Scalable and Robust Photonic Integrated Unitary Converter Based on Multiplane Light Conversion

Ryota Tanomura,* Rui Tang, Toshikazu Umezaki, Go Soma, Takuo Tanemura,† and Yoshiaki Nakano

Department of Electrical Engineering and Information Systems, School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

(Received 26 March 2021; revised 20 August 2021; accepted 5 January 2022; published 25 February 2022)

An optical unitary converter (OUC) that can convert a set of N mutually orthogonal optical modes into another set of arbitrary N orthogonal modes is expected to be a key device in diverse applications, including optical communication, deep learning, and quantum computing. While various types of OUC have been demonstrated on photonic integration platforms, the sensitivity against a slight deviation in waveguide dimensions has been the crucial issue in scaling N. Here, we demonstrate that an OUC based on the concept of multiplane light conversion (MPLC) shows outstanding robustness against waveguide deviations. Moreover, it becomes more and more insensitive to fabrication errors as we increase N, which is in clear contrast to the conventional OUC architecture, composed of 2×2 Mach-Zehnder interferometers. The physical origin behind this unique robustness and scalability is studied by considering a generalized OUC configuration. As a result, we reveal that the number of coupled modes in each stage plays an essential role in determining the sensitivity of the entire OUC. The maximal robustness is attained when all-to-all-coupled interferometers are employed, which are naturally implemented in a MPLC-based OUC.

DOI: 10.1103/PhysRevApplied.17.024071

I. INTRODUCTION

Reconfigurable unitary conversion in the optical domain is a crucial operation in various photonic signal-processing applications, in both classical and quantum systems. In particular, a large-scale integrated optical unitary converter (OUC) would be a key device in realizing low-power all-optical multi-input-multi-output processors for mode-multiplexed optical communication [1–3], optical accelerators for deep learning [4–10], and linear optical circuits for quantum computing [11–18]. For example, deep neural networks, which have been widely deployed to address a variety of problems, are suffering from the requirement for increasing amounts of computational power [10]. This issue may be solved by optical accelerators composed of OUCs, which enable linear matrix multiplication in the optical domain with minimal power consumption and latency [4]. While integrated OUCs have been experimentally demonstrated at a relatively small scale [2,4,9,11–18], larger-scale devices with more than 100 input modes are required for practical applications. As the device size increases, fabrication imperfections generally affect the overall performance. It is, therefore, essential to develop scalable and robust integrated OUCs.

An integrated $N \times N$ OUC is conventionally constructed by cascading multiple stages of 2×2 OUCs in a mesh configuration [19–21]. Each 2×2 OUC corresponds to a Mach-Zehnder interferometer (MZI), which comprises two 50:50 directional couplers (DCs) and two phase shifters. By setting all N^2 phase shifters appropriately, arbitrary $N \times N$ unitary operation can be obtained in a reconfigurable manner [19,21]. In such MZI-based OUCs, however, the accuracy of $N \times N$ unitary operation depends sensitively on the deviations of waveguide dimensions, especially when N is large [7,22–24]. This issue is related to the fact that the optical interactions in the respective MZIs are inherently local. In other words, a 2×2 MZI allows coupling only between the two adjacent modes, so that each stage of a MZI-based OUC is described by an $N \times N$ band matrix, which becomes more and more sparse as N increases. Due to this nonredundancy, each band matrix needs to be in an ideal condition to enable arbitrary $N \times N$ unitary transformation after N stages. We should note that such property is rather a general characteristic observed in other analog and quantum computing systems as well, such as coherent Ising machines, quantum annealers, and multiqubit processors [25–29], where the dense connectivity among multiple nodes has a

*ryota_tanomura@hotaka.t.u-tokyo.ac.jp
†tanemura@ee.t.u-tokyo.ac.jp

crucial impact on the performance of the entire system. To cope with this issue of MZI-based OUCs, it is proposed to insert extra redundant phase shifters to ensure the necessary full connectivity [22,23], at the cost of increased complexity, footprint, and power consumption. Various global optimization or local error correction techniques may also be employed to compensate for the device imperfectness [5,23,30–33], which, however, requires the errors to be within some acceptable range [23].

As a fundamentally different approach as compared with a MZI-based OUC, an OUC based on the multiplane light conversion (MPLC) concept has been studied recently [24,34–36]. Applications of MPLC-based OUCs to spatial-mode multiplexers [35,37,38] and multidimensional quantum gates [39–41] have also been demonstrated successfully. In this scheme, arbitrary $N \times N$ unitary conversion is obtained through cascaded stages of an N-dimensional fixed dense unitary transformation (i.e., mode-mixing layer) and a phase shifter array. MPLC-based OUCs have been implemented successfully by free-space optics [35, 37–41] and on integrated photonic platforms [3,36,42–44]. Due to the mode-mixing layer, which ensures all-to-all coupling via an $N \times N$ dense matrix, a MPLC-based OUC is expected to exhibit inherently different scalability and robustness, compared with a conventional MZI-based OUC.

In this paper, we reveal that a MPLC-based OUC with multiport DCs has unique scalability and excellent robustness against waveguide deviations without the need for redundant phase shifters. Surprisingly, the error sensitivity of the MPLC-based OUC drops rapidly at large N. This is in clear contrast to the MZI-based OUC, having the same number of phase shifters, which becomes more and more sensitive as N increases. By considering generalized circuit configurations with various degrees of coupling, we comprehensively evaluate the physical origin of the difference between the two schemes. As a result, we find that the number of coupled modes at each mode-mixing layer plays an essential role in determining the robustness of the entire OUC; the maximal robustness is obtained when all-to-all-coupled interferometers are employed, which are naturally implemented in the MPLC-based OUC.

II. OPERATING PRINCIPLE

An OUC converts an N-dimensional vector $\mathbf{a}^{\mathbf{in}}$, which describes the complex amplitudes of N optical modes at the input, into another N-dimensional vector $\mathbf{a}^{\mathbf{out}}$ at the output. Here, $\mathbf{a}^{\mathbf{out}}$ can be expressed as $\mathbf{a}^{\mathbf{out}} = \mathbf{U}\mathbf{a}^{\mathbf{in}}$, where $\mathbf{U}$ is an $N \times N$ unitary matrix.

Figure 1(a) shows the typical architecture of the MZI-based OUC [21]. It consists of 2×2 MZIs, each of which comprises two 2×2 DCs and two tunable phase shifters as shown in the inset of Fig. 1(a). The splitting ratio of the DC needs to be 50:50. Since the number of MZIs is

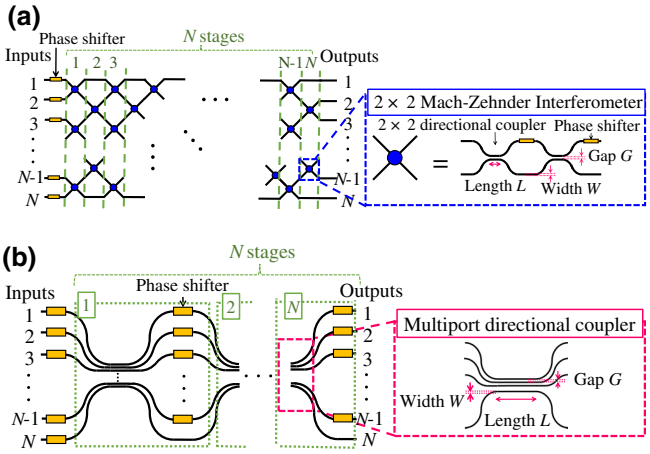

FIG. 1. Schematics of (a) $N \times N$ MZI-based OUC and (b) $N \times N$ MPLC-based OUC with multiport DCs. The MZI-based OUC is composed of a mesh of $N(N-1)/2$ MZIs. Each MZI, represented by a blue dot, consists of two series of 2×2 DC and a phase shifter. The MPLC-based OUC, in contrast, consists of N-port DCs and phase shifter arrays. Both 2×2 DC in the MZI-based OUC and multiport DC in the MPLC-based OUC are designed as a function of the waveguide width W, gap G, and coupler length L.

$N(N-1)/2$ and N additional phase shifters are required at the input stage, the total number of phase shifters is N^2. By tuning all N^2 phase shifters, arbitrary $\mathbf{U}$ is obtained [21].

In contrast, Fig. 1(b) shows the architecture of the integrated MPLC-based OUC with multiport DCs [36,45]. It consists of N stages of an N-port DC that functions as a mode-mixing layer and an N-port phase shifter array that can be controlled independently. The transfer matrix of the entire circuit is written as

$$\mathbf{U} = \mathbf{\Phi}^{(N)}\mathbf{M}^{(N)}\mathbf{\Phi}^{(N-1)} \cdots \mathbf{\Phi}^{(1)}\mathbf{M}^{(1)}\mathbf{\Phi}^{(0)}, \tag{1}$$

where $\mathbf{M}^{(j)}$ and $\mathbf{\Phi}^{(j)}$ $(j = 1, 2, \ldots, N)$ are the $N \times N$ transfer matrices of the multiport DC and the phase shifter array, respectively, at the jth stage. $\mathbf{\Phi}^{(j)}$ is a diagonal matrix, expressed as

$$\mathbf{\Phi}^{(j)} = \mathrm{diag}[\exp(i\phi^{(j)})]. \tag{2}$$

Here, $\phi^{(j)} = (\phi_1^{(j)}, \phi_2^{(j)}, \ldots, \phi_N^{(j)})$, where $\phi_m^{(j)}$ is the phase shift at the mth port of the jth stage. Note that we need N phase shifters at the input stage ($j = 0$), while the other stages require only $N-1$ phase shifters. The total number of phase shifters is, therefore, $N + N(N-1) = N^2$, which is identical to that for the MZI-based OUC. As a unique feature of the MPLC-based OUC in contrast to MZI-based OUC, the mode-mixing layers $\mathbf{M}^{(j)}$ neither need to be identical in all stages nor need to be uniform among all ports, but are only required to provide substantial coupling among all modes [24,44,45]. We should also note

that Fig. 1(b) only shows s schematic to explain the circuit architecture. In actual implementations [36,44], the waveguides between the stages can be designed in S shapes to avoid port-dependent propagation losses and wavelength sensitivity.

III. SCALABILITY ANALYSIS

The robustness and scalability of the two OUC architectures presented in Fig. 1 are compared numerically. As a test case of interest, we consider standard silicon photonic circuits with a 220-nm-thick silicon-on-insulator device layer. The waveguide width W and gap G at the DC sections (see Fig. 1 for definitions) are set to 460 and 250 nm, respectively. The 2×2 DC length in Fig. 1(a) is then fixed to 20 μm, corresponding to the 50:50 splitting length under this condition. In contrast, the coupling length of the N-port DC in Fig. 1(b) is scaled as $N \times 10$ μm. In both cases, we first derive the transfer matrix of a single 2×2 DC or a multiport DC numerically by the eigenmode expansion method (EME) [46], from which the matrix $\mathbf{M}^{(j)}$ is obtained. Then, the transfer matrix of the entire circuit $\mathbf{U}$ is calculated by Eq. (1) for the given phase shifter conditions. For simplicity, we assume lossless waveguides and phase shifters, each driven independently by an 8-bit digital-to-analog converter (DAC).

First, the sensitivity of the OUC performance is investigated against variations in W and G. For each case, all N^2 phase shifters are optimized to obtain the target unitary matrix $\mathbf{U}'$. Then, the deviation of the generated $N \times N$ transfer matrix $\mathbf{U}$ from $\mathbf{U}'$ is evaluated by calculating the mean-square error (MSE) f_{MSE}, defined as

$$f_{\mathrm{MSE}} = \frac{1}{N^2} \sum_{i,j}^{N} |U'_{ij} - U_{ij}|^2. \tag{3}$$

To ensure the reliability of the analysis, we generate 20 different Haar random $N \times N$ matrixes as the target unitary matrix $\mathbf{U}'$ and take the average of f_{MSE} in all cases of $\mathbf{U}'$. The G and W are changed in [150, 350] nm and [360, 560] nm with 10 nm step, respectively. In optimizing the MZI-based OUC, we first obtain the 2×2 transfer matrix of a single MZI for respective cases of W and G by the EME-based simulation. Using the derived transfer matrix of the 2×2 DC, the phase shifter values in each MZI are determined sequentially in the same order as for the decomposition method presented by Clements *et al.* [21]. For a MZI-based OUC with imperfect MZIs, we derive the phase shifter values that best approximate the target unitary matrix with a minimal f_{MSE}. For reference, we compare the resultant performance of OUC with that obtained by the simulated annealing algorithm and confirm that they are consistent (see Appendix A for details of the optimization procedure).

In optimizing the MPLC-based OUC, the simulated annealing algorithm [47] is employed when $N \leq 32$, whereas the adaptive moment estimation (ADAM) algorithm [48] is applied when $N = 128$ to derive the quasioptimal conditions of phase shifters. (See Appendix B for details.) Note that these general-purpose algorithms are employed for the sake of simplicity in this work, but a more efficient algorithm specific to MPLC-based OUCs may be developed in future to accelerate the convergence. Figure 2 shows f_{MSE} after optimizing all N^2 phase shifters in both architectures with increasing N. In all cases, the minimum f_{MSE} is limited to around -50 dB due to the 8-bit resolution of the DACs. (See Appendix C for the effect of the DAC resolution on the OUC performance.) For the case of the MZI-based OUC, f_{MSE} is suppressed below -45 dB only within a limited range of W and G, at which each 2×2 DC attains a precise 50:50 splitting ratio. Moreover, this regime shrinks rapidly as we increase N, meaning that arbitrary unitary conversion can no longer be obtained even after global optimization of all N^2 phase shifters. This issue would severely limit the scalability of this architecture. In contrast, for the case of the MPLC-based OUC, Fig. 2 clearly shows that f_{MSE} is less sensitive to W and G. More interestingly, this trend is enhanced as N increases to 128, where f_{MSE} is nearly insensitive to W and G.

For quantitative evaluation, the error sensitivities of both OUC configurations are plotted in Fig. 3 as a function of N. Here, the error sensitivity is defined as the largest increase in f_{MSE} when G or W deviates by 10 nm from the optimum point in Fig. 2. Figure 3 clearly indicates that MZI- and MPLC-based OUCs show contrary trends: the error sensitivity of the MZI-based OUC increases monotonically with N and approaches 20 dB/(10 nm) at $N = 128$, whereas that of the MPLC-based OUC converges to 0 as we increase N. Given that the state-of-the-art silicon photonic waveguides inevitably have nanometer-scale errors [49,50], this unique robustness of the MPLC-based OUC architecture would provide a significant advantage in constructing large-scale OUCs with $N > 100$.

The results shown in Figs. 2 and 3 imply that the accuracy of the MZI-based OUC depends sensitively on the splitting ratio of each 2×2 DC, whereas the MPLC-based OUC seems rather insensitive to the precise characteristics of the multiport DC, especially when N increases. To verify this assumption, we derive f_{MSE} of the two schemes as a function of κL, where κ is the coupling coefficient between adjacent waveguides at DCs and L is the coupling length. The results are shown in Fig. 4. We can confirm that the performance of the MZI-based OUC changes periodically with κL: when $\kappa L = (2m + 1)\pi/4$ (m is an integer), the 2×2 DC has an ideal 50:50 splitting ratio and thus f_{MSE} is suppressed, whereas f_{MSE} increases when $\kappa L = m\pi/2$. In contrast, f_{MSE} of the MPLC-based OUC quickly drops

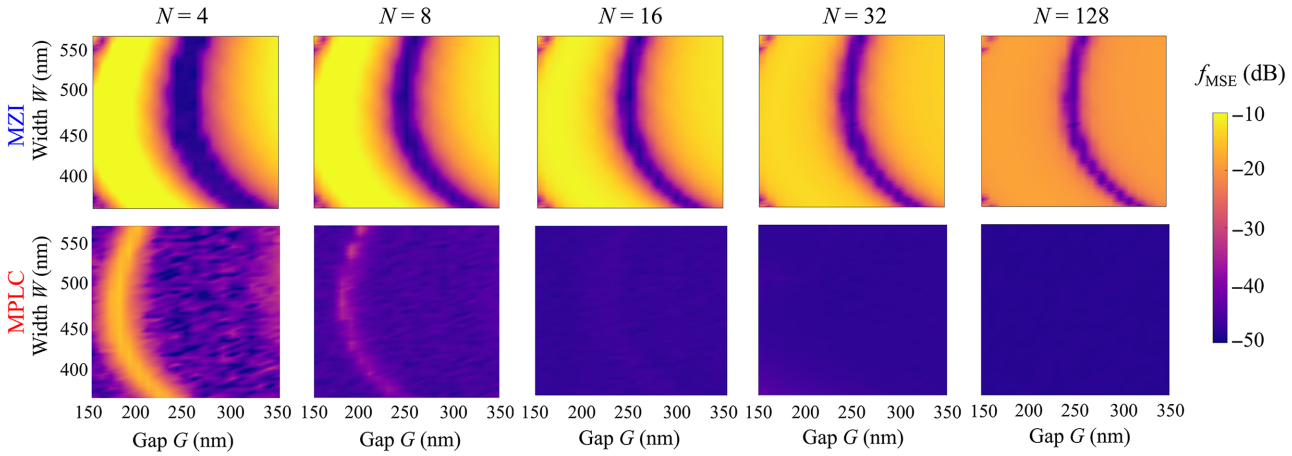

FIG. 2. Calculated f_{MSE} of MZI-based OUC (upper row) and MPLC-based OUC (lower row) after tuning all phase shifters to minimize the error. The results for $N = 4$, 8, 16, 32, and 128 are plotted as a function of width W and gap G. For the case of the MZI-based OUC, f_{MSE} is suppressed below -40 dB only within a limited range of W and G, and this range shrinks rapidly as N increases. In contrast, the MPLC-based OUC exhibits smaller f_{MSE} in a broader range, and becomes more and more insensitive to W and G as N increases.

as κL increases, and this behavior becomes more and more robust as N increases, in agreement with Fig. 3.

These results imply that the mode-mixing layers of the MPLC-based OUC function effectively as long as they offer sufficient coupling among all input modes, which is consistent with previous works [24,44,45]. Since this all-to-all coupling should be more easily achieved by transmitting through a number of stages, we can understand that the robustness of the MPLC-based OUC improves with increasing N, as shown in Fig. 3.

IV. EFFECTS OF ALL-TO-ALL COUPLING ON ERROR SENSITIVITY

From the above results and discussions, we assume that the all-to-all coupling provided at the mixing layers of the MPLC-based OUC plays an essential role in attaining the unique scalability and excellent error tolerance. To investigate this aspect in a comprehensive manner, we consider a generalized OUC configuration as shown in Fig. 5(a). The transfer matrix of this circuit is expressed as

$$\mathbf{U} = \mathbf{T}^{(N)}\mathbf{T}^{(N-1)}\cdots\mathbf{T}^{(1)}, \tag{4}$$

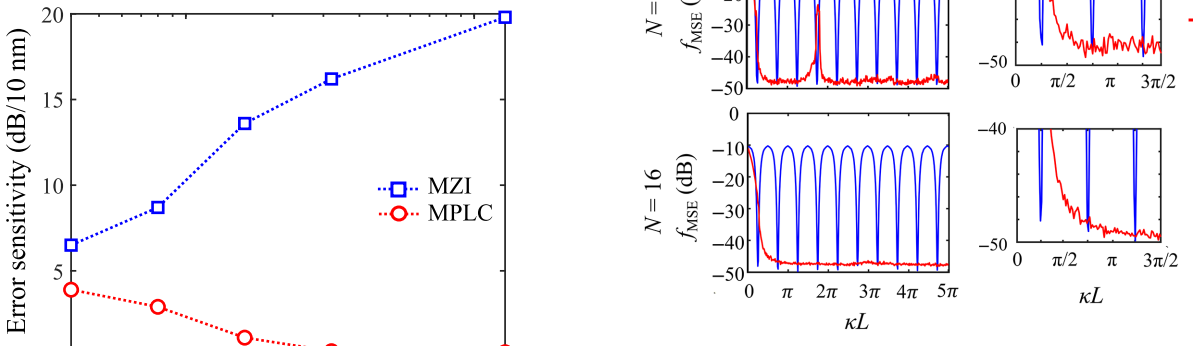

FIG. 3. Error sensitivities (defined as the largest increase in f_{MSE} when W or G deviates by 10 nm) of MZI- and MPLC-based OUCs versus N, derived at the optimal points in Fig. 2.

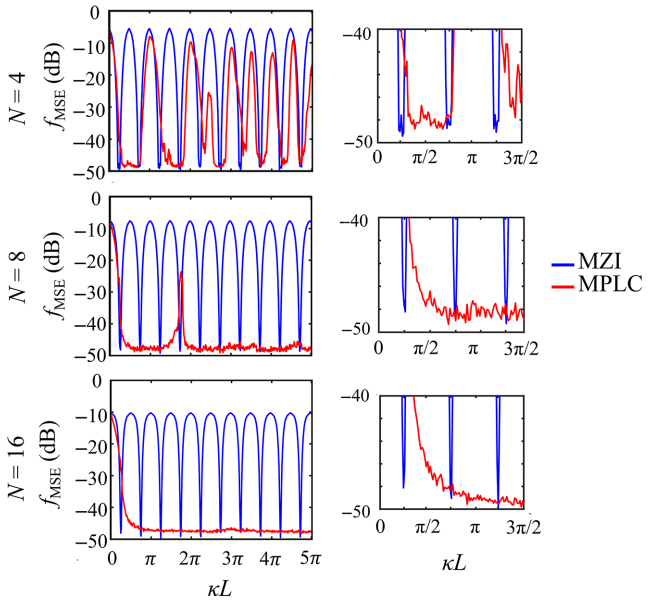

FIG. 4. f_{MSE} as a function of κL for the MZI-based OUC (blue) and MPLC-based OUC (red) with $N = 4$, 8, and 16. For the MZI-based OUC, f_{MSE} changes periodically with κL, corresponding to the change in the splitting ratio of each 2×2 DC. In contrast, f_{MSE} of the MPLC-based OUC quickly drops as κL increases, and this behavior becomes more and more robust as N increases.

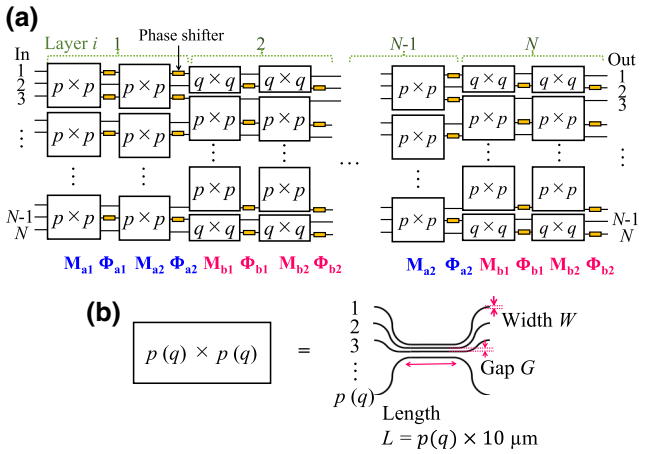

FIG. 5. Generalized N-port OUC with different degrees of coupling, represented by p and q. (a) Overall circuit configuration and (b) each $p \times p$ or $q \times q$ mode-mixing block, realized by a p- or q-port directional coupler, respectively. The MZI- and MPLC-based OUCs correspond to the cases with $(p, q) = (2, 1)$ and $(N, 0)$, respectively.

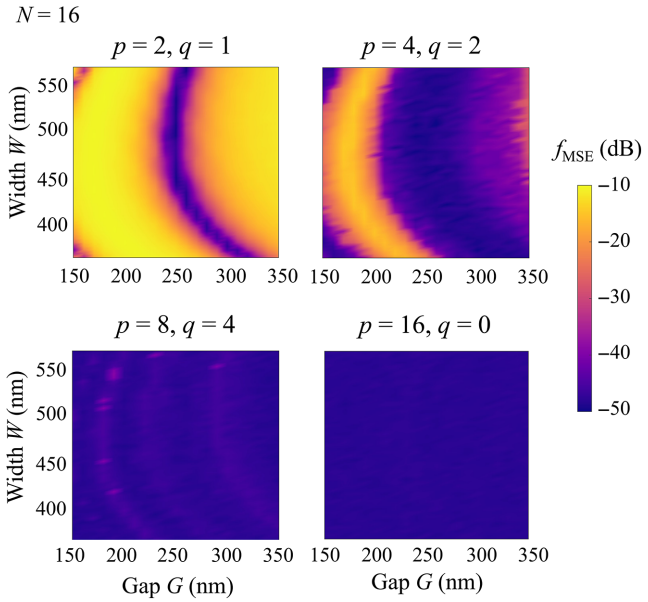

FIG. 6. f_{MSE} of a 16×16 OUC ($N = 16$) with different (p, q) as a function of width (W) and gap (G) of each DC. As p increases, f_{MSE} becomes more and more insensitive to the deviations.

where $\mathbf{T}^{(i)}$ is the transfer matrix of the ith layer and is written as

$$\mathbf{T}^{(i)} = \begin{cases} \mathbf{\Phi}_{\mathbf{b2}}^{(i)} \cdot \mathbf{M}_{\mathbf{b2}}^{(i)} \cdot \mathbf{\Phi}_{\mathbf{b1}}^{(i)} \cdot \mathbf{M}_{\mathbf{b1}}^{(i)} \text{ (when } i \text{ is odd)}, \\ \mathbf{\Phi}_{\mathbf{a2}}^{(i)} \cdot \mathbf{M}_{\mathbf{a2}}^{(i)} \cdot \mathbf{\Phi}_{\mathbf{a1}}^{(i)} \cdot \mathbf{M}_{\mathbf{a1}}^{(i)} \text{ (when } i \text{ is even)}. \end{cases} \quad (5)$$

Here, $\mathbf{\Phi}_{\boldsymbol{a}}$ ($\mathbf{\Phi}_{\boldsymbol{b}}$) describes each phase shifter stage, where the phase shifters are installed only at the odd (even) ports as shown in Fig. 5(a). On the other hand, $\mathbf{M_a}$ and $\mathbf{M_b}$ represent the mode-mixing stages. As shown in Fig. 5(a), $\mathbf{M_a}$ consists of an array of $p \times p$ multiport DCs, described in Fig. 5(b). In contrast, $\mathbf{M_b}$ consists of two blocks of $q \times q$ mode-mixing matrices at the top and the bottom channels in Fig. 5(a), whereas other ports are connected to $p \times p$ matrices. Note that p needs to be a submultiple of N. On the other hand, q is either $p/2$ or 0.

Using this generalized model, various types of OUC with different degrees of mode coupling can be analyzed in a unified manner. For example, the MZI-based OUC shown in Fig. 1(a) corresponds to a case with $(p, q) = (2, 1)$. By increasing p, we can increase the degree of coupling at each stage. As an extreme case, a circuit equivalent to the MPLC-based OUC in Fig. 1(b) is represented as $(p, q) = (N, 0)$. Through the same procedure as Fig. 2, f_{MSE} is obtained numerically as a function of W and G to investigate the dependence of the error sensitivity on (p, q).

Figure 6 shows f_{MSE} of a 16×16 OUC with various (p, q). [See Appendix E, which verifies that arbitrary unitary conversion can indeed be obtained in all cases of (p, q) with this generalized OUC configuration.] We can clearly see that as p increases, f_{MSE} becomes more and more insensitive against errors. In particular, in the extreme case of $(p, q) = (16, 0)$, the MSE is kept below -40 dB for all values of W and G. We, therefore, attribute the unique robustness of the MPLC-based OUC to the all-to-all coupling at each mode-mixing stage, which ensures the necessary full connectivity of all modes.

V. CONCLUSION

We have investigated the scalability and error tolerance of different types of OUCs. Unlike the conventional OUC architectures, composed of 2×2 MZIs, a MPLC-based OUC with multiport DCs is demonstrated to show superior robustness against waveguide deviations. Moreover, we find that it becomes more and more robust as the port count N increases, which is in clear contrast to the MZI-based OUC, whose sensitivity increases with N. More specifically, at a scale of $N = 128$, the performance of the MPLC-based OUC is demonstrated to be almost independent of the waveguide width and the gap. This robustness would provide a significant advantage in constructing a large-scale OUCs with $N > 100$, required for practical applications.

From comprehensive analyses of generalized OUC configurations, we reveal that the all-to-all coupling inside each multiport DC in the MPLC-based OUC plays an essential role in enhancing the robustness and provides the crucial difference from the MZI-based OUC. More specifically, the error sensitivity of the MZI-based OUC originates from the localized interactions at MZIs, which may result in a sparse $N \times N$ matrix at large N when they deviate from ideal conditions. In contrast, the MPLC-based

OUC naturally ensures the necessary full connectivity among all modes without the need for redundant phase shifters. The physical insight provided in this work should be useful in building large-scale OUCs for diverse applications, such as optical communication, deep learning, and quantum computing.

ACKNOWLEDGMENTS

We thank Yoshihiko Hasegawa, Quoc Hoan Tran, Tan Van Vu, and Taichiro Fukui for fruitful discussions. We also greatly appreciate the reviewers for their insightful comments. This work is in part supported by the JSPS KAKENHI (Grants No. JP20J21861, No. JP21K18168, and No. JP26000010) and MbSC2030.

APPENDIX A: OPTIMIZATION METHOD OF MZI-BASED OUC

An ideal MZI-based OUC can be optimized deterministically following the decomposition algorithm [21]. Similarly, for a MZI-based OUC with imperfect MZIs, if we have complete information about the 2×2 transfer matrix of each MZI, we can deterministically derive the condition that minimizes the error using the same procedure. When the 2×2 DC deviates from the ideal condition, it may not be possible to perfectly null each element in the matrix during the decomposition as presented in Ref. [21]. In such a case, we can still determine the phase shifter values that make the target element to zero with a minimal error. Following this procedure for all matrix elements, we can derive the condition that best approximates the target unitary matrix with a minimal f_{MSE}.

Figure 7(a) shows f_{MSE} of a 4×4 MZI-based OUC for various values of waveguide width (W) and gap (G) when the phase shifts are optimized using this method. For reference, Fig. 7(b) shows the results when we employ the simulated annealing algorithm to heuristically optimize the phase shifter conditions. We can confirm that both methods provide consistent results, with the former deterministic method giving slightly better accuracy as expected.

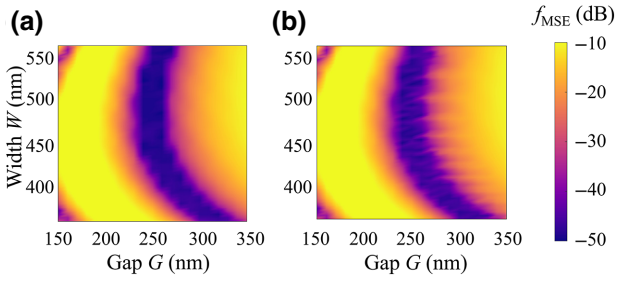

FIG. 7. f_{MSE} of a 4×4 MZI-based OUC as a function of W and G when the phase shifters are tuned (a) deterministically through the decomposition method and (b) heuristically by the simulated annealing algorithm. There is no significant difference in the OUC performance for the two cases, with the deterministic method (a) providing a slightly better accuracy.

APPENDIX B: OPTIMIZATION METHOD OF MPLC-BASED OUC

For a given target matrix, we derive the optimal condition of all N^2 phase shifters to minimize f_{MSE} defined in Eq. 3. When $N \leq 32$, we employ the simulated annealing algorithm [47] to find such a condition. The optimal condition typically converges after approximately $170N^2$ iterations. Using a standard multicore CPU (Core$^{\mathrm{TM}}$ i9-10920X) and 64 GB RAM, it takes around 20 min to reach convergence for $N = 32$. When $N = 128$, on the other hand, we use the ADAM algorithm [48] to reduce the computation time. In this algorithm, the phase shifter values are updated based on the gradient of f_{MSE}, so that we can reach convergence more efficiently (at around 18 000 iterations) compared with the simulated annealing method. Besides, the phase shifter values are updated one by one in the simulated annealing algorithm, while all phase shifters are updated at once in the ADAM algorithm.

APPENDIX C: EFFECT OF DAC RESOLUTION ON OUC PERFORMANCE

In our numerical simulations, we assume 8-bit DACs in driving the phase shifters to emulate a realistic situation. The resolution of DACs sets the upper limit of the OUC performance for both cases of MZI- and MPLC-based OUCs. Figure 8 shows the simulated performance of a 16×16 OUC as a function of the DAC resolution, where we assume $W = 460$ nm, $G = 250$ nm, and other conditions that are assumed in the main text. We can see that f_{MSE} of both types of OUC reduces as the number of

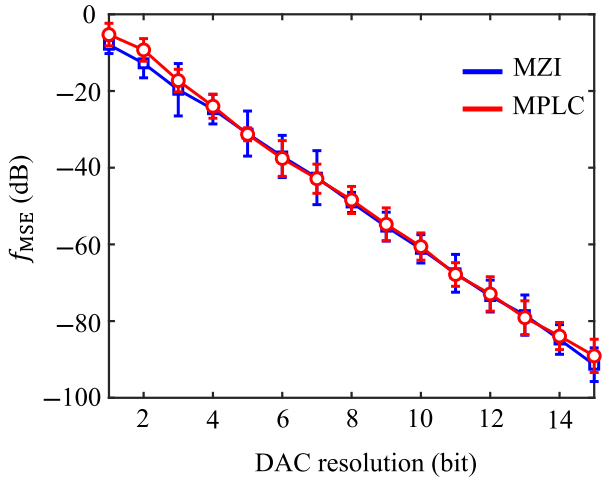

FIG. 8. f_{MSE} of 16×16 MZI- and MPLC-based OUCs with different DAC resolutions. The performance of the OUCs reduces in the same way for both schemes.

bits of DAC increases. Moreover, we can confirm that the performance of the MPLC-based OUC is comparable with that of the MZI-based OUC in all cases of DAC resolution simulated in Fig. 8.

APPENDIX D: FIDELITY OF MZI- AND MPLC-BASED OUCs

In the analyses in the main text, we use f_{MSE} for comparing the OUC performance. As an another evaluation factor, the fidelity of the matrix is also used in the literature, which is defined as [21]

$$F(\mathbf{U}', \mathbf{U}) = \left| \frac{\mathrm{tr}(\mathbf{U}^{\dagger}\mathbf{U}')}{\sqrt{N\mathrm{tr}(\mathbf{U}'^{\dagger}\mathbf{U}')}} \right|^2 . \tag{D1}$$

When the obtained matrix $\mathbf{U}$ is equal to the target matrix $\mathbf{U}'$, F is 1. In Fig. 9, we plot F of MZI- and MPLC-based OUCs calculated as a function of κL of each DC. We determine κ by setting G and W to 250 and 440 nm, respectively, and calculate F as a function of κL by varying L in increments of 1 μm. For each L, 20 matrices that are Haar random are tested and their average values are shown in Fig. 9. Similar to Fig. 4, the fidelity of the MZI-based OUC oscillates periodically depending on the splitting ratio of 2×2 DCs. On the other hand, the fidelity becomes more and more robust as N increases for the case of the MPLC-based OUC.

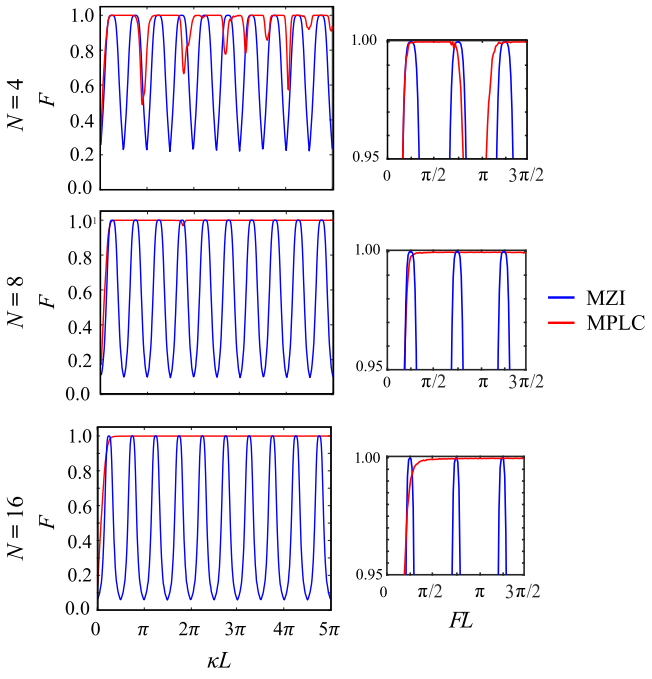

FIG. 9. Fidelity of MZI- and MPLC-based OUCs defined by Eq. (D1) as a function of κL. Similar to f_{MSE} (Fig. 4), the fidelity of the MZI-based OUC changes periodically with κL, while that of the MPLC-based OUC converges to a unity as κL increases.

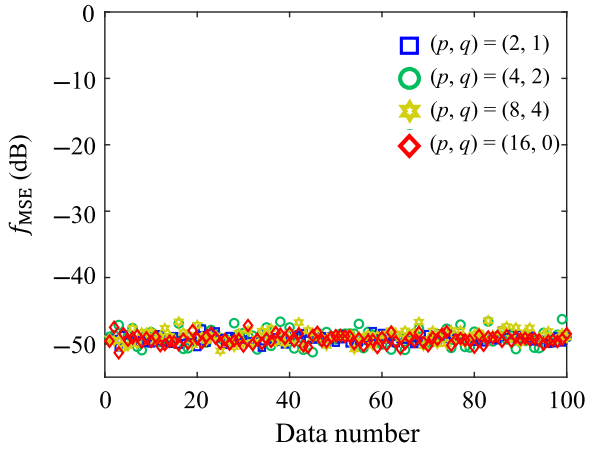

FIG. 10. f_{MSE} of 16×16 OUCs with $(p, q) = (2, 1), (4, 2), (8, 4)$, and $(16, 0)$ after programming to 100 different Haar-random unitary matrices.

APPENDIX E: ARBITRARINESS OF THE GENERALIZED OUC

To examine that the generalized OUC introduced in Fig. 4 can produce nearly arbitrary unitary matrices, we plot in Fig. 10 the MSE (f_{MSE}) of a 16×16 OUC after it is tuned to 100 randomly generated unitary matrices. We assume $W = 460$ nm, $G = 250$ nm, and all other parameters that are used in the main text. We can see that all 100 unitary conversions are realized accurately with f_{MSE} below -45 dB for all cases of (p, q).

[1] N. K. Fontaine, C. R. Doerr, M. A. Mestre, R. R. Ryf, P. J. Winzar, L. L. Buhi, Y. Sun, X. Jiang, and R. Lingle, in Optical Fiber Communication Conference (Optical Society of America, Los Angeles, 2012), p. PDP5B.1.

[2] A. Annoni, E. Guglielmi, M. Carminati, G. Ferrari, M. Sampietro, D. A. Miller, A. Melloni, and F. Morichetti, Unscrambling light—automatically undoing strong mixing between modes, Light Sci. Appl. **6**, e17110 (2017).

[3] R. Tang, T. Tanemura, S. Ghosh, K. Suzuki, K. Tanizawa, K. Ikeda, H. Kawashima, and Y. Nakano, Reconfigurable all-optical on-chip MIMO three-mode demultiplexing based on multi-plane light conversion, Opt. Lett. **43**, 1798 (2018).

[4] Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. BaehrJones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, and M. Soljačić, Deep learning with coherent nanophotonic circuits, Nat. Photonics **11**, 441 (2017).

[5] T. W. Hughes, M. Minkov, Y. Shi, and S. Fan, Training of photonic neural networks through in situ backpropagation, Optica **5**, 864 (2018).

[6] I. A. D. Williamson, T. W. Hughes, M. Minkov, B. Bartlett, S. Pai, and S. Fan, Reprogrammable electro-optic nonlinear activation functions for optical neural networks, IEEE J. Sel. Top. Quantum Electron. **26**, 7700412 (2020).

[7] M. Y.-S. Fang, S. Manipatruni, C. Wierzynski, A. Khosrowshahi, and M. R. Deweese, Design of optical neural networks with component imprecisions, Opt. Express **27**, 14009 (2019).

[8] S. Pai, I. A. D. Williamson, T. W. Hughes, M. Minkov, O. Solgaard, S. Fan, and D. A. B. Miller, Parallel programming of an arbitrary feedforward photonic network, IEEE J. Sel. Top. Quantum Electron. **26**, 6100813 (2020).

[9] H. Zhang, M. Gu, X. D. Jiang, J. Thompson, H. Cai, S. Paesani, R. Santagati, A. Laing, Y. Zhang, M. H. Yung, Y. Z. Shi, F. K. Muhammad, G. Q. Lo, X. S. Luo, B. Dong, D. L. Kwong, L. C. Kwek, and A. Q. Liu, An optical neural chip for implementing complex-valued neural network, Nat. Commun. **12**, 457 (2021).

[10] N. C. Harris, *et al.*, in Optical Fiber Communication Conference (Optical Society of America, San Diego, 2020), p. W3A.3.

[11] M. Grafe, R. Heilmann, M. Lebugle, D. Guzman-Silva, A. Perez-Leija, and A. Szameit, Integrated photonic quantum walks, J. Opt. **18**, 103002 (2016).

[12] J. Carolan, C. Harrold, C. Sparrow, E. Martín-López, N. J. Russell, J. W. Silverstone, P. J. Shadbolt, N. Matsuda, M. Oguma, M. Itoh, G. D. Marshall, M. G. Thompson, J. C. F. Matthews, T. Hashimoto, J. L. O'Brien, and A. Laing, Universal linear optics, Science **349**, 711 (2015).

[13] N. C. Harris, G. R. Steinbrecher, M. Prabhu, Y. Lahini, J. Mower, D. Bunandar, C. Chen, F. N. C. Wong, T. Baehr-Jones, M. Hochberg, S. Lloyd, and D. Englund, Quantum transport simulations in a programmable nanophotonic processor, Nat. Photonics **11**, 447 (2017).

[14] I. V. Dyakonov, I. A. Pogorelov, I. B. Bobrov, A. A. Kalinkin, S. S. Straupe, S. P. Kulik, P. V. Dyakonov, and S. A. Evlashin, Reconfigurable Photonics on a Glass Chip, Phys. Rev. Appl. **10**, 044048 (2018).

[15] J. Wang, S. Paesani, Y. Ding, R. Santagati, P. Skrzypczyk, A. Salavrakos, J. Tura, R. Augusiak, L. Mančinska, D. Bacco, D. Bonneau, J. W. Silverstone, Q. Gong, A. Acín, K. Rottwitt, L. K. Oxenløwe, J. L. O'Brien, A. Laing, and M. G. Thompson, Multidimensional quantum entanglement with large-scale integrated optics, Science **360**, 285 (2018).

[16] J. Carolan, M. Mohseni, J. P. Olson, M. Prabhu, C. Chen, D. Bunandar, M. Y. Niu, N. C. Harris, F. N. C. Wong, M. Hochberg, S. Lloyd, and D. Englund, Variational quantum unsampling on a quantum photonic processor, Nat. Phys. **16**, 322 (2020).

[17] J. M. Arrazola, *et al.*, Quantum circuits with many photons on a programmable nanophotonic chip, Nature **591**, 54 (2021).

[18] C. Taballione, R. van der Meer, H. J. Snijders, P. Hooijschuur, J. P. Epping, M. Goede, B. Kassenberg, P. Venderbosch, C. Toebes, H. Vlekkert, P. W. H. Pinkse, and J. J. Renema, A universal fully reconfigurable 12-mode quantum photonic processor, Mater. Quantum Technol. **1**, 035002 (2021).

[19] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, Experimental Realization of Any Discrete Unitary Operator, Phys. Rev. Lett. **73**, 58 (1994).

[20] D. A. B. Miller, Self-configuring universal linear optical component, Photonics Res. **1**, 1 (2013).

[21] W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley, Optimal design for universal multiport interferometers, Optica **3**, 1460 (2016).

[22] D. A. B. Miller, Perfect optics with imperfect components, Optica **2**, 747 (2015).

[23] S. Pai, B. Bartlett, O. Solgaard, and D. A. B. Miller, Matrix Optimization on Universal Unitary Photonic Devices, Phys. Rev. Appl. **11**, 064044 (2019).

[24] M. Y. Saygin, I. V. Kondratyev, I. V. Dyakonov, S. A. Mironov, S. S. Straupe, and S. P. Kulik, Robust Architecture for Programmable Universal Unitaries, Phys. Rev. Lett. **124**, 010501 (2020).

[25] P. L. McMahon, A. Marandi, Y. Haribara, R. Hamerly, C. Langrock, S. Tamate, T. Inagaki, H. Takesue, S. Utsunomiya, K. Aihara, R. L. Byer, M. M. Fejer, H. Mabuchi, and Y. Yamamoto, A fully programmable 100-spin coherent Ising machine with all-to-all connections, Science **354**, 614 (2016).

[26] S. E. Nigg, N. Lörch, and R. P. Tiwari, Robust quantum optimizer with full connectivity, Sci. Adv. **3**, e1602273 (2017).

[27] S. Puri, C. K. Andersen, A. L. Grimsmo, and A. Blais, Quantum annealing with all-to-all connected nonlinear oscillators, Nat. Commun. **8**, 15785 (2017).

[28] T. Roy, M. Chand, A. Bhattacharjee, S. Hazra, S. Kundu, K. Damle, and R. Vijay, Multimode superconducting circuits for realizing strongly coupled multiqubit processor units, Phys. Rev. A **98**, 052318 (2018).

[29] T. Roy, S. Kundu, M. Chand, S. Hazra, N. Nehra, R. Cosmic, A. Ranadive, M. P. Patankar, K. Damle, and R. Vijay, Implementation of Pairwise Longitudinal Coupling in a Three-Qubit Superconducting Circuit, Phys. Rev. Appl. **7**, 054025 (2017).

[30] D. Pérez-López, A. López, P. DasMahapatra, and J. Capmany, Multipurpose self-configuration of programmable photonic circuits, Nat. Commun. **11**, 6359 (2020).

[31] S. Bandyopadhyay, R. Harmerly, and D. Englund, Hardware error correction for programmable photonics, Optica **8**, 1247 (2021).

[32] R. Harmerly, S. bandyopadhyay, and D. Englund, Accurate self-configuration of rectangular multiport interferometers, ArXiv:2106.03249 (2021).

[33] R. Harmerly, S. bandyopadhyay, and D. Englund, Stability of self-configuring large multiport interferometers, ArXiv:2106.04363 (2021).

[34] J.-F. Morizur, L. Nicholls, P. Jian, S. Armstrong, N. Treps, B. Hage, M. Hsu, W. Bowen, J. Janousek, and H.-A. Bachor, Programmable unitary spatial mode manipulation, J. Opt. Soc. Am. A **27**, 2524 (2010).

[35] G. Labroille, B. Denolle, P. Jian, P. Genevaux, N. Treps, and J.-F. Morizur, Efficient and mode selective spatial mode multiplexer based on multi-plane light conversion, Opt. Express **22**, 15599 (2014).

[36] R. Tanomura, R. Tang, S. Ghosh, T. Tanemura, and Y. Nakano, Robust integrated optical unitary converter using multiport directional couplers, J. Lightwave Technol. **38**, 60 (2020).

[37] N. K. Fontaine, R. Ryf, H. Chen, D. T. Neilson, K. Kim, and J. Carpenter, Laguerre-Gaussian mode sorter, Nat. Commun. **10**, 1865 (2019).

[38] G. Rademacher, B. J. Puttnam, R. S. Luís, T. A. Eriksson, N. K. Fontaine, M. Mazur, H. Chen, R. Ryf, D. T. Neilson, P. Sillard, F. Achten, Y. Awaji, and H. Furukawa, Peta-bit-per-second optical communications system using a standard cladding diameter 15-mode fiber, Nat. Commun. **12**, 4238 (2021).

[39] F. Brandt, M. Hiekkamäki, F. Bouchard, M. Huber, and R. Fickler, High-dimensional quantum gates using full-field spatial modes of photons, Optica **7**, 98 (2020).

[40] M. Hiekkamäki and R. Fickler, High-Dimensional Two-Photon Interference Effects in Spatial Modes, Phys. Rev. Lett. **126**, 123601 (2021).

[41] O. Lib, K. Sulimany, and Y. Bromberg, Reconfigurable synthesizer for quantum information processing of high-dimensional entangled photons, ArXiv:2108.02258 (2021).

[42] R. Tanomura, R. Tang, T. Suganuma, K. Okawa, E. Kato, T. Tanemura, and Y. Nakano, Monolihic InP optical unitary converter based on multi-plane light conversion, Opt. Express **28**, 25392 (2020).

[43] R. Tanomura, R. Tang, T. Tanemura, and Y. Nakano, Integrated InP optical unitary converter with compact half-integer multimode interferometers, Opt. Express **29**, 43414 (2021).

[44] R. Tang, R. Tanomura, T. Tanemura, and Y. Nakano, Ten-port unitary optical processor on a silicon photonic chip, ACS Photonics **8**, 2074 (2021).

[45] R. Tang, T. Tanemura, and Y. Nakano, in OptoElectronics and Communications Conference and Photonics Global Conference (Optical Society of America, Singapore, 2017).

[46] D. F. G. Gallagher and T. P. Felici, in *Integrated Optics: Devices, Materials, and Technologies VII, Society of Photo-Optical Instrumentation Engineers 4987* (SPIE, San Jose, 2003).

[47] S. Kirkpatrick, C. D. Gelatt Jr, and M. P. Vecchi, Optimization by simulated annealing, Science **220**, 671 (1983).

[48] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization, ArXiv:1412.6980 (2014).

[49] D.-X. Xu, J. H. Schmid, G. T. Reed, G. Z. Mashanovich, D. J. Thomson, M. Nedeljkovic, X. Chen, D. Van Thourhout, S. Keyvaninia, and S. K. Selvaraja, Silicon photonic integration platform have we found the sweet spot, IEEE J. Sel. Top. Quantum Electron. **20**, 189 (2014).

[50] L. Chrostowski, X. Wang, J. Flueckiger, Y. Wu, Y. Wang, and S. Talebi Fard, in Optical Fiber Communication Conference (Optical Society of America, San Francisco, 2014), p. Th2A.37.