

FIBRE-OPTIC COMMUNICATION LINES https://doi.org/10.1070/QEL17518

Use of complex fully connected neural networks to compensate for nonlinear effects in fibre-optic communication lines

S.A. Bogdanov, O.S. Sidelnikov

Abstract. **A scheme is proposed for processing optical signals in a receiver of a communication system, based on complex fully connected neural networks. The influence of the main characteristics of the neural network on the efficiency of nonlinear distortion compensation is studied. A significant advantage of the proposed scheme over real-valued neural networks is demonstrated.**

Keywords: *optical fibre, nonlinear effects, fully connected neural networks, mathematical modelling.*

1. Introduction

The problem of compensating for signal distortions caused by nonlinear influences in optical fibre is one of the key issues that need to be addressed in order to further increase the throughput of modern information transmission systems [1–3]. To this end, various technologies for the generation and processing of an optical signal are currently being developed. Among such approaches, we should single out a family of methods based on the use of the Volterra series transfer function [4], digital methods based on elements of the perturbation theory [5, 6], a nonlinear Schrödinger filter and an algorithm utilising reception in general with bit-by-bit decision making [7], as well as optical methods using phase conjugation of a signal [8]. Machine learning methods, and neural networks (NNs) in particular, have now become especially actively used to compensate for nonlinearity in fibre-optic communication lines due to the fact that they provide high classification accuracy of received symbols at low computational complexity [9–12]. In addition, the schemes for processing received signals based on machine learning methods can be used in dynamically changing communication lines due to the possibility of periodic retraining.

In this paper, we propose a scheme based on complex fully connected neural networks to compensate for nonlinear effects in communication lines. The proposed scheme is compared with a scheme based on linear compensation and with a scheme based on real-valued fully connected neural networks.

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

Received 28 December 2020; revision received 27 February 2021
Kvantovaya Elektronika **51** (5) 459–462 (2021)
Translated by I.A. Ulitkin

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

Neural networks are powerful tools that can potentially be used to approximate almost any nonlinear function. However, without any prior knowledge of the function being approximated, this neural network can be quite cumbersome, and the learning process can take a lot of time. Therefore, at present, an approach is popular in which some preliminary knowledge about the nature of the problem being solved is incorporated into the NN architecture. The approach to compensate for nonlinear effects in the receiver of a communication line was first used in [10], where the NN architecture was designed by analogy with the digital back-propagation method [13]. In this paper, we propose to use this approach to develop a scheme based on fully connected neural networks for compensating for nonlinear distortions in fibre-optic communication lines.

The neural network in question (Fig. 1) consists of an input layer, in which several received symbols are processed simultaneously to take into account the channel memory effect; two hidden layers with the same number of neurons; and an output layer with one neuron corresponding to the predicted (transmitted) symbol. The structure of the input layer is as follows: to predict the symbol sent from the transmitter, use is made of the corresponding received symbol, as well as the values of its N subsequent and N previous neighbours at the receiver. Thus, $2N + 1$ symbols from the receiver are used to predict each transmitted symbol.

Since the Schrödinger equation describes the propagation of complex signals, the proposed NN is also complex-valued. Complex-valued neural networks are based on the description using complex numbers of both the state of the neurons themselves and the weight coefficients. Thus, each neuron of a complex-valued neural network is represented as a separate pair of real numbers, for which the corresponding complex-valued arithmetic has been implemented. This approach seems to be more natural when processing received symbols in fibre-optic communication lines, which are complex in nature.

Mathematically, a complex-valued neural network is equivalent to a real-valued neural network: one complex-valued neuron corresponds to a pair of real-valued ones. At the same time, to link two pairs of real-valued neurons, four real weight parameters (weights) are required, while two complex neurons are linked by one complex-valued weight (two real numbers). Thus, a real-valued NN with the number of pairs

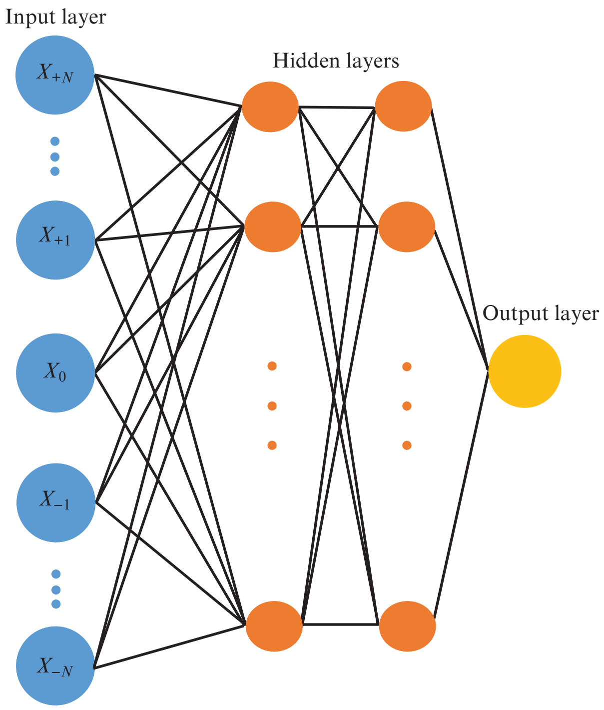

Figure 1. Scheme of a complex-valued neural network.

of real neurons equal to the number of neurons in a complex-valued network will have twice as many real weight coefficients, which will significantly complicate the learning process. However, due to the fact that one complex multiplication requires four multiplications of real numbers, the final computational complexity of both networks is the same. In addition, complex-valued neural networks allow the use of complex activation functions corresponding to nonlinear effects affecting signals during propagation through optical fibre.

The nonlinear activation function of the proposed NN corresponds to the compensation of the phase shift of the signal caused by the nonlinear interaction of signals, and has the form:

$$f(z) = e^{i\gamma_t|z|^2}z,$$

where γ_t is a parameter that is optimised during the neural network training.

The complex-valued neural network architecture was implemented using the TensorFlow 2.0 machine learning platform. To find the network weights, we used the adaptive moment estimation (Adam) optimisation algorithm, which provides an adaptive learning rate for each individual parameter. During the optimisation of the weights, the learning rate changes in accordance with the estimates of the first and second moments of the gradient. The Xavier normal distribution was used to initialise the values of the weight coefficients [the GlorotNormal() function in the TensorFlow library]. It should be noted that for each set of parameters under consideration, several runs were performed with different random initial values of all weights, and then the worst results obtained were discarded. As an error function, we used the root mean square error between the 16-QAM symbols sent from the transmitter and the symbols received after applying the NN to the training sample.

3. Mathematical modelling

The data transmission system considered in the work is schematically shown in Fig. 2. The communication line consists of a transmitter; twenty 100-km-long spans of standard single-mode fibre (SSMF); erbium optical amplifiers with a noise factor of NF = 4.5 dB, used after each span to compensate for losses; and a receiver. The transmitter generates 16-QAM signals with a symbol rate of R_s = 32 Gbaud. To shape the pulses, a root raised cosine (RRC) filter with a roll-off factor of 0.1 is used.

The propagation of signals along optical fibre is described by the nonlinear Schrödinger equation [1]:

$$\frac{\partial A}{\partial z} = -\frac{\alpha}{2}A - i\frac{\beta_2}{2}\frac{\partial^2 A}{\partial t^2} + i\gamma|A|^2 A,$$

where $A(z, t)$ is the slowly changing envelope of the optical signal; α = 0.2 dB km^{-1} is the fibre loss; $\beta_2 = -21$ ps^2 km^{-1} is the chromatic dispersion; and γ = 1.3 W^{-1} km^{-1} is the nonlinear fibre parameter. Propagation equations were solved numerically using the symmetric split-step Fourier method with a sampling rate of 16 samples per symbol.

After propagation in the channel, the received signal passed through a matched RRC filter. Next, the ideal compensation was performed for chromatic dispersion in the frequency domain. Then, nonlinear effects were compensated for using the proposed scheme based on complex fully connected neural networks. To this end, each complex symbol of the received signal was fed to a separate input node of the complex-valued NN. Finally, the signal was demodulated and the bit error rate (BER) was calculated. The proposed scheme was compared with a linear compensation scheme (a block with an NN was not used, but only the phase of the received signal was restored) and with the nonlinearity compensation method based on real-valued fully

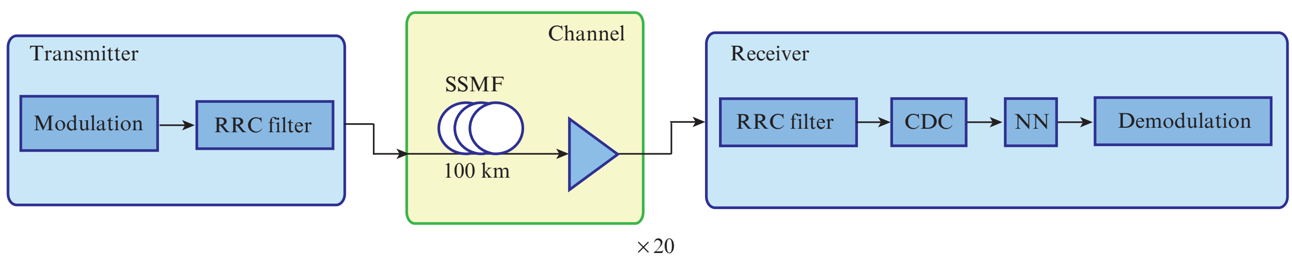

Figure 2. Schematic of a fibre-optic communication line; CDC is the chromatic dispersion compensation.

connected NNs, proposed in [11, 12]. A real-valued neural network also had two hidden layers, but, unlike a complex-valued network, it uses a hyperbolic tangent (tanh) as a nonlinear activation function. For a given neural network, the input nodes correspond to real numbers, and so each input complex symbol was divided into a pair of real numbers and fed into two different neurons at the input layer of the network; in this case, any distinction between the real and imaginary parts of the original symbol was lost. The output layer of such a network consists of two real-valued neurons, one of which corresponds to the real part of the predicted (transmitted) symbol, and the second one corresponds to the imaginary one. For a real-valued NN, the same initial weight distribution and optimisation algorithm were used as for a complex-valued network.

4. Results of applying the scheme based on complex fully connected neural networks to compensate for nonlinear effects

When examining the proposed scheme, the first step was to study the influence of the main NN characteristics on the efficiency of compensation for nonlinear distortions. Thus, the dependence of the bit error rate after applying a complex-valued NN on the number of symbols supplied to the input was investigated (Fig. 3).

It can be seen that the efficiency of the neural network application increases with increasing number of input symbols used, up to 31 symbols; then, BER changes insignificantly. This result is valid both for a complex-valued NN with 32 neurons on each of the hidden layers, and for a network with 24 neurons. From the figure, however, it follows that a neural network with 32 neurons on each hidden layer demonstrates greater efficiency with the same number of adjacent symbols used. Figure 3 also shows similar dependences for real-valued neural networks (dash-dotted curves). For correct comparison, real-valued neural networks had twice as many neurons on each hidden layer to provide computational complexity equal to that of a complex-valued neural network.

It can be seen that, except for the case when only one symbol is used at the input, complex-valued neural networks provide a significantly lower BER. In the case of using one input symbol, the degradation of the complex-valued neural network can be explained by both with the peculiarities of the error backpropagation through the implemented nonlinear activation function and by the used loss function, at which minimisation of the root mean square error does not always lead to a decrease in BER. A complex-valued neural network with one input symbol was also implemented with the tanh activation function (applied separately to the real and imaginary parts of a complex neuron); in this case, its efficiency coincided with the result for a real-valued network. Moreover, when replacing the last layer of complex- and real-valued neural networks with a classification layer (16 classes in accordance with the 16-QAM modulation format), the efficiencies of both networks coincided with the result for real-valued neural networks (see Fig. 3).

Figure 4 shows the dependences of the BER coefficient on the number of neurons on each of the hidden layers for neural networks with complex- and real-valued architectures. For real-valued neural networks (upper axis), the scale was chosen so that networks with the same computational complexity were located on the same vertical line (for example, a complex-valued neural network with 32 neurons on each of the hidden layers and a real-valued network with 64 neurons). Thus, in each case, the number of input symbols was equal to 21. It is seen that a complex-valued NN provides a lower bit error rate compared to a real-valued network under the same computational complexity. In addition, the minimum achievable bit error rate for a complex-valued network is also lower.

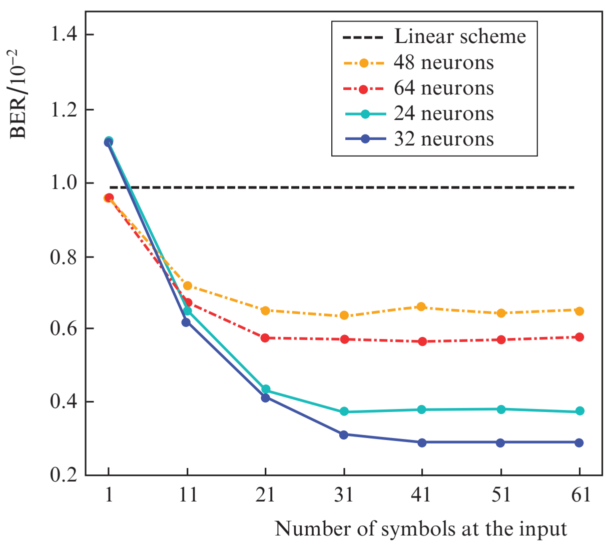

Figure 3. (Colour online) Dependences of the bit error rate on the number of symbols at the input for complex-valued networks with 32 and 24 neurons on hidden layers (solid curves), as well as for real-valued networks with 64 and 48 neurons on each hidden layer (dash-dotted curves).

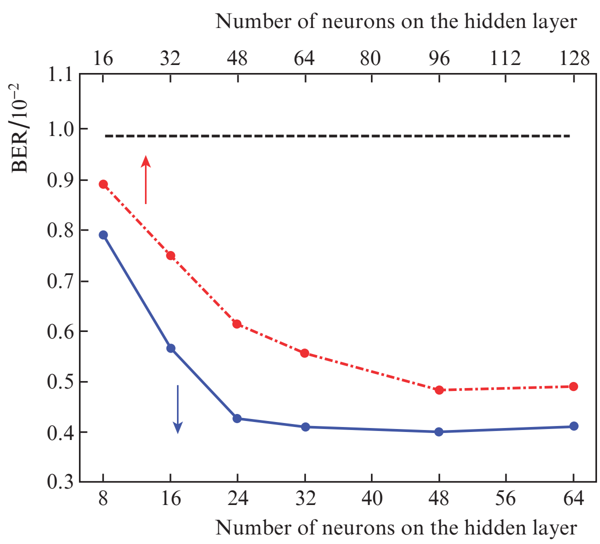

Figure 4. Dependences of the bit error rate on the number of neurons on each hidden layer for complex- (solid curve) and real-valued (dash-dotted curve) neural networks (for a real-valued network, the number of neurons was doubled; in both cases, 21 symbols were used at the input). The dashed line corresponds to the linear compensation scheme.

It should also be noted that there is some trade-off between the efficiency of a neural network and its computational complexity. This can be seen, for example, from Fig. 3. Thus, receiving 31 symbols at input, a complex-valued neural network with 24 neurons on each hidden layer provides a lower (by about 8%) efficiency in terms of BER than a neural network with 32 neurons, but has less computational complexity. This is demonstrated even more clearly in Fig. 4, where, for a fixed number of input symbols, an increase in the efficiency of the neural network is observed with an increase in the number of neurons on hidden layers, which corresponds to an increase in computational complexity.

Figure 5 shows the dependences of the bit error rate on the initial signal power for various nonlinear compensation schemes. It can be seen that, due to the effective compensation of nonlinear effects, the use of a complex-valued NN makes it possible to reduce the BER at an optimal power by 53% compared to the linear compensation scheme and by 40% compared to the real-valued NN. It should be noted that both considered neural networks had the same computational complexity.

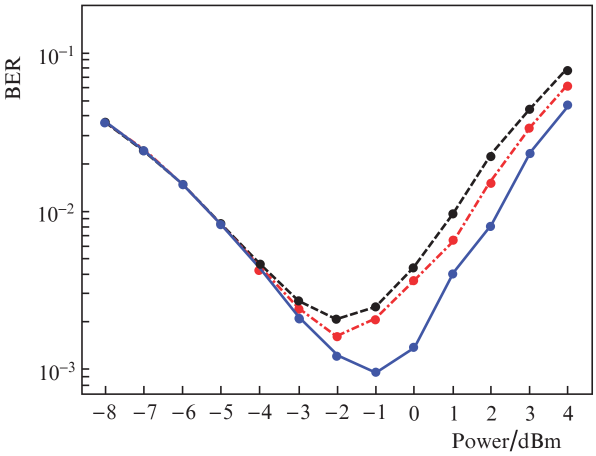

Figure 5. Dependences of the bit error rate on the initial signal power for a linear compensation scheme (dashed line), complex- (solid line) and real-valued (dash-dotted line) neural networks (64 and 32 neurons were used on each hidden layer for a real- and complex-valued network, respectively; at input there were 31 symbols).

5. Conclusions

The paper proposes a complex-valued fully connected neural network-based scheme for processing optical signals in the receiver of a communication system. This scheme is employed to study the efficiency of compensation for nonlinear effects, depending on the parameters of the neural network, i.e. the number of processed symbols at the input and the number of neurons on hidden layers. The efficiency of signal processing of the proposed scheme is compared with that of a scheme based on real-valued neural networks, and a significant advantage of complex-valued neural networks is demonstrated.

Acknowledgements. This work was supported by the RF President's Grants Council (Grant No. MK-915.2020.9).

References

1. Agrawal G. *Nonlinear Fiber Optics* (Cambridge: Academic Press, 2012).
2. Temprana E. et al. *Science*, **348**, 1445 (2015).
3. Zhitelev A.E. et al. *Quantum Electron.*, **47**, 1135 (2017) [*Kvantovaya Elektron.*, **47**, 1135 (2017)].
4. Liu L. et al. *J. Lightwave Technol.*, **30**, 310 (2012).
5. Sorokina M. et al. *Opt. Express*, **24**, 30433 (2016).
6. Redyuk A.A. et al. *Prikl. Fotonika*, **5**, 265 (2018).
7. Burdin V.A. et al. *Quantum Electron.*, **47**, 1144 (2017) [*Kvantovaya Elektron.*, **47**, 1144 (2017)].
8. Ellis A.D. et al. *Opt. Express*, **23**, 20381 (2015).
9. Wang D. et al. *IEEE Photonics Technol. Lett.*, **28**, 19 (2016).
10. Häger C. et al. *Proc. Opt. Fiber Commun. Conf. (OFC)* (San Diego, USA, 2018) paper W3A.4.
11. Sidelnikov O.S. et al. *Quantum Electron.*, **47**, 1147 (2017) [*Kvantovaya Elektron.*, **47**, 1147 (2017)].
12. Sidelnikov O. et al. *Opt. Express*, **26**, 25 (2018).
13. Ip E. *J. Lightwave Technol.*, **28**, 6 (2010).