

Neural network for calculating direct and inverse nonlinear Fourier transform

E.V. Sedov, I.S. Chekhovskoy, J.E. Prilepsky

Abstract. A neural network architecture is proposed that allows a continuous nonlinear spectrum of optical signals to be predicted and an inverse nonlinear Fourier transform (NFT) to be performed for signal modulation. The average value of the relative error in predicting the continuous spectrum by the neural network when calculating the direct NFT is found to be 2.68×10^{-3} , and the average value of the relative error in predicting the signal for the inverse NFT is 1.62×10^{-4} .

Keywords: nonlinear Schrödinger equation, inverse scattering problem method, Zakharov–Shabat problem, nonlinear Fourier transform, neural networks, machine learning.

1. Introduction

Among various factors limiting the performance of modern optical communication systems, optical channel nonlinearity is considered to be the one that most seriously degrades the characteristics of existing optical communication lines [1]. In this regard, methods aimed at increasing the throughput of communication lines are being actively investigated [2, 3]. Of many alternative methods, signal processing based on the nonlinear Fourier transform (NFT) and optical transmission methods operating with so-called nonlinear Fourier (NF) modes have recently attracted much attention. The nonlinear Fourier transform provides efficient linearization of signal evolution within a single-mode fibre model, where the propagation of light is well approximated by the nonlinear Schrödinger equation (NLSE). In the case of distributed Raman amplification or when the optical channel is approximated by a path-averaged model, the NLSE describes well the evolution of the signal along optical fibre. Problems in the practical implementation of NFT-based transmission methods are associated with deviations of the true optical channel from the NLSE model and with the sensitivity of the system to such deviations [4, 5], as well as with the presence of optical noise [6–9]. This is explained by the fact that any electronic element has its own characteristics that do not coincide with the ideal ones, and the properties of optical fibre may vary

depending on external conditions and differ from the factory characteristics. Note also that NFT can be used to analyse optical signals [10, 11] and has recently been used to identify soliton components in various photonic applications, including lasers and microcavities [12–16].

The direct and inverse NFT as applied to the focusing NLSE was first formulated in the fundamental work of Zakharov and Shabat [17]. Direct NFT associates an optical signal with its nonlinear spectrum, which in the general case may consist of discrete and continuous parts [18]; however, each part may be absent for some specific situations. The set of discrete eigenvalues corresponds to the soliton (discrete) part of the NFT spectrum of the signal, which usually exists at sufficiently high signal powers [11]. Nevertheless, the use of the continuous part of the NFT spectrum, corresponding to dispersion components, turned out to be very effective in optical data transmission [4, 9, 18, 19].

Currently, numerous ‘traditional’ numerical approaches have been proposed for calculating the nonlinear spectrum [20–23], and significant progress has been made in reducing the computational complexity of algorithms [24], as well as in increasing their accuracy [25–27]. However, when applied to complex signals with a large number of nonlinear components, difficulties can often arise associated with the stability [28]. In addition, real-time processing of NFT-based, complex signals is a challenge, which limits the ability to efficiently implement NFTs in hardware. In this case, a promising area is the use of machine learning (ML) and, in particular, the implementation of NFTs based on neural networks (NNs).

In recent years, a breakthrough has been made in the development of ML methods for solving algorithmically complex problems, such as, for example, pattern recognition and classification [29, 30]. The main steps here are training a model on a set of some data array and applying the predictive models. The first step may take longer, but the trained model is usually much faster to apply, allowing machine learning systems to be deployed on a variety of nonproductive devices. Note that ML-related methods have been successfully applied to compensate for nonlinear effects [31–34].

In recent years, ML has also been proposed to be used in NFT-based data transmission systems, but mainly for calculating soliton components (discrete spectrum) [35–37]. For systems based on a continuous spectrum, NNs were applied at the post-processing stage [38, 39]. In a recent paper [40], a simple NN was used to process a continuous nonlinear spectrum. In this work, the procedure from the MATLAB package for recognising handwritten numbers was adapted to classify points of the signal constellation, which imposes significant restrictions on the use of this approach. In this paper, we propose to implement a more advanced regression approach:

E.V. Sedov, I.S. Chekhovskoy Novosibirsk State University, ul. Pirogova 2, 630090 Novosibirsk, Russia;
e-mail: e.sedov@g.nsu.ru, i.s.chekhovskoy@nsu.ru;
J.E. Prilepsky Aston Institute of Photonic Technologies, Aston University, Birmingham B4 7ET, UK;
e-mail: y.prylepskiy1@aston.ac.uk

Received 26 October 2021
Kvantovaya Elektronika 51 (12) 1118–1121 (2021)
Translated by I.A. Ulitkin

to compute a continuous spectrum using a special NN, and also to reconstruct the original optical field (that is, to perform an inverse NFT) using a special NN architecture.

2. Nonlinear Fourier transform

The propagation of light in optical fibre in the slowly varying envelope approximation is described using NLSE, which is written in dimensionless form as

$$i \frac{\partial q}{\partial z} + \frac{1}{2} \frac{\partial^2 q}{\partial t^2} + |q|^2 q = 0. \quad (1)$$

Here $q(z, t)$ is the slowly varying optical field in fibre; z is the distance along fibre; and t is the time in the frame of reference moving with the group velocity of the wave packet. To simplify the analysis, we do not take into account the fibre gain and loss, as well as the presence of noise components [4, 41].

Equation (1) in the presented form belongs to the class of integrable equations that can be solved by the method of the inverse scattering problem. Direct NFT allows one to determine the scattering data and consists in solving the Zakharov–Shabat spectral problem using a localised ‘potential’ $q(z, t)$, which is an optical signal:

$$-\partial_t \psi_1 + q(z, t) \psi_2 = i \lambda \psi_1, \quad (2)$$

$$\partial_t \psi_2 + q^*(z, t) \psi_1 = i \lambda \psi_2.$$

Here ψ_i are auxiliary functions, and the complex parameter $\lambda = \xi + i\eta$ is a nonlinear analogue of frequency. To determine the nonlinear spectrum associated with the profile $q = q(z, t)$, it is necessary to find a special solution $\Phi(t, \lambda) = [\phi_1, \phi_2]$ of system (2), satisfying the asymptotic condition: $\Phi \rightarrow [e^{-i\xi t}, 0]$ at $t \rightarrow -\infty$. The main part of the direct NFT consists in calculating the scattering coefficients $a(\lambda)$ and $b(\lambda)$ determined using a special solution $\Phi(t, \lambda)$ as follows:

$$a(\xi) = \lim_{t \rightarrow +\infty} \phi_1(t, \xi) e^{i\xi t}, \quad b(\xi) = \lim_{t \rightarrow +\infty} \phi_2(t, \xi) e^{-i\xi t}.$$

The continuous spectrum $r(\xi) = b(\xi)/a(\xi)$ coincides with the real axis of the ξ -plane and corresponds to the dispersive wave component of the signal, while it is a direct analogue of the linear Fourier transform and converges to it at low signal powers. The discrete part of the nonlinear spectrum is not considered in our work, and below we deal only with the continuous part of the spectrum. The signals were preselected in such a way that the corresponding nonlinear spectrum did not contain discrete levels, that is, solitons. But we also found that the nonlinear effects are significant and that the continuous spectrum differs markedly from its linear Fourier spectrum.

3. Advantages of using NNs in nonlinear signal processing

Since linear, nonlinear and noise effects manifest themselves simultaneously when transmitting data over fibre-optic communication systems, these data are very suitable objects for processing using the latest advances in ML methods. Using ML, we can solve a multidimensional optimisation problem (for example, in terms of transmission quality and throughput maximisation) without manipulation of all possible parameter values. An especially urgent task is the identifica-

tion of some internal features and regularities of data. Here, NNs can be used to simulate propagation effects when a signal travels through a noisy nonlinear fibre environment. In other words, using NNs, we can simulate a nonlinear transform without computing the transform directly. The gain lies in the speed and versatility of the transform, as well as the flexibility and adaptability of NN-based operations: The neural network does not know what data it is processing; it searches for the necessary features in the data that affect the final result, and then extracts them. This process is called feature extraction. Thus, if we want to calculate a certain value of some function, then instead of (possibly) complex calculations, we can use a pre-trained NN.

Another advantage of using NN-based signal processing is that NNs can reduce noise in the data. In practice, we always have additional noise that can be critical for idealised methods. The network can filter out unnecessary information within itself, leaving only the basic features necessary for a specific task. A possible disadvantage of using NNs is their presently achievable accuracy. However, in practice, the accuracy of NNs is sufficient for most problems and, even with a sufficient set of data for training, exceeds the accuracy of existing numerical methods. As noted earlier, NNs have already demonstrated their potential in problems of processing optical signals [42] and, in particular, in NFT-based systems [35–39].

4. Results

In this paper, we use an NN to predict the nonlinear continuous spectrum of complex optical signals and convert the spectrum back to a signal (inverse NFT). For this study, we have selected examples of signals in the form of a widely used wavelength division multiplexing (WDM) format. One WDM signal is represented as the sum of independent optical carriers:

$$s(t) = \frac{1}{Q} \sum_{k=1}^M C_k e^{i\omega_k t} f(t), \quad 0 \leq t < T, \quad (3)$$

where Q is the normalisation factor that is used to control the magnitude of nonlinear effects; M is the number of WDM channels; ω_k is the carrier frequency of the k th channel; C_k are the data from the signal constellation, transmitted over the k th channel; and T is a character interval. The function $f(t)$ is the shape of the pulse, which in this work, without loss of generality, is expressed (in normalised form) as

$$f(t) = 1 - \cos \frac{4\pi t}{T} \text{ at } 0 \leq t \leq T/4 \text{ or } 3T/4 \leq t \leq T,$$

$$f(t) = 1 \text{ at } T/4 < t < 3T/4.$$

To assess the quality of the NN prediction, we use the formula

$$\eta_r(\xi) = \frac{|r_{\text{pred}}(\xi) - r_{\text{real}}(\xi)|}{\langle |r_{\text{real}}(\xi)| \rangle_\xi} \quad (4)$$

to determine the relative error for a continuous spectrum. Here the symbol $\langle \dots \rangle_\xi$ denotes averaging over the spectral interval; and the subscripts pred and real refer to the reflectance values $r(\xi)$ predicted by NNs and to the values previously calculated by standard numerical methods, respectively.

A similar formula for calculating the relative error in predicting the signal itself has the form

$$\eta_q(t) = \frac{|q_{\text{pred}}(t) - q_{\text{real}}(t)|}{\langle |q_{\text{real}}(t)| \rangle_t}, \quad (5)$$

where $q(t)$ is the signal, and the symbol $\langle \dots \rangle_t$ denotes averaging over the time interval. The relative errors $\eta_r(\xi)$ and $\eta_q(t)$ are determined at points ξ and t , respectively. We used the quantities $\langle \eta_r(\xi) \rangle_\xi$ and $\langle \eta_q(t) \rangle_t$ to estimate the overall mean of the error. We emphasise that the metric was chosen in such a way as to take into account even those areas where the value of the spectrum or signal is much less than unity.

Figure 1 shows the architecture of the NFT network, which performs NFT operations (direct and inverse transforms) with the parameters indicated in the figure. The NFT network consists of sequential convolution layers and fully connected output layers. A complex signal consisting of 1024 points arrives at the input to the network. This NN predicts only one component of the continuous NF spectrum, and so two identical NFT networks must be used to predict the real and imaginary parts of $r(\xi)$. Likewise, converting the spectrum back to a signal requires two separate NFT networks for the real and imaginary parts of the $q(t)$ signal. Each of the four neural networks with the same architecture was trained independently.

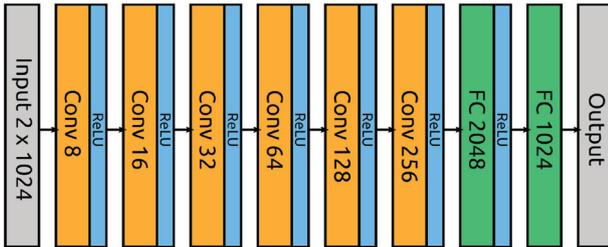


Figure 1. (Colour online) Neural network architecture for predicting the continuous nonlinear component of the spectrum of a complex signal; the same architecture was used for inverse NFT, converting the continuous spectrum back into a complex signal. To predict the real and imaginary parts of the spectrum (or signal), use is made of two separate NNs with the same structures, but with different weights.

The dataset consisted of 94035 signals, of which 9403 were used for validation and were not involved in the training process. Random data sequences encoded in quadrature phase shift keying (QPSK) format were used to generate the signals. The energy of all signals was the same and was chosen at such a level that the nonlinear effects were strong enough. At the selected energy, some signals contained a discrete spectrum, but such signals did not fall into the training dataset. The continuous spectrum for each signal was pre-computed using conventional direct NFT techniques. When training the NFT network, we used the mean square error (MSE) as a loss function, as well as the Adam (adaptive moment estimation) optimisation algorithm with a training step of 10^{-4} . On average, given the amount of data used, the learning process took 50000 epochs.

Figure 2 shows the error $\eta_r(\xi)$, that is, the difference between the predicted and actual (pre-calculated using the conventional numerical NFT method) continuous nonlinear spectrum for a specific signal, and Fig. 3 demonstrates an example of the error $\eta_q(t)$ in finding the inverse NFT. One can

see that the NN performs both direct and inverse transforms with high accuracy. A slight increase in the error at the centre for the continuous spectrum is associated with its localisation in the middle, while at the edges of the spectral interval the values of $r(\xi)$ tend to zero. There is no such feature for the signal, since it is located evenly over the entire time interval. The average value of the relative prediction error of the continuous spectrum $\langle \eta_r(\xi) \rangle_\xi$ for the NN when calculating the direct NFT is 2.7×10^{-3} . For the inverse transform, the average value of the relative signal prediction error $\langle \eta_q(t) \rangle_t = 1.6 \times 10^{-4}$. The results obtained demonstrate that neural networks can perform direct and inverse nonlinear Fourier transforms with high accuracy. To date, the practical application of NFT methods is far from commercial implementation but, however, already in laboratory conditions, their potential for transferring information at record net data rate has been demonstrated [43].

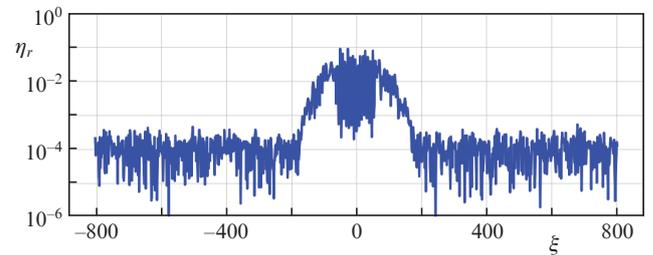


Figure 2. (Colour online) Example of the distribution of the relative error $\eta_r(\xi)$ between the pre-calculated and predicted continuous spectrum.

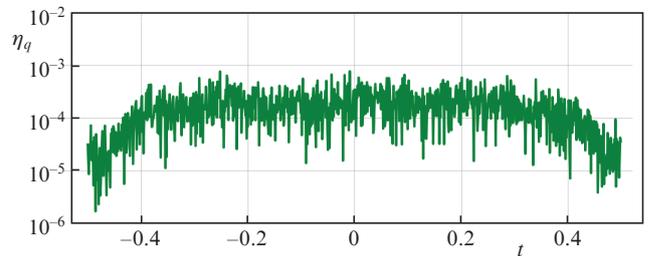


Figure 3. (Colour online) Example of the distribution of the relative error $\eta_q(t)$ between the initial and predicted signals.

5. Conclusions

Presently, ML and NN are modern technologies that are actively being investigated in applications to nonlinear signal processing and optical communication. The proposed NN architecture demonstrates the fundamental application of the NN for the analysis and (de)modulation of complex optical signals used in communications. This opens up prospects for improving existing systems without the need for a deep understanding of the internal nonlinear processes that affect the signal transmission quality. We emphasise that the method proposed in this work is only the first step in the development of methods for machine processing of optical signals. It can be used to design smart receivers with digital backpropagation algorithms based on NFTs and NNs. Our results show that the use of NNs can make it possible to study not only the internal structure, but also generate new signals using autoencoders. The fundamental possibility of using NNs for NFT can create new areas for research related

to the analysis of the nonlinear structure of the signal and its evolution characteristics.

Acknowledgements. The study was supported by the RF President's Grants Council (State Support to Young Russian Scientists Programme, Grant No. MK-677.2020.9). The work of I.S. Chekhovskoy was supported by the state assignment for fundamental research FSUS-2020-0034. The work of J.E. Prilepsky was supported by the Leverhulme Trust (Project RPG-2018-063).

References

1. Essiambre R.J. et al. *Phys. Rev. Lett.*, **101** (16), 163901 (2008).
2. Winzer P.J. et al. *Opt. Express*, **26** (18), 24190 (2018).
3. Cartledge J.C. et al. *Opt. Express*, **25** (3), 1916 (2017).
4. Le S.T. et al. *Opt. Express*, **23**, 8317 (2015).
5. Yangzhang X. et al. *J. Lightwave Technol.*, **36** (2), 485 (2018).
6. Derevyanko S.A. et al. *Nat. Commun.*, **7**, 307 (2016).
7. Civelli S. et al. *IEEE Photonics Technol. Lett.*, **29** (16), 1332 (2017).
8. Civelli S. et al. *Appl. Sci.*, **10** (24), 9099 (2020).
9. Derevyanko S. et al. *Opt. Express*, **29** (5), 6384 (2021).
10. Sedov E.V. et al. *Opt. Lett.*, **43** (24), 5985 (2018).
11. Turitsyn S.K. et al. *J. Lightwave Technol.*, **38** (2), 352 (2020).
12. Sugavanam S. et al. *Nat. Commun.*, **10** (1), 5663 (2019).
13. Ryczkowski P. et al. *Nat. Photonics*, **12** (4), 221 (2018).
14. Turitsyn S.K. et al. *Opt. Lett.*, **45** (11), 3059 (2020).
15. Wang J. et al. *Chin. Phys. B*, **29** (3), 034207 (2020).
16. Chekhovskoy I.S. et al. *Phys. Rev. Lett.*, **122** (15), 153901 (2019).
17. Zakharov V.E. et al. *Sov. Phys. JETP*, **34** (1), 62 (1972) [*Zh. Eksp. Teor. Fiz.*, **61** (1), 118 (1971)].
18. Turitsyn S.K. et al. *Optica*, **4** (3), 307 (2017).
19. Yousefi M. et al. *IEEE Trans. Inf. Theory*, **66** (1), 478 (2019).
20. Delves L.M., Lyness J.N. *Math. Comp.*, **21**, 543 (1967).
21. Boffetta G., Osborne A.R. *J. Comput. Phys.*, **102** (2), 252 (1992).
22. Burtsev S. et al. *J. Comput. Phys.*, **147** (1), 166 (1998).
23. Vasylichenkova A. et al. *Opt. Lett.*, **43** (15), 3690 (2018).
24. Wahls S., Poor H.V., in *Proc. Int. Conf. Acoustics, Speech and Signal Processing (ICASSP 2013)* (IEEE, 2013) pp 5780–5784.
25. Mullyadzhyanov R., Gelash A. *Opt. Lett.*, **44** (21), 5298 (2019).
26. Medvedev S.B. et al. *Opt. Lett.*, **45** (7), 2082 (2020).
27. Medvedev S.B. et al. *Opt. Express*, **28** (1), 20 (2020).
28. Gelash A., Mullyadzhyanov R. *Phys. Rev. E*, **101** (5), 052206 (2020).
29. Bishop C.M. *Pattern Recognition and Machine Learning* (Springer, 2006).
30. Duda R.O. et al. *Pattern Classification* (John Wiley & Sons, 2012).
31. Zibar D. et al. *Opt. Express*, **20** (26), B181 (2012).
32. Zibar D. et al. *J. Lightwave Technol.*, **34** (6), 1442 (2015).
33. Sidelnikov O.S., Redyuk A.A., Sigletos S., Fedoruk M.P. *Quantum Electron.*, **49** (12), 1154 (2019) [*Kvantovaya Elektron.*, **49** (12), 1154 (2019)].
34. Sidelnikov O.S. et al. *Opt. Express*, **26** (25), 32765 (2018).
35. Jones R.T., Gaiarin S., Yankov M.P., Zibar D., in *Optical Fiber Communication Conference, OSA Technical Digest (online)* (Optical Society of America, 2018) paper W2A.59.
36. Jones R.T. et al. *IEEE Photonics Technol. Lett.*, **30** (12), 1079 (2018).
37. Yamamoto S. et al. *IEICE Communications Express*, **8** (12), 507 (2019).
38. Kotlyar O. et al. *Opt. Lett.*, **45** (13), 3462 (2020).
39. Kotlyar O. et al. *Opt. Express*, **29** (7), 11254 (2021).
40. Zhang W.Q. et al. *Opt. Express*, **29** (8), 11591 (2021).
41. Kamalian M. et al. *J. Lightwave Technol.*, **35** (24), 5464 (2017).
42. Freire P.J. et al. arXiv preprint arXiv:2103.08212, 2021.
43. Yangzhang X. et al. *J. Lightwave Technol.*, **37** (6), 1570 (2019).