800-Gbit/s/carrier TPS-64QAM WDM Coherent Transmission over 2,400 km Utilizing Low-complexity Separated Pruning DNN-based Nonlinear Equalization

Bohan Sang(1), Miao Kong(1), Yuxuan Tan(1), Kaihui Wang(1), Li Zhao(1), Wen Zhou(1), Ze Dong(2), Bo Liu(3), Xiangjun Xin(2), Weizhang Chen(4), Bing Ye(4) and Jianjun Yu(1*)

(1) Fudan University, Shanghai 200433, China, *jianjun@fudan.edu.cn
(2) Beijing Institute of Technology, Beijing 100081, China
(3) Nanjing University of Information Science & Technology, Nanjing, 210044, China
(4) ZTE Corp, Beijing,100191, China

Abstract *We experimentally demonstrated 800-Gbit/s/carrier WDM coherent transmission over 2,400-km based on 100-GBd truncated PS-64QAM utilizing 75%-sparsity pruning DNN-based nonlinear equalization. Results show that our pruning DNN-NLE with 24%-lower complexity outperforms Volterra NLE by 20% reach improvement. ©2022 The Author(s)*

Introduction

The demand for larger capacity, higher rate and longer reach of optical fiber communication networks has become more and more urgent in the last few years. Long-haul coherent transmission above 800 Gb/s/carrier will be a prospective target for next generation optical fiber communication[1]. However, the nonlinear impairment has become a key limiting factor of transmission distance. With the development of machine learning, nonlinear equalization (NLE) based on deep neural network (DNN) has been applied to nonlinear compensation for fiber optic communication systems[2]–[4], and pruning is introduced to save complexity[5]–[7]. In our previously published work, we realized 5×800-Gbit/s WDM coherent transmission over 2,000 km based on truncated probabilistic shaped (TPS) 64QAM by utilizing 2-order MIMO Volterra NLE[1].

In this paper, we innovatively proposed the first demonstration of separated pruning DNN-based NLE (DNN-NLE) to replace MIMO Volterra NLE. Furthermore, we used pruning technique to reduce the computational complexity of DNN-NLE and finally achieved separated pruning DNN-NLE with 24%-lower complexity than VNLE. The transmission distance of 800-Gbit/s/carrier signal is extended by 20% at the same time.

Experimental setup

Our experimental setups for five-channel WDM TPS-64QAM signal at 100-Gbaud are illustrated in Fig. 1. At the Tx side, the light source at 1553.125-nm from an ECL with 100 kHz linewidth is fed into the I/Q modulator with 30-GHz 3-dB bandwidth, which is driven by the TPS-64QAM signal from two high-speed DACs with 35-GHz bandwidth and 100-GSa/s sampling rate. The entropy of TPS-64QAM is 5 bit/symbol. We have five sub-channels with 125-GHz spacing, including one measured channel (Ch. 3) and four adjacent loading channels (Ch. 1, 2, 4 and 5). After polarization multiplexing, five sub-channels are combined by an optical coupler. Then, the WDM signals are launched into a cyclic ULAF loop, consisting of four spans of 100-km ULAF amplified by a backward-pumped Raman amplifier. The optical spectra of WDM signals before and after 2,000-km ULA fiber transmission at 0.5-nm resolution are illustrated in Fig. 2 (a)

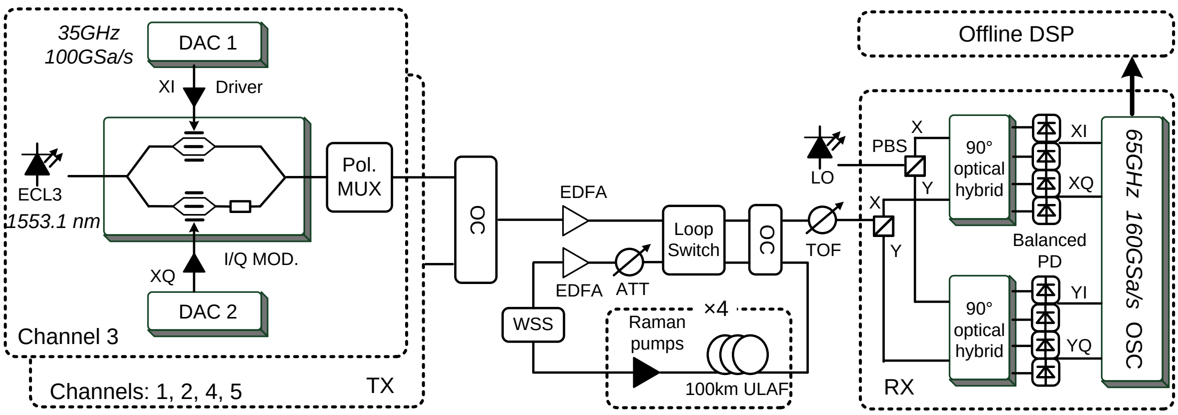

Fig. 1: Experimental setup of 5-channel WDM transmission based on 100-Gbaud PM TPS-64QAM.

and (b), respectively. After ULAF transmission, we utilize a tunable optical filter (TOF) to select the desired sub-channel before it is detected by a coherent receiver[1]. Afterwards, a digital oscilloscope (OSC) with 160-GSa/s sampling-rate and 65-GHz bandwidth is used to realize the digitization and sampling of received signals.

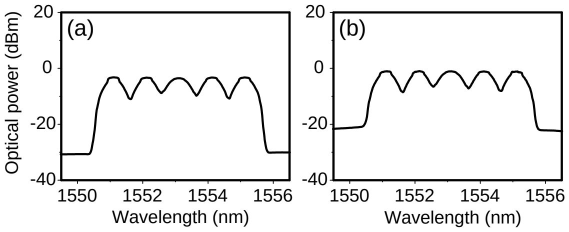

Fig. 2: The optical spectra of 5-channel WDM TPS-64QAM signals (a) before and (b) after 2,000-km ULAF transmission.

Comparison of MIMO Volterrra and DNN-based Nonlinear Equalization

The beginning steps of our DSP scheme are the same as those of common coherent transmission. We implemented NLE as the last step of the DSP after carrier phase estimation (CPE). We have two NLE schemes, which are MIMO Volterra[1] and DNN-based NLEs. The I/Q separated DNN-NLE is shown in Fig. 3. It's proved that separated NNEs can be better trained than a single NNE that adopts both I/Q signals when I/Q imbalance is well-solved by former DSP steps[8].

The parameters configuration of the NLEs is shown in Tab. 1. The linear memory length M_1 and nonlinear memory lengths M_2 of Volterra NLE are 149 and 99, respectively. For the separated DNN-NLE scheme, we used four independent DNNs, corresponding to the I/Q components of two polarizations. The memory length of DNN is 80. The number of units in first and second hidden layers is 80 and 60, respectively. Considering the uneven distribution of the signal and soft-decision LDPC, we adopted mean-square error (MSE) as loss function. Our experimental results of BER versus OSNR and transmission distance for subchannel-3 WDM TPS-64QAM signal are illustrated in Fig. 4 and 5, respectively.

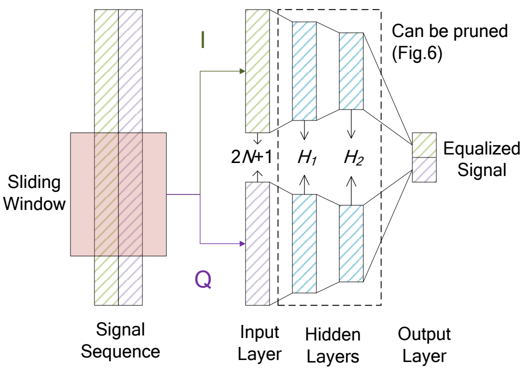

Fig. 3: Principle of separated DNN-NLE.

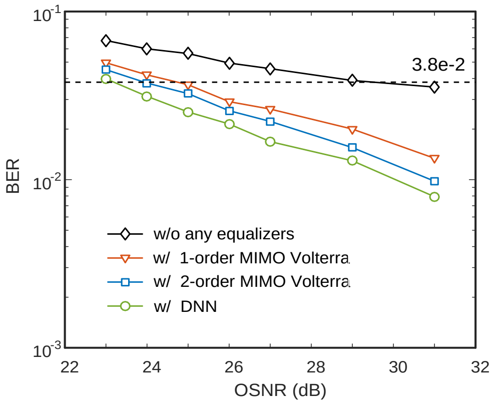

Fig. 4: BER versus OSNR under BtB condition.

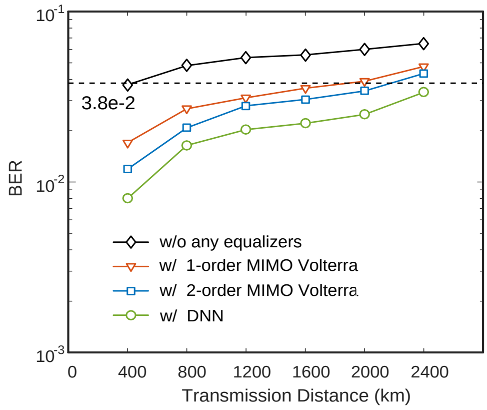

Fig. 5: BER versus ULAF transmission distance.

In order to make the nonlinear compensation performance of NLEs more clear, we added a 1-order MIMO Volterra scheme. Compared with the 1-order Volterra NLE, the 2-order Volterra NLE can bring around 0.7-dB sensitivity gain, and the transmission distance for TPS-64QAM signal can be extended from 1,600 km to 2,000 km when considering 3.8×10^{-2} LDPC threshold with 25% overhead[1]. When changing the Volterra NLE to DNN-NLE, we can obtain another 0.7-dB sensitivity gain and the transmission distance can be further improved to 2,400 km.

However, the multiply-accumulate operation (MACC) per symbol of DNN-NLE is around 3 times that of Volterra NLE. Hence, we utilized pruning technique to reduce the complexity.

Tab. 1: Computational complexity of NLE schemes.

NLEs	Parameters	MACC per symbol
Volterra NLE	M_1=149 M_2=99	$=4(M_2^2+M_2)+8M_1$ =**40,792**
DNN-NLE	N=80 H_1=80 H_2=60	$=4(2H_1(2N+1)+H_1H_2+H_2)$ =**122,480**
Pruning DNN-NLE	N=80 H_1=80 H_2=60 Sparsity=75%	=**122,480 × 25%** =**30,620**

Pruning DNN-based Nonlinear Equalization

It's shown that there are a large number of redundant neurons and weights in the neural network model, and the weights that participate in the main calculation and affect the final result only account for 5-10% of the total [9]. In long haul optical transmission, the system requires high accuracy. Variation in BER performance is critical for FEC. Therefore, it is important to control performance decay during pruning. Considering one-shot pruning can be strongly affected by noise, we pruned the model iteratively. The scheme of the pruning process is shown in Fig. 6.

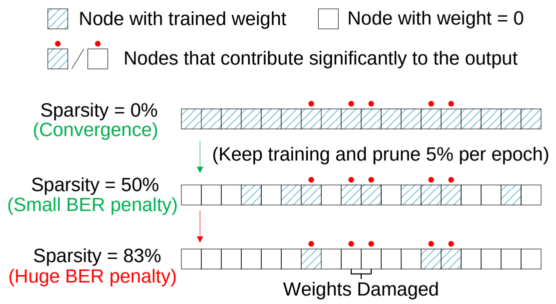

Fig. 6: Principle of pruning DNN-NLE.

After normal training (until convergence), we iteratively pruned the separated NN models and kept training the models. The sparsity of the models increased 5% and the models are trained and evaluated for the next iteration. The optimal pruning model should be picked out according to the performances.

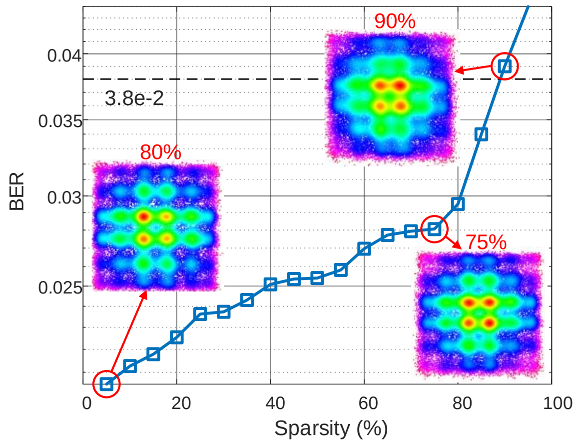

Fig. 7: BER versus pruning sparsity for TPS-64QAM signal after 1,200-km transmission.

By equalizing signals in 1,200-km transmission, the BER performances in pruning process are recorded, and the results are shown in Fig.7. We obtained a pruning scheme with 75% sparsity that combines low-complexity and acceptable performance. The pruning scheme is then used to reduce the complexity of equalizing signals in 400-2,400-km transmission.

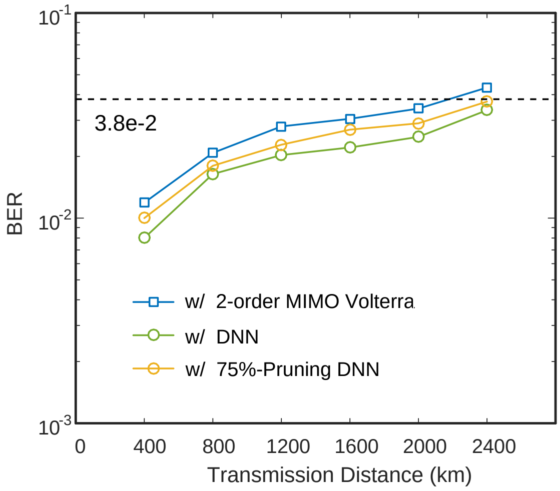

Fig. 8: BER versus ULAF transmission distance.

The BER results versus transmission distance for pruning DNN is shown in Fig. 8. When the pruning sparsity is 75%, the BER of TPS-64QAM signal after 2,400-km transmission is still lower than 3.8×10^{-2} threshold. In other words, when the complexity of DNN-NLE is reduced by 75%, the transmission distance of 800-Gbit/s/carrier TPS-64QAM signal can still reach 2,400 km. At this time, the complexity of pruning DNN-NLE is around *24%* lower than the Volterra NLE.

Conclusions

We experimentally demonstrated 800-Gbit/s/carrier WDM coherent transmission over 2,400-km based on 100-GBd truncated PS-64QAM utilizing 75%-sparsity pruning DNN-based nonlinear equalization.

We compared separated DNN-NLE with 2-order MIMO Volterra NLE. When changing the Volterra NLE to DNN-NLE, we can obtain around 0.7-dB sensitivity gain and the transmission distance can be improved from 2,000 km to 2,400 km. We also utilized pruning DNN-NLE to reduce the computational complexity. Our 75%-sparsity pruning DNN-NLE with 24%-lower complexity outperforms Volterra NLE by 20% reach improvement.

Acknowledgements

This work is supported by National Key R&D Programmes of China with number of 2018YFB1800905, and NNSF of China (61935005, 61720106015, 61835002, and 62127802).

References

[1] M. Kong, J. Shi, B. Sang, J. Ding, K. Wang, W. Li, F. Wang, C. Liu, Y. Wang, Y. Wei, B. Zhu, L. Zhao, W. Zhou and J. Yu, “800-Gb/s/carrier WDM Coherent Transmission Over 2000 km Based on Truncated PS-64QAM Utilizing MIMO Volterra Equalizer”, *Journal of Lightwave Technology*, vol. 40, no. 9, pp. 2830–2839, 2022, DOI: 10.1109/JLT.2022.3148336

[2] V. Bajaj, F. Buchali, M. Chagnon, S. Wahls, and V. Aref, “Deep Neural Network-Based Digital Pre-Distortion for High Baudrate Optical Coherent Transmission”, *Journal of Lightwave Technology*, vol. 40, no. 3, pp. 597–606, 2022, DOI: 10.1109/JLT.2021.3122161

[3] L. Yi, T. Liao, L. Huang, L. Xue, P. Li, and W. Hu, “Machine Learning for 100 Gb/s/λ Passive Optical Network”, *Journal of Lightwave Technology*, vol. 37, no. 6, pp. 1621–1630, 2019, DOI: 10.1109/JLT.2018.2888547

[4] K. Wang, C. Wang, J. Zhang, Y. Chen, and J. Yu, “Mitigation of SOA-Induced Nonlinearity with the Aid of Deep Learning Neural Networks”, *Journal of Lightwave Technology*, vol. 40, no. 4, pp. 979–986, 2022, DOI: 10.1109/JLT.2021.3126765

[5] L. Ge, W. Zhang, C. Liang, and Z. He, “Compressed Neural Network Equalization Based on Iterative Pruning Algorithm for 112-Gbps VCSEL-Enabled Optical Interconnects”, *Journal of Lightwave Technology*, vol. 38, no. 6, pp. 1323–1329, 2020, DOI: 10.1109/JLT.2020.2973718.

[6] T. Koike-Akino, Y. Wang, K. Kojima, K. Parsons, and T. Yoshida, “Zero-Multiplier Sparse DNN Equalization for Fiber-Optic QAM Systems with Probabilistic Amplitude Shaping”, in *2021 European Conference on Optical Communication (ECOC)*, Bordeaux, France, 2021, pp. 1–4, DOI: 10.1109/ECOC52684.2021.9605870.

[7] B. Sang, W. Zhou, Y. Tan, M. Kong, C. Wang, M. Wang, L. Zhao, J. Zhang and J. Yu, “Low Complexity Neural Network Equalization Based on Multi-Symbol Output Technique for 200+ Gbps IM/DD Short Reach Optical System”, *Journal of Lightwave Technology*, vol. 40, no. 9, pp. 2890–2900, 2022, DOI: 10.1109/JLT.2022.3146863

[8] C. Liu, C. Wang, W. Zhou, F. Wang, M. Kong, and J. Yu, “81-GHz W-band 60-Gbps 64-QAM wireless transmission based on a dual-GRU equalizer”, *Optical Express*, vol. 30, no. 2, pp. 2364–2377, 2022, DOI: 10.1364/OE.448845

[9] Han, Song, Huizi Mao, and William J. Dally, “Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding”, *arXiv preprint*, 2015, DOI: 10.48550/arXiv.1510.00149