



# Learning arbitrary complex matrices by interlacing amplitude and phase masks with fixed unitary operations

Matthew Markowitz

*Department of Physics, [Queens College of the City University of New York](#), Queens, New York 11367, USA  
and Physics Program, [The Graduate Center](#), City University of New York, New York, New York 10016, USA*

Kevin Zelaya <sup>\*</sup>

*Department of Physics, [Queens College of the City University of New York](#), Queens, New York 11367, USA*

Mohammad-Ali Miri <sup>†</sup>

*Department of Physics, [Queens College of the City University of New York](#), Queens, New York 11367, USA  
and Physics Program, [The Graduate Center](#), City University of New York, New York, New York 10016, USA*



(Received 27 May 2024; accepted 16 August 2024; published 3 September 2024)

Programmable photonic integrated circuits are an emerging technology that amalgamates photonics and electronics, paving the way for light-based information processing at high speeds and low power consumption. Considering its wide range of applications as one of the most fundamental mathematical operations, there has been considerable interest in developing reliably implementable programmable circuit architectures that perform matrix-vector multiplication. Recently, it was shown that discrete unitary operations can be parameterized by interlacing fixed operators with diagonal phase parameters realized with phase shifter arrays. We show that these decompositions are a special case of a broader class of factorizations that enable parametrization of arbitrary complex matrices. The proposed representation of an  $N \times N$  matrix is given by  $N + 1$  amplitude-and-phase-modulation layers interlaced with a fixed unitary layer that can be implemented, for example, via a coupled waveguide array. Thus, we introduce an architecture for physically implementing discrete linear operations, enabling the development of novel families of programmable photonic circuits for on-chip analog information processing.

DOI: [10.1103/PhysRevA.110.033501](https://doi.org/10.1103/PhysRevA.110.033501)

## I. INTRODUCTION

The prospect of realizing reconfigurable photonic circuits that perform arbitrary discrete linear operations on light has gained significant interest in the past decade [1,2]. This interest is driven by the wide range of tasks that can be characterized by matrix operations, subsequently leading to the development of a chip-scale platform to deploy general-purpose light-based applications in classical and quantum information processing [3–5]. An efficient programmable photonic matrix-vector multiplier opens numerous opportunities such as enabling on-chip photonic neural networks [6–8], novel photonic interconnects [9,10], and multistate quantum processors [11–13]. In the context of integrated photonics, previous efforts have focused on realizing circuits that perform linear unitary operations, which enforce power conservation. This field has been driven by the interesting properties of unitary groups, such as the rigorous mathematical parametrization and factorization of unitary matrices in terms of simpler unitary elements, and decompositions of general matrices. Notably, it has been shown that all unitaries can be realized with meshes of Mach-Zehnder interferometers incorporating

phase shifters in different circuit configurations [14–20]. Architectures based on mode divisor multiplexing have also proven to be another valuable resource [21,22]. On the other hand, an alternative factorization of unitaries consisting of interlaced diagonal phase modulations with a fixed intervening unitary operator was recently suggested [23–30]. In particular, we showed that arbitrary unitary operations can be represented by interlacing nonuniform waveguide couplers emulating the discrete fractional Fourier transform (DFrFT) with  $N + 1$  layers of phase shifters [28]. This architecture is robust to imperfections and defects [29], and its fixed intervening operator can be extended to a wide range of operations that subsequently allow for different physical realizations [30].

Despite numerous efforts to realize unitary operators, to the best of our knowledge, no prior work has discussed direct photonic implementation of general nonunitary matrices. A standard approach that has been utilized is the singular-value decomposition (SVD) [6], which enables the creation of general matrices by sandwiching a positive semidefinite diagonal matrix between two unitaries. This method, however, results in devices with a large optical path length and requires in its construction programmable unitaries as intermediary building blocks. On the other hand, there is a lack of direct compact factorization of arbitrary complex nonunitary matrices. Most relevant prior works showed that every complex  $2^n \times 2^n$  matrix can be factored as the finite product of circulant and diagonal

<sup>\*</sup>Contact author: [kevin.zelaya@cinvestav.mx](mailto:kevin.zelaya@cinvestav.mx)

<sup>†</sup>Contact author: [mmiri@qc.cuny.edu](mailto:mmiri@qc.cuny.edu)

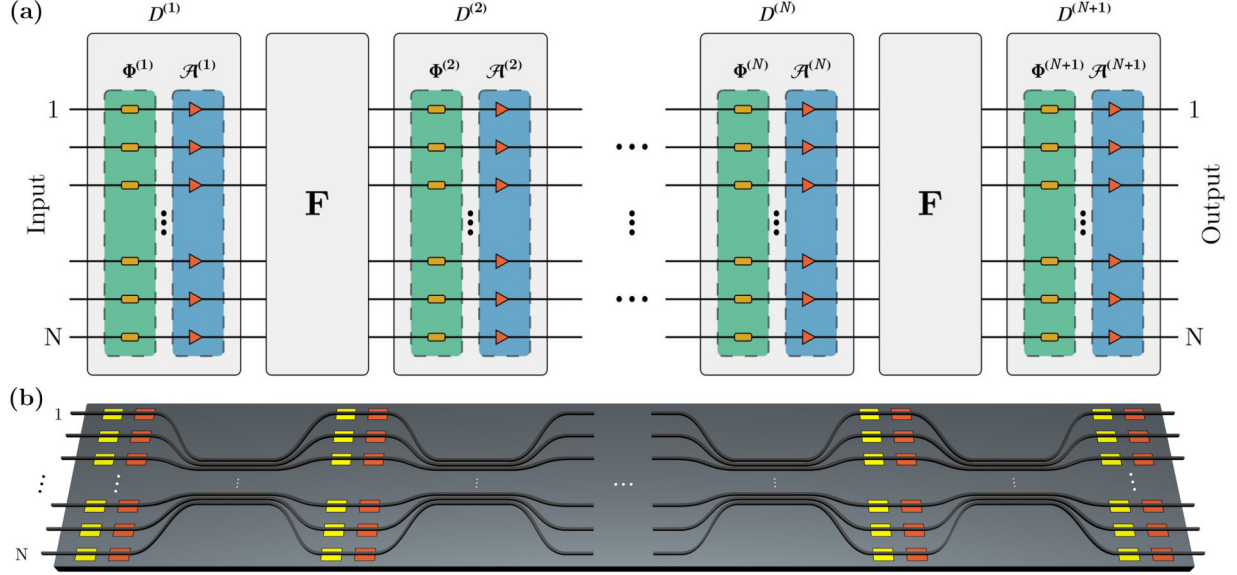


FIG. 1. (a) A block diagram sketching the proposed factorization in (1), with  $D^{(m)}$  being the complex diagonal matrices and  $F$  being the DFrFT matrix acting as the unitary interlacing matrix. Furthermore,  $D^{(m)}$  is decomposed as the product of a layer of phase shifters  $\Phi^{(m)}$  (orange rectangles) and a layer of amplitude modulators  $\mathcal{A}^{(m)}$  (red triangles). (b) A schematic of a photonic-device architecture with waveguide couplers representing the intervening operator  $F$  and with amplitude (red) and phase (yellow) modulators representing the programmable diagonal matrices.

matrices [31,32]. Furthermore, Huhtanen and Perämäki [33] demonstrated that any  $N \times N$  complex-valued matrix can be decomposed as the product of no more than  $2N - 1$  interlaced circulant and diagonal matrices. This result has been particularly useful in deep-learning models for video classification [34], optical networking [35], and frequency encoding in quantum information processing [36].

In this work, we introduce a layer-efficient factorization to represent arbitrary complex-valued matrices. This is achieved after modifying the factorization presented in [33] by limiting the maximum number of interlaced layers and replacing the interlaced discrete-Fourier-transform (DFT) layer with a passive layer performing the DFrFT operation. The latter is inspired by photonic applications, where platforms to deploy on-chip DFrFT based on waveguide arrays are known in the literature [37], as well as other optical elements to represent the factorization discussed here. Numerical experiments based on optimization algorithms reveal that such a truncation leads to numerical error as low as that obtained using the exact factorization. The case of unitary matrices is discussed as a particular limit, where the number of parameters to be optimized reduces to one half of the number required for the general problem. Remarkably, the architecture is shown to be robust against random perturbations on the passive interlacing layer, provided that the latter are within acceptable tolerance errors. Rigorous electromagnetic simulations further corroborate the realization of the proposed programmable photonic device.

## II. RESULTS

### A. Model

It is known that an arbitrary complex-valued  $N \times N$  matrix  $A$  can be parameterized in terms of at most

$2N - 1$  interlaced circulant and diagonal matrices [33]. In turn, a circulant matrix  $C^{(m)}$  can be diagonalized through the DFT and inverse-discrete-Fourier-transform (IDFT) matrices,  $\mathbb{F}$  and  $\mathbb{F}^{-1}$ , according to the relation  $C^{(m)} = \mathbb{F} D^{(m)} \mathbb{F}^{-1}$  [38]. Therefore, an arbitrary complex matrix  $A$  can be parameterized through complex diagonal matrices interlacing with the DFT matrix and its inverse as  $A = D^{(2N-1)} \mathbb{F} D^{(2N-2)} \mathbb{F}^{-1} D^{(2N-3)} \dots D^{(3)} \mathbb{F} D^{(2)} \mathbb{F}^{-1} D^{(1)}$ . Since an  $N \times N$  diagonal matrix contains  $N$  complex elements, the latter factorization involves  $4N^2 - 2N$  real parameters in total. This relation defines a vastly overparameterized problem with redundant factoring layers given that  $2N^2$  real parameters are required to characterize any  $N \times N$  complex-valued matrix.

Here, we investigate the modified factorization

$$A = D^{(M)} F D^{(M-1)} F D^{(M-2)} \dots D^{(3)} F D^{(2)} F D^{(1)}, \quad (1)$$

where  $F$  is a unitary matrix to be defined and  $D^{(m)}$  are diagonal matrices with components  $D_{p,q}^{(m)} = d_p^{(m)} e^{i\phi_p^{(m)}} \delta_{p,q}$  for  $m \in \{1, \dots, M\}$ . In the latter,  $\phi_p^{(m)} \in (-\pi, \pi]$  and  $d_p^{(m)} \geq 0$  represent the  $p$ th phase and amplitude parameters of the  $m$ th complex-valued diagonal layer, respectively. Figure 1 illustrates the block diagram of the proposed matrix decomposition, along with the corresponding photonic architecture. The unitary matrix  $F$  is realized using an array of coupled waveguides, while each diagonal matrix is composed of two contiguous layers, one containing exclusively  $N$  phase modulators ( $\Phi^{(m)}$ ) and the other containing exclusively  $N$  amplitude modulators ( $\mathcal{A}^{(m)}$ ).

In the proposed factorization, the  $M$  complex-valued diagonal matrices  $D^{(m)}$  require  $2MN$  real parameters. We numerically explore this ansatz and analyze its universality, i.e., the capability of the factorization (1) to reconstruct arbitrary  $N \times N$  complex-valued matrices, for different numbers

of complex diagonals  $M$ . We take the fixed operator  $F$  to be among the class of DFrFTs. In this regard, it is worth noting that several valid definitions for the DFrFT exist in the literature [39,40].

We adapt the definition from Ref. [37], based on the wave evolution of guided modes through a nonuniformly spaced optical lattice, referred to as the  $J_x$  lattice. This construction is useful as its physical realization is straightforward [37,41,42]. In this fashion, the DFrFT matrix of fractional order  $\alpha = \pi/2$  is given as the propagator operator  $F = \exp(iH\pi/2)$  for the normalized length  $\pi/2$ , where  $H$  is the  $J_x$  lattice Hamiltonian with matrix components  $H_{p,q} = \kappa_p \delta_{p+1,q} + \kappa_{p-1} \delta_{p-1,q}$ . The hopping rates  $\kappa_p = \frac{\tilde{\kappa}}{2} \sqrt{(N-p)p}$  for  $p \in \{1, \dots, N-1\}$  are given in terms of an experimental scaling factor  $\tilde{\kappa}$  (see the Appendix for more details). The coupling is physically achieved by using identical single-mode photonic waveguides with nonuniform spacing [37,41]. In the microwave domain, such coupling has been shown to be possible using parallel microstrip lines coupled to interdigital capacitors [42]. It is straightforward to show that  $F$  satisfies all properties of the discrete fractional Fourier transform; e.g.,  $F$  is unitary,  $F^2$  is the parity operator, and  $F^4$  is the identity matrix. Furthermore, the DFrFT defined by  $F$  converges to the DFT in the limit of large  $N$ .

### B. Optimization and universality

To demonstrate the universality of this device, we optimize the respective phases  $\phi_p^{(m)}$  and amplitudes  $d_p^{(m)}$ , where  $p \in \{1, \dots, N\}$  and  $m \in \{1, \dots, M\}$ , for an ensemble of randomly chosen target complex-valued matrices  $A_t$ . The latter are generated using SVD [43] so that each target matrix decomposes as  $A_t = U \Sigma V^\dagger$ , where  $U$  and  $V^\dagger$  are unitary matrices randomly generated in accordance with the Haar measure [44]. In turn, the diagonal matrix  $\Sigma = \text{diag}(\sigma_1, \dots, \sigma_N)$  contains the singular values of  $A_t$ , with  $\sigma_i \geq 0$  randomly taken from the interval  $[0.25, 1]$ . The goodness of approximation of the target matrices is explored against the number  $M$  of complex diagonal matrices, which can, in turn, be understood as being composed of  $M$  phase layers and  $M$  amplitude layers. The loss function, which we call the error norm, is defined as the mean square error

$$L = \frac{1}{N^2} \|A - A_t\|^2, \quad (2)$$

where  $\|A\| = \sqrt{\text{Tr}(A^\dagger A)}$  is the Frobenius norm,  $A_t$  is the complex-valued target matrix under consideration, and  $A$  is the reconstructed matrix using (1). The error norm is a multivariate function defined on the parameter space  $\{\phi_p^{(m)}, d_p^{(m)}\}_{p=1, m=1}^{N, M}$ . In such problems, the space is typically highly nonconvex, and obtained minima, no matter how low the error norm is, might not be unique. To find an acceptable solution, we define the error norm tolerance value  $L_c$  so that a local minimum is accepted as a solution whenever  $L < L_c$ .

The optimization is performed using the gradient-based Levenberg-Marquardt algorithm (LMA), which is well suited to sum-of-squares objective functions and can be used in both under- and overdetermined cases [45]. The universality of the factorization (1) is tested via 100 (500) randomly generated target matrices in the underdetermined (overdetermined) cases

for  $N = 4$  and  $N = 6$  according to the SVD decomposition previously discussed. The parameter space is randomly initialized by assigning the values  $\phi_p^{(m)} \in (-\pi, \pi]$  and  $d_p^{(m)} \in [0, \ell]$  for  $p \in \{1, \dots, N\}$  and  $m \in \{1, \dots, M\}$ , where  $\ell$  sets the upper bound for the search region of the optimized amplitudes  $d_p^{(m)}$ . In the actual implementation, an unconstrained auxiliary variable  $d_p'^{(m)}$  is used to encode  $d_p^{(m)}$  through a sigmoid function with the appropriate bounds. This is found to produce results faster and superior to those obtained using traditional constraints. Up to 100 runs of LMA are performed, which are terminated prematurely when an error norm below the prefixed tolerance  $L_c = 10^{-10}$  is obtained. For practicality, the amplitude parameters are bounded by a small value, chosen to be  $\ell = 1.5$  in most cases, which typically yields satisfactory results. In some instances this value may be increased slightly for increased reliability or optimization performance.

To establish a reference point by which to compare our findings, we also consider the motivating factorization based on interlaced layers of DFT and IDFT when truncated to  $M \leq 2N - 1$  layers. The exact factorization of this architecture for  $M = 2N - 1$  allows for appropriate analysis of the error norms for fewer layers and for comparison to DFrFT with the same total number of layers. This establishes a benchmark scenario in which performance increases can be monitored as a function of  $M$  for different  $N$ . The resulting error norms associated with the optimization process are shown in Figs. 2(a) and 2(b) for  $N = 4$  and  $N = 6$ , respectively. The exact factorization at  $M = 2N - 1$  with interlaced DFT and IDFT reveals error norms on the order of  $10^{-20}$  to  $10^{-30}$ , establishing the expected numerical precision for the optimization process. Importantly, The DFrFT and the DFT-IDFT structures are shown to provide similar results for all possible stage numbers  $M$ , providing evidence for the non-necessity of using alternating operators and their inverse. At the lower end, from  $M = N - 3$  to  $M = N$ , the error norm shows mild improvements at each stage, but with rather high values that never drop below the error threshold  $L_c$ . Between  $M = N$  and  $M = N + 1$  layers, a drastic drop of over 20 orders of magnitude on average occurs, with only a handful of outliers that meet the threshold. Furthermore, the threshold at  $M = N + 1$  is surpassed the vast majority of the time within the first few runs. This transition is analogous to that found in the unitary case with no gain or loss elements, which also occurs at the same stage number. The sharp decay between  $M = N$  and  $M = N + 1$  layers is corroborated for a larger number of ports up to  $N = 14$  in Fig. 2(c). As  $M$  is increased up to  $M = 2N - 1$ , there is no significant change in the mean or variance of the error norm. This verifies that (1) produces results numerically equivalent to the exact case discussed in Ref. [33]. We thus fix

$$M = N + 1$$

as the effective number of layers required for the factorization (1) throughout the paper.

We illustrate the nature of the optimized parameters by considering an example with  $N = 4$  and  $M = N + 1 = 5$ . A set of 500 random matrices is generated with singular values uniformly distributed so that  $\sigma_j \in [0.25, 1]$ , with  $j \in \{1, 2, 3, 4\}$ , for each target matrix, whereas  $U$  and  $V^\dagger$  are

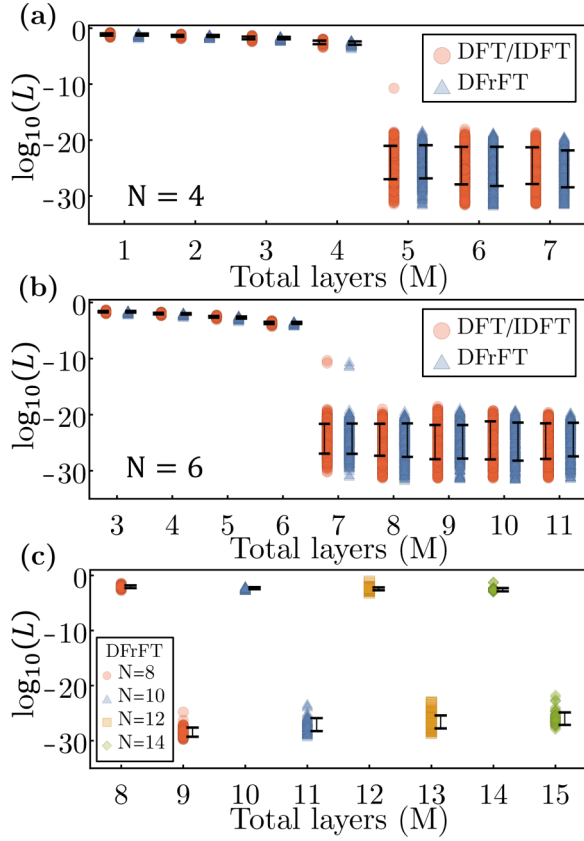


FIG. 2. Error norms  $L$  for (a)  $N = 4$  and (b)  $N = 6$  while varying the total number of layers  $M$ . One hundred targets were used for  $M \leq N$ , and 500 were used for  $M \geq N + 1$ , with LMA stopping when error tolerance  $L_c = 10^{-10}$  was achieved for the latter. Optimizations were performed using two different choices of the mixing operator  $F$  interlaced with diagonal layers. The circles correspond to alternating DFT and IDFT layers, whereas triangles correspond to using DFrFT. The upper bound on the diagonal amplitudes  $d_i^{(m)}$  was set to 1.5. (c) Numerical architecture performance for  $N > 6$  using the DFrFT layers and for  $M = N, N + 1$  to illustrate the norm transition at  $M = N + 1$ .

Haar random matrices as customary. In this form, we generate random complex-valued matrices with specific eigenvalues, and we can study the maximum bound  $\ell$  for the amplitude parameters  $d_p^{(m)}$  to be optimized. To provide more insight into the previous choice  $\ell = 1.5$ , we extend the optimization bound to  $\ell = 3$ . Furthermore, the encoded parameters  $d_p^{(m)}$  are initialized to 1, which is remarkably found to produce results and performance similar to choosing random values. This is done to find solutions with  $d_p^{(m)}$  that deviate less frequently from 1.

Figure 3(a) displays one example from the set of target matrices, with the corresponding optimized magnitudes and phases illustrated in Figs. 3(b) and 3(c), respectively. In turn, Fig. 3(d) shows the occurrence distribution frequency of the optimized amplitude parameters  $d_p^{(m)}$  for all 500 target matrices. The latter reveals that the values of  $d_p^{(m)}$  are more likely to fall within the range (0,2) when the singular values of the target matrix are less than 1. Although there may exist some  $d_p^{(m)} > 2$ , their probability of occurrence is relatively

low. Thus, one can focus mainly on the previous interval during optimization and increase the interval upper bound  $\ell$  for tuning purposes when required. This is particularly useful since the number of parameters increases quadratically with  $N$ , and the optimization becomes computationally demanding. On the other hand, amplitude parameters in the interval (1,2) appear with considerable frequency, implying that active gain elements are required in the physical realization unless the desired matrix is rescaled through a global amplitude factor, as discussed below. The latter behavior is not present for the optimization of the phase parameters  $\phi_p^{(m)}$ , as they are constrained to the compact domain  $(-\pi, \pi]$ . Figure 3(d) shows, indeed, that the optimized phases do not follow any particular pattern during the optimization of random targets, and their occurrence frequency distributes almost uniformly across  $(-\pi, \pi]$ .

### C. Unitary limit

Unitary target matrices  $A_t^u \in U(N)$  are a subset contained in the set of complex-valued  $N \times N$  matrices, and their factorization and photonic implementation have been widely explored in the literature [23–25,28,46]. It is a natural question to explore whether our architecture reproduces the results for unitaries as a special case. To address this, one can factor the diagonal matrices of (1) as  $D^{(m)} = \Phi^{(m)} \mathcal{A}^{(m)}$ , with  $\Phi^{(m)}$  and  $\mathcal{A}^{(m)}$  being diagonal matrices with components  $\Phi_{p,q}^{(m)} = e^{i\phi_p^{(m)}} \delta_{p,q}$  and  $\mathcal{A}_{p,q}^{(m)} = d_p^{(m)} \delta_{p,q}$ , where, in particular,  $\Phi^{(m)}$  is unitary. In general, the product of unitary matrices renders a unitary matrix. Following the factorization (1) and  $D^{(m)} = \Phi^{(m)} \mathcal{A}^{(m)}$ , it is thus straightforward that a unitary target matrix is obtained from the factorization if  $\mathcal{A}^{(m)} = \mathbb{I}$  for all  $m \in \{1, \dots, N\}$ , with  $\mathbb{I}$  being the  $N \times N$  identity matrix. In turn, the singular values of  $A_t^u$  are  $\sigma_p = 1$  for all  $p \in \{1, \dots, N\}$ .

The latter is a sufficient condition, but not necessary, as there might be cases in which the optimized amplitudes  $d_p^{(m)} \neq 1$  render unitary matrices. We thus numerically explore this possibility. The behavior around the vicinity of unitary matrices can be studied by first randomly constructing complex-valued matrices whose singular values  $\sigma_p \leq 1$ . By using the convention  $\sigma_p \geq \sigma_{p+1}$  for all  $p \in \{1, \dots, N-1\}$ , it is then enough to make  $\sigma_{\min} := \min(\{\sigma_p\}_{p=1}^N) = \sigma_N \approx 1$  to ensure that  $A_t$  approximates to a unitary matrix  $A_t^u$ . The unitary limit of  $A_t = U \Sigma V^\dagger$  is independent of the random unitary Haar matrices  $U$  and  $V^\dagger$  and depends exclusively on the choice of the singular values. This is illustrated by considering  $N = 4$  and randomly generating 100 target matrices so that  $\sigma_{\min} = \sigma_4$  uniformly spans the interval [0.2,1] for each target. The remaining singular values  $\sigma_{j=1,2,3}$  are randomly distributed in the interval  $[\sigma_{\min}, 1]$  for each selection of  $\sigma_{\min}$ . The eigenvalues  $\lambda$  for each of the resulting target matrices are shown in Fig. 4(a) and distribute from the inner part of the complex unit circle  $|\lambda| < 1$  ( $\sigma_{\min} < 1$ ) and converge to the unit circle in the unitary-case limit ( $\sigma_{\min} \rightarrow 1$ ). The corresponding optimization is performed for each of the random target matrices, from which the set  $r = \{d_p^{(m)}\}_{p,m=1}^{N,N+1}$  is constructed for each target so that it contains the optimized amplitude parameters exclusively.

Indeed, during the optimization process, the phases are automatically determined, but we are particularly interested in



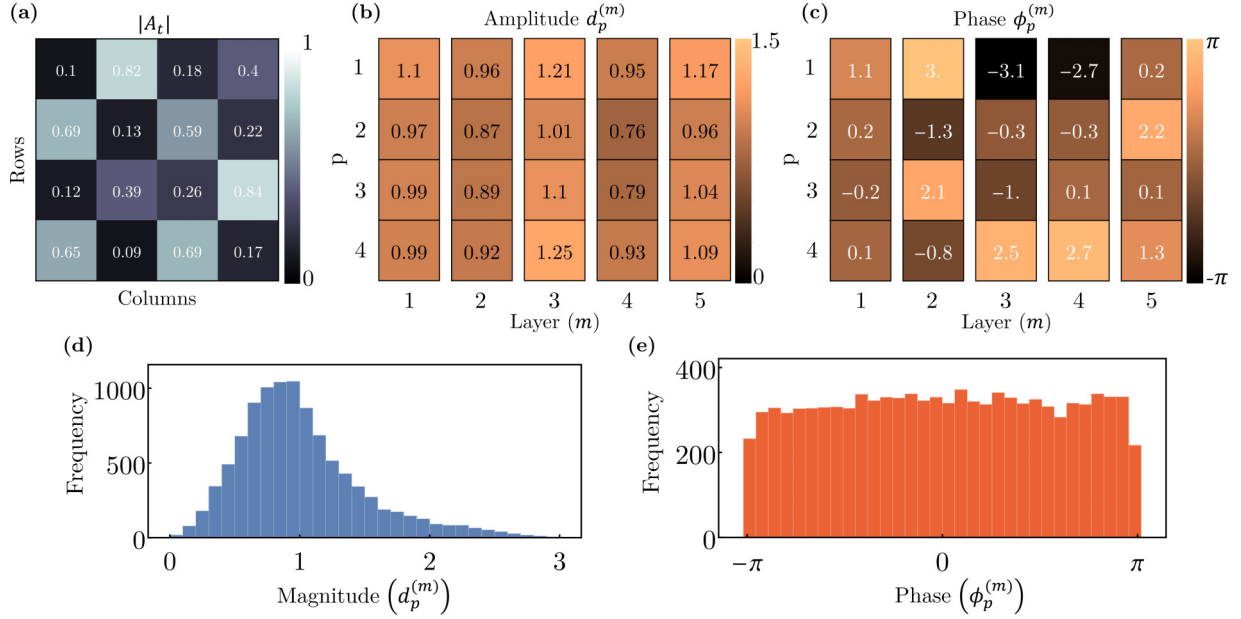


FIG. 3. (a) Example of a random target matrix for  $N = 4$  and  $M = 5$  generated using the SVD with two-Haar random unitary matrices and singular values uniformly chosen in the range  $[0.25, 1]$ . Optimized (b) amplitude and (c) phase parameters for the target matrix shown in (a). Distribution of (d) the magnitudes and (e) phases across all target matrices for  $N = 4$  and  $M = 5$ .

the amplitudes, as they provide the relevant information about the unitary limit. The mean  $\mu_r$  and standard deviation  $S_r$  of  $r$  for each target are illustrated in Fig. 4(b). The latter shows how the optimized amplitudes distribute when  $\sigma_{\min}$  increases from 0.2 to unity for two optimization scenarios. In the first case shown in Fig. 4(b) (left panel), the optimization region where the values of  $d_p^{(m)}$  are searched is bound from above to  $\ell = 1$ . Here, no particular pattern is observed for small  $\sigma_{\min} \approx 0.2$ , which is expected as the random complex-valued targets have no particular structure. In turn, a quite characteristic tendency is observed for  $\sigma_{\min} \rightarrow 1$ , where the standard deviation tends to vanish, implying that the optimized amplitude parameters start to pile up around the vicinity of  $\mu_r = 1$ . Although this numerical evidence suggests that unitaries are recovered when all the optimized amplitudes approach 1, the parameter search was performed in the interval  $0 \leq d_p^{(m)} \leq \ell = 1$ . Thus, the possibility of finding unitaries for  $d_p^{(m)} > \ell$  will not be ruled out. This is, indeed, the second case depicted in Fig. 4(b) (right panel), where the search region has been expanded,  $0 \leq d_p^{(m)} \leq \ell = 1.5$ . The transition to the unitary case is fuzzier, and although  $\mu_r \rightarrow 1$ , the standard deviation  $S_r$  does not vanish; e.g., the amplitude parameters  $d_p^{(m)}$  are strictly different from unity. To explain this, suppose that  $N - 1$  amplitude layers are equal to the identity and one is proportional to the identity, namely,  $\mathcal{A}^{(m_1)} = d_{m_1} \mathbb{I}$  for  $\lambda_1 < 1$  and some  $m_1 \in \{1, \dots, N + 1\}$ . Thus, a second diagonal amplitude layer  $\mathcal{A}^{(m_2)}$  could exist so that  $\mathcal{A}^{(m_2)} = d_{m_1}^{-1} \mathbb{I}$ , leading to a unitary matrix  $A$ . Although both cases render the required unitary matrix, the former is preferable because it is less computationally expensive.

#### D. Defects and error mitigation

So far, the universality of the architecture described by Eq. (1) has been verified. Nevertheless, one can consider

a less idealistic scenario in which the interlaced DFrFT unitary matrix  $F$  contains imperfections due to fabrication errors. Previous unitary factorizations [29] have been shown to be resilient to perturbation on the interlacing matrix  $F$  when the perturbation is considered to be unitary as well. This was achieved by perturbing the Hamiltonian defining the unitary evolution  $F = \exp(iH\pi/2)$ . Here, we account for error in a general manner by considering the perturbations  $\tilde{F}^i(\epsilon) = F^i + \epsilon R^i$ ,  $i \in 1, \dots, N$ , where  $\epsilon \ll 1$  is a perturbation strength parameter and  $R^i$  are complex-valued random matrices with components constrained as  $|R_{p,q}^i| < 1$ . This ensures that each new, nonunitary, interlacing matrix  $\tilde{F}^i$  does not significantly deviate from the DFrFT for small enough  $\epsilon$ . The value of  $\epsilon$  by itself does not provide insight into the error introduced to the system. We thus estimate deviations from the ideal model through the average relative percentage error  $E(\epsilon) := (\|\tilde{F}(\epsilon) - F\|/\|F\|) \times 100\% \equiv \epsilon (\|R\|/\|F\|) \times 100\%$ , which grows linearly with  $\epsilon$ . To visualize the error induced in the matrix  $\tilde{F}$ , we consider 100 random matrices  $R$  for each value of  $\epsilon$  and evaluate the corresponding percentage error  $E(\epsilon)$ . Figure 5(a) displays the mean (solid line) and standard deviation (shaded area) for the set of points generated for each  $\epsilon$ . Particularly, one can see that perturbations induce errors of about 3.3%, 9.77%, and 16.11% for  $\epsilon = 0.02$ ,  $\epsilon = 0.06$ , and  $\epsilon = 0.1$ , respectively. This reveals that, although  $\epsilon = 0.1$  may be considered a small perturbation, it accounts for a considerable error in  $\tilde{F}(\epsilon)$ .

From the previous considerations, we test the effects of  $\tilde{F}^i$  in the interlaced architecture. To this end, a set of 500 Haar random and complex-valued target matrices are constructed for each parameter strength  $\epsilon = 0.02, 0.06, 0.10$ . Likewise, a set of complex-valued random matrices  $R^i$  is generated for each target matrix. First, the set of phase and amplitude parameters  $\{\phi_p^{(n)}, d_p^{(n)}\}_{p=1, n=1}^{N, N+1}$  is optimized when  $\epsilon = 0$

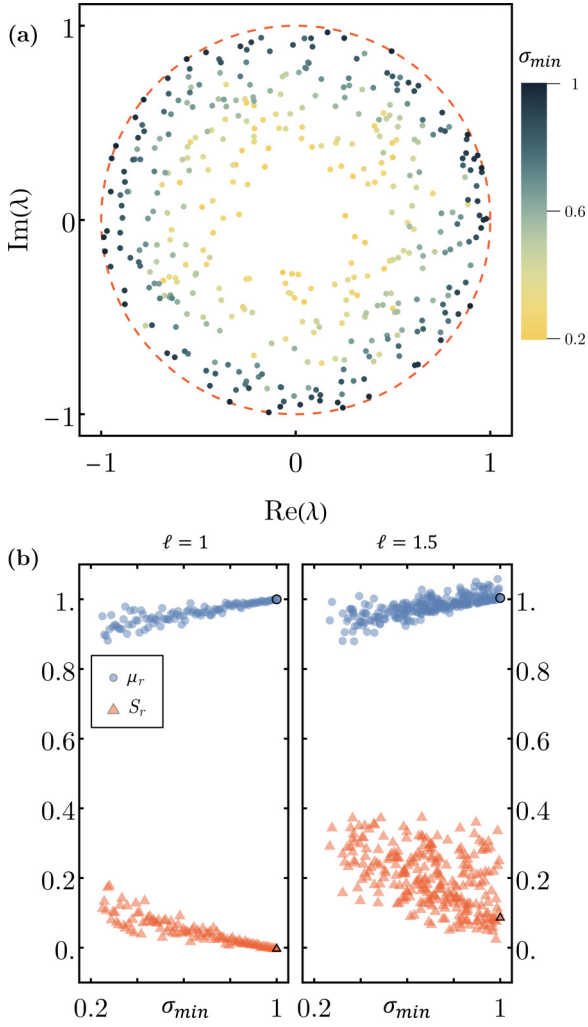


FIG. 4. (a) Eigenvalues  $\lambda$  of 100 randomly generated  $4 \times 4$  target matrices with singular values  $\sigma_{1,2,3,4} \in [\sigma_{\min}, 1]$  for increasing values of  $\sigma_{\min}$ . (b) Mean  $\mu_r$  and standard deviation  $S_r$  of the set of optimized amplitude parameters  $r = \{d_p^{(m)}\}_{p=1, m=1}^{N, M}$  for  $N = 4$ ,  $M = 5$ , and the optimization search region limited by  $\ell$  as a function of  $\sigma_{\min}$ . Black-edged markers correspond to the unitary case ( $\sigma_{\min} = 1$ ).

(unperturbed) and is then used to evaluate the interlaced architecture with  $\tilde{F}^i(\epsilon)$ . The resulting error norms (blue bars) are presented in Fig. 5(b), where it is clear that in all instances, the error norm is always higher than  $10^{-5}$  and above the fixed tolerance  $L_c$ , even for relatively small deviations of 3.3% ( $\epsilon = 0.02$ ). Despite this issue, it is possible to perform a second optimization on the interlaced architecture using the perturbed matrices  $\tilde{F}^i(\epsilon)$ . This leads to the new optimized parameters  $\{\tilde{\phi}_p^{(n)}, \tilde{d}_p^{(n)}\}_{p=1, n=1}^{N, N+1}$ . The updated error norm shown in Fig. 5(b) (orange bars) reveals that, once the second optimization is performed, the error norm drops back under the tolerance value  $L_c$ . Therefore, the architecture (1) is resilient to random defects on the DFrFT layer, which can be amended by properly tuning the phase parameters. Indeed, this holds whenever defects lie within reasonable perturbation strengths; otherwise, the resulting interlaced matrices  $\tilde{F}^i(\epsilon)$  become mostly random matrices (see insets in Fig. 5).

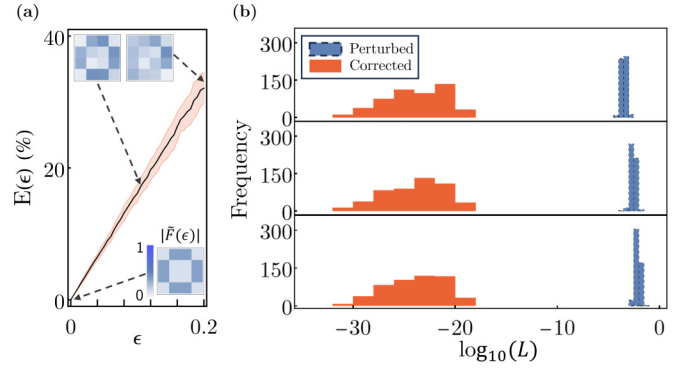


FIG. 5. (a) Mean and standard deviations of the percentage error  $E(\epsilon)$  of the perturbed matrix  $\tilde{F}(\epsilon)$  with respect to the DFrFT matrix  $F$  for  $N = 4$ . The latter was computed using 500 random target matrices per value of  $\epsilon$ . (b) Calibration replacing the mixing layers by the perturbed layers  $\tilde{F}^i = F^i + \epsilon R^i$ , where  $R^i$  are matrices with elements chosen randomly within the complex unit circle. Shown are log norms when the original amplitude and phase parameters are used (blue) and when recalibrating the amplitude and phase parameters (orange).

### E. Photonic-device design and simulation

To further verify the proposed architecture, we perform rigorous wave-propagation simulations using the finite-element method for a four-channel device ( $N = 4$ ) designed to operate at the telecommunication wavelength (1550 nm). Here, we consider silicon-on-insulator ridge waveguides with a width of 500 nm and a height of 200 nm, while we consider refractive indices of  $n_c = 3.47$  and  $n_s = 1.4711$  for the core and substrate regions, respectively. Using the effective index method, one can effectively approximate the problem as a highly accurate two-dimensional model, where the core is replaced by an effective core index  $n_{\text{eff}} = 2.7241$  for the propagation of the fundamental TE mode. In this case, we consider wave propagating along the  $x$  direction and spatially varying along the  $y$  direction, leading to an electric field of the form  $\vec{E} = \mathcal{E}(x, y)\hat{z}$ . The  $J_x$  lattice in this example comprises four waveguides with a separation of  $1.5 \mu\text{m}$  in the innermost channels and  $1.56132 \mu\text{m}$  in the outermost channels. This ensures that a DFrFT of order  $\pi/2$  is achieved for a total propagation length of  $158 \mu\text{m}$ . The phase shifters are simulated by changing the refractive index of the waveguide to achieve the desired change. These are  $20 \times 20 \mu\text{m}^2$  components [yellow bow ties in Fig. 6(a)], based on the specifications of phase-change materials (PCMs) [47,48]. In turn, the amplitude modulators [red rectangles in Fig. 6(a)] are simulated through optical media with varying imaginary parts, a property available using PCMs such as  $\text{Ge}_2\text{Sb}_2\text{Te}$  [49], which allow for a compact and relatively robust extinction rate due to strong deviations on the imaginary part of the refractive index.

An illustration of the propagation of the electric-field amplitude related to the equivalent two-dimensional structure when the architecture is independently excited at each input is presented in Fig. 6(b). For testing purposes, the phase shifters and amplitude parameters are loaded with the values shown in Figs. 3(b) and 3(c), respectively. Note that some amplitude parameters  $d_p^{(m)}$  are larger than 1, and one thus requires an intensity gain in the proposed architecture, which

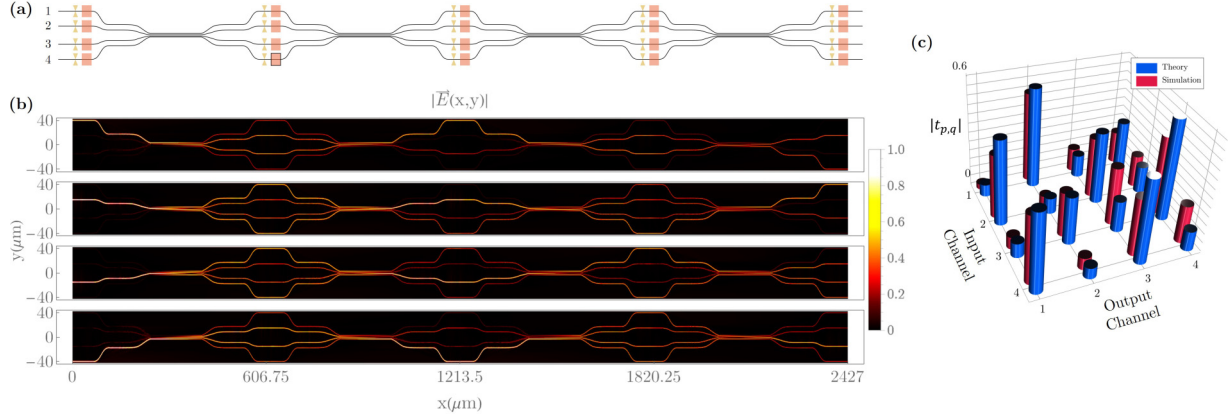


FIG. 6. (a) A sketch of the proposed architecture for  $N = 4$  based on coupled waveguide arrays and amplitude (red) and phase (yellow) modulation using phase-change materials. (b) The electric field  $|\vec{E}(x, y)|$  intensity plot for the complex-matrix factorization (1) obtained from wave simulations when the phase and amplitude parameters are those depicted in Figs. 3(b) and 3(c). The simulations are performed for the transverse-electric-field polarization. (c) Theoretical and simulated amplitudes of the field-transmission-matrix elements  $|t_{pq}|$ , where the transmission-matrix elements are defined as  $t_{pq} = b_p/a_q$ , i.e., the ratio of the electric-field complex modal amplitude measured at the output channel  $p$  over the electric-field complex modal amplitude at the input channel  $q$ . The rescaled version of the parameters in Figs. 3(b) and 3(c) was used in both theory and simulation.

is not being considered here. It is thus more convenient to reconstruct the rescaled target matrix  $\tilde{A}_t = \mathcal{D}^{-1}A_t$ , with  $\mathcal{D} = \prod_{m=1}^N \max(\{d_q^{(m)}\}_{q=1}^N)$ , which requires the amplitude parameters  $\tilde{d}_p^{(m)} = d_p^{(m)} / \max(\{d_q^{(m)}\}_{q=1}^N)$  for its reconstruction. In the latter, we have factored out the largest  $d_p^{(m)}$  for each layer  $m$ . By considering  $\tilde{A}_t$ , we can eliminate the need for amplitude-gain elements in the architecture, simplifying the design and reducing the manipulation of amplitude parameters. Specifically, one limits the manipulation of amplitude parameters to  $N - 1$  per layer instead of scaling every parameter, as the initially largest parameters per layer is now fixed to 1 while the rest are rescaled accordingly.

For comparison, the theoretical and simulated amplitudes of the field transmission matrix  $t_{p,q} = b_p/a_q$  are shown in Fig. 6(c). In this notation,  $b_p$  represents the complex modal amplitude of the electric field in the output channel  $p$ , and  $a_q$  represents the complex modal amplitude of the electric field in the input channel  $q$ . Here, the excitation is swept across all  $p$  input channels, while the output is measured at the  $q$ th channel. These results show some mild deviations of the simulation from the theory, possibly due to several factors, such as imperfect coupling in the waveguide array, amplitude modulators, and phase shifters. Despite this, the mean square error  $L$  of the simulated output with respect to the theoretical rescaled target  $\tilde{A}_t$  leads to  $L = 6.62918 \times 10^{-3}$ . The latter error can be improved by first identifying the defects on the lattice and accordingly modifying the interlacing layer  $F$ ; then, the parameters will be optimized a second time so that the defects of  $F$  are taken into account. This is, indeed, the procedure discussed in [29] for unitary architectures, which can be adapted for the present device.

### III. DISCUSSION AND CONCLUSION

In summary, we proposed an architecture for implementing arbitrary discrete linear operators with photonic

integrated circuits. The proposed architecture is built on interlacing fixed discrete-fractional-Fourier-transform layers with programmable amplitude-and-phase-modulator layers. Our results indicate that this architecture can universally represent  $N \times N$  complex-valued operations with at most  $N + 1$  controllable layers. The proposed architecture offers a fairly simple physical realization by utilizing photonic waveguide arrays in conjunction with arrays of amplitude and phase modulators. It should be noted that even if the singular values of the target matrix are less than 1, in general, optimizing the amplitude modulators requires both gains ( $d_p^{(m)} > 1$ ) and losses ( $d_p^{(m)} < 1$ ). To overcome this challenge and to bypass the demanding requirement for gain, the target matrix  $A_t$  was properly rescaled and then reconstructed. Although modulators based on Mach-Zehnder interferometers or optical ring resonators [50] are known optical implementations for amplitude modulators, such elements substantially increase the overall length and complexity of the final device, leading to potentially additional errors due to the presence of the coupling elements. Alternatively, phase-change materials based on  $\text{Ge}_2\text{Sb}_2\text{Te}$  and related structures [48,49] are emerging candidates to replace and implement the required amplitude modulation. In the microwave domain, integrated active elements that produce both amplitude gains and losses are relatively simple to implement in the architecture. Further microwave realizations of the DFrFT using microstrip and interdigital capacitors was demonstrated and experimentally verified [42].

Potential manufacturing errors on the waveguide array characterizing the DFrFT were taken into account and showed the overall resilience of the universality performance of the architecture, provided that such errors lie within allowed tolerance values, such as perturbations on  $F$  with a percentage error of up to 18%. Although the optimized parameters obtained from the unperturbed lattice produce a significant error norm  $L$  when defects are taken into account, a second optimization brings the error back to the imposed tolerance error. Therefore, defects can be mitigated overall by adequately tuning



the available active elements in the architecture. The latter brings evidence that the passive layer  $F$  could be replaced by another suitable waveguide array. Indeed, for unitary architectures, it was numerically found that interlacing layers could be replaced by matrices with an adequate density property across their matrix elements [30], and some random unitary matrices have been shown to be suitable for such a task [25].

Alternatively, the architecture can be configured to operate as a unitary matrix by fixing all the amplitude modulators to unity,  $d_p^{(m)} = 1$ , leaving only the phase elements  $\phi_p^{(m)}$  as the tunable parameters. That is, the effective number of parameters reduces from  $2N(N+1)$  to  $N(N+1)$ , rendering a functionality equivalent to that of other unitary photonic solutions reported in the literature [23,28–30,51]. The latter is by no means the only setting in which the device becomes unitary. Indeed, Fig. 4 shows numerical optimizations where amplitude parameters  $d_p^{(m)} \neq 1$  also render unitary targets. This follows from the overparametrization of the architecture, which opens the possibility for more solutions.

The universal nature of the proposed architecture permits its deployment in optical processing applications such as convolutional network accelerators [52,53]. Since the device represents a general complex-valued matrix, it can be used as a linear layer in optical neural networks, a solution explored using alternative optical elements such as Mach-Zehnder interferometers [6] and microring resonators [54]. Furthermore, given the versatility of the proposed solution, further applications can be explored by changing the material platform to those suitable for individual photon manipulation and photon entanglement generated through nonlinear effects. The latter includes silicon nitride ( $\text{Si}_3\text{N}_4$ ) and lithium niobate solutions. Applications in this regard include boson sampling in quantum computing tasks [55,56], production of NOON states [57], and on-chip production of entangled photon states [58,59].

## ACKNOWLEDGMENTS

This project is supported by U.S. Air Force Office of Scientific Research (AFOSR) Young Investigator Program (YIP) Award No. FA9550-22-1-0189 and U.S. National Science Foundation (NSF) FuSe Grant No. 2329021.

## APPENDIX: COUPLED-MODE THEORY AND PHOTONIC DFRFT

This Appendix briefly summarizes the mathematical derivation of the DFrFT. Here, we consider a dielectric waveguide composed of a main core and buried in a transparent media with refractive indexes  $n_{co}$  and  $n_{cla}$ , respectively. The waveguide has rectangular transversal geometry on the  $y$ - $z$  plane and extends longitudinally through the  $x$  axis. For design and analysis, it is convenient to approximate the waveguide system as a simpler representation using the effective index theory [60], where the three-dimensional structure collapses into an approximated two-dimensional planar model, in which the refractive index of the core is replaced by the effective index  $n_{\text{eff}}$ . Under these considerations, the electric field of the propagating wave is written as  $\mathbf{E}(x, y) =$

$\mathcal{E}(y)e^{i(\omega t - \beta x)}\mathbf{e}_z$ , with  $\beta$  being the propagation constant,  $\mathbf{e}_z$  being a Euclidean unit vector, and  $\mathcal{E}(y)$  being the electric-field amplitude determined by imposing the proper boundary conditions. For guided quasi-TE and quasi-TM modes, the modal field  $\mathcal{E}(y)$  has finite power, and the corresponding propagation constants form a discrete set  $\{\beta_m\}_{m=1}^K$ , with  $K \in \mathbb{Z}^+$ . For details see [61].

Waveguide arrays are composed of parallel, not necessarily equidistant, neighboring waveguides, each designed to allow the propagation of guided modes. In most cases, it is desired to allow only one guided mode, the fundamental mode, so that light coupling between neighbors always excites only one mode. In this scenario, light “hops” from one waveguide to its neighbor due to the overlap of the evanescent wave solutions of the guided modes. For a set of  $N$  single-mode coupled waveguides, the wave evolution across the array is ruled by coupled-mode theory [61,62], which is characterized by the discrete Schrödinger-like equation  $-id\mathbf{E}/dt = H\mathbf{E}$ , with  $\mathbf{E} := (a_1, \dots, a_N)^T \mathcal{E}(y)$  being the corresponding supermodes and  $a_p \in \mathbb{C}$  being the modal coefficient of the  $p$ th waveguide, for  $p \in \{1, \dots, N\}$ . Here, in general,  $H \in \mathbb{C}^{N \times N}$  is a tridiagonal Hermitian matrix with respect to the Euclidean inner product in  $\mathbb{C}^N$  with components  $H_{p,q} = \kappa_p \delta_{p+1,q} + \kappa_{p-1}^* \delta_{p-1,q} + v_p \delta_{p,q}$ . Here,  $\kappa_p$  and  $v_p$  are the neighbor and on-site interactions, respectively, and  $\kappa_0 = \kappa_{N+1} = 0$ . Particularly, when all waveguides have the same dimensions, the on-site term vanishes ( $v_p = 0$ ).

The  $J_x$  operator is well known in the context of angular momentum in quantum mechanics and free-space optics [39,63], with an integrated photonic analog through the coupling parameters [37]  $\kappa_p = \frac{\tilde{\kappa}}{2} \sqrt{(N-p)p}$ , with  $\tilde{\kappa}$  being a scaling coupling factor and  $v_p = 0$ . Since the couplings  $\kappa_p$  do not depend on  $x$ , the wave evolution can be written in terms of the unitary operator  $F(\alpha) := e^{-iH\alpha}$  as  $\mathbf{E}(x = \alpha) = F(\alpha)\mathbf{E}(x = 0)$ , where  $\mathbf{E}(0) \equiv \mathbf{E}$  is the initial state at the array input. This renders the eigenvalue problem  $H\mathbf{E}^{(n)} = \lambda_n \mathbf{E}^{(n)}$ , with  $\mathbf{E}^{(n)} := (a_1^{(n)}, \dots, a_N^{(n)})^T \mathcal{E}(y)$  and  $\lambda_n$  being the array supermodes and eigenvalues, respectively, given by

$$a_q^{(n)} = 2^{\frac{2q-N-1}{2}} \sqrt{\frac{(N-q)!(q-1)!}{(N-n)!(n-1)!}} P_{q-1}^{(n-q, -n-q+N+1)}(0),$$

$$\lambda_n = \frac{2n - N - 1}{2}, \quad \forall n, q \in \{1, \dots, N\}. \quad (\text{A1})$$

The upper index  $n$  denotes the supermode number,  $q$  is the waveguide (channel) number, and  $P_n^{(a,b)}(z)$  are the Jacobi polynomials. The unitary matrix is written explicitly as [37]

$$F_{p,q}(\alpha) = \sum_k e^{-i\alpha k} u_k^{(p)} u_k^{(q)} \equiv i^{p-q} \left[ \sin\left(\frac{\alpha}{2}\right) \right]^{q-p} \times \left[ \cos\left(\frac{\alpha}{2}\right) \right]^{-q-p+N+1} \sqrt{\frac{(p-1)!(N-p)!}{(q-1)!(N-q)!}} \times P_{p-1}^{(q-p, -q-p+N+1)}[\cos(\alpha)]. \quad (\text{A2})$$

The unitary operator  $F(\alpha)$  meets all the required properties for the DFrFT, that is, unitarity, the additivity rule, and the existence of a cyclic order (see [52]). Thus,  $F(\alpha)$  is deemed a photonic candidate for DFrFT of order  $\alpha$ .



- [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 (London)* **586**, 207 (2020).
- [3] J. Notaros *et al.*, Programmable dispersion on a photonic integrated circuit for classical and quantum applications, *Opt. Express* **25**, 21275 (2017).
- [4] N. C. Harris *et al.*, Quantum transport simulations in a programmable nanophotonic processor, *Nat. Photon.* **11**, 447 (2017).
- [5] S. Slussarenko and G. J. Pryde, Photonic quantum information processing: A concise review, *Appl. Phys. Rev.* **6**, 041303 (2019).
- [6] Y. Shen *et al.*, Deep learning with coherent nanophotonic circuits, *Nat. Photon.* **11**, 441 (2017).
- [7] B. J. Shastri, A. N. Tait, T. Ferreira de Lima, W. H. Pernice, H. Bhaskaran, C. D. Wright, and P. R. Prucnal, Photonics for artificial intelligence and neuromorphic computing, *Nat. Photon.* **15**, 102 (2021).
- [8] G. Wetzstein, A. Ozcan, S. Gigan, S. Fan, D. Englund, M. Soljačić, C. Denz, D. A. Miller, and D. Psaltis, Inference in artificial intelligence with deep optics and photonics, *Nature (London)* **588**, 39 (2020).
- [9] D. A. Miller, Optical interconnects to electronic chips, *Appl. Opt.* **49**, F59 (2010).
- [10] C. A. Thraskias, E. N. Lallas, N. Neumann, L. Schares, B. J. Offrein, R. Henker, D. Plettemeier, F. Ellinger, J. Leuthold, and I. Tomkos, Survey of photonic and plasmonic interconnect technologies for intra-datacenter and high-performance computing communications, *IEEE Commun. Surv. Tutorials* **20**, 2758 (2018).
- [11] J. Cariñe *et al.*, Multi-core fiber integrated multi-port beam splitters for quantum information processing, *Optica* **7**, 542 (2020).
- [12] L. S. Madsen *et al.*, Quantum computational advantage with a programmable photonic processor, *Nature (London)* **606**, 75 (2022).
- [13] L. Lu *et al.*, Three-dimensional entanglement on a silicon chip, *npj Quantum Inf.* **6**, 30 (2020).
- [14] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, Experimental realization of any discrete unitary operator, *Phys. Rev. Lett.* **73**, 58 (1994).
- [15] D. A. Miller, Self-configuring universal linear optical component, *Photon. Res.* **1**, 1 (2013).
- [16] 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).
- [17] F. Shokraneh, S. Geoffroy-Gagnon, and O. Liboiron-Ladouceur, The diamond mesh, a phase-error- and loss-tolerant field-programmable MZI-based optical processor for optical neural networks, *Opt. Express* **28**, 23495 (2020).
- [18] K. Rahbardar Mojaver, B. Zhao, E. Leung, S. Safaee, and O. Liboiron-Ladouceur, Addressing the programming challenges of practical interferometric mesh based optical processors, *Opt. Express* **31**, 23851 (2023).
- [19] S. Yu and N. Park, Heavy tails and pruning in programmable photonic circuits for universal unitaries, *Nat. Commun.* **14**, 1853 (2023).
- [20] R. Tang, R. Tanomura, T. Tanemura, and Y. Nakano, Lower-depth programmable linear optical processors, *Phys. Rev. Appl.* **21**, 014054 (2024).
- [21] R. Tang, T. Tanemura, and Y. Nakano, Integrated reconfigurable unitary optical mode converter using MMI couplers, *IEEE Photon. Technol. Lett.* **29**, 971 (2017).
- [22] Q. Ling, P. Dong, Y. Chu, X. Dong, J. Chen, D. Dai, and Y. Shi, On-chip optical matrix-vector multiplier based on mode division multiplexing, *Chip* **2**, 100061 (2023).
- [23] 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).
- [24] V. L. Pastor, J. Lundeen, and F. Marquardt, Arbitrary optical wave evolution with Fourier transforms and phase masks, *Opt. Express* **29**, 38441 (2021).
- [25] 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).
- [26] N. Skryabin, I. Dyakonov, M. Y. Saygin, and S. Kulik, Waveguide-lattice-based architecture for multichannel optical transformations, *Opt. Express* **29**, 26058 (2021).
- [27] R. Tanomura, T. Tanemura, and Y. Nakano, Multi-wavelength dual-polarization optical unitary processor using integrated multi-plane light converter, *Jpn. J. Appl. Phys.* **62**, SC1029 (2023).
- [28] M. Markowitz and M.-A. Miri, Universal unitary photonic circuits by interlacing discrete fractional Fourier transform and phase modulation, *arXiv:2307.07101*.
- [29] M. Markowitz, K. Zelaya, and M.-A. Miri, Auto-calibrating universal programmable photonic circuits: Hardware error-correction and defect resilience, *Opt. Express* **31**, 37673 (2023).
- [30] K. Zelaya, M. Markowitz, and M.-A. Miri, The Goldilocks principle of learning unitaries by interlacing fixed operators with programmable phase shifters on a photonic chip, *Sci. Rep.* **14**, 10950 (2024).
- [31] J. Müller-Quade, H. Aagedal, T. Beth, and M. Schmid, Algorithmic design of diffractive optical systems for information processing, *Phys. D (Amsterdam, Neth.)* **120**, 196 (1998).
- [32] M. Schmid, R. Steinwandt, J. Müller-Quade, M. Rötteler, and T. Beth, Decomposing a matrix into circulant and diagonal factors, *Linear Algebra Its Appl.* **306**, 131 (2000).
- [33] M. Huhtanen and A. Perämäki, Factoring matrices into the product of circulant and diagonal matrices, *J. Fourier Anal. Appl.* **21**, 1018 (2015).
- [34] A. Araujo, B. Negrevergne, Y. Chevalere, and J. Atif, Training compact deep learning models for video classification using circulant matrices, *arXiv:1810.01140*.
- [35] J. M. Lukens, H.-H. Lu, B. Qi, P. Lougovski, A. M. Weiner, and B. P. Williams, All-optical frequency processor for networking applications, *J. Lightwave Technol.* **38**, 1678 (2020).
- [36] J. M. Lukens and P. Lougovski, Frequency-encoded photonic qubits for scalable quantum information processing, *Optica* **4**, 8 (2017).

- [37] S. Weimann *et al.*, Implementation of quantum and classical discrete fractional Fourier transforms, *Nat. Commun.* **7**, 11027 (2016).
- [38] R. M. Gray, Toeplitz and circulant matrices: A review, *Found. Trends Commun. Inf. Theory* **2**, 155 (2005).
- [39] N. M. Atakishiyev and K. B. Wolf, Fractional Fourier–Kravchuk transform, *J. Opt. Soc. Am. A* **14**, 1467 (1997).
- [40] C. Candan, M. A. Kutay, and H. M. Ozaktas, The discrete fractional Fourier transform, *IEEE Trans. Signal Process.* **48**, 1329 (2000).
- [41] M. Honari-Latifpour, A. Binaie, M. A. Eftekhar, N. Madamopoulos, and M.-A. Miri, Arrayed waveguide lens for beam steering, *Nanophotonics* **11**, 3679 (2022).
- [42] R. Keshavarz, N. Shariati, and M.-A. Miri, Real-time discrete fractional fourier transform using metamaterial coupled lines network, *IEEE Trans. Microwave Theory Tech.* **71**, 3414 (2023).
- [43] T. K. Moon and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing* (Prentice Hall, Upper Saddle River, NJ, 2000).
- [44] F. Mezzadri, How to generate random matrices from the classical compact groups, [arXiv:math-ph/0609050](https://arxiv.org/abs/math-ph/0609050).
- [45] K. Levenberg, A method for the solution of certain non-linear problems in least squares, *Q. Appl. Math.* **2**, 164 (1944).
- [46] J. Zhou, J. Wu, and Q. Hu, Tunable arbitrary unitary transformer based on multiple sections of multicore fibers with phase control, *Opt. Express* **26**, 3020 (2018).
- [47] C. Ríos *et al.*, Ultra-compact nonvolatile phase shifter based on electrically reprogrammable transparent phase change materials, *Photonix* **3**, 26 (2022).
- [48] N. Youngblood, C. A. Ríos Ocampo, W. H. Pernice, and H. Bhaskaran, Integrated optical memristors, *Nat. Photon.* **17**, 561 (2023).
- [49] Y. Zhang *et al.*, Broadband transparent optical phase change materials for high-performance nonvolatile photonics, *Nat. Commun.* **10**, 4279 (2019).
- [50] Q. Xu, B. Schmidt, J. Shakya, and M. Lipson, Cascaded silicon micro-ring modulators for WDM optical interconnection, *Opt. Express* **14**, 9431 (2006).
- [51] R. Tang, R. Tanomura, T. Tanemura, and Y. Nakano, Ten-port unitary optical processor on a silicon photonic chip, *ACS Photon.* **8**, 2074 (2021).
- [52] K. Zelaya and M.-A. Miri, Integrated photonic fractional convolution accelerator, *Photon. Res.* **12**, 1828 (2024).
- [53] J. R. Ong, C. C. Ooi, T. Y. L. Ang, S. T. Lim, and C. E. Png, Photonic convolutional neural networks using integrated diffractive optics, *IEEE J. Sel. Top. Quantum Electron.* **26**, 1 (2020).
- [54] R. Tang, S. Ohno, K. Tanizawa, K. Ikeda, M. Okano, K. Toprasertpong, S. Takagi, and M. Takenaka, A symmetric silicon microring resonator optical crossbar array for accelerated inference and training in deep learning, *Photon. Res.* **12**, 1681 (2024).
- [55] N. J. Russell, L. Chakhmakhchyan, J. L. O’Brien, and A. Laing, Direct dialling of Haar random unitary matrices, *New J. Phys.* **19**, 033007 (2017).
- [56] M. Clementi *et al.*, Programmable frequency-bin quantum states in a nano-engineered silicon device, *Nat. Commun.* **14**, 176 (2023).
- [57] R. Kruse, L. Sansoni, S. Brauner, R. Ricken, C. S. Hamilton, I. Jex, and C. Silberhorn, Dual-path source engineering in integrated quantum optics, *Phys. Rev. A* **92**, 053841 (2015).
- [58] X.-W. Luo, Q.-Y. Zhang, P. Xu, R. Zhang, H.-Y. Liu, C.-W. Sun, Y.-X. Gong, Z.-D. Xie, and S.-N. Zhu, On-chip engineering of high-dimensional path-entangled states in a quadratic coupled-waveguide system, *Phys. Rev. A* **99**, 063833 (2019).
- [59] A. Belsley, T. Pertsch, and F. Setzpfandt, Generating path entangled states in waveguide systems with second-order nonlinearity, *Opt. Express* **28**, 28792 (2020).
- [60] M. Hammer and O. V. Ivanova, Effective index approximations of photonic crystal slabs: A 2-to-1-D assessment, *Opt. Quantum Electron.* **41**, 267 (2009).
- [61] A. Yariv and P. Yeh, *Photonics: Optical Electronics in Modern Communications* (Oxford University Press, Oxford, 2007).
- [62] W.-P. Huang, Coupled-mode theory for optical waveguides: An overview, *J. Opt. Soc. Am. A* **11**, 963 (1994).
- [63] L. M. Narducci and M. Orszag, Eigenvalues and eigenvectors of angular momentum operator  $J_x$  without the theory of rotations, *Am. J. Phys.* **40**, 1811 (1972).