Receiver Design for Bit-interleaved MIMO-OFDM Systems over Time-varying Channels

Fu-Hsuan Chiu[†], Sau-Hsuan Wu[‡] and C.-C. Jay Kuo[†]
[†] University of Southern California, Los Angeles, CA 90089, USA
E-mails: fchiu@usc.edu, cckuo@sipi.usc.edu
[‡] National Chiao-Tung University, Hsinchu, Taiwan
E-mails: sauhsuaw@alumni.usc.edu

***Abstract*— A complete solution to joint channel tracking and symbol detection for bit-interleaved MIMO-OFDM systems in a time-varying frequency-selective fading environment is proposed in this research. In the training mode, the Rao-Blackwellised particle filter (RBPF) is used to estimate the phase offset (PO) and the carrier frequency offset (CFO). In the data mode, based on a recursive EM procedure in conjunction with soft decoding, we develop an iterative algorithm that performs the minimum mean squared error (MMSE) channel estimation and the maximum *a posterior* (MAP) probability symbol detection jointly. The soft decoding module is composed of the list sphere decoder (LSD) and the soft-in-soft-out (SISO) channel decoder based on the BCJR algorithm. The performance of the proposed algorithm is evaluated via simulation. It is demonstrated that the new algorithm has robust performance in the presence of phase noise and the carrier frequency offset.**

I. Introduction

Space-time coding offers a bandwidth- and power-efficient solution to wireless communications. It exploits spatial diversity from multi-input multi-output (MIMO) channels to combat fading effects [1]. On the other hand, orthogonal frequency division multiplexing (OFDM) has been adopted in several commercial applications due to its robustness against frequency-selective fading. Therefore, MIMO-OFDM is a promising scheme for broadband wireless communications.

Bit-interleaved coded modulation [2] has been applied to narrowband MIMO systems over flat fading channels and to OFDM systems over frequency-selective fading channels [3] [4]. The performance analysis of bit-interleaved OFDM systems has been studied in the literatures. However, neither channel estimation nor phase/carrier frequency offsets were considered in such a context before. The phase offset (PO) and the carrier frequency offset (CFO) in OFDM systems not only result in the phase rotation of all subcarriers but also introduce the inter-carrier-interference (ICI) effect. A Kalman filter approach to suppress the phase offset was proposed in [5]. Several MUSIC-like algorithms has been adopted to estimate CFO (see [6] [7] and therein). However, a joint PO and CFO estimation scheme has not yet been presented.

The particle filter (PF) [8] simulation is a powerful sequential Monte Carlo method, where the particle representation of probability densities is used to estimate parameters of an arbitrary distribution in a state-space model. In particular, the Rao-Blackwellised particle filter (RBPF) [9] provides a reduced variance estimator of nonlinear and linear parameters by exploiting the conditional linear state-space model structure to estimate the marginal distribution. The same idea was first applied to demodulation over a fading channel [10]. Here, we propose to use RBPF to estimate nonlinear parameters, PO and CFO, in the training mode.

Several schemes were proposed for joint maximum likelihood sequence detection (MLSD) and maximum *a posterior* (MAP) or minimum mean squared error (MMSE) channel estimation in the past. Most of them are proved to be special cases of the Expectation Maximization (EM) algorithm [11]. The most well known example is the per-survivor processing (PSP) technique [12]. A main drawback of PSP is that its complexity increases exponentially with the channel memory length in time-varying fading channels. Adaptive Bayesian and EM-based detectors for single-in-single-output frequency-selective fading channels were proposed in [13], where the performance of coded systems was greatly enhanced by the soft-decision channel estimator. A unified structure based on recursive EM for joint channel tracking and symbol detection in multipath fading channels was proposed in [14]. Furthermore, in conjunction with the BCJR algorithm, the recursive procedure was presented in [14] as a message-passing scheme on a cyclic graph. This algorithm was modified and extended to MIMO channels in [15], [16]. Pioneering research on joint channel estimation and symbol detection was conducted in [17], [18], [19], where a hard-decision Kalman filter was used to track time-varying channels. A turbo receiver employing a MAP-EM demodulator and a soft decoder was proposed for space-time block-coded and LDPC-based space-time coded OFDM systems over correlated fading channels in [20] and [21], respectively. The MAP-EM demodulator treats the transmitted signal as the desired parameter and the channel as the unobserved data. Here, we extend the result in [14] to coded MIMO-OFDM systems over time-varying channels. Specifically, we consider a system with the source sequence encoded by a convolutional encoder, interleaved, QAM-modulated and then mapped to OFDM symbols followed by an MIMO encoder.

The rest of this paper is organized as follow. A general bit-interleaved MIMO-OFDM system over time-varying frequency-selective fading channels in the presence of PO and CFO is discussed in Sec. II. An efficient way to estimate PO and CFO by RBPF is given in Sec. III, and this type of information can be obtained in the training mode in the very beginning of transmission. A joint channel taps tracking and symbol detection scheme using the EM algorithm with soft decoding is proposed in Sec. IV. Soft sphere decoding based on the list sphere decoding [22] is used for MIMO soft decoding, while the BCJR algorithm [23] is used to provide soft decoding of the outer convolution codes. The performance of the proposed scheme is evaluated and compared with the Kalman filter with hard decision feedback in Sec. V. Concluding remarks are given in Sec. VI.

II. System Model

We consider a bit-interleaved MIMO-OFDM system with n_t transmit antennas and n_r receive antennas as shown in Fig. 1. The source sequences are encoded by a convolutional encoder, interleaved, QAM(or PSK) -modulated and then mapped to OFDM symbols followed by space-time code encoder.

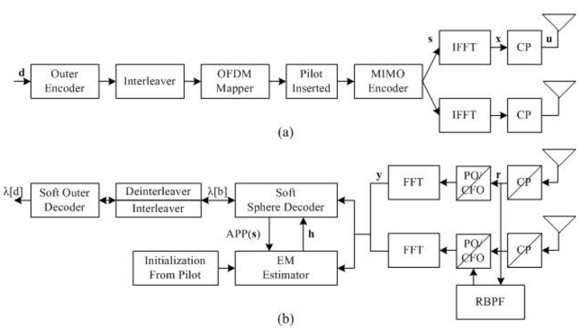

Fig. 1. (a) The transmitter of the MIMO-OFDM system and (b) the proposed receiver for joint channel estimator and symbol detection based on the EM algorithm and soft decoding.

Let N be the total number of subcarriers and $\mathbf{s}_i(m) = [s_{i,0}(m), \ldots, s_{i,N-1}(m)]^T$ the mth block of frequency-domain symbols sent by the ith antenna. The corresponding time-domain vector is given by $\mathbf{x}_i(m) = \mathbf{W}^H \mathbf{s}_i(m)$, where $\mathbf{W}$ is the N-point discrete Fourier transform (DFT) matrix. A cyclic prefix (CP) of length L_{CP} is appended in front of $\mathbf{x}_i(m)$ to eliminate the interblock interference (IBI). The resulting vector $\mathbf{u}_i(m)$ with length $L_u = N + L_{CP}$ is then transmitted over the channel. Let the discrete-time channel response from the ith antenna to jth antenna be $\mathbf{h}_{i,j}(m) = [h_{i,j,0}(m), \cdots, h_{i,j,L-1}(m)]^T$ where L is the maximum channel length. In the presence of CFO ω and PO ψ between the receiver and transmit oscillator, the OFDM signal at the receiver can be written in a compact form as

$$\mathbf{r}_m = e^{j\omega(mL_u + L_{CP}) + \psi} \Omega(\omega) \mathbf{D}(\mathbf{X}_m) \mathbf{h}_m + \mathbf{w}_m, \qquad (1)$$

where $\mathbf{r}_m = [\mathbf{r}_1^T(m), \cdots, \mathbf{r}_{n_t}^T(m)]^T$,

$$\begin{aligned}
\Omega(\omega) &= \begin{bmatrix} \mathrm{diag}(\gamma(\omega)) & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathrm{diag}(\gamma(\omega)) \end{bmatrix}, \\
\gamma(\omega) &= [1, e^{j\omega}, \cdots, e^{j(N-1)\omega}]^T, \\
\mathbf{h}_m &= [\mathbf{h}_{1,1}^T(m), \mathbf{h}_{2,1}^T(m), \cdots, \mathbf{h}_{n_t,1}, \cdots, \mathbf{h}_{n_t,n_r}^T(m)]^T,
\end{aligned}$$

and $\mathbf{D}(\mathbf{x}_m)$ is a block diagonal matrix with block matrix $\mathbf{X}_m = [\mathbf{X}_1(m), \cdots, \mathbf{X}_{n_t}(m)]$ of dimension $N \times n_t L$, where $\mathbf{X}_i(m)$ is the $N \times L$ circular matrix generated from $\mathbf{x}_i(m)$ of the ith transmit antenna with the $(p,q)th$ entry given by the $((p-q)\mathrm{mod}N)$-th entry of $\mathbf{x}_i$(m). A common way to model the time evolution of a vector process is to use the autoregressive moving average (ARMA) filter. Let us define $\underline{\mathbf{h}}_m \equiv [\mathbf{h}_m^T, ..., \mathbf{h}_{m-L_h+1}^T]^T$, where L_h is the order of the channel model. The channel dynamics can be modeled by the following state-space form

$$\begin{aligned}
\underline{\mathbf{h}}_m &= \underline{\mathbf{F}}\underline{\mathbf{h}}_{m-1} + \underline{\mathbf{B}}\mathbf{v} \\
&= \begin{bmatrix} \mathbf{F}_1 & \mathbf{F}_2 & \cdots & \mathbf{F}_{L_h} \\ \mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{I} & \mathbf{0} \end{bmatrix} \underline{\mathbf{h}}_{m-1} + \begin{bmatrix} \mathbf{B} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix} \mathbf{v}, \qquad (2)
\end{aligned}$$

where $\mathbf{F}_i$, $i = 1, \cdots, L_h$, and $\mathbf{B}$ is of dimension $n_t L \times n_t L$. Under the wide-sense-stationary-uncorrelated-scattering (WSSUS) assumption, $\mathbf{F}_i$ and $\mathbf{B}$ must be diagonal. The diagonal entries can be selected by exploiting the correlation-matching property and solving Yule-Walker's equations in order to match the physical model. We will assume $L_h = 1$ in the following. It is however easy to extend the algorithm to any case with slight modification.

III. Particle Filter in Training Mode

We propose a sequential Monte Carlo estimator based on the Rao-Blackwellised particle filter (RBPF) to estimate nonlinear parameters (*i.e.*, PO ψ and CFO ω) and linear parameters (*i.e.*, channel taps $\mathbf{h}_t$) jointly in the training mode. We change the subscript notation from m to t to emphasize the training mode. Based on (1), the joint posterior probability is

$$p(\psi, \omega, \mathbf{h}_t | \mathbf{r}_t) = p(\mathbf{h}_t | \psi, \omega, \mathbf{r}_t) p(\psi, \omega | \mathbf{r}_t), \qquad (3)$$

where $p(\mathbf{h}_t|\psi,\omega)$ is analytically trackable by the Kalman filter due to the conditional Gaussian distribution in the linear state-space model. Thus, we only need to design the particle filter to approximate the marginal posterior distribution $p(\psi,\omega|\mathbf{r}_t)$. Let $\theta = \{\psi, \omega\}$, $\mathcal{X}(\theta) = e^{j\omega(mL_u+L_{CP})+\psi}\Omega(\omega)\mathbf{D}(\mathbf{X}_t)$, and $\mathcal{R}_t = \{\mathbf{r}_i : i = 0, \cdots, t\}$. Then (1) can be written as

$$\mathbf{r}_t = \mathcal{X}(\theta)\mathbf{h}_t + \mathbf{w}_t. \qquad (4)$$

The proposed RBPF algorithm is given below.

1) Initialization. For $i = 1, \cdots, N_p$, draw initial particle $\theta_{0|-1}^{(i)}$ from the prior distribution. In our system, it is

assumed that the prior is uniform. Initialize the Kalman filter associated with each particle as $\mathbf{h}_{0|-1}^{(i)} = \hat{\mathbf{h}}_0$, and $\mathbf{P}_{0|-1}^{(i)} = \hat{\mathbf{P}}_0$.

2) Evaluating and normalizing the importance weights. For $i = 1, \cdots, N_p$, we have

$$\begin{aligned} q_m^{(i)} &= p(\mathbf{r}_t|\mathcal{R}_{t-1}, \theta_{m|t-1}^{(i)}) \sim \mathcal{N}(\mathbf{r}_{t|t-1}^{(i)}, \mathbf{U}_t^{(i)}), \\ \tilde{q}_t^{(i)} &= q_t^{(i)} / \Sigma_{j=1}^{N_p} q_t^{(j)}, \end{aligned}$$

where "$|t$" denotes using information up to time t and

$$\begin{aligned} \mathbf{r}_{t|t-1} &= \mathcal{X}(\theta_{t|t-1}^{(i)})\mathbf{h}_{t|t-1}^{(i)}, \\ \mathbf{U}_t^{(i)} &= \mathcal{X}(\theta_{t|t-1}^{(i)})\mathbf{P}_{t|t-1}^{(i)}\mathcal{X}(\theta_{t|t-1}^{(i)})^H + \sigma^2\mathbf{I}. \end{aligned}$$

3) Parameter estimate given by

$$\hat{\theta}_{t|t} = \sum_{i=1}^{N_p} \tilde{q}_t^{(i)} \theta_{t|t-1}^{(i)}. \quad (5)$$

4) Particle filter measurement update. Resample N_p particles with replacement

$$\Pr(\theta_{t|t}^{(i)} = \theta_{t|t-1}^{(j)}) = \tilde{q}_t^{(j)}.$$

5) Kalman filter measurement update. For $i = 1, \cdots, N_p$, we have

$$\begin{aligned} \mathbf{h}_t^{(i)} &= \mathbf{h}_{t|t-1}^{(i)} + \mathbf{K}_t^{(i)}(\mathbf{r}_t - \mathcal{X}(\theta_{t|t}^{(i)})\mathbf{h}_{t|t-1}^{(i)}), \\ \mathbf{P}_{t|t}^{(i)} &= \mathbf{P}_{t|t-1}^{(i)} - \mathbf{K}_t^{(i)}\mathbf{U}_t^{(i)}\mathbf{K}_t^{(i)H}, \\ \mathbf{K}_t &= \mathbf{P}_{t|t-1}^{(i)}\mathcal{X}(\theta_{t|t}^{(i)})^H \mathbf{U}_t^{(i)-H}, \\ \mathbf{U}_t &= \mathcal{X}(\theta_{t|t}^{(i)})\mathbf{P}_{t|t-1}^{(i)}\mathcal{X}(\theta_{t|t}^{(i)})^H + \sigma^2\mathbf{I}. \end{aligned}$$

6) Particle filter time update. For $i = 1, \cdots, N_p$, draw new particle $\theta_{t+1|t}^{(i)}$ from $p(\theta_{t+1|t}|\theta_{t|t}^{(i)})$.

7) Kalman filter time update. For $i = 1, \cdots, N_p$, we set

$$\begin{aligned} \mathbf{h}_{t+1|t}^{(i)} &= \mathbf{F}\mathbf{h}_{t|t-1}^{(i)}, \\ \mathbf{P}_{t+1|t}^{(i)} &= \mathbf{F}\mathbf{P}_{t|t}^{(i)}\mathbf{F}^H + \mathbf{B}\mathbf{B}^H. \end{aligned}$$

In order to obtain converged parameter estimates, a random walk evolution model with variance decaying with time for the fixed parameters is suggested in [24]. That is, we have

$$\theta_t = \alpha\theta_t + (1-\alpha)\bar{\theta}_t + \mathbf{n}_{\theta t}, \quad (6)$$

where $\alpha = (3\delta - 1)/(2\delta)$, δ is around 0.95 to 0.99, $\bar{\theta}_t$ is the Monte Carlo mean at time m, and $\mathbf{n}_{\theta t} \sim \mathcal{N}(\mathbf{0}, (1-\alpha^2)\mathbf{Q}_t)$, $\mathbf{Q}_t$ is the variance of parameter at time t. Therefore, the probability used in Step 6 is

$$p(\theta_{t+1|t}|\theta_{t|t}^{(i)}) = \mathcal{N}(\alpha\theta_{t|t}^{(i)} + (1-\alpha)\bar{\theta}_{t|t}^{(i)}, (1-\alpha^2)\mathbf{Q}_t). \quad (7)$$

The advantage of the RBPF over generic PF is that the dimension of parameters is reduced by marginalization, which results in estimates with a smaller variance. Since the Kalman filter is widely adopted to track time-varying parameters, RBPF only introduces additional complexity of sampling and memory storage (N_p Kalman filters required) when the hardware implementation is considered. Moreover, from Step 6 and (6), it is obvious that RBPF can also track dynamic nonlinear parameters.

IV. Turbo Receiver Design in Data Mode

In the data mode, we will first compensate the PO and CFO estimated from the training mode. Then, the output of DFT at the receiver $\mathbf{y}_m$ can be written as a dynamic state space model involving only linear parameters as

$$\mathbf{y}_m = \mathcal{D}(\mathbf{s}_m)\mathbf{h}_m + \mathbf{n}_m, \quad (8)$$

where $\mathbf{s}_m = [\mathbf{s}_1(m)^T, \cdots, \mathbf{s}_{n_t}(m)^T]^T$, $\mathcal{D}(\mathbf{s}_m)$ is a block diagonal matrix with $[\mathrm{diag}(\mathbf{s}_1(m))\mathbf{W}_L, \cdots, \mathrm{diag}(\mathbf{s}_{n_t}(m))\mathbf{W}_L]$ as the diagonal entry, $\mathbf{W}_L$ is the firsts $L+1$ columns of the size-N DFT matrix $\mathbf{W}$, and $\mathbf{n}_m$ is noise plus error introduced by the PO/CFO estimation error. Noise $\mathbf{n}_m$ is assumed to be white Gaussian with covariance matrix $\mathbf{C}$. It is more convenient to express (8) as an equivalent form for detection,

$$\begin{aligned} \mathbf{y}_m &= \begin{bmatrix} \mathbf{H}_{1,1}(m) & \cdots & \mathbf{H}_{n_t,1}(m) \\ \vdots & \ddots & \vdots \\ \mathbf{H}_{1,n_r}(m) & \cdots & \mathbf{H}_{n_t,n_r}(m) \end{bmatrix} \mathbf{s}_m + \mathbf{n}_m, \quad (9) \\ &= \mathcal{H}(\mathbf{h}_m)\mathbf{s}_m + \mathbf{n}_m, \quad (10) \end{aligned}$$

where $\mathbf{H}_{i,j}(m) = \mathrm{diag}(\mathbf{W}_L\mathbf{h}_{i,j}(m))$, *i.e.*, a diagonal matrix with the frequency response of $\mathbf{h}_{i,j}(m)$ as the diagonal entry. We propose an EM algorithm to track the channel. The *a posterior* probability (APP) $p(\mathbf{s}_i|\mathcal{Y}_m, \hat{\Theta}_{m|m}^{(l-1)})$ is needed to implement the recursive EM algorithm, and it can be computed by soft decoders. Thus, to conduct channel estimation and symbol detection jointly, the soft decoder and the recursive EM algorithm have to work in concert, with the soft decoder providing metrics for the EM estimator and the EM providing the likelihoods required for metric updates. This recursive procedure can be represented by a graphical model using the concept of message passing [14].

A. Pilot tones

To obtain reliable initial channel taps estimates even in the presence of deep fading, we select P pilot tones in each MIMO-OFDM block, and the identification condition is $P \geq L$. (Recall that L denotes the maximum channel length.) Since pilot tones are known at the receiver, the sequential MMSE estimator can be derived by the standard Kalman filter with observation equation (8) and evolution equation (2).

B. Soft Decoding Module

In MIMO-OFDM systems, soft demodulation is a tone-by-tone based operation due to the ISI-free property of OFDM. Let the MIMO encoder be a linear space-time mapper [22], which implies that $\mathbf{s}_m$ is a linear function of the output of the outer code encoder $\mathbf{b}_m$. Then, the detection model for a particular tone i is equivalent to a standard MIMO model

$$\mathbf{y}_m^{(i)} = \mathcal{H}_m^{(i)}\mathbf{s}_m^{(i)} + \mathbf{n}_m^{(i)}, \quad (11)$$

where $\mathbf{y}_m^{(i)}$ and $\mathcal{H}_m^{(i)}$ depend on (9) and the corresponding space-time mapper, and $\mathbf{s}_m^{(i)} = \mathrm{map}(\mathbf{b}_m^{(i)})$. Let the k-th bit of $\mathbf{b}_m^{(i)}$ be b_k. Then, the soft output of b_k can be derived as

$$\lambda[b_k] = \ln\frac{P[b_k=1]}{P[b_k=-1]} + \ln\frac{\sum_{\mathbf{b}\in\mathbb{B}_{k,+1}} p(\mathbf{y}_m^{(i)}|\mathbf{b})P(\mathbf{b}|b_k)}{\sum_{\mathbf{b}\in\mathbb{B}_{k,-1}} p(\mathbf{y}_m^{(i)}|\mathbf{b})P(\mathbf{b}|b_k)}, \tag{12}$$

where $\mathbb{B}_{k,\pm1} = \{\mathbf{b}|b_k = \pm 1\}$.

1) List Sphere Decoder (LSD): The LSD was adopted in [22] to perform soft demodulation of the MIMO-OFDM symbols and obtain the soft information of coded transmit bits. We choose the Zero-Forcing (ZF) point as the initial center since we do not have reliable covariance information of noise in the beginning. After several iterations, a reliable estimate of the covariance matrix is obtained so that we can perform whitening and use the MMSE point as the center point. The radius can be selected by [25]

$$R \simeq \left(\frac{N_{\mathcal{L}} \times |\det(\mathbb{H}_m)|}{V_s}\right)^{1/s}, \tag{13}$$

where

$$\mathbb{H}_m = \begin{bmatrix} Re\{\mathcal{H}_m\} & -Im\{\mathcal{H}_m\} \\ Im\{\mathcal{H}_m\} & Re\{\mathcal{H}_m\} \end{bmatrix}$$

is the lattice generator matrix [26], V_s is the volume of the sphere with unit radius in $\mathbb{R}^s$, s is twice the input dimension and $N_{\mathcal{L}}$ is the list size. The lattice sphere decoding algorithm is used to search the N closest vectors to the center point (ZF or MMSE). Whenever a new point is found, we add it into the list if the list is not full. If the list is full, we compare its radius with the largest radius in the list and drop the larger one. Finally, we use list $\mathcal{L}$, the soft information similar to (12), and the summation set replaced by $\mathbb{B}_{k,\pm1} \cap \mathcal{L}$) to approximate APPs. The soft output in (12) can be further simplified by the max-log approximation for computation efficiency [22], [25]. Here, we omit the simplified form due to the space limitation. It is worthwhile to emphasize that LSD is only suitable for linear MIMO-encoder mapping.

2) SISO Channel Decoder: The soft-in-soft-out (SISO) channel decoder is implemented by the well-known BCJR algorithm [23]. The soft-input of the SISO channel decoder comes from the output of LSD and the soft-output of the SISO channel decoder will serve as the prior of LSD.

C. Recursive EM Channel Estimator

Let $\mathcal{C}_m$, $\mathcal{I}_m$ and Θ_m be the complete data, incomplete data and channel parameters at time m, respectively. The two steps of the EM algorithm can be written as

1) The E-step: computing the Kullback-Leibler (K-L) measure

$$Q_m(\Theta_m|\hat{\Theta}_{m|m}^{(l-1)}) = \mathrm{E}\{\log p(\mathcal{C}_m|\Theta_m)|\mathcal{I}_m, \hat{\Theta}_{m|m}^{(l-1)}\}. \tag{14}$$

2) The M-step: maximizing the K-L measure for the new estimate

$$\hat{\Theta}_{m|m}^{(l)} = \arg\max_{\Theta_m} Q_m(\Theta_m|\hat{\Theta}_{m|m}^{(l-1)}), \tag{15}$$

where superscript (l) is the total number of iterations at each time step. Let $\mathcal{Y}_m = \{\mathbf{y}_i : i = 0, \cdots, m\}$ be the received information and $\mathcal{S}_m = \{\mathbf{s}_i : i = 0, \cdots, m\}$ be the transmitted OFDM symbols up to time m. Let $\mu_{i|m} = \mathrm{E}\{\mathbf{h}_i|\mathcal{Y}_m, \hat{\Theta}_{m|m}^{(l-1)}\}$, and $\Sigma_{i,j|m}$, $i, j = 0, \cdots, m$, be the cross-covariance matrix between $\mathbf{h}_i$ and $\mathbf{h}_j$. For the convenience of expression, we define $\hbar_m = [\mathbf{h}_m^T, \cdots, \mathbf{h}_0^T]^T$ with conditional mean $\mathcal{U}_{m|m} = [\mu_{m|m}^T, \cdots, \mu_{0|m}^T]^T$ and conditional covariance matrix $\Gamma_{i|m} = [\Sigma_{i-j,i-l}]_{j+1,l+1}$, which is a block matrix with the block at the $(j+1)$th row and $(l+1)th$ column equal to $\Sigma_{i-j,i-l}$. In this problem, the complete and incomplete data at time m are denoted by $\mathcal{C}_m = \{\mathcal{Y}_m, \hbar_m, \mathcal{S}_m\}$ and $\mathcal{I}_m = \{\mathcal{Y}_m\}$, respectively, and the parameter set of interest is $\{\mathcal{U}_{m|m}, \mathbf{C}\}$. The K-L measure with respect to $\mathcal{U}_{m|m}$ in this case is given by [14]

$$\begin{aligned} &Q_m(\mathcal{U}_{m|m}, \hat{\mathbf{C}}_{|m}|\hat{\Theta}_{m|m}^{(l-1)}) \\ &= \mathrm{E}\{\log p(\mathcal{Y}_m, \mathcal{S}_m, \hbar_m|\mathcal{U}_{m|m}, \hat{\mathbf{C}}_{|m})|\mathcal{Y}_m, \hat{\Theta}_{m|m}^{(l-1)}\} \\ &= \textstyle\sum_{i=1}^m \{-\rho_i(\tilde{\mu}_{i|m}, \mu_{i|m}) - \eta_i(\mu_{i|m}, \bar{\mu}_{i|m}) + c_i\} + e \\ &-[\mu_{0|m} - \mu_{0|-1}]^H \Sigma_{0|-1}^{-1}[\mu_{0|m} - \mu_{0|-1}] - \rho_0(\mu_{0|m}, \tilde{\mu}_{0|m}) \end{aligned} \tag{16}$$

where c_i and e are constant terms,

$$\begin{aligned} \rho_i(\tilde{\mu}_{i|m}, \mu_{i|m}) &= [\tilde{\mu}_{i|m} - \mu_{i|m}]^H \tilde{\mathbf{C}}_{i|m}[\tilde{\mu}_{i|m} - \mu_{i|m}], \\ \eta_i(\mu_{i|m}, \bar{\mu}_{i|m}) &= [\mu_{i|m} - \bar{\mu}_{i|m}]^H (\mathbf{B}\mathbf{B}^H)^{-1}[\mu_{i|m} - \bar{\mu}_{i|m}], \end{aligned}$$

and $\bar{\mu}_{i|m} = \mathbf{F}\mu_{i-1|m}$ and $\mu_{0|-1}$ and $\Sigma_{0|-1}$ are initial values obtained from the training mode. $\tilde{\mathbf{C}}_{i|m}$ and $\tilde{\mu}_{i|m}$ come from the synthetic system that is characterized by

$$\tilde{\mathbf{C}}_{i|m} = \mathrm{E}\{\mathcal{D}(\mathbf{s}_i)^H \hat{\mathbf{C}}_{|m}^{-1}\mathcal{D}(\mathbf{s}_i)|\mathcal{Y}_m, \hat{\Theta}_{m|m}^{(l-1)}\}, \tag{17}$$

$$\tilde{\mathcal{D}}_{i|m} = \mathrm{E}\{\mathcal{D}(\mathbf{s}_i)^H \hat{\mathbf{C}}_{|m}^{-1}|\mathcal{Y}_m, \hat{\Theta}_{m|m}^{(l-1)}\}, \tag{18}$$

with the synthetic log likelihood function of $\{\mathbf{h}_i, \hat{\mathbf{C}}_{|m}\}$ at time i defined as

$$\begin{aligned} &SLL_i(\mathbf{h}_i, \hat{\mathbf{C}}_{|m}|\mathcal{Y}_m, \hat{\Theta}_{m|m}^{(l-1)}) = -\log(\pi^{L_r}\det(\hat{\mathbf{C}}_{|m})) \\ &-\mathrm{E}\{(\mathbf{y}_i - \mathcal{D}(\mathbf{s}_i)\mathbf{h}_i)^H \hat{\mathbf{C}}_{|m}^{-1}(\mathbf{y}_i - \mathcal{D}(\mathbf{s}_i)\mathbf{h}_i)|\mathcal{Y}_m, \hat{\Theta}_{m|m}^{(l-1)}\} \\ &= -[\tilde{\mathbf{C}}_{i|m}^{-1}\tilde{\mathcal{D}}_{i|m}\mathbf{y}_i - \mathbf{h}_i]^H \tilde{\mathbf{C}}_{i|m}^{-1}[\tilde{\mathbf{C}}_{i|m}^{-1}\tilde{\mathcal{D}}_{i|m}\mathbf{y}_i - \mathbf{h}_i] + const., \end{aligned}$$

which can be maximized with respect to $\mathbf{h}_i$ by setting $\tilde{\mu}_{i|m} = \tilde{\mathbf{C}}_{i|m}^{-1}\tilde{\mathcal{D}}_{i|m}\mathbf{y}_i$. Therefore, $\tilde{\mu}_{i|m}$ is referred to as the synthetic ML estimate of $\mathbf{h}_i$. Notice that, in (16), except for the constant terms, all other terms are in a Gaussian quadratic form. Thus, maximizing (16) with respect to $\mathcal{U}_{m|m}$ is equivalent to seeking the recursive expression of $\hat{\mathcal{U}}_{i|m}$, which is given by a series of Kalman-like equations:

$$\hat{\mathcal{U}}_{i|m} = \hat{\mathcal{U}}'_{i|m} + \mathbf{G}_{i|m}(\tilde{\mu}_{i|m} - \mathbf{J}_i^H \hat{\mathcal{U}}'_{i|m}), \tag{19}$$

$$\hat{\Gamma}_{i|m} = \hat{\Gamma}'_{i|m} - \mathbf{G}_{i|m}\mathbf{J}_i^H \hat{\Gamma}'_{i|m}, \tag{20}$$

$$\mathbf{G}_{i|m} = \hat{\Gamma}'_{i|m}\mathbf{J}_i(\tilde{\mathbf{C}}_{i|m}^{-1} + \mathbf{J}_i^H \hat{\Gamma}'_{i|m}\mathbf{J}_i)^{-1}, \tag{21}$$

where $\mathbf{J}_i = [\mathbf{I}_{n_t n_r \times n_t n_p}, \mathbf{0}_{n_t n_p \times i n_t n_p}]^H$ and $\hat{\mathcal{U}}'_{i|m}$ and $\hat{\Gamma}'_{i|m}$ are obtained via

$$\begin{aligned}\hat{\mathcal{U}}'_{i|m} &= [\bar{\mu}^T_{i|m}, \mathcal{U}^T_{i-1|m}]^T = [(\mathbf{F}\mu_{i-1|m})^T, \mathcal{U}^T_{i-1|m}]^T,\\ \hat{\Gamma}'_{i|m} &= \begin{bmatrix} \mathbf{F}\hat{\Sigma}_{i-1,i-1|m}\mathbf{F}^H & \mathbf{F}\mathbf{V}^H_{i-1|m} \\ \mathbf{V}_{i-1|m}\mathbf{F}^H & \hat{\Gamma}_{m-1|m} \end{bmatrix} + \mathbf{J}_i\mathbf{B}\mathbf{B}^H\mathbf{J}^H_i,\end{aligned}$$

where $\mathbf{V}_{i-1|m} = \mathrm{E}\{[\mathcal{U}_{i-1|m} - \hat{\mathcal{U}}_{i-1|m}][\mu_{i-1|m} - \hat{\mu}_{i-1|m}]^H\}$. The dimension of matrix $\tilde{\mathbf{C}}^{-1}_{i|m} + \mathbf{J}^H_i \hat{\Gamma}'_{i|m}\mathbf{J}_i$, which requires inversion, is still $n_t n_r \times n_t n_r$. Although the recursive process is similar to the traditional (or Kalman) smoothing, the main difference is that a synthetic approach is used to average over the *a posterior* probabilities of $\{\mathbf{s}_m\}$. Therefore, the recursive EM algorithm can be viewed as a Kalman smoothing algorithm with soft decision feedback. Due to the interleaver between the outer channel encoder and the OFDM symbol mapper, soft decision feedback occurred at the end of the outer soft decoding of the entire data burst. That is, we use the whole M blocks as the incomplete data, and the subscript m in the above algorithm should change to $M-1$. The covariance estimator can be approximated by

$$\hat{\mathbf{C}}_{|M-1} = \frac{1}{M}\sum_{i=0}^{M-1}[\mathbf{y}_i - \hat{\mathcal{D}}_i\hat{\mathbf{h}}_i][\mathbf{y}_i - \hat{\mathcal{D}}_i\hat{\mathbf{h}}_i]^H, \qquad (22)$$

where $\hat{\mathcal{D}}_i = \mathrm{E}\{\mathcal{D}(\mathbf{s}_i)|\mathcal{Y}_{M-1}, \Theta^{(l-1)}_{i|M-1}\}$, and the initial value of $\hat{\mathbf{h}}_i$ is the sequential MMSE estimate obtained by the Kalman filter using pilot tones and updated based on the previous estimate. If we assume perfect PO/CFO compensation, the statistics of $\mathbf{n}_m$ in (8) is equal to $\mathbf{w}_m$ in (1) because of the unitary property of DFT matrix $\mathbf{W}$. Thus, if the noise level is already known at the receiver, the synthetic system (17) can be further simplified and the noise update in (22) can be omitted.

D. Summary of Proposed Algorithm

The proposed algorithm can be summarized as follows.

1) Initial estimates by pilots tones:

$$\hat{\mathbf{h}}_m = \text{Kalman filter}\{\mathcal{Y}^p_m, \mathcal{S}^p_m\}. \qquad (23)$$

2) Soft MIMO demodulation by LSD:

$$\{\lambda[b]\} = \mathrm{LSD}\{\mathcal{Y}_m, \{\mathbf{h}_m\}\} \qquad (24)$$

3) Soft decoding by SISO outer code decoder:

$$\{\lambda[d]\} = \mathrm{SISO/BCJR}\{\{\lambda[b]\}\}. \qquad (25)$$

4) Start iteration and form synthetic system (17).
5) Refine channel estimates by EM, or equivalent soft decision directed Kalman smoothing (19).
6) Repeat 2).

V. Simulation Results

We consider a bit-interleaved MIMO-OFDM system employing $N = 64$ subcarriers with $n_t = 2$ and $n_r = 2$ antennas and operating in the 5GHz frequency band. The channel response of each transmit/receive antenna pair has $L = 4$ paths and each is generated according to Jakes' model with normalized Doppler fading rate 0.04. The channel coefficients are modeled as independent complex-valued Gaussian random variables with zero-mean and an exponential power delay profile. It is assumed that the instability of the transmit/receive oscillators results in normalized CFO=0.01 within a subcarrier and PO=0.05. We use $N_p = 50$ particles to approximate the joint distribution of PO/CFO and channel taps in the training mode.

Fig. 2 and Fig. 3 show the mean square error (MSE) of the frequency and the phase offset estimates as a function of the SNR value in the training mode, respectively. Fig. 4 depicts the behavior of the steady-state MSE of channel taps estimates obtained by EM/SD in the data mode after compensating PO and CFO. Since the MSE values of channel taps are usually larger then those of PO and CFO, it is worthwhile to trade the complexity of marginalization for the performance enhancement by RBPF.

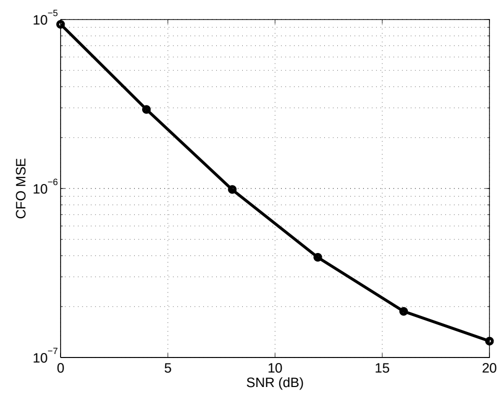

Fig. 2. The MSE of the CFO estimate as a function of SNR.

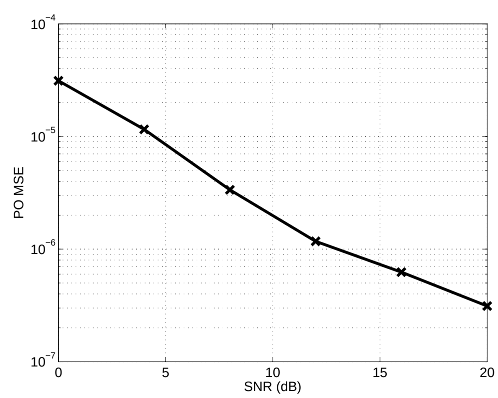

Fig. 3. The MSE of the PO estimate as a function of SNR.

In Fig. 5, we compare the performance of three different schemes: the Kalman filtering with the hard decision feedback (Kalman/HD), the data-centric MAP-EM receiver (MAP-EM/SD) [20] and the proposed algorithm (EM/(Kalman)/SD). Both soft decoding schemes outperform the hard decision feedback scheme by a gain of 3dB at the BER rate of

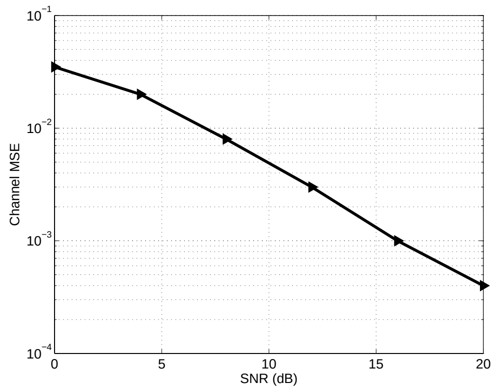

Fig. 4. The MSE of the EM/SD channel estimate as a function of SNR.

10^{-2}. The performance of the proposed algorithm is slightly better than the data-centric MAP-EM scheme due to the time correlation of the channel variation being exploited in the soft-decision Kalman filtering scheme.

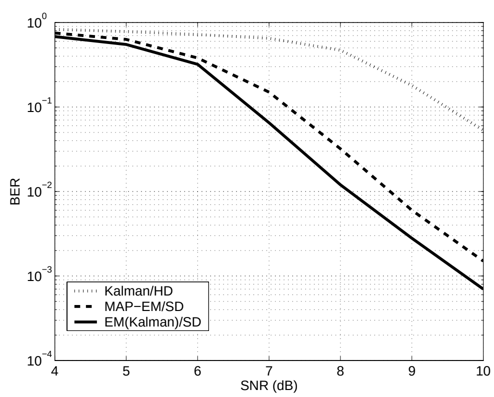

Fig. 5. The bit error rate (BER) performance comparison of three schemes.

VI. Conclusion

A solution to joint channel tracking and symbol detection for bit-interleaved MIMO-OFDM systems was examined. In the training mode, the RBPF scheme was used to estimate the PO and the CFO. In the data mode, we applied a recursive EM procedure in conjunction with soft decoding for joint minimum mean squared error (MMSE) channel estimation and the maximum *a posterior* (MAP) probability symbol detection. Simulation results showed that the proposed algorithm have robust performance in the presence of PO/CFO over time-varying channels.

References

[1] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communications: Performance criterion and code construction," *IEEE Trans. on Information Theory*, vol. 44, pp. 744–765, Mar. 1998.

[2] G. Caire, E. Taricco, and E. Biglieri, "Bit-interleaved coded modulation," *IEEE Trans. on Information Theory*, vol. 44, no. 3, pp. 927–946, May 1998.

[3] D. Rende, X. Li, and T. F. Wong, "Bit-interleaved space-frequency coded modulation with iterative decoding for OFDM systems," in *Proc. IEEE MILCOM*. OCT. 2003, 2003.

[4] I. Lee, A. M. Chan, and C. W. Sundberg, "Space-time bit-interleaved coded modulation for OFDM systems in wireless LAN applications," in *Proc. IEEE ICC*. May 2003, 2003, pp. 3413–3417.

[5] D. Petrovic, W. Rave, and G. Fettweis, "Phase noise supression in OFDM using a Kalman filter," in *Proc. WPMC 2003*.

[6] H. Liu and U. Tureli, "A high-efficiency carrier estimator for OFDM communications," *IEEE Communication Letters*, vol. 2, pp. 104–106, Apr. 1998.

[7] B. Chen, "Maximum likelihood estimation of OFDM carrier frequency offset," *IEEE Singal Processing Letters*, vol. 9, no. 4, pp. 123–126, Apr. 2002.

[8] M. Sanjeev Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. on Signal Processing*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[9] A. Doucet, N. de Freitas, K. Murphy, and S. Russell, "Rao-Blackwellised Particle Filtering for Dynamic Bayesian Networks," in *Sixteenth Conference on Uncertainty in Artificial Intelligence*. Stanford, CA, 2000, pp. 176–183.

[10] E. Punskaya, C. Andrieu, A. Doucet, and W. J. Fitzgerald, "Particle filtering for demodulation in fading channels with non-Gaussian additive noise," *IEEE Trans. on Communications*, vol. 49, no. 4, pp. 579–582, Apr. 2002.

[11] H. Zamiri-Jafarian and S. Pasupathy, "EM-based recursive estimation of channel parameters," *IEEE Trans. on Communications*, vol. 47, no. 9, pp. 1297–1302, Sept. 1999.

[12] R. Raheli, A. Polydoros, and C. K. Tzou, "Per-survivor processing: a general approach to MLSE in uncertain environment," *IEEE Trans. on Communications*, vol. 43, no. 2/3/4, pp. 354–364, 1995.

[13] M. Nissila and S. Pasupathy, "Adaptive Bayesian and EM-based detectors for frequency-selective fading channels," *IEEE Trans. on Communications*, vol. 51, no. 8, pp. 1325–1336, Aug. 2003.

[14] S.-H. Wu, U. Mitra, and C.-C. J. Kuo, "Graph representation for joint channel estimation and symbol detection," in *Proc. IEEE GLOBECOM*. Dallas, Texas, Nov 2004.

[15] F.-H. Chiu, S.-H. Wu, and C.-C. J. Kuo, "Robust Joint Channel Estimation and Symbol Detection over MIMO Channels Using EM Algorithm," in *Proc. IEEE VTC*. Dallas, Sep. 2005 (accepted to present).

[16] F.-H. Chiu, S.-H. Wu, and C.-C. J. Kuo, "Joint MIMO Channel Tracking and Symbol Detection with EM Algorithm and Soft Decoding," in *Proc. IEEE GLOBECOM*. St. Louis, Nov. 2005 (accepted to present).

[17] T.Y. Al-Naffouri, O. Awoniyi, O. Oteri, and A. Paulraj, "Receiver design for MIMO-OFDM transmission over time variant channels," in *Proc. IEEE GLOBECOM*. San Franciso, California, Nov, 2004, pp. 2487–2492.

[18] K. J. Kim and J. Yue, "Joint channel estimation and data detection algorithms for MIMO-OFDM systems," in *36th Asilomar Conference*. Montery, California, Nov, 2002, pp. 1857–1861.

[19] D. Schafhuber, G. Matz, and F Hlawatsch, "Kalman tracking of time-varying channels in wireless MIMO-OFDM systems," in *36th Asilomar Conference*. Montery, California, Nov, 2002, pp. 1261–1265.

[20] B. Lu, X. Wang, and Y. Li, "Iterative receivers for space-time block-coded OFDM systems in dispersive fading channels," *IEEE Trans. Wireless Communication*, vol. 1, no. 2, pp. 213–225, Apr. 2002.

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

[22] B. M. Hochwald and S. ten Brink, "Achieving near-capacity on a multiple-antenna channl," *IEEE Trans. on Communications*, vol. 51, no. 3, pp. 389–399, Mar. 2003.

[23] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for mimizing symbol error rate," *IEEE Trans. on Information Theory*, vol. 20, pp. 284–287, Mar. 1974.

[24] A. Doucet, N. de Freitas, and N. Gordon, *Sequential Monte Carlo Methods in Practice*, Springer-Verlag, New York, 2001.

[25] J. Boutros, N. Greeset, L. Brunel, and M. Fossorier, "Soft-input soft-output lattice sphere decoder for linear channels," in *Proc. IEEE GLOBECOM*. San Francisco, California, Nov 2003.

[26] J. H. Conway and N. J. Sloane, *Sphere packings, Lattices, and Groups*, Springer-Verlag, New York, 3rd edition, 1998.