

Optics Letters

Optimal design of error-tolerant reprogrammable multiport interferometers

S. A. FLDZHYAN, M. YU. SAYGIN,* AND S. P. KULIK

Quantum Technologies Center and Faculty of Physics, M. V. Lomonosov Moscow State University, GSP-1, Leninskie gory, Moscow 119991, Russia

*Corresponding author: saygin@physics.msu.ru

Received 12 December 2019; revised 22 March 2020; accepted 30 March 2020; posted 31 March 2020 (Doc. ID 385433); published 30 April 2020

Photonic information processing demands programmable multiport interferometers capable of implementing arbitrary transfer matrices, for which planar meshes of error-sensitive Mach–Zehnder interferometers are usually exploited. We propose an alternative design that uses a single-static beam splitter (BS) and a variable phase shift as the building blocks. The design possesses superior resilience to manufacturing errors and losses without extra elements added into the scheme. Namely, the power transmissivities of the static BSs can take arbitrary values in the range from $\approx 1/2$ to $\approx 4/5$. In this Letter, we show that the fraction of transfer matrices non-implementable by the interferometers of our design diminishes rapidly with its size. © 2020 Optical Society of America

<https://doi.org/10.1364/OL.385433>

Photonics has been progressively playing a more important role in fundamental science and applied fields, motivated by newly developed approaches to information processing that are matched well with the qualities of optics. Linear transformations between multiple optical channels are often required by these approaches, thus, making the utilization of multiport interferometer devices a necessity. For example, multiport interferometers are exploited as mode unscramblers [1] and parts of photonic neural networks [2,3].

In the recent years, optical interferometers have attracted appreciable interest by the quantum information community because of the promising quantum computing platforms that leverage linear optics and unique properties of photon discrete variables [4,5] and field continuous variables states [6,7]. Recent works have demonstrated the versatility of linear-optical quantum systems and their ability to perform quantum computing tasks, ranging from the algorithms suggested at the dawn of quantum information theory [8,9] to more specific ones that disrupt the landscape nowadays, such as the boson-sampling algorithm [10–12] and quantum deep neural networks [13,14].

Universal interferometers can be reprogrammed to implement an arbitrary linear transformation defined by a specific transfer matrix. To construct these interferometers, decomposition methods are used, which represent unitary matrices as products of simpler building blocks [15–17]. Among these

methods, the most practical are planar decompositions, since they well suit fabrication by the mature techniques of integrated photonics enabling massive production of sophisticated optical circuits [2,5,18]. Today, the most often used methods are those proposed by Reck *et al.* [15] and Clements *et al.* [16], which decompose unitary matrices into planar meshes of variable beam splitters (BSs), having triangular and rectangular forms, respectively. In these schemes, each variable BS is conveniently realized by a standard element of the Mach–Zehnder interferometer (MZI), made up of two static balanced BSs with variability provided by two phase shifts. Thus, the overall scheme is reprogrammed by setting the phase shifts [18,19].

For these schemes to be universal, it is crucial that the static BSs should be balanced. However, this condition can not be fully satisfied because of the errors that occur at realization, limiting the scheme's universality [20,21]. The negative effect of the errors progresses as the interferometer size scales up, effectively imposing stringent requirements on the fabrication tolerances and making challenging the realization of large interferometers.

There are methods that can restore the universality of the MZI-based schemes at the cost of adding extra elements into their optical schemes [21,22]. However, the common drawback of these methods is the burden of auxiliary control needed to manipulate the additional MZIs and the increased real estate occupied by the scheme on the chip. Therefore, developing more efficient designs of error-tolerant interferometers is highly demanded nowadays. Here, we propose, to the best of our knowledge, a new design of planar interferometers, which is error-tolerant to manufacturing errors and universal except for a fraction of matrices that rapidly diminishes with their size.

An N -port interferometer may be described by an $N \times N$ transfer matrix, acting on vectors of field amplitudes according to the relation $\mathbf{a}^{(\text{out})} = U\mathbf{a}^{(\text{in})}$, where $\mathbf{a}^{(\text{in})}$ and $\mathbf{a}^{(\text{out})}$ are the input and output vectors, respectively. Provided interferometers are lossless, and their transfer matrices U are unitary.

We first describe interferometers constructed with the canonical MZI-based design [16], depicted in Fig. 1(a). It is formed by N layers consisting of MZIs, each acting locally only on two neighboring channels. In this scheme, the overall number of MZIs is equal to $N(N-1)/2$. Accounting for the $N-1$ phase shifts at the input, the total number of phase shifts in the

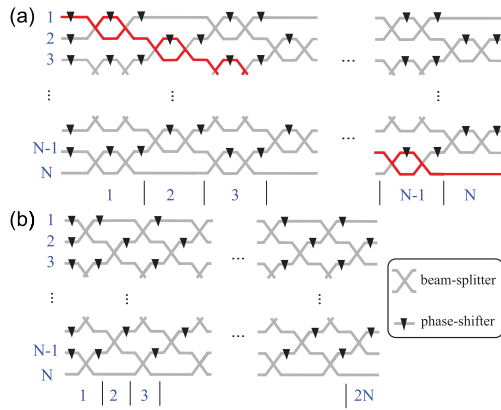


Fig. 1. Two designs of multiport interferometers: (a) conventional universal MZI-based design proposed in [16]. The elements colored in red are utilized in the rerouting operation from the first into the last port. (b) BS-based design proposed in this work.

interferometer is $N^2 - 1$, exactly the number of independent real parameters that parametrize an arbitrary unitary $N \times N$ matrix.

Therefore, the transfer matrix of the MZI-based interferometer can be written as

$$U_{\text{MZI}} = V_{\text{MZI}}^{(N)} \cdots V_{\text{MZI}}^{(2)} V_{\text{MZI}}^{(1)} \Phi, \quad (1)$$

where $V_{\text{MZI}}^{(m)}$ is the transfer matrix of m -th layer, $\Phi = \text{diag}(e^{i\varphi_1^{(\text{out})}}, \dots, e^{i\varphi_{N-1}^{(\text{out})}}, 1)$ is the diagonal matrix with $\varphi_j^{(\text{out})}$ being the phase shifts introduced in the end. In Eq. (1), layer transfer matrices $V_{\text{MZI}}^{(m)}$ are of the form $V_{\text{MZI}}^{(m)} = \prod_{j \in \Omega_m^{\text{MZI}}} T_{\text{MZI},j}^{(m)}(\varphi_{2j-1}, \varphi_{2j})$, where

$$T_{\text{MZI},j}^{(m)}(\varphi_1, \varphi_2) = \begin{pmatrix} 1 & \cdots & \cdots & \cdots & 0 \\ \vdots & \ddots & & & \vdots \\ \vdots & & B^{(m)} & & \vdots \\ \vdots & & & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 1 \end{pmatrix} \quad (2)$$

is the block matrix of a single MZI placed in the m -th layer between channels j and $j+1$, and Ω_m^{MZI} denotes the ordered sequence of MZIs in the layer with index m . Block matrix Eq. (2) has all diagonal elements 1, except those labeled $B_{1,1}^{(m)} = e^{i(\varphi_1^{(m)} + \varphi_2^{(m)})} a_j^{(m)}$ and $B_{2,2}^{(m)} = -e^{i\varphi_1^{(m)}} a_j^{(m)*}$, and all off-diagonal elements equal to 0 except those labeled $B_{1,2}^{(m)} = e^{i\varphi_1^{(m)}} b_j^{(m)}$ and $B_{2,1}^{(m)} = e^{i(\varphi_1^{(m)} + \varphi_2^{(m)})} b_j^{(m)*}$, where we introduced the shorthand notations: $a = \sin \varphi_1^{(m)} \cos(\alpha_1^{(m)} - \alpha_2^{(m)}) + i \cos \varphi_1^{(m)} \sin(\alpha_1^{(m)} + \alpha_2^{(m)})$ and $b = \cos \varphi_1^{(m)} \cos(\alpha_1^{(m)} + \alpha_2^{(m)}) + i \sin \varphi_1^{(m)} \sin(\alpha_1^{(m)} - \alpha_2^{(m)})$. The variable phase shifts φ_j are used to reconfigure the interferometer and have required ranges from 0 to 2π .

When $\alpha_l^{(m)} = 0$, the MZI-based interferometer is capable of implementing an arbitrary unitary transfer matrix; however, imbalances α 's undermine its universality. A trivial example of rerouting from port 1 into port N of an N -port interferometer is shown in Fig. 1(a). Obviously, to attain this transformation, all diagonal MZIs should be in the cross state. However, the unit

transmissivity of an MZI, $\tau = |b|^2$, can be obtained only when $\alpha_1 = \alpha_2 = 0$.

The schematic representation of our design is depicted in Fig. 1(b). Our design has rectangular placement of building blocks, each of which is a single static BS and single tunable phase shifter. Hence, the name BS-based for our design. The N -port interferometer of the BS-based design has $2N$ layers so that both the scheme depth, as quantified by the maximum static BSs crossed by the signals, and the number of phase shifts is equal to the one of the MZI-based design. The interferometer transfer matrix takes the form

$$U_{\text{BS}} = V_{\text{BS}}^{(2N)} \cdots V_{\text{BS}}^{(2)} V_{\text{BS}}^{(1)} \Phi, \quad (3)$$

in which layer transfer matrices $V_{\text{BS}}^{(m)} = \prod_{j \in \Omega_m^{\text{BS}}} T_{\text{BS},j}^{(m)}(\varphi_j^{(m)})$ have the block $T_{\text{BS}}^{(m)}(\varphi^{(m)})$ of form Eq. (2), but with $B_{1,1}^{(m)} = e^{i\varphi^{(m)}} \cos(\theta_0 + \alpha^{(m)})$, $B_{2,2}^{(m)} = \cos(\theta_0 + \alpha^{(m)})$, $B_{1,2}^{(m)} = \sin(\theta_0 + \alpha^{(m)})$, and $B_{2,1}^{(m)} = -e^{i\varphi^{(m)}} \sin(\theta_0 + \alpha^{(m)})$. Here, angle θ_0 quantifies the transmission of the static BSs, which specific value will be given below.

We consider errors as $\alpha_j \neq 0$ that tune the splitting ratios of the static BSs off required values. First, we study the effect of coherent errors at which $\alpha_j = \alpha$. For the schemes manufactured by planar lithography techniques, this type of errors is linked to the variations of waveguide's material and geometry, which is dominated as their spatial scale is usually large compared with the area occupied by the scheme [23]. These arguments can also be applied to interferometers manufactured by other methods, for example, femtosecond direct laser writing [19], as well as implementation approaches exploiting repetitively few optical elements to obtain the desired transformation [24].

We evaluate the performance of multiport interferometers by calculating the fidelity, defined as

$$F = \frac{|\text{Tr}(U^\dagger U_0)|^2}{N \text{Tr}(U^\dagger U)}, \quad (4)$$

which compares the target unitary matrix U_0 and the actual transfer matrix U realized by the interferometer, where N is the size of the matrices. Provided that the matrices U and U_0 are equal up to a complex multiplier, the fidelity Eq. (4) gets its maximum value of $F = 1$.

Generally, no analytical solution is known to derive phase shifts that maximize the fidelity Eq. (4), except for the case of error-free MZI-based interferometers, for which an analytical procedure is provided in [16]. Unfortunately, we could not find analogous procedure for the error-free BS-based interferometers.

We used a numerical optimization algorithm based on the basin-hopping algorithm. Given a unitary matrix U_0 , the algorithm was searching for a global minimum of infidelity $1 - F$ over the space of phase shifts. To decrease the chance of falling into local minima, we used multiple runs of the optimization with random initial values of the phases. Each numerical experiment involved optimization over a series of target matrices U_0 , drawn from the Haar random distribution using the method based on the QR-decomposition of random matrices [25]. With this algorithm, it took several hours to find optimal phase shifts for a single 10×10 transfer matrix, so that a multicore computer has been utilized to derive required dependencies.

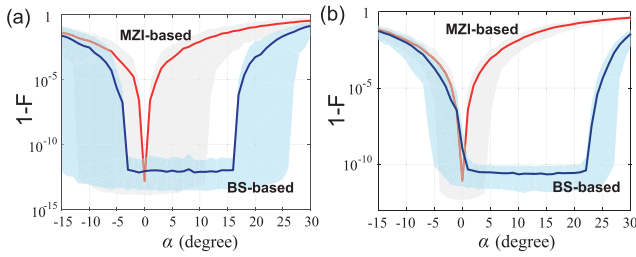


Fig. 2. Infidelity $1 - F$ as a function of error parameter α for the MZI-based and BS-based interferometers at $\theta_0 = \pi/4$ and (a) $N = 5$ and (b) $N = 10$. For each value of α , the infidelity distribution was obtained numerically using a set of 300 unitary matrices drawn randomly from uniform distribution. The solid curves correspond to the average over all samples size; the lower and upper boundaries of the shaded regions are averages for 10 infidelities with the lowest and highest values, respectively.

We understand that more efficient numerical algorithms can be developed for this specific task [26].

The obtained infidelity $1 - F$ as a function of the error parameter α is plotted in Fig. 2. The finite accuracy of the numerical algorithm sets the minimal infidelity value of $\sim 10^{-12} - 10^{-10}$, which could not be overcome neither for the MZI-based interferometer with $\alpha = 0$ where exact zero was expected. The MZI-based interferometers are equally sensitive to both positive and negative values of α with the acceptable range of errors being of the order of several degrees. For the BS-based interferometers, infidelity behaves radically different: while at $\alpha < 0$, the performance of the two are comparable, at $\alpha > 0$, the BS-based interferometers provide perfect fidelity for α as large as ~ 20 degrees, i.e., several times larger than for the MZI-based interferometers. Figure 2 suggests that when the positive and negative values of α are equiprobable, the optimal choice of the static BSs defined at the design stage is such that $\theta_0 \approx 55$ degree, corresponding to the center of the high-fidelity plateau.

The superior performance of the BS-based interferometers at positive α 's can be attributed to the interplay of two competing properties. On the one hand, more transmissive BSs with $\alpha > 0$ enable more efficient travelling of the signal amplitudes across the scheme, which has been recognized as a prerequisite for robustness to errors in other universal interferometer architectures [17]. On the other hand, the increase of the transmissivity reduces the BS interaction, which is completely absent in the limiting case of $\alpha = \pi/4$.

Second, we consider the incoherent errors at which α 's are distributed at random across the scheme. The obtained infidelities are shown in Fig. 3. Clearly, the BS-based interferometer design is ultimately tolerant to the incoherent errors.

The performed analysis cannot be complete in proving strict universality of the interferometers, since random generation of target matrices can overlook small subsets that are non-implementable error-tolerantly or non-implementable at all. We now show that such matrices do exist for the BS-based design. For this, we generate random matrices of size $N = 3$ and calculate infidelities. Figure 4(a) demonstrates that in this low-dimensional case, the relative volume occupied by the non-implementable matrices is quite large to be caught by random sampling: $\approx 6\%$ of them satisfy the condition $1 - F > 10^{-5}$ at $\theta_0 = \pi/4$, $\alpha = 0$. However, taking the results

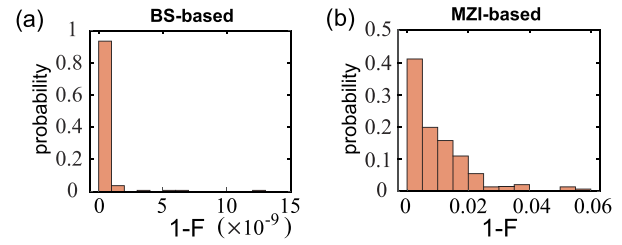


Fig. 3. Normalized histogram of infidelity $1 - F$ for the BS-based (a) and MZI-based (b) 10-port interferometers at random errors. The error angles α_j were drawn from the distribution $p(\alpha) = \exp(-\alpha^2/2\Delta^2)/\sqrt{2\pi}\Delta$ with $\Delta = 10$ deg. For the BS-based interferometer, the parameter $\theta_0 = 55$ deg, roughly corresponding to the center of the high-fidelity plateau, depicted in Fig. 2(b). The histograms are the result of the optimization of 300 randomly sampled unitary matrices.

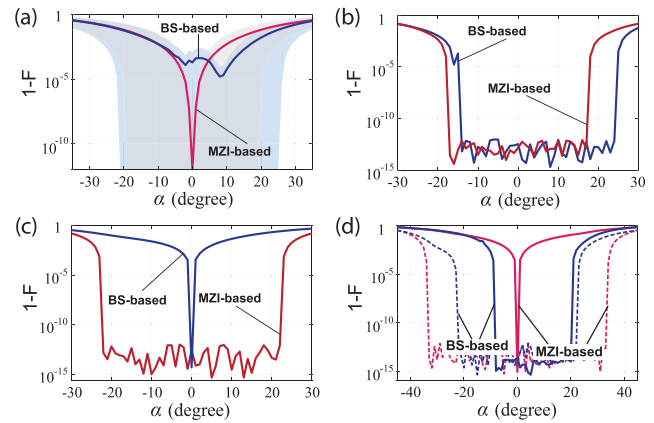


Fig. 4. Infidelity $1 - F$ as a function of error parameter α for the BS- and MZI-based interferometers implementing (a) 300 randomly generated unitary U_0 at $N = 3$; the lower and upper boundaries of the shaded regions are averages for 10 infidelities with the lowest and highest values, respectively. (b) DFT matrix at $N = 3$. (c) Swap between ports 1 and 2 at $N = 3$. (d) Hadamard transformation H_2 of ports 1 and 2 at $N = 3$ (dashed curves) and the Hadamard transformation H_4 at $N = 4$ (solid curves).

of Fig. 2 into account, where not a single non-implementable matrix was present, we conclude that the relative volume of non-implementable matrices rapidly diminishes with N .

Also, we consider some concrete examples that fall out of the picture obtained with random sampling. One example is of the discrete Fourier transform (DFT) matrix with elements $U_{mn}^{(\text{DFT})} = \exp(i2\pi(m-1)(n-1)/N)/\sqrt{N}$. Figure 4(b) illustrates the results for $N = 3$, showing that both designs are equally tolerant, which is actually true for larger N .

We next consider a block-diagonal matrix of size $N = 3$ that comprises the swap block of ports 1 and 2 [Fig. 4(c)]. In contrast to previous results, the MZI-based design works much better than the BS-based design. Moreover, we have found that the block-diagonal matrices of arbitrary sizes are not reproduced error-tolerantly by the BS-based schemes; thus, it represents a class of transformations where the MZI-based design is advantageous.

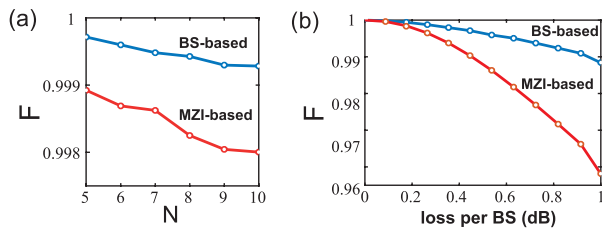


Fig. 5. (a) Average fidelity at a loss of 0.2 dB per static BS for different interferometer sizes N . (b) Average fidelity of a 10-port interferometer as a function of loss introduced by each static BS. In the BS-based design, $\theta_0 = 55$ deg, while BSs in the MZI-based design are balanced. Each average was calculated over 50 randomly sampled matrices.

Finally, we consider the Hadamard transformation. Namely, matrix $H_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ as a constituent upper-left block of a 3-by-3 matrix and matrix $H_4 = H_2 \otimes H_2$ as a whole 4-by-4 matrix. As can be seen from Fig. 4(d), both designs are good at implementing the transformation when it is a part of a larger matrix, while the BS-based design is much better in case of a whole matrix. Once again, this is the evidence that the block-diagonal matrices are better realized by the MZI-based than BS-based interferometers. In addition, we have found the permutation matrices are also better implemented by the MZI-based interferometers rather than the BS-based ones. In particular, rerouting operation depicted in Fig. 1(a) cannot be implemented perfectly by the BS-based interferometer.

In addition, we study the effect of unbalanced losses that occur because different paths through the interferometer experience different losses, which can result in poor fidelity. Typically, static BSs introduce additional losses due to waveguide bending and scattering; therefore, in both designs, unbalanced losses occur when the signals have to pass through side paths, as they contain less BSs than the inner paths. To model lossy BSs, each $B_{i,j}^{(m)}$ in the MZI- and BS-based transfer matrices was multiplied by t^2 and t , respectively, where $0 \leq t \leq 1$ is the BS transmission coefficient. Then, random target matrices were generated, and the global maxima of fidelity Eq. (4) were calculated. The obtained results are shown in Fig. 5, suggesting that the BS-based design is more loss-balanced than the MZI-based design. This can be attributed to more evenly distributed static BSs in the scheme than in the MZI-based design, where BSs are grouped in pairs. We understand that transformations might exist where our design loses the advantage; however, we have not found any outside of the set of non-implementable examples described above.

In summary, to the best of our knowledge, we propose a novel design of programmable multiport interferometers, which exhibit superior tolerance to errors and losses compared to the previously known counterparts, while not requiring redundant elements. It is noteworthy that the proposed design is less complex than the counterparts, with possible implementation by a variety of experimental platforms. For example, the static interferometers that have been used in experiments on boson sampling [12,27] have suitable placement of the passive BS elements, yet lacking programmability.

Funding. Russian Foundation for Basic Research (19-52-80034).

Disclosures. The authors declare no conflicts of interest.

REFERENCES

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).
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. T. Rudolph, *APL Photon.* **2**, 030901 (2017).
5. 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).
6. F. Lenzini, J. Janousek, O. Thearle, M. Villa, B. Haylock, S. Kasture, L. Cui, H.-P. Phan, D. V. Dao, H. Yonezawa, P. K. Lam, E. H. Huntington, and M. Lobino, *Sci. Adv.* **4**, eaat9331 (2018).
7. A. P. Lund, S. Rahimi-Keshari, and T. C. Ralph, *Phys. Rev. A* **96**, 022301 (2017).
8. A. Politi, J. C. F. Matthews, and J. L. O'Brien, *Science* **325**, 1221 (2009).
9. X.-Q. Zhou, P. Kalasuwan, T. C. Ralph, and J. L. O'Brien, *Nat. Photonics* **7**, 223 (2013).
10. 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).
11. 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).
12. H.-S. Zhong, Y. Li, W. Li, L.-C. Peng, Z.-E. Su, Y. Hu, Y.-M. He, X. Ding, W. Zhang, H. Li, L. Zhang, Z. Wang, L. You, X.-L. Wang, X. Jiang, L. Li, Y.-A. Chen, N.-L. Liu, C.-Y. Lu, and J.-W. Pan, *Phys. Rev. Lett.* **121**, 250505 (2018).
13. G. R. Steinbrecher, J. P. Olson, D. Englund, and J. Carolan, *NPJ Quantum Inf.* **5**, 60 (2019).
14. N. Killoran, T. R. Bromley, J. M. Arrazola, M. Schuld, N. Quesada, and S. Lloyd, "Continuous-variable quantum neural networks," arXiv:1806.06871 (2018).
15. M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, *Phys. Rev. Lett.* **73**, 58 (1994).
16. W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley, *Optica* **3**, 1460 (2016).
17. M. Y. Saygin, I. V. Kondratyev, I. V. Dyakonov, S. A. Mironov, S. S. Straupe, and S. P. Kulik, *Phys. Rev. Lett.* **124**, 010501 (2020).
18. C. Taballione, T. A. W. Wolterink, J. Lugani, A. Eckstein, B. A. Bell, R. Grootjans, I. Visscher, J. J. Renema, D. Geskus, C. G. H. Roeloffzen, I. A. Walmsley, P. W. H. Pinkse, and K.-J. Boller, *Frontiers in Optics/Laser Science* (Optical Society of America, 2018), paper JTU3A.58.
19. 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. Appl.* **10**, 044048 (2018).
20. J. Mower, N. C. Harris, G. R. Steinbrecher, Y. Lahini, and D. Englund, *Phys. Rev. A* **92**, 032322 (2015).
21. 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).
22. D. A. B. Miller, *Optica* **2**, 747 (2015).
23. W. Bogaerts and L. Chrostowski, *Laser Photon. Rev.* **12**, 1700237 (2018).
24. K. R. Motes, A. Gilchrist, J. P. Dowling, and P. P. Rohde, *Phys. Rev. Lett.* **113**, 120501 (2014).
25. F. Mezzadri, "How to generate random matrices from the classical compact groups," arXiv:math-ph/0609050 (2006).
26. S. Pai, B. Bartlett, O. Solgaard, and D. A. B. Miller, *Phys. Rev. Appl.* **11**, 064044 (2019).
27. A. Crespi, R. Osellame, R. Ramponi, V. Giovannetti, R. Fazio, L. Sansoni, F. De Nicola, F. Sciarrino, and P. Mataloni, *Nat. Photonics* **7**, 322 (2013).