Low-complexity Fiber Nonlinearity Compensation Based On Operator Learning Over 12075 km Single Model Fiber

Xingchen He(1), Lianshan Yan*(1), Lin Jiang(1), Anlin Yi(1), Zhengyu Pu(1), Youren Yu(1), Wei Pan(1), Bin Luo(1)

(1) School of Information Science and Technology, Southwest Jiaotong University, Chengdu 611756, Sichuan, China, lyan@home.swjtu.edu.cn
(2) Peng Cheng Laboratory, 518052, Shenzhen, China

Abstract A low-complexity fiber nonlinearity compensation algorithm is proposed, and experimentally demonstrated in a 12057 km single mode fiber transmission system. The results show that it achieves similar performance compared to 5-steps per span DBP, but its computational complexity is only 5%.

©2023 The Author(s)

Introduction

With the increase in transmission rates and device bandwidth, signals are becoming more sensitive to fiber nonlinearity distortion. The fiber nonlinearity compensation algorithms will be a key factor affecting further increases in capacity for the next generation of optical transmission systems. The most commonly used algorithm for compensating fiber nonlinearity is digital back-propagation (DBP). It is widely recognized that the multi-step per span (StpSp) DBP could pose challenges due to the limited capabilities of digital signal processing (DSP) in practical systems. In contrast, if DBP is executed under limited less StpSp, a substantial accumulation of nonlinearity residue with increasing transmission distance will reduce system performance.

Recently, a new enhanced version of DBP, named Learning DBP (LDBP) [1, 2], has been proposed by leveraging machine learning (ML). LDBP improves the performance of DBP by globally optimizing the parameters of the Finite Impulse Response (FIR) filters and reduces the computational complexity by trimming the length of FIR filters, thereby improving the compensation performance of fiber nonlinearity compensation algorithms. However, LDBP is a physics-model driven ML algorithm, and akin to DBP, its complexity augments as transmission distance rises (this is mainly manifested in the changes of the number and length of FIR filters). Therefore, it will have significant practical significance for exploring a low complexity fiber nonlinearity compensation algorithm in ultra-long haul submarine systems (≥10000km).

In recent years, using data-driven techniques to approximate partial differential equations (PDEs) is becoming increasingly prevalent as an influential paradigm. The majority of existing findings relate specifically to forward problems associated with PDEs. One popular method in this field is known as operator learning [3,4], which aims to acquire knowledge of the underlying forward solution operator for the PDE from data. Considering the broad achievements of operator learning in the domain of forward problems for PDEs, it is logical to explore their applicability in acquiring knowledge about the solutions of the corresponding inverse problems through data.

In this paper, we have extended a low-complexity operator learning framework, Koopman Neural Operator (KNO) [5, 6], and use it to solve the inverse problem of the Manakov equation, which can also be interpreted as solving the forward problem of DBP. When used in the DBP, the original KNO is limited to learn an operator that solely maps the signal at the receiver back to the transmitter for fixed link parameters. Once any of these link parameters such as dispersion (channel), transmission distance changes, the network needs to be retrained. To overcome this limit, we propose a low-complexity fiber nonlinearity compensation algorithm based on a multi-input KNO (MKNO) framework, which is experimentally verified in a 12057 km single mode fiber transmission system. The experimental results show that the compensation performance is comparable to that of the DBP with optimal step size, while the computational complexity is only around 5% of the latter.

Principle

The standard DBP algorithm seeks to reconstruct the transmitted signal $E(t, 0)$ from the received field $E(t, z)$, which can be recursively expressed as [7]

$$ E(t, 0) = \prod_{n=1}^{N_{\text{step}}} \exp\left(-\int_{z}^{\pi} \tilde{d}(\xi) d\xi\right) \exp\left(-\int_{z}^{\pi} \tilde{N}(\xi) d\xi\right) E(t, z) $$

(1)

Where $\tilde{d} = -a/2 - (j\beta_2)z^2 / (\alpha^2)$, $\tilde{N}(\xi) = j\gamma/9.E(t, z)^2)$, $\beta_2$ is the group-velocity dispersion coefficient, $\gamma$ is the nonlinear coefficient, $N_{\text{DBP}}$ denotes the
number of DBP steps, \( z_n \) and \( \Delta z_n = z_n - z_{n-1}, n = 1, \ldots, N_{DBP} \) are the propagation section and so-called DBP step size at iteration \( n \).

The objective of the neural operator is to construct a parametric map \( G_\theta: E(t,z) \rightarrow E(t,0) \) to approximates DBP algorithm, i.e. \( G_\theta(E(t,z)) \approx DBP(E(t,z)) \). Where \( \theta \) is the parameters to be optimized in the networks. The original KNO provides an explicit idea for solving \( G_\theta \), which define by

\[
G_\theta := Q \circ (FT^{-1}(\sigma K^1 \cdot \cdots \cdot \sigma K^h) FT + W) \circ P.
\]

Where \( \circ \) denotes operator composition, \( P: \mathbb{R}^L \rightarrow \mathbb{R}^h \) and \( Q: \mathbb{R} \rightarrow \mathbb{R} \) are full connection layers, \( h \) and \( L \) denote the dimension and number of hidden layers, respectively. \( W \) is one dimensional convolution. \( K \) is an integral Hilbert-Schmidt operator, whose kernel is learned from paired observations in frequency domain. \( \sigma \) is a non-linear activation function. Given that the weights of the KNO should be link parameters-dependent (taking dispersion coefficient and transmission distance as examples), we aim to learn the operator

\[
(E(t,z), \beta, \tilde{z}) \rightarrow \phi(E(t), \tilde{z})
\]

for any \( \beta, \tilde{z} \) and initial condition \( E(t,z) \). Here, the input link parameters \( \beta, \tilde{z} \) can be converted into multidimensional representations \( \phi(\beta, \tilde{z}) \) and \( \phi(\beta_2, \tilde{z}) \in \mathbb{R}^h \) for a hidden dimension \( h \) using two-layer fully connected networks that parameterized by sinusoidal embedding [8]. \( \phi(\beta, \tilde{z}) \) and \( \phi(\beta_2, \tilde{z}) \) incorporate dispersion and distance information into weights \( W \) and \( K \) through point-wise product. Fig.1(a) illustrates the MiKNO framework.

**Experimental Setup**

We adopted the experimental setup described in [9] to verify the effectiveness of MiKNO. As shown in Fig.1(b), the C+L band transoceanic transmission system consists of 257 optical channels, where the C-band ASE noise source (1529.8 nm ~ 1567.0 nm) is divided into 129 spurious optical channels with 0.3 nm interval by C-band wavelength selective switch (WSS), and the L-band ASE noise source (1570.6 nm ~ 1608.7 nm) is divided into 128 optical channels by L-band WSS. The test channel is generated using a tunable external-cavity laser with a linewidth of less than 10 kHz. The modulation process generates polarization-division multiplexing (PDM) 32GBaud PS-16QAM optical signals with oversample factor of 2 as well as varying levels of information entropy (H=3.0, 3.2, 3.5) and shapes them using a square root raised cosine (SRRC) with a roll-off factor of 0.01. The circulating loop test-bed consists of 10 spans with ~75km single mode fiber and an erbium-doped fiber amplifier (EDFA) each. Total C+L band output power launched into the 12057km transmission link is ~18 dBm. The test channel is filtered by C- or L-band WSS, detected by an integrated coherent receiver, and sampled by a real-time digital oscilloscope with 33 GHz electrical bandwidth and 80 GSa/s sampling rate. Finally, the transmission impairments of the digital signals are compensated in off-line DSP module.

The proposed MiKNO contains \( L = 4 \) hidden layers with latent dimension \( h = 32 \) and the input length per batch contains 15000 data points. We select GELU activation function and Adam for learning optimization. The training dataset is comprised solely of PS-16QAM data with H=3.5, And the testing dataset contains 3 different information entropy with H=3.5, H=3.2, and H=3.0, respectively.

**Results and discussions**

Corresponding waveforms of in-phase (I) and quadrature (Q) for DBP and MiKNO outputs are compared in Fig.2, where high waveform overlap rate demonstrates that MiKNO can provide a close approximation to DBP even over a transmission distance of 12,057 km. Furthermore, under the three different entropy values, the
NMSE between the MiKNO prediction and the DBP solution is essentially consistent, all of which are at the 8E-4 level, demonstrating the strong universality of MiKNO with respect to modulation formats.

Subsequently, we further compared the Q factors of the MiKNO, LDBP, and DBP algorithms within the range of 1530-1609 nm. During the testing process, we selected a channel every 10 nm for testing. After weighing between complexity and performance, the $S_{tpSp}$ of LDBP and standard DBP is set to 0.5 and 5 respectively. As shown in Fig.3 (a), MiKNO underperform LDBP by an average of 0.057 dB and underperform DBP by 0.02 dB respectively. Then, we further explore the complexity of these three algorithms. As shown in Fig.3 (b), the MiKNO only requires about 5,800 real number multiplications (RMs) per symbol, while DBP and LDBP require around 110,000 and 51,000 RMs per symbol. This implies that MiKNO can achieve a similar performance to 5-StpSp DBP (above 0.017 dB) and 1-StpSp LDBP (above 0.057 dB) with only 5% and 11.6% of their computational complexity, respectively. It is worth noting that MiKNO can take full advantage of parallel computing to further accelerate computation time. For example, its computation time on a GPU is only about 3 ms, while LDBP and DBP require 1.6s and 0.54s for a signal with 800,000 sampling points, respectively.

Conclusion

Based on data-driven learning, a low-complexity nonlinear compensation algorithm, MiKNO, has been proposed and experimentally validated over 12057km transmission link. Within an acceptable range of performance degradation, MiKNO exhibits computational complexity of only 5% compared to 5-StpSp DBP and 11.6% compared to LDBP. Additionally, its runtime on GPU is merely 0.1% of that of 5-StpSp DBP and 0.3% of LDBP. Furthermore, MiKNO can satisfy nonlinear compensation under different link parameters with only once training without requiring additional training or fine-tuning, which made a strong supplement to the digital signal processing library.

Acknowledgements

This research is supported by the National Key Research and Development Program of China (2021YFB2206303), and National Natural Science Foundation of China (NSFC) (62005228, 61860206006).
References


