

EM-Based Iterative Receivers for OFDM and BICM/OFDM Systems in Doubly Selective Channels

Meng-Lin Ku, *Member, IEEE*, Wen-Chuan Chen, and Chia-Chi Huang

Abstract—In this paper, we resort to the expectation-maximization (EM) algorithm to tackle the inter-carrier interference (ICI) problem, caused by time-variant multipath channels, for both orthogonal frequency division multiplexing (OFDM) systems and bit-interleaved coded modulation (BICM)/OFDM systems. We first analyze the ICI in frequency domain with a reduced set of parameters, and following this analysis, we derive an EM algorithm for maximum likelihood (ML) data detection. An ML-EM receiver for OFDM systems and a TURBO-EM receiver for BICM/OFDM systems are then developed to reduce computational complexity of the EM algorithm and to exploit temporal diversity, the main idea of which is to integrate the proposed EM algorithm with a groupwise ICI cancellation method. Compared with the ML-EM receiver, the TURBO-EM receiver further employs a soft-output Viterbi algorithm (SOVA) decoder to exchange information with a maximum *a posteriori* (MAP) EM detector through the turbo principle. Computer simulation demonstrates that the two proposed receivers clearly outperform the conventional one-tap equalizer, and the performance of the TURBO-EM receiver is close to the matched-filter bound even at a normalized maximum Doppler frequency (MDF) up to 0.2.

Index Terms—Orthogonal frequency division multiplexing, bit-interleaved coded modulation, inter-carrier interference, expectation-maximization algorithms, turbo receivers.

I. INTRODUCTION

ORTHOGONAL frequency division multiplexing (OFDM) is a promising technique to realize high data rate transmission over multipath fading channels. Due to the use of a guard interval (GI), it allows for a simple one-tap equalizer [1]. In addition, bit-interleaved coded modulation (BICM) combined with OFDM, known as BICM/OFDM, is introduced as a way to offer superior performance by exploiting frequency diversity [2]. Over the past decade, OFDM has found widespread application in several standards such as 802.16e wireless metropolitan area network (WMAN) [3]. However, in mobile radio environments, multipath channels are usually time-variant. The channel time variation destroys the orthogonality among

subcarriers, and thereby yields inter-carrier interference (ICI). The effect of ICI on the bit-error-rate (BER) performance has been intensively studied in [4], [5]. As the maximum Doppler frequency (MDF) increases, the one-tap equalizer is no longer sufficient to conquer this channel distortion. It is shown in [5] that if the MDF is larger than 8% of the subcarrier spacing, the signal-to-ICI plus noise (SINR) ratio is less than 20dB. Hence, in order to obtain reliable reception, there is a need for efficient algorithms to combat the ICI effect in a mobile OFDM receiver.

A wide variety of schemes for ICI mitigation have been proposed, mainly consisting of ICI self-cancellation, blind equalization, and ICI cancellation-based equalization [5] – [19]. At the expense of reduced bandwidth efficiency, the ICI self-cancellation scheme is simple and effective to provide good BER performance [6], [7]. The scheme, however, is not suitable for existing standards as modification to transmit formats is required. In contrast, the blind equalization scheme is efficient in saving bandwidth but it involves high computation complexity [8]. Among the three ICI mitigation schemes, the ICI cancellation-based equalization scheme is the most common [9] – [19]. Based on zero-forcing (ZF) or minimum mean-square-error (MMSE) criterion, two optimal frequency-domain equalizers are derived in [9] – [12]. To enhance the performance, successive interference cancellation (SIC) with optimal ordering can be incorporated with the MMSE equalizer [13]. Several works, like [5] and [14]–[16], are targeted toward reducing the complexity of frequency-domain equalizers. By ignoring small ICI terms, a partial MMSE equalizer is proposed in [14] to avoid the inversion of a large-size matrix, while a recursive algorithm is developed in [5] for calculation of equalizer coefficients. Moreover, [15] incorporates a partial MMSE equalizer with SIC, and [16] combines the partial MMSE equalizer with BICM. Both methods benefit greatly from time diversity gains induced by mobility. We also find two decision-feedback (DF) equalizers in [17], [18], which make use of power series expansion on time-variant frequency response. Apart from using frequency-domain equalizers, [9] and [19] consider time-domain equalizers which first achieve ICI shortening, followed by MMSE detection and parallel interference cancellation, respectively, to remove the residual ICI.

For successful implementation of the ICI cancellation-based equalization, it is essential to obtain an accurate estimate of channel variation or the equivalent ICI channel matrix. In

Manuscript received April 29, 2009; revised October 29, 2009 and May 31, 2010; accepted February 10, 2011. The associate editor coordinating the review of this paper and approving it for publication is F. Takawira.

M. L. Ku is with the Department of Communication Engineering, National Central University, Taoyuan, 320, Taiwan (e-mail: mlku@ce.ncu.edu.tw).

W. C. Chen is with the Institute for Information Industry, Taipei, 106, Taiwan (e-mail: tata9876@gmail.com).

C. C. Huang is with the Department of Electrical Engineering, National Chiao Tung University, Hsinchu, 300, Taiwan (e-mail: huangcc@faculty.nctu.edu.tw).

Digital Object Identifier 10.1109/TWC.2011.030911.090611

general, this can be accomplished through the use of embedded reference signals such as pilot symbols or pilot tones. In [13], an MMSE estimator, which demands frequent pilot symbols inserted among OFDM data symbols, is proposed to estimate time-variant channel impulse response (CIR). As complexity is concerned, most studies model the time variation of each channel tap as a polynomial function. By assuming CIR varies in a linear fashion within an OFDM symbol, [14] and [17] exploit pilot symbols for parameter estimation, whereas [18] and [20] belong to the category which uses pilot tones. It is concluded that a linear model is valid for capturing channel dynamics with the normalized MDF up to 0.1. When normalized MDF is larger than 0.1, a 2-D polynomial surface function is suggested in [10] to model time-varying channel frequency response and to gain better performance. Since most currently developed communication standards, e.g., 802.16e and long-term evolution (LTE), aim at providing mobility with the normalized MDF up to about 0.1 which is regarded as a fairly fast-fading scenario for mobile environments, we also consider using a linear model for approximating channel time variation in this paper.

The expectation-maximization (EM) algorithm can facilitate solving the maximum-likelihood (ML) estimation problem in an iterative manner which alternates between an E-step, calculating an expected complete log-likelihood (ECLL) function, and an M-step, maximizing the ECLL function with respect to some unknown parameters [21]. Recently, a few EM-based methods have been proposed for channel estimation (CE) and data detection in OFDM systems [12], [22]–[24]. The major difference among these methods lies in whether they formulate the original ML problem into a data sequence detection problem or a channel variable estimation problem. Yet, the wireless channel is assumed to be quasi-static in [22]–[24], i.e., channel gain remains constant over the duration of one OFDM symbol. Even though the EM-based channel estimation scheme in [12] can be implemented in time-varying fading channels, the proposed EM scheme can only estimate average channel gains.

In this paper, we investigate two EM-based iterative receivers for OFDM and BICM/OFDM systems in doubly selective fading channels. By assuming channel varies in a linear fashion, we first analyze the ICI effect in frequency domain and derive a data detection method based on the EM algorithm using the ML criterion. In an effort to reduce complexity, groupwise processing is adopted for the two EM-based receivers. For OFDM systems, we implement an ML-EM receiver which iterates between a groupwise ICI canceller and an EM detector. Based on this receiver structure, a TURBO-EM receiver for BICM/OFDM systems is then proposed to successively improve the performance by applying the turbo principle. Finally, for the initial setting of the two receivers, MMSE-based CE is first performed by using a few pilot tones and it is later improved via the decision feedback methodology. To the best of our knowledge, this is the first work studying an EM approach for joint ICI estimation and data detection in multipath time-varying channels. Our work differs from the previous EM-based approaches [12], [22]–[24] in two aspects. First, the works in [22]–[24] do not take into account the ICI problem, and they can only be

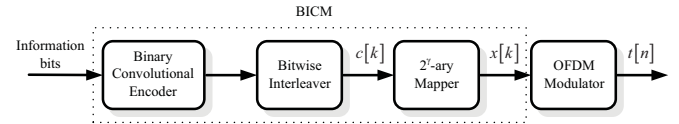


Fig. 1. BICM/OFDM systems.

applied in quasi-static channels. Second, although the work in [12] can be extended to incorporate the ICI approximation and cancellation for data detection in time-varying channels, the ICI estimation and data detection schemes therein are designed separately. In other words, the channel mean values between two consecutive OFDM symbols are first estimated to construct the ICI matrix through linear interpolation, and it is followed by an MMSE data detector.

The rest of this paper is organized as follows. In Section II, we describe the OFDM and BICM/OFDM systems, followed by the analysis of ICI in frequency domain. According to the frequency domain ICI model, an EM-based data detection method is developed in Section III. In Section IV, an ML-EM receiver and a TURBO-EM receiver are proposed. Afterwards, we describe the initialization procedure of the two receivers and discuss their computation complexity. In Section V, we present our computer simulation results. Finally, some conclusions are drawn in Section VI.

Notation: Superscripts $(\cdot)^T$ and $(\cdot)^\dagger$ stand for transpose and Hermitian transpose, respectively. Column vectors and matrices are denoted by boldface lowercase and uppercase letters, respectively. The notation $E[\cdot]$ takes expectation. We use \mathbf{I}_K to denote a $K \times K$ identity matrix and $\text{diag}\{\mathbf{x}\}$ to denote a diagonal matrix with \mathbf{x} on its diagonal. The notations $((\cdot))_N$ and \odot present the modulo- N and Hadamard product operation, respectively. The notation $\{\cdot\}$ denotes a set, e.g. a set $\mathbf{x} = \{x_1, \dots, x_N\}$. Further, we denote the vector norm as $\|\cdot\|$.

II. SYSTEM MODEL

A. Transmitted and Received Signals

Fig. 1 shows a BICM/OFDM system, where information bits are modulated by BICM along with an OFDM modulator [2]. Data symbols are generated by concatenating a binary convolutional encoder with a 2^γ -ary mapper through a bitwise interleaver (denoted as $\mathbf{\Pi}$). Throughout this paper, we only consider binary phase shift keying (BPSK) modulation ($\gamma = 1$); therefore, data symbols are one-to-one mapped from coded bits. Subsequently, these data symbols are transmitted over N_F consecutive OFDM symbols. Let $x[k]$ be the data symbols to be transmitted over the k th subcarrier for an OFDM symbol. After modulated by an N -point inverse discrete Fourier transform (IDFT) and appended with GI of length N_G , time domain samples of an OFDM symbol are given by

$$t[n] = \frac{1}{N} \sum_{k=0}^{N-1} x[k] e^{j \frac{2\pi kn}{N}} \quad (1)$$

for $n = -N_G, \dots, N - 1$, where we assume that $x[k]$ is mapped from the coded bit stream $c[k]$, and GI considered here is a cyclic prefix.

At the receiver, by removing the GI and taking the discrete Fourier transform (DFT), the demodulated signal in frequency domain is given by [20]:

$$y[k] = H[k, k]x[k] + \underbrace{\sum_{m=0, m \neq k}^{N-1} H[k, m]x[m]}_{\text{ICI term}} + z[k] \quad (2)$$

where $H[k, m] = \sum_{l=0}^{L-1} \alpha[k, m, l] e^{-j \frac{2\pi ml}{N}}$ represents the leakage term of ICI from the m th subcarrier to the k th subcarrier, $k = 0, \dots, N-1$, $\alpha[k, m, l] = (1/N) \sum_{n=0}^{N-1} h[l, n] e^{-j \frac{2\pi n((k-m))_N}{N}}$ is the DFT of a time series $h[l, n]$ corresponding to the l th channel tap at time delay n , for $l = 0, \dots, L-1$ and $n = 0, \dots, N-1$, and $z[k]$ is the additive white Gaussian noise with zero-mean and variance σ_z^2 . Moreover, we assume that the channel tap $h[l, n]$ for different l is an independent and identically distributed (i.i.d.) complex Gaussian random variable with zero mean and variance Ξ_l . From (2), we can observe that a demodulated subcarrier is affected by the ICI contributed from all the other subcarriers, and this effect severely degrades the system performance if a conventional one-tap equalizer is employed [5].

B. Modeling of ICI in Frequency Domain

We adopt a linear function to model the temporal variation of each channel tap over an OFDM symbol, as follows:

$$h[l, n] = a[l, 1]n + a[l, 0] \quad (3)$$

for $l = 0, \dots, L-1$ and $n = 0, \dots, N-1$, where $a[l, p]$ is the complex coefficient of the p th order for the l th tap. Substituting (3) into $\alpha[k, m, l]$ of (2), we can obtain

$$\alpha[k, m, l] = \begin{cases} \frac{N-1}{2} a[l, 1] + a[l, 0], & \text{for } k = m \\ \Phi[k, m] a[l, 1], & \text{otherwise} \end{cases} \quad (4)$$

where $\Phi[k, m]$ can be derived as

$$\begin{aligned} \Phi[k, m] &= \frac{1}{N} \sum_{n=0}^{N-1} n e^{-j \frac{2\pi n(k-m)}{N}} \\ &= -\frac{1}{2} + j \frac{1}{2 \tan(\pi((k-m))_N/N)} \end{aligned} \quad (5)$$

According to the fact of $1 \leq ((k-m))_N \leq N-1$, we observe that the value of $\frac{\pi((k-m))_N}{N}$ ranges from $\frac{\pi}{N}$ to $\frac{(N-1)\pi}{N}$. Apply the Maclaurin series of $\tan(x) \approx x$, for $|x| < \frac{\pi}{2}$, and after some straightforward derivation, we can represent $\Phi[k, m]$ as

$$\begin{aligned} \Phi[k, m] &\approx \begin{cases} -\frac{1}{2}, & \text{for } ((k-m))_N = \frac{N}{2} \\ -\frac{1}{2} + \frac{jN}{2\pi((k-m))_N}, & \text{for } 1 \leq ((k-m))_N < \frac{N}{2} \\ -\frac{1}{2} + \frac{jN}{2\pi((k-m))_N - N}, & \text{for } \frac{N}{2} < ((k-m))_N \leq N-1 \end{cases} \quad (6) \end{aligned}$$

From (6), it follows that $\Phi[k, m]$ is a fixed value, which only depends on $(k-m)$ modulo N and it can be calculated in advance. By using (4) and (6), (2) can be rewritten as

$$y[k] = H[k, k]x[k] + \underbrace{\sum_{m=0, m \neq k}^{N-1} \Phi[k, m] w[m] x[m]}_{\text{ICI term}} + z[k] \quad (7)$$

where $w[m] = \sum_{l=0}^{L-1} a[l, 1] e^{-j \frac{2\pi ml}{N}}$ defines a new channel variable in frequency domain. It is worthy to mention that in orthogonal frequency division multiple access (OFDMA) systems, each user merely detects a set of nearby subcarriers of interest (e.g. zones or clusters in 802.16e), instead of all the N subcarriers [3]. With the formulation of (7), one can deal only with a small number of channel variables even when the number of channel taps is large. Finally, we can rewrite (7) in a matrix notation, leading to a more compact representation:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{z} = (\mathbf{M} + \mathbf{\Phi}\mathbf{W})\mathbf{x} + \mathbf{z} = \mathbf{M}\mathbf{x} + \widehat{\mathbf{\Phi}}\mathbf{w} + \mathbf{z} \quad (8)$$

where $\mathbf{y} = [y[0], \dots, y[N-1]]^T$, the (k, m) th entry of \mathbf{H} is $H[k, m]$, $\mathbf{w} = [w[0], \dots, w[N-1]]^T$, $\mathbf{x} = [x[0], \dots, x[N-1]]^T$, $\mathbf{z} = [z[0], \dots, z[N-1]]^T$, $\mathbf{W} = \text{diag}\{\mathbf{w}\}$, $\mathbf{M} = \text{diag}\{[H[0, 0], \dots, H[N-1, N-1]]^T\}$, the (k, m) th entry of $\mathbf{\Phi}$ is just $\Phi[k, m]$, the (k, m) th entry of $\widehat{\mathbf{\Phi}}$ is given by $\Phi[k, m]X[m]$, the value of $\Phi[k, m]$ for $k = m$ is defined as zero. Moreover, we have $\mathbf{w} = \mathbf{F}\mathbf{s}$, where $\mathbf{s} = [a[0, 1], \dots, a[L-1, 1]]^T$ and \mathbf{F} is a DFT matrix of size $N \times L$, with the (m, l) th entry given by $e^{-j \frac{2\pi ml}{N}}$.

III. PROPOSED EM-BASED DATA DETECTION METHOD

From (8), the optimum ML data detection problem can be formulated as follows:

$$\begin{aligned} \mathbf{x}_{ML} &= \arg \max_{\mathbf{x} \in \{1, -1\}^N} L(\mathbf{y}|\mathbf{x}) \\ &= \arg \max_{\mathbf{x} \in \{1, -1\}^N} \log \int P(\mathbf{y}|\mathbf{w}, \mathbf{x}) P(\mathbf{w}) d\mathbf{w} \end{aligned} \quad (9)$$

where $L(\cdot)$ is a log-likelihood function, obtained by taking logarithm of the corresponding probability density function (PDF) $P(\cdot)$. Direct calculation using (9), however, involves multidimensional integration over the hidden variable \mathbf{w} . With the ability to tackle missing data models, the EM algorithm is considered as a good alternative to solve (9), and the core idea behind this algorithm is to iterate between E-step and M-step such that monotonic increase in $L(\mathbf{y}|\mathbf{x})$ is obtained. More details of the algorithm and its application can be found in [21]. The E-step and the M-step associated with the optimization problem of (9) are expressed respectively as

$$\Omega(\mathbf{x}|\mathbf{y}, \hat{\mathbf{x}}^{(m-1)}) = E_{\mathbf{w}|\mathbf{y}, \hat{\mathbf{x}}^{(m-1)}} [L(\mathbf{y}, \mathbf{w}|\mathbf{x})] \quad (10)$$

$$\hat{\mathbf{x}}^{(m)} = \arg \max_{\mathbf{x} \in \{1, -1\}^N} \Omega(\mathbf{x}|\mathbf{y}, \hat{\mathbf{x}}^{(m-1)}) \quad (11)$$

where $\hat{\mathbf{x}}^{(m)}$ denotes the hard decision of \mathbf{x} at the m th EM iteration, and $\Omega(\mathbf{x}|\mathbf{y}, \hat{\mathbf{x}}^{(m-1)})$ is known as the ECLL function, to be maximized in the M-step of (11). By using the fact that $L(\mathbf{y}, \mathbf{w}|\mathbf{x}) = L(\mathbf{y}|\mathbf{w}, \mathbf{x}) + L(\mathbf{w})$ and from (8), we can further simplify (10) as

$$\begin{aligned} \Omega(\mathbf{x}|\mathbf{y}, \hat{\mathbf{x}}^{(m-1)}) &= E_{\mathbf{w}|\mathbf{y}, \hat{\mathbf{x}}^{(m-1)}} [L(\mathbf{y}|\mathbf{w}, \mathbf{x})] + C \\ &= E_{\mathbf{w}|\mathbf{y}, \hat{\mathbf{x}}^{(m-1)}} \left[\frac{-1}{\sigma_z^2} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 \right] + C \\ &= \frac{-1}{\sigma_z^2} (\mathbf{y}^\dagger \mathbf{y} - \mathbf{y}^\dagger \tilde{\mathbf{H}}\mathbf{x} - \mathbf{x}^\dagger \tilde{\mathbf{H}}^\dagger \mathbf{y} + \mathbf{x}^\dagger \tilde{\Sigma} \mathbf{x}) + C \end{aligned} \quad (12)$$

where $\tilde{\mathbf{H}}$ and $\tilde{\Sigma}$ denote $E_{\mathbf{w}|\mathbf{y}, \hat{\mathbf{x}}^{(m-1)}} [\mathbf{H}]$ and $E_{\mathbf{w}|\mathbf{y}, \hat{\mathbf{x}}^{(m-1)}} [\mathbf{H}^\dagger \mathbf{H}]$, respectively, and the constant term

C in (12) can be dropped for simplicity. Without loss of generality, the channel state information (CSI) \mathbf{M} can be estimated through pilot tones embedded in each OFDM symbol, and we denote the estimate as $\hat{\mathbf{M}}$. By inserting $\mathbf{H} = \hat{\mathbf{M}} + \Phi\mathbf{W}$ into $\hat{\mathbf{H}}$ and $\hat{\Sigma}$, it is straightforward to calculate the two terms as

$$\tilde{\mathbf{H}} = E_{\mathbf{w}|\mathbf{y},\hat{\mathbf{x}}^{(m-1)}} \left[\hat{\mathbf{M}} + \Phi\mathbf{W} \right] = \hat{\mathbf{M}} + \Phi\tilde{\mathbf{W}} \quad (13)$$

$$\begin{aligned} \tilde{\Sigma} &= E_{\mathbf{w}|\mathbf{y},\hat{\mathbf{x}}^{(m-1)}} \left[\left(\hat{\mathbf{M}} + \Phi\mathbf{W} \right)^\dagger \left(\hat{\mathbf{M}} + \Phi\mathbf{W} \right) \right] \\ &= E_{\mathbf{w}|\mathbf{y},\hat{\mathbf{x}}^{(m-1)}} \left[\hat{\mathbf{M}}^\dagger \hat{\mathbf{M}} + \hat{\mathbf{M}}^\dagger \Phi\mathbf{W} + \mathbf{W}^\dagger \Phi^\dagger \hat{\mathbf{M}} \right. \\ &\quad \left. + \mathbf{W}^\dagger \Phi^\dagger \Phi\mathbf{W} \right] \\ &= \hat{\mathbf{M}}^\dagger \hat{\mathbf{M}} + \hat{\mathbf{M}}^\dagger \Phi\tilde{\mathbf{W}} + \tilde{\mathbf{W}}^\dagger \Phi^\dagger \hat{\mathbf{M}} \\ &\quad + E_{\mathbf{w}|\mathbf{y},\hat{\mathbf{x}}^{(m-1)}} \left[(\mathbf{w}\mathbf{w}^\dagger)^T \odot (\Phi^\dagger \Phi) \right] \\ &= \hat{\mathbf{M}}^\dagger \hat{\mathbf{M}} + \hat{\mathbf{M}}^\dagger \Phi\tilde{\mathbf{W}} + \tilde{\mathbf{W}}^\dagger \Phi^\dagger \hat{\mathbf{M}} \\ &\quad + \left(E_{\mathbf{w}|\mathbf{y},\hat{\mathbf{x}}^{(m-1)}} [\mathbf{w}\mathbf{w}^\dagger] \right)^T \odot (\Phi^\dagger \Phi) \end{aligned} \quad (14)$$

where $\tilde{\mathbf{W}}$ is defined as $E_{\mathbf{w}|\mathbf{y},\hat{\mathbf{x}}^{(m-1)}} [\mathbf{W}]$. Let $\tilde{\mathbf{W}} \triangleq \text{diag}\{\tilde{\mathbf{w}}\}$, $\tilde{\mathbf{w}} \triangleq E_{\mathbf{w}|\mathbf{y},\hat{\mathbf{x}}^{(m-1)}} [\mathbf{w}]$ and $\tilde{\Sigma}_{\mathbf{w}} \triangleq E_{\mathbf{w}|\mathbf{y},\hat{\mathbf{x}}^{(m-1)}} \left[(\mathbf{w} - \tilde{\mathbf{w}})(\mathbf{w} - \tilde{\mathbf{w}})^\dagger \right] = E_{\mathbf{w}|\mathbf{y},\hat{\mathbf{x}}^{(m-1)}} [\mathbf{w}\mathbf{w}^\dagger] - \tilde{\mathbf{w}}\tilde{\mathbf{w}}^\dagger$. We can rewrite (14) as

$$\tilde{\Sigma} = \hat{\mathbf{M}}^\dagger \hat{\mathbf{M}} + \hat{\mathbf{M}}^\dagger \Phi\tilde{\mathbf{W}} + \tilde{\mathbf{W}}^\dagger \Phi^\dagger \hat{\mathbf{M}} + \left(\tilde{\Sigma}_{\mathbf{w}} + \tilde{\mathbf{w}}\tilde{\mathbf{w}}^\dagger \right)^T \odot (\Phi^\dagger \Phi) \quad (15)$$

Also, from (8), it is observed that the conditional PDF $P(\mathbf{w}|\mathbf{y},\hat{\mathbf{x}}^{(m-1)})$ is a Gaussian distribution, with mean and covariance given by [25]

$$\tilde{\mathbf{w}} = \boldsymbol{\mu}_{\mathbf{w}} + \mathbf{C}_{\mathbf{w}\mathbf{y}} \mathbf{C}_{\mathbf{y}\mathbf{y}}^{-1} (\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}}) \quad (16)$$

$$\tilde{\Sigma}_{\mathbf{w}} = \mathbf{C}_{\mathbf{w}\mathbf{w}} - \mathbf{C}_{\mathbf{w}\mathbf{y}} \mathbf{C}_{\mathbf{y}\mathbf{y}}^{-1} \mathbf{C}_{\mathbf{y}\mathbf{w}} \quad (17)$$

where the relevant terms are defined and statistics are calculated in the following way. We first apply a first-order autoregressive (AR) channel model to compute $\boldsymbol{\mu}_{\mathbf{w}}$ and $\mathbf{C}_{\mathbf{w}\mathbf{w}}$. Details are provided in Appendix A, and the two terms can be derived as

$$\boldsymbol{\mu}_{\mathbf{w}} = E[\mathbf{w}] = \mathbf{0} \quad (18)$$

$$\mathbf{C}_{\mathbf{w}\mathbf{w}} = E \left[(\mathbf{w} - \boldsymbol{\mu}_{\mathbf{w}})(\mathbf{w} - \boldsymbol{\mu}_{\mathbf{w}})^\dagger \right] = \mathbf{F}\mathbf{C}_{\text{ss}}\mathbf{F}^\dagger \quad (19)$$

where \mathbf{C}_{ss} is a diagonal matrix with the l th diagonal entry equal to $\frac{2(1-\alpha)\Xi_l}{(N-1)^2}$, and α is the channel tap autocorrelation as defined in (A.2). Moreover, we can get

$$\boldsymbol{\mu}_{\mathbf{y}} = E[\mathbf{y}] = \hat{\mathbf{M}}\hat{\mathbf{x}}^{(m-1)} \quad (20)$$

$$\begin{aligned} \mathbf{C}_{\mathbf{y}\mathbf{y}} &= E \left[(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})^\dagger \right] \\ &= \hat{\Phi}^{(m-1)} \mathbf{C}_{\mathbf{w}\mathbf{w}} \hat{\Phi}^{(m-1)\dagger} + \sigma_z^2 \mathbf{I}_N \end{aligned} \quad (21)$$

Initialization:

Calculate $\hat{\mathbf{M}}$, choose $\hat{\mathbf{x}}^{(0)}$, and set $m = 0$

Execution of EM algorithm:

do {

$m = m + 1$

E-step:

Compute statistics

$$\mathbf{C}_{\mathbf{w}\mathbf{w}} = \mathbf{F}\mathbf{C}_{\text{ss}}\mathbf{F}^\dagger$$

$$\tilde{\mathbf{w}} = \mathbf{C}_{\mathbf{w}\mathbf{w}} \hat{\Phi}^{(m-1)\dagger} \left(\hat{\Phi}^{(m-1)} \mathbf{C}_{\mathbf{w}\mathbf{w}} \hat{\Phi}^{(m-1)\dagger} + \sigma_z^2 \mathbf{I}_N \right)^{-1} (\mathbf{y} - \hat{\mathbf{M}}\hat{\mathbf{x}}^{(m-1)})$$

$$\tilde{\Sigma}_{\mathbf{w}} = \mathbf{C}_{\mathbf{w}\mathbf{w}} - \mathbf{C}_{\mathbf{w}\mathbf{w}} \hat{\Phi}^{(m-1)\dagger} \left(\hat{\Phi}^{(m-1)} \mathbf{C}_{\mathbf{w}\mathbf{w}} \hat{\Phi}^{(m-1)\dagger} + \sigma_z^2 \mathbf{I}_N \right)^{-1} \hat{\Phi}^{(m-1)} \mathbf{C}_{\mathbf{w}\mathbf{w}}$$

$$\tilde{\mathbf{H}} = \hat{\mathbf{M}} + \Phi\tilde{\mathbf{W}}$$

$$\tilde{\Sigma} = \hat{\mathbf{M}}^\dagger \hat{\mathbf{M}} + \hat{\mathbf{M}}^\dagger \Phi\tilde{\mathbf{W}} + \tilde{\mathbf{W}}^\dagger \Phi^\dagger \hat{\mathbf{M}} + \left(\tilde{\Sigma}_{\mathbf{w}} + \tilde{\mathbf{w}}\tilde{\mathbf{w}}^\dagger \right)^T \odot (\Phi^\dagger \Phi)$$

For all $\mathbf{x} \in \{1, -1\}^N$, calculate

$$\Omega(\mathbf{x}|\mathbf{y},\hat{\mathbf{x}}^{(m-1)}) = -\frac{1}{\sigma_z^2} (\mathbf{y}^\dagger \mathbf{y} - \mathbf{y}^\dagger \tilde{\mathbf{H}}\mathbf{x} - \mathbf{x}^\dagger \tilde{\mathbf{H}}^\dagger \mathbf{y} + \mathbf{x}^\dagger \tilde{\Sigma} \mathbf{x})$$

M-step:

$$\hat{\mathbf{x}}^{(m)} = \arg \max_{\mathbf{x} \in \{1, -1\}^N} \Omega(\mathbf{x}|\mathbf{y},\hat{\mathbf{x}}^{(m-1)})$$

} while $(\hat{\mathbf{x}}^{(m)} \neq \hat{\mathbf{x}}^{(m-1)})$ and $m \leq N_{EM}$

Fig. 2. EM-based data detection method.

$$\mathbf{C}_{\mathbf{w}\mathbf{y}} = E \left[(\mathbf{w} - \boldsymbol{\mu}_{\mathbf{w}})(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})^\dagger \right] = \mathbf{C}_{\mathbf{w}\mathbf{w}} \hat{\Phi}^{(m-1)\dagger} \quad (22)$$

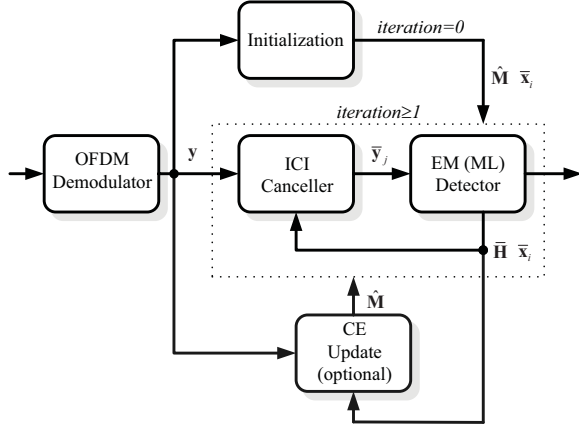
where $\mathbf{C}_{\mathbf{w}\mathbf{y}} = \mathbf{C}_{\mathbf{y}\mathbf{w}}^\dagger$, and $\hat{\Phi}^{(m-1)}$ is obtained by substituting the hard decision $\hat{\mathbf{x}}^{(m-1)}$ into $\hat{\Phi}$. Using (13)-(22), we can calculate the ECLL function of (12). The EM algorithm for data detection is then summarized in Fig. 2, and it is repeated until a stopping criterion holds. The stopping criterion is to check whether $\hat{\mathbf{x}}^{(m)} = \hat{\mathbf{x}}^{(m-1)}$ or the iteration number reaches a predefined limit N_{EM} .

IV. IMPLEMENTATION: EM-BASED ITERATIVE RECEIVERS

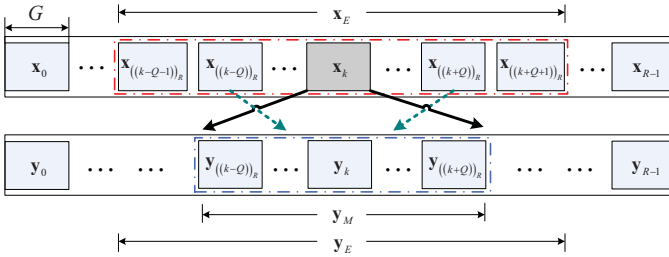
The data detection method in (10) and (11) involves not only $N \times N$ matrix inversion in the E-step but also heuristic search over 2^N lattices in the M-step, thereby making computation intractable in practice. In this section, we investigate two EM-based iterative receivers for practical implementation.

A. ML-EM Receiver for OFDM Systems

As depicted in Fig. 3, we consider an ML-EM receiver with N subcarriers partitioned into R groups, and each group consists of G subcarriers. Denote the j th group of subcarriers as $\mathbf{G}_j = \{jG, \dots, (j+1)G - 1\}$, for $j = 0, \dots, R - 1$. Next, we define the j th data group and observation group as $\mathbf{x}_j = [x[jG], \dots, x[(j+1)G - 1]]^T$ and $\mathbf{y}_j = [y[jG], \dots, y[(j+1)G - 1]]^T$, respectively. Further, we use \mathbf{B}_j to denote the set $\{((j-Q))_R, \dots, ((j+Q))_R\}$. Without loss of generality, we focus on detecting the k th data group. Assume that due to the ICI effect, the energy



(a) ML-EM receiver for OFDM systems.



1. Calculate statistics of \mathbf{w} , by using \mathbf{x}_E and \mathbf{y} .
2. Calculate ECLL function, by using \mathbf{y}_M .

(b) An illustration for group detection.

Fig. 3. (a) ML-EM receiver for OFDM systems; (b) an illustration for group detection.

of \mathbf{x}_k is spread over $2Q + 1$ observation groups of \mathbf{y}_j , for $j \in \mathbf{B}_k$, which also contains interfering energy caused by other adjacent data groups \mathbf{x}_j , for $j \in \mathbf{B}_j \setminus \{k\}$. As observed in Fig. 3(a), there is an additional iteration loop outside the EM detector, called ML iteration. Within an ML iteration, the ICI is first reconstructed and subtracted from the observation group, yielding a signal:

$$\bar{\mathbf{y}}_j = \mathbf{y}_j - \sum_{i \in \mathbf{B}_k \setminus \{k\}} \bar{\mathbf{H}}_{j,i} \bar{\mathbf{x}}_i \quad (23)$$

for $j \in \mathbf{B}_k$, where $\bar{\mathbf{x}}_i$ is the tentative decision of \mathbf{x}_i , and the (p, q) th entry of $\bar{\mathbf{H}}_{j,i}$ is given by the $(jG + p, iG + q)$ th entry of $\bar{\mathbf{H}}$, the estimate of \mathbf{H} , for $p, q = 0, \dots, G - 1$. Both $\bar{\mathbf{x}}_i$ and $\bar{\mathbf{H}}$ are obtained from the output of the EM detector at the previous ML iteration. After ICI cancellation, the EM detector is executed by applying the EM-based data detection method in Section III. Define $\mathbf{x}_E = [\bar{\mathbf{x}}_{((k-Q-1))_R}^T, \dots, \hat{\mathbf{x}}_k^{(m-1)T}, \dots, \bar{\mathbf{x}}_{((k+Q+1))_R}^T]^T$ and $\mathbf{y}_E = [\mathbf{y}_{((k-Q-1))_R}^T, \dots, \mathbf{y}_{((k+Q+1))_R}^T]^T$, where $\hat{\mathbf{x}}_k^{(m-1)}$ is the hard decision of \mathbf{x}_k at the $(m-1)$ th EM iteration within the EM detector. Particularly, for $m = 1$, we initialize $\hat{\mathbf{x}}_k^{(0)}$ as $\bar{\mathbf{x}}_k$. In the E-step, at the m th EM iteration, we replace \mathbf{x} and \mathbf{y} (in Fig. 2) with \mathbf{x}_E and \mathbf{y}_E to calculate the statistics $\bar{\mathbf{W}}$, $\bar{\mathbf{H}}$ and $\bar{\Sigma}$. The size of these three matrices now becomes $(2Q + 3)G \times (2Q + 3)G$. After that, the interference-reduced signal $\mathbf{y}_M = [\bar{\mathbf{y}}_{((k-Q)_R)}^T, \dots, \bar{\mathbf{y}}_{((k+Q)_R)}^T]^T$ is taken to compute the ECLL function, for each combination of

$\mathbf{x}_k \in \{1, -1\}^G$, as follows:

$$\Omega(\mathbf{x}_k | \mathbf{y}_M, \mathbf{y}_E, \mathbf{x}_E) = \frac{-1}{\sigma_z^2} \left(\mathbf{y}_M^\dagger \mathbf{y}_M - \mathbf{y}_M^\dagger \bar{\mathbf{H}}_k \mathbf{x}_k - \mathbf{x}_k^\dagger \bar{\mathbf{H}}_k^\dagger \mathbf{y}_M + \mathbf{x}_k^\dagger \bar{\Sigma}_k \mathbf{x}_k \right) \quad (24)$$

where the matrices $\bar{\mathbf{H}}_k$ and $\bar{\Sigma}_k$ are of size $(2Q + 1)G \times G$ and $G \times G$, with the (p, q) th entry given by the $(G + p, (Q + 1)G + q)$ th entry of $\bar{\mathbf{H}}$ and the $((Q + 1)G + p, (Q + 1)G + q)$ th entry of $\bar{\Sigma}$, respectively. Finally, the decision of \mathbf{x}_k is calculated in the M-step according to:

$$\hat{\mathbf{x}}_k^{(m)} = \arg \max_{\mathbf{x}_k \in \{1, -1\}^G} \Omega(\mathbf{x}_k | \mathbf{y}_M, \mathbf{y}_E, \mathbf{x}_E) \quad (25)$$

Within the EM detector, the above procedure is conducted to detect R groups simultaneously, i.e., we use parallel processing for group detection. Once the stopping criterion is met, the receiver proceeds to the next ML iteration until a good performance is achieved, and $\bar{\mathbf{x}}_k$ and $\bar{\mathbf{H}}$ are updated. In other words, at the end of the k th parallel processing, $\bar{\mathbf{x}}_k$ is replaced by $\hat{\mathbf{x}}_k^{(m)}$, the $(kG + j)$ th diagonal entry of $\bar{\mathbf{W}}$ is renewed by the $((Q + 1)G + j)$ th diagonal entry of $\bar{\mathbf{W}}$, for $k = 0, \dots, R - 1$ and $j = 0, \dots, G - 1$, and $\bar{\mathbf{H}}$ is calculated as $\bar{\mathbf{M}} + \Phi \bar{\mathbf{W}}$.

The intuition behind the group detection is explained as follows. While computing and maximizing the ECLL function, we can acquire the diversity gains through examining the interference-reduced signals \mathbf{y}_M , of which the energy is contributed mainly by the data group \mathbf{x}_k . Therefore, it is reasonable to expect that full diversity gain is achievable when the value of Q is sufficiently large and the ICI is perfectly cancelled out. The diversity gain we can achieve also depends upon whether a good estimate of $\bar{\mathbf{W}}$ is provided. Recall that the $2Q + 1$ observation groups (in the neighborhood of \mathbf{y}_k) are sufficient for estimating the channel variables of the k th group, but they also include ICI from adjacent data groups. For example, the $((k - Q))_R$ th observation group is interfered by the data groups \mathbf{x}_j , for $j = ((k - 2Q))_R, \dots, k$. In order to obtain accurate estimate of channel variables, it is necessary to count the ICI effect from the adjacent data groups. Hence, in the E-step, we take an enlarged cluster of original observation groups \mathbf{y}_E , as well as the corresponding data groups \mathbf{x}_E , to calculate the channel variables of the k th group. Strictly speaking, the size of \mathbf{y}_E and \mathbf{x}_E should be chosen as $4Q + 1$, but our experimental trials suggest that the choice of $2Q + 3$ is large enough to get a good result. In other words, it means that the adjacent two data groups (outside the $2Q + 1$ observation groups) cause the main power of interference to the $2Q + 1$ observation groups, and the ICI contributed from the other adjacent data groups (outside the $2Q + 1$ observation groups) have negligible influence on the accuracy of estimation of channel variables because it is relatively small.

B. TURBO-EM Receiver for BICM/OFDM Systems

Fig. 4 shows the TURBO-EM receiver for BICM/OFDM systems. The receiver implements the turbo iterations by exchanging the extrinsic information between the EM detector (after the ICI canceller) and the soft-output Viterbi algorithm (SOVA) decoder. Within each TURBO iteration, the ICI is

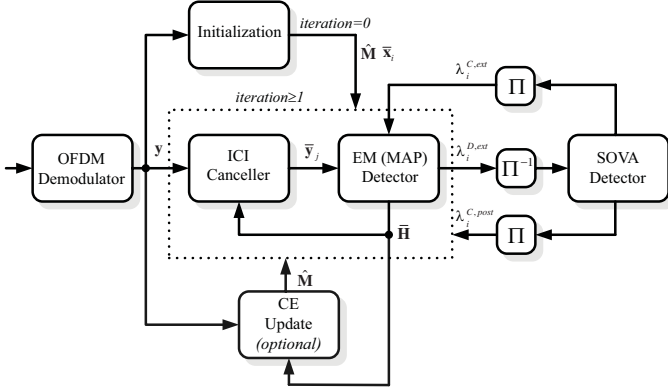


Fig. 4. TURBO-EM receiver for BICM/OFDM systems.

first reconstructed and subtracted from the observation group to obtain the interference-reduced signal \bar{y}_j by using (23), but with soft decision \bar{x}_i for interference reconstruction. In this way, we can mitigate the error propagation effect, and the soft decision for the BPSK case is given by [26]

$$\bar{x}_i = E[\mathbf{x}_i] = \tanh\left(\frac{\lambda_i^{C,post}}{2}\right) \quad (26)$$

where $\lambda_i^{C,post} = [\lambda^{C,post}[iG], \dots, \lambda^{C,post}[(i+1)G-1]]^T$ is the *a posteriori* log-likelihood ratio (LLR), associated with \mathbf{x}_i , from the SOVA decoder at the previous TURBO iteration, and the LLR of a data symbol ϑ is defined as the logarithm of the ratio of $P(\vartheta = +1)$ to $P(\vartheta = -1)$. Before applying the SOVA decoder, we execute the EM detector with several EM iterations by using the interference-reduced signal \bar{y}_j from the ICI canceller. Here, we apply the maximum *a posteriori* (MAP) EM algorithm to the EM detector, which further takes account of the *a priori* information to compute the ECLL function as follows:

$$\begin{aligned} \Omega(\mathbf{x}_k | \mathbf{y}_M, \mathbf{y}_E, \mathbf{x}_E) \\ = \frac{-1}{\sigma_z^2} \left(\mathbf{y}_M^\dagger \mathbf{y}_M - \mathbf{y}_M^\dagger \tilde{\mathbf{H}}_k \mathbf{x}_k - \mathbf{x}_k^\dagger \tilde{\mathbf{H}}_k^\dagger \mathbf{y}_M + \mathbf{x}_k^\dagger \tilde{\Sigma}_k \mathbf{x}_k \right) \\ + L(\mathbf{x}_k) \end{aligned} \quad (27)$$

where $L(\mathbf{x}_k) = \ln P(\mathbf{x}_k)$ is calculated from the interleaved extrinsic information $\lambda_k^{C,ext} = [\lambda^{C,ext}[kG], \dots, \lambda^{C,ext}[(k+1)G-1]]^T$ with respect to \mathbf{x}_k which is generated by the SOVA decoder. Under the assumption of an ideal interleaver, data symbols are independent of each other, and we obtain

$$L(\mathbf{x}_k) = \sum_{j=0}^{G-1} L(x[kG+j] = q_j) \quad (28)$$

where q_j denotes the value of $x[kG+j]$, and $L(x[kG+j] = q_j)$ is calculated by using (B.4) in Appendix B. Note that for each EM iteration, we need to choose a hard decision $\mathbf{x}_k \in \{1, -1\}^G$ to maximize the ECLL function of (27). The iterative EM procedures can help converge to the local optimal values of \mathbf{x}_k and \mathbf{w} . Finally, in order to pass soft information to the SOVA decoder, the extrinsic *a posteriori* LLR $\lambda_k^{D,ext} = [\lambda^{D,ext}[kG], \dots, \lambda^{D,ext}[(k+1)G-1]]^T$ is generated at the final EM iteration as the output of the EM

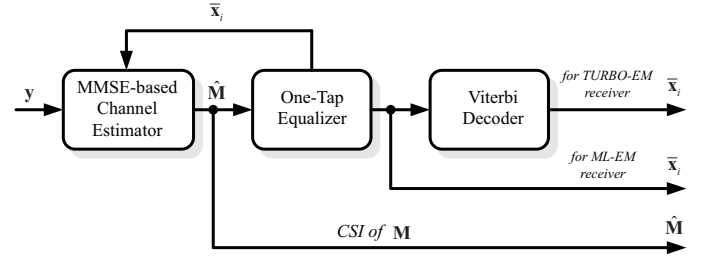


Fig. 5. Initialization procedure for ML-EM and TURBO-EM receivers.

detector. From [27] and (B.3), we get

$$\begin{aligned} \lambda^{D,ext}[kG+j] \\ = \ln \frac{P(x[kG+j]=+1|\mathbf{y}_M)}{P(x[kG+j]=-1|\mathbf{y}_M)} - \lambda^{C,ext}[kG+j] \\ \approx \max_{\mathbf{x}_k \in \Omega_j^+} \left\{ \frac{-1}{\sigma_z^2} \left\| \mathbf{y}_M - \tilde{\mathbf{H}}_k \mathbf{x}_k \right\|^2 + \frac{1}{2} \mathbf{x}_k^T \lambda_{k \setminus \{j\}}^{C,ext} \right\} \\ - \max_{\mathbf{x}_k \in \Omega_j^-} \left\{ \frac{-1}{\sigma_z^2} \left\| \mathbf{y}_M - \tilde{\mathbf{H}}_k \mathbf{x}_k \right\|^2 + \frac{1}{2} \mathbf{x}_k^T \lambda_{k \setminus \{j\}}^{C,ext} \right\} \end{aligned} \quad (29)$$

where Ω_j^+ denotes the set for which the j th entry of \mathbf{x}_k is "+1"; Ω_j^- is defined similarly, and the vectors $\mathbf{x}_{k \setminus \{j\}}$ and $\lambda_{k \setminus \{j\}}^{C,ext}$ are obtained by omitting the j th entry of \mathbf{x}_k and $\lambda_k^{C,ext}$, respectively. The extrinsic LLR $\lambda_k^{D,ext}$ is then converted into soft bits using (26), modeled as the output of an AWGN channel with unit mean and variance σ_C^2 , deinterleaved through Π^{-1} , and passed to the SOVA decoder. The variance σ_C^2 is estimated as

$$\sigma_C^2 = \frac{1}{N_I} \sum_{i=1}^{N_I} (|\mu[i] - 1|)^2 \quad (30)$$

where $\mu[i]$ indicates the soft value of the coded bits ranging between -1 and $+1$, and N_I represents the interleaver size. It is mentioned in [28] that the Gaussian assumption is not satisfied at the beginning of TURBO iterations, but it becomes a good approximation as the number of iterations increases. After the SOVA decoder produces the soft information by considering the ML path and its strongest competitor in the trellis diagram, the receiver progresses toward the next TURBO iteration until a preset maximum number of iteration, N_{TB} , is reached.

C. Initial Setting and Channel Estimation Update

Fig. 5 depicts the block diagram for initialization of the two receivers. The initial CE is performed through the use of pilot tones and improved via the decided data symbols. Let \mathbf{X}_P be a diagonal matrix whose diagonal elements are obtained from the stacked vector of J pilot tones on subcarriers $\{P_0, \dots, P_{J-1}\}$ within the OFDM symbol. Applying the MMSE-based CE method, we obtain [29]

$$\hat{\mathbf{M}} = \mathbf{F} \Psi_P^{-1} \mathbf{F}_P^\dagger \mathbf{X}_P^\dagger \mathbf{y}_P \quad (31)$$

where $\Psi_P = \mathbf{F}_P^\dagger \mathbf{X}_P^\dagger \mathbf{X}_P \mathbf{F}_P + (\sigma_z^2 + \sigma_{ICI}^2) \mathbf{I}_L$, \mathbf{y}_P and \mathbf{F}_P are defined similar to \mathbf{y} and \mathbf{F} , respectively, but here related to subcarriers $\{P_0, \dots, P_{J-1}\}$ only. By invoking central limit theorem, the ICI energy σ_{ICI}^2 is approximated as $\frac{1}{12} (2\pi f_D)^2$

TABLE I
COMPUTATION COMPLEXITY. (EXAMPLE: $G = 4$, $Q = 4$, $\gamma = 1$, $L = 6$,
 $N = 256$ AND $N_G = 64$)

ML-EM receiver and TURBO-EM receiver			
	Number of Complex Multiplications	Example	Unit
ICI canceller	$(2Q)G^2(2Q+1)$	1152	/ML (or TURBO) iteration /group
EM detector	$5(2Q+3)^3G^3 + G^2[36Q^2 + (108 + 2^{\gamma G+1})Q + (81 + 2^{\gamma G+1})] + G[(2^{\gamma G+2} + 4)Q + 3 \times 2^{\gamma G} + 6]$	447208	/ML (or TURBO) iteration /EM iteration /group
Precalculation of $\mathbf{C}_{\mathbf{w}\mathbf{w}}$ & $\Phi^\dagger\Phi$	$(2Q+3)^2(L+1)G^2 + (2Q+3)LG$	13816	/group
Eq. (29)	$2^{\gamma(G-1)}[(2Q+1)G^2 + (2Q+2)G+1]$	1480	/TURBO iteration /group
Method I and Method II in [19]			
Method I	$L^3 + 2L^2N_G + 2LN_G + 2N_G^2 + N^3 + 2N^2 + N$	16922328	/iteration
Method II	$N^3 + 2N^2 + N$	16908544	/iteration

[4]. Subsequently, an one-tap equalizer is used for data detection, and a decision-directed approach is carried out to initialize the two receivers. For the ML-EM receiver, decided data symbols together with the pilot tones are used to generate a new channel estimate $\hat{\mathbf{M}}$ by using (31) and then produce an updated decision symbol $\hat{\mathbf{x}}_i$, while for the TURBO-EM receiver, much more reliable decision symbols are generated by the Viterbi decoder. At the first TURBO iteration of the TURBO-EM receiver, $\lambda_k^{C,ext}$ is set to $\mathbf{0}$, and $\hat{\mathbf{x}}_i$ is used to replace $\tilde{\mathbf{x}}_i$ in (26). Moreover, we initialize $\hat{\mathbf{H}}$ as $\hat{\mathbf{M}}$, i.e., set $\hat{\mathbf{W}} = \mathbf{O}$, since no information on $\hat{\mathbf{W}}$ is available at the first iteration of the two receivers.

Due to ICI, the initial estimate of \mathbf{M} becomes inaccurate as f_D increases. Hence, Fig. 3(a) and Fig. 4 offer an option for CE update, in which both pilot tones and decision data symbols are used together. Particularly, for the TURBO-EM receiver, decision data symbols are fed back from the SOVA decoder. At the second and subsequent outer loop iterations, the ICI (reconstructed from N_U adjacent subcarriers in a hard or soft manner) to the subcarriers is canceled out in the received signal \mathbf{y} , and the MMSE-based CE method is again used to refine the estimate $\hat{\mathbf{M}}$ by setting $\sigma_{ICI}^2 = 0$.

D. Computation Complexity

Now let us look at the number of complex multiplications required for the two proposed receivers. Assume that the operation of $K \times K$ matrix inversion needs K^3 complex multiplications. In Table I, the first and second rows indicate the complexity of the ICI canceller and the EM detector,

respectively. The third row gives the complexity for pre-computing $\mathbf{C}_{\mathbf{w}\mathbf{w}}$ and $\Phi^\dagger\Phi$ in the EM detector. Note that the calculation of (28) in the MAP EM detector does not require any multiplications, and the number of multiplications required to calculate (29) is presented in the fourth row of the Table I. For example, in the case of $G = 4$, $Q = 4$, $\gamma = 1$ and $L = 6$, the complexity is calculated in the third column as well. Some complexity reduction can be achieved by applying the SAGE algorithm and the Viterbi algorithm, as proposed in [30] and [26] respectively, to simplify (25) and (29) when the values of G and γ are relatively large to dominate the overall computation complexity. Moreover, the complexity of the SOVA decoder is, in general, upper-bounded by two times that of the Viterbi decoder. Finally, the computation complexity of the MMSE-based CE method can be referred from [29] for details.

In addition, Table I lists the complexity measure of the two piecewise linear methods in [20] whose BER performances are demonstrated later for comparison with the ML-EM receiver. For the two methods, it is also required to calculate averaged channel impulse response, and its complexity is not included in the complexity evaluation. The number of complex multiplications for the two methods and the ML-EM receiver is mainly dominated by the values of $N^3 + 2N^2$ and $(2Q+3)^3G^3$, respectively. Since there are $R = \frac{N}{G}$ groups in the ML-EM receiver and the EM iterations number is in general small, the computation complexity of the proposed ML-EM receiver is approximately a factor $\frac{(2Q+3)^3G^2}{N^2+2N}$ of the computation complexity of the two methods.

V. COMPUTER SIMULATION

A. Simulation Parameters

Our simulation demonstrates the performance of the two EM-based receivers. The simulation parameters are defined according to the 802.16e OFDM standard [3]. The entire bandwidth, $5MHz$, is divided into $N = 256$ subcarriers among which 192 subcarriers carry data symbols, $J = 8$ subcarriers transmit pilot tones, and the remaining 56 subcarriers are virtual subcarriers. The BPSK modulation scheme is adopted for the pilot tones, and a pilot subcarrier transmits at the same power level as a data subcarrier. Each OFDM data frame is composed of $N_F = 40$ OFDM data symbols, and the length of GI is set to $N_G = 64$. For the BICM scheme, we employ a rate-1/2 convolutional code with generator polynomial (133,171) represented in octal and a block interleaver with 96 rows and 80 columns. Both a conventional two-path channel and an International Telecommunication Union (ITU) Veh.A channel are used in our simulation with path delays uniformly distributed from 0 to 50 sample periods, where the relative path power profiles are set as 0,0(dB) for the two-path channel and 0, -1, -9, -10, -15, -20(dB) for the ITU Veh.A channel [31]. The fading channel is generated with Jakes model [32]. The user-defined parameters are chosen as $N_U = 10$ and $N_{EM} = 5$. Some statistical information such as power delay profiles (PDP), Doppler frequency, and noise power is assumed to be known to the receivers. Throughout the simulation, the parameter Eb/No is defined as a ratio of

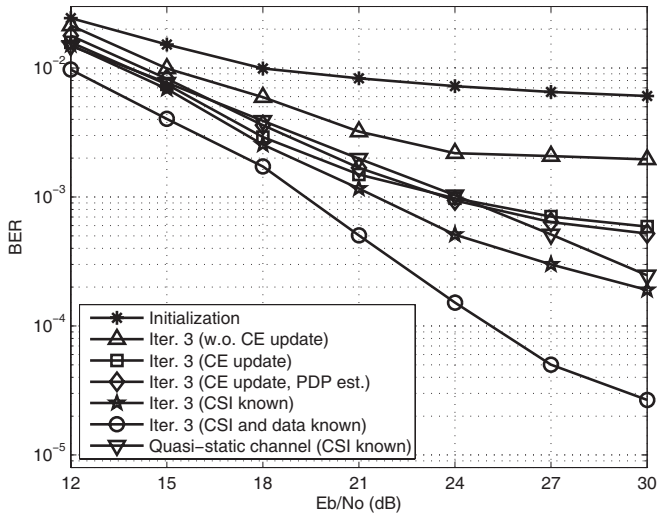


Fig. 6. BER performance of the ML-EM receiver in the ITU Veh.A channel at $f_D = 0.1$. ($N_{ML} = 3$ and $[G, Q] = [4, 4]$)

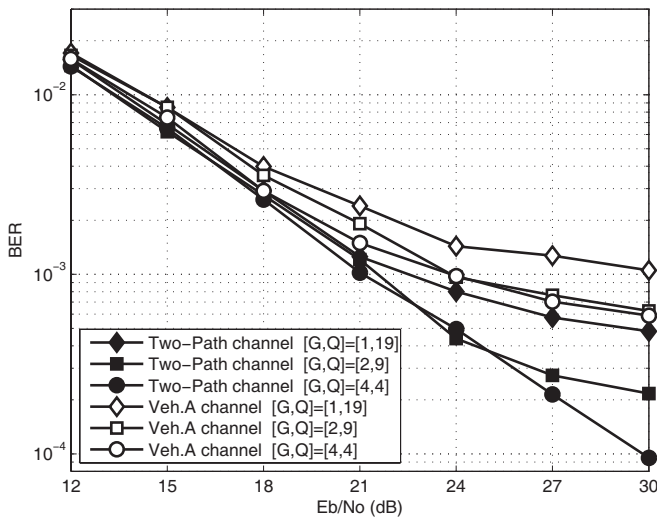


Fig. 7. BER performance of the ML-EM receiver with CE update for various $[G, Q]$. ($N_{ML} = 3$)

averaged receive bit energy to the power spectral density of noise.

As a benchmark, the performance curve with ideal initialization, labeled as CSI and data known, serves as a performance lower bound, and the results obtained with ideal CSI, denoted as CSI known, is provided for reference. We also include the performance curve of the one-tap equalizer in quasi-static channels, under the assumption of ideal CSI. In addition, the two piecewise linear methods, referred to as Method I and Method II in [20], are examined to compare with the proposed ML-EM receiver. For the two methods, the initial CSIs are found directly through IDFT-based channel estimation using pilot tones and the subsequent estimates are calculated based on the decision-feedback methodology. For fair comparison, we also simulate the BER performance of the two methods when the IDFT-based channel estimation is replaced with the EM algorithm in [23] and the MMSE-based channel estimation in [29], respectively.

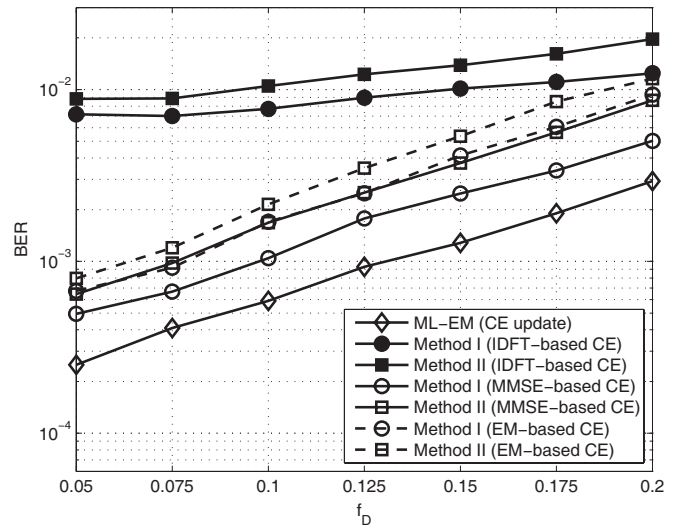


Fig. 8. BER versus f_D for the ML-EM receiver and the two piecewise linear methods in the ITU Veh.A channel. ($E_b/N_0 = 30dB$, $[G, Q] = [4, 4]$ and $N_{ML} = 3$)

B. Simulation Results

Fig. 6 shows the BER performance of the ML-EM receiver in the ITU Veh.A channel at $f_D = 0.1$. The parameter of $[G, Q]$ is set to $[4, 4]$. It is observed from Fig. 6 that after three iterations, the ML-EM receivers with or without CE update achieve much better performance than the same receiver at the initialization stage. For the ML-EM receiver with CE update, the required E_b/N_0 at $BER = 10^{-3}$ is almost the same as that for an one-tap equalizer in quasi-static channels. However, its BER performance is about $3dB$ worse than that based on ideal CSI knowledge, and for the case without CE update, an error floor occurs at $BER = 2 \times 10^{-3}$ in the high E_b/N_0 region. When compared with the lower bound, there is still an E_b/N_0 gap of $4.5dB$ at $BER = 10^{-3}$ for the ML-EM receiver with CE update. Clearly, this gap is due to the error propagation effect. It is worth noting that the performance lower bound in Fig. 6 comes very close to the theoretical matched-filter bound found in [5]. Since time-variant channels introduce diversity gains, this performance lower bound is superior to the performance of the one-tap equalizer in quasi-static channels. In realistic scenario, the PDP of wireless channels is not known a priori at the receiver. In order to check the usefulness of our proposed schemes, we demonstrate the performance of the ML-EM receiver by assuming that the CSI at the preamble of each OFDM data frame is known and the power gain of the corresponding channel impulse response is used as the PDP. Without loss of generality, the estimated CSI could be very accurate because the number of pilot tones in the preamble is usually large enough. Our simulation result shows that there is no significant performance loss when the PDP is estimated in this way. Fig. 7 addresses the impact of group size on the BER performance, and the number of Q is selected to keep $(2Q + 1)G \approx 39$ for fair comparison. As expected, joint detection with more subcarriers will ease the error propagation effect and attain better performance, and the improvement eventually saturates as the size of G increases. To further improve the performance, channel decoding is

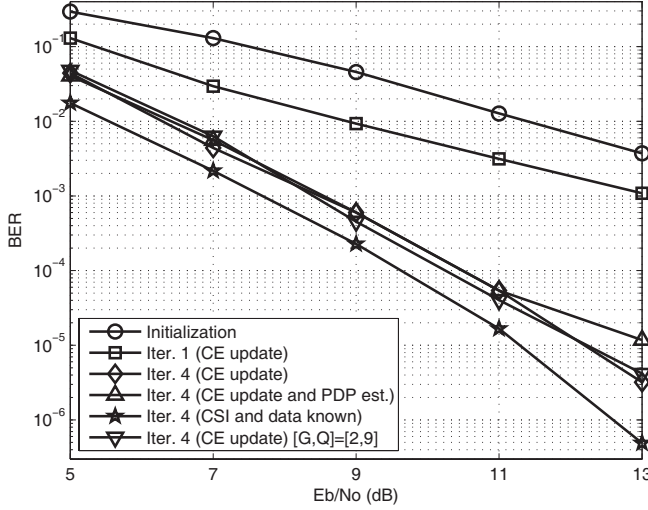


Fig. 9. BER performance of the TURBO-EM receiver in the ITU Veh.A channel at $f_D = 0.1$. ($N_{TB} = 4$ and $[G, Q] = [4, 4]$)

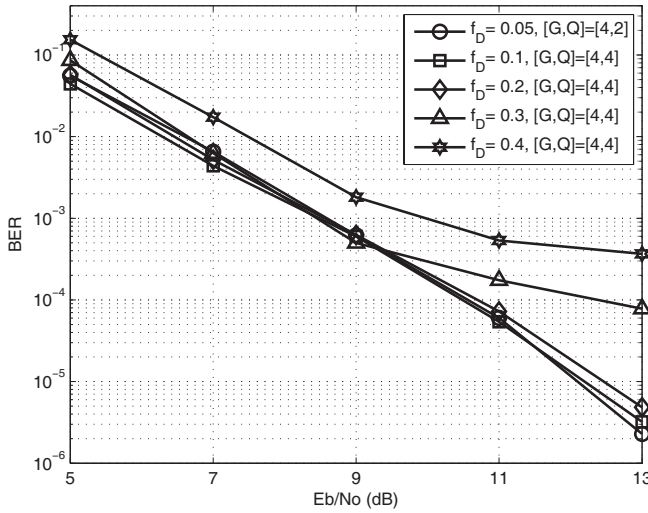


Fig. 10. BER performance of the TURBO-EM receiver in the ITU Veh.A channel for various f_D . ($N_{TB} = 4$)

needed to effectively handle the error propagation such as the method adopted in the TURBO-EM receiver. In Fig. 8, we examine the impact of f_D on the BER performance of the ML-EM receiver at $E_b/N_0 = 30dB$ in ITU-Veh.A channel, and its BER performance is compared with that of the two methods in [20]. Four iterations are carried out in these two methods, and additional iterations yield no performance improvement. Although the ML-EM receiver performs worse when Doppler effect becomes more severe, it can still achieve a BER of 3×10^{-3} at $f_D = 0.2$. Both the Method I and Method II are inferior to the ML-EM receiver no matter what kinds of channel estimation schemes are used. Obviously, due to the requirement of using some guard tones, the two methods with the IDFT-based channel estimation, originally proposed in [20], incur substantial performance loss. Even with the EM algorithm in [23] as the channel estimation scheme, the two methods perform worse than the ML-EM receiver since ICI estimation and data detection are not jointly designed in these two methods.

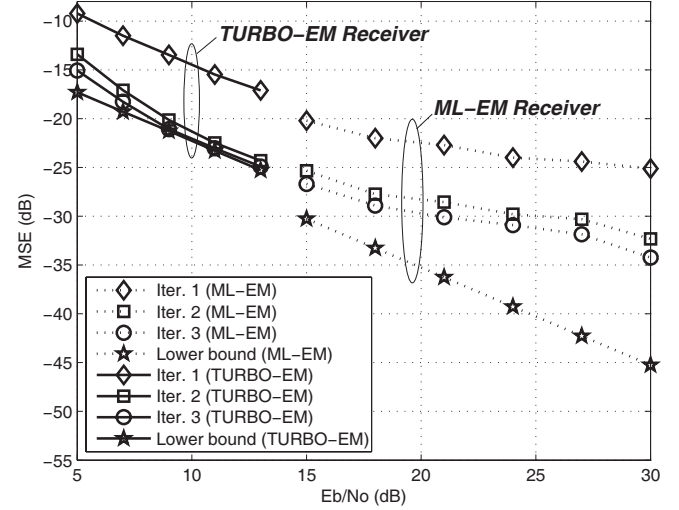


Fig. 11. MSE performance of channel estimation in the ITU Veh.A channel. ($f_D = 0.1$)

Fig. 9 shows the BER performance of the TURBO-EM receiver in the ITU Veh.A channel at $f_D = 0.1$. With the BICM, the effect of error propagation is effectively suppressed. We see from Fig. 9 that after four iterations, the receiver with CE update achieves a performance gap with respect to the lower bound by less than $1dB$ at $BER = 10^{-5}$, and the receiver exhibits a remarkable improvement as compared with the initialization stage. However, when CSI is not updated (which is not depicted here), the performance of the receiver deteriorates remarkably, although it is still better than that of the initialization stage. Hence, in order to achieve a good performance, it is necessary to refine the CSI, especially when the number of pilot tones is small and the normalized MDF is large. From Fig. 9, we also observe that the receivers with $G = 2$ and $G = 4$ have nearly identical BER performance. Besides, even when the PDP is unknown and estimated from the preamble, the TURBO-EM receiver still performs very well and there is only slight degradation in BER performance at $E_b/N_0 = 13dB$. Fig. 10 depicts the BER performance of the TURBO-EM receiver in the ITU Veh.A channel for various f_D . We note that the TURBO-EM receiver shows a robust BER performance against fading up to $f_D = 0.2$, and the performance gap, compared with the case of $f_D = 0.05$, is only about $0.3dB$ at $BER = 10^{-5}$. The receiver, however, displays an irreducible BER error floor when f_D is beyond 0.3, and a more complicated ICI model, e.g. the model suggested in [10], could be used to get better performance. Fig. 11 shows the normalized mean square error (MSE) performance of CE in the ITU Veh.A channel at $f_D = 0.1$ for the E_b/N_0 operating ranges of the two receivers. We observe that for both receivers, the MSE decreases as the number of iterations increase, and performance gain is due to the reduction of ICI in the CE update. In fact, the performance improvement in MSE is about $7dB$ to $10dB$ for the E_b/N_0 operating ranges and it eventually saturates after three iterations. For calibration purpose, the lower bound on MSE performance under the assumption of no decision error in quasi-static channels is also simulated and shown in this figure. We observe that the lower bound is

attainable for the TURBO-EM receiver, while for the ML-EM receiver, there is a significant performance gap especially at high E_b/N_0 as compared with the low bound.

VI. CONCLUSIONS

We have investigated two EM-based iterative receivers for OFDM and BICM/OFDM systems in doubly selective channels. Based on the proposed EM algorithm for data detection, both receivers use groupwise processing with ICI cancellation to reduce computation complexity and to explore time diversity inherent in time-variant channels. For OFDM systems, the ML-EM receiver significantly outperforms the conventional one-tap equalizer and the two piecewise linear methods in [20], and the BER performance at $f_D = 0.05$ even approaches the BER performance without Doppler effect. Compared with the matched-filter bound, an E_b/N_0 gap appears because of the error propagation effect. For BICM/OFDM systems, a TURBO-EM receiver, which iterates between the MAP EM detector and the SOVA decoder, is then introduced. This receiver effectively solves the error propagation problem, and it attains a performance close to the lower bound in terms of BER. Simulation results indicate that in order to attain a good performance, the CE update is required when we use low-density pilot tones at high Doppler frequencies. As a final remark, a group size of two to four is large enough to guarantee an acceptable performance under practical channel conditions.

APPENDIX A. CALCULATION OF (18) AND (19)

Let us first define an AR channel model

$$h[l, N-1] = \alpha h[l, 0] + u \quad (\text{A.1})$$

where α is the parameter of the AR model and u represents a complex Gaussian random variable, independent of $h[l, n]$, with zero mean and variance σ_u^2 . From Jakes model, α is evaluated by

$$\alpha = E[h[l, N-1]h^*[l, 0]] = J_0\left(\frac{2\pi f_D(N-1)}{N}\right) \quad (\text{A.2})$$

where $J_0(\cdot)$ is the zeroth order Bessel function of the first kind, f_D denotes the normalized MDF. Recall from (2) that we have $E[h[l, n]] = 0$ and $E[h^2[l, n]] = \Xi_l$. Following the energy conservation rule in (A.1), the variance of u can be calculated as $\sigma_u^2 = (1 - \alpha^2)\Xi_l$. According to (A.1), (A.2) and (3), we find the slope of the l th tap over the duration of one OFDM symbol as

$$\begin{aligned} a[l, 1] &= \frac{1}{N-1} (h[l, N-1] - h[l, 0]) \\ &= \frac{1}{N-1} ((\alpha - 1)h[l, 0] + u) \end{aligned} \quad (\text{A.3})$$

and its mean and variance is calculated by

$$E[a[l, 1]] = 0 \quad (\text{A.4})$$

$$\begin{aligned} E[|a[l, 1]|^2] &= \frac{(\alpha - 1)^2\Xi_l + (1 - \alpha^2)\Xi_l}{(N-1)^2} \\ &= \frac{2(1 - \alpha)\Xi_l}{(N-1)^2} \end{aligned} \quad (\text{A.5})$$

Besides, it is reasonable to assume that the slopes of channel taps are independent of each other, i.e., $E[a[l, 1]a^*[l', 1]] = 0$ if $l \neq l'$ since $h[l, n]$'s for different l 's are independent. Consequently, we get ($\mathbf{w} = \mathbf{F}\mathbf{s}$):

$$\boldsymbol{\mu}_{\mathbf{w}} = \mathbf{0} \quad (\text{A.6})$$

$$\mathbf{C}_{\mathbf{w}\mathbf{w}} = \mathbf{F}\mathbf{C}_{\mathbf{ss}}\mathbf{F}^\dagger \quad (\text{A.7})$$

where $\mathbf{C}_{\mathbf{ss}} = E[\mathbf{ss}^\dagger]$ is a diagonal matrix with the l th diagonal entry given by $\frac{2(1-\alpha)\Xi_l}{(N-1)^2}$.

APPENDIX B. CALCULATION OF (28)

Using the definition of LLR, we have

$$\begin{aligned} L(x[kG+j] = +1) \\ = \lambda^{C, ext}[kG+j] - \ln(e^0 + e^{\lambda^{C, ext}[kG+j]}) \end{aligned} \quad (\text{B.1})$$

$$L(x[kG+j] = -1) = -\ln(e^0 + e^{\lambda^{C, ext}[kG+j]}) \quad (\text{B.2})$$

The calculation of (B.1) and (B.2) can be simplified by using the rule:

$$\ln \sum_j e^{a_j} \approx \max_j a_j \quad (\text{B.3})$$

When applied, straightforward manipulation yields

$$\begin{aligned} L(x[kG+j] = q) \\ = \begin{cases} -\ln 2, & \text{if } \lambda^{C, ext}[kG+j] = 0 \\ \min(q\lambda^{C, ext}[kG+j], 0), & \text{otherwise} \end{cases} \end{aligned} \quad (\text{B.4})$$

where q is either +1 or -1.

REFERENCES

- [1] R. van Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*. Artech House, 2000.
- [2] E. Aky and E. Ayanoglu, "Achieving full frequency and space diversity in wireless systems via BICM, OFDM, STBC and Viterbi decoding," *IEEE Trans. Commun.*, vol. 54, pp. 2164-2172, Dec. 2006.
- [3] IEEE Std. 802.16e-2005 and IEEE 802.16-2004/Cor1-2005, "Part 16: Air Interface for Fixed and Mobile Broadband Wireless Access Systems," *IEEE-SA Standards Board*, Tech. Rep., 2006.
- [4] Y. Li and L. J. Cimini, Jr., "Bounds on the interchannel interference of OFDM in time-varying impairments," *IEEE Trans. Commun.*, vol. 49, pp. 401-404, Mar. 2001.
- [5] X. Cai and G. Giannakis, "Bounding performance and suppressing inter-carrier interference in wireless mobile OFDM," *IEEE Trans. Commun.*, vol. 51, pp. 2047-2056, Dec. 2003.
- [6] M.-X. Chang, "A novel algorithm of inter-subchannel interference self-cancellation for OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 2881-2893, Aug. 2007.
- [7] A. Seyedi and G. J. Saulnier, "General ICI self-cancellation scheme for OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 54, pp. 198-210, Jan. 2005.
- [8] H.-C. Wu, X. Huang, Y. Wu, and X. Wang, "Theoretical studies and efficient algorithm of semi-blind ICI equalization for OFDM," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 3791-3798, Oct. 2008.
- [9] P. Schniter, "Low-complexity equalization of OFDM in doubly selective channels," *IEEE Trans. Signal Process.*, vol. 52, pp. 1002-1011, Apr. 2004.
- [10] T. Wang, J. G. Proakis, and J. R. Zeidler, "Techniques for suppression of intercarrier interference in OFDM systems," in *Proc. IEEE Wireless Commun. and Networking Conf.*, Mar. 2005, pp. 39-44.
- [11] G. Li, H. Yang, L. Cai, and L. Gui, "A low-complexity equalization technique for OFDM system in time-variant multipath channels," in *Proc. IEEE Vehicular Technology Conf.*, Oct. 2003, pp. 2466-2470.

[12] J. Gao and H. Liu, "Low-complexity MAP channel estimation for mobile MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 774-780, Mar. 2008.

[13] Y. Choi, P. Voltz, and F. Cassara, "On channel estimation and detection for multicarrier signals in fast and selective Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 49, pp. 1375-1387, Aug. 2001.

[14] W. G. Jeon, K. H. Chang, and Y. S. Cho, "An equalization technique for orthogonal frequency-division multiplexing systems in time-variant multipath channels," *IEEE Trans. Commun.*, vol. 47, pp. 27-32, Jan. 1999.

[15] K. Kim and H. Park, "A low complexity ICI cancellation method for high mobility OFDM systems," in *Proc. IEEE Vehicular Technology Conf.*, May 2006, pp. 2528-2532.

[16] S. Kim and G. Pottie, "Robust OFDM in fast fading channel," in *Proc. IEEE GLOBECOM*, Dec. 2003, pp. 1074-1078.

[17] A. Gorokhov and J. P. Linnartz, "Robust OFDM receivers for dispersive time-varying channels: equalization and channel acquisition," *IEEE Trans. Commun.*, vol. 52, pp. 572-583, Apr. 2004.

[18] S. Tomasin, A. Gorokhov, H. Yang, and J. P. Linnartz, "Iterative interference cancellation and channel estimation for mobile OFDM," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 238-245, Jan. 2005.

[19] K. Chang, Y. Han, J. Ha, and Y. Kim, "Cancellation of ICI by Doppler effect in OFDM systems," in *Proc. IEEE Vehicular Technology Conf.*, May 2006, pp. 1411-1415.

[20] Y. Mostofi and D. C. Cox, "ICI mitigation for pilot-aided OFDM mobile systems," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 765-774, Mar. 2005.

[21] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*. Springer-Verlag, 1999.

[22] B. Lu, X. Wang, and Y. Li, "Iterative receivers for space-time block-coded OFDM systems in dispersive fading channels," *IEEE Trans. Wireless Commun.*, vol. 1, pp. 213-225, Apr. 2002.

[23] T. Y. Al-Naffouri, "An EM-based forward-backward Kalman filter for the estimation of time-variant channels in OFDM," *IEEE Trans. Signal Process.*, vol. 55, pp. 3924-3930, July 2007.

[24] K. Muraoka, K. Fukawa, H. Suzuki, and S. Suyama, "Channel estimation using differential model of fading fluctuation for EM algorithm applied to OFDM MAP detection," in *Proc. Personal, Indoor and Mobile Radio Commun.*, Sep. 2007, pp. 1-5.

[25] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, 1993.

[26] F. Peng and W. E. Ryan, "A low-complexity soft demapper for OFDM fading channels with ICI," in *Proc. IEEE Wireless Commun. and Networking Conf.*, Apr. 2006, pp. 1549-1554.

[27] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channel," *IEEE Trans. Commun.*, vol. 51, pp. 389-399, Mar. 2003.

[28] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: turbo-codes," *IEEE Trans. Commun.*, vol. 44, pp. 1261-1271, Oct. 1996.

[29] O. Edfors, M. Sandell, J. J. van de Beek, S. K. Wilson, and P. O. Borjesson, "Analysis of DFT-based channel estimators for OFDM," *Wireless Personal Commun.*, vol. 12, pp. 55-70, Jan. 2000.

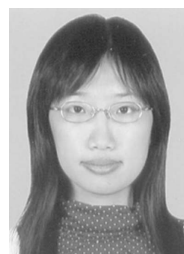
[30] J. A. Fessler and A. O. Hero, "Space-alternating generalized expectation-maximization algorithm," *IEEE Trans. Signal Process.*, vol. 42, pp. 2664-2677, Oct. 1994.

[31] J. Laiho, A. Wacker, and T. Novosad, *Radio Network Planning and Optimisation for UMTS*. Wiley, 2002.

[32] W. C. Jakes, *Microwave Mobile Communications*. Wiley, 1974.



Meng-Lin Ku was born in Taoyuan, Taiwan. He received the B.S., M.S. and Ph.D. degrees in Communication Engineering from National Chiao Tung University, Hsinchu, Taiwan, in 2002, 2003 and 2009 respectively. Between 2009 and 2010, he held post-doctoral positions at National Chiao Tung University and Harvard University. In August 2010, he joined the Department of Communication Engineering, National Central University, as an Assistant Professor. His current research interests are in the area of next-generation mobile and wireless communications, cognitive radios, and optimization for radio access.



Wen-Chuan Chen was born in Penghu, Taiwan, R.O.C. He received the B.S. and M.S. degree in the Department of Communication Engineering from Yuan Ze University in 2006 and National Chiao Tung University in 2008, respectively. She is currently working in the Institute for Information Industry, Taipei, Taiwan.



Chia-Chi Huang was born in Taiwan, R.O.C. He received the B.S. degree in electrical engineering from National Taiwan University in 1977 and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1980 and 1984, respectively.

From 1984 to 1988, he was an RF and Communication System Engineer with the Corporate Research and Development Center, General Electric Co., Schenectady, NY, where he worked on mobile radio communication system design. From 1989 to 1992, he was with the IBM T.J. Watson Research Center, Yorktown Heights, NY, as a Research Staff Member, working on indoor radio communication system design. Since 1992, he has been with National Chiao Tung University, Hsinchu, Taiwan, and currently as a Professor in the Department of Electrical Engineering.