Experimental Verification of Complex-Valued Artificial Neural Network for Nonlinear Equalization in Coherent Optical Communication Systems

Pedro J. Freire[(1)], Vladislav Neskornuik[(1)], Antonio Napoli[(2)]
Bernhard Spinnler[(2)], Nelson Costa[(3)], Jaroslaw E. Prilepsky [(1)], Emilio Riccardi [(4)], Sergei K. Turitsyn[(1)]

[(1)] Aston University, United Kingdom p.freiredecarvalhosourza@aston.ac.uk [(2)] Infinera, Munich, Germany [(3)] Infinera, Lisbon, Portugal [(4)] Telecom Italia Mobile, Torino, Italy

Abstract *We propose a novel design of neural network for mitigating the fiber nonlinearity, employing a structure based on physical modelling. The neural network achieved nearly 5 times BER reduction in a field trial when transmitting WDM 200G DP-16QAM over a 612 km legacy link.*

Introduction

Machine learning (ML) based methods have been recently considered as a highly promising solution to address various challenges arising in optical communications[1],[2]. In particular, neural network (NN) based algorithms have demonstrated high potential for Kerr nonlinearity mitigation (NLM) in optical fiber links[3]–[5]. The knowledge about the communication system structure can be used to build up advanced NN architectures that provide additional performance gains compared to the conventional NLM approaches like digital back-propagation (DBP)[3]. The ability to quickly collect abundant training datasets, needed for proper NN operation[6], from the huge amount of symbols passing through the link, constitutes an advantage of applying NNs for NLM used in several published works. First, for the regression task, works such as[7],[8] presented an NN-based NLM architecture with delay taps which allowed the equalizer to deal with the channel memory and time-dependent nonlinear response. Second, for the classification task, in[8]–[10] ML has been applied for the determination of nonlinear decision boundaries in QAM modulation formats.

In this paper, a new design for a complex-valued artificial NN performing receiver-based Kerr nonlinearity equalization (NLE) is proposed. The suggested design was numerically tested for the cases of the transmission of a single-channel dual-polarized (DP) 64QAM root-raised cosine (RRC) shaped signal with roll-off 0.06 over the links made by 6×80 km (480 km) of standard single model fiber (SSMF) or large effective area fiber (LEAF). In SSMF and LEAF testcases, respectively, the proposed NN obtains significant $\approx 2.5\times$ and $\approx 10\times$ drops in bit-error rate (BER) when compared to the electronic dispersion compensation aided by a phase/amplitude normalization. Moreover, when compared to NLE based on the dynamic neural network[7],[11] with 2 layers and 192 neurons, a BER decline of $\approx 2.5\times$ and $\approx 5\times$, correspondingly, was obtained for LEAF and SSMF. Finally, the proposed technique was experimentally validated by applying it to a 31×200G channel WDM system employing a DP-16QAM signal, transmitted over a 612 km SSMF legacy link[12]. A $\approx 5\times$ BER drop has been demonstrated on top of electronic dispersion compensation combined with phase/amplitude normalization only.

Neural Network Design

We start from the standard pass-averaged Manakov equation[13] describing the DP signal propagation along the optical channel by:

$$\frac{\partial u_{H/V}(t,z)}{\partial z} + \frac{j\beta_2}{2}\frac{\partial^2 u_{H/V}(t,z)}{\partial t^2} =$$
$$= j\frac{8\tilde{\gamma}}{9}\left[|u_H(t,z)|^2 + |u_V(t,z)|^2\right]u_{H/V}(t,z), \quad (1)$$

where $u_{H/V}(t,z)$ are the normalized optical fields of H and V polarization, respectively, β_2 is the group velocity dispersion (GVD) coefficient, $\tilde{\gamma}$ is the effective averaged nonlinearity coefficient, that includes the effective length scale $L_{\text{eff}} = (1 - e^{-\alpha L})/\alpha$ emerging due to the averaging over N_s periodic span loss and gains, with α being the fiber loss coefficient, γ is the effective nonlinear coefficient, L is the span length. In the case of no GVD $\beta_2 = 0$, the analytical solution of Eq. (1) written in terms of the transmitted x_k and received y_k symbols, can be expressed as[14]:

$$x_{H_k} = y_{H_k}\exp\left[-j\frac{8\gamma L_{\text{eff}}N_s}{9}\left(|y_{H_k}|^2 + |y_{V_k}|^2\right)\right].$$
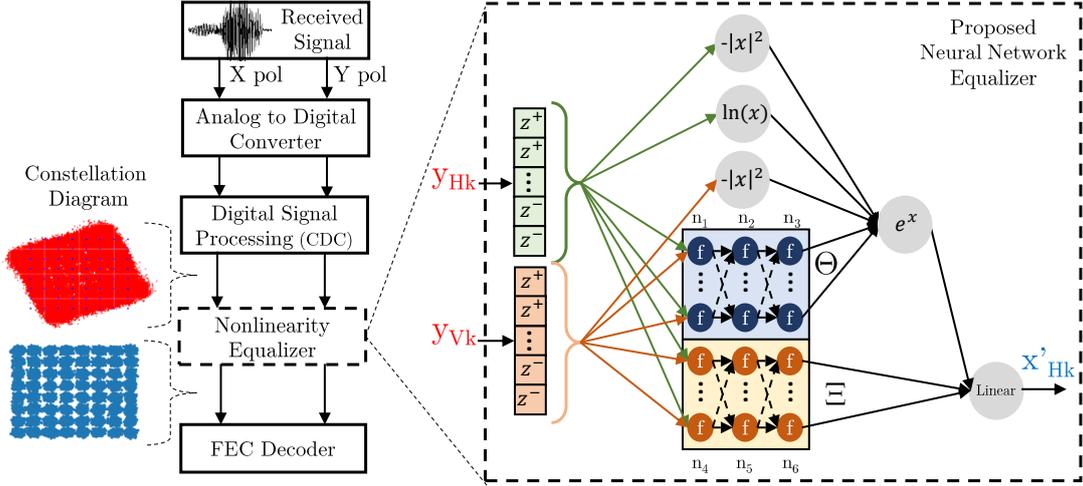$$(2)$$

Fig. 1: Structure of reveicer with the architecture of the proposed neural network.

We assume that Eq. (2), given the adaptive parameters are added, can describe the nonlinear signal distortion remaining after the chromatic dispersion compensation (CDC) in the links with the weak dispersive broadening. It leads us to the following model of nonlinear distortion:

$$x'_{H_k} = c_1 Y_{H_k} e^{\left[-c_2|Y_{H_k}|^2 - c_3|Y_{V_k}|^2 + \Theta_1\right]} + \Xi_1,$$
$$x'_{V_k} = c_4 Y_{V_k} e^{\left[-c_5|Y_{H_k}|^2 - c_6|Y_{V_k}|^2 + \Theta_2\right]} + \Xi_2,$$

(3)

where $x'_{(V/H)_k}$ is the k-th recovered transmitted symbol, $Y_{(V/H)_k}$ is a sequence of the received soft symbols described as $[y_{(V/H)_{k-N}}, ..., y_{(V/H)_k}, ..., y_{(V/H)_{k+N}}]$, $2N + 1$ is the sequence length, $\{c_1, ..., c_5\}$ are complex-valued parameter vectors with the same length $2N + 1$, and $\Theta_{1/2}$ and $\Xi_{1/2}$ are two nonlinear functions that depend on the sequences Y_{H_k} and Y_{V_k}. The sequences $Y_{(V/H)_k}$ account for the complex GVD-induced memory in the nonlinear interactions equalised by the adaptive filters represented by the c_k vectors. The nonlinear functions $\Theta_{1/2}$ and $\Xi_{1/2}$ mimic the inter-channel Kerr nonlinearity and the transciever impairments.

The proposed NN topology, implementing the nonlinear distortion model Eq. (3), is schematically depicted in Fig. 1. Since the complex-valued symbols, weights and functions are used in our approach, we design our NN as a complex-valued one[15],[16]. We use non-conventional neuron activation functions $-|x|^2$, $\ln(x)$ and e^x to make the NN resembling Eq. (3). Furthermore, we introduce Θ and Ξ as three-layer perceptrons, since a multilayer perceptron is a universal nonlinear function approximator[6]. The NN that learns Θ has three layers with n_1, n_2, and n_3 neurons, respectively, in the first, second, and third layer. In the same way, the NN representing Ξ has layers containing n_4, n_5, and n_6 neurons. All the neurons in both Θ and Ξ NNs have the following complex-valued activation function:

$$f(x = x_r + jx_i) = \frac{e^{2x_r} - 1}{e^{2x_r} + 1} + j\frac{e^{2x_i} - 1}{e^{2x_i} + 1}.$$

(4)

The NN is applied to the received soft-symbols already pre-processed by the linear digital signal processing (DSP)[17] including full frequency-domain CDC (see Fig. 1).

Tab. 1: Optimal neural network hyper-parameters found via Bayesian optimisation.

Scenario	N (taps)	n1	n2	n3
		n4	n5	n6
LEAF 480 km (numerical)	18	46	64	63
		49	27	62
SSMF 480 km (numerical)	20	51	54	53
		19	22	56
SSMF 612 km (experiment)	20	90	53	88
		56	60	64

Since the optimal NLC input size and the functions Θ and Ξ may change from case to case, we implemented a Bayesian optimization (BO) algorithm[18] to define in each scenario the best values for the following NN hyper-parameters: input memory size N and neuron numbers $n_1, ..., n_6$. A target parameter for BO was the best BER reached by the NN during 5000 training iterations. The NN itself was trained with the Adam optimizer[19] minimising the mean-squared error (MSE) between the predicted and actual transmitted symbols with the learning rate 10^{-3}, and batch size 1000. Optimal configurations found for each studied testcase, are listed in Table 1. Notably, the NN training data (both simulated and experimental ones) was randomly shuffled before every training iteration to eliminate possible dataset periodicity and avoid the risk of NN overfitting[20].
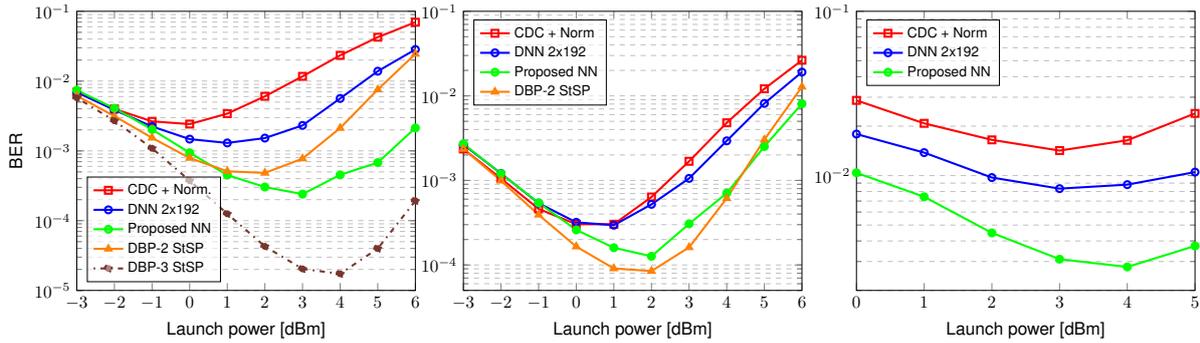
Fig. 2: Numerical test for 6×80 km LEAF   **Fig. 3:** Numerical test for 6×80 km SSMF   **Fig. 4:** Field trial for 612 km SSMF

Results and Discussion

Fistly, we numerically considered the transmission of DP-64QAM symbols shaped by an RRC with roll-off 0.06 over the metro coherent link consisting of 6×80 km LEAF or SSMF spans. The fiber loss, dispersion and effective nonlinear coefficients used were, respectively, for LEAF: α = 0.225 dB/km, D = 4.2 ps/(nm·km), γ = 1.3 W^{-1}km^{-1} and, for SSMF, $\alpha = 0.21$ dB/km, $D = 16.8$ ps/(nm·km), $\gamma = 1.14$ W^{-1}km^{-1}. Every span was followed by an ideal lumped optical amplifier (OA) fully compensating for the losses on the fiber span and injecting the additive white Gaussian noise (AWGN) with noise figure NF = 4.5 dB. The transmission was simulated by the split-step Fourier method[14]. The proposed NN was benchmarked in the role of NLC in the receiver-based DSP (see Fig. 1) against the dynamic neural network (DNN), suggested for NLC in[11], and the classic DBP[21]. To ensure a fair comparison, the employed DNN had two layers with 192 neurons each, as in[7], the same length of input symbol sequences as our NN, and operated on the symbols sequences from both polarizations Y_{H_k}, Y_{H_k}, as in[22]. The DBP operated on 2 samples/symbol signal with 2 steps per span (StPS).

Fig. 2 and Fig. 3 show the results obtained by the considered NLC strategies in the simulated LEAF and SSMF based transmission systems, respectively. For LEAF testcase, the proposed NN reduced the BER by up to $\approx 10\times$ when compared with CDC followed by a least-squares amplitude-phase shift (Norm) and up to $\approx 5\times$ when compared with a two layer DNN. The proposed NN also overcame the DBP 2 StPS performance. Additionally, the optimum launch power increased from about 0 dBm to about 3 dBm when using the proposed NN instead of CDC + norm. For the SSMF system, we have observed a reduction of the BER of up to $\approx 2.5\times$ when using the proposed NN with respect to both the CDC plus Norm. The DNN produced negligible BER improvement in this case.

The experimental setup used for the field trial is described in[12]. The transmission link consists of multiple pairs of G.652 SSMF deployed between Torino and Chivasso in Italy, leading to a total length of 612 km. The transmitted spectrum consists of the channel under test surrounded by 15×200G 16QAM neighbor channels from both the right- and left-hand sides, resulting in a 31 WDM transmission system in the 37.5 GHz grid (33.01 GBaud). Fig. 4 shows the NLC performance obtained in the field trial by the proposed NN and the DNN. The optimal launch power increased from 3 to 4 dBm and the BER fell off by up to $\approx 5\times$ and $\approx 3\times$ when compared with the CDC plus Norm and the DNN approaches, respectively. Moreover, in the experimental study, the NN led to a BER improvement in the linear regime too. This result highlights the ability of the suggested NN to mitigate the impact of not only the Kerr nonlinearity, but also of the additional limitations such as, e.g., the impact of narrow transmitter bandwidth, low-resolution digital-to-analog converters, and imperfections of the driver amplifier and the dual polarization Mach-Zehnder modulator.

Conclusions

We investigated the NLC performance of the suggested complex-valued artificial NN considering different fiber systems and compared it with the dynamic multi-layer neural network. Our results demonstrated that the proposed NN leads to significant system performance improvement when considering both simulated and experimental data. Furthermore, the NN was shown being able to compensate not only channel nonlinearities, but transceiver distortions too.

References

[1] F. Musumeci, C. Rottondi, A. Nag, I. Macaluso, D. Zibar, M. Ruffini, and M. Tornatore, "An overview on application of machine learning techniques in optical networks", *IEEE Communications Surveys & Tutorials*, vol. 21, no. 2, pp. 1383–1408, 2018.

[2] F. N. Khan, Q. Fan, C. Lu, and A. P. T. Lau, "An optical communication's perspective on machine learning and its applications", *Journal of Lightwave Technology*, vol. 37, no. 2, pp. 493–516, 2019.

[3] C. Häger and H. D. Pfister, "Nonlinear interference mitigation via deep neural networks", in *2018 Optical Fiber Communications Conference and Exposition (OFC)*, IEEE, 2018, pp. 1–3.

[4] S. Zhang *et al.*, "Field and lab experimental demonstration of nonlinear impairment compensation using neural networks", *Nature communications*, vol. 10, no. 1, pp. 1–8, 2019.

[5] E. Averyanov *et al.*, "Perturbative machine learning technique for nonlinear impairments compensation in wdm systems", in *2018 European Conference on Optical Communication (ECOC)*, IEEE, 2018, pp. 1–3.

[6] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.

[7] O. Sidelnikov, A. Redyuk, and S. Sygletos, "Equalization performance and complexity analysis of dynamic deep neural networks in long haul transmission systems", *Optics express*, vol. 26, no. 25, pp. 32 765–32 776, 2018.

[8] S. Owaki and M. Nakamura, "Equalization of optical nonlinear waveform distortion using neural-network based digital signal processing", in *2016 21st OptoElectronics and Communications Conference (OECC) held jointly with 2016 International Conference on Photonics in Switching (PS)*, IEEE, 2016, pp. 1–3.

[9] A. Amari *et al.*, "A machine learning-based detection technique for optical fiber nonlinearity mitigation", *IEEE Photonics Technology Letters*, vol. 31, no. 8, pp. 627–630, 2019.

[10] T. Koike-Akino, D. S. Millar, K. Parsons, and K. Kojima, "Fiber nonlinearity equalization with multi-label deep learning scalable to high-order dp-qam", in *Signal Processing in Photonic Communications*, Optical Society of America, 2018, SpM4G–1.

[11] O. Sidelnikov, A. Redyuk, and S. Sygletos, "Nonlinear equalization in long haul transmission systems using dynamic multi-layer perceptron networks", in *2018 European Conference on Optical Communication (ECOC)*, IEEE, 2018, pp. 1–3.

[12] G. Khanna, T. Rahman, E. De Man, E. Riccardi, A. Pagano, A. C. Piat, B. Spinnler, S. Calabro, D. Rafique, U. Feiste, *et al.*, "Comparison of single carrier 200g 4qam, 8qam and 16qam in a wdm field trial demonstration over 612 km ssmf", in *ECOC 2016; 42nd European Conference on Optical Communication*, VDE, 2016, pp. 1–3.

[13] S. Gaiarin *et al.*, "Dual-polarization nonlinear fourier transform-based optical communication system", *Optica*, vol. 5, no. 3, pp. 263–270, 2018.

[14] G. P. Agrawal, "Nonlinear fiber optics", in *Nonlinear Science at the Dawn of the 21st Century*, Springer, 2000, pp. 195–211.

[15] M. Nakamura *et al.*, "Experimental demonstration of spm compensation using a complex-valued neural network for 40-gbit/s optical 16qam signals", *IEICE Communications Express*, vol. 8, no. 8, pp. 281–286, 2019.

[16] T. Kim and T. Adali, "Fully complex multi-layer perceptron network for nonlinear signal processing", *Journal of VLSI signal processing systems for signal, image and video technology*, vol. 32, no. 1-2, pp. 29–43, 2002.

[17] M. S. Faruk and S. J. Savory, "Digital signal processing for coherent transceivers employing multilevel formats", *Journal of Lightwave Technology*, vol. 35, no. 5, pp. 1125–1141, 2017.

[18] M. Pelikan *et al.*, "Boa: The bayesian optimization algorithm", in *Proceedings of the genetic and evolutionary computation conference GECCO-99*, vol. 1, 1999, pp. 525–532.

[19] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization", *arXiv preprint arXiv:1412.6980*, 2014.

[20] T. A. Eriksson, H. Bülow, and A. Leven, "Applying neural networks in optical communication systems: Possible pitfalls", *IEEE Photonics Technology Letters*, vol. 29, no. 23, pp. 2091–2094, 2017.

[21] E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation", *Journal of Lightwave Technology*, vol. 26, no. 20, pp. 3416–3425, 2008.

[22] M. Schaedler *et al.*, "Deep neural network equalization for optical short reach communication", *Applied Sciences*, vol. 9, no. 21, p. 4675, 2019.