DTU Library

Machine learning meets photonics

Da Ros, Francesco; Yankov, Metodi P.; Zibar, Darko

Published in:
Proceedings of SPIE

Link to article, DOI:
10.1117/12.2676656

Publication date:
2023

Document Version
Publisher's PDF, also known as Version of record

Machine learning meets photonics

Francesco Da Ros, Metodi Yankov, Darko Zibar

Francesco Da Ros, Metodi P Yankov, Darko Zibar, "Machine learning meets photonics," Proc. SPIE 12655, Emerging Topics in Artificial Intelligence (ETAI) 2023, 126550C (28 September 2023); doi: 10.1117/12.2676656

Event: SPIE Nanoscience + Engineering, 2023, San Diego, California, United States

Machine learning meets photonics

Francesco Da Ros, Metodi P. Yankov, and Darko Zibar

DTU Electro, Technical University of Denmark, Kgs. Lyngby, Denmark, fdro@dtu.dk

ABSTRACT

Machine learning (ML) is becoming a ubiquitous and powerful tool helping to address challenges in countless fields. Applications of ML addressing optics challenges have been extensively studied in recent years opening up new research directions. In particular, here, we review some of our current efforts and provide examples of successful applications of ML to the characterization of photonic devices, design, and modeling of optical subsystems, and complete end-to-end optical system optimization. ML and statistical tools can yield additional insight from measurement data, e.g. by targeted filtering of noise sources. They have also been shown to assist complex or inaccurate physics-based models through black and grey-box modeling of photonics components or subsystems. Such ML-aided models have enabled easier optimization and design (including inverse design) of optical systems.

Keywords: Machine learning, photonics, neural networks, modeling

1. INTRODUCTION

The boost in available computing power, as well as decades in training data, have recently enabled a renewed interest in machine learning (ML) as a powerful set of tools to address societal and scientific challenges.[1] Photonics is one such scientific field where a number of challenges can directly benefit from the use of ML methods. In particular, ML is particularly effective in learning complex direct (input-to-output) or inverse (output-to-input) mappings. Direct models can generally rely on physics but ML approaches might allow for increased accuracy when fitting physical models to experimental data, or for increased computational speed compared to solving physical models with conventional numerical methods. Inverse models, instead, are generally non-trivial to build by relying solely on the system's physical description which is very rarely easy to invert. Furthermore, beyond learning deterministic relations, ML allows finding patterns in the data from a statistical perspective, i.e. in terms of better understanding the non-deterministic behavior of devices or systems affected by random processes such as noise.

Neural networks (NNs) and Bayesian approaches[2] have been very successfully applied to optical communication systems and networks,[3–6] to the design and optimization of photonics structures,[7] and of ultrafast laser systems,[8] as well as towards opening a more symbiotic relationship between photonics and ML where photonics provides a hardware platform for ML-oriented computations.[9]

Beyond black-box approaches, where NNs are trained to blindly learn mappings directly from a training dataset, physics-assisted ML-based models can enhance the modeling accuracy by embedding physical knowledge into the NN architecture. This both provides increased generalization capabilities as well as ease of training (e.g. in terms of training dataset size) as the NN is not required to learn the full physical relations only from training data. These models are commonly referred to as grey-box models and can span from simply including some knowledge of the functional dependence from the physics in the network[10] to learning full partial differential equations through NN.[11]

In this work, we will review a few specific applications of ML methods to (1) laser and frequency comb characterization by using Bayesian filtering to enhance measurement accuracy, (2) black- and grey-box modeling of optical amplifiers and matrix multipliers with NNs, and (3) end-to-end learning of signaling over an optical communication channel by using auto-encoders.

Emerging Topics in Artificial Intelligence (ETAI) 2023, edited by Giovanni Volpe,
Joana B. Pereira, Daniel Brunner, Aydogan Ozcan, Proc. of SPIE Vol. 12655
126550C · © 2023 SPIE · 0277-786X · doi: 10.1117/12.2676656

2. LASER PHASE NOISE CHARACTERIZATION

Understanding the noise properties of lasers and frequency combs is critical as they directly impact the performance of applications spanning from metrology and sensing to optical communications. Such an understanding could potentially lead to improved laser designs and/or novel compensation mechanisms. Conventionally the phase noise characterization of lasers and frequency combs relies on self-heterodyne interferometric approaches that are limited in frequency range and require a thorough calibration to minimize measurement noise.[12,13] Alternatively, heterodyne schemes can be applied to increase the frequency range of the measurements. However, the increased measurement bandwidth makes the characterization even more susceptible to measurement noise. The use of Bayesian filtering, such as Kalman filtering and subspace tracking[14] can provide statistical tools to extract information on the different noise sources, e.g. statistically independent sources, from measurements including measurement noise. By relying on the statistical independence between measurement noise and noise sources contributing to the overall laser/frequency comb noise, the impact of measurement noise can be limited.[15,16] Furthermore, independently tracking the phase evolution of the different lines of a frequency comb allows us to better understand how the noise evolves over frequency lines,[17] as well as uncover correlations between phase and amplitude noise processes.[18]

3. DATA-DRIVEN MODELING OF OPTICAL DEVICES

Physics-based models of optical devices are generally readily available. However, such models might suffer either from being computationally evolved to solve making iterative optimization a challenging task, or from hard-to-fit parameters which are affected by fabrication uncertainties. That is the case for optical amplifiers and for integrated optical matrix multipliers, respectively. Sub-optimal amplifiers directly impact the performance of optical communication systems, while optical multipliers with inaccurate training lead to a loss in computing precision achievable with optical hardware.[10]

3.1 Optical amplifiers

Physics-based modeling of optical amplifiers is generally available for the most mature amplification technologies such as erbium-doped fiber amplifiers (EDFAs) and Raman amplifiers, however, model parameters (such as doping concentration, fiber parameters, etc.) are not always easy to estimate accurately from direct measurements. Additionally, solving the governing equation might require solving a system of partial differential equations through numerical methods, for example in the case of Raman amplification.[19] To increase accuracy and inference speed, black-box neural-network models of Raman and EDFA have been proposed throughout the years[20–31] progressively providing higher generalizability, e.g. to multiple physical units of similar make (EDFAs[23]), to multiple input channel loads[24] and chosen fiber type.[25] As the model is required to generalize to a larger dimensional space of parameters, model complexity and availability of training data become challenging. Therefore data-augmentation and transfer learning approaches have been proposed.[25] Additionally, new directions are being explored to focus on grey-box models where physical knowledge is embedded in the network[26,27] or even nearly-white-box models where ML is used only to represent the parts of the physical model that are challenging to compute and/or optimize. As an example, in order to optimize a Raman amplifier, the Raman gain spectrum is normally not described in a form that is differentiable with respect to the pump frequency, but it can be made differentiable through fitting.[28,29]

3.2 Matrix multipliers

Matrix vector multiplication can be implemented in the optical domain with lower energy consumption, higher speed, and lower latency compared to using electronics.[9] Micro-ring resonators and Mach-Zehnder interferometers are key building components to implement the optical weights and can be scalably integrated into a photonic integrated circuit. However, as devices are brought closer together, thermal (if phase shifting relies on thermo-optic effects) and electrical (mainly due to electrical interconnections to control the individual phase-shifting elements) crosstalks become non-negligible.[32] Such effects are rather challenging to accurately describe as physical models spanning between optical, electrical and thermal domains need to be combined.[33] Additionally, the results of such complex modeling are device dependent and heavily impacted by e.g. fabrication errors. Alternatively, simple physical models can be applied but require comprehensive calibration techniques not to sacrifice accuracy.[34–37] Finally, another approach consists of applying a data-driven method where NN trained with experimental data are able to capture the deterministic impact of crosstalk.[10] In such cases, physics-rooted considerations such as device symmetry arguments,[38] and data-augmentation techniques[39] can still be applied to limit the training measurements required and increase the generalization capabilities of the models.

4. END-TO-END LEARNING OF PHASE-NOISE ROBUST COMMUNICATION

In the context of optical communication systems, parallel to the optimization of optical components (fibers, amplifiers, etc.), a significant effort is focused on optimizing the signaling techniques to maximize the system throughput or robustness to parameter uncertainties. In such a context, the use of end-to-end learning where transmitter and receiver are jointly optimized in an encoder-decoder pair can be applied.[40] The transmitter (or part thereof) can be replaced with an NN-based encoder in order to optimize the signaling choice to be robust to the channel impairments. Similarly, the receiver can be based on a NN decoder which aims at accurately reconstructing the transmitted data from the received waveform.[41,42] The joint optimization can focus on several aspects of the signaling, e.g. constellation shaping and decision regions,[43–50] pulse shaping and receiver filtering,[41,42] etc. Similarly, the optimization objective can be varied from absolute system throughput[43–47] to robustness to uncertainty in channel[41,48] or transciever[50,51] parameters. The challenge of applying end-to-end learning mainly consists of requiring a differentiable transmitter-to-receiver model in order to optimize the encoder-decoder pair through gradient-based methods. Such a requirement brings special challenges for experimental systems where no gradient can be calculated. Nevertheless, methods such as gradient-free optimizers[49] or reinforcement learning[52] show great promise to address the challenge.

5. CONCLUSIONS

We have briefly reviewed a number of different photonic problems that can be effectively addressed using ML-based tools, spanning from enhancing measurement accuracy by addressing measurement noise, to black and grey-box modeling of optical components (i.e. optical amplifiers and optical matrix multipliers), and end-to-end optimization of a communication system.

ACKNOWLEDGMENTS

This work was supported by the Villum Foundation through Villum Young Investigator OPTIC-AI project (grant no. 29334), the ERC-CoG FRECOM project (grant no. 771878), the DNRF CoE SPOC, DNRF123, and the Horizon Europe research and innovation project PROMETHEUS, grant n. 101070195.

REFERENCES

[1] Hennessy, J., "John hennessy and david patterson 2017 acm am turing award lecture," (2018).

[2] Bishop, C. M. and Nasrabadi, N. M., [*Pattern recognition and machine learning*], vol. 4, Springer (2006).

[3] Musumeci, F., Rottondi, C., Nag, A., Macaluso, I., Zibar, D., Ruffini, M., and Tornatore, M., "An overview on application of machine learning techniques in optical networks," *IEEE Communications Surveys & Tutorials* **21**(2), 1383–1408 (2018).

[4] Rafique, D. and Velasco, L., "Machine learning for network automation: overview, architecture, and applications [invited tutorial]," *Journal of Optical Communications and Networking* **10**(10), D126–D143 (2018).

[5] Khan, F. N., Fan, Q., Lu, C., and Lau, A. P. T., "An optical communication's perspective on machine learning and its applications," *Journal of Lightwave Technology* **37**(2), 493–516 (2019).

[6] Nevin, J. W., Nallaperuma, S., Shevchenko, N. A., Li, X., Faruk, M. S., and Savory, S. J., "Machine learning for optical fiber communication systems: An introduction and overview," *APL Photonics* **6**(12) (2021).

[7] Ma, W., Liu, Z., Kudyshev, Z. A., Boltasseva, A., Cai, W., and Liu, Y., "Deep learning for the design of photonic structures," *Nature Photonics* **15**(2), 77–90 (2021).

[8] Genty, G., Salmela, L., Dudley, J. M., Brunner, D., Kokhanovskiy, A., Kobtsev, S., and Turitsyn, S. K., "Machine learning and applications in ultrafast photonics," *Nature Photonics* **15**(2), 91–101 (2021).

[9] Shastri, B. J., Tait, A. N., Ferreira de Lima, T., Pernice, W. H., Bhaskaran, H., Wright, C. D., and Prucnal, P. R., "Photonics for artificial intelligence and neuromorphic computing," *Nature Photonics* **15**(2), 102–114 (2021).

[10] Cem, A., Yan, S., Ding, Y., Zibar, D., and Da Ros, F., "Data-driven modeling of mach-zehnder interferometer-based optical matrix multipliers," *Journal of Lightwave Technology* **41**, 5425–5436 (2023).

[11] Raissi, M., Perdikaris, P., and Karniadakis, G. E., "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics* **378**, 686–707 (2019).

[12] Zhao, Z., Bai, Z., Jin, D., Qi, Y., Ding, J., Yan, B., Wang, Y., Lu, Z., and Mildren, R. P., "Narrow laser-linewidth measurement using short delay self-heterodyne interferometry," *Optics Express* **30**(17), 30600–30610 (2022).

[13] Kim, J. and Song, Y., "Ultralow-noise mode-locked fiber lasers and frequency combs: principles, status, and applications," *Advances in Optics and Photonics* **8**(3), 465–540 (2016).

[14] Särkkä, S. and Svensson, L., [*Bayesian filtering and smoothing*], vol. 17, Cambridge university press (2023).

[15] Zibar, D., Chin, H.-M., Tong, Y., Jain, N., Guo, J., Chang, L., Gehring, T., Bowers, J. E., and Andersen, U. L., "Highly-sensitive phase and frequency noise measurement technique using bayesian filtering," *IEEE Photonics Technology Letters* **31**(23), 1866–1869 (2019).

[16] Zibar, D., Pedersen, J. E., Varming, P., Brajato, G., and Da Ros, F., "Approaching optimum phase measurement in the presence of amplifier noise," *Optica* **8**(10), 1262–1267 (2021).

[17] Razumov, A., Heebøll, H. R., Dummont, M., Terra, O., Dong, B., Riebesehl, J., Varming, P., Pedersen, J. E., Da Ros, F., Bowers, J. E., et al., "Subspace tracking for independent phase noise source separation in frequency combs," *arXiv preprint arXiv:2305.08681* (2023).

[18] Brajato, G., Lundberg, L., Karlsson, M., Zibar, D., et al., "Bayesian filtering framework for noise characterization of frequency combs," *Optics express* **28**(9), 13949–13964 (2020).

[19] Islam, M. N., "Raman amplifiers for telecommunications," *IEEE Journal of selected topics in Quantum Electronics* **8**(3), 548–559 (2002).

[20] Zhou, J., Chen, J., Li, X., Wu, G., Wang, Y., and Jiang, W., "Robust, compact, and flexible neural model for a fiber Raman amplifier," *Journal of lightwave technology* **24**(6), 2362 (2006).

[21] Zibar, D., Brusin, A. M. R., de Moura, U. C., Da Ros, F., Curri, V., and Carena, A., "Inverse system design using machine learning: The Raman amplifier case," *Journal of Lightwave Technology* **38**(4), 736–753 (2020).

[22] Ionescu, M., "Machine learning for ultrawide bandwidth amplifier configuration," in [*2019 21st International Conference on Transparent Optical Networks (ICTON)*], 1–4 (2019).

[23] Da Ros, F., De Moura, U. C., and Yankov, M. P., "Machine learning-based edfa gain model generalizable to multiple physical devices," in [*2020 European Conference on Optical Communications (ECOC)*], 1–4 (2020).

[24] Brusin, A. M. R., de Moura, U. C., Curri, V., Zibar, D., and Carena, A., "Introducing load aware neural networks for accurate predictions of Raman amplifiers," *Journal of Lightwave Technology* **38**(23), 6481–6491 (2020).

[25] de Moura, U. C., Zibar, D., Brusin, A. M. R., Carena, A., and Da Ros, F., "Fiber-agnostic machine learning-based Raman amplifier models," *Journal of Lightwave Technology* **41**(1), 83–95 (2022).

[26] Mahajan, A., Christodoulopoulos, K., Martínez, R., Spadaro, S., and Muñoz, R., "Modeling edfa gain ripple and filter penalties with machine learning for accurate qot estimation," *Journal of Lightwave Technology* **38**(9), 2616–2629 (2020).

[27] Jones, R. T., Bottrill, K. R., Taengnoi, N., Petropoulos, P., and Yankov, M. P., "Spectral power profile optimization of a field-deployed wavelength-division multiplexing network enabled by remote edfa modeling," *Journal of Optical Communications and Networking* **15**(8), C192–C202 (2023).

[28] Marcon, G., Galtarossa, A., Palmieri, L., and Santagiustina, M., "Model-aware deep learning method for Raman amplification in few-mode fibers," *Journal of Lightwave Technology* **39**(5), 1371–1380 (2020).

[29] Yankov, M. P., Da Ros, F., de Moura, U. C., Carena, A., and Zibar, D., "Flexible Raman amplifier optimization based on machine learning-aided physical stimulated Raman scattering model," *Journal of Lightwave Technology* **41**(2), 508–514 (2023).

[30] Minakhmetov, A., Prieur, B., Le Monnier, M., Rouvillain, D., and Lavigne, B., "Digital twin of unrepeatered line based on Raman and remote optically pumped amplifier machine learning models," in [*2023 Optical Fiber Communications Conference and Exhibition (OFC)*], 1–3 (2023).

[31] Borraccini, G., Straullu, S., Piciaccia, S., Tanzi, A., Galimberti, G., and Curri, V., "Cognitive Raman amplifier control using an evolutionary optimization strategy," *IEEE Photonics Technology Letters* **34**(4), 223–226 (2022).

[32] Pérez, D., Gasulla, I., Crudgington, L., Thomson, D. J., Khokhar, A. Z., Li, K., Cao, W., Mashanovich, G. Z., and Capmany, J., "Multipurpose silicon photonics signal processor core," *Nature Communications* **8** (12 2017).

[33] Orlandin, M., Cem, A., Curri, V., Carena, A., Da Ros, F., and Bardella, P., "Thermal crosstalk effects in a silicon photonics neuromorphic network," in [*NUSOD 2023*], (2023).

[34] Bandyopadhyay, S., Hamerly, R., and Englund, D., "Hardware error correction for programmable photonics," *Optica* **8**(10), 1247–1255 (2021).

[35] Hamerly, R., Bandyopadhyay, S., and Englund, D., "Accurate self-configuration of rectangular multiport interferometers," *Physical Review Applied* **18**(2), 024019 (2022).

[36] Mojaver, K. R. and Liboiron-Ladouceur, O., "On-chip optical phase monitoring in multi-transverse-mode integrated silicon-based optical processors," *IEEE Journal of Selected Topics in Quantum Electronics* **28**(6), 1–7 (2022).

[37] Miller, D. A., "Self-configuring universal linear optical component," *Photonics Research* **1**(1), 1–15 (2013).

[38] Cem, A., Sanchez-Jacome, D., Pérez-López, D., and Da Ros, F., "Thermal crosstalk modeling and compensation for programmable photonic processors," in [*IPC*], accepted (2023).

[39] Cem, A., Jovanovic, O., Yan, S., Ding, Y., Zibar, D., and Da Ros, F., "Data-efficient modeling of optical matrix multipliers using transfer learning," in [*CLEO: Science and Innovations*], SM3J–4 (2023).

[40] O'shea, T. and Hoydis, J., "An introduction to deep learning for the physical layer," *IEEE Transactions on Cognitive Communications and Networking* **3**(4), 563–575 (2017).

[41] Karanov, B., Chagnon, M., Thouin, F., Eriksson, T. A., Bülow, H., Lavery, D., Bayvel, P., and Schmalen, L., "End-to-end deep learning of optical fiber communications," *Journal of Lightwave Technology* **36**(20), 4843–4855 (2018).

[42] Gaiarin, S., Da Ros, F., Jones, R. T., and Zibar, D., "End-to-end optimization of coherent optical communications over the split-step fourier method guided by the nonlinear fourier transform theory," *Journal of Lightwave Technology* **39**(2), 418–428 (2020).

[43] Jones, R. T., Eriksson, T. A., Yankov, M. P., and Zibar, D., "Deep learning of geometric constellation shaping including fiber nonlinearities," in [*2018 European Conference on Optical Communication (ECOC)*], 1–3 (2018).

[44] Li, S., Häger, C., Garcia, N., and Wymeersch, H., "Achievable information rates for nonlinear fiber communication via end-to-end autoencoder learning," in [*2018 European Conference on Optical Communication (ECOC)*], 1–3 (2018).

[45] Schaedler, M., Calabrò, S., Pittalà, F., Böcherer, G., Kuschnerov, M., Bluemm, C., and Pachnicke, S., "Neural network assisted geometric shaping for 800gbit/s and 1tbit/s optical transmission," in [*2020 Optical Fiber Communications Conference and Exhibition (OFC)*], 1–3 (2020).

[46] Neskorniuk, V., Carnio, A., Bajaj, V., Marsella, D., Turitsyn, S. K., Prilepsky, J. E., and Aref, V., "End-to-end deep learning of long-haul coherent optical fiber communications via regular perturbation model," in [*2021 European Conference on Optical Communication (ECOC)*], 1–4 (2021).

[47] Oliveira, B., Neves, M., Guiomar, F., Medeiros, M., and Monteiro, P., "Autoencoder-optimized geometric constellation shaping for unamplified coherent optical links," in [*CLEO: Science and Innovations*], SW4E–7 (2022).

[48] Rode, A. and Schmalen, L., "Optimization of geometric constellation shaping for wiener phase noise channels with varying channel parameters," in [*European Conference and Exhibition on Optical Communication*], Mo3D–5 (2022).

[49] Jovanovic, O., Yankov, M. P., Da Ros, F., and Zibar, D., "Gradient-free training of autoencoders for non-differentiable communication channels," *Journal of Lightwave Technology* **39**(20), 6381–6391 (2021).

[50] Jovanovic, O., Yankov, M. P., Da Ros, F., and Zibar, D., "End-to-end learning of a constellation shape robust to channel condition uncertainties," *Journal of Lightwave Technology* **40**(10), 3316–3324 (2022).

[51] Song, J., Häger, C., Schröder, J., Amat, A. G. I., and Wymeersch, H., "Model-based end-to-end learning for wdm systems with transceiver hardware impairments," *IEEE Journal of Selected Topics in Quantum Electronics* **28**(4: Mach. Learn. in Photon. Commun. and Meas. Syst.), 1–14 (2022).

[52] Song, J., He, Z., Häger, C., Karlsson, M., i Amat, A. G., Wymeersch, H., and Schröder, J., "Over-the-fiber digital predistortion using reinforcement learning," in [*2021 European Conference on Optical Communication (ECOC)*], 1–4 (2021).