Neural Network-based Pre-Distorter with Long Memory Length for PAM-8 IM/DD Transmission

Zhiwei Chen[1,2], Wei Wang[1,2], Dongdong Zou[1,2], Weihao Ni[1,2], Mincong Deng[1,2], and Fan Li[1,2*]
[1] School of Electronics and Information Technology, Guangdong Provincial Key Laboratory of Opto-Electronic Information Processing Chips and Systems, Sun Yat-Sen University, Guangzhou 510275, China
[2] Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai 519000, China
*lifan39@mail.sysu.edu.cn

***Abstract*—Neural network-assisted pre-distortion method is proposed to generate pattern errors with memory length of more than 5 symbols in PAM-8 IM/DD transmission utilizing limited samples as training sequence, which enables an improvement on system performance.**

Keywords—Neural network, pulse-amplitude-modulation (PAM), look-up-table (LUT), pre-distortion

I. Introduction

Data center interconnect (DCI) applications, multimedia services, cloud computing and other network applications are driving demands for high capacity transmission. Intensity modulation and direct detection (IM/DD) system is widely investigated in short-reach applications due to low system complexity and low cost. Since higher symbol rate poses higher requirements on system bandwidth, advanced modulation formats, such as high-order pulse-amplitude-modulations (PAMs), are employed to achieve higher spectral efficiency [1-2]. However, high-order PAMs may be more sensitive to linear and nonlinear impairments, which limits the system performance and should be addressed with advanced digital signal processing (DSP) algorithms in the transmitter side or receiver side [3]. For short-reach applications, nonlinear distortions from devices such as nonlinear characteristics of the modulator is nonnegligible [4-5]. To alleviate or mitigate the influence of nonlinear impairments induced by modulation and square-law detection, digital pre-distortion and post-compensation techniques have been proposed such as look-up-table (LUT) pre-distortion [6-8], Volterra nonlinear equalizer (VNLE) [9] and machine learning methods [10-11]. Though the performance of LUT pre-distortion has been proved, it mainly depends on the memory length of the pattern sequences [12]. In [13], the authors compared the performance of LUTs with various memory lengths and found high-order nonlinearity can be better uncovered with greater memory length. The size of LUT grows exponentially over the memory length, which brings forward high requirements on the storage especially when high-order modulation formats are employed. Therefore, the memory length of pattern is typically limited to 5 or 3, which limits the performance. In [11], the authors propose a digital pre-distortion (DPD) method based on neural network (NN), which emulates the LUT-based approach and lessens the requirements on memory. However, they only focus on the case of high modulation orders with memory length of 3.

In this work, we propose to use artificial neural network (ANN) to generate pattern errors with longer symbol patterns. Pattern-dependent distortion value is calculated according to transmitted sequence rather than read from the memory storage, which means there is no need for excessive storage at the expense of additional computational complexity. The number of multiplications per symbol is linearly dependent on the memory length. Besides, the relations between distortion values and symbol patterns can be learned during training, which means limited sample points in sample space is enough for the training of neural network-assisted pre-distortion method. Moreover, a refined modeling of transceiver is dispensable in this approach. We compare the performance of the NN-DPD with various memory lengths, the LUT-DPD in the case of 3-symbol and the 3rd-order VNLE when modulation-induced nonlinearity exists. Simulation results show that NN-DPD outperforms VNLE in performance with a lower complexity.

II. Principle

The schematic diagram of the proposed NN-DPD is given in Fig. 1. NN is intended for nonlinear regression to figure out the relations between distortion values and symbol patterns, without the need of the transceiver model. The purpose of NN is to provide distortion values according to certain patterns, while the symbol sequences are used to calculate the distortion values rather than form the memory address. The transmitted symbol sequences are first selected by a sliding window with length of $2l+1$ which is the memory length. These sequences are fed to NN, and the outputs of NN are corresponding distortion values. The output of each layer can be expressed as (1).

$$\boldsymbol{h}_i = f(\boldsymbol{w}_i \, \boldsymbol{h}_{i-1} + \boldsymbol{b}_i) \tag{1}$$

Here $\boldsymbol{w}_i$ and $\boldsymbol{b}_i$ are weight matrix and bias vector for ith layer of NN. $f(\cdot)$ is the activation function, which is chosen to be hyperbolic tangent in case of the negative missing [14]. During training, the distortion values between the transmitted signal and the received one are used as label signal. Training sequences are not required to cover all the events in sample space which is hard to realize when extensive patterns exist, but need to contain a small number of patterns. In our simulation, training sequence is composed of 110044 randomly generated PAM-8 symbols.

In the testing phase, the distortion value is subtracted from the original transmitted symbol, and the pre-distorted symbol can be expressed as (2).

$$x'(k) = x(k) - e(k) \tag{2}$$

III. Simulation Setup

The simulation setup is shown in Fig. 2. The continuous-wave (CW) signal from a distributed feedback (DFB) laser at 1553.6 nm with linewidth of 1 MHz is modulated by 56 Gbaud PAM-8 signal via a Mach-Zehnder modulator (MZM) with peak RF voltage of 5 V. The signal is transmitted under back-to-back (BTB) case, and a variable optical attenuator (VOA) is applied to adjust the received optical power (ROP). In the

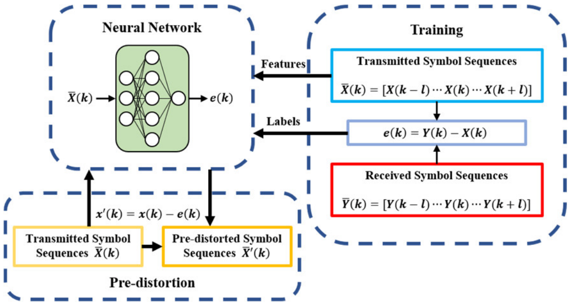

Fig. 1. Block diagram of NN pre-distortion.

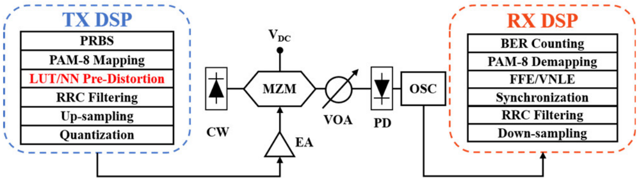

Fig. 2. Simulation setup. CW: continuous-wave; EA: electrical amplifier; VOA: variable optical power; PD: photodiode; OSC: oscilloscope.

receiver side, the signal is detected by a photodiode (PD) and captured by an oscilloscope (OSC).

At Tx DSP, pseudo-random binary sequence (PRBS) is first generated and mapped to PAM-8 symbols. Then, LUT-based or NN-based pre-distortion is applied to the signal for nonlinear compensation. The signal is then processed by root-raised-cosine (RRC) filter with a roll-off factor of 1/16 for Nyquist shaping and up-sampled to 120 GSa/s. Finally, 8-bit quantization is applied. At Rx DSP, after down-sampling, matched filtering and synchronization, the signal is equalized by 3-tap T-spaced feedforward equalizer (FFE) or 3rd-order VNLE. The 1st, 2nd, and 3rd memory lengths of the VNLE is optimized to 7, 7 and 5, respectively. After decision, bit error rate (BER) is calculated. DSP block is also given in Fig. 2.

IV. Results and Discussions

The number of hidden layer neurons is first investigated. ROP and peak-to-peak voltage (Vpp) of the electrical signal are set at -11.5 dBm and 2.6 V, respectively. The results are given in Fig. 3. When memory length increases to 9, the NN-DPD provides better performance, and 10 hidden layer neurons are enough considering the complexity. It is worth noting that the number of hidden layer neurons shouldn't be small when memory length is long, of which the nonlinear regression performance may degrade. We use 10 hidden neurons for NN-DPD in the rest of this paper.

Then we set ROP at -11.5 dBm to observe the BER performance of different schemes with different Vpp of the electrical signal, which is shown in Fig. 4. As shown in Fig. 4, the BER curves decline first since the noise influence is reduced as the Vpp increases. As the Vpp increases, the influence of nonlinearity becomes more serious, and system performance degrades when Vpp is greater than a certain value. Without pre-distortion and post-compensation, the optimal Vpp is 2.2 V, and the corresponding BER is 9×10^{-4}. Using LUT with a memory length of 3 (LUT-3) for pre-distortion helps to mitigate the effect of nonlinearity, and the optimal Vpp comes to be 2.4 V. However, its performance is limited by memory length, which is hard to increase as discussed earlier. Using VNLE instead of FFE also helps to improve the system performance, which scores higher than LUT-DPD when Vpp is small. With the NN-DPD applied, we can observe a greater improvement in BER performance compared to the above schemes, since the memory length can

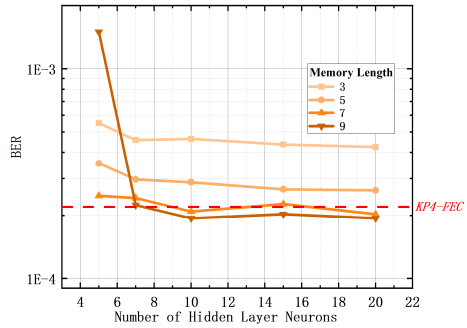

Fig. 3. BER versus the number of hidden layer neurons.

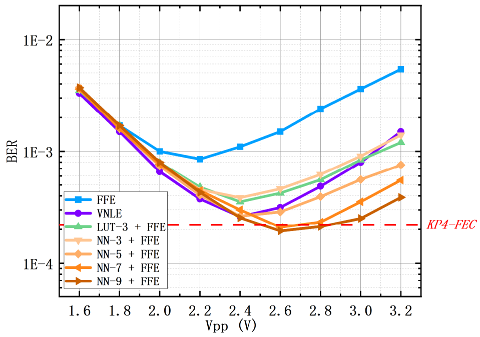

Fig. 4. BER versus Vpp in BTB transmission with ROP set at -11.5 dBm.

be increased without introducing excessive computational complexity. It can be observed that NN with a longer memory length achieves greater advance. With the assist of NN-DPD, minimum BER is achieved, which is 2×10^{-4}. We should also notice that NN with a memory length of 3 (NN-3) has a worse performance than LUT with the same memory length, which is probably due to the influence of noise on the training of NN.

We also compare different schemes with Vpp set at 2.6 V and 3.2 V, under different ROPs. The results are given in Fig. 5 (a-b). As the ROP grows, we can notice evident improvement of BER performance. Besides, longer memory length is needed when nonlinear impairments get more serious. Fig. 5 (c-f) show the receptions of signal with Vpp set at 2.6 V and ROP adjusted to -5.5 dBm, from which we can notice a most reduction of nonlinear influence with NN-9 pre-distortion.

Finally, the computational complexity is investigated between NN-based pre-distorters with different memory lengths in terms of multiplication operations per symbol, which is compared to that of the VNLE. Table I shows the multiplications per symbol needed with these methods. Here, the computational complexity of the activation function is small, which can be neglected. It is evident that the proposed NN-DPD method outperforms VNLE in respect of complexity, since the number of multiplications increases linearly with the memory length.

TABLE I. Computational Complexity of the VNLE and NN Pre-distortion Methods

Methods	Number of multiplications per symbol
3rd-order VNLE	*168 (7+28×2+35×3)*
NN-3 + FFE	*47 (10×3+10+7)*
NN-5 + FFE	*67 (10×5+10+7)*
NN-7 + FFE	*87 (10×7+10+7)*
NN-9 + FFE	*107 (10×9+10+7)*

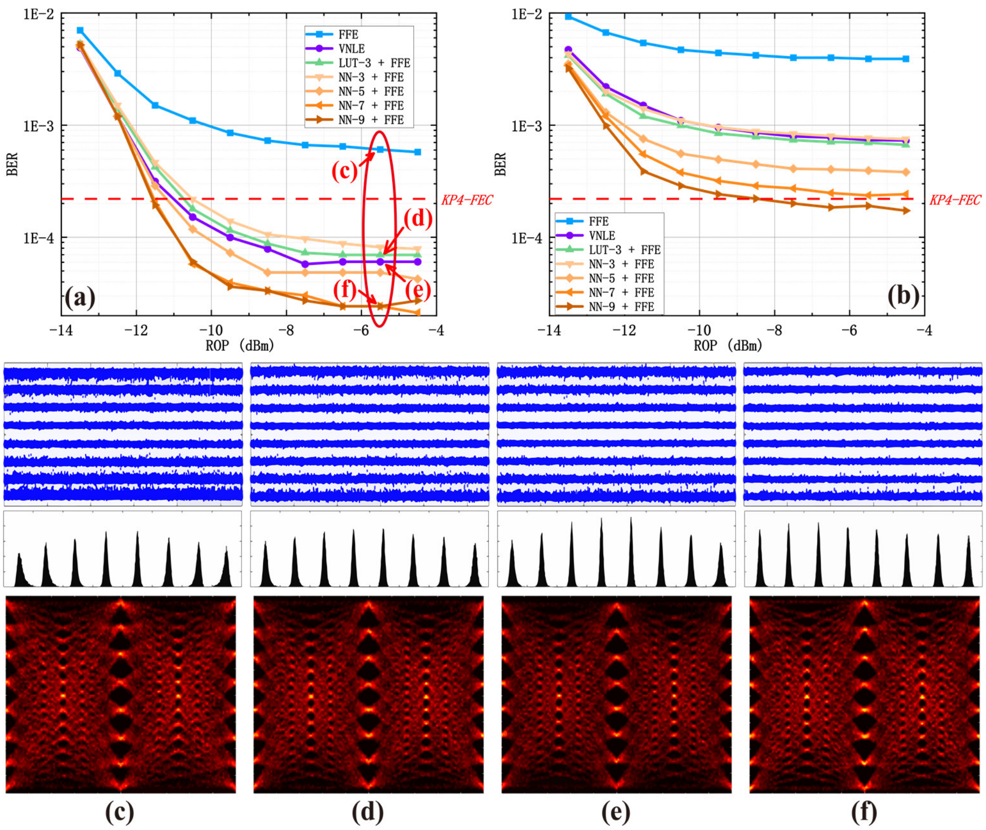

Fig. 5. (a)BER versus ROP in BTB transmission with Vpp set at 2.6 V. (b)BER versus ROP in BTB transmission with Vpp set at 3.2 V. (c)PAM-8 constellation, histogram and eye diagram of the received signal without the application of pre-distortion or post-compensation. (d)PAM-8 constellation, histogram and eye diagram of the received signal with LUT pre-distortion. (e)PAM-8 constellation, histogram and eye diagram of the received signal with the application of VNLE. (f)PAM-8 constellation, histogram and eye diagram of the received signal with NN pre-distortion.

V. Conclusion

In this paper, we propose a neural network-based pre-distortion method to mitigate the influence of nonlinear impairments, which outperforms VNLE in BER performance with less complexity. Since the relations between distortion values and patterns can be obtained by neural network, limited samples are enough for the training of NN-based pre-distorter even though numerous patterns exist. It is simple to implement NN-DPD, and simulation results have witnessed its competitive performance in nonlinear compensation.

Acknowledgment

This work is partly supported by the National Key R&D Program of China (2018YFB1800902); National Natural Science Foundation of China (U2001601, 62271517, 62035018); Local Innovation and Research Teams Project of Guangdong Pearl River Talents Program (2017BT01X121).

References

[1] X. Pang et al., "200 Gbps/lane IM/DD technologies for short reach optical interconnects," J. Lightw. Technol., vol. 38, no. 2, pp. 492-503, 15, Jan. 2020.

[2] D. Che and X. Chen, "Higher-order modulation vs faster-than-Nyquist PAM-4 for datacenter IM-DD optics: an AIR comparison under practical bandwidth limits," J. Lightw. Technol., vol. 40, no. 10, pp. 3347-3357, 15, May 2022.

[3] J. Zhang et al., "56-Gbit/s PAM-4 optical signal transmission over 100-km SMF enabled by TCNN regression model," IEEE Photon. J., vol. 13, no. 4, pp. 1-6, Aug. 2021.

[4] K. Zhong, X. Zhou, J. Huo, C. Yu, C. Lu and A. P. T. Lau, "Digital signal processing for short-reach optical communications: a review of current technologies and future trends," J. Lightw. Technol., vol. 36, no. 2, pp. 377-400, 15, Jan 2018.

[5] J. Li, X. Su, T. Ye, H. Nakashima, T. Hoshida and Z. Tao, "Characterization of nonlinear distortion in intensity modulation and direct detection systems," J. Lightw. Technol., doi: 10.1109/JLT.2023.3248624.

[6] P. Gou, L. Zhao, K. Wang, W. Zhou and J. Yu, "Nonlinear look-up table predistortion and chromatic dispersion precompensation for IM/DD PAM-4 transmission," IEEE Photon. J., vol. 9, no. 5, pp. 1-7, Oct. 2017.

[7] J. Zhang, J. Yu and H. -C. Chien, "1.6Tb/s (4×400G) unrepeatered transmission over 205-km SSMF using 65-GBaud PDM-16QAM with joint LUT pre-distortion and post DBP nonlinearity compensation," Proc. OFC 2017, pp. 1-3.

[8] J. Zhang, J. Yu and H. -C. Chien, "EML-based IM/DD 400G (4×112.5-Gbit/s) PAM-4 over 80 km SSMF based on linear pre-equalization and nonlinear LUT pre-distortion for inter-DCI applications," Proc. OFC 2017, pp. 1-3.

[9] N. Stojanovic, F. Karinou, Z. Qiang and C. Prodaniuc, "Volterra and Wiener equalizers for short-reach 100G PAM-4 applications," J. Lightw. Technol., vol. 35, no. 21, pp. 4583-4594, 1, Nov. 2017.

[10] V. Bajaj, V. Aref and S. Wahls, "Performance analysis of recurrent neural network-based digital pre-distortion for optical coherent transmission," Proc. ECOC 2022, pp. 1-4.

[11] Z. He et al., "Symbol-based supervised learning predistortion for compensating transmitter nonlinearity," Proc. ECOC 2021, pp. 1-4.

[12] H. Jiang et al., "Aggressive look-up table compression by principal component analysis for digital pre-distortion in optical communication," J. Lightw. Technol., vol. 40, no. 24, pp. 7718-7726, 15, Dec. 2022.

[13] H. Chien and J. Yu, "On single-carrier 400G line side optics using PM-256QAM," Proc. ECOC 2016, pp. 1-3.

[14] S. Qiu, X. Xu and B. Cai, "FReLU: flexible rectified linear units for improving convolutional neural networks," Proc. ICPR 2018, pp. 1223-1228