Optical Phase Conjugation With Complex-Valued Deep Neural Network for WDM 64-QAM Coherent Optical Systems

Lei Wang, Mingyi Gao, *Senior Member, IEEE*, Yongliang Zhang, Fengchu Cao, and He Huang

***Abstract*—We experimentally demonstrated a photoelectric nonlinear compensation scheme of optical phase conjugation (OPC) with complex-valued deep neural network (CVDNN) to mitigate fiber nonlinearity in wavelength division multiplexing (WDM) 64-QAM coherent optical transmission system. The factors to affect the performance of OPC and CVDNN are comprehensively considered. OPC in WDM system is experimentally optimized to alleviate the deployment requirements of strict symmetrical distributed power and chromatic dispersion. The performance penalty caused by the simplification of the OPC is further compensated by the CVDNN. The selections of the input neurons' number and the optimization algorithm are also considered to design a simple two-hidden-layer-structure CVDNN. The proposed method is experimentally verified and evaluated in a 12.5-GBd 4-channel WDM 64-QAM 160-km standard single-mode fiber (SSMF) transmission system with channel spacing of 50-GHz. The experimental results show that the proposed nonlinear equalizer based on the OPC with CDVNN has a strong robustness to the input signal power and wavelength, which can not only improve the Q factor of the signal by 1.5-dB, but also greatly increase the total launched signal power.**

***Index Terms*—Fiber nonlinearity, optical phase conjugation, complex-valued neural networks, machine learning, nonlinear equalization.**

I. Introduction

THE capacity crunch has been exacerbated by the demand for real-time data with high bandwidth and high connectivity [1]. In response to the explosive growth of traffic demand from the next generation of mobile Internet and Internet of Things, optical fiber communication system utilizes various multiplexing methods such as wavelength division multiplexing (WDM), polarization division multiplexing (PDM) and space-division multiplexing (SDM), leverages advanced modulation formats of high-order quadrature amplitude modulation (QAM) and extends the communication bandwidth to other bands and so on [2]–[5]. Nonetheless, the nonlinear Shannon limit still induces the capacity crunch. In long distance and high bandwidth backbone optical networks, the nonlinearity mainly comes from the fiber nonlinearity caused by Kerr effects in the form of self-phase modulation (SPM), cross-phase modulation (XPM), four-wave mixing (FWM), the crosstalk between them and amplified spontaneous emission (ASE) noise of Erbium-doped fiber amplifiers (EDFAs) employed in the system, which degrades the signal in the fiber transmission [6], [7].

Various digital nonlinear compensation (NLC) methods based on the digital signal processing (DSP) of digital coherent receivers have been proposed to surpass the nonlinear Shannon limit such as digital back-propagation (DBP) [8], Volterra series based nonlinear equalizer (VNLE) [9] and nonlinear symbol decision based on machine learning (ML), i.e., K-means [10], support vector machine (SVM) [11] and neural networks (NNs) [12],[13]. DBP is an efficient approach to compensate the deterministic nonlinear effects by applying split-step Fourier to solve the nonlinear Schrodinger equations of fiber inverse propagation. But its computational complexity increases with the increasing of the transmission distance and the accumulation of chromatic dispersion, which is not matched with real-time WDM systems. VNLE with lower complexity has similar issues [14], [15]. The nonlinear symbol decision based on ML can further optimize the decision margin of signals in their constellation diagrams, however, for severely degraded signals, the performance improvement is quite trivial.

Moreover, some digital NLC schemes based on artificial neural networks (ANNs), recurrent neural networks (RNNs) and convolutional neural networks (CNNs) have also been successfully proposed and applied to intensity modulation direct detection (IMDD) or coherent optical systems to equalize the distorted signal affected from transmission impairments [16]–[19]. However, for the long-haul coherent optical system with high-order QAM, these NNs need to split the transmitted complex-valued signals into the real and imaginary components for processing, which weakens the phase relationship of in-phase/quadrature (IQ) in QAM signals to some extent. The complex-valued deep neural network (CVDNN) can directly process the complex-valued transmitted signals, which ensures the signal's orthogonality in the black box of the neural network.

Manuscript received September 1, 2021; accepted September 8, 2021. Date of publication September 13, 2021; date of current version October 15, 2021. This work was supported in part by the National Key R&D Program of China under Grant 2020YFB1805805, in part by the National Natural Science Foundation of China Project under Grant 62075147, and in part by the Natural Science Foundation of Jiangsu Province of China under Grant BK20181431. *(Corresponding author: Mingyi Gao.)*

The authors are with the Jiangsu Engineering Research Center of Novel Optical Fiber Technology and Communication Network, Suzhou Key Laboratory of Advanced Optical Communication Network Technology, the School of Electronic and Information Engineering, Soochow University, Suzhou 215006, Jiangsu Province, China (e-mail: 20184228025@stu.suda.edu.cn; mygao@suda.edu.cn; 20184228024@stu.suda.edu.cn; 20204228022@stu.suda.edu.cn; hhuang@suda.edu.cn).

Digital Object Identifier 10.1109/JPHOT.2021.3111921

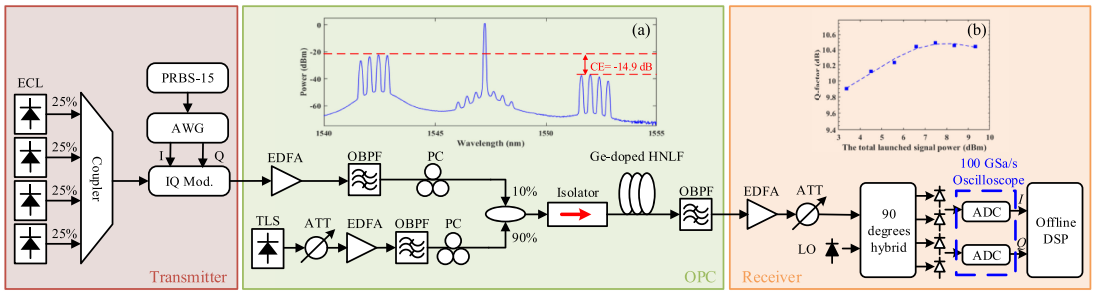


Fig. 1. Experimental setup of a 12.5-GBd WDM 64-QAM back-to-back optical phase conjugation system. Inset: (a) The output spectra (0.1-nm resolution) of Ge-doped HNLF. (b) Measured Q factor of the idler at the channel of interest versus the total input WDM signals power.

Thus, the CVDNN outperforms the conventional NN in WDM 16-QAM coherent optical communication systems [19], [20]. Nevertheless, the specific nonlinear function through six hidden layers increases the complexity of the CVDNN algorithm [19].

In spite of the improved bit error ratio (BER) performance achieved by digital NLC schemes, the launched power into fiber can't be increased significantly. In such a case, optical phase conjugation (OPC) as an NLC scheme operated in optical domain, outperforms other digital NLC schemes. Compared with digital NLC schemes, OPC can be rapidly operated over large optical bandwidths to compensate for the deterministic inter and intra nonlinearity [21]–[23]. Thus, OPC can not only increase the launched power into fiber, but also alleviate the wavelength competition in the optical network by the attached wavelength conversion. However, the high-performance OPC usually requires the strict symmetric power and chromatic dispersion (CD) profile between the two optical links because the mid-link OPC conjugates the optical field of the signals at the mid-point of the transmission link to obtain the idlers [24]. Such a requirement is quite challenging in the practical fiber transmission system, which also increases the complexity of deployment. Therefore, it is significant to simplify the OPC deployment and simultaneously mitigate the performance penalty caused by the simplification of the OPC due to the asymmetric power and CD.

In this work, we experimentally demonstrated a simple OPC assisted by a CVDNN to equalize the fiber nonlinearity in a WDM 64-QAM coherent optical communication system. In the back-to-back (BTB) WDM wavelength conversion system, high-quality idlers at low pump power are obtained by adjusting the input total power of WDM signals to achieve the optimized OPC, where the power and CD profiles of the fiber links are not symmetrically distributed. Although it can simplify the deployment requirements, the performance of compensation is limited. Therefore, a CVDNN is proposed in the digital receiver to further compensate the fiber nonlinearity and improve the compensation performance of the OPC scheme. The proposed method has a strong robustness to input power and signal wavelength and a simple two-hidden-layer structure with a general nonlinear function model to ensure the low complexity. The optimized K-means algorithm proposed in [10] is compared to experimentally verify the CVDNN outperformance.

The remainder of the paper is organized as follows. In Section II, we introduce an OPC experimental system for WDM 64-QAM signals and optimize the input signal power into the HNLF. In Section III, the structure of CVDNN is first described followed by the analysis of the optimization algorithm. Then in Section IV, we present the design of CVDNN, including the selection of significant parameters and the complexity analysis. Section V shows the experimental setup of a WDM 64-QAM transmission system with deployed mid-link OPC and discusses the experimental results with or without CVDNN or optimized K-means in detail. The paper is finally concluded in Section VI.

II. Optimization of OPC in WDM System

The experimental setup of a 12.5-GBd WDM 64-QAM BTB OPC system is shown in Fig. 1. The transmitter is a WDM signal source with four channels spaced at 50-GHz. An arbitrary waveform generator (AWG) with a 10-bit digital-to-analog converter (DAC) generates two 8-level 12.5-GBd electrical signals from the pseudo-random binary sequence (PRBS) with the length of 2^{15}–1 operating at 50-GSa/s. And the signals are de-correlated by the different delay to drive the in-phase/quadrature Mach-Zehnder modulator (I/Q MZM). The four-channel WDM signals are generated by coupled optical carriers from four external cavity lasers (ECLs) operated on a 50-GHz frequency grid extending from 194.3-THz to 194.45-THz and modulated by the I/Q MZM.

The modulated optical signals of the transmitter are launched into the OPC, where the signals are amplified by an EDFA and filtered by an optical bandpass filter (OBPF) with 1-nm bandwidth and 3.34-dB insertion loss to remove the ASE noise of the EDFA. Then, a continuous wave laser is emitted by a tunable laser source (TLS) with the wavelength of 1547.316-nm, which is amplified by another EDFA after an attenuator (ATT) and also filtered by an OBPF with 1-nm bandwidth and 1.74-dB insertion loss as the pump. In order to maximize FWM efficiency, two polarization controllers (PCs) are respectively used to adjust the polarizations of the pump and the signals. After that, an optical coupler with the ratio of 90:10 couples the pump and the signals (pump: 90% and signals: 10%). An optical isolator is used to protect the EDFAs by blocking the reflected waves induced by the stimulated Brillouin scattering (SBS). Finally, the coupled pump and signals trigger the FWM effect in the Ge-doped HNLF with the parameters in Table I, and the output of HNLF is then injected into the OBPF with 1.9-nm bandwidth and 4-dB insertion loss to filter out the idlers. Meanwhile, in the BTB OPC stage, the OBPF is set at 0.2-nm bandwidth to

TABLE I
PARAMETERS OF THE GE-DOPED HNLF

Quantity	Symbol	Value
Effective length	L_{eff}	99 m
Nonlinear coefficient	γ	11.3 $W^{-1}\cdot km^{-1}$
Attenuation	α	8.2 dB/km
Dispersion slope	β	0.017 $ps/nm^2/km$
Zero dispersion wavelength	λ_0	1545 nm
SBS threshold · effective length	SL	18 W·m

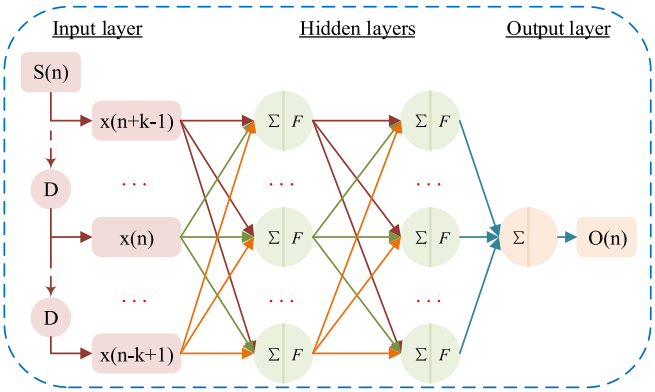


Fig. 2. The block diagram of the proposed CVDNN with two hidden layers.

directly filter out the idler of the channel of interest (COI) for the performance evaluation.

At the receiver side, the idler of COI firstly is amplified by the EDFA and attenuated to 6.5-dBm by the ATT before injected into the coherent receiver consisting of a 100-kHz- linewidth local oscillator (LO), 90 degrees optical hybrid and balanced photo detectors (BPDs). After a 100-GSa/s digital storage oscilloscope (DSO), the digitized data is then processed offline by DSP algorithms for down sampling, the correction of the delay and orthogonality, removal of inter-symbol interference (ISI) by finite impulse response (FIR) filter and carrier phase recovery.

At last, BER is calculated and the *Q*-factor is derived from BER to evaluate the performance as:

$$Q = \sqrt{2} \times erfcinv(2 \times BER)$$
$$Q(dB) = 10 \times \log_{10}(Q^2) = 20 \times \log_{10}(Q) \quad (1)$$

where *erfcinv* is the inverse of complementary error function [25].

In our previous study of the single-channel OPC system, the input signal power of the HNLF needs to be optimized to obtain high-quality idler due to the non-ideal nonlinear crosstalk, which will be more important in the WDM system [24]. The third channel with the wavelength of 1542.54-nm is selected as the COI. According to the parameters in Table I, the maximum input pump power into the HNLF is set at 23.5-dBm to avoid the SBS effect and achieve the maximum conversion efficiency (CE) of −14.9-dB. The output spectrum (0.1-nm resolution) of Ge-doped HNLF is shown in Fig. 1(a). By adjusting the total input signal power into the HNLF, the corresponding Q-factor

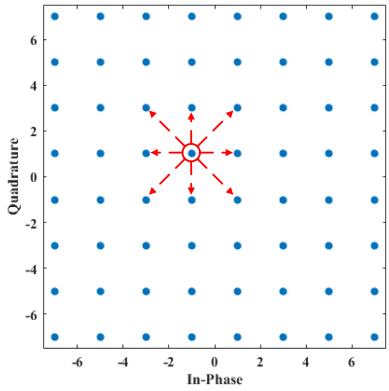


Fig. 3. Constellation of 64-QAM signal, showing the phase distortion of one symbol from adjacent symbols.

is measured, as shown in Fig. 1(b). Obviously, when the total power of four WDM signals is above 7.5-dBm, the Q-factor of the idler obtained from the COI begins to decline as a result of the non-ideal nonlinearity between the channels. Consequently, when the OPC is deployed in the transmission link to compensate the optical fiber nonlinearity combined with the proposed CVDNN, the pump and the total input signal power into the HNLF are respectively set at 23.5-dBm and 7.5-dBm.

III. THE FEATURES OF CVDNN

Fig. 2 is the block diagram of the proposed CVDNN, which is composed of an input layer, two hidden layers and an output layer. The acquired data from the digital oscilloscope in the coherent receiver is first processed offline by DSP algorithms of clock and phase recovery. Then, the processed complex-valued QAM symbol sequence S(n) is launched into the CVDNN to generate multiple input sequences *X(n)* with the form [*x(n+k-1)*, …, *x(n-1)*, *x(n)*, *x(n+1)*, …, *x(n-k+1)*. Here, a one-symbol slide shift operation is exploited and the memory length is $K = 2k$-1, to be determined to track the dependencies between symbols. Each input symbol sequence *X(n)* with the length K corresponds to one symbol output O(n).

When an input training sequence passes from the input layer to the next hidden layer, it needs to be multiplied by a complex weight coefficient plus a bias. The output of each link is first summed up at each neuron in the next layer. Then, the sum is processed by a nonlinear activation function. The input-output relationship between each layer of the CVDNN is written as:

$$h^1 = F\left(\sum_{i=1}^{K} w_i^1 x_i + b_i^1\right)$$
$$h^2 = F\left(\sum_{i=1}^{M} w_i^2 h_i^1 + b_i^2\right)$$
$$o = \sum_{i=1}^{N} w_i^3 h_i^2 + b_i^3 \quad (2)$$

where h^1, h^2 and O represent the outputs of the two hidden layers and the output layer, respectively. w^j and b^j (j = 1, 2, and 3) denote the weights and bias vectors between each layer. M and N are the numbers of two hidden layers (in general $M = N$). The activation function $F(\cdot)$ is a nonlinear function of complex-valued Tanh.

The input vectors of weights and bias are updated iteratively by means of minimizing the minimum mean square error (MSE) between the transmitted and predicted output symbols:

$$MSE = \frac{1}{B}\sum_{i=1}^{B}(o(i) - s(i))^2 \tag{3}$$

where B is the batch size of the total training samples.

A good optimization algorithm is indispensable for achieving the convergence and global optimal solution of NN. The first-order gradient descent algorithm is used commonly because of simple calculation, whereas for high-dimensional nonconvex function space, it is easily stuck at the saddle points [26]. The second-order optimization algorithm can search the optimal solution at the cost of complexity. The Newton algorithm, as a common second-order optimization algorithm, requires to solve the inverse of Hessian matrix of the objective function for each step, increasing significantly computational complexity [27]. Therefore, Broyden, Fletcher, Goldfarb and Shanno proposed a quasi-Newton algorithm, i.e., BFGS algorithm, approximating the inverse of Hessian matrix to a positive definite matrix, thus simplifying the operation complexity [28]. Whereas BFGS algorithm also requires a number of storage units for each update encountering large-scale optimization problems, which weakens the memory efficiency of BFGS. Limited-memory BFGS (LBFGS) was further introduced, and also extended to the complex-valued domain in [29]. Compared with BFGS, LBFGS only needs to store the information of the latest m iterations to approximate the inverse of Hessian matrix, releasing the pressure of storage and showing outstanding convergence performance, both in real and complex domain [30].

The memory size m is an important parameter for the stability of the LBFGS algorithm due to the selection of m being not oriented [31]. Therefore, the parameter m of the LBFGS in the proposed CVDNN has to been optimized. We set a candidate interval where different m is stored, and calculate the searching direction for each m. The final searching direction is the overlapping direction of these calculated directions, so as to avoid the impact of the various m on the algorithm performance.

IV. The Analysis of CVDNN Design

Before verifying that the proposed neural network can effectively compensate the signals' nonlinear distortion caused by the optical fiber in WDM coherent optical transmission system, the influence of the neural network's characteristics on the compensation effect needs to be analyzed.

A. Selection of Basic Parameters (Number of Input Neurons, Number of Hidden Layer Neurons)

The increase of injected power in the fiber will trigger the nonlinear Kerr effect and distort the pulse of the transmitted signal, which becomes serious in the WDM system. Nonlinear pulse phase distortion of one symbol from adjacent symbols as shown in Fig. 3, especially for symbols with high amplitude value in the periphery of the advanced modulation format.

Therefore, when we selected the N-th symbol and its k-1 preceding and k-1 succeeding symbols as input, sufficient symbols should be taken into consideration to cover 64 constellations, especially for that with high amplitude value in the periphery. Fig. 4 shows the constellations of the 64-QAM signal with 100, 200, 300 and 500 symbols selected randomly. Obviously, when the number of symbols is less than 300, there is no symbol in some symbol positions of the constellation diagrams, as shown in Fig. 4(a), Fig. 4(b) and Fig. 4(c). If the number of input neurons is less than 300, the exact functional relationship between input and output cannot be modeled, where the network training will fall into an under-fitting state and the BER is worsen. Whereas the number of input neurons is overwhelming, the calculation amount of training will be increased dramatically. Therefore, the number of input neurons in the proposed CVDNN is set to 325, and the number of both hidden layers' neurons are empirically fixed at 20. Moreover, the early stopping strategy is exploited to solve the over-fitting issue, where the CVDNN is trained by the training set and simultaneously the error of validation set is monitored. Generally, the training and validate errors decrease steadily while the iteration is in progress. Once the validate error begins to increase, the training will be terminated.

B. Complexity Analysis

As the training process only needs to be performed once offline, the calculation of the algorithm is mainly in the operational stage for the practical system. The amount of multiplications required to equalize each bit is used to evaluate the complexity of the algorithm, which for CVDNN can be calculated as:

$$\frac{4 \times (N_{in} \times N_{h1} + N_{h1} \times N_{h2})}{\log_2 M} \tag{4}$$

where, N_{in}, N_{h1}, N_{h2} and M respectively represent the number of input neurons, the first-hidden-layer neurons, the second-hidden-layer neurons and the modulation order of the QAM signal. The factor of multiplier is 4 because the multiplication of a complex value is equal to that of four real values.

It can be seen that the amount of multiplications for CVDNN is dependent on the number of neurons in the input layer and the hidden layer. For a given network discussed in section A with the number of input neurons and both hidden layers are respectively 325 and 20, the calculation amount of per bit for CVDNN is fixed, i.e., $4 \times (325 \times 20 + 20 \times 20)/6 = 4600$.

Once the structure of the network and the modulation format of the signal are determined, the multiplications of each bit for CVDNN are fixed, which does not change with the length of the transmission link, etc. On the other hand, for the DBP algorithm,

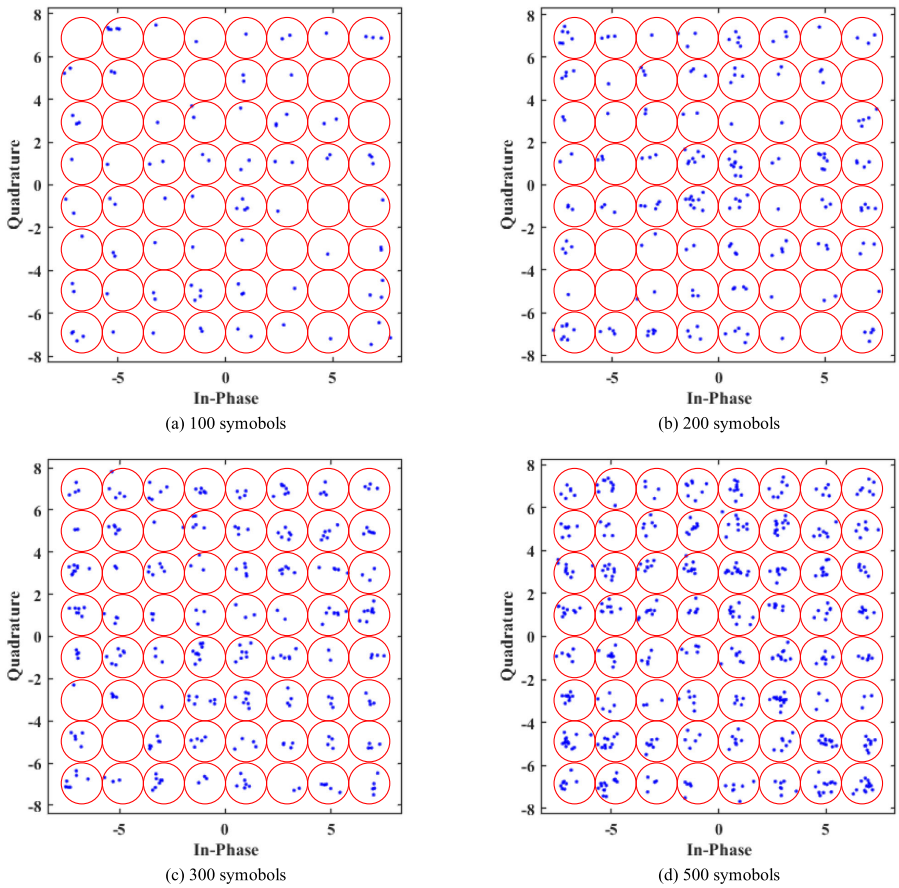


Fig. 4. Constellation of 64-QAM signal with random (a) 100 symbols, (b) 200 symbols, (c) 300 symbols, (d) 500 symbols.

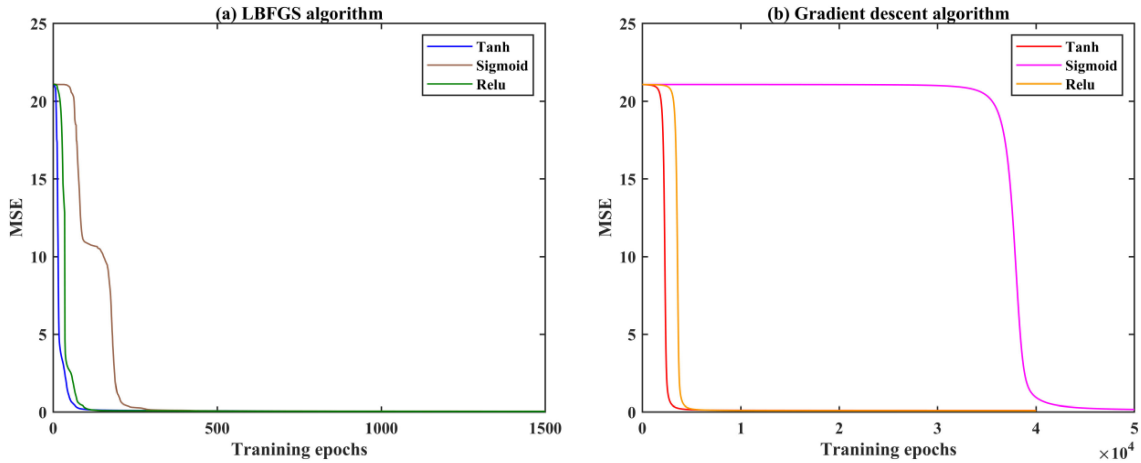


Fig. 5. MSE curve versus training epochs for different activation function with different optimization algorithms: (a) LBFGS algorithm, (b) Gradient descent algorithm.

the amount of multiplications is much affected by the transmission distance and the chromatic dispersion accumulation of link. For large-capacity long distance transmission, the size of FFT will increase, which will greatly increase the complexity of the DBP algorithm.

Meanwhile, it's worth noting that the convergence speed in the training stage with the application of LBFGS algorithm is faster than that of gradient descent algorithm whatever activation function is used, as shown in Fig. 5. On the other hand, when the network begins to converge, the MSE values for these two algorithms are different. LBFGS is approximately 0.04 and gradient descent algorithm is approximately 0.1. Because the first-order gradient descent algorithm is easily stuck at the saddle points for high-dimensional nonconvex function space, making the network fall into the local optimal state. The MSE value of 0.1 is not far enough for 64-QAM or higher order signals' nonlinearity

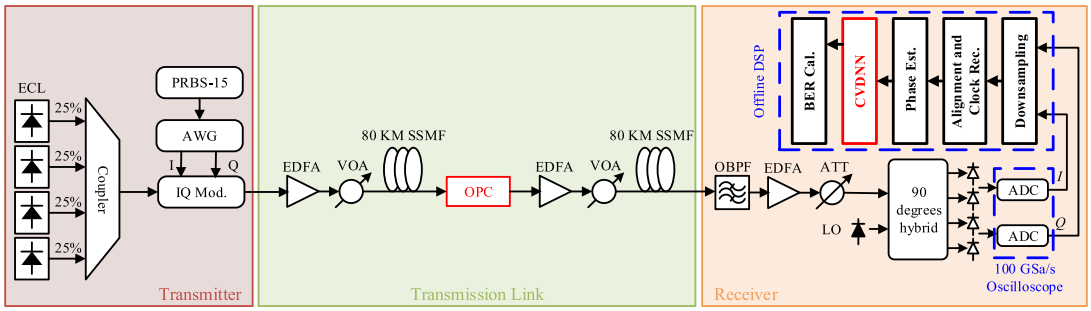


Fig. 6. Experimental setup of a 12.5-GBd WDM 64-QAM transmission system deployed mid-link OPC.

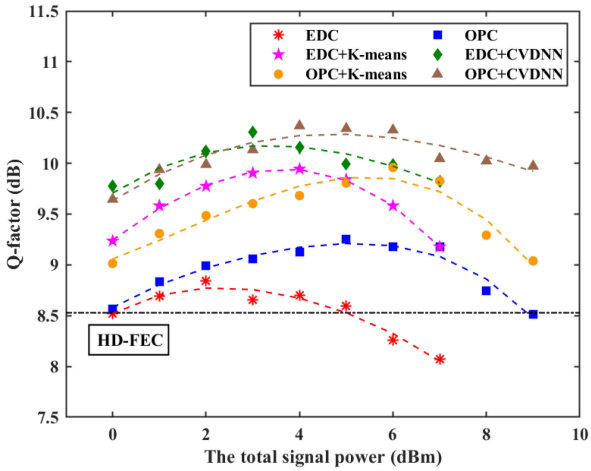


Fig. 7. Measured Q-factor of COI versus different total signal power values.

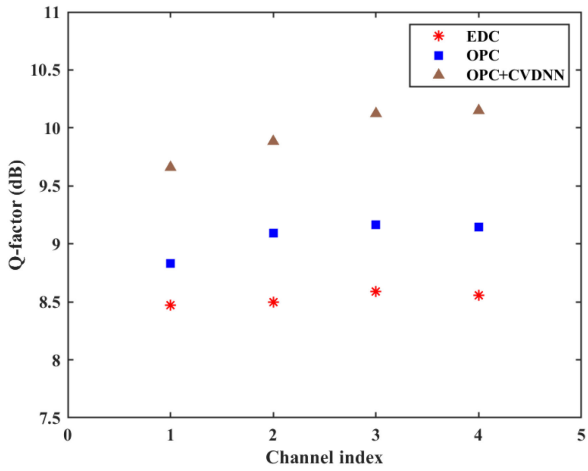


Fig. 8. Measured Q-factor of the four WDM channels.

compensation. Hence, a fast-convergent optimization algorithm is important for network training, which can significantly reduce the training time and obtain the better training network.

In our proposed CVDNN, the complex-valued Tanh activation function is used and the maximum training epoch is set as 3000. In the training stage, the total amount of multiplications can be calculated as:

$$4 \times N_t \times N_{ep} \left(N_{in} \times N_{h1} + N_{h1} \times N_{h2}\right) \tag{5}$$

where N_t and N_{ep} are the total number of training samples and the number of training iterations. In the case of a determinate network structure, the multiplication amount of training stage depends on the number of training samples and the iteration epochs.

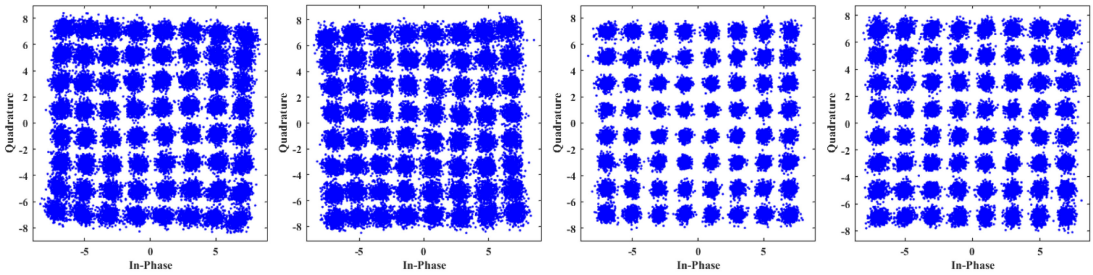
V. The Experimental Results

A. The Compensation Performance of CVDNN

The optimized OPC system presented in the Section II is deployed in the mid-link of an EDFA-amplified WDM 64-QAM coherent optical transmission experimental system, as shown in Fig. 6. The transmitter and receiver have been already described in the Section II and the transmission link consists of two 80-km standard single-mode fibers (SSMF) with a loss of 0.18-dB/km and a dispersion of 16.09-ps/nm^2. The span loss is compensated by EDFAs, and the variable optical attenuators (VOAs) are used to adjust the power launched into the SSMFs. At the receiver, the proposed CVDNN is inserted before BER calculation to further improve the accuracy of the decision in the offline DSP.

Firstly, in order to verify the performance of the proposed CVDNN, we compare the compensation performance of the proposed CVDNN and the optimized K-means algorithm in the case of without OPC, as shown in Fig. 7. Both the CVDNN and the optimized K-means algorithm can well equalize the signal distortion as a result of optical fiber nonlinearity, as shown by red asterisk-marked, magenta pentagram-marked and green diamond-marked curves. Here, electrical dispersion compensation (EDC) denotes the conventional coherent receiver with electrical dispersion compensation. Here, the conventional linear decision is exploited and the symbol decision lattices are plotted according to the original QAM mapping. In other two cases, besides the EDC, the K-means or the CVDNN algorithm is included respectively. Compared with the K-means algorithm of optimizing the symbols' decision margin, the proposed CVDNN, as a black box, can model the nonlinear functional relationship between input and output, which can effectively equalize the signal with serious nonlinearity and outperform the K-means method.

Subsequently, the compensation performance of the optimized OPC system is also compared, as shown by blue square-marked, orange circle-marked and brown triangle-marked



(a) EDC for 7 dBm (b) OPC for 9 dBm (c) EDC+CVDNN for 7 dBm (d) OPC+CVDNN for 9 dBm

Fig. 9. The constellation diagrams of COI at different launched signal powers for different methods.

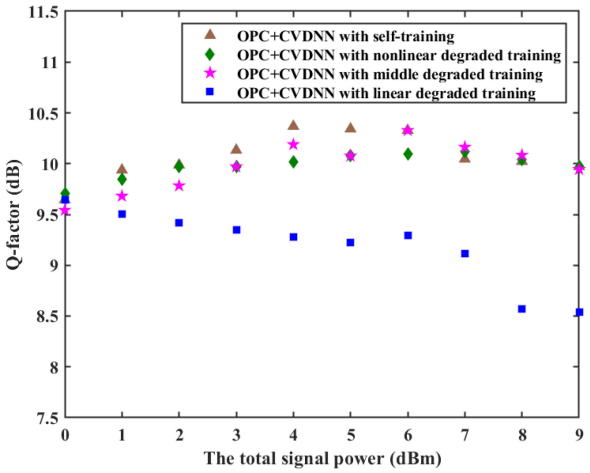


Fig. 10. Measured Q-factor versus different total signal power for different training modes.

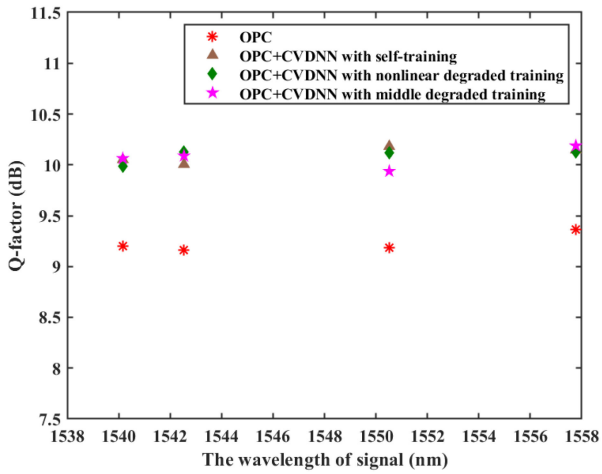


Fig. 11. Measured Q-factor of COI from different bands for different training modes.

curves in Fig. 7. With the assistance of the OPC, the Q factor of the degraded signal and the optimum launched power into SSMFs can be greatly improved, as shown by the red asterisk-marked curve and the blue square-marked curve. For the K-means algorithm, the maximum Q-factor is increased to 9.8-dB in the experiments with or without OPC. Compared with the conventional EDC with the K-means, the OPC-assisted one raises the optimum input signal power more than 2-dB, as shown by the magenta pentagram-marked curve and the orange circle-marked curve in Fig. 7, which will enable more channels launched into the fiber. On the other hand, the CVDNN algorithm enables the Q-factor of the WDM signals to increase to 10.4-dB. Simultaneously, the OPC with CVDNN has a wider range of optimum input signal power and can effectively compensate for the distorted signals, as shown by the brown triangle-marked curve.

Fig. 8 shows the measured Q-factor of the four WDM channels at the optimal launched power for the EDC, the OPC and the OPC assisted by CVDNN schemes. With the EDC, the average Q-factor of the four signals approximately is 8.5-dB, but it is increased to 10-dB after the OPC with CVDNN. The difference in the first channel is caused by the generation of harmonics in the HNLF, as shown in the spectra of inset Fig. 1(a). Furthermore, the constellations of COI at different launched power values 7-dBm and 9-dBm of various abovementioned methods EDC, OPC, EDC with CVDNN and OPC with CVDNN, are presented in Fig. 9. Nevertheless, higher input signal power aids to achieve higher optical signal noise ratio and the triggered fiber nonlinearity induces the rotation of the signal constellation, as shown in Fig. 9(a). The OPC can further increase the launched signal power and the signal constellation is shown in Fig. 9(b). The K-means algorithm can improve the BER performance by optimizing the symbols' decision margin and such a nonlinear symbol decision depends on the Euclidean distance of each symbol to the centroid of achieved constellation points. Thus, the signals' constellation diagram will not be changed. Whereas, the CVDNN can learn the knowledge from the degraded constellations and alleviate their noise to greatly improve the accuracy of the conventional linear symbols' decision, as shown in Fig. 9(c) and Fig. 9(d). To sum up, comparing to the conventional EDC, the proposed CVDNN can assist the OPC to achieve the best performance and ease the strict deployment requirements.

B. The Robustness of CVDNN

In Fig. 10, we verify the robustness of CVDNN in the OPC-assisted networks under four training conditions of different launched powers of WDM signals, i.e., linear degraded training of 0 dBm, middle degraded training of 6 dBm, nonlinear degraded training of 9 dBm and self-training of the experimentally

measured results in Fig. 7. It is obvious that the Q-factor compensated by the CVDNN, whose weights and bias are obtained from the middle degraded training and the nonlinear degraded training, has a negligible difference, compared with that from the self-training, when the total launched power of WDM signals ranges from 0 dBm to 9 dBm. However, the above results are quite different with that from the linear degraded training, because the nonlinear relation between the input and the output learned during the CVDNN training has been changed. For NNs, the training processing needs many resources and the look-up-table (LUT) in a dynamic practical system is required. It is notable that a wider training range can reduce the updating period of LUTs. Then, we evaluate the performance of the COI at different wavebands (B1: 1539.371-nm∼1540.558-nm, B2: 1541.747-nm∼1542.936-nm, B3: 1549.716-nm∼1550.918-nm, B4: 1556.959-nm∼1558.173-nm), as is shown in Fig. 11. In the OPC with CVDNN scheme, the above mentioned various degraded signals are used for training. The measured results in Fig. 11 are consistent with those in Fig. 10.

VI. Conclusion

We propose and experimentally demonstrate a simple OPC scheme with CVDNN to compensate signal degradation caused by fiber nonlinearity in WDM 64-QAM coherent optical transmission systems. In the proposed scheme, the OPC is first utilized to compensate for the fiber nonlinearity and increase the optimum launched signal power to the fiber. Next, the simple two-hidden-layer structure CVDNN is proposed to further improve the signals' performance and alleviate the requirement of the strict symmetrical distribution of power and dispersion in the OPC. The CVDNN enables the system to learn the nonlinear functional relation between the input and the output and outperforms the K-means algorithm of the optimal symbols' decision. Meanwhile, the proposed scheme has a strong robustness to the launched signal power and the signal wavelength, which dramatically reduces the complexity of the training process. The proposed scheme has a good application prospect in the future broadband dynamic optical network.

References

[1] D. Richardson, "Filling the light pipe," *Science*, vol. 330, no. 6002, pp. 327–328, Oct. 15, 2010.

[2] F. Buchali *et al.*, "DCI field trial demonstrating 1.3-Tb/s single-channel and 50.8-Tb/s WDM transmission capacity," *J. Lightw. Technol.*, vol. 38, no. 9, pp. 2710–2718, May 2020.

[3] G. Rademacher *et al.*, "10.66 Peta-bit/s transmission over a 38-core-three-mode fiber," in *Proc. Opt. Fiber Commun. Conf. Exhib. (OFC)*, San Diego, CA, USA, 2020, pp. 1–3.

[4] X. Chen, S. Chandrasekhar, J. Cho, and P. Winzer, "Transmission of 30-GBd polarization-multiplexed probabilistically shaped 4096-QAM over 50.9-km SSMF," *Opt. Exp.*, vol. 27, no. 21, pp. 29916–29923, Oct. 2019.

[5] S. Okamoto *et al.*, "A study on the effect of ultra-wide band WDM on optical transmission systems," *J. Lightw. Technol.*, vol. 38, no. 5, pp. 1061–1070, Mar. 2020.

[6] R. Essiambre, G. Kramer, P. Winzer, G. Foschini, and B. Goebel, "Capacity limits of optical fiber networks," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 662–701, Feb. 2010.

[7] A. Amari, O. Dobre, R. Venkatesan, O. Kumar, P. Ciblat, and Y. Jaouën, "A survey on fiber nonlinearity compensation for 400 Gb/s and beyond optical communication systems," *IEEE Commun. Surv. Tut.*, vol. 19, no. 4, pp. 3097–3113, Nov. 2017.

[8] E. Ip and J. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *J. Lightw. Technol.*, vol. 26, no. 20, pp. 3416–3425, Oct. 2008.

[9] F. Guiomar and A. Pinto, "Simplified Volterra series nonlinear equalizer for polarization-multiplexed coherent optical systems," *J. Lightw. Technol.*, vol. 31, no. 23, pp. 3879–3891, Dec. 2013.

[10] J. Zhang, W. Chen, M. Gao, and G. Shen, "K-means-clustering-based fiber nonlinearity equalization techniques for 64-QAM coherent optical communication system," *Opt. Exp.*, vol. 25, no. 22, pp. 27570–27580, Oct. 2017.

[11] W. Chen, J. Zhang, M. Gao, and G. Shen, "Performance improvement of 64-QAM coherent optical communication system by optimizing symbol decision boundary based on support vector machine," *Opt. Commun.*, vol. 410, pp. 1–7, Mar. 2018.

[12] S. Zhang *et al.*, "Field and lab experimental demonstration of nonlinear impairment compensation using neural networks," *Nature Commun.*, vol. 10, no. 3033, pp. 1–8, Jul. 2019.

[13] Q. Fan *et al.*, "Advancing theoretical understanding and practical performance of signal processing for nonlinear optical communications through machine learning," *Nature Commun.*, vol. 11, no. 3694, pp. 1–11, Jul. 2020.

[14] M. Secondini, E. Agrell, E. Forestieri, D. Marsella, and M. Camara, "Nonlinearity mitigation in WDM systems: Models, strategies, and achievable rates," *J. Lightw. Technol.*, vol. 37, no. 10, pp. 2270–2283, May 2019.

[15] D. Vukovic *et al.*, "Multichannel nonlinear distortion compensation using optical phase conjugation in a silicon nanowire," *Opt. Exp.*, vol. 23, no. 3, pp. 3640–3646, Feb. 2015.

[16] C. Catanese, A. Triki, E. Pincemin, and Y. Jaouën, "A survey of neural network applications in fiber nonlinearity mitigation," in *Proc. 21st Int. Conf. Transparent Opt. Netw.*, Angers, France, 2019, pp. 1–4.

[17] Z. Xu, C. Sun, T. Ji, J. Manton, and W. Shieh, "Cascade recurrent neural network-assisted nonlinear equalization for a 100-Gb/s PAM4 short-reach direct detection system," *Opt. Lett.*, vol. 45, no. 15, pp. 4216–4219, Aug. 2020.

[18] S. Deligiannidis, A. Bogris, C. Mesaritakis, and Y. Kopsinis, "Compensation of fiber nonlinearities in digital coherent systems leveraging long short-term memory neural networks," *J. Lightw. Technol.*, vol. 38, no. 21, pp. 5991–5999, Nov. 2020.

[19] P. Freire *et al.*, "Complex-valued neural network design for mitigation of signal distortions in optical links," *J. Lightw. Technol.*, vol. 39, no. 6, pp. 1696–1705, Mar. 2021, doi: 10.1109/JLT.2020.3042414.

[20] S. Liu *et al.*, "A multilevel artificial neural network nonlinear equalizer for millimeter-wave mobile fronthaul systems," *J. Lightw. Technol.*, vol. 35, no. 20, pp. 4406–4417, Oct. 2017.

[21] S. Yoshima *et al.*, "Mitigation of nonlinear effects on WDM QAM signals enabled by optical phase conjugation with efficient bandwidth utilization," *J. Lightw. Technol.*, vol. 35, no. 4, pp. 971–978, Feb. 2017.

[22] H. Zhang, Q. Zhang, C. Huang, and C. Shu, "Transmission impairment mitigation for single-sideband signals by optical phase conjugation," *IEEE Photon. Technol. Lett.*, vol. 32, no. 3, pp. 150–153, Feb. 2020.

[23] G. Saavedra, G. Liga, and P. Bayvel, "Volterra-assisted optical phase conjugation: A hybrid optical-digital scheme for fiber nonlinearity compensation," *J. Lightw. Technol.*, vol. 37, no. 10, pp. 2467–2479, May 2019.

[24] L. Wang, M. Gao, M. Liu, H. Zhu, B. Chen, and L. Xiang, "Energy-efficient all optical wavelength converter for optical phase conjugation," *Opt. Fiber Technol.*, vol. 58, Sep. 2020, Art. no. 102278.

[25] W. Freude *et al.*, "Quality metrics for optical signals: Eye diagram, Q-factor, OSNR, EVM and BER," in *Proc. 14th Int. Conf. Transparent Opt. Netw.*, Coventry, U.K., 2012, pp. 1–4.

[26] T. Nitta, "An extension of the back-propagation algorithm to complex numbers," *Neural Netw.*, vol. 10, no. 8, pp. 1391–1415, Nov. 1997.

[27] D. Xu, J. Dong, and C. Zhang, "Convergence of quasi-newton method for fully complex-valued neural networks," *Neural Process. Lett.*, vol. 46, no. 3, pp. 961–968, Dec. 2017.

[28] J. Nocedal, "Updating quasi-Newton matrices with limited storage," *Math. Computation*, vol. 35, no. 151, pp. 773–782, Jul. 1980.

[29] H. Zhang and D. Mandic, "Is a complex-valued stepsize advantageous in complex-valued gradient learning algorithms?," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 12, pp. 2730–2735, Dec. 2016.

[30] L. Sorber, M. Barel, and L. Lathauwer, "Unconstrained optimization of real functions in complex variables," *SIAM J. Optim.*, vol. 22, no. 3, pp. 879–898, Jul. 24, 2012.

[31] R. Wu *et al.*, "A L-BFGS based learning algorithm for complex-valued feedforward neural networks," *Neural Process. Lett.*, vol. 47, no. 3, pp. 1271–1284, Jun. 2018.