Optimization of a Photonic Crystal Nanocavity Using Covariance Matrix Adaptation Evolution Strategy

Kohei Takahashi and Toshihiko Baba, *Fellow, IEEE*

***Abstract*—H0-type photonic crystal nanocavities hold high quality factors Q and quite small cavity mode volumes. This study finds their ultrahigh Q structures, which allow stable operation as a nanolaser even with fabrication-induced disordering. Previously, we generated a neural network model for predicting Qs, searched for a high-Q structure and its slotted version, and found those showing $Q = 1{,}140{,}000$ and 91,600, respectively. These values were an order of magnitude higher than those obtained by manual optimizations. However, further improvement above these values was saturated because of the insufficient accuracy of the neural network model at the high Q regime. Instead of applying the model, we repeated directly calculating Qs, implementing a covariance matrix adaptation evolution strategy algorithm to search structures in this study. Consequently, Q values were increased up to 14,500,000 and 741,000 while consuming shorter calculation time. We also confirmed that these structures significantly improve robustness against structural disordering.**

***Index Terms*—Photonic crystal, nanolaaser, evolutionary calculation, machine learning.**

I. Introduction

HA0-TYPE photonic crystal nanocavities consist of slight shifts of some holes in a triangular lattice hole array in a semiconductor slab, as shown in Fig. 1. They have been investigated for nanolaser operation, biochemical sensing, nonlinear enhancement, and cavity quantum electrodynamics [1]–[7]. They simultaneously achieve a high quality factor Q and a quite small mode volume V. The nanoslot version shown in Fig. 1(a) and (b) particularly localizes the modal electric field in the nanoslot and improve the sensitivity and stability as a biochemical sensor in aqueous solution [5], [8], [9]. However, the theoretical Q of 10^5 without the nanoslot is significantly degraded to a value lower than 5000 with the nanoslot, and this value can be further reduced by the structural disordering in fabricated devices. GaInAsP semiconductors are often used for photonic crystal nanolasers due to their small nonradiative recombination, but their fine etching process is complicated and usually induces some deformation in the holes of the photonic crystal.

So far, computational optimizations have been applied to the design of various photonic devices [10]–[13]. Recently, iterative machine learning was applied to searching for better photonic crystal cavities [14]. In the machine learning process, a training dataset listing randomly generated structural parameters and corresponding Q values are collected using first-principle calculation such as finite-difference time-domain (FDTD) simulation, and based on the dataset, a neural network (NN) model is produced as an approximate function that predicts Qs much faster than the direct calculation. Then, the structural parameters are optimized by some algorithm, where Q values are predicted by the NN model and used for the optimization. Previously we applied this process to the small shifts of holes in the H0 nanocavities [15] and obtained $Q = 1140000$ and 91600 without and with the nanoslot, respectively. The laser operation does not necessarily require an ultrahigh Q over 10^5 because other parasitic losses of active semiconductors become dominant at such high-Q regime. However, theoretical higher Q is still desired to maintain a sufficient value even with the disordering, as mentioned above.

Because of this reason, we have continued searching for higher Q structures with the NN model after the previous study, but this attempt has failed consecutively. Although the NN model predicted exceptionally high values, their correct values directly calculated by FDTD were often much lower than predicted. This

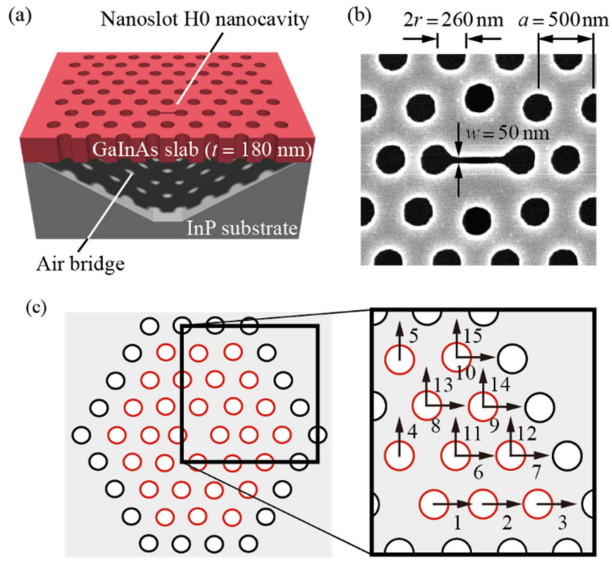

Fig. 1. GaInAsP H0-type photonic crystal nanolaser. (a) Schematic of a device with a nanoslot and (b) scanning electron micrograph of the fabricated device. (c) Cavity model (without nanoslot as an example). Hole positions around the H0 nanocavity were used for optimization. Arrows at red holes represent feature hole shifts used for optimization.

Manuscript received February 2, 2022; revised March 29, 2022; accepted April 14, 2022. Date of publication April 19, 2022; date of current version April 27, 2022. *(Corresponding author: Toshihiko Baba.)*

The authors are with the Electrical and Computer Engineering Department, Yokohama National University, Yokohama, Kanagawa 240-8501, Japan (e-mail: tozy2577@gmail.com; baba-toshihiko-zm@ynu.ac.jp).

Digital Object Identifier 10.1109/JPHOT.2022.3168543

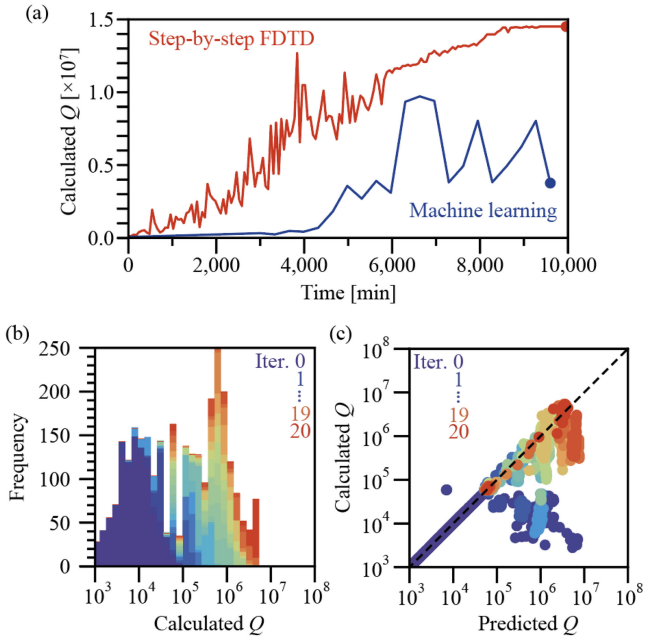

Fig. 2. Optimization of the Q value of an H0 nanocavity without a nanoslot. (a) Comparison between two optimization methods in terms of improving Q value with calculation time. Here time for collecting dataset is included for machine learning. (b) Distribution of 3000 data of calculated Qs after 20 times optimization process of machine learning. Blue shows the initial dataset, and other colors show those added in repeating the process. (c) Q correspondence between NN model prediction and FDTD calculation after the optimization process.

means that the prediction accuracy of the NN model was not sufficient at the high-Q regime due to the lack of training data. In general, it is much easier to collect low-Q data than high-Q data when relying on the random generation of parameters. This situation only guarantees accuracy within the range of existing data, and the predictions outside this range are based on extrapolation, which is hard for any NN models to maintain accuracy. To improve this situation, further adding training data and tuning hyperparameters of the NN model on each iteration might be effective. However, these processes require a manual workload and decreases the benefit of automatic learning and optimization. Besides this problem in the NN model, our previous study also had a vulnerability in the optimization algorithm; it simply searched for a better structure at each iteration based on the gradient of the prior results without implementing any stochastic or mutative randomness, which can easily be trapped to local minima.

In this study, we considered to stop using the NN model to exclude the inaccuracy problem and focused more on the optimization algorithm. We employed covariance matrix adaptation evolution strategy (CMA-ES) [16]–[18], a cutting-edge optimization method that uses stochastic random search in the nonlinear continuous parameter domain. In each iteration, CMA-ES generates the next parameters much more efficiently from previous ones based on search pass vectors represented by a covariance matrix C and standard deviation σ. The major hyperparameters are only the center and size of the distribution in the first generation and the amount of population generated

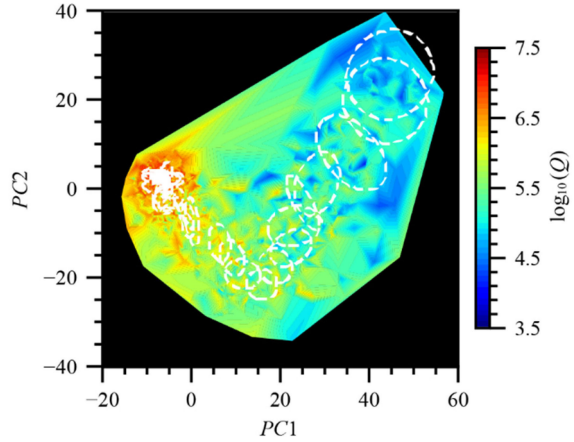

Fig. 3. Optimization process of CMA-ES without NN model, which PCA visualizes. All calculated Qs are plotted during optimization based on two principal components $PC1$ and $PC2$. Dotted circles represent the covariance matrix C for every three iterations. The scale of C is depicted five times larger than the actual size to emphasize the shape.

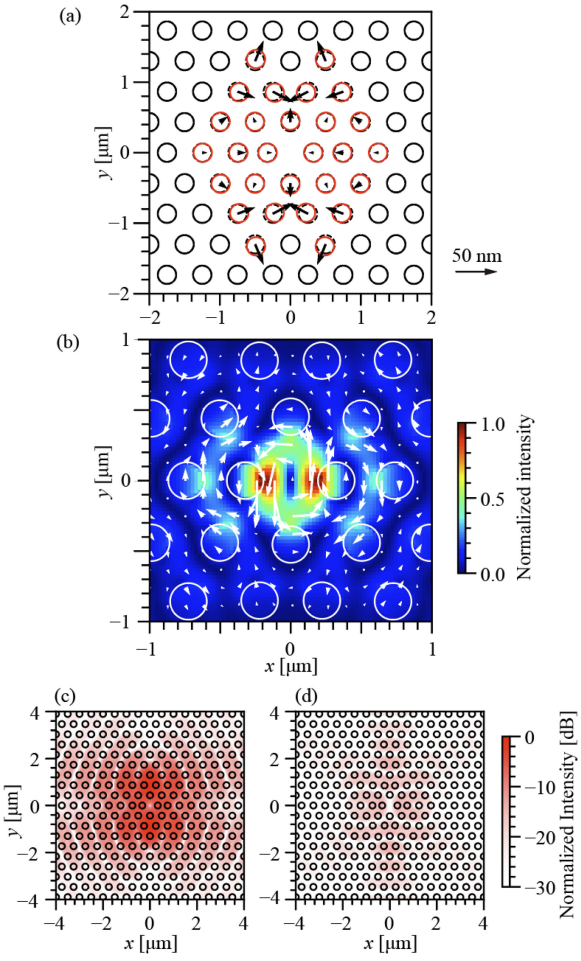

Fig. 4. Optimization results using direct Q calculation using FDTD combined with CMA-ES optimization for an H0 nanocavity without a nanoslot. (a) Hole shifts. Arrows represent shift direction and amount (×10 magnification). (b) Distribution of in-plane electric field intensity $|\boldsymbol{E}|^2$ at the slab center, which FDTD calculated. Arrows represent the direction and amplitude of the $\boldsymbol{E}$ field. (c) and (d) visualize leaky components before and after optimization, respectively. The maximum intensity of (c) is at 0 dB.

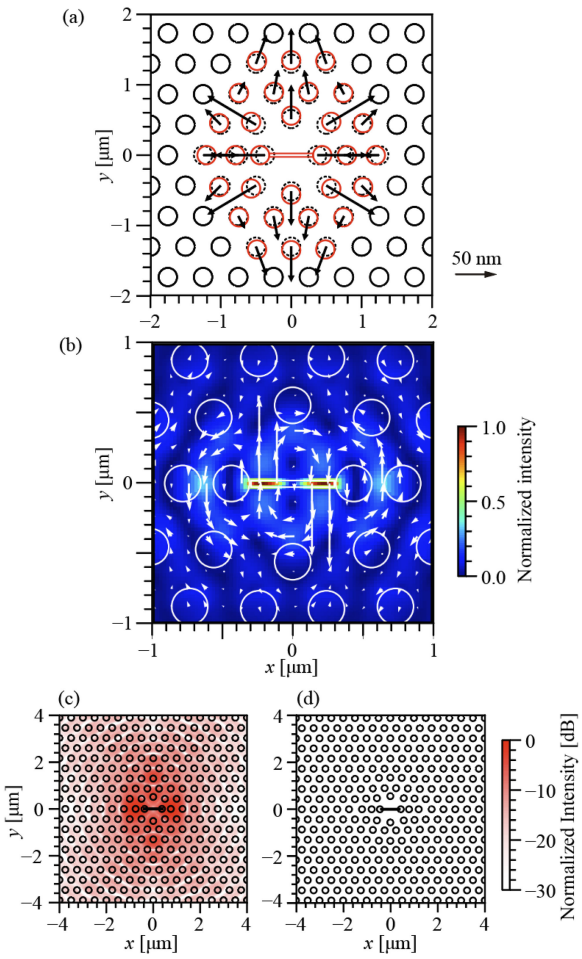

Fig. 5. Optimization results for an H0 nanoslot cavity with a nanoslot. (a)−(d) are the same as those in Fig. 4.

or maintained in subsequent generations. After setting these parameters, C and σ are automatically updated, and no manual control is needed afterward.

In this paper, we compare the results using CMA-ES with and without the NN model. For the former, we collected the training data using FDTD, generated and updated the NN model, and optimized the structure using CMA-ES with the NN model. We evaluated the accuracy as well as the calculation time of the whole process. For the latter, we optimized the structure using CMA-ES, repeating the FDTD calculations. We will show that the latter can obtain higher Qs with much higher accuracy and shorter calculation time.

II. Optimization With NN Model

H0 photonic crystal nanocavities with and without the nanoslot were assumed to be formed into a GaInAsP slab with a refractive index n_{slab} of 3.4, a thickness t_{slab} of 180 nm, and an emitting wavelength of approximately 1550 nm, for which a triangular lattice of perforated holes was designed with a lattice constant $a = 500$ nm and a hole diameter $2r = 260$ nm. Without the nanoslot, shifts of four center holes to outside, s_1 and s_4 in Fig. 1(c) were initially set at 80 and 0 nm, respectively, and the environmental refractive index n_{env} for air was 1.0. With the nanoslot, s_1 and s_4 were initially set at 120 and 80 nm, respectively, the nanoslot width w was 50 nm, and n_{env} was 1.321 for water. In the Q calculations, FDTD simulation was performed with a full three-dimensional model using 20 standard personal computers (Intel Core i7-8700K, Microsoft Windows 10 Pro) and commercial software, Ansys/Lumerical FDTD Solutions, in parallel. The Q value was determined from the temporal decay of the modal electric field, which took 60–70 min in each computer with an accuracy setting of 4 in the software.

The feature parameters of machine learning were the x and y shifts of red-colored holes in Fig. 1(c); 30 holes were shifted symmetrically against the cavity center. Eventually, 15 individual shifts were considered. We generated 1000 randomly shifted structures from the initial range within ± 30 nm with a 1 nm resolution for the training dataset. In the training process, calculated Qs were converted to $\log_{10}(Q)$ to reduce the difference in magnitude. These data were divided into training data (90%) and test data (10%), and then, the former was inputted into a fully connected sequential NN model generated by the Keras module on Python [19]. Hyperparameters for NN were manually configured as 4 hidden layers and 200 nodes in each layer, 256 batch sizes, and 100 epochs. After the training, the NN model predicted Q with a determination coefficient of $R^2 = 0.98$ for the existing test data. Then, the structure was optimized using the NN model and CMA-ES (CMA module contained within distributed evolutionary algorithm modules in Python) [20]. In each iteration, 20 structures were generated based on C and σ (initially set at 1 nm), and the NN model predicted all of their Qs within 170 ms. Then, C and σ were updated to cover the top 10 high-Q structures. After 300 iterations, the highest Q structure was chosen as an optimized one. We repeated this optimization 100 times from the beginning to eventually obtain 100 optimized candidates. Finally, we calculated their Qs using FDTD, added them to the original dataset, and performed the training again to update the NN model for next generation calculation with better accuracy.

The blue line in Fig. 2(a) shows the improvement of Q after repeating the above process 20 times. We finally obtained a structure showing $Q = 9720000$, which is 8.8 times our previous value [15], after all calculations of 6500 min. This high value is considered to be attributed to the efficient optimization of the CMA-ES. However, the maximum Q did not improve further and began to fluctuate. A histogram of the dataset explains this in Fig. 2(b), where different colors indicate different iterations. High-Q data from 10^5–10^7 were gradually added by repeating the process, but data above 10^7 were still nonexistent. This result means that the NN model predicted $Q > 10^7$ via extrapolation, which degrades the prediction accuracy significantly. Fig. 2(c) shows the correspondence between predicted and calculated Qs. The prediction was first accurate and then inaccurate below and above 10^5, respectively, and gradually improved by repeating the process. However, an invisible wall appeared to be present at 10^7, making improvement beyond this wall difficult. A similar wall was observed in the optimization of an H0 nanocavity with

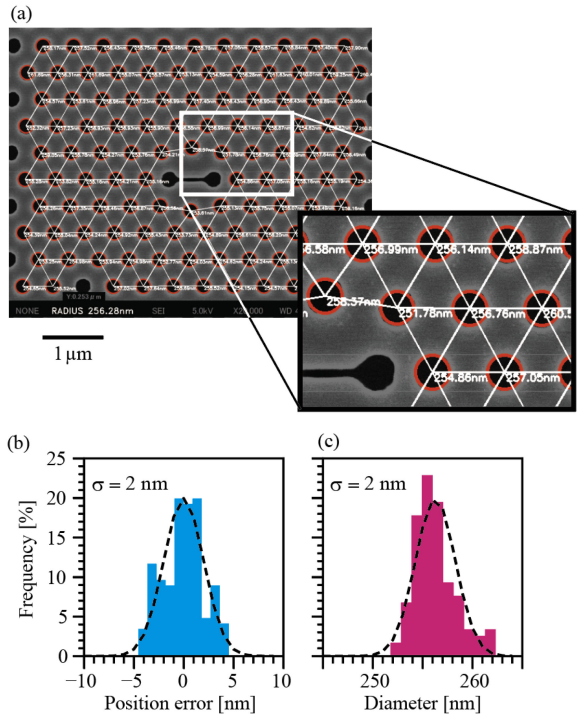

Fig. 6. Evaluation of hole diameter and position from a scanning electron micrograph. (a) Micrograph used to measure fluctuations. (b) (c) Measured fluctuation of hole position and diameter, respectively. Dotted line shows normal distributions of $\sigma = 2$ nm.

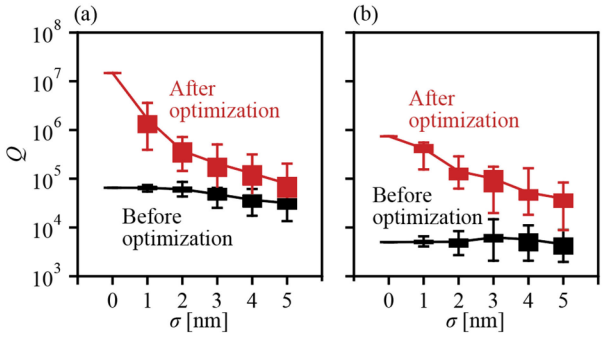

Fig. 7. Influence of structural disordering on the calculated Q factor for the structure before and after optimization. (a) H0 nanocavity without a nanoslot and (b) with a nanoslot. Bars, rectangles, and solid lines show maximum-to-minimum ranges, standard deviation ranges, and average values.

the nanoslot, for which the highest Q reached 97500 but halted advancement beyond it.

III. Optimization Without NN Model

Skipping the 1000 data collection in machine learning, we started from the Q calculations of 20 structures using FDTD and updating C and σ for the top 10 high-Q structures. The red line in Fig. 2(a) represents the improvement of Q using this method. After 160 iterations consuming 9600 min, Q reached 14500000, which is 14 times our previous value. Best and average values between 20 structures fluctuated during the optimization, but they eventually converged, which indicates that the structure reached the global best.

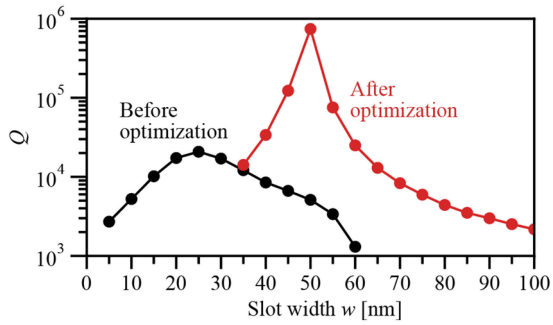

Fig. 8. Influence of fluctuated nanoslot width on calculated Q factor.

Fig. 3 also visualizes the optimization process using principal component analysis (PCA). To visualize the adaptation of C and σ for 15 hole shifts, we repeated their projection onto reduced principal axes 13 times until two axes remained using the scikit-learn module on Python [21]. The two axes were labeled as principal components 1 and 2 in the figure, and their explained variance ratio (showing how effective two axes explain the whole components) was calculated to be 64%. The values of these axes do not directly correspond to the shifts, but high- and low-Q areas can be distinguished intuitively. The center of the white circle depicting C moved almost linearly, and its shape and size of C, depicting σ, changed adaptively toward the high-Q regime.

Fig. 4 shows the hole shifts and modal characteristics of the optimized H0 nanocavity without the nanoslot. In Fig. 4(a), red-colored circles and arrows depict shifted holes and their shift direction and amount, respectively. They seem to have no noticeable tendency, but the modal electric field inside the slab was proved to be confined strongly around the cavity, as shown in Fig. 4(b). Leaky components of the mode, which do not meet the total reflection condition of the slab, were visualized by extracting the spatial Fourier components above the air light line. As shown in Fig. 4(c) and (d), they were reduced after the optimization. Even though the Q increases significantly, the mode volume V was calculated to be $0.226(\lambda/n)^3$ after the optimization, where n is the refractive index of the slab. It is comparable to or even smaller than $0.257(\lambda/n)^3$ before the optimization. This indicates that the increased Q is not attributed to a larger V but the suppression of the modal leakage.

Using the same method, we also optimized the H0 nanocavity with the nanoslot. After 200 iterations, Q drastically increased from 4600 to 741000 and was saturated. Fig. 5 shows its hole shifts and modal characteristics. The most significant shift was 71 nm, which is larger than 30 nm set as an upper limit in the optimization using the NN model. The modal electric field showed strong confinement inside the nanoslot. Compared with those before the optimization, the leakage components were significantly suppressed. Calculated V eventually decreased slightly from $0.411(\lambda/n)^3$ to $0.370(\lambda/n)^3$ after optimization. While the electric field is strongly localized in the nanoslot, its tail widespread in the semiconductor slab results in a slightly larger V than that without the nanoslot.

The methods in Sections II and III have searched $Q \approx 10^7$ structures, but the latter method achieved higher values within a

shorter calculation time. One reason for this is the higher flexibility and efficiency of the direct FDTD calculation. The search range was limited in the data collection for machine learning, and the calculation was repeated for randomly generated structures to construct the NN model. However, the calculated Qs mostly remained in the training data range and the prediction accuracy is insufficient beyond this range, even after the repetition of the optimization, adding data, and training. The direct calculation method started optimizing the structure from early iterations with guaranteed accuracy and flexibly extended or limited the search range, avoiding early saturation.

IV. Robustness

Fabricated devices usually have some disordering in their structures, which affects the cavity Q [22], [23]. In our GaInAsP cavities, the disordering mainly occurs as the deformation of holes induced during the etching process, resulting in the fluctuation of effective hole position and diameter. Fig. 6 shows a device fabricated in the same way as described in Ref. [5]. For a fabricated sample, we measured the position and diameter of 112 holes from the scanning electron micrograph. The red-colored circle was automatically fit to the contour of each hole using OpenCV Python module [24], and the effective diameter and shift were evaluated from the area and center of the hole, respectively. Their standard deviation was estimated to be 2 nm, as shown in Fig. 6(a) and (b). Therefore, robustness against this value or larger fluctuations is expected for the optimized structures.

Fig. 7 shows the degradation of the calculated $\underline{Q}$ when fluctuations of $\sigma = 0$–5 nm were randomly introduced into the hole position and diameter of the structures before and after the optimization described in Section III. After the optimization, both structures with and without the nanoslot stay above $Q = 10^5$ for $\sigma = 2$ nm, and 10^4 for $\sigma = 5$ nm, which are sufficient for laser operation. For the nanocavity with the nanoslot, the fluctuation in slot width w is also critical to degradation. Fig. 8 shows the calculated Q for these structures as a function of w. Q changes with w more drastically than the hole position and diameter particularly in the optimized structure. But still fluctuations within ± 15 nm is tolerated to ensure a Q over 10^4.

V. Conclusion

This paper described the CMA-ES optimization for H0 photonic crystal nanocavity structures toward a high Q, comparing that with an NN model constructed by machine learning and that using direct FDTD calculations. The machine learning model fast predicted Q but with low accuracy, particularly at the high-Q regime. Improving the process of the NN model was rather time-consuming, and Q could not be higher than 10^7. By thoroughly calculating all structures using FDTD, the latter method obtained Q over 10^7 faster. Finally, it reached 14500000 without the nanoslot and 741000 with the nanoslot by extending search range flexibility. We also confirmed the robustness of these structures for laser operation, which was maintained even assuming structural disordering larger than the experimental one. In this study, only 15 shifts of holes were employed as feature parameters. However, owing to the adaptation capability of CMA-ES, we will be able to further improve Q and robustness by increasing the number of holes and introducing new parameters such as hole diameter and slab thickness.

References

[1] Z. Zhang and M. Oiu, "Small-volume waveguide-section high q microcavities in 2D photonic crystal slabs," *Opt. Exp.*, vol. 12, no. 17, pp. 3988–3995, 2004.

[2] K. Nozaki, S. Kita, and T. Baba, "Room temperature continuous wave operation and controlled spontaneous emission in ultrasmall photonic crystal nanolaser," *Opt. Exp.*, vol. 15, no. 12, pp. 7506–7514, 2007.

[3] K. Nozaki *et al.*, "Sub-femtojoule all-optical switching using a photonic-crystal nanocavity," *Nature Photon.*,vol. 4, pp. 477–483, 2010.

[4] M. Nomura, K. Tanabe, S. Iwamoto, and Y. Arakawa, "High-Q design of semiconductor-based ultrasmall photonic crystal nanocavity," *Opt. Exp.*, vol. 18, no. 8, pp. 8144–8150, 2010.

[5] S. Kita *et al.*, "Photonic crystal point-shift nanolaser with and without nanoslots—Design, fabrication, lasing, and sensing characteristics," *IEEE J. Sel. Top. Quantum Electron.*, vol. 17, no. 6, pp. 1632–1647, Nov./Dec. 2011.

[6] U. P. Dharanipathy, M. Minkov, M. Tonin, V. Savona, and R. Houdré, "High-Q silicon photonic crystal cavity for enhanced optical nonlinearities," *Appl. Phys. Lett.*, vol. 105, 2014, Art. no. 101101.

[7] M. Minkov and V. Savona, "Automated optimization of photonic crystal slab cavities," *Sci. Rep.*, vol. 4, 2015, Art. no. 5124.

[8] S. Kita, S. Hachuda, K. Nozaki, and T. Baba, "Nanoslot laser," *Appl. Phys. Lett.*, vol. 97, 2010, Art. no. 161108.

[9] T. Baba, "Photonic and iontronic sensing in GaInAsP semiconductor photonic crystal nanolasers," *Photonics*, vol. 6, no. 2, 2019, Art. no. 65.

[10] J. Lu and J. Vučković, "Nanophotonic computational design," *Opt. Exp.*, vol. 21, no. 11, pp. 13351–13367, 2013.

[11] C. M. Lalau-Keraly, S. Bhargava, O. D. Miller, and E. Yablonovitch, "Adjoint shape optimization applied to electromagnetic design," *Opt. Exp.*, vol. 21, no. 18, pp. 21693–21701, 2013.

[12] L. F. Frellsen, Y. Ding, O. Sigmund, and L. H. Frandsen, "Topology optimized mode multiplexing in silicon-on-insulator photonic wire waveguides," *Opt. Exp.*, vol. 24, no. 15, pp. 16866–16873, 2016.

[13] A. Y. Piggott, J. Petykiewicz, L. Su, and J. Vučković, "Fabrication-constrained nanophotonic inverse design," *Sci. Rep.*, vol. 7, 2017, Art. no. 1786.

[14] T. Asano and S. Noda, "Iterative optimization of photonic crystal nanocavity designs by using deep neural networks," *Nanophotonics*, vol. 8, no. 12, pp. 2243–2256, 2019.

[15] R. Abe, T. Takeda, R. Shiratori, S. Shirakawa, S. Saito, and T. Baba, "Optimization of an H0 photonic crystal nanocavity using machine learning," *Opt. Lett.*, vol. 45, no. 2, pp. 319–322, 2020.

[16] N. Hansen and A. Ostermeier, "Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation," in *Proc. IEEE Int. Conf. Evol. Computation*, 1996, pp. 312–317.

[17] N. Hansen and A. Ostmeier, "Completely derandomized self-adaptation in evolution strategies," *Evol. Comput.*, vol. 9, no. 2, pp. 159–195, Jun. 2001.

[18] N. Hansen, "The CMA evolution strategy: A tutorial," 2016, *arXiv:1604.00772*.

[19] Keras API Reference, Keras SIG, Accessed: Nov. 18, 2021. [Online]. Available: https://keras.io/api/

[20] DEAP documentation, c2009-2021, DEAP Project, May 8, 2021, Accessed: Nov. 18, 2021. [Online]. Available: https://deap.readthedocs.io/en/master/index.html

[21] sklearn.decomposition.PCA, c2007-2021, scikit-learn developers, Accessed: Jan. 7, 2022. [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.decomposition.PCA.html

[22] H. Hagino, Y. Takahashi, Y. Tanaka, T. Asano, and S. Noda, "Effects of fluctuation in air hole radii and positions on optical characteristics in photonic crystal heterostructure nanocavities," *Phys. Rev. B*, vol. 79, 2009, Art. no. 085112.

[23] M. Minkov, U. P. Dharanipathy, R. Houdré, and V. Savona, "Statistics of the disorder-induced losses of high-Q photonic crystal cavities," *Opt. Exp.*, vol. 21, no. 23, pp. 28233–28245, 2013.

[24] opencv-python, opencv.org, Accessed: Dec. 27, 2021. [Online]. Available: https://github.com/opencv/opencv-python