Inverse Design of Nanophotonic Circuitry Components based on Reinforcement Learning

(Student paper)

Marco Butz[1,2,4], Alexander Leifhelm[1,2], Marlon Becker[3], Benjamin Risse[3] and Carsten Schuck[1,2,*]

[1]Institute of Physics, University of Münster, Wilhelm-Klemm-Str. 10, 48149 Münster, Germany
[2]Center for Soft Nanoscience, Busso-Peus-Str. 10, 48149 Münster, Germany
[3]Institute for Computer Science, University of Münster, Einsteinstr. 62, 48149 Münster, Germany
[4]marco.butz@uni-muenster.de, *carsten-schuck@uni-muenster.de

We introduce a reinforcement learning-based method for solving the pixel-discrete inverse-design problem. The approach is capable of producing highly efficient, yet compact photonic circuit components of any desired linear optical functionality. We apply the algorithm to design a nanophotonic waveguide-mode converter (TE_{00} —> TE_{20}) with a conversion efficiency of more than 85% and compact footprint of only a few vacuum wavelengths.
Keywords: *Inverse-Design, Reinforcement Learning, Nanophotonic Circuits*

INTRODUCTION

The complexity of photonic integrated circuits is increasing steadily as optical functionalities are replacing electronic integrated circuits and open new information processing capabilities, e.g. in quantum photonics. To sustain the growth of nanophotonic applications, novel device functionalities and compact footprints will be required. Currently, photonic integrated circuit components are predominantly designed following a bottom-up approach, where the designer aims to exploit intuitive physical effects in combination with brute-force methods for device optimization purposes. The layout for a nanophotonic waveguide mode converter, for instance, would be based on evanescent wave coupling and hybrid mode effective refractive index matching [1]. In consequence, the resulting devices suffer from large footprints and are limited in the number of possible functionalities as complex phenomena, such as multi-wave-interference effects cannot be considered, because they are not accessible by physical intuition. Consequently, the overwhelming part of the vast design space that is available with modern nanofabrication techniques, in particular those featuring sub-wavelength dimensions, is not accessible. This method is opposed by top-down approaches, where the device is seen as a pixel-discretized black-box, which is the subject of an optimization problem where known inputs and desired outputs are defined as boundary conditions. Here, the limitations mentioned above do not exist and – given a suitable computational routine – complex devices with sub-wavelength features may be found that realize previously inaccessible functionalities and smaller device footprints [2].

Generally, photonic black-box optimization problems are of non-convex nature. Thus, following inverse-design approaches that rely on gradient techniques to produce efficient device layouts on computationally feasible timescales is not a trivial task. Previously, numerous computational methods employing convex optimization or direct search methods have been demonstrated. All of them suffer from drawbacks such as exponential runtime scaling with increasing design space, solutions featuring continuous permittivity distributions that cannot be produced with common nanofabrication methods, or failure to consider stochastic processes due to their dependency on calculated gradient fields [3].

In this work we show how to phrase the nanophotonic black-box optimization problem as a reinforcement learning (RL) application. Our method overcomes previous limitations by leveraging a distributed and highly parallel computational architecture, showing stable learning behavior that scales well with increasing design space dimensionality. The direct application of general reward-based reinforcement learning algorithms enables us to optimize any pixel-discrete nanophotonic device accessible by electromagnetic simulations. We can manipulate the dataflow in different stages of the algorithm, allowing us to implement arbitrary geometry constraints, e.g. for observing limitations imposed by nanofabrication capabilities.

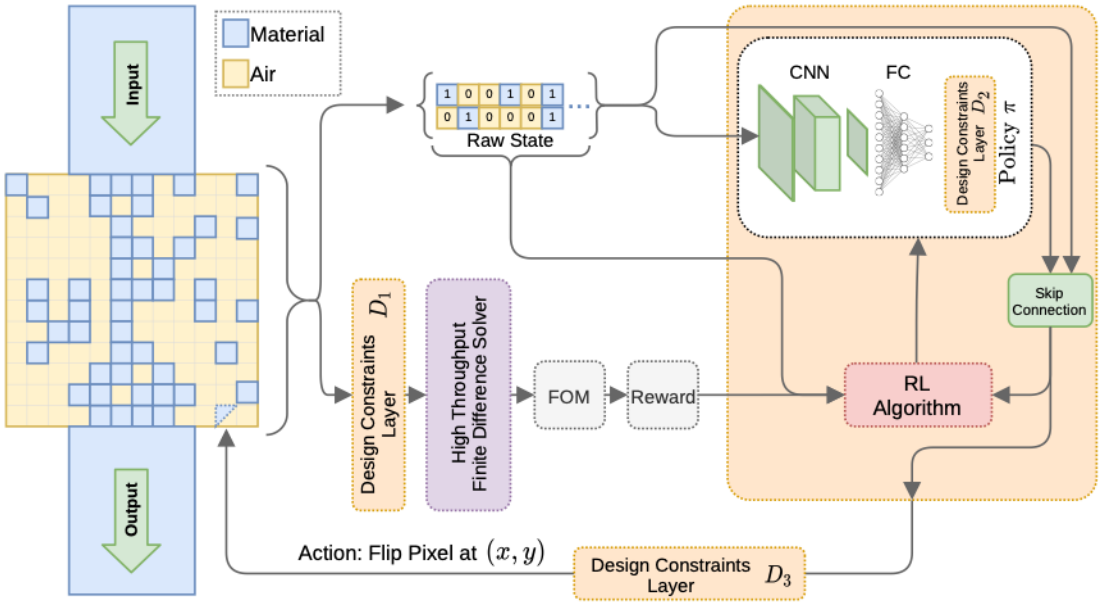

Fig. 1. Data-flow in a RL-worker. The device shown on the left is the numerical representation of the optimization problem as specified by parameters such as pixel-size, device-dimensions and the waveguide-modes under consideration. The agent (orange box) interacts with the environment while the actions and observations are processed and possibly altered in between.

METHOD AND APPLICATION

Reinforcement learning is a biology-inspired domain of artificial intelligence where an autonomously acting entity referred to as the agent explores an environment by taking a series of actions. These actions change the state of the environment, and a reward is reported back to the agent. The reward is a measure for the improvement in device performance achieved by the action, such that an optimal behavior of the agent can be obtained by maximizing the cumulative reward following a policy, understood as a mapping from states to actions. Fig. 1 shows how this concept is applied to the pixel-discrete nanophotonic inverse-design problem. The device shown on the left represents the environment, where we define the state to be the two-dimensional pixel-configuration encoding the presence of material or air as logic values ‚1' or ‚0', respectively. We refer to this representation as the „raw state", which is further processed in the policy-determining convolutional neural network (CNN) and fully connected layers (FC). To suppress the common problem of reward-hacking through repetitive behavior, we combine the network's output with the initial raw state using a specifically tailored skip connection, which bypasses the convolutional processing layers. Finally, a soft-max function is applied to the output to determine the next action from the corresponding probability distribution, i.e. the coordinates of a pixel which is to be flipped from material to air or vice versa. To calculate the associated reward, the device performance is evaluated under consideration of the surrounding waveguide geometry and involved input- and output-modes using a customized high-throughput Finite Difference Frequency Domain (FDFD) solver. The results of the simulation are used to calculate a figure of merit, which in turn is used as an input for the reward calculation, while both functions can be specified by the user in order to account for special requirements of specific devices. The resulting tuple of state, action and reward is subsequently processed by a reinforcement learning algorithm adjusting the weights in the neural network to optimize the current policy.

The workflow depicted in Fig. 1 is implemented in an isolated instance, which we refer to as a „worker". Multiple workers can be employed on remote machines, communicating with a central instance [4], thus granting scalability of our approach ranging from single desktop-PCs to whole HPC-clusters. Multiple levels of parallelism here enable us to optimize devices in large and complex environments.

In Fig. 2 we show an exemplary application of our method on a nanophotonic waveguide-mode converter, which is supposed to convert the fundamental quasi-transverse electric (TE_{00}) to the higher order quasi-TE_{20} waveguide-mode at a wavelength of 775 nm. We configure the device to have a length of 3.2 µm and a width of 1.6 µm discretized in 80 and 40 pixels, respectively, realized in 200 nm thin silicon-nitride-on-insulator. The pixels are quadratic with a side-length of 40 nm and the optimization, including all involved simulations, are conducted in full 3D. Through one of the design-constraint-layers depicted in Fig. 1, namely D_3, we enforce a symmetric layout. The final device is shown in Fig. 2 (left), while the corresponding evolution of the maximum figure of merit, as calculated by 8 parallel workers, are shown in Fig. 2 (right). The learning behavior is strong, stable and reproducible, resulting in successful mode conversion with an efficiency of 85.5% (-0.68 dB) and low crosstalk into the TE_{00}-mode of 2.8% (-15.53 dB). Our implementation is sufficiently versatile to allow for biasing the results in order obey fabrication constraints, if necessary.

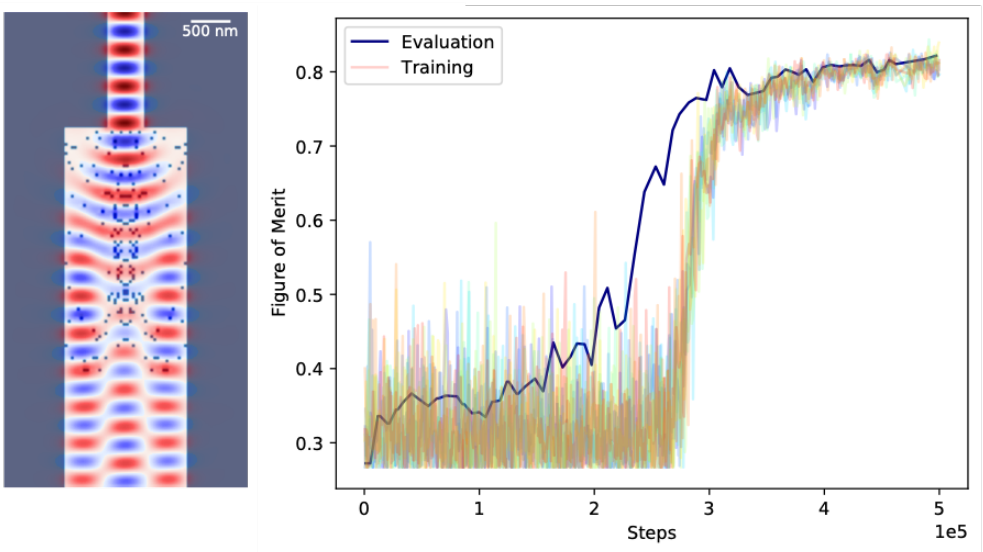

Fig. 2. Optimization of a nanophotonic waveguide-mode converter. The final device with the electric field overlay (E_y-component, perpendicular to the direction of propagation) excited by a distribution corresponding to the fundamental waveguide-mode is depicted on the left. Dark blue and white background indicate the presence of air and material, respectively. The figure of merit, which is representative of the conversion efficiency, plotted against the total number of steps taken by all workers is shown on the right.

CONCLUSION AND OUTLOOK

In this work, we apply RL, as one of the most promising approaches in artificial intelligence, to nanophotonic inverse-design. Our method is highly flexible as it allows for the optimization of any device accessible by electromagnetic simulations. Theoretical advances in the field of reinforcement learning will be directly applicable to this approach and thus we expect to immediately benefit from recent findings such as sophisticated approaches to handle delayed- or sparse-reward environments or novel and more accurate neural-network architectures to derive the current state-action-policy.

Acknowledgements: C.S. acknowledges support from the Ministry for Culture and Science of North Rhine-Westphalia (No. 421-8.03.03.02-130428).

References

1. LW. Luo, N. Ophir, C. Chen et al. *WDM-compatible mode-division multiplexing on a silicon chip*, Nature Communications 5, 3069, 2014
2. S. Molesky, Z. Lin et al. *Inverse design in nanophotonics*, Nature Photonics 12, pp. 659–670, 2018
3. Jesse Lu, Jelena Vučković, *Objective-first design of high-efficiency, small-footprint couplers between arbitrary nanophotonic waveguide modes*, Opt. Express 20, pp. 7221-7236, 2012
4. M. Volodymyr, A. P. Badia et al. *Asynchronous Methods for Deep Reinforcement Learning*, Proceedings of The 33rd International Conference on Machine Learning, PMLR 48:1928-1937, 2016