Message Passing-Aided Joint Data Detection and Estimation of Nonlinear Satellite Channels

Yikun Zhang, *Student Member, IEEE,* Bin Li, *Member, IEEE,* Nan Wu, *Member, IEEE,* Yunsi Ma, *Student Member, IEEE,* Weijie Yuan, *Member, IEEE,* and Lajos Hanzo, *Fellow, IEEE,*

***Abstract*—Satellite communication is capable of supporting seamless global coverage. However, owing to the reliance on limited-duration solar power, the high power amplifier (HPA) is often driven close to its saturation point, which leads to severe nonlinear distortion in satellite channels. Thus, mitigating the effect of the nonlinear distortion becomes essential for reliable communications. In this paper, we propose an efficient joint channel estimation and data detection method based on message passing within the associated factor graph modelling the HPA employed in nonlinear satellite channels. Then, we develop a combined belief propagation and mean field (BP-MF) method to cope with the hard constraints and dense short loops on the factor graph. In particular, the parametric message updating expressions relying on the canonical parameters are derived in the symbol detection part. To alleviate the impact of dense loops, we reformulate the system model into a compact form within the channel estimation part and then reconstruct a loop-free subgraph associated with vector-valued nodes to guarantee convergence. Furthermore, the proposed BP-MF method is also extended to the realistic scenario of having unknown noise variance. To further reduce the computational complexity of the large-scale matrix inversion of channel estimation, the generalized approximate message passing (GAMP) algorithm is employed to decouple the vector of channel coefficient estimation into a series of scalar estimations. Simulation results show that the proposed methods outperform the state-of-the-art benchmarks both in terms of bit error rate performance and channel estimation accuracy.**

***Index Terms*—Nonlinear satellite channel, Volterra series, joint channel estimation and data detection, mean field approximation, generalized approximate message passing.**

I. Introduction

Satellite communication plays an essential role in next generation wireless networks by providing seamless global coverage [1]–[5]. The onboard high power amplifier (HPA) is a key component of the transponder, which relies on limited-duration solar charging. To achieve a high power efficiency, the HPA is typically operated near its saturation point, which leads to nonlinear distortion.

The nonlinearity of the HPA has been popularly characterized by a frequency-independent memoryless model [6], [7]. The most widely used memoryless HPA model is described by the AM/AM and AM/PM functions [8]. Nonetheless, the input multiplexing (IMUX) and output multiplexing (OMUX) filters placed before and after the HPA introduce memory into these nonlinear satellite channels. Several models have been proposed for characterizing nonlinear channels exhibiting memory, such as the polynomial model [9], the Wiener model [10], the Hammerstein model [10] and Volterra series [11]. The Volterra series based model has a generic expression and hence it is widely applied for characterizing nonlinear satellite channels.

Numerous authors have investigated the design of equalizers for channels modeled by Volterra series [12]–[17]. The optimal maximum *a posteriori* (MAP) equalizer is developed in [12]. However, the complexity of the optimal MAP equalizer increases exponentially with the channel's memory length. To reduce the computational complexity, a nonlinear soft interference canceller (NL-SIC) is proposed in [13], which is an extension of the classic SIC designed for linear channels in [18]. In [14], a linear minimum mean square error (LMMSE) criterion-based equalizer is used for canceling the nonlinear distortion. However, the complexity of the LMMSE-based methods still grows cubically with the length of their processing window.

Several low-complexity iterative equalizers have also been developed based on the factor graph (FG) modelling nonlinear channels [15]–[17]. In [15], a Markov Chain Monte Carlo (MCMC) equalizer relying on different sampling methods is employed for reducing the complexity of the sum-product algorithm (SPA)-based forward-backward (FB) equalizer, while achieves the comparable performance. However, the MCMC method may suffer from convergence problems resulting from the sensitivity to the noise fluctuations [19]. In [16], a soft-input soft-output (SISO) equalizer is proposed to attain a linear computational complexity, where a simplified Volterra series model is employed for characterizing nonlinear channels, resulting in a performance erosion. In [17], a different message passing algorithm is designed based on FG, which gleans extrinsic information from the associated equalizer. By avoiding any matrix inversion operation, the computational complexity

This work was supported in part by the National Natural Science Foundation of China with Grant Nos. 62001027, 61971041, and the Beijing Institute of Technology Research Fund Program for Young Scholars. L. Hanzo would like to acknowledge the financial support of the Engineering and Physical Sciences Research Council projects EP/W016605/1 and EP/P003990/1 (COALESCE) as well as of the European Research Council's Advanced Fellow Grant QuantCom (Grant No. 789028). (Corresponding author: Bin Li)

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

W. Yuan is with the Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen 518055, China (e-mail: yuanwj@sustech.edu.cn).

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).

TABLE I
EXPLICITLY CONTRASTING OUR CONTRIBUTIONS TO THE LITERATURE.

Feature	This Paper	[13], [14]	[15]	[16]	[17]	[20]	[25]	[26]
Nonlinear Channels	✓	✓	✓	✓	✓	✓		✓
Factor Graph Framework	✓		✓	✓	✓	✓	✓	
Joint Channel Estimation and Equalization	✓						✓	✓
Parametric Nonlinear Channel Estimation	✓							
Valuable Use for Estimation Uncertainty	✓	✓			✓	✓	✓	
Without Linearization	✓	✓	✓		✓			✓
Unknown Noise Variance	✓							

of the equalizer can be significantly reduced compared to the LMMSE equalizer [14] and the FB equalizer [15]. In [20], a Gaussian message passing receiver is developed for faster-than-Nyquist signaling over nonlinear channels, where the channel model is linearized using Taylor expansion, and thus the problem of symbol detection is reformulated into a linear state-space model.

Nevertheless, both of the above-mentioned algorithms only considered the idealized simplifying assumption of perfectly known channel state information (CSI). In practice, the CSI has to be estimated by relying on sufficiently long training sequences. However, due to the inevitable radiation, ageing, temperature variations or different mismatch effects, the characteristics of HPAs are time-varying, which requires re-estimation of CSI and results in an extra training overhead. Also, with the increase of the memory length and nonlinear order of channels, the loss of spectral efficiency imposed by pilots is even higher. It has been shown [21]–[26] that joint channel estimation and data detection not only significantly reduces the number of pilots, but also improves the CSI estimation accuracy in linear channels. In [27], joint channel estimation and equalization is proposed for nonlinear MIMO channels characterized by the Volterra series, where the alternating least squares (ALS) algorithm is employed for both channel estimation and data detection in an iterative way. However, this ALS algorithm fails to reach its full potential due to failing to rely on soft information during its iterations.

Against this backdrop, we propose a joint channel estimation and data detection method for nonlinear satellite channels based on FGs. The novel contributions of this paper are explicitly itemized next and they are also clearly contrasted to the literature in Table I. Our main contributions are summarized as follows:

- We construct a FG to represent communications over nonlinear channels exhibiting memory characterized by Volterra series. To tackle the problems of message passing caused by hard constraints and dense short loops, the FG is intentionally split into belief propagation (BP) and mean field (MF) regions. Then, we invoke a combined BP-MF method for deriving messages on a FG. Due to the nonlinear terms in the system model, the messages are in non-Gaussian forms. To this end, a parametric message update with canonical parameters are derived. Moreover, by reformulating the system model, we conceive a loop-free subgraph in terms of vectors for performing channel estimation. Furthermore, the noise variance estimation using a Gamma *a priori* distribution is embedded into the proposed method.
- To circumvent the high computational complexity imposed by matrix inversion, we further resort to the generalized approximate message passing (GAMP) algorithm [28] for estimating the channel coefficients. Accordingly, the vector estimation process is transformed into scalar operations. The iterations of the GAMP algorithm may be incorporated into the iterations between the equalizer and decoder. The complexity of channel estimation based on the GAMP algorithm only grows linearly with the number of channel coefficients.

The rest of this paper is organized as follows. Section II presents the system model and problem formulation over nonlinear satellite channels. In Section III, the proposed BP-MF method of joint channel estimation and data detection over nonlinear satellite channels is derived in detail, and extended to the scenario of realistic unknown noise variance. In Section IV, our GAMP-aided joint channel estimation and data detection algorithm is developed for reducing the computational complexity. The performance of the proposed algorithms is evaluated by Monte Carlo simulations in Section V, while our conclusions are offered in Section VI.

Notations: Boldface capital and lower case letters denote matrices and vectors, respectively; The operations $(\cdot)^*$, $(\cdot)^T$ $(\cdot)^H$ and $(\cdot)^{-1}$ represent the complex conjugate, transpose, conjugate transpose, and matrix inverse, respectively; $|\cdot|$ denotes the modulus of a complex number; $\Re(x)$ denotes the real part of complex number x; The (n,m)th element of the matrix $\mathbf{X}$ is written as X_{nm}; $\mathbb{E}\{\cdot\}$ and $\mathbb{V}\{\cdot\}$ represent expectation and variance operations, respectively; $\mathrm{Tr}\{\cdot\}$ denotes the trace operation; The symbol $\propto$ denotes equality up to a constant; $\mathcal{CN}(x; m_x, V_x)$ denotes a complex Gaussian probability density function (pdf) of the random variable x with mean m_x and variance V_x.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a coded system communicating over nonlinear satellite channel exhibiting memory, as depicted in Fig. 1. At the transmitter side, the information bit sequence $\mathbf{b} \triangleq [b_1, b_2, ..., b_{N_b}]^T$ is encoded into a coded bit sequence $\mathbf{c} \triangleq [c_1, c_2, ..., c_{N_c}]^T$, where N_b and N_c are the lengths of information bit sequence and coded bit sequence, respectively. The sequence $\mathbf{c}$ is further mapped to a data symbol vector

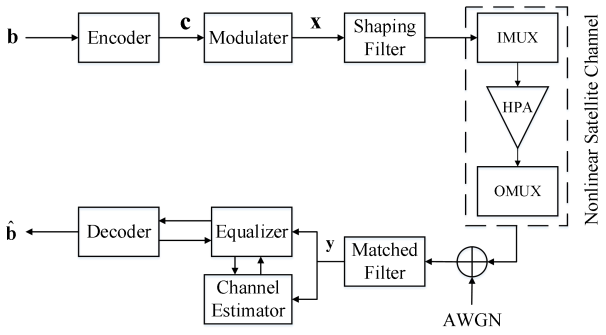


Fig. 1. System model of the coded system communicating over nonlinear satellite channels.

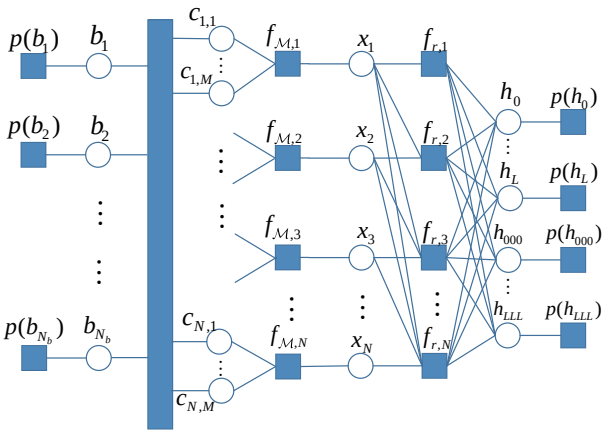


Fig. 2. Factor graph representation of joint *a posteriori* distribution in (3).

$\mathbf{x} \triangleq [x_1, x_2, ..., x_N]^T$ of length N, where we have $x_n \in \chi$ and χ denotes the 2^M-ary symbol alphabet. Then, $\mathbf{x}$ is passed through the shaping filter, whose output signal is transmitted over the nonlinear satellite channel corrupted by additive white Gaussian noise (AWGN). The satellite transponder consists of an IMUX filter, an HPA and an OMUX filter, as shown in Fig. 1. The IMUX filter aims for removing the adjacent-channel interferences, while the OMUX filter is intended for reducing the out-of-band emissions of spectral expansion. Volterra series are employed for modelling the satellite channel's nonlinearity. Due to the bandpass nature of the satellite communication channel, only the odd-order Volterra terms are considered [12]. In practice, employing a third-order Volterra series is sufficient for modelling nonlinear satellite channels [13]–[17], [20].

For a third-order Volterra series [11], the n-th discrete-time received signal at the output of the matched filter can be represented as

$$y_n = \sum_{l=0}^{L} h_l x_{n-l} + \sum_{i=0}^{L} \sum_{j\geq i}^{L} \sum_{k=0}^{L} h_{ijk} x_{n-i} x_{n-j} x_{n-k}^* + w_n, \tag{1}$$

where L describes the dispersion of the nonlinear channels, h_l and h_{ijk} represent the Volterra channel coefficients, and w_n is the AWGN having zero mean and variance of σ^2.

The joint *a posteriori* distribution of channel coefficients and information bits can be written as

$$\begin{aligned} p(\mathbf{b}, \mathbf{h}|\mathbf{y}) &\propto p(\mathbf{y}|\mathbf{h}, \mathbf{x}) p(\mathbf{h}) p(\mathbf{x}|\mathbf{c}) p(\mathbf{c}|\mathbf{b}) p(\mathbf{b}) \qquad (2) \\ &\propto \prod_{n=1}^{N} p(y_n|\mathbf{h}, x_n, ..., x_{n-L}) p(x_n|\mathbf{c}_n) \\ &\quad \times \prod_{l=0}^{L} p(h_l) \prod_{i,j,k} p(h_{ijk}) \times p(\mathbf{c}|\mathbf{b}) \times \prod_{n=1}^{N_b} p(b_n), \qquad (3) \end{aligned}$$

where $\mathbf{h} \triangleq [h_0, h_1, ..., h_L, h_{000}, h_{001}, ..., h_{00L}, ..., h_{LLL}]^T$ and $\mathbf{c}_n \triangleq [c_{n,1}, ..., c_{n,M}]^T$ with $c_{n,k} \triangleq c_{(n-1)M+k}$. The *a priori* probability $p(b_n)$ of the information bit is assumed to be uniformly distributed. Furthermore, $p(\mathbf{c}|\mathbf{b})$ and $p(x_n|\mathbf{c}_n)$ denote the deterministic coding function and symbol mapping function, respectively, while $p(h_l)$ and $p(h_{ijk})$ are the *a priori* probabilities of the Volterra channel coefficients h_l and h_{ijk}, respectively. The likelihood function $p(y_n|\mathbf{h}, x_n, ..., x_{n-L})$ can be expressed as

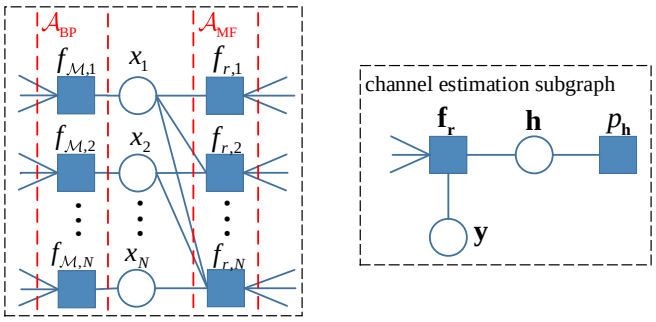
$$p(y_n|\mathbf{h}, x_n, ..., x_{n-L}) \propto \exp\left\{ -\frac{\left| y_n - \sum_{l=0}^{L} h_l x_{n-l} - \sum_{i=0}^{L} \sum_{j\geq i}^{L} \sum_{k=0}^{L} h_{ijk} x_{n-i} x_{n-j} x_{n-k}^* \right|^2}{\sigma^2} \right\}. \tag{4}$$

For the joint *a posteriori* distribution in (3), the corresponding FG representation is shown in Fig. 2, where the squares denote the factor nodes (FN) and the circles denote the variable nodes (VN). In Fig. 2, the FN $f_{\mathcal{M},n}$ corresponding to $p(x_n|\mathbf{c}_n)$ represents symbol mapping, i.e., $f_{\mathcal{M},n}(x_n, \mathbf{c}_n) \triangleq \delta[x_n - \mathcal{M}(\mathbf{c}_n)]$ with $\mathcal{M}$ being the symbol mapping function, and the FN $f_{r,n}$ corresponding to $p(y_n|\mathbf{h}, x_n, ..., x_{n-L})$ stands for the likelihood function.

To obtain the marginal distributions of channel coefficients and information bits, BP may be adopted directly. However, the FG in Fig. 2 contains short loops, which may result in the divergence of BP [29]. Although the convergence-guaranteed MF can be employed, it is not compatible with the hard constraints involved in the coding and modulation functions. To this end, we propose a combined BP-MF method for joint channel estimation and data detection in nonlinear satellite channels.

III. BP-MF for Joint Channel Estimation and Decoding in Nonlinear Channels

This section presents the proposed message passing algorithm conceived for jointly estimating the nonlinear channel's coefficients and detecting the transmitted information. Given the associated hard constraints and the FG's loopy nature, we split the FG into two regions, i.e., the BP region and the MF region. For the sake of reducing the computational complexity, a series of parametric message updating expressions are derived to obtain the estimates of data symbols. By introducing



(a) (b)

Fig. 3. Factor graphs to illustrate (a) the factor nodes classified into BP region and MF region and (b) the non-loopy subgraph with vector-valued nodes for channel estimation using the MF approximation.

the vector-valued nodes for eliminating the dense loops of the FG, a loop-free subgraph is constructed for performing channel estimation. Furthermore, the proposed algorithm is also extended to the case of realistic unknown noise variance.

A. The Proposed Joint Channel Estimation and Detection Algorithm Based on the BP-MF Method

Let us now define the BP region $\mathcal{A}_{\text{BP}}$ and the MF region $\mathcal{A}_{\text{MF}}$, where we have $\mathcal{A}_{\text{BP}} \triangleq \{f_{\mathcal{M},n} | n = 1, 2, ..., N\}$ and $\mathcal{A}_{\text{MF}} \triangleq \{f_{r,n} | n = 1, 2, ..., N\}$, as shown in Fig. 3(a). Then, we can derive the *a posteriori* beliefs of data symbols and channel coefficients using the BP-MF method.

First, we consider the message update related to the FNs in the BP region. According to [30], the message emerging from $f_{\mathcal{M},n}$ to x_n is updated by

$$\mu_{f_{\mathcal{M},n}\to x_n}(x_n) = \sum_{\mathbf{c}_n \in \{0,1\}^M} f_{\mathcal{M},n}(x_n, \mathbf{c}_n) \prod_{k=1}^{M} \mu_{c_{n,k}\to f_{\mathcal{M},n}}(c_{n,k}). \tag{5}$$

Since x_n is a discrete random variable, the message in (5) is the probability mass function (PMF) of x_n. Thus, the message $\mu_{f_{\mathcal{M},n}\to x_n}(x_n)$ has the form

$$\mu_{f_{\mathcal{M},n}\to x_n}(x_n) = \sum_{j=1}^{2^M} P(x_n = s^{(j)}) \delta(x_n - s^{(j)}), \tag{6}$$

where $s^{(j)}$ is the j-th symbol in χ.

Then, the message passed from x_n to $f_{r,i}$ is updated by

$$\mu_{x_n\to f_{r,i}}(x_n) = \mu_{f_{\mathcal{M},n}\to x_n}(x_n) \prod_{k=n}^{n+L} \mu_{f_{r,k}\to x_n}(x_n). \tag{7}$$

In order to obtain $\mu_{x_n\to f_{r,i}}(x_n)$ in (7), we have to derive the expression of $\mu_{f_{r,k}\to x_n}(x_n)$ for $k = n, ..., n + L$.

According to the MF rules, the message passed from the likelihood node $f_{r,n}$ to the symbol node x_n is updated by

$$\mu_{f_{r,n}\to x_n}(x_n) \propto \exp\Bigg\{ \sum_{\sim\{x_n\}} \int_{\mathbf{h}} \ln f_{r,n}(\mathbf{h}, x_n, ..., x_{n-L}) \times \prod_{x'_n \in N(f_{r,n})\backslash x_n} \mu_{x'_n\to f_{r,n}}(x'_n) \prod_m \mu_{h_m\to f_{r,n}}(h_m) \Bigg\}. \tag{8}$$

Substituting the expressions of $f_{r,n}$ in (4) into (8), $\mu_{f_{r,n}\to x_n}(x_n)$ can be rewritten as

$$\mu_{f_{r,n}\to x_n}(x_n) \propto \exp\Bigg\{ \mathbb{E}_{x'_n \in N(f_{r,n})\backslash x_n, h_m} \Big[-\frac{1}{\sigma^2} | y_n - \sum_{l=0}^{L} h_l x_{n-l} - \sum_{i=0}^{L} \sum_{j\geq i}^{L} \sum_{k=0}^{L} h_{ijk} x_{n-i} x_{n-j} x^*_{n-k} |^2 \Big] \Bigg\}. \tag{9}$$

Through the detailed derivations in **Appendix** A, we have

$$\mu_{f_{r,n}\to x_n}(x_n) \propto \exp\Bigg\{ - \frac{\sum_{a,b} \varphi^{(n)}_{a,b} x^a_n x^{*b}_n + 2\Re(\sum_{m,v} \varphi^{(n)}_{m,v} x^m_n x^{*v}_n)}{\sigma^2} \Bigg\}, \tag{10}$$

where $a, b, m, v \in \{1, 2, 3\}$, $\varphi^{(n)}_{a,b}$ and $\varphi^{(n)}_{m,v}$ denote the canonical parameters corresponding to the sufficient statistics $x^a_n x^{*b}_n$ and $x^m_n x^{*v}_n$. Note that the message in (10) is non-Gaussian distribution, which is different from the Gaussian distribution in linear channels. Hence, for a third-order Volterra series associated with memory length L, the corresponding canonical parameters $\varphi^{(n)}_{a,b}$ and $\varphi^{(n)}_{m,v}$ are computed according to Table II[1], where the detailed derivations of $\varphi^{(n)}_{a,b}$ and $\varphi^{(n)}_{m,v}$ are given in **Appendix** B. Note that only the canonical parameters have to be updated in each iteration, which significantly simplifies the updating of messages. Upon substituting (6) and (10) into (7), we have

$$\mu_{x_n\to f_{r,i}}(x_n) \propto \sum_{j=1}^{2^M} P(x_n = s^{(j)}) \delta(x_n - s^{(j)}) \times \exp\Bigg\{ - \frac{\sum_{k=n}^{n+L} \big[\sum_{a,b} \varphi^{(k)}_{a,b} x^a_n x^{*b}_n + 2\Re(\sum_{m,v} \varphi^{(k)}_{m,v} x^m_n x^{*v}_n) \big]}{\sigma^2} \Bigg\}. \tag{11}$$

Since we have $f_{r,i} \in \mathcal{A}_{\text{MF}}$, the belief of x_n equals to the message $\mu_{x_n\to f_{r,i}}(x_n)$. Thus, we can obtain the *a posteriori* mean m_{x_n} and variance v_{x_n} of x_n as

$$m_{x_n} = \frac{\sum_{j=1}^{2^M} s^{(j)} \mu_{x_n\to f_{r,i}}(x_n = s^{(j)})}{\sum_{j=1}^{2^M} \mu_{x_n\to f_{r,i}}(x_n = s^{(j)})}, \tag{12}$$

$$v_{x_n} = \frac{\sum_{j=1}^{2^M} |s^{(j)}|^2 \mu_{x_n\to f_{r,i}}(x_n = s^{(j)})}{\sum_{j=1}^{2^M} \mu_{x_n\to f_{r,i}}(x_n = s^{(j)})} - |m_{x_n}|^2. \tag{13}$$

[1]The term requiring the real part operation in (10) is simplified as $\Re(x^m_n x^{*v}_n)$ in Table II.

TABLE II
PARAMETERS $\varphi_{a,b}^{(n)}$ AND $\varphi_{m,v}^{(n)}$ FOR A THIRD-ORDER VOLTERRA CHANNEL WITH MEMORY LENGTH L.

sufficient statistics	canonical parameters
$\|x_n\|^2(a=b=1)$	$\mathbb{E}(\|h_0\|^2)-2\Re(y_n^*\sum_{i=1}^{L}\mathbb{E}(h_{0i0})\mathbb{E}(x_{n-i}))+2\Re(\sum_{i,j=1}^{L}\mathbb{E}(h_0^*h_{0ij}+h_ih_{0j0}^*)\mathbb{E}(x_{n-i}x_{n-j}^*))+\sum_{i,j=1}^{L}\mathbb{E}(\|h_{0ij}\|^2)\mathbb{E}(\|x_{n-i}\|^2\|x_{n-j}\|^2)$ $+2\Re(\sum_{i,j,p>i,q>j}^{L}\mathbb{E}(h_{0ij}^*h_{0pq})\mathbb{E}(x_{n-i}^*x_{n-j}x_{n-p}x_{n-q}^*))+\sum_{i,j\geq i}^{L}\mathbb{E}(\|h_{ij0}\|^2)\mathbb{E}(\|x_{n-i}\|^2\|x_{n-j}\|^2)$ $+2\Re(\sum_{i,j,p>i,q>j}^{L}\mathbb{E}(h_{ij0}^*h_{pq0})\mathbb{E}(x_{n-i}^*x_{n-j}^*x_{n-p}x_{n-q}))+2\Re(\sum_{i,j,p\geq j,q}\mathbb{E}(h_{0i0}^*h_{jpq})\mathbb{E}(x_{n-i}^*x_{n-j}x_{n-p}x_{n-q}^*))$
$\|x_n\|^4(a=b=2)$	$2\Re(\mathbb{E}(h_{000}^*)(\mathbb{E}(h_0)+\sum_{i,j=1}^{L}\mathbb{E}(h_{0ij})\mathbb{E}(x_{n-i}x_{n-j}^*)))+\sum_{i=1}^{L}\mathbb{E}(\|h_{00i}\|^2)\mathbb{E}(\|x_{n-i}\|^2)$ $+\sum_{i=1}^{L}\mathbb{E}(\|h_{0i0}\|^2)\mathbb{E}(\|x_{n-i}\|^2)+2\Re(\sum_{i,j>i}\mathbb{E}(h_{00i}^*h_{00j}+h_{0i0}h_{0j0}^*)\mathbb{E}(x_{n-i}x_{n-j}^*))$
$\|x_n\|^6(a=b=3)$	$\mathbb{E}(\|h_{000}\|^2)$
$\Re(x_n)(m=1,v=0)$	$-y_n^*\mathbb{E}(h_0)-y_n^*\sum_{i,j=1}^{L}\mathbb{E}(h_{0ij})\mathbb{E}(x_{n-i}x_{n-j}^*)+\mathbb{E}(h_0)\sum_{j=1}^{L}\mathbb{E}(h_j^*)\mathbb{E}(x_{n-j}^*)-y_n\sum_{i,j}^{L}\mathbb{E}(h_{ij0}^*)\mathbb{E}(x_{n-i}^*x_{n-j}^*)$ $+\mathbb{E}(h_0)\sum_{i,j\geq i,k}\mathbb{E}(h_{ijk}^*)\mathbb{E}(x_{n-i}^*x_{n-j}^*x_{n-k})+\sum_{i,j,k}^{L}\mathbb{E}(h_i^*h_{0jk})\mathbb{E}(x_{n-i}^*x_{n-j}x_{n-k}^*)+\sum_{i,j,k}^{L}\mathbb{E}(h_ih_{jk0}^*)\mathbb{E}(x_{n-i}x_{n-j}^*x_{n-k}^*)$ $+\sum_{i,j,k,p,q}\mathbb{E}(h_{0ij}h_{kpq}^*)\mathbb{E}(x_{n-i}x_{n-j}^*x_{n-k}^*x_{n-p}^*x_{n-q})+\sum_{i,j,k,p,q}\mathbb{E}(h_{ij0}^*h_{kpq})\mathbb{E}(x_{n-i}^*x_{n-j}^*x_{n-k}x_{n-p}x_{n-q}^*)$
$\Re(x_n^2)(m=2,v=0)$	$-y_n^*\sum_{i=1}^{L}\mathbb{E}(h_{00i})\mathbb{E}(x_{n-i}^*)+\mathbb{E}(h_0)\sum_{i,j\geq i}^{L}\mathbb{E}(h_{ij0}^*)\mathbb{E}(x_{n-i}^*x_{n-j}^*)+\sum_{i,j=1}^{L}\mathbb{E}(h_i^*h_{00j})\mathbb{E}(x_{n-i}^*x_{n-j}^*)$ $+\sum_{i,j,p,q\geq p}\mathbb{E}(h_{0ij}h_{pq0}^*)\mathbb{E}(x_{n-i}x_{n-j}^*x_{n-p}^*x_{n-q}^*)+\sum_{i,j,p\geq j,q}\mathbb{E}(h_{00i}h_{jpq}^*)\mathbb{E}(x_{n-i}^*x_{n-j}^*x_{n-p}^*x_{n-q})$
$\Re(x_n^3)(m=3,v=0)$	$\sum_{i,j,k\geq j}\mathbb{E}(h_{00i}h_{jk0}^*)\mathbb{E}(x_{n-i}^*x_{n-j}^*x_{n-k}^*)$
$\Re(x_n\|x_n\|^2)(m=2,v=1)$	$-y_n^*\mathbb{E}(h_{000})+\sum_{i=1}^{L}\mathbb{E}(h_0^*h_{00i}+h_0h_{0i0}^*+h_i^*h_{000})\mathbb{E}(x_{n-i}^*)+\sum_{i,j,k=1}^{L}\mathbb{E}(h_{00i}h_{0jk}^*)\mathbb{E}(x_{n-i}^*x_{n-j}^*x_{n-k})$ $+\sum_{i,j,k=1}^{L}\mathbb{E}(h_{0i0}^*h_{0jk})\mathbb{E}(x_{n-i}^*x_{n-j}x_{n-k}^*)+\sum_{i,j,k\geq j}^{L}\mathbb{E}(h_{0i0}h_{jk0}^*)\mathbb{E}(x_{n-i}x_{n-j}^*x_{n-k}^*)+\sum_{i,j\geq i,k}\mathbb{E}(h_{000}h_{ijk}^*)\mathbb{E}(x_{n-i}^*x_{n-j}^*x_{n-k})$
$\Re(x_n\|x_n\|^4)(m=3,v=2)$	$\sum_{i=1}^{L}\mathbb{E}(h_{000}^*h_{00i})\mathbb{E}(x_{n-i}^*)+\sum_{i=1}^{L}\mathbb{E}(h_{000}h_{0i0}^*)\mathbb{E}(x_{n-i}^*)$
$\Re(x_n^2\|x_n\|^2)(m=3,v=1)$	$\sum_{i,j\geq i}\mathbb{E}(h_{000}h_{ij0}^*)\mathbb{E}(x_{n-i}^*x_{n-j}^*)+\sum_{i,j}^{L}\mathbb{E}(h_{00i}h_{0j0}^*)\mathbb{E}(x_{n-i}^*x_{n-j}^*)$

Similarly, we can derive the messages from the likelihood function to the channel coefficient according to the MF rules. To facilitate the calculation of messages, the received signal in (1) is reformulated into the following compact form

$$\mathbf{y}=\mathbf{X}\mathbf{h}+\mathbf{w}, \tag{14}$$

where $\mathbf{w}\triangleq[w_1,w_2,\cdots,w_N]^T$, and

$$\mathbf{X}=\begin{bmatrix}(\mathbf{x}_1^L)^T & (\mathbf{x}_1^{NL})^T\\(\mathbf{x}_2^L)^T & (\mathbf{x}_2^{NL})^T\\ \vdots & \vdots\\(\mathbf{x}_N^L)^T & (\mathbf{x}_N^{NL})^T\end{bmatrix}, \tag{15}$$

with $\mathbf{x}_n^L\triangleq[x_n,x_{n-1},...,x_{n-L}]^T$, and $\mathbf{x}_n^{NL}\triangleq[x_nx_nx_n^*,x_nx_nx_{n-1}^*,...,x_nx_nx_{n-L}^*,...,x_{n-L}x_{n-L}x_{n-L}^*]^T$. In this way, the message passing can be performed in the non-loopy subgraph with vector-valued nodes, as shown in Fig. 3(b). Accordingly, the message passed from the likelihood function $\mathbf{f_r}$ to the channel coefficient $\mathbf{h}$ can be updated by

$$\mu_{\mathbf{f_r}\to\mathbf{h}}\propto\exp\{\mathbb{E}_{\mu_{\mathbf{x}\to\mathbf{f_r}}}[\ln p(\mathbf{y}|\mathbf{h},\mathbf{x})]\}. \tag{16}$$

Since $w_n\sim\mathcal{CN}(w_n;0,\sigma^2)$, $n=1,2,\cdots,N$, we have

$$p(\mathbf{y}|\mathbf{h},\mathbf{x})\propto\exp\Big\{-\frac{(\mathbf{y}-\mathbf{X}\mathbf{h})^H(\mathbf{y}-\mathbf{X}\mathbf{h})}{\sigma^2}\Big\}. \tag{17}$$

Substituting (17) into (16) yields

$$\begin{aligned}&\mu_{\mathbf{f_r}\to\mathbf{h}}\\&\propto\exp\Big\{\mathbb{E}_{\mu_{\mathbf{x}\to\mathbf{f_r}}}\big[-\frac{(\mathbf{y}-\mathbf{X}\mathbf{h})^H(\mathbf{y}-\mathbf{X}\mathbf{h})}{\sigma^2}\big]\Big\}\\&\propto\exp\Big\{-\frac{\mathbf{y}^H\mathbf{y}-\mathbf{y}^H\mathbb{E}[\mathbf{X}]\mathbf{h}-\mathbf{h}^H\mathbb{E}[\mathbf{X}^H]\mathbf{y}+\mathbf{h}^H\mathbb{E}[\mathbf{X}^H\mathbf{X}]\mathbf{h}}{\sigma^2}\Big\}\\&\propto\exp\Big\{-(\mathbf{h}-\check{\mathbf{m}}_\mathbf{h})^H\check{\mathbf{C}}_\mathbf{h}^{-1}(\mathbf{h}-\check{\mathbf{m}}_\mathbf{h})\Big\}\\&\propto\mathcal{CN}(\mathbf{h};\check{\mathbf{m}}_\mathbf{h},\check{\mathbf{C}}_\mathbf{h}),\end{aligned} \tag{18}$$

where the mean vector obeys $\check{\mathbf{m}}_\mathbf{h}=\frac{1}{\sigma^2}\check{\mathbf{C}}_\mathbf{h}\mathbb{E}[\mathbf{X}]^H\mathbf{y}$ and the covariance matrix is given by $\check{\mathbf{C}}_\mathbf{h}^{-1}=\frac{1}{\sigma^2}\big[\mathbb{E}(\mathbf{X}^H\mathbf{X})\big]$. With the estimated *a posteriori* mean and variance of x_n in (12) and (13), the mean vector $\check{\mathbf{m}}_\mathbf{h}$ and the covariance matrix $\check{\mathbf{C}}_\mathbf{h}$ can be computed straightforwardly. Assuming that the *a priori* probability $p_\mathbf{h}$ obeys a multivariate complex Gaussian pdf, i.e., $p_\mathbf{h}\propto\mathcal{CN}(\mathbf{h};\mathbf{h}^0,\mathbf{V}_\mathbf{h}^0)$, the *a posteriori* pdf of channel coefficients $\mathbf{h}$ is also a multivariate complex Gaussian pdf $\mathcal{CN}(\mathbf{h};\mathbf{m}_\mathbf{h},\mathbf{V}_\mathbf{h})$ with

$$\mathbf{m}_\mathbf{h}=\mathbf{V}_\mathbf{h}\big[(\mathbf{V}_\mathbf{h}^0)^{-1}\mathbf{h}^0+\check{\mathbf{C}}_\mathbf{h}^{-1}\check{\mathbf{m}}_\mathbf{h}\big], \tag{19}$$

$$\mathbf{V}_\mathbf{h}=\big[(\mathbf{V}_\mathbf{h}^0)^{-1}+\check{\mathbf{C}}_\mathbf{h}^{-1}\big]^{-1}. \tag{20}$$

Given (19) and (20), the channel coefficients $\mathbf{h}$ can be jointly updated.

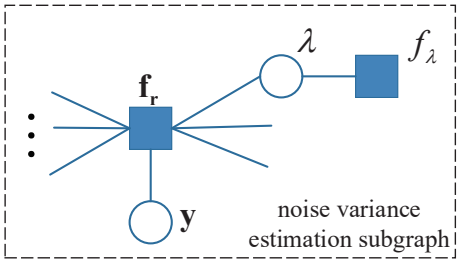


Fig. 4. The subgraph of noise variance estimation.

Finally, the message passed from the mapping function $f_{\mathcal{M},n}$ to the coded bit $c_{n,k}$ can be computed by

$$\mu_{f_{\mathcal{M},n}\to c_{n,k}}(c_{n,k}) \propto \sum_{\sim\{c_{n,k}\}} f_{\mathcal{M},n}(x_n,\mathbf{c}_n)\mu_{x_n\to f_{\mathcal{M},n}}(x_n) \times \prod_{k'\neq k}\mu_{c_{n,k'}\to f_{\mathcal{M},n}}(c_{n,k'}), \tag{21}$$

where the message passed from x_n to $f_{\mathcal{M},n}$ is updated by

$$\mu_{x_n\to f_{\mathcal{M},n}}(x_n) = \prod_{k=n}^{n+L}\mu_{f_{r,k}\to x_n}(x_n) \propto \exp\left\{-\frac{\sum_{k=n}^{n+L}\left[\sum_{a,b}\varphi_{a,b}^{(k)}x_n^a x_n^{*b} + 2\Re(\sum_{m,v}\varphi_{m,v}^{(k)}x_n^m x_n^{*v})\right]}{\sigma^2}\right\}. \tag{22}$$

The message $\mu_{f_{\mathcal{M},n}\to c_{n,k}}(c_{n,k})$ can be viewed as the extrinsic soft information passed to a SISO decoder and it is treated as the *a priori* information used for decoding. Once the SISO decoding is terminated, it will produce extrinsic soft information, which is used as the *a priori* information in (5) to determine the *a priori* probabilities of the data symbol at the next iteration. Upon using the message $\mu_{f_{\mathcal{M},n}\to c_{n,k}}(c_{n,k})$, the log-likelihood ratio (LLR) of coded bits can be computed by

$$LLR(c_{n,k}) = \ln\frac{\mu_{f_{\mathcal{M},n}\to c_{n,k}}(c_{n,k}=0)}{\mu_{f_{\mathcal{M},n}\to c_{n,k}}(c_{n,k}=1)}. \tag{23}$$

The details of the proposed combined BP-MF method of joint channel estimation and decoding are summarized in **Algorithm** 1.

B. Extension to the Case of Realistic Unknown Noise Variance

The noise variance is always unknown and has to be estimated in practical wireless communication systems. In this section, we now harness the MF for estimating the noise precision λ, which is the inverse of noise variance, i.e., $\lambda = 1/\sigma^2$. In this case, the joint *a posteriori* distribution in (2) is represented by $p(\mathbf{b},\mathbf{h},\lambda|\mathbf{y})$, the likelihood function $p(\mathbf{y}|\mathbf{h},\mathbf{x})$ becomes $p(\mathbf{y}|\mathbf{h},\mathbf{x},\lambda)$, and the *a priori* distribution $p(\lambda)$ is added to the products in (2) and (3). On the FG of Fig. 4, the additional VN λ and FN f_λ denoting the *a priori* distribution of λ are introduced. Then, according to the

Algorithm 1 The Proposed Combined BP-MF Method for Joint Channel Estimation and Decoding

1: Initialize $\mu_{c_{n,k}\to f_{\mathcal{M},n}}(c_{n,k}) = 1/2$ and the *a priori* distribution of channel coefficients is obtained using pilots based on LMMSE;
2: **for** $i_o = 1$ to I_{out} **do**
3: Calculate the messages from the mapping node to the symbol node according to (5);
4: **for** $i_t = 1$ to I_{in} **do**
5: Calculate the canonical parameters $\varphi_{a,b}^{(n)}$ and $\varphi_{m,v}^{(n)}$ according to Table II;
6: Calculate the messages from the likelihood node to the symbol node according to (10);
7: Calculate the belief of variable x_n according to (11);
8: Calculate the *a posteriori* mean and variance $\{m_{x_n}, v_{x_n}\}$ according to (12) and (13);
9: **end for**
10: Calculate the *a posteriori* mean vector and covariance matrix $\{\mathbf{m_h}, \mathbf{V_h}\}$ according to (19) and (20);
11: Calculate the messages from the mapping node to the coded bit node according to (21);
12: Calculate the extrinsic LLRs of the equalizer according to (23) and pass them to the SISO decoder;
13: Perform channel decoding algorithm and the decoder produces the extrinsic soft information to equalizer;
14: **end for**

MF rules, the message from the likelihood node $\mathbf{f_r}$ to noise precision node λ can be updated by

$$\mu_{\mathbf{f_r}\to\lambda}(\lambda) \propto \exp\left\{\int \ln p(\mathbf{y}|\mathbf{h},\mathbf{x},\lambda)\mu_{\mathbf{x}\to\mathbf{f_r}}(\mathbf{x})\mu_{\mathbf{h}\to\mathbf{f_r}}(\mathbf{h})d\mathbf{x}d\mathbf{h}\right\}, \tag{24}$$

where

$$p(\mathbf{y}|\mathbf{h},\mathbf{x},\lambda) \propto \lambda^N \exp\{-\lambda\|\mathbf{y}-\mathbf{X}\mathbf{h}\|^2\}. \tag{25}$$

Upon substituting (25) into (24), we arrive at

$$\begin{aligned}\mu_{\mathbf{f_r}\to\lambda}(\lambda) &\propto \exp\left\{N\ln\lambda - \lambda\mathbb{E}[\|\mathbf{y}-\mathbf{X}\mathbf{h}\|^2]\right\}\\ &\propto \lambda^N\exp\left\{-\lambda\mathbb{E}[\|\mathbf{y}-\mathbf{X}\mathbf{h}\|^2]\right\}\\ &\propto \lambda^N\exp\left\{-\lambda C_\lambda\right\},\end{aligned} \tag{26}$$

where $C_\lambda = \mathbb{E}[\|\mathbf{y}-\mathbf{X}\mathbf{h}\|^2] = \mathbf{y}^H\mathbf{y} - \mathbf{y}^H\mathbb{E}[\mathbf{X}]\mathbb{E}[\mathbf{h}] - \mathbb{E}[\mathbf{h}]^H\mathbb{E}[\mathbf{X}]^H\mathbf{y} + \text{Tr}\left\{\mathbb{E}[\mathbf{X}^H\mathbf{X}]\mathbb{E}[\mathbf{h}\mathbf{h}^H]\right\}$. The message in (26) is proportional to a one-dimensional complex Wishart distribution with the degree of freedom $N+1$, and covariance of C_λ^{-1}. In this case, we select a Gamma distribution as the conjugate *a priori* pdf [31], i.e., $p(\lambda) \propto \lambda^{a_p-1}\exp\{-\lambda A_p\}$. Due to the fact that the *a priori* information of the noise precision is usually unavailable at the receiver side, we select a non-informative *a priori* pdf $p(\lambda)$ associated with $a_p = 0$ and $A_p = 0$, i.e., $p(\lambda) \propto 1/\lambda$. Given $\mu_{\mathbf{f_r}\to\lambda}(\lambda)$ in (26), we obtain the belief

$$b_\lambda(\lambda) \propto \lambda^{N-1}\exp\{-\lambda C_\lambda\}. \tag{27}$$

With the belief of λ in (27), the noise precision can be calculated as

$$\hat{\lambda} = \frac{N}{C_\lambda}. \tag{28}$$

Note that for the update of data symbols $\mathbf{x}$ and channel coefficients $\mathbf{h}$, we only have to replace σ^2 by $1/\hat{\lambda}$.

IV. GAMP-Aided Complexity Reduction of BP-MF for Joint Channel Estimation and Decoding

Although the proposed BP-MF algorithm is capable of accurately estimating both channel coefficients and transmitted symbols, it still suffers from a high computational complexity on the order of $\mathcal{O}(L_h^3)$ resulting from the matrix inversion operation of (20) involved in channel estimation, where L_h denotes the number of channel coefficients. This computational complexity could be a challenging task for communication systems having a large channel dispersion. To circumvent the matrix inversion, we further resort to the computationally efficient GAMP algorithm [27] to develop a low-complexity channel estimation method. With the aid of the central limit theorem and Taylor series expansion, the GAMP decouples the vector-valued channel estimation problem into a sequence of scalar operations, and its asymptotic behavior under i.i.d. transforms can be exactly described by a simple set of state evolution equations.

To commence, we replace $\mathbf{X}$ in (14) by $\hat{\mathbf{X}}$, where $\hat{\mathbf{X}}$ is the result of replacing all the $\{x_n\}$ in $\mathbf{X}$ by its *a posteriori* mean m_{x_n}. Then (14) becomes

$$\mathbf{y} = \hat{\mathbf{X}}\mathbf{h} + \mathbf{w}. \tag{29}$$

Let us now define $z_n \triangleq \hat{\mathbf{X}}_n\mathbf{h}$, where $\hat{\mathbf{X}}_n$ denotes the nth row of matrix $\hat{\mathbf{X}}$, and the corresponding likelihood function is given by $p(y_n|z_n) = \mathcal{CN}(z_n; y_n, \sigma^2)$. Then, the approximate marginal distribution of z_n is expressed as

$$p(z_n|\mathbf{y}; \hat{p}_n, \gamma_n^p) \triangleq \frac{p(y_n|z_n)\mathcal{CN}(z_n; \hat{p}_n, \gamma_n^p)}{\int_{z_n'} p(y_n|z_n')\mathcal{CN}(z_n'; \hat{p}_n, \gamma_n^p)}, \tag{30}$$

where the quantities $\hat{p}_n(t)$ and $\gamma_n^p(t)$ at the t-th iteration are updated by

$$\hat{p}_n(t) = \sum_m \hat{X}_{nm}\hat{h}_m(t) - \gamma_n^p(t)\hat{s}_n(t-1), \tag{31}$$

$$\gamma_n^p(t) = \sum_m |\hat{X}_{nm}|^2\gamma_m^h(t). \tag{32}$$

At the first iteration, $\hat{h}_m(t)$ and $\gamma_m^h(t)$ are initialized based on the pilot-aided LMMSE estimator. The parameter $\hat{s}_n(t-1)$ related to the *a posteriori* mean of z_n is set to 0. Then, they can be updated by exploiting the equation in (31)-(41). Upon using (30), the *a posteriori* mean and variance of z_n are

$$\begin{aligned}\hat{z}_n(t) &= \mathbb{E}\{z_n|\mathbf{y}; \hat{p}_n(t), \gamma_n^p(t)\} \\ &= \hat{p}_n(t) + \frac{\gamma_n^p(t)}{\gamma_n^p(t)+\sigma^2}(y_n - \hat{p}_n(t)),\end{aligned} \tag{33}$$

$$\gamma_n^z(t) = \mathbb{V}\{z_n|\mathbf{y}; \hat{p}_n(t), \gamma_n^p(t)\} = \frac{\gamma_n^p(t)\sigma^2}{\gamma_n^p(t)+\sigma^2}. \tag{34}$$

Based on the GAMP rules, the output scalar estimation function $g_{\text{out}}(\cdot)$ is calculated by

$$\begin{aligned}g_{\text{out}}[\hat{p}_n(t), y_n, \gamma_n^p(t)] &= \frac{1}{\gamma_n^p(t)}\left[\mathbb{E}\{z_n|\mathbf{y}; \hat{p}_n(t), \gamma_n^p(t)\} - \hat{p}_n(t)\right] \\ &= \frac{y_n - \hat{p}_n(t)}{\gamma_n^p(t)+\sigma^2},\end{aligned} \tag{35}$$

and $\hat{s}_n(t) \triangleq g_{\text{out}}[\hat{p}_n(t), y_n, \gamma_n^p(t)]$. The negative partial derivative of $g_{\text{out}}(\cdot)$ becomes

$$\gamma_n^s(t) = -\frac{\partial}{\partial \hat{p}_n(t)} g_{\text{out}}[\hat{p}_n(t), y_n, \gamma_n^p(t)] = \frac{1}{\gamma_n^p(t)+\sigma^2}. \tag{36}$$

Similarly, the approximate marginal distribution of h_m is updated by

$$p(h_m|\mathbf{y}; \hat{r}_m, \gamma_m^r) \triangleq \frac{p(h_m)\mathcal{CN}(h_m; \hat{r}_m, \gamma_m^r)}{\int_{h_m'} p(h_m')\mathcal{CN}(h_m'; \hat{r}_m, \gamma_m^r)}, \tag{37}$$

where the quantities $\hat{r}_m(t)$ and $\gamma_m^r(t)$ at the t-th iteration can be formulated as

$$\hat{r}_m(t) = \hat{h}_m(t) + \gamma_m^r(t)\sum_n \hat{X}_{nm}^*\hat{s}_n(t), \tag{38}$$

$$\gamma_m^r(t) = \Big[\sum_n |\hat{X}_{nm}|^2\gamma_n^s(t)\Big]^{-1}. \tag{39}$$

The input scalar estimation function $g_{\text{in}}(\cdot)$ can be written as

$$\begin{aligned}g_{\text{in}}[\hat{r}_m(t), \gamma_m^r(t)] &= \mathbb{E}\{h_m|\mathbf{y}; \hat{r}_m(t), \gamma_m^r(t)\} \\ &= \frac{v_{h_m}^0\,\hat{r}_m(t) + \gamma_m^r(t)h_m^0}{v_{h_m}^0 + \gamma_m^r(t)},\end{aligned} \tag{40}$$

and $\hat{h}_m(t+1) = g_{\text{in}}[\hat{r}_m(t), \gamma_m^r(t)]$. Meanwhile, the partial derivative of $g_{\text{in}}(\cdot)$ satisfies

$$\begin{aligned}\gamma_m^h(t+1) &= \gamma_m^r(t)\frac{\partial}{\partial \hat{r}_m(t)} g_{\text{in}}[\hat{r}_m(t), \gamma_m^r(t)] \\ &= \mathbb{V}\{h_m|\mathbf{y}; \hat{r}_m(t), \gamma_m^r(t)\} = \frac{\nu_{h_m}^0\,\gamma_m^r(t)}{\nu_{h_m}^0 + \gamma_m^r(t)}.\end{aligned} \tag{41}$$

Finally, the outputs $\{\hat{h}_m(t+1)\}$ and $\{\gamma_m^h(t+1)\}$ are the mean and variance of the approximate marginal distribution of channel coefficients, respectively, which are invoked in the next iteration. The GAMP algorithm proposed for updating the estimates of channel coefficients can be embedded into the iteration between the equalizer and the decoder. Note that the complexity of the GAMP algorithm is dominated by the matrix-vector multiplications in (31), (32), (38), and (39), and hence its complexity is proportional to L_h. The proposed joint channel estimation and decoding algorithm based on the GAMP-aided BP-MF method is referred to as BP-MF-GAMP, which is summarized in Algorithm 2. Though the GAMP method is computationally efficient, it may diverge in the case of ill-conditioned or correlated matrices. To this end, a series of pioneering work have emerged for mitigating the convergence problem of GAMP [32]–[34]. Specifically, the memory approximate message passing algorithm proposed in [35] can circumvent the strong sensitivity of the ill-conditioned matrices to small perturbations, while it retain a low computational complexity. Note that benefiting from the

TABLE III
VOLTERRA CHANNEL COEFFICIENTS

h_0	h_1	h_2
$0.78085 + 0.41347i$	$0.40323 - 0.0064i$	$-0.15361 - 0.08961i$
h_{000}	h_{001}	h_{002}
$-0.16 - 0.036i$	$-0.14 + 0.14i$	$0.156 + 0.088i$
h_{110}	h_{220}	other h_{ijk}
$0.004 - 0.068i$	$0.072 - 0.072i$	0

convergence-guaranteed property, the proposed BP-MF algorithm with robust convergence is also appropriate for a more general transform matrix. To verify the convergence behavior of the proposed BP-MF and BP-MF-GAMP algorithms in the considered nonlinear channels, we show the convergence curves of the proposed algorithms in the following Section V.

Algorithm 2 The BP-MF-GAMP Method for Joint Channel Estimation and Decoding

1: The same as BP-MF except replacing the step (10) in **Algorithm** 1 with the following steps:
2: **Inputs**: priors $\{p(h_m)\}$, likelihood $\{p(y_n|z_n)\}$, maximum iterations I_{GAMP}, $\{\hat{X}_{nm}\}$, initial channel estimate $\{h_m^0, v_{h_m}^0\}$
3: **Initialization**: $\hat{h}_m(1) = h_m^0$, $\gamma_m^h(1) = v_{h_m}^0$, $\hat{s}_n(0) = 0$
4: **for** $t = 1$ to I_{GAMP} **do**
5: **for** each n, **do**
6: Calculate $\hat{p}_n(t)$ and $\gamma_n^p(t)$ according to (31) and (32);
7: Calculate $\hat{z}_n(t)$ and $\gamma_n^z(t)$ according to (33) and (34);
8: Calculate $\hat{s}_n(t)$ and $\gamma_n^s(t)$ according to (35) and (36);
9: **end for**
10: **for** each m, **do**
11: Calculate $\hat{r}_m(t)$ and $\gamma_m^r(t)$ according to (38) and (39);
12: Calculate the *a posteriori* mean $\hat{h}_m(t+1)$ and variance $\gamma_m^h(t+1)$ according to (40) and (41);
13: **end for**
14: **end for**

V. SIMULATION RESULTS

We consider an LDPC-coded system having the code length of $N_b = 1440$ and code rate of $R = 5/7$. The encoded bits are mapped to QPSK symbols. Moreover, the third-order Volterra series model associated with $L = 2$ is considered in Table III, the corresponding channel coefficients of which are given in [15] but with stronger linear and nonlinear ISI [17]. The initial estimates of channel coefficients are obtained from the pilots using the LMMSE method, where the ratio of pilots is 5%. In the proposed algorithms, the inner iterations within the symbol detection are embedded into the outer iterations between equalization and LDPC decoding. Unless otherwise specified, the maximum number of iterations of the proposed algorithms are set to $I_{\text{out}} = 20$ and $I_{\text{in}} = 2$, respectively, for the outer and inner iterations. Furthermore, the number of GAMP iterations is $I_{\text{GAMP}} = 20$.

The bit error rate (BER) performance of the proposed joint channel estimation and decoding methods is evaluated in Fig. 5. For comparison, the performance of ALS, LMMSE-NL-SIC and LMMSE-BP-MF are also plotted. In the LMMSE-NL-SIC and LMMSE-BP-MF algorithms, LMMSE is used for channel estimation based on pilots, and the NL-SIC as well as the BP-MF are employed for equalization, respectively. Furthermore, the BER performance of the NL-SIC and BP-MF having perfect CSI (PCSI) is considered as a benchmark. It is observed that, for the PCSI scenario, the BP-MF equalizer outperforms the NL-SIC equalizer. This is because the NL-SIC equalizer uses the Gaussian approximation for calculating the extrinsic soft information gleaned from the equalizer, and the third-order Volterra terms are ignored when calculating the variance of the estimation error. For the unknown CSI scenario, the LMMSE-NL-SIC and LMMSE-BP-MF outperform the ALS as a benefit of exploiting soft information. We can observe that, by gleaning information from both pilots and the transmitted symbols recovered in the previous iterations, the proposed BP-MF and BP-MF-GAMP perform close to the PCSI scenario, and they significantly outperform the pilot-based LMMSE-BP-MF method. The performance gap between BP-MF and BP-MF-GAMP arises due to the fact that the former explicitly takes estimation uncertainty of data detection by updating the second-order moments of transmitted symbols for calculating the parameters $\check{\mathbf{m}}_{\mathbf{h}}$ and $\check{\mathbf{C}}_{\mathbf{h}}$. By contrast, the BP-MF-GAMP directly ignores the variance v_{x_n} upon calculating the parameters among the GAMP method.

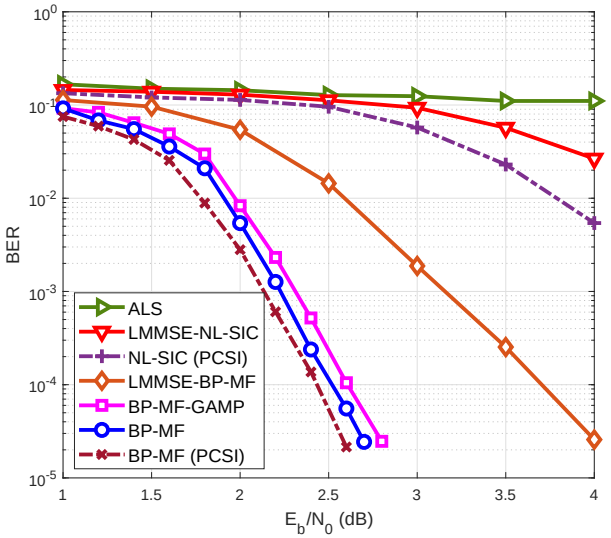


Fig. 5. BER performance of different algorithms.

In Fig. 6, the normalized mean square error (NMSE) of the proposed channel estimation algorithms is evaluated. For comparison, the performance of the ALS and LMMSE methods is also presented. Observe that the NMSE monotonically decreases as the value of E_b/N_0 increases. The proposed BP-MF and BP-MF-GAMP algorithms have a significant NMSE performance gain over the LMMSE and ALS methods. Benefiting from accounting for the symbol estimation uncertainty, the proposed BP-MF algorithm outperforms the BP-MF-GAMP in terms of NMSE performance. Moreover,

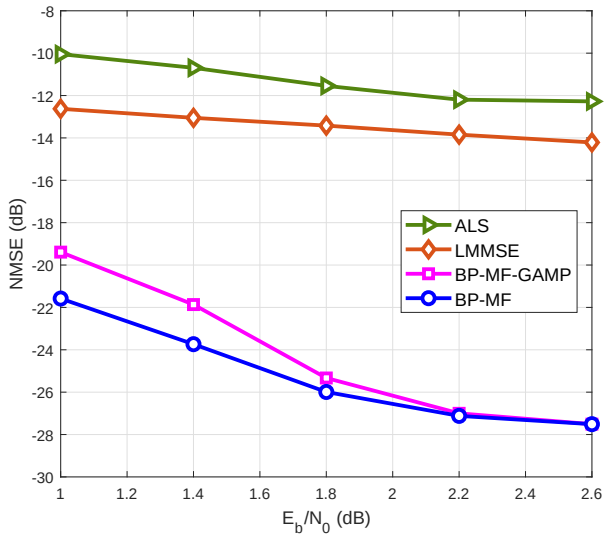


Fig. 6. NMSE performance of channel estimation of different algorithms.

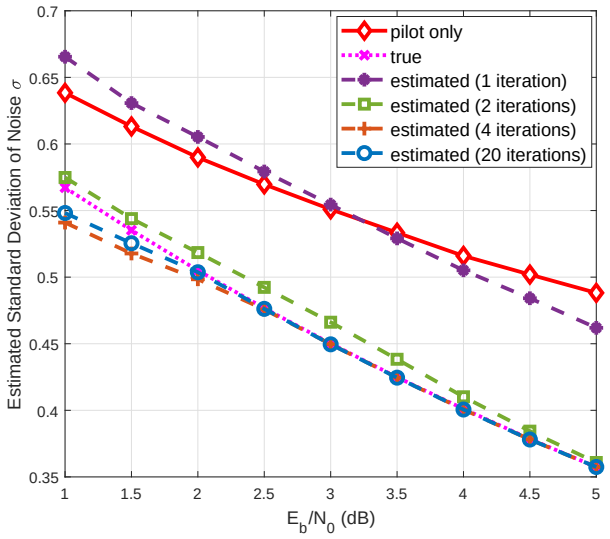


Fig. 7. Estimated average noise variance versus E_b/N_0.

with the increase of SNR, the NMSE performance of BP-MF-GAMP converges to that of the BP-MF method. This is due to the fact that the symbol detection becomes more accurate at high SNRs, and thus the performance loss imposed by ignoring the uncertainties of the estimated symbols in the BP-MF-GAMP method become negligible. Although the proposed BP-MF-GAMP method obtains the similar NMSE performance with the BP-MF at high SNRs, it still suffers from a non-negligible BER performance loss, as shown in Fig. 5. This is because superior channel estimation cannot always guarantee better BER performance.

In Fig. 7, we show the noise variance estimated by the proposed algorithm at different E_b/N_0. Since the BP-MF and the BP-MF-GAMP have similar noise variance estimation characteristics, only the results of BP-MF are given for brevity. As seen in the figure, after several iterations, the proposed noise variance estimator significantly outperforms its pure pilot-based counterpart.

The NMSE convergence behavior of the proposed BP-MF and BP-MF-GAMP algorithms by the SNR is shown in Fig. 8. Observe that both the BP-MF and BP-MF-GAMP can converge fast after about 10 iterations at high SNRs. Moreover, as a benefit of its convergence-guaranteed property as well as due to taking into account the uncertainty of data detection, the BP-MF method converges faster than the BP-MF-GAMP method. The extrinsic information transfer (EXIT) curves of the proposed BP-MF method are illustrated in Fig. 9. As expected, an open tunnel emerges between the curves of the equalizer and channel decoder at $E_b/N_0 = 2.4$dB, which demonstrates that the proposed algorithm is expected to converge.

In Fig. 10, the BER performance of the proposed BP-MF method relying on different numbers of inner and outer iterations is evaluated. It may observe that, increasing I_{out} significantly improves the BER performance. When I_{out} is higher than 16, however, the performance gain attained by

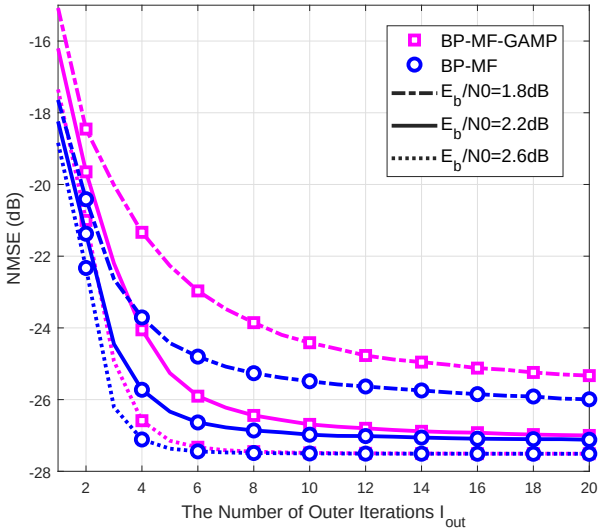


Fig. 8. NMSE convergence behavior of BP-MF and BP-MF-GAMP algorithms.

increasing I_{in} and I_{out} becomes marginal. Hence, the number of inner and outer iterations can be adjusted and optimized to strike a performance versus computational complexity trade-off.

We now further evaluate the BER and NMSE performance of the proposed BP-MF and BP-MF-GAMP algorithms in the face of stronger nonlinearity. The corresponding channel coefficients are given in Table IV. Fig. 11 shows the BER performance of the aforementioned algorithms. Compared to the results in Fig. 5, the stronger nonlinearity degrades the BER performance of all algorithms. Specifically, the ALS, LMMSE-NL-SIC and NL-SIC with PCSI fail in the face of stronger nonlinearity. The proposed BP-MF and BP-MF-GAMP methods perform close to the BP-MF having PCSI, and outperform the LMMSE-BP-MF. Fig. 12 shows the NMSE

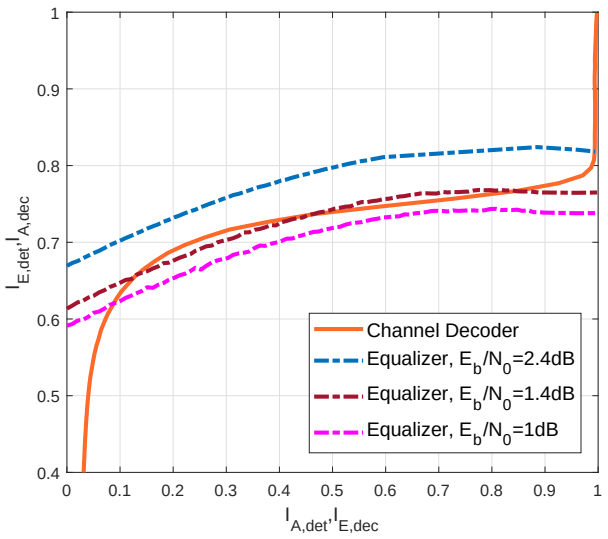


Fig. 9. The EXIT chart for the proposed BP-MF method.

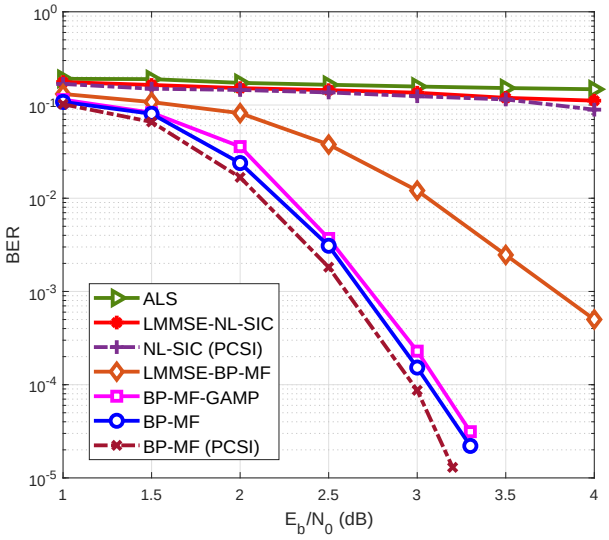


Fig. 11. BER performance of different algorithms in system with stronger nonlinearity.

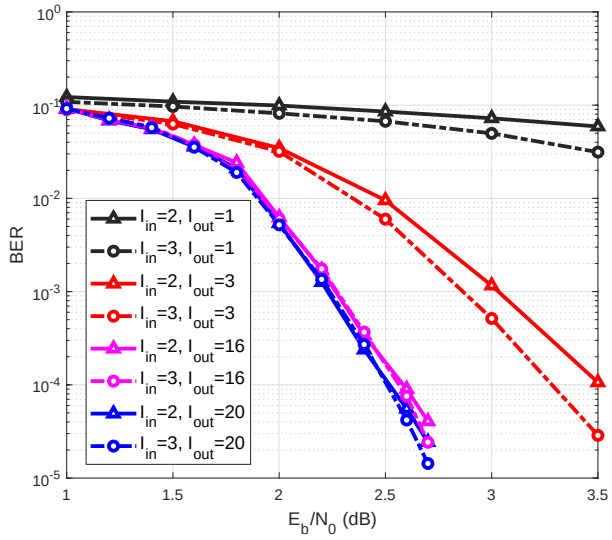


Fig. 10. BER performance of the proposed BP-MF method with different numbers of inner iteration and outer iteration.

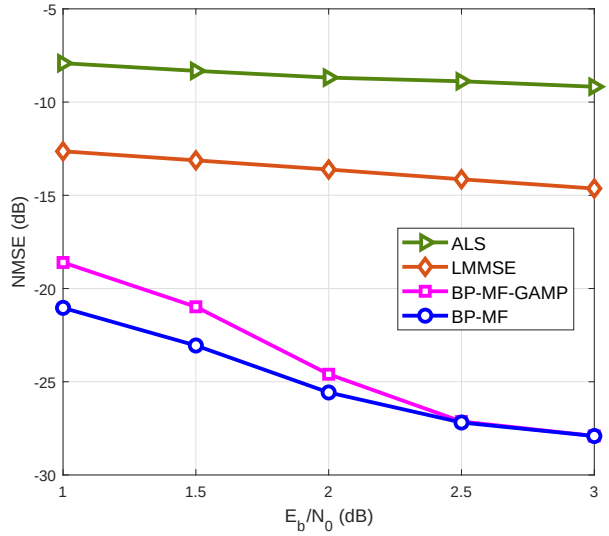


Fig. 12. NMSE performance of channel estimation of different algorithms in system with stronger nonlinearity.

TABLE IV
VOLTERRA CHANNEL COEFFICIENTS WITH STRONGER NONLINEARITY

h_0	h_1	h_2
$0.78085 + 0.41347i$	$0.40323 - 0.0064i$	$-0.15361 - 0.08961i$
h_{000}	h_{001}	h_{002}
$-0.2 - 0.045i$	$-0.175 + 0.175i$	$0.195 + 0.11i$
h_{110}	h_{220}	other h_{ijk}
$-0.005 - 0.085i$	$0.09 - 0.09i$	0

performance of channel estimation at different SNRs. We can see that the channel estimation performance of the proposed BP-MF and BP-MF-GAMP algorithms is not sensitive to the stronger channel nonlinearity, which is a benefit of the superior symbol detection performance. They still significantly outperform the ALS and LMMSE schemes.

VI. CONCLUSIONS

Message passing-aided iterative receivers were designed for joint channel estimation and data detection in nonlinear satellite channels characterized by the Volterra series. To address the problems of hard constraints and dense short loops on factor graph, we beneficially decomposed the factor graph into the BP and MF regions. Specifically, BP was exploited for calculating the messages harnessed from the factor node associated with hard constraints, while MF was employed at the observation node. The calculation of marginal distributions of the data symbols was significantly simplified by performing parametric message passing. In the channel estimation part, we reformulated the system model into a compact form and transformed the densely connected subgraph into a loop-free

subgraph associated with vector-valued nodes. The proposed BP-MF method was also extended to the realistic scenario of unknown noise variance. To avoid the high complexity of matrix inversion in the BP-MF method, we further employed GAMP to transform the vector estimation problem into a scalar one, and proposed a BP-MF-GAMP algorithm. Simulation results showed that the proposed BP-MF and BP-MF-GAMP algorithms achieved superior BER and NMSE performance compared to the state-of-the-art benchmarks. Moreover, our work provides a theoretical reference for solving the challenging joint channel estimation and data detection problem of other emerging nonlinear systems.

APPENDIX A
DERIVATIONS OF (10)

First, we expand the term $|y_n - \sum_{l=0}^{L} h_l x_{n-l} - \sum_{i=0}^{L}\sum_{j\geq i}^{L}\sum_{k=0}^{L} h_{ijk}x_{n-i}x_{n-j}x_{n-k}^*|^2$ as

$$
\begin{aligned}
&|y_n - \sum_{l=0}^{L} h_l x_{n-l} - \sum_{i=0}^{L}\sum_{j\geq i}^{L}\sum_{k=0}^{L} h_{ijk}x_{n-i}x_{n-j}x_{n-k}^*|^2 \\
&= y_n^* y_n - 2\Re(y_n^* \sum_{l=0}^{L} h_l x_{n-l} + y_n^* \sum_{i,j\geq i,k} h_{ijk}x_{n-i}x_{n-j}x_{n-k}^*) \\
&+ \sum_{l=0}^{L} h_l h_l^* x_{n-l}x_{n-l}^* + 2\Re(\sum_{i=0}^{L}\sum_{j>i}^{L} h_i^* h_j x_{n-i}^* x_{n-j}) \\
&+ 2\Re(\sum_{l=0}^{L}\sum_{i,j\geq i,k} h_l^* h_{ijk} x_{n-l}^* x_{n-i}x_{n-j}x_{n-k}^*) \\
&+ \sum_{i,j\geq i,k} h_{ijk}^* h_{ijk} x_{n-i}^* x_{n-j}^* x_{n-k}x_{n-i}x_{n-j}x_{n-k}^* \\
&+ 2\Re(\sum_{i,j,k,a,b,c} h_{ijk}^* h_{abc} x_{n-i}^* x_{n-j}^* x_{n-k}x_{n-a}x_{n-b}x_{n-c}^*).
\end{aligned}
\tag{42}
$$

In (42), $y_n^* y_n$ is a constant, and all the combinations can be split into two parts, namely those with and without real operation. Then, we take the expectation of (42) with respect to the beliefs of all x_i for $i \neq n$ and all h_l and h_{ijk}. Accordingly, these variables can be replaced by their expectations extracted from the *a posteriori* pdfs. After some further manipulations, the message $\mu_{f_{r,n}\to x_n}(x_n)$ can be written in the form of its sufficient statistics and the canonical parameters in (10), where the canonical parameters $\varphi_{a,b}^{(n)}$ and $\varphi_{m,v}^{(n)}$ consist of the combinations of expectations of the related variables.

APPENDIX B
DERIVATIONS OF PARAMETERS $\varphi_{a,b}^{(n)}$ AND $\varphi_{m,v}^{(n)}$

For the derivations of parameters $\varphi_{a,b}^{(n)}$ and $\varphi_{m,v}^{(n)}$, we take $\Re(x_n|x_n|^4)$ as an example. Considering real-valued operations and the form of this term, we focus our attention on the last term $2\Re(\sum_{i,j,k,a,b,c} h_{ijk}^* h_{abc} x_{n-i}^* x_{n-j}^* x_{n-k}x_{n-a}x_{n-b}x_{n-c}^*)$ in (42). Then, we have $2\Re(\sum_{i=1}^{L} h_{000}^* h_{00i} x_n^* x_n^* x_n x_n x_n x_{n-i}^*)$ and $2\Re(\sum_{i=1}^{L} h_{0i0}^* h_{000} x_n^* x_{n-i}^* x_n x_n x_n x_n^*)$. When the expectations are calculated, all variables except for x_n are replaced by their expectations, e.g., $\mathbb{E}(h_{000}^* h_{00i} x_n^* x_n^* x_n x_n x_n x_{n-i}^*) = \mathbb{E}(h_{000}^* h_{00i} x_{n-i}^*) x_n^* x_n^* x_n x_n x_n$. Since the moment of the product of two statistically independent random variables is equal to the product of the two moments of the corresponding independent random variables, we have $\mathbb{E}(h_{000}^* h_{00i} x_{n-i}^*) = \mathbb{E}(h_{000}^*)\mathbb{E}(h_{00i})\mathbb{E}(x_{n-i}^*)$. Accordingly, the parameters $\varphi_{m,v}^{(n)}(m=3, v=2)$ can be obtained. Given the above derivation, all the $\varphi_{a,b}^{(n)}$ and $\varphi_{m,v}^{(n)}$ can be derived.

REFERENCES

[1] Z. Zhang *et al.*, "6G wireless networks: Vision, requirements, architecture, and key technologies," *IEEE Veh. Technol. Mag.*, vol. 14, no. 3, pp. 28–41, Sep. 2019.

[2] B. Di, H. Zhang, L. Song, Y. Li, and G. Y. Li, "Ultra-dense LEO: Integrating terrestrial-satellite networks into 5G and beyond for data offloading," *IEEE Trans. Wireless Commun.*, vol. 18, no. 1, pp. 47–62, Jan. 2019.

[3] A. Guidotti *et al.*, "Architectures and key technical challenges for 5G systems incorporating satellites," *IEEE Trans. Veh. Technol.*, vol. 68, no. 3, pp. 2624–2639, Mar. 2019.

[4] M. De Sanctis, E. Cianca, G. Araniti, I. Bisio, and R. Prasad, "Satellite communications supporting Internet of remote things," *IEEE Internet Things J.*, vol. 3, no. 1, pp. 113–123, Feb. 2016.

[5] K. Kaneko, H. Nishiyama, N. Kato, A. Miura, and M. Toyoshima, "Construction of a flexibility analysis model for flexible high-throughput satellite communication systems with a digital channelizer," *IEEE Trans. Veh. Technol.*, vol. 67, no. 3, pp. 2097–2107, Mar. 2018.

[6] R. Piazza, M. R. B. Shankar, and B. Ottersten, "Data predistortion for multicarrier satellite channels based on direct learning," *IEEE Trans. Signal Process.*, vol. 62, no. 22, pp. 5868–5880, Nov. 2014.

[7] G. Colavolpe, A. Modenini, and F. Rusek, "Channel shortening for nonlinear satellite channels," *IEEE Commun. Lett.*, vol. 16, no. 12, pp. 1929–1932, Dec. 2012.

[8] A. Saleh, "Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers," *IEEE Trans. Commun.*, vol. 29, no. 11, pp. 1715–1720, Nov. 1981.

[9] J. Kim and K. Konstantinou, "Digital predistortion of wideband signals based on power amplifier model with memory," *Electron. Lett.*, vol. 37, no. 23, pp. 1417–1418, Nov. 2001.

[10] D. R. Morgan, Z. Ma, J. Kim, M. G. Zierdt, and J. Pastalan, "A generalized memory polynomial model for digital predistortion of RF power amplifiers," *IEEE Trans. Signal Process.*, vol. 54, no. 10, pp. 3852–3860, Oct. 2006.

[11] S. Benedetto, E. Biglieri, and R. Daffara, "Modeling and performance evaluation of nonlinear satellite links—A Volterra series approach," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-15, no. 4, pp. 494–507, July 1979.

[12] Y.-C. Chen and Y. T. Su, "Turbo equalization of nonlinear TDMA satellite signals," *IEICE Trans. Commun.*, vol. E92.B, no. 3, pp. 992–997, 2009.

[13] D. Ampeliotis, A. A. Rontogiannis, K. Berberidis, M. Papaleo, and G. E. Corazza, "Turbo equalization of non-linear satellite channels using soft interference cancellation," in *Proc. Adv. Satellites Mobile Syst. Conf.*, 2008, pp. 289–292.

[14] B. Benammar, N. Thomas, C. Poulliat, M.-L. Boucheret, and M. Dervin, "On linear MMSE based turbo-equalization of nonlinear Volterra channels," in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, 2013, pp. 4703–4707.

[15] F. M. Kashif, H. Wymeersch, and M. Z. Win, "Monte Carlo equalization for nonlinear dispersive satellite channels," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 2, pp. 245–255, Feb. 2008.

[16] G. Colavolpe and A. Piemontese, "Novel SISO detection algorithms for nonlinear satellite channels," *IEEE Wireless Commun. Lett.*, vol. 1, no. 1, pp. 22–25, Feb. 2012.

[17] Z. Long, N. Wu, H. Wang, and Q. Guo, "Turbo equalization based on a combined VMP-BP algorithm for nonlinear satellite channels," *IEEE Access*, vol. 6, pp. 35 492–35 500, 2018.

[18] C. Laot, A. Glavieux, and J. Labat, "Turbo equalization: Adaptive equalization and channel decoding jointly optimized," *IEEE J. Sel. Areas Commun.*, vol. 19, no. 9, pp. 1744–1752, Sep. 2001.

[19] D. J. C. MacKay, "Introduction to Monte Carlo methods," in *Learning in Graphical Models*, ser. NATO Science Series, M. I. Jordan, Ed. Kluwer Academic Press, 1998, pp. 175–204.

[20] X. Wen, W. Yuan, D. Yang, N. Wu, and J. Kuang, "Low complexity message passing receiver for faster-than-Nyquist signaling in nonlinear channels," *IEEE Access*, vol. 6, pp. 68 233–68 241, 2018.

[21] R. Prasad, C. R. Murthy, and B. D. Rao, "Joint channel estimation and data detection in MIMO-OFDM systems: A sparse Bayesian learning approach," *IEEE Trans. Signal Process.*, vol. 63, no. 20, pp. 5369–5382, Oct. 2015.

[22] M. Jiang, J. Akhtman, and L. Hanzo, "Iterative joint channel estimation and multi-user detection for multiple-antenna aided OFDM systems," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 2904–2914, Aug. 2007.

[23] O. Castaneda, T. Goldstein, and C. Studer, "VLSI designs for joint channel estimation and data detection in large SIMO wireless systems," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 65, no. 3, pp. 1120–1132, Mar. 2018.

[24] H. Vikalo, B. Hassibi, and P. Stoica, "Efficient joint maximum-likelihood channel estimation and signal detection," *IEEE Trans. Wireless Commun.*, vol. 5, no. 7, pp. 1838–1845, Jul. 2006.

[25] Y. Ma, N. Wu, W. Yuan, D. W. K. Ng, and L. Hanzo, "Joint channel estimation and equalization for index-modulated spectrally efficient frequency division multiplexing systems," *IEEE Trans. Commun.*, vol. 68, no. 10, pp. 6230–6244, Oct. 2020.

[26] J. Ma and L. Ping, "Data-aided channel estimation in large antenna systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3111–3124, Apr. 2014.

[27] C. A. R. Fernandes, "Nonlinear MIMO communication systems: Channel estimation and information recovery using Volterra models," Ph.D. dissertation, Université de Nice Sophia Antipolis, 2009.

[28] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2011, pp. 2168–2172.

[29] W. Yuan, N. Wu, Q. Guo, Y. Li, C. Xing, and J. Kuang, "Iterative receivers for downlink MIMO-SCMA: Message passing and distributed cooperative detection," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 3444–3458, May. 2018.

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

[31] L. Svensson and M. Lundberg, "On posterior distributions for signals in Gaussian noise with unknown covariance matrix," *IEEE Trans. Signal Process.*, vol. 53, no. 9, pp. 3554–3571, Sep. 2005.

[32] P. Schniter, S. Rangan, and A. K. Fletcher, "Vector approximate message passing for the generalized linear model," in *Proc. 50th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2016, pp. 1525–1529.

[33] S. Rangan, P. Schniter, and A. K. Fletcher, "Vector approximate message passing," *IEEE Trans. Inform. Theory*, vol. 65, no. 10, pp. 6664–6684, Oct. 2019.

[34] J. Ma and L. Ping, "Orthogonal AMP," *IEEE Access*, vol. 5, pp. 2020–2033, 2017.

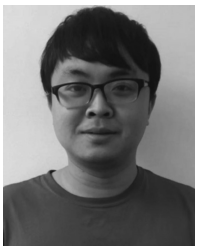
[35] L. Liu, S. Huang, and B. M. Kurkoski, "Memory approximate message passing," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Jul. 2021, pp. 1379–1384.

Yikun Zhang (Student Member, IEEE) received the B.S. degree in 2019 and the M.S. degree in 2022 from the School of Information and Electronics, Beijing Institute of Technology, Beijing, China. Her research interests include statistical inference on graphical models and its application to wireless communications.

Bin Li (Member, IEEE) received the B.S. degree in information engineering and the M.S. degree in information and communication engineering from 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. His research interests include signal processing, wireless communications, and machine learning. He serves as an Editorial Board Member of IEICE Transactions on Communications and KSII Transactions on Internet and Information Systems.

Nan Wu (Member, IEEE) received the 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, The Pennsylvania State University, USA. He is currently a Professor with the School of Information and Electronics, BIT. His research interest includes signal processing in wireless communication networks. He serves as an Editorial Board Member of IEEE Wireless Communications Letters, IEEE Access, International Journal of Electronics and Communications, and KSII Transactions on Internet and Information Systems. He was a recipient of the National Excellent Doctoral Dissertation Award by MOE of China in 2013.

Yunsi Ma (Graduate Student Member, IEEE) received the B.S. degree from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2013, the M.S. degree from the China Academy of Space Technology, Beijing, China, in 2016, and the Ph.D. degree from Beijing Institute of Technology, Beijing, China, in 2022. Her research interests include statistical inference on graphical models and its application to wireless communications.



Weijie Yuan (Member, IEEE) received the Ph.D. degree from the Beijing Institute of Technology, China, and the Ph.D. degree from the University of Technology Sydney, Australia, in 2019. He is currently a Research Associate with the School of Electrical Engineering and Telecommunications, University of New South Wales, Sydney. He has served as a Research Assistant with the University of Sydney, a Visiting Associate Fellow with the University of Wollongong, and a Visiting Fellow with the University of Southampton, from 2017 to 2019. In 2016, he was a Visiting Ph.D. Student with the Institute of Telecommunications, Vienna University of Technology, Austria. His research interest includes statistical inference on graphical models. He has served as a TPC member for several conferences.

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 is also a Fellow of the Royal Academy of Engineering (FREng), of the IET and of EURASIP. He is the recipient of the 2022 Eric Sumner Field Award. He has published 19 John Wiley research monographs and 2000+ contributions at IEEE Xplore. (http://www-mobile.ecs.soton.ac.uk, https://en.wikipedia.org/wiki/Lajos_Hanzo)