Adaptive Neural Network Equalisation Using Skip Connections for Future 100 Gbit/s/λ Passive Optical Networks

Stephen L. Murphy(1,*), Fariba Jamali(1), Paul D. Townsend(1), Cleitus Antony(1)

(1) Photonics Systems Group, Tyndall National Institute, University College Cork, Cork, Ireland
(*) stephen.murphy@tyndall.ie

Abstract *We present a novel equaliser based on a neural network with skip connections for 100 Gbit/s PAM4 SOA-preamplfied PONs which can converge within 2000 symbols on a burst-by-burst basis, and effectively compensate SOA patterning and 81.6 ps/nm fiber dispersion, enabling 21 dB system dynamic range. ©2023 The Author(s)*

Introduction

Current Passive Optical Network (PON) research is focused on pushing beyond recent 50 Gbit/s standards[1] while adhering to strict optical loss budget and dynamic range (DR) requirements. 50 Gbaud 4-level pulse amplitude modulation (PAM4) is being considered alongside 25G class opto-electronics, but is highly susceptible to nonlinearities and would require optical amplification due to its reduced signal-to-noise tolerance. Semiconductor optical amplifiers (SOAs) could provide this[2], but nonlinear gain saturation induced SOA patterning effects can limit system DR. Neural network equaliser (NNE) techniques have been widely proposed to overcome this and other nonlinear impairments, and have achieved impressive performance when compared with conventional equalisation techniques[3]–[5].

However, advanced digital signal processing is most likely to be deployed at the optical line terminal (OLT) where resources and costs can be shared among network subscribers, meaning any equalisation must be able to deal with varying levels and combinations of system and device impairments of the burst-packets received by the OLT transceiver. Adaptive NNEs would need to converge within 100s of nanoseconds, using a short sequence of training symbols located in each packet's preamble[6], making conventional NNE training schemes unsuitable as they rely on large datasets and computationally complex optimisation algorithms, such as ADAM[7], to achieve optimal performance. Many reported NNEs undergo full retraining using such algorithms for each impairment or power level considered, which is impossible under realistic PON operating conditions. Proposed solutions include creating large "universal" NNEs which are exposed to all possible packet conditions[8] during offline training, or to similarly train sub-models to differentiate each distinct packet's signal statistics[9], but these come with non-ideal, large training and data collection overhead.

In this paper, we propose a novel modular NNE architecture for PON which is trained once and then adapted to varying packet conditions using a simple LMS algorithm and only 2000 training symbols. It achieves this using skip connections which are inspired by residual networks[10], and has similar structure to the time delay NNE discussed in[11]. We apply the proposed scheme to a 100 Gbit/s PAM4 system emulating upstream PON transmission with an SOA preamplifier in continuous mode, and demonstrate its robustness to varying SOA gain saturation induced patterning and fiber dispersion up to 81.6 ps/nm. The scheme achieves 27 dB DR performance back-to-back (BtB), matching the performance observed when using full ADAM optimisation, and 21 dB DR with 81.6 ps/nm of fiber dispersion, exceeding the 19.5 dB DR of current 50G standards[1].

Adaptive Neural Network Equalisation

One approach to applying NNEs to varying PON packet impairments is to train a NNE once, exposing it to a single stressed packet scenario, and then apply to all other packet conditions without further training or adaption. While this "static" NNE fully utilizes its nonlinear capabilities for the stressed packet, it fails for disparate or less severely impaired packets, as we show below.

Fig. 1 outlines our proposed adaptive scheme. First, a NNE is trained offline and exposed to a worst-case packet, such as high dispersion or nonlinear SOA patterning, using the ADAM optimisation algorithm or similar. The NNE is then decomposed into its nonlinear hidden layers whose parameters are kept static, and an output layer consisting of a single neuron with linear activation

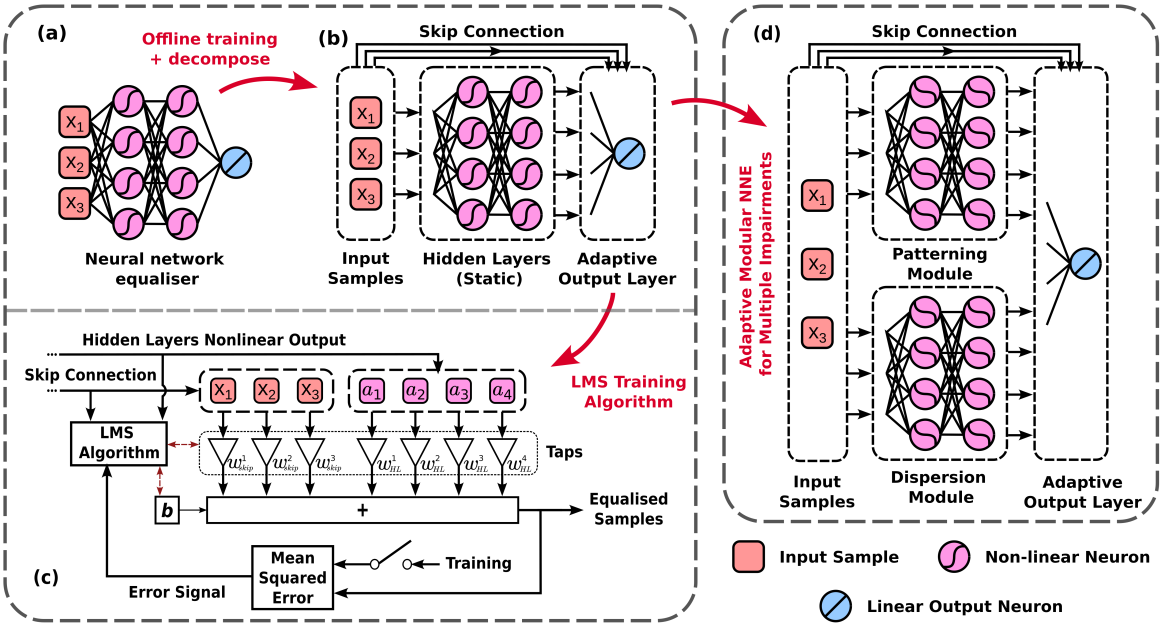

Fig. 1: (a) and (b) outline the proposed skip connection and adaptive linear output structure, which allow for rapid adaption to packet conditions outside the NNE training data using LMS, as shown in (c). In (d), the concept of combining modular NNEs for addressing multiple impairments is shown, where each "module" is exposed to a single isolated impairment during training.

which is adaptively trained packet-by-packet, as in Fig. 1 (b). Signal samples are input to the NNE as usual, but are also also fed via a skip connection to the adaptive output layer according to:

$$\hat{y} = w_{HL} \cdot a + w_{skip} \cdot x + b$$

Where $\hat{y}$, w, b are the equalised output sample, weight and bias parameters respectively. x is the equaliser inputs coming from the skip connection, while a is the nonlinear outputs from the hidden layers having processed x already. The weights of the output layer are then adaptively trained like an FFE, using the LMS algorithm as shown in Fig. 1 (c).

This architecture can be made modular as in Fig. 1 (d) to deal with multiple impairments which may appear in isolation, or combination. A NNE is trained offline for each isolated nonlinear impairment, and then combined with a single adaptive linear output and skip connection. In this way, NNEs "modules" can be added for additional impairments, without the need to retrain the other NNEs focused on different impairments.

Experimental Setup

Fig. 2 (a) shows the experimental setup. A 50 Gbaud PAM4 signal is generated using a 100 GSa/s DAC, with linear precompensation correcting for system bandwidth restrictions up to 33 GHz. A Mach Zehnder modulator combined with an EDFA booster amplifier constitutes a high-power C-band Tx. An ideal Rx composed of an EDFA preamplifier and 50 GHz photodiode is used, while waveforms are captured for offline processing and equalisation using a 200 GSa/s real time scope.

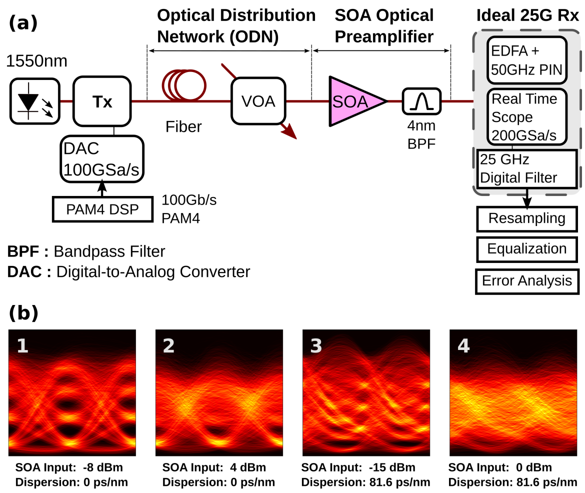

Fig. 2: (a) Experimental setup and (b) resulting eye diagrams showing range of isolated and combined signal impairments.

The SOA gain saturation induced patterning effect of the SOA is first studied BtB, without the effects of fiber dispersion. The SOA gain drops by 3 dB for -8 dBm input power, and it's input is swept from -24 to $+4$ dBm. Fiber dispersion up to 81.6 ps/nm is then introduced, using standard single mode fiber. A 50 GHz 4$^{\text{th}}$-order Bessel filter is applied in offline processing to the BtB setup in order to avoid the introduction of linear bandwidth impairments when studying the SOA patterning effect, while a 25 GHz filter emulating 25G class opto-electronics is used in the transmission case.

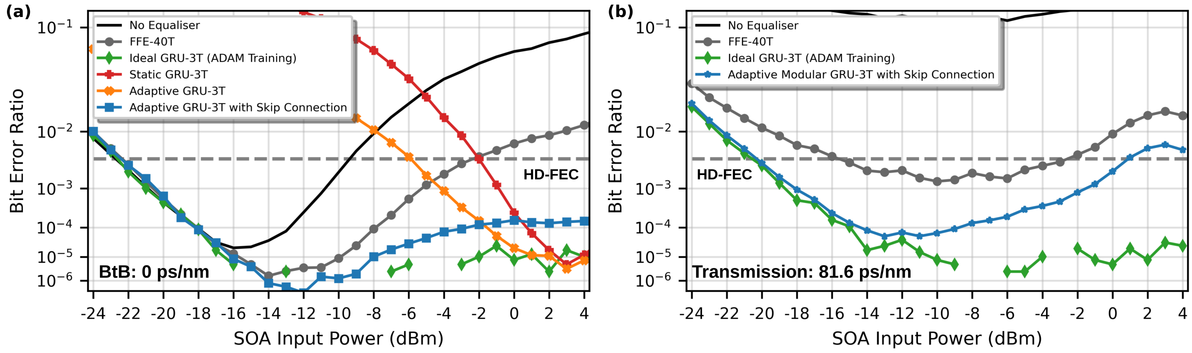

Fig. 3: In (a) the inclusion of a skip connection allows the adaptive GRU-3T model to succeed outside it's training conditions, while (b) shows the transmission performance of the same model when combined with a NNE "module" trained on dispersion.

Results

To test our scheme we use a NNE based on Gated Recurrent Units (GRU), described in[5]. It has one hidden layer composed of 6 GRU units followed by a single linear output neuron. All equalisation reported here uses 1 sample per symbol, and the considered GRU NNE uses only 3 such taps (i.e. GRU-3T). The 'Ideal GRU-3T' refers to ideal GRU-3T performance after packet-by-packet offline training using ADAM optimisation for 1000 epochs with the early stopping method. All adaptive GRU-3T equalisers and FFE are trained using the LMS algorithm and 5000 randomly generated symbols, and tested on a pseudo random quaternary sequence (PRQS15). System performance is determined using the hard decision forward error correction (HD-FEC) limit of 3.8×10^{-3} bit error ratio (BER).

Fig. 3 (a) shows BtB performance, with ideal GRU-3T achieving > 27 dB DR. However, the static GRU-3T trained at $+4$ dBm SOA input and then applied with fixed parameters to other powers clearly fails, as does adaptively training the linear output layer of the GRU-3T. In both these cases the NNEs cannot deal with less severely patterned signals below -2 and -6 dBm respectively. However, if our proposed skip connection is integrated with the adaptive GRU-3T then the ideal GRU-3T sensitivity and DR performance can be recovered, and the adaptive NNE is effective in SOA gain saturated, impairment free, and noise limited operating conditions.

Fig. 3 (b) shows the transmission performance of the proposed adaptive modular GRU-3T, where a separate GRU-3T trained on the isolated 81.6 ps/nm dispersion impairment is combined with the original GRU-3T trained on the isolated SOA patterning impairment reported above, and described in Fig. 1 (c). By switching to this modular GRU-3T, the equaliser can now adaptively equalise both varying nonlinear SOA patterning

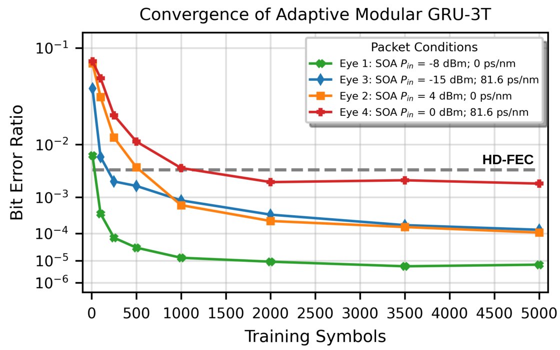

Fig. 4: The modular structure can adapt rapidly in under 2000 symbols to a wide range of packet conditions.

effect and fiber dispersion individually. Fig. 3 (b) shows the adaptive modulat GRU-3T achieves good performance in the combined patterning and dispersion regime with BER below the HD-FEC up to $+1$ dBm SOA input power with 81.6 ps/nm dispersion.

Fig 4 shows that the adaptive modular GRU-3T equaliser converges below HD-FEC in up to 2000 symbols for a wide range of packet conditions, corresponding to the eye diagrams in Fig. 2 (b). This is equivalent to 40 ns for a 50 Gbaud signal and so approaches the convergence requirements of conventional adaptive equalisers[6].

Conclusion

In summary, we propose a novel adaptive NNE architecture using skip connections which greatly simplifies NNE training for PON. It achieves 27 and 21 dB system DR for BtB and transmission cases respectively in a setup emulating a 100 Gbit/s upstream PAM4 with SOA preamplifier. These DRs are above the 19.5 dB required by recent 50G standards. Further, the adaptive NNE scheme is shown to converge for a wide range of severely impaired packet conditions using the LMS algorithm and within 2000 training symbols.

Acknowledgements

This work is supported by Science Foundation Ireland (SFI), under IPIC grant 12/RC/2276P2.

References

[1] ITU-T Recommendation G.9804.3, *50-Gigabit-capable passive optical networks (50G-PON): Physical media dependent (PMD) layer specification*, Sep. 2021.

[2] Z. Li, Y. Li, S. Luo, F. Yin, Y. Wang, and Y. Song, "SOA Amplified 100 Gb/s/λ PAM-4 TDM-PON Supporting PR-30 Power Budget with $>$18 dB Dynamic Range", *Micromachines*, vol. 13, no. 3, p. 342, Feb. 2022. DOI: 10.3390/mi13030342.

[3] K. Wang, C. Wang, J. Zhang, Y. Chen, and J. Yu, "Mitigation of SOA-Induced Nonlinearity With the Aid of Deep Learning Neural Networks", *Journal of Lightwave Technology*, vol. 40, no. 4, pp. 979–986, Feb. 2022. DOI: 10.1109/JLT.2021.3126765.

[4] X. Huang, D. Zhang, X. Hu, C. Ye, and K. Zhang, "Low-Complexity Recurrent Neural Network Based Equalizer With Embedded Parallelization for 100-Gbit/s/λ PON", *Journal of Lightwave Technology*, vol. 40, no. 5, pp. 1353–1359, Mar. 2022. DOI: 10.1109/JLT.2021.3128579.

[5] S. L. Murphy, F. Jamali, P. D. Townsend, and C. Antony, "High Dynamic Range 100G PON Enabled by SOA Preamplifier and Recurrent Neural Networks", *Journal of Lightwave Technology*, pp. 1–11, 2023. DOI: 10.1109/JLT.2023.3249046.

[6] J. Zhang, Z. Jia, M. Xu, H. Zhang, and L. Alberto Campos, "Efficient preamble design and digital signal processing in upstream burst-mode detection of 100G TDM coherent-PON", *Journal of Optical Communications and Networking*, vol. 13, no. 2, A135–A143, 2021. DOI: 10.1364/JOCN.402591.

[7] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization", *arXiv e-prints*, Dec. 2014. DOI: 10.48550/arXiv.1412.6980.

[8] V. Houtsma, E. Chou, and D. van Veen, "92 and 50 Gbps TDM-PON using Neural Network Enabled Receiver Equalization Specialized for PON", in *2019 Optical Fiber Communications Conference and Exhibition (OFC)*, 2019, pp. 1–3.

[9] C. Ye, X. Hu, D. Zhang, and K. Zhang, "A Versatile NN-Equalization for 50Gbps TDM PON Burst Uplink", in *2020 European Conference on Optical Communications (ECOC)*, Dec. 2020, pp. 1–3. DOI: 10.1109/ECOC48923.2020.9333360.

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition", *arXiv e-prints*, Dec. 2015. DOI: 10.48550/arXiv.1512.03385.

[11] M. Schädler, G. Böcherer, F. Diedolo, and S. Calabrò, "Nonlinear Component Equalization: A Comparison of Deep Neural Networks and Volterra Series", in *2022 European Conference on Optical Communication (ECOC)*, 2022, pp. 1–4.