Physics-Informed Neural Operator-based Full Wavefield Back-Propagation for Multi-span Optical Transmission

Yuchen Song, Xiaotian Jiang, Xiao Luo, Ximeng Zhang, Min Zhang, and Danshi Wang*
State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing 100876, China
Email: danshi_wang@bupt.edu.cn

Abstract: An unsupervised physics-informed neural operator-based wavefield back-propagation scheme is proposed, which enables the full wavefield information reconstruction along the link, nonlinearity compensation (0.46 dB Q-factor gain over 1StPS DBP), and fiber parameter identification.

1. Introduction

In optical fiber communication, the received signals suffer from linear and nonlinear impairments of fiber, and techniques of backward propagation enable the signal recovery and facilitate the fiber longitudinal monitoring [1]. The digital backpropagation (DBP) is the reference technique that attempts to back-propagate the received signal by solving the backward nonlinear Schrödinger equation (NLSE). However, due to the limit of complexity, only partial accumulated dispersion and nonlinearity are compensated with some coarse steps where new impairments could be induced by the DBP itself [2]. In fiber longitudinal monitoring for parameter identification, small step is required for DBP at the cost of high complexity [1]. In addition, neural networks (NNs)-based nonlinear equalizers have been proposed and shown to be effective in channel equalization [3]; however, one of their main pitfalls is the "jail window" pattern of constellations [4]. Besides, data-driven NNs treat the equalization as a data fitting problem, and the predictions are not guaranteed to satisfy the underlying physical laws, which generally has poor generalization performance on out-of-distribution predictions and extrapolation tasks.

The prior knowledge of physics is expected to empower NNs in backward modeling. In this connection, it is found that the linear and nonlinear operations of standard NNs have essentially the same formulations to the iterative linear and nonlinear operation of DBP. Leveraging this observation of physics, the filter taps can be optimized to further improve the performance of DBP, which is called the learning-DBP (L-DBP) [5]. However, only the recovered signals are obtained through supervised learning, and the detailed backward transmission process is not modeled and the whole wavefield information at any position is unavailable.

Recently, the methodology of neural operator was proposed to learn the mapping from initial conditions (ICs) to solutions for partial differential equations (PDEs). The PDEs can be used as constraints in the loss function of neural operator to steer the learning process towards finding physically consistent mappings, which constructs the physics-informed neural operator (PINO) [6]. Obtaining a solution for a new instance of the ICs requires only a forward pass of the already trained PINO, and the major computational issues incurred in classical approaches is alleviated.

In this paper, we use the backward NLSE operator to back-propagate the received signals span by span towards the transmitter side, during which both fiber linear and nonlinear effects are compensated, and the full details of wavefield back-propagation are characterized and reconstructed. The backward NLSE operator is implemented by learning the backward NLSE with no labeled data using the PINO in an unsupervised manner. In the aspect of equalization, we analyze its performance on different launch powers compared with chromatic dispersion (CD) compensation and regular DBP techniques, and a 0.46 dB Q-factor gain is observed compared to 1 steps-per-span (StPS) DBP. In order to further reduce the complexity of backward NLSE operator and demonstrate the PINO's ability in fiber parameter identification, we further investigate a backward NLSE operator with optimized hypothetical fiber parameters when multi-span step size is considered, and zero attenuation is assumed.

2. Principle of DBP and physics-informed neural operator

DBP is equivalent to passing the received signal through a fiber with parameters of the opposite sign, and the corresponding backward NLSE is solved numerically by split-step Fourier method (SSFM), which calculates the linear D and nonlinear N operation of NLSE iteratively in multiple steps.

$$\frac{\partial s}{\partial z}=(D+N)s, \qquad D=\frac{\alpha}{2}s+\frac{i\beta_2}{2}\frac{\partial^2 s}{\partial t^2}, \qquad N=-i\gamma|s|^2 s \tag{1}$$

The transmitted signal s is expressed as $s=s_I+i\cdot s_Q$, where I and Q denotes the real part (In-phase) and imaginary part (Quadrature), respectively. α, β_2, and γ denote the attenuation, group velocity dispersion, and nonlinearity coefficient. Coarse step size is usually selected in regular DBP to keep moderate complexity, which make it impossible to reconstruct the entire information of wavefield back-propagation.

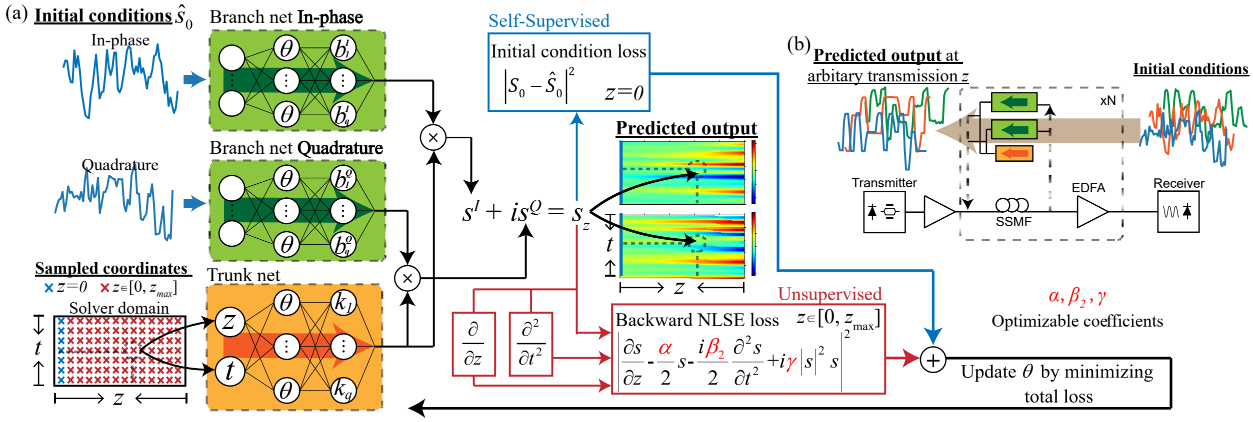

Fig. 1 (a) Schematic of the backward NLSE operator by PINO. θ denotes the parameters of NNs to be optimized. $\hat{s}_0$ denotes the initial conditions. (b) Backward NLSE operator for signal back-propagation in a span by span config.

For the structure of neural operator, we elect the DeepONet [7], which consists of two parts of networks: the branch net and the trunk net as illustrated in the Fig. 1. The trunk net samples the solver domain and get coordinates (z, t) as inputs while the branch net is further divided into the I and Q part and take respective signals of ICs $\hat{s}_0$ as input. The final signal outputs [s^I, s^Q] at coordinates (z, t) are obtained by merging two output feature embeddings between the trunk net and the corresponding brunch net by a dot product. These outputs are then connected to loss functions as shown in Fig. 1, where the backward NLSE is embedded and acts as constraints to control the outputs under the physical laws in an unsupervised way. It should be emphasized that these differential terms in the NLSE loss can be calculated efficiently using the automatic differentiation built in deep learning libraries. Furthermore, the initial condition loss ensures the PINO follow the ICs when z=0 in a self-supervised way. In conclusion, the PINO has two key advantages:

- By constraining the neural operator outputs to satisfy the backward NLSE (instead of forcing the outputs to be the labels in data-driven manner), enhanced generalization performance is guaranteed and observed [6].
- Benefiting from the trunk net, super-resolution and mesh-independent backward NLSE operator can be learned, which makes the reconstruction of full wavefield details and the identification of fiber parameters possible. The trained PINO operates much faster than SSFM under scenarios with high launch power and long sequences [8].

3. Backward NLSE operator with span-length step size

We showcase the performance of backward NLSE operator on a 10-span link with a total of 800km transmission. 16-QAM signals are transmitted at a bit rate of 56 Gbps, and the noise figure (NF) for each EDFA is 5 dB. To reconstruct the wavefield back-propagation along the entire link, all we need is the received signal (ICs here) at Rx side, and the parameters of the investigated link. The signal power is increasing in the backward transmission and is attenuated at the beginning of each span. The PINO is learned in a span by span config starting from the last span as shown in Fig. 1(b). As the input of NNs cannot be infinitely long, at each span, the long sequence of N symbols of the received signals is divided into many short sequences of M symbols for learning and testing. Overlaps in these divided short sequences are required to avoid inter-symbol interference, of which the details can be found in [8]. Only 3,000 symbols of ICs at 0 dBm launch power are used for learning.

In Fig. 2(a), we depict the distribution of wavefield back-propagation along the entire link of a 32 symbols sequence for test. The received signals at Rx side is highly distorted by the dispersion and nonlinearity, and it can be recovered to the transmitted signals at arbitrary distance with the learned PINO. Here in Fig. 2(a) the resolution of

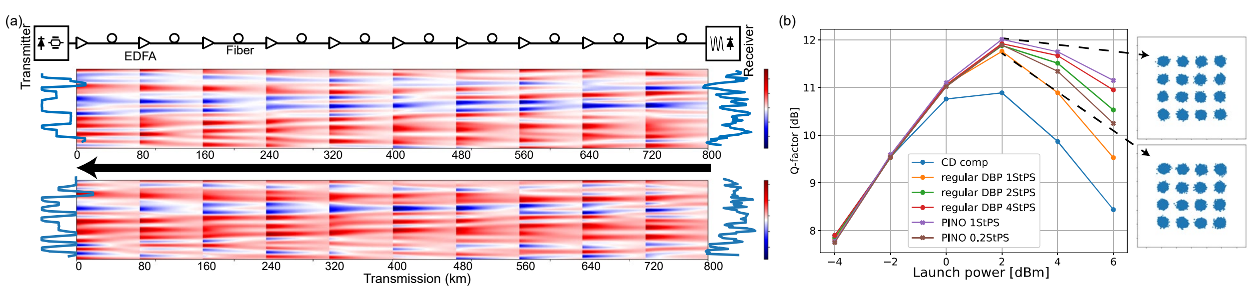

Fig. 2 (a) The distribution of wavefield back-propagation along the entire 10-span link with 32 symbols width, 0.1km resolution on transmission. (b) Q-factor vs. signal-launched power for CD compensation only, regular DBP with different StPS, and PINO based backward NLSE operator. The constellations of 1StPS PINO and 1StPS regular DBP are also presented.

transmission is 0.1km, and the PINO will have a speedup advantage of about 1,000 times over SSFM for a 2^{17} symbols sequence for 800km backward-propagation under this resolution (both under Tesla T4 GPU).

Since the essence of this approach is solving the backward NLSE using PINO, it is expected to be an alternative for DBP in the Rx side. The compensation performance of backward NLSE operator is compared with that of classical DBP and CD only compensation under different launch powers in Fig. 2(b). The Q-factor is calculated from bit error ratio (BER) through $20\log_{10}[\sqrt{2}\mathrm{erfc}^{-1}(2\mathrm{BER})]$). It can be observed that 1StPS PINO outperforms the 1StPS DBP, 2StPS DBP, and 4StPS DBP by 0.46, 0.18, and 0.1 dB, respectively, while the constellation of recovered signals by the PINO is still Gaussian-like distribution.

4. Backward NLSE operator with multi-span step size and NLSE coefficient optimization

Although the 1StPS backward NLSE operator is capable of compensating received signals, its calculation complexity is still higher than that of the classical DBP. Multi-span per step operator is investigated to further reduce its complexity. Besides, in some scenarios, the fiber parameters can be inaccurate. Benefitting from the fact that all the results are continuously obtained over the solve domain and the constraint of NLSE, the PINO is able to identify the actual fiber parameter using both signals at the Tx and Rx side.

Here, we instructively showcase the backward NLSE operator with a 5-span step size, which is equal to 0.2StPS, and the attenuation term in the NLSE is discarded. This means the received signals are back-propagated through a hypothetical fiber of 5x80km without attenuation and with modified dispersion and nonlinear index. At start, these parameters are assumed to be the default negative value of real fibers (dispersion=16ps/nm/km, nonlinearity=1.27/W/km), and the ICs (received signal) and final conditions (transmitted signal) for three different 16-symbol sequences under 3 dBm launch power are collected for the PINO to learn at z=0 and 5x80km, respectively. However, with the default assumed coefficients, the backward NLSE does not satisfy the corresponding initial and final conditions as the attenuation is zero, which makes the NLSE loss hard to decrease. After every 100 epochs for minimizing the backward NLSE loss by updating PINO parameters θ, the coefficients of dispersion and nonlinearity are also updated as shown in Fig. 3(b)(c) to minimize the NLSE loss, and the dispersion index is increased while the nonlinear index is decreased due to the discard of attenuation. To summarize, when the received signals and the NLSE are known, the transmitted signals in the solver domain can be learned; when some pairs of received and transmitted signals are known, the coefficients of NLSE can be identified.

The signal can be recovered to its previous states before 5x80km transmission as presents in Fig. 3(a). Its performance for compensation of an 800km link is also shown in Fig. 2(b). The 0.2StPS PINO is comparable to 2StPS DBP in compensation, and the calculation complexity is reduced by 5 times compared with 1StPS PINO.

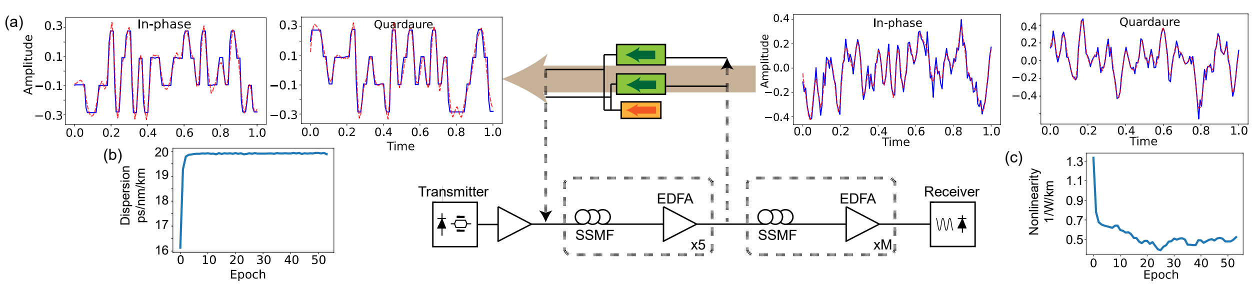

Fig. 3 (a) Backward NLSE operator for 5 spans (0.2StPS) along with 32 symbols of the received and transmitted signals, and the sampling rate is 4 samples per symbol. The optimization of (b) dispersion and (c) nonlinear index in the zero-attenuation 5-span backward NLSE.

5. Conclusion

We proposed to use unsupervised PINO to learn the backward NLSE operator and reconstructed the full wavefield along a stimulated link. In the aspect of equalization, this approach was compared with regular DBP and outperformed 2StPS DBP by 0.17 dB. This physics-informed approach is expected to empower the detailed backward modeling of fiber transmission, and the NLSE coefficient optimization demonstrates potential for fiber parameter identification from physics-data hybrid perspectives.

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

References

[1] T. Sasai et al., "Digital Longitudinal Monitoring of Optical Fiber Communication…," J. Light. Technol., vol. 40, no. 8, pp. 2390-2408, 2022.
[2] Q. Fan et al., "Advancing theoretical understanding and practical performance of signal…," Nat. Commun., vol. 11, no. 1, pp. 1-11, 2020.
[3] H. Ming et al., "Ultralow Complexity Long Short-Term Memory Network for…," J. Light. Technol., vol. 40, no. 8, pp. 2427-2434, 2022.
[4] P. J. Freire et al., "Neural Networks-Based Equalizers for Coherent…," IEEE J. Sel. Top. Quantum Electron, vol. 28, no. 4, pp. 1-23, 2022.
[5] C. Häger et al., "Nonlinear interference mitigation via deep neural networks," OFC 2018. W3A. 4.
[6] S. Wang et al., "Learning the solution operator of parametric partial differential equations…," Sci. Adv, vol. 7, no. 40, p. eabi8605, 2021.
[7] L. Lu et al., "Learning nonlinear operators via DeepONet based on the universal…," Nat. Mach. Intell, vol. 3, no. 3, pp. 218-229, 2021.
[8] Y. Song et al., "Physics-Informed Neural Operator for Fast and Scalable Optical Fiber Channel…," ECOC 2022, We5.32.