IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 17, 2017, accepted May 4, 2017, date of publication May 9, 2017, date of current version June 7, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2702571

A Hybrid BP-EP-VMP Approach to Joint Channel Estimation and Decoding for FTN Signaling over Frequency Selective Fading Channels

NAN WU[1], (Member, IEEE), WEIJIE YUAN[1], (Student Member, IEEE), QINGHUA GUO[2,3], (Member, IEEE), AND JINGMING KUANG[1]

[1]School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China
[2]School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Wollongong, NSW 2522, Australia
[3]School of Electrical, Electronic and Computer Engineering, The University of Western Australia, Perth, WA 6009, Australia

Corresponding author: Nan Wu (wunan@bit.edu.cn)

This work was supported in part by the National Science Foundation of China under Grant 61421001, Grant 61571041, and Grant 61471037, in part by the National High Technology Research and Development Program of China under Grant 2015AA01A709, and in part by the Australian Research Council's DECRA under Grant DE120101266.

ABSTRACT This paper deals with low-complexity joint channel estimation and decoding for faster-than-Nyquist (FTN) signaling over frequency selective fading channels. The inter-symbol interference (ISI) imposed by FTN signaling and the frequency selective channel are intentionally separated to fully exploit the known structure of the FTN-induced ISI. Colored noise due to the faster sampling rate than that of the Nyquist signaling system is approximated by autoregressive process. A Forney style factor graph representation of the FTN system is developed and Gaussian message passing is performed on the graph. Expectation propagation (EP) is employed to approximate the message from channel decoder to Gaussian distribution. Since the inner product between FTN symbols and channel coefficients is infeasible by belief propagation (BP), we propose to perform variational message passing (VMP) on an equivalent soft node in factor graph to tackle this problem. Simulation results demonstrate that the proposed low-complexity hybrid BP-EP-VMP algorithm outperforms the existing methods in FTN system. Compared with the Nyquist counterpart, FTN signaling with the proposed algorithm is able to increase the transmission rate by over 40%, with only negligible BER performance loss.

INDEX TERMS Faster-than-Nyquist signaling, factor graph, joint channel estimation and decoding, belief propagation, variational message passing, expectation propagation.

I. INTRODUCTION

Since the available bandwidth becomes insufficient, it is required to maximize the spectral efficiency to achieve higher data rate in mobile communication systems [1]. Amongst several methods, the faster-than-Nyquist (FTN) signaling proposed by Mazo [2] has been rediscovered recently and attracted numerous attentions since it is able to increase the transmission rate with the same bandwidth. It is known that the bit error rate (BER) performance will not be affected when the packing ratio lies above the Mazo limit [3].

It is well known that the Nyquist rate ensures intersymbol interference (ISI) free transmission. However, with FTN signaling, since the symbol period is packed, the shaping pulse is nonorthogonal with respect to the symbol interval. As a result, ISI is unavoidable. Moreover, the length of FTN-induced ISI could be very long, which leads to challenging detection problem at receiver. Several receiving techniques have been proposed to eliminate the ISI caused by FTN signaling in additive white Gaussian noise (AWGN) channels. A reduced search BCJR detector is proposed in [4]. A successive interference cancellation (SIC) detector with optimal packing ratio is devised in [5]. By taking advantage of the frequency-domain equalization (FDE), an FDE FTN detector is proposed in [6], which does not consider the colored noise imposed by the faster sampling rate of FTN receiver than that of the Nyquist signaling. Moreover, the insertion of cyclic prefix in FDE also decreases the efficiency. By using factor graph and Gaussian message passing (GMP), a graph-based linear minimum mean squared error (LMMSE) equalizer is developed in [7] for FTN signaling over AWGN channels.

Only a few studies considered receivers design for FTN signaling in fading channels. In [8], an FDE-aided iterative FTN signaling detector is proposed for frequency selective fading channels. Since the coefficients of the fading channel are assumed to be known, the ISI caused by both FTN signaling and the frequency selective channels are considered together. In [9], low-complexity FTN receivers based on two variational methods, i.e., mean field (MF) and Bethe approximations, are studied for doubly selective channels. For the receivers discussed in [8] and [9], channel state information (CSI) is assumed to be perfectly known. In practical wireless communications, however, CSI is usually unknown and has to be estimated at receiver. Many channel estimation methods have been proposed in the literature [10]–[12]. However, most of them rely on long training sequences to obtain accurate CSI, which on the other hand reduces the advantage of using FTN signaling [13]. Since data symbols and channel coefficients are coupled in observations, it is possible to jointly estimate the CSI and detect the data symbols. Joint channel estimation and decoding can not only reduce the number of pilots but enhance the accuracy of channel estimation [14]–[16].

Recently, motivated by the heuristic iterative approaches on probabilistic graphical models, several methods are derived for low-complexity iterative receivers based on message passing algorithms, e.g., belief propagation (BP) [17], variational message passing (VMP) [18], and approximate message passing (AMP) [19]. In [20], based on factor graph and BP, a frequency-domain iterative message passing receiver for FTN signaling is proposed in doubly selective channels. The algorithm is evaluated in both perfect and imperfect CSI scenarios. Nevertheless, to the best knowledge of the authors, joint channel estimation and decoding has not been investigated for FTN signaling.

This paper deals with low-complexity receiver design for FTN signaling in frequency selective fading channels. Different from the existing works in [8], [9], and [20], which assume the CSI is known at the receiver side, we consider a more practical scenario that CSI is unknown and joint channel estimation and decoding has to be performed for FTN signaling. Instead of combining the ISI imposed by FTN signaling and the frequency selective channel together as in [8], we intentionally separate them from each other, which enables us to fully exploit the known structure of the FTN-induced ISI. Considering that the packing symbol period will lead to correlation of channel taps, we use discrete Fourier transform (DFT) interpolation to obtain channel taps in FTN scenario. The colored noised due to the faster sampling rate in FTN system is approximately modeled by autoregressive (AR) process to avoid the whitening filtering. Building on this, a Forney style factor graph is constructed and GMP is employed to update messages on the graph. Considering the inner product between FTN symbols and channel coefficients is infeasible by using BP, we propose to perform VMP on an equivalent soft node instead. Different from linear minimum mean squared error (MMSE)-based channel estimator [21], the VMP algorithm also provides the uncertainty in channel estimates. Moreover, expectation propagation (EP) [22] is used to update the message from channel decoder by Gaussian distribution, which is shown to have better performance than the MMSE-based method employed in [9]. The proposed hybrid BP-EP-VMP algorithm enables effective Gaussian message approximation, thereby allowing low-complexity implementation. Simulation results demonstrate the superior performance of the proposed algorithm, and also show the advantage of using FTN signaling over the Nyquist counterpart.

Notations: We use a boldface capital letter to denote a matrix while boldface lower-case letter for a vector. The superscript $*$, T, H and -1 denote the conjugate, transpose, Hermitian, and the inverse operations, respectively; $\mathcal{N}(m, \sigma^2)$ denotes Gaussian distribution with mean m and variance σ^2; $\propto$ represents equality up to a constant normalization factor; $\mathbb{E}$ is the expectation operation; $\mathbf{A}_{:,i}$ denotes the ith row of matrix $\mathbf{A}$; $\lfloor\cdot\rceil$ denotes the rounding to integer operation; $\mathbf{A}_{ij}$ is the (i,j)th element of $\mathbf{A}$; $\overrightarrow{\cdot}$ denotes the message that passes along the direction of the edge while $\overleftarrow{\cdot}$ denotes the message passing in the opposite direction.

II. FASTER-THAN-NYQUIST SIGNALING MODEL

We consider a coded FTN signaling system illustrated in Fig. 1. At the transmitter side, the information bit sequence $\mathbf{b}$ is encoded to a coded sequence $\mathbf{c}$ and mapped to a length-N data symbol vector $\mathbf{x} = [x_0, \ldots, x_{N-1}]^T$. The data symbol block passes through the shaping filter $g(t)$, yielding transmitted signal $s(t) = \sum_i g(t - i\tau T_0)x_i$, where $0 \leq \tau \leq 1$ is the FTN packing ratio and T_0 is the symbol period under the Nyquist criterion. Obviously we can choose a smaller τ to achieve higher data rate at the cost of severer ISI. Note that an FTN symbol is interfered by neighboring symbols on both sides and the number of ISI taps is infinite. In practice, we can choose a sufficiently large number $L_{\text{FTN}} = 2L_f + 1$, then $s(t)$ is given as

$$s(t) = \sum_{i=-L_f}^{L_f} g(t - i\tau T_0)x_i. \tag{1}$$

The signal is transmitted over a frequency selective fading channel $h(t)$. For Nyquist signaling, we can model the channel with L_{nyq} independent taps $\tilde{\mathbf{h}} = [\tilde{h}_{L_{\text{nyq}}-1}, \ldots, \tilde{h}_0]^T$. However, for FTN signaling, since the symbol period is packed, the number of channel taps becomes greater, i.e, $L = \lfloor L_{\text{nyq}}/\tau \rceil$. According to [23], the channel taps in FTN system can be calculated via interpolation based on the channel taps in Nyquist counterpart. Here we exploit DFT to obtain the fading channel taps for FTN signaling. Assuming $\mathbf{D} \in \mathcal{R}^{L_{\text{nyq}} \times L_{\text{nyq}}}$ and $\mathbf{D}_F \in \mathcal{R}^{L \times L}$ are the DFT matrices, the channel tap $\mathbf{h} = [h_{L-1}, \ldots, h_l, \ldots, h_0]^T$ for FTN signaling is given as

$$\mathbf{h} = \mathbf{D}_F^H \begin{bmatrix} \mathbf{D}\tilde{\mathbf{h}} \\ \mathbf{0} \end{bmatrix}. \tag{2}$$

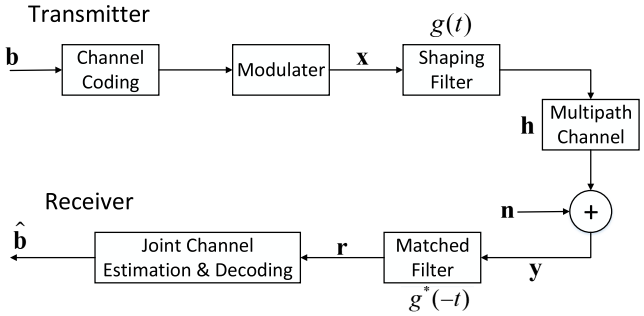


FIGURE 1. System model for considered FTN signaling system.

Note that the correlations between different channel taps will result in a nondiagonal covariance matrix of $\mathbf{h}$, which is given as

$$\mathbf{V}_h = \mathbf{D}_F^H \begin{bmatrix} \mathbf{D}\mathbf{V}_{\text{nyq}}\mathbf{D}^H & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \mathbf{D}_F,$$

where $\mathbf{V}_{\text{nyq}}$ is the covariance matrix of $\tilde{\mathbf{h}}$.

With the assumption of prefect synchronization, the received signal can be represented as

$$y(t) = \sum_{l=0}^{L-1} \sum_{i=-L_f}^{L_f} h_l g(t-(i+l)\tau T_0)x_i + n(t), \tag{3}$$

where $n(t)$ is white Gaussian noise process with power spectral density N_0. The received continuous-time signal $y(t)$ is matched-filtered by $g^*(t)$ and then sampled with the symbol rate $\frac{1}{\tau T_0}$. At the output of the matched filter, the k-th sample can be expressed as

$$\begin{aligned} r_k &= \sum_{l=0}^{L-1} \sum_{i=-L_f}^{L_f} h_l q(k\tau T_0 - (i+l)\tau T_0)x_i + \xi(k\tau T_0) \\ &= \sum_{l=0}^{L-1} \sum_{i=-L_f}^{L_f} h_l q_{k-l-i}x_i + \xi_k, \end{aligned} \tag{4}$$

where $q_{m-n} = \int g(t-m\tau T_0)g^*(t-n\tau T_0)\mathrm{d}t$ and $\xi_k = \int n(t)g^*(t-k\tau T_0)\mathrm{d}t$. Since $g(t)$ is not τT_0-orthogonal, $\{\xi_k\}$ is a colored noise process with autocorrelation function

$$\mathbb{E}[\xi_m \xi_n] = N_0 q_{m-n}. \tag{5}$$

With (4), the received symbol vector $\mathbf{r} = [r_0, \ldots, r_{N-1}]^T$ can be written as

$$\mathbf{r} = \mathbf{H}\mathbf{Q}\mathbf{x} + \boldsymbol{\xi}, \tag{6}$$

where $\mathbf{x} = [x_0, \ldots, x_{N-1}]^T$ and $\boldsymbol{\xi} = [\xi_0, \ldots, \xi_{N-1}]^T$ are the symbol and noise vector; $\mathbf{H}$ and $\mathbf{Q}$ are the matrices with respect to channel tap and FTN ISI tap as

$$\mathbf{H} = \begin{bmatrix} h_0 & & & & & \mathbf{0} \\ h_1 & h_0 & & & & \\ \vdots & & \ddots & & & \\ h_{L-1} & h_{L-2} & \cdots & h_0 & & \\ & \ddots & & & \ddots & \\ \mathbf{0} & & h_{L-1} & \cdots & h_1 & h_0 \end{bmatrix} \tag{7}$$

and

$$\mathbf{Q} = \begin{bmatrix} q_0 & q_1 & \cdots & q_{L_f} & & \mathbf{0} \\ \vdots & q_0 & & \ddots & & \\ q_{-L_f} & \cdots & q_0 & \cdots & q_{L_f} & \\ & q_{-L_f} & \cdots & q_0 & \cdots & q_{L_f} \\ & & \ddots & & & \vdots \\ \mathbf{0} & & & q_{-L_f} & \cdots & q_0 \end{bmatrix}. \tag{8}$$

The autocorrelation matrix of noise vector $\boldsymbol{\xi}$ is given as $\mathbb{E}[\boldsymbol{\xi}\boldsymbol{\xi}^H] = N_0\mathbf{Q}$.

III. MESSAGE PASSING RECEIVER DESIGN

In this section, a Gaussian message passing based iterative receiver is proposed for joint channel estimation and decoding in FTN signaling system.

A. OUTPUT LLR OF CHANNEL DECODER

The receiver performs iterative decoding by exchanging log-likelihood ratio (LLR) between the channel decoder and equalizer. For decoding, the optimal BP decoding algorithm is utilized. Then the output extrinsic LLR of channel decoder can be represented as

$$L^0(c_{n,m}) = \ln \frac{p(c_{n,m}=0)}{p(c_{n,m}=1)}, \tag{9}$$

where $c_{n,m}$ denotes the mth code bit in the nth subsequence $c_{\mathbf{n}} = [c_{n,1}, \ldots, c_{n,M}]^T$ with M being the modulation order.

B. AUTOREGRESSIVE MODEL OF COLORED NOISE

Due to the correlations between noise samples, conventional MMSE detection approaches suffer from high complexity. Some works neglect the impact of colored noise, which will cause performance loss. To overcome this problem, we employ a Pth-order AR process to approximately model the colored noise [24], i.e.,

$$\xi_k = \sum_{j=1}^{P} a_j \xi_{k-j} + w_k = \boldsymbol{a}^T \boldsymbol{\xi}_{k-1} + w_k, \tag{10}$$

where $\boldsymbol{a} = [a_1, \ldots, a_P]^T$ is the AR parameters and w_k is the white noise with zero mean and variance σ_w^2, and $\boldsymbol{\xi}_{k-1} = [\xi_{k-1}, \ldots, \xi_{k-P}]^T$ denotes the correlated noise samples. The autocorrelation parameters $\boldsymbol{a}$ can be obtained from the Yule-Walker equation as

$$N_0 q_k = \begin{cases} N_0 \sum_{j=1}^{P} a_j q_{-j} + \sigma_w^2 & k = 0 \\ N_0 \sum_{j=1}^{P} a_j q_{k-j} & \text{othertwise.} \end{cases} \tag{11}$$

C. FACTOR GRAPH REPRESENTATION

Note that the equation (4) can be reformulated as

$$s_k = \mathbf{q}^T \mathbf{x}_k, \tag{12}$$
$$r_k = \mathbf{h}^T \mathbf{s}_k + \xi_k, \tag{13}$$

where $\mathbf{x}_k = [x_{k-L_f}, \ldots, x_k, \ldots, x_{k+L_f}]^T$ and $\mathbf{s}_k = [s_{k-L+1}, \ldots, s_k]^T$. Moreover, $\mathbf{x}_k$ and $\mathbf{s}_k$ follow state transition model as

$$\mathbf{x}_k = \mathbf{G}\mathbf{x}_{k-1} + \mathbf{f}x_{k+L_f}, \tag{14}$$
$$\mathbf{s}_k = \mathbf{G}_1\mathbf{s}_{k-1} + \mathbf{f}_1^T s_k, \tag{15}$$

where the $\mathbf{G} = \begin{bmatrix} \mathbf{0}_{2L_f} & \mathbf{I}_{2L_f} \\ 0 & \mathbf{0}_{2L_f}^T \end{bmatrix}$, $\mathbf{f} = [\mathbf{0}_{2L_f}^T, 1]^T$, $\mathbf{G}_1 = \begin{bmatrix} \mathbf{0}_{L-1} & \mathbf{I}_{L-1} \\ 0 & \mathbf{0}_{L-1}^T \end{bmatrix}$ and $\mathbf{f}_1 = [\mathbf{0}_{L-1}^T, 1]^T$. Similarly, (10) can be rewritten as

$$\boldsymbol{\xi}_k = \mathbf{A}\boldsymbol{\xi}_{k-1} + \mathbf{f}_2 w_k, \tag{16}$$
$$\xi_k = \mathbf{f}_2^T \boldsymbol{\xi}_k, \tag{17}$$

with $\mathbf{f}_2 = [\mathbf{0}_{P-1}^T, 1]^T$ and $\mathbf{A} = \begin{bmatrix} 0 & \boldsymbol{a}^T \\ \mathbf{0}_{P-1} & \mathbf{I}_{P-1} \end{bmatrix}$.

According to [25], the linear state space model can be represented by a "block diagram" factor graph. Based on (12)-(17), the corresponding Forney-style factor graph is depicted in Fig. 2. On this factor graph, the edges represent variables while the factor nodes denote the local functions. The equality node can be regarded as branching points which allow different factors to share the same variables. Furthermore, a multiplier node ⊠ is introduced to denote the inner product constraint $\delta(r - \mathbf{h}^T\mathbf{s})$.

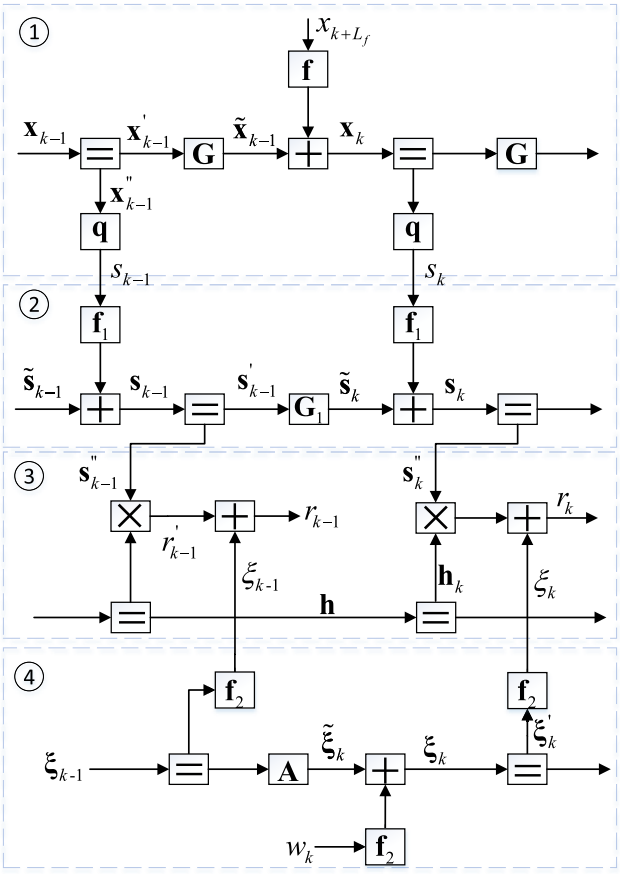


FIGURE 2. Factor graph representation for joint channel estimation and decoding for FTN system. The subgraphs denoted by ①, ②, ③ and ④ correspond to the FTN equalization, multipath channel equalization, channel estimation and colored noise process, respectively.

D. COMBINED BP-EP-VMP MESSAGE PASSING

GMP is an efficient parametric message passing algorithm in linear Gaussian system, where message on factor graph can be characterized either by the mean vector $\mathbf{m}$ and the covariance matrix $\mathbf{V}$ or by the weight matrix $\mathbf{W} = \mathbf{V}^{-1}$ and the transformed mean $\mathbf{W}\mathbf{m}$.[1] The update rules of GMP have been derived in [25].

Following the GMP rules, most messages on the factor graph can be computed. For ease of exposition, we will elaborate the message updating on the four subgraphs in Fig. 2 separately.

■ *Messages Updating for FTN Equalization (on Subgraph 1):*

Assuming that the parameters $\overrightarrow{\mathbf{W}}_{\mathbf{x}_{k-1}}$ and $\overrightarrow{\mathbf{W}}_{\mathbf{x}_{k-1}}\overrightarrow{\mathbf{m}}_{\mathbf{x}_{k-1}}$ are available, we have

$$\overrightarrow{\mathbf{V}}_{\tilde{\mathbf{x}}_{k-1}} = \mathbf{G}\left(\overleftarrow{\mathbf{W}}_{\mathbf{x}''_{k-1}} + \overrightarrow{\mathbf{W}}_{\mathbf{x}_{k-1}}\right)^{-1}\mathbf{G}^T, \tag{18}$$
$$\overrightarrow{\mathbf{W}}_{\tilde{\mathbf{x}}_{k-1}}\overrightarrow{\mathbf{m}}_{\tilde{\mathbf{x}}_{k-1}} = \mathbf{G}\left(\overrightarrow{\mathbf{W}}_{\mathbf{x}_{k-1}}\overrightarrow{\mathbf{m}}_{\mathbf{x}_{k-1}} + \overleftarrow{\mathbf{W}}_{\mathbf{x}''_{k-1}}\overleftarrow{\mathbf{m}}_{\mathbf{x}''_{k-1}}\right), \tag{19}$$

[1]It may happen frequently that the covariance matrix of a message does not exist due to the singular matrix. Under such circumstances, one may use the transformed means and weight matrices to parameterize the messages [25].

where $\overleftarrow{\mathbf{W}}_{\mathbf{x}''_{k-1}}$ and $\overleftarrow{\mathbf{W}}_{\mathbf{x}''_{k-1}}\overleftarrow{\mathbf{m}}_{\mathbf{x}''_{k-1}}$ are given as

$$\overleftarrow{\mathbf{W}}_{\mathbf{x}''_{k-1}} = \frac{\mathbf{q}\mathbf{q}^T}{V_{s_{k-1}}}, \tag{20}$$
$$\overleftarrow{\mathbf{W}}_{\mathbf{x}''_{k-1}}\overleftarrow{\mathbf{m}}_{\mathbf{x}''_{k-1}} = \frac{\mathbf{q}m_{s_{k-1}}}{V_{s_{k-1}}}. \tag{21}$$

In a similar way, the backward mean and covariance matrix $\overleftarrow{\mathbf{m}}_{\mathbf{x}_k}$ and $\overleftarrow{\mathbf{V}}_{\mathbf{x}_k}$ are obtained as

$$\overleftarrow{\mathbf{V}}_{\mathbf{x}_k} = \left(\overleftarrow{\mathbf{W}}_{\mathbf{x}'_k} + \overleftarrow{\mathbf{W}}_{\mathbf{x}''_k}\right)^{-1}, \tag{22}$$
$$\overleftarrow{\mathbf{m}}_{\mathbf{x}_k} = \overleftarrow{\mathbf{V}}_{\mathbf{x}_k}\left(\overleftarrow{\mathbf{W}}_{\mathbf{x}'_k}\overleftarrow{\mathbf{m}}_{\mathbf{x}'_k} + \overleftarrow{\mathbf{W}}_{\mathbf{x}''_k}\overleftarrow{\mathbf{m}}_{\mathbf{x}''_k}\right). \tag{23}$$

With (18)-(23), the outgoing message parameters for x_{k+L_f} are

$$\overleftarrow{m}_{x_{k+L_f}} = \mathbf{f}^T\left(\overleftarrow{\mathbf{m}}_{\mathbf{x}_k} - \overrightarrow{\mathbf{m}}_{\tilde{\mathbf{x}}_{k-1}}\right), \tag{24}$$
$$\overleftarrow{V}_{x_{k+L_f}} = \mathbf{f}^T\left(\overrightarrow{\mathbf{V}}_{\tilde{\mathbf{x}}_{k-1}} + \overleftarrow{\mathbf{V}}_{\mathbf{x}_k}\right)\mathbf{f}. \tag{25}$$

As the incoming messages are computed from $L^0(c_{n,m})$, they have discrete distributions with respect to the constellation points. In order to employ GMP, we use EP to approximate the incoming messages to be Gaussian [26]. For the

kth symbol, the incoming message can be expressed as

$$\overrightarrow{\mu}(x_k) = \sum_{\chi_i \in \mathcal{A}} p_{k,i}\delta(x_k - \chi_i), \tag{26}$$

where χ_i is the ith constellation symbol, $\mathcal{A}$ is the set of constellation symbols and $p_{k,i}$ is the probability with respect to χ_i, which is computed from the LLR $L^0(c_{n,m})$. Then the belief of x_k is obtained as a probability mass function (PMF). Based on EP, matching the first two order moments of the belief yields

$$\tilde{m}_{x_k} = \frac{1}{2\pi \overleftarrow{V}_{x_k}} \sum_{\chi_i \in \mathcal{A}} \chi_i p_{k,i} \exp\left(-\frac{(\chi_i - \overleftarrow{m}_{x_k})^2}{\overleftarrow{V}_{x_k}}\right), \tag{27}$$

$$\tilde{V}_{x_k} = \frac{1}{2\pi \overleftarrow{V}_{x_k}} \sum_{\chi_i \in \mathcal{A}} |\chi_i|^2 p_{k,i} \exp\left(-\frac{(\chi_i - \overleftarrow{m}_{x_k})^2}{\overleftarrow{V}_{x_k}}\right) - |m_{x_k}|^2. \tag{28}$$

Then the Gaussian approximation to the incoming message can be parameterized as

$$\overrightarrow{m}_{x_k} = \overrightarrow{V}_{x_k}\left(\frac{\tilde{m}_{x_k}}{\tilde{V}_{x_k}} - \frac{\overleftarrow{m}_{x_k}}{\overleftarrow{V}_{x_k}}\right), \tag{29}$$

$$\overrightarrow{V}_{x_k} = \left(\tilde{V}_{x_k}^{-1} - \overleftarrow{V}_{x_k}^{-1}\right)^{-1}. \tag{30}$$

Consequently, the outgoing messages which are passed to Subgraph 2 read

$$\overrightarrow{m}_{s_k} = \mathbf{q}^T\left(\overrightarrow{\mathbf{W}}_{\mathbf{x}_k} + \overleftarrow{\mathbf{W}}_{\mathbf{x}'_k}\right)^{-1} \times \left(\overrightarrow{\mathbf{W}}_{\mathbf{x}_k}\left(\overrightarrow{\mathbf{m}}_{\tilde{\mathbf{x}}_{k-1}} + \mathbf{f}\overrightarrow{m}_{x_{k+L_f}}\right) + \overleftarrow{\mathbf{W}}_{\mathbf{x}'_k}\overleftarrow{\mathbf{m}}_{\mathbf{x}'_k}\right), \tag{31}$$

$$\overrightarrow{V}_{s_k} = \mathbf{q}^T\left(\overrightarrow{\mathbf{W}}_{\mathbf{x}_k} + \overleftarrow{\mathbf{W}}_{\mathbf{x}'_k}\right)^{-1}\mathbf{q}, \tag{32}$$

where $\overrightarrow{\mathbf{W}}_{\mathbf{x}_k} = \left(\overrightarrow{\mathbf{V}}_{\tilde{\mathbf{x}}_{k-1}} + \mathbf{f}\overrightarrow{V}_{x_k}\mathbf{f}^T\right)^{-1}$.

■ *Messages Updating for Multipath Channel Equalization (on Subgraph 2):*

Similar to (24) and (25), the backward messages $\overleftarrow{m}_{s_k}$ and $\overleftarrow{V}_{s_k}$ are given by

$$\overleftarrow{m}_{s_k} = \mathbf{f}_1^T\left(\overleftarrow{\mathbf{m}}_{\mathbf{s}_k} - \overrightarrow{\mathbf{m}}_{\tilde{\mathbf{s}}_k}\right), \tag{33}$$

$$\overleftarrow{V}_{s_k} = \mathbf{f}_1^T\left(\overrightarrow{\mathbf{V}}_{\tilde{\mathbf{s}}_k} + \overleftarrow{\mathbf{V}}_{\mathbf{s}_k}\right)\mathbf{f}_1, \tag{34}$$

where the parameters with respect to $\mathbf{s}_k$ and $\tilde{\mathbf{s}}_k$ have similar form as in (18)-(23). According to GMP rules, the messages $\overrightarrow{\mathbf{m}}_{\mathbf{s}''_k}$ and $\overrightarrow{\mathbf{V}}_{\mathbf{s}''_k}$ can be expressed as

$$\overrightarrow{\mathbf{m}}_{\mathbf{s}''_k} = \overrightarrow{\mathbf{V}}_{\mathbf{s}''_k}\left(\overrightarrow{\mathbf{W}}_{\mathbf{s}_k}\overrightarrow{\mathbf{m}}_{\mathbf{s}_k} + \overleftarrow{\mathbf{W}}_{\mathbf{s}'_k}\overleftarrow{\mathbf{m}}_{\mathbf{s}'_k}\right), \tag{35}$$

$$\overrightarrow{\mathbf{V}}_{\mathbf{s}''_k} = \left(\overrightarrow{\mathbf{W}}_{\mathbf{s}_k} + \overleftarrow{\mathbf{W}}_{\mathbf{s}'_k}\right)^{-1}, \tag{36}$$

which are involved in the message computations in Subgraph 3.

■ *Messages Updating for Colored Noise Estimation (on Subgraph 4):*

It is noted that the correlation between colored noise samples does not affect the first-order moment. The means of messages with respect to ξ_k on Subgraph 4 are $\overrightarrow{m}_{\xi_k} = \overleftarrow{m}_{\xi_k} = \mathbb{E}[\xi_k] = 0,\ \forall k$. Therefore only the variances (covariance matrices) of messages need to be calculated. The variance $\overrightarrow{V}_{\xi_k}$ can be obtained as

$$\overrightarrow{V}_{\xi_k} = \mathbf{f}_2^T\overrightarrow{\mathbf{V}}_{\boldsymbol{\xi}'_k}\mathbf{f}_2 = \left[\overrightarrow{\mathbf{V}}_{\boldsymbol{\xi}'_k}\right]_{P,P}, \tag{37}$$

where $\overrightarrow{\mathbf{V}}_{\boldsymbol{\xi}'_k}$ is given by

$$\overrightarrow{\mathbf{V}}_{\boldsymbol{\xi}'_k} = \left(\left(\overrightarrow{\mathbf{V}}_{\tilde{\boldsymbol{\xi}}_k} + \sigma_w^2\mathbf{f}_2\mathbf{f}_2^T\right)^{-1} + \mathbf{A}^T\overleftarrow{\mathbf{W}}_{\tilde{\boldsymbol{\xi}}_{k+1}}\mathbf{A}\right)^{-1}. \tag{38}$$

■ *Messages Updating for Channel Estimation (on Subgraph 3):*

To deal with the inner product of channel vector $\mathbf{h}$ and symbol vector $\mathbf{s}_k$, we first consider the messages corresponding to the multiplier node. With $\overrightarrow{\mathbf{m}}_{\mathbf{s}''_k}$, $\overrightarrow{\mathbf{V}}_{\mathbf{s}''_k}$ and $\overrightarrow{V}_{\xi_k}$ computed on Subgraphs 2 and 4, using BP rule, the message from $\boxtimes$ to $\mathbf{h}_k$ reads

$$\begin{aligned}\overleftarrow{\mu}(\mathbf{h}_k) &\propto \int \delta(r'_k - \mathbf{h}_k^T\mathbf{s}_k)\overleftarrow{\mu}(\mathbf{s}''_k)\overrightarrow{\mu}(r'_k)\,\mathrm{d}\mathbf{s}''_k\,\mathrm{d}r'_k \\ &\propto \int \exp\left(-(\mathbf{s}''_k - \overrightarrow{\mathbf{m}}_{\mathbf{s}''_k})^H\overrightarrow{\mathbf{V}}_{\mathbf{s}''_k}^{-1}(\mathbf{s}''_k - \overrightarrow{\mathbf{m}}_{\mathbf{s}''_k})\right) \\ &\quad \times \exp\left(-\frac{(r_k - \mathbf{h}_k^T\mathbf{s}''_k)^2}{\overrightarrow{V}_{\xi_k}}\right)\mathrm{d}\mathbf{s}''_k \\ &\propto \exp\left(-\mathbf{h}_k^H\frac{\overrightarrow{\mathbf{m}}_{\mathbf{s}''_k}\overrightarrow{\mathbf{m}}_{\mathbf{s}''_k}^H}{\overrightarrow{V}_{\xi_k} + \mathbf{h}_k^H\overrightarrow{\mathbf{V}}_{\mathbf{s}''_k}\mathbf{h}_k}\mathbf{h_k}\right. \\ &\quad \left. + 2\mathbf{h}_k^H\frac{\overrightarrow{\mathbf{m}}_{\mathbf{s}''_k}r_k}{\overrightarrow{V}_{\xi_k} + \mathbf{h}_k^H\overrightarrow{\mathbf{V}}_{\mathbf{s}''_k}\mathbf{h}_k}\right).\end{aligned} \tag{39}$$

Note that it is difficult to formulate (39) to Gaussian. To this end, we resort to VMP [27] on the multiplier node to derive Gaussian messages. According to VMP update rules, the message $\overleftarrow{\mu}_{\mathbf{h}_k}$ follows

$$\overleftarrow{\mu}(\mathbf{h}_k) \propto \exp\left(\int \ln\delta(r'_k - \mathbf{h}_k^T\mathbf{s}''_k)b(\mathbf{s}''_k)b(r'_k)\,\mathrm{d}\mathbf{s}''_k\,\mathrm{d}r'_k\right), \tag{40}$$

where $b(\mathbf{s}''_k)$ and $b(r'_k)$ are the beliefs of $\mathbf{s}_k$ and r'_k.

Obviously, the logarithm of delta function involved in the integration (40) is pathological. To solve this problem, the multiplier node can be grouped with the noisy measurement to form a "soft" factor node f_k [28], as illustrated in Fig. 3. Then the message is obtained as

$$\begin{aligned}\overleftarrow{\mu}(\mathbf{h}_k) &\propto \exp\left(-\int\frac{(r_k - \mathbf{h}_k^T\mathbf{s}''_k)^2}{\overrightarrow{V}_{\xi_k}}\right. \\ &\quad \left.\times\exp\left(-(\mathbf{s}''_k - \mathbf{m}_{\mathbf{s}''_k})^H\mathbf{V}_{\mathbf{s}''_k}^{-1}(\mathbf{s}''_k - \mathbf{m}_{\mathbf{s}''_k})\right)\mathrm{d}\mathbf{s}''_k\right) \\ &\propto \exp\left(-\mathbf{h}_k^H\frac{\mathbf{V}_{\mathbf{s}''_k} + \mathbf{m}_{\mathbf{s}''_k}\mathbf{m}_{\mathbf{s}''_k}^H}{\overrightarrow{V}_{\xi_k}}\mathbf{h}_k + 2\mathbf{h}_k^H\frac{\mathbf{m}_{\mathbf{s}''_k}r_k}{\overrightarrow{V}_{\xi_k}}\right).\end{aligned} \tag{41}$$

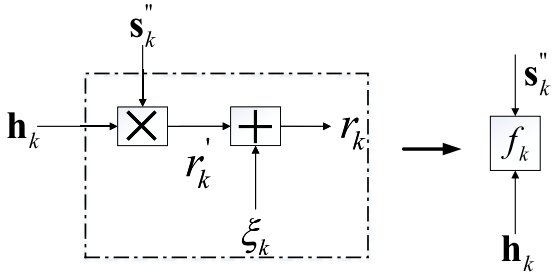


FIGURE 3. The equivalent "soft node" for multiplier node. The factor f_k denotes the probability density function of r_k conditioned on $\mathbf{s}_k$, $\mathbf{h}_k$ and ξ_k, which can be expressed as $f_k \propto \exp\left(-(r_k - \mathbf{h}_k^T \mathbf{s}''_k)^2 / \tilde{v}_{\xi_k}\right)$.

Since $\overleftarrow{\mu}(\mathbf{s}''_k)$ and $\overrightarrow{\mu}(\mathbf{s}''_k)$ have been obtained in Gaussian form, we can calculate $\mathbf{m}_{\mathbf{s}''_k}$ and $\mathbf{V}_{\mathbf{s}''_k}$ as

$$\mathbf{m}_{\mathbf{s}''_k} = \mathbf{V}_{\mathbf{s}''_k}\left(\overrightarrow{\mathbf{V}}^{-1}_{\mathbf{s}''_k}\overrightarrow{\mathbf{m}}_{\mathbf{s}''_k} + \overleftarrow{\mathbf{V}}^{-1}_{\mathbf{s}''_k}\overleftarrow{\mathbf{m}}_{\mathbf{s}''_k}\right), \tag{42}$$

$$\mathbf{V}_{\mathbf{s}''_k} = \left(\overrightarrow{\mathbf{V}}^{-1}_{\mathbf{s}''_k} + \overleftarrow{\mathbf{V}}^{-1}_{\mathbf{s}''_k}\right)^{-1}. \tag{43}$$

By noting that the incoming messages to the soft node are beliefs, (42) and (43) can be regarded as the posterior mean and variance of $\mathbf{s}_k$. Then $\overleftarrow{\mu}(\mathbf{h}_k)$ is calculated as Gaussian with mean vector and covariance matrix

$$\overleftarrow{\mathbf{V}}_{\mathbf{h}_k} = \overrightarrow{V}_{\xi_k}\left(\mathbf{V}_{\mathbf{s}''_k} + \mathbf{m}_{\mathbf{s}''_k}\mathbf{m}^H_{\mathbf{s}''_k}\right)^{-1}, \tag{44}$$

$$\overleftarrow{\mathbf{m}}_{\mathbf{h}_k} = \left(\mathbf{V}_{\mathbf{s}''_k} + \mathbf{m}_{\mathbf{s}''_k}\mathbf{m}^H_{\mathbf{s}''_k}\right)^{-1}\mathbf{m}_{\mathbf{s}''_k} r_k. \tag{45}$$

Likewise, $\overleftarrow{\mathbf{V}}_{\mathbf{s}''_k}$ and $\overleftarrow{\mathbf{m}}_{\mathbf{s}''_k}$ can be obtained as

$$\overleftarrow{\mathbf{V}}_{\mathbf{s}''_k} = \overrightarrow{V}_{\xi_k}\left(\mathbf{V}_{\mathbf{h}_k} + \mathbf{m}_{\mathbf{h}_k}\mathbf{m}^H_{\mathbf{h}_k}\right)^{-1}, \tag{46}$$

$$\overleftarrow{\mathbf{m}}_{\mathbf{s}''_k} = \left(\mathbf{V}_{\mathbf{h}_k} + \mathbf{m}_{\mathbf{h}_k}\mathbf{m}^H_{\mathbf{h}_k}\right)^{-1}\mathbf{m}_{\mathbf{h}_k} r_k, \tag{47}$$

with $\mathbf{m}_{\mathbf{h}_k}$ and $\mathbf{V}_{\mathbf{h}_k}$ computed in a similar way in (42) and (43).

E. COMPUTATION OF EXTRINSIC LLR

The equalizer calculates the extrinsic LLR based on the soft information $\overleftarrow{m}_{x_k}$ and $\overleftarrow{V}_{x_k}$.

$$\begin{aligned} L^e(c_{n,m}) &= \ln\frac{p(c_{n,m}=0|\mathbf{r})}{p(c_{n,m}=1|\mathbf{r})} - L^0(c_{n,m}) \\ &= \ln\frac{\sum_{\mathbf{d}_{i,m}=0} p(\mathbf{r}|\mathbf{c}_n=\mathbf{d}_i)p(\mathbf{c}_n=\mathbf{d}_i)}{\sum_{\mathbf{d}_{i,m}=0} p(\mathbf{r}|\mathbf{c}_n=\mathbf{d}_i)p(\mathbf{c}_n=\mathbf{d}_i)} - L^0(c_{n,m}), \end{aligned} \tag{48}$$

where $\mathbf{d}_i$ is the coded bit sequence corresponding to the constellation symbol χ_i. A concise representation of $L^e(c_{n,m})$ has been derived in [29], which reads

$$L^e(c_{n,m}) = \ln\frac{\sum_{\chi_i\in\mathcal{A}^0_m}\exp\left(-\frac{(\chi_i-\overleftarrow{m}_{x_k})^2}{\overleftarrow{V}_{x_k}}\right)\prod_{m'\neq m} p(c_{n,m'}=s_{i,m'})}{\sum_{\chi_i\in\mathcal{A}^1_m}\exp\left(-\frac{(\chi_i-\overleftarrow{m}_{x_k})^2}{\overleftarrow{V}_{x_k}}\right)\prod_{m'\neq m} p(c_{n,m'}=s_{i,m'})}, \tag{49}$$

where $\mathcal{A}^0_m$ and $\mathcal{A}^1_m$ denote the subset of $\mathcal{A}$ whose label in position m has the value 0 or 1. Then the LLRs $\{L^e(c_{n,m})\}$ are fed to the channel decoder. After decoding, the decoder outputs the extrinsic LLR and turns to the next iteration of equalization. The details of the proposed BP-EP-VMP algorithm are summarized in Algorithm 1.

Algorithm 1 The Proposed Hybrid BP-EP-VMP Approach to Joint Channel Estimation and Decoding for FTN Signaling Over Frequency Selective Channels

1: **Initialization:**
The output LLRs of channel decoder are initialized as $L^0(c_{m,n}) = 0$, i.e., $\overrightarrow{m}^0_{x_k} = 0$ and $\overrightarrow{V}^0_{x_k} = +\infty$.
The initial estimation of channel coefficients are obtained using 20 pilot symbols based on least square method.(In each turbo equalization, pilot symbols are also used to estimate the channel taps.) Then, the prior of $\mathbf{h}$ can be expressed as

$$p(\mathbf{h}) \propto \exp\left((\mathbf{h} - \overrightarrow{\mathbf{m}}^0_{\mathbf{h}})^H \overrightarrow{\mathbf{W}}^0_{\mathbf{h}} (\mathbf{h} - \overrightarrow{\mathbf{m}}^0_{\mathbf{h}})\right),$$

where $\overrightarrow{\mathbf{m}}^0_{\mathbf{h}}$ is the measured CSI and $\overrightarrow{\mathbf{W}}^0_{\mathbf{h}}$ is diagonal with the entries being $\left[\overrightarrow{\mathbf{W}}^0_{\mathbf{h}}\right]_{ii} = \sigma^2_{h_i}$. $\overrightarrow{m}^0_{x_k} = 0$, $\overrightarrow{V}^0_{x_k}$, $\overrightarrow{\mathbf{m}}^0_{\mathbf{h}}$ and $\overrightarrow{\mathbf{W}}^0_{\mathbf{h}}$ can be regarded as the prior information.
2: **for** Iter=1:I **do**
3: Calculate the messages from subgraph 1 to subgraph 2 according to (31) and (32);
4: Calculate the messages on subgraph 2 according to (33)-(36);
5: Calculate the messages on subgraph 3 according to (37) and (38);
6: Calculate the messages related to the "multiplier node" using (44)-(47);
7: Convert the outgoing messages to LLR based on (49) and feed them to channel decoder;
8: Perform BP channel decoding algorithm;
9: Calculate the incoming messages using EP as in (29) and (30);
10: **end for**

F. COMPLEXITY ANALYSIS

The complexity of the proposed algorithm is dominated by the matrix inversion operations in (18), (22), (36), (38), (43) and (44). For a non-sparse $K \times K$ matrix, the complexity for calculating its inverse is $\mathcal{O}(K^3)$. Here $\mathcal{O}$ denotes the order of time complexity. Taking (18) as an example, the total computational complexity is $\mathcal{O}(NL^3_{\text{FTN}})$ for a length-N symbol block. Then the total computational complexity of the proposed algorithm is $\mathcal{O}(N(L^3_{\text{FTN}} + L^3 + P^3))$, where L_{FTN} is the length of FTN-induced ISI considered at the receiver, L is the channel length and P is the order of AR model used to approximate the colored noise.

IV. SIMULATION RESULTS

In the simulations, we consider a 5/7-rate LDPC code with variable and check node degree distributions being $v(X) = 0.0005 + 0.2852X + 0.2857X^2 + 0.4286X^3$, $c(X) = 0.0017X^9 + 0.9983X^{10}$ [30]. The encoded bits are interleaved and mapped to a sequence of QPSK symbols. The number of transmitted symbols is $N = 2048$. The sequence of symbols passes through a root raised-cosine shaping filter with a roll-off factor $\alpha = 0.4$. The carrier frequency $f_0 = 2$GHz and the symbol period $T = 0.2\mu$s. The frequency selective fading channel for Nyquist signaling is assumed to have $L = 20$ taps and the coefficients $\{\tilde{h}_l\}$ are independently generated according to the distribution $\tilde{h}_l \sim \mathcal{N}(0, q^l)$. Then the channel taps for FTN signaling can be obtained by interpolation using DFT matrices. The normalized power delay profile is $q^l = \frac{\exp(-0.05l)}{\sum q^l}$. The number of iterations is set to $I = 10$ and the maximum number of BP decoding iterations is 50. The number of ISI taps due to FTN considered by the receiver is $L_{\text{FTN}} = 11$, unless otherwise specified. All simulation results are averaged over 1000 independent Monte Carlo trails.

We first evaluate the impact of packing factor τ. As shown in Fig. 4(a), BER performance versus the signal-to-noise ratio (SNR) of the proposed BP-EP-VMP algorithm with various τ are plotted. The performance of Nyquist signaling over the same channel is also included as a benchmark. It is observed that the proposed FTN receiver can attain the BER performance of the Nyquist signaling when the packing factor $\tau \geq 0.7$. Therefore, up to 40% of transmission rate can be increased with the same bandwidth by employing FTN signaling. Even for $\tau = 0.6$, the performance gap is less than 0.2dB, while the transmission rate in this case can be increased by more than 65%. In Fig. 4(b), we evaluate the BER performance with roll-off factor $\alpha = 0.05$. Since smaller α will lead to stronger ISI, it is seen that, compared with the Nyquist counterpart, the performance gap for $\tau = 0.6$ becomes about 0.5dB. Nevertheless, it is able to improve transmission rate up to 25% by employing FTN signaling with $\tau = 0.8$.

The complexity of the proposed algorithm depends on L_{FTN}, i.e., the length of FTN-induced ISI considered at receiver. In Fig. 5(a), BER performances with different $L_{\text{FTN}} = \{5, 11, 41\}$ are illustrated, where the roll-off factor $\alpha = 0.4$ and $\tau = 0.7$. It is seen that, $L_{\text{FTN}} = 5$ suffers from significant performance degradation due to the under-estimation of ISI induced by FTN. The increase of L_{FTN} helps to improve the BER performance, and the gain becomes marginal when $L_{\text{FTN}} \geq 11$. Therefore, the length of ISI induced by FTN signaling can be safely approximated by $L_{\text{FTN}} = 11$ in this case. We further consider a stronger packing scenario with $\tau = 0.5$ and the BER performance is illustrated in Fig. 5(b). We can see that $L_{\text{FTN}} = 11$ is not long enough to approximate the length of FTN-induced ISI and about 0.3dB performance loss can be observed compared to the $L_{\text{FTN}} = 41$ case. Therefore, In practice, we can compromise between BER performance and computational complexity by selecting a proper value of L_{FTN}.

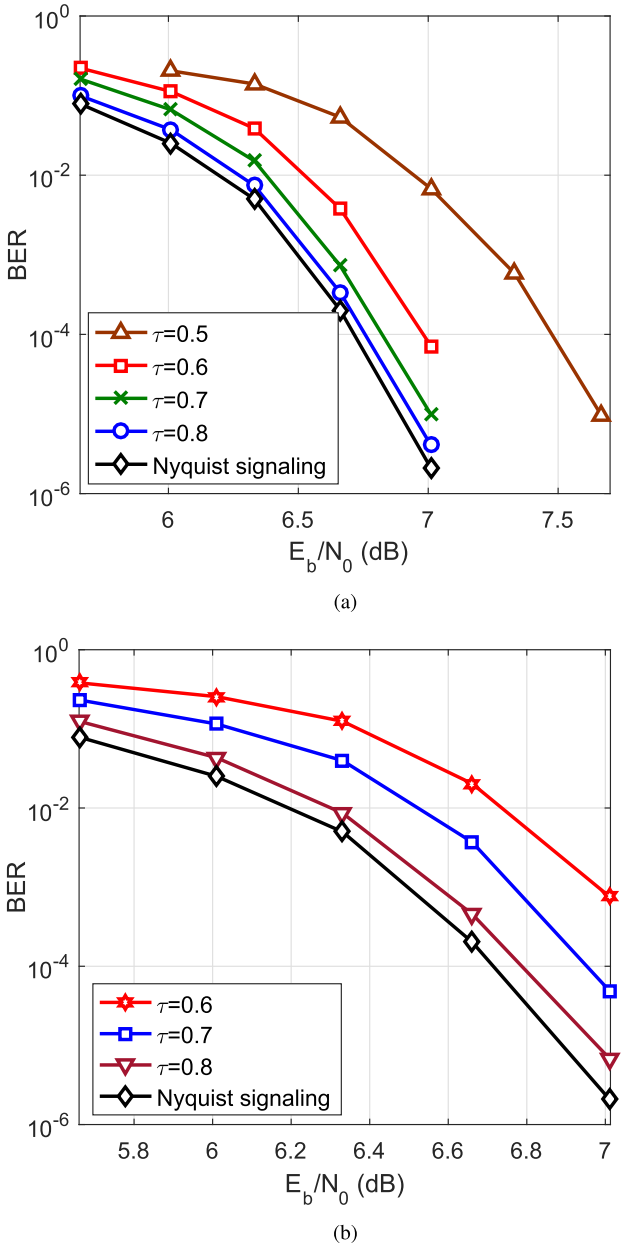
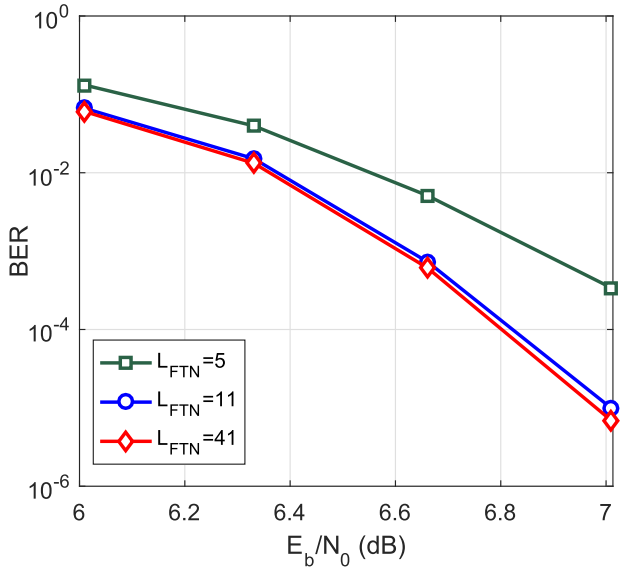
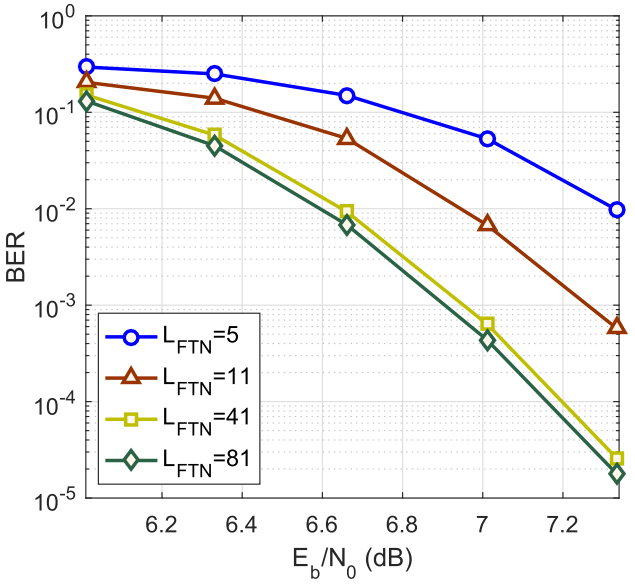


FIGURE 4. BER performance of the proposed algorithm for FTN system with different packing factor τ. The roll-off factor $\alpha = 0.4$ and $\alpha = 0.05$, respectively. (a) $\alpha = 0.4$. (b) $\alpha = 0.05$.

In Fig. 6, BER performance of the proposed BP-EP-VMP algorithm is compared with other methods. Since there is no existing work on this topic, we extend two Bayesian estimators to the FTN signaling over frequency selective channels, namely, the MMSE equalizer [14] and the variational inference (VI) method [16]. The BER performance with perfect channel information is also included as a reference. The MMSE equalizer can only treat the ISI caused by FTN signaling and the fading channel as a composite ISI channel, which leads to increased number of channel coefficients to be estimated. The VI method suffers from performance degradation due to its assumption that data symbols are conditional independent. The complexities of the MMSE equalizer and



(a)



(b)

FIGURE 5. Impact of L_{FTN} on BER performance. The roll-off factor $\alpha = 0.4$, $\tau = 0.7$ and $\tau = 0.5$, respectively. (a) $\tau = 0.7$. (b) $\tau = 0.5$.

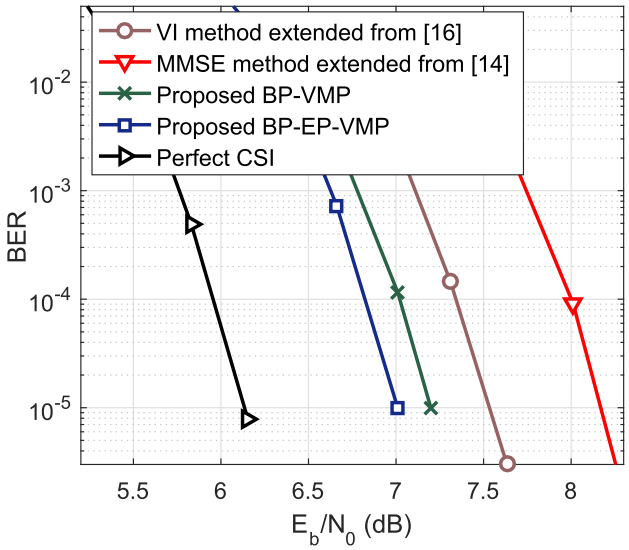


FIGURE 6. BER performance of different algorithms for considered FTN signaling system, with $\tau = 0.7$, $\alpha = 0.4$.

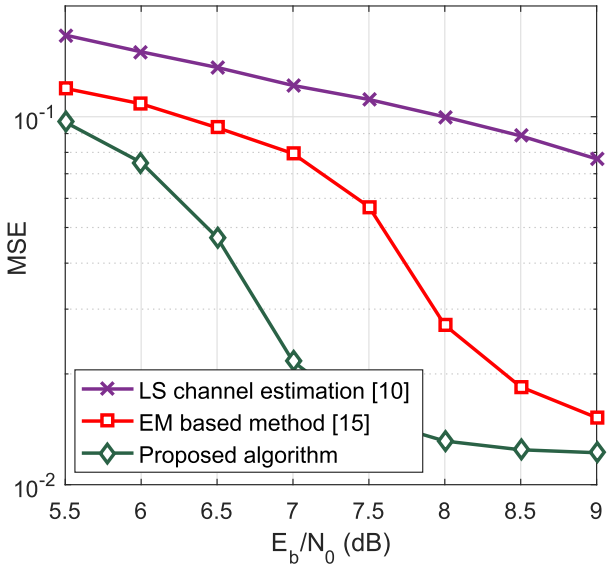


FIGURE 7. MSE of channel estimation of the proposed algorithm, with $\tau = 0.7$, $\alpha = 0.4$.

the VI method are $\mathcal{O}(N^3)$ and $\mathcal{O}(N(L_{\text{FTN}}^2 + L^2 + P^2))$, respectively. It is seen that the proposed BP-EP-VMP algorithm outperforms the other methods. When BP instead of EP is employed to update the messages from output of channel decoder to the equalizer, denoted as "BP-VMP" in In Fig. 6, about 0.2dB performance loss can be observed, which demonstrates the superior performance by employing EP.

The mean squared error (MSE) of channel estimation of the proposed BP-EP-VMP algorithm is evaluated in Fig. 7. For comparison, the performances of least square (LS) channel estimation [10] and the expectation-maximization (EM)-based method [15] are also included. The number of channel taps is 14 in $\tau = 0.7$ scenario. It is seen that the proposed BP-EP-VMP algorithm and the EM-based method significantly outperform the LS channel estimation that only uses the limited amount of pilot symbols. Moreover, the proposed BP-EP-VMP algorithm is superior to the EM-based method in [15], since the latter only provides hard channel estimation to the equalizer. Simulation results corroborate the benefits of the proposed joint channel estimation and decoding scheme.

V. CONCLUSIONS

In this paper, we proposed a low-complexity FTN receiver to perform joint channel estimation and decoding in frequency selective fading channels. The ISI imposed by FTN signaling

and that by the unknown frequency selective channel were considered separately, which enabled us to fully exploit the known structure of the FTN-induced ISI. Colored noise due to the faster sampling rate of FTN signaling was approximated by AR process to avoid using whitening filter. A Forney style factor graph, which consists four subgraphs, was constructed to represent the FTN system. We showed that using BP on the factor graph directly is infeasible, since the messages corresponding to the inner product of FTN symbol vector and the channel coefficient vector cannot be updated efficiently. We proposed to perform VMP on an equivalent ''soft node'' to tackle this problem. Moreover, EP was employed to efficiently convert the messages corresponding to FTN symbols obtained from channel decoder to Gaussian distribution. It was shown that, since the proposed hybrid BP-EP-VMP algorithm enabled effective Gaussian message approximation, the complexity only increases linearly with the block length N. Simulation results showed the superior performance of the proposed algorithm compared with the existing methods in FTN system. Compared with the Nyquist counterpart, FTN signaling with the proposed algorithm is able to increase the transmission rate over 40% in frequency selective fading channels, with negligible BER performance loss.

REFERENCES

[1] D. Tse and P. Viswanath, *Fundamentals Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[2] J. E. Mazo, ''Faster-than-Nyquist signaling,'' *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1451–1462, 1975.

[3] J. B. Anderson, F. Rusek, and V. Öwall, ''Faster-than-Nyquist signaling,'' *Proc. IEEE*, vol. 101, no. 8, pp. 1817–1830, Aug. 2013.

[4] A. Prlja and J. B. Anderson, ''Reduced-complexity receivers for strongly narrowband intersymbol interference introduced by faster-than-Nyquist signaling,'' *IEEE Trans. Commun.*, vol. 60, no. 9, pp. 2591–2601, Sep. 2012.

[5] A. Barbieri, D. Fertonani, and G. Colavolpe, ''Time-frequency packing for linear modulations: Spectral efficiency and practical detection schemes,'' *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 2951–2959, Oct. 2009.

[6] S. Sugiura, ''Frequency-domain equalization of faster-than-Nyquist signaling,'' *IEEE Wireless Commun. Lett.*, vol. 2, no. 5, pp. 555–558, Oct. 2013.

[7] P. Şen, T. Aktas, and A. O. Yilmaz, ''A low-complexity graph-based LMMSE receiver designed for colored noise induced by FTN-signaling,'' in *Proc. Wireless Commun. Network. Conf. (WCNC)*, Apr. 2014, pp. 642–647.

[8] S. Sugiura and L. Hanzo, ''Frequency-domain-equalization-aided iterative detection of faster-than-Nyquist signaling,'' *IEEE Trans. Veh. Technol.*, vol. 64, no. 5, pp. 2122–2128, May 2015.

[9] W. Yuan, N. Wu, H. Wang, and J. Kuang, ''Variational inference-based frequency-domain equalization for faster-than-Nyquist signaling in doubly selective channels,'' *IEEE Signal Process. Lett.*, vol. 23, no. 9, pp. 1270–1274, Sep. 2016.

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

[11] Y. Li, ''Pilot-symbol-aided channel estimation for OFDM in wireless systems,'' *IEEE Trans. Veh. Technol.*, vol. 49, no. 4, pp. 1207–1215, Apr. 2000.

[12] F. Qu and L. Yang, ''On the estimation of doubly-selective fading channels,'' *IEEE Trans. Wireless Commun.*, vol. 9, no. 4, pp. 1261–1265, Apr. 2010.

[13] J. Fan, S. Guo, X. Zhou, Y. Ren, G. Y. Li, and X. Chen, ''Faster-than-Nyquist signaling: An overview,'' *IEEE Access*, vol. 5, pp. 1925–1940, Sep. 2017.

[14] Y.-S. Choi, P. J. Voltz, and F. A. Cassara, ''On channel estimation and detection for multicarrier signals in fast and selective Rayleigh fading channels,'' *IEEE Trans. Commun.*, vol. 49, no. 8, pp. 1375–1387, Aug. 2001.

[15] C. Cozzo and B. L. Hughes, ''Joint channel estimation and data detection in space-time communications,'' *IEEE Trans. Commun.*, vol. 51, no. 8, pp. 1266–1270, Aug. 2003.

[16] F. Li, Z. Xu, and S. Zhu, ''Variational-inference-based data detection for OFDM systems with imperfect channel estimation,'' *IEEE Trans. Veh. Technol.*, vol. 62, no. 3, pp. 1394–1399, Mar. 2013.

[17] H. Niu, M. Shen, J. A. Ritcey, and H. Liu, ''A factor graph approach to iterative channel estimation and LDPC decoding over fading channels,'' *IEEE Trans. Wireless Commun.*, vol. 4, no. 4, pp. 1345–1350, Jul. 2005.

[18] G. E. Kirkelund and C. N. Manchön, L. P. Christensen, E. Riegler, and B. H. Fleury, ''Variational message-passing for joint channel estimation and decoding in MIMO-OFDM,'' in *Proc. IEEE Global Telecommun. Conf.*, Dec. 2010, pp. 1–6.

[19] S. Wu, L. Kuang, Z. Ni, J. Lu, D. D. Huang, and Q. Guo, ''Low-complexity iterative detection for large-scale multiuser MIMO-OFDM systems using approximate message passing,'' *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 902–915, Oct. 2014.

[20] N. Wu, W. Yuan, H. Wang, Q. Shi, and J. Kuang, ''Frequency-domain iterative message passing receiver for faster-than-Nyquist signaling in doubly selective channels,'' *IEEE Wireless Commun. Lett.*, vol. 5, no. 6, pp. 584–587, Dec. 2016.

[21] K. Takeuchi, M. Vehkapera, T. Tanaka, and R. R. Muller, ''Large-system analysis of joint channel and data estimation for MIMO DS-CDMA systems,'' *IEEE Trans. Inf. Theory*, vol. 58, no. 3, pp. 1385–1412, Mar. 2012.

[22] T. P. Minka, ''Expectation propagation for approximate Bayesian inference,'' in *Proc. 17th Conf. Uncertainty Artif. Intell.*, 2001, pp. 362–369.

[23] D. Dasalukunte, V. Owall, F. Rusek, and J. B. Anderson, *Faster Than Nyquist Signaling: Algorithms to Silicon*. The Netherlands: Springer, 2014.

[24] J. D. Gibson, B. Koo, and S. D. Gray, ''Filtering of colored noise for speech enhancement and coding,'' *IEEE Trans. Signal Process.*, vol. 39, no. 8, pp. 1732–1742, Aug. 1991.

[25] H. A. Loeliger, J. Dauwels, J. Hu, S. Korl, L. Ping, and F. R. Kschischang, ''The factor graph approach to model-based signal processing,'' *Proc. IEEE*, vol. 95, no. 6, pp. 1295–1322, Jun. 2007.

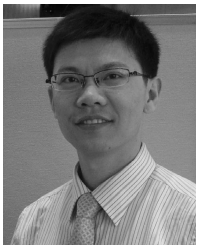
[26] J. Céspedes, P. M. Olmos, M. Sänchez-Fernández, and F. Perez-Cruz, ''Expectation propagation detection for high-order high-dimensional MIMO systems,'' *IEEE Trans. Commun.*, vol. 62, no. 8, pp. 2840–2849, Sep. 2014.

[27] J. M. Winn and C. M. Bishop, ''Variational message passing,'' *J. Mach. Learn. Res.*, vol. 6, pp. 661–694, Apr. 2005.

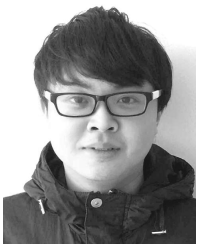
[28] J. Dauwels, A. Eckford, S. Korl, and H.-A. Loeliger. (Oct. 2009). ''Expectation maximization as message passing—Part I: Principles and Gaussian messages.'' [Online]. Available: arXiv preprint arXiv:0910.2832

[29] Q. Guo and D. D. Huang, ''A concise representation for the soft-in soft-out LMMSE detector,'' *IEEE Commun. Lett.*, vol. 15, no. 5, pp. 566–568, Sep. 2011.

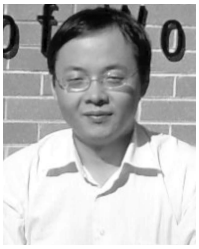
[30] J. Xu, L. Chen, I. Djurdjevic, S. Lin, and K. Abdel-Ghaffar, ''Construction of regular and irregular LDPC codes: Geometry decomposition and masking,'' *IEEE Trans. Inf. Theory*, vol. 53, no. 1, pp. 121–134, Jan. 2007.



NAN WU (M'11) received the B.S., M.S., and Ph.D. degrees from the Beijing Institute of Technology (BIT), Beijing, China, in 2003, 2005, and 2011, respectively. From 2008 to 2009, he was a Visiting Ph.D. Student with the Department of Electrical Engineering, Pennsylvania State University, USA. He is currently an Associate Professor with the School of Information and Electronics, BIT. His research interests include signal processing in wireless communication networks. He was a recipient of the National Excellent Doctoral Dissertation Award by MOE of China in 2013. He serves as an Editorial Board Member of the *International Journal of Electronics and Communications*, the *KSII Transactions on Internet and Information Systems*, and the *IEICE Transactions on Communications*.



WEIJIE YUAN (S'15) received the B.S. degree from the Beijing Institute of Technology, Beijing, China, in 2013, where he is currently pursuing the Ph.D. degree with the School of Information and Electronics. In 2016, he was a Visiting Ph.D. Student with the Institute of Telecommunications, Vienna University of Technology, Austria. His research interests include statistical inference on graphical models. He has served as a TPC Member of WCSP 2016 and VTC 2017-Spring.



QINGHUA GUO (S'07–M'08) received the B.E. degree in electronic engineering and the M.E. degree in signal and information processing from Xidian University, Xian, China, in 2001 and 2004, respectively, and the Ph.D. degree in electronic engineering from City University of Hong Kong, Hong Kong, in 2008.

He is currently a Senior Lecturer with the School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Wollongong, NSW, Australia, and also an Adjunct Associate Professor with the School of Electrical, Electronic and Computer Engineering, The University of Western Australia, Perth, WA, Australia. His research interests include signal processing and telecommunications. He was a recipient of the Australian Research Council's Discovery Early Career Researcher Award.

JINGMING KUANG received the Ph.D. degree in electrical engineering from the Technical University of Berlin, Berlin, Germany, in 1988. He is the Founder of the Ericsson Research Center of Digital Communications, Beijing Institute of Technology (BIT), which was built in 1999. From 1993 to 2007, he was the Vice-President and the President of BIT. He has been a Professor with the School of Information and Electronics, BIT, since 1989. His research interests include theory and techniques of wireless communication and digital signal processing. He was a recipient of the Excellent Returnee from Abroad, the Outstanding Scholar with Extraordinary Achievements, and the Outstanding Scholar of National Defense Department.