Joint Channel Estimation and Data Detection in MIMO-OFDM Systems Using Sparse Bayesian Learning

Ranjitha Prasad and Chandra R. Murthy
Department of ECE, Indian Institute of Science, B'lore, India.
Email:{ranjitha.p, cmurthy}@ece.iisc.ernet.in

***Abstract*—The impulse response of wireless channels between each transmit and receive antenna in a MIMO-OFDM system is known to be approximately sparse, in the sense that it has a small number of significant components relative to the channel delay spread. Moreover, it is known that the channel impulse responses in a MIMO-OFDM system are approximately group-sparse (a-group-sparse), i.e., the time-lags of the significant paths of channel impulse response between every transmit and receive antenna pair coincide. Accordingly, we cast the problem of estimating the a-group-sparse channels in the Bayesian framework, and propose novel algorithms that employ the multiple measurement vectors at the N_r receive antennas. First, we adapt the known MSBL algorithm for pilot-based a-group-sparse channel estimation in MIMO-OFDM systems. Subsequently, we generalize the MSBL algorithm to obtain a novel J-MSBL algorithm for joint a-group-sparse channel estimation and data detection. We illustrate the efficacy of the proposed techniques in terms of the mean square error and coded bit error rate performance using Monte Carlo simulations.**

***Index Terms*—Joint channel estimation and data detection, Group sparsity, MIMO, OFDM, Sparse Bayesian learning, Expectation maximization.**

I. Introduction

Multiple Input Multiple Output (MIMO) combined with Orthogonal Frequency Division Multiplexing (OFDM) is a key technology for several current and future broadband wireless systems and standards. In a MIMO-OFDM system, multiple antennas are used to exploit the diversity and multiplexing advantages of MIMO, while OFDM provides resilience to frequency-selective fading commonly encountered in a multi-path wireless environment [1]. Typically, pilot-based wireless channel estimation is employed at the MIMO-OFDM receiver to accurately decode the transmitted data bits. However, pilot-based methods in MIMO-OFDM necessitate the transmission of known pilots symbols on a set of subcarriers from each transmit antenna, resulting in a significant overhead on the bit rate. In this paper, we propose novel MIMO-OFDM channel estimation techniques using far fewer pilots compared to the conventional methods [2], [3].

In complex baseband representation, the channel impulse response between the $n_t^{\text{th}}, 0 \leq n_t \leq N_t$ transmit antenna and the $n_r^{\text{th}}, 0 \leq n_r \leq N_r$ receive antenna, denoted as $\tilde{h}_{n_t n_r}[t], t \in \mathbb{R}$, can be modeled as a stationary tapped delay line filter in the lag-domain:

$$\tilde{h}_{n_t n_r}[t] = \sum_{l=1}^{\tilde{L}} \tilde{h}_{n_t n_r, l} \delta[t - \tau_l], \tag{1}$$

where $\delta[t]$ is the Dirac delta function, $\tilde{h}_{n_t n_r, l}$ and τ_l represent the attenuation and propagation delay on the n_t^{th} transmitter and the n_r^{th} receiver path l, respectively, and $\tilde{L}$ is the number of resolvable paths [4]. It is known that the wireless channel models obtained using channel sounding experiments exhibit approximate sparsity in the lag-domain as the communication bandwidth increases, for e.g., due to the non-perfect low-pass raised cosine filtering [5]. Based on these practical considerations, we model the lag-domain filtered channel impulse response as $h_{n_t n_r}[t] = g_t[t] * \tilde{h}_{n_t n_r}[t] * g_r[t]$, where $g_t[t]$ and $g_r[t]$ represent the baseband transmit and receive filters employed at the every transmit and receive antenna of the MIMO-OFDM system, and $*$ represents the convolution operation. Then, the corresponding discrete-time channel can be represented as $h_{n_t n_r}(l) = h_{n_t n_r}[(l-1)T]$, where T is the baud interval. The overall channel is represented as $\mathbf{h}_{n_t n_r} = [h_{n_t n_r}(1), h_{n_t n_r}(2), \ldots, h_{n_t n_r}(L)]^T$. In addition, it is known that the sample-spaced representation of $\tilde{h}_{n_t n_r}[t]$ between different transmit and receive antenna pairs are group-sparse [6], [7], i.e., the locations of non-zero elements of the sparse vectors coincide. Since we assume that $g_t[t]$ and $g_r[t]$ are identical for every transmit and receive antenna, we deduce that the locations of the significant components in $\mathbf{h}_{n_t n_r}$ also coincide across the entire MIMO-OFDM system, and hence, we refer to $\{\mathbf{h}_{n_t n_r}\}, 0 \leq n_t \leq N_t, 0 \leq n_r \leq N_r$ as *approximately group-sparse* (a-group-sparse) channels.

In MIMO-OFDM systems, conventional techniques estimate the channel frequency response corresponding to the pilot subcarriers using frequency domain Least Squares (LS) or Minimum Mean Square Error (MMSE) based methods [3], and then, interpolate this to obtain the channel frequency response corresponding to the data subcarriers [2]. However, such schemes do not provide reliable estimates when the number of pilots P_b is smaller than the length of the cyclic prefix L ($P_b < L$). Further, lag-domain LS and MMSE [3] require prior knowledge of the average multipath power profile

This work was supported in part by a research grant from the Indo-UK Advanced Technology Centre and a research scholarship from the Tata Consultancy Services Research Scholar Program.

978-1-4799-2361-8/14/$31.00

measured at a particular location, also called as the Multipath Intensity Profile (MIP) of the channel [8]. A natural question arises: Is it possible to estimate the channel in the Bayesian framework using fewer pilots compared to conventional methods, especially if the MIP is not known? In the following subsection, we describe the system model, the a-group-sparse channel estimation problem, and answer the above question in the affirmative.

A. Problem Formulation and Contributions

In this subsection, we present the basic set-up of the coded MIMO-OFDM system considered in this work. We formulate the problem of a-group-sparse channel estimation, and briefly describe the contributions of this work.

We assume that the transmissions between the N_t transmit antennas and the N_r receive antennas take place through OFDM frames, where every frame consists of K OFDM symbols. Further, we consider the block-fading case, where the channel coefficients remain fixed across the OFDM frame duration and vary in an i.i.d. fashion from frame to frame. The discrete-time MIMO-OFDM system with N subcarriers is shown in Fig. 1, where the input bits $\{b\}$ are first encoded and interleaved into a new sequence of coded bits, $\{c\}$. Further, $\{c\}$ is mapped into an M-ary complex symbol sequence, which is divided into N_t streams. At every transmit antenna, P_b pilots are inserted in an OFDM frame. The pilot symbols along with data symbols $\{c\}$ are OFDM modulated and transmitted over the multipath fading channel. After OFDM demodulation, the signal received at the n_r^{th} receive antenna is given by

$$\mathbf{y}_{n_r} = \sum_{n_t=1}^{N_t} \mathbf{X}_{n_t}\mathbf{F}\mathbf{h}_{n_t n_r} + \mathbf{v}_{n_r}, \quad n_r = 1, \ldots, N_r, \qquad (2)$$

where diagonal matrices $\mathbf{X}_{n_t} \in \mathbb{C}^{N\times N}$ consists of the pilot as well as data and $\mathbf{F} \in \mathbb{C}^{N\times L}$ represents the matrix consisting of the first L columns of the $N \times N$ DFT matrix. Further, the multipath fading channel is denoted by $\mathbf{h}_{n_t n_r} \in \mathbb{C}^{L\times 1}$, as described in the previous subsection. Each component of $\mathbf{v}_{n_r} \in \mathbb{C}^{N\times 1}$ is an additive white circularly symmetric Gaussian noise with probability distribution $\mathcal{CN}(0, \sigma^2)$.

To recover the a-group-sparse channels, we cast (2) in a Multiple Measurement Vector (MMV) framework [9], [10]. Here, the observations from the N_r receivers form the observation matrix, $\mathbf{Y}$, which is related to the a-group-sparse vectors in the channel matrix, $\mathbf{H}$, through a common dictionary $\mathbf{\Phi}$. Accordingly, the a-group-sparse formulation in the MMV framework is as follows:

$$\underbrace{[\mathbf{y}_1, \ldots, \mathbf{y}_{N_r}]}_{\mathbf{Y}\in\mathbb{C}^{N\times N_r}} = \underbrace{\mathbf{X}(\mathbf{I}_{N_t} \otimes \mathbf{F})}_{\mathbf{\Phi}\in\mathbb{C}^{N\times LN_t}} \underbrace{\begin{pmatrix} \mathbf{h}_{11} & \ldots & \mathbf{h}_{1N_r} \\ \vdots & \vdots & \\ \mathbf{h}_{N_t 1} & \ldots & \mathbf{h}_{N_t N_r} \end{pmatrix}}_{\mathbf{H}\in\mathbb{C}^{LN_t\times N_r}} + \underbrace{[\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{N_r}]}_{\mathbf{V}\in\mathbb{C}^{N\times N_r}}, \qquad (3)$$

where the overall transmit data matrix $\mathbf{X} \in \mathbb{C}^{N\times NN_t}$ is given by $\mathbf{X} \triangleq [\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_{N_t}]$. Considering the pilot subcarriers, the MIMO-OFDM system model can be written as

$$\mathbf{Y}_p = \mathbf{\Phi}_p\mathbf{H} + \mathbf{V}, \qquad (4)$$

where $\mathbf{Y}_p = [\mathbf{y}_{p,1}, \ldots, \mathbf{y}_{p,N_r}]$, where $\mathbf{y}_{p,n_r}$ represents the received signal of the n_r^{th} receiver, sampled at the pilot subcarriers, $\mathbf{\Phi}_p = \mathbf{X}_p(\mathbf{I}_{N_t} \otimes \mathbf{F}_p)$, with the transmit data matrix $\mathbf{X}_p \triangleq [\mathbf{X}_{p,1}, \mathbf{X}_{p,2}, \ldots, \mathbf{X}_{p,N_t}]$ consisting of diagonal matrices $\mathbf{X}_{p,n_t}$ containing known pilot symbols along the diagonal, and $\mathbf{F}_p$ is a submatrix of $\mathbf{F}$ consisting of the rows corresponding to the pilot locations.

The goal of a MIMO-OFDM system is accurate data detection at the output of the decoder as depicted in Fig. 1. The bitwise Log Likelihood Ratios (LLRs) form the inputs to the decoder, and in turn depend on the quality of channel estimates [11]. Several papers in the MIMO channel estimation literature have proposed group-sparse based formulation [6], [7]. Compressed Sensing (CS) based sparse recovery techniques for group-sparse vectors are capable of estimating the a-group-sparse channels using $P_b < L$ pilots, when the MIP is not available [12]. These algorithms recover the magnitude and the locations of the significant channel taps by leveraging the a-group-sparse nature of the channel. However, none of the papers in the existing literature address the a-group-sparse channel estimation problem in the Bayesian framework.

In this work, we address the problem of a-group-sparse MIMO-OFDM channel estimation in the Bayesian framework. Specifically, we model the channel as $\mathbf{h}_{n_t n_r} \sim \mathcal{CN}(0, \Gamma)$ for $0 \leq n_t \leq N_t$ and $0 \leq n_r \leq N_r$, where the hyperparameter $\Gamma = \text{diag}(\gamma(1), \ldots, \gamma(L))$ is common for all the channels, i.e., if $\gamma(l) \to 0$, then the corresponding $h_{n_t n_r}(l) \to 0$ for $0 \leq n_t \leq N_t, 0 \leq n_r \leq N_r$ [13], [14]. Among the known Bayesian sparse recovery techniques [15], Sparse Bayesian Learning (SBL) exhibits the monotonicity property by virtue of the Expectation Maximization (EM) framework, and offers guarantees such as convergence to the sparsest solution when the noise variance is zero, and converging to a local minimum which is a sparse vector, irrespective of the noise variance [14]. This motivates us to employ SBL [14], [16] based algorithms for recovery of the spatially uncorrelated a-group-sparse channel in MIMO-OFDM systems. First, we propose a pilot-based technique for a-group-sparse MIMO-OFDM channel estimation using $P_b < L$, by utilizing the observations at the pilot subcarriers. However, in our previous work, we have demonstrated that incorporating the observations available at the data subcarriers into the approximately sparse (a-sparse) channel estimation using joint techniques enhance the quality of channel estimates in a SISO-OFDM system [17]. In this work, we explore such joint channel estimation and data detection techniques in the context of MIMO-OFDM systems. The contributions of this work are summarized as follows:

- We adapt the multiple response MSBL[1] algorithm [9] to the problem of pilot-based a-group-sparse channel estimation modeled in (4).

[1]The abbreviation MSBL has not been expanded by the authors in [9]. However, in MSBL or M sparse Bayesian learning, M refers to the MMV framework.

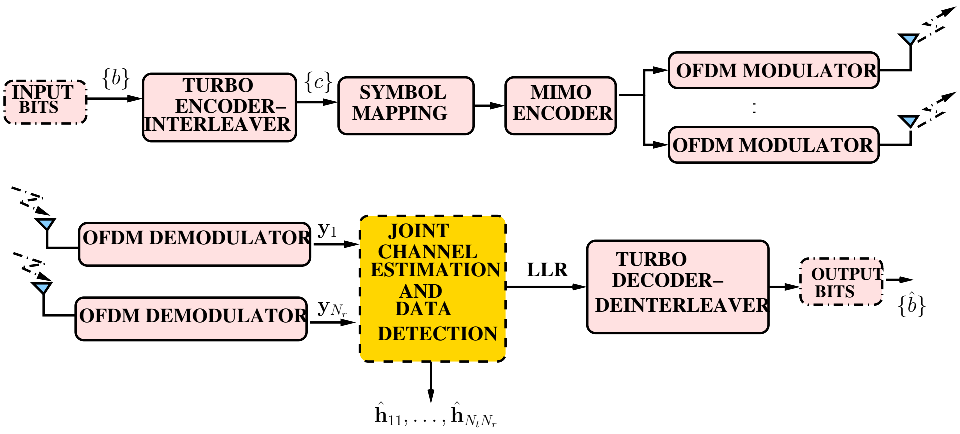

Fig. 1. MIMO-OFDM System model with the outer Turbo encoding and decoding. The dashed box (block shaded in yellow) highlights the proposed algorithms. Note that the quantities of interest are the channel estimates $\hat{\mathbf{h}}_{11}, \ldots, \hat{\mathbf{h}}_{N_t N_r}$ and output bits $\{\hat{b}\}$.

- We propose a novel J-MSBL algorithm for the joint a-group-sparse channel estimation and data detection problem modeled in (3).

SBL based algorithms employ the EM algorithm, in which the E-step provides the posterior density of $\mathbf{H}$, and hence, the maximum aposteriori estimate of $\mathbf{H}$, for a given update of Γ. The novelty of the J-MSBL algorithm lies in the fact that, in contrast to the M-step in MSBL, a joint maximization over both Γ and the data $\mathbf{X}$ is performed, and their Maximum Likelihood (ML) estimates are obtained. Incorporating the ML estimate of the transmitted data $\mathbf{X}$ into the channel estimation process enhances the quality of the channel estimates, in turn leads to better coded Bit Error Rate (BER) performance, compared to MSBL and conventional methods. Further, we demonstrate a significant improvement in the Mean Square Error (MSE) in the estimated a-group-sparse channels.

The rest of this paper is organized as follows. In Section II, we state the MSBL algorithm in the context of pilot-based a-sparse channel recovery for MIMO-OFDM systems. The novel J-MSBL algorithm is derived Section III. The efficacy of the proposed techniques are illustrated through simulation results in Section IV. We provide some concluding remarks in Section V.

Notation: Boldface small letters denote vectors and boldface capital letters denote matrices. The symbols $(\cdot)^T$ and $|\cdot|$ denote the transpose and determinant of a matrix, respectively. Also, $\text{diag}(\mathbf{a})$ denotes a diagonal matrix with entries on the diagonal given by $\mathbf{a}$. The pdf of the random variable X is represented as $p(x)$ and the random variables and deterministic parameters in the pdf are separated using a semicolon. The expectation with respect to a random variable X is denoted as $\mathbb{E}_X(\cdot)$. The $L \times L$ identity matrix is represented as $\mathbf{I}_L$ and $\mathbf{A} \otimes \mathbf{B}$ denotes the Kronecker product of $\mathbf{A}$ and $\mathbf{B}$. The i^{th} entry of a vector $\mathbf{a}$ and the $(i,j)^{\text{th}}$ entry of a matrix $\mathbf{A}$ are represented as $a(i)$ and $A(i,j)$, respectively. Throughout the paper, p as a subscript refers to pilots and (r) in the superscript refers to the iteration number.

II. MSBL for Pilot-based Channel Estimation in MIMO-OFDM Systems

In this section, we adapt the MSBL algorithm for a-group-sparse channel estimation in MIMO-OFDM systems.

Sparse Bayesian learning relies on a parameterized prior to obtain sparse solutions in regression. The parametric form of the MSBL prior can be written as

$$p(\mathbf{H}; \Gamma_b) = \prod_{n_r=1}^{N_r} p(\mathbf{h}_{n_r}; \Gamma), \tag{5}$$

where $\mathbf{h}_{n_r}$ represents the n_r^{th} column of $\mathbf{H}$, given by $\mathbf{h}_{n_r} = [\mathbf{h}_{1n_r}^T, \ldots, \mathbf{h}_{N_t n_r}^T]^T$, with a prior pdf of $\mathbf{h}_{n_r} \sim \mathcal{CN}(0, \Gamma_b)$, $\Gamma_b = \mathbf{I}_{N_t} \otimes \Gamma$ which control the variances of elements in $\mathbf{H}$. Typically, the hyperparameters in $\boldsymbol{\gamma} = [\gamma(1), \gamma(2), \ldots, \gamma(L)]$ can be estimated using the type-II ML procedure [16], i.e., by maximizing the marginalized pdf $p(\mathbf{y}_{p,n_r}; \boldsymbol{\gamma})$ at the receiver as

$$\gamma_{ML}(i) = \underset{\gamma(i) \in \mathbb{R}_+}{\arg\max}\, p(\mathbf{y}_{p,n_r}; \boldsymbol{\gamma}), \quad 1 \le n_r \le N_r, \tag{6}$$

for $1 \le i \le L$. Since the above problem cannot be solved in closed form, iterative estimators such as the EM based[2] MSBL algorithm [9] is employed. In order to use the MSBL algorithm, $\mathbf{H}$ is treated as the hidden variable and the posterior distribution of $\mathbf{H}$ is obtained in the E-step, and the ML estimate of $\boldsymbol{\gamma}$ is obtained in the M-step. The steps of the algorithm are given as

$$\text{E-step}: Q(\boldsymbol{\gamma}|\boldsymbol{\gamma}^{(r)}) = \mathbb{E}_{\mathbf{H}|\mathbf{Y}_p;\boldsymbol{\gamma}^{(r)}}[\log p(\mathbf{Y}_p, \mathbf{H}; \boldsymbol{\gamma})] \tag{7}$$

$$\text{M-step}: \gamma^{(r+1)}(i) = \underset{\gamma(i) \in \mathbb{R}_+}{\arg\max}\, Q(\boldsymbol{\gamma}|\boldsymbol{\gamma}^{(r)}), \quad 1 \le i \le L, \tag{8}$$

and these steps are iterated until convergence. The E-step requires the posterior distribution $p(\mathbf{H}|\mathbf{Y}_p; \boldsymbol{\gamma}^{(r)})$. This can be

[2]Note that all the algorithms proposed in the paper use EM-based updates, and hence, they have a convergence guarantee to a local optima, with the likelihood increasing in each iteration [18].

obtained from the likelihood at the n_r^{th} receiver given by

$$p(\mathbf{y}_{p,n_r}|\mathbf{h}_{n_r}) = (\pi\sigma^2)^{-\frac{N_r}{2}} \exp\left(-\frac{\|\mathbf{y}_{p,n_r} - \mathbf{\Phi}_p\mathbf{h}_{n_r}\|_2^2}{\sigma^2}\right). \tag{9}$$

Combining the likelihood and the prior distribution, the posterior distribution of $\mathbf{h}_{n_r}$ is given by $p\left(\mathbf{h}_{n_r}|\mathbf{y}_{p,n_r};\boldsymbol{\gamma}^{(r)}\right) \sim \mathcal{CN}(\boldsymbol{\mu}_{n_r}, \mathbf{\Sigma})$, with mean and covariance given by

$$\boldsymbol{\mu}_{n_r} = \sigma^{-2}\mathbf{\Sigma}\mathbf{\Phi}_p^H\mathbf{y}_{p,n_r} \quad \mathbf{\Sigma} = \left(\frac{\mathbf{\Phi}_p^H\mathbf{\Phi}_p}{\sigma^2} + \Gamma_b^{(r)^{-1}}\right)^{-1} \tag{10}$$

where $\Gamma_b^{(r)}$ is the estimate of the hyperparameters Γ_b in the r^{th} iteration. We represent the overall posterior mean of $\mathbf{H}$ as $\mathbf{M} = [\boldsymbol{\mu}_1 \ldots \boldsymbol{\mu}_{N_r}]$. The M-step given by (8) can be simplified as

$$\begin{aligned} Q\left(\boldsymbol{\gamma}|\boldsymbol{\gamma}^{(r)}\right) &= \mathbb{E}_{\mathbf{H}|\mathbf{Y}_p;\boldsymbol{\gamma}^{(r)}}\left[\log p(\mathbf{Y}_p, \mathbf{H}; \boldsymbol{\gamma})\right] \\ &= \mathbb{E}_{\mathbf{H}|\mathbf{Y}_p;\boldsymbol{\gamma}^{(r)}}\left[\log \prod_{n_r=1}^{N_r} p(\mathbf{y}_{p,n_r}|\mathbf{h}_{n_r}) \prod_{n_t=1}^{N_t} p(\mathbf{h}_{n_t n_r};\boldsymbol{\gamma})\right] \\ &= c' - \mathbb{E}_{\mathbf{H}|\mathbf{Y}_p;\boldsymbol{\gamma}^{(r)}}\left[\textstyle\sum_{n_r=1}^{N_r}\sum_{n_t=1}^{N_t} \mathbf{h}_{n_t n_r}^H \Gamma^{-1}\mathbf{h}_{n_t n_r}\right], \end{aligned} \tag{11}$$

where c' is a constant independent of $\boldsymbol{\gamma}$. By maximizing $Q\left(\boldsymbol{\gamma}|\boldsymbol{\gamma}^{(r)}\right)$ w.r.t. $\gamma(i)$, we obtain the update of $\gamma(i)$ as follows:

$$\begin{aligned} \gamma^{(r+1)}(i) = \frac{1}{N_t N_r}\textstyle\sum_{n_r=1}^{N_r}\sum_{n_t=0}^{N_t-1}\left(\|\mathbf{M}(i+n_tL, n_r)\|_2^2\right. \\ \left. + \mathbf{\Sigma}(i+n_tL, i+n_tL)\right). \end{aligned} \tag{12}$$

Note that, in the above equation, the a-group-sparse nature of the channel results in the update of $\boldsymbol{\gamma}$ which is averaged over the N_tN_r channels of the MIMO-OFDM system. For a SISO-OFDM system, $N_t = N_r = 1$, the above expression is simplifies to the one obtained in [19].

The MSBL requires initial estimates of the unknown parameters $\boldsymbol{\gamma}$. In practice, it is found that an initial estimate for Γ given by

$$\Gamma^{(0)} = \mathbf{I}_{L\times L}, \tag{13}$$

is sufficient for the MSBL algorithm.

In the following section, we generalize the MSBL algorithm to a J-MSBL algorithm where the unknown transmitted data bits are incorporated into the iterations, resulting in a significant performance improvement.

III. J-MSBL for Joint A-group-sparse Channel Estimation and Data Detection

In this section, we derive the J-MSBL algorithm for joint estimation of the a-group-sparse channels and the transmitted data in a MIMO-OFDM system, by generalizing the MSBL framework to the case where the measurement matrix is partially unknown due to the presence of unknown data symbols.

To derive this algorithm, we consider $\mathbf{H}$ in (3) as the hidden variable, and, in contrast to MSBL, we consider $[\boldsymbol{\gamma}, \mathbf{X}_1, \ldots, \mathbf{X}_{N_t}]$ as parameters to be estimated. The E and

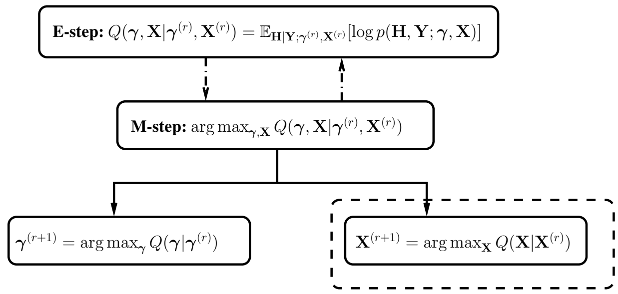

Fig. 2. The J-MSBL algorithm: E-step computes the expectation over the posterior density of $\mathbf{H}$. The joint maximization in the M-step simplifies into two independent maximizations over $\boldsymbol{\gamma}$ and $\mathbf{X}$. The dashed box indicates the novelty in the J-MSBL approach.

the M-steps of the J-MSBL algorithm can be given as

$$\begin{aligned} &\text{E-step}: Q(\boldsymbol{\gamma}, \mathbf{X}|\boldsymbol{\gamma}^{(r)}, \mathbf{X}^{(r)}) = \mathbb{E}_{\mathbf{H}|\mathbf{Y};\boldsymbol{\gamma}^{(r)}}[\log p(\mathbf{Y}, \mathbf{H}; \boldsymbol{\gamma}, \mathbf{X})] \\ &\text{M-step}: \left(\boldsymbol{\gamma}^{(r+1)}, \mathbf{X}^{(r+1)}\right) \\ &\qquad = \arg\max_{\boldsymbol{\gamma}\in\mathbb{R}_+^{L\times 1}, \mathbf{X}: x_i\in\mathcal{S}} Q(\boldsymbol{\gamma}, \mathbf{X}|\boldsymbol{\gamma}^{(r)}, \mathbf{X}^{(r)}), \end{aligned} \tag{14}$$

where x_i is an element in $\mathbf{X}$, and $\mathcal{S}$ is the constellation from which the symbol is transmitted. The E-step of J-MSBL consists of computing the posterior distribution as $p(\mathbf{H}|\mathbf{Y};\boldsymbol{\gamma}^{(r)}, \mathbf{X}^{(r)}) \sim \mathcal{CN}(\mathbf{M}, \mathbf{\Sigma})$, where an element of the matrix mean matrix $\mathbf{M} = [\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_{N_r}]$, given by $\boldsymbol{\mu}_{n_r}$, and the covariance matrix $\mathbf{\Sigma}_b$ are as follows:

$$\boldsymbol{\mu}_{n_r} = \sigma^{-2}\mathbf{\Sigma}_b\mathbf{\Phi}_b^H\mathbf{y}_{n_r k} \quad \mathbf{\Sigma}_b = \left(\sigma^{-2}\mathbf{\Phi}_b^H\mathbf{\Phi}_b + \Gamma^{(r)^{-1}}\right)^{-1} \tag{15}$$

where, for K OFDM symbols in a frame, $\mathbf{\Phi}_b = [\mathbf{\Phi}_1^T, \ldots, \mathbf{\Phi}_K^T]^T$, and for $1 \le k \le K$, $\mathbf{F}_b = \mathbf{1}_{N_t} \otimes \mathbf{F}$, $\mathbf{\Phi}_k = \mathbf{F}_b\text{blkdiag}(\mathbf{X}_{1k}^{(r)}, \ldots, \mathbf{X}_{N_t k}^{(r)})$ and $\mathbf{y}_{n_r k} = [\mathbf{y}_{1k}^T, \ldots, \mathbf{y}_{N_r k}^T]^T$.

At the outset, solving the joint optimization problem in the M-step in (14) appears to be an uphill task. However, in (14), the joint optimization problem corresponding to $\boldsymbol{\gamma}$ and the unknown data symbols $\mathbf{X}$ decouple as a sum of two independent functions of $\boldsymbol{\gamma}$ and $\mathbf{X}$, $Q(\mathbf{X}|\mathbf{X}^{(r)}) \triangleq \mathbb{E}_{\mathbf{H}|\mathbf{Y};\boldsymbol{\gamma}^{(r)},\mathbf{X}^{(r)}}[\log p(\mathbf{Y}|\mathbf{H};\mathbf{X})]$ and $Q(\boldsymbol{\gamma}|\boldsymbol{\gamma}^{(r)}) \triangleq \mathbb{E}_{\mathbf{H}|\mathbf{Y};\boldsymbol{\gamma}^{(r)},\mathbf{X}^{(r)}}[\log p(\mathbf{H};\boldsymbol{\gamma})]$, as shown in Fig. 2.[3] Further, we see that $Q(\boldsymbol{\gamma}|\boldsymbol{\gamma}^{(r)})$ of the MSBL algorithm and the J-MSBL algorithm are identical, and hence, upon optimizing $Q(\boldsymbol{\gamma}|\boldsymbol{\gamma}^{(r)})$ with respect to $\gamma(i)$, we obtain the expression for $\gamma^{(r+1)}(i)$ as in the MSBL algorithm, given by (8). The objective function to obtain $\mathbf{X}$, i.e., $Q(\mathbf{X}|\mathbf{X}^{(r)})$, can be simplified as follows:

$$\begin{aligned} Q(\mathbf{X}|\mathbf{X}^{(r)}) &= \mathbb{E}_{\mathbf{H}|\mathbf{Y};\boldsymbol{\gamma}^{(r)},\mathbf{X}^{(r)}}\left[\log\prod_{n_r=1}^{N_r} p(\mathbf{y}_{n_r}|\mathbf{h}_{n_r};\mathbf{x})\right] \\ &= -\mathbb{E}_{\mathbf{H}|\mathbf{Y};\boldsymbol{\gamma}^{(r)},\mathbf{X}^{(r)}}\left[\textstyle\sum_{n_r=1}^{N_r}\|\mathbf{y}_{n_r k} - \mathbf{\Phi}_b\mathbf{h}_{n_r}\|_2^2\right]. \end{aligned} \tag{16}$$

[3]Notice that (8) and (14) are different, since the former uses the measurement matrix containing only the known pilot symbols, $\mathbf{\Phi}_p$, whereas the latter uses measurement matrices which consist of pilot symbols along with the estimated data, together given by $\mathbf{\Phi}^{(r)}$.

and hence, the optimization problem for $\mathbf{X}$ is given by

$$X_1^{(r+1)}(i,i),\ldots,X_{N_t}^{(r+1)}(i,i) = \underset{x_1,\ldots,x_{N_t}\in\mathcal{S}}{\arg\min}\ C(i,i) + \sum_{n_r=1}^{N_r}\left|y_{n_r}(i) - \sum_{n_t=1}^{N_t} x_{n_t}\mathbf{F}_b(i,:)\boldsymbol{\mu}_{n_r}\right|^2, \quad (17)$$

where $i \in \mathcal{D}$, $\mathcal{D}$ is an index set consisting of the data subcarrier locations, $\mathbf{C} = \boldsymbol{\Phi}\boldsymbol{\Sigma}\boldsymbol{\Phi}^H$, $\mathbf{F}_b(i,:)$ is the i^{th} row of the $\mathbf{F}_b$ matrix, $\boldsymbol{\mu}_{n_r}$ and $\boldsymbol{\Sigma}$ are given in (15).

This algorithm requires initial estimates of the unknown parameters $\boldsymbol{\gamma}$ and $\mathbf{X}$. The initial estimate of Γ is taken to be the identity matrix, as in the previous section. The initialization of the $(KNN_t - P_bN_t)$ non-pilot data in turn requires an initial channel estimate. Channel estimates using methods like LS and MMSE cannot be used, as they require knowledge of MIP. Hence, the initialization of $\mathbf{X}$ can be obtained from the channel estimate obtained from a few iterations of the MSBL algorithm from the $P_b = P$ pilots (denoted as $\hat{\mathbf{h}}_{\text{MSBL}}$). The ML data detection problem for obtaining the initial data estimates is given by

$$X_1^{(0)}(i,i),\ldots,X_{N_t}^{(0)}(i,i) = \underset{x_1,\ldots,x_{N_t}\in\mathcal{S}}{\arg\min}\ |y_{n_r}(i) - \sum_{n_t=1}^{N_t} x_{n_t}\mathbf{F}_b(i,:)\hat{\mathbf{h}}_{\text{MSBL}}|^2, \quad (18)$$

for $i \in \mathcal{D}$. In order to obtain the solution of (17) and (18), we need to find the vector $[x_1,\ldots,x_{N_t}]$ that jointly minimizes their right hand side. Although we can solve this problem with moderate complexity for MIMO-OFDM systems with $N_t = 2,4$ [20], the complexity of this problem is high for large values of N_t. In such scenarios, it is desirable to use computationally efficient techniques such as sphere decoding [10].

IV. Simulation Results

In this section, we illustrate the performance of the proposed pilot-based, and joint data and channel estimation algorithms through Monte Carlo simulations. We consider the parameters in the 3GPP/LTE broadband standard [20], [21]. We use a 3MHz spatially-multiplexed MIMO-OFDM system with 256 subcarriers, with a sampling frequency of $f_s = 3.84$MHz, resulting in an OFDM symbol duration of $\sim 83.3\mu s$ with a Cyclic Prefix (CP) of $16.67\mu s$ (64 subcarriers). The length of the a-sparse channel $\mathbf{h}_{n_t n_r}$ is taken to be equal to the length of the CP. Each OFDM frame consists of $K = 7$ OFDM symbols, which is also referred to as an OFDM slot [21]. The data is transmitted using a rate $1/2$ Turbo code with QPSK modulation. For the Turbo decoding, we use the publicly available software [22], which uses a maximum of 10 Turbo iterations.

A sample instantiation of the a-sparse channel, $\mathbf{h}_{n_t n_r}$, used in the simulations and the filtered MIP are depicted in Fig. 3. The figure captures the leakage effect due to finite bandwidth sampling and practical filtering, which, in turn, leads to a-group-sparse nature of the overall channel matrix $\mathbf{H}$. To generate the plot, we have used the Pedestrian B channel model [23] with Rayleigh fading, and raised cosine filtering at the N_r receive and N_t transmit antennas, with a roll-off factor of 0.5 [21]. At the sampling frequencies considered, the number of significant channel taps are far fewer than the weak channel taps in the filtered impulse response, as seen in Fig. 3.

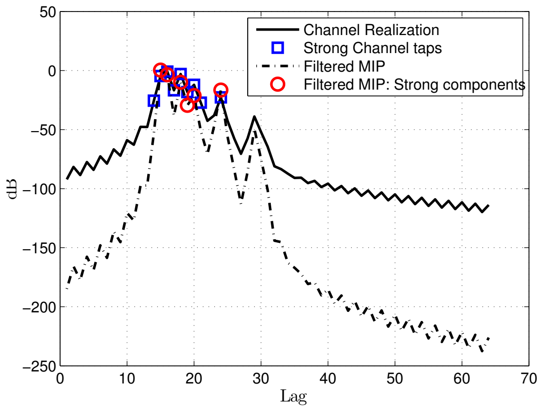

Fig. 3. Illustration of the a-sparse nature of $\mathbf{h}_{n_t n_r}$. One sample channel realization of the $\mathbf{h}_{n_t n_r}$, along with the filtered MIP, i.e., the MIP when raised cosine filters are employed at the transmitter and receiver. The plot also shows the strong (> -30 dB) channel taps and filtered-MIP components, to illustrate that the $\mathbf{h}_{n_t n_r}$ has very few strong components.

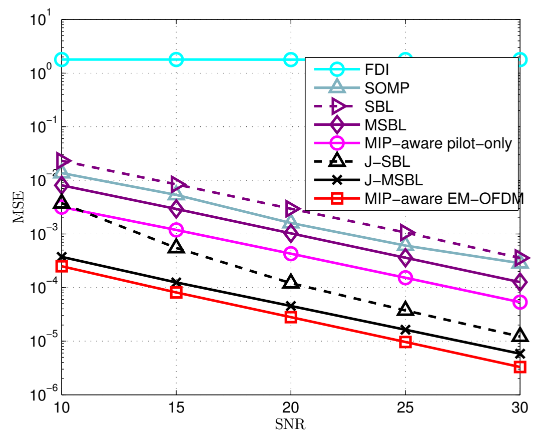

Fig. 4. MSE performance of MSBL, J-MSBL algorithms compared to FDI [2], SBL [14], Simultaneous-OMP with $P_b = 50$ [24], MIP-aware pilot-only [3] and MIP-aware EM-OFDM [10] schemes in a block-fading channel ($P_b = 44$), as a function of SNR in dB.

We consider a block-fading channel. We use $P_b = 44$ pilot subcarriers per transmit antenna, uniformly placed in each OFDM symbol [17]. We implement the MSBL and the J-MSBL algorithms, and plot their MSE performance in Fig. 4, using a convergence criteria for $\boldsymbol{\gamma}$ as $\epsilon = 10^{-8}$, i.e., $\|\boldsymbol{\gamma}^{(r)} - \boldsymbol{\gamma}^{(r-1)}\|^2 \leq \epsilon$ for both the algorithms. We also restrict the maximum number of iterations of both the algorithms to 200. We compare the MSE performance of the proposed algorithms with the MIP-unaware Frequency Domain Interpolation

(FDI) technique, Simultaneous OMP (S-OMP) [24] using 50 pilots, the MIP-aware pilot-only technique [3], and the MIP-aware joint data and channel estimation algorithm [10], which we refer to as the MIP-aware EM-OFDM algorithm. Further, we demonstrate the benefits of employing the a-group-sparse formulation by plotting the MSE performance of SBL and J-SBL for SISO-OFDM systems, as shown in [17].

From Fig. 4, we observe that the proposed algorithms perform better than the MIP-unaware, non-iterative schemes such as the Frequency Domain Interpolation (FDI) technique. Among the iterative methods, MSBL and J-MSBL algorithms are 3dB better than their SISO-OFDM counterparts and MSBL performs 2dB better than the compressed sensing based S-OMP algorithm. Further, the J-MSBL algorithm performs more than an order of magnitude better than the MSBL algorithm, while being within 3 dB from the MIP-aware EM-OFDM algorithm. The J-MSBL jointly detects the $(KNN_t - P_bN_t)$ data symbols along with the channel, resulting in a significantly lower overall MSE.

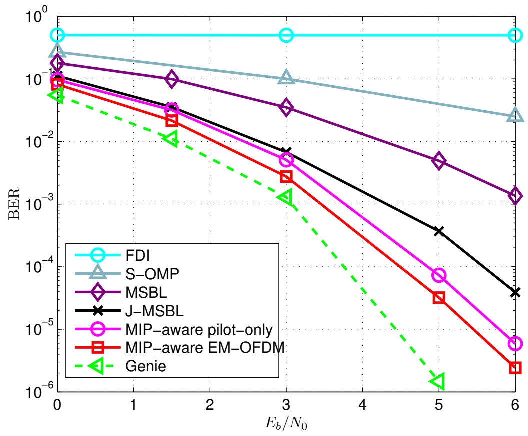

Fig. 5. BER performance of the proposed algorithms in a block-fading channel, with $P_b = 44$ pilot subcarriers, as a function of E_b/N_0.

The coded BER performance of the MIP-aware EM-OFDM, J-MSBL and a genie receiver, i.e., a receiver with perfect knowledge of the channel (labeled as `Genie`), is shown in Fig. 5. We also compare the performance with MSBL, S-OMP ($P_b = 50$) and MIP-aware pilot-only channel estimation followed by data detection. The BER performance of MSBL is superior to that of S-OMP by 2 dB. The BER performance of the J-MSBL is superior that of the MSBL by 2dB, while being less than 1dB away from the MIP-aware EM-OFDM and the Genie receiver.

V. Conclusion

In this paper, we considered SBL based pilot-based and joint channel estimation and data detection for block-fading a-group-sparse channels in MIMO-OFDM systems. To estimate such channels, we proposed the pilot-based MSBL algorithm and generalized it to obtain the J-MSBL algorithm for joint a-group-sparse channel estimation and data detection. Simulation results showed that the proposed techniques successfully exploit the a-group-sparse nature of the channel, leading to an enhanced channel estimation and data detection capability compared to the conventional and SISO-OFDM counterparts.

References

[1] L. Hanzo, M. Munster, B. J. Choi, and T. Keller, *OFDM and MC-CDMA for Broadband Multi-user communications, WLANs and Broadcasting*. John Wiley, 2003.

[2] 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, 2002.

[3] J. Van de Beek, O. Edfors, M. Sandell, S. Wilson, and P. Borjesson, “On channel estimation in OFDM systems,” in *Proc. VTC*, vol. 2, 1995, pp. 815–819.

[4] D. Tse, and P. Viswanath, *Fundamentals of wireless communication*. Cambridge University Press, 2005.

[5] A. P. Kannu and P. Schniter, “On communication over unknown sparse frequency-selective block-fading channels,” *IEEE Trans. Inf. Theory*, vol. 57, no. 10, pp. 6619–6632, 2011.

[6] A. Gilbert and J. Tropp, “Applications of sparse approximation in communications,” in *Proc. IEEE Int. Symp. Inf. Theory*. IEEE, 2005, pp. 1000–1004.

[7] Y. Barbotin, A. Hormati, S. Rangan, and M. Vetterli, “Estimation of sparse MIMO channels with common support,” *Arxiv preprint arXiv:1107.1339*, 2011.

[8] S. S. Ghassemzadeh, L. J. Greenstein, T. Sveinsson, A. Kavcic, and V. Tarokh, “UWB delay profile models for residential and commercial indoor environments,” *IEEE J. Veh. Tech.*, vol. 54, no. 4, pp. 1235–1244, 2005.

[9] D. Wipf and B. Rao, “An empirical Bayesian strategy for solving the simultaneous sparse approximation problem,” *IEEE Trans. Signal Process.*, vol. 55, no. 7, pp. 3704–3716, 2007.

[10] M. Wang, D. Quevedo, and G. Goodwin, “Joint data detection and channel estimation for MIMO-OFDM systems via EM algorithm and sphere decoding,” in *Proc. Globecom*. IEEE, Dec. 2006, pp. 1 –5.

[11] S. Schwandter, P. Fertl, C. Novak, and G. Matz, “Log-likelihood ratio clipping in mimo-bicm systems: Information geometric analysis and impact on system capacity,” in *Proc. ICASSP*. IEEE, 2009, pp. 2433–2436.

[12] W. U. Bajwa, A. Sayeed, and R. Nowak, “Compressed sensing of wireless channels in time, frequency, and space,” in *Proc. Asilomar Conf. on Signals, Syst., and Comput.* IEEE, 2008, pp. 2048–2052.

[13] M. E. Tipping, “The relevance vector machine,” in *Advances in NIPS*, vol. 12, 2000.

[14] D. P. Wipf and B. D. Rao, “Sparse Bayesian learning for basis selection,” *IEEE Trans. Signal Process.*, vol. 52, no. 8, pp. 2153–2164, 2004.

[15] S. Babacan, R. Molina, and A. Katsaggelos, “Bayesian Compressive Sensing using Laplace Priors,” *IEEE Trans. Image Process.*, vol. 19, pp. 53–64, 2010.

[16] M. E. Tipping, “Sparse bayesian learning and the relevance vector machine,” *J. of Machine Learning Research*, vol. 1, pp. 211–214, 2001.

[17] R. Prasad, C. R. Murthy, and B. D. Rao, “Joint Approximately Sparse Channel Estimation and Data Detection in OFDM Systems using Sparse Bayesian Learning,” *Submitted to IEEE Trans. Sig. Proc*, 2013.

[18] G. McLachlan and T. Krishnan, *The EM algorithm and extensions*. Wiley New York, 1997, vol. 382.

[19] R. Prasad and C. Murthy, “Bayesian learning for joint sparse OFDM channel estimation and data detection,” in *Proc. Globecom*. IEEE, 2010, pp. 1–6.

[20] J. Zyren and W. McCoy, “Overview of the 3GPP long term evolution physical layer,” *Freescale Semiconductor, Inc., white paper*, 2007.

[21] “Universal Mobile Telecommunications System (UMTS), Selection procedures for the choice of radio transmission technologies of the UMTS,” ETSI technical report, UMTS 21.01 version 3.0.1, Nov., 1997.

[22] C. Studer, C. Benkeser, S. Belfanti, and Q. Huang, “Design and implementation of a parallel turbo-decoder ASIC for 3GPP-LTE,” *IEEE J. Solid State Circuits.*, vol. 46, no. 1, pp. 8–17, 2011.

[23] Recommended ITU-R M.1225, “Guidelines for evaluation of radio transmission technologies for IMT-2000,” 1997.

[24] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, “Algorithms for simultaneous sparse approximation. part I: Greedy pursuit,” *Elsevier Signal Processing*, vol. 86, no. 3, pp. 572–588, 2006.