

Code-Aided ML Joint Synchronization and Channel Estimation for Downlink MC-CDMA

Mamoun Guenach, *Member, IEEE*, Henk Wymeersch, *Member, IEEE*, Heidi Steendam, *Member, IEEE*, and Marc Moeneclaey, *Fellow, IEEE*

Abstract—In this paper, we present a novel code-aided joint synchronization and channel estimation algorithm for downlink multicarrier code-division multiple access. The expectation-maximization algorithm is used to locate the maximum-likelihood estimate of the channel impulse response, propagation delay, and carrier frequency offset. The estimator accepts soft information from the decoder in the form of *a posteriori* probabilities of the coded symbols, and can be interpreted as performing joint estimation and data detection. The performance of the proposed algorithm is verified through computer simulations. Impressive performance gains are visible as compared with a conventional data-aided estimation scheme.

Index Terms—Code-aided estimation, estimation, synchronization.

I. INTRODUCTION

ORTHOGONAL frequency-division multiplexing (OFDM) has received intense interest from the research community during the past few decades. Its robustness to frequency-selective channels has made it one of the main candidates for high data rate transmission for current and next-generation wireless and wireline applications [1], [2].

Supporting multiple users can be achieved in a variety of ways. Popular multiple-access (MA) methods include OFDMA (where active users are assigned different subcarriers) and code-division multiple-access (CDMA)-based schemes. In CDMA, different users are distinguished based on unique spreading codes. Spreading can take place either in the frequency-domain or in the time-domain. This leads to concepts known as multicarrier-CDMA (MC-CDMA) and multicarrier direct-sequence CDMA (MC-DS-CDMA), respectively. These techniques each have different benefits and drawbacks, depending on the intended application [3], [4]. In this paper, we focus on MC-CDMA.

All OFDM-based transmission systems suffer from several sources of impairment. The first problem is related to the inherent high peak-to-average power ratio (PAPR). Saturation of

amplifiers leads to out-of-band radiation and destruction of orthogonality between carriers [5]. A second problem is synchronization [6]: accurate timing and frequency synchronization is required to maintain orthogonality among subcarriers and to avoid intersymbol interference (ISI). Additionally, the channel impulse response (CIR) must be known to coherently detect the data per subcarrier. When we consider the massive amount of research devoted to these problems (see, e.g., [6]–[15] and references therein), it becomes clear that synchronization and channel estimation are critical issues. Conventional techniques are either data-aided (i.e., exploiting training symbols in the time- or frequency-domain) [7]–[12] or blind (e.g., exploiting the presence of the cyclic prefix) [13]–[15].

With the advent of powerful error-correcting codes [including turbo- and low-density parity-check (LDPC) codes], these conventional techniques cannot always be applied successfully. Powerful codes lead to a combination of low bit-error rate (BER) at low signal-to-noise ratio (SNR), thus rendering blind techniques unreliable. Similarly, data-aided algorithms require an unreasonable amount of power and bandwidth to be devoted to training. This has spurred several research groups to consider “code-aided” or “code-aware” estimation algorithms. These algorithms iterate between data detection and estimation, thus improving both the estimates of synchronization parameters and CIR, while simultaneously performing increasingly reliable data detection. Such techniques are often inspired by the turbo-principle [16] or the expectation-maximization (EM) algorithm [17], [18]. In [19], an EM-based semi-blind technique is described that performs code-aided estimation of the CIR per multicarrier (MC) symbol. The same idea was applied in the frequency-domain in [20] for a multiantenna, multiuser system. The EM algorithm was again considered in [21] for estimation of the CIR for a time-varying multiple-input–multiple-output (MIMO)-OFDM scenario. Finally, [22] proposes an ad hoc code-aided channel estimator for time-varying OFDM systems. The reader will note that code-aided estimation of synchronization parameters has received little interest.

In the current paper, we extend our work from [23] and [24] to the problem of joint estimation of the propagation delay, carrier frequency offset (CFO), and CIR for downlink MC-CDMA. Starting from the maximum-likelihood (ML) principle, we derive an estimation algorithm based on the EM algorithm, exploiting information from the pilot symbols and coded data symbols in a systematic fashion.

This paper is organized as follows: the system model is provided in Section II, including a brief description of the detector. We continue with ML and EM estimation algorithms in Section III, paying special attention to issues related to

Manuscript received April 1, 2005; revised October 1, 2005. This work was supported in part by the Interuniversity Attraction Poles Program P5/11-Belgian Science Policy.

M. Guenach, H. Steendam, and M. Moeneclaey are with the Department of Telecommunications and Information Processing, Ghent University, B-9000 Ghent, Belgium (e-mail: guenach@telin.UGent.be; hs@telin.UGent.be; mm@telin.UGent.be).

H. Wymeersch is with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Boston, MA 02139 USA (e-mail: hwymeersch@mit.edu).

Digital Object Identifier 10.1109/JSAC.2005.864029

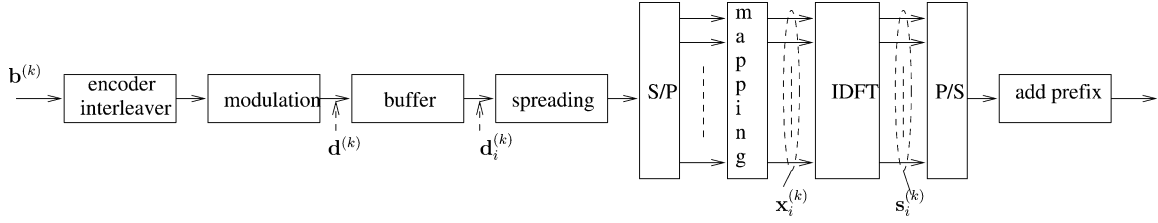


Fig. 1. Frame construction for the k th user.

computational complexity. Numerical results are provided in Section IV, before we end with conclusions in Section V.

II. MC-CDMA DOWNLINK SYSTEM MODEL

A. Transmitter

We consider the downlink of a MC-CDMA system with K_u active users. The base station transmits frames consisting of M_s MC symbols to each of the users. A frame intended for the k th user is generated as follows (see Fig. 1): a block $\mathbf{b}^{(k)} = [b_0^{(k)}, \dots, b_{N_b-1}^{(k)}]^T$ of N_b information-bits is encoded, resulting in N_c coded bits. After interleaving, these N_c coded bits are mapped to a sequence of N_d symbols, belonging to a unit-energy 2^q -point constellation Ω (with $N_d = N_c/q$). We will denote this sequence $\mathbf{d}^{(k)} = [d_0^{(k)}, \dots, d_{N_d-1}^{(k)}]^T$. Now, $\mathbf{d}^{(k)}$ is broken down into M_s blocks of length P (with $P = N_d/M_s$). The M_s blocks are buffered and converted, one at a time, to MC symbols. Let us denote the symbols in the i th MC symbol of the k th user as $\mathbf{d}_i^{(k)} = [d_{i,0}^{(k)}, \dots, d_{i,P-1}^{(k)}]^T$.

The conversion of $\mathbf{d}_i^{(k)}$ into a MC symbol is performed as follows. The p th data symbol (i.e., $d_{i,p}^{(k)}$) is first spread by a unit-energy spreading sequence $\mathbf{a}_p^{(k)}$ of length N_s , yielding a sequence of N_s chips $d_{i,p}^{(k)} \times [a_{p,0}^{(k)}, \dots, a_{p,N_s-1}^{(k)}]^T$. Performing spreading for all P symbols in the i th MC symbol results in a total of $N_s P$ chips. This chip-sequence is mapped to $N \doteq N_s P$ subcarriers (known as frequency interleaving). Hence, each of the $N_s P$ chips (say, $d_{i,p}^{(k)} a_{p,s}^{(k)}$) is assigned to a unique subcarrier (say the $n_{p,s}$ th subcarrier).¹ We end up with the block $\mathbf{x}_i^{(k)} = [x_{i,0}^{(k)}, \dots, x_{i,N-1}^{(k)}]^T$ of N interleaved chips which is provided to an N -point inverse discrete Fourier transform (IDFT).² A ν -point cyclic prefix (CP) is preappended, resulting in $N + \nu$ time-domain samples $[s_{i,-\nu}^{(k)}, \dots, s_{i,-1}^{(k)}, s_{i,0}^{(k)}, \dots, s_{i,N-1}^{(k)}]^T$, where $s_{i,l}^{(k)} = s_{i,l+N}^{(k)}$, for $l = -\nu, \dots, -1$. Hence, the length of the MC symbol is equal to $N_T = N + \nu$. We can write the m th time-domain sample ($m = -\nu, \dots, N - 1$) of the i th MC symbol of the k th user as

$$s_{i,m}^{(k)} = \sqrt{\frac{E_s}{N + \nu}} \sum_{p=0}^{P-1} \sum_{s=0}^{N_s-1} d_{i,p}^{(k)} a_{p,s}^{(k)} e^{j2\pi n_{p,s}^{(k)} m / N} \quad (1)$$

¹For notational convenience, we will later assume this mapping to be the same for all users. Extension to more general schemes is straightforward.

²Obviously, the DFT and IDFT operations are implemented through a fast Fourier transform.

where E_s denotes the energy per symbol $d_{i,p}^{(k)}$. We further define the MC symbol period T and the sampling period $T_s = T/N_T$. Finally, the signals of the K_u different users are added, shaped with a normalized transmit filter and transmitted over the channel to the mobile stations.

B. Multiplexing

A common control physical channel (CCPCH) is assumed whereby a sequence of M_p MC symbols is time-multiplexed at the header of the composite dedicated physical channels (DPCH). These symbols do not undergo spreading (i.e., $N_s = 1$ in the CCPCH). A fraction (say, $d \in [0, 1]$) of the subcarriers of the MC symbols in the CCPCH is devoted to training symbols, which serve to synchronize the mobile receivers. The remainder of the carriers (i.e., the fraction $1 - d$) is taken up with administrative data. For convenience, we will refer to these M_p MC symbols as ‘‘pilot MC symbols.’’ The pilot symbols may have a repetitive structure in the time-domain, to accommodate specific synchronization algorithms [8]. The frame structure is shown in Fig. 2.

C. Receiver

Each user has to process a total of $M = M_s + M_p$ MC symbols. From this point on, we focus on a specific mobile station, say the k' th. The signal from the base station propagates to the k' th mobile station through a channel with overall CIR $h_{\text{ch}}(t)$. This CIR incorporates the transmit filter, physical propagation channel, and receive filter (e.g., matched filter). We assume a quasi-static block-fading channel that remains constant during each frame but can vary independently from frame to frame. The CIR is assumed to have a delay spread no greater than LT_s : $h_{\text{ch}}(t) = 0$ for $t < 0$ and for $t > LT_s$. Additionally, the signal arrives at the mobile station with a certain delay τ and is affected by a CFO f_o and is corrupted by thermal noise. The propagation delay τ belongs to some interval³ $[0, \tau_{\text{max}}]$, while the CFO depends on the speed of the mobile stations and any possible mismatch between the transmit and receive oscillators. We assume the CFO to be small, compared with the bandwidth of the receiver’s matched filter. Hence, the received signal is given by

$$r(t) = s_C(t) + \sum_{k=1}^{K_u} s_D^{(k)}(t) + w(t) \quad (2)$$

where $w(t)$ is the baseband representation of the additive white Gaussian noise (AWGN) with power spectral density $N_0/2$ per

³As discussed in [7], the maximum propagation delay τ_{max} can be directly related to the cell radius R : $\tau_{\text{max}} = R/c$, where c is the speed of light.

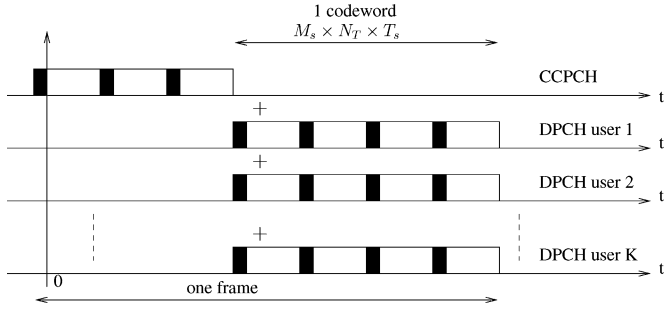


Fig. 2. Frame structure: the cyclic prefix is marked in black.

real dimension. In (2), the signal corresponding to the CCPCH is given by

$$s_C(t) = \sum_{i=0}^{M_p-1} \sum_{m=-\nu}^{N-1} s_{i,m} h_{\text{ch}}(t - mT_s - iT - \tau) e^{j2\pi f_o t} \quad (3)$$

while the DPCH signal corresponding to the k th user is

$$s_D^{(k)}(t) = \sum_{i=0}^{M_s-1} \sum_{m=-\nu}^{N-1} s_{i,m}^{(k)} h_{\text{ch}}(t - mT_s - (i + M_p)T - \tau) e^{j2\pi f_o t}. \quad (4)$$

The receiver is fully digital and samples the received signal $r(t)$ at a rate $1/T_s$ resulting in a sequence of samples $r(lT_s)$, $l = -\nu, \dots, +\infty$. Let us define the normalized CFO $F = f_o T_s N$. Following [7], we break up τ as

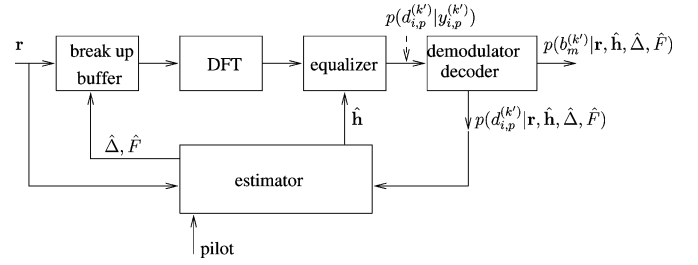
$$\tau = \Delta T_s - \delta T_s \quad (5)$$

with $\Delta \in \{0, 1, \dots, \Delta_{\max} \doteq \lceil \tau_{\max}/T_s \rceil\}$ and $\delta \in [0, 1[$. Defining $h(t) = h_{\text{ch}}(t + \delta T_s)$, we can express the samples $r(lT_s)$ as a function of $h(t)$. For instance, the contribution to $r(lT_s)$ by the k th DPCH is given by (6) shown at the bottom of the page. Since $T = N_T T_s$ is a multiple of T_s , the channel is fully characterized by the vector

$$\mathbf{h} = [h(0), h(T_s), \dots, h((L-1)T_s)]^T. \quad (7)$$

When $\nu \geq L$, no interference between successive MC symbols occurs: for each MC symbol, at least one sequence of N successive time-domain samples exists that does not suffer from interference of the preceding and the following MC symbol.⁴ As $\Delta \in \{0, 1, \dots, \Delta_{\max}\}$, samples taken prior to $t = -\nu T_s$ and later than $t = ((M_s + M_p)T + (\Delta_{\max} + L - \nu - 2)T_s)$

⁴When multiple such sequences exist, after conversion to the frequency-domain, they lead to equivalent DFT-outputs, up to a known complex rotation (which in no way affects the performance).


 Fig. 3. Detector and estimator for the k' th user.

do not depend on the current frame. Hence, the following sequence of $N_T (M_s + M_p) + \Delta_{\max} + L - 1$ time-domain samples is sufficient for data detection and synchronization

$$\mathbf{r} = [r(-\nu T_s), \dots, r((M_s + M_p)T + (\Delta_{\max} + L - \nu - 2)T_s)]^T. \quad (8)$$

The final goal of the k' th user is to recover the transmitted information bits $\mathbf{b}^{(k')}$. It is clear that this requires knowledge of both the delay shift Δ , the (normalized) CFO F and the channel taps \mathbf{h} .

The conceptual block diagram of the receiver is shown in Fig. 3. The receiver of user k' employs an estimation algorithm (e.g., by exploiting the M_p pilot MC symbols) to obtain estimates $(\hat{\Delta}, \hat{F}, \hat{\mathbf{h}})$ for the timing offset Δ , the CFO F , and the channel taps \mathbf{h} . These estimates are then used to recover the symbols $d_{i,p}^{(k')}$ contained in the M_s data MC symbols. To do this, the receiver of user k' synchronizes the sequence \mathbf{r} by shifting the sequence over $\hat{\Delta}$ samples, and correcting the phase rotation due to the CFO using the estimate \hat{F} . Then, the remaining vector is cut into blocks of length N : for each of the M_s data MC symbols of $N + \nu$ samples, the receiver disregards the ν samples corresponding to the CP, and selects the remaining N samples for further processing. The M_s blocks are converted to the frequency-domain through a N -point DFT operation. The output of the DFT at subcarrier $n_{p,s}$ for the i th MC symbol is given by (9) shown at the bottom of the page. By equalization and despreading of the samples $\{z_{i,p,s}\}_{s=0, \dots, N_s-1}$, we create a decision variable for $d_{i,p}^{(k')}$

$$y_{i,p}^{(k')} = \sum_{s=0}^{N_s-1} \alpha_{p,s} z_{i,p,s} \left(a_{p,s}^{(k')} \right)^* \quad (10)$$

$$= d_{i,p}^{(k')} \mu_{i,p}^{(k')} + v_{i,p}^{(k')} \quad (11)$$

where $\alpha_{p,s}$ is the coefficient of the one-tap equalizer at subcarrier $n_{p,s}$; the equalizer coefficients $\{\alpha_{p,s}\}$ are obtained using

$$s_D^{(k)}(lT_s) = \sum_{i=0}^{M_s-1} \sum_{m=-\nu}^{N-1} s_{i,m}^{(k)} h(lT_s - mT_s - (i + M_p)N_T T_s - \Delta T_s) e^{j2\pi F l / N} \quad (6)$$

$$z_{i,p,s} = \frac{1}{\sqrt{N}} \sum_{l=0}^{N-1} r(\hat{\Delta} T_s + lT_s + (i + M_p)T) e^{j2\pi l n_{p,s} / N} e^{-j2\pi \hat{F}(\hat{\Delta} + l + (i + M_p)N_T)}. \quad (9)$$

the estimates $\hat{\mathbf{h}}$ of the channel taps and are selected, e.g., according to the minimum mean-square error (MMSE) criterion. The decision variable $y_{i,p}^{(k')}$ can be decomposed into a useful contribution $d_{i,p}^{(k')}$, $\mu_{i,p}^{(k')}$, and a contribution $v_{i,p}^{(k')}$ that embeds the noise and interference. Both the multiplicative factor $\mu_{i,p}^{(k')}$ and the contribution $v_{i,p}^{(k')}$ depend on the estimates of the timing offset Δ , the CFO F and the channel taps \mathbf{h} . From (11), the distribution $p\left(y_{i,p}^{(k')} \middle| d_{i,p}^{(k')}\right)$ of $y_{i,p}^{(k')}$, conditioned on $d_{i,p}^{(k')}$ can be obtained.⁵ Taking into account that $p\left(d_{i,p}^{(k')} \middle| y_{i,p}^{(k')}\right) = p\left(d_{i,p}^{(k')}\right) p\left(y_{i,p}^{(k')} \middle| d_{i,p}^{(k')}\right) / p\left(y_{i,p}^{(k')}\right)$, the receiver is able to compute the probabilities $p\left(d_{i,p}^{(k')} \middle| y_{i,p}^{(k')}\right)$.

Finally, the probabilities $p\left(d_{i,p}^{(k')} \middle| y_{i,p}^{(k')}\right)$, $i = 0, \dots, M_s - 1$, $p = 0, \dots, P - 1$, are fed to a block that performs soft demodulation and decoding, operating according to the turbo principle [27], [28]. For our purposes, this block is considered as a black box that computes *a posteriori* probabilities (APPs), possibly in an iterative fashion. Upon completion, the APPs of $b_m^{(k')}$ (i.e., $p\left(b_m^{(k')} \middle| \mathbf{r}, \hat{\mathbf{h}}, \hat{\Delta}, \hat{F}\right)$) $m = 0, \dots, N_b - 1$ are used to make final decisions on the information bits. At the same time, this block also computes the APPs of the coded symbols $p\left(d_{i,p}^{(k')} \middle| \mathbf{r}, \hat{\mathbf{h}}, \hat{\Delta}, \hat{F}\right)$. For additional details, the reader is referred to [23] and [24].

III. CHANNEL ESTIMATION AND SYNCHRONIZATION

It is clear that the detector described in the previous section requires the estimates of the delay shift⁶ Δ , the CFO F , and the channel taps \mathbf{h} . In this section, we describe how these quantities may be estimated by the k' th user. We start from the ML criterion and derive a data-aided (DA) estimator. Then, capitalizing on the EM algorithm, a code-aided estimator is constructed that exploits information from both the pilot MC symbols and the data MC symbols.

A. ML Estimation

We first write our observation into a more convenient form. We start again from our observation-vector $\mathbf{r} = [r(-\nu T_s), \dots, r((M_s + M_p)T + (\Delta_{\max} + L - \nu - 2)T_s)]^T$ from (8). Note that the length of this vector is independent of Δ .

We now introduce row-vectors of length N_T : $\mathbf{s}_C^i = [s_{i,-\nu}, \dots, s_{i,N-1}]$ consists of the $N_T = N + \nu$ time-domain samples of the i th pilot MC symbol ($i = 0, \dots, M_p - 1$). Similarly, $\mathbf{s}_D^i = [s_{i,-\nu}^{(k')}, \dots, s_{i,N-1}^{(k')}]$ consists of the $N_T = N + \nu$ time-domain samples of the i th data MC symbol ($i = 0, \dots, M_s - 1$) of the k' th user. Then, a vector \mathbf{s}

⁵In [25] and [26], it is shown that under certain conditions, $p\left(y_{i,p}^{(k')} \middle| d_{i,p}^{(k')}\right)$ can be modeled as a Gaussian distribution.

⁶In correspondence with technical literature, we will name the process of determining Δ *frame synchronization*.

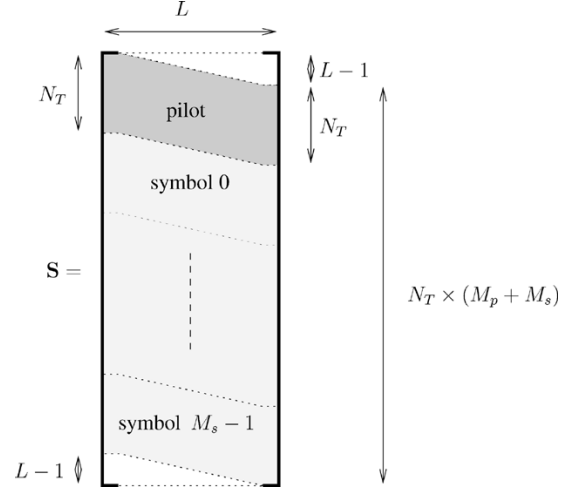


Fig. 4. Toeplitz symbol matrix \mathbf{S} for $M_p = 1$.

of length $(N_T(M_s + M_p) + 2L - 2)$ is constructed by concatenating all these time-domain samples, padded with $L - 1$ zeros at the beginning and end of the vector, leading to

$$\mathbf{s} \doteq [\mathbf{0}_{1 \times (L-1)} \mathbf{s}_C^0 \dots \mathbf{s}_C^{M_p-1} \mathbf{s}_D^0 \dots \mathbf{s}_D^{M_s-1} \mathbf{0}_{1 \times (L-1)}] \quad (12)$$

where $\mathbf{0}_{X \times Y}$ is an $X \times Y$ matrix consisting of all zeros.

Now, we define an $(N_T(M_s + M_p) + L - 1) \times L$ Toeplitz matrix \mathbf{S} as follows: the n th row of \mathbf{S} is obtained by time-reversing the n th until the $(n + L - 1)$ th sample of \mathbf{s} . For instance, the first row is given by $[s_{0,-\nu}, \mathbf{0}_{1 \times (L-1)}]$, the second row by $[s_{0,-\nu+1}, s_{0,-\nu}, \mathbf{0}_{1 \times (L-2)}]$, and so forth. The structure of this matrix is shown in Fig. 4. Note that we can write \mathbf{S} as the sum of two Toeplitz matrices of size $(N_T(M_s + M_p) + L - 1) \times L$: $\mathbf{S} = \mathbf{S}_P + \mathbf{S}_D$, where \mathbf{S}_P contains only the pilot MC symbols and \mathbf{S}_D contains only the data MC symbols.

Finally, we define an $(N_T(M_s + M_p) + \Delta_{\max} + L - 1) \times L$ matrix \mathbf{S}_Δ as

$$\mathbf{S}_\Delta = \begin{bmatrix} \mathbf{0}_{\Delta \times L} \\ \mathbf{S} \\ \mathbf{0}_{(\Delta_{\max} - \Delta + L - 1) \times L} \end{bmatrix}. \quad (13)$$

These transformations enable us to write the following simple relationship between \mathbf{r} , \mathbf{S}_Δ , and \mathbf{h} :

$$\mathbf{r} = \mathbf{F} \mathbf{S}_\Delta \mathbf{h} + \mathbf{w} \quad (14)$$

where \mathbf{F} is a diagonal matrix with $\mathbf{F}_{kk} = e^{j2\pi Fk/N}$, $k = -\nu, \dots, (M_s + M_p)N_T + \Delta_{\max} + L - \nu - 2$, and the vector \mathbf{w} embeds the thermal noise and the multiple-access interference (MAI) and is modeled as a zero-mean complex Gaussian random variable with variance σ^2 per real dimension. Note that by substituting $\mathbf{S} = \mathbf{S}_P + \mathbf{S}_D$ into (13), we can break up $\mathbf{S}_\Delta = \mathbf{S}_{\Delta,P} + \mathbf{S}_{\Delta,D}$.

The ML estimate of the delay shift and channel taps is obtained by maximizing the log-likelihood function

$$[\hat{\Delta}, \hat{F}, \hat{\mathbf{h}}] = \arg \max_{\Delta, F, \mathbf{h}} \{\log p(\mathbf{r} | \Delta, F, \mathbf{h})\} \quad (15)$$

where

$$p(\mathbf{r} | \Delta, F, \mathbf{h}) \propto \sum_{\mathbf{s}} p(\mathbf{r} | \Delta, F, \mathbf{h}, \mathbf{s}) p(\mathbf{s}) \quad (16)$$

and

$$p(\mathbf{r}|\Delta, F, \mathbf{h}, \mathbf{s}) \propto \exp\left(-\frac{1}{2\sigma^2} \|\mathbf{r} - \mathbf{F}\mathbf{S}_{\Delta}\mathbf{h}\|^2\right). \quad (17)$$

Unfortunately, the summation in (16) is intractable in practice, as it requires the evaluation of (17) over all 2^{N_b} possible codewords.

B. Data-Aided (DA) Estimation

The summation in (16) can be avoided by ignoring the unknown coded symbols. In this case

$$\begin{aligned} [\hat{\Delta}, \hat{F}, \hat{\mathbf{h}}] &= \arg \max_{\Delta, F, \mathbf{h}} \left\{ -\|\mathbf{r} - \mathbf{F}\mathbf{S}_{\Delta, P}\mathbf{h}\|^2 \right\} \\ &= \arg \max_{\Delta, F, \mathbf{h}} \left\{ -\mathbf{h}^H \mathbf{S}_{\Delta, P}^H \mathbf{S}_{\Delta, P} \mathbf{h} + 2\Re(\mathbf{r}^H \mathbf{F} \mathbf{S}_{\Delta, P} \mathbf{h}) \right\} \end{aligned} \quad (18)$$

$$(19)$$

where $\mathbf{S}_{\Delta, P}$ is obtained by replacing all unknown symbols in \mathbf{S}_{Δ} with zeros and we have used the fact that $\mathbf{F}^H \mathbf{F} = \mathbf{I}$. Now, (19) is solved as follows. We first note that for a given Δ and F , the estimate of \mathbf{h} can be found in closed form as

$$\hat{\mathbf{h}} = (\mathbf{S}_{\Delta, P}^H \mathbf{S}_{\Delta, P})^{-1} \mathbf{S}_{\Delta, P}^H \mathbf{F}^H \mathbf{r} \quad (20)$$

$$= (\mathbf{S}_P^H \mathbf{S}_P)^{-1} \mathbf{S}_{\Delta, P}^H \mathbf{F}^H \mathbf{r} \quad (21)$$

where \mathbf{S}_P was defined in the previous section. On the other hand, an estimate of Δ and F requires a two-dimensional (2-D) search

$$[\hat{\Delta}, \hat{F}] = \arg \max_{\Delta, F} \left\{ \Re(\mathbf{r}^H \mathbf{F} \mathbf{S}_{\Delta, P} \mathbf{h}) \right\}. \quad (22)$$

Back-substitution of (21) into (22) gives us

$$[\hat{\Delta}, \hat{F}] = \arg \max_{\Delta, F} \left\{ \Re\left(\mathbf{r}^H \mathbf{F} \mathbf{S}_{\Delta, P} (\mathbf{S}_P^H \mathbf{S}_P)^{-1} \mathbf{S}_{\Delta, P}^H \mathbf{F}^H \mathbf{r}\right) \right\} \quad (23)$$

and

$$\hat{\mathbf{h}} = (\mathbf{S}_P^H \mathbf{S}_P)^{-1} \mathbf{S}_{\Delta, P}^H \hat{\mathbf{F}}^H \mathbf{r}. \quad (24)$$

In the situation where the CFO is small as compared with the carrier spacing, the 2-D maximization in (23) can be reduced to two one-dimensional (1-D) maximizations as follows [7]: we first obtain a coarse timing-independent estimate of the CFO F , leading to a matrix $\tilde{\mathbf{F}}$ (e.g., setting $\tilde{\mathbf{F}} = \mathbf{I}$, effectively ignoring any CFO). We then find the delay as $\hat{\Delta} = \arg \max_{\Delta} \left\{ \Re\left(\mathbf{r}^H \tilde{\mathbf{F}} \mathbf{S}_{\Delta, P} (\mathbf{S}_P^H \mathbf{S}_P)^{-1} \mathbf{S}_{\Delta, P}^H \tilde{\mathbf{F}}^H \mathbf{r}\right) \right\}$. Based on this timing estimate, the CFO is finally estimated by maximizing the following cost function: $\hat{F} = \arg \max_F \left\{ \Re\left(\mathbf{r}^H \mathbf{F} \mathbf{S}_{\Delta, P} (\mathbf{S}_P^H \mathbf{S}_P)^{-1} \mathbf{S}_{\Delta, P}^H \mathbf{F}^H \mathbf{r}\right) \right\}$.

Note that, contrary to the DA estimator from [7], the matrix to be inverted in (23) and (24) is independent of Δ , so that $(\mathbf{S}_P^H \mathbf{S}_P)^{-1}$ can be precomputed and stored at the receiver.

One of the main drawbacks of many frame synchronization algorithms for MC systems is the presence of ambiguities: an ambiguity occurs when a cost-function such as (23): a) exhibits a plateau (with multiple consecutive values of the parameter

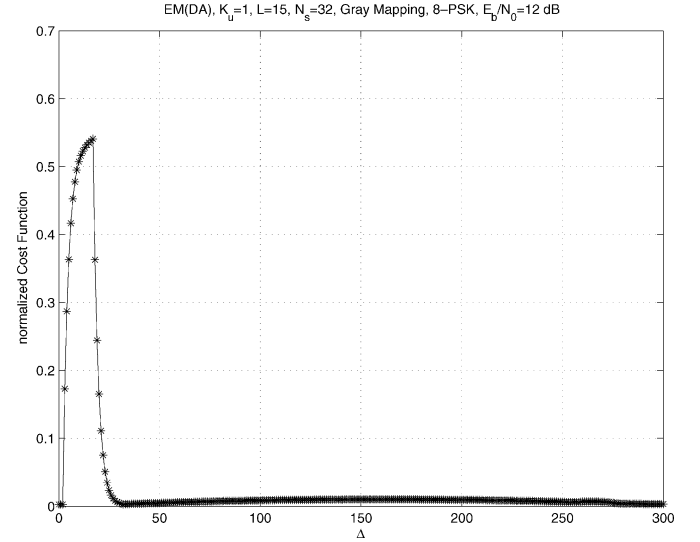


Fig. 5. DA frame synchronization: trial value of Δ versus cost function. $\Delta = 17$.

Δ giving rise to the same cost) or b) has multiple maxima, each with the same cost. For instance, the ML DA estimator from [7] is not able to estimate values of Δ beyond N_T : when $\Delta_{\max} > N_T$ ambiguities in the cost-function (23) occur. For the estimator we propose, which is based on a slightly different observation model, no ambiguity is present. To illustrate this point, we plot (with $F = 0$) a typical realization of the cost-function (23) for a single-user system in Fig. 5: there is a unique well-defined maximum of the cost-function.

While DA estimation algorithms perform well for uncoded systems, this is no longer true when error-correcting codes are concerned. Since such codes operate in low SNR regimes, many pilot symbols may be required to acquire reliable estimates. As this results in a significant loss in terms of power and bandwidth, there is great interest in developing algorithms that are also able to exploit the data MC symbols. In the following section, we describe a possible approach: the EM algorithm. It turns out that there are some very nice connections between the resulting code-aided algorithm and the conventional DA algorithm.

C. EM Estimation

1) *Principle*: The EM algorithm is a method that iteratively solves the ML problem (15) [17]. It requires us to define the so-called *complete data* \mathbf{z} . The complete data is related to the observation \mathbf{r} through some mapping $\mathbf{r} = g(\mathbf{z})$. Let us denote the parameter to be estimated θ (e.g., in our case, θ is a notational shorthand for the vector $[\Delta, F, \mathbf{h}^T]$).

The EM algorithm starts from an initial estimate of θ (say, $\hat{\theta}(0)$) and iteratively computes new estimates. At iteration ξ , the EM algorithm consists of two steps: given the current estimate $\hat{\theta}(\xi)$, we first take the expectation of the log-likelihood function of the complete data, given the observation \mathbf{r} and the current estimate of θ

$$Q(\theta|\hat{\theta}(\xi)) = E_{\mathbf{z}} \left[\log p(\mathbf{z}|\theta) \mid \mathbf{r}; \hat{\theta}(\xi) \right]. \quad (25)$$

In the second step, we maximize $Q\left(\theta|\hat{\theta}(\xi)\right)$ with respect to θ to find a new estimate

$$\hat{\theta}(\xi+1) = \arg \max_{\theta} \left\{ Q\left(\theta|\hat{\theta}(\xi)\right) \right\}. \quad (26)$$

Convergence of the EM algorithm is guaranteed in a sense that the likelihoods of the estimates are nondecreasing

$$p\left(\mathbf{r}|\hat{\theta}(\xi+1)\right) \geq p\left(\mathbf{r}|\hat{\theta}(\xi)\right) \quad (27)$$

for $\xi \geq 0$. Any value $\hat{\theta}$ for which $\hat{\theta} = \arg \max_{\theta} Q\left(\theta|\hat{\theta}\right)$ is called a *solution* of the EM algorithm. One of these solutions is the ML estimate. In order to achieve convergence to the ML estimate, a good initial estimate of θ is required.

2) *Code-Aided Estimation*: Let us take as complete data $\mathbf{z} = [\mathbf{r}, \mathbf{s}]$. In that case, since \mathbf{s} and θ are independent

$$\log p(\mathbf{z}|\theta) \propto \log p(\mathbf{r}|\theta, \mathbf{s}) \quad (28)$$

so that (25) becomes

$$Q\left(\theta|\hat{\theta}(\xi)\right) = E_{\mathbf{s}} \left[\log p(\mathbf{r}|\theta, \mathbf{s}) | \mathbf{r}, \hat{\theta}(\xi) \right]. \quad (29)$$

We know from (19) that

$$\log p(\mathbf{r}|\theta, \mathbf{s}) \propto -\mathbf{h}^H \mathbf{S}_{\Delta}^H \mathbf{S}_{\Delta} \mathbf{h} + 2\Re\left(\mathbf{r}^H \mathbf{F} \mathbf{S}_{\Delta} \mathbf{h}\right) \quad (30)$$

so that

$$Q\left(\theta|\hat{\theta}(\xi)\right) = -\mathbf{h}^H \widetilde{\mathbf{S}}_{\Delta}^H \widetilde{\mathbf{S}}_{\Delta} \mathbf{h} + 2\Re\left(\mathbf{r}^H \mathbf{F} \widetilde{\mathbf{S}}_{\Delta} \mathbf{h}\right) \quad (31)$$

where, due to the linearity of the expectation operator, $\widetilde{\mathbf{S}}_{\Delta} = E_{\mathbf{s}} \left[\mathbf{S}_{\Delta} | \mathbf{r}, \hat{\theta}(\xi) \right]$ is obtained by replacing each entry $s_{i,m}^{(k')}$ with the corresponding *a posteriori* expectation $E \left[s_{i,m}^{(k')} | \mathbf{r}, \hat{\theta}(\xi) \right]$. Similarly, $\widetilde{\mathbf{S}}_{\Delta}^H \mathbf{S}_{\Delta} = E_{\mathbf{s}} \left[\mathbf{S}_{\Delta}^H \mathbf{S}_{\Delta} | \mathbf{r}, \hat{\theta}(\xi) \right]$ is obtained by replacing the entries $s_{i,m}^{(k')} \left(s_{i',m'}^{(k')} \right)^*$ with $E \left[s_{i,m}^{(k')} \left(s_{i',m'}^{(k')} \right)^* | \mathbf{r}, \hat{\theta}(\xi) \right]$. Equation (1) tells us that

$$s_{i,m}^{(k')} = \sqrt{\frac{E_s}{N+\nu}} \sum_{p=0}^{P-1} \sum_{s=0}^{N_s-1} d_{i,p}^{(k')} a_{p,s}^{(k')} e^{j2\pi n_p s m / N} \quad (32)$$

so that

$$\begin{aligned} & E \left[s_{i,m}^{(k')} | \mathbf{r}, \hat{\theta}(\xi) \right] \\ &= \sqrt{\frac{E_s}{N+\nu}} \sum_{p=0}^{P-1} \sum_{s=0}^{N_s-1} E \left[d_{i,p}^{(k')} | \mathbf{r}, \hat{\theta}(\xi) \right] a_{p,s}^{(k')} e^{j2\pi n_p s m / N}. \end{aligned} \quad (33)$$

From Section II, we know that the detector computes the APPs of the coded symbols $p\left(d_{i,p}^{(k')} | \mathbf{r}, \hat{\theta}(\xi)\right)$. The *a posteriori* expectation $E \left[d_{i,p}^{(k')} | \mathbf{r}, \hat{\theta}(\xi) \right]$ is obtained as

$$E \left[d_{i,p}^{(k')} | \mathbf{r}, \hat{\theta}(\xi) \right] = \sum_{\omega \in \Omega} \omega \times p\left(d_{i,p}^{(k')} = \omega | \mathbf{r}, \hat{\theta}(\xi)\right) \quad (34)$$

which can be interpreted as a soft symbol decision: it is a weighted average of all possible constellation points. As $E \left[s_{i,m}^{(k')} \left(s_{i',m'}^{(k')} \right)^* | \mathbf{r}, \hat{\theta}(\xi) \right]$ is a function of $E \left[d_{i,p}^{(k')} \left(d_{i',p'}^{(k')} \right)^* | \mathbf{r}, \hat{\theta}(\xi) \right]$, it cannot be computed exactly based solely on the (marginal) APPs. However, thanks to the presence of the interleaver, the coded symbols can be assumed to be essentially uncorrelated so that

$$\begin{aligned} & E \left[d_{i,p}^{(k')} \left(d_{i',p'}^{(k')} \right)^* | \mathbf{r}, \hat{\theta}(\xi) \right] \\ & \approx E \left[d_{i,p}^{(k')} | \mathbf{r}, \hat{\theta}(\xi) \right] E \left[\left(d_{i',p'}^{(k')} \right)^* | \mathbf{r}, \hat{\theta}(\xi) \right] \end{aligned} \quad (35)$$

when $i \neq i'$ or $p' \neq p$. Additionally, if we extend the previous relation to $i = i'$ and $p = p'$, the computation of $\mathbf{S}_{\Delta}^H \mathbf{S}_{\Delta}$ can further be simplified as

$$\widetilde{\mathbf{S}}_{\Delta}^H \widetilde{\mathbf{S}}_{\Delta} \approx \widetilde{\mathbf{S}}_{\Delta}^H \widetilde{\mathbf{S}}_{\Delta}. \quad (36)$$

Finally, the updated estimates of the delay shift, the CFO, and the channel taps are given by

$$\begin{aligned} & \left[\hat{\Delta}(\xi+1), \hat{F}(\xi+1) \right] \\ &= \arg \max_{\Delta, F} \left\{ \Re \left(\mathbf{r}^H \mathbf{F} \widetilde{\mathbf{S}}_{\Delta} \left(\widetilde{\mathbf{S}}^H \mathbf{S} \right)^{-1} \widetilde{\mathbf{S}}_{\Delta}^H \mathbf{F}^H \mathbf{r} \right) \right\} \end{aligned} \quad (37)$$

and

$$\hat{\mathbf{h}}(\xi+1) = \left(\widetilde{\mathbf{S}}^H \mathbf{S} \right)^{-1} \widetilde{\mathbf{S}}_{\Delta(\xi+1)}^H \hat{\mathbf{F}}^H \mathbf{r} \quad (38)$$

where $\widetilde{\mathbf{S}}^H \mathbf{S} = E_{\mathbf{s}} \left[\mathbf{S}^H \mathbf{S} | \mathbf{r}, \hat{\theta}(\xi) \right]$ with \mathbf{S} defined in Fig. 4. Making use of the approximation (36), this leads to a very elegant interpretation: the EM-based algorithms are formally obtained by replacing in the corresponding DA algorithms, pilot symbols with *a posteriori* symbol expectations.

Note that the 2-D maximization in (23) and (37) can be reduced to two 1-D maximizations when the CFO is small [7]. Therefore, the low-complexity update rules for the timing and the CFO are

$$\begin{aligned} & \left[\hat{\Delta}(\xi+1) \right] \\ &= \arg \max_{\Delta} \left\{ \Re \left(\mathbf{r}^H \hat{\mathbf{F}}(\xi) \widetilde{\mathbf{S}}_{\Delta} \left(\widetilde{\mathbf{S}}^H \mathbf{S} \right)^{-1} \widetilde{\mathbf{S}}_{\Delta}^H \hat{\mathbf{F}}^H(\xi) \mathbf{r} \right) \right\} \end{aligned} \quad (39)$$

and

$$\begin{aligned} & \left[\hat{F}(\xi+1) \right] \\ &= \arg \max_F \left\{ \Re \left(\mathbf{r}^H \mathbf{F} \widetilde{\mathbf{S}}_{\Delta(\xi+1)} \left(\widetilde{\mathbf{S}}^H \mathbf{S} \right)^{-1} \widetilde{\mathbf{S}}_{\Delta(\xi+1)}^H \mathbf{F}^H \mathbf{r} \right) \right\}. \end{aligned} \quad (40)$$

D. Implementation Aspects

The proposed EM estimator can be modified in many ways.

- 1) Replacing $\log p(\mathbf{r}|\theta, \mathbf{s})$ with $\log p(\mathbf{r}, \theta|\mathbf{s})$ allows for performing maximum *a posteriori* (MAP) estimation. In [9], it has been shown that when the channel taps have

TABLE I
 NOTATIONS

notation	meaning
i	MC symbol index: $0 \leq i < M_s$
k	user index: $1 \leq k \leq K_u$
s	chip index: $0 \leq s < N_s$
m	time-domain index: $-\nu \leq m < N = N_s P$
p	data symbol index (for a given MC symbol): $0 \leq p < P$
ξ	EM iteration index: $0 \leq \xi$

TABLE II

COMPUTATIONAL COMPLEXITY, ASSUMING $\mathbf{S}^H \mathbf{S}$ IS A DIAGONAL MATRIX. WE HAVE OMITTED THE COMPLEXITY RELATED TO UPDATING \hat{F} AS IT DEPENDS ON THE SPECIFIC NUMERICAL TECHNIQUE USED TO SOLVE (40)

computation	complexity
$\tilde{\mathbf{S}}_\Delta$	$\mathcal{O}(NM_s \log_2 N)$
$\tilde{\mathbf{S}}^H \mathbf{S}$	$\mathcal{O}(N_T (M_s + M_p) L)$
update of $\hat{\Delta}$	$\mathcal{O}(N_T (M_s + M_p) L \Delta_{max})$
update of $\hat{\mathbf{h}}$	$\mathcal{O}(N_T (M_s + M_p) L)$

a Gaussian *a priori* distribution, MAP estimation has certain implementation advantages.

- 2) The inversion of the matrix $\tilde{\mathbf{S}}^H \mathbf{S}$ in (37) and (38) can be simplified by noticing that for large $M_p + M_s$, this matrix is roughly diagonal. A more detailed view of the computational complexity is given in Table II. Observe that the total computational cost is dominated by the Δ -update.
- 3) When the estimates of Δ and F do not change during an EM update step, all computations [with the exception of (38)] can take place in the frequency-domain.
- 4) In case the delay Δ and the CFO F are perfectly known, the code-aided estimation algorithm of the CIR can take place completely in the frequency-domain, thus avoiding FFT operations at each EM iteration. This requires a specific frequency-domain observation model. This is not pursued in the current paper.

Even with these complexity-reducing modifications, the complexity of the EM-based estimator may still be unacceptably high: let us denote by T_{EM} the time (in seconds) to compute the cost-function (31) and perform a single update to the estimates of F , Δ and \mathbf{h} , and by T_{detect} the time to detect the data, given an estimate⁷ of F , Δ , and \mathbf{h} . When I_{EM} EM iterations are performed, the total computation time is roughly $T_{tot} = I_{EM} T_{EM} + (I_{EM} + 1) T_{detect}$, so that the overhead related to estimation is given by

$$O_{EM} = \frac{T_{tot} - T_{detect}}{T_{detect}} = I_{EM} \left(1 + \frac{T_{EM}}{T_{detect}} \right)$$

which is (at least) an I_{EM} -fold complexity increase as compared with a conventional system with (noniterative) estimation followed by data detection (corresponding to an overhead of 0).

⁷Hence, T_{detect} represents the processing time for a perfectly synchronized detector with perfect channel knowledge.

Currently, many state-of-the-art detectors operate according to the so-called *turbo-principle*, which basically means that the detector operates according to some iterative procedure. Hence, we may write $T_{detect} = I_D T_D$, where I_D is the number of iterations performed within the detector, and T_D the time to perform a single iteration. We now perform the following modifications: for each EM iteration, we perform only a single iteration within the detector, but we maintain state information⁸ from one iteration to the next. This is known as *embedded* estimation [29]. In this case, the total computation time is roughly $T_{tot} = I_{EM} T_{EM} + (I_{EM} + 1) T_D$. Suppose we perform roughly $I_{EM} = I_D - 1$ EM iterations, then the overhead related to EM estimation is now

$$O_{EM} = \frac{I_{EM}}{I_{EM} + 1} \times \frac{T_{EM}}{T_D}$$

which is generally a fairly small number.⁹ Hence, embedded EM estimation is especially well-suited to detectors which are themselves iterative.

IV. NUMERICAL RESULTS

A. Simulation Parameters

To validate the proposed algorithms, we have carried out Monte Carlo simulations. We consider a system with $K_u = 5$ users, using a convolutional code with constraint length 5, rate $R = 1/2$, and polynomial generators (23)₈ and (35)₈. A block length of $N_b = 240$ information bits was chosen, leading to $N_c = 480$ coded bits. Coded bits are gray-mapped onto an 8-PSK constellation, resulting in $N_d = 160$ data symbols. This sequence of N_d 8-phase-shift keying (PSK) symbols is broken up into $M_s = 20$ blocks of $P = 8$ symbols. Spreading sequences are real-valued Walsh-Hadamard sequences, with chips belonging to $\{-1/\sqrt{N_s}, +1/\sqrt{N_s}\}$ and have a length $N_s = 32$, leading to $N = P N_s = 256$ required subcarriers. To initialize the EM algorithm, the $M_s = 20$ data MC symbols are preceded by $M_p = 1$ pilot MC symbols. Within the pilot MC symbols, only a fraction $d = 0.25$ of the subcarriers are devoted to training. The remaining 75% of the carriers is reserved for administrative data, and cannot be used during the synchronization/estimation process. The channel has length $L = 15$ and was modeled with independent components, each being a zero-mean complex Gaussian random variable with an exponential power delay profile [7]

$$E[|h(l)|^2] = \sigma_h^2 \exp\left(-\frac{2l}{5}\right), l = 0, \dots, L-1 \quad (41)$$

where we have set σ_h^2 to unity. Hence, the energy of the channel is concentrated mainly in the first few channel taps. To avoid ISI, a cyclic prefix of length $\nu = 16$ is employed. We have set $\Delta_{max} = 30$ and fix $\Delta = 17$.

Note that we have deliberately chosen a detector which is noniterative. Although this implies an I_{EM} -fold complexity overhead caused by EM estimation, a noniterative detector allows us to show the performance gains at each iteration.

⁸Sometimes known as *extrinsic* information.

⁹E.g., for $T_{EM} = T_D/2$, $O_{EM} \approx 0.5$, meaning EM estimation gives rise to a 50% computational overhead.

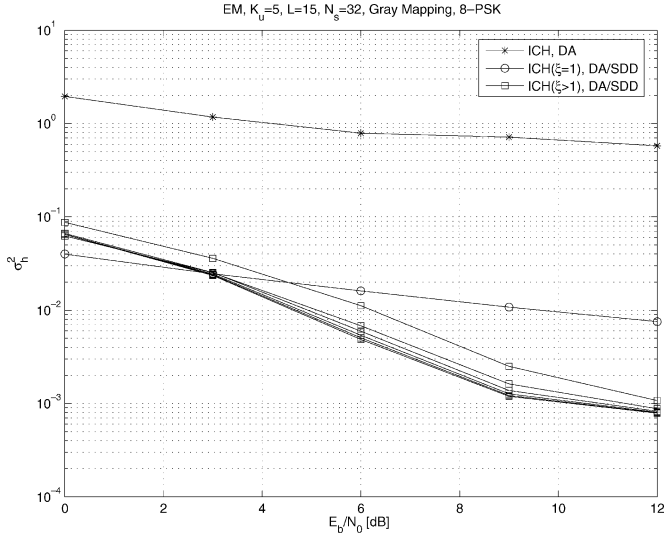


Fig. 6. CIR estimation: MSEE performance. Perfect synchronization is assumed. DA/SDD stands for data-aided/soft decision directed.

B. Performance Results

First, consider estimation of the CIR \mathbf{h} , assuming perfect knowledge of Δ and F (denoted PTS+PFS+ICH for perfect time and frequency synchronization with imperfect channel knowledge). In Fig. 6, we show the mean squared estimation error (MSEE) as a function of the SNR, for different values of ξ (the EM iteration index). The DA estimate (corresponding to $\xi = 0$, or equivalently, (24) with $\hat{\Delta} = \Delta$ and $\hat{F} = F$) results in fairly large MSEE (up to 0.6 at SNR of 12 dB). Application of the EM estimation algorithm (38) reduces the MSEE noticeably. For low SNR, the MSEE rises again slightly after the first EM iteration. This is due to the different approximations in the detector and the EM estimator, including the Gaussian approximation of the decision variables, the computation of $\mathbf{S}_{\Delta}^H \mathbf{S}_{\Delta}$, etc. After three or four EM iterations, no improvement in MSEE is visible for any SNR. To see how this translates into BER performance, consider Fig. 7: the bottom-most curve (in dashed) corresponds to the single-user bound, while the curve with square markers shows the BER performance for a genie-aided receiver (i.e., a receiver with perfect knowledge of \mathbf{h} , Δ and F). When performing DA estimation of the CIR (the top-most curve), unacceptable BER degradations are visible, as compared with the genie-aided receiver. The EM code-aided estimator is able to reduce this degradation to around 0.2 dB after $\xi = 4$ EM iterations.

Frequency synchronization [denoted (IFS): imperfect frequency synchronization] deserves some special attention. For simplicity, let us focus on a single-user environment ($K_u = 1$) with known channel \mathbf{h} and known delay Δ . It has been shown in [24] that DA ML CFO synchronization requires a fairly large amount of pilot symbols. In Fig. 8, we depict the BER as a function of the CFO. When the receiver simply ignores the CFO (marked by “IFE, $\varepsilon_F = -F$ ”), large degradations ensue. The EM-based estimator on the other hand, with $\hat{F}(\xi = 0) = 0$, is able to reduce the large degradation due to the CFO as long as the normalized CFO is smaller than $F = 4 \times 10^{-3}$. This corresponds to $f_0 \leq 1/(10 \times (M_s + M_p)T)$. This is further

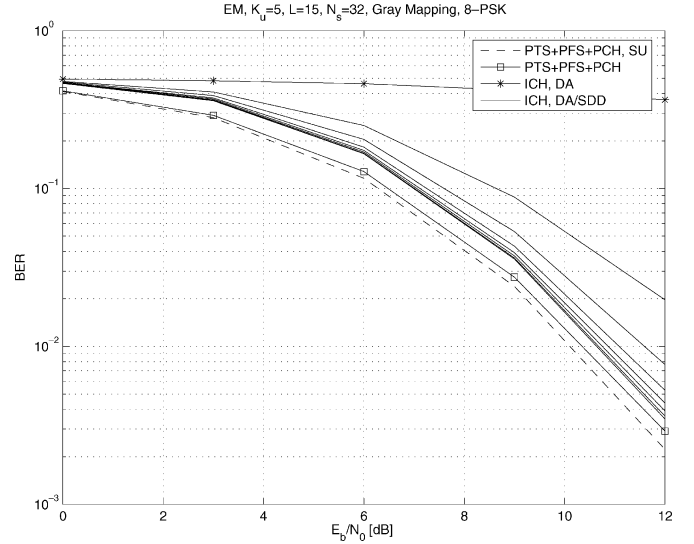


Fig. 7. CIR estimation: BER performance. Perfect synchronization is assumed. DA/SDD stands for data-aided/soft decision directed.

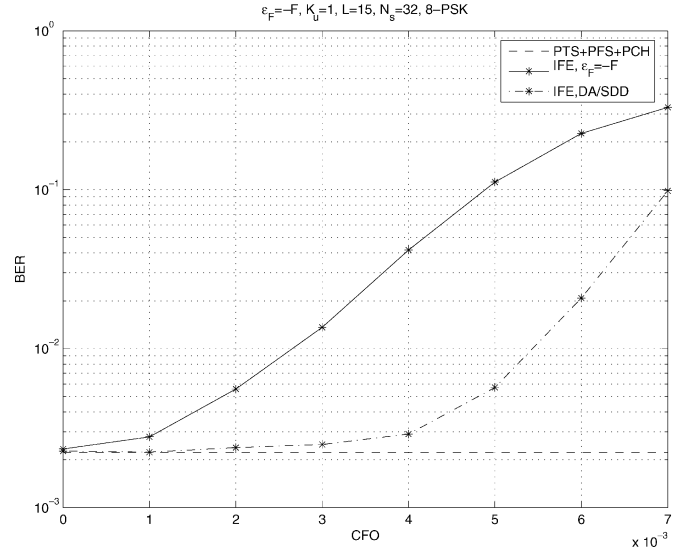


Fig. 8. Frequency synchronization: BER performance at 12 dB using $\hat{F}(\xi = 0) = 0$ as a frequency initial estimate. DA/SDD stands for data-aided/soft decision directed.

illustrated in Fig. 9 where the corresponding MSEE performance of the CFO estimate is shown (i.e., $E \left[\left| \hat{F} - F \right|^2 \right]$). Again, with $\hat{F}(\xi = 0) = 0$, the EM algorithm is able to reduce the MSEE with increasing iteration index ξ .

Let us now also take into account frame synchronization (denoted ITS+IFS+ICH for Imperfect Time and Frequency Synchronization with Imperfect Channel knowledge) for $K_u = 5$. The CFO is randomly selected such that $|F| < 4.8 \times 10^{-3}$ and the initial CFO estimate is set to 0, i.e., $\hat{F}(\xi = 0) = 0$. Note that we use the low-complexity update rules in (39) and (40) for joint frame and CFO synchronization. As the MSEE of Δ gives in general only a partial view of the behavior of $\hat{\Delta}$, we include, for a SNR of 12 dB, the simulated probability mass function (pmf) of the estimation error $\varepsilon_{\Delta} = \hat{\Delta} - \Delta$ (see Fig. 10). The DA estimator has a fairly broad pmf, with a maximum $\varepsilon_{\Delta} = 1$.

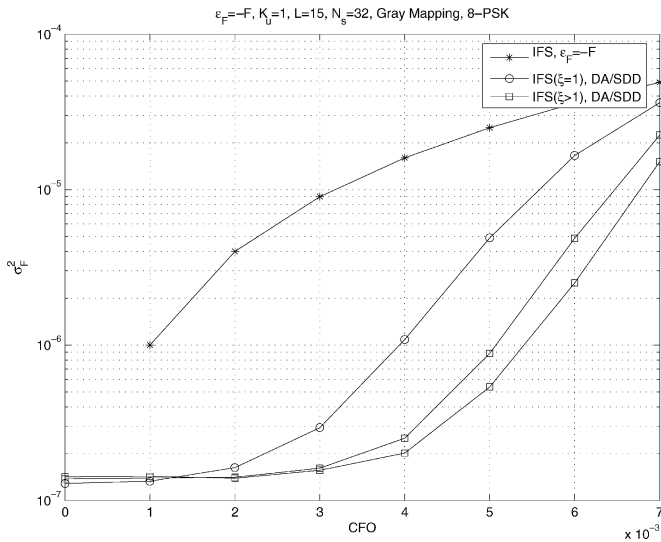


Fig. 9. Frequency synchronization: MSEE performance at 12 dB using $\hat{F}(\xi = 0) = 0$ as a frequency initial estimate. DA/SDD stands for data-aided/soft decision directed.

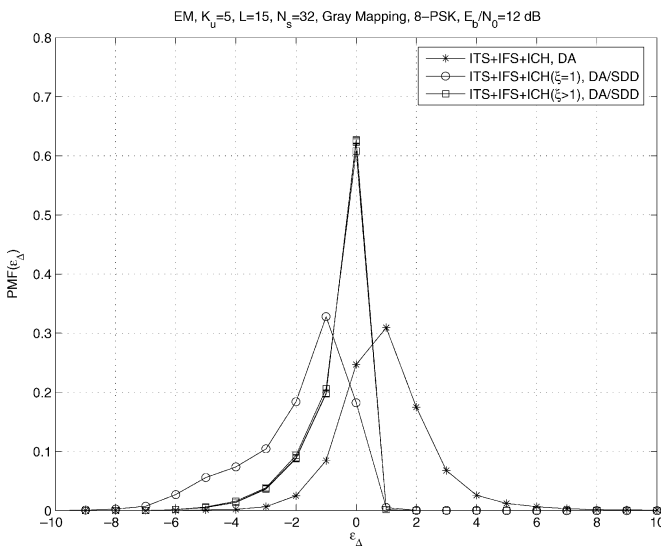


Fig. 10. Frame synchronization: pmf of estimation error at 12 dB using $\hat{F}(\xi = 0) = 0$ as a frequency initial estimate and F is selected randomly such that $|F| < 4.8 \times 10^{-3}$. DA/SDD stands for data-aided/soft decision directed.

The pmf of ε_{Δ} for the code-aided EM estimator is much more narrow, with a distinct maximum at $\varepsilon_{\Delta} = 0$. After $\xi = 2$ iterations, the pmf does not change noticeably. It should be noted that although 60% of the frames result in a correct estimate of Δ , this does not mean that the frame error rate equals 40%: when $\varepsilon_{\Delta} > 0$ (with respect to $\varepsilon_{\Delta} < 0$), ISI occurs between the current and the next (with respect to previous) MC symbol. Since the first few channel taps carry most of the energy, the situation $\varepsilon_{\Delta} < 0$ is not very critical. On the other hand, $\varepsilon_{\Delta} > 0$ should be avoided, as the estimate of \mathbf{h} will not capture the dominant components. From Fig. 10, it is clear that the latter situation occurs only rarely for the EM-based estimator.

Finally, Fig. 11 shows the BER for joint frame-frequency synchronization and channel estimation. As expected, the DA estimator gives rise to large degradations. On the other

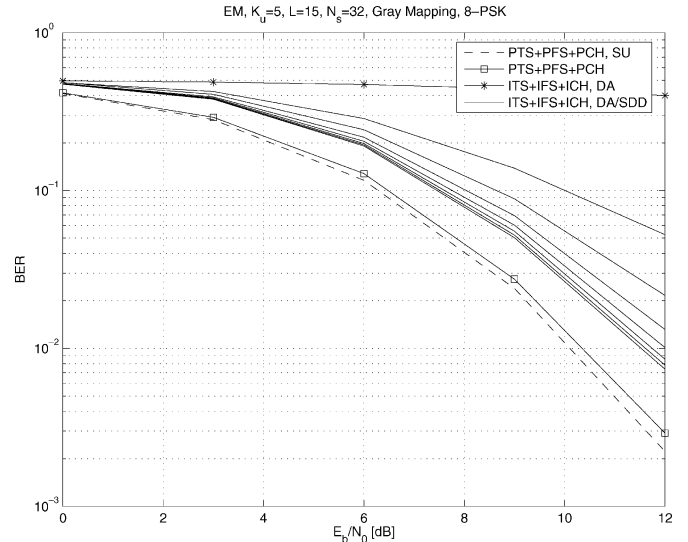


Fig. 11. Joint frame-frequency synchronization and channel estimation: BER performance using $\hat{F}(\xi = 0) = 0$ as a frequency initial estimate and F is selected randomly such that $|F| < 4.8 \times 10^{-3}$. DA/SDD stands for data-aided/soft decision directed.

hand, the EM estimator results in a BER degradation less than 1 dB as compared with the case of perfect frame-frequency synchronization and perfect channel knowledge after roughly $\xi = 4$ EM iterations.

V. CONCLUSION

We have presented a novel code-aided estimation algorithm for joint synchronization and CIR estimation for downlink MC-CDMA. Based on the EM algorithm, the receiver iterates between data detection and estimation, with the exchange of soft information in the form of a *posteriori* probabilities. Compared with a conventional data-aided algorithm, the code-aided algorithm results in impressive gains in terms of MSEE and BER performance. Although the complexity of this estimator is large, we have described how the computational load may be reduced, resulting in a practical algorithm. For iterative detectors, the proposed algorithm gives rise to a fairly small overhead.

The proposed algorithm can easily be extended to take into account other observation models (such as the aforementioned frequency-domain channel estimator).

As a final comment, we should mention that in some cases a more conventional technique may be preferred to the proposed code-aided estimator: as always, the choice of using a particular synchronization/estimation algorithm should be made on a case-by-case basis, trading performance against complexity, power consumption, etc.

REFERENCES

- [1] J. A. C. Bingman, "Multicarrier modulation: An idea whose time has come," *IEEE Commun. Mag.*, vol. 28, no. 5, pp. 5–14, May 1990.
- [2] S. Nanda, R. Walton, J. Ketchum, M. Wallace, and S. Howard, "A high-performance MIMO OFDM wireless LAN," *IEEE Commun. Mag.*, vol. 43, no. 1, pp. 53–60, Jan. 2005.
- [3] C. Ibars and Y. Bar-Ness, "Comparing the performance of coded multiuser OFDM and coded MC-CDMA over fading channels," in *Proc. IEEE Globecom*, 2001, pp. 881–885.

- [4] L.-L. Yang and L. Hanzo, "Multicarrier DS-CDMA: A multiple access scheme for ubiquitous broadband wireless communications," *IEEE Commun. Mag.*, vol. 41, no. 10, pp. 116–124, Oct. 2003.
- [5] V. Tarokh and H. Jafarkhani, "On the computation and reduction of peak-to-average power ratio in multicarrier communications," *IEEE Trans. Commun.*, vol. 41, no. 1, pp. 37–44, Jan. 2000.
- [6] H. Steendam and M. Moeneclaey, "Synchronization sensitivity of multicarrier systems," *Eur. Trans. Commun.*, (ETT Special Issue on Multi-Carrier Spread Spectrum), vol. 52, no. 5, pp. 834–844, May 2004.
- [7] M. Morelli, "Timing and frequency synchronization for the uplink of an OFDMA system," *Trans. Commun.*, vol. 52, pp. 296–306, Feb. 2004.
- [8] T. M. Schmidl and D. C. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Trans. Commun.*, vol. 45, pp. 1613–1621, Dec. 1997.
- [9] 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.
- [10] M. Speth, S. A. Fechtel, G. Fock, and H. Meyr, "Optimum receiver design for wireless broadband systems using OFDM. I," *IEEE Trans. Commun.*, vol. 47, no. 11, pp. 1668–1677, Nov. 1999.
- [11] M. Speth, S. Fechtel, G. Fock, and H. Meyr, "Optimum receiver design for OFDM-based broadband transmission II. A case study," *IEEE Trans. Commun.*, vol. 49, no. 4, pp. 571–579, Apr. 2001.
- [12] E. G. Larsson, G. Liu, J. Li, and G. B. Giannakis, "Joint symbol timing and channel estimation for OFDM based WLAN's," *IEEE Commun. Lett.*, vol. 5, no. 8, pp. 325–327, Aug. 2001.
- [13] Y. Yao and G. B. Giannakis, "Blind carrier frequency offset estimation in SISO, MIMO, and multiuser OFDM systems," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 173–183, Jan. 2005.
- [14] M. Tanda, "Blind symbol-timing and frequency offset estimation in OFDM systems with real data symbols," *IEEE Trans. Commun.*, vol. 52, no. 10, pp. 1609–1612, Oct. 2004.
- [15] M. C. Necker and G. L. Stuber, "Totally blind channel estimation for OFDM on fast varying mobile radio channels," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1514–1525, Sep. 2004.
- [16] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo codes," in *Proc. IEEE ICC'94*, 1994, pp. 1064–1070.
- [17] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Royal Statist. Soc.*, ser. B, vol. 39, no. 1, pp. 1–38, 1977.
- [18] N. Noels, V. Lottici, A. Dejonghe, H. Steendam, M. Moeneclaey, M. Luise, and L. Vandendorpe, "A theoretical framework for soft information based synchronization in iterative (Turbo) receivers," *EURASIP J. Wireless Commun. Netw. JWCN* (Special Issue on Advanced Signal Processing Algorithms for Wireless Communications), 2005, to be published.
- [19] G. A. Al-Rawi, T. Y. Al-Naffouri, A. Bahai, and J. Cioffi, "Exploiting error-control coding and cyclic prefix in channel estimation for coded OFDM systems," *IEEE Commun. Lett.*, vol. 7, no. 7, pp. 388–390, Jul. 2003.
- [20] C. H. Aldana, E. de Carvalho, and J. M. Cioffi, "Channel estimation for multicarrier multiple input single output systems using the EM algorithm," *IEEE Trans. Signal Processing*, vol. 51, no. 12, pp. 3280–3292, Dec. 2003.
- [21] T. Y. Al-Naffouri, O. Awoniyi, O. Oteri, and A. Paulraj, "Receiver design for MIMO-OFDM transmission over time variant channels," in *Proc. IEEE Global Commun. Conf. (Globecom)*, vol. 4, Dec. 2004, pp. 436–440.
- [22] S. Y. Park, Y. G. Kim, and C. G. Kang, "Iterative receiver for joint detection and channel estimation in OFDM systems under mobile radio channels," *IEEE Trans. Veh. Technol.*, vol. 53, no. 2, pp. 450–460, Mar. 2004.
- [23] M. Guenach, H. Wymeersch, and M. Moeneclaey, "On channel parameter estimation in a space-time bit-interleaved coded modulation systems for multi-path DS-CDMA uplink with receive diversity," *IEEE Trans. Veh. Technol.*, vol. 54, no. 5, pp. 1747–1758, Sep. 2005. [Online]. Available: <http://telin.ugent.be/~hwymeersch/publications.html>.
- [24] M. Guenach, F. Simoens, H. Wymeersch, and M. Moeneclaey, "Code-aided joint channel and frequency offset estimation for DS-CDMA," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 1, pp. 181–189, Jan. 2006.
- [25] S. Kaiser, "OFDM code-division multiplexing in fading channels," *IEEE Trans. Commun.*, vol. 50, pp. 1266–1273, Aug. 2002.
- [26] X. Wang and H. V. Poor, "Space-time multiuser detection in multipath CDMA channels," *IEEE Trans. Commun.*, vol. 47, pp. 2356–2374, 1999.
- [27] S. ten Brink, J. Speidel, and R.-H. Yan, "Iterative demapping and decoding for multilevel modulation," in *Proc. IEEE Global Commun. Conf. (Globecom)*, Sydney, Australia, Nov. 1998, pp. 579–584.
- [28] F. Kschischang, B. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, pp. 498–519, Feb. 2001.
- [29] V. Lottici and M. Luise, "Embedding carrier phase recovery into iterative decoding of turbo-coded linear modulations," *IEEE Trans. Commun.*, vol. 52, pp. 661–669, Apr. 2004.



Mamoun Guenach (S'99–M'03) received the degree of engineer in electronics and communications from the Ecole Mohamadia d'Ingenieurs, Morocco. In 1997, he moved to the faculty of applied sciences at the Universite Catholique de Louvain, Louvain, Belgium, where he received the M.S. degree in electricity and the Ph.D. degree in applied sciences in 1998 and 2002, respectively.

Currently, he is working as a Postdoctoral Researcher at Ghent University, Gent, Belgium. His main research interests are turbo equalization, synchronization and channel estimations for multiuser CDMA and MC systems.



Henk Wymeersch (S'99–M'05) received the Diploma degree in computer science engineering and the Ph.D. degree in applied sciences from Ghent University, Ghent, Belgium, in 2001 and 2005, respectively.

From 2001 to 2005, he was with the Department of Telecommunications and Information Processing, Ghent University. Currently, he is a Postdoctoral Fellow of the Belgian-American Educational Foundation and is affiliated with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology (MIT), Cambridge. His research interests include channel coding, synchronization, iterative processing, and sensor networks.



Heidi Steendam (M'00) received the M.Sc. degree in electrical engineering and the Ph.D. degree in applied sciences from Ghent University, Ghent, Belgium, in 1995 and 2000, respectively.

Since October 2002, she has been a full-time Professor with the Digital Communications (DIGCOM) Research Group, Department of Telecommunications and Information Processing (TELIN), Ghent University. She is the author of more than 70 scientific papers in international journals and conference proceedings. Her main research interests are in statistical communication theory, carrier and symbol synchronization, bandwidth-efficient modulation and coding, spread-spectrum (multicarrier spread-spectrum), satellite, and mobile communication.



Marc Moeneclaey (M'93–SM'99–F'02) received the Diploma and the Ph.D. degrees from Ghent University, Ghent, Belgium, in 1978 and 1983, respectively, both in electrical engineering.

He is currently a Professor in the Department of Telecommunications and Information Processing, Ghent University. He coauthored *Digital Communication Receivers—Synchronization, Channel Estimation, and Signal Processing* (New York: Wiley, 1998) with H. Meyr (RWTH Aachen) and S. Fechtel (Siemens AG). He is the author of about 250 scientific papers in international journals and conference proceedings. His main research interest are in statistical communication theory, carrier and symbol synchronization, bandwidth-efficient modulation and coding, spread spectrum, satellite, and mobile communication.