Learning Optimal PAM Levels for VCSEL-based Optical Interconnects

Learning Optimal PAM Levels for VCSEL-based Optical Interconnects

Muralikrishnan Srinivasan[(1)], Jinxiang Song[(1)], Christian Häger[(1)], Krzysztof Szczerba[(2)], Henk Wymeersch[(1)], and Jochen Schröder[(3)] *

[(1)] Dept. of Electrical Engineering, Chalmers Univ. of Technology, Gothenburg, Sweden
[(2)] Juniper Networks, Sunnyvale, CA, United States
[(3)] Dept. of Microtechnology and Nanoscience, Chalmers Univ. of Technology, Gothenburg, Sweden
mursri@chalmers.se

Abstract *An auto-encoder that optimizes a VCSEL-based fiber-optic system end-to-end and provides a 1.5dB sensitivity gain at higher temperatures is trained, utilizing a neural network that models the response of a VCSEL for a range of operating temperatures.*

Introduction

The uncooled GaAs-based Vertical-Cavity Surface-Emitting Laser (VCSEL) is the light source of choice for short-reach optical interconnects (OIs)[1]. It can be directly modulated at high speed and enables miniature footprint transceiver modules, and the most energy and cost-efficient interconnect[2]. Intensity modulation formats such as 4-PAM can be used to provide high data rates in OIs[3],[4]. However, the non-trivial nonlinear VCSEL response limits OIs modulated with equidistant PAM levels[5]. As optics moves closer to the board and OIs find uses in diverse applications, and settings like consumer devices, cars, or planes, operation over broader temperature ranges, e.g., from -40 to $125℃$ becomes imperative. Therefore, it is essential to model the VCSEL and adapt the modulation format to maximize the data throughput from the lowest to the highest temperature for the OIs.

Machine learning (ML) techniques have been used to adapt the performance of OIs, especially for equalization[6]–[8]. ML has also been applied to optimize a specific function in the fiber-optic system, e.g., coding, modulation, and equalization[9]–[11]. End-to-end learning using autoencoders (AE) for jointly optimizing the transmitter and receiver has the potential to achieve an optimal performance[12]–[17]. However, there is scant literature on applying AE to VCSEL links, in particular, when including temperature dependence. The differential equations governing the VCSEL are not directly amenable to gradient back-propagation and depend on the differential equations solver[18]. Therefore, incorporating the rate equations in the end-to-end optimization of AEs is challenging. Moreover, equivalent circuit models such as[19] are also not sufficient for incorporating into the AE optimization. Therefore, to fully capture the nonlinear behavior of VCSELs and optimize the PAM levels using learning techniques, a neural network (NN) equivalent that captures the dynamics of the rate equations of VCSELs is highly advantageous. Initial work along these lines was presented in[20] but failed to include the temperature dependence.

Our contributions in this paper are two-fold:

1. We first develop a VCSEL NN model (referred to as VNet) that models the nonlinear dynamic response of the VCSEL to accurately reproduce the optical waveforms generated by the VCSEL under different operating temperatures.
2. We incorporate the VNet in a low-complexity AE that models an end-to-end fiber-optic system, and we use it to determine optimized input PAM levels that provide improved bit error rate performance.

VCSEL Neural Network - VNet

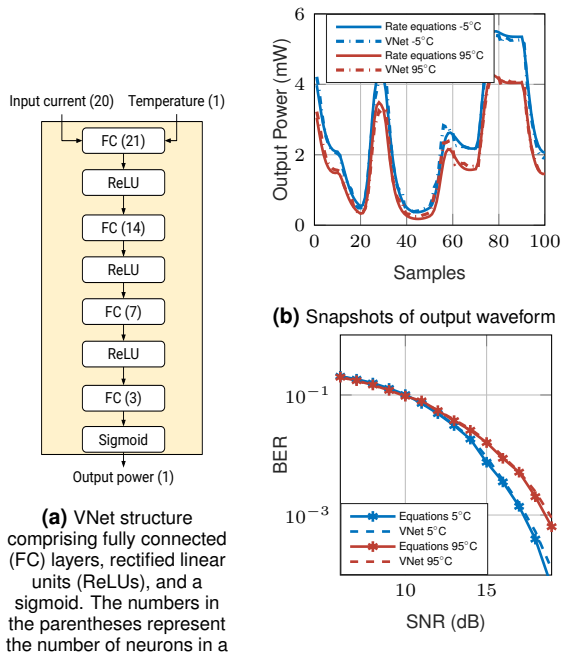
For this first proof-of-concept demonstration, we utilize the single-mode laser rate equations of a simple VCSEL model[20]

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \frac{I}{qV} - \frac{c}{n_{\text{geff}}} gS - \frac{N}{\tau_n} \tag{1}$$

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \Gamma \frac{c}{n_{\text{geff}}} gS + \Gamma\beta \frac{N}{\tau_n} - \frac{S}{\tau_p}, \tag{2}$$

where N denotes carrier density, S the photon density, I the injection current, V the active volume, n_{geff} the effective refractive index, g the

*Presented at European Conference on Optical Communication 2022, Early Access



(a) VNet structure comprising fully connected (FC) layers, rectified linear units (ReLUs), and a sigmoid. The numbers in the parentheses represent the number of neurons in a specific layer

(b) Snapshots of output waveform

(c) BER performance

Fig. 1: VNet structure, and equivalence to the rate equations in terms of response and BER.

gain per unit length, Γ the internal quantum efficiency, τ_n the carrier lifetime and τ_p the photon lifetime. Here we consider Γ, g, τ_n and τ_p to be the temperature-dependent terms. Note that the model does not account for parasitics, gain compression, and self-heating of the VCSEL.

By generating input-current and output-power samples obtained from (1)–(2) using a 28 GHz white Gaussian noise sequence with 8 mA mean current and 6 mA standard deviation for temperatures -5℃ to 95℃ in steps of 5℃, the VNet NN (structure shown in Fig. 1a) is trained in a supervised fashion to capture the VCSEL dynamics in the relevant parameter ranges. The input current sequence is over-sampled by a factor 10 and shaped into rectangular pulses. The current sequence is then partitioned into 20 consecutive samples, which, along with the temperature, serve, after normalization, as the NN's input, thus avoiding the need for retraining for different temperatures in[20].

The trained network is tested with a 28 Gbaud 4-PAM current sequence ranging between 2 mA and 12 mA, with equidistant modulated levels. A snapshot of the matching output power sequences from the rate equations and the VNet for two different temperatures is shown in Fig. 1b. To further validate the accuracy of VNet, the bit error rate (BER) vs. the signal-to-noise ratio (SNR) obtained from VNet and rate equations for a 4-PAM input sequence over an additive white Gaussian noise (AWGN) channel are shown in Fig. 1c. The

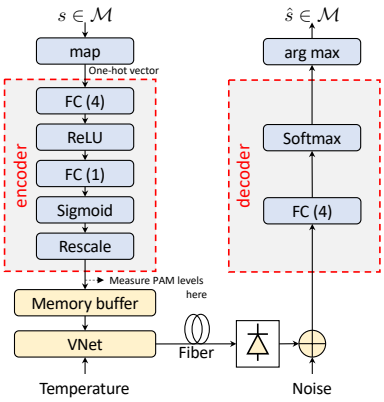


Fig. 2: AE structure containing two encoder layers, the transmission blocks and a single decoder layer. The numbers in the parentheses represent the number of neurons in a specific layer.

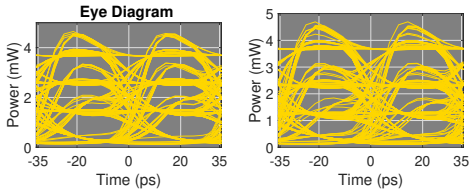
performance again matches at different temperatures (shown only for 5℃ and 95℃, for clarity).

Autoencoder for Optimized PAM Levels

This section implements the fiber-optic communication system and transmission chain, including transmitter, receiver, and channel, as a complete end-to-end AE. All the components of the transceiver chain, as well as the channel model, are depicted in detail in Fig. 2. A message $s \in [1, \ldots, M] \triangleq \mathcal{M}$ is encoded into a one-hot vector of size M (here, $M = 4$), denoted by $\mathbf{x}$, where the s-th element equals 1 and the other elements are 0. Then the encoder is applied (see Fig. 2, left side). The output of the sigmoid layer, which is between $[0, 1]$, is appropriately scaled and shifted to the chosen dynamic range, here chosen to be $[2, 12]$ mA, which ensures that the input to VNet is above the VCSEL threshold current for all ambient temperatures and avoids that the AE arbitrarily increases bias current, which would be penalized by self-heating in a real system[20]. The output of the encoder is over-sampled by a factor of 10 (additional low-pass filtering can be applied here) and stored in memory buffers to accumulate 20 samples, which corresponds to 2 symbol durations, concatenated with the temperature, and passed to the VNet (described in Fig. 1). The fiber channel is modeled by AWGN, though the proposed architecture can be readily extended to account for additional low-pass filtering and other forms of dispersion and complex circuitry, like output driver circuits. The photodiode output is processed using a single fully-connected layer with softmax activation. The output probability vector $\mathbf{y} = [y_1, \ldots, y_M]^\top$, and estimated message is $\hat{s} = \arg\max_i [\mathbf{y}]_i$. The loss function minimized is the categorical cross-entropy, given

Tab. 1: AE determined PAM levels (in mA) vs temperature, where the first row corresponds to equidistant PAM.

T °C	PAM-1	PAM-2	PAM-3	PAM-4
–	2.0	5.33	8.67	12.0
5	2.0	5.06	8.56	12.0
25	2.0	5.05	8.50	12.0
45	2.0	5.05	8.41	12.0
65	2.0	4.86	8.21	12.0
85	2.0	4.76	8.15	12.0
95	2.0	4.73	7.87	12.0



(a) For Equidistant PAM **(b)** For AutoEncoder PAM
Fig. 3: Eye Diagrams of output power for 95℃.

by $-\sum_{i=1}^{M} x_i \log(y_i)$. This loss can be related to an achievable information rate using arguments from mismatched decoding[15].

Performance Analysis

The training is performed over a dataset of 2.5×10^4 randomly chosen message symbols with a batch size of 50 for an SNR of 21 dB for 1500 epochs. At the end of the training procedure, the optimized PAM levels are measured at the output of the rescaling layer.

The PAM levels obtained from the AE for various temperatures are shown in Tab. 1 along with equidistant PAM levels. At 5℃, the PAM levels are close to that of the equidistant PAM. However, at 95℃, the spacing between the first two levels is 2.73 mA, and between the second and the third is 3.14 mA. At higher temperatures, the AE obtains an almost equal vertical eye-opening of the output power for all four levels by compressing the lower levels. The equal eye-opening for all the levels for the AE-optimized PAM can also be observed in the eye diagrams shown in fig. 3. The chosen driver current range is only for proof-of-concept, and the AE can be trained to obtain a different extinction ratio.

For a fair comparison of the cross-entropy performance of the AE-obtained PAM levels against that of the equidistant PAM levels, we have trained a stand-alone decoder. The cross-entropy vs. SNR results of the AE and the equidistant-decoder (ED) are plotted in Fig. 4. The cross-entropy of AE is lower than that of ED for all temperatures. For example, at 95℃, the cross-entropy of AE is 5 times smaller than that of ED. Even at lower temperatures, the cross-entropy of AE is 2 times smaller than that of ED. The BER performance of both AE and ED are shown in

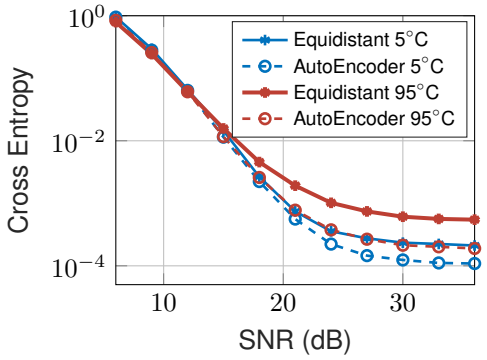


Fig. 4: Cross Entropy vs SNR for AE and ED.

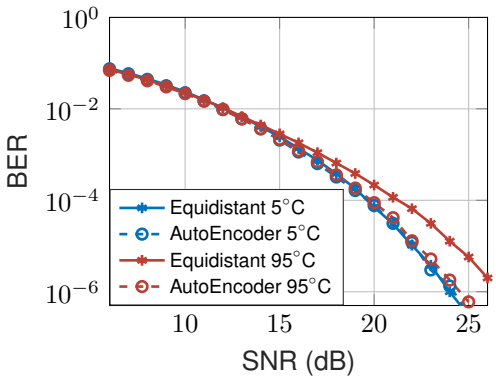


Fig. 5: BER vs SNR of AE and ED.

Fig. 5. The advantage of AE is that even though the loss function minimized is the cross-entropy, the BER of AE is also lower than that of ED at higher temperatures. At 95℃, the sensitivity improvement is about 1.5dB, validating the idea that equidistant PAM levels do not attain an optimal BER. Also, the training and computational complexity of the proposed AE method are low when compared to existing ML-based nonlinear equalization approaches[8]. Also, in a real system, training would likely be performed in a calibration step, and optimized PAM levels could be integrated into the driver and receiver hardware.

Conclusions

We demonstrated AE-based end-to-end optimization of VCSEL-based fiber-optic system under different operating temperatures utilizing a NN-based VCSEL model . The optimized 4-PAM input levels to the VCSEL are shown to have about 1.5 dB sensitivity improvement when compared to that of equidistant 4-PAM, at higher temperatures. Also, the optimized PAM levels are shown to have an improved cross-entropy across a range of operating temperatures.

Acknowledgements

This work was supported by the Swedish Foundation for Strategic Research (SSF, HOT-OPTICS

Project). The work of C. Häger was also supported by the Swedish Research Council (grant no. 2020-04718).

References

[1] D. Mahgerefteh *et al.*, "Techno-economic Comparison of Silicon Photonics and Multimode VCSELs," *Journal of Lightwave Technology*, vol. 34, no. 2, pp. 233–242, 2016.

[2] A. Larsson *et al.*, "VCSEL Modulation Speed: Status and Prospects," in *Vertical-Cavity Surface-Emitting Lasers XXIII*, SPIE, vol. 10938, 2019, p. 1 093 802.

[3] K. Szczerba *et al.*, "30 Gbps 4-PAM Transmission over 200m of MMF using an 850-nm VCSEL," *Optics express*, vol. 19, no. 26, B203–B208, 2011.

[4] F. Karinou *et al.*, "IM/DD vs. 4-PAM using a 1550-nm VCSEL over Short-range SMF/MMF links for Optical Interconnects," in *OFC*, 2013.

[5] C.-Y. Chuang *et al.*, "Employing Deep Neural Network for High Speed 4-PAM Optical Interconnect," in *2017 European Conference on Optical Communication (ECOC)*, IEEE, 2017, pp. 1–3.

[6] L. Ge *et al.*, "Compressed Neural Network Equalization based on Iterative Pruning Algorithm for 112-Gbps VCSEL-enabled Optical Interconnects," *Journal of Lightwave Technology*, vol. 38, no. 6, pp. 1323–1329, 2020.

[7] F. Tian and C. Yang, "Deep Belief Network-hidden Markov Model based Nonlinear Equalizer for VCSEL based Optical Interconnect," *Science China Information Sciences*, vol. 63, no. 6, pp. 1–9, 2020.

[8] Q. Zhou *et al.*, "Low Computationally Complex Recurrent Neural Network for High Speed Optical Fiber Transmission," *Optics Communications*, vol. 441, pp. 121–126, 2019.

[9] J. Thrane *et al.*, "Machine Learning Techniques for Optical Performance Monitoring from Directly Detected PDM-QAM Signals," *Journal of Lightwave Technology*, vol. 35, no. 4, pp. 868–875, 2016.

[10] D. Zibar *et al.*, "Machine Learning Techniques in Optical Communication," *Journal of Lightwave Technology*, vol. 34, no. 6, pp. 1442–1452, 2015.

[11] F. N. Khan *et al.*, "Modulation Format Identification in Heterogeneous Fiber-optic Networks using Artificial Neural Networks," *Optics express*, vol. 20, no. 11, pp. 12 422–12 431, 2012.

[12] B. Karanov *et al.*, "End-to-end Deep Learning of Optical Fiber Communications," *Journal of Lightwave Technology*, vol. 36, no. 20, pp. 4843–4855, 2018.

[13] W. Jiang *et al.*, "End-to-end Learning based Bit-wise Autoencoder for Optical OFDM Communication System," in *Asia Communications and Photonics Conference*, Optical Society of America, 2021, T4A–64.

[14] 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, 2017.

[15] S. Li *et al.*, "Achievable Information Rates for Nonlinear Fiber Communication via End-to-end Autoencoder Learning," in *2018 European Conference on Optical Communication (ECOC)*, IEEE, 2018, pp. 1–3.

[16] O. Jovanovic *et al.*, "End-to-end Learning of a Constellation Shape Robust to Channel Condition Uncertainties," *Journal of Lightwave Technology*, 2022.

[17] V. Aref and M. Chagnon, "End-to-end Learning of Joint Geometric and Probabilistic Constellation Shaping," in *2022 Optical Fiber Communications Conference and Exhibition (OFC)*, IEEE, 2022, pp. 1–3.

[18] R. T. Chen *et al.*, "Neural Ordinary Differential Equations," *Advances in neural information processing systems*, vol. 31, 2018.

[19] A. Grabowski *et al.*, "Large-Signal Equivalent Circuit for Datacom VCSELs," *Journal of Lightwave Technology*, vol. 39, no. 10, pp. 3225–3236, 2021.

[20] K. Szczerba and C. Kocot, "Behavioral Modeling of VCSELs for High-speed Optical Interconnects," in *Vertical-Cavity Surface-Emitting Lasers XXII*, International Society for Optics and Photonics, vol. 10552, 2018, p. 1 055 204.