

Time-Domain Learned Digital Back-Propagation

(Invited Paper)

Eric Sillekens, Wenting Yi, Daniel Semrau,
Alessandro Ottino, Boris Karanov, Domanić Lavery,
Lidia Galdino, Polina Bayvel and Robert I. Killey

Optical Networks Group
Dept. Electronic & Electrical Engineering
UCL (University College London)
WC1E 7JE London, U.K.

Sujie Zhou, Kevin Law and Jack Chen
Huawei Chengdu Research Institute U1
Chengdu, Sichuan Province, P.R. China
Postal Code 611731

Abstract—Performance for optical fibre transmissions can be improved by digitally reversing the channel environment. When this is achieved by simulating short segment by separating the chromatic dispersion and Kerr nonlinearity, this is known as digital back-propagation (DBP). Time-domain DBP has the potential to decrease the complexity with respect to frequency domain algorithms. However, when using finer step in the algorithm, the accuracy of the individual smaller steps suffers. By adapting the chromatic dispersion filters of the individual steps to simulated or measured data this problem can be mitigated. Machine learning frameworks have enabled the gradient-descent style adaptation for large algorithms. This allows to adopt many dispersion filters to accurately represent the transmission in reverse.

The proposed technique has been used in an experimental demonstration of learned time-domain DBP using a four channel 64-GBd dual-polarization 64-QAM signal transmission over a 10 span recirculating loop totalling 1014 km. The signal processing scheme consists of alternating finite impulse response filters with nonlinear phase shifts, where the filter coefficient were adapted using the experimental measurements. Performance gains to linear compensation in terms of signal-to-noise ratio improvements were comparable to those achieved with conventional frequency-domain DBP. Our experimental investigation shows the potential of digital signal processing techniques with learned parameters in improving the performance of high data rate long-haul optical fibre transmission systems.

I. INTRODUCTION

The non-linear fibre channel has a limited capacity due to increasing nonlinear signal distortions with increasing transmission power, leading to a peak in achievable information rate (AIR) for a fixed bandwidth. One approach to increase this maximum AIR is to mitigate the non-linear distortion. This can be achieved with digital signal processing (DSP) by solving the differential equation that describes the non-linear fibre response backwards, using the received signal as the initial condition. This method, known as split-step Fourier method (SSFM) based digital back-propagation (DBP)[1], [2], has been shown to allow increased data throughput and transmission reach [3], [4].

A significant drawback of the DBP technique is its computational complexity, making it challenging to implement in real-time systems. In conventional frequency-domain DBP (FD-DBP), dispersion compensation is performed in the frequency domain and nonlinear phase shifts are corrected in the time

domain, requiring repeated conversions of the signal between the time and frequency domains using fast Fourier transforms. This leads to high computational complexity, particularly when small step-sizes, and hence a large number of steps, are used to achieve high accuracy. To reduce complexity, both dispersion and nonlinearity could be compensated in the time-domain (TD-DBP), with the dispersion compensation being carried out with tap-and-delay finite impulse response (FIR) filters [5]. However, low-order FIR filters are fundamentally unable to accurately compensate small amounts of dispersion. An approach to overcome this drawback was proposed in [6], [7], and involves applying machine-learning techniques to optimise the combined response of all the cascaded filters. The approach leverages the similarities between time-domain DBP and deep feed-forward neural networks; in both structures, linear filters and nonlinear functions are interleaved. The recent rapid advances in algorithms, and readily available software packages, allow implementation of these algorithms for the optical transmission channel.

Machine learning has been subject to a lot of research interest in recent years for communication systems [8], [9], in particular optical fibre systems[10], [11], [12]. However, often the investigated techniques are entirely substituting the function of the conventional digital processing modules with a general universal function approximator enabled by an artificial neural network. In contrast, this work utilises to a higher extend prior domain knowledge of the task being performed and individual sub-modules are designed to have specific tractable tasks, e.g., receiver side chromatic dispersion and Kerr phase shift compensation.

In this work, we experimentally demonstrate, for the first time, learned time-domain digital back-propagation. First, the method of training the required time-domain filter weights is explained. Next, the performance of the learned TD-DBP is assessed for 4-channel 64 GBd polarisation division multiplexing (PDM)-64QAM transmission over 1014 km, and compared with the performance of conventional frequency-domain DBP. Finally, the resulting filter tap weights and frequency response of the FIR filters are analysed. We observed performance improvements over linear compensation comparable to those obtained using the conventional FD-DBP implementation.

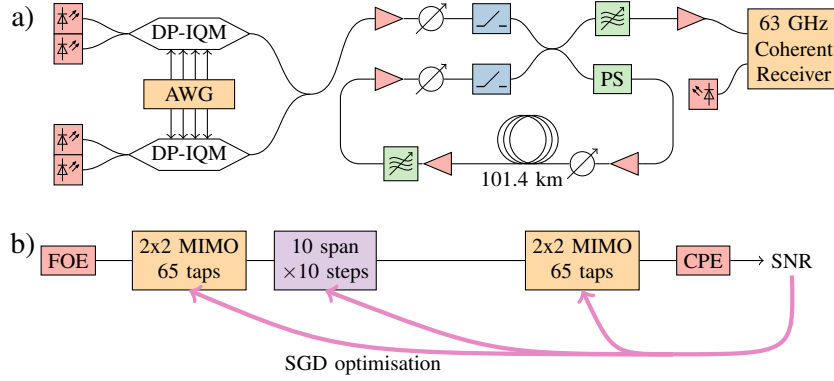


Fig. 1. a) Experimental configuration with 4×30 GbD channels and 101.4 km recirculating loop. b) Function diagram of the receiver DSP for the L-TDDBP.

II. DIGITAL BACK-PROPAGATION

DBP implements the non-linear Schrödinger equation (NLSE) for each step in two parts. For a step starting at distance z along the fibre, firstly, the chromatic dispersion and loss are applied, in the frequency domain (in the case of the conventional FD-DBP implementation) as a linear operator, and the non-linear phase shift is performed in the time domain, using the respective transformations[13, Ch. 8];

$$E(\omega, z + \Delta z) = e^{\alpha \Delta z} e^{jK(\omega T)^2} E(\omega, z) \quad (1)$$

$$E(t, z + \Delta z) = e^{-j\gamma \Delta z |E_z(t)|^2} E(t, z), \quad (2)$$

where α is fibre loss, Δz is fibre step length, $K = \frac{\beta_2 \Delta z}{2T^2}$, ω angular frequency, T sampling period, and β_2 group velocity dispersion, γ the nonlinearity coefficient and $|E_z(t)|^2$ the normalised, step-averaged, instantaneous optical power.

In the time-domain DBP approach, the chromatic dispersion part of each step is applied using a time-domain finite impulse response (FIR) filter (tap-and-delay filter). The full-band least-squares FIR filter design from [14] could be used with a total number of taps given by $N_c \leq 2[2\pi K] + 1$. However, as described in [6], if the filter tap weights given by [14, Eq. (13)] are used for the multiple cascaded low-order FIR filters in the DBP, the ripples introduced into the frequency response result in large performance penalties, negating the gains achieved through the non-linearity mitigation. The solution proposed in [6] is to update all the FIR filter weights simultaneously using algorithms which have been developed to update the weights in deep feed-forward neural networks. The method consists of implementing the dispersion as a convolutional layer and the fibre phase shift as a non-linear activation function.

In this work, the deep learning of the filter weights (equivalent to a neural network's layer parameters) is implemented in Tensorflow using the RAdam optimiser[15]. Identical complex FIR filter weights are applied to both polarisations in each step, reducing the overall number of weights to be optimized. Initialisation of the time-domain filter taps was carried out via numerical simulation of the fibre transmission link for a single channel. The forward propagation was modelled using a small NLSE fibre step size (100 m) at a launch power of 5 dBm

(beyond optimum launch power for linear compensation). Starting from the least squares solution[14], a set of 10 filters was designed using a 10 span simulation. Note, these filters purely compensate fibre transmission and no transceiver impairments. Further training of the filters was carried on the experimental waveforms, before the performance was tested.

III. EXPERIMENTAL SETUP

The experimental setup is shown in Fig. 1(a). A fibre transmission distance of 1014 km was emulated using a recirculating loop. The waveform of the 64-QAM 64-GBd channel under test (CUT) was generated offline and sent to two channels of a 33-GHz 92-GSa/s arbitrary waveform generator (AWG) and, using a dual-polarization IQ modulator (IQM), modulated the outputs of two <100 kHz external cavity lasers (ECLs). Two additional 64-QAM 64-GBd aggressor channels were modulated using an additional AWG with a dual polarization IQM onto two ECLs and interleaved with the other channels to achieve uncorrelated sequences between neighbouring WDM channels. The recirculating loop with a 101.4 km span, a polarisation scrambler (PS) and three Erbium doped fibre amplifiers (EDFA) and an optical band-pass filter had the signal circulating 10 times, totalling a 1014 km transmission. At the receiver an optical band pass filter followed by an EDFA extracts the CUT for detection with a coherent receiver employing 63-GHz bandwidth 160-GS/s analogue-to-digital converters.

For this experimental demonstration, 10 steps per span were chosen for the learned TD-DBP. To take fibre loss into account, non-uniform FIR filter lengths were employed, implementing steps with equal power differences between their inputs and outputs. For the fibre nonlinearity compensation of 10 spans of 101.4 km each, the parameters were α of 0.16 dB/km, β_2 of -20.18 ps²/km and γ_{DBP} of 0.8 1/W/km. The 10 FIR filters used in the TD-DBP employed a total of 270 complex-valued tap weights at a sampling rate of 128 GSa/s. For the FD-DBP, 50 equidistant steps/span were used. This requires $2 \times 10 \times 50$ FFT operations per polarisation, while in the TD-DBP scheme the use of FFT operations is circumvented, lowering the computational complexity.

Next, for the processing of experimental data, the filter weights from simulation were used for initialisation. To prevent the dispersion filters from learning the response of the transceiver impairments, an additional 2x2 multiple input, multiple output (MIMO) filter was added before applying digital back propagation, as shown in Fig. 1(b). Thus, the resulting structure has two linear MIMO equalisers, compensating for PMD and transmitter and receiver impairments. Note that in this way, using the automatic differentiation in Tensorflow, the filter that is applied prior the link compensation is also optimized through gradient descent. A root-raised cosine (RRC) filter was applied before the MIMO blocks. The carrier phase estimation was achieved by inserting pilot symbols. One in 32 symbol was a known quadrature phase shift keyed symbol (QPSK). Interpolation of the phase between the pilot symbols was performed using a Wiener filter [16, Eq. (32)], following which a mean-squared-error cost is calculated.

During the training procedure, first the linear filters at both sides of the link compensation were optimised. Subsequently, all filters were updated on each optimisation step. For the FD-DBP, the α , γ and launched power were swept for optimisation, after which pilot-aided DSP was applied. For the experimental waveform, a single randomly generated 2^{16} -symbol waveform was used. We split the bit sequence and corresponding received waveforms into two datasets. The first 52224 symbols (80%) were used as training data for updating the filter weights. The remaining 13312 symbols (20%) were used as testing data, to obtain results reported in the figures presented.

IV. RESULTS

The launched power was increased with 1 dB increments from -6 to +8 dBm per channel. The resulting SNR, defined as $\frac{E[|X|^2]}{E[|X-Y|^2]}$, where X and Y are the transmitted and received signal respectively. Fig. 2(a) shows a comparison of achieved SNR for TD-DBP, FD-DBP and EDC. The TD-DBP and FD-DBP are implemented using 10 and 50 steps per span respectively. Both schemes provide similar performance improvements from non-linearity compensation, with slightly higher accuracy in the high power regime for the conventional FD-DBP scheme, due to the larger number of steps used. However, the TD-DBP achieves a higher SNR in the low power regime, suggesting a better linear compensation. Fig. 2(b) compares TD-DBP with two learned linear compensation strategies. The figure shows the learned DBP performance for two cases, the proposed non-linear mitigation scheme, and the same scheme with $\gamma_{\text{DBP}} = 0$, i.e., providing only linear compensation and a scheme where the whole chromatic dispersion is compensated in a single filter. For a low launched power into the fibre, the first two schemes show comparable performance, while a non-linearity mitigation gain of 0.3 dB is achieved at optimal launch powers. Using a single filter achieves better linear gain, but converges to the $\gamma_{\text{DBP}} = 0$ method in the high launched power regime.

To confirm that the algorithm is performing digital back-propagation, i.e., approximating the SSFM model, the ampli-

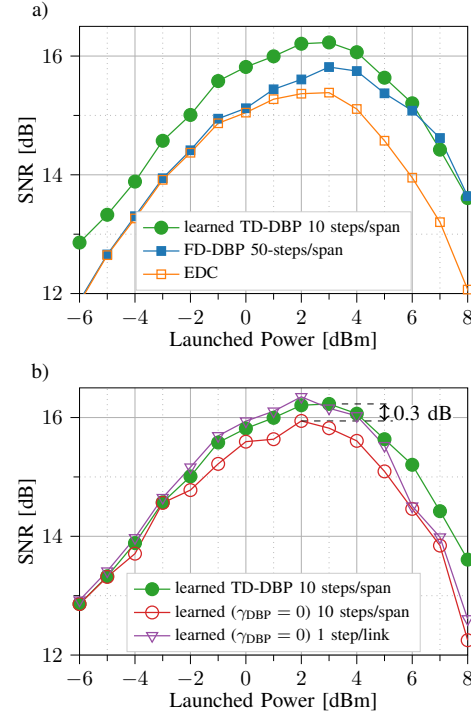


Fig. 2. a) SNR vs. launched power for learned TD-DBP compared to conventional FD-DBP and linear EDC only methods. b) TD-DBP compared to the same structure with $\gamma_{\text{DBP}} = 0$ with the same number of steps and a single step for the whole link.

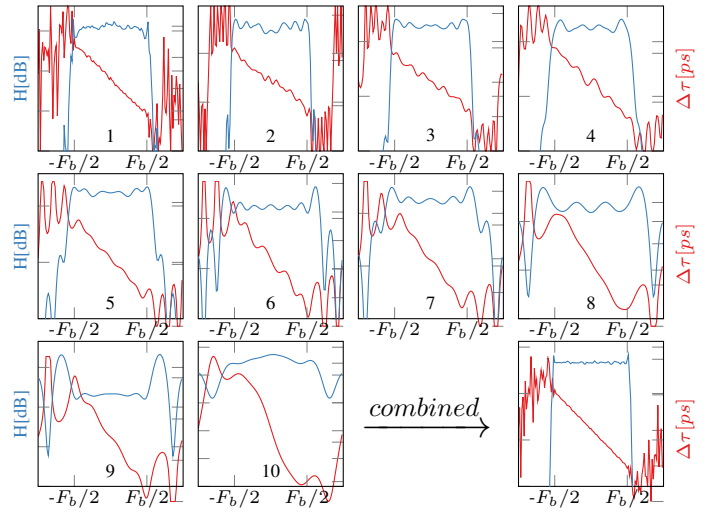


Fig. 3. Amplitude response and group delay of the 10 individual filters used every span. Bottom right: combined response of all 10 cascaded filters.

tude response and group delay of the 10 individual filter used each span are plotted in Fig. 3. The expected response is an all-pass filter (H) with a linear group delay ($\Delta\tau$), compensation for chromatic dispersion. It can be seen that, while the individual filters have significant ripples, the combined filter, depicted as the last subplot of Fig. 3, has an almost perfect response within the signal bandwidth, with a flat amplitude response and a smooth group delay.

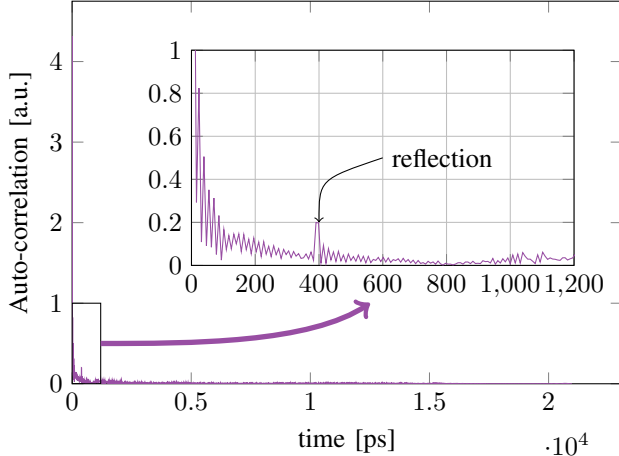


Fig. 4. The auto-correlation of the learned single filter used to compensate chromatic dispersion.

V. CONCLUSION

We experimentally demonstrated learned time-domain digital back-propagation. Our experiment concentrates on wide-band multi-channel long-haul optical transmission systems. The learned algorithm was verified to approximate the NLSE model, showing flat amplitude response and smooth group delay for the cascaded filters. Furthermore, we highlight the improvements stemming from linear and non-linear compensation. We show that the investigated method achieves an SNR improvement of 0.3 dB due to non-linearity mitigation, which is comparable to the conventional mitigation algorithms.

ACKNOWLEDGEMENTS

Financial support by EPSRC TRANSNET grant EP/R035342/1 and EU COIN programme under the Marie-Sklodowska-Curie grant No.676448. Eric Sillekens was funded by EPSRC grant EP/M507970/1.

REFERENCES

- [1] X. Li, X. Chen, G. Goldfarb, E. Mateo, I. Kim, F. Yaman, and G. Li, "Electronic post-compensation of wdm transmission impairments using coherent detection and digital signal processing," *Opt. Express*, vol. 16, no. 2, pp. 880–888, Jan 2008.
- [2] E. Ip and J. M. Kahn, "Compensation of dispersion and nonlinear impairments using digital backpropagation," *J. Lightw. Technol.*, vol. 26, no. 20, pp. 3416–3425, Oct 2008.
- [3] L. Galdino, D. Semrau, D. Lavery, G. Saavedra, C. B. Czegledi, E. Agrell, R. I. Killey, and P. Bayvel, "On the limits of digital backpropagation in the presence of transceiver noise," *Opt. Express*, vol. 25, no. 4, pp. 4564–4578, Feb 2017.
- [4] G. Liga, T. Xu, A. Alvarado, R. I. Killey, and P. Bayvel, "On the performance of multichannel digital backpropagation in high-capacity long-haul optical transmission," *Opt. Express*, vol. 22, no. 24, pp. 30 053–30 062, Dec 2014.
- [5] C. Fougstedt, M. Mazur, L. Svensson, H. Eliasson, M. Karlsson, and P. Larsson-Edefors, "Time-domain digital back propagation: Algorithm and finite-precision implementation aspects," in *Optical Fiber Communications Conference and Exhibition (OFC)*, 2017, p. W1G.4.
- [6] C. Häger and H. D. Pfister, "Nonlinear interference mitigation via deep neural networks," in *Optical Fiber Communications Conference and Exhibition (OFC)*, 2018, p. W3A.4.
- [7] —, "Deep learning of the nonlinear schrödinger equation in fiber-optic communications," in *IEEE ISIT*, Jun. 2018, pp. 1590–1594.
- [8] O. Simeone, "A very brief introduction to machine learning with applications to communication systems," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 4, pp. 648–664, 2018.
- [9] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Transactions on Cognitive Communications and Networking*, vol. 3, no. 4, pp. 563–575, Dec 2017.
- [10] F. N. Khan, Q. Fan, C. Lu, and A. P. T. Lau, "An optical communication's perspective on machine learning and its applications," *Journal of Lightwave Technology*, vol. 37, no. 2, pp. 493–516, Jan 2019.
- [11] D. Zibar, M. Piels, R. Jones, and C. G. Schäeffler, "Machine learning techniques in optical communication," *Journal of Lightwave Technology*, vol. 34, no. 6, pp. 1442–1452, March 2016.
- [12] B. Karanov, M. Chagnon, F. Thouin, T. A. Eriksson, H. Bülow, D. Lavery, P. Bayvel, and L. Schmalen, "End-to-end deep learning of optical fiber communications," *Journal of Lightwave Technology*, vol. 36, no. 20, pp. 4843–4855, Oct 2018.
- [13] G. P. Agrawal, *Fiber-Optic Communication Systems*, 4th ed. John Wiley and Sons, 1992.
- [14] A. Eghbali, H. Johansson, O. Gustafsson, and S. J. Savory, "Optimal least-squares fir digital filters for compensation of chromatic dispersion in digital coherent optical receivers," *J. Lightw. Technol.*, vol. 32, no. 8, pp. 1449–1456, Apr 2014.
- [15] L. Liu, H. Jiang, P. He, W. Chen, X. Liu, J. Gao, and J. Han, "On the variance of the adaptive learning rate and beyond," 2019, arXiv:1908.03265.
- [16] E. Ip and J. M. Kahn, "Feedforward carrier recovery for coherent optical communications," *J. Lightw. Technol.*, vol. 25, no. 9, pp. 2675–2692, Sep 2007.