Neural Operator-based Fiber Channel Modeling for WDM Optical Transmission System

Ximeng Zhang
State Key Laboratory of Information Photonics and Optical Communications
Beijing University of Posts and Telecommunications (BUPT)
Beijing, China
zhangxm@bupt.edu.cn

Danshi Wang*
State Key Laboratory of Information Photonics and Optical Communications
Beijing University of Posts and Telecommunications (BUPT)
Beijing, China
danshi_wang@bupt.edu.cn

Yuchen Song
State Key Laboratory of Information Photonics and Optical Communications
Beijing University of Posts and Telecommunications (BUPT)
Beijing, China
songyc@bupt.edu.cn

Xiaotian Jiang
State Key Laboratory of Information Photonics and Optical Communications
Beijing University of Posts and Telecommunications (BUPT)
Beijing, China
jxt@bupt.edu.cn

Jin Li
State Key Laboratory of Information Photonics and Optical Communications
Beijing University of Posts and Telecommunications (BUPT)
Beijing, China
lijin930629@163.com

Min Zhang
State Key Laboratory of Information Photonics and Optical Communications
Beijing University of Posts and Telecommunications (BUPT)
Beijing, China
mzhang@bupt.edu.cn

***Abstract*—Deep operator network (DeepONet) is introduced to solve NLSE for multi-channel transmission systems modeling in multiple formats. Results show DeepONet is accurate, easy-training and more flexible than pure data-driven models, and much faster than SSFM.**

***Keywords*—DeepONet, optical fiber, WDM system**

I. Introduction

Modeling optical fiber in long-haul wavelength division multiplexing (WDM) systems is essential to simulate, evaluate and optimize the performance of transmission systems. Generally, both linear and nonlinear effects in optical fiber are characterized by the nonlinear Schrodinger equation (NLSE). By solving NLSE, abundant information including time domain waveforms, spectra, and generalized signal-to-noise ratio (GSNR) can be obtained. Split-step Fourier method (SSFM) is widely used to calculate the approximate solutions of NLSE by handling nonlinear and linear effects separately in a single step. However, the calculation of SSFM strongly depends on optical fiber parameters and is always time-consuming to achieve accurate numerical solutions which need a small step size. Since the computational complexity of SSFM is close to the square of the number of steps, it becomes unacceptable in long-haul transmission with multiple channels. Compared with SSFM, machine learning-based models are time-saving and have been widely used in recent years. The data-driven neural network (NN) based models were studied and applied in different transmission systems [1-5], which requires more powerful ability in flexibility and interpretability.

Recently, techniques of physics-informed machine learning attract lots of attention from various areas, where physics-informed neural network (PINN) [6] is one of the most famous algorithms that incorporate the physical knowledge into loss function to solve partial differential equations (PDE) directly, and is also introduced into optical fiber communications [7-10]. However, its applications are typically limited to simple original conditions, such as short pulse or low-order format. To address the issue, the deep operator network (DeepONet) is newly developed to solve PDE in data and knowledge-hybrid-driven manner [11], which has special architecture called branch net and trunk net. The structure of DeepONet can be treated as an inverse Fourier transform (IFT) and learns input-output pairs from physical meaning instead of data only. If applying DeepONet to signal prediction, it is expected that the signal waveforms at arbitrary locations can be obtained and transmission length is as flexible as SSFM, but with lower computation complexity. Stacked branch nets allow various combinations of channels while trunk net allows solutions with higher sample rates than input, which is suitable for WDM systems.

In this paper, we utilize the DeepONet-based optical fiber model to solve the Manakov equation for WDM systems over 800km involving three modulation formats. Both spectra and time domain waveforms are used as labels, which effectively increases generalization ability and reduces the requirement of the dataset scale. Furthermore, predictions of received signals with higher sample rates after any transmission length are obtained owing to the special architecture of DeepONet. Accurate time domain waveforms and constellations are obtained at any span, and the mean square error (MSE) is lower than 0.0012 at any transmission distance, which proves the effectiveness of the model. Additionally, compared with the bidirectional long-short step memory (BiLSTM) based pure data-driven method, DeepONet has lower training costs and achieves better performance.

II. Principle

The architecture of the stacked-DeepONet is shown in Fig. 1(b), which consists of several branch nets and a trunk net whose output layer has activation functions. The branch nets encode related coefficients or initial functions u of PDE to be solved, while the trunk net encodes the location y to evaluate the output function [12]. By multiplexing and adding the outputs of branch nets and trunk net as (1), the value of the solutions f at the specified location y is obtained as the final output result. For each input function u, same number of evaluations at the same scattered sensors are required, while we do not enforce any constraints on the number or locations for the evaluation y at output function.

$$f(y_m) = \sum_{j=1}^{P} S_j^{branch}(u) \cdot \sigma_{acti}(k_j^{trunk}(y_m)) \tag{1.1}$$

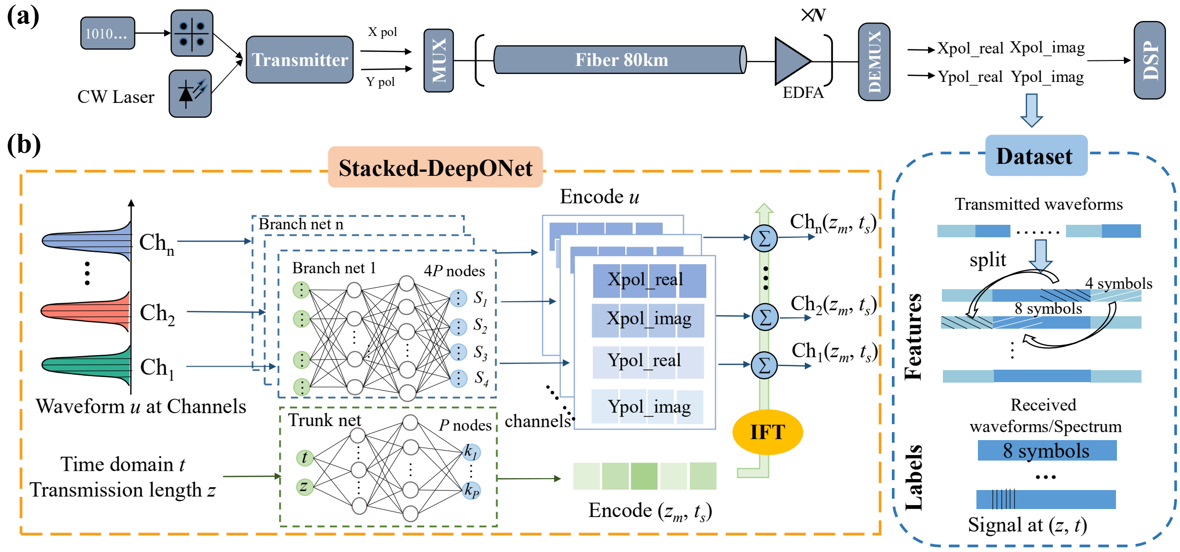

Fig.1. (a) Schematic of 5-channel WDM systems. (b) The architecture of dataset and the stacked-DeepONet model which has one trunk network and n stacked branch networks.

In our study, we apply stacked-DeepONet model for fiber channel modeling and train it to learn the Manakov equation. The waveforms of transmitted signals at each channel serve as input of branch nets. The space-time auxiliary coordinates (z, t), representing the transmission position in longitudinal fiber and sample position at time-domain waveform are fed into the trunk net. In this case, DeepONet has ability to predict signals after any transmission distance and at any time-domain locations. Overall, the branch nets achieve mapping time domain waveforms at the transmitter to an unknown function space, and the outputs of the trunk net turn that space to received waveforms. If we adopt this unknown function space as spectral space and choose exponential form functions as the activation function of the output layer of trunk net, the final step of DeepONet will be seen as IFT in series form as (2). The ability of branch net to map transmitted waveforms to the frequency space of received signals can be proved by the theory mentioned in [13].

$$f(t_m)=\sum_{j=1}^{n-1}F_n\cdot e^{(2\pi i t_m f_j)} \tag{1.2}$$

Fully connected networks are chosen as our branch nets and trunk net, and each branch net calculates signals at one channel. It is known from IFT that the spectra of both real and imaginary parts of dual-polarization signals can share the same values calculated by trunk net, so the ratio of branch net nodes to trunk net nodes are set to 4:1 to avoid increasing computational complexity and training time. When modeling fiber with a fixed transmission length, we introduce spectra data into labels additionally to increases the utilization of the dataset and improves the performance of our model. It can be seen as a form of data augmentation in neural network training. In dataset, we split dual-polarization complex signals into real parts and imaginary parts, and sample 8 points per symbol. As shown in Fig. 1(b), every 8 symbols are arranged sequentially in order of X/Y polarization direction as a calculation duration. Considering the influence of signals in the time domain, 4 symbols before and after the calculation duration are inputted at the same time to ensure the accuracy of the model.

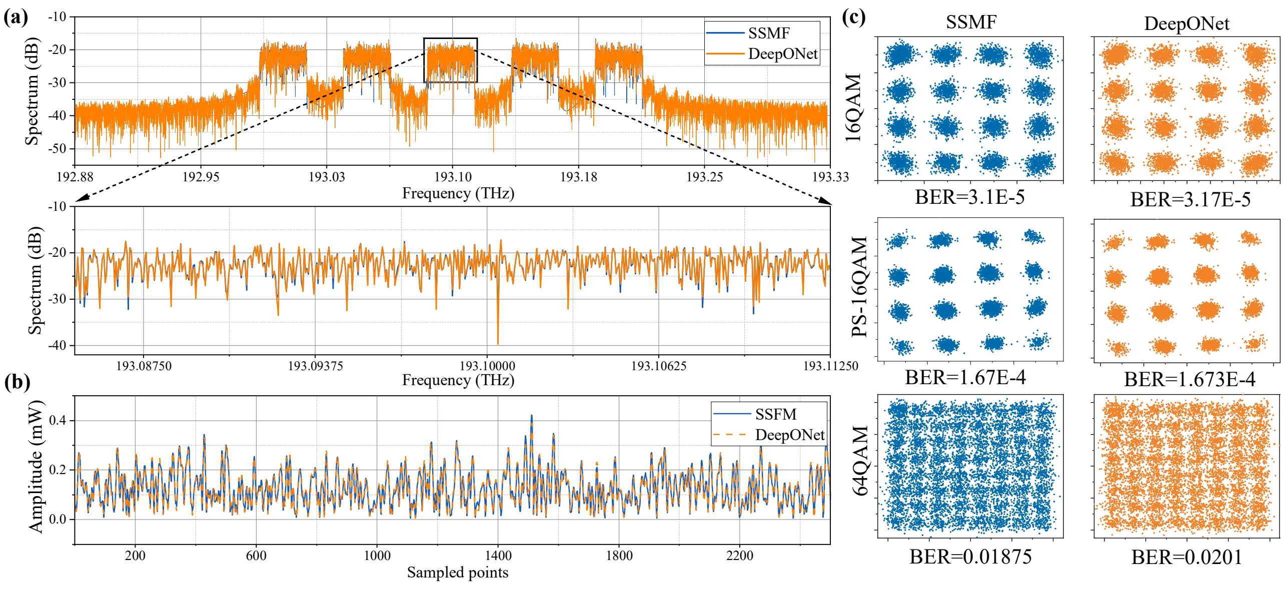

Fig.2. Signals transmitted after 800 km solved by SSFM and DeepONet. (a) The spectrum of 16QAM signals. (b)Waveforms of 16QAM signals at 193.10 THz. (c) Constellations of different modulation formats after digital signal processing (DSP).

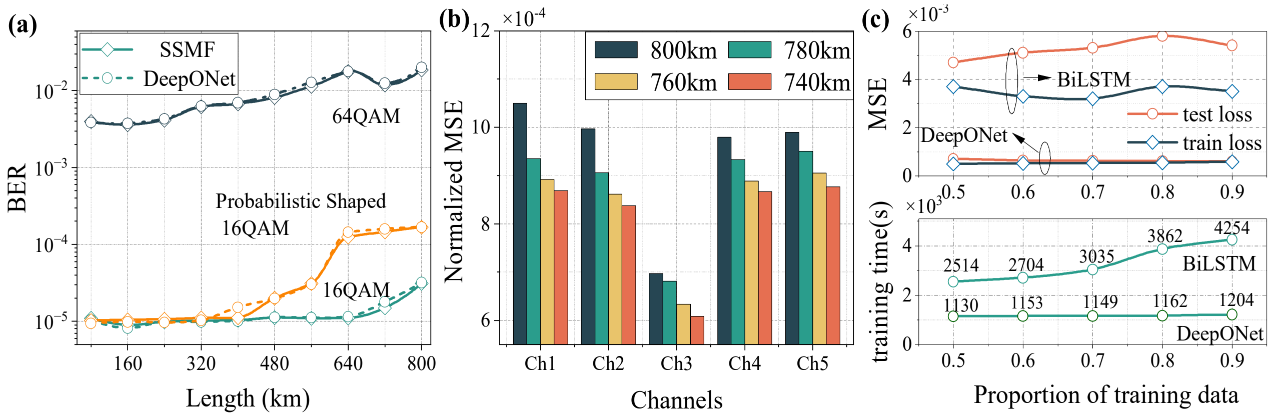

Fig.3. (a) The BER of signals calculated by SSFM and DeepONet under different modulation formats at different distances. (b) The error between outputs of SSFM and DeepONet versus distance when calculating signals at different channel. (c) Comparison of training time and the normalized MSE of BiLSTM and DeepONet, trained on Tesla P100 16GB GPU.

III. Demonstration and Results

In this research, we collect the simulation signals that are transmitted on a 10-span optical fiber in the WDM system as the dataset. As shown in Fig. 1 (a), the WDM system consists of 5 channels loading 28 GBaud dual-polarization signals whose launch power is 0 dBm and the channel spacing is 50 GHz. At the transmitter, a root-raised cosine (RRC) filter with a roll-off of 0.02 is used for pulse shaping. In each span, an erbium-doped fiber amplifier (EDFA) without amplifier spontaneous emission (ASE) noise and a standard single mode fiber (SSMF) are applied, of which dispersion, attenuation, nonlinear index and length are 16.7 ps/(nm·km), 0.2 dB/km, 1.3 /W·km and 80 km, respectively. The received waveforms are solved by SSFM with a small step size of 10 m to provide a precise dataset. To investigate the generalization performance of DeepONet, three different modulation formats are studied: standard 16QAM, probabilistic shaped (PS) 16QAM, and 64QAM signals.

The capability of the DeepONet to solve the Manakov-PMD equation of optical fiber is demonstrated in several dimensions including constellations, optical waveforms, and spectra, as shown in Fig. 2. The mean square error (MSE) of the model is 0.0012 at most. According to the spectra, the error between outputs of DeepONet and SSFM is equivalent to noise, but it is ignorable compared to signal amplitude. At the central channel of the WDM system, 3 modulation formats with the same launch power and symbol rate are included. For different modulation formats, the constellations of results of SSFM and DeepONet have few differences as shown in Fig. 2(c), which verifies the outstanding generalization ability for both channels and modulation formats.

To further analyze the accuracy of the model, the error distributions are measured in Fig. 3 (a)-(b). The BER of received signals calculated by DeepONet and SSFM are identical. The MSE at different transmission lengths in the final span has the same curves according to channels, channel 3 whose central frequency is 193.1 THz achieves optimal performance since its central frequency is as same as the reference frequency of fiber and the waveforms are less affected by the receiver. In addition, DeepONet not only performs significantly on waveform prediction but also has lower training difficulty than the BiLSTM-based data-driven model as shown in Fig. 3 (c). The dataset used to compare the performance includes 24576 samples, and are spilt into training and testing datasets at different ratios. The main computation complexity of both DeepONet and BiLSTM concentrates on parameter matrix multiplications. Thus, the special structure “gate” with more parameters in BiLSTM results in approximately 8 times more complex than that of DeepONet with the same number of nodes. In general, compared to BiLSTM, DeepONet can calculate signals with higher sample rates at any transmission distance, predict waveforms more accurately, and have lower training difficulty.

IV. Conclusion

In this paper, we proposed a stacked-DeepONet model to address the Manakov equation for dual-polarized signals with multiple modulation formats in WDM systems. The waveforms provided by DeepONet are not only as accurate as that of other data-driven models, but also available for flexible transmission distances and have higher sample rates. The special architecture of DeepONet is treated as IFT and the spectra are also introduced into labels, which increases the generalization ability and the efficiency of the dataset. Furthermore, a comparison with BiLSTM is demonstrated, which shows the stacked-DeepONet has lower computational complexity and better performance than those of other widely used models.

Acknowledgment

National Natural Science Foundation of China (No. 62171053 and 61975020).

References

[1] Yang, Hang, et al. "Fast and accurate optical fiber channel modeling using generative adversarial network." Journal of Lightwave Technology 39.5 (2020): 1322-1333.

[2] Qiu Q, Lun H, Liu X, et al. Fourier Neural Operator Based Fibre Channel Modelling for Optical Transmission[C]//European Conference and Exhibition on Optical Communication. Optica Publishing Group, 2022: We3B. 2.

[3] Gautam N, Choudhary A, Lall B. Comparative study of neural network architectures for modelling nonlinear optical ulse propagation[J]. Optical Fiber Technology, 2021, 64: 102540..

[4] Zang Y, Yu Z, Xu K, et al. Multi-span long-haul fiber transmission model based on cascaded neural networks with multi-head attention mechanism[J]. Journal of Lightwave Technology, 2022, 40(19): 6347-6358.

[5] Wang D, Song Y, Li J, et al. Data-driven optical fiber channel modeling: A deep learning approach[J]. Journal of Lightwave Technology, 2020, 38(17): 4730-4743.

[6] Karniadakis G E, Kevrekidis I G, Lu L, et al. Physics-informed machine learning[J]. Nature Reviews Physics, 2021, 3(6): 422-440.

[7] Wang D, Jiang X, Song Y, et al. Applications of physics-informed neural network for optical fiber communications[J]. IEEE Communications Magazine, 2022, 60(9): 32-37.

[8] Jiang, Xiaotian, et al. "Physics - Informed Neural Network for Nonlinear Dynamics in Fiber Optics." Laser & Photonics Reviews 16.9 (2022): 2100483.

[9] Jiang, Xiaotian, et al. "Physics-Informed Neural Network for Optical Fiber Parameter Estimation From the Nonlinear Schrödinger Equation." Journal of Lightwave Technology 40.21 (2022): 7095-7105.

[10] Song, Yuchen, et al. "Physics-Informed Neural Operator for Fast and Scalable Optical Fiber Channel Modelling in Multi-Span Transmission." 2022 European Conference on Optical Communication (ECOC). IEEE, 2022.

[11] Lu L, Jin P, Karniadakis G E. Deeponet: Learning nonlinear operators for identifying differential equations based on the universal approximation theorem of operators[J]. arXiv preprint arXiv:1910.03193, 2019.

[12] Lu L, Jin P, Pang G, et al. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators[J]. Nature machine intelligence, 2021, 3(3): 218-229.

[13] Gallant A R, White H. On learning the derivatives of an unknown mapping with multilayer feedforward networks[J]. Neural Networks, 1992, 5(1): 129-138.