Global calibration of large-scale photonic integrated circuits

Jin-Hao Zheng,[1, 2] Qin-Qin Wang,[1, 3] Lan-Tian Feng,[1, 3, 4, *] Yu-Yang Ding,[5] Xiao-Ye
Xu,[1, 3, 4, †] Xi-Feng Ren,[1, 3, 4] Chuan-Feng Li,[1, 3, 4, ‡] and Guang-Can Guo[1, 3, 4]

[1]*CAS Key Laboratory of Quantum Information, University of Science and Technology of China, Hefei 230026, China*
[2]*School of the Gifted Young, University of Science and Technology of China, Hefei 230026, China*
[3]*CAS Center for Excellence in Quantum Information and Quantum Physics,*
University of Science and Technology of China, Hefei 230026, China
[4]*Hefei National Laboratory, University of Science and Technology of China, Hefei 230088, China*
[5]*Hefei Guizhen Chip Technologies Co., Ltd., Hefei 230000, China*
(Dated: October 12, 2024)

The advancing maturity of photonic integrated circuit (PIC) fabrication technology enables the high integration of an increasing number of optical components onto a single chip. With the incremental circuit complexity, the calibration of active phase shifters in a large-scale PIC becomes a crucially important issue. The traditional one-by-one calibration techniques encounter significant hurdles with the propagation of calibration errors, and achieving the decoupling of all phase shifters for independent calibration is not straightforward. To address this issue, we propose a global calibration approach for large-scale PIC. Our method utilizes a custom network to simultaneously learn the nonlinear phase-current relations for all thermo-optic phase shifters on the PIC by minimizing the negative likelihood of the measurement datasets. Moreover, the reflectivities of all static beamsplitter components can also be synchronizedly extracted using this calibration method. As an example, a quantum walk PIC with a circuit depth of 12 is calibrated, and a programmable discrete-time quantum walk is experimentally demonstrated. These results will greatly benefit the applications of large-scale PICs in photonic quantum information processing.

I. INTRODUCTION

The diverse applications of photonic integrated circuits (PICs)[1] in the field of classical and quantum information processing[2–9] have recently experienced significant growth, as they offer key attributes such as compactness, stability and scalability[10]. In addition, PICs can be programmable to reconfigure the chip's functionality on demand, making them more versatile and flexible than application-specific static circuits[11]. The programmability feature of PICs is typically achieved by actively controlling the tunable phase shifters. Controlled phase shifts have been introduced in many ways[12], and the most commonly adopted approach is the thermo-optic effect[13], which utilizes electrically driven heaters to reversibly modify the refractive index of waveguides to impose phase shifts. So far, a variety of experiments based on the programmable PICs have been reported[14–18]. However, as the increasing complexity of practical information processing tasks demands large-scale PICs, it becomes challenging to calibrate the significant number of active phase shifters on the chip.

The traditional calibration method requires the decoupling of all phase shifters followed by individual calibration for each one[19–23]. However, this approach isn't ideal for large-scale PICs due to two main challenges. Firstly, while the traditional method can be used to calibrate the standard triangular "Rech" PICs[24, 25], rect-angular "Clements" PICs[26–29], and PICs with cascaded phase shifters[30, 31] by following a specific decoupling order, it is difficult to find an effective and direct way to decouple all phase shifters in the intricate mesh structures. Secondly, deviations in the calibration values of phase shifters in preceding levels may also spread errors in the calibration of subsequent phase shifters. As a result, calibration errors accumulate progressively with the increasing number of phase shifters, eventually becoming significant and non-negligible.

To address these issues, we present an efficient calibration method developed to globally calibrate all optical components on the PICs, thereby mitigating the accumulation of errors inherent in the traditional one-by-one calibration approach. The global calibration method utilizes a custom network with simulation layers that creates a virtual duplicate of the physical PICs, modelling their practical non-unitary processes involving optical loss. Using the gradient-descent approach, we can determine the optimal values of parameters that need to be calibrated, ensuring that the outputs of the network closely resemble the distributions of measurement data. As an example, on a large-scale quantum walk (QW) PIC with a circuit depth of 12, we find that the trained custom network enables the simultaneous extraction of not only the nonlinear phase-current relations of all the thermo-optic phase shifters but also the reflectivities of all the beamsplitters. Moreover, we further experimentally demonstrate the silicon PIC calibration and implement the on-chip QW[32–40].

This article is organized as follows: In Sec. II, we will introduce our custom-architecture global calibration method, provide a concise overview of the QW PIC, and

* fenglt@ustc.edu.cn
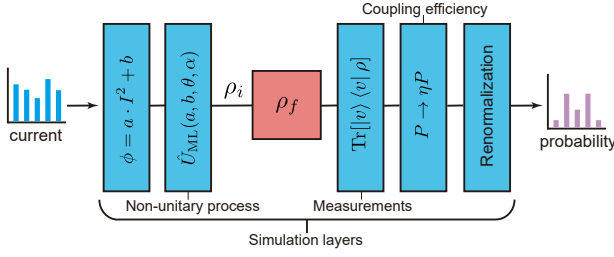† xuxiaoye@ustc.edu.cn
‡ cfli@ustc.edu.cn

FIG. 1. **Illustration of the custom architecture for global calibration.** The input data consists of the current distribution applied to the thermo-optic phase shifters. The first simulation layer of our custom network transforms the distribution of the current $\boldsymbol{I}$ into the distribution of the phase shift $\boldsymbol{\phi}$, with a subsequent layer outputting the parameterized final state ρ_f for a given initial state ρ_i. The third simulation layer provides the corresponding measurement probability distribution of the final state $\mathrm{Tr}[|v\rangle\langle v|\,\rho_f]$ in the computational basis $|v\rangle$. The fourth simulation layer introduces a parameter $\boldsymbol{\eta}$ into this probability distribution, which accounts for the relative fiber-chip coupling efficiency between output ports. The last layer renormalizes the probability distribution so that the sum of the probabilities over all output ports equals one, and a parameterized network distribution $\boldsymbol{P}_{\mathrm{ML}}(\boldsymbol{\lambda})$ is generated thereafter.

evaluate the performance of our calibration method using synthetic datasets featuring Gaussian noise. In Sec. III, we will showcase the calibration results of a silicon PIC utilizing noisy experimental datasets and discuss the performance regarding the experimental implementation of QW. Finally, Sec. IV presents the summary of our work and engages in a discussion.

II. THEORETICAL IDEA

In this section, we will detail our method for achieving precise and fast calibration of PICs. Our method is based on a custom network with simulation layers situated between the input and output layers, as depicted in Fig. 1. The entire custom network is trained using gradient-based optimization techniques to align its virtual process with the real physical process of the PICs, thereby allowing it to estimate the chip parameters that need to be calibrated.

A. Global calibration method

PICs for linear optics comprise a cascade of passive beamsplitters and active phase shifters arranged in various mesh structures [15]. The 2×2 unitary matrix for each on-chip beamsplitter can be expressed as:

$$\hat{U}_{\mathrm{BS}} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \qquad (1)$$

where $\sin^2\theta$ denotes the reflectivity of the beamsplitter. Considering the inevitable optical loss in a practical chip, the transfer matrix for a lossy BS becomes non-unitary and is given by [41]:

$$\hat{U}_{\mathrm{LBS}} = \begin{pmatrix} \sqrt{\alpha} & 0 \\ 0 & \sqrt{\alpha} \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \qquad (2)$$

where α denotes the transfer efficiency of a single BS, and $1 - \alpha$ represents its optical loss. The phase shifter provides a relative phase difference ϕ between the two output paths of the beamsplitter, and its matrix representation is $\hat{U}_{\mathrm{PS}} = \mathrm{diag}\{1, e^{i\phi}\}$. For the widely-used thermo-optic phase shifter, the relative phase ϕ can be tuned by the current I applied to the phase shifter, according to the following formula [13]:

$$\phi = aI^2 + b, \qquad (3)$$

where a and b are hardware parameters to be determined. Then, the non-unitary process $\hat{U}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\alpha})$ of PICs is the matrix product of its individual components. Moreover, considering the mode-field diameter mismatch between the output ports of the PICs and the optical fiber arrays, it is necessary to introduce another parameter, $\boldsymbol{\eta}$, to indicate the fiber-chip coupling efficiency. Consequently, $\boldsymbol{\lambda} = \{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\eta}\}$ are all undetermined parameters for a linear photonic circuit.

In contrast to the traditional one-by-one localized calibration method [19–23], we aim to globally calibrate all optical components on the PICs to reduce the accumulation of errors and avoid complex algorithms for decoupling the phase shifters. To achieve this, we utilize a custom network with simulation layers that builds a virtual duplicate of the physical PICs, as illustrated in Fig. 1. The simulation layers of the network leverage prior knowledge of the structure of the PICs to be calibrated. As a result, by utilizing the probability distributions of output ports, global calibration of the PICs could be implemented, thereby mitigating the high demand for complex quantum state and process tomography.

The learning process necessitates a number of training data, comprising randomly generated current distributions $\{\boldsymbol{I}\}_n$ applied to all the phase shifters and the target probability distributions $\{\boldsymbol{P}_{\mathrm{tar}}\}_n$ of the output ports on the chip. The index $n = 1, \cdots, N$ represents each training sample, and N is the total number of training datasets. We set the current distribution $\boldsymbol{I}$ as inputs, which are then fed to the input layer of our custom network. The custom-network-parameterized non-unitary process $\hat{U}_{\mathrm{ML}}(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\alpha})$ can be determined using Eqs. (2) and (3). For a given initial state ρ_i, the final state is given by $\rho_f = \hat{U}_{\mathrm{ML}}\rho_i\hat{U}_{\mathrm{ML}}^{\dagger}$, and the probability distribution $\mathrm{Tr}[|v\rangle\langle v|\,\rho_f]$ can be obtained. Considering the relative fiber-coupling efficiency $\boldsymbol{\eta}$ between the output ports of the PICs, the final probability distribution is normalized, and then the parameterized network distribution $\boldsymbol{P}_{\mathrm{ML}}(\boldsymbol{\lambda})$ is generated.

We define the cost function $\mathcal{L}$ as the summation of the L1-norm distance [42] between the target probability
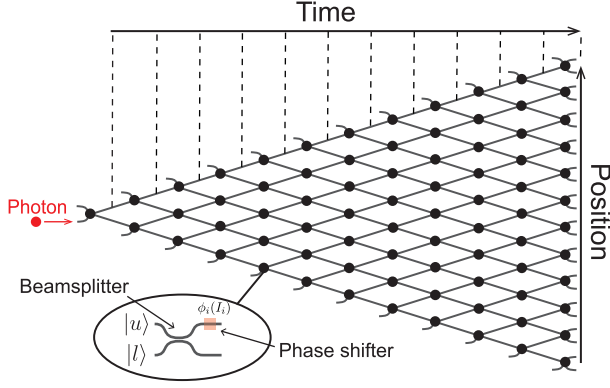
FIG. 2. **Conceptual diagram of the QW PIC.** Each black circle means a beamsplitter element followed by a tuneable thermo-optic phase shifter (see inset), and black solid lines represent the waveguides. At the last discrete time, the column of black circles represents only a beamsplitter without a following phase shifter, as it does not affect the output probability distribution.

distributions in the training datasets $\{\boldsymbol{P}_{\text{tar}}\}_n$ and the parameterized network distributions $\{\boldsymbol{P}_{\text{ML}}(\boldsymbol{\lambda})\}_n$: $\mathcal{L}(\boldsymbol{\lambda}) = \sum_n \frac{1}{2}|\boldsymbol{P}_{\text{ML}}(\boldsymbol{\lambda}) - \boldsymbol{P}_{\text{tar}}|$. The aim of training the custom network is to find the optimal parameters $\boldsymbol{\lambda}^*$ by minimizing the cost function such that the parameterized network distributions approximate the target ones. Then, the parameterized non-unitary process $\hat{U}_{\text{ML}}$ together with the relative coupling efficiency $\boldsymbol{\eta}^*$ for the trained network can well reflect the target physical process that generates the distributions $\{\boldsymbol{P}_{\text{ML}}(\boldsymbol{\lambda}^*)\}_n \simeq \{\boldsymbol{P}_{\text{tar}}\}_n$.

B. Model of on-chip quantum walks

Herein, we consider a QW PIC with a circuit depth of 12, illustrated in Fig. 2, as an example to validate our global calibration method. We note that our method can also be applied to other mainstream types of PICs implementing universal unitary operations, such as the triangular "Rech" and rectangular "Clements" PICs [43]. The QW PIC is calibration hard and shows difficulty in decoupling all phase shifters. The whole physical process of the QW PIC can be mapped onto the discrete-time QW on a one-dimensional (1D) line [44–46]. The discrete-time evolution of a lossy QW can be described by the non-unitary process $\hat{U} = \prod_{t=1}^{T}(\hat{S}\hat{M}\hat{C})$. Here, T denotes the number of discrete-time steps, $\hat{M}$ is the loss operator, $\hat{S}$ is the shift operator, which signifies the movement of the walker, and $\hat{C}$ is the coin operator, representing the coin-tossing process. In the tensor product basis $|v\rangle = |x\rangle \otimes |c\rangle$, where $|x\rangle$ ($x \in \mathbb{Z}$) and $|c\rangle$ ($c = \downarrow, \uparrow$) express the position and coin space, respectively, $\hat{C}$, $\hat{M}$, and $\hat{S}$ can be written as [47]:

$$\hat{C} = \sum_x |x\rangle \langle x| \otimes \begin{pmatrix} \cos\theta & e^{i\gamma}\sin\theta \\ -e^{i\beta}\sin\theta & e^{i(\gamma+\beta)}\cos\theta \end{pmatrix}, \quad (4)$$

$$\hat{M} = \sum_x |x\rangle \langle x| \otimes \begin{pmatrix} \sqrt{\alpha} & 0 \\ 0 & \sqrt{\alpha} \end{pmatrix}, \quad (5)$$

$$\hat{S} = \sum_x (|x-1\rangle \langle x| \otimes |\downarrow\rangle \langle\downarrow| + |x+1\rangle \langle x| \otimes |\uparrow\rangle \langle\uparrow|). \quad (6)$$

Here, the parameter α determines the loss strength and varies depending on both the position and time step. $\cos^2\theta$ and $\sin^2\theta$ denote the probabilities of a walker moving backwards and forwards on the 1D line, respectively. Parameters β and γ introduce the relative phase difference. The coin operator $\hat{C}$ can be position- and time-dependent, a feature necessary for specific tasks in QW-based quantum simulations [35] and quantum algorithms [14].

For the QW PIC, the coin state $\{|\uparrow\rangle, |\downarrow\rangle\}$ in the QW model can be cast on the path state (inset of Fig. 2): the upper path of the beamsplitter corresponds to $|u\rangle$, while the lower path corresponds to $|l\rangle$. The localized coin state at each position is described by $|c\rangle = c_1|u\rangle + c_2|l\rangle$, which can be interpreted as a superposition state located on the two paths of the corresponding beamsplitter. Each beamsplitter with a splitting ratio of 50:50 and phase shifter work together (black circle in Fig. 2) to realize the coin operator $\hat{C}$ in the sense of path states, corresponding to Eq. (4) with $\theta = \pi/4$ and $\beta = 0$. The interwoven mesh architecture of the PIC, composed of the on-chip waveguides (black solid lines), realizes the shift operators $\hat{S}$. After T-step non-unitary dynamics of the QW PIC, the size of the position space is T, with each position having two paths representing the coin space. Therefore, the total number of output ports of the QW PIC is $2T$.

C. Numerical benchmarks

Before training the custom network with noisy experimental results, we begin by numerically benchmarking the global calibration method using synthetic datasets. To effectively simulate noisy experimental environments and assess the robustness of our calibration method against external noise, we introduce Gaussian noise into the data generation process for the synthetic datasets [48]. Here, the Gaussian noise is added to the output probabilities to mimic experimental measurement errors, such as small vibrations of the optical fibers coupled to the output ports of the PICs. The noise is sampled from a zero-mean Gaussian distribution $\mathcal{N}(0, \epsilon)$ with a standard deviation of $\epsilon = 0.01$ defining the noise degree. Then, the calibration process aims to minimize the cost function $\mathcal{L}(\boldsymbol{\lambda}) = \sum_n \frac{1}{2}|\boldsymbol{P}_{\text{ML}}(\boldsymbol{\lambda}) - \boldsymbol{P}_{\text{th}}|$ for given datasets of noisy synthetic probability distributions $\{\boldsymbol{P}_{\text{th}}\}_n$.

The QW PIC comprises 66 phase shifters, 78 beamsplitters, and 24 input/output ports. The initial state is fixed to be a localized state $\rho_i = |\psi_i\rangle\langle\psi_i|$ with $|\psi_i\rangle = |x_0\rangle \otimes |l\rangle$, corresponding to the lower path of the first waveguide. The target values for the reflectivity of the
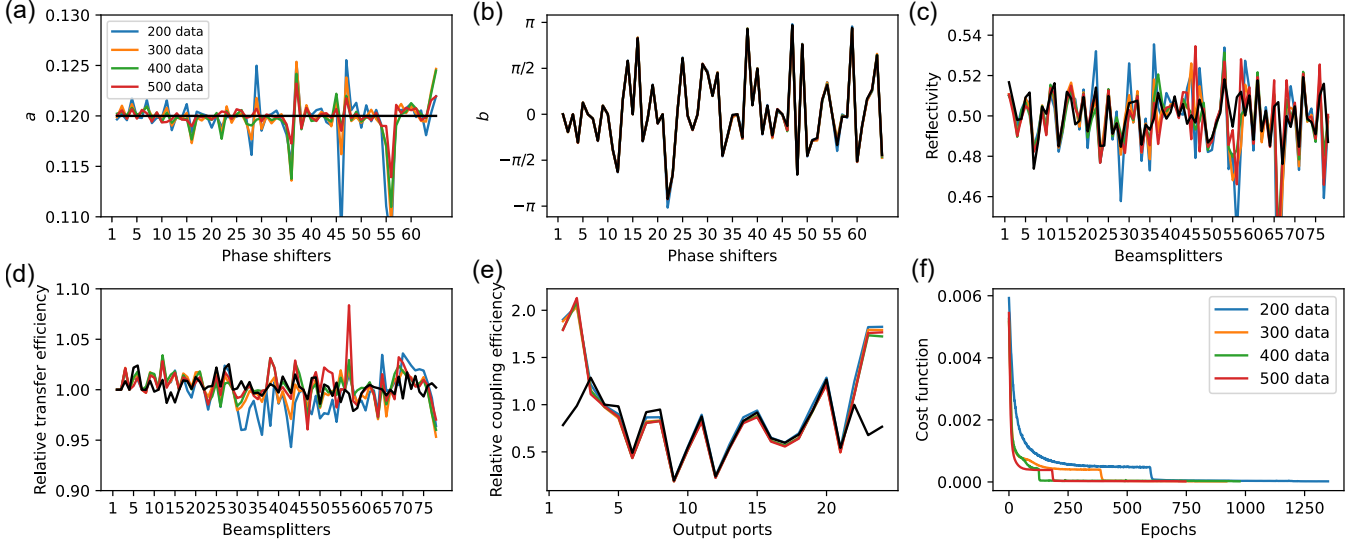
FIG. 3. **Calibration results using synthetic datasets.** (a)∼(e) exhibit the calibration values of a, b, $\sin^2 \theta$, α, and η, respectively, given different numbers of training samples. Black solid lines in (a-e) are the preset target values in theory. The indices of the phase shifters and beamsplitters on the QW PIC are arranged in order from the first column on the far left to the last column, with each column arranged from top to bottom. The indices of the output ports are ordered from top to bottom. The significant deviations in the relative coupling efficiencies η of the edge output ports at the top and bottom positions are attributed to the extremely low output probability (nearly 10^{-4}) at these four ports. (f) The values of the cost function versus the number of epochs.
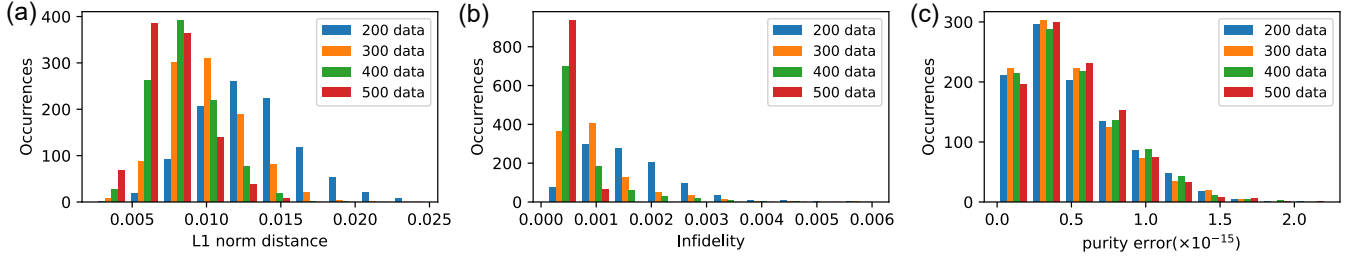

FIG. 4. **Benchmarking the performance of the global calibration method.** Histogram of the values of (a) L1-norm distance, (b) infidelity and (c) purity error for 1000 test samples, given different numbers of training samples.

BSs $\sin^2 \theta$, the relative transfer efficiency α between BSs, and the relative output coupling efficiency η between output ports are set randomly near their ideal values of 0.5, 1, and 1, respectively. The target values for the coefficient a are set to an empirical value of 0.12, while the values for b are randomly selected from the interval $[-\pi, \pi]$. The cost function $\mathcal{L}(\boldsymbol{\lambda})$ is minimized with Adam optimizer that updates the network parameters $\boldsymbol{\lambda} = \{a, b, \theta, \alpha, \eta\}$ for each epoch. The Adam optimizer is configured as follows[49]: the initial learning rate is 0.01 and decreases by multiplication with a factor of 0.99 after each epoch. All other core parameters of Adam are maintained at their default values.

During the iteration process, simultaneously updating all network parameters using the gradient-descent method is susceptible to getting trapped in a local minimum of $\mathcal{L}$. To mitigate this issue and escape local minima, we utilize a non-simultaneous parameter update method, updating only partial parameters at each epoch while keeping the remaining parameters fixed. The relative coupling efficiency η is configured to be constantly updated. The reflectivity parameter θ and the relative transfer efficiency α are frozen at first. We then alternatively train one of the two parameters $\{a, b\}$ while the remaining one is kept frozen and untrainable. Within a single alternating training process, the training for each parameter ceases upon meeting the cut-off condition, where five consecutive decreases in the loss function between adjacent epochs are each less than 10^{-5}. After eight rounds of alternating training, we proceed to train the last two parameters, θ and α, while maintaining the fixed values of a and b. In the end, following the completion of the non-simultaneous parameter updates described above, we proceed to simultaneously update
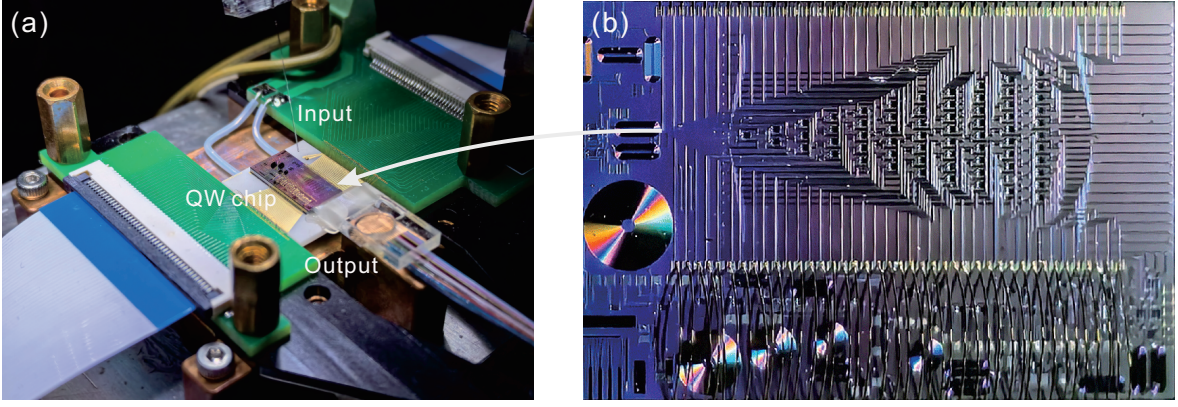
FIG. 5. **Experimental on-chip QWs and external setup.** (a) The experimental device includes a laser input connected to the photonic QW chip and an array of optical fibers coupled to the output ports. (b) A close-up view of the chip depicted in (a). The chip consists of 66 phase shifters, 78 beamsplitters, 24 input ports, and 24 output ports.

all parameters $\boldsymbol{\lambda}$ until the cut-off condition is met.

We examine the impact of varying training sample numbers N on the performance of our calibration method, as depicted in Fig. 3. The optimal parameters $\boldsymbol{\lambda}^*$ of the trained custom network are illustrated in Figs. 3(a)-3(e) for $N = 200$ (blue), 300 (orange), 400 (green), and 500 (red). It is observed that as the number of training samples increases, the values of the optimal parameters approach the preset target ones. The corresponding training process of the custom network is displayed in Fig. 3(f), where the cost function is plotted as a function of the training epochs. Compared to the slow-descent regions, each distinctly sharp drop in the cost function occurs as a result of one round of the alternating training of $\boldsymbol{a}$ and $\boldsymbol{b}$.

To quantify the performance of the calibrated PIC, we first compute the L1-norm distance $d(\boldsymbol{\lambda}^*) = \frac{1}{2}|\boldsymbol{P}_{\mathrm{ML}}(\boldsymbol{\lambda}^*) - \boldsymbol{P}_{\mathrm{tar}}|$ between the target distributions and the network's predictions for 1000 test samples. Furthermore, to fully characterize the output states, we also provide infidelity $F_{\mathrm{in}} = 1 - \mathrm{Tr}(\sqrt{\sqrt{\rho_{\mathrm{tar}}}\rho_{\mathrm{ML}}(\boldsymbol{\lambda}^*)\sqrt{\rho_{\mathrm{tar}}}})$ as a measurement of the distance between the target and network-predicated output quantum states. The values of the L1-norm distance and infidelity range from 0, indicating a perfect match, to 1, indicating a complete mismatch. Figs. 4(a)-4(c) show the histograms of the L1-norm distance, infidelity and purity error $|\mathrm{Tr}[\rho_{\mathrm{ML}}^2(\boldsymbol{\lambda}^*)] - \mathrm{Tr}[\rho_{\mathrm{tar}}^2]|$ for the test datasets, using different numbers of training samples N, respectively. We can see that nearly all the values of the three quantities are less than 1.5×10^{-2}, 1×10^{-3}, and 2×10^{-15} for $N = 500$ (red bars), respectively. These results indicate that the QW PIC with a circuit depth of 12 can be well calibrated by using several hundred training data samples. We note that our calibration method requires only about a minute for the entire training and validation processes of the network on a laptop equipped with an AMD Ryzen 7 5800H CPU. The training and validation processes are performed using the PyTorch package for Python program [50].

III. EXPERIMENTAL SETUP AND RESULTS

A. Experimental setup

In our experiment, the conceptual diagram of the QW PIC depicted in Fig. 2 is implemented through silicon waveguide circuits fabricated using the complementary metal–oxide semiconductor (CMOS)-compatible silicon photonics process [51], as illustrated in Fig. 5. The $6.8\,\mathrm{mm} \times 3.4\,\mathrm{mm}$ chip comprises an array of 66 individually controlled thermo-optic phase shifters with 78 waveguide beamsplitters to form the cascade of the interference network of a 12-step on-chip QW. The controlled phase shifter is implemented by electrically heating a resistor to change the refractive index of the waveguide to deform the optical path length [52]. Then, the phase shifter has a linear phase-power relation, and its phase change $[0, 2\pi]$ can be tuned by changing the imposed current according to the nonlinear phase-current relation in Eq. (3). To program all these phase shifters, we have developed a multi-channel electronic control system. The beamsplitter elements are realized using multimode interference couplers. On-chip grating couplers [53] are used for fiber-chip coupling, and the coupling loss is 4.5 dB per facet. A continuous-wave laser (TLX1, Thorlabs) operating at a wavelength of 1550 nm and with a power output of 5 mW is coupled into the QW PIC through a single-mode optical fiber. The input fiber is precisely aligned with the lower path of the first waveguide beamsplitter using a 3-axis motorized positioning stage (KWL06050, Suruga), and the polarization of the input laser is fine-tuned using the 3-paddle fiber polarization controller (FPC030, Thorlabs). We simultaneously measure the light intensity from all output ports by coupling them to an array of optical fibers and subsequently directing them to Germanium-amplified detectors (PDA30B2, Thorlabs). For each output port, we detect the light intensity five times and store the average value. Consequently, the measured intensity distributions are normalized to the
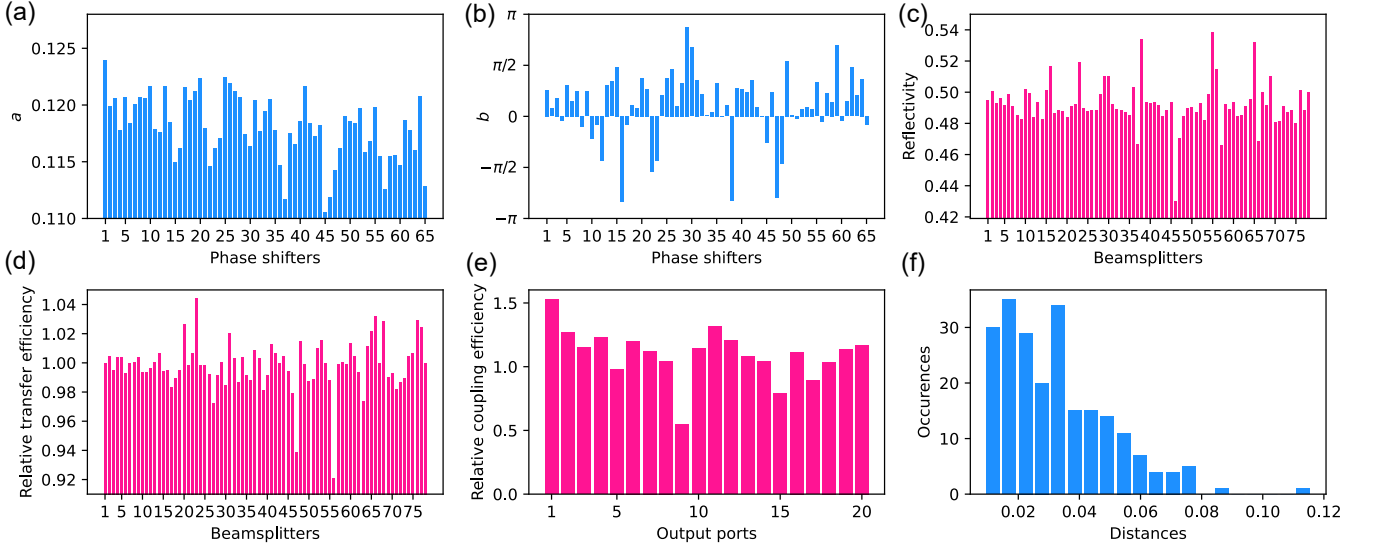
FIG. 6. **Experimental calibration results.** (a) and (b) depict the predicted values of the parameters $\boldsymbol{a}$ and $\boldsymbol{b}$ in the phase-current relation for each phase shifter, respectively. (c) Predicted values of the reflectivity $\sin^2\boldsymbol{\theta}$ for individual on-chip beamsplitter. (d) Predicted relative transfer efficiency $\boldsymbol{\alpha}$ between the beamsplitters. (e) Predicted relative coupling efficiency $\boldsymbol{\eta}$ between the output ports. (f) Histogram of the measured values of the L1-norm distance $d(\boldsymbol{\lambda}^*)$, which represent the difference between the test dataset's output light probability distributions $\{\boldsymbol{P}_{\exp}^{\text{test}}\}_n$ and the trained network's predictions $\{\boldsymbol{P}_{\text{ML}}(\boldsymbol{\lambda}^*)\}_n$.

experimental probability distribution $\boldsymbol{P}_{\exp}$. In the experiment, the QW PIC is temperature-controlled at 22°C.

B. Experimental results

Herein, the calibration of the QW PIC is implemented and tested by training the custom network directly on actual experimental datasets $\{\boldsymbol{P}_{\exp}\}_n$. In order to generate the experimental training and test datasets, we randomly generate 1500 sets of current distributions $\{\boldsymbol{I}\}_n$ ranging from $0\,\text{mA}$ to $7\,\text{mA}$ imposed on the phase shifters on our silicon PIC. We obtain 1500 sets of experimental probability distributions $\{\boldsymbol{P}_{\exp}\}_n$ after normalizing the measured light intensity over all output ports. The experimental datasets are divided into two parts according to an 85:15 ratio, where 1275 sets of datasets are used to train the custom network, and the remaining 225 sets of datasets are used to test the trained network. With these experimental training sets, the parameters $\boldsymbol{\lambda} = \{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\eta}\}$ of the custom network in Fig. 1 are then pre-trained.

The trained network can characterize the phase-current relation, assumed to be of the form $\phi = aI^2 + b$, the reflectivity parameter $\boldsymbol{\theta}$ and the relative transfer efficiency $\boldsymbol{\alpha}$ of BSs, as well as the relative output coupling efficiency $\boldsymbol{\eta}$ between the output ports. Here, we provide the relative transfer efficiency between all BSs and the relative coupling efficiency between all output ports because we utilize the normalized probability distributions to calibrate the QW PIC. Even so, we only need to measure the absolute transfer efficiency for one BS and the

absolute coupling efficiency for one output port. The efficiencies of all the remaining BSs and output ports can then be directly calculated.

The characterization results for 65 phase shifters, 78 beamsplitters and 20 output ports are shown in Figs. 6(a)-6(e). Note that we only experimentally detect 20 output ports in the middle, neglecting two output ports at the top and bottom positions, as the output light intensities at these four ports are not affected by the phase shifter settings. The phase shifter in the upper right corner of the QW PIC is also omitted because it has no impact on the output light intensity. Due to the influence of the chip fabrication errors, the coefficients $\boldsymbol{a}$ and $\boldsymbol{b}$ in Figs. 6(a) and 6(b) for each phase shifter are often not the same, and the reflectivity of some beamsplitters in Fig. 6(c) deviates from the ideal value of 0.5.

After completing the calibration of the photonic chip using the training datasets, we utilize the experimental test datasets to further verify the accuracy of the calibration. We apply the remaining 225 sets of distributions of random currents, and compare the experimentally measured distributions $\{\boldsymbol{P}_{\exp}^{\text{test}}\}_n$ with the predicted probability distributions $\{\boldsymbol{P}_{\text{ML}}(\boldsymbol{\lambda}^*)\}_n$ of the trained network with the optimal hyperparameters $\boldsymbol{\lambda}^*$. The difference between the two distributions for each test sample n is defined by L1-norm distance: $d(\boldsymbol{\lambda}^*) = \frac{1}{2}|\boldsymbol{P}_{\text{ML}}(\boldsymbol{\lambda}^*) - \boldsymbol{P}_{\exp}^{\text{test}}|$. We observe that the distance values for 85% of the test samples are less than 0.05, as shown in Fig. 6(f). The average distance for all test samples is $\bar{d} = 0.033 \pm 0.018$, indicating that our calibration results for the phase shifters and beamsplitters are in good agreement with the actual physical process. The slight deviation suggests the pres-
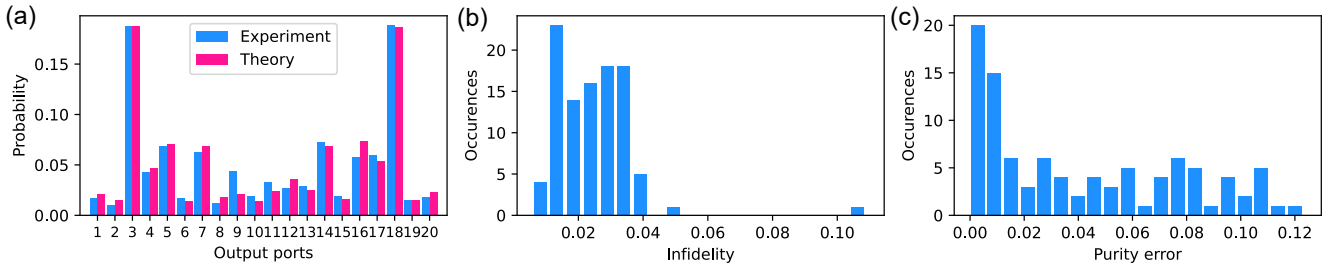
FIG. 7. **Validation of the QW PIC.** (a) The measured probability distribution $P_{\mathrm{exp}}^{\mathrm{QW}}$ (blue bars) for 12-step Hadamard QWs. The red bars mean the distribution $P_{\mathrm{ML}}(\boldsymbol{\lambda}^*)$ predicted by the trained network. (b) and (c) display histograms of the measured values of infidelity and purity error for 100 output quantum states after 9-step QW dynamics, respectively.

ence of additional chip imperfections that are not considered by our custom network, such as the non-dominant thermal crosstalk effect [54]. The error bar of $\bar{d}$ represents one standard deviation.

After verifying the performance on numerous random validation samples using the calibrated photonic chip, we proceed to implement the conventional 1D Hadamard QWs, where the coin operators are fixed as the Hadamard gate and independent of time and position. We inject 1550 nm light into the lower path of the first waveguide beamsplitter followed by setting the phase of the first phase shifter to $\pi/2$. Therefore, we prepare a balanced superposition initial state between two paths with a phase difference of $\pi/2$ through one-step QW dynamics [55]. The phases of the remaining phase shifters are set to 0. We measure the light intensity coupled to the optical fiber arrays, and the experimental probability distribution $P_{\mathrm{exp}}^{\mathrm{QW}}$ for a 12-step QW is obtained using the calibrated fiber-coupling efficiency for compensation, as depicted in Fig. 7(a). The resulting L1-norm distance with respect to the QW distribution predicted by the trained network is $d = 0.057$, indicating their good agreement.

Beyond the position distributions, we then perform the quantum state tomography for the calibrated QW PIC with inherent high-dimensional structures [56–59]. To achieve this, in addition to the computational basis $|v\rangle$, extra measurement bases need to be introduced to fully characterize the output quantum state. We divide the QW PIC into two sections: the first part, with a circuit depth of 9, is used to implement arbitrary 9-step QW dynamics, while the second part, with a circuit depth of 3, serves as a high-dimensional measurement device [60–63]. For each 9-step QW dynamics, we impose 14 sets of current distributions into the measurement device and measure the output probability distributions. Here, we perform standard maximum-likelihood reconstruction with the estimator taking the form $\rho_{\mathrm{m}} = TT^{\dagger}/\mathrm{Tr}(TT^{\dagger})$, where T is a lower triangular matrix [64]. The quantum states after the 9-step QW can be reconstructed by maximizing the likelihood of 14 sets of experimental distributions. We arbitrarily generate 100 types of QW dynamics, and the histograms of the infidelity and purity error of the

reconstructed quantum states are displayed in Figs. 7(b) and 7(c), respectively. The corresponding average values of the infidelity and purity error are 0.025 ± 0.012 and 0.042 ± 0.037, respectively, indicating that the experimental states align well with the theoretical expectations.

IV. CONCLUSION

In conclusion, we introduce a global calibration method for the integrated photonic circuits. To achieve this, we construct a network with custom layers to model the physical process of the chip and train it using the gradient-descent-based tool, Adam. We experimentally validate our calibration method using a photonic QW chip with a circuit depth of 12. The trained network can globally calibrate the nonlinear phase-current relations of the thermo-optical phase shifters and the reflectivities of beamsplitters on the QW chip, which efficiently bypasses the error accumulation problem associated with the traditional one-by-one calibration method. Consequently, a low average distance of 0.033 between the measured chip distributions and the network predictions is achieved.

Our calibration method does not require a meticulous design of the decoupling scheme to calibrate phase shifters one by one, making it suitable for photonic chips with a prior physical structure. Moreover, this method offers high degrees of freedom, allowing any parameters involved in the physical process to be added to the custom network and learned through gradient-descent-based optimization. For instance, thermal crosstalk parameters between phase shifters could be further characterized by computing the nondiagonal terms in a given matrix phase-current relation. Thus, our method shows promise as a potent tool with potential applications in large-scale chip calibration and optimal control problems [65–68].

[1] R. Osgood Jr and X. Meng, *Principles of Photonic Integrated Circuits* (Springer, 2021).

[2] J. Wang, F. Sciarrino, A. Laing, and M. G. Thompson, Integrated photonic quantum technologies, Nat. Photonics **14**, 273–284 (2020).

[3] A. W. Elshaari, W. Pernice, K. Srinivasan, O. Benson, and V. Zwiller, Hybrid integrated quantum photonic circuits, Nat. Photonics **14**, 285–298 (2020).

[4] L.-T. Feng, G.-C. Guo, and X.-F. Ren, Progress on integrated quantum photonic sources with silicon, Adv. Quantum Technol. **3**, 1900058 (2020).

[5] E. Pelucchi, G. Fagas, I. Aharonovich, D. Englund, E. Figueroa, Q. Gong, H. Hannes, J. Liu, C.-Y. Lu, N. Matsuda, *et al.*, The potential and global outlook of integrated photonics for quantum technologies, Nat. Rev. Phys. **4**, 194–208 (2022).

[6] L. Feng, M. Zhang, J. Wang, X. Zhou, X. Qiang, G. Guo, and X. Ren, Silicon photonic devices for scalable quantum information applications, Photon. Res. **10**, A135–A153 (2022).

[7] G. Moody, V. J. Sorger, D. J. Blumenthal, P. W. Juodawlkis, W. Loh, C. Sorace-Agaskar, A. E. Jones, K. C. Balram, J. C. F. Matthews, A. Laing, *et al.*, 2022 roadmap on integrated quantum photonics, J. Phys.: Photonics **4**, 012501 (2022).

[8] T. Giordani, F. Hoch, G. Carvacho, N. Spagnolo, and F. Sciarrino, Integrated photonics in quantum technologies, Riv. Nuovo Cim. **46**, 71–103 (2023).

[9] W. Luo, L. Cao, Y. Shi, L. Wan, H. Zhang, S. Li, G. Chen, Y. Li, S. Li, Y. Wang, S. Sun, M. F. Karim, H. Cai, L. C. Kwek, and A. Q. Liu, Recent progress in quantum photonic chips for quantum communication and internet, Light: Sci. Appl. **12**, 175 (2023).

[10] Y. Su, Y. He, X. Guo, W. Xie, X. Ji, H. Wang, X. Cai, L. Tong, and S. Yu, Scalability of large-scale photonic integrated circuits, ACS Photonics **10**, 2020–2030 (2023).

[11] J. Capmany and D. Pérez, *Programmable Integrated Photonics* (Oxford University Press, 2020).

[12] H. Sun, Q. Qiao, Q. Guan, and G. Zhou, Silicon photonic phase shifters and their applications: A review, Micromachines **13**, 1509 (2022).

[13] S. Liu, J. Feng, Y. Tian, H. Zhao, L. Jin, B. Ouyang, J. Zhu, and J. Guo, Thermo-optic phase shifters based on silicon-on-insulator platform: state-of-the-art and a review, Front. Optoelectron. **15**, 9 (2022).

[14] N. C. Harris, G. R. Steinbrecher, M. Prabhu, Y. Lahini, J. Mower, D. Bunandar, C. Chen, F. N. C. Wong, T. Baehr-Jones, M. Hochberg, S. Lloyd, and D. Englund, Quantum transport simulations in a programmable nanophotonic processor, Nat. Photonics **11**, 447–452 (2017).

[15] 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**, 1623–1631 (2018).

[16] W. Bogaerts, D. Pérez, J. Capmany, D. A. Miller, J. Poon, D. Englund, F. Morichetti, and A. Melloni, Programmable photonic circuits, Nature **586**, 207–216 (2020).

[17] W. Bogaerts and A. Rahim, Programmable photonics: An opportunity for an accessible large-volume pic ecosystem, IEEE J. Sel. Top. Quantum Electron. **26**, 1–17 (2020).

[18] T. Giordani, R. Wagner, C. Esposito, A. Camillini, F. Hoch, G. Carvacho, C. Pentangelo, F. Ceccarelli, S. Piacentini, A. Crespi, *et al.*, Experimental certification of contextuality, coherence, and dimension in a programmable universal photonic processor, Sci. Adv. **9**, eadj4249 (2023).

[19] P. J. Shadbolt, M. R. Verde, A. Peruzzo, A. Politi, A. Laing, M. Lobino, J. C. Matthews, M. G. Thompson, and J. L. O'Brien, Generating, manipulating and measuring entanglement and mixture with a reconfigurable photonic circuit, Nat. Photonics **6**, 45–49 (2012).

[20] H. W. Li, J. Wabnig, D. Bitauld, P. Shadbolt, A. Politi, A. Laing, J. L. O'Brien, and A. O. Niskanen, Calibration and high fidelity measurement of a quantum photonic chip, New J. Phys. **15**, 063017 (2013).

[21] D. A. B. Miller, Analyzing and generating multimode optical fields using self-configuring networks, Optica **7**, 794–801 (2020).

[22] C. Alexiev, J. C. C. Mak, W. D. Sacher, and J. K. S. Poon, Calibrating rectangular interferometer meshes with external photodetectors, OSA Continuum **4**, 2892–2904 (2021).

[23] J. Cao, Z. Chen, Y. Wang, T. Feng, Z. Li, Z. Xing, H. Li, and X. Zhou, Experimental demonstration of a fast calibration method for integrated photonic circuits with cascaded phase shifters, Chin. Phys. B **31**, 114204 (2022).

[24] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, Experimental realization of any discrete unitary operator, Phys. Rev. Lett. **73**, 58–61 (1994).

[25] J. Carolan, C. Harrold, C. Sparrow, E. Martín-López, N. J. Russell, J. W. Silverstone, P. J. Shadbolt, N. Matsuda, M. Oguma, M. Itoh, G. D. Marshall, M. G. Thompson, J. C. F. Matthews, T. Hashimoto, J. L. O'Brien, and A. Laing, Universal linear optics, Science **349**, 711–716 (2015).

[26] W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley, Optimal design for universal multiport interferometers, Optica **3**, 1460–1465

(2016).

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

[28] C. Pentangelo, N. D. Giano, S. Piacentini, R. Arpe, F. Ceccarelli, A. Crespi, and R. Osellame, High-fidelity and polarization-insensitive universal photonic processors fabricated by femtosecond laser writing, Nanophotonics **13**, 2259–2270 (2024).

[29] N. Maring, A. Fyrillas, M. Pont, E. Ivanov, P. Stepanov, N. Margaria, W. Hease, A. Pishchagin, A. Lemaître, I. Sagnes, *et al.*, A versatile single-photon-based quantum computing platform, Nat. Photonics **18**, 603–609 (2024).

[30] X. Qiang, X. Zhou, J. Wang, C. M. Wilkes, T. Loke, S. O'Gara, L. Kling, G. D. Marshall, R. Santagati, T. C. Ralph, J. B. Wang, J. L. O'Brien, M. G. Thompson, and J. C. F. Matthews, Large-scale silicon quantum photonics implementing arbitrary two-qubit processing, Nat. Photonics **12**, 534–539 (2018).

[31] J. Bao, Z. Fu, T. Pramanik, J. Mao, Y. Chi, Y. Cao, C. Zhai, Y. Mao, T. Dai, X. Chen, *et al.*, Very-large-scale integrated quantum graph photonics, Nat. Photonics **17**, 573–581 (2023).

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

[33] A. Peruzzo, M. Lobino, J. C. F. Matthews, N. Matsuda, A. Politi, K. Poulios, X.-Q. Zhou, Y. Lahini, N. Ismail, K. Wörhoff, Y. Bromberg, Y. Silberberg, M. G. Thompson, and J. L. OBrien, Quantum walks of correlated photons, Science **329**, 1500–1503 (2010).

[34] L. Sansoni, F. Sciarrino, G. Vallone, P. Mataloni, A. Crespi, R. Ramponi, and R. Osellame, Two-particle bosonic-fermionic quantum walk via integrated photonics, Phys. Rev. Lett. **108**, 010502 (2012).

[35] A. Crespi, R. Osellame, R. Ramponi, V. Giovannetti, R. Fazio, L. Sansoni, F. De Nicola, F. Sciarrino, and P. Mataloni, Anderson localization of entangled photons in an integrated quantum walk, Nat. Photonics **7**, 322–328 (2013).

[36] F. Caruso, A. Crespi, A. Ciriolo, F. Sciarrino, and R. Osellame, Fast escape of a quantum walker from an integrated photonic maze, Nat. Commun. **7**, 11682 (2016).

[37] H. Tang, X.-F. Lin, Z. Feng, J.-Y. Chen, J. Gao, K. Sun, C.-Y. Wang, P.-C. Lai, X.-Y. Xu, Y. Wang, L.-F. Qiao, A.-L. Yang, and X.-M. Jin, Experimental two-dimensional quantum walk on a photonic chip, Sci. Adv. **4**, eaat3174 (2018).

[38] X. Qiang, Y. Wang, S. Xue, R. Ge, L. Chen, Y. Liu, A. Huang, X. Fu, P. Xu, T. Yi, F. Xu, M. Deng, J. B. Wang, J. D. A. Meinecke, J. C. F. Matthews, X. Cai, X. Yang, and J. Wu, Implementing graph-theoretic quantum algorithms on a silicon photonic quantum walk processor, Sci. Adv. **7**, eabb8375 (2021).

[39] Y. Wang, Y. Liu, J. Zhan, S. Xue, Y. Zheng, R. Zeng, Z. Wu, Z. Wang, Q. Zheng, D. Wang, *et al.*, Large-scale full-programmable quantum walk and its applications (2022), arXiv:2208.13186 [quant-ph].

[40] R. He, Y. Zhao, C. Sheng, J. Duan, Y. Wei, C. Sun, L. Lu, Y.-X. Gong, S. Zhu, and H. Liu, Experimental realization of quantum walks near synthetic horizons on

[41] R. Uppu, T. A. W. Wolterink, T. B. H. Tentrup, and P. W. H. Pinkse, Quantum optics of lossy asymmetric beam splitters, Opt. Express **24**, 16440–16449 (2016).

[42] M. A. Broome, A. Fedrizzi, B. P. Lanyon, I. Kassal, A. Aspuru-Guzik, and A. G. White, Discrete single-photon quantum walks with tunable decoherence, Phys. Rev. Lett. **104**, 153602 (2010).

[43] See Supplemental Material at URL for the calibration results of the triangular "Rech" and rectangular "Clements" chips.

[44] Y. Aharonov, L. Davidovich, and N. Zagury, Quantum random walks, Phys. Rev. A **48**, 1687 (1993).

[45] X.-Y. Xu, Q.-Q. Wang, W.-W. Pan, K. Sun, J.-S. Xu, G. Chen, J.-S. Tang, M. Gong, Y.-J. Han, C.-F. Li, and G.-C. Guo, Measuring the winding number in a large-scale chiral quantum walk, Phys. Rev. Lett. **120**, 260501 (2018).

[46] K. L. Fenwick, F. Bouchard, G. S. Thekkadath, D. England, P. J. Bustard, K. Heshami, and B. Sussman, Photonic quantum walk with ultrafast time-bin encoding, Optica **11**, 1017–1023 (2024).

[47] S.-J. Tao, Q.-Q. Wang, Z. Chen, W.-W. Pan, S. Yu, G. Chen, X.-Y. Xu, Y.-J. Han, C.-F. Li, and G.-C. Guo, Experimental optimal generation of hybrid entangled states in photonic quantum walks, Opt. Lett. **46**, 1868–1871 (2021).

[48] S. Ahmed, F. Quijandría, and A. F. Kockum, Gradient-descent quantum process tomography by learning kraus operators, Phys. Rev. Lett. **130**, 150402 (2023).

[49] D. P. Kingma and J. Ba, Adam: A method for stochastic optimization (2017), arXiv:1412.6980 [cs.LG].

[50] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, Pytorch: An imperative style, high-performance deep learning library (2019), arXiv:1912.01703 [cs.LG].

[51] L.-T. Feng, M. Zhang, X. Xiong, D. Liu, Y.-J. Cheng, F.-M. Jing, X.-Z. Qi, Y. Chen, D.-Y. He, G.-P. Guo, G.-C. Guo, D.-X. Dai, and X.-F. Ren, Transverse mode-encoded quantum gate on a silicon photonic chip, Phys. Rev. Lett. **128**, 060501 (2022).

[52] N. C. Harris, Y. Ma, J. Mower, T. Baehr-Jones, D. Englund, M. Hochberg, and C. Galland, Efficient, compact and low loss thermo-optic phase shifter in silicon, Opt. Express **22**, 10487–10493 (2014).

[53] G. Son, S. Han, J. Park, K. Kwon, and K. Yu, High-efficiency broadband light coupling between optical fibers and photonic integrated circuits, Nanophotonics **7**, 1845–1864 (2018).

[54] F. Ceccarelli, S. Atzeni, C. Pentangelo, F. Pellegatta, A. Crespi, and R. Osellame, Low power reconfigurability and reduced crosstalk in integrated photonic circuits fabricated by femtosecond laser micromachining, Laser Photonics Rev. **14**, 2000024 (2020).

[55] Q.-Q. Wang, S.-J. Tao, W.-W. Pan, Z. Chen, G. Chen, K. Sun, J.-S. Xu, X.-Y. Xu, Y.-J. Han, C.-F. Li, and G.-C. Guo, Experimental verification of generalized eigenstate thermalization hypothesis in an integrable system, Light: Sci. Appl. **11**, 194 (2022).

[56] Q.-Q. Wang, S. Dong, X.-W. Li, X.-Y. Xu, C. Wang, S. Han, M.-H. Yung, Y.-J. Han, C.-F. Li, and G.-C. Guo,

Efficient learning of mixed-state tomography for photonic quantum walk, Sci. Adv. **10**, eadl4871 (2024).

[57] J. G. Titchener, A. S. Solntsev, and A. A. Sukhorukov, Two-photon tomography using on-chip quantum walks, Opt. Lett. **41**, 4079–4082 (2016).

[58] D. Zia, R. Checchinato, A. Suprano, T. Giordani, E. Polino, L. Innocenti, A. Ferraro, M. Paternostro, N. Spagnolo, and F. Sciarrino, Regression of high-dimensional angular momentum states of light, Phys. Rev. Res. **5**, 013142 (2023).

[59] A. Suprano, D. Zia, L. Innocenti, S. Lorenzo, V. Cimini, T. Giordani, I. Palmisano, E. Polino, N. Spagnolo, F. Sciarrino, G. M. Palma, A. Ferraro, and M. Paternostro, Experimental property reconstruction in a photonic quantum extreme learning machine, Phys. Rev. Lett. **132**, 160802 (2024).

[60] L. Innocenti, H. Majury, T. Giordani, N. Spagnolo, F. Sciarrino, M. Paternostro, and A. Ferraro, Quantum state engineering using one-dimensional discrete-time quantum walks, Phys. Rev. A **96**, 062326 (2017).

[61] T. Giordani, E. Polino, S. Emiliani, A. Suprano, L. Innocenti, H. Majury, L. Marrucci, M. Paternostro, A. Ferraro, N. Spagnolo, and F. Sciarrino, Experimental engineering of arbitrary qudit states with discrete-time quantum walks, Phys. Rev. Lett. **122**, 020503 (2019).

[62] Y.-Y. Zhao, N.-K. Yu, P. Kurzyński, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, Experimental realization of generalized qubit measurements based on quantum walks, Phys. Rev. A **91**, 042101 (2015).

[63] X. Wang, X. Zhan, Y. Li, L. Xiao, G. Zhu, D. Qu, Q. Lin, Y. Yu, and P. Xue, Generalized quantum measurements on a higher-dimensional system via quantum walks, Phys. Rev. Lett. **131**, 150803 (2023).

[64] J. Shang, Z. Zhang, and H. K. Ng, Superfast maximum-likelihood reconstruction for quantum tomography, Phys. Rev. A **95**, 062336 (2017).

[65] R. Burgwal, W. R. Clements, D. H. Smith, J. C. Gates, W. S. Kolthammer, J. J. Renema, and I. A. Walmsley, Using an imperfect photonic network to implement random unitaries, Opt. Express **25**, 28236–28245 (2017).

[66] M. Jacques, A. Samani, E. El-Fiky, D. Patel, Z. Xing, and D. V. Plant, Optimization of thermo-optic phase-shifter design and mitigation of thermal crosstalk on the soi platform, Opt. Express **27**, 10456–10471 (2019).

[67] S. Bandyopadhyay, R. Hamerly, and D. Englund, Hardware error correction for programmable photonics, Optica **8**, 1247–1255 (2021).

[68] A. Fyrillas, O. Faure, N. Maring, J. Senellart, and N. Belabas, Scalable machine learning-assisted clear-box characterization for optimally controlled photonic circuits, Optica **11**, 427–436 (2024).