Technical University of Denmark

DTU

Low complexity variational bayes iterative reviver for MIMO-OFDM systems

Xiong, Chunlin; Wang, Hua; Zhang, Xiaoying; Wei, Jibo; Tang, Chaojing

Link back to DTU Orbit

DTU Library
Technical Information Center of Denmark

Low Complexity Variational Bayes Iterative Receiver for MIMO-OFDM Systems

Chun-lin Xiong[†], Hua Wang[◇], Xiao-ying Zhang[†], Ji-bo Wei[†], Chao-jing Tang[†]

[†]School of Electronic Science and Engineering, National University of Defense Technology, Changsha, China

[◇]Department of Communications, Optics & Materials Technical University of Denmark, Lyngby, Denmark

Email: [†]xchlzju@nudt.edu.cn, [◇]huw@com.dtu.dk

Abstract—A low complexity iterative receiver is proposed in this paper for MIMO-OFDM systems in time-varying multi-path channel based on the Variational Bayes (VB) method. According to the VB method, the estimation algorithms of the signal distribution and the channel distribution are derived for the receiver. With the aid of the soft-output QRD-M algorithm, whose complexity is fixed and relatively low, the signal distribution can be obtained conveniently. In particular, a sequential channel estimation algorithm, which completely avoids the computation of matrix inversion and multiplication, is introduced for the channel distribution estimation. Moreover, the distribution estimations of the signals and the channels are performed in a cyclical iteration way. The simulation results show that the performance loss of the proposed receiver is only 1dB for fast varying channels and less than 0.5dB for slow varying channels at the bit error rate of 10^{-4} after 3 iterations, compared with the optimum receiver with perfect channel state information.

I. INTRODUCTION

MIMO-OFDM offers a promising technique for broadband wireless communications [1]. The optimal receiver for coded MIMO-OFDM systems performs joint signal detection and channel estimation on the maximum a-posterior (MAP) criterion. Since the MAP receiver is infeasible due to the prohibitively high computational complexity [2], it is a challenging task to design a practical receiver.

Recently, the Variational Bayes (VB) method, which has been widely applied in research fields such as artificial intelligence, machine learning and statistical inference [3∼5], is employed as an effective pathway to the design of tractable signal processing algorithms for parametric inference in various communication systems [6∼10]. In [11], a VBEM iterative receiver was first derived for MIMO-OFDM systems. However, the channel estimation algorithms of [11] require intensive computation of matrix inversion and multiplication.

Based on the VB method, a practical iterative receiver with very low complexity is proposed for MIMO-OFDM systems in this paper. By factorizing of a joint distribution into the approximate marginal distributions, the VB method converts the joint distribution estimation of the parameters into the cyclic iterative estimations of the approximate marginal distributions [5]. Therefore, the proposed receiver performs the approximate marginal distribution estimations of the signals and the channels in a cyclical iteration way. With the aid of the soft-output QRD-M algorithm, whose complexity is fixed and relatively low, the signal distribution can be obtained conveniently. In particular, a sequential channel estimation
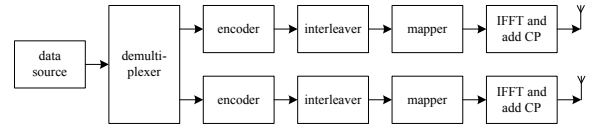

Fig. 1. Transmitter Scheme

algorithm, which completely avoids the computation of matrix inversion and multiplication, is presented for the channel distribution estimation. Simulation results demonstrate that the proposed low complexity receiver exhibits excellent performance.

The remainder of this paper is organized as follows. In section II, the system model and notations are introduced. The proposed low complexity iterative receiver is derived in section III. The simulation results are presented in section IV and some conclusions are drawn in section V.

Notations: $(\cdot)^*, (\cdot)^T, (\cdot)^H, \otimes$, and $E(\cdot)$ denote conjugate, transpose, Hermitian transpose, Kronecker product and statistical expectation, respectively. $\mathrm{diag}(\mathbf{X})$ represents a diagonal matrix with vector $\mathbf{X}$ on the main diagonal, and $\mathbf{I}_N$ is a $N \times N$ identity matrix.

II. SYSTEM DESCRIPTION

Let us consider a coded MIMO-OFDM system with N_t transmitters, N_r receivers and K subcarriers as shown in Fig.1. At the transmitter, serial data stream is demultiplexed into N_t parallel streams. Each stream is individually encoded, interleaved, and mapped into QAM symbols from a complex-valued finite alphabet set Ψ. And then, the QAM symbols are grouped into OFDM symbols, each of which is modulated by a K-point inverse discrete Fourier transform (IDFT). Finally, the modulated OFDM symbols are transmitted through the corresponding antenna after proper cyclical prefix (CP) insertion.

We assume that the channel is subject to the uncorrelated multipath Rayleigh fading and is spatially uncorrelated. The channel impulse response (CIR) between the t-th transmit antenna and r-th receive antenna at the n-th OFDM symbol is described as $\mathbf{h}_{r,t}(n) = [h_{r,t}(n,0), h_{r,t}(n,1), \cdots, h_{r,t}(n, L-1)]^T$ for $1 \le t \le N_t$ and $1 \le r \le N_r$, where L is the channel memory length. The channel is further assumed to be constant during an OFDM symbol period, but vary from one OFDM symbol to another. The time-variant behavior of the CIR $\mathbf{h}_{r,t}(n)$ is approximated by a first-order autoregressive (AR)

model [12]

$$\mathbf{h}_{r,t}(n) = J_0(2\pi f_d)\mathbf{h}_{r,t}(n-1) + \mathbf{v}(n), \quad (1)$$

where $J_0(\cdot)$ is the zero-order Bessel function of the first kind, f_d is the normalized Doppler frequency, and $\mathbf{v}(\mathbf{n})$ is the stimulus noise with covariance matrix $\mathbf{Q_v}$.

Let $\underline{\mathbf{X}}_t(n) = [X_t(n,0), X_t(n,1), \cdots, X_t(n, K-1)]^T$ denote the n-th frequency domain OFDM symbol at the t-th transmit antenna. With perfect time and frequency synchronization, a K-point discrete Fourier transform (DFT) is performed at the receiver after the removal of CP. Then, the n-th OFDM symbol at the r-th receive antenna $\underline{\mathbf{Y}}_r(n) = [Y_r(n,0), Y_r(n,1), \cdots, Y_r(n, K-1)]^T$ can be expressed as

$$\underline{\mathbf{Y}}_r(n) = \sum_{t=1}^{N_t} \text{diag}[\underline{\mathbf{X}}_t(n)]\mathbf{F}\mathbf{h}_{r,t}(n) + \underline{\mathbf{W}}_r(n). \quad (2)$$

Here, $\mathbf{F}$ is the first L columns of $K \times K$ DFT matrix, $\underline{\mathbf{W}}_r(n) = [W_r(n,0), W_r(n,1), \cdots, W_r(n, K-1)]^T$ represents the zero-mean complex additional Gaussian white noise (AWGN) vector with covariance matrix $\sigma^2\mathbf{I}_K$, at the r-th receive antenna. By collecting the n-th received OFDM symbols of all receive antennas into a $K \times N_r$ matrix $\mathbf{Y}(n) = [\underline{\mathbf{Y}}_1(n), \underline{\mathbf{Y}}_2(n), \cdots, \underline{\mathbf{Y}}_{N_r}(n)]$, equation (2) becomes

$$\mathbf{Y}(n) = \mathbf{D}(n)\mathbf{T}\mathbf{h}(n) + \mathbf{W}(n), \quad (3)$$

where $\mathbf{D}(n) = [\text{diag}(\underline{\mathbf{X}}_1(n)), \text{diag}(\underline{\mathbf{X}}_2(n)), \cdots, \text{diag}(\underline{\mathbf{X}}_{N_t}(n))]_{K \times KN_t}$, $\mathbf{W}(n) = [\underline{\mathbf{W}}_1(n), \underline{\mathbf{W}}_2(n), \cdots, \underline{\mathbf{W}}_{N_r}(n)]_{K \times N_r}$,

$$\mathbf{T} = \mathbf{I}_{N_t} \otimes \mathbf{F}, \quad \mathbf{h}(n) = \begin{bmatrix} \mathbf{h}_{1,1}(n) & \cdots & \mathbf{h}_{N_r,1}(n) \\ \vdots & \ddots & \vdots \\ \mathbf{h}_{1,N_t}(n) & \cdots & \mathbf{h}_{N_r,N_t}(n) \end{bmatrix}.$$

III. VARIATIONAL BAYES ITERATIVE RECEIVER

The Variational Bayes (VB) method of distribution approximation originates in statistical physics, in the area known as Mean Field Theory. The central idea of the VB method is to factorize the exact but intractable joint distribution into a product of tractable approximate marginal distributions, and the best such approximation is found by minimization of a Kullback-Leibler divergence (KLD) [5]. Let $P(\mathbf{Z}_1, \mathbf{Z}_2|\mathbf{Y})$ be the joint distribution of parameters $\mathbf{Z}_1$ and $\mathbf{Z}_2$ given observation $\mathbf{Y}$, $q(\mathbf{Z}_1)$ and $q(\mathbf{Z}_2)$ be the approximate marginal distributions, then the optimal $q(\mathbf{Z}_1)$ and $q(\mathbf{Z}_2)$, which satisfies $P(\mathbf{Z}_1, \mathbf{Z}_2|\mathbf{Y}) \approx q(\mathbf{Z}_1)q(\mathbf{Z}_2)$, can be obtained by minimizing the KLD defined as follows [5]

$$\begin{aligned} &KL[q(\mathbf{Z}_1)q(\mathbf{Z}_2)||P(\mathbf{Z}_1, \mathbf{Z}_2|\mathbf{Y})] \\ &= \int q(\mathbf{Z}_1)q(\mathbf{Z}_2) \ln \frac{q(\mathbf{Z}_1)q(\mathbf{Z}_2)}{P(\mathbf{Z}_1, \mathbf{Z}_2|\mathbf{Y})} d\mathbf{Z}_1 d\mathbf{Z}_2. \end{aligned} \quad (4)$$

Since equation (4) usually does not lead to a closed-form solution, it can be minimized in a cyclical iteration way. At the i-th iteration, the distribution $q(\mathbf{Z}_1)$ and $q(\mathbf{Z}_2)$ are updated with the following two steps [5]

$$\ln q^{(i)}(\mathbf{Z}_1) = E_{q^{(i-1)}(\mathbf{Z}_2)}[\ln P(\mathbf{Z}_1, \mathbf{Z}_2|\mathbf{Y})], \quad (5)$$

$$\ln q^{(i)}(\mathbf{Z}_2) = E_{q^{(i)}(\mathbf{Z}_1)}[\ln P(\mathbf{Z}_1, \mathbf{Z}_2|\mathbf{Y})]. \quad (6)$$
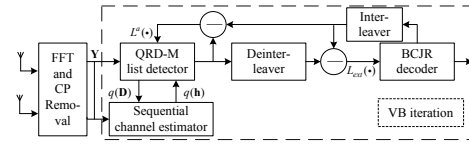

Fig. 2. Receiver Structure

The VB method provides a systematic approach to an iterative receiver design. In the considered context, the transmitted signal $\mathbf{D}(n)$ and the CIR $\mathbf{h}(n)$ are both treated as the parameters to be estimated. Since $\mathbf{D}(n)$ and $\mathbf{h}(n)$ are mutually independent, we have

$$\begin{aligned} &P(\mathbf{D}(n), \mathbf{h}(n)|\mathbf{Y}(n)) \\ &\propto P(\mathbf{Y}(n)|\mathbf{D}(n), \mathbf{h}(n))P(\mathbf{D}(n))P(\mathbf{h}(n)), \end{aligned} \quad (7)$$

where $P(\mathbf{D}(n))$ and $P(\mathbf{h}(n))$ is the prior distribution of $\mathbf{D}(n)$ and $\mathbf{h}(n)$. Based on the VB method, we derive the VB iterative receiver for MIMO-OFDM systems with joint signal detection and channel estimation, as shown in Fig.2. Since the derivation only uses one OFDM symbol, the time index n is dropped in the following. For simplicity, uncorrelated constants of all probability expressions are also omitted.

A. Signal Detection

Considering signal detection, we rewrite equation (7) as

$$P(\mathbf{D}, \mathbf{h}|\mathbf{Y}) \propto P(\mathbf{Y}|\mathbf{D}, \mathbf{h})P(\mathbf{D}). \quad (8)$$

By setting $\mathbf{Z}_1 = \mathbf{D}$, $\mathbf{Z}_2 = \mathbf{h}$, and substituting equation (8) into equation (5), we have

$$\ln Q^{(i)}(\mathbf{D}) = \ln P(\mathbf{D}) + E_{q^{(i-1)}(\mathbf{h})}[\ln P(\mathbf{Y}|\mathbf{D}, \mathbf{h})]. \quad (9)$$

In equation (9), $P(\mathbf{D})$ is approximated as

$$P(\mathbf{D}) = \prod_{k=0}^{K-1} P(\mathbf{X}_k) \propto \prod_{k=0}^{K-1} \exp\left(\frac{1}{2}\mathbf{B}_k^T L_a^{(i-1)}(\mathbf{B}_k)\right). \quad (10)$$

Here, $\mathbf{X}_k = [X_1(n,k), X_2(n,k), \cdots, X_{N_t}(n,k)]^T$ denotes the signals of all transmit antennas at the k-th subcarrier, $\mathbf{B}_k$ is the bit vector of $\mathbf{X}_k$, and $L_a^{(i-1)}(\mathbf{B}_k)$ is the prior information sent by the BCJR decoder [13]. Moreover, the term $\ln P(\mathbf{Y}|\mathbf{D}, \mathbf{h})$ in equation (9) can be derived as

$$\ln P(\mathbf{Y}|\mathbf{D}, \mathbf{h}) \propto -\frac{1}{\sigma^2}\sum_{k=0}^{K-1} \|\mathbf{Y}_k - \mathbf{H}_k\mathbf{X}_k\|^2 \quad (11)$$

with $\mathbf{Y}_k = [Y_1(n,k), Y_2(n,k), \cdots, Y_{N_r}(n,k)]^T$, $\mathbf{H}_k = (\mathbf{\Phi}_k\mathbf{h})^T$, and $\mathbf{\Phi}_k = \mathbf{I}_{N_t} \otimes \mathbf{u}_k^T\mathbf{F}$, where $\mathbf{u}_k$ represents the $K \times 1$ unit column vector with only one non-zero element at position k. Then, we can get

$$\begin{aligned} \ln Q^{(i)}(\mathbf{D}) = -\sum_{k=0}^{K-1} &\left(\frac{1}{\sigma^2}\left\|\mathbf{Y}_k - \bar{\mathbf{H}}_k^{(i-1)}\mathbf{X}_k\right\|^2 + \right. \\ &\left. \frac{1}{\sigma^2}\mathbf{X}_k^H\mathbf{\Omega}_k^{(i-1)}\mathbf{X}_k - \frac{1}{2}\mathbf{B}_k^T L_a^{(i-1)}(\mathbf{B}_k)\right). \end{aligned} \quad (12)$$

Here, $\bar{\mathbf{H}}_k^{(i-1)} = (\mathbf{\Phi}_k\bar{\mathbf{h}}^{(i-1)})^T$, $\mathbf{\Omega}_k^{(i-1)} = \mathbf{\Phi}_k E[(\mathbf{h} - \bar{\mathbf{h}}^{(i-1)})(\mathbf{h} - \bar{\mathbf{h}}^{(i-1)})^H]\mathbf{\Phi}_k^H$ and $\bar{\mathbf{h}}^{(i-1)}$ is the estimated mean

of $\mathbf{h}$ at the $(i-1)$-th iteration. Assuming that $Q^{(i)}(\mathbf{D})$ can be written as $Q^{(i)}(\mathbf{D}) = \prod_{k=0}^{K-1} Q^{(i)}(\mathbf{X}_k)$ with

$$
\ln Q^{(i)}(\mathbf{X}_k) = -\frac{1}{\sigma^2} \left\| \mathbf{Y}_k - \bar{\mathbf{H}}_k^{(i-1)} \mathbf{X}_k \right\|^2
$$
$$
- \frac{1}{\sigma^2} \mathbf{X}_k^H \mathbf{\Omega}_k^{(i-1)} \mathbf{X}_k + \frac{1}{2} \mathbf{B}_k^T L_a^{(i-1)}(\mathbf{B}_k). \tag{13}
$$

Once all $Q^{(i)}(\mathbf{X}_k)(k = 0, \cdots, K-1)$ are estimated, $Q^{(i)}(\mathbf{D})$ is obtained accordingly. Note that $\mathbf{X}_k$ takes discrete values, the distribution $q^{(i)}(\mathbf{X}_k)$ should be discrete-valued

$$
q^{(i)}(\mathbf{X}_k) = \sum_{c=1}^{|\Psi|^{N_t}} Q^{(i)}(\mathbf{X}_k^{(c)}) \delta(\mathbf{X}_k - \mathbf{X}_k^{(c)}), \tag{14}
$$

where $\mathbf{X}_k^{(c)}$ is the possible transmit vector specified by index c, $|\Psi|$ denotes the cardinality of set Ψ, and $\delta(\cdot)$ is the Dirac delta function.

Since there are $|\Psi|^{N_t}$ possible combinations of $\mathbf{X}_k$, the computation complexity of evaluating $q^{(i)}(\mathbf{X}_k)$ is prohibitively high for large systems. In practice, we only concentrate on a small subset of the signal space Ψ^{N_t}, i.e. candidate list. Instead of using the list sphere decoder [11], we introduce the QRD-M tree search algorithm with fixed and relatively low complexity as the list detector [14]. In the QRD-M algorithm, the path metric $PM(\mathbf{X}_k)$ is defined as

$$
PM(\mathbf{X}_k) = \frac{1}{\sigma^2} \left\| \mathbf{Y}_k - \bar{\mathbf{H}}_k^{(i-1)} \mathbf{X}_k \right\|^2
$$
$$
+ \frac{1}{\sigma^2} \mathbf{X}_k^H \mathbf{\Omega}_k^{(i-1)} \mathbf{X}_k - \frac{1}{2} \mathbf{B}_k^T L_a^{(i-1)}(\mathbf{B}_k). \tag{15}
$$

The list $C(\mathbf{X}_k) = \left\{ \mathbf{X}_k^{(c)} | c = 1, 2, \cdots, M \right\}$ of size $M(M < |\Psi|^{N_t})$ is then identified by M paths with minimum metrics, and $q^{(i)}(\mathbf{X}_k)$ can be expressed as

$$
q^{(i)}(\mathbf{X}_k) \approx \sum_{c=1}^{M} P_c \delta(\mathbf{X}_k - \mathbf{X}_k^{(c)}), \tag{16}
$$

$$
P_c = \frac{Q^{(i)}(\mathbf{X}_k^{(c)})}{\sum_{c=1}^{M} Q^{(i)}(\mathbf{X}_k^{(c)})} = \frac{\exp(-PM(\mathbf{X}_k^{(c)}))}{\sum_{c=1}^{M} \exp(-PM(\mathbf{X}_k^{(c)}))}. \tag{17}
$$

At the same time, with max-log approximation, the extrinsic information $L_{ext}(b_j)$ sent to the BCJR decoder can be derived as

$$
L_{ext}(b_j) \approx \max_{\mathbf{X}_k \in C_{j,-1}(\mathbf{X}_k)} \{ PM(\mathbf{X}_k) + \frac{1}{2} b_j L_a^{(i-1)}(b_j) \},
$$
$$
- \max_{\mathbf{X}_k \in C_{j,+1}(\mathbf{X}_k)} \{ PM(\mathbf{X}_k) + \frac{1}{2} b_j L_a^{(i-1)}(b_j) \}, \tag{18}
$$

where b_j is the j-th bit of $\mathbf{B}_k$ and $C_{j,+1}(\mathbf{X}_k)(C_{j,-1}(\mathbf{X}_k))$ denotes the subset of the candidates with $b_j = +1(b_j = -1)$ in the candidate list.

B. Channel Estimation

To begin with, we first rewrite equation (7) as

$$
P(\mathbf{D}, \mathbf{h} | \mathbf{Y}) \propto P(\mathbf{Y} | \mathbf{D}, \mathbf{h}) P(\mathbf{h}). \tag{19}
$$

And then, based on equation (6), given $q^{(i)}(\mathbf{X}_k)$, we can update $q^{(i)}(\mathbf{h})$ through the following form

$$
\ln q^{(i)}(\mathbf{h}) = \ln P(\mathbf{h}) + E_{q^{(i)}(\mathbf{D})} \left[\ln P(\mathbf{Y} | \mathbf{D}, \mathbf{h}) \right]
$$
$$
= \ln P(\mathbf{h}) - \frac{1}{\sigma^2} \sum_{k=0}^{K-1} E_{q^{(i)}(\mathbf{X}_k)} \left[\| \mathbf{Y}_k - \mathbf{H}_k \mathbf{X}_k \|^2 \right]. \tag{20}
$$

Without any prior knowledge about the channel, the prior distribution $P(\mathbf{h})$ is decoupled and approximated as

$$
\ln P(\mathbf{h}) \approx -\sum_{r=1}^{N_r} (\mathbf{h}_r - \bar{\mathbf{h}}_r^{(i-1)})^H (\mathbf{R}_{h_r}^{(i-1)})^{-1} (\mathbf{h}_r - \bar{\mathbf{h}}_r^{(i-1)}), \tag{21}
$$

where $\mathbf{h}_r$ is the r-th column of matrix $\mathbf{h}$, $\bar{\mathbf{h}}_r^{(i-1)}$ is the estimated mean of $\mathbf{h}_r$ and $\mathbf{R}_{h_r}^{(i-1)}$ is the corresponding error covariance matrix at iteration $(i-1)$. Based on equation (16), the second term $E_{q^{(i)}(\mathbf{X}_k)} \left[\| \mathbf{Y}_k - \mathbf{H}_k \mathbf{X}_k \|^2 \right]$ of equation (20) can be expanded as

$$
E_{q^{(i)}(\mathbf{X}_k)} \left[\| \mathbf{Y}_k - \mathbf{H}_k \mathbf{X}_k \|^2 \right]
$$
$$
= \sum_{r=1}^{N_r} \sum_{c=1}^{M} \left| Y_r(n, k) - (\mathbf{X}_k^{(c)T}) \mathbf{\Phi}_k \mathbf{h}_r \right|^2 P_c. \tag{22}
$$

Using equation (21), (22) in (20), we have

$$
\ln q^{(i)}(\mathbf{h}) =
$$
$$
- \sum_{r=1}^{N_r} \left\{ (\mathbf{h}_r - \bar{\mathbf{h}}_r^{(i-1)})^H (\mathbf{R}_{h_r}^{(i-1)})^{-1} (\mathbf{h}_r - \bar{\mathbf{h}}_r^{(i-1)}) \right.
$$
$$
\left. + \sum_{k=0}^{K-1} \sum_{c=1}^{M} \frac{1}{\sigma^2} \left| Y_r(n, k) - \mathbf{X}_k^{(c)T} \mathbf{\Phi}_k \mathbf{h}_r \right|^2 P_c \right\}. \tag{23}
$$

Under the assumption of independent channels, $\ln q^{(i)}(\mathbf{h})$ can be written as $\ln q^{(i)}(\mathbf{h}) = \sum_{r=1}^{N_r} \ln q^{(i)}(\mathbf{h}_r)$ with

$$
\ln q^{(i)}(\mathbf{h}_r) = -(\mathbf{h}_r - \bar{\mathbf{h}}_r^{(i-1)})^H (\mathbf{R}_{h_r}^{(i-1)})^{-1} (\mathbf{h}_r - \bar{\mathbf{h}}_r^{(i-1)})
$$
$$
- \sum_{k=0}^{K-1} \sum_{c=1}^{M} \frac{1}{\sigma^2} \left| Y_r(n, k) - \mathbf{X}_k^{(c)T} \mathbf{\Phi}_k \mathbf{h}_r \right|^2 P_c. \tag{24}
$$

Note that we can independently estimate the distribution of each column of matrix $\mathbf{h}$. Solving the zero-gradient point of equation (24) with respect to $\mathbf{h}_r$ yields

$$
\bar{\mathbf{h}}_r^{(i)} = \mathbf{R}_{h_r}^{(i)} \left[(\mathbf{R}_{h_r}^{(i-1)})^{-1} \bar{\mathbf{h}}_r^{(i-1)} \right.
$$
$$
\left. + \frac{1}{\sigma^2} \sum_{k=0}^{K-1} \sum_{c=1}^{M} \mathbf{\Phi}_k^H \mathbf{X}_k^{(c)*} P_c Y_r(n, k) \right], \tag{25}
$$

$$
(\mathbf{R}_{h_r}^{(i)})^{-1} = (\mathbf{R}_{h_r}^{(i-1)})^{-1}
$$
$$
+ \frac{1}{\sigma^2} \sum_{k=0}^{K-1} \sum_{c=1}^{M} \mathbf{\Phi}_k^H \mathbf{X}_k^{(c)*} \mathbf{X}_k^{(c)T} \mathbf{\Phi}_k P_c, \tag{26}
$$

and the distribution of $\mathbf{h}$ at iteration i is obtained as

$$
\ln q^{(i)}(\mathbf{h}) = -\sum_{r=1}^{N_r} (\mathbf{h}_r - \bar{\mathbf{h}}_r^{(i)})^H (\mathbf{R}_{h_r}^{(i)})^{-1} (\mathbf{h}_r - \bar{\mathbf{h}}_r^{(i)}). \tag{27}
$$

Finally, the mean and the covariance of the CIR in time domain are converted to frequency domain for signal detection, i.e. $\bar{\mathbf{H}}_k^{(i)} = (\mathbf{\Phi}_k \bar{\mathbf{h}}^{(i)})^T$ and $\mathbf{\Omega}_k^{(i)} = \mathbf{\Phi}_k (\sum_{r=1}^{N_r} \mathbf{R}_{h_r}^{(i)}) \mathbf{\Phi}_k^H$. The final estimation of CIR at the $(n-1)$-th OFDM symbol duration is exploited as the initial values at the n-th symbol duration.

C. Low Complexity Sequential Channel Estimation

Note that a $LN_t \times LN_t$ matrix inversion is required in equation (25) and (26) to solve the channel estimation problem. To reduce the computational complexity, a low complexity sequential channel estimation (LCSCE) algorithm is introduced based on equation (25) and (26).

We first approximate equation (24) as

$$
\ln q^{(i)}(\mathbf{h}_r) \approx -\sum_{k=0}^{K-1} \frac{1}{\sigma^2} \left| Y_r(n,k) - \bar{\mathbf{X}}_k^T \mathbf{\Phi}_k \mathbf{h}_r \right|^2 \\
- (\mathbf{h}_r - \bar{\mathbf{h}}_r^{(i-1)})^H (\mathbf{R}_{h_r}^{(i-1)})^{-1} (\mathbf{h}_r - \bar{\mathbf{h}}_r^{(i-1)}),
\tag{28}
$$

where $\bar{\mathbf{X}}_k = \sum_{c=1}^M \mathbf{X}_k^{(c)} P_c$. This is reasonable in high signal-to-noise ratios (SNR), as the candidate with minimum metric almost dominates in the candidate list. With the definition of partial sum $\mathbf{\Delta}_r^{(i)}(m)$ and $\mathbf{U}_r^{(i)}(m) (m=0,\cdots,K-1)$ through

$$
\mathbf{\Delta}_r^{(i)}(m) = (\mathbf{R}_{h_r}^{(i-1)})^{-1} + \frac{1}{\sigma^2} \sum_{k=0}^m \mathbf{\Phi}_k^H \bar{\mathbf{X}}_k^* \bar{\mathbf{X}}_k^T \mathbf{\Phi}_k,
\tag{29}
$$

$$
\mathbf{U}_r^{(i)}(m) = (\mathbf{R}_{h_r}^{(i-1)})^{-1} \bar{\mathbf{h}}_r^{(i-1)} + \sum_{k=0}^m \frac{\mathbf{\Phi}_k^H \bar{\mathbf{X}}_k^* Y_r(n,k)}{\sigma^2},
\tag{30}
$$

we can get the following two recursive equations

$$
\mathbf{\Delta}_r^{(i)}(m) = \mathbf{\Delta}_r^{(i)}(m-1) + \frac{1}{\sigma^2} \mathbf{\Phi}_m^H \bar{\mathbf{X}}_m^* \bar{\mathbf{X}}_m^T \mathbf{\Phi}_m,
\tag{31}
$$

$$
\mathbf{U}_r^{(i)}(m) = \mathbf{U}_r^{(i)}(m-1) + \frac{\mathbf{\Phi}_m^H \bar{\mathbf{X}}_m^* Y_r(n,m)}{\sigma^2}.
\tag{32}
$$

Let's define

$$
\gamma_r^{(i)}(m) = \mathbf{\Delta}_r^{(i)}(m)^{-1} \mathbf{U}_r^{(i)}(m).
\tag{33}
$$

Substituting equation (31) and (32) into (33), after some manipulations with the matrix inversion lemma [15], we have

$$
\mathbf{G}_r^{(i)}(m) = \frac{\mathbf{\Xi}_r^{(i)}(m-1) \mathbf{\Phi}_m^H \bar{\mathbf{X}}_m^*}{\sigma^2 + \bar{\mathbf{X}}_m^T \mathbf{\Phi}_m \mathbf{\Xi}_r^{(i)}(m-1) \mathbf{\Phi}_m^H \bar{\mathbf{X}}_m^*},
\tag{34}
$$

$$
\mathbf{\Xi}_r^{(i)}(m) = (\mathbf{I}_{LN_t} - \mathbf{G}_r^{(i)}(m) \bar{\mathbf{X}}_m^T \mathbf{\Phi}_m) \mathbf{\Xi}_r^{(i)}(m-1),
\tag{35}
$$

$$
\gamma_r^{(i)}(m) = \gamma_r^{(i)}(m-1) + \\
\mathbf{G}_r^{(i)}(m)(Y_r(n,m) - \bar{\mathbf{X}}_m^T \mathbf{\Phi}_m \gamma_r^{(i)}(m-1)),
\tag{36}
$$

where $\mathbf{\Xi}_r^{(i)}(m) = \mathbf{\Delta}_r^{(i)}(m)^{-1} (m=0,\cdots,K-1)$. Note that equation (33) with $m = K-1$ is the zero-gradient point of equation (28), which means $\bar{\mathbf{h}}_r^{(i)} = \gamma_r^{(i)}(K-1)$ and $\mathbf{R}_{h_r}^{(i)} = \mathbf{\Xi}_r^{(i)}(K-1)$. Performing the LCSCE algorithm on all receive antennas, we can obtain the channel distribution $q^{(i)}(\mathbf{h})$.

As we can see from equations (34) $\sim$ (36), matrix inversion is totally avoided in the LCSCE algorithm. However, matrix multiplications are still required. In order to further reduce the computational complexity, we replace $\mathbf{\Xi}_r^{(i)}(m-1)$ by a

diagonal matrix $\mathbf{\Theta}_r^{(i)}(m-1)$, which can be directly obtained just by setting all off-diagonal elements of $\mathbf{\Xi}_r^{(i)}(m-1)$ to zero. Accordingly, we rewrite equation (34) and (35) as

$$
\mathbf{G}_r^{(i)}(m) = \frac{\mathbf{\Theta}_r^{(i)}(m-1) \mathbf{\Phi}_m^H \bar{\mathbf{X}}_m^*}{\sigma^2 + \bar{\mathbf{X}}_m^T \mathbf{\Phi}_m \mathbf{\Theta}_r^{(i)}(m-1) \mathbf{\Phi}_m^H \bar{\mathbf{X}}_m^*},
\tag{37}
$$

$$
\mathbf{\Theta}_r^{(i)}(m) = \mathbf{\Theta}_r^{(i)}(m-1) \times \\
\left[\mathbf{I}_{LN_t} - \mathrm{diag}\left(\mathbf{G}_r^{(i)}(m) \right) \mathrm{diag}\left(\bar{\mathbf{X}}_m^T \mathbf{\Phi}_m \right) \right].
\tag{38}
$$

Since $\mathbf{\Theta}_r^{(i)}(m-1)$ is a diagonal matrix, $\mathbf{\Theta}_r^{(i)}(m)$ and $\mathbf{\Phi}_m \mathbf{\Theta}_r^{(i)}(m-1) \mathbf{\Phi}_m^H$ both are diagonal. Therefore, we completely avoid matrix inversion and matrix multiplication in the proposed sequential channel estimation algorithm, which significantly reduces the computational complexity. In addition, we can see that matrix $\mathbf{\Omega}_k^{(i)} = \mathbf{\Phi}_k (\sum_{r=1}^{N_r} \mathbf{R}_{h_r}^{(i)}) \mathbf{\Phi}_k^H$ is also diagonal. As a result, the additional complexity, introduced by considering the channel estimation error in data detection, is quite small.

IV. SIMULATION RESULTS

In this section, we illustrate the performance of the proposed algorithm by simulating a MIMO-OFDM system with $N_t = 2$ transmit antennas, $N_r = 2$ receive antennas, and $K = 64$ subcarriers. The length of CP is set to be 16. The channel encoder is a rate $1/2$ convolutional encoder with generators 7 and 5 in octal notation and the coded bits are mapped into QPSK symbols. Each transmission frame has 11 OFDM symbols, where the first one contains training symbols designed in [16] for initial channel estimation. We use $L = 3$ paths channels with $f_d = 0.02$ and $f_d = 0.002$ in the simulation. Furthermore, the channels are assumed to have exponential power delay profile similar to [16].

Fig.3 and 4 show the normalized MSE (NMSE) performance of the proposed channel estimation algorithms versus SNR with $f_d = 0.02$ and $f_d = 0.002$ respectively. The bit-error-rate (BER) performance is depicted in Fig.5 and 6. In these figures, "INVCE" represents the proposed channel estimation algorithm (or receiver) associated with equations (25) and (26), while "Optimal" denotes the optimal MAP receiver with perfect channel state information. "Iter" is the index of the iteration.

As shown in Fig.3 and 4, the INVCE algorithm slightly outperforms the LCSCE algorithm in low SNR region. However, the NMSE gap between the two algorithms decreases dramatically when the SNR increases.

In Fig.5 and 6, we can see that, in both relatively fast ($f_d = 0.02$, see Fig.5) and low ($f_d = 0.002$, see Fig.6) varying multi-path environment, the receivers with the two proposed channel estimation algorithms achieve almost the same performance at the same iteration number. Compared with the optimal receiver, only about 1 dB performance loss of the proposed receivers is observed in fast varying channels after 3 iterations, at BER $= 10^{-4}$, while less than 0.5 dB loss in slow varying channels. In addition, performance gain is obtained from the VB iterations. From these results, we can
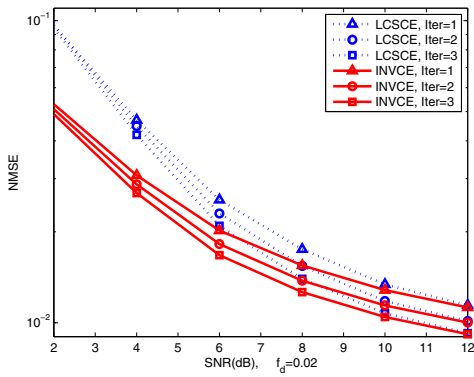
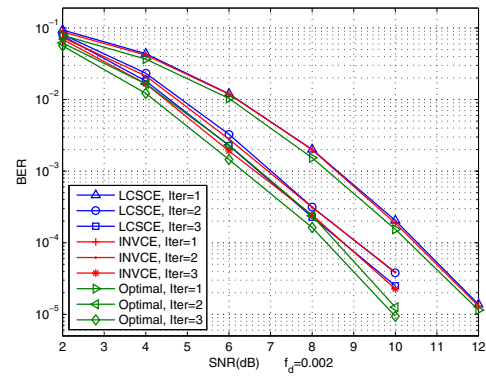Fig. 3. NMSE of the proposed channel estimation algorithms versus SNR ($f_d = 0.02$).
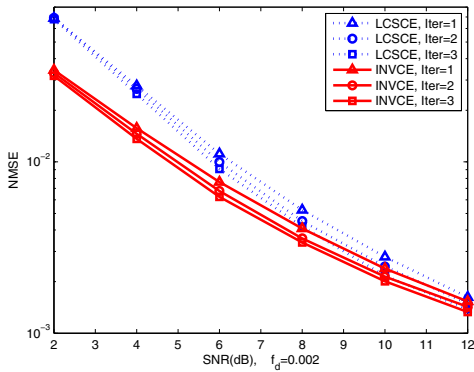

Fig. 4. NMSE of the proposed channel estimation algorithms versus SNR ($f_d = 0.002$).
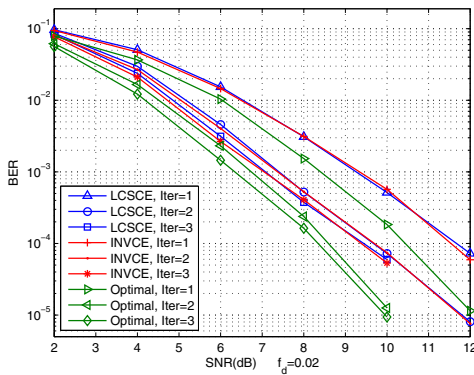

Fig. 5. BER performance of the proposed receivers and the optimal receiver versus SNR ($f_d = 0.02$).

conclude that the proposed iterative receiver exhibits excellent BER performance with extremely low complexity.

V. CONCLUSION

This paper proposes a low complexity iterative receiver for MIMO-OFDM systems based on the Variational Bayes method. Both the signal detection algorithm and the channel estimation algorithm are derived. By making some appropriate approximations, we totally avoid matrix inversion and matrix multiplication in the proposed low complexity channel estima-

Fig. 6. BER performance of the proposed receivers and the optimal receiver versus SNR ($f_d = 0.002$).

tion algorithm. Simulation results show that the proposed low complexity receiver can achieve near optimal performance, thus is quite feasible in practical systems.

REFERENCES

[1] Hong-Wei Yang, "A Road to Future Broadband Wireless Access: MIMO-OFDM Based Air Interface," *IEEE Commun. Magazine*, 43(1):53-60, 2005.
[2] T. Kashima, K. Fukawa, and H. Suzuki, "Adaptive MAP Receiver via the EM Algorithm and Message Passing for MIMO-OFDM Mobile Communications," *IEEE J. Select. Areas Commun.*, 24(3):437-447, 2006.
[3] David J.C. Mackay, Information Theory, Inference, and Learning Algorithms. Cambridge University Press, 2003.
[4] M. Beal, Variational Algorithms for Approximate Bayesian Inference. Ph.D. thesis, Gatsby Computational Neuroscience Unit, London's Global University, UK, 2003.
[5] V. Smidl and A. Quinn, The Variational Bayes Method in Signal Processing. Springer, 2006.
[6] L.P.B. Christensen, J. Larsen, "On Data and Parameter Estimation Using the Variational Bayesian EM-Algorithm for Block-Fading Frequency-Selective MIMO Channels," *IEEE ICASSP 2006*, 4:465-468.
[7] B. Hu, I. Land, R. Piton, and B.H. Fleury, "A Bayesian Framework for Iterative Channel Estimation and Multiuser Decoding in Coded DS-CDMA," *IEEE GLOBECOM 2007*, pp.1582-1586.
[8] D.D. Lin and T.J. Lim, "The Variational Inference Approach to Joint Data Detection and Phase Noise Estimation in OFDM," *IEEE Trans. on Signal Processing*, 55(1):1862-1874, 2007.
[9] Stefanatos, K. Stelios, and K. Aggelos, "Joint Data Detection and Channel Tracking for OFDM Systems Using the Variational Bayes Method," *IEEE SPAWC 2007*, pp.1-5.
[10] N. Mauri and P. Subbarayan, "Iterative Receivers for Digital Communications via Variational Inference and Estimation," Ph.D. thesis, Department of Electrical and Information Engineering, INFORTECH OULU, University of OULU, Finland, 2008.
[11] Xiao-Ying Zhang, De-gang Wang and Ji-Bo Wei, "Joint Symbol Detection and Channel Estimation for MIMO-OFDM Systems via Variational Bayes EM algorithm," *IEEE WCNC 2008*, pp.13-17.
[12] T. Y . Al-Naffouri, "An EM Based Forward-Backward Kalman Filter for the Estimation of Time Variant Channels in OFDM," *IEEE Trans. on Signal Processing*, 55(7):3924-3930, 2007.
[13] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate," *IEEE Trans. On Information Theory*, vol. 20, pp. 284-287, 1974.
[14] K.J. Kim, J. Yue, R.A .Iltis and J. D .Gibson, "A QRD-M/Kalman Filter Based Detection and Channel Estimation Algorithm for MIMO-OFDM Systems," *IEEE Trans. on Wireless Commun.*, 3(2):710-721, 2005.
[15] Zhang Xianda. Matrix Analysis and Applications. Tsinghua University Press, Beijing, 2004.
[16] H. Minn and N. Al-Dhahir, "Optimal Training Signals for MIMO OFDM Channel Estimation," *IEEE Trans. on Wireless Commun.*, 5(5):1158-1168, 2006.