EM-type algorithms are employed to perform frequency synchronization, channel estimation and data detection in a joint fashion. The interference induced by residual CFOs is mitigated directly at the BS without the need of returning CFO estimates back to mobile units to achieve frequency synchronization. This significantly reduces the system overhead and solves the problem of outdated information that occurs in rapidly-varying channels due to feedback delay. Simulations indicated that the proposed receiver is robust against strong-powered interfering users and outperforms other existing solutions in the presence of large CFO values. Although these advantages are achieved at the price of an increased computational load, the overall complexity seems compatible with the technological constraints of future wireless communications systems.

APPENDIX

In this appendix we highlight the major steps leading to the computation of the Q-function defined in (12), which can be rewritten as

$$
\begin{aligned}
Q\left(\tilde{\boldsymbol{\theta}}_m \,\Big|\, \hat{\boldsymbol{\theta}}_m^{(i,c)}\right) =& \int_\Omega \ln\left[p\left(\hat{\boldsymbol{y}}_m^{(i)}\,\Big|\,\boldsymbol{h}_m,\tilde{\boldsymbol{\theta}}_m\right)\right]\cdot \\
& p\left(\hat{\boldsymbol{y}}_m^{(i)}\,\Big|\,\boldsymbol{h}_m,\hat{\boldsymbol{\theta}}_m^{(i,c)}\right)p\left(\boldsymbol{h}_m\right)\,d\boldsymbol{h}_m,
\end{aligned}
\tag{A1}
$$

where $\tilde{\boldsymbol{\theta}}_m = \left[\begin{array}{cc}\tilde{\boldsymbol{s}}_m^T & \tilde{\epsilon}_m\end{array}\right]^T$, $\hat{\boldsymbol{\theta}}_m^{(i,c)} = \left[\begin{array}{cc}\hat{\boldsymbol{s}}_m^{(i,c)T} & \hat{\epsilon}_m^{(i,c)}\end{array}\right]^T$ and $p\left(\boldsymbol{h}_m\right)$ is the a-priori pdf of $\boldsymbol{h}_m$. The major assumptions are that $i)$ $\boldsymbol{h}_m$ is a circularly symmetric Gaussian vector with zero mean (Rayleigh fading) and covariance matrix $\boldsymbol{C}_m = E\left\{\boldsymbol{h}_m\boldsymbol{h}_m^H\right\}$; $ii)$ the disturbance $\boldsymbol{\eta}_m^{(i)}$ in (10) is nearly Gaussian with zero-mean and covariance matrix $\sigma_\eta^2(i)\boldsymbol{I}_N$. Thus, we may write

$$
p\left(\boldsymbol{h}_m\right) = \frac{1}{\pi^{L_h}\det(\boldsymbol{C}_m)}\exp\left\{-\boldsymbol{h}_m^H\boldsymbol{C}_m^{-1}\boldsymbol{h}_m\right\}
\tag{A2}
$$

$$
\begin{aligned}
&p\left(\hat{\boldsymbol{y}}_m^{(i)}\,\Big|\,\boldsymbol{h}_m,\hat{\boldsymbol{\theta}}_m^{(i,c)}\right) \approx \\
&\qquad\frac{1}{\left(\pi\sigma_\eta^2(i)\right)^N}\exp\left\{-\frac{1}{\sigma_\eta^2(i)}\left\|\hat{\boldsymbol{y}}_m^{(i)}-\hat{\boldsymbol{z}}_m^{(i,c)}\right\|^2\right\},
\end{aligned}
\tag{A3}
$$

$$
\begin{aligned}
&\ln\left[p\left(\hat{\boldsymbol{y}}_m^{(i)}\,\Big|\,\boldsymbol{h}_m,\tilde{\boldsymbol{\theta}}_m\right)\right] \approx \\
&\qquad -N\ln\left(\pi\sigma_\eta^2(i)\right)-\frac{1}{\sigma_\eta^2(i)}\left\|\hat{\boldsymbol{y}}_m^{(i)}-\tilde{\boldsymbol{z}}_m\right\|^2,
\end{aligned}
\tag{A4}
$$

with $\hat{\boldsymbol{z}}_m^{(i,c)} = \boldsymbol{\Gamma}(\hat{\epsilon}_m^{(i,c)})\boldsymbol{F}^H\boldsymbol{D}(\hat{\boldsymbol{s}}_m^{(i,c)})\boldsymbol{W}\boldsymbol{h}_m$ and $\tilde{\boldsymbol{z}}_m = \boldsymbol{\Gamma}(\tilde{\epsilon}_m)\boldsymbol{F}^H\boldsymbol{D}(\tilde{\boldsymbol{s}}_m)\boldsymbol{W}\boldsymbol{h}_m$. Substituting (A2)-(A4) into (A1) and skipping the additive and multiplicative terms independent of $\tilde{\boldsymbol{\theta}}_m$ yields (A5), where we have used the identity $\|\tilde{\boldsymbol{z}}_m\|^2 = \operatorname{tr}\left\{\tilde{\boldsymbol{z}}_m\tilde{\boldsymbol{z}}_m^H\right\}$.

To proceed further, we observe that

$$
\begin{aligned}
&\frac{1}{\sigma_\eta^2(i)}\left\|\hat{\boldsymbol{y}}_m^{(i)}-\tilde{\boldsymbol{z}}_m\right\|^2+\boldsymbol{h}_m^H\boldsymbol{C}_m^{-1}\boldsymbol{h}_m = \\
&\frac{1}{\sigma_\eta^2(i)}\left\{\left[\boldsymbol{h}_m-\hat{\boldsymbol{h}}_{m,\mathrm{MMSE}}\right]^H\boldsymbol{P}\left(\hat{\boldsymbol{s}}_m^{(i,c)}\right)\left[\boldsymbol{h}_m-\hat{\boldsymbol{h}}_{m,\mathrm{MMSE}}\right]\right. \\
&\left.+\left\|\hat{\boldsymbol{y}}_m^{(i)}\right\|^2-\hat{\boldsymbol{h}}_{m,\mathrm{MMSE}}\hat{\boldsymbol{h}}_{m,\mathrm{MMSE}}^H\right\},
\end{aligned}
\tag{A6}
$$

where $\hat{\boldsymbol{h}}_{m,\mathrm{MMSE}}$ and $\boldsymbol{P}\left(\hat{\boldsymbol{s}}_m^{(i,c)}\right)$ are defined in (14) and (15), respectively. Note that the functional dependence of $\hat{\boldsymbol{h}}_{m,\mathrm{MMSE}}$ on $\hat{\boldsymbol{\theta}}_m^{(i,c)}$ in (A6) has been neglected for notational simplicity. By substituting (A6) into (A5), integrating over Ω and neglecting irrelevant multiplicative terms, we are led to (A7).

Finally, by subtracting the term $\left\|\hat{\boldsymbol{y}}_m^{(i)}\right\|^2$ from the RHS of (A7) and recalling that $\boldsymbol{F}^H\boldsymbol{F} = \boldsymbol{\Gamma}^H\left(\tilde{\epsilon}_m\right)\boldsymbol{\Gamma}\left(\tilde{\epsilon}_m\right) = \boldsymbol{I}_N$, we see that $Q_2\left(\tilde{\boldsymbol{\theta}}_m\,\Big|\,\hat{\boldsymbol{\theta}}_m^{(i,c)}\right)$ is equivalent to $\Phi\left(\tilde{\boldsymbol{\theta}}_m\,\Big|\,\hat{\boldsymbol{\theta}}_m^{(i,c)}\right)$ in (13).

REFERENCES

[1] H. Yaghoobi, "Scalable OFDMA physical layer in IEEE 802.16 WirelessMAN," *Intel Technology J.*, vol. 8, no. 3, pp. 201–212, Aug. 2004.

[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, no. 2-4, pp. 887–895, Feb.-Apr. 1995.

[3] H. Sari and G. Karam, "Orthogonal frequency-division multiple access and its application to CATV networks," *European Trans. Telecommun.*, vol. 9, pp. 507–516, Dec. 1998.

[4] P. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Trans. Commun.*, vol. 42, no. 10, pp. 2908–2914, Oct. 1994.

[5] J. van de Beek, M. Sandell, and P. Borjesson, "ML estimation of timing and frequency offset in OFDM systems," *IEEE Trans. Signal Processing*, vol. 45, no. 7, pp. 1800–1805, July 1997.

[6] T. Schmidl and D. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Trans. Commun.*, vol. 45, no. 12, pp. 1613–1621, Dec. 1997.

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

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

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

[10] Z. Cao, U. Tureli, and Y. D. Yao, "Efficient structure-based carrier frequency offset estimation for interleaved OFDMA uplink," in *Proc. IEEE ICC*, May 2003, pp. 3361–3365.

[11] M. Pun, S. Tsai, and C.-C. J. Kuo, "Joint maximum likelihood estimation of carrier frequency offset and channel for uplink OFDMA systems," in *Proc. IEEE Globecom*, Nov. 2004, pp. 3748–3752.

[12] M. Pun, M. Morelli, and C.-C. J. Kuo, "Joint synchronization and channel estimation in uplink OFDMA systems," in *Proc. IEEE ICASSP*, Mar. 2005, pp. 857–860.

[13] Z. Cao, U. Tureli, Y. D. Yao, and P. Honan, "Frequency synchronization for generalized OFDMA uplink," in *Proc. IEEE Globecom*, Nov. 2004, pp. 1071–1075.

[14] J. Choi, C. Lee, H. Jung, and Y. Lee, "Carrier frquency offset compensation for uplink of OFDM-FDMA systems," *IEEE Commun. Lett.*, vol. 4, no. 12, pp. 414–416, Dec. 2000.

[15] A. Tonello, "Multiuser detection and turbo multiuser decoding for asynchronous multitone multiple access systems," in *Proc. IEEE Veh. Technol. Conf.*, Sep. 2002, pp. 970–974.

[16] D. Huang and K. Letaief, "An interference-cancellation scheme for carrier frequency offsets correction in OFDMA systems," *IEEE Trans. Commun.*, vol. 53, no. 7, pp. 1155–1165, July 2005.

[17] J. Fessler and A. Hero, "Space-alternating generalized expectation-maximization algorithm," *IEEE Trans. Signal Processing*, vol. 42, no. 10, pp. 2664–2677, Oct. 1994.

[18] G. J. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*. New York: John Wiley & Sons, Inc., 1997.

[19] B. Haindl, M. Sajatovic, C. Rihacek, J. Prinz, M. Schnell, E. Haas, and I. Cosovic, "B-VHF: A multi-carrier based broadband VHF communications concept for air traffic management," in *Proc. IEEE Aerospace Conf.*, Mar. 2005, pp. 1894-1904.

[20] M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 4, pp. 477–489, Apr. 1988.

[21] M. Morelli and U. Mengali, "A comparison of pilot-aided channel estimation methods for OFDM systems," *IEEE Trans. Signal Processing*, vol. 49, no. 12, pp. 3065–3073, Dec. 2001.

Man-On Pun (S'05) received the BEng. (Hon.) degree in electronic engineering from the Chinese University of Hong Kong in 1996, the MEng. degree in information sciences from University of Tsukuba, Japan in 1999 and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, in 2006, respectively.

From 1999 to 2001, he was with the Sony Corporation in Tokyo, Japan. His current research interests are in the area of statistical signal processing for wireless communications.

He is a recipient of the Japanese Government (Monbusho) Scholarship, the Sir Edward Youde Memorial fellowship for Overseas Studies and the Croucher postdoctoral fellowship.

Michele Morelli (M'04) received the Laurea (cum laude) degree in electrical engineering and the "Premio di Laurea SIP" from the University of Pisa, Pisa, Italy, in 1991 and 1992, respectively. From 1992 to 1995, he was with the Department of Information Engineering, University of Pisa, where he received the Ph.D. degree in electrical engineering. In September 1996, he joined the Centro Studi Metodi e Dispositivi per Radiotrasmissioni (CSMDR) of the Italian National Research Council (CNR) in Pisa, where he held the position of Research Assistant.

Since 2001, he has been with the Department of Information Engineering, University of Pisa, where he is currently an Associate Professor of Telecommunications. His research interests are in wireless communication theory, with emphasis on synchronization algorithms and channel estimation in multiple-access communication 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 Department of Electrical Engineering-Systems and the Signal and Image Processing Institute at the University of Southern California, where he currently has a joint appointment as Professor of Electrical Engineering and Mathematics. His research interests are in the areas of digital signal and image processing, audio and video coding, multimedia communication technologies and delivery protocols, and embedded system design. Dr. Kuo has guided about 75 students to their Ph.D. degrees and supervised 15 postdoctoral research fellows. He is co-author of seven books and more than 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.