

Nonlinear Equalization in Long Haul Transmission Systems Using Dynamic Multi-Layer Perceptron Networks

Oleg Sidelnikov^(1,2), Alexey Redyuk^(1,2), Stylianos Sygletos⁽³⁾

⁽¹⁾ Novosibirsk State University, Novosibirsk, Russia, o.s.sidelnikov@gmail.com

⁽²⁾ Institute of Computational Technologies SB RAS, Novosibirsk, Russia

⁽³⁾ Aston Institute of Photonic Technologies, Aston University, Birmingham, UK

Abstract In this paper we investigate the application of dynamic multi-layer perceptron networks for long haul transmission systems showing performance improvement and significant superiority of neural network complexity over digital back-propagation method.

Introduction

Electronic mitigation of nonlinear transmission impairments can bring a substantial capacity increase in future fibre transmission systems. Most of the proposed schemes emulate an inverse fibre link propagation by means of a split-step Fourier (SSF)¹ method or a Volterra Series Transfer Function². However, these approaches require multiple computational steps along the link and can be applied only in static connections as they need prior knowledge of the optical path's parameters. The field of machine learning offers powerful statistical tools for the development of adaptive equalizers capable of dealing with the nonlinear transmission effects. Contrary to back-propagation based reception, in machine learning the signal equalization and demodulation processes are treated jointly as a classification problem by mapping the baseband signal onto a space determined by the direct interpretation of a known training sequence. This can bring an efficient adaptive performance and a significant reduction in the required number of computational steps. Although machine learning based equalization has been extensively studied in wireless systems their application in the optical transmission channel characterized by long memory depth and high modulations rates has not been adequately explored³.

In this paper we investigate the performance of multi-layer perceptron networks in long haul transmission systems and we compare it with the linear compensation, as well as with digital back-propagation. Our results show that a static neural networks (NN) equalizer is unable to compensate the nonlinear channel response and outperform the linear equalizer. On the contrary, when using a dynamic neural network architecture, we were able to calculate a Q^2 -factor improvement of 1.5

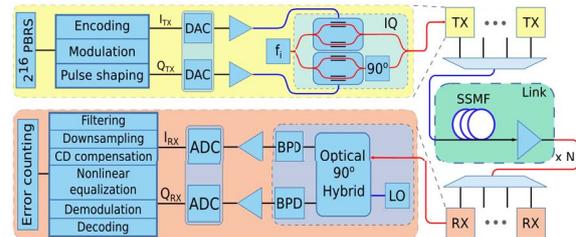
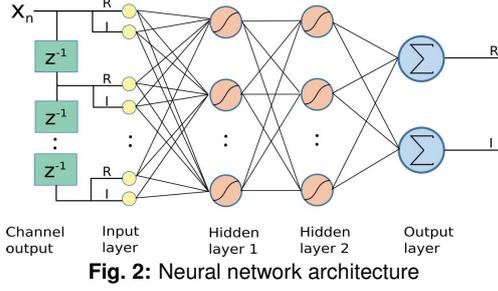


Fig. 1: Scheme of simulated transmission link

dB for single channel transmission and of 1.4 dB for multi channel transmission, along a 1000 km fibre link. The required number of taps at the input of the NN has been identified as a function of the number of fibre spans. Complexity calculations have been also performed showing the superior performance of the neural network scheme against conventional methods.

Transmission system model

The simulated transmission link is depicted in Fig. 1. Both single channel and five channel transmission scenarios were investigated. Each transmitter generated 16-QAM modulated root raised cosine pulses at 32 GBaud, with a roll-off of 0.001 and an oversampling factor of 16. In the multi-channel case the spacing was equal to the reciprocal of the baudrate. The central wavelength of the emitted signal band was at $\lambda = 1550$ nm. The generated signals were subsequently launched into a transmission link that consisted of 10 spans of 100 km single mode fibre each. An EDFA of 4.5 dB noise figure was used to compensate the losses of each span. Signal propagation was simulated with a typical symmetrized split-step Fourier method. The rest of the fibre link's parameters were: fibre loss $\alpha = 0.2$ dB/km, dispersion $D = 17$ ps/(km-nm) and nonlinear factor $\gamma = 1.4$ W⁻¹km⁻¹. After the transmission the signals were coherently detected. Each channel was selected by a root raised cosine filter

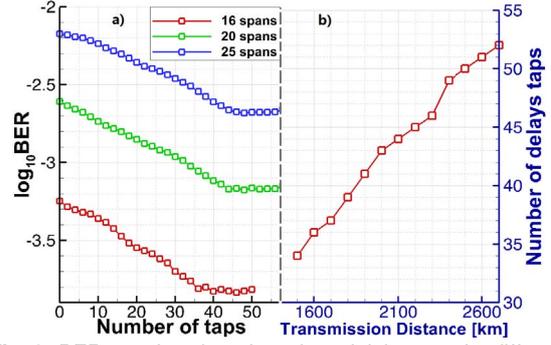


of the same roll-off factor as that of the transmitter and down-sampled to 2 samples per symbol. Then a linear equalization stage enabling ideal compensation of chromatic dispersion effects was added. After down-conversion to single sample per symbol the nonlinear equalization took place by means of a multi-layer dynamic neural network architecture.

Fig. 2 shows the neural network architecture used in the work. Contrary to previous approaches³ that employed separate neural networks for the real and imaginary part of the signal, here both signal features are fed into the same topology, reducing significantly the computational complexity. The number of neurons on the input layer is $2(N_{del} + 1)$, where N_{del} is the number of delay blocks. The network had also two hidden layers of 16 neurons each and an output layer of two neurons, i.e. one for the real and one for the imaginary output. The hidden layers had hyperbolic tangent sigmoid transfer functions, whereas the output layers had linear transfer function. Training was based on the Riedmiller's resilient-back propagation (Rprop) algorithm⁴. Each BER calculated point was averaged from 15 signal block transmissions, each containing 2^{16} symbols. From each block 2^{12} symbols were used for training and the remaining for the error rate calculation.

Numerical Results

First, we investigate the impact of the number of delay blocks on the equalization performance. Fig. 3a depicts the calculated BER for single channel transmission as a function of the number of delay taps for 16, 20 and 25 spans. The calculations have been performed at the optimum launched power. As we increase the number of delay taps the equalization performance improves until a starting point of a BER floor region where there is no need to further increase the complexity of the equalizer. Repeating the same optimization procedure we have mapped the required number of time-delay filter taps as a function of the transmission link's length, see Fig. 3b, where a linear



dependence is clearly noticed. So for a 1600 km link we need at least 36 taps, while for 2600 km the required number of taps becomes 51.

Having defined the required number of delay taps for each distance we subsequently characterized the achieved equalization performance. In Fig. 4a we have plotted the Q^2 -factor as a function of the span length for single channel transmission and for different compensation methods at the receiver. The depicted Q^2 -factors have been extrapolated from the calculated BER according to: $Q^2 = 20 \log_{10} [\sqrt{10} \operatorname{erfc}^{-1}(8BER/9)]$, at the point of maximum performance (i.e. for the launched optical power that minimized the corresponding BER). Obviously, the system with the linear compensator is the worst performing. However, when using a static neural network (i.e. without delay blocks), the achieved improvement was extremely small. A dynamic NN with an optimally dimensioned number of delay taps gave an improvement between 1 dB and 1.5 dB when the system length varied between 1500 km and 2700 km. This was a slightly higher than the performance of a symmetric digital back-propagation algorithm with 2 calculation steps per span and 2 samples per symbol. Similar conclusions can be drawn for the 5-channel Nyquist-WDM transmission scenario in Fig. 4b, where an optimal dynamic NN can give 1.4 dB Q^2 -improvement when compared to linear compensation, and slightly better than a single channel DBP of 2 steps per span.

Subsequently we compared the computational complexity of the proposed dynamic NN architecture with the DBP method. Since proposed neural network works with the real numbers, for comparing we will consider the amount of real multiplications per bit needed by DBP and scheme based on the dynamic NN. Since for DBP we use complex numbers and one complex multiplication is equal to four real multiplications we should multiply complexity of the DBP by 4. So for the DBP al-

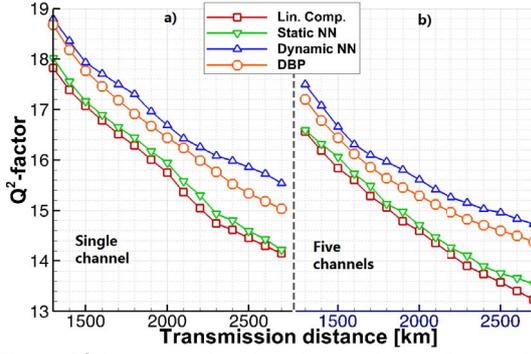


Fig. 4: Q^2 -factor as a function of transmission distance for single channel (a) and five channels (b) transmission

gorithm the complexity per bit can be easily evaluated according to the following formula⁶: $C_{DBP} = 4 \cdot N_{Span} \cdot N_{StpSp} \left(\frac{N \cdot (\log_2 N + 1) \cdot n_s}{(N - N_D + 1) \cdot \log_2 M} + n_s \right)$, where N_{Span} is the number of spans and N_{StpSp} is the number of steps per span, N is the FFT size, n_s is the oversampling factor, M is the constellation order and $N_D = n_s \tau_D / T$, where τ_D corresponds to the dispersive channel impulse response and T is the symbol duration.

In the case of the NN-based receiver a fair comparison should also include the complexity of the frequency domain equalizer for chromatic dispersion, i.e. $C_{FDE} = 4 \cdot \frac{N \cdot (\log_2 N + 1) \cdot n_s}{(N - N_D + 1) \cdot \log_2 M}$, apart from the complexity of dynamic neural network architecture. The complexity of neural network will consist of the complexity of NN training $C_{train} = N_{ep}(N_{TS} (n_i n_{hid} + n_o^2 n_{hid} + n_o n_{hid})) + 5n_o n_{hid} + n_o n_{hid}^2 + 5n_{hid}^2 + n_{hid}^2 n_i + 5n_{hid} n_i$ and the complexity of NN prediction $C_{predict} = (n_i n_{hid} + n_{hid}^2 + n_{hid} n_o) / \log_2 M$, where N_{ep} is the number of completed epochs required for training, N_{TS} is the number of samples in training set, n_i , n_{hid} and n_o are the number of neurons on the input, hidden and output layers, respectively. It should be noted that the training doesn't transmit any useful information and since it has a rather high complexity, for small number of transmitted symbols (N_S) the use of the scheme based on the dynamic NN will be inefficient. The complexity per transmitted bit for the complete scheme based on the dynamic NN will be: $C_{dNN} = C_{train} / (N_S \log_2 M) + C_{predict}$. So with the increase N_S , the overall complexity of the scheme will decrease.

Fig. 5 shows number of operations per bit as a function of transmission distance for DBP and NN with different number of transmitted symbols. As we can see, when transmitting 2^{16} symbols (including 2^{12} for training NN), the complexity of the methods turns out to be comparable. But already at $N_S = 2^{17}$ the complexity of dynamic NN will be much less. When transmitting a very large

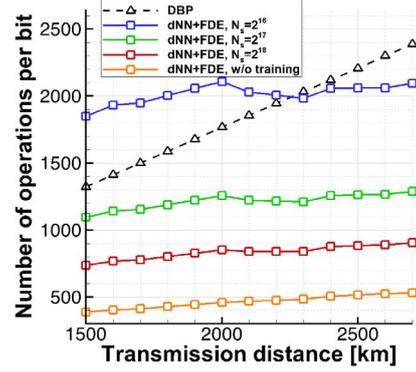


Fig. 5: The dependency of complexity of DBP and dynamic NN on transmission distance

number of symbols (in this case the complexity of training can be neglected), a scheme based on the neural network shows a significant superiority over the DBP. Moreover in all cases NN complexity grows at much slower slope with the transmission distance.

Conclusions

In this paper we investigate the performance dynamic multi-layer perceptron networks for long haul transmission and we compare it with the linear compensation and digital back propagation. Our results show dynamic neural networks is a perspective path for future systems offering significant complexity reduction when compared with traditional digital back propagation schemes of the same physical layer performance.

Acknowledgements

The work of O.S. and A.R. was supported by the Russian Science Foundation (Grant No. 17-72-30006).

References

- [1] E. Ip, "Nonlinear compensation using back-propagation for polarization multiplexed transmission," J. Lightwave Technol., Vol. **28**, no. 6, p. 939 (2010).
- [2] K. V. Peddanarappagari et al., "Volterra series transfer function of single-mode fibers," J. Lightwave Technol., Vol. **15**, no. 12, p. 2232–2241 (1997).
- [3] M. A. Jarajreh et al., "Artificial Neural Network Nonlinear Equalizer for Coherent Optical OFDM," Photon. Technol. Lett., Vol. **27**, no. 4, p. 387–390 (2015).
- [4] M. Riedmiller et al., "A direct adaptive method for faster backpropagation learning: the RPROP algorithm," Proc. ICNN, p. 586–591, San Francisco (1993).
- [5] R. A. Shafik et al., "On the extended relationships among EVM, BER, and SNR as performance metrics," Proc. 4th ICECE, p. 408–411, Dhaka (2006).
- [6] A. Napoli et al., "Reduced complexity digital back-propagation methods for optical communication systems," J. Lightwave Technol., Vol. **32**, no. 7, p. 1351–1362 (2014).