Neural Network-Enhanced Optical Phase Conjugation for Nonlinearity Mitigation

Morteza Kamalian-Kopae, Abdallah A. Ali, Karina Nurlybayeva, A. Ellis, S. Turitsyn
Aston Institute of Photonic Technologies
m.kamalian-kopae1@aston.ac.uk

Abstract: Using a multi-layer perceptron to equalise the residual nonlinearity from the transmission of PDM 28 Gbaud 64QAM over 400km of SSMF employing midlink optical phase conjugation, we demonstrate 12-fold reduction in the BER.

1. Introduction

Techniques to mitigate the impact of Kerr nonlinearity in standard single mode fiber (SSMF) can be categorised into two sets of digital signal processing (DSP) methods such as digital back propagation [1], Volterra series equaliser [2], and machine learning [3] or optical methods such as optical phase conjugation (OPC) [4]. Whilst the DSP methods are limited to the compensation of intra-channel effects, the OPC introduces simultaneous compensation of the dispersion and nonlinearity of inter- and intra-channel nonlinear effects. However, there are some impairments resulting from the experimental design of the OPC device [5]. The merging of the DSP methods with OPC can offer an opportunity for combining advantages of both methods while overcoming the drawbacks. Examples of this approach can be reducing the OPC power symmetry requirement with integrating Volterra equaliser with the OPC [6], or relaxing the OPC design using machine learning [7]. One of the simplest and widely used machine learning techniques in equalising communication signals is the multi-layer perceptron (MLP). MLP is a feed-forward neural network (NN) consisting of several layers each with neurons (represented by nonlinear activation functions) and linear connection between layers. The MLP is an attractive method for an equaliser due to two main reasons. First is the simple structure of MLP with low computational complexity (CC). This is mainly due to the fact that numerical implementation of the nonlinear activation functions used in the MLP is based on a look-up-table with minimal computational load. This should be compared to the more powerful, but computationally expensive units such as Gated recurrent units (GRU) and Long Short-Term Memory (LSTM) [8]. Another advantage of the MLP is that there are several ways to implement an MLP in the optical domain using photonic devices that are inherently power efficient and fast. The basic building blocks of an MLP are the weighted summation and nonlinear function. The former can be implemented using optical components [9, 10] but the latter is more challenging yet there are proposals with experimental demonstration of their capabilities [11, 10]. For such a simple NN, the number of layers and neurons in each layer determine the CC, which in turn, governs other important traits of the equaliser such as power consumption, required memory, and induced processing delay. Among the parameters of the NN the input vector length is of particular importance and is varied in our study to investigate the memory impact of the OPC-aided communication system. This is an important number because in a fully connected NN such as the one we considered here, the first layers of the network play the role of a feature generator taking into account all the combinations of the samples in the input vector. This means that an input vector of length N samples can accommodate a channel with the maximum memory of N samples. This length impacts the CC of the equaliser and the processing delay and needs to be kept as small as possible. The effective remaining channel memory is a complicated quantity influenced by many parameters of the optical link [12]. An optical solutions mitigating fibre distortions can help to reduce this figure and decrease cross-talk between channels and symbols especially in high data rate and long distance transmissions [13]. In a fully in-line dispersion-compensated link or a symmetric OPC system, the effective channel memory is expected to be small. However, due to nonlinearity of the fibre and asymmetry of the link some residual memory will be always present. This makes the equalisation process more challenging and limits the achievable improvement as will be demonstrated in the following. In this paper, we use a simple MLP equaliser in a fibre communication system and investigate the impact of the OPC as an optical technique to mitigate nonlinearity of the SSMF on the performance of this equaliser and show that this combination reduces the BER below the %7 overhead HD-FEC threshold. We also demonstrate how the OPC helps to reduce the residual memory of the link to the extent that only 7 signal samples is required to equalise the sample of interest.

2. Experimental setup

The experimental setup is shown in Fig.1 consisting of a polarisation division multiplexing (PDM) 28-Gbaud 64QAM transmitter, 4 spans of SSMF (α= 0.2 dB/km, dispersion = 17 ps/nm/km,γ=1.3 /W/km) (with Mid-link OPC) and a coherent receiver. A CW laser tuned at 1555.75 nm is used as the signal source for a 45 Gbaud IQ modulator. The IQ modulator is driven by 4 channels AWG where the symbols (28Gbaud baudrate) are generated. The symbols were time-multiplexed with 5% quadrature phase-shift keying pilot symbols periodically to be used in the data-aided algorithm of the receiver DSP. The symbols over-sampled and loaded in the AWG (sampling rate 56GSa/s). An EDFA (6-dB noise figure) was added at the beginning of each span to change the power. In the middle of the link, either the signal passed through the OPC, was amplified by an EDFA with fixed output power (15 dBm) and its conjugate was propagated in the second half of the link, or the signal bypassed the OPC device. A variable optical attenuator with a WSS was employed to emulate the insertion penalty of the OPC where the OSNR was set to be the same in the second half of the link for the two cases (with and without the OPC). A dual pump polarisation independent OPC is used to generate the conjugate [5] where two orthogonal pumps at 1540.5 nm and 1560.1 nm (line-widths of $<$10 kHz and$<$100 kHz respectively) were generated using counter-dithered (using two RF tones at 60 and 600MHz) CW lasers and high power EDFAs, then filtered and combined with the signal in 100m of a highly nonlinear fiber (HNLF) (zero-dispersion wavelength = 1550 nm,α=1.2 dB/km, γ=21.4 /W/km, and dispersion slope = 0.041 ps/nm^2/km at 1550 nm). An optical band-pass filter with WSS is used to suppress the pumps and extract the conjugate. The coherent receiver consists of a local oscillator (100-kHz linewidth) whose signal was mixed with the received signal in a 90° optical hybrid. Four balanced photo-diodes were connected to the hybrid outputs, and a real-time sampling scope (100-GS/s sampling rate, 33-GHz 3-dB bandwidth) was used as analogue-to-digital converter. Offline data-aided digital signal processing was performed on a desktop computer and the machine learning is applied before symbols de-mapping. The configuration of the NN including the number of layers and neurons in each layer is optimised for all combinations of launched power and input vector sizes using the Bayesian optimisation. Here, all layers are considered to have the same number of neurons so the resulting configuration is sub-optimal. The input layer consists of a sequence of consecutive samples from the two polarisations separated into real and imaginary parts with various sizes to study the memory reduction effect of the OPC. An important element of the MLP is the activation function. Here we used *tanh* function for all layers except the last one where we deployed a linear function. This makes our equaliser a regression ML solution where the aim is to reduce the impact of noise–that is a complicated and nonlinear intermix of noise and signal–from the received QAM symbol. The training procedure consists of 500 epochs with batch size of 2000 using 2^{18} pairs of complex-valued dual polarisation samples and the associate target transmitted QAM symbol. The training, validation and test sets are 70, 20, and 10 per cent of the dataset, respectively. The loss function used for the training is mean square error and Adam optimiser with learning rate of 0.001 is chosen for the back-propagation. To reduce the possibility of over-fitting l_2 weight regularisation is applied.

3. Experimental results

We applied the optimum (at each power and input vector length) MLP on the samples from PDM 28 Gbaud 64QAM signal with and without OPC. The results shown in Fig. 2 are the BER with and without OPC with no nonlinear compensation and are plotted with square markers. Fig. 2 also shows the improvement in the performance obtained by using an MLP with 7 and 15 sample input vector which demonstrates considerable reduction in the final BER as much as 12 fold for the case of OPC-aided MLP equaliser. This improvement is achieved through a noise reduction mechanism as can be seen from the received constellation of the selected points in the side panels of Fig. 2. As is evident from Fig. 2, this improvement is translated in 4 dB increase in the optimum launch power (slightly lower than the 7% overhead HD-FEC BER) when an OPC-aided MLP-based nonlinear compensation scheme is in place. As the signal power increases and the nonlinear effects become dominant, the amount of improvement by including the MLP equaliser also grows as the nonlinearity compensation through the MLP is more effective at higher powers, see Fig. 3a. In this figure, the improvement ratio of the BER for cases with and without OPC for 7 and 15 sample input vector to the MLP is depicted. From this figure we conclude that the equaliser is in fact tackling the nonlinearity effects of the communication link. Another interesting observation is the impact of the input vector size on the performance of the MLP equaliser. Figure 3b illustrates the BER at the

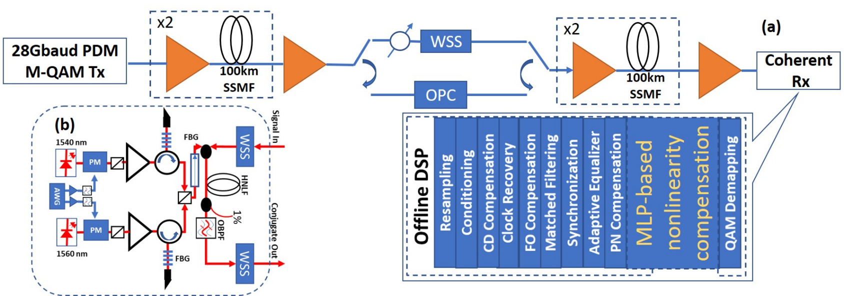

Fig. 1. Block diagram of the experimental setup and the place of the MLP equaliser.

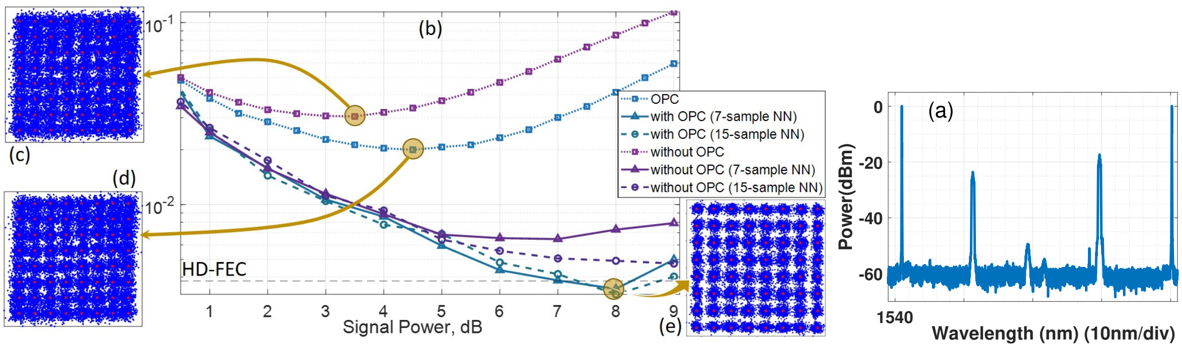

Fig. 2. BER versus power and the equalised 64QAM constellation w/ and w/o OPC/NN equaliser.

optimum (for the case of OPC but not far from the optimum power for the no OPC case as well) power of 8 dBm as a function of the input vector size. We varied the size from 3 samples (equivalent to 0.1 ns channel memory shown on the top) up to 21 samples (equivalent to 0.6 ns). The equivalent delay indicates how long will the MLP wait to collect enough samples in order to equalise the sample of interest during the inference stage. Figure 3b also shows that the OPC-aided system performance saturates beyond 7 sample input vector size which is not the case for the system without OPC. This displays the memory reduction effect of the OPC as was expected. This memory reduction in addition to channel containment has the potential of significantly reducing the complexity of nonlinear noise reduction especially in the WDM and high bandwidth transmissions. As can be seen from Fig. 3b, even in the OPC-aided system, the MLP equaliser still needs some information about 3 samples before and after the sample of interest in order to be able to restore it to the originally transmitted symbol. This none-zero effective memory of the OPC system can be explained by the asymmetry of the link and imperfections of the devices and equipment which could be reduced by using improved symmetric link with dual pump Raman amplification instead of lumped amplification.

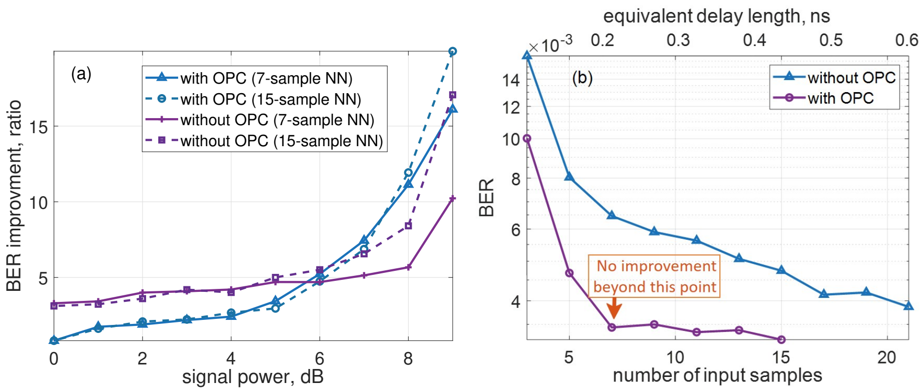

Fig. 3. a) BER improvement w/ and w/o OPC with input vector length of 7, 15, b) BER at optimum power with NN equaliser vs the size of the input vector for two cases with and without OPC.

4. Conclusion

We have implemented an optimised MLP-based equaliser using Bayesian approach, achieving a 12x BER reduction in a PDM 28 Gbaud OPC-aided fibre-optic communication in a 400 km transmission. We have demonstrated that the equaliser can reduce the residual nonlinear impairments. We established that the OPC reduces the effective channel memory that is made evident by the smaller required size and computational complexity of the equaliser.

5. Acknowledgement

This work was supported by the UK ESPRC TRANSNET EP/R035342/1, EEMC EP/S016171/1 and PHOS EP/S003436/1.The authors would like to thank Furukawa Electric Co., Ltd. for providing the HNLF.

References

1. E. Ip *et al.*, JLT, **26**, 3416–3425 (2008).
2. F. Guiomar *et al.*, Opt. Express, **20**, 1360–1369 (2012).
3. T. T. Nguyen *et al.*, JLT, IEEE, **39**, 388–399 (2020).
4. A. A. I. Ali *et al.*, Th2A.11, Proc. OFC, (2019).
5. E. Agrell *et al.*, JLT, **32**, 2862–2876 (2014).
6. E. Shapiro *et al.*, Quantum Electron, **47** (2017).
7. A. A. I. Ali *et al.*, F1B.4 , Proc. OFC, (2021).
8. S. K. Turitsyn *et al.*, Comptes Rendus Physique, **4**, 145–161 (2003).
9. G. Saavedra *et al.*, JLT, **37**, 2467–2479 (2019).
10. L. Wang *et al.*, IEEE Photonics J.,**13**, 1–8 (2021).
11. Y. Zuo *et al.*, Optica, **6**, 1132–1137 (2019).
12. X. Lin *et al.*, Science, **361**, 1004–1008 (2018).
13. Y. Shen *et al.*, Nat. Photonics,**11**, 441–446 (2017).
14. X. Miscuglio *et al.*, Opt. Mater. Express,**11** (2018).
15. P.J. Freire *et al.*, arXiv:2109.08711 (2021).