Robust Architecture for Programmable Universal Unitaries

M. Yu. Saygin[1,*] I. V. Kondratyev,[1] I. V. Dyakonov,[1] S. A. Mironov,[2,3,4] S. S. Straupe,[1] and S. P. Kulik[1]

[1]*Quantum Technologies Center, Faculty of Physics, Lomonosov Moscow State University, Leninskie gory 1, building 35, Moscow 119991, Russia*
[2]*Institute for Nuclear Research of the Russian Academy of Sciences, 60th October Anniversary Prospect, 7a, Moscow 117312, Russia*
[3]*Institute for Theoretical and Experimental Physics, Bolshaya Cheriomyshkinskaya, 25, Moscow 117218, Russia*
[4]*Moscow Institute of Physics and Technology, Institutski pereulok, 9, Dolgoprudny 141701, Russia*

The decomposition of large unitary matrices into smaller ones is important because it provides ways to the realization of classical and quantum information processing schemes. Today, most of the methods use planar meshes of tunable two-channel blocks; however, the schemes turn out to be sensitive to fabrication errors. We study a novel decomposition method based on multichannel blocks. We have shown that the scheme is universal even when the block's transfer matrices are chosen at random, making it virtually insensitive to errors. Moreover, the placement of the variable elements can be arbitrary, so that the scheme is not bound to specific topologies. Our method can be beneficial for large-scale implementations of unitary transformations by techniques, which are not of wide proliferation today or have yet to be developed.

Linear transformations of multiple channels are ubiquitous in many applied fields, as well as in fundamental research. The implementations of linear multichannel transformations are particularly important in the optical context, in which they are used as indispensable components of communication and computing systems, in particular, those based on novel approaches to information processing. For example, universal optical interferometers —devices capable of performing arbitrary linear transformations—are exploited as mode unscramblers [1] or as a part of photonic neural networks [2,3].

In addition to classical applications, universal optical interferometers play an important role in the implementations of quantum algorithms. In particular, multichannel interferometers are a necessary part of the promising quantum computing platform that leverages linear-optical circuits and nonclassical properties of photons to realize quantum algorithms [4,5]. Recent works have demonstrated the versatility of linear-optical quantum systems and their ability to perform a number of quantum computing tasks ranging from the well-known algorithms [6,7] to the more specific ones, such as boson sampling [8–10] and variational algorithms [11–13]. The efficient realization of the later two classes of algorithms was made possible by the universal programmable interferometers [14].

The complexity of the multichannel scheme, which is quantified by the number of channels in the interferometer and its programmability, defines its practical utility. These characteristics, however, are largely determined by a particular architecture used to construct the interferometer. Nowadays, the most widely used universal architectures belong to different planar varieties. One reason for this is the fact that planar schemes are easier to realize by the mature integrated photonics technology. On the other hand, methods have been devised that enable the decomposition of large unitary matrices into planar meshes of smaller building blocks.

The decomposition methods of wide use today are obtained from the unitary matrix factorization theorem proposed in [15], which was adopted to construct planar reconfigurable schemes, e.g., in [16,17]. In the linear-optical context, such a multichannel scheme has a form of a mesh, constructed out of tunable Mach-Zehnder interferometers (MZIs), thereby enabling the whole interferometer to be reconfigured by tuning the variable phase elements [16,17]. For these schemes to be universal, the beam splitters that constitute the MZIs should necessarily be balanced, therefore, imposing strict requirements on the fabrication tolerances. However, in practical implementations this condition is not fully satisfied, so that the multichannel transformation is universal only to a certain degree [18]. This negative effect becomes more pronounced as the interferometer scheme is scaled up. Therefore, the realization of sophisticated information processing algorithms with optics calls for the development of novel architectures for multichannel interferometers that are more resilient to the implementation errors.

In this Letter we explore an alternative approach to building large-scale universal interferometers, which is drastically different from the known methods based on meshes of two-channel blocks. It is based on sequences of fixed multichannel mixing blocks instead of arrays of balanced beam splitters. As a result, the interferometric scheme provides the freedom to choose the transfer matrix

of the blocks out of the continuous unitary space. Surprisingly, such schemes turn out to be extremely robust to perturbations in the transfer matrices of the blocks, in fact, as we show below, they may be chosen at random, and the desired transformation may be dialed later on the manufactured device by tuning the phase shifts only. In addition, the placement of the variable phase shifts in the scheme can also be chosen arbitrary. Since no relation to the scheme topology is implied in the analysis, a wide variety of technologies can be used to realize it in practice, for example, those based on frequency and temporal encoding, that may potentially enlarge the scale of the universal interferometric schemes.

Methods of decomposition of unitary matrices.—Several types of parametrization for unitary matrices are known [15,19–21]; however, only a few of them give a recipe to build up large devices out of elementary blocks, which is convenient to realize in practice. Linear optics is the most representative domain of practical application for multichannel transformations, where their implementations are employed to transform the input vector of field amplitudes $\mathbf{a}^{(\mathrm{in})}$ into the output vector $\mathbf{a}^{(\mathrm{out})}$ by transformation, described by a matrix U: $\mathbf{a}^{(\mathrm{out})} = U\mathbf{a}^{(\mathrm{in})}$. Following the decomposition methods, the required matrix U, that can be regarded as a $\mathrm{SU}(N)$ transformation, is obtained by setting appropriate values of some parameters, characterizing the elements in the actual physical network [16,17].

A natural building block for multichannel schemes in the context of optics is a MZI consisting of a pair of static balanced beam splitters and two variable phase shifters. Adding a third phase to the MZI allows us to cover an entire $\mathrm{SU}(2)$ group [22] with the 2×2 transformation matrix:

$$e^{i\psi} U_{\mathrm{MZI}}(\theta, \varphi) = e^{i(\psi-\pi)/2 X} e^{i(\theta+\pi)Z} e^{i(\varphi-\pi)/2 X}, \quad (1)$$

where X and Z are the Pauli matrices, φ, θ, and ψ are the phase parameters.

Although, in principle any fixed beam splitter transformation mixing a pair of optical modes densely generates $\mathrm{SU}(N)$ [23], a more efficient way is to implement a network scheme allowing us to program the required unitary using tunable MZIs. The experiments reported in [4,14,24,25] have demonstrated the capabilities of such a network design to experimentally approximate arbitrary unitaries. However, the performance of the linear optical network heavily depends on the quality of individual MZI elements. For example, the work [18] studied the effect of individual beam splitter errors on the overall fidelity and concluded that the beam splitter reflectivity errors of few percents diminish the quality of the unitary transformations significantly. This makes a subset of unitary transformations unavailable for the programmable linear optical network composed of imperfect optical elements [26].

As an illustrative example, consider the decomposition proposed in [17], which is widely used today to construct
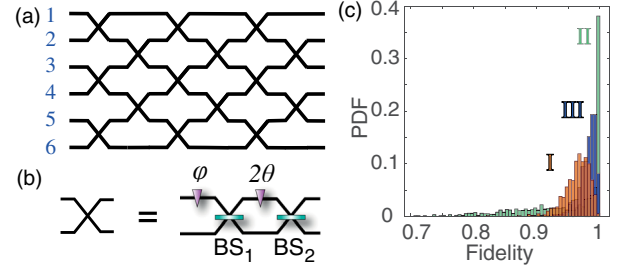

FIG. 1. The planar universal design of a multichannel decomposition proposed in [17]. (a) The scheme layout for $N = 6$. (b) The building block of the scheme—an MZI with two tunable phase shifts, θ and φ, and static balanced beam splitters, BS_1 and BS_2. (c) The histograms illustrating fidelity of transformations [see formula (4)] for multichannel schemes with $N = 10$, corresponding to three error models: (I) all static beam splitters are biased by an equal angle α, chosen randomly in the range $[0;$ $20]$ degrees (red); (II) the beam splitters are biased independently by angles α_j distributed in the range $[-20; 20]$ degrees (green); (III) the beam splitters are biased independently by angles α_j distributed in the range $[0; 20]$ degrees (blue). Each histogram was obtained by an optimization procedure that found a global maximum of fidelity with respect to the set of phase shifts, collected over 1000 randomly generated unitary matrices.

reconfigurable planar optical schemes. The corresponding scheme layout is illustrated in Fig. 1(a) for a number of inputs and outputs $N = 6$. Its main building block—an MZI is depicted in Fig. 1(b). Figure 1(c) illustrates the negative effect of the beam splitter imbalances on the quality of the transformation, clearly witnessing the sensitivity of this decomposition to errors.

Layered decompositions with static multichannel blocks.—The decomposition we study in this Letter is based on static multichannel blocks, rather than two-channel balanced beam splitters. The schematic of the decomposition is shown in Fig. 2. It is built up of multiple layers, that come in two types, which are stacked alternately. The variable layer consists of independent single phase shifts φ_j, so that its N-channel transfer matrix has the diagonal form: $\Phi(\vec{\varphi}) = \mathrm{diag}[\exp(i\vec{\varphi})]$, where $\vec{\varphi} = (\varphi_1, \varphi_2, \ldots, \varphi_N)$ is a vector of phase shifts; thus, we call it the phase layer. The other type of layer, in the following referred to as the mixing layer, introduces interaction between the channels that is required for multichannel
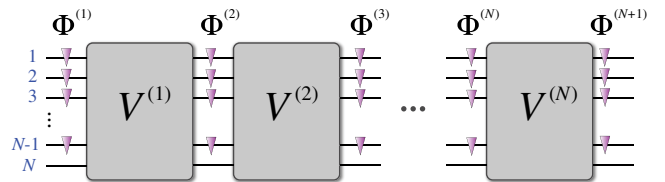

FIG. 2. A schematic of a matrix decomposition (2) the number of layers $K = N$ corresponding to $N - 1$ phase shifts in each layer.

interference. It is the aim of our work to find out what transfer matrices $V^{(m)}$ should these mixing layers have in order for the whole scheme to be universal. Moreover, we are interested in a general case, when the mixing layers can have different transfer matrices $V^{(m)}$, so that the overall transformation has the following form:

$$U = \Phi^{(K+1)}V^{(K)}\Phi^{(K)}\cdot, \dots, \cdot V^{(1)}\Phi^{(1)} \qquad (2)$$

where $\Phi^{(m)} = \Phi[\vec{\varphi}^{(m)}]$ with $\vec{\varphi}^{(m)}$ being the set of phases describing the layer with index m, K is the number of mixing layers. The decomposition (2) enables a variety of schemes, each having different number of layers K at a given N. Here, we describe the case of $K = N$ [27].

Previous works considered special cases of decomposition (2) with specific choice of $V^{(m)}$. A design of the form (2) with the mixing layers described by the discrete Fourier transform matrix:

$$U_{mn}^{(\text{DFT})} = \frac{1}{\sqrt{N}}\exp\left(i\frac{2\pi}{N}(m-1)(n-1)\right), \qquad (3)$$

$m, n = 1, \dots, N$, implemented in a planar integrated photonic circuit was studied in [33]. Work [34] deals with the mixing layers implemented by regions of coupled waveguides. The authors of these works provided some numerical evidence in favor of universality of this specific types of circuits. In both these works, the transfer matrices of the mixing layers were essentially defined with a specific system in mind. Besides, the number of phase shifts per layer was $N - 1$, which is a maximum possible number. Relevant ideas concerning decomposition (2) can be found in [33,35,36]. In this Letter, we show that the class of transfer matrices $V^{(m)}$ suitable for universal programmable schemes is not bounded to some specific instances, to the contrary, it occupies a large part of the SU(N) group volume.

We first notice that by simple counting of independent parameters necessary to define an arbitrary unitary matrix, a universal scheme performing an SU(N) transformation should have at least $N^2 - 1$ tunable phase shifts. We will restrict our attention to the schemes using exactly this minimal set of phases.

We quantify performance of the decomposition (2) by fidelity, defined as:

$$F(U, U_s) = \frac{1}{N^2}|\text{Tr}(U^\dagger U_s)|^2, \qquad (4)$$

which compares a target unitary matrix U_s and an actual transfer matrix realized by the decomposition U (2). Provided that the matrices U and U_s are equal up to a global phase, fidelity (4) gets its maximum value $F = 1$.

In contrast to the known decomposition of [17], which comes with an analytical procedure allowing us to obtain a set of phases at which $F = 1$ for any given U_s, we could not find an analytical solution for the general decomposition (2). Therefore, a numerical optimization procedure has been used. Our numerical algorithm is based on the basin-hopping algorithm realized using the SciPy python library. Given a unitary matrix U_s, the algorithm was searching for a global minimum of infidelity $1 - F$ over the space of phase vectors $\vec{\varphi}^{(m)}$ ($m = 1, \dots, K + 1$). To decrease the chance of sticking into local minima, we used multiple runs of the basin-hopping routine with random initial values of the phases. The algorithm has reasonable efficiency and has subexponential runtime dependence on the interferometer size [37]. This way, we have achieved the error of calculating the global minimum on the level of $\sim 10^{-9}$ [37]. Each numerical experiment involved optimization over a series of 1000 matrices U_s.

Let us now discuss the choice of the transfer matrices of the mixing layers, used in construction of multichannel transformations. Surprisingly, not only specific fixed unitaries are suitable, but almost any unitary matrix will work as a mixing layer. Moreover, the mixing transformation need not be fixed, it can vary from layer to layer. To demonstrate that, we chose every matrix $V^{(m)}$ in (2) at random using the very same approach, described above for the generation of U_s. Figure 3 presents the results of optimization for decomposition depicted in Fig. 2(a) for the case of a fixed transfer matrix, namely a DFT matrix (3), and random transfer matrices of the mixing layers. In the optimization process we varied the number of phase layers, ranging from small values containing less phase shifts than required for universality, which is done for illustrative purposes, to the proper number containing in total $N^2 - 1$ phase shifts. The worst value of infidelity corresponding to the latter case for all matrix sizes presented was not greater than $1 - F < 10^{-9}$ and its nonvanishing value is attributed to the finite accuracy of the implementation of the numerical algorithm.

To make sure, that our result is not a mere artifact of high-dimensional random matrix generation, we have considered a specific example of an SU(6) transformation corresponding to a probabilistic linear optical CNOT gate [38]. We were able to reproduce this particular unitary with infidelity of $\sim 10^{-8}$ using random mixing layers [27].

To demonstrate that our architecture is resilient to practical constraints and errors that can undermine the universality of the interferometer, we have added random perturbations to the mode mixing layers according to the following procedure. First, we construct the parametrized matrix

$$T_\alpha^{(m)} = (1 - \alpha)V_0^{(m)} + \alpha R, \qquad (5)$$

which is a weighted sum of an initial matrix $V_0^{(m)}$, and a Haar-random perturbation R, modeling imperfections. Since $T_\alpha^{(m)}$ is nonunitary at $0 < \alpha < 1$, we then apply the singular
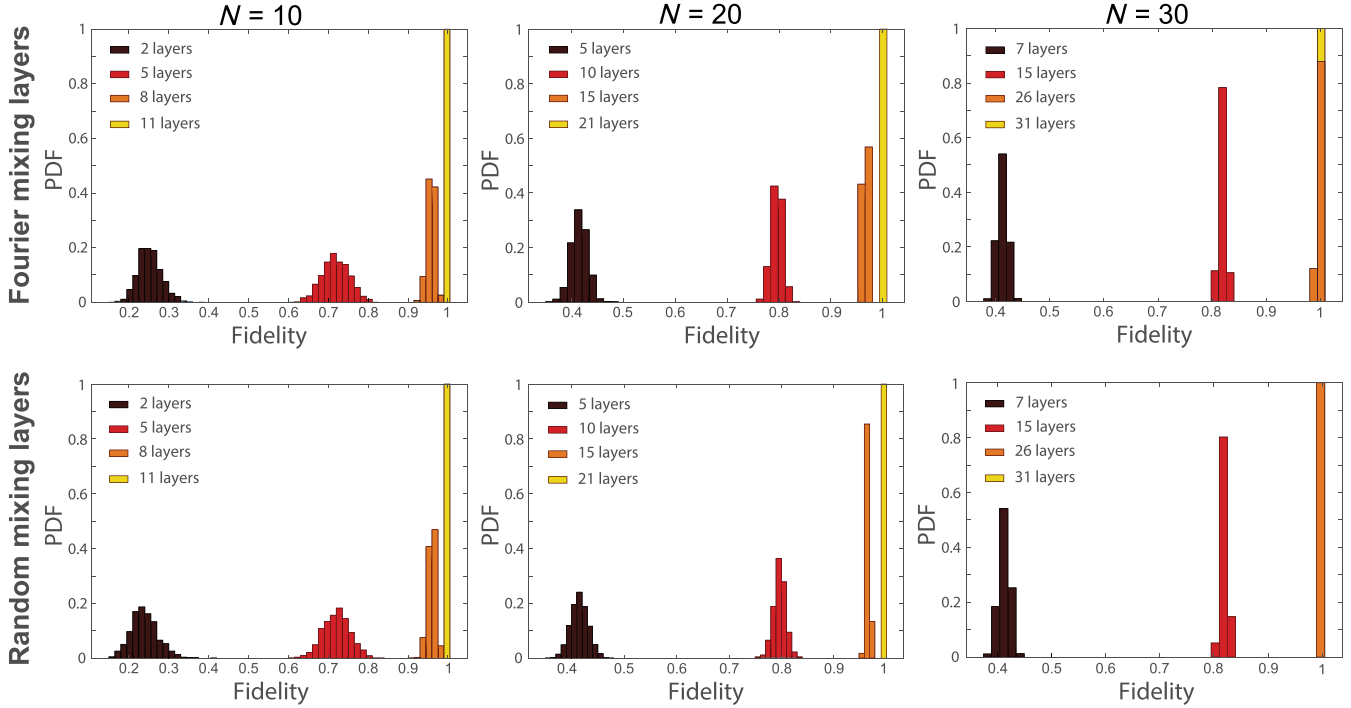
FIG. 3. Fidelity of transformation for the decomposition depicted in Fig. 2(a) for $N = 10, 20$, and 30 and different choice of the mixing layer transfer matrices. The figures of the top row correspond to probability density function (PDF) at transfer matrices $V^{(m)} = U^{(\mathrm{DFT})}$, while the bottom row corresponds to all $V^{(m)}$ taken at random. Each histogram is a collection of 1000 fidelity values, corresponding to randomly sampled target unitary matrix. Different numbers of phase layers were used to illustrate the convergence of the schemes to universal, which correspond to the maximum phase layers of $N + 1$.

value decomposition: $T_\alpha^{(m)} = W_\alpha^{(m,L)} D_\alpha^{(m)} W_\alpha^{(m,R)}$, where $W_\alpha^{(m,L)}$ and $W_\alpha^{(m,R)}$ are unitary matrices and $D_\alpha^{(m)}$ is a diagonal matrix. The unitary part of (5) as $V_\alpha^{(m)} = W_\alpha^{(m,L)} W_\alpha^{(m,R)}$. Even for a rather high value of $\alpha = 1$ we have observed essentially no effect of the perturbation on the infidelity with an ideal CNOT matrix.

There are obvious examples of matrices which are not suitable for mixing layers, such as an identity matrix I or permutation matrices. However, as our numerical results suggest, the relative volume of these matrices is negligible for large N. Parametrized families of matrices $V_\alpha^{(m)}$ allow us to analyze the worst case performance of our scheme. Changing R in (5) with I we can see how close to a worst case we can get before the scheme becomes nonuniversal. For each value of α a series of 300 target matrices U_s and block transfer matrices $V_0^{(m)}$ were generated at random from a Haar-uniform distribution and fidelity (4) was maximized over the space of phase shifts. Figure 4 demonstrates the infidelity of total transformation $1 - F$ as a function of parameter α along with the infidelity for blocks $1 - S$, where $S = \sum_{m=1}^{N} F(V_0^{(m)}, V_\alpha^{(m)})/N$ is fidelity between the initial and the perturbed block transfer matrices. As can be seen from the figure, while S for each block grows monotonically, the dependence of infidelity for the whole transformation has a clear threshold behavior; namely, it

stays on the constant level of accuracy of the numerical algorithm as far as α does not exceed value ≈ 0.5 after which it grows abruptly. We did not study the dependence of the threshold value of α on the interferometer size N in details, but results for $N = 5$ and $N = 10$ suggest, that it may slightly increase with increasing N [27].

Theoretical framework.—The strict proof of the universality implies showing that the transformations $U(\{\vec{\varphi}^{(m)}\})$ should form a group under matrix multiplication and then
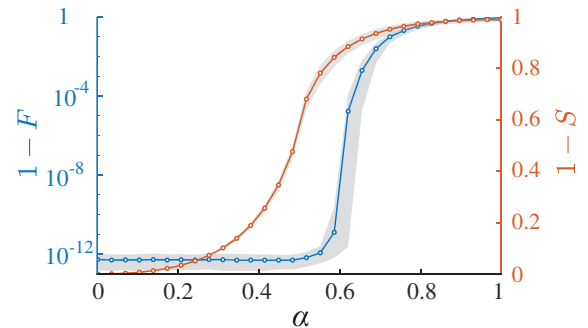

FIG. 4. Infidelity of transformation $1 - F$ and dissimilarity of block transfer matrices $1 - S$, where $S = \sum_{m=1}^{N} F(V_0^{(m)}, V_\alpha^{(m)})/N$, as a function of parameter α at $N = 10$. The circles correspond to the average values; the lower and upper boundaries of the shaded regions are the averages of the 10 best and worst infidelities, respectively.

prove that this group densely covers the $SU(N)$ group of the corresponding dimension. We could not find the rigorous proof neither for the simplest nontrivial case of $N = 3$ nor for the general case of arbitrary dimension. However, we have worked out preliminary considerations on the structure of the manifold described by (2) [27].

Conclusion.—Our work demonstrates that multichannel layered schemes provide an alternative architecture for universal linear-optical unitaries. This architecture is not limited to specific strictly predefined transfer matrices, such as, for example, a discrete Fourier transformation as in previous proposals [35,39]. Namely, we have shown that the transfer matrices of the static blocks can be chosen from a continuous class of unitary matrices at random without sacrificing the quality of approximation for an arbitrary target unitary. This also comes immediately with resilience to errors that usually occur in implementations. In practice, the interferometer manufactured with imperfect mixing layers may be tuned to implement the desired transformation postfactum by tuning the phase-shifts only and optimizing the measured fidelity with the desired transformation [25]. Therefore, these results, showing that one does not necessarily need to carefully engineer the building blocks of the schemes to be able to reach almost any matrix in the unitary space, are of primary importance for experimental applications.

---

[*]saygin@physics.msu.ru

[1] A. Annoni, E. Guglielmi, M. Carminati, G. Ferrari, M. Sampietro, D. A. Miller, A. Melloni, and F. Morichetti, Light Sci. Appl. **6**, e17110 (2017), original article.

[2] Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, and M. Soljacic, Nat. Photonics **11**, 441 (2017).

[3] T. W. Hughes, M. Minkov, Y. Shi, and S. Fan, Optica **5**, 864 (2018).

[4] 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, Optica **5**, 1623 (2018).

[5] T. Rudolph, APL Photonics **2**, 030901 (2017).

[6] A. Politi, J. C. F. Matthews, and J. L. O'Brien, Science **325**, 1221 (2009).

[7] X.-Q. Zhou, P. Kalasuwan, T. C. Ralph, and J. L. O'Brien, Nat. Photonics **7**, 223 (2013).

[8] J. B. Spring, B. J. Metcalf, P. C. Humphreys, W. S. Kolthammer, X.-M. Jin, M. Barbieri, A. Datta, N. Thomas-Peter, N. K. Langford, D. Kundys, J. C. Gates, B. J. Smith, P. G. R. Smith, and I. A. Walmsley, Science **339**, 798 (2013).

[9] M. A. Broome, A. Fedrizzi, S. Rahimi-Keshari, J. Dove, S. Aaronson, T. C. Ralph, and A. G. White, Science **339**, 794 (2013).

[10] A. Crespi, R. Osellame, R. Ramponi, D. J. Brod, E. F. Galvão, N. Spagnolo, C. Vitelli, E. Maiorino, P. Mataloni, and F. Sciarrino, Nat. Photonics **7**, 545 (2013).

[11] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, Nat. Commun. **5**, 4213 (2014).

[12] N. Wiebe and C. Granade, Phys. Rev. Lett. **117**, 010503 (2016).

[13] J. Wang, S. Paesani, R. Santagati, S. Knauer, A. A. Gentile, N. Wiebe, M. Petruzzella, J. L. O'Brien, J. G. Rarity, A. Laing, and M. G. Thompson, Nat. Phys. **13**, 551 (2017).

[14] J. Carolan, C. Harrold, C. Sparrow, E. Martin-Lopez, 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, Science **349**, 711 (2015).

[15] A. Hurwitz, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse **1897**, 71 (1897).

[16] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, Phys. Rev. Lett. **73**, 58 (1994).

[17] W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walsmley, Optica **3**, 1460 (2016).

[18] R. Burgwal, W. R. Clements, D. H. Smith, J. C. Gates, W. S. Kolthammer, J. J. Renema, and I. A. Walmsley, Opt. Express **25**, 28236 (2017).

[19] C. Jarlskog, J. Math. Phys. (N.Y.) **46**, 103508 (2005).

[20] P. Dita, J. Phys. A **15**, 3465 (1982).

[21] P. A. Ivanov and N. V. Vitanov, Phys. Rev. A **77**, 012335 (2008).

[22] T. Tilma and E. C. G. Sudarshan, J. Phys. A **35**, 10467 (2002).

[23] A. Bouland and S. Aaronson, Phys. Rev. A **89**, 062316 (2014).

[24] N. C. Harris, G. R. Steinbrecher, M. Prabhu, Y. Lahini, J. Mower, D. Bunandar, C. Chen, F. N. C. Wong, T. Baehr-Jones, M. Hochberg, S. Lloyd, and D. Englund, Nat. Photonics **11**, 447 (2017).

[25] I. V. Dyakonov, I. A. Pogorelov, I. B. Bobrov, A. A. Kalinkin, S. S. Straupe, S. P. Kulik, P. V. Dyakonov, and S. A. Evlashin, Phys. Rev. Applied **10**, 044048 (2018).

[26] N. J. Russell, L. Chakhmakhchyan, J. L. O'Brien, and A. Laing, New J. Phys. **19**, 033007 (2017).

[27] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.124.010501 for other possible variants of the interferometer schemes, for the analysis of the efficiency of the numerical algorithm runtime and the effect of finite phase shift precision on transformation quality, which includes Ref. [28], for a particular example of a linear-optical CNOT gate and robustness analysis of the interferometer architecture under consideration, for the performance analysis of the interferometer architecture under consideration, and for the analytical results, which includes Refs. [29–32].

[28] S. Pai, B. Bartlett, O. Solgaard, and D. A. B. Miller, Phys. Rev. Applied **11**, 064044 (2019).

[29] D. Melnikov, A. Mironov, S. Mironov, A. Morozov, and A. Morozov, Nucl. Phys. **B926,** 491 (2018).

[30] A. M. Kowalevicz, V. Sharma, E. P. Ippen, J. G. Fujimoto, and K. Minoshima, Opt. Lett. **30,** 1060 (2005).

[31] A. Yariv, IEEE J. Quantum Electron. **9,** 919 (1973).

[32] N. Spagnolo, L. Aparo, C. Vitelli, A. Crespi, R. Ramponi, R. Osellame, P. Mataloni, and F. Sciarrino, Sci. Rep. **2** (2012).

[33] R. Tang, T. Tanemura, and Y. Nakano, IEEE Photonics Technol. Lett. **29,** 971 (2017).

[34] J. Zhou, J. Wu, and Q. Hu, Opt. Express **26,** 3020 (2018).

[35] H.-H. Lu, J. M. Lukens, N. A. Peters, O. D. Odele, D. E. Leaird, A. M. Weiner, and P. Lougovski, Phys. Rev. Lett. **120,** 030502 (2018).

[36] H.-H. Lu, J. M. Lukens, B. P. Williams, P. Imany, N. A. Peters, A. M. Weiner, and P. Lougovski, npj Quantum Inf. **5,** 24 (2019).

[37] The code is available on request.

[38] T. C. Ralph, N. K. Langford, T. B. Bell, and A. G. White, Phys. Rev. A **65,** 062324 (2002).

[39] M. Huhtanen, Linear Algebra Appl. **424,** 304 (2007).