Enhancing Generalization in Neural Channel Model for Optical Fiber WDM Transmission through Learned Encoding of System Parameters

Chuyan Zeng, Zekun Niu, Hang Yang, Minghui Shi, Weisheng Hu, and Lilin Yi*
State Key Lab of Advanced Communication Systems and Networks, School of Electronic Information and Electrical Engineering
Shanghai Jiao Tong University
Shanghai, China
e-mail: lilinyi@sjtu.edu.cn

Abstract: We propose a learned encoding method to enhance neural channel model generalization by integrating system parameters as side information. This approach achieves large-scale generalization, encompassing optical fiber transmission launch power and distance. © 2024 The Author(s)

1. Introduction

Neural network (NN) has drawn wide attention in field of optical fiber channel modeling, due to its strong fitting capability and high running efficiency [1-3], which greatly facilitate system optimizations and the verification of the digital signal processing (DSP) algorithms [4]. Nevertheless, the generalization ability of the neural model become increasingly significant considering their flexible application to different channel conditions.

For the long-haul wavelength-division multiplexing (WDM) optical fiber transmission, the launch power and transmission distance are two key system parameters. The capacity for flexible simulations, allowing the configuration of these parameters with arbitrary values, carries great significance for practical applications. However, few research of fiber-channel modeling has concerned with the generalization ability of these data-driven methods. Although the Feature-Decoupling Distributed (FDD) [2] and Fourier Neural Operator (FNO) [3] models have achieved certain degree of power generalization, the ability is limited, mainly within the weak nonlinear region and cannot exceed the interval of the training data. Regarding transmission distance, only FDD realizes varying distance modeling by iterative calculation, but the flexibility requires further improved since the span length is fixed.

In this paper, we propose a novel approach aimed at generalization enhancement by incorporating system parameters as side information in a learned encoding form. We elucidate that the encoding process is necessary for obtaining the implicit representations of the extra information so as to be distinguished from the signals to be modeled. Leveraging the learned encoding method, the neural channel model achieves a large-scale and flexible generalization with respect to both launch power and transmission distance, together with an overall modeling accuracy improvement. The numerical modeling error and the Q prediction error is reduced by an average of 59.2% and 0.62dB, respectively, within power range of -2-6 dBm/channel for 800km 5-channel transmission compared with the traditional generalization scheme. Furthermore, the span length can be adjusted arbitrarily from 30km to 100km. The Q prediction error is less than 0.25 dB for 1200 km,15 span transmission with the span length adjusted arbitrarily from 30km to 100km. The average numerical modeling error and Q error is reduced by 25% and 0.59 dB, respectively, compared to the original model. The proposed scheme is effective and easily to implement, which can be a good reference for improving the generalization effect of other NN techniques in optical communication research, such as the nonlinearity compensation [5].

2. The proposed learned encoding framework

The structure of the neural channel model is illustrated in Fig. 1(a). The proposed approach contains a fully-connected (FC) layer for parameters encoding which enables NN understand the deep relationships between signals and system parameters, and the feature decoupling distributed bi-directional long-short term memory ((FDD-BiLSTM) framework for channel modeling. In FDD-BiLSTM, NN only models the nonlinear features, and linear features are modeled by a principle-driven model. To realize more flexible distance generalization, the span length can be independently configured, ranging from 30 to 100km.

The whole channel modeling process is described in details below. A dual-polarization WDM transmission system is considered. The input conditions including power and the length of each span, which are measured by milliwatt and kilometers, respectively. As for the input signals, the real and imaginary parts of the two polarizations of the complex samples get separated and reshaped into a one-dimensional vector. The system parameters are fed into the FC layer for encoding and then appended to the input signal matrix. This process is denoted as learned

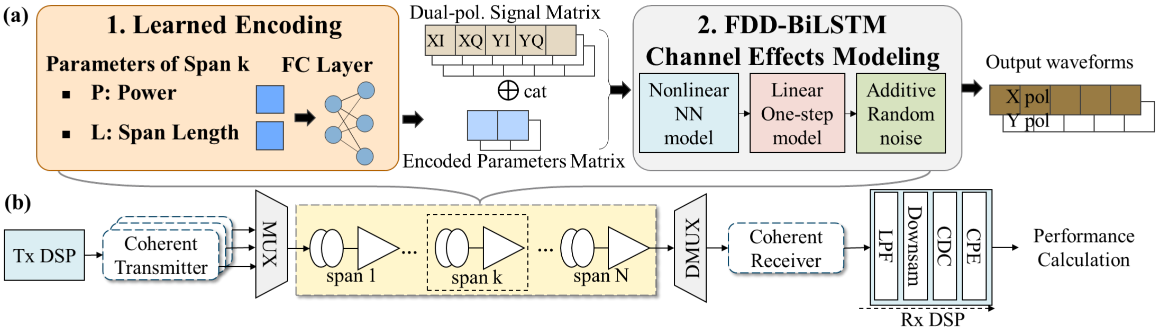


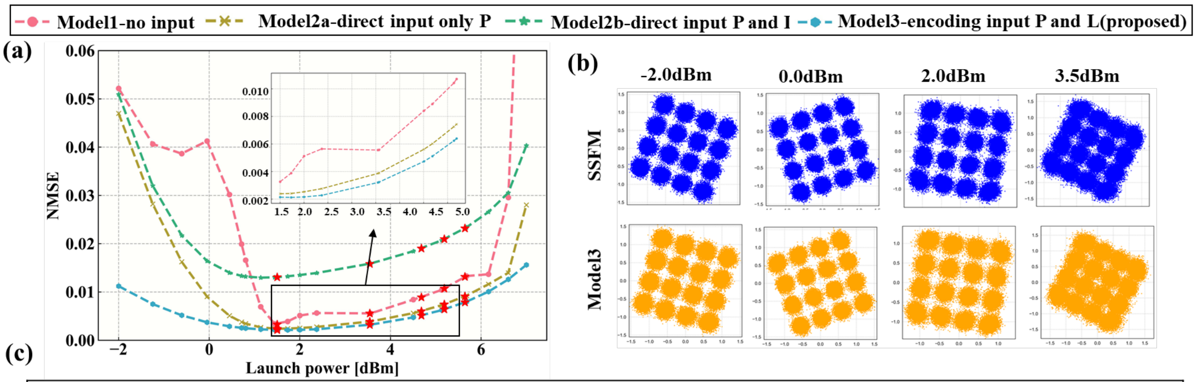
Fig. 1. The structure of the simulation WDM optic-fiber channel and the single-span channel model. (a)The single-span FDD channel model with parameters embedding encoding layer. (b) The block diagram of the optical fiber WDM transmission system.

encoding, since both the FC layer and the FDD-BiLSTM model are jointly trained and optimized. The concatenated matrix is then fed to NN for channel modeling. Finally, the model outputs the complex signal waveforms representing a single span transmission.

To realize generalization training, we create dataset encompassing various combinations of power and span length conditions. Specifically, we choose power values larger than the optimal launch power so that the nonlinear features of the signals are more apparent and easier captured by NN. The data first get shuffled and then divided into batches, each containing 500 symbols. The encoding layer and the NN for nonlinearity modeling are co-trained. The Adam algorithm is used for learning optimization with an initial learning rate of 0.002.

3. Results and Discussion

We simulate a 5-channel WDM system with 30GBaud symbol rate and 50GHz channel space based on the Split-step Fourier method (SSFM) for training data collection and evaluating model performance. The block diagram of the system is illustrated in Fig. 1(b). At the transmitter, the signals get modulated, up-sampled, pulse-shaped by a root raised-cosine filter with a roll-off factor of 0.1, and then get multiplexed. The fiber link contains multiple spans, with an erbium-doped fiber amplifier (EDFA) employed after each span. After channel transmission, the signals get demultiplexed, then passed through a low-pass filter and down-sampled to 2 samples per symbol, followed by receiver DSP to compensate for signal distortions. Lastly, performance metrics such as bit-error rates (BERs) and Q-factors are calculated.



Q factor error (dB)			
P (dBm)	Model1	Model2b	Model3
-1.3 (lower than the training range)	1.167	1.268	**0.473**
3.5 (within the training range)	1.068	0.453	**0.445**
6.1 (higher than the training range)	1.685	0.688	**0.495**

Fig. 2. The power generalization test results. (a)The NMSE performance at the launch power range of -2.0 to 6.0 dBm/channel at a distance of 800km and 80km per span. The power conditions contained in the training dataset are marked with red stars. (b) The output constellations of SSFM and Model3 under the power conditions of -2.0,2.0,3.5dBm/channel, respectively. (c) The Q prediction errors of the Model1, Model2b, Model3 after DSP at the launch power of -1.3, 3.5,6.1dBm/channel.

We evaluate the modeling performance by comparison with SSFM. Four models are considered: Model1 which does not input extra information, Model2a which directly inputs either of the two parameters and focuses on single-dimension generalization, Model2b which directly inputs two parameters without the encoding process, and Model3 which is the proposed one, with both power and length parameters input in an encoded form. As for the metrics of evaluating modeling performance, we employ normalized mean square error (NMSE) to quantify the waveform error. Typically, an NMSE value below 0.02 is considered acceptable for accurate waveform modeling [1]. And the Q Factor is also employed for assessing the performance prediction accuracy for long-haul transmission.

We first compare the generalization ability to the launch power under 800km transmission with 80km per span. Fig. 2(a) shows the NMSEs at power range of -2dBm to 5dBm/channel, which fully covers the linear and high nonlinearity power region, with the training conditions marked by red stars. It can be discovered that Model3 achieves an overall highest accuracy, with an average of 59.2% modeling errors reduction to Model1. It is noteworthy that Model2b, though with the same conditions input as Model3, has the lowest accuracy while Model2a performs relatively better, indicating that NN fails to understand affluent extra information without encoding preprocessing. Notably, only Model3 generalizes to the powers exceeding the training powers. The output constellations of Model3 and SSFM presented in Fig2(b) are highly consistent under different conditions. The Q prediction errors of the 4 models under different power conditions in Fig. 2(c) align with the NMSE results, in which the proposed Model3 is the lowest and almost unaffected by the change of power conditions, while Model2a and Model2b appear significant deterioration out of the training range.

Regarding the span length generalization, we set total length to 1200 km and span num to 15 while varying the length of each span from 30km to 100km. Fig. 3(a) shows the Q prediction errors after 1200km transmission under 5 groups of random length configurations, and Fig. 3(b) shows the NMSEs after each span under Config5 in (a). It can be discovered that Model3 still performs the best under different length configurations, with all Q Errors below 0.25 dB under random configurations, achieving an average of 0.59dB reduction in Q prediction Error compared to Model1, demonstrating its flexibility in accommodating to the change of span length.

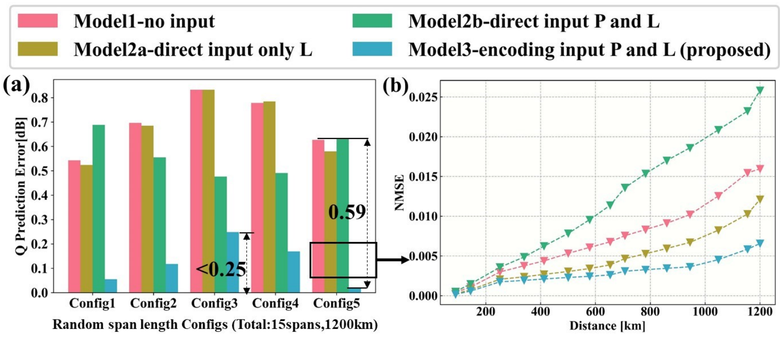


Fig. 3. The distance generalization test results. (a) Q prediction errors of the four models after 1200km transmission under 5 random span length configurations. (b) NMSE vs Distance of the three models under Config5 in (a).

Synthesizing the above results, system parameters input is demonstrated effective for enhancing the generalization ability of NN. And to achieve a large-scale and multidimensional generalization, the preprocess encoding of the affluent extra information is of great significance.

5. Conclusion

In this paper, we propose a learned encoding scheme for utilizing the system parameters to improve the generalization ability of the neural fiber channel model. The encoded input of the launch power and the transmission distance conditions enables large-scale generalization of the NN channel with respect to the two channel conditions. The proposed method can also be a valuable reference to other NN applications in optic-fiber communication research.

References

[1] D. Wang et al., "Data-driven Optical Fiber Channel Modeling," J. Lightw.Technol, vol. 38, no. 17, pp. 4730-4743(2020)
[2] H.Yang et al., "Fast and Accurate Waveform Modeling," J. Lightw.Technol, vol. 40, no. 14, pp. 4571-4580 (2022)
[3] Q. Qiu, et al., "Fourier Neural Operator Based Fibre Channel Modelling for Optical Transmission," ECOC, pp. 1-4(2022)
[4] B. Karano, et al., "End-to-End Deep Learning of Optical Fiber Communications," in J. Lightw.Technol, vol. 36, no. 20, pp. 4843-4855 (2018)
[5] Fan, Q. et al., "Advancing theoretical understanding and practical performance," Nat Commun 11, 3694 (2020).