

## Lower-depth programmable linear optical processors

Rui Tang<sup>1</sup>,\* Ryota Tanomura<sup>1</sup>, Takuo Tanemura, and Yoshiaki Nakano<sup>1</sup>

*Department of Electrical Engineering and Information Systems, The University of Tokyo, Tokyo 113-8656, Japan*



(Received 10 June 2023; revised 12 October 2023; accepted 11 January 2024; published 26 January 2024)

Programmable linear optical processors (LOPs) can have widespread applications in computing and information processing due to their capabilities of implementing reconfigurable on-chip linear transformations. A conventional LOP that uses a mesh of Mach-Zehnder interferometers (MZIs) requires  $2N + 3$  stages of phase shifters for  $N \times N$  matrices. However, it is beneficial to reduce the number of phase-shifter stages to realize a more compact and lower-loss LOP, especially when long and lossy electro-optic phase shifters are used. In this work, we propose a structure for LOPs that can implement arbitrary matrices as long as they can be realized by previous MZI-based schemes. Through numerical analysis, we further show that the number of phase-shifter stages in the proposed structure can be reduced to  $N + 2$  and  $N + 3$  for a large number of random dense matrices and sparse matrices, respectively. This work contributes to the realization of compact, low-loss, and energy-efficient programmable LOPs.

DOI: [10.1103/PhysRevApplied.21.014054](https://doi.org/10.1103/PhysRevApplied.21.014054)

### I. INTRODUCTION

Programmable linear optical processors (LOPs) capable of implementing reconfigurable on-chip linear transformations have attracted increasing attention in recent years due to their promising applications in computing and information processing [1–5]. By programming the transfer matrix of optical processors, parallel tasks such as matrix multiplication can be directly performed in the optical domain, which may significantly reduce latency and lower energy consumption compared to their electronic counterparts [6–9].

So far, to implement an arbitrary matrix on a LOP, the matrix first needs to be decomposed into the product of two unitary matrices and a diagonal matrix via singular-value decomposition [10], as illustrated in Fig. 1(a). The unitary matrices are realized using the universal multiport interferometer architectures, which consist of a mesh of tunable Mach-Zehnder interferometers (MZIs) [11–15]. Bell *et al.* have proposed a compact MZI-based structure that requires  $N + 2$  stages of phase shifters for realizing arbitrary  $N \times N$  unitary matrices [14], as shown in Fig. 1(b). The diagonal matrix is realized using either a gain or absorber array or an MZI array. In practice, it is preferred to construct the entire LOP only using MZIs for easier fabrication and control. In this case, the structure of a LOP using the Bell structure is schematically shown in Fig. 1(c). Here, the last phase-shifter stage in the section

for  $\mathbf{V}$  and the first phase-shifter stage in the section for  $\mathbf{U}$  have been absorbed into the MZI array for  $\mathbf{\Sigma}$ . Therefore, for a conventional  $N \times N$  LOP only using MZIs, in its most compact form,  $2N + 3$  phase-shifter stages are required.

While thermo-optic (TO) phase shifters are widely used in existing devices [3,16], a TO phase shifter typically consumes more than 1 mW of power and is therefore not suitable for large-scale devices [17–19]. To realize a high-speed and low-power LOP, electro-optic (EO) phase shifters are highly desirable, since a gigahertz modulation bandwidth and ultralow power consumption during static operations can be easily achieved [20–23]. However, these phase shifters tend to have non-negligible insertion loss or can be millimeters in length, severely hindering their applications in LOPs. Therefore, it would be highly beneficial to reduce the number of phase-shifter stages for the realization of compact and low-loss LOPs using EO phase shifters.

In this work, we propose a structure for programmable LOPs based on the concept of multiplane light conversion (MPLC) [24–26]. We show that using  $N$  input-output ports out of an  $N' \times N'$  universal multiport interferometer ( $N' \geq 2N$ ), an arbitrary  $N \times N$  matrix can be obtained as long as it can be realized by previous MZI-based schemes. Through numerical analysis, we further show that the number of phase-shifter stages in the proposed structure can be reduced to  $N + 2$  and  $N + 3$  for a large number of random dense matrices and sparse matrices, respectively. This work contributes to the realization of compact, low-loss, and energy-efficient programmable LOPs.

\*ruitang@mosfet.t.u-tokyo.ac.jp

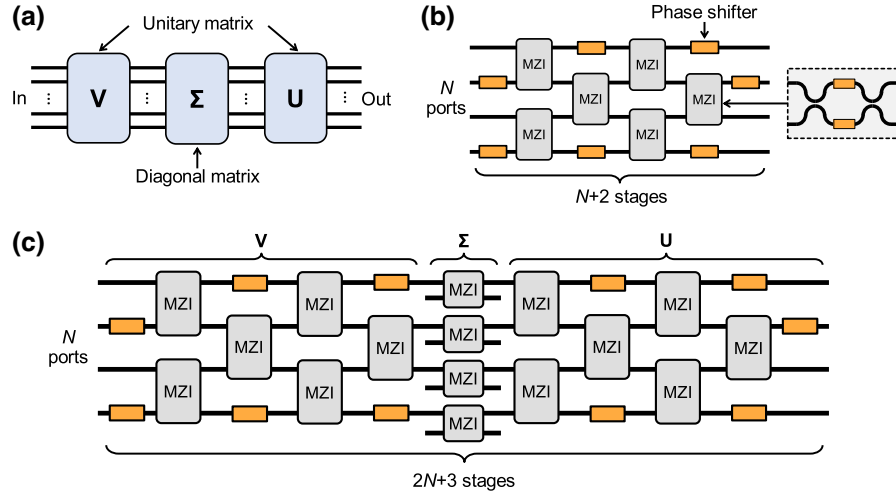


FIG. 1. (a) Conventional programmable LOPs based on singular-value decomposition. The target matrix is first decomposed into the product of two unitary matrices and a diagonal matrix. (b) A compact universal multiport interferometer proposed by Bell *et al.* for implementing arbitrary unitary matrices [14].  $N + 2$  stages of phase shifters are needed ( $N = 4$  in this figure). (c) The structure of a LOP using the Bell structure. The last phase-shifter stage in the section for  $\mathbf{V}$  and the first phase-shifter stage in the section for  $\mathbf{U}$  have been absorbed into the MZI array for  $\Sigma$ .  $2N + 3$  stages of phase shifters are needed ( $N = 4$  in this figure).

## II. PRINCIPLE

Our proposed structure is illustrated in Fig. 2. We use  $N$  input-output ports out of an  $N' \times N'$  universal multiport interferometer, which consists of cascaded phase-shifter arrays and  $N' \times N'$  couplers. The complex-valued transfer matrix can be calculated by multiplying the transfer matrices of all phase-shifter arrays and couplers. The couplers can be multimode interference (MMI) couplers [27–29], or multiport directional couplers (MDCs) [30–33]. As explained in the previous works [31,32,34,35], for schemes based on MPLC, the coupler is not required to have a specific transfer matrix as long as it properly scrambles the transmitted light.

For an  $N \times N$  complex matrix  $\mathbf{S}$  that can be realized in previous schemes only using MZIs, as shown in Fig. 1(c),

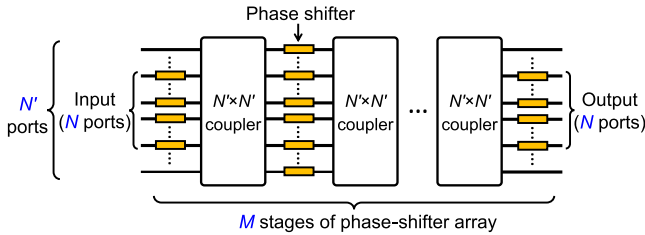


FIG. 2. The proposed LOP structure. While the whole device has  $N'$  ports, only  $N$  ports are used as input-output ports. For unused input and output ports, phase shifters are not needed and thus are omitted in the first and last stages. For the stages in between two  $N' \times N'$  couplers, all  $N'$  phase shifters are used. The couplers can be multimode interference (MMI) couplers or multiport directional couplers.  $M$  is the number of phase-shifter arrays.

it can be written in the form of

$$\mathbf{S} = \mathbf{U}\Sigma\mathbf{V}, \quad (1)$$

where  $\mathbf{U}$  and  $\mathbf{V}$  are two  $N \times N$  unitary matrices and  $\Sigma$  is a  $N \times N$  non-negative real diagonal matrix. Since we have assumed that  $\Sigma$  is realized by an MZI array, it can be written in the form of

$$\Sigma = \begin{pmatrix} \sin \alpha_1 & & & \\ & \sin \alpha_2 & & \\ & & \ddots & \\ & & & \sin \alpha_N \end{pmatrix}, \quad (2)$$

where  $\alpha_i \in [0, \pi/2]$  ( $i = 1, 2, \dots, N$ ).

We now show that using  $N$  input-output ports of a  $N' \times N'$  universal multiport interferometer, which can implement arbitrary  $N' \times N'$  unitary matrix, an arbitrary  $\mathbf{S}$  in the form of Eq. (1) can be realized provided that  $N' \geq 2N$ . We first construct two  $N' \times N'$  unitary matrices  $\mathbf{U}'$  and  $\mathbf{V}'$  from  $\mathbf{U}$  and  $\mathbf{V}$ , respectively:

$$\mathbf{U}' = \begin{pmatrix} \mathbf{U} & \mathbf{O} \\ & \mathbf{O} \end{pmatrix}, \quad (3)$$

$$\mathbf{V}' = \begin{pmatrix} \mathbf{V} & \mathbf{P} \\ & \mathbf{P} \end{pmatrix}, \quad (4)$$

where  $\mathbf{O}$  and  $\mathbf{P}$  are arbitrary unitary matrices with  $N' - N$  rows and columns. We then assume  $N' \geq 2N$  and construct a  $N' \times N'$  matrix  $\Sigma'$  in the form of

$$\Sigma' = \left( \begin{array}{ccc|ccc} \sin \alpha_1 & & & j \cos \alpha_1 & & \\ & \sin \alpha_2 & & & j \cos \alpha_2 & \\ & & \ddots & & & \ddots \\ & & & \sin \alpha_N & & j \cos \alpha_N \\ \hline j \cos \alpha_1 & & & \sin \alpha_1 & & \\ & j \cos \alpha_2 & & & \sin \alpha_2 & \\ & & \ddots & & & \ddots \\ & & & j \cos \alpha_N & & \sin \alpha_N \\ \hline & & & & & \mathbf{Q} \end{array} \right), \quad (5)$$

where  $j$  is the imaginary unit and  $\mathbf{Q}$  is an arbitrary unitary matrix with  $N' - 2N$  rows and columns.  $\mathbf{Q}$  vanishes at the boundary case ( $N' = 2N$ ). It is easy to verify that  $\Sigma'$  is a unitary matrix and can be written as

$$\Sigma' = \begin{pmatrix} \Sigma & \mathbf{D} \\ \mathbf{D}^\top & \mathbf{E} \end{pmatrix}, \quad (6)$$

where

$$\mathbf{D} = \begin{pmatrix} j \cos \alpha_1 & & & 0 & \dots \\ & j \cos \alpha_2 & & 0 & \dots \\ & & \ddots & \vdots & \\ & & & j \cos \alpha_N & 0 & \dots \end{pmatrix}, \quad (7)$$

$$\mathbf{E} = \begin{pmatrix} \Sigma & \\ & \mathbf{Q} \end{pmatrix}, \quad (8)$$

and  $\mathbf{D}^\top$  is the transpose of  $\mathbf{D}$ .

Now, we can see that the product of  $\mathbf{U}'$ ,  $\Sigma'$ , and  $\mathbf{V}'$  is

$$\begin{aligned} \mathbf{S}' &= \mathbf{U}' \Sigma' \mathbf{V}' = \begin{pmatrix} \mathbf{U} & \\ & \mathbf{O} \end{pmatrix} \begin{pmatrix} \Sigma & \mathbf{D} \\ \mathbf{D}^\top & \mathbf{E} \end{pmatrix} \begin{pmatrix} \mathbf{V} \\ \\ & \mathbf{P} \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{S} & \mathbf{UDP} \\ \mathbf{OD}^\top \mathbf{V} & \mathbf{OEP} \end{pmatrix}, \end{aligned} \quad (9)$$

where  $\mathbf{S}$  is included as a submatrix. Since  $\mathbf{U}'$ ,  $\Sigma'$ , and  $\mathbf{V}'$  are all unitary matrices,  $\mathbf{S}'$  must also be a unitary matrix, which by definition can be realized by an  $N' \times N'$  universal multiport interferometer. Because  $\mathbf{S}$  is a submatrix of  $\mathbf{S}'$ , we reach the conclusion that using  $N$  input-output ports of a  $N' \times N'$  universal multiport interferometer, an arbitrary  $\mathbf{S}$  in the form of Eq. (1) can be realized provided that  $N' \geq 2N$ . On the contrary, if  $N' < 2N$ , a unitary  $\Sigma'$  can no longer be constructed and therefore  $\mathbf{S}'$  cannot be realized. It also can be verified that selecting  $N$  ports from  $N'$  ports can be arbitrary and does not have to be consecutive as in the above case, since permuting rows and columns in  $\mathbf{S}'$  does not affect its unitarity. Meanwhile, we note that other

submatrices in  $\mathbf{S}'$  can be arbitrary, since  $\mathbf{O}$ ,  $\mathbf{P}$ , and  $\mathbf{Q}$  are all arbitrary. This suggests that there are redundant degrees of freedom in  $\mathbf{S}'$  and that we may need fewer phase-shifter stages to realize  $\mathbf{S}$ . Considering the necessary degrees of freedom ( $2N^2 + N$ ), the stage number  $M$  should be no less than  $N + 2$ .

### III. NUMERICAL ANALYSIS

We use two different types of couplers—MMI couplers and MDCs, respectively—to investigate the necessary phase-shifter stages in this scheme. For an ideal MMI coupler in which insertion loss and power imbalance are ignored, its transfer matrix is given in Ref. [36]. For an MDC, we use the coupled mode theory to derive its transfer matrix [37]. The details are provided in the Appendix. In real devices, the MMI coupler should be engineered to have a low insertion loss, which typically involves the use of waveguide tapers; the MDC should be designed to have a large coupling entropy [32].

#### A. Dense matrices

We consider the cases of various  $N$ ,  $N'$ , and  $M$ . For each  $N$ , we generate 100 random  $\mathbf{S}$  as target matrices by multiplying randomly generated  $\mathbf{U}$ ,  $\Sigma$ , and  $\mathbf{V}$ , where  $\mathbf{U}$  and  $\mathbf{V}$  are Haar-random unitary matrices and the diagonal elements of  $\Sigma$  are randomly sampled from the uniform distribution [0, 1]. In this way, all generated  $\mathbf{S}$  are dense matrices with nonzero elements. For each target matrix, we construct several LOPs with different  $N'$  and  $M$  and then optimize all the phase shifts to obtain the target matrix, using a covariance-matrix-adaptation evolution-strategy (CMA-ES) algorithm [38]. An open-source PYTHON package developed for the CMA-ES optimization is used [39]. For simplicity, the  $N$  ports are chosen to be the middle ports of the  $N'$  ports. After optimization, the difference between the target matrix  $\mathbf{S}$  and the obtained matrix  $\mathbf{S}^{\text{ob}}$  is evaluated using the normalized squared error (NSE):

$$\text{NSE} = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N \left| S_{ij} - S_{ij}^{\text{ob}} \right|^2. \quad (10)$$

The NSE is also used as the cost function during optimizations and the stopping criterion is set as either  $\text{NSE} < 10^{-12}$  or upon reaching a specific iteration number. For all optimizations, we set the initial phase shifts to  $\pi$  and the update step size (parameter “sigma” in the algorithm) to 2. In most cases, the algorithm converges successfully and returns the desired results.

Figures 3(a) and 3(b) show the average NSEs when the proposed LOPs are used to implement 100 random dense  $4 \times 4$  matrices ( $N = 4$ ), assuming the use of MMI couplers and MDCs, respectively. The error bar represents the range between the maximum and minimum value among the 100 cases. The dashed line indicates the NSE of  $10^{-12}$ . We can see that very similar results are obtained for the two different couplers. As predicted in Sec. II, an arbitrary matrix cannot be realized if  $N' < 2N$ . Therefore, for  $N' = 6$  and 7, the average NSE does not decrease significantly even when  $M$  is increased. On the contrary, for  $N' = 8$ , sufficiently small NSEs ( $< 10^{-12}$ ) are obtained for all the 100 cases when  $M \geq N + 2$ , indicating that all the desired matrices are almost perfectly realized. Figure 4 further shows the cases of various  $N$ . The error band represents the range between the maximum and minimum value among the 100 cases. For all desired matrices, sufficiently small NSEs are obtained for both types when  $N' = 2N$  and  $M = N + 2$ . For  $M = N + 1$ , although the NSE decreases slightly with increasing  $N$ , it cannot be sufficiently suppressed compared to the case of  $M = N + 2$ . There is no notable difference between the LOP using MMI couplers and MDCs in terms of NSE. Therefore, while a rigorous

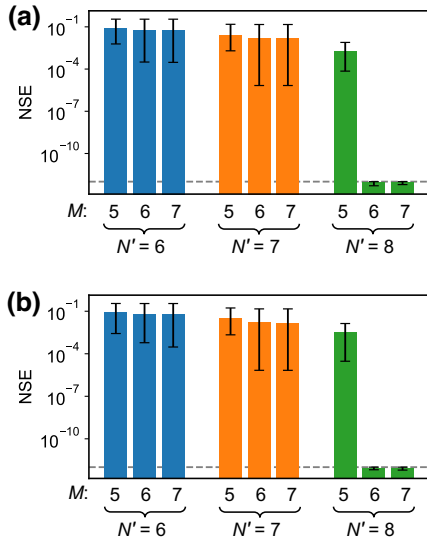


FIG. 3. Average NSEs when the proposed LOPs are used to implement 100 random dense matrices with nonzero elements. The error bar represents the range between the maximum and minimum values among the 100 cases. The dashed line indicates the NSE of  $10^{-12}$ . (a) LOPs using MMI couplers ( $N = 4$ ). (b) LOPs using MDCs ( $N = 4$ ).

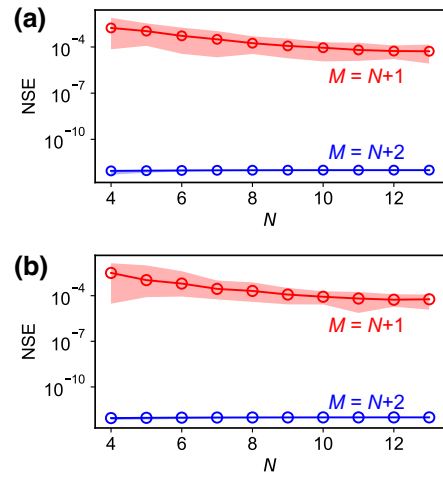


FIG. 4. Average NSEs for random dense matrices with nonzero elements. The error band represents the range between the maximum and minimum values among the 100 cases. (a) LOPs using MMI couplers ( $N' = 2N$ ). (b) LOPs using MDCs ( $N' = 2N$ ).

proof is still lacking, the numerical analysis shows that  $N + 2$  stages of phase shifters are sufficient in this scheme for a large number of random dense matrices with nonzero elements.

## B. Sparse matrices

We then investigate the necessary stage number in this scheme for sparse matrices. Random dense matrices with nonzero elements are first generated following the same procedures as described in Sec. III A. Subsequently, sparse matrices with only one nonzero element are created by randomly setting elements to 0 in the dense matrices. We construct various LOPs to implement these sparse matrices. Figure 5 shows the average NSEs after optimizing the phase shifts. Each point again represents the average value of 100 random cases. It can be seen that all these sparse matrices are realized sufficiently well using  $N + 3$  stages. Although an extra stage is needed compared with the case of dense matrices, this increase is still acceptable from a practical standpoint. Additionally, we can note the differences in Figs. 5(a) and 5(b) for  $N = 4$  and 5, which are attributed to the differences in the coupling entropy between the two couplers.

## C. Hardware-induced computational errors

We further compare the hardware-induced computational error in this scheme and that in the conventional MZI-based scheme shown in Fig. 1(c). Computational errors can arise from many factors, such as fabrication errors and finite-phase control resolutions. Since it is beyond the scope of this paper to investigate the effects of different error sources, here we focus solely on the

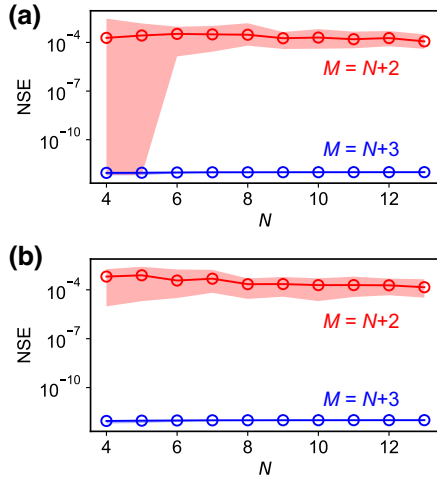


FIG. 5. Average NSEs for random sparse matrices with one nonzero element. The error band represents the range between the maximum and minimum values among the 100 cases. (a) LOPs using MMI couplers ( $N' = 2N$ ). (b) LOPs using MDCs ( $N' = 2N$ ).

phase-quantization error, which is caused by the finite-phase control resolution in practical devices. In the results shown in Figs. 3–5, all phase shifts have been assumed to have sufficiently fine resolutions. Now, we assume that the resolution of phase control is 10 bits and calculate the induced errors in both schemes. We use the same 100 random matrices in Fig. 4 as the targets and calculate the ideal phase configurations for each scheme, respectively. The phase configurations for the MZI-based scheme are derived using the decomposition algorithm proposed in Ref. [14]. Then, these ideal phase configurations are replaced by 10-bit approximations and the new matrices are calculated. Figure 6 shows the average NSEs between the target and obtained matrices. For this scheme, the results of using MMI couplers and using MDCs are almost the same. The MZI-based scheme has larger average NSEs

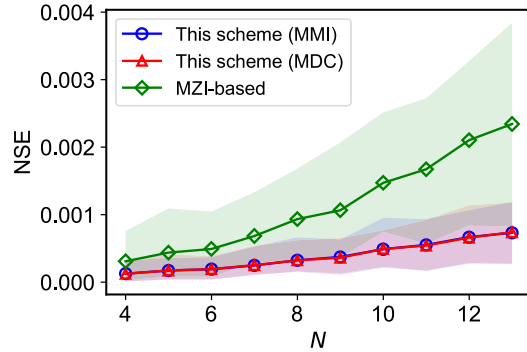


FIG. 6. Average NSEs of this scheme ( $N' = 2N, M = N + 2$ ) and the MZI-based scheme when all phase shifts are assumed to have 10-bit resolutions.

than this scheme and the difference increases with increasing  $N$ . This arises from the fact that an error occurring at an earlier stage affects all the following stages. The error accumulates along with the propagation of light in the circuit and, therefore, a deeper LOP has a larger error than a shallower one.

#### IV. DISCUSSION

It is natural to wonder if the proposed method can be applied to conventional schemes based on MZI meshes. Here, we show that although an arbitrary  $S$  can be realized by increasing the port number to  $2N$ , the number of phase-shifter stages can only be reduced by 1. We use  $N = 3$  and  $N' = 6$  as the example, as illustrated in Fig. 7. Without loss of generality, we assume that ports 1–3 are used as the input-output ports. The transfer matrix of this  $6 \times 6$  MZI mesh  $\mathbf{T}_{\text{mesh}}$  can be calculated by sequentially multiplying the matrices of all components:

$$\begin{aligned} \mathbf{T}_{\text{mesh}} = & \Phi_4^8 \Phi_2^8 \Phi_6^7 \mathbf{T}_{4,5}^7 \mathbf{T}_{2,3}^7 \Phi_1^7 \mathbf{T}_{5,6}^6 \mathbf{T}_{3,4}^6 \mathbf{T}_{1,2}^6 \Phi_6^5 \mathbf{T}_{4,5}^5 \mathbf{T}_{2,3}^5 \Phi_1^5 \mathbf{T}_{5,6}^4 \\ & \times \mathbf{T}_{3,4}^4 \mathbf{T}_{1,2}^4 \Phi_6^3 \mathbf{T}_{4,5}^3 \mathbf{T}_{2,3}^3 \Phi_1^3 \mathbf{T}_{5,6}^2 \mathbf{T}_{3,4}^2 \mathbf{T}_{1,2}^2 \Phi_6^1 \Phi_4^1 \Phi_2^1, \end{aligned} \quad (11)$$

where  $\mathbf{T}_{i,i+1}^k$  corresponds to the MZI between the  $i$ th and  $(i+1)$ th ports on the  $k$ th stage [14],  $\Phi_i^k$  corresponds to the phase shifter outside MZIs on the  $i$ th port and  $k$ th stage and takes the form of

$$\Phi_i^k = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & e^{j \cos \phi_i^k} & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}, \quad (12)$$

where  $\phi_i^k$  is the phase shift. For an arbitrary unitary matrix, each MZI matrix and phase shifter is determined recursively using the decomposition algorithm proposed by Bell *et al.* [14]. Since this MZI mesh can realize arbitrary  $6 \times 6$  unitary matrices, according to the conclusion reached in Sec. II, an arbitrary  $3 \times 3$  matrix can also be realized if it can be written in the form of Eq. (1). Meanwhile, it is known that multiplying a matrix with  $\mathbf{T}_{i,i+1}^j$  only affects the associated elements in the  $i$ th and  $(i+1)$ th rows and columns. Therefore,  $\mathbf{T}_{4,5}$ ,  $\mathbf{T}_{5,6}$  and  $\Phi_6$  on all stages are redundant, as they do not affect the matrix elements of our interest. It then becomes obvious that the required stages of phase shifters are  $2N + 2$ . The distinct difference in the required number of phase-shifter stages between the MZI mesh and our proposed structure is attributed to the  $N' \times N'$  couplers. The MMI coupler splits the input light from any input waveguide equally into all output waveguides. The MDC does not split the input light equally but

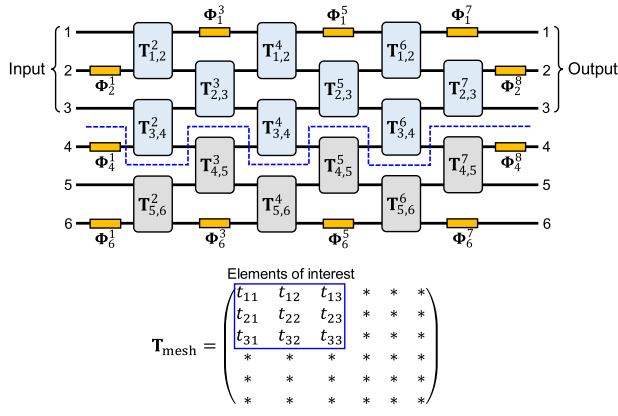


FIG. 7. An example of using a  $6 \times 6$  MZI-based universal multipoint interferometer to implement  $3 \times 3$  matrices. All MZIs and phase shifters below the dashed line are redundant, since they have no effect on the matrix elements of interest.  $2N + 2$  phase-shifter stages are required ( $N = 3$  in this example).

scrambles it in a way such that the light is not localized in only a few waveguides [31]. Therefore, the phase shifters following an  $N' \times N'$  coupler change the matrix elements globally, since each phase shifter affects the light from all ports in the previous stage. By contrast, a phase shifter in the MZI mesh changes the matrix elements locally, since it only affects the light from two ports in the previous stage. This leads to the result that a large number of phase shifters in Fig. 6 are redundant.

Although the structure of the universal multipoint interferometer is not new, the method of reducing the circuit depth by encoding a complex matrix as a block in a unitary matrix has not been proposed in previous works. While the number of ports has doubled in our scheme, the required number of phase-shifter stages has been reduced from  $2N + 3$  in the previous scheme to  $N + 2$  or  $N + 3$ , which brings significant advantages in terms of insertion loss and footprint, especially when  $N$  is large. In addition, the doubling of ports may not be a serious issue, since there is no thermal crosstalk between the EO phase shifters, allowing them to be placed very close to each other. For unused output ports, optical power monitors can be integrated to provide real-time monitoring signals.

## V. CONCLUSIONS

We have proposed a structure for programmable LOPs based on the concept of MPLC. We have shown that using  $N$  input-output ports out of an  $N' \times N'$  universal multipoint interferometer ( $N' \geq 2N$ ), an arbitrary  $N \times N$  matrix can be obtained as long as it can be realized by conventional schemes based on MZI meshes. While a rigorous proof is yet to be provided, our numerical analysis suggests that the number of phase-shifter stages in the proposed structure can be significantly reduced to  $N + 2$  for a large number of dense matrices and  $N + 3$  for a large number of

sparse matrices. We have further demonstrated that the same level of reduction cannot be achieved for the conventional scheme. This work contributes to the realization of compact, low-loss, and energy-efficient programmable LOPs.

## ACKNOWLEDGMENTS

This work was partly supported by Japan Society for the Promotion of Science (JSPS) KAKENHI Grants No. JP22K14298 and No. JP21K18168. R. Tang acknowledges fruitful discussions with Ryan Hamerly and Bryn Bell.

## APPENDIX: TRANSFER MATRIX OF THE MULTIPOINT DIRECTIONAL COUPLER (MDC)

We consider an MDC consisting of  $N$  parallel straight waveguides with the same width and spacing.  $N$  in this appendix should not be confused with the  $N$  in the main text. We assume that the perturbation is weak so that the propagation constant in each waveguide is approximately the same as that in a single waveguide and only coupling from nearest waveguides needs to be considered. Under these assumptions, the amplitudes of light in the MDC can be described by the coupled equations [37]:

$$\frac{da_1}{dz} = -j\beta a_1 - j\kappa_{1,2}a_2, \quad (\text{A1a})$$

$$\frac{da_i}{dz} = -j\kappa_{i,i-1}a_{i-1} - j\beta a_i - j\kappa_{i,i+1}a_{i+1}, \quad (\text{A1b})$$

$$\frac{da_N}{dz} = -j\kappa_{N,N-1}a_{N-1} - j\beta a_N, \quad (\text{A1c})$$

where  $z$  is the distance along the propagation direction,  $a_i$  is the amplitude in the  $i$ th waveguide ( $1 < i < N$ ),  $\beta$  is the propagation constant, and  $\kappa_{i,i-1}$  is the coupling coefficient from the  $(i-1)$ th to the  $i$ th waveguide. Since all the waveguides are assumed to have the same width and spacing, the coupling coefficient between each waveguide pair is the same and will hereafter be denoted as  $\kappa$ . Equations (A1a)–(A1c) can be written in matrix form as

$$\frac{d\mathbf{A}}{dz} = -j\mathbf{H}\mathbf{A}, \quad (\text{A2})$$

where

$$\mathbf{A} = (a_1 \ a_2 \ \dots \ a_N)^\top, \quad (\text{A3})$$

$$\mathbf{H} = \begin{pmatrix} \beta & \kappa & & & & \\ \kappa & \beta & \kappa & & & \\ & & & \ddots & & \\ & & & & \kappa & \beta & \kappa \\ & & & & & \kappa & \beta \end{pmatrix}. \quad (\text{A4})$$

Since  $\mathbf{H}$  is a real Hermitian matrix, we can find a unitary matrix  $\mathbf{B}$  to diagonalize it into a diagonal matrix  $\mathbf{C}$ :

$$\mathbf{C} = \mathbf{B}^{-1}\mathbf{H}\mathbf{B}. \quad (\text{A5})$$

Then, by letting  $\mathbf{W} = \mathbf{B}^{-1}\mathbf{A}$  and substituting it into Eq. (A2), we obtain

$$\frac{d\mathbf{W}}{dz} = -j\mathbf{C}\mathbf{W}. \quad (\text{A6})$$

It is easy to see that the solution of Eq. (A6) is

$$\mathbf{W} = \begin{pmatrix} e^{-jC_{11}z} & e^{-jC_{22}z} & \dots & e^{-jC_{NN}z} \end{pmatrix}^T, \quad (\text{A7})$$

where  $C_{ii}$  is the  $i$ th diagonal element of  $\mathbf{C}$ . It follows that

$$\mathbf{A}(z) = \mathbf{B} \begin{pmatrix} e^{-jC_{11}z} & e^{-jC_{22}z} & \dots & e^{-jC_{NN}z} \end{pmatrix}^T, \quad (\text{A8a})$$

$$\mathbf{A}(0) = \mathbf{B} \begin{pmatrix} 1 & 1 & \dots & 1 \end{pmatrix}^T. \quad (\text{A8b})$$

Equations (A8) can be rewritten as

$$\mathbf{A}(z) = \mathbf{B} \begin{pmatrix} e^{-jC_{11}z} & & & \\ & e^{-jC_{22}z} & & \\ & & \ddots & \\ & & & e^{-jC_{NN}z} \end{pmatrix} \mathbf{B}^{-1}\mathbf{A}(0), \quad (\text{A9})$$

Therefore, for an MDC with a length of  $L$ , the transfer matrix is given by

$$\mathbf{T} = \mathbf{B} \begin{pmatrix} e^{-jC_{11}L} & & & \\ & e^{-jC_{22}L} & & \\ & & \ddots & \\ & & & e^{-jC_{NN}L} \end{pmatrix} \mathbf{B}^{-1}. \quad (\text{A10})$$

In this work, we assume the use of silicon waveguides with core dimensions of  $500 \times 220 \text{ nm}^2$ . From numerical simulation,  $\beta$  is found to be  $9.91 \text{ rad}/\mu\text{m}$  for the transverse-electric (TE) mode at  $1550\text{-nm}$  wavelength. For an  $N$ -port MDC, we assume  $\kappa = 0.05$  and select an  $L$  so that small deviations in  $\kappa$  and  $L$  do not affect the overall performance of the LOP. Specifically, for  $N = 8, 10, 12, 14, 16, 18, 20, 22, 24,$  and  $26$ , we choose  $L$  to be  $50, 60, 75, 85, 100, 120, 130, 140, 150,$  and  $160 \mu\text{m}$ , respectively. Regardless of the value of  $N$ , two parameters—the waveguide spacing and the length—can be swept when designing an MDC. The design goal is to thoroughly scramble the light propagating through the MDC, which can be quantitatively described by the coupling entropy [32]. There exists a wide design region where the coupling entropy is

large and does not change abruptly. The design parameters should be chosen from this region so that the device is robust to fabrication errors.

The above method yields unitary transfer matrices that simplify our analysis in this work. However, if the distance between adjacent waveguides is small enough that the assumption of weak perturbation no longer holds, numerical approaches should be used to obtain more precise results [40].

- 
- [1] N. C. Harris, J. Carolan, D. Bunandar, M. Prabhu, M. Hochberg, T. Baehr-Jones, M. L. Fanto, A. M. Smith, C. C. Tison, P. M. Alsing, and D. Englund, Linear programmable nanophotonic processors, *Optica* **5**, 1623 (2018).
  - [2] W. Bogaerts, D. Pérez, J. Capmany, D. A. Miller, J. Poon, D. Englund, F. Morichetti, and A. Melloni, Programmable photonic circuits, *Nature* **586**, 207 (2020).
  - [3] Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, and M. Soljačić, Deep learning with coherent nanophotonic circuits, *Nat. Photon.* **11**, 441 (2017).
  - [4] D. Pérez, I. Gasulla, L. Crudgington, D. J. Thomson, A. Z. Khokhar, K. Li, W. Cao, G. Z. Mashanovich, and J. Capmany, Multipurpose silicon photonics signal processor core, *Nat. Commun.* **8**, 636 (2017).
  - [5] M. Dong, G. Clark, A. J. Leenheer, M. Zimmermann, D. Dominguez, A. J. Menssen, D. Heim, G. Gilbert, D. Englund, and M. Eichenfield, High-speed programmable photonic circuits in a cryogenically compatible, visible–near-infrared 200 mm CMOS architecture, *Nat. Photon.* **16**, 59 (2022).
  - [6] M. A. Nahmias, T. F. De Lima, A. N. Tait, H.-T. Peng, B. J. Shastri, and P. R. Prucnal, Photonic multiply-accumulate operations for neural networks, *IEEE J. Sel. Top. Quantum Electron.* **26**, 1 (2019).
  - [7] H. Zhou, J. Dong, J. Cheng, W. Dong, C. Huang, Y. Shen, Q. Zhang, M. Gu, C. Qian, H. Chen, Z. Ruan, and X. Zhang, Photonic matrix multiplication lights up photonic accelerator and beyond, *Light Sci. Appl.* **11**, 30 (2022).
  - [8] R. Tang, M. Okano, K. Toprasertpong, S. Takagi, D. Englund, and M. Takenaka, Two-layer integrated photonic architectures with multipoint photodetectors for high-fidelity and energy-efficient matrix multiplications, *Opt. Express* **30**, 33940 (2022).
  - [9] A. N. Tait, Quantifying power in silicon photonic neural networks, *Phys. Rev. Appl.* **17**, 054029 (2022).
  - [10] D. A. Miller, Self-configuring universal linear optical component, *Photon. Res.* **1**, 1 (2013).
  - [11] W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley, Optimal design for universal multipoint interferometers, *Optica* **3**, 1460 (2016).
  - [12] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, Experimental realization of any discrete unitary operator, *Phys. Rev. Lett.* **73**, 58 (1994).
  - [13] 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).
- [14] B. A. Bell and I. A. Walmsley, Further compactifying linear optical unitaries, *APL Photonics* **6**, 070804 (2021).
- [15] S. Pai, B. Bartlett, O. Solgaard, and D. A. B. Miller, Matrix optimization on universal unitary photonic devices, *Phys. Rev. Appl.* **11**, 064044 (2019).
- [16] 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).
- [17] N. C. Harris, Y. Ma, J. Mower, T. Baehr-Jones, D. Englund, M. Hochberg, and C. Galland, Efficient, compact and low loss thermo-optic phase shifter in silicon, *Opt. Express* **22**, 10487 (2014).
- [18] H. Qiu, Y. Liu, C. Luan, D. Kong, X. Guan, Y. Ding, and H. Hu, Energy-efficient thermo-optic silicon phase shifter with well-balanced overall performance, *Opt. Lett.* **45**, 4806 (2020).
- [19] T. Kita and M. Mendez-Astudillo, Ultrafast silicon MZI optical switch with periodic electrodes and integrated heat sink, *J. Light. Technol.* **39**, 5054 (2021).
- [20] P. Dong, S. Liao, H. Liang, R. Shafiha, D. Feng, G. Li, X. Zheng, A. V. Krishnamoorthy, and M. Asghari, Submilliwatt, ultrafast and broadband electro-optic silicon switches, *Opt. Express* **18**, 25225 (2010).
- [21] J.-H. Han, F. Boeuf, J. Fujikata, S. Takahashi, S. Takagi, and M. Takenaka, Efficient low-loss InGaAsP/Si hybrid MOS optical modulator, *Nat. Photon.* **11**, 486 (2017).
- [22] H. Tang, Q. Li, C. P. Ho, J. Fujikata, M. Noguchi, S. Takahashi, K. Toprasertpong, S. Takagi, and M. Takenaka, Modulation bandwidth improvement of III-V/Si hybrid MOS optical modulator by reducing parasitic capacitance, *Opt. Express* **30**, 22848 (2022).
- [23] Y. Zheng, H. Zhong, H. Zhang, L. Song, J. Liu, Y. Liang, Z. Liu, J. Chen, J. Zhou, Z. Fang, M. Wang, L. Li, R. Wu, and Y. Cheng, Electro-optically programmable photonic circuits enabled by wafer-scale integration on thin-film lithium niobate, *Phys. Rev. Research* **5**, 033206 (2023).
- [24] 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).
- [25] N. K. Fontaine, R. Ryf, H. Chen, D. T. Neilson, K. Kim, and J. Carpenter, Laguerre-Gaussian mode sorter, *Nat. Commun.* **10**, 1865 (2019).
- [26] P. Boucher, A. Goetschy, G. Sorelli, M. Walschaers, and N. Treps, Full characterization of the transmission properties of a multi-plane light converter, *Phys. Rev. Res.* **3**, 023226 (2021).
- [27] R. Tang, T. Tanemura, and Y. Nakano, Integrated reconfigurable unitary optical mode converter using MMI couplers, *IEEE Photon. Technol. Lett.* **29**, 971 (2017).
- [28] 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).
- [29] R. Tanomura, R. Tang, T. Sukanuma, K. Okawa, E. Kato, T. Tanemura, and Y. Nakano, Monolithic InP optical unitary converter based on multi-plane light conversion, *Opt. Express* **28**, 25392 (2020).
- [30] R. Tang, T. Tanemura, and Y. Nakano, in *2017 Opto-Electronics and Communications Conference (OECC) and Photonics Global Conference (PGC)* (IEEE, Singapore, 2017), p. 1.
- [31] R. Tanomura, R. Tang, S. Ghosh, T. Tanemura, and Y. Nakano, Robust integrated optical unitary converter using multiport directional couplers, *J. Light. Technol.* **38**, 60 (2020).
- [32] R. Tang, R. Tanomura, T. Tanemura, and Y. Nakano, Ten-port unitary optical processor on a silicon photonic chip, *ACS Photonics* **8**, 2074 (2021).
- [33] S. Kuzmin, I. Dyakonov, and S. Kulik, Architecture agnostic algorithm for reconfigurable optical interferometer programming, *Opt. Express* **29**, 38429 (2021).
- [34] R. Tanomura, R. Tang, T. Umezaki, G. Soma, T. Tanemura, and Y. Nakano, Scalable and robust photonic integrated unitary converter based on multiplane light conversion, *Phys. Rev. Appl.* **17**, 024071 (2022).
- [35] 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).
- [36] M. Bachmann, P. A. Besse, and H. Melchior, General self-imaging properties in  $N \times N$  multimode interference couplers including phase relations, *Appl. Opt.* **33**, 3905 (1994).
- [37] W.-P. Huang, Coupled-mode theory for optical waveguides: An overview, *J. Opt. Soc. Am. A* **11**, 963 (1994).
- [38] N. Hansen, The CMA evolution strategy: A tutorial, [arXiv:1604.00772](https://arxiv.org/abs/1604.00772) [cs.LG] (2023).
- [39] <https://pypi.org/project/cmaes/>.
- [40] M. L. Cooper and S. Mookherjea, Numerically-assisted coupled-mode theory for silicon waveguide couplers and arrayed waveguides, *Opt. Express* **17**, 1583 (2009).