# Experimental VCSEL Digital Twin modeling for net 100 Gb/s/$\lambda$ nonlinear Digital Pre-Distortion

Leonardo Minelli
*Dipartimento di Elettronica e Telecomunicazioni*
*Politecnico di Torino*
Torino, Italy
leonardo.minelli@polito.it

Fabrizio Forghieri
*CISCO Photonics*
Vimercate, Italy

Roberto Gaudino
*Dipartimento di Elettronica e Telecomunicazioni*
*Politecnico di Torino*
Torino, Italy

***Abstract*—We experimentally model a VCSEL-based optical transmitter for high speed intra data center interconnects using a convolutional neural network digital twin. The device is able to effectively reproduce the VCSEL linear and nonlinear distortions on PAM4 signals transmitted at 107.2 Gbps, thus enabling the optimization of nonlinear VCSEL-MMF digital pre-distorters.**

***Index Terms*—VCSEL, Neural Networks, Digital Pre-Distortion, Data Center Intra-connects**

## I. Introduction

Today's prevailing solution for short-reach optical intra-Data Center Inter-Connects (DCI) consists of Intensity-Modulation Direct-Detection (IM-DD) optical links. Specifically, for DCI up to 100 meters, Multi-Mode Fibers (MMF) and Vertical-Cavity Surface-Emitting Lasers (VCSEL) are utilized. As the next generation of VCSEL-MMF DCI is expected to provide a net 100 Gb/s rate per wavelength ($\lambda$), these links will deploy quaternary pulse amplitude modulation (PAM4), using Digital Signal Processing (DSP) at the transmitter (TX) and the receiver (RX) side to support the signal reliability [1]. The use of nonlinear Digital Pre-Distorters (DPD) has recently been investigated to compensate for bandwidth limitations and nonlinear VCSEL distortions at such high data rates. Figure 1 shows the experimental eye-diagrams when transmitting a PAM4 signal at 107.2 Gbps through an 850 nm VCSEL with 22 GHz bandwidth. Without applying DPD (Fig.1.a), the

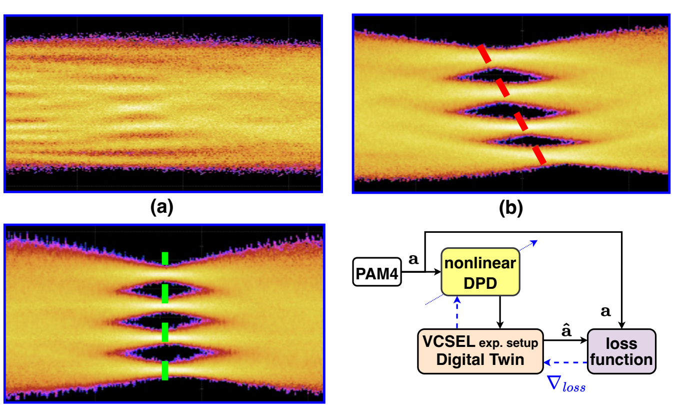


Fig. 1: Experimental PAM4 107.2 Gbps VCSEL output eye-diagrams (a) without applying DPD; (b) using linear DPD; (c) using nonlinear DPD. (d) DLA scheme for VCSEL DPD optimization.

PAM4 eye is fully closed at the VCSEL output, while applying linear DPD can compensate for the bandwidth limitations, as shown in Fig.1.b, only nonlinear DPD (Figure 1.c) is also able to remove the asymmetric time-domain eye-skew caused by the VCSEL typical nonlinear behaviors (see red dashed line in Fig.1.b). However, nonlinear DPDs for VCSEL-MMF links require sophisticated optimization techniques, such as the Direct Learning Architecture (DLA) [2], [3]. The latter constists of a convolutional neural network (CNN) made by the cascade of the DPD and a digital twin of the VCSEL-based TX, and trained by updating only the DPD coefficients. Figure 1.d shows a the baseline schematics of the DLA-based optimization. Alternative approaches, such as End-to-end (E2E) learning, have been also applied for VCSEL-MMF links [4], [5]. However, both DLA and E2E approaches requires a digital twin of the TX experimental setup as a key "building block" for optimizing the nonlinear DPD. In this paper, we obtain this surrogate model from an experimental VCSEL-MMF TX setup: we train a CNN on the input-output response of a VCSEL-based TX, enabling the optimization of nonlinear DPDs for net 100 Gbps/$\lambda$ rate [2].

## II. Digital Twin optimization Methodology

We experimentally model using a CNN the back-to-back VCSEL setup shown in Figure 2. It consists of a 107.2 GSa/s Arbitrary Waveform Generator (AWG), driving with 500 mVpp an 850 nm probed VCSEL ($B_{3dB}$=22 GHz), supplied by 8 mA of DC current. The emitted optical waveforms are then collimated into 2 meters of MMF, to get then received by a Digital Communication Analyzer: this comprises a PhotoDiode (PD), an Electrical Amplifier (EA) and a Digital Sampling Oscilloscope (DSO). The CNN is built using a transposed 1D convolution layer, followed by $ReLU$ activation functions alternated to 1D convolution layers. Each hidden convolution layer has 30 channels [2], and kernel sizes are designed to get a CNN with a memory equivalent to a Finite Impulse Response (FIR) filter with 150 taps. The CNN is trained to predict the output signal $\mathbf{y}$ obtained from the experimental setup (see Figure 2), given the same input TX signal $\mathbf{x}$. During the training, the TX signal $\mathbf{x}$ is a linearly pre-distorted PRBS PAM4 signal at 53.6 GBaud, with periodicity of $2^{16}$ symbols: this allows to model the digital twin in a VCSEL driving condition close to the actual nonlinear DPD signal transmission. Before optimizing the digital twin, the signal $\mathbf{y}$ is processed by the DCA to mitigate the noise impairment, i.e. averaging the RX signal over the PRBS period. It is then resampled to a 10 samples-per-symbol ratio, to then be aligned with the TX sequence and normalized to zero mean and unitary variance. The training consists of an iterative procedure where the CNN predicts the experimental output $\hat{\mathbf{y}}$ :

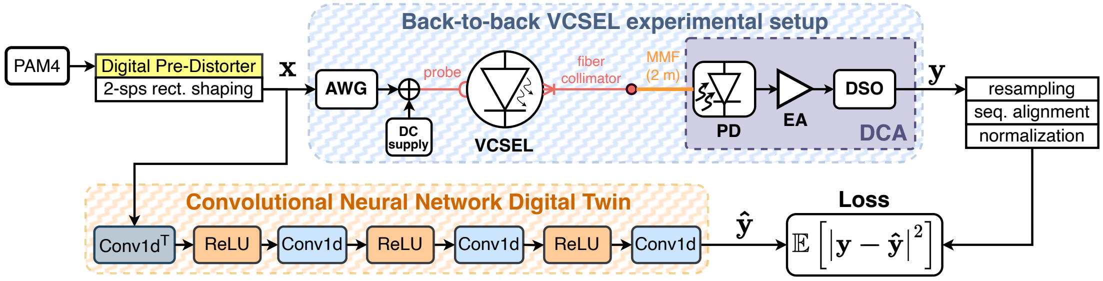


Fig. 2: Experimental modeling of a back-to-back VCSEL setup using a Convolutional Neural Network Digital Twin. $Conv1d^{\top}$: Transpose 1D convolution layer; $Conv1d$: 1D convolution layer

the Mean Square Error (MSE) between $\hat{\mathbf{y}}$ and $\mathbf{y}$ is computed, and the CNN coefficients are updated based on their MSE gradient [2]. This process is performed on a random mini-batch of 1000 symbols of $\mathbf{y}$ and repeated for 700 iterations. With the provided methodology, the obtained digital twin effectively models the deterministic response of the VCSEL when transmitting pre-distorted PAM4 signals at the target 107.2 Gbps rate, disregarding random noise impairments: these are instead efficiently introduced semi-analytically during the DPD optimization as an additive regularization term in the loss function [4].

## III. Experimental Results

After the optimization of the digital twin, we evaluate the goodness of the fit with the experimental setup, using a different experimental acquisition to prevent overfitting. The test TX signal $\mathbf{x}$ is a nonlinearly pre-distorted PAM4 signal at 53.6 GBaud, generated from a PRBS different from the training one, with periodicity of $2^{16}$: this also allows to assess the digital twin reliability in the target optimal VCSEL driving condition (i.e., as in Figure1.c). To evaluate also the improvements induced by the nonlinear model, we compare its fit performance with a linear FIR filter digital twin, with the same memory of the CNN. The predicted output eyediagrams for the two surrogate models are shown in Figures 3.a and 3.b. Noticeably, the FIR filter digital twin shows an inverse time-

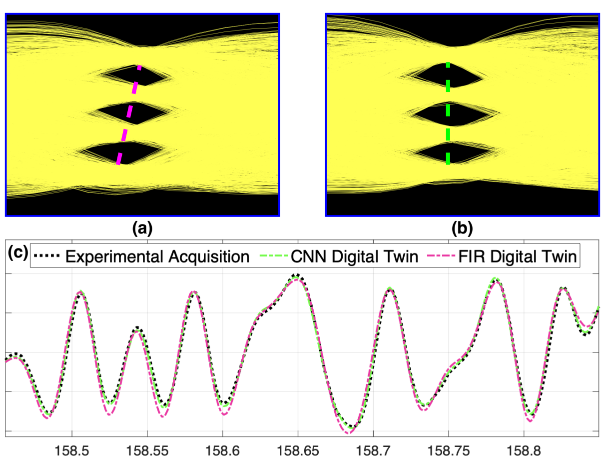


Fig. 3: Simulated VCSEL output (a) eye-diagram using FIR digital twin; (b) eye-diagram using CNN digital twin. (c) Time-domain comparison between simulated and actual VCSEL output

domain eye-skew (see magenta-dashed line in Fig.3.a): this is due to the inability of the FIR to synthesize the VCSEL nonlinear effect, thus mantaining the pre-compensated skew induced by the nonlinear DPD. The CNN digital twin is instead able to fully reproduce the VCSEL behavior. Figure 3.c then compares the two digital twins predictions with the actual experimental output: the CNN digital twin clearly shows to better fit the real signal than the FIR filter. In order to quantify the model accuracies, we report in Table I the normalized root MSE (NRMSE) between $\hat{\mathbf{y}}$ and $\mathbf{y}$ both in dB and in terms of fit percentage, expressed as follows:

$$fitness\ (\%) = 100 \cdot \left(1 - \frac{||\mathbf{y}-\hat{\mathbf{y}}||}{||\mathbf{y}-\mathbb{E}[\mathbf{y}]||}\right) \tag{1}$$

TABLE I:

| Digital Twin Architecture | NRMSE fitness (%) | NRMSE [db] |
|---|---|---|
| Linear FIR filter | 88.9 % | -19.1 dB |
| Convolutional Neural Network | 93.7 % | -24.0 dB |

As shown in the Table, the CNN shows to accurately model the VCSEL TX, with a NRMSE equal to -24 dB. It also significantly outperforms the FIR digital twin, with a gain of 4.9 dB in terms of NRMSE and a 4.8 % of fitness improvement.

## IV. Conclusion

In this paper, we experimentally demonstrate a CNN able to reliably emulate the VCSEL linear and nonlinear distortions on PAM4 signals at 107.2 Gbps. The effective optimization definitively enables a reliable nonlinear DPD optimization for VCSEL-MMF optical links.

## 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," in IEEE Std 802.3db-2022, pp.1-73, 2022.

[2] 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," 2023 Optical Fiber Communications Conference and Exhibition (OFC), San Diego, CA, USA, 2023, pp. 1-3

[3] G. Paryanti, H. Faig, L. Rokach and D. Sadot, "A Direct Learning Approach for Neural Network Based Pre-Distortion for Coherent Nonlinear Optical Transmitter," in Journal of Lightwave Technology, vol. 38, no. 15, pp. 3883-3896, 2020

[4] 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" in Journal of Lightwave Technology, vol. 41, no. 2, pp. 420-431, 2023

[5] 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," 2022 European Conference on Optical Communication (ECOC), Basel, Switzerland, 2022, pp. 1-4.