Neural Networks based Equalization of Experimental Transmission using the Nonlinear Fourier Transformation

Jonas Koch(1), Ken Chan(2), Sebastian Kühl(1), Christian G. Schaeffer(2), Stephan Pachnicke(1)

(1) Chair of Communications, Kiel University, Kiel, jonas.koch@tf.uni-kiel.de
(2) Chair of High Frequency Technology, Helmut Schmidt University, Hamburg

Abstract *We use cross-correlations between continuous and discrete parts of the nonlinear Fourier spectrum to train an equalizer based on neural networks. In a dual-polarization transmission experiment reach improvements of more than 2000 km and up to beyond 5000 km were achieved.*

Introduction

For exploiting the full capacity of the fiber optic communication channel, high signal powers have to be used. Nonlinear impairments caused by the inherent Kerr nonlinearity of the fiber limit the spectral efficiency. With the approach of the Nonlinear Fourier Transformation (NFT) a linearization of the lossless Nonlinear Schrödinger Equation (NLSE) in the case of single polarization transmission[1] can be achieved. If both polarizations are to be considered, the Manakov equations[2] need to be employed. Using NFT a signal can be back propagated by a linear transmission function without nonlinear interference[1]. In contrast to the conventional Fourier Transformation, the NFT consists of two spectra, namely the continuous spectrum $\hat{q}(\lambda_c) = \frac{b(\lambda_c)}{a(\lambda_c)}, \lambda_c \in \mathbb{R}$ and the discrete spectrum $\tilde{q}(\lambda_j) = b(\lambda_j) / \frac{\mathrm{d}a(\lambda_j)}{\mathrm{d}\lambda}|_{\lambda_j}, \lambda_j \in \mathbb{C}^+, \quad j \in \mathbb{N}$.The eigenvalues λ do not change during ideal transmission. b and a are the so-called NFT coefficients, which are used for data modulation and propagate linearly. The finite number of eigenvalues λ_j in the discrete spectrum represents solitons, while the continuous spectrum describes dispersive parts of the wave. If the NFT is expanded to two polarizations, a b-coefficient for both polarizations can be used. For the discrete spectrum the two polarizations have to be modulated jointly, since each $b_k(\lambda_j), k \in \{1,2\}$ is related to the same λ_j.

One drawback of the NFT is that due to the impact of loss along the link, the NLSE becomes non-integrable. Therefore, the lossless path averaged (LPA) method is used[3]. However, loss and noise lead to variations of the eigenvalues and consequently their corresponding coefficients during transmission, which disturbs the orthogonality of the system. Furthermore, also due to numeric errors parts of the solitonic signal-power shift into the dispersive continuous spectrum. The resulting cross-correlations between eigenvalues and scattering coefficients of the discrete spectrum have been investigated in multiple publications[4],[5]. If the full nonlinear spectrum (NFDM) is used cross-correlations between the discrete and the continuous spectrum occur[6].

In our system only the discrete spectrum is used, i.e. the continuous part is zero at the Tx. As soon as there is more than infinitesimal loss, the equilibrium between linear and nonlinear influences is disturbed and a bleeding of energy into the continuous spectrum takes place. For the first time of our knowledge we propose to use these cross-correlations to equalize the coefficients $b_k(\lambda_j)$ of the discrete spectrum. The signals energy which crossed over from the discrete- into the continuous spectrum can be used by a novel feed forward deep neural network (NN) equalizer, to vastly improve the detection performance in experiments. A well optimized NN is able to use the identified deviations of the discrete eigenvalues to lessen the impact of noise and other disturbances (such as loss or Tx/Rx imbalances) on the information carrying NFT coefficients $b(\lambda_j)$[7]. Here, this is extended to the continuous spectrum to further improve upon the performance of the solely discrete NN equalizer.

Experimental Setup

The experimental data and setup is based on[8] and is depicted in Fig. 1. A dual polarization transmission using two eigenvalues was chosen. The main hardware components and NFT parameters are summarized in Tab. 1. Blocks of 2718 solitons were modulated differentially[9] and transmitted. Additionally, 256 solitons were used to recover the polarization. The frequency responses of transmitter and receiver were compensated. Furthermore, the phase of $b_k(\lambda_2)$ was rotated by $\pi/4$, which leads to favouring solitons in terms of PAPR, shape and bandwidth[10]. In total two eigenvalues were QPSK modulated on two polarizations and transmitted at a rate of 1 GBd, resulting in a total data rate of 8 Gb/s, including FEC-overhead.

Cross-Correlations of the NFT Spectra

As described above theoretical orthogonality of different NFT carriers of the continuous and the

Tab. 1: Experimental Settings

DAC sampling rate	88 GS/s
ADC sampling rate	40 GS/s
Baud rate	1 GBd
Carrier laser wavelength	1550.12 nm
Carrier laser linewidth	<1 kHz
Normalization time T_0	45.45 ps
Discrete eigenvalues λ_i	$\mp$0.15+0.3i
Cont. eigenvalues λ_c	-4 ... 4
Number of λ_c	161
Modulation format	QPSK
Amplitudes $\lvert b(\lambda_k)\rvert$	[0.5 2.5]
LPA launch power	-5 dBm
Span length	50.3 km
Effective nonlinearity γ_{eff}	0.96 W^{-1}km^{-1}
Fiber attenuation α	0.2 dB/km
Chromatic dispersion β_2	-5.75 ps^2/km

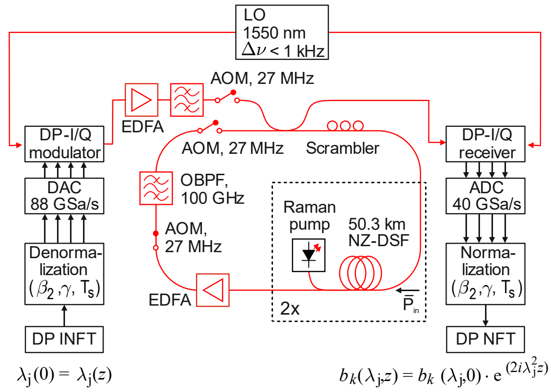

Fig. 1: Experimental Setup.

discrete spectrum does not hold, if hardware imperfections, attenuation and noise are considered. For this work, only the discrete spectrum was modulated. However, a continuous spectrum can still be computed at the receiver side. This spectrum is mainly filled by noise. In addition, parts of the solitonic spectrum can shift into the continuous spectrum, which can be seen

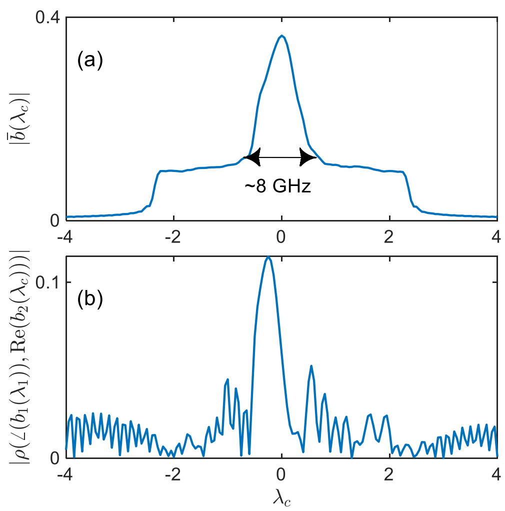

Fig. 2: (a) Averaged absolute received continuous spectrum after 3018 km transmission. (b) Absolute linear pearson correlation between the phase deviation of $b_1(\lambda_1)$ and the real part of the continuous spectrum $\hat{q}_2(\lambda_c)$.

in Fig. 2(a). Here a peak of the continuous spectrum can be observed at the continuous eigenvalues, which are near to $\text{Re}\{\lambda_j\}$. Furthermore, the correlation between the deviation of the phase of $b_1(\lambda_1)$ (on the X-polarization) and the continuous spectrum of the Y-polarization $(q_2(\lambda))$ is depicted in Fig. 2(b). Note that a correlation of the deviations of discrete and continuous spectra of different polarizations can be seen, especially for continuous eigenvalues close to the real-part of λ_1.

Neural Networks

The identified deviations of e.g. λ_j can be used to equalize the information carrying coefficients $b(\lambda_j)$ by taking advantage of the aforementioned correlations. This can be implemented by using a simple linear MMSE equalizer[4] or neural networks[7]. Feed forward neural networks have been shown to work for the equalization and classification of a solitonic NFT transmission[7],[11],[12]. In this work, in addition to the deviations of λ_j the received continuous spectrum is used by a neural network to equalize the discrete spectrum. In order to employ a NN for equalization, the problem was formulated as a classification task, mapping the features calculated by the NFT $\mathbf{n}_{\text{NN}} = [b_{1,2}(\lambda_{1,2})\ \Delta\text{d}a(\lambda_{1,2})\ \Delta\lambda_{1,2}\ b_{1,2}(\lambda_c)]$ to the class of the corresponding transmitted bits. This results in a total of four complex valued received discrete NFT coefficients of the discrete eigenvalues, two deviations of the known transmitted NFT coefficient $\Delta\text{d}a(\lambda_{1,2})$ and two deviations of the eigenvalues $\lambda_{1,2}$ themselves. In addition, the feature vector is filled with the continuous spectrum's NFT parameter $b_{1,2}(\lambda_c)$ of both polarizations. As aforementioned the output layer consisted of $M = 256$ nodes which represent one of the 2^8 possible transmitted bit combinations using the soft-max

$$g(x_n) = \frac{\exp(x_n)}{\sum_{m=1}^{256} \exp(x_m)} \qquad (1)$$

as activation function for the output layer. Here, x_n is the weighted sum of the outputs of the last hidden layer. Hence, the output vector of the NN can be understood as a discrete probability distribution for each class. The symbol referring to the class with $\max_{1\le n\le 256} g(x_n)$ is interpreted as estimated symbol. Furthermore 2 hidden layers with 132 and 596 nodes determined by a random search algorithm were employed respectively using the Rectified linear unit (ReLU) activation function. The training of the NN was based on minimizing the cross-entropy loss between the one-hot encoded class of the transmitted bit and the output vector by using Adam's algorithm[13].

Results and Discussion

In order to determine the training complexity of training needed for this network the amount of training symbols and training episodes have been varied. The resulting BERs are depicted in Fig. 3. From this it can be deduced that the amount of episodes needed is not depending on the amount of used training symbols. Furthermore, after 500 episodes the gain of using more episodes is small.

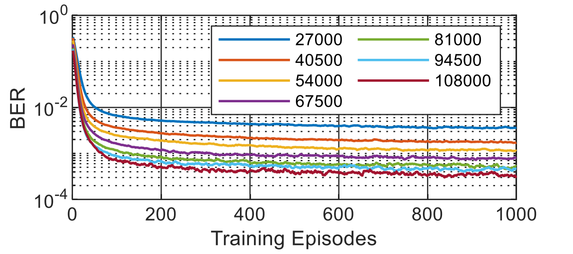

Fig. 3: BER for 4828 km transmission depending on the number of Training Episodes for different amounts of training symbols, denoted by the legend.

Fig. 4 shows the resulting BER depending on the amount of used training symbols and different transmission distances. Here, one can see that for relatively shorter transmission distances (here: 4024 km) there is no gain of increasing the amount of training symbols beyond 90,000, whereas for longer distances increasing this number can lead to further improved BERs.

To show the gain of using the cross-correlations of continuous and discrete spectra in NN equalizers, Fig. 5 depicts the BER depending on the transmission distance. Here, the BER without equalization, a simple LMMSE equalizer, an NN equalizer which takes only the discrete spectrum into account and an NN equalizer which takes all cross-correlations into account are compared.

If no equalizer is used, transmissions up to 3018 km are possible with BERs below the HD-FEC limit of 3.8e-3. The LMMSE equalizer used in this case was employed on each eigenvalue individually. Hence, no cross-correlations were taken into account. This increased the achievable reach by 400 km up to 3400 km. To take full advantage of the cross-correlations between

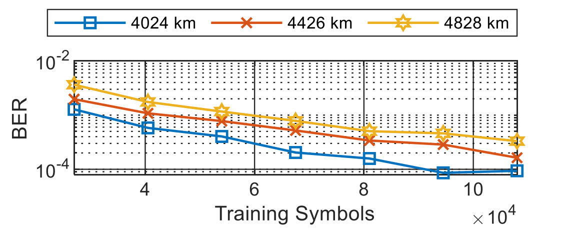

Fig. 4: BER depending on the amount of used training symbols for different transmission distances.

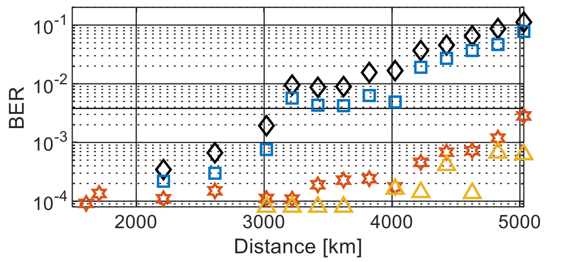

Fig. 5: BER depending on the transmission distance using different equalizers. No equalization (black, diamond), LMMSE (blue,square) using 2000 training symbols, discrete NN (red, star) using 45,000 training symbols and discrete+continuous NN (yellow, triangle) using 67,500 training symbols. For readability the HD-FEC limit is included (black line).

eigenvalues and NFT-coefficients of the discrete spectrum, an NN equalizer was built, which equalized all discrete eigenvalues of both polarizations jointly[14]. This leads to large improvements of achievable transmission distance of up to 5000 km (below the HD-FEC limit). Although, this equalizer shows an error floor at a BER of around 1e-4. This can lead to higher BERs for short transmission distances, when compared to the LMMSE or a transmission without any equalizer. If the cross-correlations of the continuous spectrum are taken into account, the NN equalizer is able to further reduce the BER for high transmission distances (here: BER of 6.3e-4 for 5030 km). Additionally, this equalizer has no error floor and can lead to almost error free transmission up to 3621 km.

Conclusions

We have shown a novel approach to use neural networks to significantly improve the detection performance of NFT transmission experiments using the discrete spectrum on both polarizations. The amount of required training has been evaluated regarding the number of training symbols and training epochs. Taking the cross-correlations of the whole NFT spectrum into account, an increase in possible transmission distance of more than 2000 km can be reached, if the HD-FEC is considered. This leads to further implications, if and how discrete eigenvalues can be used to improve a transmission based on the continuous spectrum. This way the potentially higher spectral efficiency of the continuous spectrum[15] can be used, while also transmitting discrete eigenvalues for equalization purposes.

Acknowledgements

This work was supported by Deutsche Forschungsgemeinschaft (DFG) in the framework of the priority program "Electronic-Photonic Integrated Systems for Ultrafast Signal Processing" by grant PA 1705/1-1.

References

[1] M. Yousefi and F. Kschischang, *Information Transmission Using the Nonlinear Fourier Transform, Part I-III*, IEEE Transactions on Information Theory, Vol. 60 No. 7, August 2014.

[2] S. Gaiarin, A. M. Perego, E. P. da Silva, F. Da Ros, and D. Zibar, *Dual polarization nonlinear Fourier transform-based optical communication system,* Optica, vol. 5, pp. 263–270, Mar. 2018.

[3] S. T. Le, J. E. Prilepsky, P. Rosa, J. D. Ania-Castañón, and S. K. Turitsyn, *Nonlinear inverse synthesis for optical links with distributed Raman amplification,* IEEE J. Lightw. Technol., vol. 34, no. 8, pp. 1778–1786, Apr. 15, 2016.

[4] T. Gui, T. H. Chan et al., *Alternative Decoding Methods for Optical Communications Based on Nonlinear Fourier Transform*, IEEE J. Lightw. Technol., Vol. 35, No. 9, May, 2017.

[5] J. Koch, A. Moscoso-Martir, J. Mueller, F. Merget, S. Pachnicke, J. Witzens*, Silicon Photonics DWDM NLFT Soliton Transmitter*, arXiv preprint arXiv:2004.03183, 2020.

[6] V. Aref, S. T. Le and H. Buelow, *Modulation Over Nonlinear Fourier Spectrum: Continuous and Discrete Spectrum*, IEEE J. Lightw. Technol., Vol. 36, No. 6, March, 2018.

[7] J. Koch, R. Weixer, S. Pachnicke, *Equalization of Soliton Transmission Based on Nonlinear Fourier Transform using Neural Networks*, ECOC 2019; 45th European Conference on Optical Communication, Dublin, Ireland, 22-26 September, 2019.

[8] K. Chan, A. Geisler, et al., *Experimental Demonstration of Differential Coding Gain for Dual Polarization Two-Eigenvalue Signals*, IEEE Photon. Technol. Lett., Vol. 31, No. 24, Dec. 2019.

[9] A. Span, V. Aref et al., *Efficient Precoding Scheme for Dual-Polarization Multi-Soliton Spectral Amplitude Modulation*, IEEE Transactions on Communications, vol. 67, no. 11, pp. 7604-7615, Nov. 2019.

[10] S. Gaiarin, A. M. Perego, E. P. da Silva, F. Da Ros, and D. Zibar, *Dual polarization nonlinear Fourier transform-based optical communication system,* Optica, vol. 5, pp. 263–270, Mar. 2018.

[11] R. T. Jones, S. Gaiarin, M. P. Yankov and D. Zibar, *Time-Domain Neural Network Receiver for Nonlinear Frequency Division Multiplexed Systems,* IEEE Photon. Technol. Lett., Vol. 30, No. 12, June 2018.

[12] S. Gaiarin, F. Da Ros, N. De Renzis, R. T. Jones, D. Zibar, *Experimental demonstration of nonlinear frequency division multiplexing transmission with neural network receiver,* IEEE J. Lightw. Technol., 2020.

[13] I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning*, MIT Press, 2017.

[14] J. Koch, K. Chan, C. G. Schaeffer, S. Pachnicke, *Signal Processing Techniques for Optical Transmission Based on Eigenvalue Communication (Invited),* IEEE J. Sel. Top. in Quant. Electr., submitted, (under review).

[15] X. Yangzhang, V. Aref, S. T. Le, H. Buelow, D. Lavery, P. Bayvel, *Dual-Polarization Non-Linear Frequency-Division Multiplexed Transmission With b-Modulation*, IEEE J. Lightw. Technol., Vol. 37, No. 6, March 2019.