SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

April 2018, Vol. 61 042301:1–042301:13
doi: 10.1007/s11432-016-9056-5

Turbo equalization based on joint Gaussian, SIC-MMSE and LMMSE for nonlinear satellite channels

Zheren LONG, Hua WANG*, Nan WU & Jingming KUANG

School of Information and Electronics, Beijing Institute of Technology, Beijing 100081*, China*

Received 18 October 2016/Accepted 21 March 2017/Published online 20 November 2017

Abstract The nonlinear distortion of wideband signal due to the filtering and efficiently operated high power amplifiers limits the performance of satellite communications. Volterra series can be used to describe the nonlinear satellite channels effectively. Most existing equalizers simply ignore the nonlinear terms or treat all the nonlinear combinations of symbols as interference. In this study, by properly exploiting information from nonlinear terms, we propose three turbo equalizers for nonlinear satellite channels, namely, joint Gaussian (JG), soft interference cancellation-minimum mean square error (SIC-MMSE) and linear minimum mean square error (LMMSE) equalizers. In JG and SIC-MMSE-based equalizers, both the linear and nonlinear terms that contain the symbol of interest are considered as desired signals. Accordingly, the required statistics are calculated based on the a priori probabilities of coded bits from output of channel decoder. For LMMSE-based equalizer, we propose to calculate the extrinsic information from output of equalizer by excluding the prior information in both the linear and nonlinear terms. Simulation results demonstrate that the proposed equalizers significantly outperform the method which ignores the presence of nonlinear interferences. Moreover, the nonlinear terms that contain the symbol of interest can be exploited to further improve the performance of turbo equalization.

Keywords nonlinear satellite channel, turbo equalization, joint Gaussian, soft interference cancellation-minimum mean square error (SIC-MMSE), linear minimum mean square error (LMMSE)

Citation Long Z R, Wang H, Wu N, et al. Turbo equalization based on joint Gaussian, SIC-MMSE and LMMSE for nonlinear satellite channels. Sci China Inf Sci, 2018, 61(4): 042301, doi: 10.1007/s11432-016-9056-5

1 Introduction

To meet the increased power demand, high power amplifiers (HPA), such as traveling-wave tube amplifiers (TWTA) and solid-state power amplifiers (SSPA), are employed in satellite communication systems [1]. In order to obtain better energy efficiency, these HPAs are driven to work as close as possible from the saturation point, which leads to severe nonlinear distortion. Together with pre-filter and post-filter, both linear inter-symbol interference (ISI) and nonlinear ISI degrade the performance of satellite communication systems [2]. It is shown in [3] that nonlinear satellite channels can be effectively modeled by Volterra series representation, which has been widely adopted in many researches [4–7].

Although predistortion technique can be used to mitigate the impact of nonlinear distortion at the earth station transmitter [8], due to the low signal-to-noise ratio (SNR) of the feedback loop, it is generally difficult to compensate the nonlinearity of HPA in the satellite directly. Alternatively, at the receiver

* Corresponding author (email: wanghua@bit.edu.cn)

side, equalization can be employed to eliminate the distortion. Many studies have been performed to develop equalizer in linear channels [9–17]. In conventional solutions, equalization and channel decoding work separately, which results in performance loss [18]. Thanks to the invention of Turbo codes and their iterative decoding methods, turbo equalization has been proposed in [19]. By combining a maximum likelihood sequence estimator (MLSE) or a maximum a posteriori (MAP) detector [20] with channel decoder, the performance of receiver has been proved to be near optimum in many transmission channels. However, the complexity of the above methods increases exponentially with the length of channel. To solve this problem, matched filter-soft interference cancellation (MF-SIC) equalizer [21], soft interference cancellation-minimum mean square error (SIC-MMSE) equalizer [22] and linear minimum mean square error (LMMSE) equalizer [23, 24] are investigated for turbo equalization. In [25], a joint Gaussian (JG) method is proposed for linear code division multiple access (CDMA) system, which is proved in [26] to be equivalent to LMMSE equalizer with a concise expression. Although the above equalizers can be applied in nonlinear channels by simply ignoring the nonlinear terms, this approximation will obviously lead to performance loss [27].

For nonlinear channels, a Volterra filter is proposed in [28] to compensate the signal distortion in uncoded satellite communication systems. In [29], the BCJR algorithm is employed to obtain the exact a posteriori probability of symbols. Although this method is optimal, the computational complexity is very high. Markov chain Monte Carlo methods are proposed in [27] to obtain the marginal posterior probability with lower computational and storage complexity. Nonlinear ISI estimator is employed in [30] to estimate and eliminate the interference. The hard decision used in equalizer leads to performance degradation. In [31], soft interference cancellation is adopted in the equalizer, which is an extension of [21], where cancellation filter is excited by not only the feedback symbols but also their nonlinear combinations. LMMSE-based turbo equalizer is studied in nonlinear satellite channel in [32]. However, the a priori probabilities of symbols are taking into account only for the linear terms.

In this paper, we study turbo equalization for nonlinear satellite channels described by Volterra model. Three equalizers based on JG, SIC-MMSE and LMMSE methods are considered, respectively. Different from existing works, both the JG-based and SIC-MMSE-based equalizers exploit information of interested symbol from not only the linear terms but also the nonlinear combinations. For LMMSE-based turbo equalizer, we propose to calculate the extrinsic information from the output of equalizer by fully taking into account the nonlinear terms. The performance gain of the proposed algorithms compared with the conventional and the state-of-the-art methods is evaluated by Monte Carlo simulations.

The organization of this paper is as follows. System model is introduced in Section 2. Three equalizers based on JG, SIC-MMSE and LMMSE are proposed in Section 3, respectively. Statistics are derived in detail and the information exchanged between equalizer and channel decoder are discussed. Simulation results and discussion are shown in Section 4. Finally, conclusion is drawn in Section 5.

2 System model

Consider a coded linearly modulated signal transmitted over a nonlinear satellite channel illustrated in Figure 1. Due to the high SNR of uplink, only downlink noise is taking into account. At the transmitter, the information bit sequence $\boldsymbol{b} = [b_1, b_2, \ldots, b_K]^{\mathrm{T}} \in \{0, 1\}^K$ is encoded by a channel encoder with coding rate $r = K/M$, yielding coded bit sequence $\boldsymbol{c} = [c_1, c_2, \ldots, c_M]^{\mathrm{T}} \in \{0, 1\}^M$. Then, the coded bits are interleaved to $\boldsymbol{c}^I = [c_1^I, c_2^I, \ldots, c_M^I]^{\mathrm{T}}$ and partitioned into subsequence $\boldsymbol{c}_n^I = [c_{n,1}^I, c_{n,2}^I, \ldots, c_{n,P}^I]^{\mathrm{T}}$. Each $\boldsymbol{c}_n^I$ is mapped to a symbol $x_n \in \chi$, where χ is the constellation set $\{s_i\}_{i=1,2,\ldots,2^P}$, with size $\mathcal{M}$. After pulse shaping of symbol sequence $\boldsymbol{x} = [x_1, x_2, \ldots, x_N]^{\mathrm{T}}$, the signal is sent to the channel.

The nonlinear satellite channel consists of pre-filter, power amplifier and post-filter. Pre-filter is used to remove the interference from adjacent channels, while post-filter restrains the extended spectrum caused by the nonlinearity of power amplifier. Volterra model can be used to describe the characteristic of

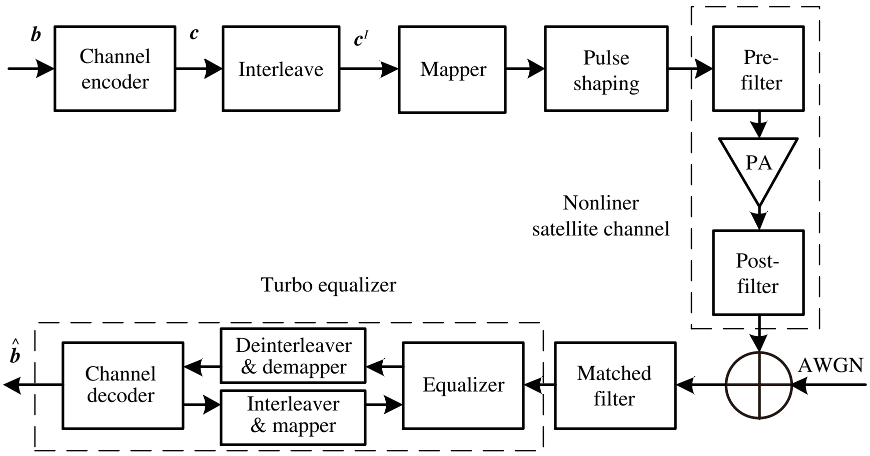

Figure 1 Block diagram of the considered satellite communication system.

nonlinear satellite channel, i.e.,

$$y(n)=\sum_{v=1}^{\infty}\sum_{n_1}\cdots\sum_{n_{2v-1}} h^{2v-1}_{n_1,\ldots,n_{2v-1}}x(n-n_1)\cdots x(n-n_v)x^*(n-n_{v+1})\cdots x^*(n-n_{2v-1})+w(n), \quad (1)$$

where $x(n)$ and $y(n)$ are, respectively, the input and output baseband equivalent signal, $h^{2v-1}_{n_1,\ldots,n_{2v-1}}$ is the kernel of Volterra series, $w(n)$ is the circularly-symmetric white Gaussian noise with zero mean and variance σ^2. Since even-order terms fall out of frequency band of interest [5], only odd terms are reserved in (1). Generally, a third order Volterra model is adequate to describe a mild nonlinear channel. Then, Eq. (1) becomes

$$y(n)=\sum_{l=0}^{L_0}h(l)x(n-l)+\sum_i\sum_j\sum_k h_{ijk}x(n-i)x(n-j)x^*(n-k)+w(n), \quad (2)$$

where L_0 is the linear ISI length of the channel[1)]. The channel model in (2) can be written in a matrix form, i.e.,

$$\boldsymbol{y}=\boldsymbol{H}_0\boldsymbol{x}_0+\underbrace{\sum_i\sum_j\sum_k \boldsymbol{H}_{ijk}\boldsymbol{x}_{ijk}}_{\text{nonlinear interference}}+\boldsymbol{w}, \quad (3)$$

where $\boldsymbol{y}=[y_1,y_2,\ldots,y_{N+L}]^{\mathrm{T}}$, $\boldsymbol{x}_0=[x_1,x_2,\ldots,x_N]^{\mathrm{T}}$, $\boldsymbol{x}_{ijk}=[x_{1-i}x_{1-j}x^*_{1-k},\ldots,x_{N-i}x_{N-j}x^*_{N-k}]^{\mathrm{T}}$, $\boldsymbol{H}_0$ is a Toeplitz matrix, which characterizes the linear ISI, the Volterra kernel matrix $\boldsymbol{H}_{ijk}$ is a diagonal matrix with element h_{ijk}, N is the block size. It is seen from (3) that, different from the traditional ISI channel, both linear and nonlinear interference have to be considered in nonlinear satellite channel.

3 Joint Gaussian, SIC-MMSE and LMMSE based turbo equalization for Volterra channel

Since the nonlinear channel modeled by Volterra series can be described by trellis, BCJR algorithm can be employed to obtain $p(\boldsymbol{y}|x_n)$. However, the computational complexity increases exponentially with the channel length. We propose three alternative algorithms based on JG, SIC-MMSE and LMMSE. In JG method, interference is approximated by additive Gaussian noise, hence, $p(\boldsymbol{y}|x_n)$ can be expressed in Gaussian form. In SIC-MMSE and LMMSE, the soft symbol estimation $\hat{x}_n$ is calculated first and then $p(\boldsymbol{y}|x_n)$ is approximated by $p(\hat{x}_n|x_n)$. We remark that, different from existing works, those nonlinear interference terms that contain the interested symbol x_n will be exploited.

1) The length of the nonlinear channel L is the maximum value of ISI length of linear and nonlinear part.

3.1 Joint Gaussian-based equalizer

We can rewrite (3) to characterize different terms as follows:

$$\boldsymbol{y} = \underbrace{\boldsymbol{h}_0^n x_0^n}_{\text{desired signal}} + \underbrace{\sum_{q\neq n}\boldsymbol{h}_0^q x_0^q}_{\text{linear ISI}} + \underbrace{\sum_{n_{ijk}}\sum_{i,j,k}\boldsymbol{h}_{ijk}^{n_{ijk}} x_{ijk}^{n_{ijk}}}_{\text{nonlinear interference with } x_n} + \underbrace{\sum_{q_{ijk}\neq n_{ijk}}\sum_{i,j,k}\boldsymbol{h}_{ijk}^{q_{ijk}} x_{ijk}^{q_{ijk}}}_{\text{nonlinear interference without } x_n} + \boldsymbol{w}, \tag{4}$$

where $\boldsymbol{h}_0^n$ and $\boldsymbol{h}_0^q$ are the nth and qth column of $\boldsymbol{H}_0$, respectively, $\boldsymbol{h}_{ijk}^{n_{ijk}}$ and $\boldsymbol{h}_{ijk}^{q_{ijk}}$ are the n_{ijk}th and q_{ijk}th column of $\boldsymbol{H}_{ijk}$, respectively, x_0^n, x_0^q, $x_{ijk}^{n_{ijk}}$ and $x_{ijk}^{q_{ijk}}$ are the nth, qth, n_{jik}th and q_{ijk}th elements in $\boldsymbol{x}_0$ and $\boldsymbol{x}_{ijk}$. n and n_{ijk} are the indexes of elements that contain x_n[2)].

It can be observed from (4) that both the linear term $\boldsymbol{h}_0^n x_0^n$ and nonlinear terms $\boldsymbol{h}_{ijk}^{n_{ijk}} x_{ijk}^{n_{ijk}}$ are related to x_n. Depending on whether the nonlinear terms $\boldsymbol{h}_{ijk}^{n_{ijk}} x_{ijk}^{n_{ijk}}$ is considered as useful signal or not, in the following, we will derive two JG-based equalizers.

When only the linear term $\boldsymbol{h}_0^n x_0^n$ is considered as useful signal, the other terms $\sum_{q\neq n}\boldsymbol{h}_0^q x_0^q + \sum_{i,j,k}\boldsymbol{H}_{ijk}\boldsymbol{x}_{ijk} + \boldsymbol{w}$ are taken as interference, which is assumed to be Gaussian distributed, with mean $\boldsymbol{m}_n$ and variance $\boldsymbol{v}_n$ given as

$$\boldsymbol{m}_n = \mathrm{E}\left(\sum_{q\neq n}\boldsymbol{h}_0^q x_0^q + \sum_{i,j,k}\boldsymbol{H}_{ijk}\boldsymbol{x}_{ijk} + \boldsymbol{w}\right) = \boldsymbol{H}_0\mathrm{E}(\boldsymbol{x}_0') + \sum_{i,j,k}\boldsymbol{H}_{ijk}\mathrm{E}(\boldsymbol{x}_{ijk}), \tag{5}$$

and

$$\begin{aligned}\boldsymbol{v}_n =& \mathrm{cov}\left(\sum_{q\neq n}\boldsymbol{h}_0^q x_0^q + \sum_{i,j,k}\boldsymbol{H}_{ijk}\boldsymbol{x}_{ijk} + \boldsymbol{w}, \sum_{q\neq n}\boldsymbol{h}_0^q x_0^q + \sum_{i,j,k}\boldsymbol{H}_{ijk}\boldsymbol{x}_{ijk} + \boldsymbol{w}\right)\\ =& \mathrm{cov}\left(\boldsymbol{H}_0\boldsymbol{x}_0' + \sum_{i,j,k}\boldsymbol{H}_{ijk}\boldsymbol{x}_{ijk}, \boldsymbol{H}_0\boldsymbol{x}_0' + \sum_{i,j,k}\boldsymbol{H}_{ijk}\boldsymbol{x}_{ijk}\right) + \sigma^2\boldsymbol{I}\\ =& \boldsymbol{H}_0\mathrm{cov}(\boldsymbol{x}_0', \boldsymbol{x}_0')\boldsymbol{H}_0^{\mathrm{H}} + \sum_{i,j,k}\boldsymbol{H}_0\mathrm{cov}(\boldsymbol{x}_0', \boldsymbol{x}_{ijk})\boldsymbol{H}_{ijk}^{\mathrm{H}} + \left(\sum_{i,j,k}\boldsymbol{H}_0\mathrm{cov}(\boldsymbol{x}_0', \boldsymbol{x}_{ijk})\boldsymbol{H}_{ijk}^{\mathrm{H}}\right)^{\mathrm{H}}\\ &+ \sum_{i,j,k}\sum_{i',j',k'}\boldsymbol{H}_{ijk}\mathrm{cov}(\boldsymbol{x}_{ijk}, \boldsymbol{x}_{i'j'k'})\boldsymbol{H}_{i'j'k'}^{\mathrm{H}} + \sigma^2\boldsymbol{I},\end{aligned} \tag{6}$$

where $\boldsymbol{x}_0' = [x_1, \ldots, x_{n-1}, 0, x_{n+1}, \ldots, x_N]^{\mathrm{T}}$.

Based on (4)–(6), we have

$$p(\boldsymbol{y}|x_n) \sim \mathcal{N}(\boldsymbol{h}_0^n x_n + \boldsymbol{m}_n, \boldsymbol{v}_n). \tag{7}$$

Since only the linear term $\boldsymbol{h}_0^n x_n$ is considered as useful signal, we refer this method as 'JG-L'.

On the other hand, when both the linear term $\boldsymbol{h}_0^n x_n$ and the nonlinear terms $\boldsymbol{h}_{ijk}^{n_{ijk}} x_{ijk}^{n_{ijk}}$ are considered as useful signal, the rest of the terms is regarded as interference. Similarly, we have the mean and covariance given as

$$\boldsymbol{m}_n = \mathrm{E}\left(\sum_{q\neq n}\boldsymbol{h}_0^q x_0^q + \sum_{q_{ijk}\neq n_{ijk}}\sum_{i,j,k}\boldsymbol{h}_{ijk}^{q_{ijk}} x_{ijk}^{q_{ijk}}\right), \tag{8}$$

and

$$\boldsymbol{v}_n = \mathrm{cov}\left(\sum_{q\neq n}\boldsymbol{h}_0^q x_0^q + \sum_{q_{ijk}\neq n_{ijk}}\sum_{i,j,k}\boldsymbol{h}_{ijk}^{q_{ijk}} x_{ijk}^{q_{ijk}} + \boldsymbol{w}, \sum_{q\neq n}\boldsymbol{h}_0^q x_0^q + \sum_{q_{ijk}\neq n_{ijk}}\sum_{i,j,k}\boldsymbol{h}_{ijk}^{q_{ijk}} x_{ijk}^{q_{ijk}} + \boldsymbol{w}\right)$$

2) The elements in $\boldsymbol{x}_{ijk}$ that contain x_n are in the form of $x_n x_j x_k^*$, $x_i x_n x_k^*$, $x_i x_j x_n^*$, $x_n x_n x_k^*$, $x_n x_j x_n^*$, $x_i x_n x_n^*$ and $x_n x_n x_n^*$, where i, j, k are integers but not n.

$$
\begin{aligned}
&=\mathrm{cov}\left(\boldsymbol{H}_0\boldsymbol{x}'_0+\sum_{i,j,k}\boldsymbol{H}_{ijk}\boldsymbol{x}'_{ijk},\boldsymbol{H}_0\boldsymbol{x}'_0+\sum_{i,j,k}\boldsymbol{H}_{ijk}\boldsymbol{x}'_{ijk}\right)+\sigma^2\boldsymbol{I}\\
&=\boldsymbol{H}_0\mathrm{cov}(\boldsymbol{x}'_0,\boldsymbol{x}'_0)\boldsymbol{H}_0^{\mathrm{H}}+\sum_{i,j,k}\boldsymbol{H}_0\mathrm{cov}(\boldsymbol{x}'_0,\boldsymbol{x}'_{ijk})\boldsymbol{H}_{ijk}^{\mathrm{H}}+\left(\sum_{i,j,k}\boldsymbol{H}_0\mathrm{cov}(\boldsymbol{x}'_0,\boldsymbol{x}'_{ijk})\boldsymbol{H}_{ijk}^{\mathrm{H}}\right)^{\mathrm{H}}\\
&\quad+\sum_{i,j,k}\sum_{i',j',k'}\boldsymbol{H}_{ijk}\mathrm{cov}(\boldsymbol{x}'_{ijk},\boldsymbol{x}'_{i'j'k'})\boldsymbol{H}_{i'j'k'}^{\mathrm{H}}+\sigma^2\boldsymbol{I},
\end{aligned}
\tag{9}
$$

where $\boldsymbol{x}'_{ijk}$ has the same form as $\boldsymbol{x}_{ijk}$ except that the elements which contain x_n are zero.

Based on (4), (8) and (9), we have

$$
p(\boldsymbol{y}|x_n)\sim\mathcal{N}\left(\boldsymbol{h}_0^n x_n+\sum_{n_{ijk}}\sum_{i,j,k}\boldsymbol{h}_{ijk}^{n_{ijk}}\mathrm{E}(x_{ijk}^{n_{ijk}}|x_n)+\boldsymbol{m}_n,\boldsymbol{v}_n\right). \tag{10}
$$

Since not only the linear term $\boldsymbol{h}_0^n x_0^n$ but also the nonlinear terms $\boldsymbol{h}_{ijk}^{n_{ijk}}x_{ijk}^{n_{ijk}}$ are considered as useful signal, we refer this method as 'JG-NL'.

The expectations and covariances in (5), (6), (8) and (9) are calculated based on the a priori probabilities (APPs) of coded bits given by the channel decoder. Moreover, the covariances are derived based on the expectations of symbols and their nonlinear combinations, e.g., $\mathrm{cov}(\boldsymbol{x}_{ijk},\boldsymbol{x}_{i'j'k'})=\mathrm{E}(\boldsymbol{x}_{ijk}\boldsymbol{x}_{i'j'k'}^{\mathrm{H}})-\mathrm{E}(\boldsymbol{x}_{ijk})\mathrm{E}(\boldsymbol{x}_{i'j'k'})^{\mathrm{H}}$. The expectations of matrix and vectors are calculated element by element. Taking $\mathrm{E}(\boldsymbol{x}_{ijk}\boldsymbol{x}_{i'j'k'}^{\mathrm{H}})$ as example, the variables are in the form of $x_ix_jx_k^*x_{i'}^*x_{j'}^*x_{k'}$. Due to the discrete characteristic of symbols and independent relationship between different symbols, $\mathrm{E}(x_ix_jx_k^*x_{i'}^*x_{j'}^*x_{k'})$ is calculated as follows:

$$
\mathrm{E}(x_ix_jx_k^*x_{i'}^*x_{j'}^*x_{k'})=\prod_m\mathrm{E}(x_m^u x_m^{*v}), \tag{11}
$$

$$
\mathrm{E}(x_m^u x_m^{*v})=\sum_{i=1}^{\mathcal{M}}p(x_m=s_i)s_i^u s_i^{*v}, \tag{12}
$$

$$
p(x_m=s_i)=\prod_p p(c_{m,p}^I), \tag{13}
$$

where $p(c_{m,p}^I)$ is the APP of $c_{m,p}^I$ that is fed back from the channel decoder, u and v are the numbers that x_m and x_m^* appear in the term $x_ix_jx_k^*x_{i'}^*x_{j'}^*x_{k'}$. Other expectations can be calculated in a similar way.

Given the expression of $p(\boldsymbol{y}|x_n)$ in (7) and (10), the extrinsic log-likelihood ratio (LLR) of coded bits from the JG-based equalizer and demapper to the deinterleaver is given by

$$
L^e(c_{n,p}^I|\boldsymbol{y})=\ln\frac{p(c_{n,p}^I=0|\boldsymbol{y})}{p(c_{n,p}^I=1|\boldsymbol{y})}-L(c_{n,p}^I)=\ln\frac{\sum_{\forall x_n:c_{n,p}^I=0}p(\boldsymbol{y}|x_n)\prod_{p'\neq p}p(c_{n,p'}^I=d_{n,p'})}{\sum_{\forall x_n:c_{n,p}^I=1}p(\boldsymbol{y}|x_n)\prod_{p'\neq p}p(c_{n,p'}^I=d_{n,p'})}, \tag{14}
$$

where $c_{n,p}^I=i$ means the pth bit of $\boldsymbol{c}_n^I$ is i, and $x_n:c_{n,p}^I=i$ refers to the symbol x_n corresponding to $\boldsymbol{c}_n^I$ whose pth bit is i.

3.2 SIC-MMSE-based equalizer

SIC-MMSE equalizer consists of two components, i.e., soft interference canceller and linear filter. The interference from other symbols are eliminated roughly by using their expectations, and then the residual interference is processed by a linear filter.

In a similar way to that of the JG-based equalizer, we are able to derive two SIC-MMSE equalizers depending on whether the nonlinear terms which contain x_n are regarded as useful signal or not.

When only the linear term related to x_n is considered as useful signal and all the nonlinear terms are considered as interference, the output of soft interference canceller is given by

$$\tilde{\boldsymbol{y}} = \boldsymbol{y} - \mathrm{E}\left(\sum_{q\neq n} \boldsymbol{h}_0^q x_0^q + \sum_{i,j,k} \boldsymbol{H}_{ijk}\boldsymbol{x}_{ijk}\right), \tag{15}$$

where the expectation is calculated based on the a priori information from the channel decoder.

A linear filter is then employed to eliminate the residual interference. The final soft decision symbol is

$$\hat{x}_n = \boldsymbol{c}^{\mathrm{H}}\tilde{\boldsymbol{y}}, \tag{16}$$

where $\boldsymbol{c}^{\mathrm{H}}$ is the coefficient vector given by

$$\boldsymbol{c}^{\mathrm{H}} = \mathrm{E}(x_n\tilde{\boldsymbol{y}}^{\mathrm{H}})\mathrm{E}(\tilde{\boldsymbol{y}}\tilde{\boldsymbol{y}}^{\mathrm{H}})^{-1}. \tag{17}$$

The first term $\mathrm{E}(x_n\tilde{\boldsymbol{y}}^{\mathrm{H}})$ in (17) is derived as

$$\begin{aligned}\mathrm{E}(x_n\tilde{\boldsymbol{y}}^{\mathrm{H}}) =& \mathrm{E}[x_n(\boldsymbol{H}_0(\boldsymbol{x}_0 - \mathrm{E}(\boldsymbol{x}_0')))^{\mathrm{H}}] + \mathrm{E}\left[x_n\sum_{i,j,k}(\boldsymbol{H}_{ijk}(\boldsymbol{x}_{ijk} - \mathrm{E}(\boldsymbol{x}_{ijk})))^{\mathrm{H}}\right] + \mathrm{E}(x_n\boldsymbol{w}^{\mathrm{H}})\\ =& \mathrm{E}(x_nx_n^*)\boldsymbol{s}_n^{\mathrm{H}}\boldsymbol{H}_0^{\mathrm{H}} + \sum_{i,j,k}\mathrm{cov}(x_n, \boldsymbol{x}_{ijk})\boldsymbol{H}_{ijk}^{\mathrm{H}},\end{aligned} \tag{18}$$

where $\boldsymbol{s}_n = [0,\ldots,0,1,0,\ldots,0]^{\mathrm{T}}$, i.e., the nth element is 1, others are all zero. The second term $\mathrm{E}(\tilde{\boldsymbol{y}}\tilde{\boldsymbol{y}}^{\mathrm{H}})$ in (17) can be derived as

$$\begin{aligned}\mathrm{E}(\tilde{\boldsymbol{y}}\tilde{\boldsymbol{y}}^{\mathrm{H}}) =& \mathrm{E}[(\boldsymbol{H}_0(\boldsymbol{x}_0 - \mathrm{E}(\boldsymbol{x}_0')))(\boldsymbol{H}_0(\boldsymbol{x}_0 - \mathrm{E}(\boldsymbol{x}_0')))^{\mathrm{H}}]\\ &+ \mathrm{E}\left[(\boldsymbol{H}_0(\boldsymbol{x}_0 - \mathrm{E}(\boldsymbol{x}_0')))\sum_{i,j,k}(\boldsymbol{H}_{ijk}(\boldsymbol{x}_{ijk} - \mathrm{E}(\boldsymbol{x}_{ijk})))^{\mathrm{H}}\right]\\ &+ \mathrm{E}\left[\sum_{i,j,k}(\boldsymbol{H}_{ijk}(\boldsymbol{x}_{ijk} - \mathrm{E}(\boldsymbol{x}_{ijk})))(\boldsymbol{H}_0(\boldsymbol{x}_0 - \mathrm{E}(\boldsymbol{x}_0')))^{\mathrm{H}}\right]\\ &+ \sum_{i,j,k}\sum_{i',j',k'}\boldsymbol{H}_{ijk}\mathrm{cov}(\boldsymbol{x}_{ijk}, \boldsymbol{x}_{i'j'k'})\boldsymbol{H}_{i'j'k'}^{\mathrm{H}} + \sigma^2\boldsymbol{I}\\ =& \boldsymbol{H}_0\boldsymbol{V}_0'\boldsymbol{H}_0^{\mathrm{H}} + \left(\sum_{i,j,k}\boldsymbol{H}_0\mathrm{cov}(\boldsymbol{x}_0, \boldsymbol{x}_{ijk})\boldsymbol{H}_{ijk}^{\mathrm{H}}\right) + \left(\sum_{i,j,k}\boldsymbol{H}_0\mathrm{cov}(\boldsymbol{x}_0, \boldsymbol{x}_{ijk})\boldsymbol{H}_{ijk}^{\mathrm{H}}\right)^{\mathrm{H}}\\ &+ \sum_{i,j,k}\sum_{i',j',k'}\boldsymbol{H}_{ijk}\mathrm{cov}(\boldsymbol{x}_{ijk}, \boldsymbol{x}_{i'j'k'})\boldsymbol{H}_{i'j'k'}^{\mathrm{H}} + \sigma^2\boldsymbol{I},\end{aligned} \tag{19}$$

where $\boldsymbol{V}_0' = \mathrm{diag}[\mathrm{var}(x_1),\ldots,\mathrm{var}(x_{n-1}),\mathrm{E}(|x_n|^2),\mathrm{var}(x_{n+1}),\ldots,\mathrm{var}(x_N)]$.

Combining (15)–(19), we can obtain the soft symbol estimation $\hat{x}_n$. We refer this method as 'SIC-MMSE-L'.

As we have discussed, nonlinear terms $\boldsymbol{h}_{ijk}^{n_{ijk}}x_{ijk}^{n_{ijk}}$ also contain x_n and therefore can be reserved in useful signal[3]. Based on this observation, the interference canceller can be written as

$$\tilde{\boldsymbol{y}} = \boldsymbol{y} - \mathrm{E}\left(\sum_{q\neq n}\boldsymbol{h}_0^q x_0^q + \sum_{q_{ijk}\neq n_{ijk}}\sum_{i,j,k}\boldsymbol{h}_{ijk}^{q_{ijk}}\boldsymbol{x}_{ijk}^{q_{ijk}}\right). \tag{20}$$

3) Note that nonlinear terms that have only one x_n will be reserved.

Then, the two terms to calculate the coefficient vector $\boldsymbol{c}$ in (17) are given by

$$\mathrm{E}(x_n\tilde{\boldsymbol{y}}^{\mathrm{H}}) = \mathrm{E}(|x_n|^2)(\boldsymbol{H}_0\boldsymbol{s}_n)^{\mathrm{H}} + \sum_{i,j,k}\triangle_{n,ijk}\boldsymbol{H}_{ijk}^{\mathrm{H}}, \tag{21}$$

and

$$\begin{aligned}\mathrm{E}(\tilde{\boldsymbol{y}}\tilde{\boldsymbol{y}}^{\mathrm{H}}) = \boldsymbol{H}_0\boldsymbol{V}_0'\boldsymbol{H}_0^{\mathrm{H}} + \sum_{i,j,k}\boldsymbol{H}_0\triangle_{0,ijk}\boldsymbol{H}_{ijk}^{\mathrm{H}} + \left(\sum_{i,j,k}\boldsymbol{H}_0\triangle_{0,ijk}\boldsymbol{H}_{ijk}^{\mathrm{H}}\right)^{\mathrm{H}} \\ + \sum_{i,j,k}\sum_{i',j',k'}\boldsymbol{H}_{ijk}\triangle_{ijk,i'j'k'}\boldsymbol{H}_{i'j'k'}^{\mathrm{H}} + \sigma^2\boldsymbol{I},\end{aligned} \tag{22}$$

where $\triangle_{n,ijk} = \mathrm{E}[(x_n)(\boldsymbol{x}_{ijk} - \mathrm{E}(\boldsymbol{x}'_{ijk}))]$, $\triangle_{0,ijk} = \mathrm{E}[(\boldsymbol{x}_0 - \mathrm{E}(\boldsymbol{x}'_0))(\boldsymbol{x}_{ijk} - \mathrm{E}(\boldsymbol{x}'_{ijk}))^{\mathrm{H}}]$, and $\triangle_{ijk,i'j'k'} = \mathrm{E}[(\boldsymbol{x}_{ijk} - \mathrm{E}(\boldsymbol{x}'_{ijk}))(\boldsymbol{x}_{i'j'k'} - \mathrm{E}(\boldsymbol{x}'_{i'j'k'}))^{\mathrm{H}}]$. The element having only one x_n in $\mathrm{E}(\boldsymbol{x}'_{ijk})$ and $\mathrm{E}(\boldsymbol{x}'_{i'j'k'})$ is zero, and others are calculated based on the APPs of coded bits obtained from channel decoder.

Combining (16), (17), and (20)–(22), we name this method 'SIC-MMSE-NL', since nonlinear terms that contain x_n are also exploited.

We approximate the conditional distribution of SIC-MMSE equalizer output as Gaussian distribution [33], i.e.,

$$p(\hat{x}_n|x_n) \sim \mathcal{N}(\mathrm{E}(\hat{x}_n|x_n), \mathrm{var}(\hat{x}_n|x_n)), \tag{23}$$

where the mean and variance are expressed as

$$\mathrm{E}(\hat{x}_n|x_n) = \boldsymbol{c}^{\mathrm{H}}\mathrm{E}(\tilde{\boldsymbol{y}}|x_n), \tag{24}$$

and

$$\begin{aligned}\mathrm{var}(\hat{x}_n|x_n) &= \mathrm{E}(\hat{x}_n\hat{x}_n^{\mathrm{H}}|x_n) - \mathrm{E}(\hat{x}_n|x_n)\mathrm{E}(\hat{x}_n|x_n)^{\mathrm{H}} \\ &= \boldsymbol{c}^{\mathrm{H}}\mathrm{E}(\tilde{\boldsymbol{y}}\tilde{\boldsymbol{y}}^{\mathrm{H}}|x_n)\boldsymbol{c} - |\mathrm{E}(\hat{x}_n|x_n)|^2.\end{aligned} \tag{25}$$

Then, the extrinsic information from SIC-MMSE equalizer to the demapper and deinterleaver is given by

$$L^e(c_{n,p}^I|\hat{x}_n) = \ln\frac{p(c_{n,p}^I = 0|\hat{x}_n)}{p(c_{n,p}^I = 1|\hat{x}_n)} - L(c_{n,p}^I) = \ln\frac{\sum_{\forall x_n:c_{n,p}^I=0} p(\hat{x}_n|x_n)\prod_{p'\neq p} p(c_{n,p'}^I = d_{n,p'})}{\sum_{\forall x_n:c_{n,p}^I=1} p(\hat{x}_n|x_n)\prod_{p'\neq p} p(c_{n,p'}^I = d_{n,p'})}. \tag{26}$$

3.3 LMMSE-based equalizer

LMMSE equalizer is an affine transform [23,34], i.e., $\hat{x}_n = \boldsymbol{a}_n^{\mathrm{H}}\boldsymbol{y} + b_n$, where $\boldsymbol{a}_n^{\mathrm{H}}$ is the coefficient vector and b_n is the bias. In linear Gaussian model, when the symbol obeys Gaussian distribution, it is equivalent to minimum mean square error (MMSE)/MAP detector. Specifically, when b_n is set to zero, it becomes Wiener equalizer. The expressions of LMMSE equalizer and Wiener equalizer are given as follows:

For LMMSE equalizer,

$$\hat{x}_n = \mathrm{E}(x_n) + \mathrm{cov}(x_n, \boldsymbol{y})\mathrm{cov}(\boldsymbol{y}, \boldsymbol{y})^{-1}(\boldsymbol{y} - \mathrm{E}(\boldsymbol{y})). \tag{27}$$

For Wiener equalizer,

$$\hat{x}_n = \mathrm{E}(x_n\boldsymbol{y}^{\mathrm{H}})\mathrm{E}(\boldsymbol{y}\boldsymbol{y}^{\mathrm{H}})^{-1}\boldsymbol{y}. \tag{28}$$

In turbo equalization, the extrinsic information from output of equalizer should be calculated based on $\hat{x}_n$ that does not depend on $L(c_{n,q}^I)$ [23]. Therefore, based on turbo principle, Eq. (27) can be rewritten as

$$\hat{x}_n = \boldsymbol{a}_n^{\mathrm{H}}(\boldsymbol{y} - \mathrm{E}(\boldsymbol{y}))|_{p(x_n)=1/\mathcal{M}} = \mathrm{cov}(x_n, \boldsymbol{y})\mathrm{cov}(\boldsymbol{y}, \boldsymbol{y})^{-1}(\boldsymbol{y} - \mathrm{E}(\boldsymbol{y}))|_{p(x_n)=1/\mathcal{M}}, \tag{29}$$

where the expectation and covariance are calculated based on the condition that the possible values of x_n are equal-probable.

We derive the statistics in (29) term by term. Different from the linear ISI channel discussed in [23], due to the nonlinear combination of symbols in $\boldsymbol{y}$, there is no concise expression for these statistics.

The first term $\mathrm{cov}(x_n, \boldsymbol{y})$ is given by

$$\begin{aligned}\mathrm{cov}(x_n, \boldsymbol{y})|_{p(x_n)=1/\mathcal{M}} &= \left.\left[\mathrm{cov}(x_n, \boldsymbol{H}_0\boldsymbol{x}_0) + \mathrm{cov}\left(x_n, \sum_{i,j,k} \boldsymbol{H}_{ijk}\boldsymbol{x}_{ijk}\right)\right]\right|_{p(x_n)=1/\mathcal{M}} \\ &= \boldsymbol{s}_n^{\mathrm{H}}\boldsymbol{H}_0^{\mathrm{H}} + \sum_{i,j,k} \mathrm{cov}(x_n, \boldsymbol{x}_{ijk})\boldsymbol{H}_{ijk}^{\mathrm{H}}\Big|_{p(x_n)=1/\mathcal{M}}. \end{aligned} \tag{30}$$

The second term is derived as

$$\begin{aligned}\mathrm{cov}(\boldsymbol{y}, \boldsymbol{y})|_{p(x_n)=1/\mathcal{M}} = &\left[\boldsymbol{H}_0\mathrm{cov}(\boldsymbol{x}_0, \boldsymbol{x}_0)\boldsymbol{H}_0^{\mathrm{H}} + \sum_{i,j,k} \boldsymbol{H}_0\mathrm{cov}(\boldsymbol{x}_0, \boldsymbol{x}_{ijk})\boldsymbol{H}_{ijk}^{\mathrm{H}} + \left(\sum_{i,j,k} \boldsymbol{H}_0\mathrm{cov}(\boldsymbol{x}_0, \boldsymbol{x}_{ijk})\boldsymbol{H}_{ijk}^{\mathrm{H}}\right)^{\mathrm{H}}\right. \\ &\left.\left. + \sum_{i,j,k}\sum_{i',j',k'} \boldsymbol{H}_{ijk}\mathrm{cov}(\boldsymbol{x}_{ijk}, \boldsymbol{x}_{i'j'k'})\boldsymbol{H}_{i'j'k'}^{\mathrm{H}} + \sigma^2\boldsymbol{I}\right]\right|_{p(x_n)=1/\mathcal{M}}, \end{aligned} \tag{31}$$

where $\mathrm{cov}(\boldsymbol{x}_0, \boldsymbol{x}_0)$ is a diagonal matrix, the nonzero elements are the corresponding variance of symbols.

The last term in (29) is derived as

$$\boldsymbol{y} - \mathrm{E}_{p(x_n)=1/\mathcal{M}}(\boldsymbol{y}) = \boldsymbol{y} - \boldsymbol{H}_0\mathrm{E}(\boldsymbol{x}_0') - \sum_{i,j,k} \boldsymbol{H}_{ijk}\mathrm{E}(\boldsymbol{x}_{ijk})|_{p(x_n)=1/\mathcal{M}}. \tag{32}$$

Substituting (30)–(32) into (29), we have the LMMSE equalizer for nonlinear channel. Note that for each x_n, the statistics related to the nonlinear combination of symbols in (30)–(32), e.g., $\mathrm{cov}(\boldsymbol{x}_0, \boldsymbol{x}_{ijk})$, have to be recalculated by setting $p(x_n) = 1/\mathcal{M}$. To reduce complexity, one can ignore this constraint and using prior information from channel decoder output directly for the nonlinear terms. Accordingly, for different x_n, the parts of the statistics related to the nonlinear terms are not needed to be recalculated. This approximation leads to the result in [32], which will be also evaluated in Section 4.

In a similar way, we can calculate the statistics in (28) for Wiener equalizer, i.e.,

$$\mathrm{E}(x_n\boldsymbol{y}_n^{\mathrm{H}})|_{p(x_n)=1/\mathcal{M}} = \left.\left[\mathrm{E}(x_n\boldsymbol{x}_0^{\mathrm{H}})\boldsymbol{H}_0^{\mathrm{H}} + \sum_{i,j,k} \mathrm{E}(x_n\boldsymbol{x}_{ijk}^{\mathrm{H}})\boldsymbol{H}_{ijk}^{\mathrm{H}}\right]\right|_{p(x_n)=1/\mathcal{M}}, \tag{33}$$

and

$$\begin{aligned}\mathrm{E}(\boldsymbol{y}_n\boldsymbol{y}_n^{\mathrm{H}})|_{p(x_n)=1/\mathcal{M}} = &\left[\boldsymbol{H}_0\mathrm{E}(\boldsymbol{x}_0\boldsymbol{x}_0^{\mathrm{H}})\boldsymbol{H}_0^{\mathrm{H}} + \sum_{i,j,k} \boldsymbol{H}_0\mathrm{E}(\boldsymbol{x}_0\boldsymbol{x}_{ijk}^{\mathrm{H}})\boldsymbol{H}_{ijk}^{\mathrm{H}} + \left(\sum_{i,j,k} \boldsymbol{H}_0\mathrm{E}(\boldsymbol{x}_0\boldsymbol{x}_{ijk}^{\mathrm{H}})\boldsymbol{H}_{ijk}^{\mathrm{H}}\right)^{\mathrm{H}}\right. \\ &\left.\left. + \sum_{i,j,k}\sum_{i',j',k'} \boldsymbol{H}_{ijk}\mathrm{E}(\boldsymbol{x}_{ijk}\boldsymbol{x}_{i'j'k'}^{\mathrm{H}})\boldsymbol{H}_{i'j'k'}^{\mathrm{H}} + \sigma^2\boldsymbol{I}\right]\right|_{p(x_n)=1/\mathcal{M}}. \end{aligned} \tag{34}$$

As in [23, 35], we approximate $p(\hat{x}_n|x_n)$ as Gaussian distribution. Then, for LMMSE equalizer, the mean and covariance are given as follows:

$$\mathrm{E}(\hat{x}_n|x_n) = \boldsymbol{a}_n^{\mathrm{H}}(\mathrm{E}(\boldsymbol{y}_n|x_n) - \mathrm{E}(\boldsymbol{y}_n)|_{p(x_n)=1/\mathcal{M}}), \tag{35}$$

$$\mathrm{cov}(\hat{x}_n, \hat{x}_n|x_n) = \boldsymbol{a}_n^{\mathrm{H}}\mathrm{cov}(\boldsymbol{y}_n, \boldsymbol{y}_n|x_n)\boldsymbol{a}_n. \tag{36}$$

Table 1 Computational complexity per symbol per iteration

Algorithm	BCJR	LMMSE [23]	LMMSE [32]	Proposed JG	Proposed SIC-MMSE	Proposed LMMSE
Complexity	$\mathcal{O}(\mathcal{M}^L)$	$\mathcal{O}(S^3+S\mathcal{M})$	$\mathcal{O}(S^3+\mathcal{L}^2S^2K\mathcal{M})$	$\mathcal{O}(S^3+\mathcal{L}^2S^2K\mathcal{M})$	$\mathcal{O}(S^3+\mathcal{L}^2S^2K\mathcal{M}^2)$	$\mathcal{O}(S^3+\mathcal{L}^2S^2K\mathcal{M}^2)$

For Wiener equalizer, the mean and covariance are

$$\mathrm{E}(\hat{x}_n|x_n) = \boldsymbol{a}_n^{\mathrm{H}}\mathrm{E}(\boldsymbol{y}_n|x_n), \tag{37}$$

$$\mathrm{cov}(\hat{x}_n, \hat{x}_n|x_n) = \boldsymbol{a}_n^{\mathrm{H}}\mathrm{cov}(\boldsymbol{y}_n, \boldsymbol{y}_n|x_n)\boldsymbol{a}_n. \tag{38}$$

Extrinsic information can be achieved by taking (35)–(38) into (26). We can find that the expectation and covariance should be recalculated for each constellation point, which is too complex for high order modulation. Inspired by [23,35], $p(\hat{x}_n|x_n)$ can be approximated by the probability density function of the output error, $p(\hat{x}_n|x_n) \approx p(e_n|x_n)$, where $e_n = \hat{x}_n - kx_n$. After some derivations and approximations, we can find that, with $p(x_n) = 1/\mathcal{M}$, $k = \mathrm{cov}(\hat{x}_n, x_n) = \boldsymbol{a}_n^{\mathrm{H}}\mathrm{cov}(\boldsymbol{y}, x_n) = \mathrm{cov}(x_n, \boldsymbol{y})\mathrm{cov}(\boldsymbol{y}, \boldsymbol{y})^{-1}\mathrm{cov}(\boldsymbol{y}, x_n)$, and the corresponding mean and variance of e_n are given as follows:

$$\mathrm{E}(e_n) = 0, \tag{39}$$

$$\mathrm{var}(e_n) = \mathrm{cov}(\hat{x}_n, \hat{x}_n) - \mathrm{cov}(\hat{x}_n, x_n)\mathrm{cov}(\hat{x}_n, x_n)^{\mathrm{H}} = \mathrm{cov}(\hat{x}_n, x_n)(1 - \mathrm{cov}(\hat{x}_n, x_n)^{\mathrm{H}}). \tag{40}$$

In fact, the covariance matrixes have been obtained during calculating the soft decision symbol $\hat{x}_n$. By simply replacing k with $\mathrm{E}(x_n\tilde{\boldsymbol{y}}^{\mathrm{H}})\mathrm{E}(\tilde{\boldsymbol{y}}\tilde{\boldsymbol{y}}^{\mathrm{H}})^{-1}\mathrm{E}(x_n\tilde{\boldsymbol{y}}^{\mathrm{H}})^{\mathrm{H}}$, this method is also applicable to SIC-MMSE-based equalizer and the corresponding $p(\hat{x}_n|x_n)$ in (26) is Gaussian distribution with mean kx_n and variance $k(1-k)^*$.

3.4 Complexity analysis

Similar to the traditional LMMSE equalizer in [23, 32], sliding window is used in the proposed methods to reduce the complexity. It is shown in [36] that, for a suitable window size S, performance degradation due to the sliding window can be ignored. Only the computational complexity per symbol per iteration of equalizer is considered here.

Consider a Kth order nonlinear Volterra channel with memory length L and the number of nonlinear terms is $\mathcal{L}$. The size of constellation is $\mathcal{M}$. The complexity of BCJR algorithm depends on the number of trellis states, which is $\mathcal{O}(\mathcal{M}^L)$ per symbol per iteration. The complexity of other algorithms consists of two parts: calculating inverse matrix and statistical matrices. Only taking the linear terms into consideration, the complexity of the traditional LMMSE equalizer [23] is $\mathcal{O}(S^3 + S\mathcal{M})$ per symbol per iteration, where $\mathcal{O}(S^3)$ arises from the calculation of inverse matrix and $\mathcal{O}(S\mathcal{M})$ is due to the calculation of statistical matrix. The complexity of JG-based equalizer is the same as the LMMSE equalizer in [32], i.e., $\mathcal{O}(S^3+\mathcal{L}^2S^2K\mathcal{M})$ per symbol per iteration. Where $\mathcal{L}^2$ is the number of statistical matrices related to pure nonlinear terms. The complexity of exact SIC-MMSE and LMMSE-based equalizer is $\mathcal{O}(S^3+\mathcal{L}^2S^2K\mathcal{M}^2)$ per symbol per iteration, when the approximation method is used for calculating extrinsic information, it reduces to $\mathcal{O}(S^3+\mathcal{L}^2S^2K\mathcal{M})$. The computational complexities of different algorithms are summarized in Table 1.

3.5 Relationship between different equalizers

In linear ISI channel, JG and LMMSE equalizers are proved to be equivalent [26]. For nonlinear memory channel, however, no such equivalence holds due to the nonlinear interferences. It is difficult to find any direct relationships between JG equalizer and others. As to SIC-MMSE and LMMSE equalizers, they can be written in the form of $\hat{x} = \boldsymbol{a}^{\mathrm{H}}(\boldsymbol{y} - \boldsymbol{b})$, where $\boldsymbol{a}^{\mathrm{H}}$ is the coefficient vector and $\boldsymbol{b}$ is the bias vector. Both of them can be considered as a canceller followed by a filter. Therefore, the same structure can be used for the implementation of LMMSE and SIC-MMSE equalizers, but with different cancellation terms and coefficients of filters. Also, we can find that the proposed SIC-MMSE-L method is equivalent to the LMMSE equalizer in [32] when the prior probability of decision symbol is set to be equally-likely only for the linear terms.

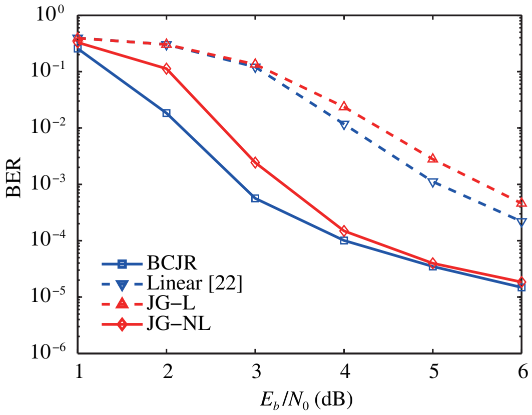

Figure 2 (Color online) BER performance of the proposed JG-based turbo equalizers for BPSK signal.

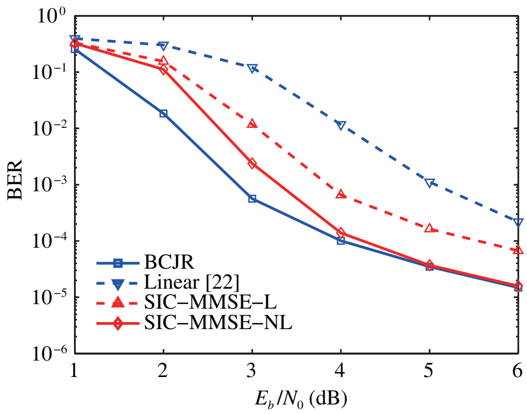

Figure 3 (Color online) BER performance of the proposed SIC-MMSE-based turbo equalizers for BPSK signal.

4 Simulation results

The proposed turbo equalizers based on JG, SIC-MMSE, and LMMSE are evaluated by means of Monte Carlo simulations. The number of information bits in each frame is set to 1024, the rate-1/2 (117, 155) convolutional code with truncated termination is used for channel coding. 16-random interleaver is employed to scramble the coded bits in order to reduce burst error and produce the independence between symbols.

The nonlinear Volterra model is $y_n = 0.407(n)+0.815(x_n-1)+0.407x(n-2)+0.4x(n)x(n-1)x^*(n-2)$. The memory depth of channel is 2, and the size of sliding window is 11. The number of turbo iteration is 4, unless otherwise specified.

Bit error rate (BER) performance of the proposed JG-based turbo equalizers is evaluated in Figure 2 for BPSK signal. For comparison purpose, the results by BCJR algorithm and that by ignoring the nonlinear interference [23, 27] (named as 'Linear') are also plotted. It can be observed that the proposed JG-NL equalizer, which takes the nonlinear terms that contain the symbol to be detected as useful signal, performs very close to BCJR algorithm at high SNRs and significantly outperforms the 'Linear' method. Interestingly, although nonlinear channel instead of linear channel is used and the statistics of nonlinear interference is considered, JG-L equalizer performs worse than the 'Linear' method. This may be due to the large deviation of the mean and covariance obtained in JG-L.

BER performance of the proposed SIC-MMSE and LMMSE equalizers using exact extrinsic information for BPSK signal is illustrated in Figures 3 and 4, respectively. It is seen that all the proposed algorithms have superior performance than the 'Linear' method, which reveals that not only linear ISI but also nonlinear interference have been eliminated. By properly reserving the nonlinear terms as useful signal in the interference cancellation step, about 1dB performance gain can be obtained by using SIC-MMSE-NL compared with SIC-MMSE-L at BER $= 10^{-4}$. In Figure 4, we can observe that the proposed LMMSE and Wiener equalizers have superior performance than the LMMSE equalizer in [32] at medium to high SNRs. This is due to the fact that ref. [32] excludes only the prior information of the linear term when calculating the extrinsic information from the output of equalizer, which is an approximation of the proposed algorithms. Moreover, LMMSE equalizer outperforms Wiener equalizer via including nonzero bias. We remark that although LMMSE and Wiener equalizers are based on linear transforming, when they are properly designed for Volterra channel, nonlinear interference can also be eliminated.

Convergence behavior for the proposed algorithms is shown in Figure 5 for BPSK signal at $E_b/N_0 =$ 4 dB. We can see that all the proposed equalizers are able to converge. After 4 iterations, the performance gain by increasing the number of iterations becomes negligible.

As discussed in Section 3, for SIC-MMSE and LMMSE equalizers, the extrinsic information from equalizer to demapper has to be calculated for all the possible values of the symbol, which leads to

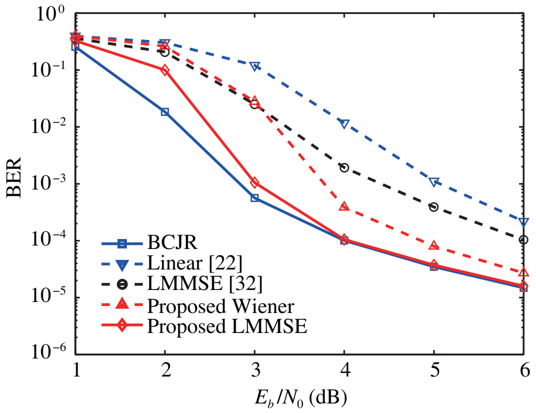

Figure 4 (Color online) BER performance of the proposed LMMSE-based turbo equalizers for BPSK signal.

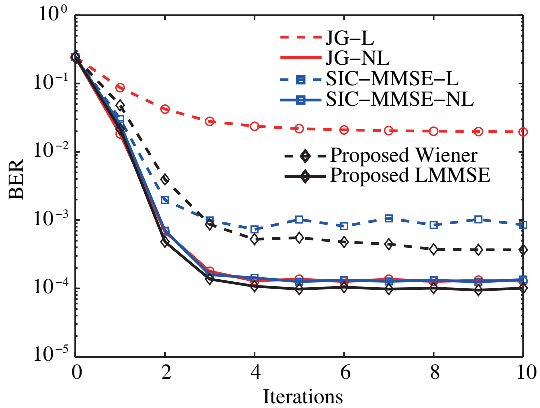

Figure 5 (Color online) Convergence behavior for the proposed algorithms for BPSK signal.

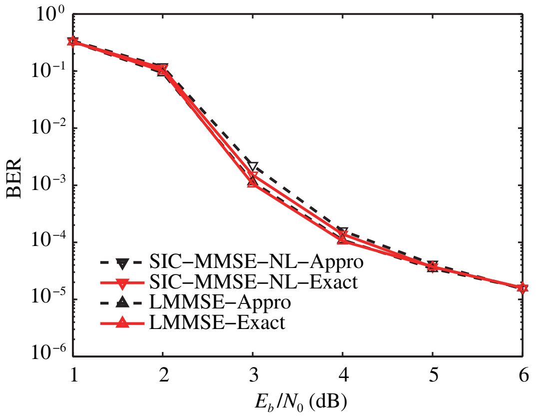

Figure 6 (Color online) Performance comparisons by using the approximation in (39) and (40) for BPSK signal.

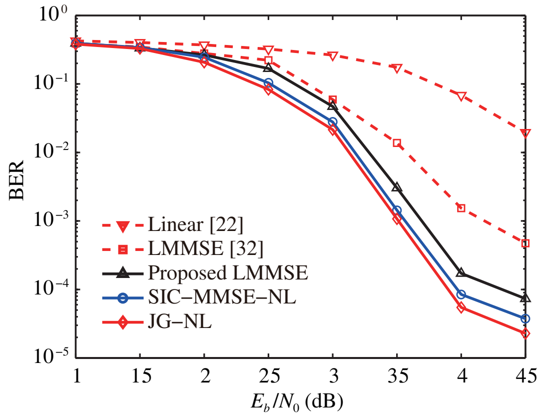

Figure 7 (Color online) BER performance of the proposed algorithms for QPSK signal.

high complexity for high order modulations. This problem can be solved by using (39) and (40) to approximate the distribution of the output error of equalizer. The performance of the proposed SIC-MMSE and LMMSE equalizers by using this approximation is evaluated in Figure 6 for BPSK signal, and compared with that using the exact solutions. It is seen that performance loss due to this approximation for SIC-MMSE and LMMSE equalizers is negligible, which provides a low complexity implementation in practice.

The proposed algorithms are evaluated for QPSK signal in Figure 7. For clarity purpose, only the equalizers that properly consider the information in nonlinear terms, i.e., JG-NL, SIC-MMSE-NL and the proposed LMMSE equalizers, are plotted. Extrinsic information obtained by approximation method is used for SIC-MMSE-NL and LMMSE equalizer in order to reduce the complexity. We can see that the proposed equalizers perform close to each other, although JG-NL slightly outperforms the other two. Moreover, the proposed equalizers have superior performance than [23], which simply ignores the nonlinear term, and also outperforms the LMMSE method in [32], which does not exclude the prior information of nonlinear terms when calculating of extrinsic information.

5 Conclusion

In this paper, we studied turbo equalization for nonlinear satellite channels represented by Volterra model. Based on the observation that the interested symbol is present not only in the linear terms but also in the nonlinear combinations of symbols, three equalizers based on JG, SIC-MMSE and LMMSE were proposed. Depending on whether the nonlinear terms that contain the interested symbol were treated as useful signals, JG-L, JG-NL, and SIC-MMSE-L, SIC-MMSE-NL were derived. For LMMSE-based equalizer, we proposed to calculate the extrinsic information from output of equalizer by excluding the prior information in both the linear and nonlinear terms. Based on this method, LMMSE and Wiener-based turbo equalizers were presented. Simulation results showed that all the proposed equalizers significantly outperformed the method that ignored the nonlinear terms. The proposed JG-NL and SIC-MMSE-NL have superior performance than JG-L and SIC-MMSE-L, which revealed that nonlinear terms that contain the symbol of interest should be considered as useful signals. Moreover, the proposed LMMSE and Wiener equalizers outperformed the existing LMMSE method, which verified the necessary of excluding the prior information in nonlinear terms when calculating the extrinsic information from output of equalizer.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 61471037, 61571041).

Conflict of interest The authors declare that they have no conflict of interest.

References

1 Benedetto S, Garello R, Montorsi G, et al. MHOMS: high-speed ACM modem for satellite applications. IEEE Wirel Commun, 2005, 12: 66–77
2 Casini E, Gaudenzi R D, Ginesi A. DVB-S2 modem algorithms design and performance over typical satellite channels. Int J Satell Commun Netw, 2004, 22: 281–318
3 Benedetto S, Biglieri E, Daffara R. Modeling and performance evaluation of nonlinear satellite links-a volterra series approach. IEEE Trans Aerosp Electron Syst, 1979, 4: 494–507
4 Karam G, Sari H. Analysis of predistortion, equalization, and ISI cancellation techniques in digital radio systems with nonlinear transmit amplifiers. IEEE Trans Commun, 1989, 37: 1245–1253
5 Gutierrez A, Ryan W E. Performance of Volterra and MLSD receivers for nonlinear band-limited satellite systems. IEEE Trans Commun, 2000, 48: 1171–1177
6 Malone J, Wickert J. Practical Volterra equalizers for wideband satellite communications with TWTA nonlinearities. In: Proceedings of Digital Signal Processing Workshop and IEEE Signal Processing Education Workshop, Sedona, 2011. 48–53
7 Deleu T, Horlin F, Dervin M. Turbo-equalization of the remaining interference in a pre-distorted non-linear satellite channel. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, Florence, 2014. 1946–1950
8 Qian H, Yao S, Huang H, et al. A low-complexity digital predistortion algorithm for power amplifier linearization. IEEE Trans Broadcast, 2014, 60: 670–678
9 Chen S, Tan S, Xu L, et al. Adaptive minimum error-rate filtering design: a review. Signal Process, 2008, 88: 1671–1697
10 Cai Y, de Lamare R C. Space-time adaptive MMSE multiuser decision feedback detectors with multiple-feedback interference cancellation for CDMA systems. IEEE Trans Veh Tech, 2009, 58: 4129–4140
11 Wu J, Zhong J, Cai Y, et al. New detection algorithms based on the jointly Gaussian approach and successive interference cancelation for iterative MIMO systems. Int J Commun Syst, 2014, 27: 1964–1983
12 Xing C, Ma S, Zhou Y. Matrix-monotonic optimization for MIMO systems. IEEE Trans Signal Process, 2015, 63: 334–348
13 Gong C, Xu Z. Channel estimation and signal detection for optical wireless scattering communication with inter-symbol interference. IEEE Trans Wirel Commun, 2015, 14: 5326–5337
14 Xing C, Gao F, Zhou Y. A framework for transceiver designs for multi-hop communications with covariance shaping constraints. IEEE Trans Signal Process, 2015, 63: 3930–3945
15 Zhou W, Zhang S. The decision delay in finite-length MMSE-DFE systems systems. Wirel Pers Commun, 2015, 83: 1–15
16 Xing C, Ma Y, Zhou Y et al. Transceiver optimization for multi-hop communications with per-antenna power constraints. IEEE Trans Signal Process, 2016, 64: 1519–1534
17 Huang G M, Gillin D, Zhou D, et al. An efficient and robust method to determine the optimal tap coefficients of high speed FIR equalizer. Sci China Inf Sci, 2017, 60: 022401
18 Muller R R, Gerstacker W H. On the capacity loss due to separation of detection and decoding. IEEE Trans Inf Theory, 2004, 50: 1769–1778
19 Douillard C, Jezequel M, Berrou C. Iterative correction of intersymbol interference: turbo equalization. Eur Trans Telecommun, 1995, 6: 507–511

20 Schlegel C B, Perez L C. Trellis and Turbo Coding: Iterative and Graph-Based Error Control Coding. Hoboken: John Wiley & Sons, 2015

21 Laot C, Glavieux A, Labat J. Turbo equalization: adaptive equalization and channel decoding jointly optimized. IEEE J Sel Areas Commun, 2001, 19: 1744–1752

22 Reynolds D, Wang X. Low-complexity turbo-equalization for diversity channels. Signal Process, 2001, 81: 989–995

23 Tuchler M, Koetter R, Singer A C. Turbo equalization: principles and new results. IEEE Trans Commun, 2002, 50: 754–767

24 Zhong W, Lu A A, Gao X Q. MMSE SQRD based SISO detection for coded MIMO-OFDM systems. Sci China Inf Sci, 2014, 57: 042311

25 Liu L, Leung W, Ping L. Simple iterative chip-by-chip multiuser detection for CDMA systems. In: Proceedings of IEEE Vehicular Technology Conference, Jeju Island, 2003. 2157–2161

26 Guo Q, Ping L. LMMSE turbo equalization based on factor graphs. IEEE J Sel Areas Commun, 2008, 26: 311–319

27 Kashif F M, Wymeersch H, Win M Z. Monte carlo equalization for nonlinear dispersive satellite channels. IEEE J Sel Areas Commun, 2008, 26: 245–255

28 Benedetto S, Biglieri E. Nonlinear equalization of digital satellite channels. IEEE J Sel Areas Commun, 1983, 1: 57–62

29 Chen Y C, Su Y T. Turbo equalization of nonlinear TDMA satellite signals. IEICE Trans Commun, 2009, 92: 992–997

30 Burnet C E, Cowley W G. Performance analysis of turbo equalization for nonlinear channels. In: Proceedings of International Symposium on Information Theory, Adelaide, 2005. 2026–2030

31 Ampeliotis D, Rontogiannis A, Berberidis K et al. Turbo equalization of non-linear satellite channels using soft interference cancellation. Adv Sat Mobile Syst, 2008, 124: 289–292

32 Benammar B, Thomas N, Poulliat C, et al. On linear MMSE based turbo-equalization of nonlinear volterra channels. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, Vancouver, 2013. 4703–4707

33 Wang X, Poor H V. Iterative (turbo) soft interference cancellation and decoding for coded CDMA. IEEE Trans Commun, 1999, 47: 1046–1061

34 Xing C, Ma S, Wu Y C. Robust joint design of linear relay precoder and destination equalizer for dual-hop amplify-and-forward MIMO relay systems. IEEE Trans Signal Process, 2010, 58: 2273–2283

35 Tuchler M, Singer A C. Turbo equalization: an overview. IEEE Trans Inf Theory, 2011, 57: 920–952

36 Liu L, Ping L. An extending window MMSE turbo equalization algorithm. IEEE Signal Process Lett, 2004, 11: 891–894