Machine learning applied to inverse systems design

Uiara C. de Moura, Francesco Da Ros and Darko Zibar
DTU Fotonik
Technical University of Denmark
Kongens Lyngby, Denmark
fdro@fotonik.dtu.dk

Ann Margareth Rosa Brusin and Andrea Carena
DET
Politecnico di Torino
Turin, Italy

***Abstract*—In this work, we will give an overview of some of the most recent and successful applications of machine learning-based inverse system designs in photonic systems. Then, we will focus on our recent research on the Raman amplifier inverse design. We will show how the machine learning framework is optimized to generate on-demand arbitrary Raman gain profiles in a controlled and fast way and how it can become a key feature for future optical communication systems.**

***Index Terms*—inverse design, machine learning, photonic systems, optical amplifiers**

I. Introduction

The recent renewed interest in machine learning (ML) has been motivated by the massive amount of information generated in our modern society and the recent evolution of high efficient computers. Based on the idea that the underlying features in a given data set can be learned, ML is used to model complex functions to make decisions/predictions for unseen data. These ML models have been successfully applied to address some of society's biggest and most complex problems in different fields such as business [1], healthcare [2], and astronomy [3]. Recently, there has been an increasing amount of research applying the concepts of ML in the field of optical communications. Some examples are quality of transmission estimation [4], modulation format recognition [5], optical performance monitoring [6], and most recently the inverse system design of photonic structures [7]–[11] and devices [12]–[17].

The inverse system design consists in finding the optimum set of design parameters that provides a desirable system response. The traditional procedure to design optical devices starts with an initial set of parameters (normally based on the designer's previous knowledge) and performs some parameter sweep around this initial condition to find the desired response. This human-controlled design approach has two key drawbacks: it is time-consuming/work-intensive and tends to ignore solutions that could have better performance, but are far from the initial guess. A way to speed up this process is by using simplified models for the function $f(\cdot)$ that relates the system parameter to its response. This can be done by using a ML model to learn $f(\cdot)$. Such a model can go inside an iterative optimization routine such as gradient descent [22], [23] or an evolutionary strategy [24] to search the optimum set of parameters.

Another interesting way of applying ML to solve the inverse system design is to learn the inverse function $f(\cdot)^{-1}$ relating the system response to the design parameters. Such an "inverse" system model provides the direct retrieval of the optimum set of parameters for a target response. This idea has been applied for the inverse design of photonic integrated structures [7]–[11], optical fibers [25], and optical amplifiers [12]–[17]. In these works, an artificial neural network (NN) model receives the target device performance and provides an optimized set of parameters more straightforwardly when compared to the iterative optimization routines.

In this work, we review the recent progress in ML-based approaches for the inverse design and optimization of photonic structures and devices. Then, we will discuss our recently proposed ML framework for the Raman amplifier inverse design and how it is applied to provide on-demand gain profiles in a controlled way.

II. ML-enabled inverse design approaches

Applying ML to the inverse system design is not a recent idea. Since the 90's it has been used to design bipolar junction transistors [18], [19], microwave filters [20], and microstrip antennas [21]. These works explore the versatility of NNs in learning $f(\cdot)$ (forward models) and $f(\cdot)^{-1}$ (inverse models). They show how forward models are very useful to reduce the design time by replacing computational-expensive and time-consuming numerical simulation tools [19], [21]. They also show how inverse models can reduce even more the design time by instantaneously providing the physical structure given a target response [18], [20].

One of the first works applying the NN-based inverse models to solve the design problem in photonic systems shows that, when applied to the design of complex nanophotonic structures that require large training data, the NN has problems to learn the inverse mapping due to one-to-many mappings [7]. This is a fundamental challenge in problems where the same system response can be created by different designs. To solve this issue, the authors propose a cascaded network structure with an inverse NN model followed by a pre-trained forward NN. In this auto-encoder-like structure, the input and output are the target and predicted system performances, respectively.

This project has received funding from the European Research Council through the ERC-CoG FRECOM project (grant agreement no. 771878), the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 754462, the Villum Foundations (VYI OPTIC-AI grant no. 29344), and Ministero dell'Istruzione, dell'Università e della Ricerca (PRIN 2017, project FIRST).

978-3-903176-44-7

The design parameter is retrieved between the NN models. During the inverse NN training, the one-to-many mapping problem is eliminated by minimizing the errors between target and predicted system responses. Once trained, the authors in [7] used the cascaded NN to design dielectric multilayers films (SiO2 and Si3O4) to achieve a certain transmission response.

Following works have also explored the NN-based inverse model idea. In [8], the authors use a deep NN to predict the geometry of plasmonic nanostructures based on far-field measurements. This deep NN is then used in a sensing application to find the nanostructure configuration that best interacts with a given molecule. In [9], the authors propose an adaptive normalized NN for the inverse design of graphene-based metamaterials. Finally, the authors in [25] applied the inverse NN model to effectively and instantaneously optimize the structural parameters of ring-assisted few-mode fibers with weak coupling optimization. These are just a few examples of how ML is revolutionizing the design process in photonics by avoiding designer guesses and time-consuming computation for Maxwell's equations. A comprehensive state-of-the-art of ML applied to design photonic structures and devices can be found in [10], [11].

III. Raman amplifier inverse design

The ability to shape the gain profile in a controlled way is an exclusive feature of Raman amplifiers [27]. This is done by properly adjusting the Raman pump powers to achieve the target gain profile. However, due to the complex interactions between pumps and signals, this adjustment is not a trivial task and has been referred to as the Raman amplifier design.

The concept of inverse NN models was first applied to the Raman amplifier design by [12]. This work evolved to reach a comprehensive ML framework [26] and it is illustrated in Fig. 1(a). The framework consists of two neural networks NN_{fwd} and NN_{inv}, for the forward and inverse system models, respectively. For the Raman amplifier case, NN_{fwd} learns the direct (forward) mapping for the Raman amplifier relating the Raman pump parameters $\mathbf{P}$ to the Raman gain profile response $\mathbf{G}$, i.e., $\mathbf{G} = f(\mathbf{P})$. NN_{inv} learns the inverse mapping $\mathbf{P} = f^{-1}(\mathbf{G})$. Here the function $f(\cdot)$ is a set of non-linear ordinary differential equations describing the Raman amplifier process. NN_{fwd} and NN_{inv} are trained using supervised learning. Therefore, they need a data set with uniformly distributed examples of $\mathbf{P}$ and their corresponding $\mathbf{G}$. A thorough description of the data set generation and the NNs training can be found in [26].

The Raman amplifier design illustrated in Fig. 1(a) consists in applying NN_{inv} to provide the pump configuration $\tilde{\mathbf{P}}$ given a target gain profile $\mathbf{G_T}$ at its input. As an optional step, $\tilde{\mathbf{P}}$ can be fine optimized. This fine design process applies NN_{fwd} in a gradient descent (GD) routine to minimize the mean squared error (MSE) between predicted $\tilde{\mathbf{G}}$ and target $\mathbf{G_T}$ gain profiles. This is possible because NN_{fwd} is differentiable, which is not the case for $f(\cdot)$. Moreover, the optimized pump parameters $\mathbf{P_{opt}}$ are obtained after a few iterations since the process started from a close to optimum solution provided by $\tilde{\mathbf{P}}$.

The robustness of the ML framework for different input signal spectral profiles is covered by [29]. Its generalization properties to different fiber types and lengths are experimentally evaluated by [30], where a general model is proposed. The proposed ML framework was also updated to consider noise figure prediction during the design [28]. All these works consider 4-pumps C-band distributed Raman amplifiers.

In this work, we will show the experimental validation for the design of an ultra-wideband discrete Raman amplifier covering the S, C, and L bands [27]. The signal has 148 frequency channels spaced by 100 GHz and covering 19.4 THz. Their spectral allocation is shown on the top of Fig. 1(c). The gaps in the spectrum are due to pump laser allocation overlap and the lack of signal lasers on that region. In this example, the Raman amplifier has 8 pumps equally spaced in frequency. The gain is measured as the difference between the output optical spectrum with the pump lasers turned on and off. Details about the experimental setup, the neural networks training, and their individual performance can be found in [27].

In the design stage, we consider three cases of target gain profiles: flat, tilted, and arbitrary (illustrated in Fig. 1(a)). Flat and tilted gain profiles range from 14 to 20 dB with a 1 dB step (total of 7 cases). Tilted gain profiles consider a -0.2-dB/THz slope coefficient. The arbitrary gains are feasible gains, i.e. they are experimentally measured for 1025 uniformly distributed pump power configurations. The experimental validation illustrated in Fig. 1(b) applies both $\tilde{\mathbf{P}}$ and $\mathbf{P_{opt}}$ to the experimental Raman amplifier setup. The measured gain $\mathbf{G_M}$ is then compared to the target $\mathbf{G_T}$ by calculating the absolute error along the frequency channels $Error = |G_T - G_M|$.

Fig. 1(c) and (d) show the absolute errors per frequency for the flat and tilted gain profiles, respectively. To better visualize the results, each box plot considers all channels in a 600-GHz bandwidth (6 channel slots). These designs were obtained by applying $\mathbf{P_{opt}}$. The following analysis excludes the lowest frequency channel. Tilted and flat gain designs have very similar performances, with a high design error for high frequencies. This is because higher frequency channels have contributions from a higher number of pump lasers due to the pumps' non-symmetric Raman gain spectrum. This makes the design more complex in this region, especially for flat and tilted gains [27].

Fig. 1(e) shows the absolute errors versus frequency over 1025 arbitrary gain profiles. These designs do not need the fine design and are, therefore, obtained by applying just NN_{inv} outcome $\tilde{\mathbf{P}}$. Since the error bars are too small, we plot the mean error ($\overline{Error}$) on a separate curve in the right y-axis (also excluding the lowest frequency channel). In this case, the absolute errors are almost constant along the frequencies. High errors are again for the high-frequency channels, and may be related to the complexity in learning $f(\cdot)^{-1}$ with more pump contributions in high frequency.

The highest errors observed for the lowest frequency channel is due to instabilities observed after the amplification process, which are consequence of the channel position isolated

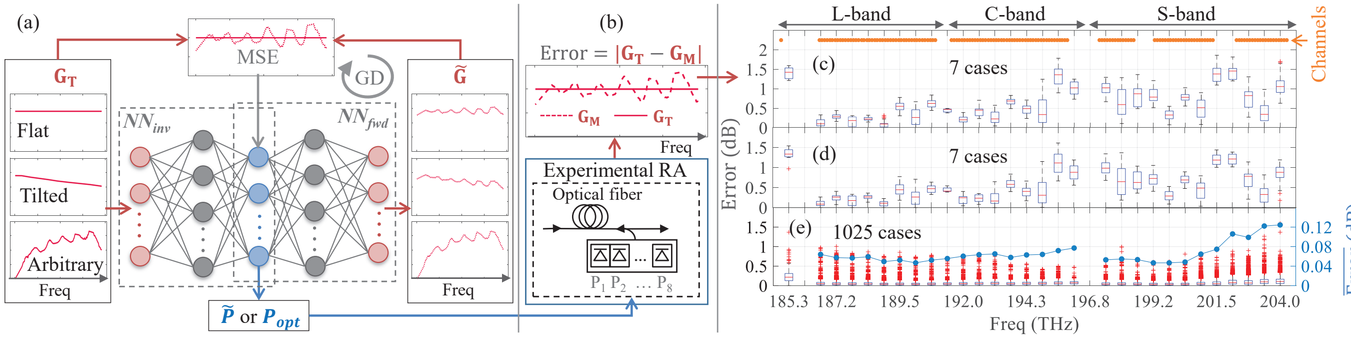

Fig. 1. (a) Full machine learning framework (i.e. inverse NN_{inv} and forward NN_{fwd} neural network models) for the design and gradient descent-based fine design; (b) experimental design validation procedure applying the pump configurations from (fine) design to the experimental Raman amplifier, comparing the corresponding measured gain ($\mathbf{G_M}$) to the target gain ($\mathbf{G_T}$), and the error ($|\mathbf{G_T} - \mathbf{G_M}|$) along the frequency for (c) flat, (d) tilted and (e) arbitrary gains.

on the edge of the spectra.

The ability of NN in learning the complex relations between pump and signal as an inverse system model was also evaluated by other works considering different scenarios, such as for hybrid amplifiers [13], [14] and few-mode Raman amplifiers [15], [16]. Finally, in [17], they apply a convolutional neural network to find the pump powers and wavelengths of a distributed Raman amplifier required for a target signal power evolution in both frequency and distance along the fiber.

IV. Conclusions

This work gave a brief overview of some recent works applying machine learning to solve the inverse system design problem in photonics. We focused on works applying neural networks to learn the inverse system function, mapping the system response to design parameters. Such data-driven models are highly accurate and can solve the design problem almost instantaneously. This is a brand new field of research that is totally transforming the way we engineer and with the potential to have a high impact beyond optics and photonics.

References

[1] I. Bose, R. K. Mahapatra, "Business data mining — a machine learning perspective," Information & Management **39**, 211 (2001).
[2] PH.C. Chen, Y. Liu, L. Peng, "How to develop machine learning models for healthcare," Nat. Mater. **18** 410 (2019).
[3] Ž. Ivezić et al., "Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data", Princeton University Press, 2014.
[4] Y. Pointurier, "Machine learning techniques for quality of transmission estimation in optical networks", J. Opt. Commun. Netw. **13**, B60 (2021).
[5] D. Wang et al., "Modulation Format Recognition and OSNR Estimation Using CNN-Based Deep Learning," in IEEE Phot. Tech. Lett. **29**, 1667 (2017).
[6] J. Thrane et al., "Machine Learning Techniques for Optical Performance Monitoring From Directly Detected PDM-QAM Signals," J. Light. Technol. **35**, 868 (2017).
[7] D. Liu et al., "Training deep neural networks for the inverse design of nanophotonic structures," ACS Photon. **5**, 1365 (2018).
[8] I. Malkiel et al. "Plasmonic nanostructure design and characterization via deep learning", Light Sci. Appl. 7, 60 (2018).
[9] Y. Chen et al., "Smart inverse design of graphene-based photonic metamaterials by an adaptive artificial neural network," Nanoscale **11**, 9749 (2019).
[10] Y. Xu et al., "Interfacing photonics with artificial intelligence: an innovative design strategy for photonic structures and devices based on artificial neural networks," Photon. Res. **9**, B135 (2021).
[11] W. Ma et al. "Deep learning for the design of photonic structures", Nat. Photonics **15**, 77 (2021).
[12] D. Zibar et al., "Machine learning-based Raman amplifier design," in Proc. Opt. Fiber Commun. Conf., 2019, p. M1J.1.
[13] M. Ionescu, "Machine Learning for Ultrawide Bandwidth Amplifier Configuration," in Proc. 21th Int. Conf. Transp. Opt. Netw., 2019.
[14] X. Ye et al. "Experimental Prediction and Design of Ultra-Wideband Raman Amplifiers Using Neural Networks," in Proc. Opt. Fiber Commun. Conf., 2020, p. W1K.3.
[15] Y. Chen et al. "Intelligent gain flattening of FMF Raman amplification by machine learning based inverse design," in Proc. Opt. Fiber Commun. Conf., 2020, p. T4B.1.
[16] G. Marcon et al. "Model-aware deep learning method for Raman amplification in few-mode fibers," J. Light. Technol., **39**, 1371 (2020).
[17] M. Soltani et al., "Inverse design of a Raman amplifier in frequency and distance domains using convolutional neural networks," Opt. Lett. **46**, 2650 (2021).
[18] R. Ferguson, D. J. Roulston, "Neural Networks for the Design and Reverse Engineering of BJTs," in Proc. Eur. Solid State Device Res. Conf., 1996, pp. 969-972.
[19] M. M. Vai et al., "Reverse modeling of microwave circuits with bidirectional neural network models," in IEEE Trans. Microw. Theory Techn. , **46**, 1492, (1998).
[20] H. Kabir et al., "Neural Network Inverse Modeling and Applications to Microwave Filter Design," in IEEE Trans. Microw. Theory Techn. , **56** 867 (2008).
[21] V. V. Thakare, P. Singhal, "Microstrip antenna design using artificial neural networks", Int. J. RF Microw. Comput.-Aided Eng. **20**, 76 (2010).
[22] J. Zhou et al., "Robust, Compact, and Flexible Neural Model for a Fiber Raman Amplifier," J. Light. Technol., **24**, 2362 (2006).
[23] J. Peurifoy et al., "Nanophotonic particle simulation and inverse design using artificial neural networks," Sci. Adv. **4**(6), r4206 (2018).
[24] J. Chen and H. Jiang, "Optimal Design of Gain-Flattened Raman Fiber Amplifiers Using a Hybrid Approach Combining Randomized Neural Networks and Differential Evolution Algorithm," IEEE Photonics J. **10**, 7101915 (2018).
[25] Z. He, et al. "Machine learning aided inverse design for few-mode fiber weak-coupling optimization," Opt. Express **28**, 21668 (2020).
[26] D. Zibar et al. "Inverse System Design Using Machine Learning: The Raman Amplifier Case," J. Light. Technol., **38**, 736 (2020).
[27] U. C. de Moura et al. "Multi–Band Programmable Gain Raman Amplifier," J. Light. Technol., **39**, 429 (2021).
[28] U. C. de Moura et al., "Simultaneous gain profile design and noise figure prediction for Raman amplifiers using machine learning," Opt. Lett. **46**, 1157 (2021).
[29] U. C. de Moura et al., "Experimental Characterization of Raman Amplifier Optimization Through Inverse System Design," J. Light. Technol., **39**, 1162 (2021).
[30] U. C. de Moura et al., "Generalization Properties of Machine Learning-based Raman Models," in Proc. Opt. Fiber Commun. Conf., 2021, p. Th1A.28.