Bayesian Optimization for Nested Adversarial Variational Autoencoder in Tunable Nanophotonic Device Design

Koike-Akino, Toshiaki; Jung, Minwoo; Chakrabarty, Ankush; Wang, Ye; Kojima, Keisuke; Brand, Matthew

TR2023-036 May 09, 2023

Abstract

We propose a new device optimization framework based on Bayesian optimization for efficient latent sampling of adversarial generative neural networks to expedite a complex inverse design of tunable nanophotonic wavelength splitters. Our design, at broadband telecomm-wavelengths, is electrically switchable via liquid crystal tuning.

Conference on Lasers and Electro-Optics (CLEO) 2023

Bayesian Optimization for Nested Adversarial Variational Autoencoder in Tunable Nanophotonic Device Design

Toshiaki Koike-Akino[1,*], Minwoo Jung[1,2], Ankush Chakrabarty[1], Ye Wang[1], Keisuke Kojima[1], Matthew Brand[1]

[1] *Mitsubishi Electric Research Laboratories (MERL), 201 Broadway, Cambridge, MA 02139*
[2] *Department of Physics, Cornell University, Ithaca, NY 14853*

[*] *koike@merl.com*

Abstract: We propose a new device optimization framework based on Bayesian optimization for efficient latent sampling of adversarial generative neural networks to expedite a complex inverse design of tunable nanophotonic wavelength splitters. Our design, at broadband telecomm-wavelengths, is electrically switchable via liquid crystal tuning.

1. Introduction

In recent years, generative deep neural networks (DNN) such as generative adversarial network (GAN) and variational autoencoder (VAE) have been successfully applied to time-efficient inverse design of photonic devices [1,2]. Generative DNN models typically use random numbers as latent variables to generate multiple devices with the expectation that some of them give very good performance [2,3]. In this paper, we introduce Bayesian optimization (BO) [4,5] for efficient latent sampling of generative DNNs. In order to compensate for the mismatch from the desired spectrum, we use BO for sampling target spectrum values besides the latent variables, unlike a typical BO-based latent space optimization [7,8]. We demonstrate that the BO-assisted nACVAE can accelerate the optimization trials, realizing an extinction ratio greater than 14.5 dB.

2. LC-Tunable Wavelength splitter

Our target device structure is based on a compact on-chip wavelength de-multiplexer which is electrically tunable with a liquid crystal (LC) over nanophotonic circuits, such that the outputs are swapped when the LC is on [6]. Fig. 1 shows a cross-sectional view of the tunable photonic device. We use silicon nitride waveguide core on insulator covered with LC. Using the adjoint method [9] provided by Lumerical, we first generate several good device structures given different target responses of the dual-state wavelength splitter. We tried to maximize the extinction ratio of the splitter at two wavelength λ_1 and λ_2, depending on the LC condition of either ON (*e*-axis along out-of-plane direction) or OFF (*e*-axis perpendicular to the input waveguide). Fig. 2(a) shows an example device topology optimized for $\bar{\lambda} = (\lambda_1 + \lambda_2)/2 = 1517$ nm and $\Delta\lambda = |\lambda_1 - \lambda_2| = 47$ nm, and the resulting device response is shown in Fig. 2(b).

3. DNN Model and Design Method

The goal of the inverse design in this device is to generate a useful device topology (denoted as T), given a desired transmission spectra (denoted as S), however, the users may not know or care about the entire spectral response. The users' demand would be in the form of partial information of the spectra (denoted as S'), such as transmission

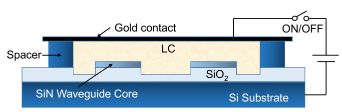

Fig. 1: Cross-sectional view of LC tunable photonic devices.

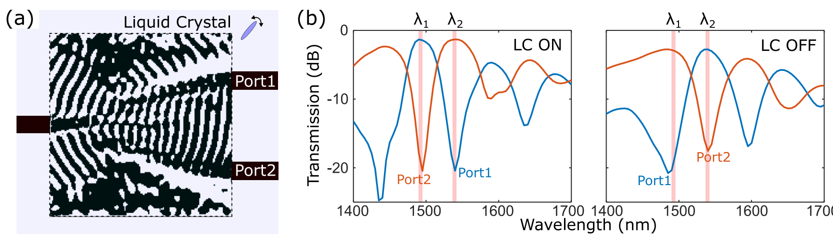

Fig. 2: (a) device topology (black: silicon nitride, pink: LC, dashed box: optimization area) and (b) transmission spectra of an example device generated by DNN, optimized for high ER at λ_1 and λ_2.

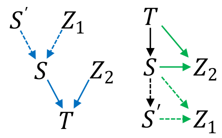

Fig. 3: Nested ACVAE architecture [6].

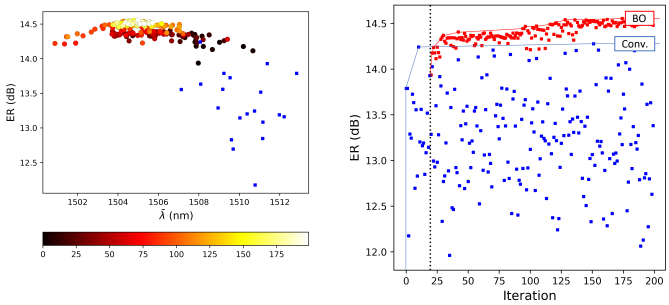

Fig. 4: ER performance while exploring latent space of nested ACVAE by GP-BO and conventional sampling.

levels at specific wavelengths $S' = [\bar{\lambda}, \Delta\lambda, ER]$. The 'Nested ACVAE' [6] model assumes a step-by-step statistical dependence from S' to S and to T.

The DNN training dataset was prepared by starting from the center ($\bar{\lambda} = 1550$ nm and a fixed separation $\Delta\lambda = 45$ nm), and use them as the initial condition for the neighboring grid points. This ensure smoothness in T across varying S' within a series of such cascades. The adjoint optimization produces many ($\sim$100) intermediate sub-optimal results en route to the final optimal design.

4. BO-assisted nested ACVAE and Results

The original nested ACVAE design method uses a random Gaussian sampling in latent variables. To accelerate the exploration, we propose to use Bayesian optimization (BO), which has shown potential for global optimization of blackbox functions in a sample-efficient manner [4]. BO requires designing two components: a probabilistic map from the latent variables to the figure of merit, and an acquisition function that guides the selection of the next optimizer candidate given the available data points. Classically, BO methods leverage Gaussian process (GP) regression for the task of providing a probabilistic map, while it scales cubically with the number of available data points and the dimension [5]. In this paper, we use GP-BO to optimize the latent variables of nACVAE, specifically z_1 and z_2, instead of standard Gaussian sampling. Furthermore, we optimize the target values of ($\bar{\lambda}$, $\Delta\lambda$, ER) in addition to the latent variables so that we can compensate for the misalignment issue. We use an expected improvement (EI) as an acquisition function for GP-BO.

We consider 4 dimensions for the latent space of z_1 and z_2, and 3 additional dimensions for target values for $\bar{\lambda}$, $\Delta\lambda$, and ER. The GP-BO explores spaces at more hopeful regions by analyzing the landscape and prediction uncertainty. Fig. 4 shows the optimization trajectory. We first generated 20 random sample devices for training GP-BO, and latent space exploration was caried out by GP-BO on the fly to validate the performance with 3-dimensional finite-difference time-domain (FDTD) simulations. It is verified that the BO can significantly outperform the conventional random sampling. To design one device using the adjoint method takes about 1.5 day using a computing cluster. Even though it takes 4 hours to train the DNN model, to generate and validate the 410 devices takes about 9 hours. This shows that DNN has the potential to cover the whole target parameter space in a short time, without using the adjoint method to design each device individually.

5. Summary

We demonstrated that BO-assisted DNN sampling technique can accelerate an inverse design of tunable nanophotonic wavelength splitters. Specifically, the nested ACVAE model found in AutoBayes showed superior performance when latent variables are sampled by GP-BO, achieving an extinction ratio of about 14.5 dB over wide wavelengths in a small varidation iterations.

References

1. Z. Liu et al, Bulletin of the American Physical Society, vol. 64, 2019.
2. Y. Tang et al., Laser & Photonics Reviews, vol. 14, no. 12, p. 2 000 287, 2020.
3. A. K. Zhaxylyk et al., Applied Physics Reviews, vol. 7, p.021407, 2020.
4. J. Snoek et al., Advances in neural information processing systems, vol. 25, 2012.
5. J. Snoek et al., in International conference on machine learning, PMLR, 2015, pp. 2171–2180.
6. M. Jung et al, CLEO, paper SW5E.6, 2022.
7. Y. Zhang et al, Scientific reports, vol. 10, no. 1, pp. 1–13, 2020.
8. T. Tripp et al, Advances in Neural Information Processing Systems, vol. 33, pp. 11 259–11 272, 2020.
9. Y. Piggott et al., ature Photonics, vol. 9, no. 6, pp. 374–377, 2015.