TDECQ optimization of VCSEL-MMF nonlinear digital pre-distorters using end-to-end learning

TDECQ optimization of VCSEL-MMF nonlinear Digital Pre-Distorters using End-to-end Learning

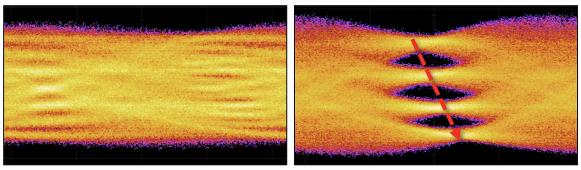
Leonardo Minelli(1), Fabrizio Forghieri(2), Ali Shahpari(3), Tong Shao(3), Roberto Gaudino(1)

(1)Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino. leonardo.minelli@polito.it
(2)CISCO Photonics Italy, Vimercate
(3)CISCO Optical GmbH, Nuremberg

Abstract *We optimize nonlinear Digital Pre-Distorters for VCSEL-MMF links using an End-to-end (E2E) learning architecture focused on TDECQ IEEE specifications for 100 Gbps/λ[1]. We experimentally demonstrate that our E2E training improves the TDECQ performance by more than 0.8 dB compared to Direct Learning.*

Introduction

Today's most widely adopted low-cost solutions for Data Center Intra-connects (DCI) up to few hundred meters are based on Intensity Modulation-Direct Detection (IM-DD) optical links, leveraging directly-modulated Vertical Cavity-Surface Emitting Lasers (VCSEL) and Multi-Mode Fibers (MMF)[2]. Next generation DCI, targeting 100 Gbps net per wavelength (λ) using PAM4 over these links, have made it necessary to establish requirements on quality measures for the transmitters, such as the Transmitter Dispersion Eye Closure Quaternary (TDECQ), to ensure vendor interoperability[1]. However, deploy VCSEL-MMF links at such high data-rate becomes challenging: in fact, TX signals are strongly impaired by the limited bandwidth of electro-optical components (see eyediagram in Fig.1.a) which bring to noise enhancement when compensated at RX[3]. Moreover, when linear Digital Pre-Distorsion (DPD) is applied at TX, nonlinear eye-skews caused by VCSELs (see red dashed line in Fig.1.b) still distort the signal[4].



(a) (b)

Fig. 1: Experimental 53.6 GBaud PAM4 signal at the VCSEL output (a) without DPD applied; (b) with linear DPD applied

As these effects strongly penalize the TDECQ performance[3], solutions such as nonlinear DPD have been investigated in order to improve the TX quality by using DSP[4]–[6]. In particular, there is strong interest in nonlinear DPDs working at 1 sample-per-symbol (sps) ratio: pre-distorted TX symbols, once calibrated at factory level, can be pre-stored inside Look-up Tables, thereby eliminating the need for complex nonlinear compensation at the RX side[4]. Still, the optimization of nonlinear DPDs remains an open research topic. In this paper, we propose a novel algorithm which explicitely focuses the optimization of the nonlinear DPDs to improve the TDECQ measure. We upgrade the Direct Learning Architecture (DLA) used in[4] into an End-to-end (E2E) learning[7] architecture, which explicitely models the DSP system specified for TDECQ standard measure[1]. Using an 850 nm VCSEL driven at 107.2 Gbps (for a net 100G net data rate), under different nonlinear conditions, we experimentally demonstrate that our TDECQ-based E2E optimization architecture significantly outperforms the DLA. By training nonlinear DPDs with both algorithms, we show a TDECQ improvement of more than 0.8 dB when comparing VCSEL output signals with the same TX average power and Optical Modulation Amplitude (OMA).

Experimental Methodology

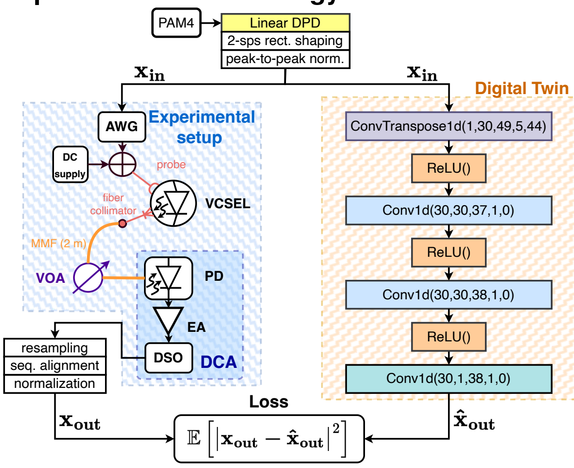


Fig. 2: Schematics of the experimental setup (left), modeled by a CNN Digital Twin for nonlinear DPD optimization (right). PD: PhotoDiode, EA: Electrical Amplifier.

We describe in this section the nonlinear DPD optimization using DLA and TDECQ-based E2E learning approaches. We test the two algorithms for PAM4 transmission at 107.2 Gbps, a conservatively higher rate than the 106.25 Gbps required by the standard[1]. We perform the experiments on the experimental setup illustrated in Fig.2: this consists of a 107.2 GSa/s Arbitrary Waveform Generator (AWG), a probed 850 nm VSCEL (Bandwidth: 22 GHz) with temperature fixed at 25 °C from a Peltier cell, a Variable Optical Attenuator (VOA) and a Keysight DCA-M N1092A Digital Communication Analyzer (DCA), able to acquire the RX signal through a Digital Sampling Oscilloscope and evaluate TDECQ and TECQ TX signal

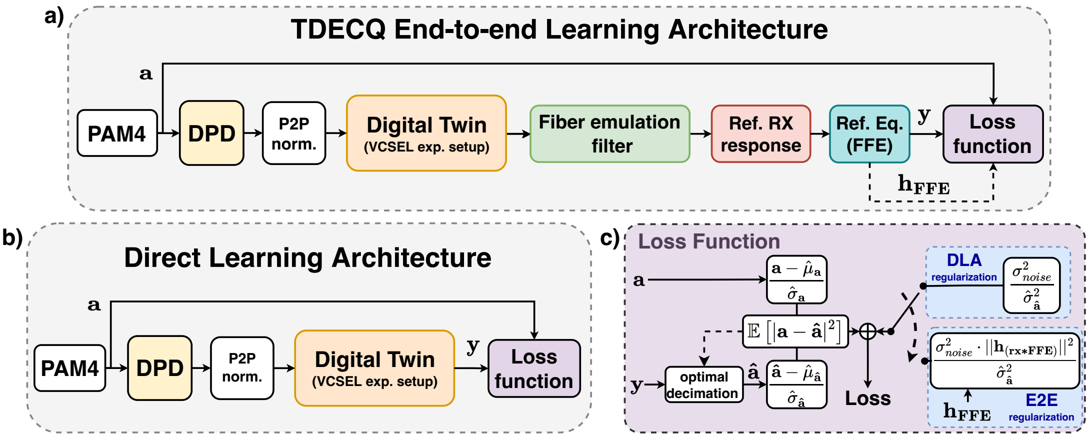


Fig. 3: Block schematics for (a) TDECQ End-to-end Architecture (b) Direct Learning Architecture (c) Loss function computation

quality metrics with an official software compliant to the standards[1]. A back-to-back configuration was adopted, as MMF impairments are digitally emulated by the DCA's TDECQ test software[1]. Both DLA and E2E approaches require two main optimization steps.

In the first one, a Convolutional Neural Network (CNN) models the input-output response of the experimental setup (see Fig.2), in order to get a digital twin of the transmitter: it is trained by minimizing the Mean Square Error between the predicted output $\hat{\mathbf{x}}_{\mathbf{out}}$ and the actual output signal of the real system $\mathbf{x}_{\mathbf{out}}$, given the same input signal vector $\mathbf{x}_{\mathbf{in}}$ transmitted through the setup (a PAM4 linearly pre-distorted PRBS pattern)[4]. The CNN is built using, in *PyTorch* notation, *Conv1d*$(ch_{in}, ch_{out}, k, s, p)$ and *ConvTranspose1d*$(ch_{in}, ch_{out}, k, s, p)$ layers ($ch_{in/out}$= input/output channels, k= kernel size, s= stride and p= padding), alternated by $ReLU()$ activation functions (see Fig.2 for layers specs). This new digital twin was able to emulate the experimental setup with an improved[4] accuracy of -24 dB of MSE (normalized w.r.t. signal power[6]) on the SSPRQ pattern[1].

In the second step, the digital twin and the nonlinear DPD to optimize are combined together to form a unique neural network, which serves as learning architecture, as illustrated in Fig3. The DLA (Fig3.b) consists of cascading the DPD to the digital twin, inserting a peak-to-peak (P2P) normalization layer at the DPD output to fulfill the VCSEL-AWG input dynamics constraints[4][8]. The TDECQ-based E2E architecture (Fig3.b) upgrades the DLA by introducing the DSP blocks required for the TDECQ measure test[1]:

1. A 18 GHz bandwidth low-pass filter emulating a worst-case MMF optical channel[1]
2. A low-pass filter with bandwidth equal to 0.5 of the Baud Rate, emulating the response of a reference RX[1], implemented as FIR filter with impulse response $\mathbf{h_{rx}}$
3. A Reference Equalizer, implemented as Feed-Forward Equalizer (FFE) with 9 taps response $\mathbf{h_{FFE}}$[1],

The DPD optimization consists of propagating (for 2000 iterations) a random PAM4 sequence $\mathbf{a}$ (3000 symbols) through the learning architecture, to then back-propagate the gradients with respect to a loss function throught the system. While in the DLA the gradients solely update the DPD coefficients, the End-to-End (E2E) system also optimizes the FFE, jointly training the TX and the RX[6]. The loss function (see Fig.3.c) main component is the MSE between the normalized system input $\mathbf{a}$ and the signal $\hat{\mathbf{a}}$: the latter is obtained by decimating the system output $\mathbf{y}$ to 1 sps at the optimum giving the lowest MSE[4]. Regularization terms are added to the DLA and E2E losses as follows:

$$L_{DLA} = \mathbb{E}\left[|\mathbf{a}-\hat{\mathbf{a}}|^2\right] + \frac{\sigma^2_{noise}}{\hat{\sigma}^2_{\hat{\mathbf{a}}}} \tag{1}$$

$$L_{E2E} = \mathbb{E}\left[|\mathbf{a}-\hat{\mathbf{a}}|^2\right] + \frac{\sigma^2_{noise} \cdot ||\mathbf{h_{(rx*FFE)}}||^2}{\hat{\sigma}^2_{\hat{\mathbf{a}}}} \tag{2}$$

where L_{DLA} and L_{E2E} are the two output losses for the DLA and E2E approaches, and the vector $\mathbf{h}^{\top}_{\mathbf{(rx*FFE)}}$ is the impulse response of the discrete linear convolution between $\mathbf{h_{rx}}$ and $\mathbf{h_{FFE}}$[6]. In Equation 1, the regularization term is introduced to the Mean Square Error (MSE) to analytically account for the effects of Gaussian noise with power σ^2_{noise} at the output of the Digital Twin[4]. The E2E loss (Equation 2) also exploits a regularization term accounting for the RX noise injection (here assumed to be inserted after the fiber emulation filter[1]), but it also models the noise enhancement induced by the FFE[6] cascaded to the RX filter. Unlike the DLA, which only targets the compensation of the TX distortions, the TDECQ-based E2E approach focuses the nonlinear DPD on improving the signal at the output of the worst-case communication system designed for the TDECQ measure.

Experimental results

In this section we report the experimental results obtained by training using both E2E and DLA optimizations two different nonlinear DPDs (the same adopted in [4]): a Volterra Nonlinear Equalizer (VNLE) and a Convolutional Neural Network

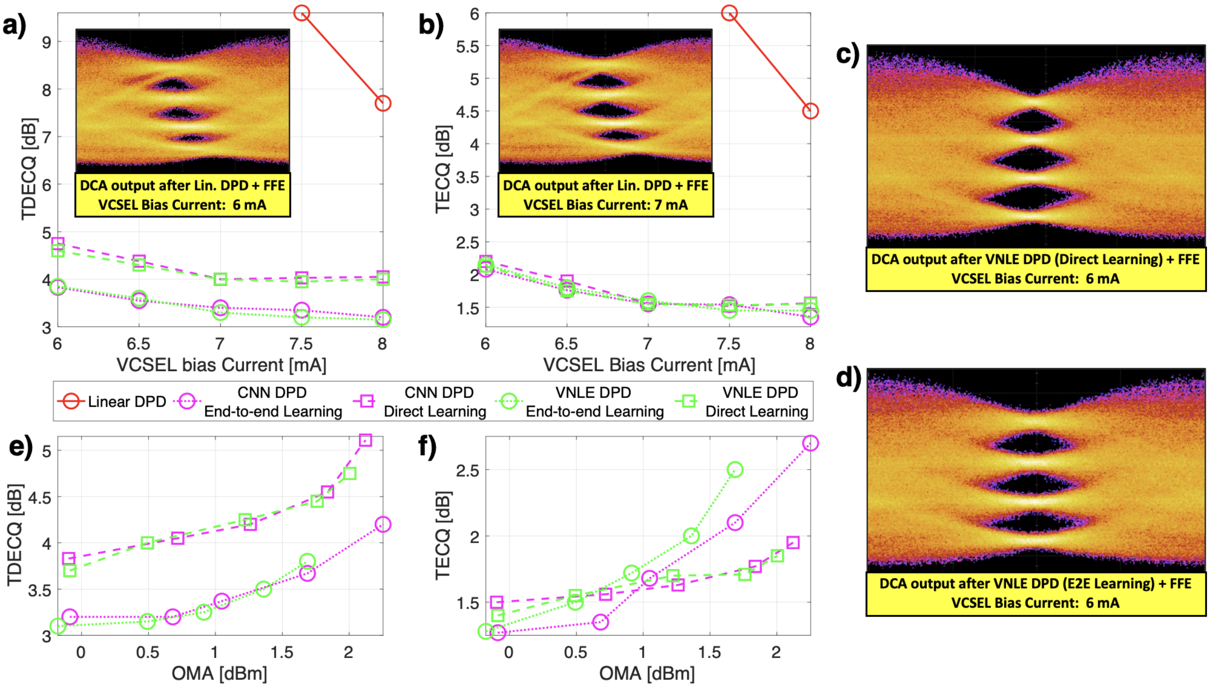


Fig. 4: Experimental performance results: (a) TDECQ vs VCSEL bias current (b) TECQ vs VCSEL bias current(e) TDECQ vs VCSEL OMA (f) TECQ vs OMA. Eyediagrams at DCA output for low bias currents showed in (a, inset),(b,inset),(c) and (d)

(CNN). We compare the performance also with respect to a linear DPD. All the DPDs had a total memory of 7 symbols. We tested the DPDs by evaluating TDECQ and TECQ through the DCA on the experimental setup (see Fig.2), using SSPRQ PAM4 pattern at 107.2 Gbps[1]. In Fig.4.a and 4.b we show the TDECQ and TECQ performance when driving the VCSEL with a AWG P2P modulation voltage set to 500 mV, varying the bias current along the typical power-efficient range expected for these lasers (6-8 mA[9]). The RX average power at the DCA input is set to 2 dBm. In all considered cases, without DPD it was not possible to achieve the target BER=2.4e-4[1], making the TDECQ measure unfeasible. Moreover, for VCSEL bias current lower than 7.5 mA, TDECQ was unmeasurable even using linear DPD, due to the severe nonlinear distortions (see inset images of Fig. 4.a and Fig. 4.b). Nonlinear DPDs instead are able to fully compensate for the nonlinear eye-skew even at the lowest bias current considered, as shown in Fig. 4.c and Fig. 4.d. The TDECQ performance using TDECQ-based E2E learning significantly improves compared to using the DLA, with a consistent gain of approximately 0.8 dB as shown in Fig. 4.a. However, E2E learning does not improve the TECQ performance of DLA-trained nonlinear DPDs, as shown in Fig. 4.b. Both DLA and E2E achieve indeed equivalent TECQ results, underscoring the specificity of the proposed E2E approach in focusing the DPD on the TDECQ optimization. Moreover, VNLE and CNN nonlinear DPDs exhibit equivalent performances either in TDECQ and TECQ, indicating that using even more sophisticated structures may not yield significant improvements. Finally, in Fig.4.e and Fig.4.f we show the TDECQ vs OMA and TECQ vs OMA results when driving the VCSEL with bias current set to 8 mA (average RX power set to 2 dBm). We vary the AWG P2P modulation voltage from 400 mV to 800 mV, to then measure the OMA together with the TDECQ and the TECQ through the DCA software. As it can be seen, using both nonlinear DPDs, the E2E outperforms the DLA TDECQ performance even when comparing signals with same OMA and RX power, with a gain of more than 0.8 dB with OMA=0.5 dBm (see Fig.4.e). Also in this case, E2E learning does not provide improvements in terms of TECQ for different OMA (Fig.4.f). However, the important achievement demonstrated by experimental results is that with the proposed E2E approach nonlinear DPD allows to meet with margin the IEEE requirements (TDECQ and TECQ less than 4.4 dB[1]) even with strong VCSEL nonlinearities caused by low input bias currents or high input modulation swings.

Conclusion

In this paper, we propose a TDECQ-based E2E learning approach for optimizing nonlinear DPDs on VCSEL+MMF links. The new approach experimentally shows to significantly improve the TDECQ performance with respect to DLA, demonstrating that nonlinear DPD optimization focused on TDECQ can improve this metric.

Acknowledgements

This work was carried out under a research contract with Cisco Photonics. We also acknowledge the PhotoNext Center at Politecnico di Torino (http://www.photonext.polito.it/) and Cisco Optical GmbH at Nuremberg.

References

[1] "IEEE Standard for Ethernet - Amendment 3: Physical Layer Specifications and Management Parameters for 100 Gb/s, 200 Gb/s, and 400 Gb/s Operation over Optical Fiber using 100 Gb/s Signaling", *IEEE Std 802.3db-2022 (Amendment to IEEE Std 802.3-2022 as amended by IEEE Std 802.3dd-2022 and IEEE Std 802.3cs-2022)*, pp. 1–73, 2022. DOI: 10.1109/IEEESTD.2022.9988984.

[2] P. Torres-Ferrera, G. Rizzelli, A. Nespola, *et al.*, "Statistical Analysis of 100 Gbps per Wavelength SWDM VCSEL-MMF Data Center Links on a Large Set of OM3 and OM4 Fibers", *Journal of Lightwave Technology*, vol. 40, no. 4, pp. 1018–1026, 2022. DOI: 10.1109/JLT.2021.3129455.

[3] S. Echeverri-Chacón, J. J. Mohr, J. J. V. Olmos, *et al.*, "Transmitter and Dispersion Eye Closure Quaternary (TDECQ) and Its Sensitivity to Impairments in PAM4 Waveforms", *Journal of Lightwave Technology*, vol. 37, no. 3, pp. 852–860, 2019. DOI: 10.1109/JLT.2018.2881986.

[4] L. Minelli, F. Forghieri, T. Shao, A. Shahpari, and R. Gaudino, "Net 100 gb/s/λ VCSEL+MMF nonlinear Digital Pre-Distortion using Convolutional Neural Networks", in *Optical Fiber Communication Conference (OFC) 2023*, Optica Publishing Group, 2023, Tu3I.4. [Online]. Available: https://opg.optica.org/abstract.cfm?URI=OFC-2023-Tu3I.4.

[5] J. Zhang, P. Gou, M. Kong, *et al.*, "PAM-8 IM/DD Transmission Based on Modified Lookup Table Nonlinear Predistortion", *IEEE Photonics Journal*, vol. 10, no. 3, pp. 1–9, 2018. DOI: 10.1109/JPHOT.2018.2828869.

[6] L. Minelli, F. Forghieri, A. Nespola, S. Straullu, and R. Gaudino, "A Multi-Rate Approach for Nonlinear Pre-Distortion Using End-to-End Deep Learning in IM-DD Systems", *Journal of Lightwave Technology*, vol. 41, no. 2, pp. 420–431, 2023. DOI: 10.1109/JLT.2022.3216591.

[7] B. Karanov, M. Chagnon, F. Thouin, *et al.*, "End-to-end deep learning of optical fiber communications", *Journal of Lightwave Technology*, vol. 36, no. 20, pp. 4843–4855, 2018. DOI: 10.1109/JLT.2018.2865109.

[8] L. Minelli, F. Forghieri, and R. Gaudino, "Nonlinear Pre-distortion through a Multi-rate End-to-end Learning Approach over VCSEL-MMF IM-DD Optical Links", in *2022 European Conference on Optical Communication (ECOC)*, 2022, pp. 1–4.

[9] J. A. Tatum *et al.*, "VCSEL-Based Interconnects for Current and Future Data Centers", *Journal of Lightwave Technology*, vol. 33, no. 4, pp. 727–732, 2015. DOI: 10.1109/JLT.2014.2370633.