POLITECNICO DI TORINO
Repository ISTITUZIONALE

Machine Learning Aided Control of Ultra-Wideband Indium Phosphide IQ Mach-Zehnder Modulators

Original
Machine Learning Aided Control of Ultra-Wideband Indium Phosphide IQ Mach-Zehnder Modulators / D'Ingillo, R., D'Amico, A., Usmani, F., Borraccini, G., Straullu, S., Siano, R., Belmonte, M., Curri, V.. - ELETTRONICO. - (2023), pp. 1-3. (2023 International Conference on Photonics in Switching and Computing (PSC) Mantova, Italy 26-29 September 2023) [10.1109/PSC57974.2023.10297214].

Availability:
This version is available at: 11583/2983886 since: 2023-11-20T14:49:55Z

Publisher:
IEEE

Published
DOI:10.1109/PSC57974.2023.10297214

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
IEEE postprint/Author's Accepted Manuscript

©2023 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

19 November 2026

Machine Learning Aided Control of an Ultra-Wideband Indium Phosphide IQ Mach-Zehnder Modulator

Rocco D'Ingillo[(1)], Andrea D'Amico[(1)], Fehmida Usmani[(1),(2)], Giacomo Borraccini[(1)], Stefano Straullu[(3)], Rocco Siano[(4)], Michele Belmonte[(4)], Vittorio Curri[(1)]

[(1)]Politecnico di Torino, Italy, rocco.dingillo@polito.it;
[(2)]National University of Sciences & Technology (NUST), Pakistan; [(3)]LINKS Foundation, Italy; [(4)]Lumentum, Italy

Abstract *A digital model of a dual-polarization IQ ultra-wideband indium phosphide Mach-Zehnder modulator is obtained through machine learning techniques. The model is used to test optimization algorithms that automatically set the modulator control voltages under different operative conditions finding the optimum bias point.* ©2023 The Author(s)

Introduction

Mach-Zehnder Modulators (MZMs) play a critical role in optical communication, and their performance heavily relies on device bias point control. Indium Phosphide (InP) technology is promising for MZM fabrication as it can be easily integrated with other photonics components, enabling a more consistent usage of photonic integrated circuits (PICs)[1]. On the other hand, the material characteristics make MZM control more complex due to the nonlinear behavior of the device[2], making the search for the optimal bias point challenging. There are two main techniques to optimize the operating bias point control: optical power-based[3]–[5] and dither-based techniques[6]–[8]. Power-based techniques use photodetectors (PDs) to monitor the modulator's output optical power, adjusting the bias voltage until the desired optical power is achieved. Conversely, dither-based techniques use a sinusoidal voltage signal to modulate the bias voltage and monitor the modulator output to maintain a constant output by adjusting the bias voltage[9]. The power-based technique is simpler and can be implemented by means of a minimum laboratory equipment, while the dither-based technique is more robust to noise and optical power fluctuations at the price of more complex circuitry. This paper presents a novel approach to search and control the bias point of an InP ultra-wideband (UWB) dual polarization (DP) IQ-MZM using machine learning (ML) techniques to generate a digital model of the MZM. In recent years, the use of ML and artificial intelligence (AI) has spread widely in the photonics industry[10],[11]. In this study, an AI-based system learns the behavior of the MZM and applies optimization algorithms on the predicted model to find its optimal bias point automatically, following a promising strategy also proposed in recent works[12],[13]. The same techniques have been applied to the actual component in laboratory, following a ML-aided power-based bias point control technique. The algorithm has been validated for a structure-agnostic approach by observing different modulator settings aiming to reach optimal performance using a limited number of measurements and minimum laboratory equipment. The presented ML approach aims to be the starting point for modulator control algorithms moving towards a full real-time digital twin aided control of the component through reinforcement or transfer learning techniques[14].

Mach-Zehnder Modulator Bias Point Control

Fig. 1 shows the simplified internal structure of the InP DP-MZM under test, which is a Lumentum high-bandwidth coherent driver modulator (HB-CDM) with four-channel modulator driver integrated circuits and two nested modulators designed to modulate amplitude and phase of the input light in both polarization states. Monitoring PDs are present on the output of both X and Y polarization and after the Polarization Beam Combiner (PBC). The module has one input and one output fiber and control of the MZMs is managed via analog differential voltage inputs to the module through the bias signals *I*, *Q* and *Ph*, which are applied to the DC electrodes of the corresponding MZMs. The radio frequency (RF) input contacts are located at the rear of the package[15]. To achieve Quadrature Phase Shift Key-

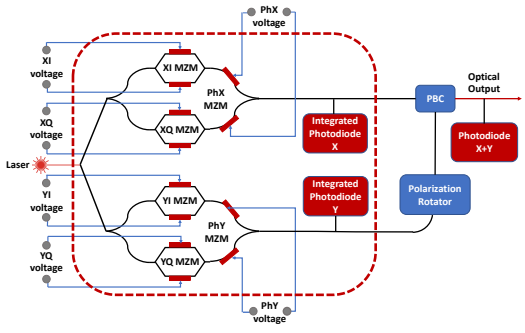


Fig. 1: Internal structure of the DP IQ MZM; the integrated components are highlighted in the red dashed box.

ing (QPSK) and Quadrature Amplitude Modulation (QAM) modulation formats, the bias signals are evaluated in order to reach a *Min/Min/Quad* bias condition of the DP-MZM[5],[16]. In this condition, both internal *I* and *Q* MZMs are biased at the minimum of their respective response function and the phase bias signal *Ph* is adjusted to maintain the quadrature between the two modulated branch signals[17],[18].

Experimental Setup

To study and test different bias point search and control strategies, a data collection of power measurements for different input combinations of XI, XQ, and PhX control voltages has been performed using the experimental setup shown in Fig. 2. The Y polarization modulator has been turned off, keeping only the X polarization on, for all the duration of the data collection. To observe the optical carrier behavior, the HB-CDM module is connected to a Lumentum evaluation board controlled by a workstation via USB-to-SPI communication. The modulating signal is managed through a 8 bit 92 GS/s digital to analog converter (DAC) controlled by the workstation. The optical input is generated by a continuous wave (CW) laser centered at 1550 nm with a launch power of 10 dBm. The optical output is split between a Teledyne Lecroy Optical Modulation Analyzer (OMA) and a Keysight 8163A optical power meter through a 50:50 optical splitter for observing the constellation diagrams and power measurements, respectively. The power meter is also connected to workstation via GPIB-USB interface to monitor the power measurements through the control terminal.

Modulator Machine-Learning Model

Due to the fine granularity of each input bias voltage (1 mV over a range of more than 5 V) and the sensitivity of the HB-CDM output to variations in input bias control, 300 000 random input combinations of XI, XQ, and PhX bias voltages have been sent from the workstation to the HB-CDM via evaluation board for data collection; a relatively limited subset of the 100 billion possible realizations. The power measurements have been

ML Evaluation Metrics	Normalized	dB
Mean Squared Error	9.39×10^{-4}	0.17
Mean Absolute Error	0.02	0.32
Maximum Absolute Error	0.16	1.57
R2 Score	98.22%	

Tab. 1: ML modulator power model evaluation metrics, expressed in normalized units and in dBm.

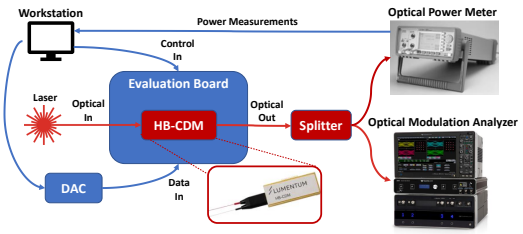


Fig. 2: Experimental setup used for data collection for machine learning training and for bias point search and control.

obtained from the power meter and saved in the terminal. The collected data are then used to train a deep-learning-based model[19], utilizing a deep neural network (DNN) architecture consisting of four dense layers and one dropout layer with 3 input features (XI, XQ and PhX) and one output label (output power). The first three dense layers contained 1024 neurons each and used the ReLU activation function[20], while the final dense layer contained a single neuron with a linear activation function. The Adam optimizer[21] is used to optimize the model's performance on the training set (65% of the dataset), and the Early Stopping technique is used to prevent over-fitting[22]. The model is compiled using the Adam optimizer and monitored using the mean squared error (MSE) loss function as a metric for evaluating the model's performance during training[23] and monitoring a validation set (15% of the original dataset). Then, the model has been evaluated using the remaining test set (20% of the original dataset), and the resulting ML evaluation metrics are presented in Tab. 1, which includes important performance evaluation metrics such as MSE, Mean Absolute Error (MAE)[24], Maximum Absolute Error[25], expressed in both normalized units and dB, and the R2 Score parameter (R2S)[26], expressed in percentage. The high R2S percentage ($> 90\%$) and low error statistics values validate the accuracy of the ML predicted model, which has been utilized for data augmentation to cover all the cases not measured during the data collection and for pre-

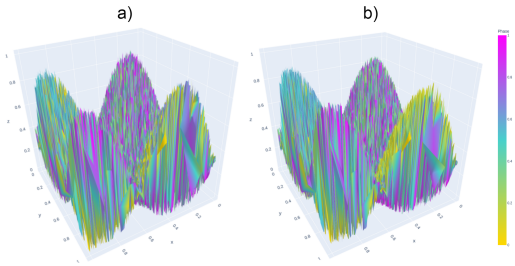


Fig. 3: Power profile of machine-learning predicted results (a) versus power measurements (b). X-axis represents XI input values, Y-axis XQ input values, Z-axis the output power values. On the right, the legend presents the Phase X input values. All values are normalized between 0 and 1.

dicting the performance of the bias point search and control on the modelled polarization branch of the modulator. A graphical comparison between the measured and predicted power values is shown in Fig. 3; in particular, Fig. 3-a) shows the normalized power profile of ML predicted results and Fig. 3-b) the actual power measurements, also normalized, considering as input the normalized XI, XQ and PhX input values. The comparison of the two figures confirms the accurate match of the predicted values with the measured data, also demonstrating the highly nonlinear behavior of the HB-CDM response as the input bias voltages vary.

Results

A bisection optimization method has been utilized to determine the optimal *Ph*, *I*, and *Q* voltages to reach the Min/Min/Quad bias condition of the modulator. Due to the strong non-linearity of the modulator power model, which results in significant output power variations for slight changes in the bias control voltage, and its ability to locate global minima and maxima of continuous functions[27], the bisection optimization performs better than other methods such as gradient descent and simulated annealing. This algorithm finds and sets the optimal *Ph* bias point first, finding through bisection the -3 dB power relative to its maximum, with *I* and *Q* voltages set at 0 V. Then, *I* minimum power is identified and subsequently *Q* minimum power is reached and set, both through bisection approach, in order to achieve the Min/Min/Quad bias condition. Initially, the algorithm is tested on the digital model of the X polarization power profile of the modulator and, subsequently, it is used on the HB-CDM module, employing the same experimental setup depicted in Fig. 2, using power meter and internal photodiode X+Y feedback for power measurements. An average of 129 total ML simulations is required to reach the predicted optimal bias point on the same sample, for the X polarization, whereas an average of 134 total measurements have been taken on the same polarization of the real component, resulting in a MAE of 0.39 dB between the predicted and measured optimal power values. The algorithm's efficacy as a structure-agnostic solution has been validated applying the same approach to both X and Y polarization branches of the HB-CDM at three different wavelengths $\lambda 1$, $\lambda 2$, $\lambda 3$ equal to 1550 nm, 1565 nm and 1535 nm, respectively. The search algorithm requires a similar number of measurements for all the polarization branches, as for the $\lambda 1$ simulation case. To verify the presented results, 16-QAM modulating signals at a symbol rate of 32 GBd have been sent through the DAC to the modulator, generating constellation diagrams visible on the OMA as illustrated in Fig. 4. The Figure shows that, except for a necessary minimum bias adjustment to balance the power between the two X and Y polarization branches[18], the optimum bias points found through the bisection algorithm are correct and meet the Min/Min/Quad condition allowing for a precise usage of QAM and QPSK modulation.

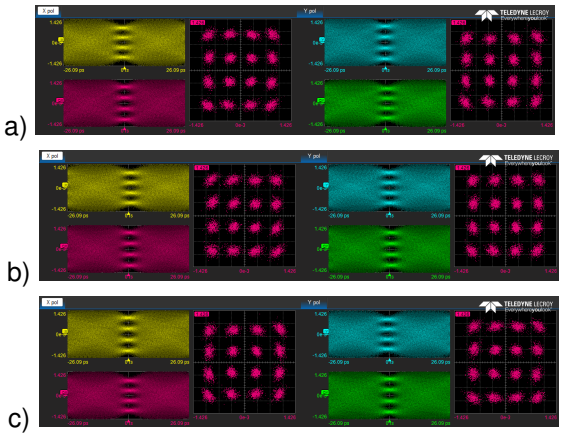


Fig. 4: OMA acquisitions of 16-QAM constellation diagrams for a) $\lambda 1$, b) $\lambda 2$ and c) $\lambda 3$ wavelengths under test.

Conclusions

A novel machine learning aided approach for bias point control of an InP DP IQ-MZM is presented in this work. Generating a digital model of the MZM and using an AI-based system to learn the modulator behavior, it is shown that the proposed bisection approach is effective in automatically finding the optimal bias point. The technique has been validated in laboratory for a structure-agnostic approach by observing different modulator settings. The results demonstrate the effectiveness of the proposed strategy in achieving optimal performance using a limited number of measurements with simple laboratory equipment thanks to the power-based bias technique, finding the optimum bias solution using a fast and simple method, adjusting only one polarization at a time. The digital model-based proposed approach shows significant potentiality for real-case applications in photonics components. It reduces the complexity of bias point control and it can be further enhanced evolving into a full digital twin-based approach through the application of reinforcement or transfer learning techniques.

Acknowledgements

This research has been made possible thanks to valuable contribution by LINKS Foundation and Lumentum.

References

[1] M. Smit, K. Williams, and J. van der Tol, "Past, present, and future of inp-based photonic integration", *APL Photonics*, vol. 4, no. 5, p. 050 901, 2019.

[2] R. A. Griffin, S. K. Jones, N. Whitbread, S. C. Heck, and L. N. Langley, "Inp mach–zehnder modulator platform for 10/40/100/200-gb/s operation", *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 19, no. 6, pp. 158–166, 2013.

[3] K. Sekine, C. Hasegawa, N. Kikuchi, and S. Sasaki, "A novel bias control technique for mz modulator with monitoring power of backward light for advanced modulation formats", *OFC/NFOEC 2007 - 2007 Conference on Optical Fiber Communication and the National Fiber Optic Engineers Conference*, pp. 1–3, 2007.

[4] H. Choi, Y. Takushima, H. Choi, J. Chang, and Y. C. Chung, "Modulation-format-free bias control technique for mz modulator based on differential phasor monitor", in *National Fiber Optic Engineers Conference*, Optica Publishing Group, 2011, JWA033.

[5] P. S. Cho, J. B. Khurgin, and I. Shpantzer, "Closed-loop bias control of optical quadrature modulator", *IEEE photonics technology letters*, vol. 18, no. 21, pp. 2209–2211, 2006.

[6] T. Yoshida, T. Sugihara, K. Uto, *et al.*, "A study on automatic bias control for arbitrary optical signal generation by dual-parallel mach-zehnder modulator", in *36th European Conference and Exhibition on Optical Communication*, IEEE, 2010, pp. 1–3.

[7] H. Kawakami, T. Kobayashi, E. Yoshida, and Y. Miyamoto, "Auto bias control technique for optical 16-qam transmitter with asymmetric bias dithering", *Optics express*, vol. 19, no. 26, B308–B312, 2011.

[8] H. Kawakami, E. Yoshida, and Y. Miyamoto, "Auto bias control technique based on asymmetric bias dithering for optical qpsk modulation", *Journal of lightwave technology*, vol. 30, no. 7, pp. 962–968, 2012.

[9] Y. Li, Y. Zhang, and Y. Huang, "Any bias point control technique for mach–zehnder modulator", *IEEE photonics technology letters*, vol. 25, no. 24, pp. 2412–2415, 2013.

[10] G. Genty, L. Salmela, J. M. Dudley, *et al.*, "Machine learning and applications in ultrafast photonics", *Nature Photonics*, vol. 15, no. 2, pp. 91–101, 2021.

[11] I. Khan, L. Tunesi, M. U. Masood, *et al.*, "Optimized management of ultra-wideband photonics switching systems assisted by machine learning", *Opt. Express*, vol. 30, no. 3, pp. 3989–4004, Jan. 2022.

[12] Y. Bae, S. Jang, S. Yoo, M. Yi, J. Ryoo, and J. Shin, "Automatic bias control technique of dual-parallel mach–zehnder modulator based on simulated annealing algorithm for quadrupled signal generation", in *Photonics*, MDPI, vol. 8, 2021, p. 80.

[13] H. Pang, Q. Zhu, S. An, J. Li, and Y. Su, "Flexible bias control for a mach-zehnder modulator based on a two-layer neural network algorithm", in *Asia Communications and Photonics Conference*, Optical Society of America, 2019, M4A–14.

[14] J. Li, D. Wang, M. Zhang, and S. Cui, "Digital twin-enabled self-evolved optical transceiver using deep reinforcement learning", *Optics Letters*, vol. 45, no. 16, pp. 4654–4657, 2020.

[15] Lumentum, *High bandwidth coherent driver modulator*, `https://www.lumentum.com/en/products/high-bandwidth-coherent-driver-modulator`, Accessed: 2023.

[16] M. Sotoodeh, Y. Beaulieu, J. Harley, and D. L. McGhan, "Modulator bias and optical power control of optical complex e-field modulators", *Journal of Lightwave Technology*, vol. 29, no. 15, pp. 2235–2248, 2011.

[17] G. Mak, Y. Beaulieu, and M. Sotoodeh, *Optimum modulator bias systems and methods in coherent optical transmitters*, US Patent 9,059,805, Jun. 2015.

[18] M. Sotoodeh, G. Mak, and Y. Beaulieu, *Quadrature power balance control in optical transmitters*, US Patent 9,124,364, Sep. 2015.

[19] T. Wang, C.-K. Wen, H. Wang, F. Gao, T. Jiang, and S. Jin, "Deep learning for wireless physical layer: Opportunities and challenges", *China Communications*, vol. 14, no. 11, pp. 92–111, 2017.

[20] A. F. Agarap, "Deep learning using rectified linear units (relu)", *arXiv preprint arXiv:1803.08375*, 2018.

[21] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization", *arXiv preprint arXiv:1412.6980*, 2014.

[22] R. Caruana, S. Lawrence, and C. Giles, "Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping", *Advances in neural information processing systems*, vol. 13, 2000.

[23] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? a new look at signal fidelity measures", *IEEE signal processing magazine*, vol. 26, no. 1, pp. 98–117, 2009.

[24] J. Qi, J. Du, S. M. Siniscalchi, X. Ma, and C.-H. Lee, "On mean absolute error for deep neural network based vector-to-vector regression", *IEEE Signal Processing Letters*, vol. 27, pp. 1485–1489, 2020.

[25] J. Giguere, E. Manoukian, and J. Roy, "Maximum absolute error for bartlett's chi-square approximation", *Journal of Statistical Computation and Simulation*, vol. 15, no. 2-3, pp. 109–117, 1982.

[26] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation", *PeerJ Computer Science*, vol. 7, e623, 2021.

[27] S. Tanweer, F. A. Khasawneh, and E. Munch, "Robust zero-crossings detection in noisy signals using topological signal processing", *arXiv preprint arXiv:2301.07703*, 2023.