

# Data-Driven Method for Nonlinear Optical Fiber Channel Modeling Based on Deep Neural Network

Rui Jiang , Ziling Fu , Yansheng Bao, Huiying Wang, Xin Ding, and Zhi Wang

**Abstract**—Recently, data-driven fiber channel modeling methods based on deep learning have been proposed in optical communication system simulations. We investigate a new data-driven method based on the deep neural network (DNN) to model the nonlinear fiber channel with the characteristics of attenuation, chromatic dispersion, amplified spontaneous emission noise, self-phase modulation (SPM), and cross-phase modulation (XPM). Demonstration in multiple dimensions, including constellations, optical waveforms, spectra, and the normalized mean square error, shows that DNN can approach the transfer function of the fiber channel accurately. Additionally, the DNN shows good generalization for modulation formats and wavelength schemes. Besides, the time complexity of DNN-based method for modeling nonlinear fiber channel is reduced significantly (96.5%) compared to the conventional model-driven method, which is based on the split-step Fourier method. This work demonstrates that the DNN can model accurately the nonlinear fiber channel that takes account of both SPM and XPM. Therefore, it can contribute to the application of data-driven methods in modern optical communication system simulations and designs.

**Index Terms**—Data-driven, deep learning, deep neural network, fiber channel modeling, fiber optics systems, and split-step fourier method.

## I. INTRODUCTION

**S**IMULATIONS are vital in optical communication system designs [1], [2]. Conventional optical communication system simulations are based on a series of blocks that are characterized by rigorous numerical models, including a laser, modulator, fiber channel, optical amplifier, filter, detector, and analyzer. [3] Therefore, for the model-driven method, it is a systematic engineering task that requires expert knowledge to construct a comprehensive and complete optical communication system simulation. This is why business optical communication simulation software are usually non-open and expensive. Furthermore, the computation complexity of conventional simulations can be very high due to the nested-function structure and the repeated iterative operations, especially the split-step Fourier method (SSFM) which is performed to model the fiber channel

by solving the nonlinear Schrödinger equation (NLSE) [1], [4], [5]. Therefore, a method for optical fiber channel modeling with relatively low computation complexity is quite valuable.

Deep learning (DL) is a powerful tool that has dramatically improved the state-of-the-art in many domains, such as speech recognition, visual object recognition, object detection, and drug discovery [6]. DL has also been widely applied in the field of optical communication [7], including optical performance monitoring [8], [9], nonlinearity compensation [10], [11], equalizer [12]–[14], predistortion [15], [16], software-defined networking [17], [18], and photonic device design [19], [20]. Recently, some methods that are based on DL have been proposed for optical fiber channel modeling. These methods can be divided into two categories. One is the principle-driven method [21]–[23], and the other is the data-driven method [1], [2]. The principle-driven method views the fiber modeling problem as a partial differential equation solving problem and fully considers the prior knowledge including the essential mathematical equations, physical theories, and the corresponding constraint conditions of the target problems, meaning that it requires much human expertise. In addition, it is usually applied in the pulse evolution task rather than the signal transmission task which is the focus of this work [23]. Thus, the data-driven method that is designed for the signal transmission task is taken into our consideration in the following content. The data-driven method which regards the fiber modeling problem as a data regression task is proposed based on the fact that the deep neural network (DNN) can be considered as a universal approximator for both linear and nonlinear functions [24]. DNN can approximate the channel transfer functions after being trained on the data set which is composed of the channel input and output signals. The data-driven method does not require complex mathematical theories along with expert knowledge and has a distinct advantage in computation complexity for the fact that no complicated operation, such as fast Fourier transform (FFT), is involved.

The data-driven fiber channel modeling method based on DL is first introduced by Danshi Wang [1]. In their work, bidirectional long short-term memory (BiLSTM) neural networks are built to model optical fiber channels for on-off keying (OOK) and pulse amplitude modulation 4 (PAM4) signals. The BiLSTM has learned the approximate transfer function of the fiber channel, and the computation time of fiber channel modeling by their data-driven method is reduced by 80% compared with the model-driven method. However, the modulation format is limited to OOK and PAM4, meaning that no advanced modulation format is studied. In [2], Hang Yang builds a generative

Manuscript received March 21, 2022; revised June 7, 2022; accepted June 15, 2022. Date of publication June 17, 2022; date of current version June 30, 2022. This work was supported in part by Beijing Natural Science Foundation under Grant L201021, and in part by National Key R&D Plan under Grant 2021YFB2900700. (Corresponding author: Zhi Wang.)

The authors are with the Institute of Optical Information, School of Science, Beijing Jiaotong University, Beijing 100044, China (e-mail: 20118044@bjtu.edu.cn; 19118022@bjtu.edu.cn; 19118021@bjtu.edu.cn; 17118463@bjtu.edu.cn; 20121649@bjtu.edu.cn; zhiwang@bjtu.edu.cn).

Digital Object Identifier 10.1109/JPHOT.2022.3184354

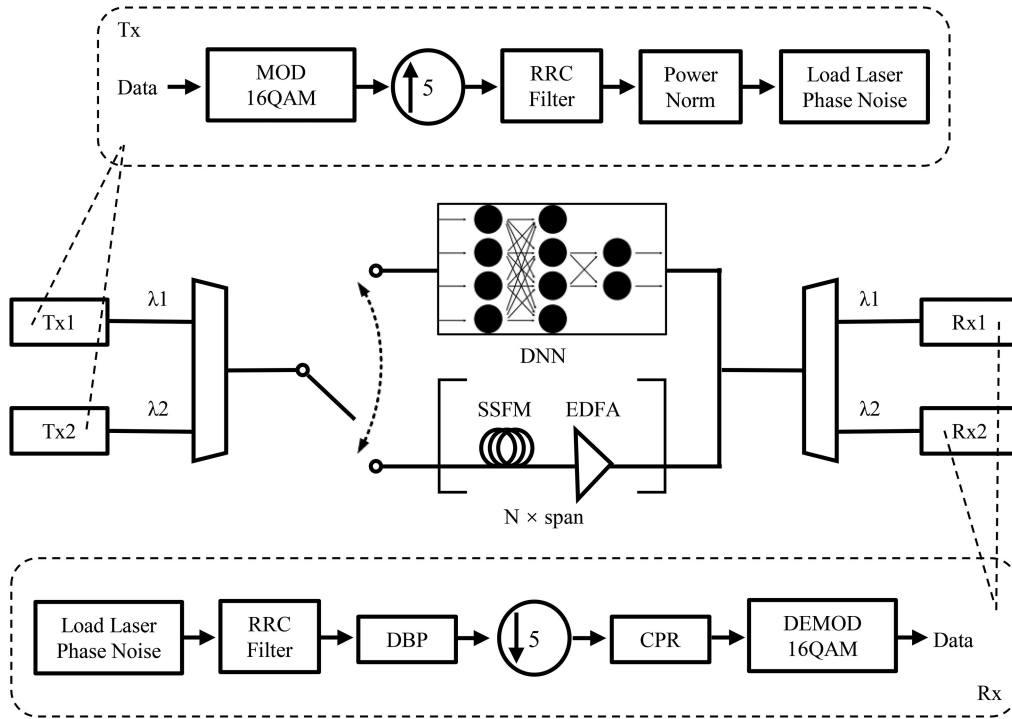


Fig. 1. Optical communication system setup.

adversarial network (GAN) to learn the distribution of the fiber channel transfer function. In their investigation, many channel effects have been taken into consideration, including attenuation, chromatic dispersion (CD), self-phase modulation (SPM), and amplified spontaneous emission (ASE) noise which is induced by erbium-doped fiber amplifier (EDFA). GAN has successfully learned the accurate transfer function of the fiber channel and reduces the complexity of the fiber channel modeling remarkably. However, with only SPM involved, other nonlinear effects, such as cross-phase modulation (XPM) which is a quite important impairment during transmission in the modern optical communication system, have not been studied. Apart from this, GAN can be difficult to train and it is often observed in practice that gradient descent based GAN optimization does not lead to convergence [25].

In this research, we choose the DNN built with fully connected neural networks, which can be trained much easier than GAN, to model the nonlinear optical fiber channel with the characteristics of attenuation, CD, ASE caused by EDFA, SPM, and XPM. The capability of DNN to approach the transfer function of fiber channel is demonstrated in multiple dimensions, including constellations, optical waveforms, spectra, and the normalized mean square error (MSE). Results show that the DNN can model the nonlinear fiber channel accurately. Besides, after being trained on the training data set that contains only 16 quadrature amplitude modulation (QAM) symbols, the DNN-based fiber model made a precise prediction of the fiber channel output signals with other modulation formats, such as quadrature phase-shift keying (QPSK), indicating that the DNN has a good generalization ability for modulation formats. The analysis also shows that

the DNN generalizes well for wavelength schemes. In addition, the time complexity of the DNN-based method is analyzed as well, indicating that the computing time is relatively reduced by 96.5% compared with the SSFM-based method. Therefore, this method can be an auxiliary tool for future optical communication system simulations.

## II. OPTICAL FIBER COMMUNICATION SYSTEM STRUCTURE

Throughout this work, we focus on several impairments in the optical fiber channel. Two effects of Kerr nonlinearity, SPM and XPM effects, are among our consideration. SPM is an intra-channel nonlinear effect, while the inter-channel XPM effect involves two-channel interactions [5]. Considering that the nonlinear interference contributions of multiple wavelength division multiplexing (WDM) channels can be added up independently, the analysis is performed with only two channels [26]. [27]. A WDM system, whose setup is shown in Fig. 1, is simulated to demonstrate the capability of DNN to model the nonlinear fiber channel. The system consists of transmitters, (de)multiplexers, an optical fiber channel, and receivers. The optical fiber channel is affected by the attenuation, CD, ASE noise, SPM, and XPM. The transmitter is assumed to use 16QAM, so all the symbols and samples in this system are complex-valued. In the transmitter, following the modulation is five times up-sampling and root raised cosine (RRC) filter which is to shape the signal. Then power normalization is used to control the power of the optical signal, and for simplicity, the power values of the two channels are set to be the same. The laser phase noise, which is modeled by a Wiener process [28], is loaded before the two optical carriers

TABLE I  
DETAILED PARAMETERS OF SSFM-BASED FIBER CHANNEL

Parameters	Value
Lambda 1	1550.12 nm
Lambda 2	1549.32 nm
Symbol rate 1 & 2	30G Baud
Symbol length 1 & 2	$1.6 \times 10^6$
Attenuation	0.2dB/km
Dispersion	17 ps/(nm·km)
Nonlinear coefficient	$2e-3$ /W/m
Step distance	20 m
Span length	80 km
EDFA noise figure	5 dB
RRC roll-off factor	0.1

being multiplexed into the optical fiber channel that is composed of standard single-mode fiber (SSMF) and EDFA. The main fiber channel parameters are shown in Table I.

When two optical fields propagate simultaneously inside the fiber, they interact with each other through XPM. On condition that the wavelengths of the two optical beams are so close to each other that the group-velocity mismatch is negligible (i.e.,  $v_{g1} \approx v_{g2}$ ), such as in DWDM systems, the propagation of two optical fields through a single-mode fiber can be governed by the following set of two coupled NLSE [5]:

$$\begin{cases} \frac{\partial A_1}{\partial z} = -\frac{\alpha}{2} A_1 - \frac{i}{2} \beta_{21} \frac{\partial A_1^2}{\partial t^2} + i\gamma \left( |A_1|^2 + 2|A_2|^2 \right) A_1 \\ \frac{\partial A_2}{\partial z} = -\frac{\alpha}{2} A_2 - \frac{i}{2} \beta_{22} \frac{\partial A_2^2}{\partial t^2} + i\gamma \left( |A_2|^2 + 2|A_1|^2 \right) A_2 \end{cases} \quad (1)$$

where  $A_1$  and  $A_2$  are the complex envelope of the slowly varying optical fields and  $z$  is the propagation distance. Parameter  $\alpha$ ,  $\beta_{21}$ ,  $\beta_{22}$ , and  $\gamma$  represent the propagation attenuation, dispersion of the two wavelengths, and nonlinear coefficient, respectively. SSFM is the most common numerical approach to solving the NLSE [1], [2], [5]. It can be expressed by:

$$A_i(z+h, t) = \exp\left(\frac{h}{2}\hat{D}_i\right) \times \exp\left\{h\hat{N}_i\left[A_i\left(z+\frac{h}{2}, t\right)\right]\right\} \exp\left(\frac{h}{2}\hat{D}_i\right) \quad (2)$$

where  $\hat{D}_i$  and  $\hat{N}_i$  denote the linear operator and nonlinear operator of  $i$ th ( $i = 1, 2$ ) optical field, respectively;  $h$  denotes the step size.

After each span, an EDFA is applied to compensate for the attenuation, which would introduce ASE noise simultaneously.

In the receiver, after demultiplexing, the laser phase noise is loaded, following which a matched RRC filter is used, and then multichannel digital backward propagation (DBP) algorithm is performed to compensate for the CD, SPM, and XPM. Then down-sampling is performed and carrier phase recovery (CPR) is used to compensate for the laser phase noise, which is followed by demodulation.

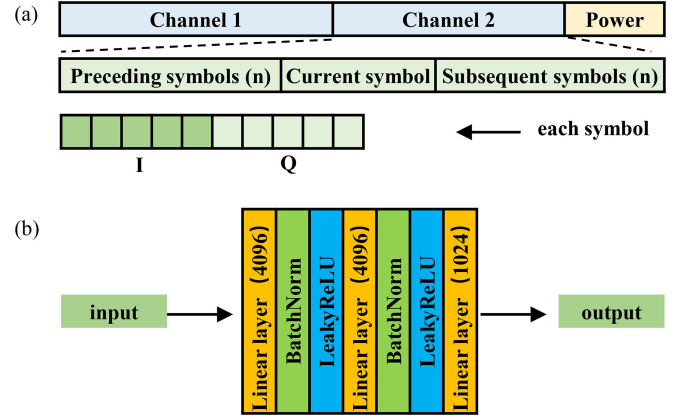


Fig. 2. (a) The input vector structure of DNN. (b) the architecture of the DNN. Linear layer is a fully connected layer. BatchNorm is Batch Normalization. LeakyReLU is an activation function.

Note that, the channel transfer function has no analytical expression, therefore the differences between the nonlinear channel modeled by SSFM and DNN cannot be compared directly. As a result, similar to the work of Hang Yang *et al.* [2], the DBP compensation is utilized in this work as an auxiliary verification method. Let's denote the transfer functions of the nonlinear fiber channel modeled by SSFM and DNN as  $f(\cdot)$  and  $g(\cdot)$  respectively. DBP compensation can be regarded as the inverse function of the channel transfer function  $f^{-1}(\cdot)$ . If the DBP compensated output by DNN is similar to the channel input,  $g(\cdot)$  and  $f^{-1}(\cdot)$  can be regarded as a pair of inverse functions, meaning that  $g(\cdot)$  and  $f(\cdot)$  can be regarded as equivalent, which proves that the DNN has learned the linear and nonlinear characteristics of the fiber channel and can estimate the distribution of the channel transfer function.

### III. DEEP NEURAL NETWORK ARCHITECTURE

A DNN is defined and trained to model the nonlinear fiber channel. As is shown in Fig. 2, the input vector of DNN is defined to improve modeling accuracy and flexibility, which is similar to the definition of the condition vector in [2]. First of all, two channels are simulated in the optical communication system, therefore both of them must be taken into consideration. Secondly, as for each channel, considering inter-symbol interference (ISI) caused by CD, the input vector must include not only the current transmitted symbol but also the preceding symbols and subsequent symbols. The number of the preceding symbols and subsequent symbols, which is denoted by  $n$  in Fig. 2, is set to eight for each span in this paper, after taking account of the ISI and transmission rate of the signal. Thirdly, considering that the up-sampling rate is 5 in simulation, each symbol in the input vector of DNN includes five samples. The real part and imaginary part of the samples, i.e., the in-phase (I) and quadrature (Q) parts of the fiber channel input signals are concatenated so that the input vector of DNN can be real-valued, as is shown in Fig. 2. Finally, the optical launch power is also considered and its value is appended at the rear of the input

vector of DNN. The length of the output vector of DNN is 20, meaning that one symbol per channel is generated by DNN.

The architecture of the DNN is shown in Fig. 2 as well. The input vector must be normalized before being fed to the neural network to avoid slowing down the convergence of the model and control the average value of the input around 1 [2]. Adam optimizer is used and the learning rate is set to 0.001. The loss function is MSE Loss. The normalization is applied to initialize the weights of the linear layers, which would help with the convergence of deep models with ReLu-like activation functions [29]. All the biases are initialized to 0. The batch size is set to 500.

#### IV. DEMONSTRATION AND RESULTS

In the simulation, dispersion and nonlinearity coefficient are kept constant, while transmission distance and power are changed to control the dispersion and nonlinear intensity. To prove the validity of the simulation by DNN, a comparison between the optical signals generated by SSFM- and DNN-based method is made in multiple dimensions, including constellations, optical waveforms, spectra, and the normalized MSE. The constellations after DBP compensation are plotted to present the characteristics of dispersion and nonlinearity, which can verify the accuracy of the channel transfer function modeled by DNN [2]. Optical waveforms and spectra are to demonstrate the accuracy of the simulation by DNN in the time- and frequency- domain. The normalized MSE is to evaluate the similarities between the two simulation methods quantitatively. MSE denotes the average of squares of the amplitude errors, i.e., the average squared deviation between the amplitude values of SSFM-generated and DNN-generated waveforms or spectra. Considering that optical communication systems were simulated with different optical launch powers and that the absolute MSE may increase with the power values, the normalized MSE rather than the absolute one is adopted [1]. The normalized MSE is defined as below:

$$MSE_{nor} = \frac{\sum_i^m (\bar{y} - y)^2}{\sum_i^m \bar{y}^2} \quad (3)$$

where  $m$  is the sample size,  $\bar{y}$  is the output label signal, and  $y$  is the output signal generated by DNN. As is mentioned in [1], [2], the acceptable upper limit of  $MSE_{nor}$  is set to 0.02.

To model the optical fiber channel by DNN, the data set is built by collecting the channel input and output when simulating via SSFM. The input vector, i.e., the training sample for DNN, is generated from the data of fiber input. Then it is fed to DNN, aiming to get an output that approximates the corresponding optical signals of fiber output. The training data set size is  $1 \times 10^6$ .

##### A. The Nonlinear Fiber Channel Modeling Capability of DNN

We first studied the capability of DNN to model the nonlinear fiber channel. As a demonstration, WDM systems whose setup is shown in Fig. 1 were simulated. Then we trained DNNs with the corresponding data sets. For comparison purposes, BiLSTM, the DL algorithm which has been employed in single-channel

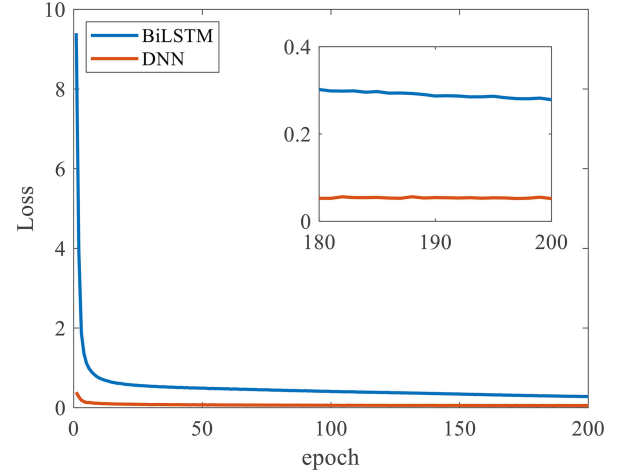


Fig. 3. The training loss of DNN and BiLSTM.

modeling [1], is also implemented by PyTorch. The input size of BiLSTM is 20 ( $5 \times 2 \times 2$ ), which means a symbol of channel 1 is concatenated by another symbol of channel 2. Then every 17 adjacent symbols constitute a group. These groups are sent into BiLSTM in chronological order. The hidden size is 100. Finally, the output of BiLSTM is mapped by a fully connected layer with 20 neurons into 2 symbols, one of which for channel 1 and the other for channel 2.

The losses of DNN and BiLSTM during training are shown in Fig. 3. As is implied in Fig. 3, the losses of DNN and BiLSTM both decrease rapidly. However, the DNN achieves a faster convergence rate. This result is quite different from that of Wang's work, in which the BiLSTM achieves a faster convergence rate [1]. This may attribute to the unsatisfying capability of BiLSTM to learn the XPM effect of fiber channel.

After DNN and BiLSTM are trained, the constellations after DBP compensation of SSFM-generated signals and DNN/BiLSTM-generated signals are plotted for comparison, as is shown in Fig. 4. Note that, the system transmission distance is 80 km. The optical launch power of the two channels is 4 dBm. The format for both channels is 16QAM. The constellations of DNN-generated signals are very similar to that of SSFM-generated signals for both channels, meaning that DNN can mimic the effect of CD, SPM, XPM, and EDFA ASE noise to approach the transfer function of the nonlinear fiber channel. However, the constellations of BiLSTM-generated signals are not very similar to that of SSFM-generated signals. Thus it can be concluded that BiLSTM can not approach the transfer function of the nonlinear fiber channel which takes account of both SPM and XPM as well as DNN.

To demonstrate more diversely and accurately, Fig. 5 shows amplitudes of the optical waveforms and spectra of SSFM-generated signals and DNN-generated signals. Note that, the system transmission is 240 km. The optical launch power is 0 dBm. The format for both channels is 16QAM as well. Considering that the two channels have the same transmission distance, optical launch power, and modulation format, we just present the results of one channel, as is shown in Fig. 5. It

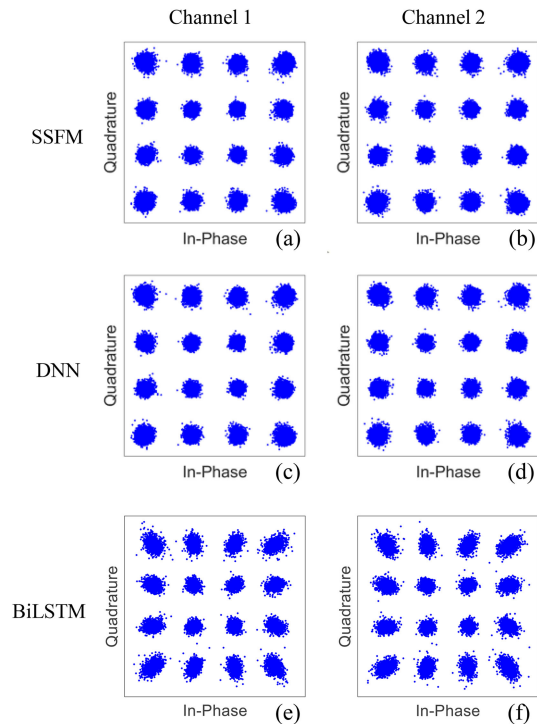


Fig. 4. Constellations after DBP compensation of SSFM-generated signals and DNN/BiLSTM-generated signals. Note that, the transmission distance is 80km and the optical launch power of the two channels is 4dBm. (a): Channel 1, (b): Channel 2 of the SSFM-based fiber channel. (c): Channel 1, (d): Channel 2 of the DNN-based fiber channel. (e): Channel 1, (f): Channel 2 of the BiLSTM-based fiber channel.

can be observed that the two waveforms are virtually identical from an overall perspective. Even after zooming in, they are still substantially identical. The  $MSE_{nor}$  can measure the distance between the two waveforms or spectra, therefore, we calculated the  $MSE_{nor}$  for quantitative analysis. As for channel 1,  $MSE_{nor}$  in the time-domain is 0.0032, which is the same as that of channel 2. The two  $MSE_{nor}$  are much less than the upper limit of 0.02. The situation is very similar for the optical spectra. Fig. 5 also shows the spectra of SSFM-generated signals and DNN-generated signals within 30GHz which is the same as the transmission baud rate. The two spectra are nearly the same in both overall view and enlarged view. And the  $MSE_{nor}$  in the frequency-domain of channel 1(2) is 0.0004(0.0005), which is quite less than 0.02 as well. The waveforms (or spectra) of SSFM-generated signals and DNN-generated signals overlap highly, meaning that the DNN has learned the time- and frequency-domain characteristics of signals. That is to say, the fiber channel has been modeled accurately by the DNN.

### B. Generalization of DNN

In this section, the generalization of the trained DNN is also studied. The ability of DNN to generalize means how good the trained neural network is when applying the learnt information from training data sets to make accurate predictions on new, previously unseen data. This is quite important for the practical

applications of fiber channel modeling. We analyzed the generalization of our model for the input with different modulation formats and wavelength schemes.

The training data set contains only 16QAM symbols. Therefore, other modulation formats, including QPSK, 8PSK, and 8QAM, are taken into account in our test data set. Fig. 6 shows the constellations after DBP compensation of SSFM-generated signals and DNN-generated signals with different modulation formats. In Fig. 6, the left ones are quite similar to the right ones, indicating visually that the DNN model has good generalization for modulation formats.

We also calculate the  $MSE_{nor}$  between the SSFM-generated signals and DNN-generated signals in both time- and frequency- domain, as is shown in Table II. In addition, different wavelength schemes are also considered in the test data set and the corresponding  $MSE_{nor}$  are shown in Table II as well. Note that, when testing the generalization for different modulation formats, we used the DNN model that is trained on the data set which is built under the case (transmission distance = 80 km, optical launch power = 4 dBm, format = ‘16QAM’,  $\lambda_1$  is 1549.32 nm and  $\lambda_2$  is 1550.12 nm). To demonstrate that the DNN model which is trained under a different case has good generalization ability as well, we apply the DNN model that is trained on the data set which is built under another case (transmission distance = 160 km, optical launch power = 4dBm, format = ‘16QAM’,  $\lambda_1$  is 1549.32 nm and  $\lambda_2$  is 1550.12 nm) to test the generalization for different wavelength schemes. As is illustrated in Table II, the  $MSE_{nor}$  are all much less than the upper limit of 0.02. Although the  $MSE_{nor}$  increase when the wavelengths become farther from 1549.32(1550.12) nm, they are still far less than 0.02. Therefore, we believe that the DNN model has a good generalization ability for the input with different modulation formats and wavelength schemes.

### C. Time Complexity

Lastly, the time complexity of the DNN-based method is analyzed. In the SSFM-based method, the running time of simulation can be very long because the FFT operation, which is the principal calculation amount of SSFM [2], is performed repeatedly. In the DNN-based method, however, no FFT operation is performed, and only multiplication between neurons is involved. Therefore, the major advantage of DNN-based method compared with SSFM-based method is its low time complexity.

We measured the computing time of SSFM-based fiber model and DNN-based fiber model. Note that, the computing time is only the time consumed by simulating the nonlinear fiber channel, excluding the time consumption of transmitter, receiver, etc. In fairness, we running the codes of SSFM-based method and DNN-based method on the same device (intel@ core i5-10210U). The absolute computing time is strongly dependent upon the hardware performance and computing resources. Therefore, it is reasonable to compare the relative computing times of the two methods. The computing time of DNN-based method with  $2^{16}$  symbols at 80 km is selected as the basic reference, meaning that all computing time data are normalized by this time value. It is acknowledged that the advantage of

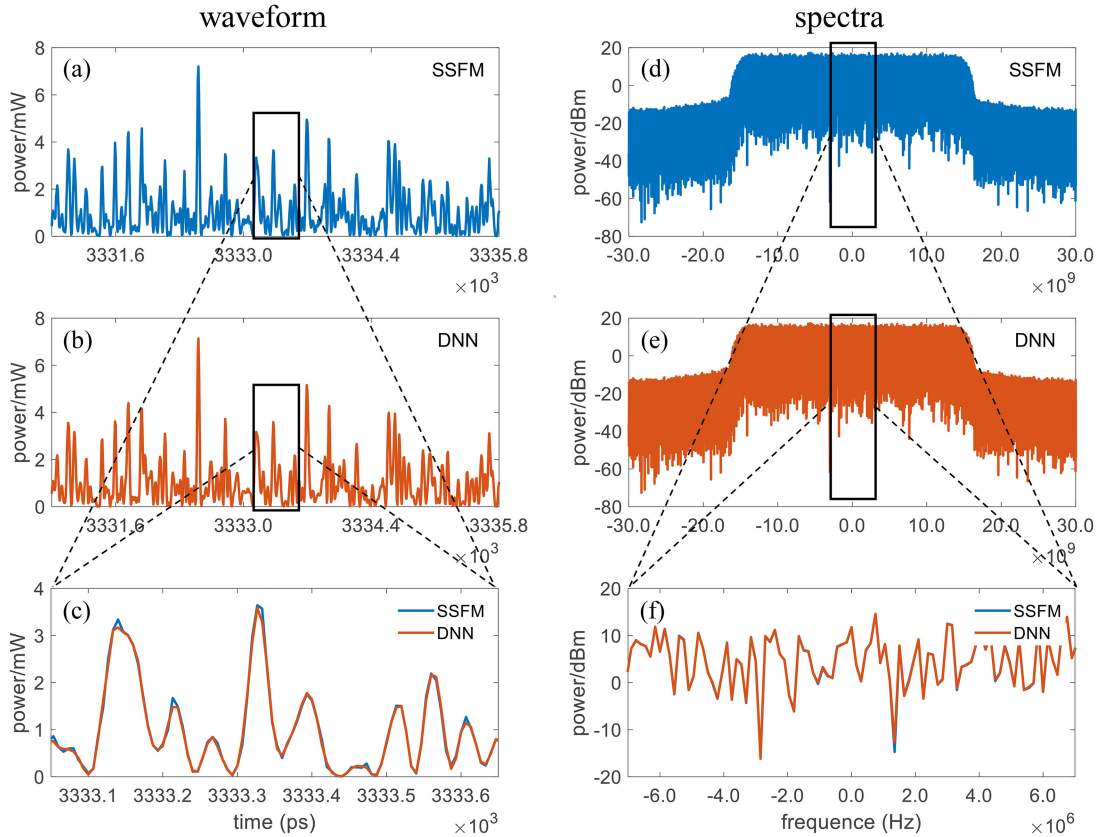


Fig. 5. Amplitudes of waveforms of channel 1 of (a) SSFM-based fiber channel and (b) DNN-based fiber channel. (c) an enlarged view of the sections in black rectangles in (a) and (b). Optical spectra of channel 1 of (d) SSFM-based fiber channel and (e) DNN-based fiber channel. (f) an enlarged view of the sections in black rectangles in (d) and (e).

TABLE II  
NORMALIZED MSE BETWEEN THE SSFM-GENERATED SIGNALS AND DNN-GENERATED SIGNALS IN TIME- AND FREQUENCY- DOMAIN WITH DIFFERENT MODULATION FORMATS AND WAVELENGTH SCHEMES

		MSE_nor_t_1	MSE_nor_t_2	MSE_nor_f_1	MSE_nor_f_2
Modulation formats	QPSK	0.0017	0.0017	0.00067	0.00070
	8PSK	0.0017	0.0017	0.00066	0.00069
	8QAM	0.0020	0.0023	0.00094	0.0013
Wavelength schemes	WS1	0.0034	0.0026	0.00042	0.00038
	WS2	0.0019	0.0023	0.00041	0.00037
	WS3	0.0016	0.0018	0.00040	0.00036
	WS4	0.0018	0.0015	0.00040	0.00036
	WS5	0.0021	0.0017	0.00041	0.00036
	WS6	0.0030	0.0038	0.00042	0.00038

Notes:

a).  $WS_i$  denotes  $i$ th wavelength scheme.

WS1:  $\lambda_1 = 1542.94$  nm,  $\lambda_2 = 1543.73$  nm. WS2:  $\lambda_1 = 1546.12$  nm,  $\lambda_2 = 1546.92$  nm.

WS3:  $\lambda_1 = 1547.72$  nm,  $\lambda_2 = 1548.51$  nm. WS4:  $\lambda_1 = 1550.92$  nm,  $\lambda_2 = 1551.72$  nm.

WS5:  $\lambda_1 = 1552.52$  nm,  $\lambda_2 = 1553.33$  nm. WS6:  $\lambda_1 = 1555.75$  nm,  $\lambda_2 = 1556.55$  nm.

b). MSE\_nor\_t(f)\_1(2) represents MSE\_nor value in time(frequency) domain of channel 1(2).

DNN-based method can be enlarged if the codes are run on the graphics processing unit. However, it is deemed sufficient to illustrate the advantage in the time complexity of DNN-based method with the relative computing time on the same device, though that device is a central processing unit.

As is shown in Fig. 7, with the transmission distance and symbol length increasing, the computing time of SSFM-based method grows rapidly. Nevertheless, the computing time of DNN-based method rises slowly with the symbol length and even keeps almost stable when the transmission distance

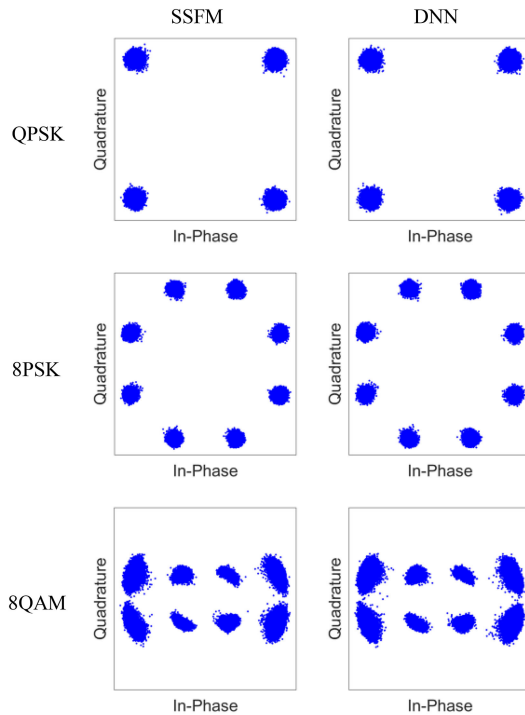


Fig. 6. Constellations after DBP compensation of SSFM-generated signals and DNN-generated signals with different modulation formats that are not included in the training data set.

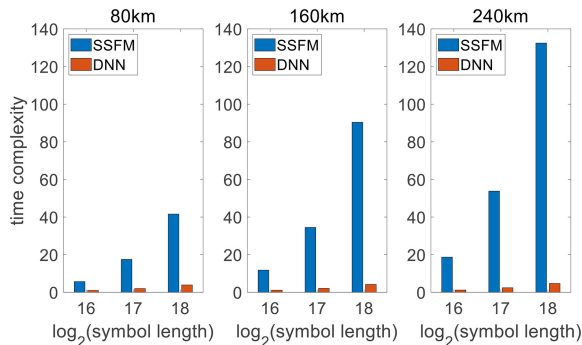


Fig. 7. Comparison of the time complexity of SSFM-based and DNN-based method at different symbol lengths and transmission distances.

increases. It is worth mentioning that, as for the best case (distance = 240 km, symbol length =  $2^{18}$ ) shown in Fig. 7, the computing time is relatively reduced by 96.5% with the DNN-based method.

## V. CONCLUSION

This work demonstrates the capability of DNN to model the nonlinear fiber channel. Based on the comprehensive analysis of the constellations after DBP compensation, optical waveforms, spectra, and the normalized MSE, it is concluded that the DNN can learn well the linear and nonlinear effects in the fiber channel, such as CD, attenuation, ASE noise, SPM, and XPM. Additionally, it is also found that DNN has a good generalization ability for modulation formats and wavelength schemes, which provides flexibilities and versatility for fiber channel modeling.

Besides, the simulation with DNN-based method can reduce the time complexity significantly (96.5%) compared to SSFM-based method. Therefore, the DNN is a good candidate to model the optical fiber channel in optical communication system simulations. This work is an initial exploration of the DNN-based method applied in the WDM system simulations. Considering that there is no limitation on the input signals of DNN proposed in this work, this DNN-based method is capable of modeling a more complicated communication system with more channels. However, a new input vector structure and parameters of DNN may be required for the reason that the inter-channel interferences become more complex. Moreover, since DNN can be viewed as a universal approximator for both linear and nonlinear functions [24], it has the latent capacity theoretically to approximate the channel transfer function with a longer transmission distance after being trained on the corresponding data set which is composed of the channel input and output signals. Thus, this DNN-based method has the potential to be extended to model a more complicated optical communication system that takes account of polarization effect or more wavelengths or has a longer transmission distance.

## ACKNOWLEDGMENT

Rui Jiang would like to thank Hsutung King of Peking University and Yuan Young of Beihang University for their help in learning artificial intelligence. Rui Jiang would also like to thank Yang Cao of Beijing Foreign Studies University for her help in refining the language of this article.

## REFERENCES

- [1] D. Wang *et al.*, "Data-driven optical fiber channel modeling: A deep learning approach," *J. Lightw. Technol.*, vol. 38, no. 17, pp. 4730–4743, 2020, doi: [10.1109/jlt.2020.2993271](https://doi.org/10.1109/jlt.2020.2993271).
- [2] H. Yang *et al.*, "Fast and accurate optical fiber channel modeling using generative adversarial network," *J. Lightw. Technol.*, vol. 39, no. 5, pp. 1322–1333, 2021, doi: [10.1109/jlt.2020.3037905](https://doi.org/10.1109/jlt.2020.3037905).
- [3] G. P. Agrawal, *Fiber-Optic Communication Systems*, 4th ed., Hoboken, NJ, USA: Wiley, 2010.
- [4] O. V. Sinkin, R. Holzlohner, J. Zweck, and C. R. Menyuk, "Optimization of the split-step Fourier method in modeling optical-fiber communications systems," (in English), *J. Lightw. Technol.*, vol. 21, no. 1, pp. 61–68, Jan. 2003, doi: [10.1109/jlt.2003.808628](https://doi.org/10.1109/jlt.2003.808628).
- [5] G. P. Agrawal, *Nonlinear Fiber Optics*, 3rd ed., Cambridge, MA, USA: Academic Press, 2001.
- [6] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: [10.1038/nature14539](https://doi.org/10.1038/nature14539).
- [7] F. N. Khan, Q. Fan, C. Lu, and A. P. T. Lau, "An optical communication's perspective on machine learning and its applications," *J. Lightw. Technol.*, vol. 37, no. 2, pp. 493–516, Jan. 2019, doi: [10.1109/jlt.2019.2897313](https://doi.org/10.1109/jlt.2019.2897313).
- [8] H. Z. Lun *et al.*, "GAN based soft failure detection and identification for long-haul coherent transmission systems," (in English), in *Proc. Opt. Fiber Commun. Conf. Expo.*, 2021, pp. 1–3, doi: [10.1364/OFC.2021.Th4J.2](https://doi.org/10.1364/OFC.2021.Th4J.2).
- [9] P. Safari, B. Shariati, G. Bergk, and J. K. Fischer, "Deep convolutional neural network for network-wide QoT estimation," (in English), in *Proc. Opt. Fiber Commun. Conf. Expo.*, 2021, doi: [10.1364/OFC.2021.Th4J.3](https://doi.org/10.1364/OFC.2021.Th4J.3).
- [10] S. Fujisawa *et al.*, "Nonlinear impairment compensation using neural networks," (in English), in *Proc. Opt. Fiber Commun. Conf. Expo.*, 2021, doi: [10.1364/OFC.2021.M5F.1](https://doi.org/10.1364/OFC.2021.M5F.1).
- [11] M. Schaedler, F. Pittala, S. Calabro, G. Bocherer, C. Bluemm, and S. Pachnicke, "Recurrent neural network soft demapping for mitigation of fiber nonlinearities and ISI," (in English), in *Proc. Opt. Fiber Commun. Conf. Expo.*, 2021, doi: [10.1364/OFC.2021.M5F.4](https://doi.org/10.1364/OFC.2021.M5F.4).
- [12] P. J. Freire *et al.*, "Experimental study of deep neural network equalizers performance in optical links," (in English), in *Proc. Opt. Fiber Commun. Conf. Expo.*, 2021, pp. 1–3, doi: [10.1364/OFC.2021.M3H.2](https://doi.org/10.1364/OFC.2021.M3H.2).

- [13] X. A. Huang, D. X. Zhang, X. F. Hu, C. H. Ye, and K. B. Zhang, "Recurrent neural network based equalizer with embedded parallelization for 100 Gbps/lambda PON," (in English), in *Proc. Opt. Fiber Commun. Conf. Expo.*, 2021, pp. 1–3, doi: [10.1364/OFC.2021.M3G.2](https://doi.org/10.1364/OFC.2021.M3G.2).
- [14] Z. Zhai *et al.*, "An interpretable mapping from a communication system to a neural network for optimal transceiver-joint equalization," *J. Lightw. Technol.*, vol. 39, no. 17, pp. 5449–5458, 2021, doi: [10.1109/jlt.2021.3086301](https://doi.org/10.1109/jlt.2021.3086301).
- [15] H. M. Nguyen, S. C. Huang, C. C. Wei, C. Y. Chuang, and J. Chen, ">55-Gbps and 30-dB loss budget LR-OFDM PON downstream enabled by ANN-based predistortion," (in English), in *Proc. Opt. Fiber Commun. Conf. Expo.*, 2021, pp. 1–3, doi: [10.1364/OFC.2021.M3G.4](https://doi.org/10.1364/OFC.2021.M3G.4).
- [16] Q. Zhou *et al.*, "Digital predistortion enhancement by convolutional neural network for probabilistic shaped discrete multi-tone signal transmission in passive optical network," (in English), in *Proc. Opt. Fiber Commun. Conf. Expo.*, 2021, pp. 1–3, doi: [10.1364/OFC.2021.M3G.3](https://doi.org/10.1364/OFC.2021.M3G.3).
- [17] B. W. Bao *et al.*, "Node-oriented traffic prediction and scheduling based on graph convolutional network in metro optical networks," (in English), in *Proc. Opt. Fiber Commun. Conf. Expo.*, 2021, doi: [10.1364/OFC.2021.F2G.2](https://doi.org/10.1364/OFC.2021.F2G.2).
- [18] D. Andreoletti, C. Rottondi, A. Bianco, and S. Giordano, "A machine learning framework for scalable routing and wavelength assignment in large optical networks," (in English), in *Proc. Opt. Fiber Commun. Conf. Expo.*, 2021, doi: [10.1364/OFC.2021.F2G.3](https://doi.org/10.1364/OFC.2021.F2G.3).
- [19] Y. C. Song *et al.*, "Physical information-embedded deep learning for forward prediction and inverse design of nanophotonic devices," *J. Lightw. Technol.*, vol. 39, no. 20, pp. 6498–6508, Oct. 2021, doi: [10.1109/jlt.2021.3103758](https://doi.org/10.1109/jlt.2021.3103758).
- [20] Y. C. Song, D. S. Wang, H. Ye, J. Qin, and M. Zhang, "Wavelength controllable forward prediction and inverse design of nanophotonic devices using deep learning," in *Proc. Eur. Conf. Opt. Commun.*, Electr. Network, Dec. 6–10 2020, pp. 1–4, doi: [10.1109/ecoc48923.2020.9333397](https://doi.org/10.1109/ecoc48923.2020.9333397).
- [21] X. Jiang, D. Wang, Q. Fan, M. Zhang, C. Lu, and A. P. T. Lau, "Physics-informed neural network for nonlinear dynamics in fiber optics," 2021, *arXiv:2109.00526*.
- [22] X. T. Jiang *et al.*, "Solving the nonlinear schrodinger equation in optical fibers using Physics-informed neural network," in *Proc. Opt. Fiber Commun. Conf. Exhib.*, Electr. Network, Jun. 6–11 2021, doi: [10.1364/OFC.2021.M3H.8](https://doi.org/10.1364/OFC.2021.M3H.8).
- [23] Y. Zang *et al.*, "Principle-driven fiber transmission model based on PINN neural network," *J. Lightw. Technol.*, vol. 40, no. 2, pp. 404–414, Jan. 2022, doi: [10.1109/jlt.2021.3139377](https://doi.org/10.1109/jlt.2021.3139377).
- [24] K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," (in English), *Neural Netw.*, vol. 2, no. 5, pp. 359–366, 1989, doi: [10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8).
- [25] L. Mescheder, A. Geiger, and S. Nowozin, "Which training methods for GANs do actually converge?," 2018, *arXiv:1801.04406*.
- [26] M. Secondini and E. Forestieri, "On XPM mitigation in WDM fiber-optic systems," *IEEE Photon. Technol. Lett.*, vol. 26, no. 22, pp. 2252–2255, Nov. 2014, doi: [10.1109/lpt.2014.2353217](https://doi.org/10.1109/lpt.2014.2353217).
- [27] C. Li, F. Zhang, Y. Zhu, M. Jiang, Z. Chen, and C. Yang, "Fiber non-linearity mitigation in single carrier 400 G and 800 G Nyquist-WDM systems," *J. Lightw. Technol.*, vol. 36, no. 17, pp. 3707–3715, Sep. 2018, doi: [10.1109/jlt.2018.2844467](https://doi.org/10.1109/jlt.2018.2844467).
- [28] M. Qiu, Q. Zhuge, M. Chagnon, F. Zhang, and D. V. Plant, "Laser phase noise effects and joint carrier phase recovery in coherent optical transmissions with digital subcarrier multiplexing," *IEEE Photon. J.*, vol. 9, no. 1, Feb. 2017, Art. no. 7901013, doi: [10.1109/jphot.2016.2647221](https://doi.org/10.1109/jphot.2016.2647221).
- [29] K. M. He, X. Y. Zhang, S. Q. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on imagenet classification," in *Proc. IEEE Int. Conf. Comput. Vis.*, Santiago, CHILE, Dec. 2015, pp. 1026–1034, doi: [10.1109/iccv.2015.123](https://doi.org/10.1109/iccv.2015.123).