A Joint Channel Estimation and Data Detection Receiver for Multiuser MIMO IFDMA Systems

The-Hanh Pham, Ying-Chang Liang, *Senior Member, IEEE*, and A. Nallanathan, *Senior Member, IEEE*

***Abstract*—In this paper, we are interested in the problem of joint channel estimation and data detection for multi-input multi-output (MIMO) interleaved frequency division multiple access (IFDMA) systems. Although IFDMA is free from the multiple access interference (MAI), it suffers from intersymbol interference (ISI). MIMO-IFDMA system suffers from both ISI and multi-stream interference (MSI). The block iterative generalized decision feedback equalizer (BI-GDFE) is an iterative and effective interference cancelation scheme which could provide near maximum likelihood (ML) performance with very low complexity. However, BI-GDFE needs the channel state information (CSI). In this paper, we utilize the soft estimates of the transmitted symbols provided by the BI-GDFE to estimate the channel via an Expectation Maximization (EM)-based algorithm. By doing so, a joint channel estimation and data detection receiver is developed. To evaluate the performance of the proposed channel estimation algorithm, we derive the Cramér-Rao Lower Bound (CRLB). Computer simulations show that the bit error rate (BER) performance of the proposed joint channel estimation and signal detection receiver can reach the performance of the BI-GDFE with perfect CSI.**

***Index Terms*—Multi-input Multi-ouput (MIMO), interleaved frequency division multiple access (IFDMA), channel estimation, expectation maximization (EM) algorithm.**

I. Introduction

INTERLEAVED frequency division multiple access (IFDMA) [1, 2] has been accepted as the uplink air-interface for the third generation long term evolution cellular mobile systems (3G LTE) [3, 4]. IFDMA is a new spread spectrum multiple access scheme which combines the advantages of spread spectrum and multi-carrier transmission. In the multi-user scenarios, a set of sub-carriers is assigned to each user. These sets of sub-carriers are orthogonal to each other. Hence, no multiple access interference (MAI) arises even when the transmission is under a severe frequency-selective fading channel. At the receiver, user discrimination is accomplished by using frequency division multiple access. By selecting the sub-carriers associated to a particular user from the set of equally-spaced sub-carriers, IFDMA is a single carrier based system in nature; thus, it enjoys a lower peak-to-average power ratio (PAPR) as compared to multi-carrier systems.

Paper approved by N. Al-Dhahir, the Editor for Space-Time, OFDM and Equalization of the IEEE Communications Society. Manuscript received October 24, 2007; revised January 30, 2008 and March 28, 2008.

T.-H. Pham is with the Department of Electrical & Computer Engineering, National University of Singapore, Singapore 119260 (e-mail: g0301029@nus.edu.sg).

Y.-C. Liang is with the Institute for Infocomm Research (I2R), A*STAR, Singapore (e-mail: ycliang@i2r.a-star.edu.sg).

A. Nallanathan is with the Division of Engineering, King's College London, United Kingdom (e-mail: nallanathan@ieee.org).

Digital Object Identifier 10.1109/TCOMM.2009.06.070555

Although IFDMA system is free of MAI, it still suffers from inter-symbol interference (ISI). Moreover, multi-antenna transmission over multi-input multi-output (MIMO) channels has been proven to be effective in increasing the channel capacity [5, 6]. Therefore, the MIMO system can be incorporated into IFDMA to satisfy the increasing demands of high data rate applications. However, MIMO-IFDMA suffers not only from ISI but also heavily from multi-stream interference (MSI) resulting from the superposition of multiple transmitted signal streams at the receiver. As a consequence, advanced equalizers are needed to recover the transmitted signals.

Linear receivers such as zero-forcing (ZF) and minimum mean square error (MMSE) equalizers [7] have been proposed to detect the transmitted symbols. Although simple to implement, performance of the linear equalizers is far from that of the ML detection bound. To improve the performance, nonlinear equalizers such as the generalized decision feedback equalizer (GDFE) [8] have been introduced. However, the performance of these receivers may not reach the ML bound due to the error propagation. In [9], a MMSE-based iterative soft interference cancelation (MMSE-SIC) receiver has been proposed. In this work, the interference is canceled by using the soft estimates of the symbols. These soft estimates are the log-likelihood ratios (LLRs) of the symbols. The complexity of MMSE-SIC is low compared to the ML detection. However, it still incurs complex computations in the calculation of filter weights. To reduce the computational complexity further, the Block-Iterative GDFE (BI-GDFE) receiver has been introduced in [10]. The BI-GDFE receiver iteratively and simultaneously (but not jointly) detects the transmitted symbols by utilizing the decisions from the previous iteration to cancel out the interference. Based on a statistical reliability factor of hard decisions for each iteration, namely the input-decision correlation (IDC), interference is reduced iteratively, and this improves the bit error rate (BER) performance of the system. The low complexity of the BI-GDFE receiver relies on the fact that the equalizer coefficients can be determined in an off-line manner. However, BI-GDFE needs the channel state information (CSI). In this paper, we utilize the hard decision of the transmitted symbols and the IDC provided by the BI-GDFE for channel estimation via an EM-based algorithm. The combination of the transmitted symbols' hard decisions together with the IDC can be viewed as the soft information of the transmitted symbols provided by the BI-GDFE.

When the number of repetitions in IFDMA systems equals to 1, MIMO-IFDMA model reduces to a single-user MIMO single carrier cyclic-prefix system [11, 12] (MIMO-SCCP). Therefore, our proposed receiver can also be used to a MIMO-

SCCP system. To evaluate the performance of channel estimation, we also derive the Cramér-Rao Lower Bound (CRLB). Because of the difficulty in finding a closed-form expression of the exact CRLB, we rely on the so-called *modified* CRLB [13]. The obtained closed-form modified CRLB agrees with the exact CRLB through simulations; therefore, it is used to evaluate the mean square error (MSE) performance of the estimated parameters.

The rest of the paper is organized as follows. Section II describes the system model of the multi-user MIMO-IFDMA and the BI-GDFE receiver. The details of the EM-based channel estimation algorithm is given in Section III. CRLB derivation is presented in Section IV. In Section V, computer simulations are provided. Finally, conclusions are drawn in Section VI and necessary derivations are given in the Appendix.

Notations: Capital bold letters denote matrices and small bold letters denote row/column vectors; transpose, conjugate and Hermitian conjugate of a vector/matrix are denoted by $(\cdot)^{\mathcal{T}}$, $(\cdot)^*$ and $(\cdot)^{\mathcal{H}}$, respectively. The $(m,n)^{th}$ element of a matrix $\boldsymbol{A}$ is denoted by $(\boldsymbol{A})_{m,n}$. $\mathrm{tr}\{\boldsymbol{A}\}$ is the trace of the matrix $\boldsymbol{A}$. $\mathrm{diag}\{\boldsymbol{A}\}$ is a diagonal matrix which takes the diagonal elements from the matrix $\boldsymbol{A}$; $\mathrm{diag}\{\boldsymbol{a}\}$ is a diagonal matrix whose diagonal elements are from the vector $\boldsymbol{a}$. $\boldsymbol{I}_N$ denotes the identity matrix of size N. $\boldsymbol{0}_N$ is a null matrix of size $N \times N$. The Kronecker product is denoted by $\otimes$.

II. System Model and Preliminaries

In this section, we provide an overview of the single-input single-output IFDMA (SISO-IFDMA), multi-input multi-output IFDMA (MIMO-IFDMA) systems and the BI-GDFE receiver. In the literature, there are two equivalent implementations of IFDMA systems: one is in the time domain [1, 14] and the other is in the frequency domain [2]. In this paper, we adopt the time domain approach to illustrate the principle of IFDMA.

A. SISO-IFDMA

The block diagram of a SISO-IFDMA system is shown in Fig. 1. In this figure, a sequence of data symbols $\{s^{(i)}(n)\}_{n=0}^{TN-1}$ from the i^{th} user is first serial-to-parallel (S/P) converted into a frame of T blocks, each block with size N. The t^{th} block is denoted by the vector $\boldsymbol{s}^{(i)}(t) = [s^{(i)}(t;0)\ s^{(i)}(t;1)\ \cdots\ s^{(i)}(t;N-1)]^{\mathcal{T}}$, $t = 1, 2, \ldots, T$ where $s^{(i)}(t;n) = s^{(i)}(n + tN)$. The t^{th} block is compressed and repeated $R-1$ times. The resultant expression for the t^{th} IFDMA symbol is given by

$$\boldsymbol{c}^{(i)}(t) = \frac{1}{\sqrt{R}}\Big[s^{(i)}(t;0)\ s^{(i)}(t;1)\ \cdots\ s^{(i)}(t;N-1); \\ \cdots\ ; s^{(i)}(t;0)\ s^{(i)}(t;1)\ \cdots\ s^{(i)}(t;N-1)\Big]^{\mathcal{T}}. \quad (1)$$

Before transmission, the IFDMA symbol is modified by a phase vector $\boldsymbol{b}^{(i)} = [b^{(i)}(0)\ \cdots\ b^{(i)}(NR-1)]^{\mathcal{T}}$ of size NR of which the p^{th} element is given by

$$b^{(i)}(p) = \exp\{-j \cdot p \cdot \phi^{(i)}\},\ p = 0, 1, \ldots, NR-1, \quad (2)$$

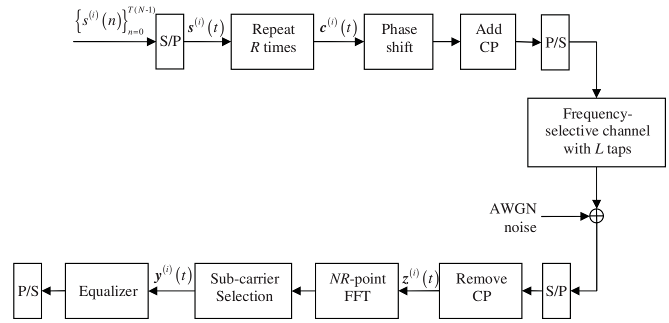

Fig. 1. The block diagram of the SISO-IFDMA system.

where $\phi^{(i)} = i\frac{2\pi}{NR}$ is called the user-dependence phase shift. The element-wise multiplication of vector $\boldsymbol{c}^{(i)}(t)$ and $\boldsymbol{b}^{(i)}$ assures that every set of frequencies assigned to each user is orthogonal to each other. The resulted signal from the i^{th} user can be written as

$$\boldsymbol{x}^{(i)}(t) = \big[c^{(i)}(t;0)\ \ c^{(i)}(t;1)e^{-j\phi^{(i)}}\ \ c^{(i)}(t;2)e^{-j2\phi^{(i)}} \\ \cdots\ \ c^{(i)}(t;NR-1)e^{-j(NR-1)\phi^{(i)}}\big]^{\mathcal{T}}. \quad (3)$$

The last P symbols of $\boldsymbol{x}^{(i)}(t)$ are used for the cyclic prefix (CP) which is added in front of $\boldsymbol{x}^{(i)}(t)$, resulting in a new block of size $(NR + P)$ to be transmitted.

This block is transmitted over a frequency-selective fading channel which is characterized by the channel vector of length L, $\boldsymbol{h}^{(i)} = [h_0^{(i)}\ h_1^{(i)}\ \cdots\ h_{L-1}^{(i)}]^{\mathcal{T}}$. We assume that the channel is static over the entire transmission of a frame. Moreover, in order to prevent interblock interference, the CP length satisfies $P \geq L$.

At the receiver side, after the S/P conversion and CP removal, the received signal $\boldsymbol{z}^{(i)}(t)$ corresponding to the t^{th} block can be represented by

$$\boldsymbol{z}^{(i)}(t) = \Big(\boldsymbol{W}_{NR}^{\mathcal{H}}\boldsymbol{\Lambda}_{NR}^{(i)}\boldsymbol{W}_{NR}\Big)\boldsymbol{x}^{(i)}(t) + \bar{\boldsymbol{n}}^{(i)}(t),\ t = 1, 2, \ldots, T, \quad (4)$$

where $\boldsymbol{\Lambda}_{NR}^{(i)} = \mathrm{diag}\big\{\lambda_0^{(i)}\ \lambda_1^{(i)}\ \cdots\ \lambda_{NR-1}^{(i)}\big\}$ with $\lambda_p^{(i)} = \sum_{l=0}^{L-1} h_l^{(i)} \exp\{-j\frac{2\pi}{NR}pl\}$ denoting the frequency response for the p^{th} subcarrier of the channel; $\boldsymbol{W}_{NR}$ is NR-point discrete Fourier transform (DFT) matrix and $\bar{\boldsymbol{n}}^{(i)}(t)$ is a realization of a zero-mean complex Gaussian random vector with covariance matrix $\sigma_n^2\boldsymbol{I}_{NR}$.

Our objective here is to detect the transmitted signal $\boldsymbol{s}^{(i)}(t)$. After the NR-point DFT operation on $\boldsymbol{z}^{(i)}(t)$ and through subcarrier selection, the transformed signal of the desired user is given by

$$\boldsymbol{y}^{(i)}(t) = \boldsymbol{\Lambda}_N^{(i)}\boldsymbol{W}_N\boldsymbol{s}^{(i)}(t) + \boldsymbol{n}^{(i)}(t),\ t = 1, 2, \ldots, T, \quad (5)$$

where $\boldsymbol{\Lambda}_N^{(i)} = \mathrm{diag}\big\{\lambda_i^{(i)}\ \lambda_{i+R}^{(i)}\ \cdots\ \lambda_{i+(N-1)R}^{(i)}\big\}$, $\boldsymbol{W}_N$ is the N-point DFT matrix and $\boldsymbol{n}^{(i)}(t)$ is a realization of zero-mean complex Gaussian random vector with covariance matrix $\sigma_n^2\boldsymbol{I}_N$.

Equation (5) can also be written in a different but equivalent form as

$$\boldsymbol{y}^{(i)}(t) = \boldsymbol{\mathcal{D}}^{(i)}(t)\boldsymbol{F}^{(i)}\boldsymbol{h}^{(i)} + \boldsymbol{n}^{(i)}(t),\ t = 1, 2, \ldots, T, \quad (6)$$

where $\boldsymbol{\mathcal{D}}^{(i)}(t) = \mathrm{diag}\{\boldsymbol{W}_N \boldsymbol{s}^{(i)}(t)\}$ and $\boldsymbol{F}^{(i)}$ is a matrix constructed from the first L columns and $i^{th}, (i+R)^{th}, \cdots, (i+(N-1)R)^{th}$ rows from $\sqrt{NR}\boldsymbol{W}_{NR}$.

When there are U ($U \leq R$) users in the system, the received signal at the base station is the superposition of signals from all users. Each user is assigned with a different user-dependent phase shift $\phi^{(i)}$, $i = 1, 2, \ldots, U$. We assume that the signals from all users are received synchronously within the CP window and each user occupies a set of sub-carriers which is orthogonal to the sub-carrier sets of other users. Thus, the users do not interfere with each other. Due to this orthogonality among users, without loss of generality, the superscript $^{(i)}$ as well as the subscript N is dropped from $\boldsymbol{\Lambda}_N^{(i)}$ and $\boldsymbol{W}_N$. Hence, (5) can be simplified as

$$\boldsymbol{y}(t) = \boldsymbol{\Lambda W s}(t) + \boldsymbol{n}(t), \tag{7}$$

and (6) is written as

$$\boldsymbol{y}(t) = \boldsymbol{\mathcal{D}}(t)\boldsymbol{F h} + \boldsymbol{n}(t). \tag{8}$$

B. MIMO-IFDMA

In this subsection, we consider a system of up to R users in which each user is equipped with N_T transmit antennas. The receiver is equipped with N_R receive antennas. In order not to exceed the total transmission bandwidth, N_T independent data streams belonging to a user are multiplexed to occupy the same set of sub-carriers. Due to the orthogonality among the sets of sub-carriers, which is explained in the previous subsection, the user index superscript is omitted.

With these assumptions, the received signal at the k^{th} receive antenna can be written as

$$\boldsymbol{y}_k(t) = \sum_{l=1}^{N_T} \boldsymbol{\Lambda}_{k,l} \boldsymbol{W s}_l(t) + \boldsymbol{n}_k(t),\ t = 1, 2, \ldots, T, \tag{9}$$

where $\boldsymbol{\Lambda}_{k,l}$ is the diagonal matrix consisting of the frequency responses at the appropriate sub-carriers of the channel from the l^{th} transmit antenna to the k^{th} receive antenna; $\boldsymbol{s}_l(t)$ is the t^{th} signal block transmitted from the l^{th} transmit antenna; $\boldsymbol{n}_k(t)$ is the additive noise at the k^{th} receive antenna which is a realization of a zero-mean complex Gaussian random vector with covariance matrix of $\sigma_n^2 \boldsymbol{I}_N$. If we collect N_R received vectors $\boldsymbol{y}_k(t)$, $k = 1, 2, \ldots, N_R$ to form a vector $\boldsymbol{y}(t) = [\boldsymbol{y}_1^{\mathcal{T}}(t)\ \boldsymbol{y}_2^{\mathcal{T}}(t) \cdots\ \boldsymbol{y}_{N_R}^{\mathcal{T}}(t)]^{\mathcal{T}}$, this vector can be written as

$$\boldsymbol{y}(t) = \tilde{\boldsymbol{\Lambda}}\tilde{\boldsymbol{W}}\boldsymbol{s}(t) + \boldsymbol{n}(t),\ t = 1, 2, \ldots, T, \tag{10}$$

where

$$\tilde{\boldsymbol{\Lambda}} = \begin{bmatrix} \boldsymbol{\Lambda}_{1,1} & \boldsymbol{\Lambda}_{1,2} & \cdots & \boldsymbol{\Lambda}_{1,N_T} \\ \boldsymbol{\Lambda}_{2,1} & \boldsymbol{\Lambda}_{2,2} & \cdots & \boldsymbol{\Lambda}_{2,N_T} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\Lambda}_{N_R,1} & \boldsymbol{\Lambda}_{N_R,2} & \cdots & \boldsymbol{\Lambda}_{N_R,N_T} \end{bmatrix}, \tag{11}$$

$$\tilde{\boldsymbol{W}} = \begin{bmatrix} \boldsymbol{W} & \boldsymbol{0}_N & \cdots & \boldsymbol{0}_N \\ \boldsymbol{0}_N & \boldsymbol{W} & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{0}_N & \cdots & \cdots & \boldsymbol{W} \end{bmatrix}, \tag{12}$$

$$\boldsymbol{s}(t) = [\boldsymbol{s}_1^{\mathcal{T}}(t)\ \boldsymbol{s}_2^{\mathcal{T}}(t)\ \cdots\ \boldsymbol{s}_{N_T}^{\mathcal{T}}(t)]^{\mathcal{T}}, \tag{13}$$

$$\boldsymbol{n}(t) = [\boldsymbol{n}_1^{\mathcal{T}}(t)\ \boldsymbol{n}_2^{\mathcal{T}}(t)\ \cdots\ \boldsymbol{n}_{N_T}^{\mathcal{T}}(t)]^{\mathcal{T}}. \tag{14}$$

Denote $\boldsymbol{h}_{k,l}$ as the channel vector from the l^{th} transmit antenna to the k^{th} receive antenna and $\boldsymbol{h}_k$ as the vector containing the channel vectors from all N_T transmit antennas to the k^{th} receive antenna, i.e., $\boldsymbol{h}_k = [\boldsymbol{h}_{k,1}^{\mathcal{T}}\ \boldsymbol{h}_{k,2}^{\mathcal{T}}\ \cdots\ \boldsymbol{h}_{k,N_T}^{\mathcal{T}}]^{\mathcal{T}}$. Relying on (8), (10) can be written in a different form as

$$\boldsymbol{y}(t) = \boldsymbol{\mathcal{S}}(t)\boldsymbol{h} + \boldsymbol{n}(t),\ t = 1, \cdots, T, \tag{15}$$

where

$$\boldsymbol{\mathcal{S}}(t) = \boldsymbol{I}_{N_R} \otimes [\boldsymbol{\mathcal{D}}_1(t)\boldsymbol{F}\ \boldsymbol{\mathcal{D}}_2(t)\boldsymbol{F}\ \cdots\ \boldsymbol{\mathcal{D}}_{N_T}(t)\boldsymbol{F}], \tag{16}$$

$$\boldsymbol{h} = [\boldsymbol{h}_1^{\mathcal{T}}\ \boldsymbol{h}_2^{\mathcal{T}}\ \cdots\ \boldsymbol{h}_{N_R}^{\mathcal{T}}]^{\mathcal{T}}, \tag{17}$$

$$\boldsymbol{\mathcal{D}}_l(t) = \mathrm{diag}\{\boldsymbol{W s}_l(t)\},\ l = 1, 2, \ldots, N_T, \tag{18}$$

$$\boldsymbol{n}(t) = [\boldsymbol{n}_1^{\mathcal{T}}(t)\ \boldsymbol{n}_2^{\mathcal{T}}(t)\ \cdots\ \boldsymbol{n}_{N_R}^{\mathcal{T}}(t)]^{\mathcal{T}}, \tag{19}$$

and $\boldsymbol{n}(t)$ is a realization of a zero-mean complex Gaussian random vector with covariance matrix $\sigma_n^2 \boldsymbol{I}_{NN_R}$.

The model of (10) will be used in the BI-GDFE algorithm which is reviewed in Section II-C and the model of (15) will assist us in the EM-based channel estimation which will be proposed in Section III.

Note that if the number of repetition in IFDMA systems equals to 1, i.e., $R = 1$, (10) and (15) become the model of a single user MIMO-SCCP. The only difference lies in the matrix $\boldsymbol{F}$. In the MIMO-SCCP, the matrix $\boldsymbol{F}$ is the first L columns of $\sqrt{N}\boldsymbol{W}_N$. Hence, what we propose in this paper can also be applied to the single user MIMO-SCCP system.

C. BI-GDFE receiver

For simplicity, the signal model (10) is written as

$$\boldsymbol{y} = \boldsymbol{H s} + \boldsymbol{n}, \tag{20}$$

where $\boldsymbol{y}$, $\boldsymbol{s}$ and $\boldsymbol{n}$ are the received signal vector, transmitted signal vector and noise vector, respectively. $\boldsymbol{H}$ is the channel matrix. In (20), we assume that $\boldsymbol{y}$, $\boldsymbol{s}$ and $\boldsymbol{n}$ are all vectors of size $N \times 1$; $\boldsymbol{H}$ is a matrix of size $N \times N$. We further assume that $E\{\boldsymbol{s}\boldsymbol{s}^{\mathcal{H}}\} = \sigma_s^2 \boldsymbol{I}_N$, $E\{\boldsymbol{n}\boldsymbol{n}^{\mathcal{H}}\} = \sigma_n^2 \boldsymbol{I}_N$.

For model (20), linear equalizers, such as the zero-forcing or MMSE equalizers can be applied to detect the transmitted signal vector $\boldsymbol{s}$. However, in spite of the easy implementation, linear equalizers are single user detectors in the sense that when detecting one symbol, the others are treated as interference.

The block diagram of the BI-GDFE receiver is presented in the BI-GDFE box in Fig. 2. The BI-GDFE receiver is an iterative one. It detects the symbols simultaneously (but not jointly). For each iteration, the symbols decided from the previous iteration is used to re-construct the ISI. Then, the ISI effect is canceled out from the received signal vector such that the statistical SINRs are maximized for all detected symbols.

We denote the superscript v as the v^{th} iteration of the BI-GDFE. In the v^{th} iteration, the received signal vector $\boldsymbol{y}$ is filtered by the feed-forward equalizer (FFE), $\boldsymbol{K}_v$. Meanwhile, the hard decision signal vector from the $(v-1)^{th}$ iteration, $\hat{\boldsymbol{s}}^{[v-1]}$, is also passed through the feedback equalizer (FBE), $\boldsymbol{D}_v$. The outputs of the FFE and the FBE are then combined to form the v^{th} decision variable vector $\boldsymbol{z}^{[v]}$, which is further used by the decision device to obtain the next hard decision signal vector, $\hat{\boldsymbol{s}}^{[v]}$. We have

$$\boldsymbol{z}^{[v]} = \boldsymbol{K}_v^{\mathcal{H}} \boldsymbol{y} - \boldsymbol{D}_v \hat{\boldsymbol{s}}^{[v-1]}. \tag{21}$$

The optimal design of FFE and FBE for BI-GDFE receiver to maximize the SINR at the v^{th} iteration are given by:

$$\begin{aligned} \boldsymbol{K}_v &= \sigma_s^2 \left[\sigma_s^2 (1 - (\rho^{[v-1]})^2) \boldsymbol{H}\boldsymbol{H}^{\mathcal{H}} + \sigma_n^2 \boldsymbol{I}_N \right]^{-1} \boldsymbol{H}, \\ \boldsymbol{D}_v &= \rho^{[v-1]} \big(\boldsymbol{K}_v^{\mathcal{H}} \boldsymbol{H} - \boldsymbol{A}_v \big), \end{aligned} \tag{22}$$

where we assume $E\{\hat{\boldsymbol{s}}(\hat{\boldsymbol{s}}^{[v-1]})^{\mathcal{H}}\} = \rho^{[v-1]} \boldsymbol{I}_N$, $E\{\hat{\boldsymbol{s}}^{[v-1]}(\hat{\boldsymbol{s}}^{[v-1]})^{\mathcal{H}}\} = \sigma_s^2 \boldsymbol{I}_N$ and $\boldsymbol{A}_v$ is defined as $\boldsymbol{A}_v = \mathrm{diag}\{\boldsymbol{K}_v^{\mathcal{H}} \boldsymbol{H}\}$.

The details of BI-GDFE applied to MIMO-IFDMA with perfect channel state information can be found in [15].

III. Joint Channel Estimation and Data Detection

The block diagram of the proposed joint channel estimation and data detection receiver is shown in Fig. 2. In this figure, the data detection is accomplished by using the BI-GDFE method and the channel estimation is completed by using the EM-based algorithm. After the v^{th} iteration of the BI-GDFE which uses the CSI from the $(m-1)^{th}$ iteration of the EM-based algorithm, we obtain the IDC and the hard decisions of the transmitted signal blocks for the whole frame. These information would be used by the EM-based algorithm to update the CSI. Below are the details of the m^{th} iteration of the EM-based algorithm.

We start the EM-based channel estimation algorithm by considering the following model from (15)

$$\boldsymbol{y}(t) = \boldsymbol{\mathcal{S}}(t)\boldsymbol{h} + \boldsymbol{n}(t), \ t = 1, 2, \ldots, T. \tag{23}$$

According to the EM terminology, the set of $\{\boldsymbol{y}(t)\}_{t=1}^T$ is the *incomplete data space*. The parameter to be estimated is $\boldsymbol{h}$. We define a *complete data space*, $\boldsymbol{X} = \big(\{\boldsymbol{y}(t)\}_{t=1}^T, \{\boldsymbol{\mathcal{S}}(t)\}_{t=1}^T\big)$, for the parameter we want to estimate. The probability density function of $\boldsymbol{X}$ as a function of $\boldsymbol{h}$ is

$$\begin{aligned} f(\boldsymbol{X}|\boldsymbol{h}) &= f\big(\{\boldsymbol{y}(t)\}_{t=1}^T, \{\boldsymbol{\mathcal{S}}(t)\}_{t=1}^T|\boldsymbol{h}\big) \\ &= f\big(\{\boldsymbol{y}(t)\}_{t=1}^T|\{\boldsymbol{\mathcal{S}}(t)\}_{t=1}^T, \boldsymbol{h}\big) f\big(\{\boldsymbol{\mathcal{S}}(t)\}_{t=1}^T|\boldsymbol{h}\big). \end{aligned} \tag{24}$$

Since $\boldsymbol{y}(t), \ t = 1, 2, \ldots, T$ are independent, we can write $f\big(\{\boldsymbol{y}(t)\}_{t=1}^T|\{\boldsymbol{\mathcal{S}}(t)\}_{t=1}^T, \boldsymbol{h}\big)$ as

$$\begin{aligned} f\big(\{\boldsymbol{y}(t)\}_{t=1}^T|\{\boldsymbol{\mathcal{S}}(t)\}_{t=1}^T, \boldsymbol{h}\big) &= \prod_{t=1}^T f\big(\boldsymbol{y}(t)|\boldsymbol{\mathcal{S}}(t), \boldsymbol{h}\big) \\ &= \prod_{t=1}^T \frac{1}{|\pi\sigma_n^2 \boldsymbol{I}_{NN_R}|} \exp\Big\{-\frac{1}{\sigma_n^2}\|\boldsymbol{y}(t) - \boldsymbol{\mathcal{S}}(t)\boldsymbol{h}\|^2\Big\} \\ &= \frac{1}{(\pi\sigma_n^2)^{NN_RT}} \exp\Big\{-\frac{1}{\sigma_n^2}\sum_{t=1}^T\|\boldsymbol{y}(t) - \boldsymbol{\mathcal{S}}(t)\boldsymbol{h}\|^2\Big\}. \end{aligned} \tag{25}$$

One iteration of the EM algorithm consists of two steps: the first one is the E-step and the other is the M-step.

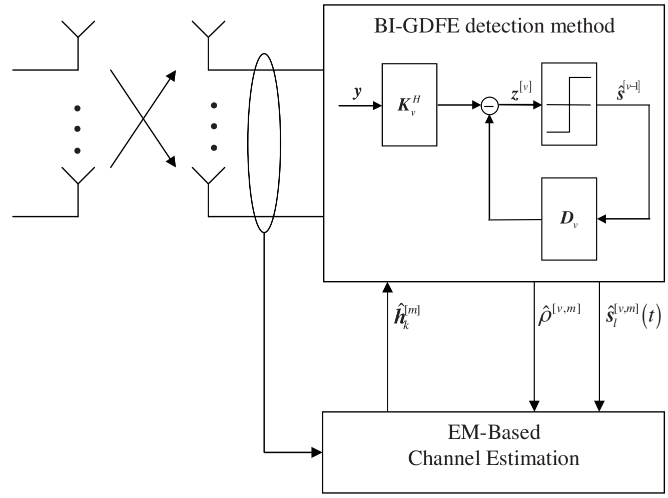

Fig. 2. The block diagram of the proposed joint channel estimation and data detection receiver.

A. E-step

In this step, we calculate

$$Q\big(\boldsymbol{h}|\hat{\boldsymbol{h}}^{[m]}\big) = E\big\{\log f\big(\boldsymbol{X}|\boldsymbol{h}\big)|\{\boldsymbol{y}(t)\}_{t=1}^T, \hat{\boldsymbol{h}}^{[m]}\big\}. \tag{26}$$

Because of the independence between the signal vectors (in the form of $\{\boldsymbol{\mathcal{S}}(t)\}_{t=1}^T$) and the channel vector $\boldsymbol{h}$, the probability density function $f\big(\{\boldsymbol{\mathcal{S}}(t)\}_{t=1}^T|\boldsymbol{h}\big)$ is not a function of $\boldsymbol{h}$. Hence, it is bypassed when we consider (26). Substituting (25) into (26), and dropping some terms that do not relate to the parameter $\boldsymbol{h}$, we have:

$$\begin{aligned} Q\big(\boldsymbol{h}|\hat{\boldsymbol{h}}^{[m]}\big) &= E\big\{\log f\big(\boldsymbol{X}|\boldsymbol{h}\big)|\{\boldsymbol{y}(t)\}_{t=1}^T, \hat{\boldsymbol{h}}^{[m]}\big\} \\ &= C_1 - E\Big\{\sum_{t=1}^T \|\boldsymbol{y}(t) - \boldsymbol{\mathcal{S}}(t)\boldsymbol{h}\|^2 \big|\{\boldsymbol{y}(t)\}_{t=1}^T, \hat{\boldsymbol{h}}^{[m]}\Big\} \\ &= C_2 + \boldsymbol{h}^{\mathcal{H}}\Big(\sum_{t=1}^T E\big\{\boldsymbol{\mathcal{S}}^{\mathcal{H}}(t)|\boldsymbol{y}(t), \hat{\boldsymbol{h}}^{[m]}\big\}\boldsymbol{y}(t)\Big) \\ &\quad + \Big(\sum_{t=1}^T \boldsymbol{y}^{\mathcal{H}}(t) E\big\{\boldsymbol{\mathcal{S}}(t)|\boldsymbol{y}(t), \hat{\boldsymbol{h}}^{[m]}\big\}\Big)\boldsymbol{h} \\ &\quad - \boldsymbol{h}^{\mathcal{H}}\Big(\sum_{t=1}^T E\big\{\boldsymbol{\mathcal{S}}^{\mathcal{H}}(t)\boldsymbol{\mathcal{S}}(t)|\boldsymbol{y}(t), \hat{\boldsymbol{h}}^{[m]}\big\}\Big)\boldsymbol{h}. \end{aligned} \tag{27}$$

where C_1 and C_2 are two constant that do not relate to the parameter $\boldsymbol{h}$.

B. M-step

M-step is to find the $\boldsymbol{h}$ that maximizes (27) and this value of $\boldsymbol{h}$ is denoted by $\hat{\boldsymbol{h}}^{[m+1]}$. We have the equation to find $\hat{\boldsymbol{h}}^{[m+1]}$ as follows:

$$\hat{\boldsymbol{h}}^{[m+1]} = \arg\max_{\boldsymbol{h}} \ Q\big(\boldsymbol{h}|\hat{\boldsymbol{h}}^{[m]}\big). \tag{28}$$

Differentiating (27) with respect to $\boldsymbol{h}$ [16], we obtain

$$\begin{aligned} \frac{\partial Q\big(\boldsymbol{h}|\hat{\boldsymbol{h}}^{[m]}\big)}{\partial \boldsymbol{h}} &= -\Big(\sum_{t=1}^T E\big\{\boldsymbol{\mathcal{S}}^{\mathcal{H}}(t)|\boldsymbol{y}(t), \hat{\boldsymbol{h}}^{[m]}\big\}\boldsymbol{y}(t)\Big)^* \\ &\quad + \Big(\Big(\sum_{t=1}^T E\big\{\boldsymbol{\mathcal{S}}^{\mathcal{H}}(t)\boldsymbol{\mathcal{S}}(t)|\boldsymbol{y}(t), \hat{\boldsymbol{h}}^{[m]}\big\}\Big)\boldsymbol{h}\Big)^*. \end{aligned} \tag{29}$$

Equating (29) to $\mathbf{0}$, we evaluate $\hat{\boldsymbol{h}}^{[m+1]}$ as follows:

$$\hat{\boldsymbol{h}}^{[m+1]} = \Big(\sum_{t=1}^{T} E\big\{\boldsymbol{\mathcal{S}}^{\mathcal{H}}(t)\boldsymbol{\mathcal{S}}(t)|\boldsymbol{y}(t),\hat{\boldsymbol{h}}^{[m]}\big\}\Big)^{-1} \times\Big(\sum_{t=1}^{T} E\big\{\boldsymbol{\mathcal{S}}^{\mathcal{H}}(t)|\boldsymbol{y}(t),\hat{\boldsymbol{h}}^{[m]}\big\}\boldsymbol{y}(t)\Big). \quad (30)$$

Equation (30) is calculated as shown below.

- Define $(\hat{\boldsymbol{\mathcal{S}}}^{[m]}(t))^{\mathcal{H}} = E\big\{\boldsymbol{\mathcal{S}}^{\mathcal{H}}(t)|\boldsymbol{y}(t),\hat{\boldsymbol{h}}^{[m]}\big\}$. From (16), we have

$$\hat{\boldsymbol{\mathcal{S}}}^{[m]}(t) = \boldsymbol{I}_{N_R}\otimes \big[\hat{\boldsymbol{\mathcal{D}}}_1^{[m]}(t)\boldsymbol{F}\ \hat{\boldsymbol{\mathcal{D}}}_2^{[m]}(t)\boldsymbol{F}\ \cdots\ \hat{\boldsymbol{\mathcal{D}}}_{N_T}^{[m]}(t)\boldsymbol{F}\big], \quad (31)$$

where $\hat{\boldsymbol{\mathcal{D}}}_l^{[m]}(t) = \mathrm{diag}\big\{\boldsymbol{W}\hat{\rho}^{[v,m]}\hat{\mathbf{s}}_l^{[v,m]}(t)\big\}$, $l = 1,2,\ldots,N_T$; $\hat{\rho}^{[v,m]}$ and $\hat{\mathbf{s}}_l^{[v,m]}(t)$, $t = 1,\cdots,T$ are the IDC and the hard decisions after v iterations of the BI-GDFE method using the channel estimate $\hat{\boldsymbol{h}}^{[m]}$, respectively. Hence, the second quantity in (30) can be determined as

$$\Big(\sum_{t=1}^{T} E\big\{\boldsymbol{\mathcal{S}}^{\mathcal{H}}(t)|\boldsymbol{y}(t),\hat{\boldsymbol{h}}^{[m]}\big\}\boldsymbol{y}(t)\Big) = \Big[(\hat{\boldsymbol{B}}_1^{[m]})^{\mathcal{T}}\ (\hat{\boldsymbol{B}}_2^{[m]})^{\mathcal{T}}\ \cdots\ (\hat{\boldsymbol{B}}_{N_R}^{[m]})^{\mathcal{T}}\Big]^{\mathcal{T}}, \quad (32)$$

where

$$\hat{\boldsymbol{B}}_k^{[m]} = \begin{bmatrix} F^{\mathcal{H}}\sum_{t=1}^{T}(\hat{\boldsymbol{\mathcal{D}}}_1^{[m]}(t))^{\mathcal{H}}\boldsymbol{y}_k(t) \\ F^{\mathcal{H}}\sum_{t=1}^{T}(\hat{\boldsymbol{\mathcal{D}}}_2^{[m]}(t))^{\mathcal{H}}\boldsymbol{y}_k(t) \\ \vdots \\ F^{\mathcal{H}}\sum_{t=1}^{T}(\hat{\boldsymbol{\mathcal{D}}}_{N_T}^{[m]}(t))^{\mathcal{H}}\boldsymbol{y}_k(t) \end{bmatrix}, \quad (33)$$

for $k = 1,2,\ldots,N_R$.

- We have

$$\boldsymbol{\mathcal{S}}^{\mathcal{H}}(t)\boldsymbol{\mathcal{S}}(t) = \boldsymbol{I}_{N_R}\otimes\begin{bmatrix}\boldsymbol{F}^{\mathcal{H}}\boldsymbol{\mathcal{D}}_1^{\mathcal{H}}(t)\\ \vdots \\ \boldsymbol{F}^{\mathcal{H}}\boldsymbol{\mathcal{D}}_{N_T}^{\mathcal{H}}(t)\end{bmatrix}[\boldsymbol{\mathcal{D}}_1(t)\boldsymbol{F}\ \cdots\ \boldsymbol{\mathcal{D}}_{N_T}(t)\boldsymbol{F}]$$
$$= \boldsymbol{I}_{N_R}\otimes \begin{bmatrix}\boldsymbol{F}^{\mathcal{H}}\boldsymbol{\mathcal{D}}_1^{\mathcal{H}}(t)\boldsymbol{\mathcal{D}}_1(t)\boldsymbol{F} & \cdots & \boldsymbol{F}^{\mathcal{H}}\boldsymbol{\mathcal{D}}_1^{\mathcal{H}}(t)\boldsymbol{\mathcal{D}}_{N_T}(t)\boldsymbol{F}\\ \vdots & \ddots & \vdots\\ \boldsymbol{F}^{\mathcal{H}}\boldsymbol{\mathcal{D}}_{N_T}^{\mathcal{H}}(t)\boldsymbol{\mathcal{D}}_1(t)\boldsymbol{F} & \cdots & \boldsymbol{F}^{\mathcal{H}}\boldsymbol{\mathcal{D}}_{N_T}^{\mathcal{H}}(t)\boldsymbol{\mathcal{D}}_{N_T}(t)\boldsymbol{F}\end{bmatrix} \quad (34)$$

Hence, the first quantity in (30) can be determined as

$$\sum_{t=1}^{T} E\big\{\boldsymbol{\mathcal{S}}^{\mathcal{H}}(t)\boldsymbol{\mathcal{S}}(t)|\boldsymbol{y}(t),\hat{\boldsymbol{h}}^{[m]}\big\} = \boldsymbol{I}_{N_R}\otimes\hat{\boldsymbol{A}}^{[m]}, \quad (35)$$

where $\hat{\boldsymbol{A}}^{[m]}$ is given in (36) on the top of next page.

Equipped with (33) and (35), we can, after some simple derivations, easily see that the updating process of the whole channel vector $\boldsymbol{h}$ is decomposed into N_R smaller updating processes in which the N_R channels from all transmit antennas to the k^{th} receive antenna are determined by

$$\hat{\boldsymbol{h}}_k^{[m+1]} = (\hat{\boldsymbol{A}}^{[m]})^{-1}\hat{\boldsymbol{B}}_k^{[m]},\ k = 1,2,\ldots,N_R. \quad (37)$$

For the EM-based channel estimation algorithm, we need to know the initial estimate of the channel, i.e., $\hat{\boldsymbol{h}}^{[0]}$. To do so, pilot symbols should be inserted in each frame to obtain a good initial channel estimation.

IV. Cramér-Rao Lower Bound (CRLB)

The CRLB is an important criterion to evaluate how good an unbiased estimator can be since the CRLB provides the MSE lower bound for all unbiased estimators. In this section, the CRLB for the channel vector $\boldsymbol{h}$ in the model of (15) is presented.

The CRLB for the channel vector $\boldsymbol{h}$ is given

$$\mathrm{CRLB}(\boldsymbol{h}) = \mathrm{tr}\{\boldsymbol{I}^{-1}(\boldsymbol{h})\}, \quad (38)$$

where, from the Appendix, we have

$$\boldsymbol{I}(\boldsymbol{h}) = \frac{1}{\sigma_n^2}(\boldsymbol{I}_{N_R}\otimes\boldsymbol{C}) \quad (39)$$

in which $\boldsymbol{C}$ is given in (40) on the top of next page.

Hence,

$$\mathrm{CRLB}(\boldsymbol{h}) = N_R\sigma_n^2\mathrm{tr}\{\boldsymbol{C}^{-1}\}. \quad (41)$$

It is clear from (41) that the CRLB changes from a frame of T blocks to another because of different signal blocks transmitted. Hence, we have the average CRLB [16], denoted by $\mathrm{aCRLB}(\boldsymbol{h})$, as follows

$$\mathrm{aCRLB}(\boldsymbol{h}) = E\left\{\mathrm{CRLB}(\boldsymbol{h})\right\}, \quad (42)$$

where the expectation is performed with respect to the transmitted signal blocks in the frame of length T. Finding $\mathrm{CRLB}(\boldsymbol{h})$ or $\mathrm{aCRLB}(\boldsymbol{h})$ is not an easy task because it is not straightforward to obtain an explicit expression for the inversion of $\boldsymbol{I}(\boldsymbol{h})$.

Another CRLB is called the modified CRLB [13], which is denoted by mCRLB. It is defined as:

$$\mathrm{mCRLB}(\boldsymbol{h}) = \sum_{p=1}^{N_RN_TL}\frac{1}{E\big\{\big(\boldsymbol{I}(\boldsymbol{h})\big)_{p,p}\big\}} = N_R\sigma_n^2\sum_{p=1}^{N_TL}\frac{1}{E\left\{(\boldsymbol{C})_{p,p}\right\}} = \frac{N_RN_TL\sigma_n^2}{TN\sigma_s^2}. \quad (43)$$

We observe that the mCRLB is inversely proportional to the number of observed blocks T and SNR. Note that the mCRLBs do not depend on the probabilistic model of channel vector $\boldsymbol{h}$; it depends on the length of $\boldsymbol{h}$. Hence, it can be applied to any multipath fading channel.

Reference [13] showed that $\mathrm{aCRLB}(\boldsymbol{h}) \geq \mathrm{mCRLB}(\boldsymbol{h})$, by using the Cauchy-Schwarz inequality. In other words, the $\mathrm{aCRLB}(\boldsymbol{h})$ is always tighter than the $\mathrm{mCRLB}(\boldsymbol{h})$. However, as shown in Section V, the $\mathrm{aCRLB}(\boldsymbol{h})$ and $\mathrm{mCRLB}(\boldsymbol{h})$ curves agree to each other, and this justifies the use of the $\mathrm{mCRLB}(\boldsymbol{h})$ as a performance measure for unbiased channel estimation algorithms in the interested system.

$$\hat{\boldsymbol{A}}^{[m]} = \begin{bmatrix} \boldsymbol{F}^{\mathcal{H}}\big(\sum_{t=1}^{T}(\hat{\boldsymbol{\mathcal{D}}}_1^{[m]}(t))^{\mathcal{H}}\hat{\boldsymbol{\mathcal{D}}}_1^{[m]}(t)\big)\boldsymbol{F} & \cdots & \boldsymbol{F}^{\mathcal{H}}\big(\sum_{t=1}^{T}(\hat{\boldsymbol{\mathcal{D}}}_1^{[m]}(t))^{\mathcal{H}}\hat{\boldsymbol{\mathcal{D}}}_{N_T}^{[m]}(t)\big)\boldsymbol{F} \\ \vdots & \ddots & \vdots \\ \boldsymbol{F}^{\mathcal{H}}\big(\sum_{t=1}^{T}(\hat{\boldsymbol{\mathcal{D}}}_{N_T}^{[m]}(t))^{\mathcal{H}}\hat{\boldsymbol{\mathcal{D}}}_1^{[m]}(t)\big)\boldsymbol{F} & \cdots & \boldsymbol{F}^{\mathcal{H}}\big(\sum_{t=1}^{T}(\hat{\boldsymbol{\mathcal{D}}}_{N_T}^{[m]}(t))^{\mathcal{H}}\hat{\boldsymbol{\mathcal{D}}}_{N_T}^{[m]}(t)\big)\boldsymbol{F} \end{bmatrix} \quad (36)$$

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{F}^{\mathcal{T}}\big(\sum_{t=1}^{T}\boldsymbol{\mathcal{D}}_1^{\mathcal{T}}(t)\boldsymbol{\mathcal{D}}_1^{*}(t)\big)\boldsymbol{F}^{*} & \cdots & \boldsymbol{F}^{\mathcal{T}}\big(\sum_{t=1}^{T}\boldsymbol{\mathcal{D}}_1^{\mathcal{T}}(t)\boldsymbol{\mathcal{D}}_{N_T}^{*}(t)\big)\boldsymbol{F}^{*} \\ \vdots & \ddots & \vdots \\ \boldsymbol{F}^{\mathcal{T}}\big(\sum_{t=1}^{T}\boldsymbol{\mathcal{D}}_{N_T}^{\mathcal{T}}(t)\boldsymbol{\mathcal{D}}_1^{*}(t)\big)\boldsymbol{F}^{*} & \cdots & \boldsymbol{F}^{\mathcal{T}}\big(\sum_{t=1}^{T}\boldsymbol{\mathcal{D}}_{N_T}^{\mathcal{T}}(t)\boldsymbol{\mathcal{D}}_{N_T}^{*}(t)\big)\boldsymbol{F}^{*} \end{bmatrix} \quad (40)$$

V. Simulation Results

Computer simulations are carried out to evaluate the performance of the proposed joint channel estimation and data detection for a multi-user MIMO-IFDMA system. We consider the case where the number of repetitions is R, i.e., the system can support up to R users. Each user is equipped with N_T transmit antennas and the receiver is equipped with N_R receive antennas. We assume that the channel between each pair of transmit and receive antennas is frequency-selective with order L and the power delay profile is uniform. Here, we make a standard assumption that the $N_T N_R$ channels (for each user) in our system are independent of each other. Furthermore, channels of all users are also independent. As stated in Section II, when $R = 1$, we have a single-user MIMO-SCCP. Therefore, this section also presents the results of our proposed receiver for MIMO-SCCP. For all systems, QPSK modulation is used. The ML bound of the interested systems cannot be presented because of the large signal size. Instead, the single user matched filter bound (SUMFB) [7] is used as the lowest bound for evaluating the performance of the proposed receiver. First of all, we compare the aCRLB and mCRLB derived in Section IV.

A. Comparison of aCRLB and mCRLB

Fig. 3 illustrates the aCRLB and mCRLB for the multi-user MIMO-IFDMA. The number of repetitions is $R = 4$, i.e., the system can support up to 4 users. Because of the orthogonality among the users at the receiver, we only present, without loss of generality, the results for the first user. We consider three cases: $N_T = N_R = 1$, $N_T = N_R = 2$ and $N_T = N_R = 3$. The channel order is $L = 17$, the block size is $N = 64$, and the number of blocks per frame is $T = 10$. We observe from the figure that the aCRLB and the mCRLB agree with each other very well when the system has a small number of transmit/receive antennas. When the number of transmit/receive antennas becomes larger, they only have a very small difference. This figure justifies our usage of mCRLB in the evaluation of the MSE performance of $\boldsymbol{h}$.

B. Multi-user MIMO-IFDMA

We consider in this section a multi-user MIMO-IFDMA system with $N_T = N_R = 2$ and $R = 4$. The channel order is still $L = 17$. We assume that the channel is unchanged over a frame of $T = 10$ signal blocks; each block is of size $N = 64$. The first signal block is devoted to pilot symbols

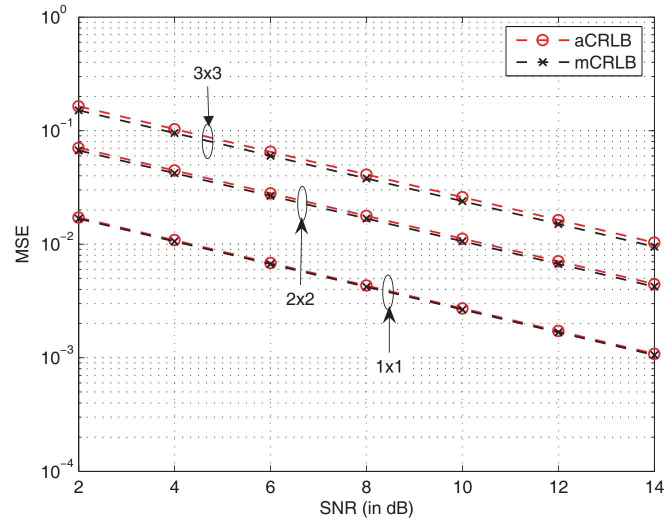

Fig. 3. Comparison of the average CRLB with the modified CRLB.

and the remaining blocks are left for data symbols. As in the previous subsection, we only present the results for the first user. Fig. 4 provides the BER performance when the number of iterations of the BI-GDFE method is 2 and 6 while the number of iterations of EM-based algorithm is 3. We can see that the BER decreases greatly after the first iteration of EM-based algorithm. However, the BER gain is marginal after the second iteration. The gap between the proposed receiver (using 6 iterations of BI-GDFE method and 3 iterations of EM-based algorithm and the BI-GDFE using perfect CSI is around 0.1dB. We also observe that the BER improvement is more significant when the number of iterations of the BI-GDFE method is larger.

In order to see the relationship between the BER performance and the number of iterations of the BI-GDFE method/ EM-based algorithm, we examine Fig. 5 in which the BER is plotted as a function of the number of iterations of the BI-GDFE method/ EM-based algorithm. In the left hand side sub-figure, the BER as a function of the number of iterations of BI-GDFE method is provided, assuming 3 iterations of the EM-based algorithm. We can see that the BER does not change if we use 5 or 6 iterations of BI-GEFE method. Comparing different SNR values, BER drops faster for the larger SNR. Moving to the right hand side where BER is plotted as a function of the number of iterations of the EM-based algorithm, we can see that after 2 or 3 iterations of the EM-based algorithm, the BER is stable. In this sub-

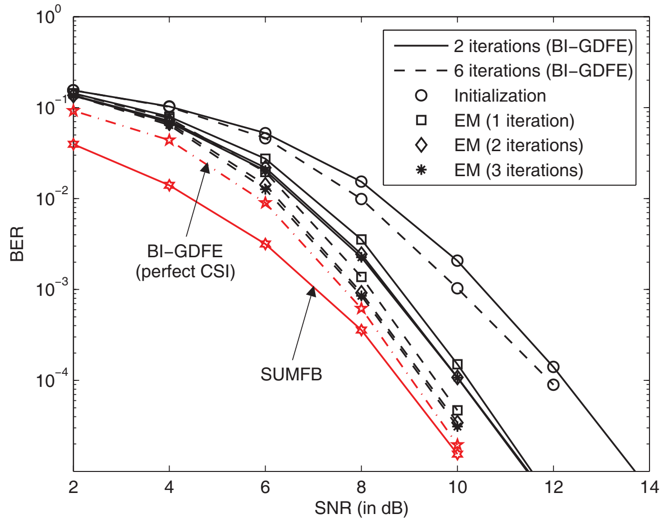

Fig. 4. BER v.s. SNR for different iterations of BI-GDFE and EM for MIMO-IFDMA.

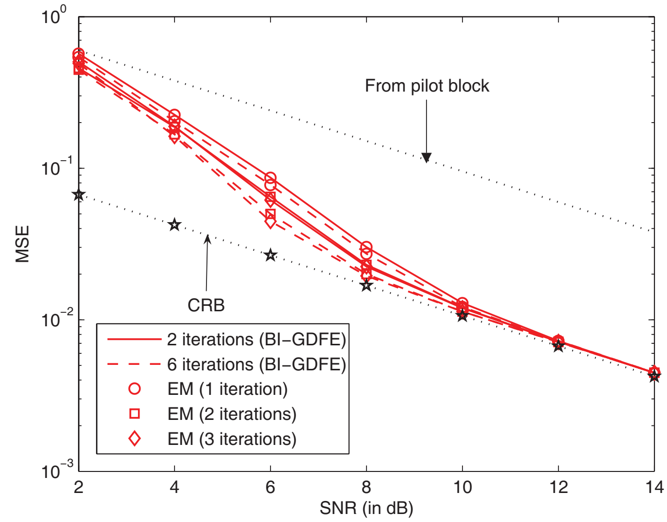

Fig. 6. The MSE performance of MIMO-IFDMA.

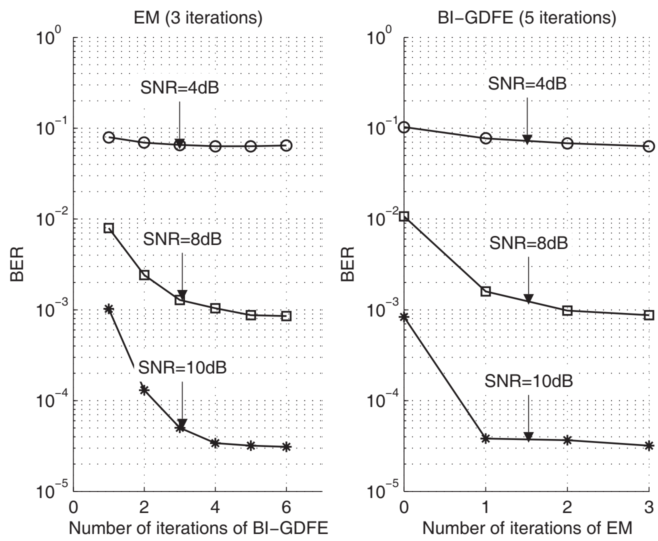

Fig. 5. The left-hand side figure: BER v.s. number of iterations of BI-GDFE while the number of EM-based algorithm is fixed at 3. The right-hand side figure: BER v.s. number of iterations of EM while the number of iterations of BI-GDFE is fixed at 5. Both are for MIMO-IFDMA.

figure, the results corresponding to 0 iterations of the EM-based algorithm is the performance of the BI-GDFE using the CSI obtained from the pilot block. These coincide with the initialization curve in Fig. 4. A great performance gain can be observed with the use of the EM-based channel estimation for larger SNR.

To examine the MSE performance of the channel estimation, Fig. 6 provides the MSE of $\boldsymbol{h}$ for different scenarios. We can see that the MSE given only by the pilot block is far away from the CRLB bound. The MSE decreases dramatically with the use of the proposed channel estimation. We do not see a big gap in the MSE performance when we use different number of iterations of the BI-GDFE as compared with the BER performance. The proposed channel estimation algorithm provides a performance very close the CRLB and the gap becomes smaller with an increase of SNR.

C. *Single-user MIMO-SCCP*

In this subsection we consider a single-user MIMO-SCCP system in which the user has $N_T = 2$ transmit antennas and the receiver is equipped with $N_R = 2$ receive antennas. We assume the channel order between each pair of transmit/receive antennas is $L = 15$. The channel is static over a frame of $T = 10$ signal blocks in which the block size is $N = 64$. Fig. 7 provides the BER when the number of iterations of BI-GDFE is 2 and 6 while the number of iterations of the EM-based algorithm is 3. We can see that the BER decreases dramatically after the first iteration of the EM-based algorithm. However, the BER gap is small after the second iteration. The gap between our proposed receiver with 6 iterations of BI-GDFE and 3 iterations of the EM-based algorithm and the BI-GDFE with perfect CSI is only around 0.1dB. We also see that the larger the number of iterations of BI-GDFE, the better the BER improvement.

Fig. 8 presents the BER performance as a function of the number of iterations of BI-GDFE method/ EM-based algorithm. This figure suggests that if we use more than 6 iterations of BI-GDFE and 3 iterations of the EM-based algorithm, the gain we obtain is marginal.

Fig. 9 illustrates the MSE performance. We can again observe that the performance provided by our proposed algorithm is much better than that given by the pilot block only.

VI. Conclusions

In this paper, a joint channel estimation and data detection algorithm is proposed for the multi-user MIMO-IFDMA system and its special case, single-user MIMO-SCCP. The detection is based on BI-GDFE, which yields a near-ML performance with low complexity. The BI-GDFE provides not only the hard decision of transmitted signals but also the IDC coefficient, which constitutes the soft information of transmitted symbols. This soft information is used for the EM-based channel estimation algorithm. Simulation results show that the proposed joint channel estimation and data detection receiver approaches the performance of BI-GDFE with perfect CSI. To facilitate the evaluation of the channel estimation, the

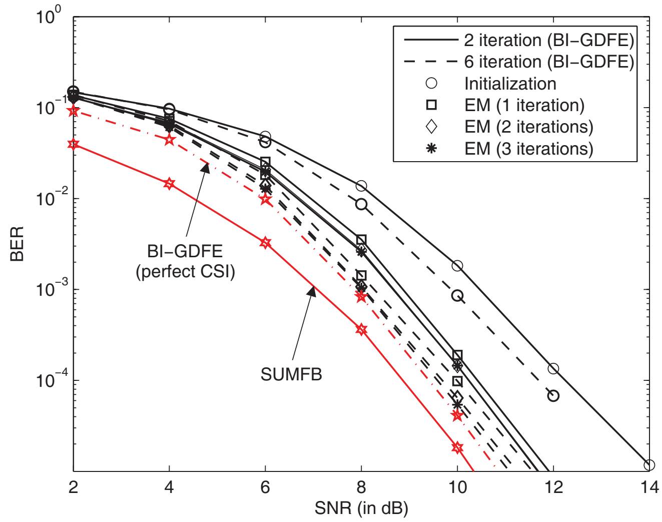

Fig. 7. BER v.s. SNR for different number of iterations of BI-GDFE and EM for MIMO-SCCP.

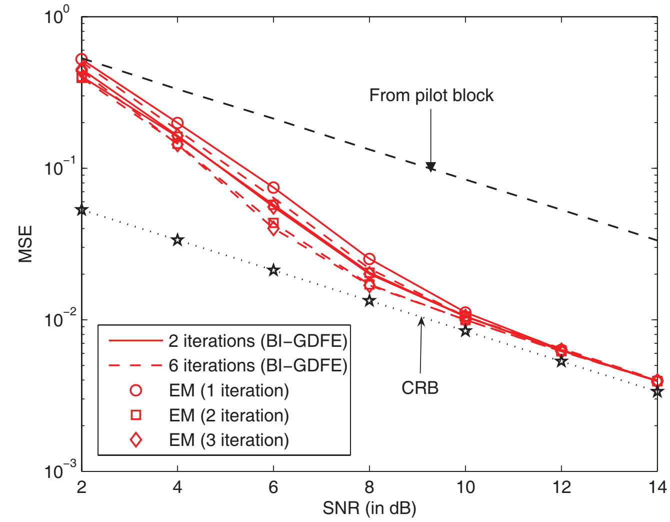

Fig. 9. MSE performance of MIMO-SCCP.

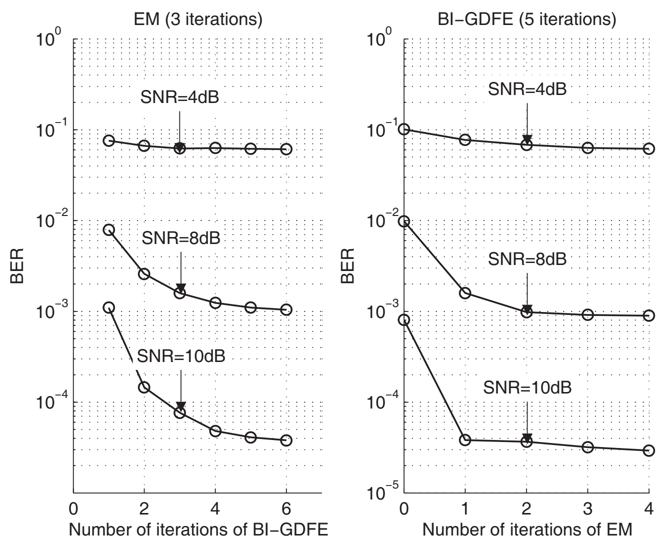

Fig. 8. The left-hand side figure: BER v.s. number of iterations of BI-GDFE while the number of EM-based algorithm is fixed at 3. The right-hand side figure: BER v.s. number of iterations of EM while the number of iterations of BI-GDFE is fixed at 5. Both are for MIMO-SCCP.

CRLB is also addressed in this paper. Due to the difficulty of deriving the exact CRLB, the modified CRLB is obtained with a closed-form solution. The modified CRLB agrees with the exact CRLB and is used to evaluate the performance of the proposed channel estimation. Simulations have shown the MSE performance obtained with our proposed algorithm is close to the theoretical CRLB.

APPENDIX

In this appendix, we derive the CRLB for the channel vector $\boldsymbol{h}$ in the model of (15) over the T blocks. Instead of separating the complex vector $\boldsymbol{h}$ into real and imaginary parts, we simplify the derivation by applying derivatives with respect to the complex vector $\boldsymbol{h}$ itself. The CRLB states the lower bound for the vector $\boldsymbol{h}$ as follows

$$\mathrm{CRLB}(\boldsymbol{h}) = \mathrm{tr}\{\boldsymbol{I}^{-1}(\boldsymbol{h})\}, \tag{A-1}$$

where $\boldsymbol{I}(\boldsymbol{h})$ is the Fisher information matrix [16]

$$\boldsymbol{I}(\boldsymbol{h}) = E\Big\{\frac{\partial}{\partial \boldsymbol{h}} \log f\big(\{\boldsymbol{y}(t)\}_{t=1}^{T}|\boldsymbol{h}\big) \Big(\frac{\partial}{\partial \boldsymbol{h}} \log f\big(\{\boldsymbol{y}(t)\}_{t=1}^{T}|\boldsymbol{h}\big)\Big)^{\mathcal{H}}\Big\}. \tag{A-2}$$

From (15), we have the conditional probability density function of $\{\boldsymbol{y}(t)\}_{t=1}^{T}$ given $\boldsymbol{h}$:

$$f\big(\{\boldsymbol{y}(t)\}_{t=1}^{T}|\boldsymbol{h}\big) = \frac{1}{(\pi\sigma_n^2)^{NN_RT}} \exp\Big\{-\frac{1}{\sigma_n^2}\sum_{t=1}^{T}\|\boldsymbol{y}(t) - \boldsymbol{\mathcal{S}}(t)\boldsymbol{h}\|^2\Big\}, \tag{A-3}$$

where we assume that the signal is known (in the form of the matrix $\boldsymbol{\mathcal{S}}(t),\ t = 1, 2, \ldots, T$). Therefore, the probability density function is not conditioned on $\boldsymbol{\mathcal{S}}(t),\ t = 1, 2, \ldots, T$. After differentiating the logarithm of (A-3) with respect to $\boldsymbol{h}$, we obtain

$$\begin{aligned}\frac{\partial}{\partial \boldsymbol{h}} \log f\big(\{\boldsymbol{y}(t)\}_{t=1}^{T}|\boldsymbol{h}\big) &= \frac{\partial}{\partial \boldsymbol{h}}\Big(-\frac{1}{\sigma_n^2}\sum_{t=1}^{T}\|\boldsymbol{y}(t) - \boldsymbol{\mathcal{S}}(t)\boldsymbol{h}\|^2\Big)\\ &= \frac{1}{\sigma_n^2}\Big(\sum_{t=1}^{T}\boldsymbol{\mathcal{S}}^{\mathcal{H}}(t)\boldsymbol{y}(t)\Big)^* - \frac{1}{\sigma_n^2}\Big(\sum_{t=1}^{T}\boldsymbol{\mathcal{S}}^{\mathcal{H}}(t)\boldsymbol{\mathcal{S}}(t)\boldsymbol{h}\Big)^*\\ &= \frac{1}{\sigma_n^2}\sum_{t=1}^{T}\Big[\big(\boldsymbol{y}^{\mathcal{H}}(t) - \boldsymbol{h}^{\mathcal{H}}\boldsymbol{\mathcal{S}}^{\mathcal{H}}(t)\big)\boldsymbol{\mathcal{S}}(t)\Big]^{\mathcal{T}}.\end{aligned} \tag{A-4}$$

Then the Fisher information matrix in (A-2) can be obtained as follows

$$\begin{aligned}\boldsymbol{I}(\boldsymbol{h}) &= E\Big\{\frac{\partial}{\partial \boldsymbol{h}} \log f\big(\{\boldsymbol{y}(t)\}_{t=1}^{T}|\boldsymbol{h}\big) \Big(\frac{\partial}{\partial \boldsymbol{h}} \log f\big(\{\boldsymbol{y}(t)\}_{t=1}^{T}|\boldsymbol{h}\big)\Big)^{\mathcal{H}}\Big\}\\ &= E\Big\{\frac{1}{\sigma_n^2}\sum_{t=1}^{T}\Big[\big(\boldsymbol{y}^{\mathcal{H}}(t) - \boldsymbol{h}^{\mathcal{H}}\boldsymbol{\mathcal{S}}^{\mathcal{H}}(t)\big)\boldsymbol{\mathcal{S}}(t)\Big]^{\mathcal{T}} \times\\ &\qquad \frac{1}{\sigma_n^2}\sum_{t=1}^{T}\Big[\big(\boldsymbol{y}^{\mathcal{H}}(t) - \boldsymbol{h}^{\mathcal{H}}\boldsymbol{\mathcal{S}}^{\mathcal{H}}(t)\big)\boldsymbol{\mathcal{S}}(t)\Big]^*\Big\}\end{aligned}$$

$$= \frac{1}{(\sigma_n^2)^2} E\left\{ \sum_{t=1}^{T} \boldsymbol{\mathcal{S}}^{\mathcal{T}}(t)\big(\boldsymbol{y}^{\mathcal{H}}(t) - \boldsymbol{h}^{\mathcal{H}}\boldsymbol{\mathcal{S}}^{\mathcal{H}}(t)\big) \big(\boldsymbol{y}^{\mathcal{H}}(t) - \boldsymbol{h}^{\mathcal{H}}\boldsymbol{\mathcal{S}}^{\mathcal{H}}(t)\big)^{*}\boldsymbol{\mathcal{S}}^{*}(t)\right\}$$

$$= \frac{1}{\sigma_n^2}\sum_{t=1}^{T} \boldsymbol{\mathcal{S}}^{\mathcal{T}}(t)\boldsymbol{\mathcal{S}}^{*}(t). \qquad \text{(A-5)}$$

Replacing (16) into (A-5), we have (39).

Acknowledgement

The authors would like to thank the anonymous reviewers for their critical comments that greatly improved this paper.

References

[1] U. Sorger, I. D. Broeck, and M. Schnell, "Interleaved FDMA—new spread-spectrum multiple-access scheme," in *Proc. IEEE ICC'1998*, vol. 2, June 1998, pp. 1013-1017.

[2] R. Dinis, D. Falconer, C. T. Lam, and M. Sabbaghian, "A multiple access scheme for the uplink of broadband wireless systems," in *Proc. IEEE GLOBECOM'2004*, vol. 6, Nov. 2004, pp. 3808-3812.

[3] H. Ekstrom, A. Furuskar, J. Karlsson, M. Meyer, S. Parkvall, J. Torsner, and M. Wahlqvist, "Technical solutions for 3G long-term evolution," *IEEE Commun. Mag.*, vol. 42, no. 3, pp. 38-45, Mar. 2006.

[4] Y. Ofuji, K. Higuch, and M. Sawahashi, "Frequency domain channel-dependent scheduling employing an adaptive transmission bandwidth for pilot channel in uplink single-carrier-FDMA radio access," in *Proc. IEEE VTC'2006-Spring*, May 2006, pp. 334-338.

[5] E. Telatar, "Capacity of multi-antenna gaussian channels," *AT&T Bell Labs Intern. Tech. Memo.*, July 1995.

[6] G. J. Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41-59, 1996.

[7] E. A. Lee and D. G. Messerschmitt, *Digital Communication*, 2nd ed. Kluwer Academic Publishing, 1993.

[8] G. Ginis and J. M. Cioffi, "On the relation between V-BLAST and the GDFE," *IEEE Commun. Lett.*, vol. 5, pp. 364-366, Sept. 2001.

[9] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 1046-1061, July 1999.

[10] Y.-C. Liang, S. Sun, and C. K. Ho, "Block-iterative generalized decision feedback equalizers for large MIMO systems: algorithm design and asymptotic performance analysis," *IEEE Trans. Signal Processing*, vol. 54, no. 6, pp. 2035-2048, June 2006.

[11] D. Falconer, S. L. Ariyavisitakul, A. Benyamin-Seeyar, and B. Edison, "Frequency domain equalization for single-carrier broadband wireless systems," *IEEE Commun. Mag.*, vol. 40, no. 4, pp. 58-66, Apr. 2002.

[12] J. Coon, J. Siew, M. Beach, A. Nix, S. Armour, and J. McGreehan, "A comparison of MIMO-OFDM and MIMO-CSFDE in WLAN environments," in *Proc. IEEE GLOBECOM'2003*, vol. 6, Dec. 2003, pp. 3296-3301.

[13] A. N. D'Andrea, U. Meganli, and R. Reggiannini, "The modified Cramér-Rao bound and its application to synchronization problems," *IEEE Trans. Commun.*, vol. 42, no. 234, pp. 1391-1399, 1994.

[14] M. Schnell and I. D. Broeck, "Application of IFDMA to mobile radio transmission," in *Proc. IEEE ICUPC'1998*, Oct. 1998, pp. 1267-1272.

[15] Y. Pei and Y.-C. Liang, "Subcarrier-based block-iterative GDFE (BI-GDFE) receivers for MIMO interleaved FDMA," in *Proc. IEEE VTC'2007-Spring*, Apr. 2007, pp. 2033-2037.

[16] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*. Englewood Cliffs, NJ: Prentice Hall, 1993.

The-Hanh Pham received the B.Eng. degree in Electronics and Telecommunications from Hanoi University of Technology, Hanoi, Vietnam, in 2001, and is currently working toward the Ph.D. degree at the Department of Electrical Engineering, National University of Singapore. His research interests are in signal processing for communications, MIMO systems.

Ying-Chang Liang (SM'00) is currently a Senior Scientist with the Institute for Infocomm Research (I2R), A*STAR, Singapore, where he has been leading the research activities in the area of cognitive radio and cooperative communications. He also holds adjunct associate professorship positions in Nanyang Technological University (NTU) and National University of Singapore (NUS), both in Singapore, and adjunct professorship position with University of Electronic Science & Technology of China (UESTC). He has been teaching graduate courses in NUS since 2004. From Dec 2002 to Dec 2003, Dr Liang was a visiting scholar with the Department of Electrical Engineering, Stanford University, CA, USA. His research interest includes cognitive radio, dynamic spectrum access, reconfigurable signal processing for broadband communications, space-time wireless communications, wireless networking, information theory and statistical signal processing.

Dr Liang is now an Associate Editor of IEEE Transactions on Vehicular Technology and Lead Guest-Editor of IEEE Journal on Selected Areas in Communications, Special Issue on Cognitive Radio Networking and Communications. He was an Associate Editor of IEEE Transactions on Wireless Communications from 2002 to 2005, Lead Guest-Editor of IEEE Journal on Selected Areas in Communications, Special Issue on Cognitive Radio: Theory and Applications, and Guest-Editor of Computer Networks Journal (Elsevier) Special Issue on Cognitive Wireless Networks. He received the Best Paper Awards from IEEE 50th Vehicular Technology Conference in 1999, the IEEE 16th International Symposium on Personal, Indoor, and Mobile Radio Communications, in 2005, and the 2007 Institute of Engineers Singapore (IES) Prestigious Engineering Achievement Award. Dr Liang has served as technical committee member and/or organizing committee member for various IEEE conferences.

Arumugam Nallanathan (S'97-M'00-SM'05) received the B.Sc. with honors from the University of Peradeniya, Sri-Lanka, in 1991, the CPGS from the Cambridge University, United Kingdom, in 1994 and the Ph.D. from the University of Hong Kong, Hong Kong, in 2000, all in Electrical Engineering. He was an Assistant Professor in the Department of Electrical and Computer Engineering, National University of Singapore, Singapore from August 2000 to December 2007. Currently, he is a Senior Lecturer in the Department of Electronic Engineering at King's College London, United Kingdom. His research interests include MIMO-OFDM systems, ultra-wide bandwidth (UWB) communication and localization, cooperative diversity techniques and cognitive radio. In these areas, he has published over 150 journal and conference papers. He is a co-recipient of the Best Paper Award presented at 2007 IEEE International Conference on Ultra-Wideband (ICUWB'2007).

He is now an editor for the IEEE Transactions on Wireless Communications and an associate editor for the IEEE Transactions on Vehicular Technology. He was a Guest Editor for EURASIP Journal of Wireless Communications and Networking: Special issue on UWB Communication Systems-Technology and Applications. He currently serves as the Secretary for the Signal Processing and Consumer Electronics Technical Committee of IEEE Communications Society. He was a technical program committee member for more than 50 IEEE international conferences. He also served as the General Track Chair for IEEE VTC'2008-Spring, Co-Chair for the IEEE GLOBECOM'2008 Signal Processing for Communications Symposium. He currently serves as the Co-Chair for the IEEE ICC'2009 Wireless Communications Symposium.