

# Joint Decision-Directed Channel and Noise-Variance Estimation for MIMO OFDM/SDMA Systems Based on Expectation-Conditional Maximization

Jiankang Zhang, *Student Member, IEEE*, Lajos Hanzo, *Fellow, IEEE*, and Xiaomin Mu

**Abstract**—A joint channel impulse response (CIR) and noise-variance estimation scheme is proposed for multiuser multiple-input–multiple-output (MIMO) orthogonal frequency-division multiplexing/space-division multiple access (OFDM/SDMA) systems, which is based on the expectation-conditional maximization (ECM) algorithm. Multiple users communicating over fading channels exhibiting a range of different characteristics are considered in this paper. Channel estimation becomes quite challenging in this scenario since an increased number of independent transmitter–receiver links having different statistical characteristics have to be simultaneously estimated for each subcarrier. To cope with this scenario, we design an ECM-based joint CIR and noise-variance estimator for multiuser MIMO OFDM/SDMA systems, which is capable of simultaneously estimating diverse CIRs and noise variance. Furthermore, we propose a forward error code (FEC)-aided decision-directed channel estimation scheme based on the ECM algorithm, which further improves the ECM algorithm by exploiting the error correction capability of an FEC decoder for iteratively exchanging information between the decoder and the ECM algorithm.

**Index Terms**—Channel estimation, expectation-conditional maximization (ECM), multiple-input–multiple-output (MIMO), orthogonal frequency-division multiplexing (OFDM), space-division multiple access (SDMA).

## I. INTRODUCTION

ORTHOGONAL frequency-division multiplexing (OFDM) [1] constitutes a promising technique of combating the detrimental effects of multipath-induced delay spread in high-data-rate transmission. In recent years, various smart antenna designs have attracted substantial research interests because they are capable of mitigating the deleterious effects of multipath fading on the desired signal and of suppressing the inter-

fering signals, thereby increasing the achievable performance of wireless systems [2]. Specifically, smart antenna-assisted space-division multiple access (SDMA) is capable of achieving high spectral efficiency by supporting multiplicity of users within the same frequency band and facilitating the separation of their signals based on their unique user-specific channels.

As a beneficial combination, OFDM/SDMA systems have attracted substantial interests [1]–[4]. Typical OFDM/SDMA systems employ an array of antennas at the base station (BS), which detects the received signal of multiple single-antenna-aided user terminals. As a result, a substantially improved system capacity is achieved despite employing low-complexity user terminals [2], [5]. However, the performance of these systems is critically dependent on the precision of the channel knowledge. Furthermore, while exploiting the joint benefits of OFDM and SDMA, their combination faces new challenges because a significantly increased number of independent transmitter–receiver channel links have to be estimated simultaneously for each subcarrier, whereas the interfering signals of the other transmitters have to be suppressed [3].

Over the past decade, intensive research efforts have been devoted to developing effective approaches for both channel estimation and symbol detection in transmitter- and/or receiver-diversity-aided systems. For example, [6]–[8] proposed optimal training sequence design and optimal pilot tone allocation using the mean-square error (MSE) metric for channel estimation in multiple-input–multiple-output (MIMO) OFDM systems. Barhumy *et al.* [7] demonstrated that the optimal pilot sequences should have equi-powered, equi-spaced, and phase shift orthogonal pilots, whereas Zhang *et al.* [8] discussed the corresponding necessary conditions for correlated fading channels. Li *et al.* [9] developed a channel estimator by exploiting the time- and frequency-domain correlations of the channel impulse response (CIR) and Frequency-Domain Channel Transfer Function (FD-CHTF), respectively, which was further simplified and enhanced in [6], [10], and [11]. As the affordable hardware capacity is increasing, it becomes more feasible to implement iterative receivers allowing for substantial improvements of the physical layer functions. The iterative expectation–maximization (EM) algorithm [12] and the various derivatives of this algorithm have been shown to strike an attractive tradeoff between the performance attained and the complexity imposed.

A classic EM-based channel estimation algorithm and the so-called space-alternating generalized EM-based channel

Manuscript received May 5, 2010; revised February 2, 2011; accepted April 15, 2011. Date of publication May 5, 2011; date of current version June 20, 2011. This work was supported in part by the China Scholarship Council and the European Union under the auspices of the OPTIMIX project and in part by the Engineering and Physical Sciences Research Council, U.K. The review of this paper was coordinated by Prof. E. Bonek.

J. Zhang was with the School of Electronics and Computer Science, University of Southampton, SO17 1BJ Southampton, U.K. He is now with the School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China (e-mail: jz09v@ecs.soton.ac.uk).

L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, SO17 1BJ Southampton, U.K. (e-mail: lh@ecs.soton.ac.uk).

X. Mu is with the School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China (e-mail: iexmmu@zzu.edu.cn).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TVT.2011.2148184

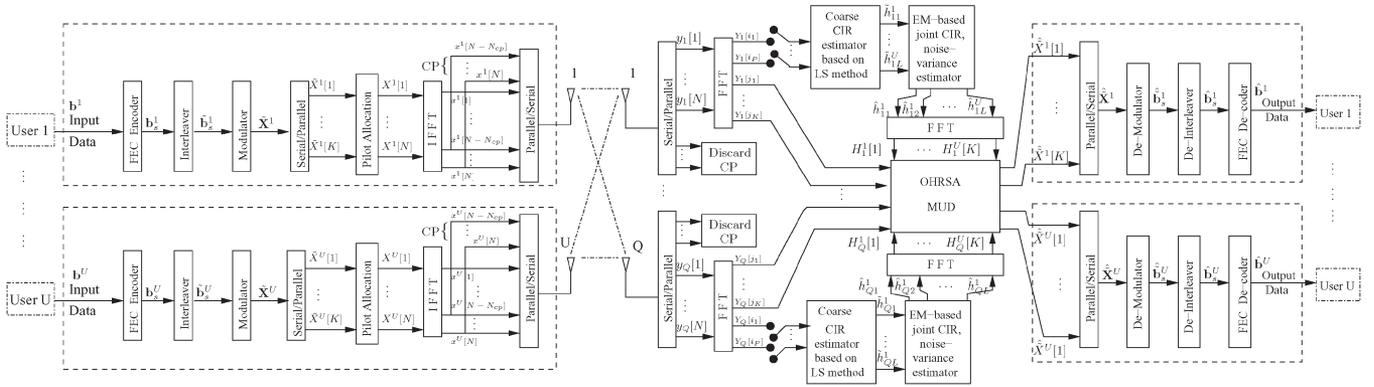


Fig. 1. Block diagram of a multiuser MIMO SDMA/OFDM UL system.

estimation algorithm were designed in [13] for OFDM systems invoking transmitter diversity, and their convergence rates were compared. In [14], the so-called unbiased EM and the unbiased expectation-conditional maximization channel estimator were designed by exploiting the similarity between the families of EM-type CIR estimators and least-square (LS) estimators. The authors of [15] derived an EM algorithm using low-rank approximation to avoid inverting large matrices, which substantially reduced the receiver’s complexity. Choi [16] developed a robust EM-based channel estimation method that was applicable to diverse MIMO systems and operating without requiring the probability density function (pdf) of the channel parameters.

The channel estimation techniques found in the open literature were typically developed under the assumption that all the channels are statistically similar to each other: either time invariant or time variant and either Rayleigh or Rician fading. First, instead of considering a whole range of vehicular velocities, i.e., Doppler frequencies, we considered the two extreme scenarios of static and uncorrelated fading, which model stationary and high-speed scenarios. However, in practical multiuser systems, the channel of each mobile station (MS) is different and independent of that of the others. More specifically, the following scenario may be encountered in multiuser systems: 1) Some of them are stationary MSs, whereas the others may be roaming. 2) The Doppler shift is also different for the different MSs since the velocity is different for the different MSs. 3) Line-of-sight communication links may exist between some of the MSs and the BS, whereas this may not be the case for other MSs. 4) Even if all the MSs are stationary, the surroundings of the MSs may vary, which will also result in different channels. Hence, we solve the open problem of jointly estimating the channels of both stationary and roaming OFDM/SDMA users having time-invariant and time-variant channels with different fading characteristics, respectively. By exploiting the statistical characteristics of the different channel types, we design an expectation-conditional maximization (ECM)-based joint CIR and noise-variance estimator for multiuser MIMO OFDM/SDMA systems, which is capable of simultaneously estimating the diverse CIRs and the noise variance. Furthermore, we design a forward error code (FEC)-aided decision-directed (FEC-A-DD) estimation technique based on the ECM algorithm, which further improves the ECM-based channel estimation by exploiting the error cor-

rection capability of an FEC decoder for iteratively exchanging information between the decoder and the ECM algorithm.

The rest of this paper is organized as follows: The classic system model of multiuser MIMO OFDM/SDMA is described in Section II. A brief review of the ECM algorithm is provided in the first part of Section III, and then the proposed ECM-based joint CIR and noise-variance estimation scheme proposed for a multiuser MIMO OFDM/SDMA system is elaborated in the second part of Section III. Furthermore, we proposed an FEC-A-DD channel estimation scheme based on the ECM algorithm in the third part of Section III. The computational complexity and the Cramer–Rao lower bound (CRLB) of CIR estimation is discussed in Section IV. Our computer simulation results and discussions are presented in Section V, and our conclusions are provided in Section VI.

## II. SYSTEM MODEL

The detailed schematic of the FEC-coded multiuser MIMO OFDM/SDMA UpLink (UL) system considered in this paper is shown in Fig. 1. The OFDM/SDMA system considered supports  $U$  UL users simultaneously transmitting to the BS. Each of the users has a single antenna, whereas the BS has an array of  $Q$  antennas. It is assumed that a time-division multiple-access protocol manages the division of the available time-domain (TD) resources into OFDM/SDMA time slots. The  $U$  MSs simultaneously transmit their streams of OFDM-modulated symbols to the SDMA BS. The respective UL streams are separated by the BS upon exploiting their unique user-specific “spatial signature” [2], [5].

### A. Transmitter of the Multiuser MIMO OFDM/SDMA UL System

The left part of Fig. 1 illustrates the transmitter of a multiuser MIMO SDMA/OFDM system. All of the  $U$  users transmit independent data streams, which are denoted by  $\mathbf{b}^u$ ,  $u = 1, 2, \dots, U$ . Each user  $u$  employs OFDM modulation having  $N$  subcarriers and a cyclic prefix (CP) of length  $N_{cp}$ . The information block  $\mathbf{b}^u$  is first encoded by a user-specific FEC encoder and interleaved by the interleaver. The information bits output by the interleaver are grouped into  $J$ -bit symbols and mapped to a stream of modulated data symbols, each

forming a complex number. The value of  $J$  is determined by the modulation scheme used. For example, we have  $J = 2$  for 4-quadrature-amplitude modulation (4-QAM) and  $J = 4$  for 16-QAM. The modulated data  $\tilde{\mathbf{X}}^u[k]$ ,  $k = 1, 2, \dots, K$  are then serial-to-parallel converted, and the frequency-domain (FD) pilots are embedded into certain subcarriers. The parallel modulated data (including the pilots) are further processed by inverse fast Fourier transform (FFT) to form a set of OFDM symbols. The basedband TD model of the  $n$ th sample of the  $m$ th OFDM symbol of user  $u$  can be formulated as

$$x^u[m, n] = \frac{1}{N} \sum_{i=1}^N X^u[m, i] e^{j2\pi k n t_s / T} \quad (1)$$

where  $t_s$  is the OFDM sampling interval, and  $T_s = N t_s$  is the time duration of an OFDM symbol without the CP. After concatenating the CP of  $N_{\text{cp}}$  samples, the TD signal data are transmitted through a multipath fading channel and contaminated by the receiver's additive white Gaussian noise (AWGN).

### B. Channel Model

We assume that the MIMO channel link between the MS and BS antennas is subject to independent multipath Rayleigh fading or multipath Rician fading and is spatially uncorrelated with each other, whereas  $h_q^u(t; \tau)$  denotes the CIR between the  $u$ th UL user's antenna and the  $q$ th antenna at the BS, which is a function of the delay  $\tau$  at time instant  $t$ . The CIR  $h_q^u(t; \tau)$  can be described by [2], [9]

$$h_q^u(t; \tau) = \sum_{l=1}^{L_q^u} \alpha_{q,l}^u(t) \delta(\tau - \tau_l) \quad (2)$$

where  $L_q^u$  is the number of taps of the channel link between the  $u$ th transmit user's antenna and the  $q$ th BS receive antenna at the BS, whereas  $\tau_l = l t_s$  is the delay of the  $l$ th path, and  $\alpha_{q,l}^u(t)$  is a zero-mean complex Gaussian random variable having a power-delay profile  $\theta(\tau_l)$ , with  $l$  being the propagation path index between the  $u$ th user's antenna and the  $q$ th antenna of the BS at time  $t$ .

### C. BS Receiver of the Multiuser MIMO SDMA/OFDM UL System

The optimized hierarchy reduced search algorithm (OHRSA) [17] is a sphere-decoder-like multiuser detector (MUD) that is used at the BS, as shown in the right part of Fig. 1. The received signals are, again, conventional OFDM signals [5]. The CP is discarded from every OFDM symbol, and the resultant signal is fed into the corresponding FFT-based receiver. Let  $Y_q[s, n]$  denote the signal received by the  $q$ th receiver antenna element in the  $n$ th subcarrier of the  $s$ th OFDM symbol, which is given as the superposition of the different users' channel-impaired received signal contributions plus the AWGN, which is expressed as [2]

$$Y_q[s, n] = \sum_{u=1}^U H_q^u[s, n] X^u[s, n] + W_q[s, n] \quad (3)$$

where  $H_q^u[s, n]$  denotes the FD-CHTF of the channel link between the  $u$ th user and the  $q$ th receiver antenna in the  $n$ th subcarrier of the  $s$ th OFDM symbol, which can be expressed as

$$H_q^u[s, n] = \sum_{l=1}^{L_q^u} h_q^u[s, l] F_N^m \quad (4)$$

where  $h_q^u[s, l] = h_q^u(T_f, n(T_s/N))$ , and  $F_N = \exp(-j(2\pi/N))$ . In the foregoing expression,  $T_f$  is the block length given by  $T_f = T_s + T_g$ , with  $T_g$  being the duration of the CP.

Upon invoking vector notations, the set of equations constituted by (3) for  $n = 1, 2, \dots, N$  can be rewritten as

$$\mathbf{Y}_q[s] = \mathbf{X}^T[s] \mathbf{H}_q[s] + \mathbf{W}_q[s] \quad (5)$$

where the superscript  $T$  of  $[\cdot]^T$  denotes the transpose, whereas  $\mathbf{Y}_q[s] \in \mathbb{C}^{N \times 1}$  and  $\mathbf{W}_q[s] \in \mathbb{C}^{N \times 1}$  are column vectors hosting the subcarrier-related variables  $Y_q[s, n]$  and  $W_q[s, n]$ , respectively. Furthermore, for the inner-product-based representation of (5), we have defined  $\mathbf{X}[s] \in \mathbb{C}^{UN \times N}$  and  $\mathbf{H}_q[s] \in \mathbb{C}^{UN \times 1}$ , which are given by

$$\mathbf{X}[s] = [\mathbf{X}^1[s], \mathbf{X}^2[s], \dots, \mathbf{X}^U[s]]^T \quad (6)$$

$$\mathbf{H}_q[s] = [\mathbf{H}_q^{1T}[s], \mathbf{H}_q^{2T}[s], \dots, \mathbf{H}_q^{UT}[s]]^T. \quad (7)$$

In (6),  $\mathbf{X}^u[s] \in \mathbb{C}^{N \times N}$  is a diagonal matrix with elements given by  $X^u[s, n]$ ,  $n = 1, 2, \dots, N$ .

To simplify our notation without any loss of generality, we will omit the receiver antenna's index  $q$  from now on, and the discrete model of the received signal associated with one of the BS antennas can be rewritten as

$$\mathbf{Y}[s] = \mathbf{X}^T[s] \mathbf{H}[s] + \mathbf{W}[s]. \quad (8)$$

## III. EXPECTATION CONDITIONAL MAXIMIZATION-BASED CHANNEL ESTIMATION FOR MULTIUSER MULTIPLE-INPUT-MULTIPLE-OUTPUT ORTHOGONAL FREQUENCY-DIVISION MULTIPLEXING/SPACE-DIVISION MULTIPLE ACCESS SYSTEMS

### A. ECM Algorithm

The EM algorithm [12] constitutes an iterative technique of finding the maximum-likelihood (ML) estimates of parameters that is particularly attractive when direct access to the data necessary to make an estimate is unavailable or when some of the data are missing. However, if ML estimation of the complete data is rather complicated, then the EM algorithm becomes less attractive because the maximization step (M-Step) is computationally unattractive. The ECM algorithm [18] is an extension of the EM algorithm, which simplifies the ML estimation of the complete data by replacing the complex M-Step of the EM algorithm by several computationally simpler conditional M-Steps (CM-Steps).

To elaborate a little further, we let  $\mathcal{B} = [\mathcal{B}_1, \mathcal{B}_2, \dots, \mathcal{B}_M]^T$  denote a possibly vector-valued parameter to be estimated from the observation vector  $\mathcal{Y}$ , which has the pdf of  $f(\mathcal{Y}|\mathcal{B})$ . Let  $\mathcal{X}$  denote the "complete" data set, which can be separated

into two components as  $\mathcal{X} = (\mathcal{Y}, \mathcal{Z})$ , where the observation vector  $\mathcal{Y}$  is referred to as the ‘‘incomplete’’ data set within the ECM framework, and  $\mathcal{Z}$  is called the ‘‘missing’’ data set. The expectation step (E-Step) of the ECM algorithm is the same as the EM algorithm, which is given as follows:

E-Step: Determine the conditional expectation of the log-likelihood function (LLF) of the complete data as follows:

$$Q(\mathcal{B}|\mathcal{B}^{(p)}) = E\left\{\log f(\mathcal{X}|\mathcal{B})|\mathcal{Y}, \mathcal{B}^{(p)}\right\} \quad (9)$$

where  $E\{\cdot\}$  represents the expectation operation, the superscript  $(p)$  denotes the iteration index, and  $\mathcal{B}^{(p)}$  is the estimate of  $\mathcal{B}$  at the  $p$ th iteration. The CM-Steps are carried out over reduced-dimensional spaces since the maximization is conditioned on specific parameter values.

CM-Step: Conditionally maximize the average LLF of the complete data over all possible values of  $\mathcal{B}_m$ ,  $m = 1, 2, \dots, M$ , whereas  $\mathcal{B}_v$ ,  $v \neq m$ , are fixed at their most recently updated values, which are formulated as

$$\hat{\mathcal{B}}_m^{(p+1)} = \arg \max_{\mathcal{B}_m} Q(\mathcal{B}|\mathcal{B}^{(p)}) \Big|_{\mathcal{B}_v = \hat{\mathcal{B}}_v^{(p)}, v \neq m}. \quad (10)$$

### B. ECM-Based Channel Estimation Scheme Proposed for Multiuser MIMO OFDM/SDMA UL Systems

We assume that the channels of the user set  $\mathbb{Q} = \{q_c | q_c = q_1, q_2, \dots, q_C\}$  are time invariant over the duration of  $S$  consecutive OFDM symbols, whereas the channels of the users set  $\mathbb{R} = \{r_d | r_d = r_1, r_2, \dots, r_D\}$  are time variant with a maximum Doppler frequency  $f_d$ , where  $C$  and  $D$  are the number of time-invariant and -variant channel links, respectively. The classification of these channels is based on their rough estimate, which was acquired, for example, with the aid of the simplified LS channel estimator of [6]. Upon invoking (4),  $\mathbf{H}[s]$  in (8) may be expressed in a vectorial form as

$$\mathbf{H}[s] = \mathbf{F}\mathbf{h}[s] \quad (11)$$

where  $\mathbf{h}[s] \in \mathbb{C}^{L \times 1}$ ,  $L$  is the total number of the CIR taps in  $\mathbf{h}[s]$ , which is given by  $L = \sum_{u=1}^U L^u$ , and  $L^u$  represents the number of CIR taps in  $\mathbf{h}^u[s]$ . The elements  $\mathbf{h}^u[s] \in \mathbb{C}^{L^u \times 1}$ ,  $u = 1, 2, \dots, U$ , are defined as

$$\mathbf{h}[s] = [\mathbf{h}^{1T}[s], \mathbf{h}^{2T}[s], \dots, \mathbf{h}^{UT}[s]]^T \quad (12)$$

$$\mathbf{h}^u[s] = [h^u[s, 1], h^u[s, 2], \dots, h^u[s, L^u]]^T \quad (13)$$

while the block diagonal matrix  $\mathbf{F} \in \mathbb{C}^{UN \times L}$  of (11) is defined as  $\mathbf{F} = \text{diag}(\mathbf{F}^1, \mathbf{F}^2, \dots, \mathbf{F}^U)$ , where  $\mathbf{F}^u$  is an  $(N \times L^u)$ -

element matrix with  $\mathbf{F}^u[n, l] = e^{-j2\pi(n-1)(l-1)/N}$ ,  $1 \leq n \leq N$ ,  $1 \leq l \leq L^u$ .

Considering that there are two types of channels, we can rewrite (8) using (11) as

$$\mathbf{Y}[s] = \sum_{q_c=q_1}^{q_C} \mathbf{X}^{q_c}[s] \mathbf{F}^{q_c} \mathbf{h}^{q_c}[s] + \sum_{r_d=r_1}^{r_D} \mathbf{X}^{r_d}[s] \mathbf{F}^{r_d} \mathbf{h}^{r_d}[s] + \mathbf{W}[s]. \quad (14)$$

In (14), we can treat  $\mathbf{h}^{q_c}[s]$  as a  $(L_0 \times 1)$ -element vector of fixed elements to be estimated over the period of  $S$  consecutive OFDM symbols, where  $L_0$  is the assumed numbers of CIR taps. The number of CIR taps  $L_0$  is assumed to be larger than the actual number of CIR taps, regardless of the user index. Again,  $\mathbf{h}^{q_c}[s] = \mathbf{h}^{q_c}$  for  $s = 1, 2, \dots, S$ .  $\mathbf{h}^{r_d}[s]$  is treated as a  $(L_0 \times 1)$ -element vector having random Gaussian elements, whose distribution obeys  $N(0, \mathbf{\Omega}^{r_d})$ . We emphasize that  $\mathbf{h}^{r_d}[s]$  is independent of the noise vector  $\mathbf{W}[s]$ , which has a distribution of  $N(0, \sigma_n^2 \mathbf{I}_P)$ . Here,  $\sigma_n^2$  is the noise variance to be estimated, and  $\mathbf{I}_P$  is the  $(P \times P)$ -element identity matrix having values of unity on the main diagonal and zeros elsewhere.

Following the terminology of the ECM algorithm, which has been briefly introduced in Section III-A, we view  $\mathcal{Y} = [\mathbf{Y}^T[1], \mathbf{Y}^T[2], \dots, \mathbf{Y}^T[S]]^T$  as the ‘‘incomplete’’ data,  $\mathcal{X} = [\mathcal{X}^T[1], \mathcal{X}^T[2], \dots, \mathcal{X}^T[S]]^T$  as the ‘‘complete’’ data, whereas  $\mathcal{Z} = [\mathbf{h}^{r_dT}[1], \mathbf{h}^{r_dT}[2], \dots, \mathbf{h}^{r_dT}[S]]^T$  is the ‘‘missing’’ data vector, respectively. More explicitly, the ‘‘complete’’ data element  $\mathcal{X}[s]$  is defined as (15), shown at the bottom of the page, where  $\mathbf{A}^{q_c}[s] = \mathbf{X}^{q_c}[s] \mathbf{F}^{q_c}$ ,  $q_c = q_1, q_2, \dots, q_C$ ,  $\mathbf{A}^{r_d}[s] = \mathbf{X}^{r_d}[s] \mathbf{F}^{r_d}$ ,  $r_d = r_1, r_2, \dots, r_D$ , and  $\hat{\mathbf{h}}^{r_v}[s]$  represents the most recent estimates from the previous iteration of the ECM algorithm. We can see from this equation that the ‘‘complete’’ data element  $\mathcal{X}[s]$  is the linear transformation of the mutually independent Gaussian random vectors  $\mathbf{h}^{r_d}[s]$  and  $\mathbf{W}[s]$ . Hence,  $\mathcal{X}[s]$  is also a Gaussian random vector, which has a multivariate normal distribution with a mean of

$$\mu_{\mathcal{X}[s]} = \begin{bmatrix} \sum_{q_c=q_1}^{q_C} \mathbf{A}^{q_c}[s] \mathbf{h}^{q_c} + \sum_{\substack{r_v=r_1 \\ r_v \neq r_d}}^{r_D} \mathbf{A}^{r_v}[s] \hat{\mathbf{h}}^{r_v}[s] \\ \mathbf{0} \end{bmatrix} \quad s = 1, 2, \dots, S \quad (16)$$

and a covariance matrix of

$$\Sigma_s = \begin{bmatrix} \Sigma_{\mathcal{Y}[s]} & \Sigma_{\mathcal{Y}[s] \mathbf{h}^{r_d}[s]} \\ \Sigma_{\mathbf{h}^{r_d}[s] \mathcal{Y}[s]} & \Sigma_{\mathbf{h}^{r_d}[s]} \end{bmatrix}, s = 1, 2, \dots, S \quad (17)$$

$$\begin{aligned} \mathcal{X}[s] &= \begin{bmatrix} \mathbf{Y}[s] \\ \mathbf{h}^{r_d}[s] \end{bmatrix} = \begin{bmatrix} \sum_{q_c=q_1}^{q_C} \mathbf{A}^{q_c}[s] \mathbf{h}^{q_c} + \sum_{\substack{r_v=r_1 \\ r_v \neq r_d}}^{r_D} \mathbf{A}^{r_v}[s] \hat{\mathbf{h}}^{r_v}[s] + \mathbf{A}^{r_d}[s] \mathbf{h}^{r_d}[s] + \mathbf{W}[s] \\ \mathbf{h}^{r_d}[s] \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{A}^{r_d}[s] & \mathbf{I}_P \\ \mathbf{I}_{L_0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{h}^{r_d}[s] \\ \sum_{q_c=q_1}^{q_C} \mathbf{A}^{q_c}[s] \mathbf{h}^{q_c} + \sum_{\substack{r_v=r_1 \\ r_v \neq r_d}}^{r_D} \mathbf{A}^{r_v}[s] \hat{\mathbf{h}}^{r_v}[s] + \mathbf{A}^{r_d}[s] \mathbf{h}^{r_d}[s] + \mathbf{W}[s] \end{bmatrix} \end{aligned} \quad (15)$$

where we have

$$\Sigma_{\mathbf{Y}[s]} = \sum_{r_d=r_1}^{r_D} \mathbf{A}^{r_d}[s] \mathbf{\Omega}^{r_d} \mathbf{A}^{r_d H}[s] + \sigma_n^2 \mathbf{I}_P \quad (18)$$

$$\Sigma_{\mathbf{Y}[s] \mathbf{h}^{r_d}[s]} = \mathbf{A}^{r_d}[s] \mathbf{\Omega}^{r_d} \quad (19)$$

$$\Sigma_{\mathbf{h}^{r_d}[s] \mathbf{Y}[s]} = \Sigma_{\mathbf{Y}[s] \mathbf{h}^{r_d}[s]}^H = \mathbf{\Omega}^{r_d} \mathbf{A}^{r_d H}[s] \quad (20)$$

$$\Sigma_{\mathbf{h}^{r_d}[s]} = \mathbf{\Omega}^{r_d}. \quad (21)$$

The vector  $\mathcal{B}$  consists of  $\mathbf{h}^{q_c}$ ,  $\sigma_n^2$ , and  $\mathbf{\Omega}^{r_d}$ . Hence, the logarithmic conditional pdf of the ‘‘complete’’ data is given by

$$\begin{aligned} \log f(\mathcal{X}|\mathcal{B}) &= \sum_{s=1}^S \log f(\mathcal{X}[s]|\mathcal{B}) \\ &= -\frac{(P+L_0)S}{2} \log(2\pi) \\ &\quad -\frac{1}{2} \sum_{s=1}^S \left[ \log |\Sigma_s| + (\mathcal{X}[s] - \mu_{\mathcal{X}[s]})^H \right. \\ &\quad \left. \times \Sigma_s^{-1} (\mathcal{X}[s] - \mu_{\mathcal{X}[s]}) \right]. \quad (22) \end{aligned}$$

To compute the E-Step of the ECM algorithm that determines the expectation of  $\log f(\mathcal{X}|\mathcal{B})$  with respect to  $\mathbf{h}^{r_d}$  conditioned on  $\mathcal{Y}$  and the latest estimate of  $\mathcal{B}$ ,  $\mathcal{B}^{(p)}$ , it can be seen from (16)–(22) that we require the following conditional moments of the missing data  $\mathbf{h}^{r_d}$ :

$$E \left\{ \mathbf{h}^{r_d}[s] | \mathbf{Y}[s], \mathcal{B}^{(p)} \right\} \quad (23)$$

$$E \left\{ \mathbf{h}^{r_d}[s] \mathbf{h}^{r_d H}[s] | \mathbf{Y}[s], \mathcal{B}^{(p)} \right\} \quad (24)$$

which may be directly obtained from the classic results of multivariate theory [19].

According to the Bayesian Gauss–Markov theorem [19], the linear minimum MSE (LMMSE) estimator of the variance matrix  $\mathbf{\Omega}^{r_d}$ , the noise variance  $\sigma_n^2$ , and the CIR  $\mathbf{h}^{r_d}[s]$  are given by

$$\hat{\mathbf{\Omega}}^{r_d(p)} = \frac{1}{S} \sum_{s=1}^S E \left\{ \mathbf{h}^{r_d}[s] \mathbf{h}^{r_d H}[s] | \mathbf{Y}[s], \mathcal{B}^{(p-1)} \right\} \quad (25)$$

$$\hat{\sigma}_n^{2(p)} = \frac{1}{SP} \sum_{s=1}^S E \left\{ \mathbf{W}^H[s] \mathbf{W}[s] | \mathbf{Y}[s], \mathcal{B}^{(p-1)} \right\} \quad (26)$$

$$\hat{\mathbf{h}}^{r_d(p)}[s] = E \left\{ \mathbf{h}^{r_d}[s] | \mathbf{Y}[s], \mathcal{B}^{(p)} \right\}. \quad (27)$$

After further manipulations (see Appendix A), we have the following more concrete expressions derived from (25)–(27):

$$\begin{aligned} \hat{\mathbf{\Omega}}^{r_d(p)} &= \frac{1}{S} \sum_{s=1}^S \left( \Psi^{r_d(p-1)-1}[s] + \hat{\mathbf{h}}^{r_d(p-1)}[s] \hat{\mathbf{h}}^{r_d(p-1)H}[s] \right) \quad (28) \end{aligned}$$

$$\begin{aligned} \hat{\sigma}_n^{2(p)} &= \frac{1}{SP} \sum_{s=1}^S \left( \sum_{r_d=r_1}^{r_D} \text{tr} \left\{ \mathbf{A}^{r_d H}[s] \Psi^{r_d(p-1)-1}[s] \cdot \mathbf{A}^{r_d}[s] \right\} \right. \\ &\quad \left. + \widehat{\mathbf{W}}^{(p-1)H}[s] \widehat{\mathbf{W}}^{(p-1)}[s] \right) \quad (29) \end{aligned}$$

$$\begin{aligned} \hat{\mathbf{h}}^{r_d(p)}[s] &= \hat{\sigma}_n^{-2(p)} \Psi^{r_d(p)-1}[s] \mathbf{A}^{r_d H}[s] \\ &\quad \times \left( \mathbf{Y}[s] - \sum_{q_c=q_1}^{q_C} \mathbf{A}^{q_c}[s] \hat{\mathbf{h}}^{q_c(p)} - \sum_{\substack{r_v=r_1 \\ r_v \neq r_d}}^{r_D} \mathbf{A}^{r_v}[s] \hat{\mathbf{h}}^{r_v(p-1)}[s] \right) \quad (30) \end{aligned}$$

where we have

$$\Psi^{r_d(p)}[s] = \hat{\sigma}_n^{(p)-2} \mathbf{A}^{r_d H}[s] \mathbf{A}^{r_d}[s] + \widehat{\mathbf{\Omega}}^{r_d(p)-1} \quad (31)$$

$$\begin{aligned} \widehat{\mathbf{W}}^{(p-1)}[s] &= \left( \mathbf{Y}[s] - \sum_{q_c=q_1}^{q_C} \mathbf{A}^{q_c}[s] \hat{\mathbf{h}}^{q_c(p-1)} \right. \\ &\quad \left. - \sum_{r_d=r_1}^{r_D} \mathbf{A}^{r_d}[s] \hat{\mathbf{h}}^{r_d(p-1)}[s] \right). \quad (32) \end{aligned}$$

Having obtained  $\widehat{\mathbf{\Omega}}^{r_d(p)}$ ,  $\hat{\sigma}_n^{2(p)}$ , and  $\hat{\mathbf{h}}^{r_d(p)}[s]$ , we can expand the E-Step of (9) as

$$\begin{aligned} Q(\mathcal{B}|\mathcal{B}^{(p)}) &= -\frac{1}{2} \sum_{s=1}^S \left[ \log |\widehat{\Sigma}_s^{(p)}| + (\mathcal{X}[s] - \hat{\mu}_{\mathcal{X}[s]}^{(p)})^H \right. \\ &\quad \left. \times \widehat{\Sigma}_s^{(p)-1} (\mathcal{X}[s] - \hat{\mu}_{\mathcal{X}[s]}^{(p)}) \right] \quad (33) \end{aligned}$$

where we can omit the expected value of the constant  $-((P+L_0)S/2) \log(2\pi)$  because it does not depend on  $\mathbf{h}^{q_c}$ . In (33),  $\hat{\mu}_{\mathcal{X}[s]}^{(p)}$  and  $\widehat{\Sigma}_s^{(p)}$  are given by

$$\begin{aligned} \hat{\mu}_{\mathcal{X}[s]}^{(p)} &= \begin{bmatrix} \sum_{q_c=q_1}^{q_C} \mathbf{A}^{q_c}[s] \hat{\mathbf{h}}^{q_c(p)} + \sum_{\substack{r_v=r_1 \\ r_v \neq r_d}}^{r_D} \mathbf{A}^{r_v}[s] \hat{\mathbf{h}}^{r_v(p)}[s] \\ \mathbf{0} \end{bmatrix} \\ s &= 1, 2, \dots, S \quad (34) \end{aligned}$$

$$\widehat{\Sigma}_s^{(p)} = \begin{bmatrix} \mathbf{A}^{r_d}[s] \widehat{\mathbf{\Omega}}^{r_d(p)} \mathbf{A}^{r_d H}[s] + \hat{\sigma}_n^{2(p)} \mathbf{I}_P & \mathbf{A}^{r_d}[s] \widehat{\mathbf{\Omega}}^{r_d(p)} \\ \widehat{\mathbf{\Omega}}^{r_d(p)} \mathbf{A}^{r_d H}[s] & \widehat{\mathbf{\Omega}}^{r_d(p)} \end{bmatrix}. \quad (35)$$

The M-Step of (10) aims at calculating the new estimates for the channel taps  $\hat{\mathbf{h}}^{q_c(p+1)}$  of the  $(p+1)$ st iteration that maximizes  $Q(\mathcal{B}|\mathcal{B}^{(p)})$  given  $\mathcal{B}^{(p)}$ . The  $(p+1)$ st iteration estimates for  $\hat{\mathbf{h}}^{q_c(p+1)}$  can be obtained by direct differentiation of  $Q(\mathcal{B}|\mathcal{B}^{(p)})$ , which may be expressed as

$$\begin{aligned} \hat{\mathbf{h}}^{q_c(p+1)} &= \left( \sum_{s=1}^S \mathbf{A}^{q_c H}[s] \mathbf{A}^{q_c}[s] \right)^{-1} \sum_{s=1}^S \mathbf{A}^{q_c H}[s] \\ &\quad \times \left( \mathbf{Y}[s] - \sum_{\substack{q_v=q_1 \\ q_v \neq q_c}}^{q_C} \mathbf{A}^{q_v}[s] \hat{\mathbf{h}}^{q_v(p)} - \sum_{r_v=r_1}^{r_D} \mathbf{A}^{r_v}[s] \hat{\mathbf{h}}^{r_v(p)}[s] \right). \quad (36) \end{aligned}$$

### C. FEC-A-DD Estimation Based on the ECM Algorithm

The structure of the proposed FEC-A-DD estimation scheme, which is based on the ECM algorithm’s philosophy, is illustrated in Fig. 2. We refer to this scheme as the FEC-A-DD + ECM arrangement. The proposed channel estimation scheme

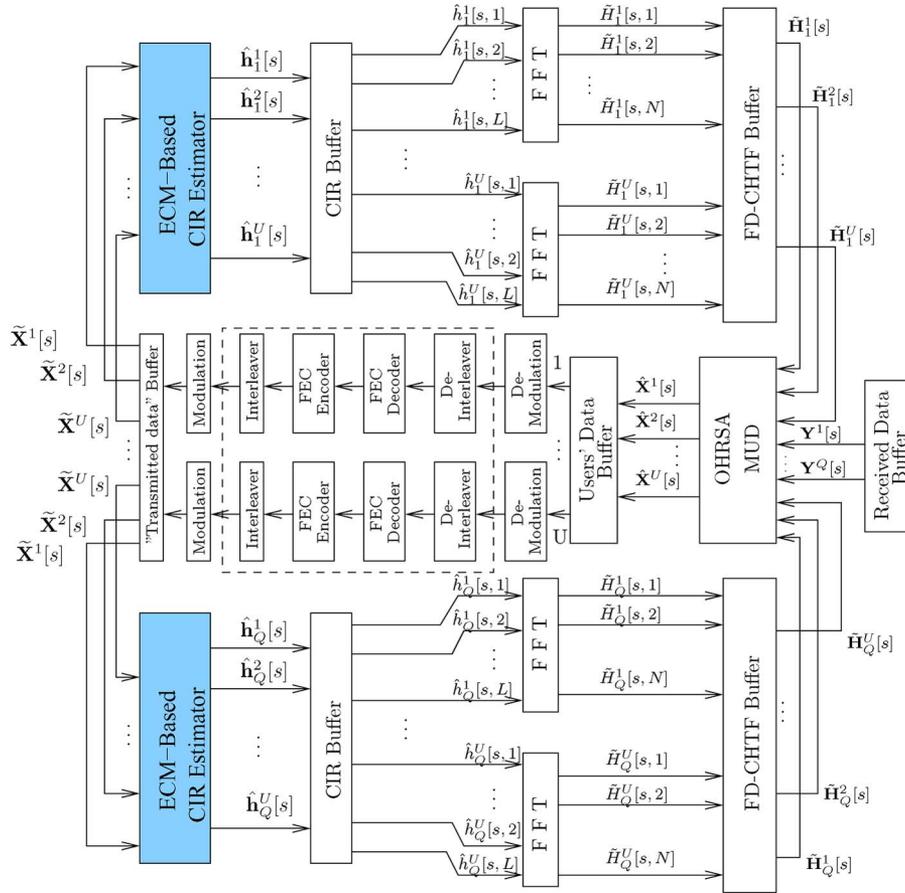


Fig. 2. Structure of the proposed FEC-A-DD FD-CHTF estimator scheme based on the ECM algorithm.

exploits the error correction capability of an arbitrary FEC decoder to mitigate the effects of noise and residual errors. It should be noted that our proposed solution is different from the conventional DD technique [20], which uses the current OFDM symbol's FD-CHTF estimate as the initial FD-CHTF estimate for the next OFDM symbol and then exploits the error correction capability of a FEC decoder to generate a more accurate FD-CHTF.<sup>1</sup> By contrast, the proposed scheme employs the FEC-A-DD technique for updating the channel estimates during the consecutive iterations of the ECM algorithm, as seen in Fig. 2. More specifically, the operation of the FEC-A-DD estimation scheme is detailed as follows.

Step 1) Activate the OHRSA-based MUD using the initial FD-CHTFs  $\hat{\mathbf{H}}^{(1)}$ .

Step 2) Reliable estimation of the transmitted signal is achieved by exploiting the error correction capability of the FEC decoder in Fig. 2. The bit stream output by the FEC decoder is not delivered to the user before the ECM-based estimator's convergence; instead, it is reencoded and remodulated to generate  $\tilde{\mathbf{X}}^{(p)}$  of Fig. 2.

Step 3) The reencoded and remodulated signal  $\tilde{\mathbf{X}}^{(p)}$  is then used in the "feedback loop" of Fig. 2 to perform joint CIR and noise-variance estimation based on the ECM algorithm.

Step 4) The CIR estimate  $\hat{\mathbf{h}}^{(p+1)}$  is then transformed to the FD by the FFT, as shown in Fig. 2. The resultant FD-CHTF  $\hat{\mathbf{H}}^{(p+1)}$  is then fed to the OHRSA MUD, according to Step 1, so that the process may continue iteration by iteration.

#### IV. DISCUSSIONS ON COMPUTATIONAL COMPLEXITY, CRAMER–RAO LOWER BOUND, AND CONVERGENCE

##### A. Analysis of the Computational Complexity

Throughout this section, the complexity is quantified in terms of the number of operations within  $S$  consecutive OFDM symbols for each iteration. More specifically, the implementational complexity is evaluated by counting the required number of complex multiplications and additions. It is noted that divisions are treated as multiplications. This may underestimate the complexity of divisions, but it has modest impact on the overall complexity estimates. Subtractions are treated as additions. Table I summarizes the number of complex multiplications and additions needed for  $S$  consecutive OFDM symbols.

Specifically, we consider a four-user four-antenna scenario, where the channels of two users' are time invariant, whereas the other two users' channels are time variant, that is, we have  $C = D = 8$ . The time-invariant channels have a constant

<sup>1</sup>We also note that other DD alternatives are also attractive, where the initial FD-CHTF estimate is derived using a low pilot-overhead, which is then exploited within the same OFDM symbol for an improved second detection. In the absence of decision errors during the first tentative detection. We may assume the presence of 100% pilots [21].

TABLE I  
 COMPARISON OF THE COMPUTATIONAL COMPLEXITY

Method: Proposed channel estimation scheme		
Multiplications	$NSQ [L_0^2(U + 2D) + L_0(U^2 + U + 2D) + U + D] + L_0^2Q(C + 2DS) + L_0^3DS$	
Additions	$L_0^2Q [C + 3DS(N + 1) + SU(N - 1)] + L_0Q [NS(U^2 + ND) - DS - SU] - NSQ(U^2 - U + ND) + (N - 1)DQ$	
Method: EM channel estimation scheme in [13]		
Multiplications	$(5L_0 + 2)NUQS$	
Additions	$(3N - 1)L_0UQS$	
Method: EM-MAP channel estimation scheme in [15]		
Multiplications	$(2N^2 + NL_0 + 4L_0 + 4)NUQS$	
Additions	$(2N^3 - 2N^2 + N^2L_0 + 2N + NL_0 - L_0)UQS$	

fading envelope for  $S = 72$  consecutive OFDM symbols, and the number of subcarriers is  $N = 64$ . The number of CIR taps is assumed to be  $L_0 = 6$ . For this special case, the numbers of multiplications and additions of the proposed scheme are 17 764 992 and 67 503 456, respectively, whereas the number of multiplications and additions are 2 359 296 and 1 320 192 for the EM-based method of [13] AND 634 355 712 and 623 437 056 for the EM-MAP based method of [15], respectively.

### B. CRLB Analysis

The CRLB [19] characterizes the best achievable performance of an unbiased estimator. Since the ML estimator is an unbiased estimator and the proposed ECM-based channel estimation scheme constitutes an iterative method of finding the ML estimate of parameters, we may characterize the proposed scheme by comparing the average MSE to the average CRLB. In the following, we derive the CRLB for the CIR of a multi-user MIMO OFDM/SDMA system to evaluate the achievable performance of the proposed ECM-based channel estimation scheme.

Upon recalling (8) and that we have  $\mathbf{A}^u[s] = \mathbf{X}^u[s]\mathbf{F}^u$ ,  $u = 1, 2, \dots, U$ , right below (15), we can write the received signal model of the multiuser MIMO OFDM/SDMA system for a single receive antenna at the BS, which supports  $U$  users as

$$\mathbf{Y}[s] = \mathbf{A}[s]\mathbf{h}[s] + \mathbf{W}[s] \quad (37)$$

where  $\mathbf{A}[s] \in \mathbb{C}^{N \times L}$ ,  $\mathbf{h}[s] \in \mathbb{C}^{L \times 1}$ ,  $\mathbf{h}[s]$  is defined by (12), and  $\mathbf{A}[s] = \mathbf{X}^T[s]\mathbf{F}$ .

Then, the estimated parameter vector is the CIR  $\mathbf{h}[s] = [\mathbf{h}^{1T}[s], \mathbf{h}^{2T}[s], \dots, \mathbf{h}^{UT}[s]]^T$ . The lower bound of the variance for the unbiased estimate of the CIR  $\mathbf{h}[s]$  is given by the CRLB as follows:

$$\text{CRLB}(\mathbf{h}_i) = \mathbf{I}^{-1}(\mathbf{h})_{ii}, i = 1, 2, \dots, L \quad (38)$$

where  $L$  is the total number of  $U$  users' channel taps, and  $\mathbf{I}(\mathbf{h})$  is the Fisher information matrix, which may be formulated as [19]

$$\begin{aligned} \mathbf{I}(\mathbf{h}) &\stackrel{\text{def}}{=} -E \left\{ \frac{\partial^2 \log f(\mathbf{Y}|\mathbf{h})}{\partial \mathbf{h} \partial \mathbf{h}^H} \right\} \\ &= -E \left\{ \frac{\partial}{\partial \mathbf{h}^H} \left( \frac{\partial \log f(\mathbf{Y}|\mathbf{h})}{\partial \mathbf{h}^H} \right)^H \right\} \end{aligned} \quad (39)$$

where  $\log f(\mathbf{Y}|\mathbf{h})$  is defined as

$$\begin{aligned} \log f(\mathbf{Y}|\mathbf{h}) &\stackrel{\text{def}}{=} \sum_{s=1}^S \log f(\mathbf{Y}[s]|\mathbf{h}) \\ &= \frac{1}{(2\pi\sigma_n^2)^N} \exp \left( -\frac{1}{2\sigma_n^2} \|\mathbf{Y} - \mathbf{A}[s]\mathbf{h}[s]\|^2 \right). \end{aligned} \quad (40)$$

Here, we assumed that the channels of the users in the set  $\mathbb{Q} = \{q_c | q_c = q_1, q_2, \dots, q_C\}$  are time invariant over  $S$  consecutive OFDM symbols and the channels of the users in  $\mathbb{R} = \{r_d | r_d = r_1, r_2, \dots, r_D\}$  are time variant over the same duration. Then, upon recalling (39) and (40), we arrive at

$$\mathbf{I}(\mathbf{h}^{q_c}) = \frac{1}{2\sigma_n^2} \sum_{s=1}^S \mathbf{A}^{q_c H}[s] \mathbf{A}^{q_c}[s], c = 1, 2, \dots, C \quad (41)$$

$$\mathbf{I}(\mathbf{h}^{r_d}[s]) = \frac{1}{2\sigma_n^2} \mathbf{A}^{r_d H}[s] \mathbf{A}^{r_d}[s], s = 1, 2, \dots, S, d = 1, 2, \dots, D \quad (42)$$

where we have  $\mathbf{A}^{q_c}[s] = \mathbf{X}^{q_c}[s]\mathbf{F}^{q_c}$ , and  $\mathbf{A}^{r_d}[s] = \mathbf{X}^{r_d}[s]\mathbf{F}^{r_d}$ . Then, we arrive at the CRLB of the estimated CIR in the form of

$$\text{CRLB}(\mathbf{h}^{q_c}) = \text{trace} \{ \mathbf{I}^{-1}(\mathbf{h}^{q_c}) \} \quad (43)$$

$$\text{CRLB}(\mathbf{h}^{r_d}[s]) = \text{trace} \{ \mathbf{I}^{-1}(\mathbf{h}^{r_d}[s]) \}. \quad (44)$$

### C. Characterizing the Convergence of the ECM Algorithm

In general, neither EM-type nor any other optimization algorithms are guaranteed to converge to a global or local maximum. The ECM algorithm, which is an extension of the EM algorithm, also fails to eliminate this deficiency. Nonetheless, Meng and Rubin [18] had discussed conditions under which the ECM algorithm will converge to a local maximum. In the following, we will characterize the convergence behavior of the ECM algorithm developed in this paper. Here, we only consider the convergence speed of  $\hat{\mathbf{h}}^{q_c(p)}$ , but these discussions are also applicable to  $\hat{\mathbf{h}}^{r_d(p)}[s]$  and  $\hat{\sigma}_n^{2(p)}$ .

We may establish the relationship between  $\hat{\mathbf{h}}^{q_c(p)}$  and  $\hat{\mathbf{h}}^{q_c(p-1)}$  with the aid of the following equation:

$$\hat{\mathbf{h}}^{q_c(p)} = \mathbf{g}(\hat{\mathbf{h}}^{q_c(p-1)}) \quad (45)$$

TABLE II  
DETAILED CHARACTERISTICS OF THE FADING CHANNELS

MS	Antenna of BS	Fading	Average path gains (dB)	$F_D$ (unless specified)	$K$ -factor
MS-1	Antenna-1	Rayleigh	[0, -3, -6, -9, -12, -15]	0.03	—
	Antenna-2	Rayleigh	[0, -5, -8, -12, -14, -17]	0.03	—
	Antenna-3	Rayleigh	[0, -4, -10, -15, -17, -19]	0.03	—
	Antenna-4	Rayleigh	[0, -7, -10, -17, -20, -25]	0.03	—
MS-2	Antenna-1	Rician	[0, -2, -4, -10, -13, -16]	0	10
	Antenna-2	Rician	[0, -3, -5, -8, -15, -17]	0	10
	Antenna-3	Rician	[0, -3, -6, -15, -18, -20]	0	10
	Antenna-4	Rician	[0, -2, -5, -10, -14, -18]	0	10
MS-3	Antenna-1	Rician	[0, -5, -10, -15, -20, -25]	0.03	5
	Antenna-2	Rician	[0, -3, -6, -10, -13, -18]	0.03	5
	Antenna-3	Rician	[0, -7, -10, -14, -18, -22]	0.03	5
	Antenna-4	Rician	[0, -2, -4, -10, -17, -20]	0.03	5
MS-4	Antenna-1	Rayleigh	[0, -4, -7, -15, -19, -25]	0	—
	Antenna-2	Rayleigh	[0, -2, -8, -10, -14, -19]	0	—
	Antenna-3	Rayleigh	[0, -7, -10, -13, -19, -23]	0	—
	Antenna-4	Rayleigh	[0, -4, -8, -13, -18, -22]	0	—
MS-5	Antenna-1	Rician	[0, -3, -7, -12, -16, -22]	0	7
	Antenna-2	Rician	[0, -5, -8, -11, -17, -23]	0	7
	Antenna-3	Rician	[0, -4, -6, -10, -17, -20]	0	7
	Antenna-4	Rician	[0, -3, -9, -12, -18, -26]	0	7

where  $\mathbf{g}(\cdot)$  is an appropriate mapping function. Using the first term of the Taylor series expansion of  $\mathbf{g}(\cdot)$ , we have

$$\begin{aligned} \hat{\mathbf{h}}^{q_c(p+1)} - \hat{\mathbf{h}}^{q_c(p)} &= \mathbf{g} \left( \hat{\mathbf{h}}^{q_c(p)} \right) - \mathbf{g} \left( \hat{\mathbf{h}}^{q_c(p-1)} \right) \\ &= \mathbf{J}^{(p-1)} \left( \hat{\mathbf{h}}^{q_c(p)} - \hat{\mathbf{h}}^{q_c(p-1)} \right) \end{aligned} \quad (46)$$

where  $\mathbf{J}^{(p-1)}$  is the matrix of partial derivatives  $\mathbf{J} = \partial \mathbf{g}(\mathbf{h}^{q_c}) / \partial \mathbf{h}^{q_c}$  evaluated at  $\mathbf{h}^{q_c} = \hat{\mathbf{h}}^{q_c(p-1)}$ , which is formulated as  $\mathbf{J} = \partial \mathbf{g}(\mathbf{h}^{q_c}) / \partial \mathbf{h}^{q_c} |_{\mathbf{h}^{q_c} = \hat{\mathbf{h}}^{q_c(p-1)}}$ . It is challenging to provide an explicit formula for  $\mathbf{J}$  since we do not have a universal expression for the appropriate mapping function  $\mathbf{g}(\cdot)$ . However, at the end of the iterations, (46) may be approximated as follows [13]:

$$\hat{\mathbf{h}}^{q_c(p+1)} - \hat{\mathbf{h}}_{ML}^{q_c} = \tilde{\mathbf{J}} \left( \hat{\mathbf{h}}^{q_c(p)} - \hat{\mathbf{h}}_{ML}^{q_c} \right) \quad (47)$$

where  $\hat{\mathbf{h}}_{ML}^{q_c}$  is the optimal ML estimate of  $\mathbf{h}^{q_c}$ , and the largest eigenvalue magnitude of the matrix  $\tilde{\mathbf{J}}$  predetermines the rate of convergence.

Similar to the derivation of the EM algorithm's rate of convergence provided in [13], after some further manipulations, we have the following streamlined expression for  $\tilde{\mathbf{J}}$ :

$$\begin{aligned} \tilde{\mathbf{J}} &= \left( \sum_{s=1}^S \mathbf{A}^{q_c H} [s] \mathbf{A}^{q_c} [s] \right)^{-1} \sum_{s=1}^S \left( \mathbf{A}^{q_c H} [s] \right. \\ &\quad \left. \cdot \sum_{r_v}^{r_D} \left( \mathbf{A}^{r_v} [s] \hat{\sigma}_n^{-2(p)} \Psi^{r_d(p-1)} [s] \mathbf{A}^{r_v H} [s] \right) \mathbf{A}^{q_c} [s] \right). \end{aligned} \quad (48)$$

## V. SIMULATION RESULTS AND DISCUSSIONS

We constructed a multiuser MIMO OFDM/SDMA UL system to demonstrate the efficiency of our proposed schemes. At the BS, we employ the OHRSA-MUD of [17] to separate the signals of the simultaneous users. The parameters of each UL transmitter are set to values similar to those of the IEEE

802.11n WLAN using  $N = 64$  subcarriers and a single sample per subcarrier. To avoid the dispersion-induced interference of consecutive OFDM symbols, a CP of 16 samples is employed as the guard interval for each OFDM symbol. Different users may employ different modulation schemes, but for simplicity, in this paper, we assume that all users employ 16-QAM modulation. Moreover, all of the users' data were protected by a 2/3-rate convolutional FEC encoder. More specifically, the constraint length was [4, 3], whereas the octally represented generator polynomials were [4, 5, 17; 7, 4, 2].

Both Rayleigh and Rician fading channels having different Doppler frequencies of  $F_D = f_d T_f$  normalized to the OFDM symbol duration were considered in our simulations, where  $f_d$  and  $T_f$  are the maximum Doppler frequency and the OFDM symbol duration including the CP, respectively. The detailed characteristics of the fading channels used in our simulations were outlined in Table II. The  $K$ -factor seen in Table II denotes the ratio of the specular power to the diffuse power of Rician fading channels. The stopping criterion of the iterations is that both  $|\hat{\mathbf{h}}^{q_c(p+1)} - \hat{\mathbf{h}}^{q_c(p)}|^2 \leq 10^{-4}$  and  $|\mathbf{h}^{r_d(p+1)} - \mathbf{h}^{r_d(p)}|^2 \leq 10^{-4}$  were met.

Fig. 3 shows the attainable MSE performance for a time-invariant slow-fading channel having a constant fading envelope for 72 consecutive OFDM symbols versus the SNR for different channel estimation schemes. The average MSE of the time-invariant channel's estimated CIRs is defined as

$$\overline{MSE}_{TI} = \frac{1}{CL_{TI}} \sum_{c=1}^C E \left\{ \|\hat{\mathbf{h}}^{q_c} - \mathbf{h}^{q_c}\|^2 \right\} \quad (49)$$

where  $C$  is the number of users, whose channel is time invariant within  $S = 72$  consecutive OFDM symbols. As expected, the proposed ECM-based joint channel estimation scheme performs close to the CRLB, which is significantly better than that of the existing EM channel estimation methods. There is an error floor for  $E_b/N_0 > 21$  dB, which may be attributable to the residual error induced by the Doppler shift of the time-variant channel links, since the signals received via the

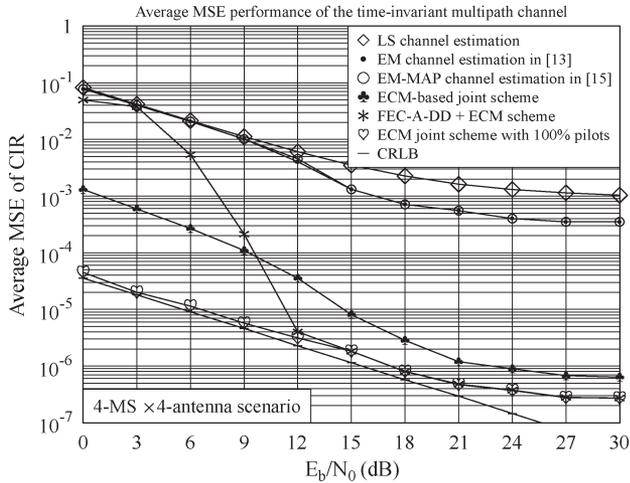


Fig. 3. MSE performance for time-invariant channels, which have a constant envelope for 72 consecutive OFDM symbols, whereas the time-variant channels have a normalized Doppler frequency of  $F_D = 0.03$ . The fading characteristics are outlined in Table II.

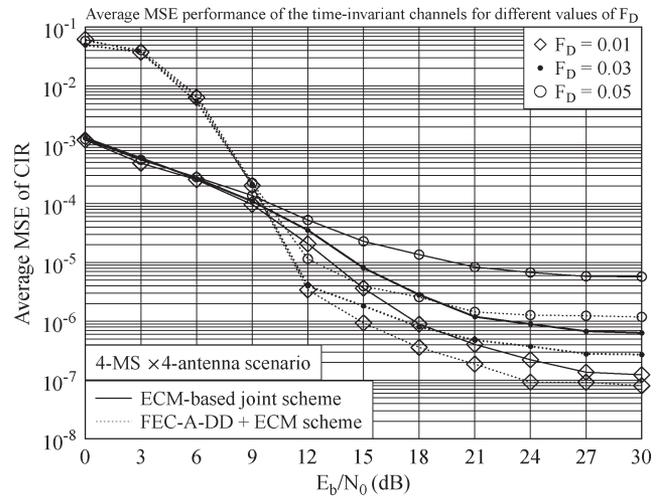


Fig. 5. MSE performance for time-invariant channels. The values of the Doppler frequency  $F_D$  marked in the figure refer to the time-variant channels. The fading characteristics are outlined in Table II.

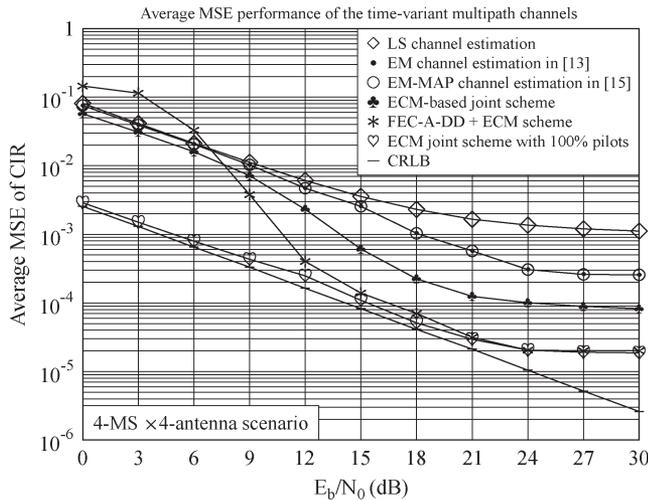


Fig. 4. MSE performance for time-variant channels, which has a normalized Doppler frequency of  $F_D = 0.03$ . The fading characteristics are outlined in Table II.

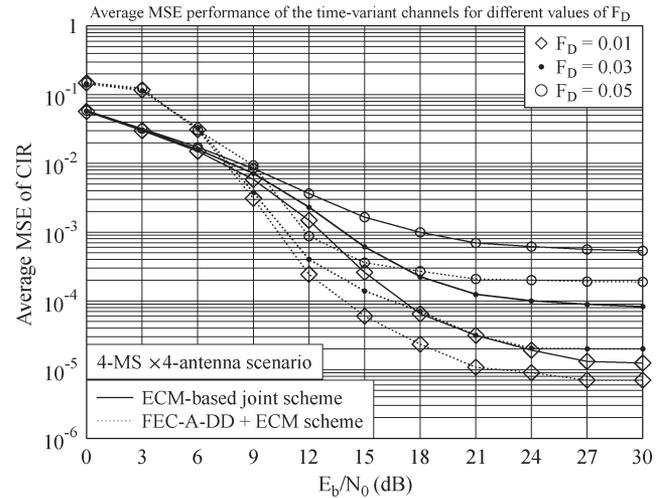


Fig. 6. MSE performance for time-variant channels at the values of the Doppler frequency  $F_D = 0.01, 0.03$ , and  $0.05$ . The fading characteristics are outlined in Table II.

time-invariant and -variant channels are superimposed on each other at the receiver antennas, hence inflicting interference.

By contrast, in Fig. 4, we characterize the average MSE performance for the time-variant channel of the four-user four-antenna MIMO OFDM/SDMA UL system. The MSE of the time-variant channel is defined as

$$\overline{MSE}_{TV} = \frac{1}{SDL_{TV}} \sum_{d=1}^D \sum_{s=1}^S E \left\{ \left\| \hat{\mathbf{h}}^{rd}[s] - \mathbf{h}^{rd}[s] \right\|^2 \right\} \quad (50)$$

where  $D$  is the number of users, whose channel is time variant. Observe from Fig. 4 that the proposed ECM-based joint channel estimation scheme obtains a slim improvement compared with the algorithms in [13] and [15]; however, similar to the time-invariant scenario in Fig. 3, the proposed FEC-A-DD + ECM scheme is capable of approaching the ideal case associated with 100% pilots by exploiting the error capability of the FEC decoder.

In Figs. 5 and 6, we portrayed the average MSE performance of the time-invariant and -variant channels, respectively, recorded for several different values of  $F_D$ . We can see from these two figures that the average MSE performance was degraded upon increasing the Doppler frequency  $F_D$ . Observe from Fig. 5 that the average MSE performance recorded for time-invariant channels was also affected by the associated Doppler shift. The reason for this detrimental influence may be that the signals received via both time-invariant and time-variant channels are superimposed on each other at the receiver antennas, which results in interference between them.

To give an overall impression, we evaluated the attainable bit error ratio (BER) in Fig. 7 both with and without convolutional FEC coding, which are shown using solid and dashed lines, respectively. Observe in Fig. 7 that although there is still a 5-dB gap to the ideal case associated with perfect channel, the BER performance of our ECM-based joint estimation scheme has a significant improvement compared with the algorithms in

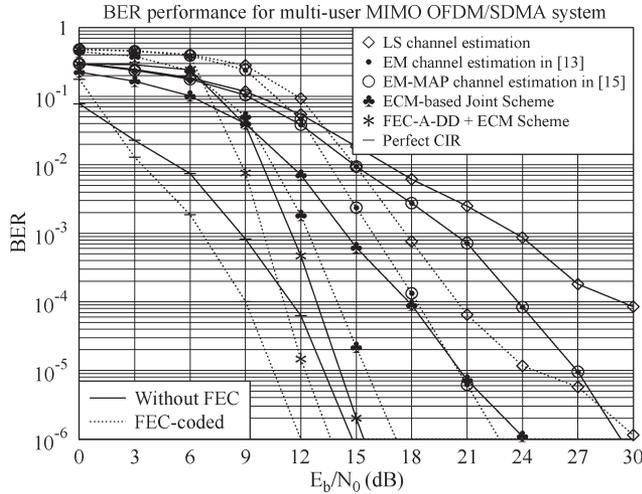


Fig. 7. BER performance for various channel estimation algorithms when two of the user’s channels are assumed to be time invariant over a frame duration of 72 consecutive OFDM symbols, whereas the other two user’s channels are assumed to be time variant with a Doppler frequency of  $F_D = 0.03$ . The system’s BER performance recorded both with and without convolutional FEC coding, namely, at the output of the demodulation and FEC decoder, respectively, when using the schematic in Fig. 1.

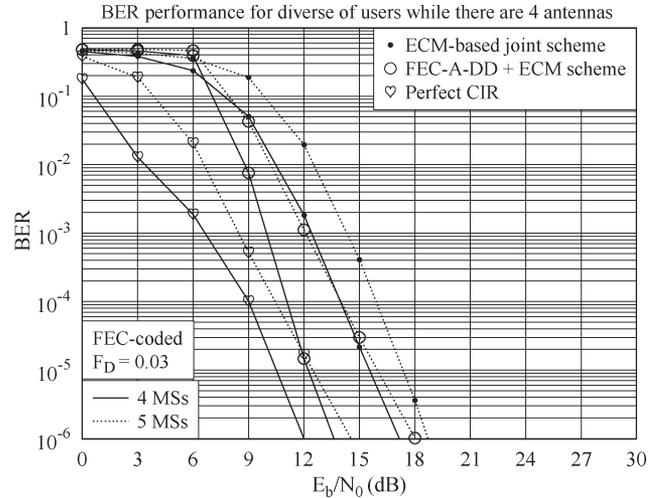


Fig. 9. BER performance versus  $E_b/N_0$  curves for four to five simultaneous users while the BS employs an array of  $Q = 4$  antennas. For the five transmitter user scenario, three of the user’s channels are assumed to be time invariant over a frame duration of 72 consecutive OFDM symbols, whereas the other two user’s channels are assumed to be time variant with a normalized Doppler frequency of  $F_D = 0.03$ . The fading characteristics are outlined in Table II. The system’s BER performance was recorded at the output of the FEC decoder when using the schematic in Fig. 1.

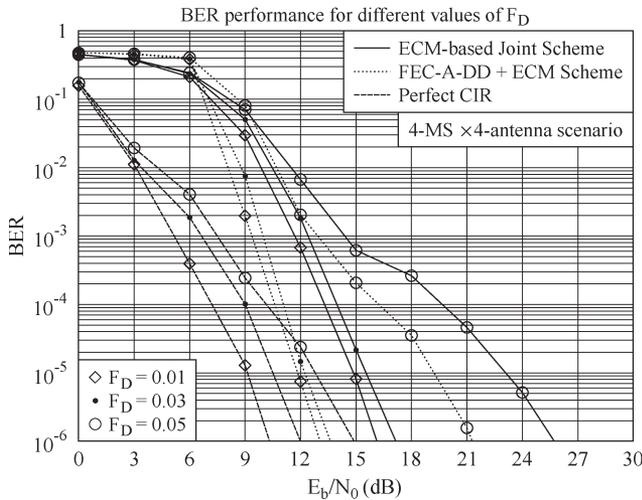


Fig. 8. BER performance at different values of Doppler frequency  $F_D$ . The system’s BER performance recorded was with convolutional FEC coding, namely, at the output of FEC decoder, respectively, when using the schematic in Fig. 1.

[13] and [15] and the initial LS channel estimation. As seen in Fig. 7, the BER performance of our FEC-A-DD + ECM scheme only had a 1.5-dB gap between the ideal case associated with perfect channel information recorded at the output of the FEC decoder. Furthermore, the BER performance of our FEC-A-DD + ECM scheme may even become better than that of the ideal case associated with perfect channel information but using independent FEC decoding, since the FEC decoder was invoked within the iterative loop. Naturally, this beneficial effect was achieved by activating the decoder more than once.

The BER performances of the proposed ECM-based joint estimation scheme and the FEC-A-DD + ECM scheme recorded at different Doppler frequencies  $F_D$  were portrayed in Fig. 8, which demonstrated that the BER performance improved upon reducing the Doppler frequency.

It is also important to consider the challenging rank-deficient scenario, when we have  $U > Q$ , because more users would like to access the system than the number of UL BS receiver antennas  $Q$ . To demonstrate the robustness of the proposed scheme for the rank-deficient scenario, in Fig. 9, we show the resultant BER versus SNR curves for  $u = 4$  and 5 simultaneous users, whereas the BS employs an array of  $Q = 4$  antennas. Observe from this figure that the proposed FEC-A-DD + ECM scheme consistently approaches the ideal case associated with a perfect CIR, regardless of the number of users supported, which demonstrates the robustness of the proposed schemes. An important additional observation is that the system employing  $Q = 4$  antennas achieved a substantial performance gain as a benefit of its increased spatial diversity when comparing the BER performance of the systems supporting four and five UL users.

The relationship between the number of EM iterations required for reliable convergence at different SNRs is shown in Fig. 10. The number of iterations required for the EM-based and EM-MAP channel estimator of [13] and [15] obeys an approximately linear reduction upon increasing the SNR, whereas the number of iterations required for the proposed ECM-based joint estimation schemes is lower for these benchmarks at both low and high SNRs. The maximum number of iterations was set to 10 in our simulations for the FEC-A-DD + ECM scheme to avoid perpetual iterations in the range of low SNRs, since the error correction capability of the FEC decoder became limited at low SNRs. Observe from Fig. 4 that the proposed FEC-A-DD + ECM scheme needs less iterations in the SNR range above  $\text{SNR} = 9$  dB, where the FEC decoder’s feedback provided a more accurate estimate of the transmitted signal. Recall from the computational complexity analysis in Section IV-A that again the proposed scheme had a nonnegligible computational complexity for each iteration; it had the lowest overall complexity since it required a reduced number of iterations.

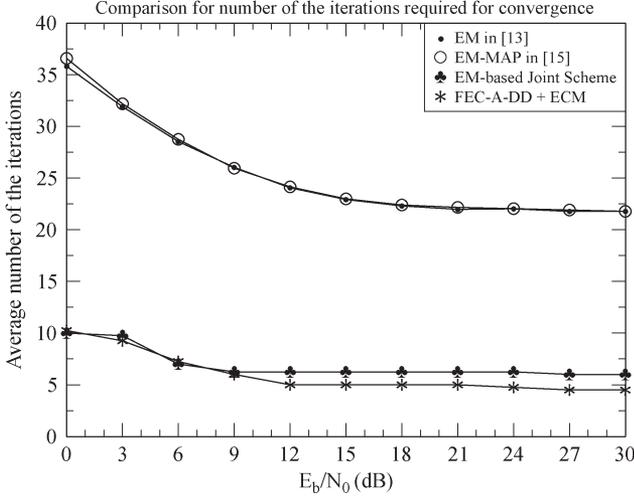


Fig. 10. Number of iterations versus  $E_b/N_0$  for the considered channel estimation schemes.

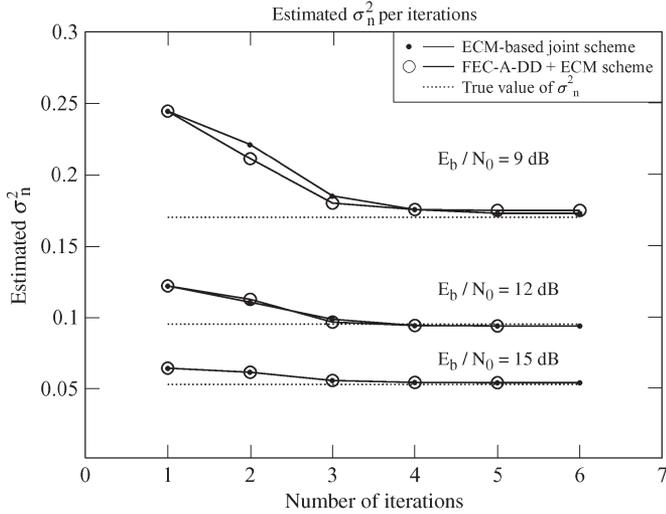


Fig. 11. Iterative convergence for the estimated  $\sigma_n^2$  versus iteration index.

In Fig. 11, we characterize the estimated  $\hat{\sigma}_n^2$  versus the number of iterations at  $E_b/N_0$  values of 9, 12, and 15 dB. In our investigations, the initial value of  $\sigma_n^2$  is set higher than the actual  $\sigma_n^2$  value. We observe that the estimated  $\hat{\sigma}_n^2$  rapidly converges to the true value of  $\sigma_n^2$ , which is indicated by the dashed line in Fig. 11.

## VI. CONCLUSION

In this paper, we have considered a realistic scenario, where some users are stationary and hence their channels are time invariant, whereas the CIRs of the other users are time variant. To cope with this scenario, we proposed a joint CIR and noise-variance estimation scheme based on the ECM algorithm designed for multiuser MIMO OFDM/SDMA UL systems. The proposed FEC-A-DD + ECM channel estimation technique is capable of approaching the CRLB, for SNR > 9 dB, especially for time-invariant channels. Our simulations demonstrated that the proposed ECM-based channel estimation scheme is capable of narrowing the SNR gap with respect to the ideal scenario associated with perfect channel information by about 5 dB

compared with the algorithms in [13] and [15], both with and without FEC coding. Furthermore, the BER performance curve of our FEC-A-DD + ECM scheme exhibited a 1.5-dB gap with respect to the ideal case. Our simulation results also demonstrated that the estimated noise variance  $\hat{\sigma}_n^2$  rapidly converges to the true value of  $\sigma_n^2$ .

## APPENDIX A

### DETAILED DERIVATION OF $\hat{\Omega}^{(p)}$ , $\hat{\sigma}_n^{2(p)}$ , AND $\hat{\mathbf{h}}^{r_d(p)}[s]$

Since  $\mathcal{X}[s] = [\mathbf{Y}^{r_d}[s]]$  are jointly complex Gaussian distributed, then the conditional pdf  $f(\mathbf{h}^{r_d}[s]|\mathbf{Y}[s], \mathcal{B}^{(p)})$  is also complex Gaussian type with mean vector and variance matrix following

$$E \left\{ \mathbf{h}^{r_d}[s] | \mathbf{Y}[s], \mathcal{B}^{(p)} \right\} = \mu_{\mathbf{h}^{r_d}[s]} + \mathbf{\Omega}^{r_d} \mathbf{A}^{r_d H}[s] \left( \mathbf{A}^{r_d}[s] \mathbf{\Omega}^{r_d} \mathbf{A}^{r_d H}[s] + \sigma_n^2 \mathbf{I}_{L_0} \right)^{-1} \cdot \left( \mathbf{Y}[s] - \sum_{q_c=q_1}^{q_C} \mathbf{A}^{q_c}[s] \hat{\mathbf{h}}^{q_c(p)} - \sum_{\substack{r_v=r_1 \\ r_v \neq r_d}}^{r_D} \mathbf{A}^{r_v}[s] \hat{\mathbf{h}}^{r_v(p)}[s] \right) \quad (51)$$

$$\text{Cov} \left\{ \mathbf{h}^{r_d}[s] | \mathbf{Y}[s], \mathcal{B}^{(p)} \right\} = \mathbf{\Omega}^{r_d} - \mathbf{\Omega}^{r_d} \mathbf{A}^{r_d H}[s] \left( \mathbf{A}^{r_d}[s] \mathbf{\Omega}^{r_d} \mathbf{A}^{r_d H}[s] + \sigma_n^2 \mathbf{I}_{L_0} \right)^{-1} \times \mathbf{A}^{r_d}[s] \mathbf{\Omega}^{r_d}. \quad (52)$$

According to the Bayesian Gauss–Markov theorem [19], the LMMSE estimator of the self-variance matrix  $\mathbf{\Omega}^{r_d}$  and the CIR  $\mathbf{h}^{r_d}[s]$  is given by

$$\mathbf{\Omega}^{r_d(p)} = \frac{1}{S} \sum_{s=1}^S E \left\{ \mathbf{h}^{r_d}[s] \mathbf{h}^{r_d H}[s] | \mathbf{Y}[s], \mathcal{B}^{(p-1)} \right\} = \frac{1}{S} \sum_{s=1}^S \left[ \text{Cov} \left\{ \mathbf{h}^{r_d}[s] | \mathbf{Y}[s], \mathcal{B}^{(p-1)} \right\} + \hat{\mathbf{h}}^{r_d(p-1)}[s] \hat{\mathbf{h}}^{r_d(p-1) H}[s] \right] \quad (53)$$

$$\hat{\mathbf{h}}^{r_d(p)}[s] = E \left\{ \mathbf{h}^{r_d}[s] | \mathbf{Y}[s], \mathcal{B}^{(p)} \right\}. \quad (54)$$

It is easy to obtain that

$$\text{Cov} \left\{ \mathbf{h}^{r_d}[s] | \mathbf{Y}[s], \mathcal{B}^{(p-1)} \right\} = \hat{\mathbf{\Omega}}^{(p-1)} - \hat{\mathbf{\Omega}}^{r_d(p-1)} \mathbf{A}^{r_d H}[s] \times \left( \mathbf{A}^{r_d}[s] \hat{\mathbf{\Omega}}^{r_d(p-1)} \mathbf{A}^{r_d H}[s] + \hat{\sigma}_n^{2(p-1)} \mathbf{I}_{L_0} \right)^{-1} \times \mathbf{A}^{r_d}[s] \hat{\mathbf{\Omega}}^{r_d(p-1)} \quad (55)$$

$$E \left\{ \mathbf{h}^{r_d}[s] | \mathbf{Y}[s], \mathcal{B}^{(p)} \right\} = \hat{\mathbf{\Omega}}^{r_d(p)} \mathbf{A}^{r_d H}[s] \left( \mathbf{A}^{r_d}[s] \hat{\mathbf{\Omega}}^{r_d(p)} \mathbf{A}^{r_d H}[s] + \hat{\sigma}_n^{2(p)} \mathbf{I}_{L_0} \right)^{-1} \cdot \left( \mathbf{Y}[s] - \sum_{q_c=q_1}^{q_C} \mathbf{A}^{q_c}[s] \hat{\mathbf{h}}^{q_c(p)}[s] - \sum_{\substack{r_v=r_1 \\ r_v \neq r_d}}^{r_D} \mathbf{A}^{r_v}[s] \hat{\mathbf{h}}^{r_v(p)}[s] \right). \quad (56)$$

Recalling the matrix inversion lemma, we can obtain that

$$\begin{aligned}
& \widehat{\mathbf{\Omega}}^{r_d(p-1)} - \widehat{\mathbf{\Omega}}^{r_d(p-1)} \mathbf{A}^{r_dH} [s] \\
& \times \left( \mathbf{A}^{r_d} [s] \widehat{\mathbf{\Omega}}^{r_d(p-1)} \mathbf{A}^{r_dH} [s] + \hat{\sigma}_n^{2(p-1)} \mathbf{I}_{L_0} \right)^{-1} \\
& \times \mathbf{A}^{r_d} [s] \widehat{\mathbf{\Omega}}^{r_d(p-1)} \\
& = \left( \hat{\sigma}_n^{(p-1)-2} \mathbf{A}^{r_dH} [s] \mathbf{A}^{r_d} [s] + \widehat{\mathbf{\Omega}}^{r_d(p-1)-1} \right)^{-1} \\
& = \Psi^{r_d(p-1)-1} [s]. \tag{57}
\end{aligned}$$

Substituting (53), (55), and (57) into (25), it is easy to obtain (28).

$E\{\mathbf{W}[s]\mathbf{W}^H[s]|\mathbf{Y}[s], \mathcal{B}^{(p-1)}\}$  in (26) can be rewritten as

$$\begin{aligned}
& E\left\{\mathbf{W}^H[s]\mathbf{W}[s]|\mathbf{Y}[s], \mathcal{B}^{(p-1)}\right\} \\
& = \text{tr}\left\{E\left\{\mathbf{W}[s]\mathbf{W}^H[s]|\mathbf{Y}[s], \mathcal{B}^{(p-1)}\right\}\right\} \\
& = \text{tr}\left\{\text{Cov}\left\{\mathbf{W}[s]|\mathbf{Y}[s], \mathcal{B}^{(p-1)}\right\}\right\} \\
& \quad + \widehat{\mathbf{W}}^{(p-1)H} [s] \widehat{\mathbf{W}}^{(p-1)} [s]. \tag{58}
\end{aligned}$$

Upon invoking the definition of  $\widehat{\mathbf{W}}^{(p-1)} [s]$  in (32),  $\text{Cov}\{\mathbf{W}[s]|\mathbf{Y}[s], \mathcal{B}^{(p-1)}\}$  can be rewritten as

$$\begin{aligned}
& \text{Cov}\left\{\mathbf{W}[s]|\mathbf{Y}[s], \mathcal{B}^{(p-1)}\right\} \\
& = \sum_{r_d=r_1}^{r_D} \text{Cov}\left\{\mathbf{A}^{r_d} [s] \mathbf{h}^{r_d} [s] |\mathbf{Y}[s], \mathcal{B}^{(p-1)}\right\} \\
& = \sum_{r_d=r_1}^{r_D} \mathbf{A}^{r_d} [s] \text{Cov}\left\{\mathbf{h}^{r_d} [s] |\mathbf{Y}[s], \mathcal{B}^{(p-1)}\right\} \mathbf{A}^{r_dH} [s]. \tag{59}
\end{aligned}$$

Substituting (57)–(59) into (26), we can obtain (29). To obtain (30), we make use of the identity [19]

$$\begin{aligned}
& \widehat{\mathbf{\Omega}}^{r_d(p)} \mathbf{A}^{r_dH} [s] \left( \mathbf{A}^{r_d} [s] \widehat{\mathbf{\Omega}}^{r_d(p)} \mathbf{A}^{r_dH} [s] + \hat{\sigma}_n^{2(p)} \mathbf{I}_{L_0} \right)^{-1} \\
& = \hat{\sigma}_n^{(p)-2} \left( \hat{\sigma}_n^{(p)-2} \mathbf{A}^{r_dH} [s] \mathbf{A}^{r_d} [s] + \widehat{\mathbf{\Omega}}^{r_d(p)-1} \right)^{-1} \mathbf{A}^{r_dH} [s]. \tag{60}
\end{aligned}$$

Substituting (56), (57), and (60) into (27), we can obtain (30).

## REFERENCES

- [1] M. Jiang and L. Hanzo, "Multiuser MIMO-OFDM for next-generation wireless systems," *Proc. IEEE*, vol. 95, no. 7, pp. 1430–1469, Jul. 2007.
- [2] L. Hanzo, M. Münster, B. J. Choi, and T. Keller, *OFDM and MC-CDMA for Broadband Multi-User Communications, WLANs, and Broadcasting*. Piscataway, NJ: IEEE, 2003.
- [3] M. Jiang, J. Akhtman, and L. Hanzo, "Iterative joint channel estimation and multi-user detection for multiple-antenna aided OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 2904–2914, Aug. 2007.
- [4] X. Dai, "Pilot-aided OFDM/SDMA channel estimation with unknown timing offset," *Proc. Inst. Elect. Eng.—Commun.*, vol. 153, no. 3, pp. 392–398, Jun. 2006.

- [5] P. Vandenameele, L. Van Der Perre, M. Engels, B. Gyselinckx, and H. De Man, "A combined OFDM/SDMA approach," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 11, pp. 2312–2321, Nov. 2000.
- [6] Y. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 67–75, Jan. 2002.
- [7] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," *IEEE Trans. Signal Process.*, vol. 51, no. 6, pp. 1615–1624, Jun. 2003.
- [8] H. Zhang, Y. Li, A. Reid, and J. Terry, "Optimum training symbol design for MIMO OFDM in correlated fading channels," *IEEE Trans. Wireless Commun.*, vol. 5, no. 9, pp. 2343–2347, Sep. 2006.
- [9] Y. Li, N. Seshadri, and S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 461–471, Mar. 1999.
- [10] H. Minn, D. Kim, and V. Bhargava, "A reduced complexity channel estimation for OFDM systems with transmit diversity in mobile wireless channels," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 799–807, May 2002.
- [11] Y. Li, J. Winters, and N. Sollenberger, "MIMO-OFDM for wireless communications: Signal detection with enhanced channel estimation," *IEEE Trans. Commun.*, vol. 50, no. 9, pp. 1471–1477, Sep. 2002.
- [12] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Stat. Soc. Ser. B (Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.
- [13] Y. Z. Xie and C. N. Georghiades, "Two EM-type channel estimation algorithms for OFDM with transmitter diversity," *IEEE Trans. Commun.*, vol. 51, no. 1, pp. 106–115, Jan. 2003.
- [14] X. Wautelet, C. Herzet, A. Dejonghe, J. Louveaux, and L. Vandendorpe, "Comparison of EM-based algorithms for MIMO channel estimation," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 216–226, Jan. 2007.
- [15] J. Gao and H. Liu, "Low-complexity MAP channel estimation for mobile MIMO-OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 3, pp. 774–780, Mar. 2008.
- [16] J. Choi, "An EM based joint data detection and channel estimation incorporating with initial channel estimate," *IEEE Commun. Lett.*, vol. 12, no. 9, pp. 654–656, Sep. 2008.
- [17] J. Akhtman, A. Wolfgang, S. Chen, and L. Hanzo, "An optimized-hierarchy-aided approximate log-MAP detector for MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 5, pp. 1900–1909, May 2007.
- [18] X. Meng and D. Rubin, "Maximum likelihood estimation via the ECM algorithm: A general framework," *Biometrika*, vol. 80, no. 2, pp. 267–278, Jun. 1993.
- [19] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1993.
- [20] V. Mignone and A. Morello, "CD3-OFDM: A novel demodulation scheme for fixed and mobile receivers," *IEEE Trans. Commun.*, vol. 44, no. 9, pp. 1144–1151, Sep. 1996.
- [21] M. Münster and L. Hanzo, "Parallel-interference-cancellation-assisted decision-directed channel estimation for OFDM systems using multiple transmit antennas," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2148–2162, Sep. 2005.



**Jiankang Zhang** (S'08) received the B.S. degree in mathematics and applied mathematics from the Beijing University of Posts and Telecommunications, Beijing, China, in 2006. He is currently working toward the Ph.D. degree with Zhengzhou University, Zhengzhou, China.

From September 2009 to September 2011, he was a Visiting Student with the School of Electronics and Computer Science, University of Southampton, Southampton, U.K. His research interests are in the areas of wireless communications and signal processing, including space-division multiple access, orthogonal frequency-division multiplexing, channel estimation, and multiuser detection.



**Lajos Hanzo** (M'91–SM'92–F'04) received the degree in electronics in 1976 and the doctorate degree in 1983. In 2009, he received the honorary doctorate “Doctor Honaris Causa” degree from the Technical University of Budapest, Budapest, Hungary.

During his 35-year career in telecommunications, he has held various research and academic posts in Hungary, Germany and the U.K. Since 1986, he has been with the School of Electronics and Computer Science, University of Southampton, U.K., where he holds the chair in telecommunications. He has co-authored 20 John Wiley/IEEE Press books on mobile radio communications totalling in excess of 10 000 pages, published about 990 research entries on IEEE Xplore, acted as Technical Program Committee Chair of IEEE conferences, presented keynote lectures, and been awarded a number of distinctions. Currently, he is directing an academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council (EPSRC) U.K., the European IST Programme, and the Mobile Virtual Centre of Excellence (VCE), U.K. He is an enthusiastic supporter of industrial and academic liaison, and he offers a range of industrial courses. He is also a Governor of the IEEE Vehicular Technology Society. Since 2008, he has been the Editor-in-Chief of the IEEE Press and, since 2009, a Chaired Professor, also at Tsinghua University, Beijing, China. For further information on research in progress and associated publications, see <http://www-mobile.ecs.soton.ac.uk>.



**Xiaomin Mu** received the M.S. degree from the Beijing Institute of Technology, Beijing, China, in 1982.

She is currently a Full Professor with the School of Information Engineering, Zhengzhou University, Zhengzhou, China. She has published various papers in the field of signal processing and coauthored two books. Her research interests include signal processing in communications system, wireless communications, and cognitive radio.