IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 7, 2018, accepted June 12, 2018, date of publication June 20, 2018, date of current version July 19, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2849147

Turbo Equalization Based on a Combined VMP-BP Algorithm for Nonlinear Satellite Channels

ZHEREN LONG[1,2], NAN WU[1], (Member, IEEE), HUA WANG[1], (Member, IEEE), AND QINGHUA GUO[3,4], (Member, IEEE)

[1]School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China
[2]Institute of Telecommunication Satellite, China Academy of Space Technology, Beijing 100094, China
[3]School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Wollongong, NSW 2522, Australia
[4]School of Electrical, Electronic and Computer Engineering, The University of Western Australia, Perth, WA 6009, Australia

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

This work was supported in part by the National Science Foundation of China under Grant 61471037 and Grant 61571041 and in part by the Australian Research Council's DECRA under Grant DE120101266.

ABSTRACT Close to saturation operation of high power amplifier (HPA) leads to strong nonlinear and dispersive characteristic of satellite channels. At the receiver, the observation signals are distorted by not only the linear inter-symbol interference (ISI) but also the nonlinear ones, which makes it challenge to perform optimal detection. In this paper, we study factor graph (FG)-based turbo equalization for nonlinear satellite channels characterized by Volterra series. Factor nodes on FG are classified into belief propagation (BP) set and variational message passing (VMP) set to enable low complexity combined message passing implementation while with high performance. BP is used on the hard constraint nodes, such as demapping and decoding, while VMP is employed to update messages of the likelihood function node. It is shown that, without any approximation on the Volterra series channel model, messages can be expressed in a closed form via canonical parameters, and the extrinsic information from equalizer to decoder is derived in an explicit way. Simulation results demonstrate the superior performance of the proposed combined VMP-BP algorithm with low computational complexity.

INDEX TERMS Nonlinear channel, Volterra series, turbo equalization, factor graph, belief propagation, variational message passing.

I. INTRODUCTION

High power amplifier (HPA) is an essential component in the transponder of satellite communication system [1]–[5]. When operating near the saturation point, HPA will lead to strong nonlinear distortion. Together with the input demultiplexer (IMUX) filter and output multiplexer (OMUX) filter, the transponder introduces both linear and nonlinear inter-symbol interference (ISI), leading to a significant performance degradation. At the receiver side, equalizer can be used to eliminate the distortion [3]–[5].

Volterra series is shown to be able to model the nonlinear satellite channels [1]. In [3], equalizer based on Volterra filter is adopted to compensate for the distortion of satellite communication system and the corresponding Wiener solution is achieved. In [4], the least-mean square (LMS) algorithm is modified to update the coefficients of Volterra filter. Inspired by the structure of turbo codes, turbo equalizer is proposed in [6] for coded system, which improves the system performance significantly. Many low-complexity turbo equalizers are subsequently proposed for linear channels in [7]–[11]. Turbo equalizers for nonlinear channels can be extend from equalizers in linear channels. In [12], cancellation filters for nonlinear terms are included to extend the method in [7] to nonlinear channel. Based on [8], the filters in [13] is designed to be excited by both the original symbols and their nonlinear combinations. In [14]–[16], the method in [9] is extended by taking the nonlinear ISI into consideration in the standard affine transform. The distribution of interference is assumed to be Gaussian in the above equalizers when calculating the extrinsic information, which are heuristic.

Factor graph (FG) can be used to efficiently represent the factorization of joint probability function [17]. Together with belief propagation (BP) algorithm [18], iterative receivers for

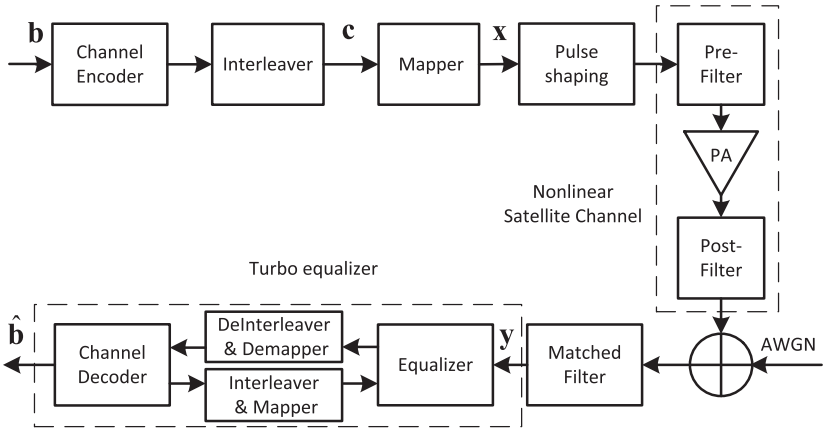


FIGURE 1. Block diagram of a coded satellite communication system.

linear channels can be designed in a unified way [19], [20]. FG with BP algorithm has been employed to design iterative equalizers for linear ISI channels in [21]–[26]. For nonlinear satellite channels, forward/backward equalizer and Markov Chain Monte Carlo (MCMC) equalizer are proposed in [27] based on FG. In fact, the forward/backward equalizer is the same as the BCJR equalizer in [28], whose complexity becomes unacceptable for high order modulations and/or channels with large memory depth. In [27], the *a posteriori* probability of symbols is obtained based on sampling methods. However,'opening node' [29] operation is not performed, i.e., the equalizer part is considered as one factor node without any factorizations, which requires significantly high computational complexity. In [30], by ignoring some terms of the Volterra kernels in a continuous domain model, the likelihood function is further factorized. Although equalizer with linear complexity can be obtained based on the above FG, the approximation leads to performance degradation. Moreover, for MPSK signals, this approximation results in a linear ISI channel.

In this paper, we study turbo equalization based on a combined variational message passing (VMP)-BP algorithm on FG for nonlinear satellite channels modeled by Volterra series. Different from [27] and [30], Volterra series expression is employed without any approximation, and the nonlinear observation function is factorized directly. By employing VMP algorithm on the likelihood function node, the extrinsic information from equalizer is derived in a rigorous way, which can be expressed by canonical parameters. BP algorithm is performed on the hard constraints nodes, e.g., decoding and demapping. Based on the proposed combined VMP-BP algorithm, all the messages on FG can be derived in closed form and parametric message passing can be performed, which significantly reduces the computational complexity. Simulation results demonstrate the superior performance of the proposed algorithm compared with the state-of-the-art methods.

The organization of this paper is as follows. System model is introduced in Section II. Factor graph of the probabilistic model for nonlinear satellite channels and the message passing algorithms are given in Section III. Receiver based on a combined VMP-BP algorithm for nonlinear satellite channel is derived in detail in Section IV. Simulation results and discussions are shown in Section V. Finally, conclusions are drawn in Section VI.

II. SYSTEM MODEL

We consider coded linearly modulated signals transmitted over nonlinear satellite channel as illustrated in Fig. 1. At the transmitter, the information bit sequence $b \triangleq [b_1, b_2, \ldots, b_K]^{\mathrm{T}}$ is converted to coded bits by channel encoder with coding rate $R = K/M$, and interleaved to bit sequence $c \triangleq [c_1, c_2, \ldots, c_M]^{\mathrm{T}}$. Then, the interleaved coded bits are mapped into symbols $x \triangleq [x_1, x_2, \ldots, x_N]^{\mathrm{T}}$, where $x_n \in \chi$, and χ is the constellation set with size $|\chi| = 2^P$. After pulse-shaping, the signal is transmitted over the nonlinear satellite channel.

Pre-filter, power amplifier and post-filter are the three components of transponder. Unexpected signals from adjacent channels are removed by pre-filter, while the extended spectrum caused by the nonlinearity of power amplifier is restrained by post-filter. The equivalent discrete-time baseband model at symbol rate can be described by Volterra series. The received signal vector after matched filter is $y \triangleq [y_1, y_2, \ldots, y_N]^{\mathrm{T}}$. The relationship between the input signal x_n and the output signal y_n is given as

$$y_n = \sum_{v=1}^{\infty} \sum_{n_1} \cdots \sum_{n_{2v-1}} h^{2v-1}_{n_1,\ldots,n_{2v-1}} x_{n_1,\ldots,n_{2v-1}} + w_n, \qquad (1)$$

where

$$x_{n_1,\ldots,n_{2v-1}} \triangleq x_{n-n_1} \ldots x_{n-n_v} x^*_{n-n_{v+1}} \ldots x^*_{n-n_{2v-1}},$$

and $h_{n_1,\ldots,n_{2v-1}}$ is the kernel of Volterra series, w_n is the circularly-symmetric white Gaussian noise with variance σ^2. Due to the bandpass nature of the channel, only odd terms are included [2].

Generally, a third-order Volterra series with a certain memory depth is adequate to describe practical nonlinear satellite channels [13], [14]. In that way, (1) becomes

$$\begin{aligned} y_n &= \sum_{l=0}^{L} h_l x_{n-l} + \sum_{i=0}^{L}\sum_{j\geq i}^{L}\sum_{k=0}^{L} h_{ijk} x_{n-i} x_{n-j} x_{n-k}^{*} + w_n \\ &= \mathbf{h}^{\mathrm{T}}\mathbf{x}_n + \mathbf{h}'^{\mathrm{T}}\mathbf{x}'_n + w_n, \end{aligned} \tag{2}$$

where $\mathbf{h} \triangleq [h_0, h_1, \ldots, h_L]^{\mathrm{T}}$, $\mathbf{x}_n \triangleq [x_n, x_{n-1}, \ldots, x_{n-L}]^{\mathrm{T}}$, $\mathbf{h}' \triangleq [h_{000}, h_{001}, \ldots, h_{00L}, \ldots, h_{LLL}]^{\mathrm{T}}$, $\mathbf{x}'_n \triangleq [x_n x_n x_n^{*}, x_n x_n x_{n-1}^{*}, \ldots, x_n x_n x_{n-L}^{*}, \ldots, x_{n-L} x_{n-L} x_{n-L}^{*}]^{\mathrm{T}}$ and L is the channel length. The elements in $\mathbf{h}$ and $\mathbf{h}'$ could be zero, and the symmetric characteristic of Volterra series has been taken into consideration.

At the receiver side, the maximum *a posteriori* probability (MAP) bit-by-bit detector is give by

$$\hat{b}_k = \underset{b_k\in\{0,1\}}{\operatorname{argmax}}\, p(b_k|\mathbf{y}) = \underset{b_k\in\{0,1\}}{\operatorname{argmax}} \sum_{\sim\{b_k\}} p(\mathbf{b}|\mathbf{y}), \tag{3}$$

where $\sim\{b_k\}$ denotes the summation over all the variables in information bit sequence $\mathbf{b}$, except b_k. However, due to the coupling between variables, the direct marginalization in (3) is intractable. We introduce FG and message passing algorithms to solve this problem.

III. FACTOR GRAPH AND MESSAGE PASSING ALGORITHMS

A. FACTOR GRAPH FOR NONLINEAR SATELLITE CHANNEL

Due to the conditional independence, given the observation vector $\mathbf{y}$, the *a posteriori* probability of information bits $\mathbf{b}$, interleaved coded bits $\mathbf{c}$ and transmitted symbols $\mathbf{x}$ is

$$\begin{aligned} p(\mathbf{b},\mathbf{c},\mathbf{x}|\mathbf{y}) &\propto p(\mathbf{b})p(\mathbf{c}|\mathbf{b})p(\mathbf{x}|\mathbf{c})p(\mathbf{y}|\mathbf{x}) \\ &= \prod_{k=1}^{K} f(b_k) f_c(\mathbf{b},\mathbf{c}) f_m(\mathbf{c},\mathbf{x}) \\ &\quad \times \prod_{n=1}^{N} f_{o_n}(x_{n-L},\ldots,x_n,y_n), \end{aligned} \tag{4}$$

where $f(b_k)$ is the *a priori* probability of the information bit, $f_c(\mathbf{b},\mathbf{c})$ and $f_m(\mathbf{c},\mathbf{x})$ are the coding&interleaving and mapping functions, respectively,

$$f_m(\mathbf{c},\mathbf{x}) = \prod_{n=1}^{N} f_{m_n}(c_{(n-1)P+1},\ldots,c_{nP},x_n) \tag{5}$$

$$f_{m_n}(\cdot) = \delta(G(c_{(n-1)P+1},\ldots,c_{nP}) - x_n) \tag{6}$$

with the mapping function G, and $f_{o_n}(\cdot)$ is the likelihood function, i.e.,

$$\begin{aligned} f_{o_n}(\cdot) &= p(y_n|x_{n-L},\ldots,x_n) \\ &\propto \exp\left(-\frac{|y_n - (\mathbf{h}^{\mathrm{T}}\mathbf{x}_n + \mathbf{h}'^{\mathrm{T}}\mathbf{x}'_n)|^2}{\sigma^2}\right). \end{aligned}$$

The corresponding factor graph of the probabilistic model in (4) for nonlinear satellite channel is shown in Fig. 2.

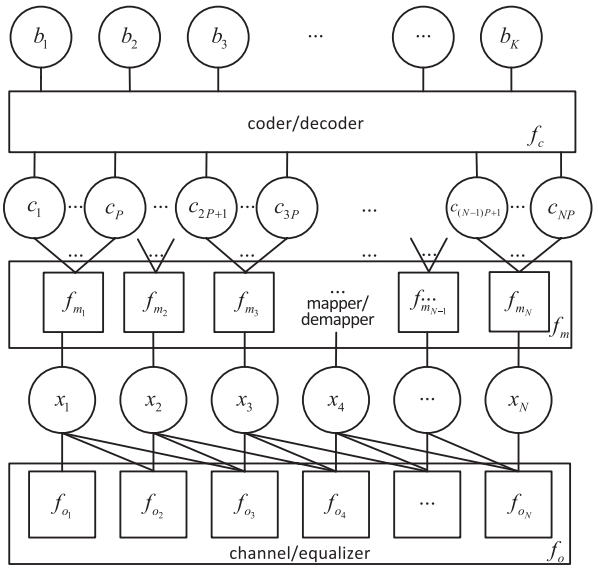


FIGURE 2. Factor graph representation of the probabilistic model in (4).

B. MESSAGE PASSING ALGORITHM

BP and VMP are two message passing algorithms on FG. BP is derived from the equations of the stationary points of the constrained Bethe free energy. VMP is obtained by minimizing the variational free energy subject to the mean-field approximation constraint [31]. Partitioning the FG into BP and VMP parts, a combined VMP-BP algorithm is proposed by minimizing the region-based free energy [32]–[34].

For a FG consisting of a set of factor nodes $\mathbf{F}$ and a set of variable nodes $\mathbf{V}$, according to the combined VMP-BP algorithm, the factor nodes can be assigned to two sets: BP set $\mathbf{F}^{\mathrm{BP}}$ and VMP set $\mathbf{F}^{\mathrm{VMP}}$, $\mathbf{F}^{\mathrm{BP}} \cup \mathbf{F}^{\mathrm{VMP}} = \mathbf{F}$, $\mathbf{F}^{\mathrm{BP}} \cap \mathbf{F}^{\mathrm{VMP}} = \emptyset$. The message from factor node $f \in \mathbf{F}$ to its neighbor variable node $v \in \mathcal{N}(f) \subseteq \mathbf{V}$ is denoted as $m_{f\to v}(v)$, and the message from variable node v to its neighbor factor node $f \in \mathcal{N}(v) \subseteq \mathbf{F}$ is denoted as $n_{v\to f}(v)$.

The updating rule related to the factor node f_b in the BP part on FG ($f_b \in \mathbf{F}^{\mathrm{BP}}$) is

$$m^{\mathrm{BP}}_{f_b\to v}(v) \propto \sum_{\sim\{v\}} f_b(\mathbf{v}) \prod_{v'\in\mathcal{N}(f_b)\backslash v} n_{v'\to f_b}(v'). \tag{7}$$

The updating rule related to the factor node f_a in the VMP part on FG ($f_a \in \mathbf{F}^{\mathrm{VMP}}$) is

$$m^{\mathrm{VMP}}_{f_a\to v}(v) \propto \exp\left(\sum_{\sim\{v\}} \ln(f_a(\mathbf{v})) \prod_{v'\in\mathcal{N}(f_a)\backslash v} n_{v'\to f_a}(v')\right). \tag{8}$$

The updating rule related to the variable node v is

$$n_{v\to f_c}(v) \propto \prod_{f_a\in\mathcal{N}(v)\cap\mathbf{F}^{\mathrm{VMP}}} m^{\mathrm{VMP}}_{f_a\to v}(v) \prod_{f_b\in\mathcal{N}(v)\cap\mathbf{F}^{\mathrm{BP}}\backslash f_c} m^{\mathrm{BP}}_{f_b\to v}(v). \tag{9}$$

$$(\mathbf{h}^{\mathrm{T}}\mathbf{x}_n+\mathbf{h}'^{\mathrm{T}}\mathbf{x}'_n)^{\mathrm{H}}(\mathbf{h}^{\mathrm{T}}\mathbf{x}_n+\mathbf{h}'^{\mathrm{T}}\mathbf{x}'_n)=\sum_{i=0}^{L}h_i^*h_ix_{n-i}^*x_{n-i}-\sum_{i=0}^{L}\sum_{j\geq i}^{L}\sum_{k=0}^{L}h_{ijk}^*h_{ijk}x_{n-i}^*x_{n-j}^*x_{n-k}x_{n-i}x_{n-j}x_{n-k}^*$$
$$+2\Re\Big(\sum_{i=0}^{L}\sum_{o=0}^{L}\sum_{p\geq o}^{L}\sum_{q=0}^{L}h_i^*x_{n-i}^*h_{opq}x_{n-o}x_{n-p}x_{n-q}^*\Big)+2\Re\Big(\sum_{i=0}^{L}\sum_{j>i}^{L}h_i^*h_jx_{n-i}^*x_{n-j}\Big)$$
$$+2\Re\Big(\sum_{i=0}^{L}\sum_{j\geq i}^{L}\sum_{k=0}^{L}\sum_{o\geq i}^{L}\sum_{p\geq o,p\geq j}^{L}\sum_{q\geq k}^{L}h_{ijk}^*h_{opq}x_{n-i}^*x_{n-j}^*x_{n-k}x_{n-o}x_{n-p}x_{n-q}^*\Big) \tag{14}$$

The belief of variable v is

$$b(v)\propto\prod_{f_a\in\mathcal{N}(v)\cap\mathbf{F}^{\mathrm{VMP}}}m_{f_a\to v}^{\mathrm{VMP}}(v)\prod_{f_b\in\mathcal{N}(v)\cap\mathbf{F}^{\mathrm{BP}}}m_{f_b\to v}^{\mathrm{BP}}(v). \tag{10}$$

IV. RECEIVER BASED ON COMBINED VMP-BP FOR NONLINEAR SATELLITE CHANNELS

Due to the hard constraints of coding, interleaving and mapping, factor nodes related to these functions, i.e., f_c and f_{m_i}, are classified into the BP set, while the factor nodes of likelihood functions, i.e., f_{o_n}, are classified into the VMP set. Although the subgraph between symbol variable nodes x_i and likelihood function nodes f_{o_n} contains cycles with girth four, the convergence is guaranteed by the VMP algorithm [35].

A. EQUALIZATION BASED ON VMP ALGORITHM

The messages from likelihood function nodes to symbol nodes are computed as follows

$$m_{f_{o_n}\to x_i}(x_i)\propto\exp\left(\sum_{\sim\{x_i\}}\ln(f_{o_n}(\cdot))\prod_{x_j\in\mathcal{N}(f_{o_n})\backslash x_i}n_{x_j\to f_{o_n}}(x_j)\right)$$
$$=\exp\left(\mathbb{E}_{x_j\in\mathcal{N}(f_{o_n})\backslash x_i}\{-\frac{|y_n-(\mathbf{h}^{\mathrm{T}}\mathbf{x}_n+\mathbf{h}'^{\mathrm{T}}\mathbf{x}'_n)|^2}{\sigma^2}\}\right). \tag{11}$$

The term $|y_n-(\mathbf{h}^{\mathrm{T}}\mathbf{x}_n+\mathbf{h}'^{\mathrm{T}}\mathbf{x}'_n)|^2$ in (11) can be decomposed as

$$|y_n-(\mathbf{h}^{\mathrm{T}}\mathbf{x}_n+\mathbf{h}'^{\mathrm{T}}\mathbf{x}'_n)|^2$$
$$=y_n^*y_n-2\Re(y_n^*(\mathbf{h}^{\mathrm{T}}\mathbf{x}_n+\mathbf{h}'^{\mathrm{T}}\mathbf{x}'_n))$$
$$+(\mathbf{h}^{\mathrm{T}}\mathbf{x}_n+\mathbf{h}'^{\mathrm{T}}\mathbf{x}'_n)^*(\mathbf{h}^{\mathrm{T}}\mathbf{x}_n+\mathbf{h}'^{\mathrm{T}}\mathbf{x}'_n), \tag{12}$$

The first term $y_n^*y_n$ can be considered as a constant. The second and the third terms are expanded in (13) and (14), as shown in top of this page, respectively.

$$y_n^*(\mathbf{h}^{\mathrm{T}}\mathbf{x}_n+\mathbf{h}'^{\mathrm{T}}\mathbf{x}'_n)=\sum_{i=0}^{L}y_n^*h_ix_{n-i}+\sum_{i=0}^{L}\sum_{j\geq i}^{L}\sum_{k=0}^{L}y_n^*h_{ijk}x_{n-i}x_{n-j}x_{n-k}^*. \tag{13}$$

The likelihood function node f_{o_n} is connected to $L+1$ symbols, i.e., $x_n, x_{n-1},\ldots x_{n-L}$. When $m_{f_{o_n}\to x_i}(x_i)$ is calculated, x_i is considered as variable while other symbols are taken place by their expectations, e.g.,

$$\sum_{\sim x_n}h_{001}^*h_{012}x_n^*x_n^*x_{n-1}x_nx_{n-1}x_{n-2}^*\prod_{j=n-1,n-2}n_{x_j\to f_{o_n}}(x_j)$$
$$=h_{001}^*h_{012}\mathbb{E}\{x_{n-1}x_{n-1}\}\mathbb{E}\{x_{n-2}^*\}x_n^*x_n^*x_n, \tag{15}$$

We can find that the message $m_{f_{o_n}\to x_i}(x_i)$ belongs to the exponential family. Therefore, the sufficient statistics [36] can be chosen as the combinations of the bases of Volterra series, which can be split into two parts, i.e., with/without real operation. They are in the form of $x_ix_i^*$, $x_ix_ix_i^*x_i^*$, $x_ix_ix_ix_i^*x_i^*x_i^*$, $\Re(x_i)$, $\Re(x_ix_i)$, $\Re(x_ix_ix_i^*)$, $\Re(x_ix_ix_i)$, $\Re(x_ix_ix_ix_i^*)$, $\Re(x_ix_ix_ix_i^*x_i^*)$ as in (13) and (14).[1] Then, the message $m_{f_{o_n}\to x_i}(x_i)$ can be written as

$$m_{f_{o_n}\to x_i}(x_i)\propto\exp\left(-\frac{\sum_{m,n}c_{m,n}^{o_n}x_i^mx_i^{*n}+2\Re(\sum_{p,q}c_{p,q}^{o_n}x_i^px_i^{*q})}{\sigma^2}\right) \tag{16}$$

where $c_{m,n}^{o_n}$ and $c_{p,q}^{o_n}$ are the canonical parameters, $x_i^mx_i^{*n}$ and $x_i^px_i^{*q}$ are the sufficient statistics.[2] Obviously, only the canonical parameters [37] have to be calculated during the message updating.

Taking a third-order Volterra series with memory depth one as an example, the sufficient statistics and corresponding canonical parameters of message $m_{f_{o_n}\to x_n}(x_n)$ are shown in Table 1.

For a third-order Volterra channel with larger memory depth, the sufficient statistics are the same as that in Table. 1, while more terms are involved in canonical parameters. For Volterra channel with higher order and larger memory depth, there will be more sufficient statistics and more terms in canonical parameters. Nevertheless, the canonical parameters are calculated in a similar way. Moreover, due to the sparsity characteristic of nonlinear satellite channels [38], many terms in the canonical parameters will be zeros.

[1] The index of the conjugate symbol is set to be less than that of the original symbol, in order to reduce the number of sufficient statistics.

[2] Due to the real part operation, $x_i^px_i^{*q}$ is marked as $\Re(x_i^px_i^{*q})$ in Table 1

TABLE 1. Sufficient statistics and canonical parameters of $m_{f_{o_n}\to x_n}(x_n)$ for a third-order Volterra channel with memory depth 1.

Sufficient statistics	Canonical parameters ($c^{o_n}_{m,n}, c^{o_n}_{p,q}$)
$\Re(x_n)$	$-y_n^* h_0$ $-y_n h_{110}^* \mathbb{E}\{x_{n-1}^* x_{n-1}^*\}$ $-y_n^* h_{011} \mathbb{E}\{x_{n-1} x_{n-1}^*\}$ $+h_0 h_1^* \mathbb{E}\{x_{n-1}^*\}$ $+h_0 h_{111}^* \mathbb{E}\{x_{n-1}^* x_{n-1}^* x_{n-1}\}$ $+h_1^* h_{011} \mathbb{E}\{x_{n-1} x_{n-1}^* x_{n-1}^*\}$ $+h_1 h_{110}^* \mathbb{E}\{x_{n-1}^* x_{n-1}^* x_{n-1}\}$ $+h_{011} h_{111}^* \mathbb{E}\{x_{n-1}^* x_{n-1}^* x_{n-1}^* x_{n-1} x_{n-1}\}$ $+h_{110}^* h_{111} \mathbb{E}\{x_{n-1} x_{n-1} x_{n-1}^* x_{n-1}^* x_{n-1}^*\}$
$\Re(x_n x_n)$	$-y_n^* h_{001} \mathbb{E}\{x_{n-1}^*\}$ $+h_0 h_{110}^* \mathbb{E}\{x_{n-1}^* x_{n-1}^*\}$ $+h_1^* h_{001} \mathbb{E}\{x_{n-1}^* x_{n-1}^*\}$ $+h_{001} h_{111}^* \mathbb{E}\{x_{n-1}^* x_{n-1}^* x_{n-1}^* x_{n-1}\}$ $+h_{011} h_{110}^* \mathbb{E}\{x_{n-1} x_{n-1}^* x_{n-1}^* x_{n-1}^*\}$
$\Re(x_n x_n x_n^*)$	$-y_n^* h_{000}$ $+h_0^* h_{001} \mathbb{E}\{x_{n-1}^*\}$ $+h_0 h_{010}^* \mathbb{E}\{x_{n-1}^*\}$ $+h_1^* h_{000} \mathbb{E}\{x_{n-1}^*\}$ $+h_{000} h_{111}^* \mathbb{E}\{x_{n-1}^* x_{n-1}^* x_{n-1}\}$ $+h_{001} h_{011}^* \mathbb{E}\{x_{n-1}^* x_{n-1}^* x_{n-1}\}$ $+h_{010}^* h_{011} \mathbb{E}\{x_{n-1} x_{n-1}^* x_{n-1}^*\}$ $+h_{010} h_{110}^* \mathbb{E}\{x_{n-1} x_{n-1}^* x_{n-1}^*\}$
$\Re(x_n x_n x_n)$	$h_{001} h_{110}^* \mathbb{E}\{x_{n-1}^* x_{n-1}^* x_{n-1}^*\}$
$\Re(x_n x_n x_n x_n^*)$	$+h_{001} h_{010}^* \mathbb{E}\{x_{n-1}^* x_{n-1}^*\}$ $+h_{000} h_{110}^* \mathbb{E}\{x_{n-1}^* x_{n-1}^*\}$
$\Re(x_n x_n x_n x_n^* x_n^*)$	$+h_{000}^* h_{001} \mathbb{E}\{x_{n-1}^*\}$ $+h_{000} h_{010}^* \mathbb{E}\{x_{n-1}^*\}$
$x_n x_n^*$	$+h_0^* h_0$ $+h_{110}^* h_{110} \mathbb{E}\{x_{n-1} x_{n-1} x_{n-1}^* x_{n-1}^*\}$ $+h_{011}^* h_{011} \mathbb{E}\{x_{n-1} x_{n-1} x_{n-1}^* x_{n-1}^*\}$ $-2\Re(y_n^* h_{010} \mathbb{E}\{x_{n-1}\})$ $+2\Re(h_0^* h_{011} \mathbb{E}\{x_{n-1} x_{n-1}^*\})$ $+2\Re(h_1^* h_{010} \mathbb{E}\{x_{n-1} x_{n-1}^*\})$ $+2\Re(h_{010}^* h_{111} \mathbb{E}\{x_{n-1} x_{n-1} x_{n-1}^* x_{n-1}^*\})$
$x_n x_n x_n^* x_n^*$	$+h_{001}^* h_{001} \mathbb{E}\{x_{n-1} x_{n-1}^*\}$ $+h_{010}^* h_{010} \mathbb{E}\{x_{n-1} x_{n-1}^*\}$ $+2\Re(h_0^* h_{000})$ $+2\Re(h_{000}^* h_{011} \mathbb{E}\{x_{n-1} x_{n-1}^*\})$
$x_n x_n x_n x_n^* x_n^* x_n^*$	$h_{000}^* h_{000}$

Then, the messages from symbol nodes to the likelihood function nodes, i.e., $n_{x_i\to f_{o_n}}(x_i)$, are calculated as follows

$$n_{x_i\to f_{o_n}}(x_i) = \prod_{k=i}^{i+L} m_{f_{o_k}\to x_i}(x_i) m_{f_{m_i}\to x_i}(x_i), \tag{17}$$

where $m_{f_{m_i}\to x_i}(x_i)$ is the information from mapping node, which is the probability mass function (PMF) of x_i, i.e.,

$$m_{f_{m_i}\to x_i}(x_i) = \sum_j p(x_i = s_j)\delta(x_i - s_j). \tag{18}$$

Substituting (16) and (18) into (17), we have

$$\begin{aligned} &n_{x_i\to f_{o_n}}(x_i) \\ &\propto \sum_j p(x_i = s_j)\delta(x_i - s_j) \\ &\quad \times \exp\left(-\frac{\sum_{k=i}^{i+L} (\sum_{m,n} c^{o_k,x_i}_{m,n} s_j^m s_j^{*n} + 2\Re(\sum_{p,q} c^{o_k,x_i}_{p,q} s_j^p s_j^{*q}))}{\sigma^2} \right) \end{aligned} \tag{19}$$

where $c^{o_k,x_i}_{m,n}$ and $c^{o_k,x_i}_{p,q}$ are the canonical parameters of message from the likelihood function node f_{o_k} to symbol node x_i.

It is seen from Table 1 that the (raw) moments [39] with different orders can be used to calculate the canonical parameters. Due to the discrete characteristic of symbols, enumeration method is used to obtain the corresponding moments e.g.,

$$\mathbb{E}\{x_i^2 x_i^{*2}\} = \sum_j n_{x_i\to f_{o_n}}(x_i = s_j) s_j^2 s_j^{*2}. \tag{20}$$

Although there are cycles with girth four between the likelihood function nodes and symbol nodes, the convergence is ensured by VMP algorithm.

B. DEMAPPING BASED ON BP ALGORITHM

The relationship of mapping belongs to hard constraint. The corresponding function node is $f_{m_n} = \delta(G(c_{(n-1)P+1}, \ldots c_{nP}) - x_n)$, where the function is one when there is a valid relationship between $c_{(n-1)P+1}, \ldots, c_{nP}$ and x_n, otherwise it is zero. Mapping factor nodes belong to BP part, so the message $n_{x_n\to f_{m_n}}(x_n)$ from symbol node x_n to mapping factor node f_{m_n} is the *extrinsic* information, i.e.,

$$\begin{aligned} &n_{x_n\to f_{m_n}}(x_n) \\ &= \prod_{i=n}^{n+L} m_{f_{o_i}\to x_n}(x_n) \\ &\propto \exp\left(-\frac{\sum_{i=n}^{n+L} (\sum_{a,b} c^{o_i,x_n}_{a,b} x_n^a x_n^{*b} + 2\Re(\sum_{c,d} c^{o_i,x_n}_{c,d} x_n^c x_n^{*d}))}{\sigma^2} \right). \end{aligned} \tag{21}$$

Since the message is in the same form of exponential distribution as $m_{f_{o_n}\to x_i}$, only the canonical parameters are required to be updated.

Then, the message from mapping nodes f_{m_n} to coded bit nodes c_i, i.e., $m_{f_{m_n}\to c_i}(c_i)$, is given by

$$\begin{aligned} &m_{f_{m_n}\to c_i}(c_i) \\ &\propto \sum_{k\neq i} \int f_{m_n}(\mathbf{c}_n, x_n) n_{x_n\to f_{m_n}}(x_n) dx_n \prod_{k\neq i} n_{c_k\to f_{m_n}}(c_k) \\ &= \sum_{l=0}^{1} \sum_{s_l\in\gamma_i^l} n_{x_n\to f_{m_n}}(x_n = s_l) \prod_{k\neq i} n_{c_k\to f_{m_n}}(c_k = p_k^l)\delta(c_i - l), \end{aligned} \tag{22}$$

where p_k^l and l could be 0 or 1, when c_k is p_k^l and c_i is l, the corresponding symbol x_n is s_l. Usually, the log-likelihood ratio (LLR) of coded bits are used instead of $m_{f_{m_n}\to c_i}(c_i)$ for simplicity, i.e.,

$$\begin{aligned} LLR(c_i) &= \ln \frac{m_{f_{m_n}\to c_i}(c_i = 0)}{m_{f_{m_n}\to c_i}(c_i = 1)} \\ &= \ln \frac{\sum_{s_0\in\gamma_i^0} n_{x_n\to f_{m_n}}(x_n = s_0) \prod_{k\neq i} n_{c_k\to f_{m_n}}(c_k = p_k^0)}{\sum_{s_1\in\gamma_i^1} n_{x_n\to f_{m_n}}(x_n = s_1) \prod_{k\neq i} n_{c_k\to f_{m_n}}(c_k = p_k^1)}. \end{aligned} \tag{23}$$

The message from mapping factor nodes to symbol variable nodes is

$$\begin{aligned} & n_{f_{m_n}\to x_n}(x_n) \\ &= \sum_{k} \delta(G(\mathbf{c}_n) - x_n) \prod_{k=(n-1)P+1}^{nP} n_{c_k\to f_{m_n}}(c_k) \\ &= \sum_{j} \prod_{k=(n-1)P+1}^{nP} n_{c_k\to f_{m_n}}(c_k = q_{k_j})\delta(x_n - s_j), \end{aligned} \tag{24}$$

where $c_{(n-1)P+1}, \ldots . c_{nP}$ are mapped to $x_n = s_j$ when each component c_k is q_{k_j}.

C. DECODING BASED ON BP ALGORITHM

The coding factor node belongs to BP part due to its hard constraints. Classical algorithms are employed, e.g., sum-product algorithm (SPA) for LDPC code, BCJR algorithm for convolutional code. Both SPA and BCJR can be derived based on BP algorithm on FG [18], and we do not discuss the decoding algorithm in this paper.

D. SCHEDULING OF MESSAGES ON FG

Obviously, the FG in Fig. 2 contains cycles. Therefore, the proposed VMP-BP algorithm has to be performed iteratively. The iterative message updating within nonlinear channel equalization is named 'inner loop', while the iterative message updating between equalization and decoding is named 'outer loop'. The proposed VMP-BP algorithm is summarized in Algorithm 1.

Algorithm 1 Turbo Equalization Based on VMP-BP Algorithm

Require: Prior information of interleaved coded bits $n_{c_i\to f_{m_n}}(c_i)$.

Ensure: Extrinsic information of interleaved coded bits $m_{f_{m_n}\to c_i}(c_i)$.

1: Initialization: $n_{c_i\to f_{m_n}}(c_i) = 1/2$, $n_{x_i\to f_{o_n}}(x_i) = 1/2^P$, $m_{f_{m_n}\to x_n}(x_n) = 1/2^P$;
2: Update the message from the mapping factor node to the symbol variable node $n_{f_{m_n}\to x_n}(x_n)$ using (24)
3: **repeat** inner loop, with fixed $m_{f_{m_n}\to x_n}(x_n)$
4: Update the message from the symbol variable node to observation factor node $n_{x_i\to f_{o_n}}(x_i)$ using (19)
5: Update the canonical parameters illustrated in Table 1
6: Update the message from the likelihood function node to symbol variable node $m_{f_{o_n}\to x_i}(x_i)$ using (16)
7: **until** the number of iterations for equalization reaches its maximum
8: Update the message from the symbol variable node to mapping factor node $n_{x_n\to f_{m_n}}(x_n)$ using (21)
9: Update the message from mapping factor node to the interleaved coded bit variable node $m_{f_{m_n}\to c_i}(c_i)$ using (22)

E. COMPLEXITY ANALYSIS

The complexity of the proposed turbo equalization based on VMP-BP algorithm are compared with three exsiting methods, namely, MMSE equalizer in [9], linear MMSE-based equalizer in [14] and forward/backward (FB) equalizer in [27]. A third-order Volterra channel with memory length L and $\mathcal{L}$ nonlinear terms is considered. The size of signal constellation is $\mathcal{M}$. A sliding window with size S is used in [9] and [14] to reduce the complexity.

The complexity of FB equalizer depends on the number of trellis states of the channel, which is $\mathcal{O}(\mathcal{M}^L)$. The calculation of matrix inverse and the covariance of the observation vector dominate the complexity of equalizers in [9] and [14]. By ignoring the nonlinear terms, the complexity of the traditional MMSE equalizer [9] is $\mathcal{O}(S^3 + S\mathcal{M})$. Taking the nonlinear terms into consideration, the complexity of [14] increases rapidly due to the calculation of the covariance of nonlinear terms, which is $\mathcal{O}(S^3 + \mathcal{L}^2S^2\mathcal{M})$. Different from the LMMSE based equalizers, only the expectations of different combination of symbols are required in the proposed algorithm. Without any operation of matrix inverse, the complexity of the proposed VMP-BP equalizer is $\mathcal{O}(\mathcal{L}^2\mathcal{M})$.

V. SIMULATION RESULTS

The performance of the proposed turbo equalizer is evaluated by Monte Carlo simulations. We consider a rate-1/2 (5,7) convolutional code with truncated termination. The block length of information bits is 2048. A 16-random interleaver is employed to scramble the coded bits in order to reduce burst error and produce the independence between symbols. QPSK and 16QAM signals with Gray mapping are considered. For the proposed VMP-BP equalizer, the maximum number of iterations for the inner loop and outer loop are set to $I_{\text{inner}} = 5$ and $I_{\text{outer}} = 10$, respectively, unless otherwise specified.

The nonlinear Volterra model is $y_n = h_0x_n + h_1x_{n-1} + h_2x_{n-2} + h_{000}x_nx_nx_n^* + h_{001}x_nx_nx_{n-1}^* + h_{002}x_nx_nx_{n-2}^* + h_{110}x_{n-1}x_{n-1}x_n^* + h_{220}x_{n-2}x_{n-2}x_n^*$, with coefficients given in Table 3 [12], [27].

The convergence behavior of the proposed VMP-BP equalizer at different E_b/N_0 is shown in Fig. 3. It is seen that

TABLE 2. Complexity analysis.

Algorithm	Complexity
MMSE equalizer [9]	$\mathcal{O}(S^3 + S\mathcal{M})$
LMMSE equalizer [14]	$\mathcal{O}(S^3 + \mathcal{L}^2S^2\mathcal{M})$
FB equalizer [27]	$\mathcal{O}(\mathcal{M}^L)$
Proposed VMP-BP equalizer	$\mathcal{O}(\mathcal{L}^2\mathcal{M})$

TABLE 3. Coefficients of volterra channel.

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}	
-0.005-0.085i	0.09-0.09i	

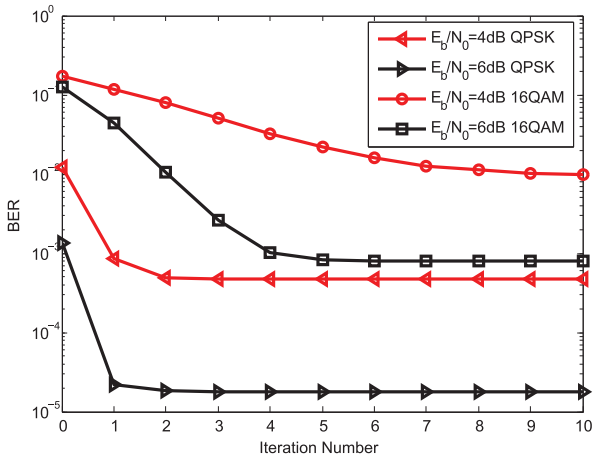


FIGURE 3. Convergence behavior of the proposed VMP-BP equalizer.

the proposed algorithm converges well in both QPSK and 16QAM cases, but the convergence speed of QPSK signal is much faster than that of the 16QAM. When E_b/N_0 is 6dB, only 2 iterations are required for QPSK, while 5 iterations are required for 16QAM to converge.

Bit error rate (BER) performance of the proposed VMP-BP equalizer is illustrated in Fig. 4. Three existing methods are also illustrated for comparison, which are the MMSE equalizer proposed in [9] and [40] (denoted as 'Linear'), LMMSE equalizer proposed in [14] (denoted as 'LMMSE'), and forward/backward equalizer proposed in [27] (denoted as 'FB'). The number of iterations between decoder and equalizer is set to 10 to ensure that all equalizers can converge. It is observed that, due to the ignorance of the nonlinear terms in Volterra channels, Linear equalizer has the worst performance as expected. The LMMSE equalizer outperforms the Linear one by taking part of nonlinear terms into consideration. However, its performance is much worse than that of the proposed VMP-BP equalizer. This is because all the

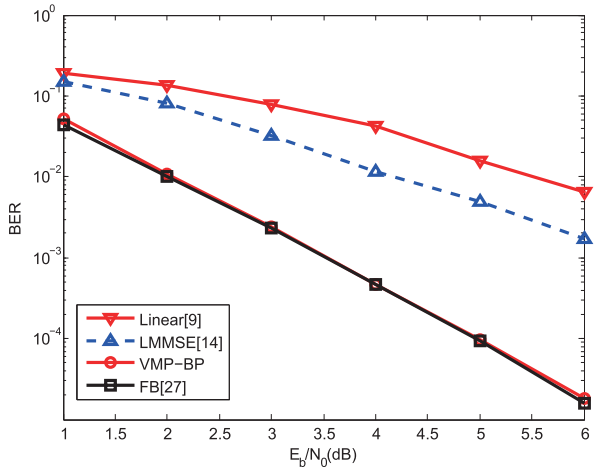


FIGURE 4. BER performance of different algorithms for QPSK signal.

nonlinear terms are taken into consideration in the proposed VMP-BP algorithm, and Gaussian approximation is not used when calculating the message (*extrinsic* information) from equalizer to decoder. It is seen that the proposed VMP-BP equalizer performs very close to the FB equalizer, with much lower computational complexity.

The performance of the proposed VMP-BP equalizer with different number of iterations for the inner loop and outer loop is shown in Fig. 5. The curve marked $I_{\text{outer}} = 0$ refers to the configuration that there is no iteration between decoder and equalizer. It can be observed that when I_{outer} is small, increasing I_{inner} can improve the BER performance. However, when I_{outer} is large, e.g., $I_{\text{outer}} = 10$, performance gap between different I_{inner} is negligible, which motivates us to set $I_{\text{inner}} = 1$ and embed the iteration of equalizer into the iteration of decoding and equalization.

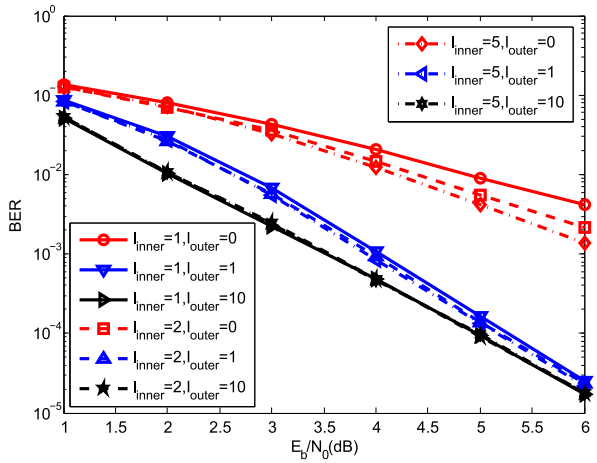


FIGURE 5. BER performance of the proposed VMP-BP equalizer with different number of iterations.

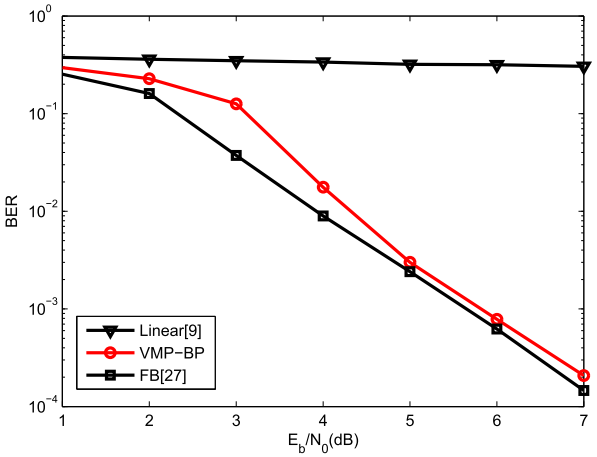


FIGURE 6. BER performance of different algorithms for 16QAM signal.

The performance of different equalizers for 16QAM signal are shown in Fig. 6. Due to the high complexity of calculating the covariance matrix of observation vector, the results of LMMSE equalizer are not included. Due to the non-constant

modulus property of the constellations, the nonlinear distortion of 16QAM signal is much more serious than that of QPSK signal. It is seen that, due to the ignorance of the nonlinear terms, Linear equalizer failed for 16QAM signal. The proposed VMP-BP equalizer performs well, and the gap between VMP-BP equalizer and FB equalizer is small when E_b/N_0 is in the range from 5dB to 7dB. Compared with FB equalizer, the proposed VMP-BP equalizer has an acceptable BER performance while having a much lower complexity.

VI. CONCLUSION

In this paper, turbo equalizer for nonlinear satellite channel modeled by Volterra series was studied. The probabilistic model of system was represented by FG without approximation on the nonlinear channels. BP was applied to the hard constraint nodes, such as demapping and decoding, while VMP was employed for the equalization of nonlinear channel. It was shown that messages from the equalizer belong to the exponential family, and only the canonical parameters have to be updated, which significantly reduced the computational complexity. Simulation results demonstrated that, for both QPSK and 16QAM signals, the proposed VMP-BP algorithm performed very close to the forward/backword equalizer and significantly outperformed the LMMSE equalizer which only considered part of the nonlinear ISI. The iterations within channel equalization can be fully embedded into the turbo processing between equalization and channel decoding, which further reduced the receiver complexity.

REFERENCES

[1] 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, Jul. 1979.

[2] G. Karam and H. Sari, "Analysis of predistortion, equalization, and ISI cancellation techniques in digital radio systems with nonlinear transmit amplifiers," *IEEE Trans. Commun.*, vol. COM-37, no. 12, pp. 1245–1253, Dec. 1989.

[3] S. Benedetto and E. Biglieri, "Nonlinear equalization of digital satellite channels," *IEEE J. Sel. Areas Commun.*, vol. SAC-1, no. 1, pp. 57–62, Jan. 1983.

[4] A. Gutierrez and W. E. Ryan, "Performance of Volterra and MLSD receivers for nonlinear band-limited satellite systems," *IEEE Trans. Commun.*, vol. 48, no. 7, pp. 1171–1177, Jul. 2000.

[5] E. Casini, R. De Gaudenzi, and A. Ginesi, "DVB-S2 modem algorithms design and performance over typical satellite channels," *Int. J. Satell. Commun. Netw.*, vol. 22, no. 3, pp. 281–318, Jun. 2004.

[6] C. Douillard *et al.*, "Iterative correction of intersymbol interference: Turbo-equalization," *Eur. Trans. Telecommun.*, vol. 6, no. 5, pp. 507–511, 1995.

[7] 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.

[8] D. Reynolds and X. Wang, "Low-complexity Turbo-equalization for diversity channels," *Signal Process.*, vol. 81, no. 5, pp. 989–995, 2001.

[9] M. Tuchler, R. Koetter, and A. C. Singer, "Turbo equalization: Principles and new results," *IEEE Trans. Commun.*, vol. 50, no. 5, pp. 754–767, May 2002.

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

[11] 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.

[12] 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.

[13] D. N. Liu and M. P. Fitz, "Iterative equalization in non-linear satellite channels," in *Proc. Int. Symp. Turbo Codes Iterative Inf. Process. (ISTC)*, 2012, pp. 220–224.

[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] Z. Long, H. Wang, N. Wu, W. Song, and D. Yang, "LMMSE based turbo equalization for nonlinear memory channel," in *Proc. Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2016, pp. 1–5.

[16] Z. Long, H. Wang, N. Wu, and J. Kuang, "Turbo equalization based on joint Gaussian, SIC-MMSE and LMMSE for nonlinear satellite channels," *Sci. China Inf. Sci.*, vol. 61, no. 4, p. 042301, 2018.

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

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

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

[20] S. Tong, D. Huang, Q. Guo, Y. Yu, and J. Xi, "Low complexity optimal soft-input soft-output demodulation of MSK based on factor graph," *IEEE Commun. Lett.*, vol. 18, no. 7, pp. 1139–1142, Jul. 2014.

[21] G. Colavolpe and G. Germi, "On the application of factor graphs and the sum-product algorithm to ISI channels," *IEEE Trans. Commun.*, vol. 53, no. 5, pp. 818–825, May 2005.

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

[23] Q. Guo and L. Ping, "LMMSE turbo equalization based on factor graphs," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 2, pp. 311–319, Feb. 2008.

[24] G. Colavolpe, D. Fertonani, and A. Piemontese, "SISO detection over linear channels with linear complexity in the number of interferers," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 8, pp. 1475–1485, Dec. 2011.

[25] P. Sun, C. Zhang, Z. Wang, C. N. Manchón, and B. H. Fleury, "Iterative receiver design for ISI channels using combined belief- and expectation-propagation," *IEEE Signal Process. Lett.*, vol. 22, no. 10, pp. 1733–1737, Oct. 2015.

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

[27] 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.

[28] M.-C. Chiu, Y.-C. Chen, and Y. T. Su, "Turbo equalization of nonlinear TDMA satellite signals," in *Proc. IEEE Global Telecomm. Conf.*, vol. 3, Nov. 2002, pp. 2860–2864.

[29] H. Wymeersch, *Iterative Receiver Design*. Cambridge, U.K.: Cambridge Univ. Press, 2007.

[30] 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.

[31] C. N. Manchón, G. E. Kirkelund, E. Riegler, L. P. B. Christensen, and B. H. Fleury. (Nov. 2011). "Receiver architectures for MIMO-OFDM based on a combined VMP-SP algorithm." [Online]. Available: https://arxiv.org/abs/1111.5848

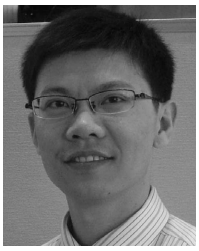
[32] E. Riegler, G. E. Kirkelund, C. N. Manchón, M.-A. Badiu, and B. H. Fleury, "Merging belief propagation and the mean field approximation: A free energy approach," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 588–602, Jan. 2013.

[33] Z. Yuan, C. Zhang, Z. Wang, Q. Guo, and J. Xi, "An auxiliary variable-aided hybrid message passing approach to joint channel estimation and decoding for MIMO-OFDM," *IEEE Signal Process. Lett.*, vol. 24, no. 1, pp. 12–16, Jan. 2017.

[34] 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.

[35] E. Riegler, G. E. Kirkelund, C. N. Manchón, M.-A. Badiu, and B. H. Fleury, "Merging belief propagation and the mean field approximation: A free energy approach," in *Proc. Int. Symp. Turbo Codes Iterative Inf. Process. (ISTC)*, 2010, pp. 256–260.
[36] C. P. Robert, *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*. New York, NY, USA: Springer-Verlag, 2001.
[37] F. Nielsen and V. Garcia. (Nov. 2009). "Statistical exponential families: A digest with flash cards." [Online]. Available: https://arxiv.org/abs/0911.4863
[38] J. Reina-Tosina, M. Allegue-Martínez, C. Crespo-Cadenas, C. Yu, and S. Cruces, "Behavioral modeling and predistortion of power amplifiers under sparsity hypothesis," *IEEE Trans. Microw. Theory Techn.*, vol. 63, no. 2, pp. 745–753, Feb. 2015.
[39] A. Winkelbauer. (Sep. 2012). "Moments and absolute moments of the normal distribution." [Online]. Available: https://arxiv.org/abs/1209.4340
[40] 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.

ZHEREN LONG received the B.S. degree from Nanchang University, Jiangxi, China, in 2010, and the Ph.D. degree from the Beijing Institute of Technology, Beijing, China, in 2017. He currently holds the position of engineer with the Institute of Telecommunication Satellite, China Academy of Space Technology, Beijing. His research interests include parameter estimation, channel predistortion, equalization, and system testing in satellite communication.



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

HUA WANG received the Ph.D. degree from the Beijing Institute of Technology (BIT), Beijing, China, in 1999. He is currently a Professor with the School of Information and Electronics, BIT. From 2009 to 2010, he was a Visiting Professor with the Department of Electrical Engineering, Arizona State University, USA. His research interests are in the fields of communication theory and signal processing, wireless networking, modem design, and implementation for satellite communication.

QINGHUA GUO (S'07–M'08) received the B.E. degree in electronic engineering and the M.E. degree in signal and information processing from Xidian University, Xian, China, in 2001 and 2004, respectively, and the Ph.D. degree in electronic engineering from the City University of Hong Kong, Hong Kong, in 2008. He is currently a Senior Lecturer with the School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Wollongong, NSW, Australia, and also an Adjunct Associate Professor with the School of Electrical, Electronic and Computer Engineering, The University of Western Australia, Perth, WA, Australia. His research interests include signal processing and telecommunications. He was a recipient of the Australian Research Council's Discovery Early Career Researcher Award.