Design Aspects of Frequency-Domain Learned MIMO Volterra Equalisers

Nelson Castro[*], Sonia Boscolo, Andrew D. Ellis, and Stylianos Sygletos
Aston Institute of Photonic Technologies, Aston University, Birmingham B4 7ET, United Kingdom
[*]*cast1901@aston.ac.uk*

Abstract: We numerically demonstrate a frequency-domain learned multiple-input multiple-output Volterra nonlinear equaliser and reveal the impact of practical implementation parameters on the per-channel performance.

1. Introduction

Digital compensation of nonlinear transmission impairments is a promising solution for the capacity enhancement of optical communication systems [1]. However, the dominance of inter-channel impairments in wavelength division multiplexing (WDM) systems makes multi-channel operation an indispensable feature of future nonlinear equalisers (NLEs). The development of multiple-input multiple-output (MIMO) digital back-propagation (DBP) based models [2] has significantly brought down the algorithmic complexity compared to earlier full-field approaches [3]. Furthermore, machine learning (ML) has recently enabled the development of versatile multi-channel NLE algorithms with enhanced estimation capabilities and low complexity, capable of operating without complete knowledge of the transmission link parameters [4]. In particular, it has already been shown that ML can drastically simplify pure time-domain (TD) implementations of NLEs by enabling joint optimisation of their parameters in the linear and nonlinear stages [5]. In this respect, we have recently introduced a simplified Volterra-based MIMO equalisation scheme with fully trainable TD stages [6]. Despite the considerable reduction in finite impulse response (FIR) filter size achieved for the linear stages (i.e., ~ 40 taps for $\sim$ 50-km computational step length at 32 Gbaud), implementing such TD filters is still challenging. Alternatively, a frequency-domain (FD) approach using block-processing techniques, such as overlap-and-save [7], may offer a more feasible implementation of the linear stages for long-haul transmission scenarios, whilst sacrificing the adaptive operation advantage.

Building upon our previous work [6], this paper introduces a learned Volterra-based MIMO NLE in which static FD equalisers represent the linear stages, whereas the TD FIR filters of the nonlinear stages are the adaptive parts of the algorithm. Contrary to [8], we use stochastic gradient descent-based optimisation, enabling fast model convergence. Through extensive numerical simulations, we have established practical design rules for optimising the equalization performance. A 5×5 MIMO equaliser with 4 computational steps achieves $\sim$ 2-dB improvement in effective signal-to-noise ratio (SNR) over a single-channel equaliser with the same number of computational steps and requires 2048 overlapping samples to ensure uniform equalisation across a 160-GHz WDM bandwidth for 1000-km transmission.

2. Methods, Results and Discussion

The architecture of our proposed MIMO simplified learned inverse Volterra series transfer function (L-simIVSTF) NLE is shown in Fig. 1(a) for the example of a 2×2 interconnection with two computational steps. Further advancing our previous TD model [6], the linear steps addressing the per-span chromatic dispersion (CD) were solved in the frequency domain (FD). Direct and inverse fast Fourier transforms were used to switch between the time and frequency representations of the signal. The desired linear filtering was ensured by using adequate overlapped framing. The same block size was used throughout the structure. The H_{CD} transfer functions, otherwise identical for each step and wavelength channel, included an appropriate delay term to account for the walk-off among channels. The parallel branches deriving from the filter cascade offset the impact of nonlinearity by reversing the Kerr-induced nonlinear transformation in each fibre section. Following an enhanced transfer function approach successfully applied within DBP schemes [8], we considered the filtered impact of adjacent samples in both self-phase modulation (SPM) and cross-phase modulation (XPM) transformations. Therefore, the SPM and XPM activation functions of the nth channel at the tth sampling instant were realised as $\sigma(t)_{n,\mathrm{SPM}} = -i\gamma L y_n \sum_{c=-l}^{l} \alpha_c |y_n^{(t+c)}|^2$ and $\sigma(t)_{n,\mathrm{XPM}} = -2i\gamma L y_n \sum_{k\neq n}^{N_{\mathrm{ch}}} \sum_{c=-m}^{m} \beta_{c,k} |y_k^{(t+c)}|^2$, respectively, where N_{ch} is the total number of channels, γ is the fibre nonlinear coefficient, L is the effective length of each fibre step, and the instantaneous signal powers are convolved with SPM and XPM FIR filters having real-valued coefficients α_c and β_c and lengths $2l+1$ and $2m+1$, respectively. Similarly to [6], α_c and β_c were jointly optimised using gradient-based back-propagation.

We considered the transmission of 5 single-polarisation wavelength channels along a $10\times$ 100-km standard single-mode fibre link (dispersion parameter $D = 17\,\mathrm{ps/(nm\cdot km)}$, $\gamma = 1.3\,(\mathrm{W\cdot km})^{-1}$, loss coefficient $\alpha = 0.2\,\mathrm{dB/km}$). Erbium-doped fibre amplifiers of 4.5-dB noise figure compensated for the span losses. Each channel carrier was modulated with 2^{18} 64-QAM symbols at a rate of 32 Gbaud. The channel spacing was 40 GHz.

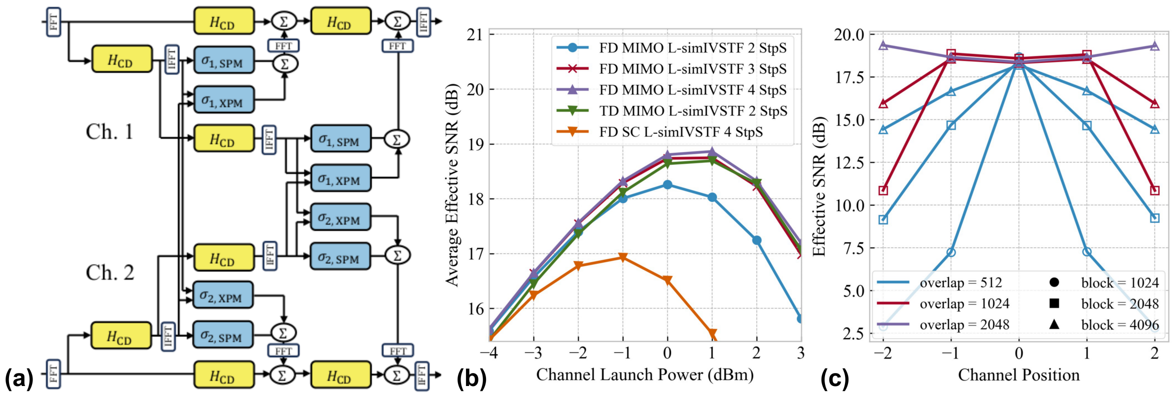

Fig. 1: (a) Block diagram of the proposed FD L-simIVSTF NLE for 2×2 MIMO operation with 2 computational steps. (b) Performance of different realisations of the model for 5×5 MIMO operation versus channel launched power. (c) Per-channel performance of the 5×5 MIMO model with 4 computational steps for different overlap and block lengths.

Ideal optical filters were assumed for channel multiplexing/demultiplexing. At the receiver, after demodulation and down-sampling to 2 samples per symbol, the channels were processed by a 5×5 MIMO NLE. We used block-by-block processing, efficiently implemented by an overlap-and-save method [7]. For the optimisation of the SPM and XPM filters, we used a training data set of 2^{20} symbols, while the validation and testing data sets had a 2^{18}-symbol size. Model training was performed for a total of 1500 epochs, after which no further performance improvement was observed. The filters were initialised with 0's except the middle taps, which were set to 1. The filter lengths were set to the minimum required values found for the TD model [6], i.e., 7 taps for the SPM filters and 31 taps for the XPM filters. It is worth noting that increasing the number of steps per span in the model can further reduce the length requirement for these filters. Nevertheless, our findings suggested that the model's performance remains unaffected, provided that sufficient training iterations are performed.

Figure 1(b) shows the performance of different implementations of a 5×5 MIMO L-simIVSTF NLE in terms of the average effective SNR of the channels as a function of the channel launched power, where the SNR for each channel was derived from the corresponding bit-error rate. The performance of the FD model is shown for 2, 3, and 4 steps per span, and compared to those of the previously proposed TD MIMO L-simIVSTF model [6] and a FD single-channel L-simIVSTF NLE operated at 4 steps per span. We can see that the model needs to be operated at more than 3 steps per span to achieve similar performance to the TD model at 2 steps per span. Operating the model at 4 steps per span brings about 0.3-dB SNR improvement over its 2-steps-per-span TD counterpart and $\sim$ 2-dB improvement over the equivalent single-channel model. The requirement of a larger number of steps per span by the FD implementation of the MIMO model stems from the absence of any training applied to its linear steps. Furthermore, we studied the relationship between block length, overlap length and performance of the FD MIMO model. We considered a 5×5 configuration operated at 4 steps per span. A separate model was trained for each combination of overlap and block lengths. The results are summarised in Fig. 1(c). They evidence that tuning these parameters is essential to ensure the convergence of the MIMO equaliser so that it provides nonlinearity equalisation to all channels. Whilst an overlap length of 1024 samples is sufficient to achieve inter-channel compensation of 3 channels only, 2048 overlapping samples are required for equal compensation of all 5 channels. The necessary overlap length is dictated by the walk-off between the outermost channels. Conversely, the block length is constrained by the overlap length and stands as an optimisable parameter, potentially allowing for complexity reduction.

3. Conclusion

We have presented an FD learned Volterra MIMO NLE and studied the impact of practical implementation parameters on its performance. We have established the minimum required overlap and block lengths to attain effective inter-channel nonlinearity compensation. **Acknowledgments:** This work was partly supported by the UK EPSRC grants TRANSNET (EP/R035342/1) and CREATE (EP/X019241/1), and the RAEng RCSR fellowship.

References

1. A.D. Ellis, A.A.I. Ali, M. Tan, N. Salgado, S. Sygletos, Proc. Optica Adv. Photon. Congress 2022, NeTu3D.4.
2. E.F. Mateo, F. Yaman, G. Li, Opt. Express **18**, 15144–15154 (2010).
3. G. Liga, T. Xu, A. Alvarado, R. I. Killey and P. Bayvel, Opt. Express **22**, 30053–30062 (2014).
4. T. Inoue, R. Matsumoto, A.S. Namiki, Opt. Express **30**, 14851–14872 (2022).
5. C. Häger, H.D. Pfister, Proc. 2018 ECOC, 1–3.
6. N. Castro, S. Sygletos, Proc. ACP 2023, 1–4.
7. V. Oppenheim, R.V. Schafer, *Digital Signal Processing* (Prentice-Hall, 1975).
8. M. Secondini, S. Rommel, G. Meloni, F. Fresi, E. Forestieri, L. Potì, Photon. Netw. Commun. **31**, 493–502 (2016).