Factor Graph EM Algorithm for Joint Channel Tracking and MAP Detection of MIMO-OFDMA in Fading Channels

Sau-Hsuan Wu
Department of Communication Engineering
National Chiao Tung University, Hsinchu, Taiwan
E-mail : sauhsuan@cm.nctu.edu.tw

Abstract— **Based on the expectation maximization (EM) algorithm, a joint channel tracking and maximal a posteriori symbol detection method is proposed for orthogonal frequency division multiple access systems employing multiple transmit and receive antennas (MIMO). Using the Forney-style factor graph (FFG), a forward-and-backward signal processing method is developed to recursively solve the EM problem for joint estimation and detection (JED) in MIMO fading channels. By proper message scheduling and passing over the FFG, simulations show that the proposed algorithm is effective for joint MIMO channel tracking and symbol detection in fast fading channels.**

Keywords— **Factor Graph, MIMO, OFDM, OFDMA, EM, Joint Estimation and Detection.**

I. Introduction

Orthogonal frequency division multiplexing/multiple access (OFDM/OFDMA) using multiple input and multiple output (MIMO) transceiving antennas is considered promising technology for next-generation broadband mobile wireless access (BMWA) due to its flexible architecture for multiple access and outstanding ability to combat severe signal degradation in BMWA channels. To fully exploit the advantage of OFDM for BMWA, joint channel estimation and detection (JED) is considered an effective method to perform blind/semi-blind data detection in fast fading channels. However, the complexity of JED for all channel coefficients of a MIMO-OFDM system is very high for fast fading channels. To reduce the complexity, there has been some research results available in the literature, among which [1, 2], *e.g.*, perform JED with hard symbol decision feedbacks while [3,4] use soft information feedbacks according to the expectation-maximization (EM) principle. It has been shown in [4, 5], among others, that the EM-based JED can outperform algorithms using hard decision feedbacks.

Despite the complexity of JED for OFDM, a direct application of the JED of OFDM to OFDMA may not be feasible and will suffer from severe performance degradation due to the multiple access interference (MAI) present in the time-domain OFDMA received signals. In addition, the MAI will aggravate when multiple transmit and receive antennas are used in the OFDMA system. Exploiting the signal architecture of OFDMA, in this work, we propose a hybrid JED scheme that can perform MIMO channel tracking for OFDMA without the influence of MAI. Moreover, to reduced the algorithm complexity, a forward-and-backward signal processing procedure is developed to recursively solve the EM problem for JED, using the max-sum algorithm on the Forney-style factor graph (FFG). By proper message scheduling over the FFGs, the proposed algorithm presents promising results of JED for MIMO-OFDMA in fast fading channels, using only one out 256 OFDM symbols for pilots.

This research has been funded in part by ZyXEL Cooperation, Taiwan, under the grant 95C169 and in part by the National Science Council, Taiwan, under the grant NSC 95-2219-E-009-011 and NSC 95-2219-E-009-016.

II. System Model

We consider a MIMO OFDMA system employing N_t transmit antennas and N_r receive antennas. There are a total of K users sharing the available subcarriers of the system. The number of the overall subcarriers is denoted by N. Modulated symbols of a user are first multiplexed into N_t transmission streams, each corresponding to a transmit antenna. Then, each sub-stream is mapped to the designated subcarriers and transformed with the inverse Fourier transformation (IFFT). The outputs of IFFT are cyclic-prefixed and transmitted through the corresponding antenna over a time-varying multiple-path fading channel. The discrete-time channel response between the i-th receive and j-th transmit antenna pair is denoted by $\mathbf{h}_{i,j}(t) = \{h^1_{i,j}(t), \ldots, h^L_{i,j}(t)\}$. The number of channel taps, L, for each user is considered smaller than or equal to the length of the cyclic-prefix, L_{cp}, which is a system parameter designed to combat inter-symbol interference. Throughout this paper, we assume perfect synchronization is reached at the receiver and the channel coefficients remain unchanged within each OFDM block of length $N + L_{cp}$ and fade from block to block according to the Jake's model for Rayleigh fading channels. As a result, there is no inter-carrier interference (ICI) in the system.

The received time-domain signal at the i-th receive antenna for the m-th OFDM symbol interval is denoted by $\mathbf{y}_{i,m} = \{y_{i,(m-1)(N+L_{cp})+L_{cp}}, \cdots, y_{i,m(N+L_{cp})-1}\}$, after removing the cyclic prefix. Transforming $\mathbf{y}_{i,m}$ with FFT, the frequency-domain received signal $Y^n_{i,m}$ at the n-th frequency tone, $n = 0, \ldots, N-1$, is given by

$$Y^n_{i,m} = \underline{X}^n_m \underline{H}^n_{i,m} + N^n_{i,m} \tag{1}$$

where $\underline{X}^n_m \triangleq [X^n_{1,m}, \ldots, X^n_{N_t,m}]$ is the frequency-domain transmitted symbol vector, and $\underline{H}^n_{i,m} \triangleq [H^n_{i,1,m}, \cdots, H^n_{i,N_t,m}]^T$ in which $H^n_{i,j,m}$ is the frequency-domain channel response between the ith receive antenna and the jth transmit antenna. In addition, the noise vector $N_{i,m}$ is considered zero-mean and additive-white complex Gaussian (AWGN) distributed, denoted by $\sim \mathcal{CN}(\mathbf{0}, \sigma^2_n)$. We note that in this paper a bold-faced lower-case character $\mathbf{x}$ represents a vector in the time-domain, while an underlined upper-case $\underline{X}$ stands for

a vector in the frequency domain. In addition, a bold-faced $\mathbf{X}$ denotes a matrix either in the time or frequency domain.

In an OFDMA system, a single user is unlikely to occupy the entire bandwidth of the system. Therefore, without loss of generality, we assume the set of frequency bins allocated to the kth user is defined as $\{f_1^k, \ldots, f_{N_k}^k\}$, where $1 \leq f_n^k \leq N$ and N_k being the number of tones employed by user k. Collecting the frequency-domain received signals with respect to (*w.r.t.*) user k, we have $\underline{Y}_m^k \triangleq [Y_{1,m}(f_1^k), \ldots, Y_{N_r,m}(f_1^k), \cdots, Y_{1,m}(f_{N_k}^k), \ldots, Y_{N_r,m}(f_{N_k}^k)]^T$, which, under AWGN, can be modeled as

$$\underline{Y}_m^k = \mathbb{X}_m^k \underline{\mathcal{H}}_m^k + \underline{N}_m^k \tag{2}$$

with $\underline{\mathcal{H}}_m^k \triangleq [\underline{H}_{1,m}^{f_1^k}, \ldots, \underline{H}_{N_r,m}^{f_1^k}, \cdots, \underline{H}_{1,m}^{f_{N_k}^k}, \ldots, \underline{H}_{N_r,m}^{f_{N_k}^k}]^T$, $\mathbb{X}_m^k \triangleq \text{diag}\{\mathbf{I}_{N_r} \otimes \underline{X}_m^{f_1^k}, \ldots, \mathbf{I}_{N_r} \otimes \underline{X}_m^{f_{N_k}^k}\}$, $\otimes$ being the Kronecker product, and $\underline{N}_m^k \sim \mathcal{CN}(0, \sigma_n^2 \mathbf{I}_{N_r N_k})$. Since we assume no ICI in the system, users of the system operate in orthogonal frequency channels. The derivations corresponding to a certain user also apply to other users in the system. Therefore, for convenience of expression, we drop the user index, k, in the sequel.

A. Time-Domain Channel Dymanics

To track the time-varying channel coefficients, we model the time-domain Rayleigh multipath fading coefficients as a first-order auto-regressive (AR) random process. Let $h_{i,j,m}^\ell$ be the ℓ-th path time-domain channel coefficient between the i-th receive antenna and the j-th transmit antenna at time m. We define for the ℓ-th path of the multipath MIMO channel a time-domain channel vector $\mathbf{h}_m^\ell \triangleq [h_{1,1,m}^\ell, \ldots, h_{1,N_t,m}^\ell, \ldots, h_{N_r,1,m}^\ell, \ldots, h_{N_r,N_t,m}^\ell]^T$. Based on the AR model, the multipath MIMO channel coefficients, $\underline{\mathbf{h}}_m \triangleq [(\mathbf{h}_m^1)^T, \ldots, (\mathbf{h}_m^L)^T]^T$, can be expressed as

$$\underline{\mathbf{h}}_m = \mathbb{F}\underline{\mathbf{h}}_{m-1} + \mathbb{B}\mathbf{v}_m \tag{3}$$

where for wide-sense stationary uncorrelated scattering (WSSUS) channels, we have $\mathbb{F} = \text{diag}\{\mathbf{F}_1, \ldots, \mathbf{F}_L\}$ and $\mathbb{B} = \text{diag}\{\mathbf{B}_1, \ldots, \mathbf{B}_L\}$, each of which being a block diagonal matrix of dimension $N_r N_t L \times N_r N_t L$. Furthermore, we also assume that the MIMO channel coefficients are spatially uncorrelated. As a result, $\mathbf{F}_\ell = \alpha \mathbf{I}_{N_r N_t}$ and $\mathbf{B}_\ell = \sigma_\ell^2(1-\alpha^2)\mathbf{I}_{N_r N_t}$, $\ell = 1, \ldots, L$, with $\alpha = J_0(2\pi f_d T_s)$ and $\sigma_\ell^2 = E\{\|h_{i,j,m}^\ell\|^2\}$. The function $J_0(\cdot)$ is the zeroth-order Bessel function of the first kind. And $f_d T_s$ is the Doppler frequency shift, f_d, normalized by the sampling time T_s. The noise vector, $\mathbf{v}_m \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{N_r N_t L})$.

Since our goal is to recover $\mathbb{X}_m$ in (2), we should establish the relationship between $\underline{\mathcal{H}}_m$ and $\underline{\mathbf{h}}_m$. Let $\mathbf{W}$ denote the FFT matrix of dimension $N \times N$. The frequency-domain channel of $\mathbf{h}_{i,j,m} \triangleq [\mathbf{h}_{i,j,m}^1, \ldots, \mathbf{h}_{i,j,m}^L]^T$ is equal to $\mathbf{W}(:, 1:L)\mathbf{h}_{i,j,m}$. According to the channel allocation made for user k, the frequency-domain channel vector, $\underline{\mathcal{H}}_{i,j,m}$, corresponding to the i-th receive and j-th transmit antenna pair of user k is given by

$$\underline{\mathcal{H}}_{i,j,m} = \mathbf{S}_k^T \mathbf{W}(:, 1:L)\mathbf{h}_{i,j,m} \triangleq \mathbf{W}_k \mathbf{h}_{i,j,m} \tag{4}$$

where $\mathbf{S}_k \triangleq [\mathbf{e}_{f_1^k}, \cdots, \mathbf{e}_{f_{N_k}^k}]$, and $\mathbf{e}_n$ is a $N \times 1$ vector with a '1' at the n-th entry and all the other entries 0's. The matrix $\mathbf{W}_k$ is of dimension $N_k \times L$, downsampled from $\mathbf{W}$. Given the time-frequency relationship in (4), it can be shown that

$$\underline{\mathcal{H}}_m = (\mathbf{W}_k \otimes \mathbf{I}_{N_r N_t})\underline{\mathbf{h}}_m. \tag{5}$$

Substituting this expression back into (2) leads to

$$\begin{aligned} \underline{Y}_m &= \mathbb{X}_m(\mathbf{W}_k \otimes \mathbf{I}_{N_r N_t})\underline{\mathbf{h}}_m + \underline{N}_m \\ &\triangleq \mathbb{X}_m \mathbb{W}_k \underline{\mathbf{h}}_m + \underline{N}_m. \end{aligned} \tag{6}$$

III. Joint Channel Tracking and Symbol Detection over MIMO fading Channels

We now present joint channel estimation and symbol detection (JED) for $\underline{\mathbf{h}}_m$ and $\mathbb{X}_m$ of each user according to the channel dynamic model (3) and the system model (6).

Under the EM framework in [6], we define for user k the observation up to time M as $\underline{Y}_1^M \triangleq \{\underline{Y}_1, \cdots, \underline{Y}_M\}$ and the unknown parameter set as $\underline{\mathbf{h}}_1^M \triangleq \{\underline{\mathbf{h}}_1, \cdots \underline{\mathbf{h}}_M\}$. The hidden state of the system is $\underline{\mathcal{X}}_1^M \triangleq \{\underline{\mathcal{X}}_1, \cdots, \underline{\mathcal{X}}_M\}$, where $\underline{\mathcal{X}}_m \triangleq \{\underline{X}_m^{f_1^k}, \ldots, \underline{X}_m^{f_{N_k}^k}\}$. It is clear that the complete data for estimating the parameter set $\underline{\mathbf{h}}_1^M$ is $\{\underline{Y}_1^M, \underline{\mathcal{X}}_1^M\}$. In the absence of $\underline{\mathcal{X}}_1^M$, the log likelihood (LLK) of $\underline{\mathbf{h}}_1^M$ is given by

$$\log P(\underline{Y}_1^M, \underline{\mathbf{h}}_1^M) = \log E_{\underline{\mathcal{X}}_1^M}\{P(\underline{Y}_1^M, \underline{\mathcal{X}}_1^M, \underline{\mathbf{h}}_1^M)\}, \tag{7}$$

where $E_{\underline{\mathcal{X}}_1^M}\{\cdot\}$ is the expectation w.r.t. $\underline{\mathcal{X}}_1^M$.

It is in general difficult to estimate $\underline{\mathbf{h}}_1^M$ directly from the LLK. To reduce the complexity, we use the EM algorithm to approximate it iteratively. To this end, we first define a Kullback-Liebler (K-L) measure of $\underline{\mathbf{h}}_1^M$ at iteration ℓ as

$$\begin{aligned} &Q_M(\underline{\mathbf{h}}_1^M | \widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1}) \triangleq \\ &\sum_{\{\mathbb{X}_1^M\}} \log\{P(\underline{Y}_1^M, \underline{\mathbf{h}}_1^M, \underline{\mathcal{X}}_1^M)\} P(\underline{\mathcal{X}}_1^M | \underline{Y}_1^M, \widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1}), \end{aligned} \tag{8}$$

where $\widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1}$ is the estimate of $\underline{\mathbf{h}}_1^M$ at iteration $\ell-1$. The EM procedure is summarized in the following two major tasks:

E-step: Compute $Q_M(\underline{\mathbf{h}}_1^M | \widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1})$;

M-step: $\widehat{\underline{\mathbf{h}}}_{1:M}^{\ell} = \arg\max_{\underline{\mathbf{h}}_1^M} Q_M(\underline{\mathbf{h}}_1^M | \widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1})$.

The LLK of $\widehat{\underline{\mathbf{h}}}_{1:M}^{\ell}$ is guaranteed non-decreasing through the EM procedure [7].

Even if the complexity of JED can be drastically reduced using the EM algorithm, the computational complexities for the corresponding E-step and M-step are still high. In the sequel, we will develop a low-complexity recursive procedure to process these two tasks based on the max-sum rule of the factor graph algorithm in [8].

According to the Markovian property presented in (3), the joint probability $P(\underline{Y}_1^M, \underline{\mathbf{h}}_1^M, \underline{\mathcal{X}}_1^M)$ can be factorized into

$$\begin{aligned} &P(\underline{Y}_1^M, \underline{\mathbf{h}}_1^M, \underline{\mathcal{X}}_1^M) = P(\underline{Y}_1^{m-1}, \underline{\mathbf{h}}_1^{m-1}, \underline{\mathcal{X}}_1^{m-1}) \\ &P(\underline{Y}_m, \underline{\mathcal{X}}_m | \underline{\mathbf{h}}_m) P(\underline{\mathbf{h}}_m | \underline{\mathbf{h}}_{m-1}) P(\underline{Y}_{m+1}^M, \underline{\mathbf{h}}_{m+1}^M, \underline{\mathcal{X}}_{m+1}^M | \underline{\mathbf{h}}_m). \end{aligned} \tag{9}$$

Given $P(\underline{\mathcal{X}}_m | \underline{Y}_1^M, \widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1})$, $m = 1, \ldots, M$, the E-step can be rewritten as

$$\begin{aligned} &Q_M(\underline{\mathbf{h}}_1^M | \widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1}) \equiv Q_M(\underline{\mathbf{h}}_1^{m-1} | \widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1}) \\ &+ E_{\underline{\mathcal{X}}_m}\left\{\|\underline{Y}_m - \mathbb{X}_m \mathbb{W}_k \underline{\mathbf{h}}_m\|^2 / \sigma_n^2 | \underline{Y}_1^M; \widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1}\right\} \\ &+ \log P(\underline{\mathbf{h}}_m | \underline{\mathbf{h}}_{m-1}) + Q_M(\underline{\mathbf{h}}_{m+1}^M | \widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1}) \end{aligned} \tag{10}$$

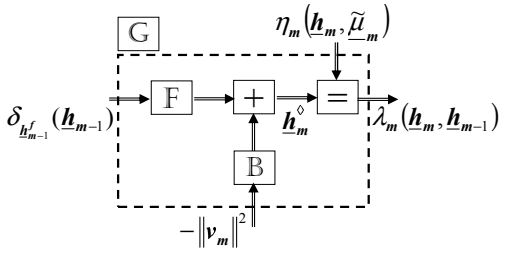

Fig. 1. FFG representation of the summation of (11) and (14), where $\delta_{\underline{\mathbf{h}}^f_{m-1}}(\underline{\mathbf{h}}_{m-1}) = -\gamma\|\underline{\mathbf{h}}^f_{m-1} - \underline{\mathbf{h}}_{m-1}\|^2$, $\gamma \to \infty$.

where $E_{\underline{\mathcal{X}}_m}\{\cdot|\underline{Y}_1^M; \widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1}\}$ is the expectation w.r.t. $\underline{\mathcal{X}}_m$ in $\mathbb{X}_m$, using $P(\underline{\mathcal{X}}_m|\underline{Y}_1^M, \widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1}) \propto P(\underline{Y}_m|\underline{\mathcal{X}}_m, \widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1})$ obtained in the frequency domain. By some manipulations, we have

$$E_{\underline{\mathcal{X}}_m}\left\{\|\underline{Y}_m - \mathbb{X}_m\mathbb{W}_k\underline{\mathbf{h}}_m\|^2/\sigma_n^2|\underline{Y}_1^M; \widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1}\right\} \equiv$$
$$-[\underline{\mathbf{h}}_m - \widetilde{\mathbb{C}}_m^{-1}\mathbb{W}_k^H\widetilde{\mathbb{X}}_m^H\underline{Y}_m]^H\widetilde{\mathbb{C}}_m[\underline{\mathbf{h}}_m - \widetilde{\mathbb{C}}_m^{-1}\mathbb{W}_k^H\widetilde{\mathbb{X}}_m^H\underline{Y}_m] \quad (11)$$

where we have defined the posterior mean

$$\widetilde{\mathbb{X}}_m \triangleq (1/\sigma_n^2)E_{\underline{\mathcal{X}}_m}\{\mathbb{X}_m|\underline{Y}_1^M; \widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1}\} \quad (12)$$

and its corresponding transformed correlation matrix

$$\widetilde{\mathbb{C}}_m \triangleq (1/\sigma_n^2)\mathbb{W}_k^H E_{\underline{\mathcal{X}}_m}\{\mathbb{X}_m^H\mathbb{X}_m|\underline{Y}_1^M; \widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1}\}\mathbb{W}_k. \quad (13)$$

We note that the dimension of $E_{\underline{\mathcal{X}}_m}\{\mathbb{X}_m^H\mathbb{X}_m|\underline{Y}_1^M; \widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1}\} \leq N_rN_k$, and the dimension of $\mathbb{W}_k$ is $N_kN_rN_t \times LN_rN_t$. For $\widetilde{\mathbb{C}}_m^{-1}$ to exist, $\min\{\dim(E_{\underline{\mathcal{X}}_m}\{\mathbb{X}_m^H\mathbb{X}_m|\underline{Y}_1^M; \widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1}\}), N_rN_k\} \geq LN_rN_t$. Therefore, we must at least have $N_k \geq LN_t$.

On the other hand, based on the channel dynamic model (3), we have

$$\log P(\underline{\mathbf{h}}_m|\underline{\mathbf{h}}_{m-1}) = -[\underline{\mathbf{h}}_m - \mathbb{F}\underline{\mathbf{h}}_{m-1}]^H\mathbb{B}^{-H}\mathbb{B}^{-1}[\underline{\mathbf{h}}_m - \mathbb{F}\underline{\mathbf{h}}_{m-1}]. \quad (14)$$

Both (11) and (14) are of Gaussian quadratic form. Therefore, the summation of these two is still a Gaussian quadratic function of $\underline{\mathbf{h}}_m$ and $\underline{\mathbf{h}}_{m-1}$. This can be more conveniently expressed and evaluated using a forney-style factor graph.

We define for (11) a Gaussian quadratic function of

$$\eta_m(\underline{\mathbf{h}}_m, \widetilde{\underline{\mu}}_m) = \quad -[\underline{\mathbf{h}}_m - \widetilde{\underline{\mu}}_m]^H\widetilde{\mathbb{C}}_m[\underline{\mathbf{h}}_m - \widetilde{\underline{\mu}}_m] \quad (15)$$

where $\widetilde{\underline{\mu}}_m = \widetilde{\mathbb{C}}_m^{-1}\mathbb{W}_k^H\widetilde{\mathbb{X}}_m^H\underline{Y}_m$. By the max-sum rule in [8], the summation of (11) and (14) is equivalent to the constrained maximization

$$\lambda_m(\underline{\mathbf{h}}_m, \underline{\mathbf{h}}_{m-1}) =$$
$$\max_{\underline{\mathbf{h}}_m^\diamond}\max_{\mathbf{v}_m}\max_{\underline{\mathbf{h}}^f_{m-1}}\Big\{\delta_{\underline{\mathbf{h}}^f_{m-1}}(\underline{\mathbf{h}}_{m-1}) - \|\mathbf{v}_m\|^2 + \eta_m(\underline{\mathbf{h}}_m, \widetilde{\underline{\mu}}_m)$$
$$+\delta_{\underline{\mathbf{h}}_m^\diamond}(\mathbb{F}\underline{\mathbf{h}}^f_{m-1} + \mathbb{B}\mathbf{v}_m) + \delta_{\underline{\mathbf{h}}_m}(\underline{\mathbf{h}}_m^\diamond)\Big\}$$
$$\equiv -[\underline{\mathbf{h}}_m - \underline{\mu}_m]^H\Sigma_m^{-1}[\underline{\mathbf{h}}_m - \underline{\mu}_m] \quad (16)$$

where $\delta_x(z) = -\gamma\|x - z\|^2$, $\gamma \to \infty$. The Forney-style factor (FFG) representation of the above constrained maximization is given in Fig. 1. The constraint imposed by the system, denoted by $\mathbb{G}$ in the figure, is given in the third line. Using the update rules of [8] for each individual node in $\mathbb{G}$, we obtain $\Sigma_m = (\mathbb{B}^{-H}\mathbb{B}^{-1} + \widetilde{\mathbb{C}}_m)^{-1}$ and $\underline{\mu}_m = \Sigma_m(\mathbb{B}^{-H}\mathbb{B}^{-1}\mathbb{F}\underline{\mathbf{h}}_{m-1} + \mathbb{W}_k^H\widetilde{\mathbb{X}}_m^H\underline{Y}_m)$.

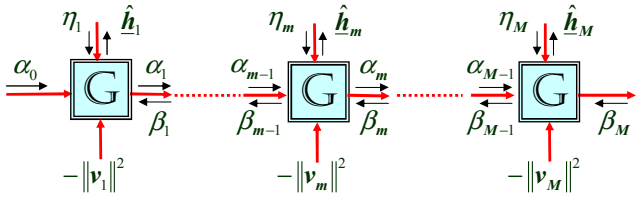

Fig. 2. FFG representation of the forward-and-backward signal processing for $\arg\max_{\underline{\mathbf{h}}_1^M} Q_M(\underline{\mathbf{h}}_1^M|\widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1})$, where the functional node $\mathbb{G}$ is defined in Fig. 1.

IV. Recursive Maximization of the EM Algorithm

Equations (9), (11) and (16) together imply that the K-L measure (10) can be expressed as a summation of two recursion forms, $Q_M(\underline{\mathbf{h}}_1^{m-1}|\widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1})$ and $Q_M(\underline{\mathbf{h}}_{m+1}^M|\widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1})$, plus the Guassian quadratic form $\lambda_m(\underline{\mathbf{h}}_m, \underline{\mu}_m)$. More specifically, by the factorization of

$$P(\underline{Y}_1^{m-1}, \underline{\mathbf{h}}_1^{m-1}, \underline{\mathcal{X}}_1^{m-1}) = P(\underline{Y}_1^{m-2}, \underline{\mathbf{h}}_1^{m-2}, \underline{\mathcal{X}}_1^{m-2})$$
$$P(\underline{Y}_{m-1}|\underline{\mathbf{h}}_{m-1})P(\underline{\mathbf{h}}_{m-1}|\underline{\mathbf{h}}_{m-2})P(\underline{\mathcal{X}}_{m-1}) \quad (17)$$

in (9), we have a forward recursion form of

$$Q_M(\underline{\mathbf{h}}_1^{m-1}|\widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1}) \equiv Q_M(\underline{\mathbf{h}}_1^{m-2}|\widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1}) + \lambda_{m-1}(\underline{\mathbf{h}}_{m-1}, \underline{\mathbf{h}}_{m-2}) \quad (18)$$

according to the results from (10) to (16). Similarly, we have a backward recursion of

$$Q_M(\underline{\mathbf{h}}_{m+1}^M|\widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1}) \equiv \lambda_{m+1}(\underline{\mathbf{h}}_{m+1}, \underline{\mathbf{h}}_m) + Q_M(\underline{\mathbf{h}}_{m+2}^M|\widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1}). \quad (19)$$

Successively applying the recursions of (18) and (19) to (10) yields

$$Q_M(\underline{\mathbf{h}}_1^M|\widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1}) \equiv \sum_{m=1}^{M}\lambda_m(\underline{\mathbf{h}}_m, \underline{\mathbf{h}}_{m-1}). \quad (20)$$

This completes the E-Step of the EM algorithm. As we can see from (12) and (13) that the evaluation relies on the computations for the posterior mean of $\mathbb{X}_m$ and its corresponding correlation matrix for all m. According to the system model (6), the *a posteriori* probability (APP) of $\mathbb{X}_m$ can be obtained with

$$P(\mathbb{X}_m|\underline{Y}_1^M; \widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1}) \propto P(\underline{Y}_m|\mathbb{X}_m, \widehat{\underline{\mathbf{h}}}_m^{\ell-1}). \quad (21)$$

For pilot symbols where $\mathbb{X}_m$ is given, there is no need for the evaluations.

A. Forward-and-Backward Recursive Maximization

Equation (20) and (16) suggest that the maximization of (20) *w.r.t.* $\underline{\mathbf{h}}_1^M$ can be done recursively using the max-sum rule in [8]. In the sequel, we present the message passing procedure for $\arg\max_{\underline{\mathbf{h}}_1^M} Q_M(\underline{\mathbf{h}}_1^M|\widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1})$. The details for the corresponding message updates can be found in [8].

Using Fig. 1, we can rewrite

$$\arg\max_{\underline{\mathbf{h}}_1^M} Q_M(\underline{\mathbf{h}}_1^M|\widehat{\underline{\mathbf{h}}}_{1:M}^{\ell-1}) = \arg\max_{\underline{\mathbf{h}}_m}\Big\{\max_{\underline{\mathbf{h}}_m^b}\max_{\underline{\mathbf{h}}_m^\diamond}\max_{\mathbf{v}_m}\max_{\underline{\mathbf{h}}^f_{m-1}}$$
$$\Big\{\alpha(\underline{\mathbf{h}}^f_{m-1}, \underline{\mu}^f_{m-1}) - \|\mathbf{v}_m\|^2 + \delta_{\underline{\mathbf{h}}_m^\diamond}(\mathbb{F}\underline{\mathbf{h}}^f_{m-1} + \mathbb{B}\mathbf{v}_m) +$$
$$\delta_{\underline{\mathbf{h}}_m}(\underline{\mathbf{h}}_m^\diamond) + \eta_m(\underline{\mathbf{h}}_m, \widetilde{\underline{\mu}}_m) + \delta_{\underline{\mathbf{h}}_m}(\underline{\mathbf{h}}_m^b) + \beta(\underline{\mathbf{h}}_m^b, \underline{\mu}_m^b)\Big\}\Big\} \quad (22)$$

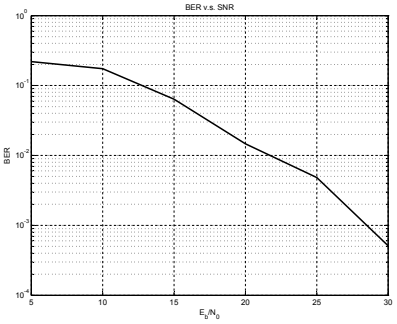

Fig. 3. BER of joint estimation and detection.

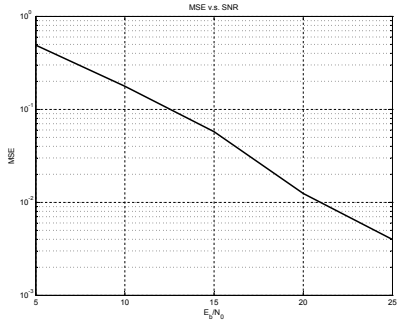

Fig. 4. MSE for the estimates of channel amplitudes.

for $m = 1, \ldots, M$, where we have the forward recursion of

$$\alpha(\underline{\mathbf{h}}_{m-1}, \underline{\mu}^f_{m-1}) \triangleq \max_{\underline{\mathbf{h}}^\diamond_{m-1}} \max_{\mathbf{v}_{m-1}} \max_{\underline{\mathbf{h}}_{m-2}} \Big\{ \alpha(\underline{\mathbf{h}}_{m-2}, \underline{\mu}^f_{m-2}) - \|\mathbf{v}_{m-1}\|^2 + \delta_{\underline{\mathbf{h}}^\diamond_{m-1}}(\mathbb{F}\underline{\mathbf{h}}_{m-2} + \mathbb{B}\mathbf{v}_{m-1}) + \delta_{\underline{\mathbf{h}}_{m-1}}(\underline{\mathbf{h}}^\diamond_{m-1}) + \eta_{m-1}(\underline{\mathbf{h}}_{m-1}, \underline{\widetilde{\mu}}_{m-1}) \quad (23)$$

and the backward recursion of

$$\beta(\underline{\mathbf{h}}_m, \underline{\mu}^b_m) \triangleq \max_{\underline{\mathbf{h}}^\diamond_{m+1}} \max_{\mathbf{v}_{m+1}} \max_{\underline{\mathbf{h}}_{m+1}} \max_{\underline{\mathbf{h}}^b_{m+1}} \Big\{ \beta(\underline{\mathbf{h}}^b_{m+1}, \underline{\mu}^b_{m+1}) + \delta_{\underline{\mathbf{h}}_{m+1}}(\underline{\mathbf{h}}^b_{m+1}) + \eta_{m+1}(\underline{\mathbf{h}}_{m+1}, \underline{\widetilde{\mu}}_{m+1}) + \delta_{\underline{\mathbf{h}}^\diamond_{m+1}}(\underline{\mathbf{h}}_{m+1}) - \|\mathbf{v}_{m+1}\|^2 + \delta_{\underline{\mathbf{h}}_m}(\mathbb{F}^{-1}[\underline{\mathbf{h}}^\diamond_{m+1} - \mathbb{B}\mathbf{v}_{m+1}]). \quad (24)$$

The maximization is essentially performed in 3 steps. First, pass messages from $t = 0$ to $M-1$ using (23) with an initial message $\alpha(\underline{\mathbf{h}}_0, \underline{\mu}^f_0) = -[\mathbf{h}_0 - \underline{\mu}^f_0]^H \Sigma^f_0 [\mathbf{h}_0 - \underline{\mu}^f_0]$. Second, pass messages from $t = M$ to 1 using (24) with $\beta(\underline{\mathbf{h}}_M, \underline{\mu}^b_M) = \mathbf{0}$. Finally, combine all messages with (22) resulting in $\widehat{\underline{\mathbf{h}}}_{1:M}$. The FFG representation for the forward-and-backward signal processing procedure is shown in Fig. 2.

V. Message Scheduling and Simulation Results

We now present the message passing method and the simulation results to demonstrate the performance of the proposed JED scheme for spatial multiplexed MIMO-OFDMA systems. The size of FFT is $N = 128$. Both the numbers of transmit and receive antennas are equal to 2. The number of tones for each user is 24. The cyclic prefix is $L_{cp} = 8$, and the number of path for each user is also $L = 8$. The MIMO channel coefficients for each user are independently generated using Jake's model with the normalized Doppler frequency shift $f_d T_s = 0.01$, where T_s is the OFDM symbol time. The size of each processing block for JED is 256 OFDM symbols with the first OFDM symbol of each block used as pilots. User data are modulated with DQPSK to combat the phase ambiguity resulting from JED.

To start JED, first APP($\mathbb{X}_1$) is evaluated with (21) by setting $\underline{\mathbf{h}}_1 = \mathbf{0}$. Given APP($\mathbb{X}_1$), we calculate $\widetilde{\mathbb{X}}_1$ and $\widetilde{\mathbb{C}}_1$ for $\eta_1(\underline{\mathbf{h}}_1, \underline{\widetilde{\mu}}_1)$ in (15) with (12) and (13), respectively. We next apply the forward update rule to obtain $\alpha(\underline{\mathbf{h}}_1, \underline{\mu}^f_1)$ in (23) by setting $\underline{\mu}^f_0 = \mathbf{0}$ and $\Sigma^f_0 = \mathbf{I}$ for $\alpha(\underline{\mathbf{h}}_0, \underline{\mu}^f_0)$. Having done that, we move to $t = 2$ and evaluate APP($\mathbb{X}_2$) by setting $\underline{\mathbf{h}}_2 = \underline{\mu}^f_1$. Repeating the aforementioned steps yields $\eta_2(\underline{\mathbf{h}}_2, \underline{\widetilde{\mu}}_2)$ and $\alpha(\underline{\mathbf{h}}_2, \underline{\mu}^f_2)$. Repeatedly applying these message passing steps results in $\eta_m(\underline{\mathbf{h}}_m, \underline{\widetilde{\mu}}_m)$ and $\alpha(\underline{\mathbf{h}}_m, \underline{\mu}^f_m)$, $\forall m = 1, \ldots, M$.

Given $\eta_m(\underline{\mathbf{h}}_m, \underline{\widetilde{\mu}}_m)$, we can start the backward message passing for $\beta(\underline{\mathbf{h}}_m, \underline{\mu}^b_m)$, with $\beta(\underline{\mathbf{h}}_M, \underline{\mu}^b_M) = 0$. Finally, combing $\alpha(\underline{\mathbf{h}}_m, \underline{\mu}^f_m)$ and $\beta(\underline{\mathbf{h}}_m, \underline{\mu}^b_m)$ with (22) for all m results in $\widehat{\underline{\mathbf{h}}}^1_{1:M}$ for the first iteration. The estimate $\widehat{\underline{\mathbf{h}}}^1_{1:M}$ can be used to evaluate APP($\mathbb{X}_m$) again, $\forall m = 1, \ldots, M$, hence starting another iteration of the forward-and-backward message passing for $\widehat{\underline{\mathbf{h}}}^2_{1:M}$. Fig. 3 presents the bit error rate (BER) of JED after 5 iterations of message passing, and the mean squared error (MSE) for the estimates of channel amplitudes is given in Fig. 4. These results show that the proposed method is effective for JED in MIMO fading channels.

VI. Conclusions

We presented a factor graph EM algorithm and the corresponding message-passing scheme for JED of MIMO-OFDMA systems. Simulation results show that the proposed scheme is effective for JED in MIMO fading channels.

References

[1] T.Y. Al-Naffouri, O. Awoniyia, O. Oteri, and A. Paulraj, "Receiver design for MIMO-OFDM transmission over time variant channels," in *Proc. of IEEE GLOBECOM*. San Franciso, California, Nov 2004.

[2] D. Schafhuber, G. Matz, and F. Hlawatsch, "Kalman tracking of timevarying channels in wireless MIMO-OFDM systems," in *Proc. of 36th Asilomar Conference on Signals, Systems and Computers*. Montery, California, Nov 2002.

[3] B. Lu, X. Wang, and K. R. Narayanan, "LDPC-Based space-time Coded OFDM systems over correlated fading channels: performance analysis and receiver design," *IEEE Trans. on Communications*, vol. 50, no. 1, pp. 74–88, Jan. 2002.

[4] F.-H. Chiu, S.-H. Wu, and C.-C. J. Kuo, "Receiver design for bit-interleaved MIMO-OFDM Systems over time-varying channels," in *Proc. of IEEE Wireless Communications and Networking Conference*. Las Vegas, Nevada, April 2006.

[5] S.-H. Wu, U. Mitra, and C.-C. J. Kuo, "Graph representation for joint channel estimation and symbol detection," in *Proc. IEEE Globecom*. Dallas, TX, Dec. 2004.

[6] G. K. Kaleh and R. Vallet, "Joint parameter estimation and symbol detection for linear or nonlinear unknown channels ," *IEEE Trans. on Communications*, vol. 42, no. 7, pp. 2406–2413, July 1994.

[7] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum-likelihood from incomplete data via the EM algorithm," *J. Roy. Statist. Soc.*, vol. 39, pp. 1–17, 1977.

[8] H.-A. Loeliger, J. Dauwels, J. Hu, S. Korl, L. Ping, and F. R. Kschischang, "The factor graph approach to model-based signal processing," *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1295–1322, 2007.