

Parametric Bilinear Iterative Generalized Approximate Message Passing Reception of FTN Multi-Carrier Signaling

Yunsi Ma, *Student Member, IEEE*, Nan Wu, *Member, IEEE*, J. Andrew Zhang, *Senior Member, IEEE*, Bin Li, *Member, IEEE*, and Lajos Hanzo, *Fellow, IEEE*

Abstract—A low-complexity parametric bilinear generalized approximate message passing (PBiGAMP)-based receiver is conceived for multi-carrier faster-than-Nyquist (MFTN) signaling over frequency-selective fading channels. To mitigate the inherent ill-conditioning problem of MFTN signaling, we construct a segment-based frequency-domain received signal model in the form of a block circulant linear transition matrix, which can be efficiently calculated by applying a two dimensional fast Fourier transform. Based on the eigenvalue decomposition of the block circulant matrices, we can diagonalize the covariance matrix of the complex-valued colored noise process imposed by the associated two dimensional non-orthogonal matched filtering. Building on this model, a PBiGAMP-based parametric joint channel estimation and equalization (JCEE) algorithm is proposed for MFTN systems. In this algorithm, we introduce a pair of additive terms for characterizing the interferences arising from adjacent segments and employ the exact discrete *a priori* probabilities of the transmitted symbols for improving the bit error rate (BER) performance. To further enhance the system's robustness in the presence of ill-conditioned matrices, we develop a refined PBiGAMP-based JCEE algorithm by introducing a series of scaled identity matrices. Moreover, the proposed PBiGAMP-based JCEE algorithms may be readily decomposed into GAMP-based equalization algorithms, when the channel state information is perfectly known. The overall complexity of the proposed algorithms only increases logarithmically with the total number of transmitted symbols. Our simulation results demonstrate the benefits of the proposed PBiGAMP-based iterative message passing receiver conceived for MFTN signaling.

Index Terms—Multicarrier faster-than-Nyquist signaling, channel estimation, frequency-domain equalization, parametric bilinear generalized approximate message passing, complex-valued colored noise.

I. INTRODUCTION

Faster-than-Nyquist (FTN) signaling constitutes an innovative spectrally efficient non-orthogonal signaling solution

This work was supported by the “National Science Foundation of China (NSFC)” (Grant Nos. 61971041, 62001027). L. Hanzo would like to acknowledge the financial support of the Engineering and Physical Sciences Research Council projects EP/P034284/1 and EP/P003990/1 (COALESCE) as well as of the European Research Council's Advanced Fellow Grant QuantCom (Grant No. 789028) (*Corresponding Author: Nan Wu.*)

Y. Ma, N. Wu and B. Li are with the School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China (e-mail: yunsi@bit.edu.cn; wunan@bit.edu.cn; binli@bit.edu.cn).

J. Andrew Zhang is with the Global Big Data Technologies Centre and the School of Electrical and Data Engineering, University of Technology Sydney, Sydney, NSW 2007, Australia (e-mail: andrew.zhang@uts.edu.au).

L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, U.K. (e-mail: lh@ecs.soton.ac.uk).

for high-speed wireless communication systems [1]–[3]. By abandoning the orthogonality condition of different modulated waveforms, FTN signaling becomes capable of improving the transmission rate without requiring additional bandwidth. Hence, it achieves higher capacity than Nyquist signaling [4]. Given its benefits, the concept of FTN has also been extended to two dimensional (2D) multi-carrier transmission systems for further improving its the bandwidth efficiency (BE) [5].

Explicitly, multi-carrier faster-than-Nyquist (MFTN) further improves the BE by ‘squeezing’ the Nyquist interval of the signaling pulse in the time domain (TD) and the minimum orthogonal frequency spacing in the frequency domain (FD). Hence, we can optimize the 2D distance between the signals for maximizing the achievable BE [6]. In recent years, MFTN signaling has been investigated both in satellite systems [7] and high-capacity optical systems [8], [9]. However, the severe inherent intersymbol interferences (ISIs) and intercarrier interferences (ICIs) imposed by time-frequency packing lead to a potentially excessive detection complexity, which may limit the practical applications of MFTN systems.

Nonetheless, substantial research efforts have been dedicated to designing efficient receivers for MFTN systems. In [5], a maximum *a posteriori* (MAP) probability receiver based on the Bahl-Cocke-Jelinek-Raviv (BCJR) algorithm [10] was proposed for approaching the bit error rate (BER) performance of Nyquist systems, albeit at the expense of an exponentially escalating complexity. An attraction 2D minimum mean-squared error (MMSE)-based equalizer was designed for the interference mitigation of MFTN signaling in [11], where authors succeed in rendering the complexity independent of both the number of subcarriers and of the constellation size. However, its BER performance was significantly degraded in severe ICI scenarios. To tackle this problem, a one dimensional MMSE equalizer coupled with successive interference cancellation (SIC) was also designed in [11] for approaching the MAP-SIC equalizer's performance in the high signal-to-noise ratio (SNR) region. However, the scheme of [11] still exhibited a high computational complexity owing to the associated multiple matrix inversion operations. Note that the aforementioned equalizers cannot deal with the colored noise induced by the non-orthogonal matched filter of MFTN systems. Recently, a parametric Gaussian message passing receiver based on a novel truncated 2D trellis structure was proposed for eliminating the severe 2D interferences encountered by MFTN signaling in frequency-selective channels [12],

TABLE I: Boldly and explicitly contrasting our proposed PBiGAMP-based receiver to the typical FTN receivers, where \checkmark and the blank indicate the ability to deal with the challenges.

Themes of Challenges	This Paper	[5]	[11]	[12]	[13]	[15]	[16]
ISIs & ICIs	\checkmark	\checkmark	\checkmark	\checkmark			
Fading Channels	\checkmark			\checkmark	\checkmark	\checkmark	\checkmark
Colored Noise	\checkmark			\checkmark	\checkmark	\checkmark	\checkmark
Discrete Priors	\checkmark					\checkmark	
Ill-Conditioned Problem	\checkmark						
JCEE	\checkmark				\checkmark	\checkmark	\checkmark
Equalization	\checkmark						

where a realistic non-diagonal colored noise covariance matrix was used for updating the messages on the associated factor graph. Nevertheless, an inevitable BER performance loss is caused by the Gaussian approximation of the discrete *a priori* distributions of the transmitted symbols.

All of the above-mentioned MFTN receivers were investigated under the idealized simplifying assumption of perfectly known channel state information (CSI). In practice, perfect CSI is unattainable, but it may be approached by joint channel estimation and equalization (JCEE) techniques even in frequency-selective fading channels. Having said that, no JCEE solutions have been reported in the open literature for MFTN systems. This is because it is difficult to extend the existing JCEE algorithms of single-carrier FTN systems [13]–[16] to MFTN systems due to the intractable coupling of the Doppler spreads, frequency selectivity and severe 2D interferences. Moreover, the existing JCEE algorithms of single-carrier FTN systems can only deal with the one-dimensional interferences and cannot alleviate the ill-conditioning problem imposed by time-frequency packing.

Motivated by tackling these challenges, we appropriately adapt the parametric bilinear generalized approximate message passing (PBIGAMP) algorithm of [17] for solving the high-dimensional random bilinear mixing recovery problem. The associated ill-conditioning problem¹ caused by the non-orthogonal nature of MFTN signaling often prevents the PBIGAMP algorithm for readily converge, but still no solutions have been found for improving the convergence of the PBIGAMP algorithm. Hence, we develop a low-complexity PBIGAMP-based FD iterative message passing receiver, which is robust to the above-mentioned impediments of MFTN systems operating in frequency-selective fading channels. With the objective of tackling these challenges of MFTN signaling, TABLE I boldly and explicitly contrasts the novel features of our PBIGAMP-based receiver to those of the existing FTN receivers. The main contributions of this work are related to innovative signal modeling and efficient algorithm development, which are elaborated on below:

- We derive a novel segment-based FD received signal model through inserting a few cyclic postfixes in the FD and through truncating the inherent 2D interferences. By

¹When the minimum eigenvalue of the interference matrix approaches zero, the linear system tends to become ill-conditioned [18]. As discussed in [19], their ill-conditioning cannot be completely overcome by any unbiased estimator. Hence, we can only mitigate the effect of ill-conditioning on the BER performance, but cannot completely eliminate it.

reformulating the linear transition matrix into one having a block circulant structure, we can reduce its condition number and then mitigate the ill-conditioning problem of MFTN signaling. Moreover, we reconstruct a block circulant covariance matrix for the equivalent colored noise process induced by the non-orthogonal matched filter and hence diagonalize it based on the eigenvalue decompositions of the block circulant matrices.

- Based on the proposed segment-based FD received signal model, we develop a pair of PBIGAMP-based FD JCEE algorithms for MFTN systems. With the aid of several average approximations, the refined version not only circumvents the strong sensitivity of the ill-conditioned matrices to small perturbations, but also reduces the computational complexity. Instead of using the typical Gaussian approximations of the transmitted symbols that rely on the moment matching method, the exact discrete *a priori* probabilities are used for improving the symbol detection performance. To eliminate the 2D interferences engendered by the adjacent segments, we introduce two additive terms for improving the accuracy of the “pseudo” *a priori* mean vector of the noiseless measurements. In the case of known CSI, the proposed JCEE algorithms can be readily decomposed into two FD equalization (FDE) algorithms.

The rest of the paper is organized as follows. Section II presents the system model of MFTN signaling in frequency-selective fading channels. In Section III, we construct a segment-based FD received signal model and then discuss its effects on the ill-conditioning problem of MFTN signaling. The proposed PBIGAMP-based iterative receiver is described in Section IV. The BER and NMSE performance results of the proposed algorithms are evaluated and compared to the state-of-the-art algorithms in Section V. Finally, our conclusions are offered in Section VI.

Notations: Boldface capital and small letters represent matrices and vectors, respectively. Specifically, x_n and $X_{m,n}$ denote the n -th element of the vector \mathbf{x} and the (m,n) -th element of the matrix \mathbf{X} , respectively. \mathbf{F}_N is the $N \times N$ normalized discrete Fourier transform (DFT) matrix with $F_{m,n} = N^{-1/2} \exp(-j2\pi(m-1)(n-1)/N)$. \mathbf{I}_N and $\mathbf{0}_{M \times N}$ denote an identity matrix of size $N \times N$ and all-zero matrix of size $M \times N$, while $\mathbf{0}_N$ and $\mathbf{1}_N$ denote all-zero and all-one column-vectors of length N , respectively. Furthermore, $\mathcal{BC}_{M,N}$ denotes the set of $MN \times MN$ block circulant matrices

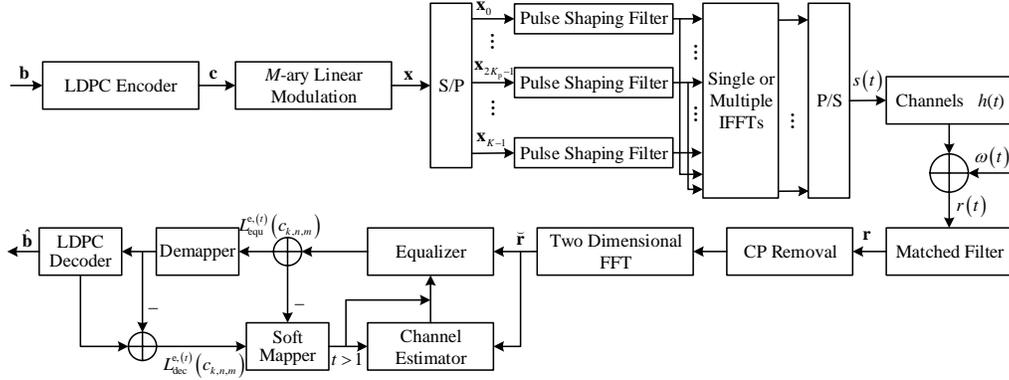


Fig. 1: Block diagram of LDPC-coded MFTN transceiver.

associated with M^2 arbitrary square matrices. $\mathcal{D}(\mathbf{x})$ or $\mathcal{D}(\mathbf{X})$ represents the diagonal matrix constructed from the vector \mathbf{x} or from the main diagonal vector of the square matrix \mathbf{X} . The operations $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$, $(\cdot)^{-1}$ and $\text{tr}(\cdot)$ denote the complex conjugate, transpose, conjugate transpose, matrix inverse and trace operator. The operators $*$, \otimes , \odot , \oslash , \otimes , and \propto denote linear convolution, cyclic convolution, element-wise product, element-wise division, Kronecker product, and equality up to a constant normalization factor. The operation $\lfloor \cdot \rfloor$ returns the nearest integer less than or equal to the element. The function $\text{rem}(a, b)$ returns the remainder after division of a by b . $\mathcal{CN}(x; m_x, v_x)$ represents a complex Gaussian distribution of variable x with mean m_x and variance v_x . Moreover, $\mathbb{E}\{\cdot\}$ and $\mathbb{V}\{\cdot\}$ denote the expectation and variance operations. For readability, the key variable notations are summarized in TABLE II.

TABLE II: List of Variable Notations

N_b	The number of the uncoded input bits \mathbf{b}
N_c	The number of the coded bits \mathbf{c}
N	The number of symbols for each subcarrier
K	The number of subcarriers without cyclic postfixes for each symbol
K_p	One half of the number of cyclic postfixes for each symbol
N_I or K_I	The truncated ISI or ICI intervals
L_h	The length of channel memory
L_q	The number of segments
L_m	The number of the transmitted symbols in each segment
N_s	The number of the received symbols in each segment
x, \mathbf{x}	M -ary linear modulated symbol in the FD
$s(t)$	Baseband transmitted signal in the TD
$h(t, \iota)$	Channel impulse response in the TD
$r(t), \mathbf{r}$ $\tilde{\mathbf{r}}, \check{\mathbf{r}}$	Received MFTN signal in the TD or FD
$\omega(t), \boldsymbol{\omega}$	Communication additive noise in the TD
\mathbf{v}	Equivalent additive colored noise in the TD or FD

II. SYSTEM MODEL

The general transceiver structure of low density parity check (LDPC)-coded MFTN systems is illustrated in Fig. 1. The input bits $\mathbf{b} = [b_0, \dots, b_{N_b-1}]^T$ are encoded to obtain N_c LDPC-coded bits $\mathbf{c} = [c_0, \dots, c_{N_c-1}]^T$. Considering a MFTN system having K subcarriers and each subcarrier containing N symbols in the TD, the coded bits are mapped onto M -ary linear modulation constellation points to generate the independently and identically distributed modulated symbols. Then, the serial modulated symbol sequence \mathbf{x} is partitioned into K parallel subsequences, where $\mathbf{x} = [\mathbf{x}_0^T, \dots, \mathbf{x}_{K-1}^T]^T$ and $\mathbf{x}_k = [x_{k,0}, \dots, x_{k,N-1}]^T$ contains N modulated symbols of the k -th subcarrier. Furthermore, the modulated symbols are passed through a pulse shaping filter $p(t)$ associated with the symbol interval τT_0 , where $\tau \in (0, 1]$ is the time packing factor and T_0 is the Nyquist time interval. After that, the symbols of the first $2K_p$ subcarriers are concatenated to the original symbol sequence as cyclic postfixes in the FD. To further improve the BE, the modulated symbols are transmitted on K overlapped subcarriers with the frequency spacing νF_0 , where $\nu \in (0, 1]$ is the frequency packing factor and F_0 is the minimum orthogonal frequency spacing. This operation can be efficiently implemented by single or multiple inverse fast Fourier transform (IFFT) modules [20]. After the parallel to serial (P/S) conversion, the baseband transmitted MFTN signal is expressed as

$$s(t) = \sum_{k=0}^{\tilde{K}-1} \sum_{n=0}^{N-1} x_{k,n} p(t - n\tau T_0) e^{j2\pi k\nu F_0 t}, \quad (1)$$

where $\tilde{K} = K + 2K_p$, k and n are the subcarrier index and the time index, respectively.

At the output of a frequency-selective fading channel having a channel impulse response (CIR) $h(t, \iota)$, the received signal is given by

$$r(t) = \int_{-\infty}^{+\infty} h(t, \iota) s(t - \iota) d\iota + \omega(t), \quad (2)$$

where $h(t, \iota) = 0$ for $\iota < 0$ and $\iota > \iota_{\max}$, ι_{\max} is the maximum delay spread, and $\omega(t)$ is an additive white Gaussian noise (AWGN) process with zero mean and variance σ_0^2 .

We assume perfect synchronization between the transmitter and the receiver. The output of the non-orthogonal matched

filter is derived as

$$\begin{aligned}
r_{k_r, n_r} &= \int_{-\infty}^{\infty} r(t)p(t - n_r\tau T_0)e^{-j2\pi k_r\nu F_0 t} dt \\
&= \sum_{l=0}^{L_h-1} h_{n_r, l} \sum_{k_t=0}^{\bar{K}-1} \sum_{n_t=0}^{N-1} x_{k_t, n_t} e^{-j2\pi k_t l \nu F_0 \tau T_0} \int_{-\infty}^{\infty} p(t - n_r\tau T_0) \\
&\quad \times p(t - (n_t + l)\tau T_0) e^{-j2\pi(k_r - k_t)\nu F_0 t} dt + \omega_{k_r, n_r} \\
&= \sum_{l=0}^{L_h-1} h_{n_r, l} \sum_{k_t=0}^{\bar{K}-1} \sum_{n_t=0}^{N-1} A_p((n_\Delta + l)\tau T_0, k_\Delta \nu F_0) \psi_{k_\Delta, n_r}^{k_t, l} x_{k_t, n_t} \\
&\quad + \omega_{k_r, n_r},
\end{aligned} \tag{3}$$

where $n_\Delta = n_t - n_r$, $k_\Delta = k_t - k_r$, k_r and n_r are the indices of the received symbols, k_t and n_t are the indices of the transmitted symbols, $h_{n_r, l}$ is the l -th channel coefficient at the time index n_r , $A_p(\tau T_0, \nu F_0) = \int_{-\infty}^{+\infty} p(t)p(t - \tau T_0)e^{j2\pi\nu F_0 t} dt$ is the ambiguity function, $\psi_{k_\Delta, n_r}^{k_t, l} = e^{j2\pi(k_\Delta n_r - k_t l)\nu F_0 \tau T_0}$ and finally $\omega_{k_r, n_r} = \int_{-\infty}^{\infty} \omega(t)p(t - n_r\tau T_0)e^{-j2\pi k_r \nu F_0 t} dt$ is the colored noise process having a covariance matrix \mathbf{R}_ω with the elements of $\mathbb{E}\{\omega_{k_1, n_1} \omega_{k_2, n_2}^*\} = \sigma_\omega^2 A_p[(n_1 - n_2)\tau T_0, (k_1 - k_2)\nu F_0]$.

III. SIGNAL MODELING OF FREQUENCY-DOMAIN SEGMENT-BASED RECEIVER

Since the 2D interference matrix of (3) suffers from potential ill-conditioning problem, the solutions of the bilinear JCEE problem are sensitive to small perturbations. The condition number can be regarded as an indicator of whether the linear transition matrix is ill-conditioned or well-conditioned. The condition number of the 2D interference matrix in (3) increases with the reduction of the packing factors and hence the negative effects of the ill-conditioning become worse. In this section, we construct a novel received signal model assisting in mitigating the ill-conditioning of MFTN signaling.

According to the received signal model of (3), the inevitable ISIs and ICIs of MFTN signaling can be characterized by the ambiguity function $A_p(\tau T_0, \nu F_0)$. The inherent 2D interferences imposed by time-frequency packing depend on both the shaping pulse and on the packing factors, but they reduced upon increasing the intervals of the symbols and subcarriers. In this paper, we employ a root raised cosine (RRC) pulse shaping filter having a roll-off factor β . Considering that the ISIs and ICIs are dominated by the neighboring symbols and subcarriers, we can reduce the demodulation complexity of MFTN systems by employing a truncated 2D interference model, i.e., $A_p[(n_\Delta + l)\tau T_0, k_\Delta \nu F_0] = 0$ for $|n_\Delta + l| > N_1$ or $|k_\Delta| > K_1$, where N_1 and K_1 are the truncated ISI and ICI intervals.

To develop an efficient MFTN receiver for frequency-selective fading channels, we partition the transmitted symbols of the k -th subcarriers \mathbf{x}_k into L_q short segments of length L_m given by $\mathbf{x}_{k_i, q} = [x_{k_i, q L_m}, \dots, x_{k_i, (q+1)L_m-1}]^T$, where $q = 0, \dots, L_q - 1$, $N = L_q L_m$, $L_g - 1 \leq L_m \leq N$, and $L_g = L_h + 2N_1$. For the time-invariant or quasi-static channels, L_m can be set to N . Without loss of generality, we assume that the CSI is fixed for the duration of a single segment [15]. The observation vector of the k -th subcarrier in the q -th segment,

i.e., $\mathbf{r}_{k_r, q} = [r_{k_r, q L_m}, \dots, r_{k_r, q L_m + N_s - 1}]^T$, $N_s = L_m + L_g - 1$, can be expressed as

$$\begin{aligned}
\mathbf{r}_{k_r, q} &= \sum_{l=0}^{L_h-1} h_{q, l} \sum_{k_t=k_r-K_1}^{k_r+K_1} \mathcal{D}(\psi_{k_\Delta, q}^{k_t, l})(\mathbf{g}_{k_\Delta, l} * \mathbf{x}_{k_t, q}) \\
&\quad + \mathbf{z}_{k_r, q-1}^{\text{tail}} + \mathbf{z}_{k_r, q+1}^{\text{head}} + \omega_{k_r, q},
\end{aligned} \tag{4}$$

where $h_{q, l}$ is the l -th channel coefficient in the q -th segment, $\psi_{k_\Delta, q}^{k_t, l} = [\psi_{k_\Delta, q L_m}^{k_t, l}, \dots, \psi_{k_\Delta, q L_m + N_s - 1}^{k_t, l}]^T$, $\mathbf{g}_{k_\Delta, l} = [\mathbf{0}_l^T \quad \mathbf{g}_{k_\Delta}^T \quad \mathbf{0}_{(L_h-1-l)}^T]^T$ represents the significant ISIs and ICIs within one segment, $\mathbf{g}_{k_\Delta} = [A_p(N_1\tau T_0, k_\Delta \nu F_0), \dots, A_p(-N_1\tau T_0, k_\Delta \nu F_0)]^T$, and $\omega_{k_r, q}^T = [\omega_{k_r, q L_m}, \dots, \omega_{k_r, q L_m + N_s - 1}]^T$ is the truncated colored noise vector. Moreover, $\mathbf{z}_{k_r, q-1}^{\text{tail}}$ and $\mathbf{z}_{k_r, q+1}^{\text{head}}$ characterize the 2D interferences imposed by the tail of the $(q-1)$ -th segment and by the head of the $(q+1)$ -th segment, which are given by

$$\begin{aligned}
\mathbf{z}_{k_r, q-1}^{\text{tail}} &= \Xi_t \left[\sum_{l=0}^{L_h-1} h_{q-1, l} \sum_{k_t=k_r-K_1}^{k_r+K_1} \mathcal{D}(\psi_{k_\Delta, q-1}^{k_t, l})(\mathbf{g}_{k_\Delta, l} * \mathbf{x}_{k_t, q-1}) \right], \\
\mathbf{z}_{k_r, q+1}^{\text{head}} &= \Xi_h \left[\sum_{l=0}^{L_h-1} h_{q+1, l} \sum_{k_t=k_r-K_1}^{k_r+K_1} \mathcal{D}(\psi_{k_\Delta, q+1}^{k_t, l})(\mathbf{g}_{k_\Delta, l} * \mathbf{x}_{k_t, q+1}) \right],
\end{aligned} \tag{5}$$

where $\Xi_t = [\mathbf{0}_{(L_g-1) \times L_m} \quad \mathbf{I}_{L_g-1}; \quad \mathbf{0}_{L_m \times L_m} \quad \mathbf{0}_{L_m \times (L_g-1)}]$, $\Xi_h = [\mathbf{0}_{L_m \times (L_g-1)} \quad \mathbf{0}_{L_m \times L_m}; \quad \mathbf{I}_{L_g-1} \quad \mathbf{0}_{(L_g-1) \times L_m}]$, and $\mathbf{z}_{k_r, q-1}^{\text{tail}} = \mathbf{z}_{k_r, L_q}^{\text{head}} = \mathbf{0}_{N_s}$.

By appending the required number of zeros to both the truncated interference vector $\mathbf{g}_{k_\Delta, l}$ and the transmitted symbol vector $\mathbf{x}_{k_t, q}$, the linear convolution in (4) can be converted to cyclic convolution as

$$\mathbf{r}_{k_r, q} = \sum_{l=0}^{L_h-1} h_{q, l} \sum_{k_t=k_r-K_1}^{k_r+K_1} \mathcal{D}(\psi_{k_\Delta, q}^{k_t, l})(\mathbf{g}_{k_\Delta, l} \circledast \Phi_m \mathbf{x}_{k_t, q}) + \mathbf{v}_{k_r, q}, \tag{7}$$

where $\Phi_g = [\mathbf{I}_{L_g} \quad \mathbf{0}_{L_g \times (N_s - L_g)}]^T$ is used for inserting $N_s - L_g$ zeros into the truncated interference vector, $\Phi_m = [\mathbf{I}_{L_m} \quad \mathbf{0}_{L_m \times (N_s - L_m)}]^T$ is exploited for inserting $N_s - L_m$ zeros into the transmitted data vector, and $\mathbf{v}_{k_r, q} = \mathbf{z}_{k_r, q-1}^{\text{tail}} + \mathbf{z}_{k_r, q+1}^{\text{head}} + \omega_{k_r, q}$ is the equivalent colored noise process.

According to the properties of cyclic convolution, the received signal in (7) is rewritten as

$$\mathbf{r}_{k_r, q} = \sum_{l=0}^{L_h-1} h_{q, l} \sum_{k_t=k_r-K_1}^{k_r+K_1} \mathcal{D}(\psi_{k_\Delta, q}^{k_t, l}) \mathbf{F}_{N_s}^H \mathcal{D}(\bar{\mathbf{g}}_{k_\Delta, l}) \mathbf{F}_{N_s} \Phi_m \mathbf{x}_{k_t, q} + \mathbf{v}_{k_r, q}, \tag{8}$$

where $\bar{\mathbf{g}}_{k_\Delta, l} = \sqrt{N_s} \mathbf{F}_{N_s} \Phi_g \mathbf{g}_{k_\Delta, l}$. By removing the first and last K_1 received subcarriers, the received symbols in the q -th segment are expressed as

$$\begin{aligned}
\mathbf{r}_q &= \sum_{l=0}^{L_h-1} h_{q, l} \Phi_l \Theta_q (\mathbf{I}_K \otimes \mathbf{F}_{N_s})^H \bar{\mathbf{G}}_l (\mathbf{I}_K \otimes \mathbf{F}_{N_s}) (\mathbf{I}_K \otimes \Phi_m) \mathbf{x}_q \\
&\quad + \mathbf{v}_q,
\end{aligned} \tag{9}$$

where we have $\mathbf{r}_q = [\mathbf{r}_{K_1, q}^T, \dots, \mathbf{r}_{K-K_1-1, q}^T]^T$, $\Phi_l = \mathcal{D}(\phi_l) \otimes \mathbf{I}_{N_s}$ is a $KN_s \times KN_s$ diagonal matrix, $\phi_l = [1, e^{-j2\pi l \nu F_0 \tau T_0}, \dots, e^{-j2\pi(K-1)l \nu F_0 \tau T_0}]^T$, Θ_q and $\bar{\mathbf{G}}_l$ are

$KN_s \times KN_s$ block circulant matrices with diagonal blocks, $\mathbf{x}_q = [\mathbf{x}_{0,q}^T, \dots, \mathbf{x}_{K-1,q}^T]^T$, $\mathbf{v}_q = [\mathbf{v}_{K_1,q}^T, \dots, \mathbf{v}_{K-K_1-1,q}^T]^T$. Specifically, the 2D interference matrix $\bar{\mathbf{G}}_l \in \mathcal{BC}_{K,N_s}$ is given by

$$\bar{\mathbf{G}}_l = \begin{bmatrix} \mathcal{D}(\bar{\mathbf{g}}_{-K_1,l}) & \cdots & \mathcal{D}(\bar{\mathbf{g}}_{K_1,l}) & \mathbf{0} \\ \mathbf{0} & \ddots & \ddots & \ddots \\ \mathcal{D}(\bar{\mathbf{g}}_{K_1,l}) & \mathbf{0} & \mathcal{D}(\bar{\mathbf{g}}_{-K_1,l}) & \cdots & \mathcal{D}(\bar{\mathbf{g}}_{K_1-1,l}) \\ \vdots & \ddots & \mathbf{0} & \ddots & \vdots \\ \mathcal{D}(\bar{\mathbf{g}}_{-K_1+1,l}) & \cdots & \mathcal{D}(\bar{\mathbf{g}}_{K_1,l}) & \mathcal{D}(\bar{\mathbf{g}}_{-K_1,l}) \end{bmatrix}, \quad (10)$$

where the subscripts of the all-zero matrices are omitted for simplicity. Moreover, $\Theta_q \in \mathcal{BC}_{K,N_s}$ has the similar expression of (10), where $\mathcal{D}(\bar{\mathbf{g}}_{k_\Delta,l})$ is replaced by $\mathcal{D}(\boldsymbol{\theta}_{k_\Delta,q})$ with $\boldsymbol{\theta}_{k_\Delta,q} = [e^{j2\pi k_\Delta q L_m \nu F_0 \tau T_0}, \dots, e^{j2\pi k_\Delta (qL_m + N_s - 1) \nu F_0 \tau T_0}]^T$.

Note that $\bar{\mathbf{G}}_l \in \mathcal{BC}_{K,N_s}$ consists of a series of diagonal matrices. According to the properties of block circulant matrices, $\bar{\mathbf{G}}_l$ satisfies the decomposition of

$$\bar{\mathbf{G}}_l = (\mathbf{F}_K \otimes \mathbf{I}_{N_s})^H \boldsymbol{\Lambda}_{\bar{\mathbf{G}}_l} (\mathbf{F}_K \otimes \mathbf{I}_{N_s}), \quad (11)$$

where $\boldsymbol{\Lambda}_{\bar{\mathbf{G}}_l}$ is a diagonal matrix and the derivations are given in **Appendix A**. According to the Theorem 5.8.1 of [21], we introduce another decomposition of the block circulant matrices for simplifying the received signal model of (9). Then, we obtain $\Phi_l = (\mathbf{F}_K \otimes \mathbf{F}_{N_s})^H \boldsymbol{\Lambda}_{\Phi_l} (\mathbf{F}_K \otimes \mathbf{F}_{N_s})$ and $\Theta_q = (\mathbf{F}_K \otimes \mathbf{F}_{N_s})^H \boldsymbol{\Lambda}_{\Theta_q} (\mathbf{F}_K \otimes \mathbf{F}_{N_s})$. After left multiplying \mathbf{r}_q in (9) by a unitary matrix $(\mathbf{F}_K \otimes \mathbf{F}_{N_s})$, we have the received symbol vector in the q -th segment as

$$\begin{aligned} \check{\mathbf{r}}_q &= (\mathbf{F}_K \otimes \mathbf{F}_{N_s}) \mathbf{r}_q \\ &= \sum_{l=0}^{L_h-1} h_{q,l} \boldsymbol{\Lambda}_{\Phi_l} \boldsymbol{\Lambda}_{\Theta_q} (\mathbf{F}_K \otimes \mathbf{F}_{N_s}) (\mathbf{F}_K \otimes \mathbf{F}_{N_s})^H \\ &\quad \times \boldsymbol{\Lambda}_{\bar{\mathbf{G}}_l} (\mathbf{F}_K \otimes \mathbf{F}_{N_s}) (\mathbf{I}_K \otimes \Phi_m) \mathbf{x}_q + \check{\mathbf{v}}_q \\ &= \sum_{l=0}^{L_h-1} h_{q,l} \boldsymbol{\Lambda}_{q,l} \ddot{\mathbf{F}}_{KN_s} \mathbf{x}_q + \check{\mathbf{v}}_q, \end{aligned} \quad (12)$$

where $\boldsymbol{\Lambda}_{q,l} = \boldsymbol{\Lambda}_{\Phi_l} \boldsymbol{\Lambda}_{\Theta_q} \boldsymbol{\Lambda}_{\bar{\mathbf{G}}_l}$, $\ddot{\mathbf{F}}_{KN_s} = (\mathbf{F}_K \otimes \mathbf{F}_{N_s}) (\mathbf{I}_K \otimes \Phi_m)$ denotes the $KN_s \times KN_s$ normalized 2D DFT matrix having inserted zeros, while $\check{\mathbf{v}}_q = (\mathbf{F}_K \otimes \mathbf{F}_{N_s}) \mathbf{v}_q$ represents the FD colored noise process obeying Gaussian distribution with mean vector \mathbf{z}_q^v and covariance matrix $\boldsymbol{\Lambda}_q^v$. The relevant derivations of the equivalent colored noise are shown in **Appendix B**. Note that the product of $(\mathbf{F}_K \otimes \mathbf{F}_{N_s})$ and any vector in (12) can be efficiently calculated via the 2D FFT with complexity $KN_s \log(KN_s)$. Hence, the proposed segment-based FD received signal model is equivalently obtained by a 2D FFT module, as shown in Fig. 1.

To emphasize the benefits of the proposed segment-based FD received signal model on the ill-conditioning problem of MFTN signaling in Fig. 2, we compare its condition number to that of the original received signal model of (3) in the case of various packing factor combinations. The simulation parameters are as follows. The truncated ISI and ICI intervals are $N_1 = 12$ and $K_1 = 1$, respectively. The length of each segment is $L_m = 42$, and we employ an RRC shaping pulse having a roll-off factor of $\beta = 0.3$. Considering a frequency-

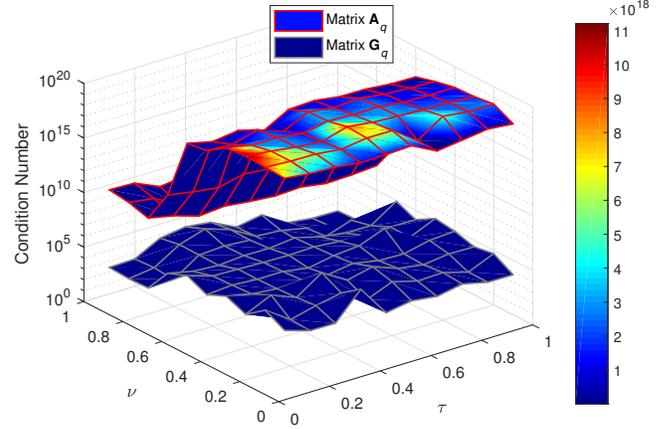


Fig. 2: Comparison of the condition number of linear transition matrices in the received signal models of Equation (2).

selective fading channel of $L_h = 8$ taps, the l -th power delay profile (PDP) satisfies $\sigma_{h_l}^2 = \exp(-0.1l) / (\sum_l \sigma_{h_l}^2)$.

The linear transition matrix of the proposed segment-based FD received signal model of (12) is defined as $\mathbf{G}_q = \sum_{l=0}^{L_h-1} h_{q,l} (\mathbf{F}_K \otimes \mathbf{F}_{N_s})^H \boldsymbol{\Lambda}_{q,l} (\mathbf{F}_K \otimes \mathbf{F}_{N_s})$ in the q -th segment. Accordingly, the linear transition matrix of the original received signal model in Section II is $\mathbf{A}_q = \sum_{l=0}^{L_h-1} h_{q,l} \dot{\mathbf{A}}_l$, where the (m,n) -th element of $\dot{\mathbf{A}}_l$ is $A_p([\text{rem}(n, N_s) - \text{rem}(m, N_s)]\tau T_0, ([n/N_s] - [m/N_s])\nu F_0)$ in (3). As shown in Fig. 2, for a fixed BE, we can obtain a linear transition matrix having a minimum condition number by jointly optimizing the combination of the packing factors. For the original model, the condition number of the linear transition matrix \mathbf{A}_q significantly increases with the reduction of the frequency packing factor ν . By contrast, the condition number of the proposed model varies gently with the packing factors. Given the values of packing factor, the condition number of \mathbf{G}_q is much smaller than that of \mathbf{A}_q , which is an explicit benefit of the block circulant structure associated with the circulant blocks in \mathbf{G}_q . Hence, the proposed segment-based FD received signal model is expected to significantly mitigate the inherent ill-conditioning problem of MFTN signaling, which will be verified by our simulations.

IV. FREQUENCY-DOMAIN JOINT CHANNEL ESTIMATION AND EQUALIZATION ALGORITHMS

Building on the above-mentioned segment-based FD received signal model, we can rewrite the received symbol vector in (12) as

$$\check{\mathbf{r}}_q = \check{\mathbf{z}}_q + \mathbf{w}_q = \sum_{l=0}^{L_h-1} h_{q,l} \boldsymbol{\Lambda}_{q,l} \ddot{\mathbf{F}}_{KN_s} \mathbf{x}_q + \mathbf{z}_q^v + \mathbf{w}_q, \quad (13)$$

where $\check{\mathbf{z}}_q = \sum_{l=0}^{L_h-1} h_{q,l} \boldsymbol{\Lambda}_{q,l} \ddot{\mathbf{F}}_{KN_s} \mathbf{x}_q + \mathbf{z}_q^v$ denotes the noiseless observation vector, \mathbf{z}_q^v is the mean vector of the colored noise process $\check{\mathbf{v}}_q$, while \mathbf{w}_q is the equivalent colored noise process with zero mean and the covariance matrix of $\boldsymbol{\Lambda}_q^v$.

Assuming that the likelihood function is component-wise independent, the *a posteriori* probability of the channel coef-

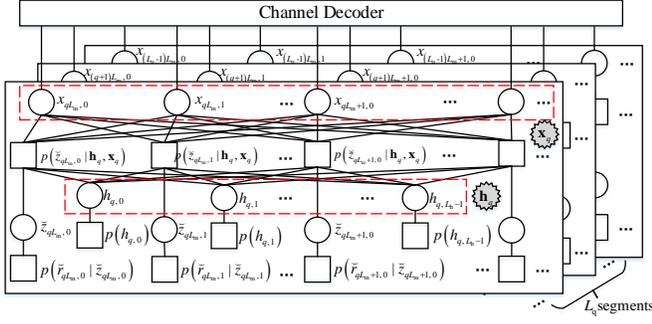


Fig. 3: Factor graph representation of the segment-based FD received signal model.

ficients and the transmitted symbols can be factorized as

$$\begin{aligned} p(\mathbf{x}_q, \mathbf{h}_q | \check{\mathbf{r}}_q) &\propto p(\check{\mathbf{r}}_q | \check{\mathbf{z}}_q) p(\check{\mathbf{z}}_q | \mathbf{h}_q, \mathbf{x}_q) p(\mathbf{h}_q) p(\mathbf{x}_q) \\ &\propto \prod_{k,n} p(\check{r}_{k,n} | \check{z}_{k,n}) p(\check{z}_{k,n} | \mathbf{h}_q, \mathbf{x}_q) \\ &\quad \times \prod_l p(h_{q,l}) \prod_{k',n'} p(x_{k',n'}), \end{aligned} \quad (14)$$

where $p(\check{\mathbf{r}}_q | \check{\mathbf{z}}_q)$ is the likelihood function, $p(h_{q,l})$ is the *a priori* probability of the l -th channel coefficient in the q -th segment, and $p(x_{k',n'})$ is the *a priori* probability of the n' -th TD transmitted symbol of the k' -th subcarrier, $k' = 0, \dots, K-1, n' = qL_m, \dots, (q+1)L_m - 1$.

According to the factorization of the *a posteriori* probability in (14), we construct a factor graph to characterize the statistical relationships between variables shown in Fig. 3. Since the factor graph contains dense short loops, the typical loopy belief propagation (LBP) over the factor graph leads to an excessive computational complexity. Although Gaussian BP can be used for reducing the complexity of the LBP, the Gaussian approximations of the discrete *a priori* messages of the transmitted symbols may result in an additional performance loss. Our goal is to simultaneously estimate the channel coefficients and the transmitted symbols from the noisy received observations. This is equivalent to solving the high-dimensional random bilinear mixing recovery problem. Hence, we can resort to the computationally efficient PBiGAMP algorithm for decoupling the bilinear estimation problem into a series of scalar computations, where all messages can be characterized and updated by their mean and variance vectors.

A. PBiGAMP-Based Frequency-Domain JCEE Algorithm

The *a posteriori* probability of the noiseless measurement $\check{z}_{k,n}$ is proportional to the product of the message propagated from the factor node (FN) $p(\check{z}_{k,n} | \mathbf{h}_q, \mathbf{x}_q)$ to the variable node (VN) $\check{z}_{k,n}$ and the conditional probability density function $p(\check{r}_{k,n} | \check{z}_{k,n})$. Hence, the former can be regarded as the pseudo *a priori* probability of the noiseless measurement $\check{z}_{k,n}$. As shown in Fig. 3, this message is calculated by integrating the product of the messages propagated from all VNs \mathbf{x}_q and \mathbf{h}_q to the FN $p(\check{z}_{k,n} | \mathbf{h}_q, \mathbf{x}_q)$ and the function $p(\check{z}_{k,n} | \mathbf{h}_q, \mathbf{x}_q)$ over all the VNs in \mathbf{x}_q and \mathbf{h}_q . Applying central-limit-theorem (CLT) to the messages gleaned from the FNs $p(\check{\mathbf{z}}_q | \mathbf{h}_q, \mathbf{x}_q)$ to the VNs $\check{\mathbf{z}}_q$, we can approximate the pseudo *a priori* probabilities of the noiseless measurements $\check{\mathbf{z}}_q$ by independent Gaussian

distributions having the mean vector of $\hat{\mathbf{p}}_q(t)$ and the variance vector of $\boldsymbol{\nu}_q^p(t)$ [22]. The pseudo *a priori* variance vector $\boldsymbol{\nu}_q^p(t)$ of the noiseless measurements $\check{\mathbf{z}}_q$ in the t -th iteration is derived as

$$\boldsymbol{\nu}_q^p(t) = \bar{\boldsymbol{\nu}}_q^p(t) + \sum_{l=0}^{L_h-1} \nu_{q,l}^h(t) |\boldsymbol{\Lambda}_{q,l} \ddot{\mathbf{F}}_{KN_s}|^2 \boldsymbol{\nu}_q^x(t), \quad (15)$$

with

$$\begin{aligned} \bar{\boldsymbol{\nu}}_q^p(t) &= \sum_{l=0}^{L_h-1} \nu_{q,l}^h(t) |\boldsymbol{\Lambda}_{q,l} \ddot{\mathbf{F}}_{KN_s} \hat{\mathbf{x}}_q(t)|^2 \\ &\quad + \left| \sum_{l=0}^{L_h-1} \hat{h}_{q,l}(t) \boldsymbol{\Lambda}_{q,l} \ddot{\mathbf{F}}_{KN_s} \right|^2 \boldsymbol{\nu}_q^x(t), \end{aligned} \quad (16)$$

where $\hat{h}_{q,l}(t)$ and $\nu_{q,l}^h(t)$ are the *a posteriori* mean and variance of the l -th channel estimate in the q -th segment.

Note that the noiseless measurement contains the colored noise process of (13) having non-zero mean. To improve the accuracy of the pseudo *a priori* mean vector of the noiseless measurement, we can introduce two additive terms characterizing the 2D interferences arising from the adjacent segments for improving its updating expression. Accordingly, the pseudo *a priori* mean of $\check{\mathbf{z}}_q$ is given by

$$\begin{aligned} \hat{\mathbf{p}}_q(t) &= \sum_{l=0}^{L_h-1} \hat{h}_{q,l}(t) \boldsymbol{\Lambda}_{q,l} \ddot{\mathbf{F}}_{KN_s} \hat{\mathbf{x}}_q(t) - \bar{\boldsymbol{\nu}}_q^p(t) \odot \hat{\mathbf{s}}_q(t-1) \\ &\quad + (\mathbf{F}_K \otimes \mathbf{F}_{N_s}) [(\mathbf{I}_K \otimes \boldsymbol{\Xi}_t) \hat{\mathbf{z}}_{q-1}(t) + (\mathbf{I}_K \otimes \boldsymbol{\Xi}_h) \hat{\mathbf{z}}_{q+1}(t)], \end{aligned} \quad (17)$$

where

$$\hat{\mathbf{z}}_q(t) = (\mathbf{F}_K \otimes \mathbf{F}_{N_s})^H \sum_{l=0}^{L_h-1} \hat{h}_{q,l}(t) \boldsymbol{\Lambda}_{q,l} \ddot{\mathbf{F}}_{KN_s} \hat{\mathbf{x}}_q(t). \quad (18)$$

By taking the product of the likelihood function $p(\check{\mathbf{r}}_q | \check{\mathbf{z}}_q) \propto \mathcal{CN}[\check{\mathbf{z}}_q; \check{\mathbf{r}}_q, \boldsymbol{\Lambda}_q^v(t)]$ and the pseudo *a priori* message $\mathcal{CN}[\check{\mathbf{z}}_q; \hat{\mathbf{p}}_q(t), \mathcal{D}(\boldsymbol{\nu}_q^p(t))]$, the *a posteriori* probability of $\check{\mathbf{z}}_q$ is modelled by the Gaussian distribution as $p(\check{\mathbf{z}}_q | \check{\mathbf{r}}_q) \propto \mathcal{CN}[\check{\mathbf{z}}_q; \hat{\mathbf{z}}_q(t), \mathcal{D}(\boldsymbol{\nu}_q^z(t))]$ with

$$\hat{\mathbf{z}}_q(t) = [\check{\mathbf{z}}_q \odot \boldsymbol{\nu}_q^p(t) + \boldsymbol{\Lambda}_q^v(t) \hat{\mathbf{p}}_q(t)] \odot [\boldsymbol{\nu}_q^p(t) + \boldsymbol{\Lambda}_q^v(t) \mathbf{1}_{KN_s}], \quad (19)$$

$$\boldsymbol{\nu}_q^z(t) = (\boldsymbol{\Lambda}_q^v(t) \boldsymbol{\nu}_q^p(t)) \odot (\boldsymbol{\nu}_q^p(t) + \boldsymbol{\Lambda}_q^v(t) \mathbf{1}_{KN_s}). \quad (20)$$

Note that the covariance matrix $\boldsymbol{\Lambda}_q^v(t)$ of the equivalent colored noise depends not only on the covariance matrix of the 2D interferences arising from the adjacent segments but also on the variance of the AWGN process, which is updated according to the derivations in **Appendix B**.

We can update the first derivative $\hat{\mathbf{s}}_q$ and second derivative $\boldsymbol{\nu}_q^s$ based on a Taylor series expansion of the messages passed from the likelihood function nodes passed to the channel VNs or the transmitted symbol VNs as

$$\begin{aligned} \hat{\mathbf{s}}_q(t) &= [\hat{\mathbf{z}}_q(t) - \hat{\mathbf{p}}_q(t)] \odot \boldsymbol{\nu}_q^p(t) \\ &= [\check{\mathbf{r}}_q - \hat{\mathbf{p}}_q(t)] \odot [\boldsymbol{\nu}_q^p(t) + \boldsymbol{\Lambda}_q^v(t) \mathbf{1}_{KN_s}], \end{aligned} \quad (21)$$

$$\begin{aligned} \boldsymbol{\nu}_q^s(t) &= [\boldsymbol{\nu}_q^p(t) - \boldsymbol{\nu}_q^z(t)] \odot [\boldsymbol{\nu}_q^p(t)]^2 \\ &= \mathbf{1}_{KN_s} \odot [\boldsymbol{\nu}_q^p(t) + \boldsymbol{\Lambda}_q^v(t) \mathbf{1}_{KN_s}]. \end{aligned} \quad (22)$$

We introduce γ_q to represent the product of all outgoing messages from the likelihood function nodes to the transmitted symbol VNs. The vector γ_q can be regarded as the noisy observation of the transmitted symbol \mathbf{x}_q in an AWGN channel having the noise variance of ν_q^γ . The marginal *a posteriori* probability of the unknown transmitted symbols is given by

$$p(\mathbf{x}_q|\check{\mathbf{r}}_q) = \frac{p(\mathbf{x}_q)\mathcal{CN}[\mathbf{x}_q; \hat{\gamma}_q(t), \mathcal{D}(\nu_q^\gamma(t))]}{\int_{\mathbf{x}_q} p(\mathbf{x}_q)\mathcal{CN}[\mathbf{x}_q; \hat{\gamma}_q(t), \mathcal{D}(\nu_q^\gamma(t))]}, \quad (23)$$

where the outgoing mean vector $\hat{\gamma}_q(t)$ and variance vector $\nu_q^\gamma(t)$ are updated as

$$\nu_q^\gamma(t) = \mathbf{1}_{KL_m} \odot \left[|\ddot{\mathbf{F}}_{KN_s}^H \left(\sum_{l=0}^{L_h-1} \hat{h}_{q,l}(t) \mathbf{\Lambda}_{q,l} \right)^H|^2 \nu_q^s(t) \right], \quad (24)$$

$$\hat{\gamma}_q(t) = \hat{\mathbf{x}}_q(t) + \nu_q^\gamma(t) \odot \left[\sum_{l=0}^{L-1} \hat{h}_{q,l}(t) \mathbf{\Lambda}_{q,l} \ddot{\mathbf{F}}_{KN_s} \right] \hat{\mathbf{s}}_q(t) \quad (25)$$

$$- \nu_q^\gamma(t) \odot \hat{\mathbf{x}}_q(t) \odot \sum_{l=0}^{L-1} \nu_{q,l}^h(t) |\ddot{\mathbf{F}}_{KN_s}^H \mathbf{\Lambda}_{q,l}|^2 \nu_q^s(t).$$

The *a priori* messages of the transmitted symbols are calculated based on soft extrinsic information gleaned from the channel decoder. Let us assume that the \bar{m} -th extrinsic log-likelihood ratio (LLR) of $x_{k,n}$ acquired from the channel decoder output is $L_{\text{dec}}^{e,(t)}(c_{k,n,\bar{m}})$, the *a priori* probability of the coded bit $c_{k,n,\bar{m}} \in \{0, 1\}$, $m = 0, \dots, \log_2 M - 1$ in the t -th iteration is $p(c_{k,n,\bar{m}}) = \frac{1}{2} [1 + (-1)^{c_{k,n,\bar{m}}} \tanh(\frac{1}{2} L_{\text{dec}}^{e,(t)}(c_{k,n,\bar{m}}))]$. Since each transmitted symbol belongs to the constellation set $\mathcal{S} \triangleq \{\mathcal{S}_1, \dots, \mathcal{S}_M\}$, we can obtain the discrete *a priori* probabilities $\{p(x_{k,n})\}$ according to the constellation mapping rules. Hence, the conditional *a posteriori* probability of each transmitted symbol $x_{k,n}$ in the q -th segment is $p(x_{k,n}|\check{\mathbf{r}}_q) \propto p(x_{k,n})\mathcal{CN}[x_{k,n}; \hat{\gamma}_{k,n}(t), \nu_q^\gamma(t)]$. The corresponding conditional *a posteriori* expectation and variance are given by

$$\hat{x}_{k,n}(t+1) = \mathbb{E}\{x_{k,n}|\gamma_{k,n} = \hat{\gamma}_{k,n}(t); \nu_q^\gamma(t)\} \quad (26)$$

$$\propto \sum_m \mathcal{S}_m p(x_{k,n} = \mathcal{S}_m) \exp\left(-\frac{|\mathcal{S}_m - \hat{\gamma}_{k,n}(t)|^2}{\nu_q^\gamma(t)}\right),$$

$$\nu_{k,n}^x(t+1) = \mathbb{V}\{x_{k,n}|\gamma_{k,n} = \hat{\gamma}_{k,n}(t); \nu_q^\gamma(t)\} \quad (27)$$

$$\propto \sum_m |\hat{x}_{k,n}(t+1) - \mathcal{S}_m|^2 p(x_{k,n} = \mathcal{S}_m) \times \exp\left(-\frac{|\mathcal{S}_m - \hat{\gamma}_{k,n}(t)|^2}{\nu_q^\gamma(t)}\right).$$

Note that the *a posteriori* variance of $\nu_{k,n}^x(t+1)$ is used for characterizing the estimation accuracy of the transmitted symbol in the t -th iteration and then for updating the “pseudo” *a priori* variance of the noiseless measurement in the $(t+1)$ -st iteration.

Similar to the above-mentioned characterization of the symbol detection, the outgoing messages of the unknown channel coefficients are approximated by the Gaussian distribution $\mathcal{CN}(\hat{\mathbf{h}}_q; \hat{\mathbf{q}}_q, \nu_q^e)$ based on the CLT. Hence, the l -th element of the outgoing variance vector ν_q^e and mean vector $\hat{\mathbf{q}}_q$ are

expressed respectively as

$$1/\nu_{q,l}^e(t) = |(\mathbf{\Lambda}_{q,l} \ddot{\mathbf{F}}_{KN_s} \hat{\mathbf{x}}_q(t))^H|^2 \nu_q^s(t), \quad (28)$$

$$\hat{q}_{q,l}(t) = \hat{h}_{q,l}(t) + \nu_{q,l}^e(t) [\mathbf{\Lambda}_{q,l} \ddot{\mathbf{F}}_{KN_s} \hat{\mathbf{x}}_q(t)]^H \hat{\mathbf{s}}_q(t) \quad (29)$$

$$\times \hat{h}_{q,l}(t) \nu_{q,l}^e(t) \nu_q^x(t) \mathbf{1}_{KN_s}^T |\ddot{\mathbf{F}}_{KN_s}^H \mathbf{\Lambda}_{q,l}|^2 \nu_q^s(t).$$

Under the assumption of the Gaussian *a priori* probability, the *a posteriori* distribution of $h_{q,l}$ is given by $p(h_{q,l}|\check{\mathbf{z}}_q) \propto \mathcal{CN}[h_{q,l}; \hat{h}_{q,l}(t+1), \nu_{q,l}^h(t+1)]$ with

$$\hat{h}_{q,l}(t+1) = [\hat{q}_{q,l}(t) \bar{\nu}_{q,l}^h + \bar{h}_{q,l} \nu_{q,l}^e(t)] / [\bar{\nu}_{q,l}^h + \nu_{q,l}^e(t)], \quad (30)$$

$$\nu_{q,l}^h(t+1) = \bar{\nu}_{q,l}^h \nu_{q,l}^e(t) / [\bar{\nu}_{q,l}^h + \nu_{q,l}^e(t)], \quad (31)$$

where the *a priori* probability $\mathcal{CN}(h_{q,l}; \bar{h}_{q,l}, \bar{\nu}_{q,l}^h)$ can be obtained by pilot-based least-squared estimation (LSE) [23]. Similarly, the *a posteriori* variance $\nu_{q,l}^h(t+1)$ is also utilized for quantifying the accuracy of the channel estimate and then for updating the “pseudo” *a priori* variance of the noiseless measurement in the next iteration.

According to the aforementioned derivations, the PBiGAMP-based JCEE algorithm proposed for MFTN signaling is summarized in **Algorithm 1**.

Algorithm 1 The Proposed PBiGAMP-based FD JCEE Algorithm (PBiGAMP-FDJCEE)

- 1: **Initialization:** Set the *a posteriori* mean and variance of the transmitted symbols to be $\hat{x}_{k,n}(1) = 0$ and $\nu_{k,n}^x(1) = 1$, $k = 0, \dots, K-1$, $n = 0, \dots, N-1$. The initialized channel coefficients $\hat{\mathbf{h}}_q(1)$ and $\nu_q^h(1)$ are obtained by a pilot-based LSE. Set $\hat{\mathbf{s}}_q(0) = \mathbf{0}_{N_s}$ for $q = 0, \dots, L_q - 1$ as described in [17].
 - 2: **for** $t = 1$ to T_{ex} **do**
 - 3: **for** $q = 0$ to $L_q - 1$ **do**
 - 4: The TD received symbols are transformed to FD via 2D FFT in (13).
 - 5: Compute $\nu_q^p(t)$ and $\hat{\mathbf{p}}_q(t)$ using (15) and (17).
 - 6: Compute $\hat{\mathbf{s}}_q(t)$ and $\nu_q^s(t)$ using (21) and (22).
 - 7: Compute the outgoing messages $\nu_q^\gamma(t)$ and $\hat{\gamma}_q(t)$ of the transmitted symbols using (24) and (25).
 - 8: Compute each element of the *a posteriori* messages $\hat{\mathbf{x}}_q(t+1)$ and $\nu_q^x(t+1)$ of the transmitted symbols using (26) and (27).
 - 9: Compute each element of the outgoing messages $\nu_q^e(t)$ and $\hat{\mathbf{q}}_q(t)$ of the channel coefficients using (28) and (29).
 - 10: Compute each element of the *a posteriori* messages $\hat{\mathbf{h}}_q(t+1)$ and $\nu_q^h(t+1)$ of the channel coefficients using (30) and (31).
 - 11: **end for**
 - 12: Compute the extrinsic LLRs of the equalizer based on the outgoing messages of the transmitted symbols and then feed them to the channel decoder.
 - 13: Perform BCJR channel decoding and feed the soft extrinsic information to the equalizer.
 - 14: **end for**
-

B. Refined PBiGAMP-Based Frequency-Domain JCEE Algorithm

According to the derivations of the PBiGAMP-FDJCEE algorithm proposed in Section IV.A, the equations (15), (16), (24), (25), (28), and (29) contain the squared modulus of the FD linear transition matrix, which may increase the condition numbers. To solve this problem, we introduce a series of scaled identity matrices for approximating the variance vectors of the transmitted symbols in (27) and the second derivative term in (22). Moreover, the scaled identity matrix corresponding to $\boldsymbol{\nu}_q^s(t)$ is also harnessed for improving the accuracy of channel estimation. The detailed derivations are as follows.

The pseudo *a priori* variance of the noiseless measurement in (15) can be rewritten as

$$\begin{aligned} \boldsymbol{\nu}_q^p(t) &= \bar{\boldsymbol{\nu}}_q^p(t) + \sum_{l=0}^{L_h-1} \nu_{q,l}^h(t) |\boldsymbol{\Lambda}_{q,l} \ddot{\mathbf{F}}_{KN_s}|^2 \boldsymbol{\nu}_q^x(t) \\ &= \bar{\boldsymbol{\nu}}_q^p(t) + \sum_{l=0}^{L_h-1} \nu_{q,l}^h(t) \mathcal{D}[\boldsymbol{\Lambda}_{q,l} \ddot{\mathbf{F}}_{KN_s} \mathcal{D}(\boldsymbol{\nu}_q^x(t)) \ddot{\mathbf{F}}_{KN_s}^H \boldsymbol{\Lambda}_{q,l}^H] \mathbf{1}_{KN_s}, \end{aligned} \quad (32)$$

where we have $\ddot{\mathbf{F}}_{KN_s} = (\mathbf{F}_K \otimes \mathbf{F}_{N_s})(\mathbf{I}_K \otimes \Phi_m)$. To simplify the second term in the above equation, we can introduce a scaled identity matrix $\nu_q^x(t) \mathbf{I}_{KN_s}$ as the average approximation of the term $(\mathbf{I}_K \otimes \Phi_m) \mathcal{D}(\boldsymbol{\nu}_q^x(t)) (\mathbf{I}_K \otimes \Phi_m)^H$, where $\nu_q^x(t) = 1/(KN_s) \mathbf{1}_{KN_s}^T (\mathbf{I}_K \otimes \Phi_m) \mathcal{D}(\boldsymbol{\nu}_q^x(t)) (\mathbf{I}_K \otimes \Phi_m)^H \mathbf{1}_{KN_s}$ is the average variance of the transmitted symbols in the q -th segment. Then, the pseudo *a priori* variance vector of the noiseless measurement vector is expressed as

$$\begin{aligned} \boldsymbol{\nu}_q^p(t) &\approx \bar{\boldsymbol{\nu}}_q^p(t) + \sum_{l=0}^{L_h-1} \nu_{q,l}^h(t) \mathcal{D}[\boldsymbol{\Lambda}_{q,l} (\mathbf{F}_K \otimes \mathbf{F}_{N_s}) \\ &\quad \times \nu_q^x(t) \mathbf{I}_{KN_s} (\mathbf{F}_K \otimes \mathbf{F}_{N_s})^H \boldsymbol{\Lambda}_{q,l}^H] \mathbf{1}_{KN_s} \\ &= \bar{\boldsymbol{\nu}}_q^p(t) + \nu_q^x(t) \sum_{l=0}^{L_h-1} \nu_{q,l}^h(t) \boldsymbol{\Lambda}_{q,l} \boldsymbol{\Lambda}_{q,l}^H \mathbf{1}_{KN_s}. \end{aligned} \quad (33)$$

Moreover, the second term of $\bar{\boldsymbol{\nu}}_q^p(t)$ in (16) can also be simplified using the scaled identity matrix $\nu_q^x(t) \mathbf{I}_{KN_s}$. Accordingly, (16) is approximated as

$$\begin{aligned} \bar{\boldsymbol{\nu}}_q^p(t) &\approx \sum_{l=0}^{L_h-1} \nu_{q,l}^h(t) \boldsymbol{\Lambda}_{q,l} \boldsymbol{\Lambda}_{q,l}^H |\ddot{\mathbf{F}}_{KN_s} \hat{\mathbf{x}}_q(t)|^2 \\ &\quad + \nu_q^x(t) \left(\sum_{l=0}^{L_h-1} \hat{h}_{q,l}(t) \boldsymbol{\Lambda}_{q,l} \right) \left(\sum_{l'=0}^{L_h-1} \hat{h}_{q,l'}(t) \boldsymbol{\Lambda}_{q,l'} \right)^H \mathbf{1}_{KN_s}. \end{aligned} \quad (34)$$

In (24), we can employ a scaled identity matrix $\varsigma_q^\gamma(t) \mathbf{I}_{KN_s}$ for replacing the middle term between the 2D Fourier operations. Then the outgoing variance vector of the transmitted symbols can be calculated as

$$\begin{aligned} \boldsymbol{\nu}_q^\gamma(t) &= \mathbf{1}_{KL_m} \otimes \mathcal{D} \left(\ddot{\mathbf{F}}_{KN_s}^H \left(\sum_{l=0}^{L_h-1} \hat{h}_{q,l}(t) \boldsymbol{\Lambda}_{q,l} \right)^H \right. \\ &\quad \left. \times \mathcal{D}(\boldsymbol{\nu}_q^s(t)) \left(\sum_{l'=0}^{L_h-1} \hat{h}_{q,l'}(t) \boldsymbol{\Lambda}_{q,l'} \right) \ddot{\mathbf{F}}_{KN_s} \right) \\ &\approx \mathbf{1}_{KL_m} \otimes \mathcal{D}(\ddot{\mathbf{F}}_{KN_s}^H \varsigma_q^r(t) \mathbf{I}_{KN_s} \ddot{\mathbf{F}}_{KN_s}) = 1/\varsigma_q^\gamma(t) \mathbf{1}_{KL_m}, \end{aligned} \quad (35)$$

where the scaled value is given by

$$\varsigma_q^\gamma(t) = \frac{1}{KN_s} \mathbf{1}_{KN_s}^H \left(\sum_{l=0}^{L_h-1} \hat{h}_{q,l}(t) \boldsymbol{\Lambda}_{q,l} \right) \left(\sum_{l'=0}^{L_h-1} \hat{h}_{q,l'}(t) \boldsymbol{\Lambda}_{q,l'} \right) \boldsymbol{\nu}_q^s(t). \quad (36)$$

Note that all outgoing variances in (35) are identical. Hence, we only compute a scalar outgoing variance $\nu_q^\gamma(t) = \varsigma_q^\gamma(t)$.

Furthermore, the corresponding outgoing mean vector $\hat{\gamma}_q(t)$ in (25) is approximated as

$$\begin{aligned} \hat{\gamma}_q(t) &\approx \hat{\mathbf{x}}_q(t) + \nu_q^\gamma(t) \left[\sum_{l=0}^{L_h-1} \hat{h}_{q,l}(t) \boldsymbol{\Lambda}_{q,l} \ddot{\mathbf{F}}_{KN_s} \right]^H \hat{\mathbf{s}}_q(t) \\ &\quad - \nu_q^\gamma(t) \hat{\mathbf{x}}_q(t) \odot \sum_{l=0}^{L_h-1} \nu_{q,l}^h(t) |\ddot{\mathbf{F}}_{KN_s}^H \boldsymbol{\Lambda}_{q,l}^H|^2 \boldsymbol{\nu}_q^s(t) \\ &= \hat{\mathbf{x}}_q(t) + \nu_q^\gamma(t) \left[\sum_{l=0}^{L_h-1} \hat{h}_{q,l}(t) \boldsymbol{\Lambda}_{q,l} \ddot{\mathbf{F}}_{KN_s} \right]^H \hat{\mathbf{s}}_q(t) - \nu_q^\gamma(t) \\ &\quad \times \hat{\mathbf{x}}_q(t) \odot \sum_{l=0}^{L_h-1} \nu_{q,l}^h(t) \mathcal{D}[\ddot{\mathbf{F}}_{KN_s}^H \boldsymbol{\Lambda}_{q,l}^H \mathcal{D}(\boldsymbol{\nu}_q^s(t)) \boldsymbol{\Lambda}_{q,l} \ddot{\mathbf{F}}_{KN_s}] \\ &\approx \hat{\mathbf{x}}_q(t) + \nu_q^\gamma(t) \left[\sum_{l=0}^{L_h-1} \hat{h}_{q,l}(t) \boldsymbol{\Lambda}_{q,l} \ddot{\mathbf{F}}_{KN_s} \right]^H \hat{\mathbf{s}}_q(t) \\ &\quad - \nu_q^r(t) \sum_{l=0}^{L_h-1} \nu_{q,l}^h(t) \varsigma_{q,l}^s(t) \hat{\mathbf{x}}_q(t), \end{aligned} \quad (37)$$

where the scaled identity matrix $\varsigma_{q,l}^s(t) \mathbf{I}_{KN_s}$ is introduced to replace $\boldsymbol{\Lambda}_{q,l}^H \mathcal{D}(\boldsymbol{\nu}_q^s(t)) \boldsymbol{\Lambda}_{q,l}$ by

$$\varsigma_{q,l}^s(t) = \frac{1}{KN_s} \mathbf{1}_{KN_s}^T \boldsymbol{\Lambda}_{q,l}^H \boldsymbol{\Lambda}_{q,l} \boldsymbol{\nu}_q^s(t). \quad (38)$$

As for channel estimation, by introducing the scaled identity matrix $\varsigma_{q,l}^s(t) \mathbf{I}_{KN_s}$, the outgoing variance of the l -th unknown channel coefficient in (28) can be simplified as

$$\begin{aligned} 1/\nu_{q,l}^o(t) &= (\hat{\mathbf{x}}_q(t))^H \ddot{\mathbf{F}}_{KN_s}^H \boldsymbol{\Lambda}_{q,l}^H \mathcal{D}(\boldsymbol{\nu}_q^s(t)) \boldsymbol{\Lambda}_{q,l} \ddot{\mathbf{F}}_{KN_s} \hat{\mathbf{x}}_q(t) \\ &\approx \varsigma_{q,l}^s(t) \|\hat{\mathbf{x}}_q(t)\|^2. \end{aligned} \quad (39)$$

Moreover, the outgoing mean of the l -th channel coefficient in (29) is approximated as

$$\begin{aligned} \hat{q}_{q,l}(t) &\approx \hat{h}_{q,l}(t) + \nu_{q,l}^o(t) (\boldsymbol{\Lambda}_{q,l} \ddot{\mathbf{F}}_{KN_s} \hat{\mathbf{x}}_q(t))^H \hat{\mathbf{s}}_q(t) \\ &\quad - KN_s \hat{h}_{q,l}(t) \nu_{q,l}^o(t) \nu_q^x(t) \varsigma_{q,l}^s(t). \end{aligned} \quad (40)$$

According to the above-mentioned approximations, the proposed refined PBiGAMP-based JCEE procedure of MFTN signaling is summarized in **Algorithm 2**.

C. GAMP-Based FDE Algorithms

In the special case of perfectly known CSI, the channel coefficients no longer have to be estimated. Hence, we can replace the channel estimates by the exact channel coefficients. The proposed PBiGAMP-based receivers are readily simplified to the GAMP-FDE and R-GAMP-FDE algorithms, which are summarized in **Algorithm 3** and **Algorithm 4**.

Algorithm 2 The Refined PBiGAMP-based FD JCEE Algorithm (R-PBiGAMP-FDJCEE)

- 1: Run the same initialization and calculation steps in **Algorithm 1** except for the following three modifications:
 - 2: **for** $t = 1$ to T_{ex} **do**
 - 3: **for** $q = 0$ to $L_q - 1$ **do**
 - 4: Compute the *a priori* variance vector of the noiseless measurements $\nu_q^p(t)$ using (33) and (34).
 - 5: Compute the outgoing messages $\nu_q^\gamma(t)$ and $\hat{\gamma}_q(t)$ of the transmitted symbols using (35)-(37).
 - 6: Compute each element of the outgoing messages $\nu_q^o(t)$ and $\hat{\rho}_q(t)$ of the channel coefficients using (39) and (40).
 - 7: **end for**
 - 8: **end for**
-

Algorithm 3 The GAMP-Based FDE Algorithm (GAMP-FDE) For Perfectly Known CSI

- 1: **Initialization:** Set the *a posteriori* mean and variance of the transmitted symbols to be $\hat{x}_{k,n}(1) = 0$ and $\nu_{k,n}^x(1) = 1$, $k = 0, \dots, K - 1$, $n = 0, \dots, N - 1$. Set $\hat{\mathbf{s}}_q(0) = \mathbf{0}_{N_s}$, $\hat{h}_{q,l}(t) = h_{q,l}$, and $\nu_{q,l}^h(t) = 0$ for $q = 0, \dots, L_q - 1$, $l = 0, \dots, L_h - 1$.
 - 2: Run Step 2-14 in **Algorithm 1** except Step 9 and 10.
-

D. Complexity Analysis

In this section, we analyze the computational complexity of the proposed PBiGAMP-based algorithms and compare them to that of the existing algorithms, as summarized in TABLE III. Since all methods perform standard soft information calculations and BCJR decoding, we only discuss the computational complexity of the equalizer and channel estimator in a single iteration. Moreover, the complexity of FDE algorithms is dominated by the number of complex multiplications and 2D FFT/IFFT operations, which are analyzed independently in TABLE III.

Under the assumption of perfectly known CSI, we compare the proposed GAMP-based algorithms to three typical TDE algorithms: MMSE-TDE, GAMP-TDE, and AMP with unitary transformation (UTAMP)-TDE. The MMSE-TDE algorithm of [24] needs to calculate the inversion of the equivalent channel matrix, resulting in a complexity order of $\mathcal{O}(K^3N^3)$. As described in [25], the worst-case total complexity of GAMP per iteration is $\mathcal{O}(K^2N^2)$ when the size of the equivalent channel matrix is $KN \times KN$. As a modified version of AMP, UTAMP exploits singular value decomposition (SVD) of the $KN \times KN$ equivalent channel matrix for improving the convergence of AMP [26]. The complexity of SVD is $\mathcal{O}(K^3N^3)$, while the complexity of message passing relying on the non-structured equivalent channel matrix is $\mathcal{O}(K^2N^2)$. For GAMP-FDE, nine KN_s -point 2D FFT/IFFT operations are needed for the calculations of (12), (16)-(18), (24), and (25). Note that $\tilde{\mathbf{F}}_{KN_s} \hat{\mathbf{x}}_q(t)$ in (17) and (18) only has to be calculated once. For its refined version, four KN_s -point 2D FFT/IFFT operations imposed by (16) and (24) are omitted by introducing a series of scaled identity matrices in (34)

Algorithm 4 The Refined GAMP-Based FDE Algorithm (R-GAMP-FDE) For Perfectly Known CSI

- 1: Run the same initialization and calculation steps in **Algorithm 3** except for the following two modifications:
 - 2: **for** $q = 0$ to $L_q - 1$ **do**
 - 3: **for** $t = 1$ to T_{ex} **do**
 - 4: Compute the *a priori* variance vector of the noiseless measurements $\nu_q^p(t)$ using (33) and (34).
 - 5: Compute the outgoing messages $\nu_q^\gamma(t)$ and $\hat{\gamma}_q(t)$ using (35)-(37).
 - 6: **end for**
 - 7: **end for**
-

and (35). Hence, five KN_s -point 2D FFT/IFFT operations are required for message updating of R-GAMP-FDE per segment. The remaining complexity mainly depends on the complex multiplications of a diagonal matrix and a vector or of vectors, resulting in a complexity order of $\mathcal{O}(KN_s)$ per segment.

For MFTN systems operating in the absence of CSI, we extend both the LS-based channel estimator of [23] and the MMSE-based equalizer of [24] labeled as ‘‘LS-MMSE’’, by conceiving the proposed segment-based FD received signal model for MFTN systems. Since the interferences of MFTN signaling only depend on the shaping pulse and packing factors, the weighted matrix of the LS-based channel estimator can be calculated off-line and stored in advance. This substantially reduces the computational complexity in practice. Hence, we only consider the complexity associated with the complex multiplication of the weighted matrix and the received observations. The complexity order of LS-based channel estimator is $\mathcal{O}(L_h N_p)$ per segment, with N_p being the number of pilots, while that of MMSE-based equalizer is $\mathcal{O}(K^3 N_s^3)$ per segment dominated by the matrix inversion of the non-diagonal covariance matrix of the FD transmitted symbols. For PBiGAMP-FDJCEE, thirteen KN_s -point 2D FFT/IFFT operations are needed for the calculations of (12), (15)-(18), (24), and (25) for equalization. Note that (28) and (29) characterizing the channel estimation can utilize the outputs of the corresponding FFT/IFFT operations the equalizer, without any additional complexity. For R-PBiGAMP-FDJCEE, a set of eight KN_s -point 2D FFT/IFFT operations imposed by (15)-(16) and (24)-(25) are omitted due to the simplifications in (32)-(38). Hence, only five KN_s -point 2D FFT/IFFT operations are required for the message updating of R-PBiGAMP-FDJCEE per segment. Similarly, the remaining complexity order of the PBiGAMP-based algorithms is $\mathcal{O}(KN_s)$ for equalization and $\mathcal{O}(L_h KN_s)$ for channel estimation per segment, respectively.

V. SIMULATION RESULTS

In this section, we evaluate both the BER and the normalized mean square error (NMSE) of the proposed GAMP-FDE, R-GAMP-FDE, PBiGAMP-FDJCEE, and R-PBiGAMP-FDJCEE algorithms designed for MFTN signaling in frequency-selective fading channels. To verify their efficiency, we extend the MMSE of [24], the UTAMP algorithm

TABLE III: Complexity Analysis

Algorithm	Equalization		Channel Estimation
	Complex Multiplications	FFT/IFFT Operations	
MMSE-TDE	$\mathcal{O}(K^3 N^3)$	-	-
GAMP-TDE	$\mathcal{O}(K^2 N^2)$	-	-
UTAMP-TDE	$\mathcal{O}(K^3 N^3)$	-	-
GAMP-FDE	$\mathcal{O}(L_q K N_s)$	$9L_q K N_s \log(K N_s)$	-
R-GAMP-FDE	$\mathcal{O}(L_q K N_s)$	$5L_q K N_s \log(K N_s)$	-
LS-MMSE	$\mathcal{O}(L_q K^3 N_s^3)$	-	$\mathcal{O}(L_q L_h N_p)$
PBiGAMP-FDJCEE	$\mathcal{O}(L_q K N_s)$	$13L_q K N_s \log(K N_s)$	$\mathcal{O}(L_q L_h K N_s)$
R-PBiGAMP-FDJCEE	$\mathcal{O}(L_q K N_s)$	$5L_q K N_s \log(K N_s)$	$\mathcal{O}(L_q L_h K N_s)$

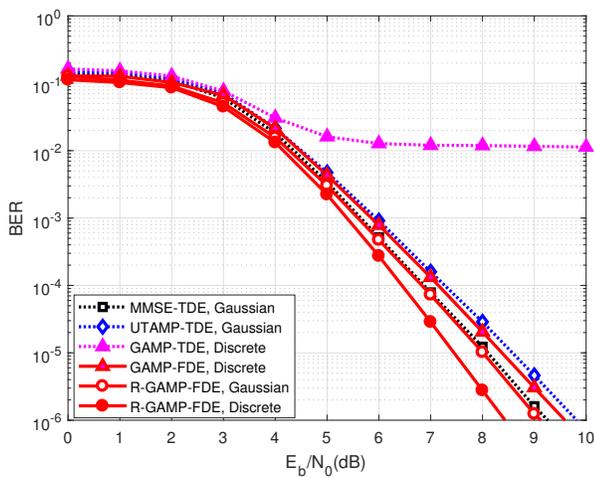


Fig. 4: BER performance of different equalization algorithms for MFTN systems with $\tau = 0.9, \nu = 0.8$ in the case of perfectly known CSI.

of [26], and the GAMP algorithm of [25] to TDE of MFTN systems and then compare them to the proposed algorithms.

We consider an LDPC code of rate $R_c = 23/28$ and code length of $L_c = 4032$ bits for our MFTN systems. The number of subcarriers is $K = 64$ and each subcarrier has $N = 512$ symbols employing quadrature phase shift keying (QPSK) modulation. The carrier frequency and symbol period are 2 GHz and 2 μ s, respectively. The roll-off factor of the RRC shaping filter is $\beta = 0.3$ and the truncated lengths of the inherent 2D interferences are $N_1 = 12$ and $K_1 = 1$, respectively. The fixed truncated lengths have been verified to be able to cover the dominant 2D interferences in the following simulations. Moreover, the length of each segment is $L_m = 42$. For frequency-selective fading channels, the number of taps is set to be $L_h = 8$ and the l -th PDP satisfies $\sigma_{h_l}^2 = \exp(-0.1l) / (\sum_l \sigma_{h_l}^2)$. We consider a car traveling at 80 km/h and a high-speed trains at 270 km/h. The corresponding Doppler spreads are $f_d = 150$ Hz and $f_d = 500$ Hz, respectively, while the normalized fading rates are 0.0003 and 0.001. The number of external iterations between the equalizer and the channel decoder is $T_{\text{ex}} = 50$ and the number of internal LDPC decoding iterations is $T_{\text{in}} = 50$.

In Fig. 4, we compare the BER performance of the proposed GAMP-FDE and R-GAMP-FDE algorithms to those of the extended equalization algorithms at a velocity of 80 km/h,

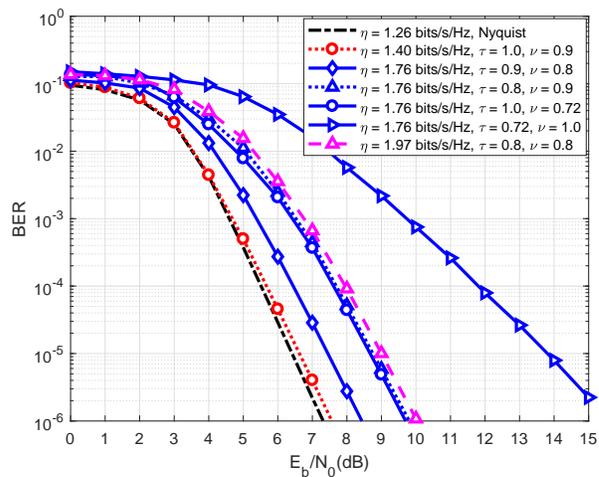


Fig. 5: BER performance of the proposed R-GAMP-FDE algorithm for MFTN systems with various time and frequency packing factors in the case of perfectly known CSI.

where the Doppler spread is $f_d = 150$ Hz. The GAMP-TDE algorithm has an error floor in the high E_b/N_0 region due to the divergence of GAMP induced by the ill-conditioning problem of MFTN systems. Compared to GAMP-TDE, both UTAMP-TDE² and GAMP-FDE significantly improve the BER performance, where GAMP-FDE obtains an additional gain by exploiting the exact discrete *a priori* distributions of the transmitted symbols. Since GAMP-based algorithms rely on the MMSE approximations, the MMSE-TDE algorithm³ outperforms both the UTAMP-TDE and GAMP-FDE. Moreover, the proposed R-GAMP-FDE algorithm is superior to the MMSE-TDE algorithm in terms of BER. The performance gain of R-GAMP-FDE can be attributed to the average approximations of $\nu_q^x(t)$ and $\nu_q^s(t)$ in (33)-(38), which circumvents the strong sensitivity of the ill-conditioned matrices to small perturbations. Moreover, compared to its counterpart using the exact discrete *a priori* probabilities, R-GAMP-FDE employing the Gaussian *a priori* distribution approximation suffers from an additional 0.7 dB E_b/N_0 loss at $\text{BER} = 10^{-6}$.

²Since the UTAMP algorithm is guaranteed to converge in case of Gaussian priors, the UTAMP-TDE algorithm only exploits the Gaussian approximations of the *a priori* probabilities of the transmitted symbols.

³To avoid an intractable computational complexity imposed by the discrete *a priori* distributions of the transmitted symbols, the MMSE-TDE algorithm only employ its Gaussian approximations.

In Fig. 5, we evaluate the BER performance of MFTN systems employing R-GAMP-FDE at different combinations of time and frequency packing factors, where the Doppler spread is $f_d = 150$ Hz. The BER performance of their Nyquist counterpart (i.e., $\tau = 1.0, \nu = 1.0$) with perfect CSI is also included as a benchmark. It is seen that at $\tau = 1.0, \nu = 0.9$, MFTN signaling improves the BE⁴ by up to 11% at a negligible BER performance degradation. When further reducing τ and ν , our MFTN systems based on the proposed R-GAMP-FDE algorithm obtain about 39% and 56% higher transmission rates, for $\tau = 0.9, \nu = 0.8$ and $\tau = 0.8, \nu = 0.8$, respectively, albeit at the cost of about 1.1 dB and 2.6 dB E_b/N_0 losses at $\text{BER} = 10^{-6}$. Moreover, for a given pulse shaping filter and a fixed BE, we can significantly improve the BER performance of MFTN signaling by optimizing τ and ν . Observe that MFTN systems having $\tau = 1.0, \nu = 0.72$, $\tau = 0.9, \nu = 0.8$, $\tau = 0.8, \nu = 0.9$ and $\tau = 0.72, \nu = 1.0$ achieve the same BE, while with $\tau = 0.9, \nu = 0.8$ they achieve better BER than other packing factors. This is because MFTN systems having $\tau = 0.9, \nu = 0.8$ suffer from a lower inherent interference energy⁵ than those of other cases, when we employ an RRC shaping filter, as shown in Fig. 6(a).

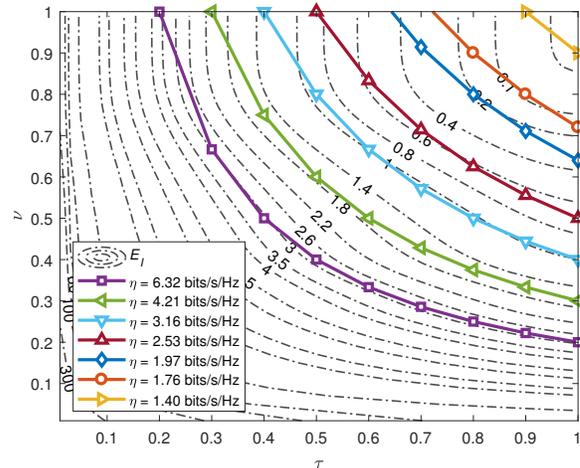
To clarify the relationship between the packing factor combinations and the pulse shaping filter in Fig. 6, we evaluate the inherent interference energy E_I and BE η of MFTN systems employing RRC and Gaussian pulses⁶ with different packing factor combinations. Observe that the effects of τ and ν on E_I depend on the choice pulse shaping filters for a fixed BE. For an RRC pulse, the reduction of τ leads to higher interference energy than that of ν , when we keep the BE constant. For example, the MFTN systems having $\tau = 0.9, \nu = 0.8$ and $\tau = 0.72, \nu = 1.0$ have the same BE, while the interference energy for $\tau = 0.72, \nu = 1.0$ is much higher than that for $\tau = 0.9, \nu = 0.8$. The corresponding BER performance is depicted in Fig. 5, where the MFTN system with $\tau = 0.72, \nu = 1.0$ suffers from an additional 6.5 dB E_b/N_0 erosion at $\text{BER} = 10^{-5}$ compared to its counterpart using $\tau = 0.9, \nu = 0.8$. However, for a Gaussian pulse, the negative influence of ν on E_I is more grave than that of τ . The interference energy remains nearly constant within a certain range of τ . Then, we obtain a higher BE at a negligible BER performance degradation via reducing τ in the case of Gaussian pulses. It is concluded that we can improve the BE by jointly optimizing τ and ν , instead of only reducing either τ or ν .

In Fig. 7 and Fig. 8, we compare both the BER and NMSE performance of the proposed JCEE algorithms to those of the LS-MMSE method. In the following simulations, we consider MFTN systems in frequency-selective fading channels at a velocity of 270 km/h, where the corresponding Doppler spread is $f_d = 500$ Hz. The BER and NMSE performance of

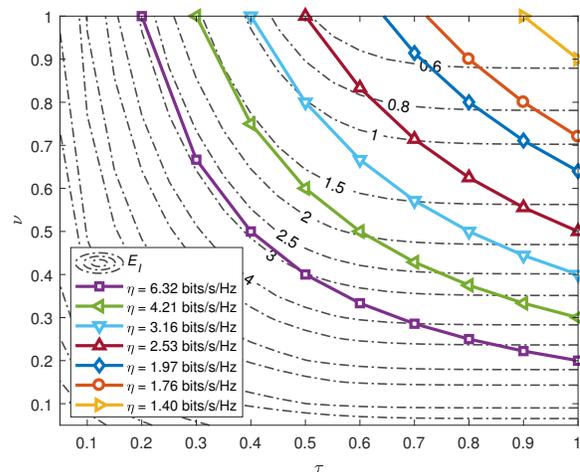
⁴The BE of MFTN signaling can be computed as $\eta = R_c \log_2 M / [\tau\nu(1 + \beta)]$ bits/s/Hz [1].

⁵In [27], the inherent interference energy of MFTN signaling is defined as $E_I = E_s \sum_n \sum_k |A_p(n\tau T, k\nu F)|^2 - 1$, where E_s denotes the energy of the transmitted symbols.

⁶The Gaussian pulse is expressed as $p_G(t) = (2/\rho)^{1/4} e^{-\pi t^2/\rho}$ with $\rho = 0.3$.



(a) RRC



(b) Gaussian

Fig. 6: Interference energy and BE of MFTN systems with various pulse shaping filters and packing factors.

their Nyquist-based counterpart with unknown CSI are also included as benchmarks. For channel estimation, the proposed PBiGAMP-based algorithms obtain more accurate channel estimates than LS-MMSE, since the former exploits both the pilots and the data symbols for improving the accuracy of channel estimation. However, superior channel estimation cannot always guarantee better BER performance, especially in severe inherent interference scenarios. Compared to the Nyquist-based counterpart, our MFTN systems employing the proposed algorithms obtain the similar NMSE performance, but still suffer from a non-negligible BER performance loss. For example, MFTN signaling employing the proposed R-PBiGAMP-FDJCEE algorithm can increase the transmission rate by up to 39%, at the cost of about 1.2 dB E_b/N_0 loss at $\text{BER} = 10^{-6}$. That means that the inherent 2D interferences induced by time-frequency packing are more crucial than those imposed by channels in this case. For equalization, as an efficient approximation of MMSE equalization [17], PBiGAMP-FDJCEE can only approaches the LS-MMSE algorithm in terms of BER, while the complexity of PBiGAMP-

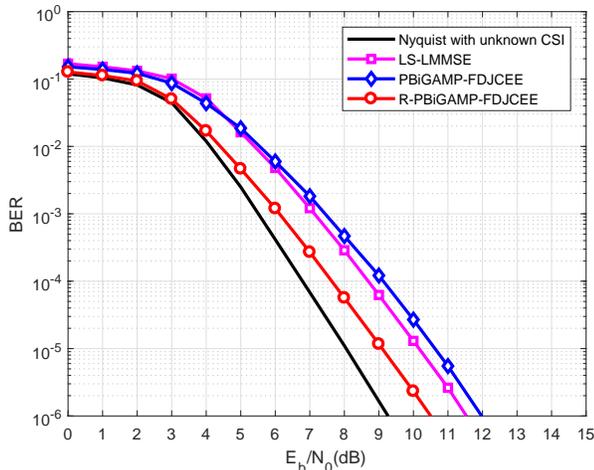


Fig. 7: BER performance of different algorithms for MFTN systems with $\tau = 0.9, \nu = 0.8$ in the absence of CSI.

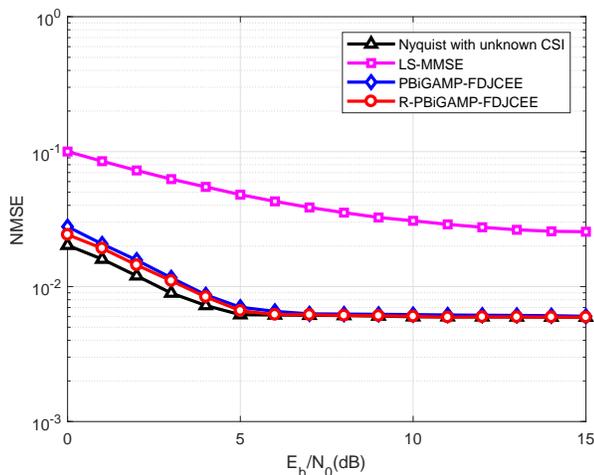


Fig. 8: NMSE performance of different algorithms for MFTN systems with $\tau = 0.9, \nu = 0.8$ in the absence of CSI.

FDJCEE is much smaller than that of LS-MMSE as shown in TABLE III. Note that the proposed R-PBiGAMP-FDJCEE can still improve the BER performance, compared to LS-MMSE. In Fig. 7, we observe about 1.0 dB E_b/N_0 gain at $\text{BER} = 10^{-5}$. The performance gain of R-PBiGAMP-FDJCEE arises from the efficient average approximations of $\nu_q^x(t)$ and $\nu_q^s(t)$ in (33)-(40), which improves the robustness of the ill-conditioned matrices to small perturbations. Moreover, as shown in Fig. 8, the NMSE performance of different algorithms tends to remain constant upon increasing E_b/N_0 .

Considering that the channel coefficients are assumed to be constant in each segment, the segment length of L_m is crucial for its performance versus computational complexity. In Fig. 9, we investigate the impact of L_m on the BER of the proposed R-PBiGAMP-FDJCEE algorithm. It is observed that BER degrades upon increasing L_m for a fixed Doppler spread. When the segment length increases to $L_m = 126$, a significant BER loss can be observed due to the inaccurate quasi-static channel assumption. By contrast, we can improve the BER of the proposed R-PBiGAMP-FDJCEE algorithm by reducing the segment length, L_m , at the cost of increased computational

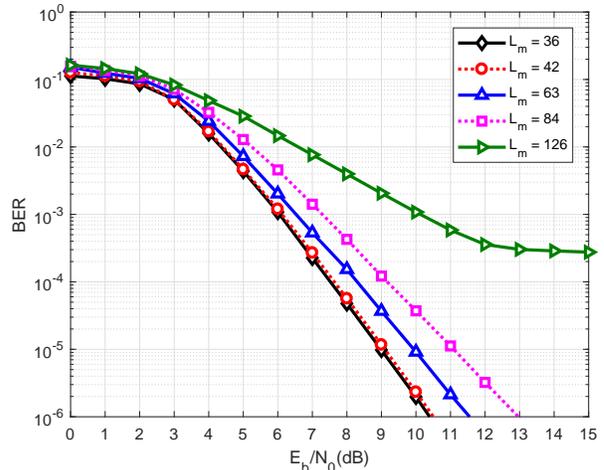


Fig. 9: BER performance of the proposed R-PBiGAMP-FDJCEE algorithm employing various length of segment for MFTN systems with $\tau = 0.9, \nu = 0.8$.

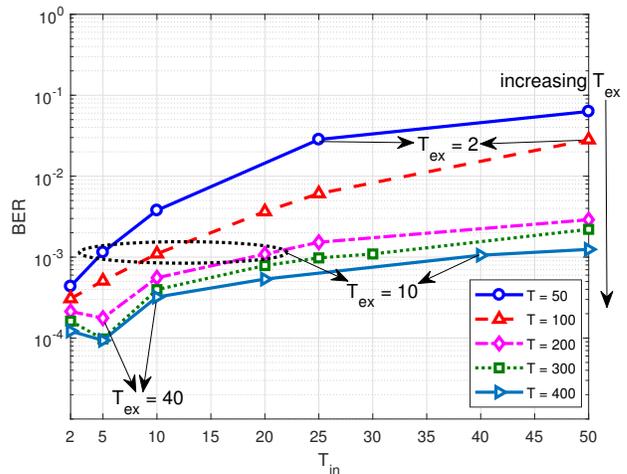


Fig. 10: BER performance of the proposed R-PBiGAMP-FDJCEE algorithm employing different number of iterations for MFTN systems with $\tau = 0.9, \nu = 0.8$ at $E_b/N_0 = 7$ dB.

complexity. Compared to the cases with $L_m = 63$ and $L_m = 84$, R-PBiGAMP-FDJCEE using $L_m = 42$ obtains an additional 1.2 dB and 2.7 dB E_b/N_0 gains at $\text{BER} = 10^{-6}$ in Fig. 9. When we further reduce the segment length, its BER performance gain becomes marginal, while the number of segments L_q significantly increases. Hence, $L_m = 42$ is an appropriate option for R-PBiGAMP-FDJCEE in the MFTN system considered. In practice, we can adjust the segment length of L_m to strike a performance versus complexity trade-off.

In Fig. 10, we evaluate the BER performance of the proposed R-PBiGAMP-FDJCEE algorithm with various combinations of T_{in} and T_{ex} at $E_b/N_0 = 7$ dB, where the total number of iterations $T = T_{in} \cdot T_{ex}$. It is seen that, given T_{in} , BER performance is improved with the increase of T_{ex} . On the contrary, given T_{ex} , e.g., $T_{ex} = 2$ and $T_{ex} = 10$, increasing T_{in} does not improve the BER performance. Moreover, when we fix the total number of iterations T , for $T_{in} > 10$, increasing T_{in} (or decreasing T_{ex}) deteriorates the BER performance.

Hence, we can conclude that the external iterations between the equalizer and the channel decoder are much more crucial than the internal iterations of LDPC decoder. Observe that $T_{\text{in}} = 5$ is appropriate for the studied MFTN systems. In fact, increasing T_{in} may even worsen the BER performance. For $T_{\text{ex}} = 40$, there is a significant BER loss when T_{in} varies from 5 to 10. This is because the belief output from LDPC decoder could be overconfident when T_{in} becomes large. Hence, T_{in} and T_{ex} should be optimized to compromise between the BER performance and the computational complexity.

VI. CONCLUSIONS

In this paper, we proposed a low-complexity PBiGAMP-based iterative message passing receiver for MFTN systems operating in frequency-selective fading channels. By inserting a cyclic postfix in the FD and truncating the severe 2D interferences, we are able to construct a segment-based FD received signal model in the form of a block circulant linear transition matrix. It was shown that the proposed model reduce the condition number of the ill-conditioned linear transition matrices, which can potentially lead to improved BER performance. Building on this model, we proposed the PBiGAMP-based parametric JCEE algorithm for simultaneously estimating the modulated symbols and channel coefficients. For improving the robustness of the proposed JCEE algorithm to small perturbations, we further developed a refined JCEE algorithm by inserting a series of scaled identity matrices. The proposed JCEE algorithms can be readily decomposed into two FDE algorithms when the CSI is perfectly known. The simulation results verified the benefits of the proposed PBiGAMP-based receiver for MFTN systems operating in frequency-selective fading channels. Compared to its Nyquist-signaling counterpart, MFTN systems employing the proposed refined algorithms achieve about 39% higher transmission rate at the cost of 1.2 dB BER E_b/N_0 loss. Moreover, our work provides a theoretical reference for solving the challenging channel estimation problem of other emerging systems [28]–[30].

APPENDIX A DERIVATIONS OF (11)

According to Theorem 5.6.2 in [21], a block circulant matrix $\mathbf{A} \in \mathcal{BC}_{m,n}$ having the the first n rows of $[\mathbf{A}_1 \ \mathbf{A}_2 \ \cdots \ \mathbf{A}_m]$ satisfies $\mathbf{A} = \sum_{k=0}^{m-1} (\mathbf{\Pi}_m^k \otimes \mathbf{A}_{k+1})$, where $\mathbf{\Pi}_m^k$ is an $m \times m$ circulant permutation matrix with the first row vector $\boldsymbol{\lambda}_{\mathbf{\Pi}_m^k} = [0_k^T \ 1 \ 0_{(m-k-1)}^T]$, \mathbf{A}_k is an $n \times n$ matrix. Hence, the 2D interference matrix $\bar{\mathbf{G}}_l \in \mathcal{BC}_{K,N_s}$ can be expressed as

$$\bar{\mathbf{G}}_l = \sum_{k=0}^{2K_1} (\mathbf{\Pi}_K^k \otimes \mathcal{D}(\bar{\mathbf{g}}_{K_1-k,l})). \quad (41)$$

Since $\mathbf{\Pi}_K^k$ is a circulant matrix, we can diagonalize it using the FFT operation. Equation (41) can be rewritten as

$$\begin{aligned} \bar{\mathbf{G}}_l &= \sum_{k=0}^{2K_1} (\mathbf{F}_K^H \boldsymbol{\Lambda}_{\mathbf{\Pi}_K^k} \mathbf{F}_K) \otimes \mathbf{I}_{N_1}^H (\mathbf{I}_{N_1} \mathcal{D}(\bar{\mathbf{g}}_{K_1-k,l}) \mathbf{I}_{N_1}^H) \mathbf{I}_{N_1} \quad (42) \\ &= (\mathbf{F}_K \otimes \mathbf{I}_{N_1})^H \left[\sum_{k=0}^{2K_1} \boldsymbol{\Lambda}_{\mathbf{\Pi}_K^k} \otimes \mathcal{D}(\bar{\mathbf{g}}_{K_1-k,l}) \right] (\mathbf{F}_K \otimes \mathbf{I}_{N_1}) \\ &= (\mathbf{F}_K \otimes \mathbf{I}_{N_1})^H \boldsymbol{\Lambda}_{\bar{\mathbf{G}}_l} (\mathbf{F}_K \otimes \mathbf{I}_{N_1}), \end{aligned}$$

where $\boldsymbol{\Lambda}_{\mathbf{\Pi}_K^k} = \mathcal{D}(\sqrt{K} \mathbf{F}_K^H \boldsymbol{\lambda}_{\mathbf{\Pi}_K^k} \mathbf{F}_K)$ and $\boldsymbol{\Lambda}_{\bar{\mathbf{G}}_l} = \sum_{k=0}^{2K_1} \boldsymbol{\Lambda}_{\mathbf{\Pi}_K^k} \otimes \mathcal{D}(\bar{\mathbf{g}}_{K_1-k,l})$ are diagonal matrices having different sizes.

APPENDIX B

DERIVATIONS OF THE PROBABILITY DISTRIBUTION OF THE COLORED NOISE IN (12)

In (12), the colored noise term is given by

$$\begin{aligned} \check{\mathbf{v}}_q &= (\mathbf{F}_K \otimes \mathbf{F}_{N_s}) (\mathbf{z}_{q-1}^{\text{tail}} + \mathbf{z}_{q+1}^{\text{head}} + \boldsymbol{\omega}_q) \quad (43) \\ &= (\mathbf{F}_K \otimes \mathbf{F}_{N_s}) [(\mathbf{I}_K \otimes \mathbf{\Xi}_t) \mathbf{z}_{q-1} + (\mathbf{I}_K \otimes \mathbf{\Xi}_h) \mathbf{z}_{q+1} + \boldsymbol{\omega}_q], \end{aligned}$$

where \mathbf{z}_q denotes the noiseless measurement vector in the q -th segment and $\boldsymbol{\omega}_q$ contains the zero mean colored noise samples. Hence, the mean vector of the colored noise $\check{\mathbf{v}}_q$ is derived as

$$\mathbf{z}_q^v = (\mathbf{F}_K \otimes \mathbf{F}_{N_s}) [(\mathbf{I}_K \otimes \mathbf{\Xi}_t) \hat{\mathbf{z}}_{q-1} + (\mathbf{I}_K \otimes \mathbf{\Xi}_h) \hat{\mathbf{z}}_{q+1}], \quad (44)$$

where the estimated noiseless measurement vector is $\hat{\mathbf{z}}_q = (\mathbf{F}_K \otimes \mathbf{F}_{N_s})^H \sum_{l=0}^{L_b-1} \hat{h}_{q,l} \boldsymbol{\Lambda}_{q,l} \ddot{\mathbf{F}}_{K N_s} \hat{\mathbf{x}}_q$, and $\hat{\mathbf{x}}_q$ represents the transmitted symbol estimates.

Moreover, the covariance matrix of the FD colored noise imposed by a non-orthogonal matched filter is expressed as

$$\begin{aligned} \mathbf{R} &= (\mathbf{F}_K \otimes \mathbf{F}_{N_s}) \mathbf{R}_{\boldsymbol{\omega}} (\mathbf{F}_K \otimes \mathbf{F}_{N_s})^H \quad (45) \\ &= (\mathbf{F}_K \otimes \mathbf{F}_{N_s}) (\ddot{\mathbf{R}}_{\boldsymbol{\omega}} - \dot{\mathbf{R}}_{\boldsymbol{\omega}}) (\mathbf{F}_K \otimes \mathbf{F}_{N_s})^H, \end{aligned}$$

where $\mathbf{R}_{\boldsymbol{\omega}}$ is the covariance matrix of the colored noise in the TD, $\ddot{\mathbf{R}}_{\boldsymbol{\omega}}$ is a block circulant matrix, its first N_s rows is $[\ddot{\Theta}_{-K_1} \ \cdots \ \ddot{\Theta}_{K_1} \ \mathbf{0}_{N_s \times (K-2K_1-1)N_s}]$, $\ddot{\Theta}_k$ is an $N_s \times N_s$ circulant matrix having the first row vector $\check{\boldsymbol{\lambda}}_{\ddot{\Theta}_k} = [A_p(0, k), \dots, A_p(N_1 \tau T_0, k), 0, \dots, 0, A_p(-N_1 \tau T_0, k), \dots, A_p(-\tau T_0, k)]$. Similarly, the first N_s rows of the block circulant matrix $\dot{\mathbf{R}}_{\boldsymbol{\omega}}$ is $[\dot{\Theta}_{-K_1} \ \cdots \ \dot{\Theta}_{K_1} \ \mathbf{0}_{N_s \times (K-2K_1-1)N_s}]$, where $\dot{\Theta}_k$ is an $N_s \times N_s$ Toeplitz matrix having the first row vector $\check{\boldsymbol{\lambda}}_{\dot{\Theta}_k} = [0, \dots, 0, A_p(-N_1 \tau T_0, k), \dots, A_p(-\tau T_0, k)]$ and the first column vector $\check{\boldsymbol{\zeta}}_{\dot{\Theta}_k} = [0, \dots, 0, A_p(N_1 \tau T_0, k), \dots, A_p(\tau T_0, k)]^T$.

According to the properties of block circulant matrices in [21], we can obtain the block diagonal matrices via 2D FFT operations. As a special case, the block circulant matrices associated with circulant matrices can be diagonalized by a 2D FFT. Hence, (45) is rewritten as

$$\mathbf{R} = \ddot{\mathbf{\Lambda}} - \dot{\mathbf{\Upsilon}} \approx \ddot{\mathbf{\Lambda}}_{\boldsymbol{\omega}} - \mathcal{D}(\dot{\mathbf{\Upsilon}}), \quad (46)$$

where $\ddot{\mathbf{\Lambda}} = (\mathbf{F}_K \otimes \mathbf{F}_{N_s}) \ddot{\mathbf{R}}_{\boldsymbol{\omega}} (\mathbf{F}_K \otimes \mathbf{F}_{N_s})^H$ is a diagonal matrix, $\dot{\mathbf{\Upsilon}} = (\mathbf{F}_K \otimes \mathbf{F}_{N_s}) \dot{\mathbf{R}}_{\boldsymbol{\omega}} (\mathbf{F}_K \otimes \mathbf{F}_{N_s})^H$ is a block diagonal matrix. Moreover, the approximation error can be controlled by appropriately selecting the packing factor in the TD [15].

Considering that the variances of the transmitted symbols are averaged in the proposed PBiGAMP-based iterative message passing receiver, the average variance of the 2D interferences imposed by the adjacent segments is formulated as

$$\nu_q^v = \frac{2K_1 + 1}{N_s} \left[\sum_{l=0}^{L_h-1} l |\hat{h}_{q-1,l}|^2 \nu_{q-1}^x + \sum_{l=0}^{L_h-1} (L_h - 1 - l) |\hat{h}_{q+1,l}|^2 \nu_{q+1}^x \right]. \quad (47)$$

According to the above derivations, the covariance matrix of $\check{\mathbf{v}}_q$ is expressed as

$$\mathbf{\Lambda}_q^v = \nu_q^v \mathbf{I}_{KN_s} + \mathbf{R}. \quad (48)$$

REFERENCES

- [1] P. Banelli, S. Buzzi, G. Colavolpe, A. Modenini, F. Rusek, and A. Ugolini, "Modulation formats and waveforms for 5G networks: Who will be the heir of OFDM?: An overview of alternative modulation schemes for improved spectral efficiency," *IEEE Signal Process. Mag.*, vol. 31, no. 6, pp. 80–93, Nov. 2014.
- [2] S. Li, J. Yuan, B. Bai, and N. Benvenuto, "Code-based channel shortening for faster-than-Nyquist signaling: Reduced-complexity detection and code design," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 3996–4011, Jul. 2020.
- [3] W. Yuan, N. Wu, Q. Guo, D. W. K. Ng, J. Yuan, and L. Hanzo, "Iterative joint channel estimation, user activity tracking, and data detection for FTN-NOMA systems supporting random access," *IEEE Trans. Commun.*, vol. 68, no. 5, pp. 2963–2977, May 2020.
- [4] F. Rusek and J. B. Anderson, "Constrained capacities for faster-than-Nyquist signaling," *IEEE Trans. Inf. Theory*, vol. 55, no. 2, pp. 764–775, Feb. 2009.
- [5] —, "Multistream faster than Nyquist signaling," *IEEE Trans. Commun.*, vol. 57, no. 5, pp. 1329–1340, May 2009.
- [6] 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.
- [7] A. Piemontese, A. Modenini, G. Colavolpe, and N. S. Alagha, "Improving the spectral efficiency of nonlinear satellite systems through time-frequency packing and advanced receiver processing," *IEEE Trans. Commun.*, vol. 61, no. 8, pp. 3404–3412, Aug. 2013.
- [8] G. Colavolpe and T. Foggi, "Time-frequency packing for high-capacity coherent optical links," *IEEE Trans. Commun.*, vol. 62, no. 8, pp. 2986–2995, Aug. 2014.
- [9] M. Secondini, T. Foggi, F. Fresi, G. Meloni, F. Cavaliere, G. Colavolpe, E. Forestieri, L. Pot, R. Sabella, and G. Prati, "Optical timefrequency packing: Principles, design, implementation, and experimental demonstration," *J. Lightwave Technol.*, vol. 33, no. 17, pp. 3558–3570, Sep. 2015.
- [10] L. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inf. Theory*, vol. 20, no. 2, pp. 284–287, Mar. 1974.
- [11] S. Peng, A. Liu, X. Liu, K. Wang, and X. Liang, "MMSE turbo equalization and detection for multicarrier faster-than-Nyquist signaling," *IEEE Trans. Veh. Technol.*, vol. 67, no. 3, pp. 2267–2275, Mar. 2018.
- [12] Y. Ma, F. Tian, N. Wu, B. Li, and X. Ma, "A low-complexity receiver for multicarrier faster-than-Nyquist signaling over frequency selective channels," *IEEE Commun. Lett.*, vol. 24, no. 1, pp. 81–85, Jan. 2020.
- [13] N. Wu, W. Yuan, Q. Guo, and J. Kuang, "A hybrid BP-EP-VMP approach to joint channel estimation and decoding for FTN signaling over frequency selective fading channels," *IEEE Access*, vol. 5, pp. 6849–6858, 2017.
- [14] B. Cai, A. Liu, X. Tong, and F. Cheng, "Training sequences design for channel estimation in FTN system using discrete Fourier transform techniques," *IET Commun.*, vol. 11, no. 17, pp. 2561–2565, Dec. 2017.
- [15] Q. Shi, N. Wu, X. Ma, and H. Wang, "Frequency-domain joint channel estimation and decoding for faster-than-Nyquist signaling," *IEEE Trans. Commun.*, vol. 66, no. 2, pp. 781–795, Feb. 2018.
- [16] W. Yuan, N. Wu, A. Zhang, X. Huang, Y. Li, and L. Hanzo, "Iterative receiver design for FTN signaling aided sparse code multiple access," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 915–928, Feb. 2020.
- [17] J. T. Parker and P. Schniter, "Parametric bilinear generalized approximate message passing," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 4, p. 114, Jun. 2016.
- [18] H. J. Kim, K. Choi, H. B. Lee, H. K. Jung and S. Y. Hahn, "A new algorithm for solving ill conditioned linear systems," *IEEE Trans. Magn.*, vol. 32, no. 3, pp. 1373–1376, May 1996.
- [19] T. Piotrowski and I. Yamada, "MV-PURE estimator: Minimum-variance pseudo-unbiased reduced-rank estimator for linearly constrained ill-conditioned inverse problems," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3408–3423, Aug. 2008.
- [20] S. Isam and I. Darwazeh, "Simple DSP-IDFT techniques for generating spectrally efficient FDM signals," in *Proc. 7th Int. Symp. Commun. Syst. Netw. Digital Signal Process.*, Jul. 2010, pp. 20–24.
- [21] P. J. Davis, *Circulant Matrices*. John Wiley & Sons, Inc, 1979.
- [22] L. Liu, Y. Li, C. Huang, C. Yuen, and Y. L. Guan, "A new insight into GAMP and AMP," *IEEE Trans. Veh. Technol.*, vol. 68, no. 8, pp. 8264–8269, Aug. 2019.
- [23] C. Huang, L. Liu, C. Yuen, and S. Sun, "Iterative channel estimation using LSE and sparse message passing for mmWave MIMO systems," *IEEE Trans. Signal Process.*, vol. 67, no. 1, pp. 245–259, Jan. 2019.
- [24] M. Tuchler, A. C. Singer, and R. Koetter, "Minimum mean squared error equalization using a priori information," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 673–683, Mar. 2002.
- [25] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2011, p. 21682172.
- [26] Q. Guo and J. Xi, "Approximate message passing with unitary transformation," *arXiv preprint, arXiv:1504.04799*, 2015.
- [27] S. Peng, A. Liu, X. Tong, X. Liang, and K. Wang, "On the practical benefit of hexagonal multicarrier faster-than-Nyquist signaling," in *Proc. 9th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Dec. 2017, pp. 1–6.
- [28] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Debbah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4157–4170, Aug. 2019.
- [29] C. Huang, S. Hu, G. C. Alexandropoulos, A. Zappone, C. Yuen, R. Zhang, M. D. Renzo, and M. Debbah, "Holographic MIMO surfaces for 6G wireless networks: Opportunities, challenges, and trends," *IEEE Wireless Commun.*, vol. 27, no. 5, pp. 118–125, Oct. 2020.
- [30] H. Liu, X. Yuan, and Y.-J. A. Zhang, "Matrix-calibration-based cascaded channel estimation for reconfigurable intelligent surface assisted multiuser MIMO," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2621–2636, Nov. 2020.



Yunsi Ma (S'19) received the B.S. degree from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2013, and the M.S. degree from China Academy of Space Technology, Beijing, China, in 2016. She is currently working toward the Ph.D. degree in Beijing Institute of Technology and University of Technology Sydney. Her research interests include statistical inference on graphical models and its application to wireless communications.



Nan Wu (M'11) received his B.S., M.S. and Ph.D. degrees from 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 a 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 IEEE Wireless Communications Letters, the IEEE Access, the International Journal of Electronics and Communications, and the KSII Transactions on Internet and Information Systems.



Dr. J. Andrew Zhang (M'04-SM'11) received the B.Sc. degree from Xi'an JiaoTong University, China, in 1996, the M.Sc. degree from Nanjing University of Posts and Telecommunications, China, in 1999, and the Ph.D. degree from the Australian National University, in 2004.

Currently, Dr. Zhang is an Associate Professor in the School of Electrical and Data Engineering, University of Technology Sydney, Australia. He worked as a researcher with ZTE Corp., Nanjing, China; NICTA, Canberra, Australia; and Data61, CSIRO,

Sydney, Australia. Dr. Zhang's research interests are in the area of signal processing for wireless communications and sensing. He has published more than 210 papers in leading international Journals and conference proceedings, and has won 5 best paper awards. He is a recipient of CSIRO Chairman's Medal and the Australian Engineering Innovation Award in 2012 for exceptional research achievements in multi-gigabit wireless communications. He is serving as an Editor for IEEE Transactions on Communications.



Bin Li received the B.S. degree in information engineering and the M.S. degree in information and communication engineering from the Beijing Institute of Technology, Beijing, China, in 2012 and 2015, respectively, and the Ph.D. degree in electrical and electronic engineering from The University of Hong Kong, Hong Kong, in 2019. Currently, he is an assistant professor with the School of Information and Electronics, Beijing Institute of Technology, Beijing, China. His research interests include signal processing, wireless communications, and machine

learning. He serves as an Editorial Board Member for IEICE Transactions on Communications, and KSII Transactions on Internet and Information Systems.



Lajos Hanzo (<http://www-mobile.ecs.soton.ac.uk>, https://en.wikipedia.org/wiki/Lajos_Hanzo) (FIEEE'04) received his Master degree and Doctorate in 1976 and 1983, respectively from the Technical University (TU) of Budapest. He was also awarded the Doctor of Sciences (DSc) degree by the University of Southampton (2004) and Honorary Doctorates by the TU of Budapest (2009) and by the University of Edinburgh (2015). He is a Foreign Member of the Hungarian Academy of Sciences and a former Editor-in-Chief of the IEEE

Press. He has served several terms as Governor of both IEEE ComSoc and of VTS. He has published 1970 contributions at IEEE Xplore, 19 Wiley-IEEE Press books and has helped the fast-track career of 123 PhD students. Over 40 of them are Professors at various stages of their careers in academia and many of them are leading scientists in the wireless industry. He is also a Fellow of the Royal Academy of Engineering (FREng), of the IET and of EURASIP.