Low-Complexity MAP Channel Estimation for Mobile MIMO-OFDM Systems

Jie Gao and Huaping Liu, *Member, IEEE*

***Abstract*—This paper presents a reduced-complexity maximum *a posteriori* probability (MAP) channel estimator with iterative data detection for orthogonal frequency division multiplexing (OFDM) systems over mobile multiple-input multiple-output channels. The optimal MAP estimator needs to invert an $NN_T \times NN_T$ data-dependent matrix each in OFDM symbol interval, where N is the number of subcarriers and N_T is the number of transmit antennas. We derive an expectation maximization (EM) algorithm with low-rank approximation to avoid inverting large-size matrices, and thus drastically reduce the receiver complexity. In the iterative process, channel parameters are initially obtained by a least square (LS) estimator for temporary symbol decisions. Then, inter-carrier interference (ICI) due to fast fading is approximated and canceled. Finally, the temporary symbol decisions and the ICI-canceled received signals are processed by the EM-based MAP estimator to refine the channel state information for improved detection. The proposed scheme achieves about 2 dB gain over the LS scheme in channels with medium to high normalized Doppler shifts.**

***Index Terms*—Multiple-input multiple-output, orthogonal frequency division multiplexing, fast fading, inter-carrier interference, maximum *a posteriori* probability estimation, low-rank approximation.**

I. INTRODUCTION

ORTHOGONAL frequency division multiplexing (OFDM) is widely used for reliable data transmission over frequency-selective channels without requiring a complex equalizer. Multiple-input multiple-output (MIMO) antennas and OFDM can be implemented to achieve a low error rate and/or high data rate by flexibly exploiting the diversity gain and/or the spatial multiplexing gain [1]–[3]. Realizing these gains requires the channel state information (CSI) at the receiver, which is often obtained through channel estimation.

There exist two main types of channel estimation schemes: pilot-assisted schemes, in which a portion of the bandwidth is allocated to training symbols [4], [5], and blind approaches, which can be implemented by exploiting the statistical properties [6] or the deterministic information of the transmitted symbol (e.g., finite alphabet, constant modulus, etc.) [7], [8]. For pilot-assisted schemes, CSI can be estimated by exploiting the frequency correlation and/or the time correlation of the pilot and data symbols [9]. The estimates are in general reliable, but pilot symbols increase signaling overhead. On the other hand, blind estimation requires a long data observation interval; the slow convergence rate makes it difficult to apply the statistical approach in fast-fading channels, and the high computational complexity required to solve the maximization problem makes the deterministic approach appropriate only for certain applications. In [10], algorithms based on comb-type pilots with improvement using interpolation at data frequencies are studied. Performance bound of a pilot-assisted, least-square (LS) channel estimator over a slowly fading channel is derived in [11]. A Kalman filter based scheme to estimate the state-transition matrix of time-varying MIMO-OFDM channels and a scheme based on minimizing the mean-square error (MSE) of a cost function are developed in [12], [13] and [14], [15], respectively. To enhance the LS channel estimation for MIMO-OFDM systems, optimal pilot sequences and optimal placements of pilot tones are derived in [16]. Two expectation-maximization (EM) algorithms, the classical EM and space-alternating generalized EM (SAGE), are compared in terms of their convergence rates in [17], [18]. In fast-fading channels, inter-carrier interference (ICI) in OFDM systems could be severe. In order to mitigate ICI, various detection structures are proposed in [4] and an iterative channel estimator with ICI cancellation to maximize the signal-to-noise-plus-ICI ratio is derived in [5].

Maximum *a posteriori* probability (MAP) channel estimation algorithms generate optimal results. When applied to MIMO-OFDM systems, however, its complexity could be prohibitively high for most applications. This paper develops an iterative channel estimation and data detection scheme for mobile MIMO-OFDM systems for which ICI may not be neglected. The main contribution is on the derivation of a reduced-complexity MAP channel estimator while maintaining a high data-detection performance. In the proposed scheme, the LS algorithm that operates on pilot symbols only is applied to obtain initial channel estimates for temporary symbol decisions. Then, the ICI component is approximated and canceled from the received signals. In fast-fading channels, LS estimates exploiting pilot symbols only might not be sufficient to provide a high detection performance. With the temporary data decisions and channel estimates, performance could be significantly improved by applying a MAP estimator. The major problem with the MAP estimator is that it requires inversion and multiplication of matrices of size $NN_T \times NN_T$ for each OFDM symbol, where N is the number of subcarriers and N_T is the number of transmit antennas. We derive an EM-based MAP estimator, which, by exploiting the channel statistical information and employing a low-rank approximation, practically eliminates the need of frequent matrix inversions. We show that the proposed scheme performs the same as the MAP estimator while its complexity is significantly lower.

Manuscript received December 27, 2005; revised December 26, 2006 and July 11, 2007; accepted November 1, 2007. The associate editor coordinating the review of this paper and approving it for publication was J. Garcia-Frias.

J. Gao was with the School of Electrical Engineering and Computer Science, Oregon State University, Corvallis, OR 97331, USA. She is now with Intel Corporation, Santa Clara, CA 95052, USA (e-mail: jie.gao@intel.com).

H. Liu is with the School of Electrical Engineering and Computer Science, Oregon State University, Corvallis, OR 97331, USA (e-mail: hliu@eecs.oregonstate.edu).

Digital Object Identifier 10.1109/TWC.2008.051072.

II. System Model

Consider a system with N_T transmit antennas and N_R receive antennas. Sub-channels are defined as the spatial channels from the u-th $(1 \le u \le N_T)$ transmit antenna to the v-th $(1 \le v \le N_R)$ receive antenna. In the transmitter, data are serial-to-parallel (S/P) converted and sent to N_T transmit antennas for simultaneous transmission. Each sub-channel consists of $L+1$ paths, and each OFDM symbol consists of N subcarriers. Let $h_{v,u}(l)$ denote the tap gain of path l for the sub-channel from transmit antenna u to receive antenna v. The channel has an exponentially decaying multipath power-delay profile which determines the power distribution among the taps, and the maximum tap delay is assumed to be shorter than the OFDM guard interval. The power-delay profile is also assumed identical for all independent sub-channels, which is represented as [19]

$$E\left\{h_{v,u}(l_1)h^*_{v,u}(l_2)\right\} = \epsilon e^{-l_1/L}\delta_{l_1 l_2} \tag{1}$$

where $\epsilon = \frac{1-e^{-1/L}}{1-e^{-(L+1)/L}}$ is a normalization factor to ensure $\epsilon \sum_l e^{-l/L} = 1$, δ denotes the Kronecker delta function, $E\{\cdot\}$ denotes statistical expectation, and $\{\cdot\}^*$ represents complex conjugate. Let $\mathcal{H}(l)$ denote the $N_R \times N_T$ spatial channel matrix whose (v,u)-th element is $h_{v,u}(l)$. From Eq. (1), it is easy to obtain $E\left\{\text{vec}\{\mathcal{H}(l_1)\}\text{vec}^H\{\mathcal{H}(l_2)\}\right\} = \mathbf{0}_{N_T N_R},\ \forall l_1 \ne l_2$, where $(\cdot)^H$ denotes Hermitian transpose, $\text{vec}\{\mathcal{H}(l)\}$ is an $N_R N_T \times 1$ vector constructed by stacking the columns of $\mathcal{H}(l)$, and $\mathbf{0}_{N_T N_R}$ represents the $N_T N_R \times N_T N_R$ zero matrix. We write the discrete-time multipath channel coefficients along the delay path as $\boldsymbol{h}_{v,u} = [h_{v,u}(0), h_{v,u}(1), \cdots, h_{v,u}(L)]^T$, where $(\cdot)^T$ denotes transpose. We also define

$$H_{v,u}(k) = \sum_{l=0}^{L} h_{v,u}(l)e^{-j2\pi kl/N}, \quad 0 \le k \le N-1 \tag{2}$$

which represents the frequency response of the channel for the k-th subcarrier. In a vector-matrix form, $H_{v,u}(k)$ is the k-th element of the $N \times 1$ vector $\boldsymbol{F}\boldsymbol{h}_{v,u}$, where $\boldsymbol{F}$ is an $N \times (L+1)$ matrix with $\boldsymbol{F}[k,l] = e^{-j2\pi kl/N}$, $0 \le k \le N-1$, $0 \le l \le L$.

The discrete-time transmitted signal at the n-th sampling interval from antenna u is expressed as $s_u(n) = \sqrt{\frac{E_s}{N}} \sum_{k=0}^{N-1} d_u(k)e^{j2\pi nk/N}$, where $d_u(k)$ is the transmitted symbol at the k-th subcarrier from the u-th antenna and E_s is the data symbol energy per subcarrier. Since the guard interval is not shorter than the maximum delay of the channel, there is no intersymbol interference. For simplicity of notation, we let the symbol energy per subcarrier be normalized to 1. The received signal on the k-th subcarrier of antenna v is given by

$$Y_v(k) = \sum_{u=1}^{N_T} d_u(k)H_{v,u}(k) + W_v(k) \tag{3}$$

where $W_v(k) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} w_v(n)e^{-j2\pi nk/N}$ and $w_v(n)$ is the zero-mean additive white Gaussian noise (AWGN) with variance N_0.

III. Low-complexity MAP channel estimation with iterative data detection

In this section, we first derive the proposed scheme assuming a quasi-static channel. Then we extend the analysis to time-varying fading channels, where ICI needs to be considered.

A. MAP channel estimation with low-rank approximation and EM implementation

1) LS estimation: As mentioned in Section I, the LS scheme is a pilot-assisted approach to obtain the initial estimates. Let p_s and M denote the pilot subcarrier spacing and the number of subcarriers dedicated to pilot symbols, respectively. The received pilot vector at receive antenna v, $\underline{Y}_{v(p)}$, and the transmitted pilot matrix $\boldsymbol{D}_{u(p)}$ from antenna u are written as $\underline{Y}_{v(p)} = [Y_v(0), Y_v(p_s), \cdots, Y_v((M-1)p_s)]^T$, $\boldsymbol{D}_{u(p)} = \text{diag}[d_u(0), d_u(p_s), \cdots, d_u((M-1)p_s)]$, where $\text{diag}[\cdot]$ denotes a diagonal matrix. Applying Eq. (3) to pilot signals, we obtain

$$\underline{Y}_{v(p)} = \boldsymbol{Q}_{(p)}\boldsymbol{h}_v + \underline{W}_{v(p)} \tag{4}$$

where $\boldsymbol{Q}_{(p)} = [\boldsymbol{D}_{1(p)}\boldsymbol{F}_{(p)}, \boldsymbol{D}_{2(p)}\boldsymbol{F}_{(p)}, \cdots, \boldsymbol{D}_{N_T(p)}\boldsymbol{F}_{(p)}]$, $\boldsymbol{h}_v = [\boldsymbol{h}^T_{v,1}, \boldsymbol{h}^T_{v,2}, \cdots, \boldsymbol{h}^T_{v,N_T}]^T$, and $\boldsymbol{F}_{(p)}$ is an $M \times (L+1)$ matrix with $\boldsymbol{F}_{(p)}[k,l] = e^{-j2\pi kl/N}$, $k = 0, p_s, 2p_s, \cdots, (M-1)p_s$, $0 \le l \le L$, $W_v(k)$ is the k-th element of $M \times 1$ vector $\underline{W}_{v(p)}$.

The LS estimate of $\boldsymbol{h}_v$ is simply obtained as $\hat{\boldsymbol{h}}_v = (\boldsymbol{Q}_{(p)})^+\underline{Y}_{v(p)}$, where $(\cdot)^+$ denotes the pseudo-inverse. Since $\boldsymbol{Q}_{(p)}$ is an $M \times N_T(L+1)$ matrix, a unique LS solution exists if the number of pilot subcarriers M is not less than N_T times the number of channel delay taps $(L+1)$. Calculating the inverse of an $N_T(L+1) \times N_T(L+1)$ matrix could be computationally extensive. Thus, it is favorable to ignore the channel taps whose magnitudes are small, like the method of significant tap catching (STC) proposed in [14]. With L_r significant taps $(L_r < L+1)$, the required computation is reduced to the inversion of an $N_T L_r \times N_T L_r$ matrix. However, an irreducible error floor is introduced since the power-delay profile cannot be completely represented by the L_r taps [18], [20].

An EM-based scheme that provides a more reliable channel estimate than the STC scheme while avoiding the inversion of large-size matrices is introduced and compared with the SAGE algorithm in terms of convergence rate in [17], [18]. This algorithm transforms the estimation process of multiple-input channels into the estimation of a series of independent single-input single-output (SISO) channels. In the *E-step*, $\hat{\underline{Y}}^{(\kappa)}_{v,u(p)} = \boldsymbol{D}_{u(p)}\boldsymbol{F}_{(p)}\hat{\boldsymbol{h}}^{(\kappa)}_{v,u}$ and $\hat{\boldsymbol{r}}^{(\kappa)}_{v,u(p)} = \hat{\underline{Y}}^{(\kappa)}_{v,u(p)} + \beta_u\left[\underline{Y}_{v(p)} - \sum_{u=1}^{N_T}\hat{\underline{Y}}^{(\kappa)}_{v,u(p)}\right]$ are computed for $u = 1, 2, \ldots, N_T$, where superscript $^{(\kappa)}$ represents the κ-th sub-iteration and $\sum_{u=1}^{N_T}\beta_u = 1$. Typically, $\beta_u, u = 1, \cdots, N_T$, are chosen as $\beta_1 = \cdots = \beta_{N_T}$. In the *M-step*, channel coefficients are estimated as $\hat{\boldsymbol{h}}^{(\kappa+1)}_{v,u} = \boldsymbol{F}^H_{(p)}\boldsymbol{D}^{-1}_{u(p)}\hat{\boldsymbol{r}}^{(\kappa)}_{v,u(p)}$. Since $\boldsymbol{D}_{u(p)}$ is a diagonal matrix, $\boldsymbol{D}^{-1}_{u(p)}$ can be obtained via division only. The channel estimates can be initially set as $\hat{\boldsymbol{h}}^{(0)}_{v,u} = \mathbf{1}_{L+1}, (1 \le v \le N_R, 1 \le u \le N_T)$, where $\mathbf{1}_{L+1}$ is an $(L+1) \times 1$ vector whose elements are all 1's.

With the channel coefficients $\hat{\boldsymbol{h}}_{v,u}$ estimated by the LS algorithm, the estimate of the channel frequency response for subcarrier k, $\hat{H}_{v,u}(k)$, is simply the k-th element of $\boldsymbol{F}\hat{\boldsymbol{h}}_{v,u}$, where $\boldsymbol{F}$ is an $N \times (L+1)$ matrix given below (2).

The received signals across all receive antennas on the k-th subcarrier, $\underline{Y}(k)$, is expressed in a vector form as

$\underline{Y}(k) = [Y_1(k), Y_2(k), \cdots, Y_{N_R}(k)]^T$. With $\hat{H}_{v,u}(k)$, data symbols $d_1(k), \cdots, d_{N_T}(k)$ can be detected after spatial demultiplexing using a linear zero-forcing (ZF) filter or minimum mean square error (MMSE) scheme.

2) MAP estimation: Although the transmitted data could be detected by employing the LS estimates of the channel coefficients, the performance will be significantly improved by employing a MAP channel estimator. The received signal vector on all subcarriers at antenna v, $\underline{Y}_v = [Y_v(0), \cdots, Y_v(N-1)]^T$, can be expressed as

$$\underline{Y}_v = \boldsymbol{D}\boldsymbol{H}_v + \underline{W}_v, \quad v = 1, \cdots, N_R \tag{5}$$

where $\boldsymbol{D} = [\boldsymbol{D}_1, \cdots, \boldsymbol{D}_{N_T}]$, $\boldsymbol{D}_u = \mathrm{diag}[d_u(0), \cdots, d_u(N-1)]$, $\underline{W}_v = [W_v(0), \cdots, W_v(N-1)]^T$, and $\boldsymbol{H}_v = [(\boldsymbol{H}_{v,1})^T, \cdots, (\boldsymbol{H}_{v,N_T})^T]^T$, whose u-th block is expressed as $\boldsymbol{H}_{v,u} = \boldsymbol{F}\boldsymbol{h}_{v,u}$.

As mentioned in Section I, the major problem with the MAP estimator is that it needs to invert an $NN_T \times NN_T$ data-dependent correlation matrix and an $N \times (NN_T)$ data matrix $\boldsymbol{D}$, where the data-dependency requires the inversion be carried out for each OFDM symbol. The computational load becomes prohibitively high when NN_T is large. This partly motivates us to apply the EM algorithm to be described at the end of Section III to decompose the MIMO channel into N_T SISO channels. After the decomposition, the data matrix reduces to an $N \times N$ diagonal matrix whose inversion is trivial. The data-dependent correlation matrix also reduces to size $N \times N$, and we derive a low-rank approximation to practically avoid matrix inversion. It is more convenient to write the received vector at antenna v as: $\underline{Y}_v = \sum_{u=1}^{N_T} \underline{Y}_{v,u}$ and $\underline{Y}_{v,u} = \boldsymbol{D}_u\boldsymbol{H}_{v,u} + W_{v,u}$. Since the EM algorithm can decompose the N_T spatialy multiplexed channels given in $\underline{Y}_v$, we develop the MAP estimator based on above expression.

With the received signal at antenna v, the optimal MAP estimator maximizes the probability density function (pdf) of $\boldsymbol{H}_{v,u}$ conditioned on the received signal and the transmitted data matrix as [21]

$$\begin{aligned}\hat{\boldsymbol{H}}_{v,u} &= \arg\max_{\boldsymbol{H}_{v,u}} f\left(\boldsymbol{H}_{v,u}|\underline{Y}_{v,u}, \boldsymbol{D}_u\right) \\ &= \arg\max_{\boldsymbol{H}_{v,u}} f\left(\underline{Y}_{v,u}|\boldsymbol{H}_{v,u}, \boldsymbol{D}_u\right) f\left(\boldsymbol{H}_{v,u}|\boldsymbol{D}_u\right)\end{aligned} \tag{6}$$

where

$$f(\underline{Y}_{v,u}|\boldsymbol{H}_{v,u}, \boldsymbol{D}_u) = \pi^{-1}|\boldsymbol{R}_N|^{-1} \cdot \exp(-(\underline{Y}_{v,u} - \boldsymbol{D}_u\boldsymbol{H}_{v,u})^H \boldsymbol{R}_N^{-1}(\underline{Y}_{v,u} - \boldsymbol{D}_u\boldsymbol{H}_{v,u})) \tag{7a}$$

$$f(\boldsymbol{H}_{v,u}|\boldsymbol{D}_u) = \pi^{-1}|\boldsymbol{R}_H|^{-1}\exp(-\boldsymbol{H}_{v,u}^H\boldsymbol{R}_H^{-1}\boldsymbol{H}_{v,u}). \tag{7b}$$

It was shown in [21], [22] that the MAP estimate of $\boldsymbol{H}_{v,u}$ can be expressed as

$$\hat{\boldsymbol{H}}_{v,u} = \boldsymbol{\mu} + \boldsymbol{R}_H\boldsymbol{D}_u^H\left(\boldsymbol{D}_u\boldsymbol{R}_H\boldsymbol{D}_u^H + \boldsymbol{R}_N\right)^{-1}\left(\underline{Y}_{v,u} - \boldsymbol{D}_u\boldsymbol{\mu}\right) \tag{8}$$

where $\boldsymbol{R}_N$ is the correlation matrix of the zero-mean noise vector, and $\boldsymbol{\mu}$ and $\boldsymbol{R}_H$ denote, respectively, the mean and corrrelation matrix of $\boldsymbol{H}_{v,u}$. In a quasi-static channel, $\boldsymbol{R}_N$ is expressed as $\boldsymbol{R}_N = \sigma^2\boldsymbol{I}_N$, where $\boldsymbol{I}_N$ is the $N \times N$ identity matrix and $\sigma^2 = \sigma^2_{\mathrm{AWGN}}/N_T$. Considering that the channel coefficients have a mean zero (i.e., $\boldsymbol{\mu} = \boldsymbol{0}$), it is straightforward to rewrite (8) as

$$\hat{\boldsymbol{H}}_{v,u} = \boldsymbol{R}_H(\boldsymbol{R}_H + \sigma^2(\boldsymbol{D}_u^H\boldsymbol{D}_u)^{-1})^{-1}\boldsymbol{D}_u^{-1}\underline{Y}_{v,u}. \tag{9}$$

Clearly, the inversion of the $N \times N$ data-dependent correlation matrix and the multiplication of two $N \times N$ matrices must be done for *all sub-channels* during each OFDM symbol interval.

3) Complexity reduction via low-rank approximation: It was shown in [24] that matrix $(\boldsymbol{D}_u^H\boldsymbol{D}_u)^{-1}$ could be replaced by $E\{(\boldsymbol{D}_u^H\boldsymbol{D}_u)^{-1}\}$ at the expense of a slight performance degradation. Assuming a normalized constellation power and equally probable constellation points and independent data symbols, we can easily show that $E\{(\boldsymbol{D}_u^H\boldsymbol{D}_u)^{-1}\} = \alpha\boldsymbol{I}_N$, where α equals $1, 1.8889,$ and 2.6854 for QPSK, 16-QAM, and 64-QAM, respectively. Thus, Eq. (9) can be approximated as

$$\hat{\boldsymbol{H}}_{v,u} \approx \boldsymbol{R}_H(\boldsymbol{R}_H + \sigma^2\alpha\boldsymbol{I}_N)^{-1}\boldsymbol{D}_u^{-1}\underline{Y}_{v,u}. \tag{10}$$

The approximation in (10) effectively avoids the frequent inversion and multiplication of $N \times N$ matrices for every OFDM symbol (note again that $\boldsymbol{D}_u^{-1}$ is a diagonal matrix).

The complexity can be further reduced by exploiting low-rank approximation to the matrices involved in the MAP estimation process. Let $\boldsymbol{\Gamma} = \boldsymbol{R}_H(\boldsymbol{R}_H + \sigma^2\alpha\boldsymbol{I}_N)^{-1}$. It was shown in [25] that $\boldsymbol{\Gamma}$ can be optimally approximated by an $N \times N$ matrix $\boldsymbol{\Gamma}_m$ with low rank. The optimal rank reduction can be achieved by minimizing the trace of the extra covariance as $\min_{\boldsymbol{\Gamma}_m} tr[(\boldsymbol{\Gamma} - \boldsymbol{\Gamma}_m)(\boldsymbol{R}_H + \sigma^2\alpha\boldsymbol{I}_N)(\boldsymbol{\Gamma} - \boldsymbol{\Gamma}_m)^T]$. The solution will make $\boldsymbol{\Gamma}_m(\boldsymbol{R}_H + \sigma^2\alpha\boldsymbol{I}_N)^{1/2}$ the best low-rank approximation for $\boldsymbol{\Gamma}(\boldsymbol{R}_H + \sigma^2\alpha\boldsymbol{I}_N)^{1/2}$.

Since the correlation matrix $\boldsymbol{R}_H$ is Hermitian and positive semidefinite, we can write $\boldsymbol{R}_H = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^H$, where $\boldsymbol{U}$ is a unitary matrix and $\boldsymbol{\Lambda}$ is a diagonal matrix whose diagonal elements λ_m, $m = 0, 1, \cdots, N{-}1$, are the eigen values of $\boldsymbol{R}_H$. The MAP channel estimate given in (10) can be expressed as $\hat{\boldsymbol{H}}_{v,u} = \boldsymbol{U}\boldsymbol{\Lambda}(\boldsymbol{\Lambda} + \sigma^2\alpha\boldsymbol{I}_N)^{-1}\boldsymbol{U}^H\boldsymbol{D}_u^{-1}\underline{Y}_{v,u}$. We also have $\boldsymbol{\Gamma}(\boldsymbol{R}_H + \sigma^2\alpha\boldsymbol{I}_N)^{1/2} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^H(\boldsymbol{U}(\boldsymbol{\Lambda} + \sigma^2\alpha\boldsymbol{I}_N)\boldsymbol{U}^H)^{-1/2} = \boldsymbol{U}\boldsymbol{\Lambda}\left(\boldsymbol{\Lambda} + \sigma^2\alpha\boldsymbol{I}_N\right)^{-1/2}\boldsymbol{U}^H$, where we have applied the property $(\boldsymbol{U}(\boldsymbol{\Lambda} + \sigma^2\alpha\boldsymbol{I}_N)\boldsymbol{U}^H)^{1/2} = \boldsymbol{U}(\boldsymbol{\Lambda} + \sigma^2\alpha\boldsymbol{I}_N)^{1/2}\boldsymbol{U}^H$ in obtaining the second equality.

Let $\boldsymbol{\Delta} = \boldsymbol{\Lambda}(\boldsymbol{\Lambda} + \sigma^2\alpha\boldsymbol{I}_N)^{-1}$. The optimal low-rank approximation for $\boldsymbol{\Delta}$ would then be

$$\boldsymbol{\Delta}_m = \mathrm{diag}\left[\frac{\lambda_0}{\lambda_0 + \sigma^2\alpha} \;\cdots\; \frac{\lambda_L}{\lambda_L + \sigma^2\alpha} \;\; 0 \;\cdots\; 0\right]. \tag{11}$$

Therefore, the low-rank approximated channel estimate based on (10) is expressed as

$$\hat{\boldsymbol{H}}_{v,u} \approx \boldsymbol{U}\boldsymbol{\Delta}_m\boldsymbol{U}^H\boldsymbol{D}_u^{-1}\underline{Y}_{v,u}. \tag{12}$$

As shown in the Appendix, given the channel length L, the (m,n)-th element of the correlation matrix $\boldsymbol{R}_H$ can be derived to be $[\boldsymbol{R}_H]_{m,n} = \epsilon\sum_{l=0}^{L} e^{-l/L}e^{-j2\pi(m-n)l/N}$.

4) Decomposition using EM algorithm: To estimate the coefficients for the channel vectors from all transmit antennas, the EM algorithm to decompose the MIMO channel into SISO channels can be efficiently implemented as

E-step: for $u = 1, 2, \ldots, N_T$,

$$\underline{\hat{Y}}_{v,u}^{(g)} = \boldsymbol{D}_u \boldsymbol{U} \boldsymbol{\Delta}_m^{+} \boldsymbol{U}^H \boldsymbol{F} \hat{\boldsymbol{h}}_{v,u}^{(g)} \tag{13a}$$

$$\hat{\boldsymbol{r}}_{v,u}^{(g)} = \underline{\hat{Y}}_{v,u}^{(g)} + \beta_u \left[\underline{Y}_v - \sum_{u=1}^{N_T} \underline{\hat{Y}}_{v,u}^{(g)} \right] \tag{13b}$$

where, as the EM-LS scheme, superscript $^{(g)}$ represents the g-th sub-iteration and $\beta_u, u = 1, \cdots, N_T$, satisfy $\sum_{u=1}^{N_T} \beta_u = 1$ and are typically chosen as $\beta_1 = \cdots = \beta_{N_T}$. Note that the $L+1$ non-zero diagonal elements of $\boldsymbol{\Delta}_m^{+}$ is easily obtained to be $\frac{\lambda_m + \sigma^2 \alpha}{\lambda_m}, m = 0, 1, 2, \cdots, L$.

M-step: in order to minimize the detection error, the estimated channel coefficients are updated as

$$\hat{\boldsymbol{h}}_{v,u}^{(g+1)} = \boldsymbol{F}^H \boldsymbol{U} \boldsymbol{\Delta}_m \boldsymbol{U}^H \boldsymbol{D}_u^{-1} \hat{\boldsymbol{r}}_{v,u}^{(g)}. \tag{14}$$

Since $\boldsymbol{D}_u$ is a diagonal matrix, $\boldsymbol{D}_u^{-1}$ can be obtained by division only; thus practically no matrix inversion is required for the proposed EM-based MAP channel estimator. Also, for practical scenarios, the SVD of $\boldsymbol{R}_H$ can be calculated in advance and updated infrequently. Increasing the number of sub-iterations G will result in better quality of the channel estimates. However, as will be shown in Section IV, for most common MIMO-OFDM configurations and fading rates, the performance saturates quickly as G increases; thus, a large G is typically unnecessary.

B. Implementation in time-varying fading channels

The proposed scheme can be easily extended and applied to the time-varying channels. In this case the channel coefficients at time nT_s, $h_{v,u}(n,l)$, are assumed to be constant in one *sampling interval* T_s, which is related to the OFDM symbol duration T as $T_s = T/N$, and change over different sampling intervals according to the channel correlation property $E\left\{h_{v,u}(n_1,l_1) h_{v,u}^*(n_2,l_2)\right\} = \epsilon J_0\left(2\pi f_d T_s (n_2 - n_1)\right) e^{-l_1/L} \delta_{l_1 l_2}$, where $J_0(\cdot)$ is the zeroth order Bessel function of the first kind, f_d is the maximum Doppler shift of the channel.

Time-varying fading causes ICI; thus (3) must be modified as

$$Y_v(k) = \sum_{u=1}^{N_T} d_u(k) \bar{H}_{v,u}(k) + \zeta_v(k) + W_v(k) \tag{15}$$

where $\bar{H}_{v,u}(k) = \frac{1}{N} \sum_{n=0}^{N-1} H_{v,u}(k,n)$ denotes the *mean value* of the channel response for the k-th subcarrier, $H_{v,u}(k,n) = \sum_{l=0}^{L} h_{v,u}(n,l) e^{-j2\pi kl/N}$, $\zeta_v(k) = \sum_{u=1}^{N_T} \zeta_{v,u}(k)$, and $\zeta_{v,u}(k) = \frac{1}{N} \sum_{\substack{m=0 \\ m \neq k}}^{N-1} d_u(m) \sum_{n=0}^{N-1} H_{v,u}(m,n) e^{j2\pi n(m-k)/N}$ represents the ICI component. For most common OFDM systems, the number of subcarrier is a large number (e.g., 128 or greater). Therefore, the ICI component $\zeta_{v,u}(k)$ can be approximated as a zero-mean Gaussian random variable by invoking the central limit theorem [23]. The variance of $\zeta_{v,u}(k)$ is given as [4]

$$\sigma_{\text{ICI}}^2 = \frac{1}{N^2} \sum_{m=0,\ m \neq k}^{N-1} \left[N + 2 \sum_{n=1}^{N-1} (N-n) J_0\left(2\pi f_d T_s n\right) \cos\left(2\pi (m-k) \frac{n}{N}\right) \right]. \tag{16}$$

Therefore, with the modification of the noise variance in (9)−(11) as $\sigma^2 = \sigma_{\text{ICI}}^2 + \sigma_{\text{AWGN}}^2 / N_T$, the proposed scheme can be readily applied for time-varying channels.

When ICI is severe, it is necessary to cancel it in the detection process. The channel transfer function can be approximated using the first-order Taylor series expansion as [5] $H_{v,u}(k,n) = H_{v,u}(k,n_0) + H_{v,u}^{'}(k,n_0)(n - n_0)$. The ICI component $\zeta_{v,u}(k),\ k = 0, 1, \cdots, N-1$, defined in (15) can be rewritten as

$$\zeta_{v,u}(k) = \sum_{m=0}^{N-1} H_{v,u}(m,n_0)^{'} \Xi_k(m) d(m) \tag{17}$$

where $\Xi_k(m) = \frac{1}{N} \sum_{n=0}^{N-1} (n - n_0) e^{j2\pi n(m-k)/N}$. Let $\boldsymbol{\Xi}$ be an $N \times N$ matrix whose (k,m)-th element is $\Xi_k(m), k, m = 0, 1, \cdots, N-1$. With the initial estimate of the channel and the temporary symbol decisions for all the subcarriers, the ICI component is approximated as

$$\hat{\boldsymbol{\zeta}}_{v,u} = \boldsymbol{\Xi} \boldsymbol{H}_{v,u}^{'} \hat{\boldsymbol{d}}_u \tag{18}$$

where $\hat{\boldsymbol{\zeta}}_{v,u} = [\hat{\zeta}_{v,u}(0), \hat{\zeta}_{v,u}(1), \cdots, \hat{\zeta}_{v,u}(N-1)]^T$, $\hat{\boldsymbol{d}}_u = [\hat{d}_u(0), \hat{d}_u(1), \cdots, \hat{d}_u(N-1)]^T$, and $\boldsymbol{H}_{v,u}^{'} = \text{diag}[H_{v,u}^{'}(0,n_0), H_{v,u}^{'}(1,n_0), \cdots, H_{v,u}^{'}(N-1,n_0)]$. The first-order derivative of the channel response, $H_{v,u}^{'}(k,n_0)$, can be estimated by calculating the difference of $\bar{H}_{v,u}(k)$ between two consecutive OFDM symbols [5]. The ICI component is then canceled before the next iteration of data detection as $\underline{\hat{Y}}_v = \underline{Y}_v - \sum_{u=1}^{N_T} \hat{\boldsymbol{\zeta}}_{v,u}$, where $\underline{Y}_v$ was given in (5). Once the ICI component is canceled from the received signal, both channel estimation and data detection should be significantly improved.

The block diagram of the proposed MAP channel estimator with iterative data detection for MIMO-OFDM systems is shown in Fig. 1. The receiver employs an LS channel estimator to obtain the *initial* estimate of the channel coefficients for all the sub-channels by exploiting only the pilot signals, followed by an MMSE data detector. Once the temporary data decisions ($\boldsymbol{D}_u$) are available, the ICI component can be approximated and canceled from the received signal. The received signals after ICI cancellation, the temporary symbol decisions, and the statistical information of the channel are then processed by the proposed MAP estimator to obtain more accurate channel parameters ($\hat{\boldsymbol{H}}_{v,u}$). In the next iteration, the temporary data decisions are used to estimate the ICI component given by (18), which is subsequently canceled from the received signal. The channel parameters are then updated following (13) and (14).

The MAP estimator can be derived using (5). Its complexity is approximately as follows. For *each OFDM symbol*, the pseudo-inverse of a data matrix of size $N \times NN_T$: $\mathcal{O}(N^3)$); inversion of a data-dependent correlation matrix of

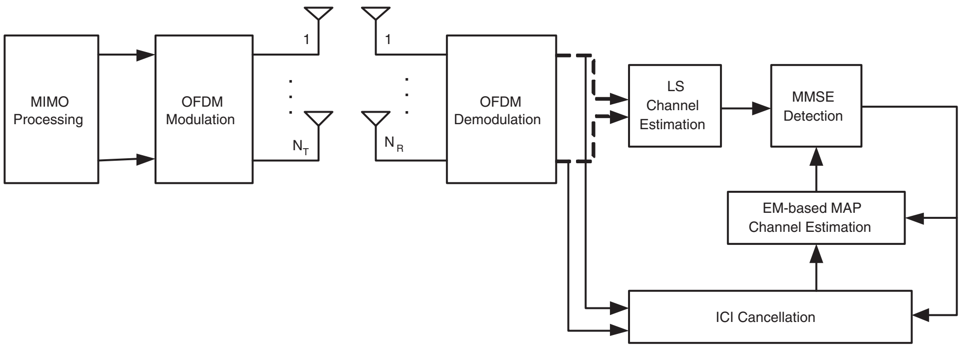

Fig. 1. Block diagram of the iterative channel estimation and data detection scheme.

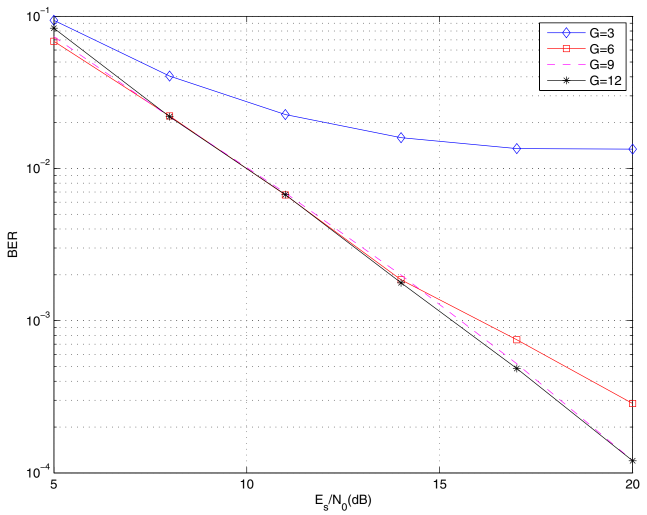

Fig. 2. The effect of the number of sub-iterations in the EM process ($N_T = 2,\ N_R = 3$).

size $NN_T \times NN_T$: $\mathcal{O}((NN_T)^3)$; and multiplication of two $NN_T \times NN_T$ matrices: $\mathcal{O}((NN_T)^3)$. The significance of (12) as a result of the approximation in (10) and the low-rank approximation and the EM algorithm is that the inversion and multiplication of matrices mentioned above are practically eliminated (note that $\boldsymbol{D}_u$ and $\boldsymbol{\Delta}_m$ are diagonal matrices). Additionally, as mentioned at the end of Section III-A.3, even if f_d changes, an SVD to re-calculate $\boldsymbol{U}$ is not needed, as long as f_dT is not unrealistically large. The only change needed to reflect the change in f_d is to re-calculate $\boldsymbol{\Delta}_m$ given in (11), which is trivial due to the special form of $\boldsymbol{\Delta}_m$, once the variance of ICI as a result of fast fading is estimated. However, in the EM process, steps given in (13) and (14) need to be executed g times (e.g., $g = 9$) for each OFDM symbol. Although the exact complexity of these steps are difficult to quantify, it is far lower than $\mathcal{O}((NN_T)^3))$ since all matrices involved are either fixed (e.g., $\boldsymbol{U}$ and $\boldsymbol{F}$, which do not need to be updated on a per OFDM symbol basis) or diagonal (e.g., $\boldsymbol{\Delta}_m$ and $\boldsymbol{D}_u$).

IV. Simulation Results and Discussion

Simulation results are obtained for MIMO-OFDM systems with $N = 128$ subcarriers employing QPSK modulation. A cyclic prefix of 16 samples is inserted at the beginning of each OFDM symbol. The pilot subcarrier spacing p_s is 4; thus the absolute pilot spacing in the frequency domain equals $4/(NT_s)$. Note that data symbol energy E_s for all simulation results is the energy spent in information-bearing symbols only and is not adjusted by the energy spent in pilots. Since the multipath spread of the channel is assumed to be LT_s, the channel coherence bandwidth is approximately equal to $1/(LT_s)$. Let R be the ratio of the absolute pilot spacing to the channel coherence bandwidth, i.e., $R = \frac{4L}{N}$. Since the grid density of the pilot symbols must satisfy the 2-D sampling theorem in order to recover channel parameters [26], the pilot spacing must be less than or equal to half of the coherence bandwidth of the channel, which results in $R \leq 0.5$. The channel is assumed to have 15 multipath components ($L = 14$).

We use the normalized Doppler shift f_dT to measure fading rates. In terrestrial digital video broadcasting 2k mode (DVB-T) systems, if the vehicle speed is $v = 134.8$ km/h, the normalized Doppler shift is obtained to be $f_dT = 0.02$ [5] (the maximum absolute Doppler shift $f_d = f_c v/c$ is applied), which is considered to be a fairly fast-fading scenario for mobile environments.

Fading processes are piecewise-constant approximated, allowing the channel coefficients to be constant in one *sampling interval* and change over different sampling intervals within one OFDM symbol period according to the correlation function described in Section III-B. There are many methods to generate the fading coefficients [27]–[30], and we adopt the one described in [28] since it gives a better autocorrelation property of the fading process than the one in [27].

The number of sub-iterations in the EM process, G, affects the receiver bit error rate (BER). Fig. 2 shows that, for the common set of system parameters chosen, BER does not further improve after $G = 9$ sub-iterations. Hence, in obtaining the rest of the simulation results, $G = 9$ with step sizes of $\beta_1 = \beta_2 = \ldots = \beta_{N_T} = 1/N_T$ will be adopted. Fig. 3 compares the error performance of receivers that employ the LS, the MAP, and the proposed scheme for $(2, 2)$ (i.e., $N_T = 2,\ N_R = 2$) and $(2, 3)$ MIMO-OFDM systems. Performance of the ideal case that assumes perfect CSI is used as the baseline performance. The proposed EM-MAP scheme performs almost the same as the normal MAP scheme; the performance degradation as a result of reducing

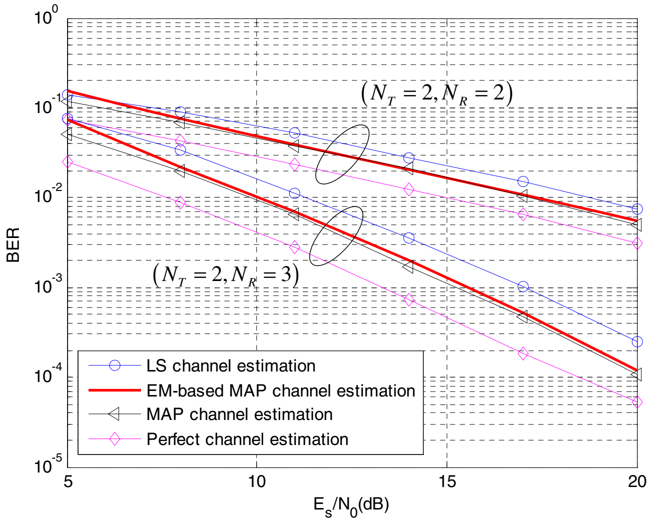

Fig. 3. Performance comparison of various channel estimation schemes ($f_d T = 0.02$, pilot subcarrier spacing $p_s = 4$).

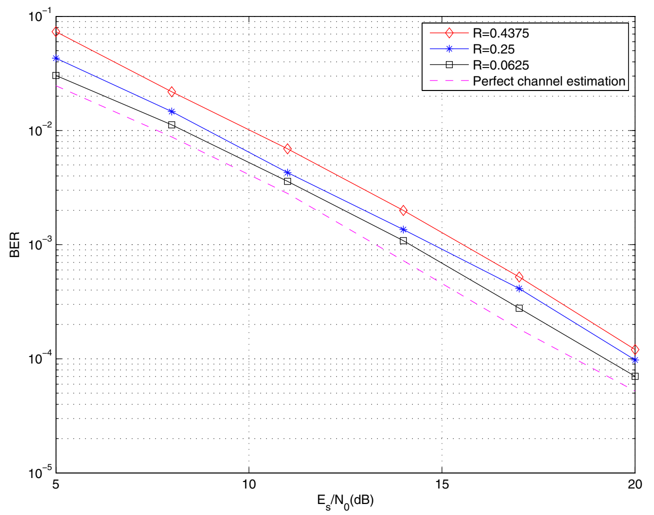

Fig. 4. The effect of the ratio of pilot spacing to channel coherence bandwidth ($N_T = 2,\ N_R = 3$).

complexity using (10) and the low-rank approximation in (12) is negligible. Compared with the LS estimator, the proposed scheme achieves an improvement of approximately 2 dB. For the cases simulated, the performance gap between the proposed scheme and the ideal case is within about 2 dB.

The effect of the ratio of pilot spacing to the channel coherent bandwidth on the BER performance of a $(2, 3)$ MIMO-OFDM system operating at a fading rate of $f_d T = 0.02$ is shown in Fig. 4. As expected, the lower the ratio, the better the performance. However, with the same channel scenario and system configuration, a lower value of R results in a lower spectral efficiency. The performance gap between the proposed scheme and the ideal case increases as R increases.

V. Conclusion

We have derived an EM-based MAP channel estimator with iterative estimation and detection for MIMO-OFDM systems, which works well in fast-fading channels. The estimator does not need to invert large-size matrices, resulting in a much lower complexity than existing schemes, while achieving the optimal performance of MAP channel estimation. The iterative process also enables one to incorporate approximation and cancellation of the ICI component for high-performance detection. Simulation results demonstrate the great robustness of the proposed scheme to fast time-varying fading. With a fading rate as high as $f_d T = 0.02$, the proposed scheme achieves an error performance that is 2 dB better than the LS estimator and is about 2 dB worse than the ideal case that assumes perfect channel estimates.

Appendix: Approximation of the Channel Correlation Matrix

Recall that $\bar{H}_{v,u}(k) = \frac{1}{N}\sum_{n=0}^{N-1} H_{v,u}(k,n)$ and $H_{v,u}(k,n) = \sum_{l=0}^{L} h_{v,u}(n,l)e^{-j2\pi kl/N}$. Applying Eq. (1), we obtain the (m,n)-th element of the correlation matrix $\boldsymbol{R}_H$ as

$$\begin{aligned}
[\boldsymbol{R}_H]_{m,n} &= E\left\{\bar{H}_{v,u}(m)\bar{H}^*_{v,u}(n)\right\} \\
&= \frac{1}{N^2}E\left\{\sum_{l_1=0}^{L}\sum_{n_1=0}^{N-1} e^{\frac{-j2\pi m l_1}{N}} h_{v,u}(n_1,l_1)\cdot \right. \\
&\qquad \left. \sum_{l_2=0}^{L}\sum_{n_2=0}^{N-1} h_{v,u}(n_2,l_2)e^{\frac{j2\pi n l_2}{N}}\right\} \\
&= \epsilon\sum_{l=0}^{L} e^{-l/L}e^{\frac{-j2\pi(m-n)l}{N}}\frac{1}{N^2}\cdot \\
&\qquad \sum_{n_1=0}^{N-1}\sum_{n_2=0}^{N-1} J_0\left(2\pi f_d T_s(n_2-n_1)\right). \qquad (19)
\end{aligned}$$

For most application scenarios, the normalized Doppler shift $f_d T$ is smaller than 0.05. In such a case, and noting that $T_s = T/N,\ \ J_0\left(2\pi f_d T_s(n_2-n_1)\right) \approx 1$. Thus

$$[\boldsymbol{R}_H]_{m,n} \approx \epsilon\sum_{l=0}^{L} e^{-l/L}e^{\frac{-j2\pi(m-n)l}{N}}. \qquad (20)$$

References

[1] H. Bolckei, D. Gesbert and A. J. Paulraj, "On the capacity of OFDM-based spatial multiplexing systems," *IEEE Trans. Commun.*, vol. 50, pp. 225–234, Feb. 2002.

[2] J. Akhtar and D. Gesbert, "A closed-form precoder for spatial multiplexing over correlated MIMO channels," in *IEEE Proc. Globecom'03*, vol. 4, pp. 1847–1851, Dec. 2003.

[3] D. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multi-element antenna systems," *IEEE Trans. Commun.*, vol. 48, pp. 502–513, Mar. 2000.

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

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

[6] B. Muquet and M. de Courville, "Blind and semi-blind channel identification methods using second order statistics for OFDM systems," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing* , vol. 5, pp. 2745–2748, Mar. 1999.

[7] M. C. Necker and G. L. Stuber, "Totally blind channel estimation for OFDM on fast varying mobile radio channels," *IEEE Trans. Wireless Commun.*, vol. 3, pp. 1514–1525, Sept. 2004.

[8] N. Chotikakamthorn and H. Suzuki, "On identifiability of OFDM blind channel estimation," in *Proc. IEEE Veh. Technol. Conf.*, vol. 4, pp. 2358–2361, Sept. 1999.

[9] D. Gesbert, H. Bolcskei, D. Gore, and A. Paulraj, "Outdoor MIMO wireless channels: Models and performance prediction," *IEEE Trans. Commun.*, vol. 50, pp. 1926–1934, Dec. 2002.

[10] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," *IEEE Trans. Broadcast.*, vol. 48, no. 3, pp. 223–229, Sept. 2002.

[11] T. L. Tung and K. Yao, "Channel estimation and optimal power allocation for a multiple-antenna OFDM system," *EURASIP J. Applied Signal Processing* , vol. 2002, pp. 330–339, Mar. 2002.

[12] D. Schafhuber, G. Matz, and F. Hlawatsch, "Kalman tracking of time-varying channels in wirelss MIMO-OFDM systems," in *Proc. 37th Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, vol. 2, pp. 1261–1265, Nov. 2003.

[13] K. Han, S. Lee, J. Lim, and K. Sung, "Channel estimation for OFDM with fast fading channels by modified Kalman filter," *IEEE Trans. Consum. Electron.*, vol. 50, pp. 443–449, May 2004.

[14] 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.

[15] G. L. Stuber, J. R. Barry, S. W. McLaughlin, Y. Li, M. A. Ingram, and T. G. Pratt, "Broadband MIMO-OFDM wireless communications," *Proc. IEEE*, vol. 92, pp. 271–294, Feb. 2004.

[16] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," *IEEE Trans. Signal Processing*, vol. 51, pp. 1615–1624, June 2003.

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

[18] Y. Xie and C. N. Georghiades, "Two EM-type channel estimation algorithms for OFDM with transmitter diversity," *IEEE Trans. Commun.*, vol. 51, pp. 106–115, Jan. 2003.

[19] T. Lv, H. Li, and J. Chen, "Joint estimation of symbol timing and carrier frequency offset of OFDM signals over fast time-varying multipath channels," *IEEE Trans. Signal Processing*, vol. 53, pp. 4526–4535, Dec. 2005.

[20] Y. Xie and C. N. Georghiades, "An EM-based channel estimation algorithm for OFDM with transmitter diversity," in *Proc. IEEE Globecom'01*, San Antonio, TX, vol. 2, pp. 871–875, Nov. 2002.

[21] H. Meyr, M. Moeneclaey, and S. A. Fechtel, *Digital Communication Receivers: Synchronization, Channel Estimation, and Signal Processing*. Wiley Series in Telecommuni. and Signal Processing, 1998.

[22] X. Deng, A. M. Haimovich, and J. Garcia-Frias, "Decision directed iterative channel estimation for MIMO systems," in *Proc. IEEE ICC'03*, vol. 4, pp. 2326–2329, May, 2003.

[23] M. R. Spiegel, *Theory and Problems of Probability and Statistics*. New York: McGraw-Hill, 1992.

[24] O. Edfors, M. Sandell, J. Beek, and S. Wilson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, pp. 931–939, July 1998.

[25] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Aanlysis*. Addison-Wesley, 1991.

[26] Y. Li, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *IEEE Trans. Veh. Technol.*, vol. 49, pp. 1207–1215, July 2000.

[27] W. C. Jakes, *Microwave mobile communications*. New York: John Wiley and Sons, Inc., 1974.

[28] P. Dent, G. E. Bottomley, and T. Croft, "Jakes fading model revisited," *Electron. Lett.*, vol. 29, pp. 1162–1663, June 1993.

[29] Y. Zheng and C. Xiao, "Simulation models with correct statistical properties for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 51, pp. 920–928, June 2003.

[30] Y. X. Li and X. Huang, "The simulation of independent Rayleigh faders," *IEEE Trans. Commun.*, vol. 50, pp. 1503–1514, Sept. 2002.