

Received March 15, 2020, accepted April 12, 2020, date of publication April 20, 2020, date of current version May 5, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2989081

A Novel Nonlinear Noise Power Estimation Method Based on Error Vector Correlation Function Using Artificial Neural Networks For Coherent Optical Fiber Transmission Systems

TAO YANG¹, AIYING YANG¹, (Member, IEEE), PENG GUO¹,
YAOJUN QIAO², (Member, IEEE), AND XIANGJUN XIN³

¹Key Laboratory of Photonics Information Technology, Ministry of Industry and Information Technology, School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China

²State Key Laboratory of Information Photonics and Optical Communications, School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

³State Key Laboratory of Information Photonics and Optical Communications, School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

Corresponding author: Aiyang Yang (yangaiying@bit.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61427813, in part by the Open Fund of IPOC (BUPT) under Grant IPOC2018B003, and in part by the State Key Laboratory of Advanced Optical Communication Systems and Networks, China.

ABSTRACT In this paper, we propose a promising nonlinear noise power estimation method based on correlation functions using artificial neural networks (ANN), which is robust against both amplifier spontaneous emission noise and the symbol patterns. Error vector correlation (EVC), together with the amplitude noise correlation (ANC) and the phase noise correlation (PNC), is used as the input of ANN in the proposed method. 378 cases of 224 Gb/s polarization-multiplexed 16-quadrature amplitude modulated (PM-16-QAM) signal transmitted over a wide range of conditions by varying launch power, OSNR and transmission distance are used to train and test the ANN. With the launch power varying from 0 to 8 dBm and the transmission distance as long as 2400 km, the results of tested samples demonstrate that the maximum absolute deviation (MAD), mean absolute error (MAE) and root mean square error (RMSE) of the estimated nonlinear noise power are 0.65 dB, 0.20 dB and 0.27 dB, respectively. In order to verify the independence of symbol patterns, 150 new cases with 30 pseudo-random binary sequence (PRBS) seeds are used to test the trained ANN. The results show that the MAE, MAD and RMSE of estimated nonlinear noise power are 0.86 dB, 0.25 dB and 0.33 dB respectively, meaning that the trained ANN is valid even if the test samples are not covered by the trained process. The results in this work verified that the ANN with EVC, ANC and PNC as input can make the trained model feasible to accurately estimate the nonlinear noise power in high speed optical coherent communication systems.

INDEX TERMS Fiber nonlinearity, error vector correlation function, artificial neural networks.

I. INTRODUCTION

The upsurge of network traffic has caused an increase in capacity of coherent fiber transmission systems. These systems with high bit rates are more vulnerable to linear and nonlinear fiber impairments. Since the linear fiber impairments such as chromatic dispersion (CD) and polarization mode

dispersion (PMD) can be compensated at the receiver by equalization algorithms with negligible penalties [1], [2], the nonlinear distortion caused by fiber Kerr effect has become a critical factor limiting the increase of system capacity [3], [4]. Therefore, many efforts have been made on fiber nonlinearity (NL) research in recent years. On the one hand, NL modeling and NL compensation have been extensive research topics [5]–[9]. On the other hand, there are also some studies on nonlinear noise power estimation. The capability to

The associate editor coordinating the review of this manuscript and approving it for publication was Tianhua Xu.

accurately estimate signal nonlinear distortion has potential application to nonlinear compensation as well as system capacity prediction. However, it is challenging to distinguish between nonlinear distortion and optical amplifier spontaneous emission (ASE) noise when estimating nonlinear noise power [10]. There are mainly two ways to separate NL from ASE noise. One way is utilizing the distribution difference of ASE and NL with the special designed pilot [11]–[13]. Authors in [11] proposed an inter-channel cross phase modulation (XPM) power estimation method based on polarization diversified pilot tones. After the angular noise squeezing, the pilot carried amplitude noise featured a notched spectrum. Then, the authors used the notched spectrum that is characterized to separate ASE noise and XPM noise to estimate the XPM noise power. In [12], linear-frequency modulated (LFM) signal acted as time domain pilot in front of the payload to estimate inter-channel nonlinear noise power. The LFM pilot is transformed into a fractional domain by an optimal fractional Fourier transform. After removing the signal peaks in the fractional domain, the sum of ASE noise and nonlinear noise can be calculated. And the power of inter-channel nonlinear noise power can be obtained by subtracting the ASE power measured in the frequency domain. Moreover, a method in [13] employed low-frequency pilot tone modulation and zero-power gaps in the amplitude of the transmitted signal. The signal power modulated with a pilot tone leads to a modulation on the nonlinear noise power. Since the ASE noise is not modulated, the pilot tone amplitude in the zero-power gaps can be used to estimate nonlinear noise power. These pilot-aided methods will reduce the system spectral efficiency when the pilots are typically time-multiplexed with information symbols. Moreover, the methods in [12] need modification of the transmitter.

The other way to calculate the nonlinear noise power is utilizing the statistical difference between ASE noise and nonlinear noise in time domain [14]–[16]. Authors in [14] used amplitude noise correlation (ANC) among neighboring symbols to characterize nonlinear noise power. However, a distance-dependent calibration factor is needed as a quantitative estimate of the nonlinear noise power. In [15], the estimation accuracy of nonlinear noise power was improved by considering the amplitude noise correlation functions between two polarizations. Authors in [16] used a new parameter depending on both transmission distances and launch powers to more accurately estimate nonlinear noise power. The methods mentioned in [14]–[16] all required a complex calibration process and need to know the link information. Therefore, it is not easy to apply these methods to practical systems.

In recent years, machine learning is a hot research area and has been applied to optical performance monitoring [17]–[19]. In [20], the method to monitor nonlinear signal-to-noise ratio (SNR_{nl}) used artificial neural networks (ANN) where the ANC proposed in [14] was used for training. After training with various systems, the ANN

model for SNR_{nl} monitoring achieved high accuracy. In [21], authors proposed that the tangential and normal statistical components of the constellation can be used to characterize NL for 16QAM signals. These statistics, together with ANC, are used in ANN training to estimate SNR_{nl}. In [22], another statistical characteristic phase noise correlation (PNC) was added to train the ANN model in [20], which made the method more accurate. However, it is found that the estimation accuracy depends on the symbol patterns. If the test symbol patterns are not included in the trained stage, the estimation error will be greater. For methods in [20]–[22], the characteristics used in these methods reflect amplitude or phase information separately. It is known that nonlinear distortion is manifested both in amplitude and phase of the signal, so the error vector containing both amplitude and phase information can be utilized. In this paper, we use a new feature, the error vector correlation (EVC), to characterize NL, which is robust to both ASE noise and symbol patterns. Then, we proposed a nonlinear noise power estimation method based on ANN where the EVC, ANC and PNC are used as the input of estimator.

The remaining of the paper is organized as follows. In section 2, the properties of the EVC are presented and the ANN-based nonlinear noise power estimation method considered in this paper is described in detail. In section 3, the effectiveness of our method is demonstrated in 224 Gb/s PM-16-quadrature amplitude modulation (PM-16-QAM) simulation system, which is followed by the conclusion in section 4.

II. THEORETICAL PRINCIPLE

A. EVC: A NEW FEATURE TO CHARACTERIZE FIBER NONLINEARITY

In coherent optical fiber transmission systems, signals are subject to various link impairments. These impairments are caused by the factors including ASE noise, Kerr effect, CD, PMD, laser phase noise, etc. Typically, received symbols need to be processed by a standard signal processing flow i.e. normalization, resampling, CD compensation, PMD compensation, frequency offset estimation and phase recovery. As a result, the residual impairments on symbols processed by the above operations mainly consist of ASE noise and distortion caused by Kerr nonlinearity. In this case, the received k^{th} symbol \hat{S}_k can be represented as:

$$\hat{S}_k = S_k + N_k + NL_k \quad (1)$$

where S_k is the k^{th} reference symbol at the transmitting end, N_k and NL_k are the ASE noise distortion and Kerr nonlinear distortion to the k^{th} symbol, respectively. Obviously, \hat{S}_k , S_k , N_k and NL_k are all plural in (1). Then, we obtained the error vector by:

$$\Delta_k = \hat{S}_k - S_k = N_k + NL_k \quad (2)$$

We defined the error vector correlation (EVC) function of m delay symbols as:

$$\begin{aligned} EVC(m) &= |E(\Delta_k \Delta_{k+m}^*)| \\ &= |E[(N_k + NL_k)(N_{k+m} + NL_{k+m})^*]| \\ &= |E(N_k N_{k+m}^* + E(N_k NL_{k+m}^*) + E(NL_k N_{k+m}^*) \\ &\quad + E(NL_k NL_{k+m}^*)| \end{aligned} \quad (3)$$

where $E(\cdot)$ denotes expectation.

It is known that ASE noise is a band-limited complex circularly symmetric zero-mean Gaussian random process, and the correlation of ASE noise should be expressed as:

$$E(N_k N_{k+m}^*) = P_N * \delta(m) \quad (4)$$

where P_N is the power spectral density of ASE noise, $\delta(m)$ is the Delta function. In addition, it can be considered that the correlation between nonlinear distortion and ASE noise is negligible. Therefore, when m is not equal to zero, we can simplify the EVC function into:

$$EVC(m) = |E(NL_k \cdot NL_{k+m}^*)| \quad (5)$$

Obviously, it is noted by Eq. (5) that the EVC function characterizes the correlation between nonlinear distortions of adjacent symbols. Moreover, different m values can reflect the correlation between different delay symbols. Figure 1(a) shows the $EVC(m)$ calculated for the PM-16QAM signals after 1600 km transmission. Firstly, we can notice from Figure 1(a) that, when fixing the value of m , the value of EVC is proportional to the launch power of the signal, which is consistent with the theoretical prediction of the fiber Kerr effect. Further, Figure 1(b) shows that, for the same launch power, the value of EVC does not fluctuate substantially with the change of OSNR, which embodies that EVC is robust to ASE noise. In addition, when m increases, the EVC value under the same transmission condition gradually decreases. The strongest correlation appears when the relative time index $m = 1$ ($m = 0$ is for an autocorrelation and not included), so $EVC(1)$ is used in our proposed method to characterize fiber nonlinearity. For the simplicity of expression, EVC refers to $EVC(1)$ in the following text. We can conclude from Figure 1 that EVC is insensitive to the ASE noise with a wide OSNR range from 18-36 dB and proportional to the launch power. Therefore, EVC is capable of characterizing NL as (5) indicated.

In previous work [20]–[22], two other features, amplitude noise correlation (ANC), $ANC(m) = |E(\Delta_{A(k)} \Delta_{A(k+m)})|$ and phase noise correlation (PNC), $PNC(m) = |E(\Delta_{P(k)} \Delta_{P(k+m)})|$ are defined and used to estimate nonlinear noise power. Where $\Delta_{A(k)}$ and $\Delta_{P(k)}$ are the amplitude noise and phase noise of the k^{th} symbol, respectively. However, it is noted from the definition of ANC that:

$$\begin{aligned} \Delta_{A(k)} &= |\hat{S}_k| - |S_k| \\ &= |S_k + N_k + NL_k| - |S_k| \\ &\neq \text{Sgn}(|\hat{S}_k| - |S_k|) \cdot |N_k + NL_k| \end{aligned} \quad (6)$$

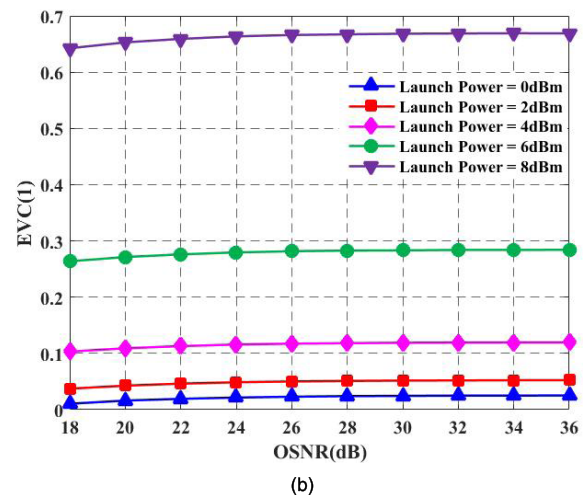
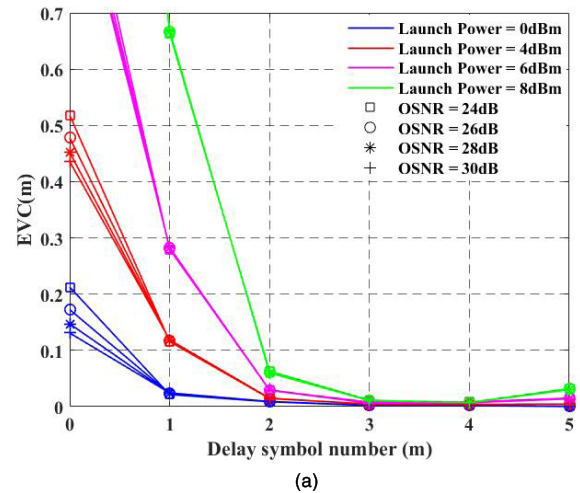


FIGURE 1. Error vector correlation for 224 Gb/s PM-16-QAM signals after 1600 km transmission with various signal launched powers and OSNR values. (a): Value of EVC with different delay symbols m . (b): $EVC(1)$ over a wide OSNR range and launch power range.

where $\text{Sgn}(\cdot)$ is the sign function. From (6), it is not hard to understand that $\Delta_{A(k)}$ is not equal to the noise amplitude, and depends on the pattern of transmission symbols, which causes the value of ANC to be related to the transmitted symbol pattern. Similarly, PNC also depends on the pattern of transmission symbols. Figure 2 compares the dependence of symbol patterns for ANC, PNC and EVC with 30 pseudo-random binary sequence (PRBS) seeds of PM-16QAM signal under the same conditions. In figures 2(a), (b) and (c), it is clear that ANC and PNC fluctuate significantly with PRBS seeds, while EVC hardly fluctuates, which is consistent with the above analysis. Further, we define the normalized root mean square error (NRMSE) as the root characterizing the fluctuation to indicate such pattern dependence. NRMSE is calculated by:

$$NRMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y})^2}{n\hat{y}^2}} \quad (7)$$

where n is the number of samples, y_i is the i_{th} sample and \hat{y} is the mean value of n samples. As shown in Figure 2(d), the NRMSE of EVC is lower than that of ANC and PNC at each launch power. This indicates that EVC is more robust to the pattern dependence than the other two features, which is preferred for the nonlinear noise power estimation methods based on machine learning. Therefore, we calculate EVC and input it to the ANN to make the trained network independent of the symbol patterns.

B. ANN-BASED NONLINEAR NOISE POWER ESTIMATION

As one of the categories of supervised learning, given a set of features and a target, ANN can learn a non-linear function approximator for classification or regression. Figure 3 shows a one hidden layer ANN with scalar output. Obviously, the layers in the network are fully connected (each neuron in each layer is connected to all neurons in the previous layer). The top layer, known as the input layer, consists of a set of neurons representing the input features $X = [x_1, x_2, x_3, \dots, x_n]$. The middle and bottom layers are the hidden layer and the output layer, respectively. The value of i^{th} neuron h_i in the hidden layer is obtained by linearly weighted summing the values of the previous layer and using a non-linear function:

$$h_i = g(W_i^{(1)}X^T + b_i^{(1)}) \tag{8}$$

where $g(\cdot)$ is the non-linear activation function, $W_i^{(1)} = [w_{1,i}^{(1)}, w_{2,i}^{(1)}, w_{3,i}^{(1)}, \dots, w_{n,i}^{(1)}]$ and $b_i^{(1)}$ are the weight vector and the bias for h_i , respectively. Then, the output layer receives the values from the hidden layer and transforms them into output value F , which can be expressed as:

$$F = W^{(2)}H^T + b^{(2)} \tag{9}$$

where $H = [h_1, h_2, \dots, h_m]$ is the vector of hidden layer neurons, $W^{(2)} = [w_1^{(2)}, w_2^{(2)}, w_3^{(2)}, \dots, w_m^{(2)}]$ and $b^{(2)}$ are the weight vector and the bias for the output layer. The output F of the network has a difference compared with the real reference value, and the mathematical relationship between them constitutes the loss function of the network. During the training phase using the back-propagation algorithm, all weight vectors and biases will be optimized in an iterative process until the loss function changes less than a set threshold.

In our work, we use an ANN with only one hidden layer containing 10 neurons. The activation functions of the hidden layer and the output layer are the rectified linear unit function and the identity function, respectively. The ANN uses the square error as the loss function and the Limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm as the solver for weight optimization.

Based on the analysis of ANC, PNC and EVC, three features ANC, PNC and EVC, are used as the ANN neural network input, without priori information of fiber link parameters, to train the network and estimate nonlinear noise power. Figure 4. shows the flowchart of our method. Firstly, collect

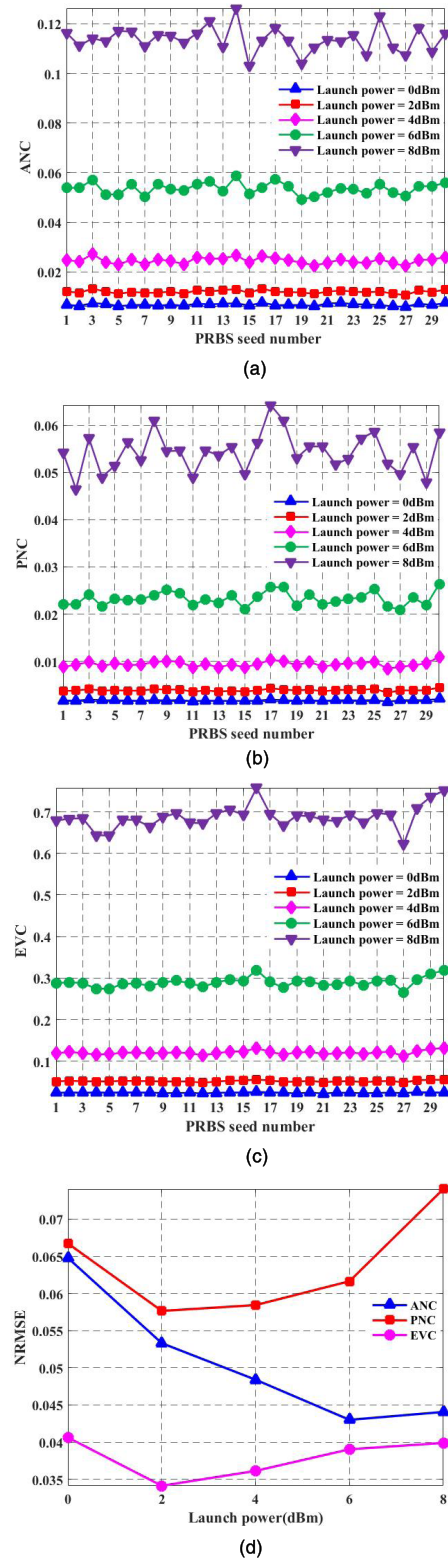


FIGURE 2. Comparison of extracted features, ANC, PNC and EVC, from the PM-16QAM signals after 1600 km transmission with various signal launched powers and PRBS seeds (OSNR = 26 dB). (a) ANC. (b) PNC. (c) EVC. (d) NRMSE of ANC, PNC and EVC.

the signal data sets for various cases (such as launch powers, transmission distances, OSNR values and PRBS seeds).

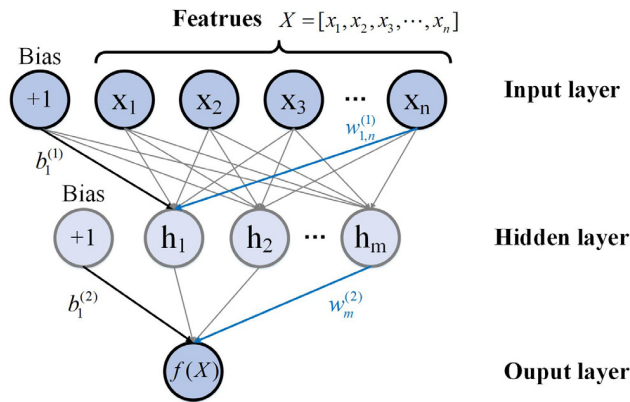


FIGURE 3. Schematic diagram of single hidden layer ANN.

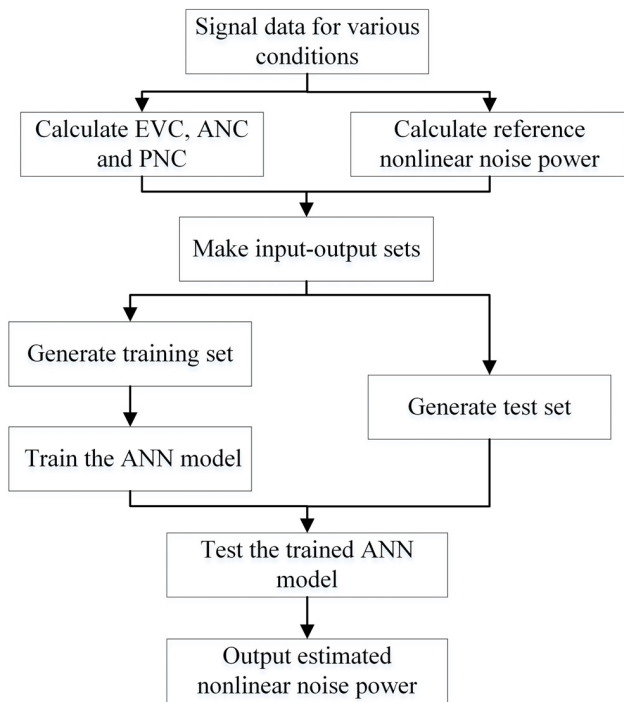


FIGURE 4. The flowchart of ANN based nonlinear noise power estimation method.

Then, calculate the features (EVC, ANC, PNC) and obtain the reference nonlinear noise power to make input-output sets. These input-output sets are randomly divided into training set and test set in the next step. The training set is used to train the ANN to achieve an optimal performance. Finally, the trained ANN is used to estimate the nonlinear noise power as output.

III. SIMULATION SETUP AND RESULTS ANALYSIS

In order to verify the feasibility of the proposed method, we use the commercial software Virtual Photonics Inc. (VPI) to simulate the 28 GBaud PM-16-QAM systems as shown in Figure 5. The black arrow and the red arrow represent the electrical signals and the optical signals respectively. At the transmitter side, the optical signals are mapped into 16-QAM format and modulate two orthogonal linearly polarized laser

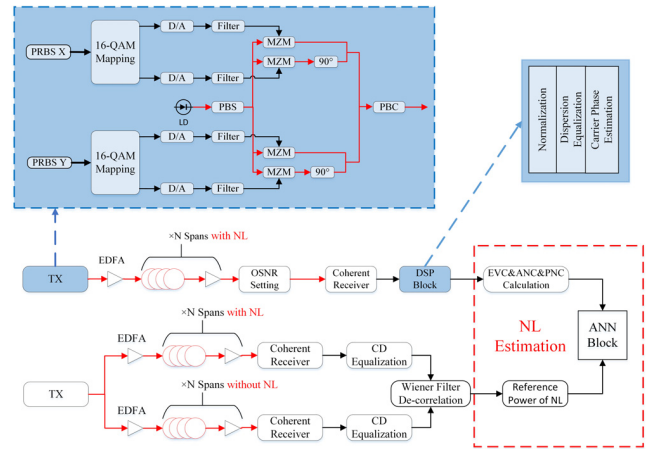


FIGURE 5. Schematic diagram of nonlinear noise power estimation method for 224 Gb/s PM-16-QAM optical fiber transmission systems.

by Mach-Zehnder modulator (MZM). For pulse shaping, a Bessel electrical low-pass filters with filter order 4 are used, the bandwidth of which is 21 GHz. Then, a polarization beam coupler (PBC) is used to combine the two beam optical signals for transmission. The modulated optical signals are amplified using an erbium-doped fiber amplifier (EDFA) and launch into the optical fiber link. Each span consists of an 80 km standard single-mode fiber (SSMF) and an EDFA used to compensate the energy loss caused by the fiber attenuation. The linewidth of the laser is set to 100 kHz. The CD and the nonlinear coefficient are set to $17e^{-6}s/m^2$ and $1.3(W \cdot km)^{-1}$ respectively. There is no ASE noise added in each EDFA and the OSNR is controlled by 'OSNR Setting' module at the end of the fiber link.

To obtain signal data sets for various situations, we sweep the parameters of interest. At the transmitter end, we use three PRBS seeds and make the launch power vary from 0-8 dBm in 2 dB steps. The transmission length of the link has three values: 1600, 2000, and 2400 km. The OSNR varies in the range of 18-36 dB with a step of 2 dB. In digital signal processing (DSP) block, we used overlapped frequency-domain equalizer (OFDE) algorithm for CD compensation and QPSK partition algorithm for carrier phase estimation, respectively. Assume the transmitted sequence is known, 30000 symbols are used to calculate the EVC, ANC and PNC for each simulation case. Besides, the received signals with and without NL are de-correlated by Wiener filter to calculate the reference value of nonlinear noise power [23], [24].

In order to describe the strength of nonlinear noise power, we use ASE noise to nonlinear noise ratio (ANR) defined in Ref [21] as:

$$ANR = 10 \log_{10}(P_{ASE}/P_{NL}) \tag{10}$$

where P_{ASE} and P_{NL} are the power of ASE noise and nonlinear noise, respectively. We select the cases where the ANR is less than 10 dB to make the data set, i.e., the nonlinear noise power is greater than 1/10 of the ASE noise power. For cases below this limit, nonlinear noise is negligible. In

TABLE 1. The estimation error of different ANN models for the 114 test cases.

ANN model	Input	MAD(dB)	MAE(dB)	RMSE(dB)
1	ANC	2.89	0.67	1.00
2	PNC	1.35	0.45	0.57
3	EVC	1.19	0.41	0.50
4	ANC&&PNC	0.91	0.22	0.33
5	ANC&&EVC	1.16	0.38	0.45
6	PNC&&EVC	1.09	0.23	0.31
7	ANC&&PNC&&EVC	0.65	0.20	0.27

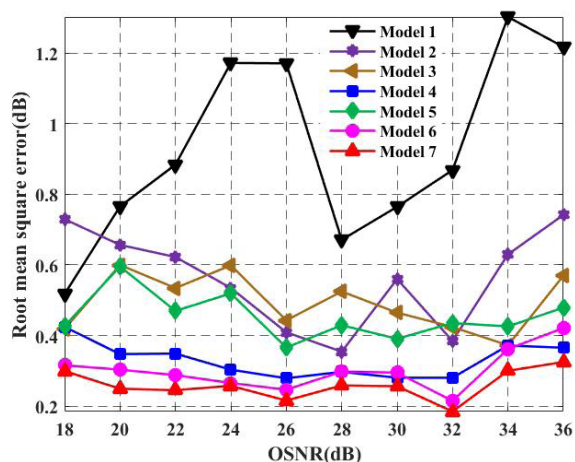


FIGURE 6. The estimation RMSE of different ANN models at different OSNR values.

this way, we obtained 378 input-output data sets composed of features and reference values of nonlinear noise power. The random 264 input-output sets are used for training. Then, the remaining 114 cases are used for testing and their estimation errors are calculated. In order to analyze the contribution of three individual features to the performance of our proposed method, we also take each single feature as the input of ANN in both training and testing. Seven models (numbered 1-7 respectively) with different inputs are tried. To describe the estimation accuracy of nonlinear noise power, the maximum absolute deviation (MAD), mean absolute error (MAE) and root mean square error (RMSE) are calculated. Table 1 compares the estimation error of nonlinear noise power of respective models.

When a single feature is used as the input of the ANN, the model using EVC as input has the least MAD, MAE, and RMSE. If 2 features are input to the ANN, all the estimation errors are less than the ANN models with one single feature input. The ANN model with three features as input has the highest accuracy, and the MAD, MAE and RMSE are 0.65 dB, 0.20 dB and 0.27 dB, respectively. Figure 6 shows the RMSE of estimated nonlinear noise power of models 1-7 at different OSNR values. It can be found that Model 7 using the three features as ANN input has the least RMSE at

TABLE 2. The estimation error of different ANN models for verifying the independence of symbol patterns.

ANN model	Input	MAD(dB)	MAE(dB)	RMSE(dB)
1	ANC	3.11	0.98	1.21
2	PNC	1.69	0.45	0.57
3	EVC	1.07	0.31	0.37
4	ANC&&PNC	1.60	0.37	0.47
5	ANC&&EVC	1.54	0.39	0.56
6	PNC&&EVC	1.03	0.31	0.39
7	ANC&&PNC&&EVC	0.86	0.25	0.33

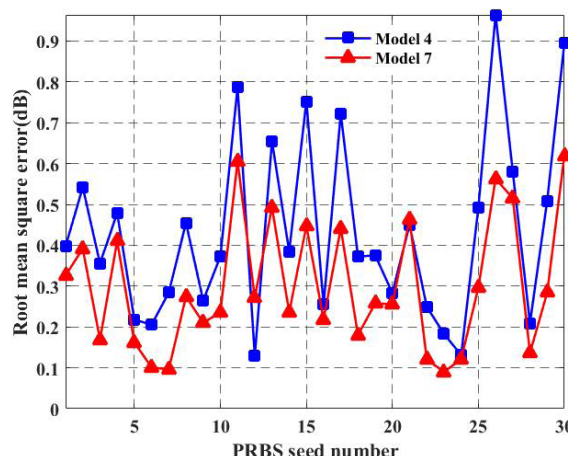


FIGURE 7. The estimation RMSE of different ANN models for different PRBS seeds.

all OSNR values compared to others, and the RMSE of Model 7 only fluctuates slightly within a range of 0.19 - 0.33 dB when the OSNR is from 18 to 36 dB. This indicates that our method is robust to ASE noise.

Further, in order to verify that our method is independent of symbol patterns, we use another 30 PRBS seeds to simulate 150 new cases and calculate three features from received signals. In these cases, the transmission distance and OSNR are fixed at 1600 km and 26 dB, respectively. The launch power varies in the range of 0-8 dBm with a step of 2 dB. Then, all 150 new cases are used to test the trained models. The estimation errors of nonlinear noise power by seven models are shown in Table 2. Model 3 with EVC as input has least error if only one feature inputs to ANN. Model 6 performs best among models using two features. Model 7 has best performance for estimating nonlinear noise power, in which MAE, MAD and RMSE of 150 cases are 0.86 dB, 0.25 dB and 0.33dB, respectively. It is because that both Model 3, Model 6 and Model 7 use EVC as one input of the ANN. The results demonstrate that it is EVC that makes the nonlinear noise power estimation independent of symbol patterns. The maximum estimation error is 1.07dB and 0.86dB for ANN model 3 and 7 respectively. The difference is about 0.2dB between the two models, while the latter has greater

computation complexity. Therefore, ANN model 3 with EVC as ANN input can be preferred if the computing speed is limited.

To analyze the contribution of EVC to the model's independence of symbol pattern, we compare the RMSE of estimated nonlinear noise power with model 4 and model 7 in Figure 7. Model 7 differs from Model 4 in that EVC is added as one input. For the cases of 30 PRBS seeds, Model 7, with RMSE ranging from 0.10 to 0.62 dB, is more insensitive to symbol patterns than Model 4 with RMSE ranging from 0.13 to 0.96 dB. Therefore, it can be concluded that models using EVC are more independent of symbol patterns.

IV. CONCLUSION

In this paper, we firstly propose a new feature EVC to characterize nonlinear noise power in optical fiber communication systems, which is robust to both ASE noise and symbol patterns. Then, the nonlinear noise power estimation method based on correlation functions and ANN is proposed, in which EVC, ANC and PNC are used as the input of ANN. The method is demonstrated to estimate nonlinear noise power accurately in 224 Gb/s PM-16-QAM system. With the launch power varying from 0 to 8 dBm and the transmission distance as long as 2400 km, the results demonstrate that the MAE, MAD and RMSE of estimated nonlinear noise power are 0.65dB, 0.20 dB and 0.27 dB, respectively. Moreover, to verify that our method is independent of symbol patterns, 150 new cases containing 30 different PRBS seeds are used to test the trained model. The result show that the MAE, MAD and RMSE for these cases are 0.86 dB, 0.25 dB and 0.33 dB, respectively. Finally, we analyze that it is EVC as ANN input that makes the nonlinear noise power estimation independent of symbol patterns. ANN model with EVC as ANN input can be preferred if the computing speed is limited. It can be concluded that our method based on EVC and ANN is a promising candidate for accurately estimating nonlinear noise power in long haul optical fiber communication systems.

REFERENCES

- [1] C. R. S. Fludger, T. Duthel, D. van den Borne, C. Schullien, E.-D. Schmidt, T. Wuth, J. Geyer, E. De Man, K. Giok-Djan, and H. de Waardt, "Coherent equalization and POLMUX-RZ-DQPSK for robust 100-GE transmission," *J. Lightw. Technol.*, vol. 26, no. 1, pp. 64–72, Jan. 1, 2008.
- [2] M. Kuschnerov, F. N. Hauske, K. Piyawanno, B. Spinnler, M. S. Alfiad, A. Napoli, and B. Lankl, "DSP for coherent single-carrier receivers," *J. Lightw. Technol.*, vol. 27, no. 16, pp. 3614–3622, Aug. 15, 2009.
- [3] R. Dar, M. Feder, A. Mecozzi, and M. Shtaif, "Time varying ISI model for nonlinear interference noise," in *Proc. Opt. Fiber Commun. Conf. Exhibit.*, 2014, pp. 1–3.
- [4] M. C. Tan, F. N. Khan, W. H. Al-Arashi, Y. Zhou, and A. P. T. Lau, "Simultaneous optical performance monitoring and modulation format/bit-rate identification using principal component analysis," *J. Opt. Commun. Netw.*, vol. 6, no. 5, pp. 441–448, May 2014.
- [5] A. Napoli, Z. Maalej, V. A. J. M. Steiffer, M. Kuschnerov, D. Rafique, E. Timmers, B. Spinnler, T. Rahman, L. D. Coelho, and N. Hanik, "Reduced complexity digital back-propagation methods for optical communication systems," *J. Lightw. Technol.*, vol. 32, no. 7, pp. 1351–1362, Apr. 1, 2014.
- [6] F. P. Guomar and A. N. Pinto, "Simplified volterra series nonlinear equalizer for polarization-multiplexed coherent optical systems," *J. Lightw. Technol.*, vol. 31, no. 23, pp. 3879–3891, Dec. 1, 2013.
- [7] M. Sorokina, S. Sygletos, and S. Turitsyn, "Sparse identification for nonlinear optical communication systems: SINO method," *Opt. Express*, vol. 24, no. 26, pp. 30433–30443, Dec. 2016.
- [8] A. D. Ellis, M. E. McCarthy, M. A. Z. Al-Khateeb, and S. Sygletos, "Capacity limits of systems employing multiple optical phase conjugators," *Opt. Express*, vol. 23, no. 16, pp. 20381–20393, Aug. 2015.
- [9] O. S. Sidelnikov, A. A. Redyuk, and S. Sygletos, "Dynamic neural network-based methods for compensation of nonlinear effects in multimode communication lines," *Quantum Electron.*, vol. 47, no. 12, pp. 1147–1149, Dec. 2017.
- [10] L. Dou, T. Yamauchi, X. Su, Z. Tao, S. Oda, Y. Aoki, T. Hoshida, and J. C. Rasmussen, "An accurate nonlinear noise insensitive OSNR monitor," in *Proc. Opt. Fiber Commun. Conf. Exhibit.*, 2016, pp. 1–3, Paper W3A.5.
- [11] Y. Zhao, Z. Tao, S. Oda, Y. Aoki, and T. Hoshida, "Pilot based cross phase modulation power estimation," in *Proc. Opt. Fiber Commun. Conf. Exhibit.*, 2017, pp. 1–3, Paper W1G.2.
- [12] W. Wang, A. Yang, P. Guo, Y. Lu, and Y. Qiao, "Joint OSNR and interchannel nonlinearity estimation method based on fractional Fourier transform," *J. Lightw. Technol.*, vol. 35, no. 20, pp. 4497–4506, Oct. 15, 2017.
- [13] Z. Jiang and X. Tang, "Nonlinear noise monitoring in coherent systems using amplitude modulation pilot tone and zero-power gap," in *Proc. Opt. Fiber Commun. Conf. Exhibit.*, vol. 2019, pp. 1–3, Paper Th2A.34.
- [14] Z. Dong, A. P. T. Lau, and C. Lu, "OSNR monitoring for QPSK and 16-QAM systems in presence of fiber nonlinearities for digital coherent receivers," *Opt. Express*, vol. 20, no. 17, pp. 19520–19534, Aug. 2012.
- [15] H. G. Choi, J. H. Chang, H. Kim, and Y. C. Chung, "Nonlinearity-tolerant OSNR estimation technique for coherent optical systems," in *Proc. Opt. Fiber Commun. Conf. Exhibit.*, 2015, pp. 1–3, Paper W4D.2.
- [16] Z. Wang, A. Yang, P. Guo, Y. Lu, and Y. Qiao, "Nonlinearity-tolerant OSNR estimation method based on correlation function and statistical moments," *Opt. Fiber Technol.*, vol. 39, pp. 5–11, Dec. 2017.
- [17] D. Wang, M. Zhang, J. Li, Z. Li, J. Li, C. Song, and X. Chen, "Intelligent constellation diagram analyzer using convolutional neural network-based deep learning," *Opt. Express*, vol. 25, no. 15, pp. 17150–17166, Jul. 2017.
- [18] F. N. Khan, K. Zhong, X. Zhou, W. H. Al-Arashi, C. Yu, C. Lu, and A. P. T. Lau, "Joint OSNR monitoring and modulation format identification in digital coherent receivers using deep neural networks," *Opt. Express*, vol. 25, no. 15, pp. 17767–17776, Jul. 2017.
- [19] Z. Wang, A. Yang, P. Guo, and P. He, "OSNR and nonlinear noise power estimation for optical fiber communication systems using LSTM based deep learning technique," *Opt. Express*, vol. 26, no. 16, pp. 21346–21357, Aug. 2018.
- [20] A. S. Kashi, Q. Zhuge, J. C. Cartledge, A. Borowiec, D. Charlton, C. Laperle, and M. O'Sullivan, "Fiber nonlinear noise-to-signal ratio monitoring using artificial neural networks," in *Proc. Eur. Conf. Opt. Commun. (ECOC)*, Sep. 2017, pp. 1–3, Paper M2F.2.
- [21] F. J. V. Caballero, D. Ives, Q. Zhuge, M. O'Sullivan, and S. J. Savory, "Joint estimation of linear and non-linear signal-to-noise ratio based on neural networks," in *Proc. Opt. Fiber Commun. Conf. Exhibit.*, 2018, pp. 1–3, Paper M2F.4.
- [22] A. S. Kashi, Q. Zhuge, J. C. Cartledge, S. A. Etamad, A. Borowiec, D. W. Charlton, C. Laperle, and M. O'Sullivan, "Nonlinear signal-to-noise ratio estimation in coherent optical fiber transmission systems using artificial neural networks," *J. Lightw. Technol.*, vol. 36, no. 23, pp. 5424–5431, Dec. 1, 2018.
- [23] W. Wang, A. Yang, P. Guo, Y. Lu, and Y. Qiao, "Monitoring interchannel nonlinearity based on differential pilot," *Opt. Commun.*, vol. 417, pp. 24–29, Jun. 2018.
- [24] J. Wu, J. Droppo, L. Deng, and A. Acero, "A noise-robust ASR front-end using Wiener filter constructed from MMSE estimation of clean speech and noise," in *Proc. IEEE Workshop Automat. Speech Recognit. Understand.*, Nov./Dec. 2013, pp. 321–326.



TAO YANG received the B.E. degree from the Beijing Institute of Technology, China, in 2017, where he is currently pursuing the master's degree. He is a member of the Key Laboratory of Photonics Information Technology, China. His main research interests include optical fiber communication systems and nonlinear noise power monitoring.



AIYING YANG (Member, IEEE) received the B.S. degree in physics from Jilin University, China, in 1997, and the Ph.D. degree in information and communication systems from Peking University, China, in 2003. She is currently a Professor with the School of Optics and Photonics, Beijing Institute of Technology, China. She is also a member of OSA. Her current researches mainly focus on optical fiber communications and visible light communications.



PENG GUO received the B.S. and Ph.D. degrees from Peking University, in 2007 and 2013, respectively. He is currently an Associate Professor with the School of Optics and Photonics, Beijing Institute of Technology, China. His current researches mainly focus on optical fiber communications and visible light communications.



YAOJUN QIAO (Member, IEEE) received the B.S. degree from Hebei Normal University, Shijiazhuang, China, in 1994, the M.S. degree from Jilin University, Jilin, China, in 1997, and the Ph.D. degree from the Beijing University of Posts and Telecommunications, Beijing, China, in 2000. He was with Lucent and Fujitsu, from 2000 to 2007. In 2007, he joined the Beijing University of Posts and Telecommunications, where he is currently a Professor with the School of Information and Communication Engineering. His research interests include optical and wireless communications.



XIANGJUN XIN received the Ph.D. degree in electromagnetic field and microwave technology from the Beijing University of Posts and Telecommunication (BUPT), Beijing, China, in 2004. He is currently a Professor with the School of Electronic Engineering and a member of the State Key Laboratory of Information Photonics and Optical Communications, BUPT. He has more than 100 publications in prestigious journals and conferences. His main research interests include high-speed fiber communication systems, broadband optical transmission technologies, and all-optical networks.

...