Quantum Electronics **51** (12) 1076–1080 (2021)

https://doi.org/10.1070/QEL17656

Application of complex fully connected neural networks to compensate for nonlinearity in fibre-optic communication lines with polarisation division multiplexing

S.A. Bogdanov, O.S. Sidelnikov, A.A. Redyuk

Abstract. **A scheme is proposed to compensate for nonlinear distortions in extended fibre-optic communication lines with polarisation division multiplexing, based on fully connected neural networks with complex-valued arithmetic. The activation function of the developed scheme makes it possible to take into account the nonlinear interaction of signals from different polarisation components. This scheme is compared with a linear one and a neural network that processes signals of different polarisations independently, and the superiority of the proposed neural network architecture is demonstrated.**

Keywords: *fibre-optic communication systems, nonlinearity of optical fibre, fully connected neural networks, polarisation division multiplexing, compensation of nonlinear distortions.*

1. Introduction

Currently, due to the regular appearance of new multimedia applications and services, there is a constantly growing demand for information transmission systems with increased capacity [1]. The development of coherent long-haul communication systems based on polarisation-division multiplexing (PDM) technology makes it possible to double the data transmission rate due to the simultaneous propagation of optical signals along two polarisation fibre components. However, the operation of such systems involves an increase in the total signal power in fibre, which leads to an increase in the influence of nonlinear propagation effects that are known to be one of the key factors limiting the capacity of modern information transmission systems [2, 3]. To overcome this limitation, various technologies for generating and processing an optical signal have recently been proposed.

Among the techniques used for processing optical signals, we can single out methods based on perturbation theory [4, 5] and phase conjugation of signals [6], as well as methods using the Volterra functional series [7] and the Schrödinger nonlinear filter [8]. These approaches are currently competing with machine learning (ML) methods, in particular artificial neural networks (NNs), which allow transmitted symbols to be predicted with high accuracy, while maintaining low computational complexity [9–11].

S.A. Bogdanov, O.S. Sidelnikov, A.A. Redyuk Novosibirsk State University, ul. Pirogova 2, 630090 Novosibirsk, Russia;
e-mail: s.bogdanov@g.nsu.ru

Received 26 October 2021
Kvantovaya Elektronika **51** (12) 1076–1080 (2021)
Translated by I.A. Ulitkin

ML methods are a powerful tool that has many applications for the analysis of complex nonlinear systems. Therefore, it is quite natural that ML has become widely used to improve the efficiency of complex modern fibre-optic communication systems. Due to the accumulation of large arrays of data required for analysis and the emergence of easy-to-use software, ML methods are currently used in almost all areas of optical communication. For example, Zibar et al. [12], using a combination of two multilayer NNs, determine the optimal pump powers and wavelengths to achieve the required broadband Raman gain profile.

In fibre-optic communication lines, a special NN architecture – an autoencoder – is widely used. Using this architecture, Jones et al. [13] solved the problem of optimising the shape of the signal constellation, taking into account the channel nonlinearity and providing the maximum bandwidth of the communication line with wavelength division multiplexing. This approach, in addition to the fibre nonlinearity, allows one to account for the distortions introduced by the components of the transmitter and receiver, and the resulting structure of the signal constellation differs significantly from that in the case of standard optimisation strategies. Another application of an autoencoder for communication lines was demonstrated by Karanov et al. [14], who presented a fibre-optic communication line in the form of an end-to-end deep NN and determined the optimal profile of a quadrature amplitude modulated optical signal, which ensures the best quality of data transmission in a communication system.

ML methods are widely used for failure management in communication lines [15]. A large number of works are presented in which various ML algorithms are used to identify, prevent or eliminate failures caused by various reasons in the operation of optical networks; these ML algorithms are random decision forests [16], artificial NNs [17], support vector machines (SVMs) [18], Gaussian processes [19], Bayesian networks [20], etc.

One of the main directions of application of ML algorithms in communication systems is the development of methods for suppression and compensation of nonlinear signal distortions. For example, Redyuk et al. [21] proposed to combine the perturbation theory-based model and linear regression methods to solve this problem. By using this approach together with quantising the perturbation coefficients and introducing a circular buffer, the authors managed to increase the efficiency of nonlinear distortion compensation, while maintaining low computational complexity. Many works are devoted to the study of methods for processing optical signals, based on artificial neural networks. Note that most of the well-known neural network architectures are currently used for this purpose. Thus, for example, using a neural network, Häger et al. [9] modelled the digital back-propagation (DBP) method, which made it possible to achieve high efficiency of compensation of nonlinear distortions. Further development of this approach was proposed in [22], where digital back-propagation in a communication system with wavelength division multiplexing was simulated using convolutional neural networks with complex-valued arithmetic. Due to the symmetry of dispersion filters and an improved nonlinear activation function, this scheme makes it possible to significantly improve the quality of signal transmission, while maintaining low computational complexity. Deligiannidis et al. [23] demonstrated a signal processing scheme in a receiver of a communication line based on long short-term memory (LSTM) elements, that is, on a variety of recurrent neural networks. Due to the peculiarities of this architecture of neural networks, the methods obtained on their basis have low computational complexity.

The most popular NN architecture, on the basis of which various nonlinear distortion compensation schemes are developed, is fully connected neural networks. A large number of works are devoted to signal processing using such an NN architecture in communication lines with pulse amplitude modulation [24, 25]. In papers [10, 11], using fully connected NNs, nonlinearity is compensated for in systems with 16-QAM format. On the basis of this type of NNs, schemes for processing OFDM signals (orthogonal frequency-division multiplexing) have also been proposed [26]. However, in almost all of these works, the architecture of such networks is implemented using real numbers. The application of the Bayesian optimisation method to complex NNs is demonstrated in [27].

In this paper, we extend the nonlinear distortion compensation scheme based on fully connected neural networks with complex-valued arithmetic, proposed in [11], to the case of communication systems with polarisation division multiplexing. The following sections of the paper describe the fibre-optic communication system in question, the architecture of the proposed NN and the results of applying the developed scheme to compensate for nonlinearity in the receiver of the communication line and to predict the transmitted symbols.

2. Communication line under study

The data transmission system in question is schematically shown in Fig. 1. The communication line consists of a transmitter; twenty 100-km-long spans of standard single-mode fibre (SSMF); erbium-doped optical amplifiers with a noise factor NF = 4.5 dB, used after each span to compensate for losses; and a receiver. The transmitter generates 16-QAM-PDM signals with a symbol rate of 32 Gbaud. A root raised cosine (RRC) filter with a roll-off factor of 0.1 is used to shape the pulses.

The nonlinear propagation of signals along an optical fibre is described by Manakov's system of nonlinear equations [28]:

$$\begin{aligned}\frac{\partial A_x}{\partial z} &= -\frac{\alpha}{2}A_x - \mathrm{i}\frac{\beta_2}{2}\frac{\partial^2 A_x}{\partial t^2} + \mathrm{i}\gamma\frac{8}{9}(|A_x|^2 + |A_y|^2)A_x,\\ \frac{\partial A_y}{\partial z} &= -\frac{\alpha}{2}A_y - \mathrm{i}\frac{\beta_2}{2}\frac{\partial^2 A_y}{\partial t^2} + \mathrm{i}\gamma\frac{8}{9}(|A_x|^2 + |A_y|^2)A_y,\end{aligned} \tag{1}$$

where $A_x(z,t)$ and $A_y(z,t)$ are the x- and y-components of the slowly varying envelope of the optical signal with orthogonal polarisations, respectively; $\alpha = 0.2$ dB km^{-1} is the fibre loss; $\beta_2 = -21$ ps^2 km^{-1} is the chromatic dispersion; and $\gamma = 1.3$ W^{-1} km^{-1} is a nonlinear fibre parameter. Propagation equations were solved numerically using the symmetric split-step Fourier method with a sampling rate of 16 samples per symbol.

In the receiver, after separation of the polarisation components, the signal passed through a matched RRC filter. Then, the accumulated chromatic dispersion was accurately compensated for in the frequency domain and the sampling rate was downsampled to 1 sample per symbol. Next, nonlinear effects were compensated for using the proposed scheme based on a fully connected NN with complex-valued arithmetic. To this end, each complex symbol of the received signal was fed to a separate input node of the NN. Then the signal was demodulated and the bit error rate (BER) was calculated.

3. Fully connected neural network-based scheme of compensation for nonlinear effects

The architecture of the complex NN proposed in this work is shown in Fig. 2. The network consists of two fully connected subnets, each of which processes a signal of one of the polarisations. In this architecture, subnets are interconnected through nonlinear layers. The neural network with complex-valued arithmetic is based on the description using complex numbers of both the state of the neurons themselves and the weight coefficients. Thus, each neuron of the considered NN is represented as a pair of numbers corresponding to the real and imaginary parts of the symbols for which complex-valued arithmetic was implemented. This approach looks more natural when processing received symbols in coherent fibre-optic communication lines, which are complex in nature. In addition, NNs with complex-valued arithmetic make it possible to use complex activation functions corresponding to nonlinear effects affecting signals when they propagate through an optical fibre.

The transmitted signals of two polarisations are received at the input of the NN at a sampling rate of one sample per

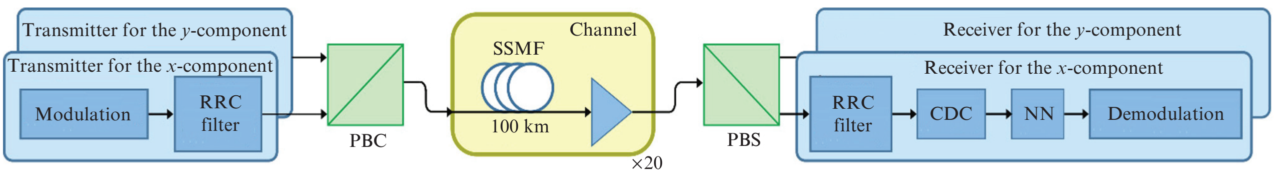

Figure 1. (Colour online) Schematic of a fibre-optic communication line: (CDC) chromatic dispersion compensation; (PBC) polarisation beam combiner; (PBS) polarisation beam splitter.

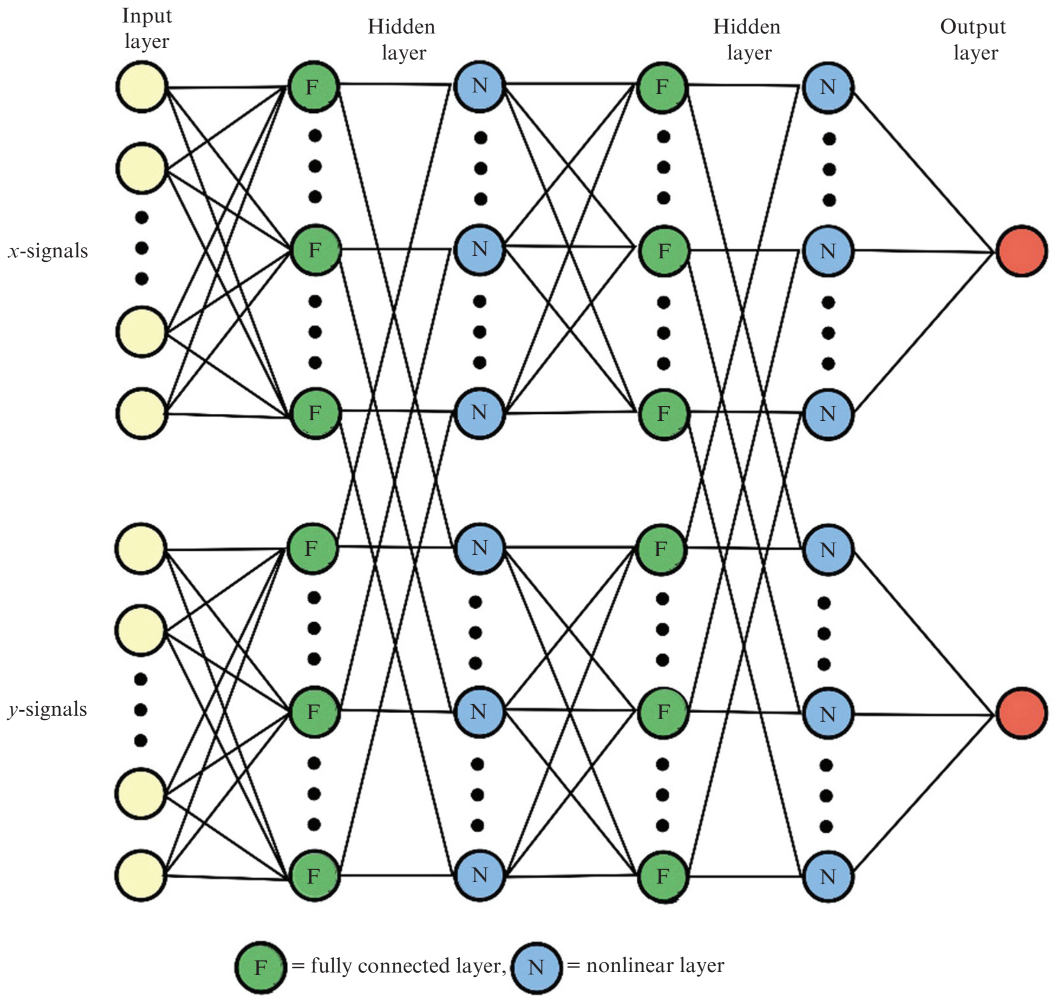

Figure 2. (Colour online) Scheme of a complex fully connected NN for joint processing of data obtained for signals of two polarisations.

symbol, which then propagate through the hidden NN layers. When processing each received symbol, N previous and N subsequent symbols for each polarisation are simultaneously fed to the input, which makes it possible to take into account the channel memory effect. Thus, the total number of input complex-valued symbols for both polarisations is $2 \times (2N + 1)$. Each subnetwork of the proposed NN consists of an input layer, two hidden fully connected layers and an output layer corresponding to the predicted transmitted symbol for this polarisation. The number of symbols N and the number of neurons on the hidden layers were optimised during the study to improve the efficiency of nonlinear distortion compensation. Note that the architecture under consideration is an extension of the complex fully connected NN proposed in [11] for the case of communication systems with polarisation division multiplexing.

The linear part of each hidden layer is the result of multiplying the vector of neuron values obtained on the previous layer by a matrix with complex-valued learning elements of size $M \times P$, where M is the number of neurons on the previous layer, and P is the size of the current hidden layer. Note that symbols for different polarisations propagate along linear layers in parallel and independently of each other. To take into account the nonlinear interaction of signals of different polarisation components with each other, we used the following nonlinear activation function:

$$f(z_1) = \exp[\mathrm{i}(\gamma_1|z_1|^2 + \gamma_2|z_2|^2)]z_1,$$

where z_1 and z_2 are the values of the neurons of the linear layers for the first and second polarisations, respectively; and γ_1 and γ_2 are trainable parameters. It should be noted that the form of the activation function corresponds to a nonlinear step in the split-step method when solving Manakov's equations (1). Thus, symbols for different polarisations with their parallel propagation along the NN interact only 'inside' the nonlinear activation functions of hidden layers.

The presented NN was implemented using the TensorFlow 2.0 library. The initial distribution of the weight coefficients was specified by the Xavier function [GlorotNormal() function in the TensorFlow library]. To find the optimal values of the weight coefficients, we used the Adam algorithm (adaptive moment estimation). As a loss function, the root-mean-square error was chosen between the 16-QAM symbols sent by the transmitter and the symbols received at the NN output.

To assess the efficiency of the proposed scheme, the results of its operation were compared with the results of the operation of the following nonlinear compensation schemes: a linear scheme, in which the block with the NN was not used, but only the phase of the received signal was restored; schemes based on a complex NN, in which signals of both polarisations were processed independently and, therefore, their influence on each other was not taken into account in any way (in this case, the NN subnets were not connected); a complex NN, in which the symbols for two polarisations, in addition to the connection through activation functions (nonlinear layers), as in the proposed scheme, were also connected through hidden linear layers, which means that all hidden neurons of the subnetwork for one polarisation were connected by the trained parameters with each neuron hidden subnet layer for a different polarisation.

4. Results of applying the proposed scheme to compensate for nonlinear distortion

The main tasks in the study of the proposed scheme include the search for the optimal NN characteristics, which, on the one hand, would provide the greatest efficiency of its operation, and on the other, would lead to the minimum computational complexity of signal processing. Thus, one of the studied parameters was the number of processed symbols at the NN input.

Figure 3 shows the nonlinearity compensation efficiency in terms of BER as a function of the number of processed symbols at the input of each NN subnetwork. The case of independent data processing for each polarisation is compared with the cases when the subnets were connected either only through nonlinear layers, or through linear and nonlinear layers simultaneously. For each architecture under consideration, 32 neurons were used on each hidden layer. One can see that, with the exception of the case when there is only one symbol at the input, complex NNs provide significantly lower BER compared to the linear compensation scheme. In addition, for all the considered implementations of neural networks, BER decreased with an increase in the number of input neurons until it reached 31. Further changes in BER were insignificant. It can also be seen that independent data processing yielded the worst result in comparison with other NN implementations. Obviously, unconnected subnets for different polarisations are deprived of the opportunity to take into account their influence on each other. The best result is

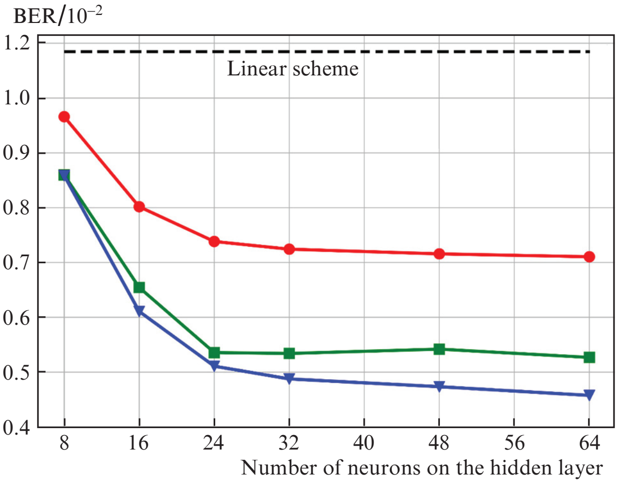

Figure 4. (Colour online) Dependences of BER on the number of neurons on each hidden NN layer for cases when symbols for both polarisations are processed independently (●), are connected only on nonlinear layers (▼) or on nonlinear and linear layers simultaneously (■).

obtained when connecting subnets through nonlinear layers. While additional connection through line layers should potentially provide a larger number of degrees of freedom, this case gives higher BER due to too many weights, which makes it impossible to effectively train the network.

Figure 4 shows the dependence of BER on the number of neurons on each hidden NN layer for the three network architectures described above. In this case, 21 symbols are supplied to the input for each polarisation. One can see that the NN efficiency increases with increasing number of neurons on the hidden layers up to 32, and a further change in BER is insignificant. In this case, an NN with connection only on nonlinear layers also shows the best efficiency of nonlinear distortion compensation, while the efficiency of a network with additional connection on linear layers is slightly inferior to it.

Figure 5 shows the dependence of BER on the input signal power for various nonlinear compensation schemes. One can see that, due to the effective compensation of the nonlinear

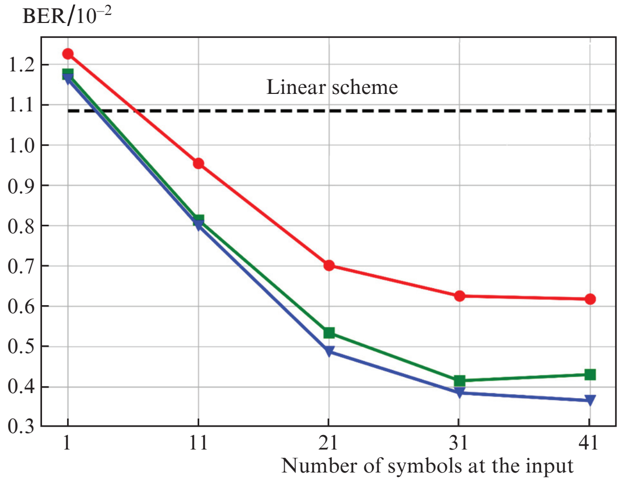

Figure 3. (Colour online) Dependences of BER on the number of symbols at the input of each NN subnetwork for the cases when the symbols for both polarisations are processed independently (●), are connected only on nonlinear layers (▼) or on nonlinear and linear layers simultaneously (■).

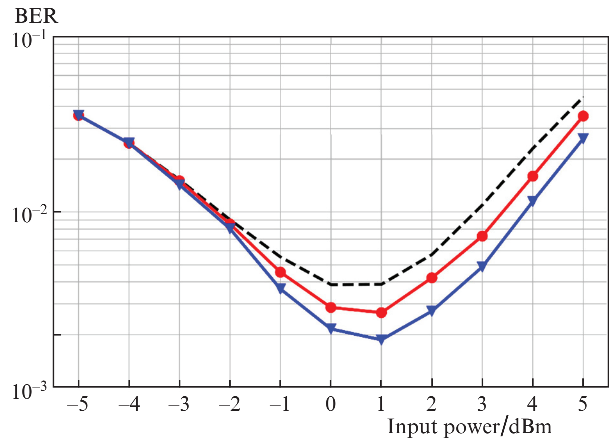

Figure 5. (Colour online) Dependences of BER on the input signal power for a linear compensation scheme (dashed curve), an NN, in which the symbols for both polarisations are processed independently (●), and an NN, in which the polarisations are connected only on nonlinear layers (▼).

interaction of signals of two polarisations, the complex NN proposed in this work makes it possible to reduce BER at an input power of 1 dBm by 52% in comparison with the linear compensation scheme and by 30% in comparison with the NN, in which the symbols for both polarisations are processed independently.

5. Conclusions

We have proposed a scheme based on a fully connected NN with complex-valued arithmetic for processing optical signals in a receiver of a communication system with polarisation division multiplexing. The activation function of the developed scheme makes it possible to take into account the nonlinear interaction of signals from different polarisation components. For this scheme, the efficiency of compensation for nonlinear effects has been investigated as a function the NN parameters: the number of processed symbols at the input and the number of neurons on hidden layers. The efficiencies of the scheme in question, a linear scheme and a complex NN, which processes signals of different polarisations independently, have been compared, and the superiority of the proposed NN architecture has been demonstrated.

Acknowledgements. The work was supported by the RF President's Grants Council (State Support to Young Russian Scientists Programme, Grant No. MK-915.2020.9). The work of S.A. Bogdanov was supported by the state assignment for fundamental research (FSUS-2020-0034). The work of A.A. Redyuk was supported by the Ministry of Science and Higher Education of the Russian Federation (Project No. FSUS-2021-0015).

References

1. Temprana E. et al. *Science*, **348**, 1445 (2015).
2. Agrawal G. *Nonlinear Fiber Optics* (Cambridge: Academic Press, 2012).
3. Zhitelev A.E. et al. *Quantum Electron.*, **47**, 1135 (2017) [*Kvantovaya Elektron.*, **47**, 1135 (2017)].
4. Sorokina M. et al. *Opt. Express*, **24**, 30433 (2016).
5. Redyuk A.A. et al. *Prikl. Fotonika*, **5**, 265 (2018).
6. Ellis A.D. et al. *Opt. Express*, **23**, 20381 (2015).
7. Liu L. et al. *J. Lightwave Technol.*, **30**, 310 (2012).
8. Burdin V.A. et al. *Quantum Electron.*, **47**, 1144 (2017) [*Kvantovaya Elektron.*, **47**, 1144 (2017)].
9. Häger C. et al. *Proc. Opt. Fiber Commun. Conf. (OFC)* (San Diego, USA, 2018) paper W3A.4.
10. Sidelnikov O. et al. *Opt. Express*, **26**, 25 (2018).
11. Bogdanov S.A. et al. *Quantum Electron.*, **51**, 459 (2021) [*Kvantovaya Elektron.*, **51**, 459 (2021)].
12. Zibar D. et al. *J. Lightwave Technol.*, **38**, 736 (2020).
13. Jones R.T. et al. *Proc. Eur. Conf. Opt. Commun. (ECOC)* (Dublin, Ireland, 2019) paper W.1.B.3.
14. Karanov B. et al. *J. Lightwave Technol.*, **36**, 4843 (2018).
15. Musumeci F. et al. *J. Lightwave Technol.*, **37**, 4125 (2019).
16. Shahkarami S. et al. *Proc. Opt. Fiber Commun. Conf. (OFC)* (San Diego, USA, 2018) paper M3A.5.
17. Jargon J.A. et al. *Opt. Express*, **18**, 4931 (2010).
18. Wang Z. et al. *Opt. Express*, **25**, 18553 (2017).
19. Meng F. et al. *Proc. Opt. Fiber Commun. Conf. (OFC)* (San Diego, USA, 2018) paper W4F.5.
20. Ruiz M. et al. *Proc. Eur. Conf. Opt. Commun. (ECOC)* (Dusseldorf, Germany, 2016) paper Th.2.P.2.
21. Redyuk A. et al. *J. Lightwave Technol.*, **38**, 1250 (2020).
22. Sidelnikov O. et al. *J. Lightwave Technol.*, **39**, 2397 (2021).
23. Deligiannidis S. et al. *J. Lightwave Technol.*, **38**, 5991 (2020).
24. Luo M. et al. *Proc. Opt. Fiber Commun. Conf. (OFC)* (San Diego, USA, 2018) paper M2F.2.
25. Reza A.G. et al. *IEEE Photonics Technol. Lett.*, **30**, 1416 (2018).
26. Jarajreh M.A. et al. *IEEE Photonics Technol. Lett.*, **27**, 387 (2014).
27. Freire P.J. et al. *J. Lightwave Technol.*, **39**, 1696 (2020).
28. Marcuse D. et al. *J. Lightwave Technol.*, **15**, 1735 (1997).