2014 IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)

EXPECTATION PROPAGATION APPROACH TO JOINT CHANNEL ESTIMATION AND DECODING FOR OFDM SYSTEMS

Sheng Wu[1,4], *Linling Kuang*[2,4], *Zuyao Ni*[3], *Jianhua Lu*[1], *Defeng (David) Huang*[5], *Qinghua Guo*[5,6] *

[1]Department of Electronic Engineering, Tsinghua University
[2]Tsinghua Space Center, Tsinghua University
[3]Research Institute of Information Technology, Tsinghua University
wu-s08@mails.tsinghua.edu.cn, {kll, lhh-dee, nzy}@tsinghua.edu.cn
[4]Shenzhen Key Laboratory of Wireless Broadband Communication and Signal Processing
[5]School of Electrical, Electronic and Computer Engineering, The University of Western Australia
{huangdf, qhguo}@ee.uwa.edu.au
[6]School of Electrical, Computer and Telecommunications Engineering, University of Wollongong

ABSTRACT

We propose a message-passing algorithm of joint channel estimation and decoding for OFDM systems, where expectation propagation is exploited to deal with channel estimation. Specially, the message updating is formulated into a recursive form. As a result, for system with K subcarriers and L channel taps, only $\mathcal{O}(K+L)$ messages need to be tracked, and meanwhile they can be efficiently calculated using FFT with complexity $\mathcal{O}(K|\mathcal{A}|+K\log_2 K)$, where $|\mathcal{A}|$ denotes the constellation size. Numerical experiments show that our algorithm achieves BER performance within 0.5 dB of the known-channel bound.

Index Terms— Expectation propagation, joint channel estimation and decoding, message passing, OFDM

1. INTRODUCTION

Factor graph is convenient to define the structure of receiver performing joint channel estimation and decoding [1]. However, exact sum-product algorithm (SPA) [2] for joint channel estimation and decoding is computationally infeasible, as the representation of continuous channel states increases exponentially. To overcome this problem, different approximate approaches have been proposed in [3–11]. Monte Carlo methods were used to represent the distribution of channel states, but at high cost [3, 4]. Various approximate message-passing algorithms have been proposed in [5–10]. Recently, Riegler *et al* have derived a generic message-passing algorithm that merges belief propagation (BP) and the mean-field (MF) approximation (BP-MF) [11], and applied it to channel estimation in OFDM systems [12]. Although its performance is excellent, the BP-MF has a high complexity as large matrices need to be inverted. We also note that a low-complexity version of the BP-MF algorithm has been proposed in [13]; however, its performance is degraded in the meantime.

*This work was partially supported by the National Nature Science Foundation of China (Grant Nos. 91338101, 61021001, and 61132002), the National Basic Research Program of China (Grant No. 2013CB329001), the Co-innovation Laboratory of Aerospace Broadband Network Technology, and also by the Australian Research Council's Discovery Projects DP1093000, DECRA Grant DE120101266.

In this paper, we utilize the principle of expectation propagation [14] to make the messages of channel state tractable. To further reduce complexity, the message updating for the channel estimation is reformulated as recursions. It leads to a reduced number of messages need to be tracked from $\mathcal{O}(KL)$ to $\mathcal{O}(K+L)$ and efficient message updating using fast Fourier transform (FFT), where K denotes the number of subcarriers and L denotes the channel taps length. The resulting algorithm has a computational complexity of only $\mathcal{O}(K|\mathcal{A}|+K\log_2 K)$, where $|\mathcal{A}|$ denotes the constellation size.

Throughout the paper, we use the following notations. The superscript T and $*$ denote the transpose operation and the conjugate operation, respectively. Also, $\ln(\cdot)$ denotes the natural logarithm; $\mathcal{N}_{\mathbb{C}}(x;m,v)\triangleq(\pi v)^{-1}\exp(-|x-m|^2/v)$ denotes a complex Gaussian function; and $\boldsymbol{h}\setminus h_l$ denotes all the components in $\boldsymbol{h}$ with h_l excluded. Furthermore, $\mathbb{E}_{p(x)}[\cdot]$ denotes the statistical expectation operation with respect to the distribution $p(x)$.

2. SYSTEM MODEL

Consider an OFDM system with K subcarriers, each modulated by a symbol chosen from a 2^Q-ary constellation set $\mathcal{A}$. At the transmitter, a bit vector $\boldsymbol{b}$ is encoded by a rate-R code, interleaved by a random interleaver, and multiplexed with some training bits $\boldsymbol{c}^{(p)}$, resulting in a vector $\boldsymbol{c}=[\boldsymbol{c}[1],\ldots,\boldsymbol{c}[N]]^{\mathsf{T}}$, where $\boldsymbol{c}[i]=[\boldsymbol{c}_1[i],\ldots,\boldsymbol{c}_K[i]]^{\mathsf{T}}, 1\le i\le N$, $\boldsymbol{c}_k[i]=[c_k^1[i],\ldots,c_k^Q[i]]^{\mathsf{T}}, 1\le k\le K, c_k^q\in\{0,1\}, 1\le q\le Q$ and N denotes the number of OFDM symbols, and then K symbols, $\boldsymbol{x}[i]=[x_1[i],\ldots,x_K[i]]^{\mathsf{T}}$, are generated by mapping each sub-vector $\boldsymbol{c}_k[i]$ onto a symbol $x_k[i]\in\mathcal{A}$. Before sent through a channel $\boldsymbol{h}[i]=[h_1[i],\ldots,h_L[i]]^{\mathsf{T}}$, the symbol vector $\boldsymbol{x}[i]$ is OFDM modulated and a cyclic prefix is added. At the receiver, the frequency-domain observations $\boldsymbol{y}[i]=[y_1[i],\ldots,y_K[i]]^{\mathsf{T}}$ with respect to the ith OFDM symbol are written as

$$y_k[i]=x_k[i]\sum_{l=1}^{L}\Phi_{k,l}h_l[i]+n_k[i], k=1,\ldots,K, \qquad (1)$$

where $\Phi_{k,l}=\exp(-j2\pi kl/K)$ denotes the (k,l)th entry in the discrete Fourier transform matrix $\boldsymbol{\Phi}\in\mathbb{C}^{K\times K}$, and $n_k[i]$ denotes the i.i.d Gaussian noise with zero mean and variance σ_n^2. For notational simplicity, the OFDM symbol index i is henceforth omitted.

3. FACTOR GRAPH REPRESENTATION

Based on the frequency-domain observations $\boldsymbol{y}$ and the *a priori* log likelihood ratios (LLRs) $\{\lambda^a(c_k^q) \triangleq \ln[p(c_k^q=1)/p(c_k^q=0)]\}$ fed back from the decoder or specified by the training bits, our task is to generate the LLRs $\{\lambda^e(c_k^q) = \ln[p(c_k^q=1|\boldsymbol{y})/p(c_k^q=0|\boldsymbol{y})] - \lambda^a(c_k^q)\}$, where the *a posteriori* marginal probability $p(c_k^q|\boldsymbol{y})$ is obtained by

$$p\left(c_k^q|\boldsymbol{y}\right) \propto \sum_{\boldsymbol{x},\boldsymbol{c}\backslash c_k^q} \int_{\boldsymbol{h}} p\left(\boldsymbol{c},\boldsymbol{x},\boldsymbol{y},\boldsymbol{h}\right). \tag{2}$$

However, the exact evaluation of $p(c_k^q|\boldsymbol{y})$ is computationally prohibitive for the problem sizes of interest, thus we will evaluate it approximately by message-passing algorithm. For the presentation of factor graph and the message-passing algorithm, we will use the same convention as in [15], to which we refer the reader for an in-depth review.

As $\boldsymbol{c} \to \boldsymbol{x} \to \boldsymbol{y} \leftarrow \boldsymbol{h}$ is a Markov chain, the joint probability $p(\boldsymbol{c},\boldsymbol{x},\boldsymbol{y},\boldsymbol{h})$ can be factorized into

$$p\left(\boldsymbol{c},\boldsymbol{x},\boldsymbol{y},\boldsymbol{h}\right) = p(\boldsymbol{c})p(\boldsymbol{x}|\boldsymbol{c})p(\boldsymbol{y}|\boldsymbol{h},\boldsymbol{x})p(\boldsymbol{h}). \tag{3}$$

Using the independence of the symbols associated with different subcarriers, $p(\boldsymbol{x}|\boldsymbol{c})$ in (3) can be further factorized into $p(\boldsymbol{x}|\boldsymbol{c}) = \prod_k p(x_k|\boldsymbol{c}_k)$, where $p(x_k|\boldsymbol{c}_k) = \delta(\psi(\boldsymbol{c}_k) - x_k)$ denotes the deterministic mapping $x_k = \psi(\boldsymbol{c}_k)$ and $\delta(\cdot)$ is the Kronecker delta function. Channel taps $\boldsymbol{h}$ are assumed to be independent Gaussian random variables, thus the *a priori* probability $p(\boldsymbol{h})$ can be factorized into $p(\boldsymbol{h}) = \prod_l \mathcal{N}_{\mathbb{C}}(h_l; m_{h_l}^a, v_{h_l}^a)$, where $m_{h_l}^a$ and $v_{h_l}^a$ denote the *a priori* mean and variance of the lth channel tap h_l, respectively. Finally, the channel transition function $p(\boldsymbol{y}|\boldsymbol{h},\boldsymbol{x})$ can be factorized into $p(\boldsymbol{y}|\boldsymbol{h},\boldsymbol{x}) = \prod_k f_k(y_k|\boldsymbol{h},x_k)$, where $f_k(y_k|\boldsymbol{h},x_k) = \mathcal{N}_{\mathbb{C}}(y_k; x_k\sum_l \Phi_{k,l}h_l, \sigma_n^2)$. The probabilistic structure characterized by (3) is illustrated by Fig. 1, where Ψ_k denotes the mapping constraint $p(x_k|\boldsymbol{c}_k)$, f_k denotes the channel transition function $f_k(y_k|\boldsymbol{h},x_k)$, and "=" denotes the cloning node.

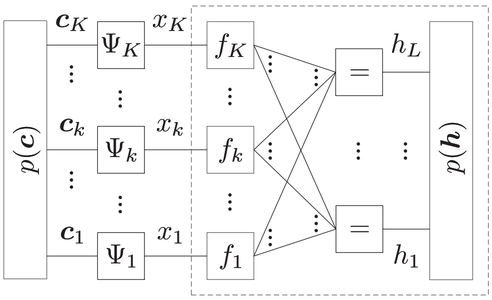

Fig. 1. Factor graph of an OFDM system with K subcarrier and L channel taps.

As the factor graph in Fig. 1 is loopy, we consider a message passing from the left to the right and then back again as one turbo iteration. During a single turbo iteration, there may be several inner iterations within the dashed block in Fig. 1 and the decoder. Specially, an inner iteration within the dashed block is also defined as a message passing from the nodes $\{f_k\}$ to the cloning nodes and back again. To achieve joint channel estimation and decoding, the messages are passed over the global factor graph including the code structure, namely, the inner messages of decoder calculated in the last turbo iteration are available for the decoding in the next turbo iteration.

4. MESSAGE PASSING BASED ON EXPECTATION PROPAGATION

We will formulate the message passing in a single turbo iteration, where the inner iteration in the dashed block are indexed by t. According to the SPA, the message passed rightward from the mapping node Ψ_k is given by

$$\mu_{\Psi_k\to f_k}(x_k) = \prod_q p(c_k^q), \tag{4}$$

where $p(c_k^q) = \exp(c_k^q\lambda^a(c_k^q))/[1+\exp(\lambda^a(c_k^q))]$. Given the message $\mu_{\Psi_k\to f_k}(x_k)$ and the messages $\{\mu_{h_l\to f_k}^{t-1}(h_l)\}$, a local belief of h_l is defined at the node f_k:

$$\beta_k^t(h_l) = \frac{1}{\eta_k^t}\int_{\boldsymbol{h}\backslash h_l}\sum_{x_k\in\mathcal{A}} f_k(x_k,\boldsymbol{h})\mu_{\Psi_k\to f_k}(x_k)\prod_{l'}\mu_{h_{l'}\to f_k}^{t-1}(h_{l'}), \tag{5}$$

where $\mu_{h_{l'}\to f_k}^{t-1}(h_{l'}) = \mathcal{N}_{\mathbb{C}}(m_{h_{l'}\to f_k}^{t-1}, v_{h_{l'}\to f_k}^{t-1})$ is shown later as (18). From $\int_{h_l}\beta_k^t(h_l) = 1$, the normalization constant η_k^t is given by

$$\eta_k^t = \sum_{x_k\in\mathcal{A}}\mu_{\Psi_k\to f_k}(x_k)\mathcal{N}_{\mathbb{C}}\big(0; x_k z_{f_k}^t(x_k), |x_k|^2\tau_{f_k}^t(x_k)\big), \tag{6}$$

where

$$z_{f_k}^t(x_k) = \frac{y_k}{x_k} - \sum_l \Phi_{k,l} m_{h_l\to f_k}^{t-1}, \tag{7}$$

$$\tau_{f_k}^t(x_k) = \frac{\sigma_n^2}{|x_k|^2} + \sum_l v_{h_l\to f_k}^{t-1}. \tag{8}$$

By completing the square in the exponent of Gaussian functions, $\beta_k^t(h_l)$ is written as a mixture of Gaussian functions, i.e.,

$$\beta_k^t(h_l) = \sum_{x_k\in\mathcal{A}} q(x_k)\mathcal{N}_{\mathbb{C}}(\Phi_{k,l}h_l; m_{\Phi_{k,l}h_l}^t(x_k), v_{\Phi_{k,l}h_l}^t(x_k)), \tag{9}$$

where the mixture weight $q\left(x_k\right)$ and the component parameters are given by

$$q(x_k) = \frac{1}{\eta_k^t}\mu_{\Psi_k\to f_k}(x_k)\mathcal{N}_{\mathbb{C}}\big(0; x_k z_{f_k}^t(x_k), |x_k|^2\tau_{f_k}^t(x_k)\big), \tag{10}$$

$$m_{\Phi_{k,l}h_l}^t(x_k) = v_{h_l\to f_k}^{t-1}\epsilon_k^t(x_k) + \Phi_{k,l}m_{h_l\to f_k}^{t-1}, \tag{11}$$

$$v_{\Phi_{k,l}h_l}^t(x_k) = \big(1 - v_{h_l\to f_k}^{t-1}/\tau_{f_k}^t(x_k)\big)v_{h_l\to f_k}^{t-1}, \tag{12}$$

where $\epsilon_k^t(x_k) \triangleq z_{f_k}^t(x_k)/\tau_{f_k}^t(x_k)$. To keep the messages in the dashed block analytical, we project the local belief $\beta_k^t(h_l)$ into a Gaussian function $\hat{\beta}_k^t(h_l) \triangleq \mathcal{N}_{\mathbb{C}}\big(h_l; \hat{m}_{\Phi_{k,l}h_l}, \hat{v}_{\Phi_{k,l}h_l}\big)$. When the criterion of minimum KL divergence, $\mathrm{KL}\big(\beta_k^t(h_l), \hat{\beta}_k^t(h_l)\big)$, is employed, the projection reduces to matching the moments of $\hat{\beta}_k^t(h_l)$ and $\beta_k^t(h_l)$ [16], i.e.,

$$\begin{aligned}\hat{m}_{\Phi_{k,l}h_l}^t &= \mathbb{E}_{q(x_k)}\big[m_{\Phi_{k,l}h_l}^t(x_k)\big]\\ &= \Phi_{k,l}m_{h_l\to f_k}^{t-1} + v_{h_l\to f_k}^{t-1}\mathbb{E}_{q(x_k)}\big[\epsilon_k^t(x_k)\big]\end{aligned} \tag{13}$$

$$
\begin{aligned}
\hat{v}^t_{\Phi_{k,l}h_l} &= \mathbb{E}_{q(x_k)}\left[v^t_{\Phi_{k,l}h_l}(x_k) + |m^t_{\Phi_{k,l}h_l}(x_k)|^2\right] - |\hat{m}^t_{\Phi_{k,l}h_l}|^2 \\
&= v^{t-1}_{h_l\to f_k}(1-\zeta_k^t v^{t-1}_{h_l\to f_k}),
\end{aligned} \tag{14}
$$

where $\zeta_k^t \triangleq \mathbb{E}_{q(x_k)}\left[1/\tau^t_{f_k}(x_k) - |\epsilon_k^t(x_k)|^2\right] + \left|\mathbb{E}_{q(x_k)}[\epsilon_k^t(x_k)]\right|^2$. With the approximate belief $\hat{\beta}_k^t(h_l)$, the message $\mu^t_{f_k\to h_l}(h_l)$ are then updated by $\hat{\beta}_k^t(h_l)/\mu^{t-1}_{h_l\to f_k}(h_l)$:

$$
\mu^t_{f_k\to h_l}(h_l) \propto \mathcal{N}_{\mathbb{C}}(\phi_{k,l}h_l; \overline{z}^t_{f_k\to h_l}, \overline{\tau}^t_{f_k\to h_l}), \tag{15}
$$

where

$$
\overline{\tau}^t_{f_k\to h_l} = 1/\zeta_k^t - v^{t-1}_{h_l\to f_k}, \tag{16}
$$

$$
\overline{z}^t_{f_k\to h_l} = \mathbb{E}_{q(x_k)}\left[\epsilon_k^t(x_k)\right]/\zeta_k^t + \Phi_{k,l}m^{t-1}_{h_l\to f_k}. \tag{17}
$$

As the approximate belief $\hat{\beta}_k^t(h_l)$ tries to cover all modes of the Gaussian mixture $\beta_k^t(h_l)$ [16], the variance of $\hat{\beta}_k^t(h_l)$, namely $\hat{v}^t_{\Phi_{k,l}h_l}$, may become so large that $\overline{\tau}^t_{f_k\to h_l} < 0$. If $\overline{\tau}^t_{f_k\to h_l} < 0$ happens, we set $\zeta_k^t = \mathbb{E}_{q(x_k)}\left[1/\tau^t_{f_k}(x_k)\right]$, which forces $\hat{\beta}_k^t(h_l)$ to cover less modes and guarantees $\overline{\tau}^t_{f_k\to h_l} > 0$. This is just a heuristic, but in our simulations it indeed avoids the instability of expectation propagation in general.

At the cloning node of h_l, the message $\mu^t_{h_l\to f_k}(h_l)$ is updated by

$$
\begin{aligned}
\mu^t_{h_l\to f_k}(h_l) &= p(h_l)\prod_{k'\neq k}\mu^t_{f_{k'}\to h_l}(h_l) \\
&= \mathcal{N}_{\mathbb{C}}(h_l; m^t_{h_l\to f_k}, v^t_{h_l\to f_k}).
\end{aligned} \tag{18}
$$

where

$$
v^t_{h_l\to f_k} = \left(1/v^t_{h_l} - 1/\overline{\tau}^t_{f_k\to h_l}\right)^{-1}, \tag{19}
$$

$$
m^t_{h_l\to f_k} = m^t_{h_l} + \frac{v^t_{h_l}\left(m^t_{h_l} - \Phi^*_{k,l}\overline{z}^t_{f_k\to h_l}\right)}{\overline{\tau}^t_{f_k\to h_l} - v^t_{h_l}}, \tag{20}
$$

and $v^t_{h_l}$ and $m^t_{h_l}$ are defined by

$$
v^t_{h_l} \triangleq \left(1/v^a_{h_l} + \sum_k 1/\overline{\tau}^t_{f_k\to h_l}\right)^{-1}, \tag{21}
$$

$$
m^t_{h_l} \triangleq v^t_{h_l}\left(m^a_{h_l}/v^a_{h_l} + \sum_k \Phi^*_{k,l}\overline{z}^t_{f_k\to h_l}/\overline{\tau}^t_{f_k\to h_l}\right). \tag{22}
$$

After the inner iteration is terminated, the message $\mu_{f_k\to\Psi_k}(x_k)$ is updated by

$$
\begin{aligned}
\mu_{f_k\to\Psi_k}(x_k) &= \int_{\boldsymbol{h}} f_k(x_k,\boldsymbol{h})\prod_l \mu^{t_{max}}_{h_l\to f_k}(h_l) \\
&\propto \mathcal{N}_{\mathbb{C}}\Big(y_k; x_k\sum_l \Phi_{k,l}m^{t_{max}}_{h_l\to f_k}, \sigma_n^2 + |x_k|^2\sum_l v^{t_{max}}_{h_l\to f_k}\Big),
\end{aligned} \tag{23}
$$

where t_{max} is the maximum number of inner iterations. Finally, the LLRs of coded bits corresponding to the symbol x_k are calculated by

$$
\lambda^e(c_k^q) = \ln\frac{\sum_{x_k\in\mathcal{A}_q^1}\mu^{t_{max}}_{f_k\to\Psi_k}(x_k)\prod_{q'}p(c_k^{q'})}{\sum_{x_k\in\mathcal{A}_q^0}\mu^{t_{max}}_{f_k\to\Psi_k}(x_k)\prod_{q'}p(c_k^{q'})} - \lambda^a(c_k^q), \tag{24}
$$

for $q = 1,\ldots,Q$. We summarize the message passing algorithm using expectation propagation for the inner iteration in the Alg. 1, which will be referred to as "EP".

Algorithm 1 The EP algorithm for the tth inner iteration.

Initialization: If $t = 1$, set $m^0_{h_l\to f_k} = m^a_{h_l}, v^0_{h_l\to f_k} = v^a_{h_l}, \forall k, \forall l$.
for $k = 1 \to K$ **do** $z^t_{f_k}(x_k) = \frac{y_k}{x_k} - \sum_l \Phi_{k,l}m^{t-1}_{h_l\to f_k}$,
  $\tau^t_{f_k}(x_k) = \frac{\sigma_n^2}{|x_k|^2} + \sum_l v^{t-1}_{h_l\to f_k}$, $\epsilon_k^t(x_k) = \frac{z^t_{f_k}(x_k)}{\tau^t_{f_k}(x_k)}$,
  $\zeta_k^t = \mathbb{E}_{q(x_k)}\left[\frac{1}{\tau^t_{f_k}(x_k)} - |\epsilon_k^t(x_k)|^2\right] + \left|\mathbb{E}_{q(x_k)}[\epsilon_k^t(x_k)]\right|^2$,
  for $l = 1 \to L$ **do** $\overline{\tau}^t_{f_k\to h_l} = 1/\zeta_k^t - v^{t-1}_{h_l\to f_k}$,
    $\overline{z}^t_{f_k\to h_l} = \mathbb{E}_{q(x_k)}\left[\epsilon_k^t(x_k)\right]/\zeta_k^t + \Phi_{k,l}m^{t-1}_{h_l\to f_k}$.
  end for
end for
for $l = 1 \to L$ **do** $v^t_{h_l} = \left(1/v^a_{h_l} + \sum_k 1/\overline{\tau}^t_{f_k\to h_l}\right)^{-1}$,
  $m^t_{h_l} = v^t_{h_l}\left(m^a_{h_l}/v^a_{h_l} + \sum_k \Phi^*_{k,l}\overline{z}^t_{f_k\to h_l}/\overline{\tau}^t_{f_k\to h_l}\right)$,
  for $k = 1 \to K$ **do** $v^t_{h_l\to f_k} = \left(1/v^t_{h_l} - 1/\overline{\tau}^t_{f_k\to h_l}\right)^{-1}$,
    $m^t_{h_l\to f_k} = m^t_{h_l} + \frac{v^t_{h_l}\left(m^t_{h_l} - \Phi^*_{k,l}\overline{z}^t_{f_k\to h_l}\right)}{\overline{\tau}^t_{f_k\to h_l} - v^t_{h_l}}$.
  end for
end for

4.1. LOW-COMPLEXITY IMPLEMENTATION

In the above EP algorithm, the number of messages need to be calculated is $\mathcal{O}(KL)$ and the complexity of one inner iteration is $\mathcal{O}(K|\mathcal{A}| + KL)$. By formulating the message passing into a recursive form, we can reduce the number of messages to $\mathcal{O}(K+L)$ and efficiently calculate them using FFT.

The parameter $v^t_{h_l\to f_k}$ shown in (19) can be approximated by $v^t_{h_l}$ shown in (21), then $\tau^t_{f_k}(x_k)$ becomes $\tau^t_{f_k}(x_k) = \frac{1}{|x_k|^2}\sigma_n^2 + \sum_l v^{t-1}_{h_l}$. Define $\overline{\tau}^t_{f_k} \triangleq \frac{1}{\zeta_k^t}$, then $\overline{\tau}^t_{f_k\to h_l}$ can be approximated by $\overline{\tau}^t_{f_k}$, and $v^t_{h_l}$ is finally written as $v^t_{h_l} = \left(\frac{1}{v^a_{h_l}} + \sum_k \frac{1}{\overline{\tau}^t_{f_k}}\right)^{-1}$. As a result, $m^t_{h_l\to f_k}$ and $m^t_{h_l}$ become

$$
m^t_{h_l\to f_k} = m^t_{h_l} - v^t_{h_l}\Phi^*_{k,l}\overline{z}^t_{f_k\to h_l}/\overline{\tau}^t_{f_k}, \tag{25}
$$

$$
m^t_{h_l} = v^t_{h_l}\left(m^a_{h_l}/v^a_{h_l} + \xi_l^t\right), \tag{26}
$$

where ξ_l^t is defined by $\xi_l^t \triangleq \sum_k \frac{1}{\overline{\tau}^t_{f_k}}\Phi^*_{k,l}\overline{z}^t_{f_k\to h_l}$. Define $\gamma_k^t \triangleq \frac{1}{\overline{\tau}^t_{f_k}}\sum_l v^t_{h_l}\overline{z}^t_{f_k\to h_l}$ and $\overline{z}^t_{f_k} \triangleq \frac{1}{\zeta_k^t}\mathbb{E}_{q(x_k)}[\epsilon^t_{f_k}(x_k)]$. Using (25), $z^t_{f_k}(x_k)$ and $\overline{z}^t_{f_k\to h_l}$ can be rewritten as

$$
z^t_{f_k}(x_k) = \frac{y_k}{x_k} + \gamma_k^{t-1} - \sum_l \Phi_{k,l}m^{t-1}_{h_l}, \tag{27}
$$

$$
\overline{z}^t_{f_k\to h_l} = \overline{z}^t_{f_k} + \Phi_{k,l}m^{t-1}_{h_l} - v^{t-1}_{h_l}\overline{z}^{t-1}_{f_k\to h_l}/\overline{\tau}^{t-1}_{f_k}. \tag{28}
$$

With $\overline{z}^t_{f_k\to h_l}$ shown in (28), ξ_l^t and γ_k^t can be expressed recursively as

$$
\xi_l^t \approx \left(\sum_k \Phi^*_{k,l}\frac{\overline{z}^t_{f_k}}{\overline{\tau}^t_{f_k}}\right) + \left(\sum_{k'}\frac{m^{t-1}_{h_l}}{\overline{\tau}^t_{f_{k'}}}\right) - \frac{Kv^{t-1}_{h_l}\xi_l^{t-1}}{\sum_{k''}\overline{\tau}^t_{f_{k''}}}, \tag{29}
$$

$$
\gamma_k^t \approx \frac{\overline{z}^t_{f_k}\sum_l v^t_{h_l}}{\overline{\tau}^t_{f_k}} + \frac{\sum_{l'}\Phi_{k,l'}v^t_{h_{l'}}m^{t-1}_{h_{l'}}}{\overline{\tau}^t_{f_k}} - \frac{\gamma_k^{t-1}\sum_{l''}v^t_{h_{l''}}}{L\overline{\tau}^t_{f_k}}. \tag{30}
$$

Finally, $\mu_{f_k\to\Psi_k}(x_k)$ are given by

$$
\mathcal{N}_{\mathbb{C}}\Big(y_k; x_k(\sum_l \Phi_{k,l}m^{t_{max}}_{h_l} - \gamma_k^{t_{max}}), \sigma_n^2 + |x_k|^2\sum_l v^{t_{max}}_{h_l}\Big). \tag{31}
$$

We summarize the simplified message-passing algorithm for the inner iteration in Alg. 2, which will be referred to as "EP-LC". Note that $z_{f_k}^t(x_k), \gamma_k^t, k = 1, \ldots, K$, and $\xi_l^t, l = 1, \ldots, L$, can be efficiently calculated using FFT and inverse FFT, respectively.

Algorithm 2 The EP-LC algorithm for the tth inner iteration.

Initialization: If $t = 1$, set $\gamma_k^0 = 0, \forall k$, $m_{h_l}^0 = m_{h_l}^a, v_{h_l}^0 = v_{h_l}^a, \xi_l^t = 0, \forall l$.

for $k = 1 \to K$ **do** $z_{f_k}^t(x_k) = \frac{y_k}{x_k} + \gamma_k^{t-1} - \sum_l \Phi_{k,l} m_{h_l}^{t-1}$,

$\quad \tau_{f_k}^t(x_k) = \frac{1}{|x_k|^2}\sigma_n^2 + \sum_l v_{h_l}^{t-1}$, $\epsilon_k^t(x_k) = \frac{z_{f_k}^t(x_k)}{\tau_{f_k}^t(x_k)}$,

$\quad \zeta_k^t = \mathbb{E}_{q(x_k)}\left[\frac{1}{\tau_{f_k}^t(x_k)} - |\epsilon_k^t(x_k)|^2\right] + \left|\mathbb{E}_{q(x_k)}[\epsilon_k^t(x_k)]\right|^2$,

$\quad \overline{z}_{f_k}^t = \frac{1}{\zeta_k^t}\mathbb{E}_{q(x_k)}[\epsilon_{f_k}^t(x_k)]$, $\overline{\tau}_{f_k}^t = \frac{1}{\zeta_k^t}$.

end for

for $l = 1 \to L$ **do**

$\quad \xi_l^t \approx \left(\sum_k \Phi_{k,l}^* \frac{\overline{z}_{f_k}^t}{\overline{\tau}_{f_k}^t}\right) + \left(\sum_{k'} \frac{m_{h_l}^{t-1}}{\overline{\tau}_{f_{k'}}^t}\right) - \frac{K v_{h_l}^{t-1} \xi_l^{t-1}}{\sum_{k''} \overline{\tau}_{f_{k''}}^t}$,

$\quad m_{h_l}^t = v_{h_l}^t\left(m_{h_l}^a / v_{h_l}^a + \xi_l^t\right)$, $v_{h_l}^t = \left(\frac{1}{v_{h_l}^a} + \sum_k \frac{1}{\overline{\tau}_{f_k}^t}\right)^{-1}$.

end for

for $k = 1 \to K$ **do**

$$\gamma_k^t \approx \frac{\overline{z}_{f_k}^t \sum_l v_{h_l}^t}{\overline{\tau}_{f_k}^t} + \frac{\sum_{l'} \Phi_{k,l'} v_{h_{l'}}^t m_{h_{l'}}^{t-1}}{\overline{\tau}_{f_k}^t} - \frac{\gamma_k^{t-1} \sum_{l''} v_{h_{l''}}^t}{L \overline{\tau}_{f_k}^t}.$$

end for

5. NUMERICAL RESULTS

5.1. SIMULATION SETUP

We examine the proposed algorithm using the $K = 512$ OFDM with 16QAM. The channel taps are assumed to change from one OFDM symbol to another but be constant within an OFDM symbol. The number of the channel taps is $L = 32$, and the channel power delay profile is $\{v_{h_l}^a = 1/L\}$. A $R = 9216/16128$ LDPC code with codeword length 16128 bits and average column weight 3 is employed, and 2304 training bits (all '1's) were uniformly multiplexed with the code-word. As a result, the spectral efficiency is 2 bits per subcarrier use. In a single turbo iteration, both the number of inner iteration in the dashed block and the LDPC decoder is set to 1.

5.2. RESULTS

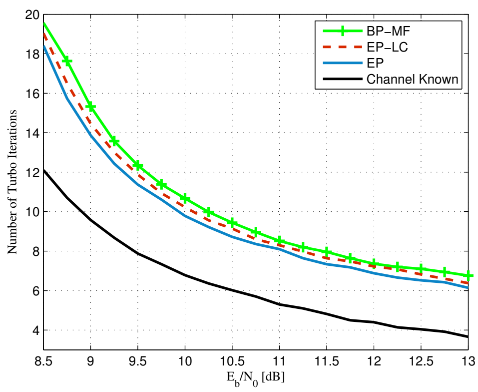

Fig. 4. Average number of turbo iterations versus E_b/N_0.

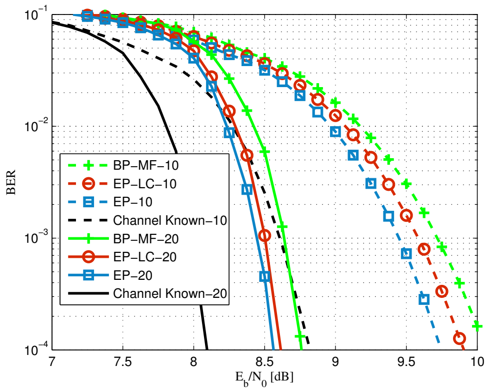

Fig. 2. BER versus E_b/N_0; dashed curves refer to 10 turbo iterations; solid curves refer to 20 turbo iterations.

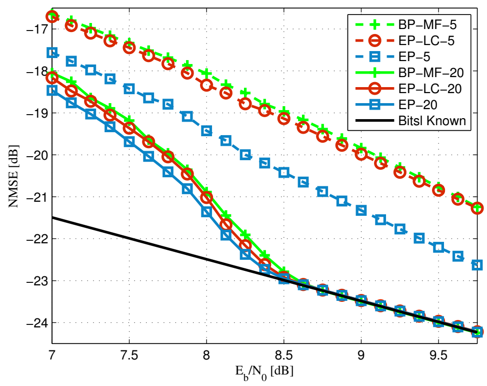

Fig. 3. Channel taps NMSE versus E_b/N_0; dashed curves refer to 5 turbo iterations; solid curves refer to 20 turbo iterations.

Fig. 2 shows that both the proposed EP algorithm and EP-LC algorithm perform close to the bound of known channel within 0.5 dB, and outperform the BP-MF algorithm proposed in [12]. Fig. 3 shows the normalized mean squared error (NMSE) of channel estimates versus E_b/N_0. We also consider the case of known bits, which serves as a lower bound for the channel estimation. It is shown that our algorithms outperform BP-MF algorithm in the low-E_b/N_0 region or when only a few turbo iterations are performed. Fig. 4 shows the average number of turbo iterations need for successful decoding. The number of turbo iteration performed by our algorithms is less than that of the BP-MF algorithm, although the BP-MF algorithm has a higher complexity per turbo iteration.

6. CONCLUSION

In this paper, we presented a message passing approach to joint channel estimation and decoding for OFDM systems. The complexity of our algorithm is $\mathcal{O}(K|\mathcal{A}| + K\log_2 K)$ per turbo iteration, quite suitable for the system with many subcarriers and long channel memory. Numerical experiments demonstrated that our algorithm achieved BER performance within 0.5 dB of the known-channel bound and outperformed the BP-MF algorithm proposed by Riegler *et al* in terms of both BER performance and complexity.

7. REFERENCES

[1] A. P. Worthen and W. E. Stark, "Unified design of iterative receivers using factor graphs," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 843–849, 2001.

[2] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 498–519, Feb. 2001.

[3] R. Chen, X. Wang, and J. S. Liu, "Adaptive joint detection and decoding in flat-fading channels via mixture Kalman filtering," *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 2079–2094, 2000.

[4] C. J. Bordin and M. G. Bruno, "Particle filters for joint blind equalization and decoding in frequency-selective channels," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2395–2405, 2008.

[5] Y. Liu, L. Brunel, and J. J. Boutros, "Joint channel estimation and decoding using Gaussian approximation in a factor graph over multipath channel," in *Proc. Int. Symp. on Personal, Indoor and Mobile Radio Commun. (PIMRC)*, 2009, pp. 3164–3168.

[6] G. E. Kirkelund, C. N. Manchón, L. P. B. Christensen, E. Riegler, and B. H. Fleury, "Variational message-passing for joint channel estimation and decoding in MIMO-OFDM," in *Proc. IEEE Global Telecomm. Conf. (GLOBECOM)*, 2010, pp. 1–6.

[7] Q. Guo and D. D. Huang, "EM-based joint channel estimation and detection for frequency selective channels using gaussian message passing," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 4030–4035, 2011.

[8] P. Schniter, "A message-passing receiver for BICM-OFDM over unknown clustered-sparse channels," *IEEE J. Sel. Topics in Signal Process.*, vol. 5, no. 8, pp. 1462–1474, 2011.

[9] C. Knievel, P. A. Hoeher, A. Tyrrell, and G. Auer, "Multi-dimensional graph-based soft iterative receiver for MIMO-OFDM," *IEEE Trans. Commun.*, vol. 60, no. 6, pp. 1599–1609, June 2012.

[10] C. Novak, G. Matz, and F. Hlawatsch, "IDMA for the multiuser MIMO-OFDM uplink: A factor graph framework for joint data detection and channel estimation," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 4051–4066, Aug. 2013.

[11] E. Riegler, G. E. Kirkelund, C. N. Manchón, M.-A. Badiu, and B. H. Fleury, "Merging belief propagation and the mean field approximation: A free energy approach," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 588–602, Jan. 2013.

[12] M.-A. Badiu, G. E. Kirkelund, C. N. Manchón, E. Riegler, and B. H. Fleury, "Message-passing algorithms for channel estimation and decoding using approximate inference," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, 2012, pp. 2376–2380.

[13] M.-A. Badiu, C. Manchón, and B. Fleury, "Message-passing receiver architecture with reduced-complexity channel estimation," *IEEE Commun. Lett.*, vol. 17, no. 7, pp. 1404–1407, Jul. 2013.

[14] T. P. Minka, "A family of algorithms for approximate Bayesian inference," Ph.D. dissertation, Massachusetts Institute of Technology, 2001.

[15] H.-A. Loeliger, "An introduction to factor graphs," *IEEE Signal Process. Mag.*, vol. 21, no. 1, pp. 28–41, Jan. 2004.

[16] T. P. Minka, "Divergence measures and message passing," *Microsoft Research Cambridge, Tech. Rep.*, 2005.