



**HAL**  
open science

# Scalable machine learning-assisted clear-box characterization for optimally controlled photonic circuits

Andreas Fyrrillas, Olivier Faure, Nicolas Maring, Jean Senellart, Nadia Belabas

## ► To cite this version:

Andreas Fyrrillas, Olivier Faure, Nicolas Maring, Jean Senellart, Nadia Belabas. Scalable machine learning-assisted clear-box characterization for optimally controlled photonic circuits. *Optica*, 2024, 11 (3), pp.427. <10.1364/OPTICA.512148>. <hal-04621227>

**HAL Id: hal-04621227**

**<https://hal.science/hal-04621227v1>**

Submitted on 23 Jun 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization



# Scalable machine learning-assisted clear-box characterization for optimally controlled photonic circuits

ANDREAS FYRILLAS,<sup>1,2,\*</sup>  OLIVIER FAURE,<sup>1</sup> NICOLAS MARING,<sup>1</sup>  JEAN SENELLART,<sup>1</sup> AND NADIA BELABAS<sup>2</sup> 

<sup>1</sup>Quandela, 7 Rue Léonard de Vinci, 91300 Massy, France

<sup>2</sup>Centre for Nanosciences and Nanotechnologies, CNRS, Université Paris-Saclay, UMR 9001, 10 Boulevard Thomas Gobert, 91120, Palaiseau, France  
\*andreas.fyrillas@quandela.com

Received 10 November 2023; revised 4 February 2024; accepted 16 February 2024; published 19 March 2024

Photonic integrated circuits offer a compact and stable platform for generating, manipulating, and detecting light. They are instrumental for classical and quantum applications. Imperfections stemming from fabrication constraints, tolerances, and operation wavelength impose limitations on the accuracy and thus utility of current photonic integrated devices. Mitigating these imperfections typically necessitates a model of the underlying physical structure and the estimation of parameters that are challenging to access. Direct solutions are currently lacking for mesh configurations extending beyond trivial cases. We introduce a scalable and innovative method to characterize photonic chips through an iterative machine learning-assisted procedure. Our method is based on a clear-box approach that harnesses a fully modeled virtual replica of the photonic chip to characterize. The process is sample-efficient and can be carried out with a continuous-wave laser and powermeters. The model estimates individual passive phases, crosstalk, beamsplitter reflectivity values, and relative input/output losses. Building upon the accurate characterization results, we mitigate imperfections to enable enhanced control over the device. We validate our characterization and imperfection mitigation methods on a 12-mode Clements-interferometer equipped with 126 phase shifters, achieving beyond state-of-the-art chip control with an average 99.77% amplitude fidelity on 100 implemented Haar-random unitary matrices. © 2024

Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

<https://doi.org/10.1364/OPTICA.512148>

## 1. INTRODUCTION

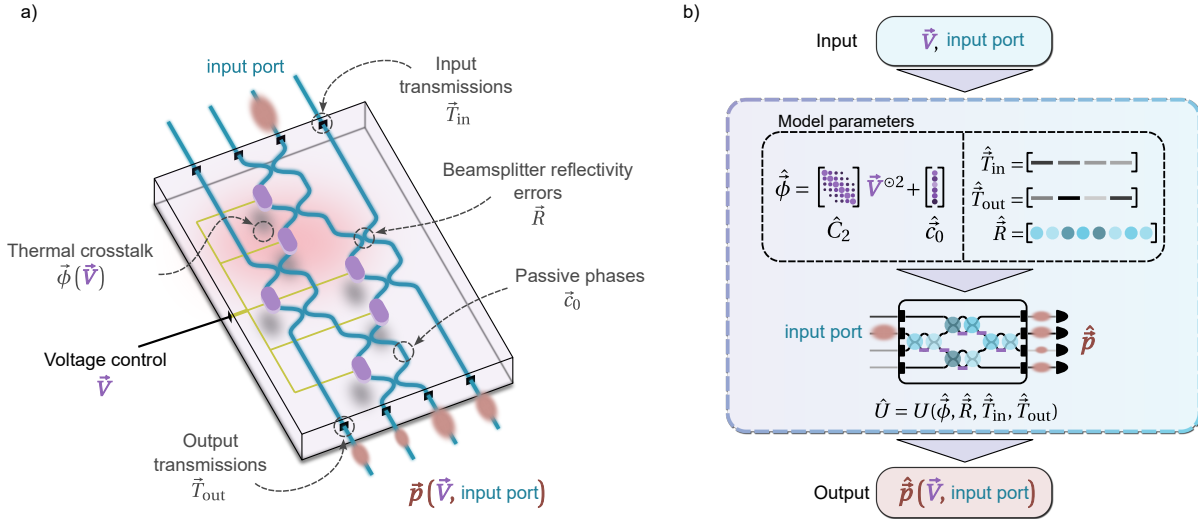
Photonic integrated circuits (PICs) incorporate optical components on a compact substrate, enabling the generation, manipulation, and detection of light [1]. PICs have emerged as a compelling and versatile platform to manipulate light, thanks to an unprecedented stability, compactness, and capability for scaling up. These miniature devices have showcased their potential to revolutionize photonic quantum computing [2,3], quantum communication [4], quantum cryptography [5], and quantum sensing [6]. Beyond the quantum realm, PICs find utility in classical domains such as microwave photonics, optical beamforming, and high-precision sensing [7]. Furthermore, PICs can be used to natively perform matrix-vector multiplications, offering the potential to propel the field of artificial intelligence forward [8].

PICs are in particular widely used to perform linear operations on light, featuring components such as static beamsplitters and tunable phase shifters, which are controlled by voltages or electric currents. Nevertheless, these elements exhibit imperfections in current photonic devices that cannot be overlooked [see Fig. 1(a)]. In terms of practical applications, these defects lead for instance to a severe degradation in performance of optical neural networks

[9] and a marked decrease in the fidelity of quantum gates [10]. Optimal control of PICs despite deviation from ideal devices is thus a primary challenge for quantum and classical applications.

Self-configuration protocols for PICs exist and mitigate imperfections without requiring detailed knowledge about the device. PIC self-configuration is a viable option in specific use cases, requiring for instance to route light from one input to one output [11,12], or when dealing with particular photonic circuits [13]. Otherwise, the light amplitude and phase transformation implemented by a PIC must be measured [14–17] and the phase shifters reconfigured in a trial-and-error approach until the targeted transformation is reached. This is however a costly and experimentally cumbersome operation hindering taking full advantage of PIC reconfigurability. Self-configuration may additionally break down in the presence of inhomogeneous output losses and crosstalk between phase shifters.

A promising strategy is to leverage the capabilities of machine learning. Neural networks have been successfully trained in a black-box approach to connect the measured single-photon statistics produced by a 3-mode PIC to the voltages applied on 2 phase shifters [18]. Neural networks were also used in intermediate gray-box approaches where the algorithm only learns the Hamiltonian



**Fig. 1.** Photonic chip imperfections modeled in a virtual replica. (a) Physical photonic integrated circuits (PICs) often exhibit various imperfections resulting from fabrication constraints, tolerances, and operation wavelength, illustrated here on a simplified PIC. In general, input and output ports have different optical transmissions, stored in vectors  $\vec{T}_{in}$  and  $\vec{T}_{out}$ . In addition, the real beamsplitter reflectivity values  $\vec{R}$  deviate from the target. Phase shifters (purple components) dissipating heat entail a phase-voltage relation of the type  $\vec{\phi} = \vec{\phi}(\vec{V})$  between all the physical phase shifts  $\vec{\phi}$  and applied voltages  $\vec{V}$ . In addition, optical path variations lead to non-zero phase shifts even without any voltages applied, i.e.,  $\vec{\phi}(\vec{0}) = \vec{c}_0 \neq \vec{0}$ . When sending light into the PIC, here represented by a laser pulse, the output light intensity distribution  $\vec{p}$  depends on the applied voltages and the chosen input port. (b) Our characterization process uses a virtual replica of the physical PIC. Hardware imperfections are modeled in the replica following Section 2. The model parameters represent the replica current knowledge of the physical PIC characteristics: matrix phase-voltage relation  $\hat{\phi} = \hat{C}_2 \cdot \vec{V}^{\circ 2} + \hat{c}_0$  [see Eq. (2)], optical input/output transmissions  $\hat{T}_{in}$  and  $\hat{T}_{out}$ , and beamsplitter reflectivities  $\hat{R}$ , where the hat notation indicates predicted quantities. When given a list of voltages  $\vec{V}$ , the model predicts the implemented phases  $\hat{\phi}$  on the virtual PIC and generates the matrix  $\hat{U} = U(\hat{\phi}, \hat{R}, \hat{T}_{in}, \hat{T}_{out})$  that encapsulates the virtual PIC action on light. The model then computes the predicted output light intensity distribution  $\hat{p}$ , normalized such that its elements sum to 1, resulting from light injected into a single input port.

of the photonic device and the measurement probabilities are computed according to the laws of quantum mechanics [19]. In both cases, scalability is a major issue as the number of required data samples to train the neural network grows heavily with the complexity of the physical PIC.

In the light of self-configuration and neural network-based approaches, an ideal method for achieving PIC optimal control should possess the dual characteristics of adaptability to address various defect types and scalability concerning the number of components within a PIC. To that end, we adopt here a clear-box approach relying on a model of the PIC and its imperfections derived from physical intuition. In clear-box methods, the model is constrained to learn only the parameters of interest, which is a promise for enhanced sample efficiency. However, the efficiency of this transparent approach hinges on the precision and faithfulness of the modeling of imperfections present in the physical system. Operating in the clear-box paradigm implies that each imperfection type must be addressed with a tailored mitigation strategy. Errors in beamsplitter reflectivity, for instance, can be compensated by computing rectified phase shifts either optimized globally via gradient-based methods [10,20] or optimized locally with faster deterministic schemes [21,22]. Similarly, compensation of crosstalk between phase shifters can in theory be achieved for optical neural networks [23], and has been demonstrated in simple experimental cases [24]. For clear-box imperfection mitigation, an accurate prior modeling and characterization of the PIC imperfections is essential. However, accessing directly the values of individual model parameters may be very challenging depending on the specific PIC architecture. This is especially

true for universal-scheme PICs (Reck [25], Clements [26], Bell [27]), which are notoriously hard to characterize, yet constitute the backbone of near-term photonic quantum processors [28].

We present an iterative machine learning-assisted PIC characterization process. We harness the sample-efficiency of the clear-box approach, which paves the way for scalability in the characterization of increasingly bigger PIC architectures. We also exploit the large data processing abilities of machine learning to handle the large resulting number of physically meaningful parameters and complex interferometer meshes. A comparable characterization strategy has recently been mentioned in [29], with limited elaboration on the methodology. Our method offers valuable insights into the physics of the device, which can then be used to improve fabrication processes, in contrast to neural network models that lack interpretability. In addition, we require only a laser and powermeters, or alternatively a single-photon source and single-photon detectors. The results of the characterization process are subsequently harnessed by a custom imperfection mitigation. We achieved unparalleled optimal control on a 12-mode Clements universal interferometer, one of the largest PICs in terms of number of components currently available (see state of the art in Supplement 1 A).

- In Section 2, we model the physical linear PIC to characterize and the relevant imperfections to take into account.
- Section 3 presents the different stages of the characterization protocol that allow us to finetune the modelled imperfection parameters. The protocol is then simulation benchmarked to

demonstrate its effectiveness to converge to the true parameter values.

- Harnessing the knowledge gained by the characterization step, Section 4 details our imperfection mitigation that translates targeted unitary matrices or sets of targeted phase shifts into voltages/electric currents and implements the target with high fidelity on the PIC.

- In Section 5, we experimentally validate our characterization process on a 12-mode Clements-interferometer PIC featuring 126 thermo-optic phase shifters and 132 directional couplers. We characterize passive phases, thermal crosstalk, beamsplitter reflectivity errors, and relative input/output losses. Using our imperfection mitigation, we implement unitary operations on single-photons and demonstrate beyond state-of-the-art 99.77% fidelity to the target.

We focus in the following on PICs for linear optics featuring beamsplitters and phase shifters, but our approach is generalizable to PICs manipulating for instance polarization of photons and featuring non-linear optical elements.

## 2. MODELLING PHOTONIC INTEGRATED CIRCUIT IMPERFECTIONS

The relevant imperfections in PICs for linear optics are as shown in Fig. 1(a):

- Beamsplitter reflectivity deviating from the target value. On-chip beamsplitters are realized by directional couplers [30]. Fabrication introduces random and systematic errors. Systematic errors also occur due to deviations of the light wavelength of the user from the fabrication wavelength. For universal-scheme PICs, which can in practice implement any unitary matrix acting on the spatial input modes, beamsplitter errors reduce the number of implementable unitary matrices [20,31].

- Passive phases [32] due to waveguide length differences and inhomogeneities of the refractive index. As a consequence, phase shifters induce a non-zero phase even when no voltage/electric current is applied. Passive phases add a layer of difficulty to the characterization process.

- Crosstalk induced by reconfigurable components. For instance, heat produced by a thermo-optic phase shifter [33] diffuses and distorts the action of other phase shifters [34]. Crosstalk is also expected in the case of phase shifters harnessing strain-induced birefringence [35] or the electro-optic effect [36]. Electric crosstalk can occur if the phase shifters share a common electric ground.

- Inhomogeneous input and output optical transmissions, due to differences in the optical coupling to the PIC or in light detection efficiencies. In practice, absorption losses in the PIC waveguides may not be entirely homogeneous. Nevertheless, we will assume uniformity of internal absorption losses.

The number of modes  $m$  of a PIC is the number of input/output ports. The physical action of a PIC is encapsulated in an  $m \times m$  matrix  $U$  whose element  $|u_{i,j}|^2$  represents the probability that photons injected in the input port  $j$  exit through the output port  $i$ . This picture is valid both in classical electrodynamics and in quantum optics.  $U$  may be non unitary to account for losses in the system. We use the following convention for an on-chip beamsplitter of reflectivity  $R$ :

$$\begin{bmatrix} \sqrt{R} & i\sqrt{1-R} \\ i\sqrt{1-R} & \sqrt{R} \end{bmatrix}. \quad (1)$$

The matrix  $U(\vec{\phi}, \vec{R})$  implemented by a PIC with phases  $\vec{\phi}$  on its phase shifters and beamsplitters of reflectivity  $\vec{R}$  is the matrix product of its individual components. If the  $i^{\text{th}}$  input (resp. output) port has optical transmission  $T$ , we model this by multiplying the  $i^{\text{th}}$  column (resp. row) of  $U(\vec{\phi}, \vec{R})$  by  $\sqrt{T}$ . The lists of optical input and output transmissions are denoted  $\vec{T}_{\text{in}}$  and  $\vec{T}_{\text{out}}$ . We normalize the maximum value of  $\vec{T}_{\text{out}}$  to 1; hence,  $\vec{T}_{\text{in}}$  contains both the input inhomogeneities and the uniform global setup losses. The resulting  $m \times m$  matrix modelling the PIC is denoted  $U(\vec{\phi}, \vec{R}, \vec{T}_{\text{in}}, \vec{T}_{\text{out}})$ .

In the following, we consider that the PIC phase shifters are voltage-controlled, but the case of electric current control is treated in an analog way. To model crosstalk between phase shifters, we use a phase-voltage relation linking the physically implemented phases  $\vec{\phi}$  and the applied voltages  $\vec{V}$  by a matrix relation of the type

$$\vec{\phi} = \sum_{k \geq 1} C_k \cdot \vec{V}^{\odot k} + \vec{c}_0 \quad (2)$$

where  $\vec{c}_0$  is a vector with  $n_{\text{PS}}$  entries containing the passive phases ( $n_{\text{PS}}$  is the number of on-chip phase shifters),  $\odot$  represents element-wise exponentiation, and the  $C_k$  are  $n_{\text{PS}} \times n_{\text{PS}}$  matrices. Off-diagonal elements in the  $C_k$  account for crosstalk between phase shifters. In principle it is sufficient for thermo-optic phase shifters to keep only the passive phases and the  $\vec{V}^{\odot 2}$  term, because of the  $V^2$ -dependence of Joule heating. Optionally, adding a  $\vec{V}^{\odot 4}$  term takes into account the change of heater resistance with temperature. We discuss in Supplement 1 B the case of electric crosstalk that is possibly present in PICs with voltage-controlled phase shifters. In the following, we will use without loss of generality a phase-voltage relation of the form  $\vec{\phi} = C_2 \cdot \vec{V}^{\odot 2} + \vec{c}_0$ .

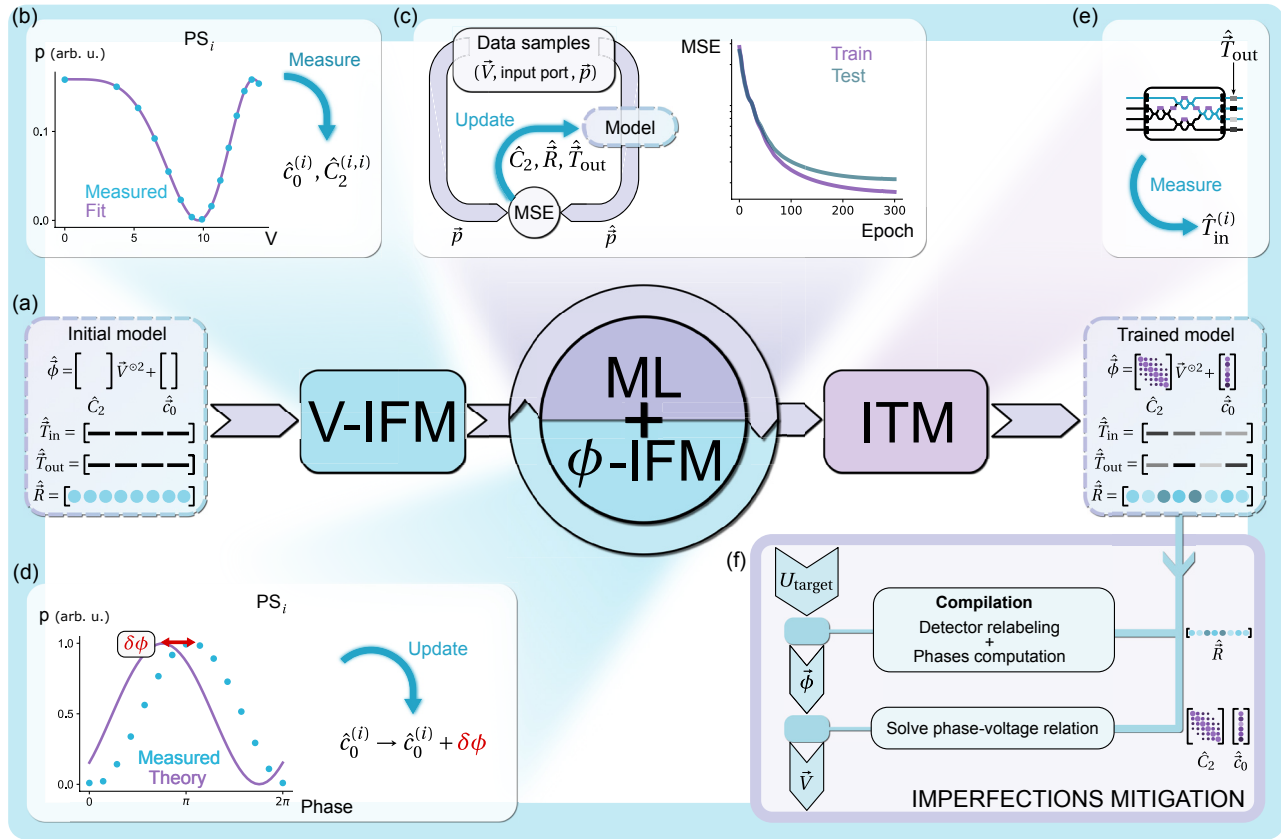
## 3. ITERATIVE MACHINE LEARNING-ASSISTED PIC CHARACTERIZATION

### A. Virtual PIC Replica

Our method relies on a virtual replica model of the physical PIC to characterize. The replica depicted on Fig. 1(b) is endowed with imperfections modeled as described in Section 2. The model parameters are the matrix  $\hat{C}_2$ , and the passive phase vector  $\hat{c}_0$  included in the phase-voltage relation Eq. (2), the optical input/output transmission vectors  $\hat{T}_{\text{in}}$  and  $\hat{T}_{\text{out}}$ , and the beamsplitter reflectivities  $\hat{R}$ . The hat notation indicates predicted quantities.

The virtual replica predicts the output light intensity distribution  $\hat{p}(\vec{V}, i)$  expected at the output of the physical PIC when light is injected in input  $i$  and voltages  $\vec{V}$  are applied. To do so, the replica uses the learned model parameters to compute the predicted phase shifts  $\hat{\phi}$  and construct the matrix  $\hat{U} = U(\hat{\phi}, \hat{R}, \hat{T}_{\text{in}}, \hat{T}_{\text{out}})$  describing the replica action on light. The resulting output light intensity distribution  $\hat{p}$  is computed from  $\hat{U}$  and the input port index and normalized afterwards to sum to 1.

The model parameters are optimized along our characterization process with the aim that the difference between the measured output light intensity distributions  $\vec{p}(\vec{V}, i)$  and the predicted  $\hat{p}(\vec{V}, i)$  is minimized. The replica is initialized as in Fig. 2(a):  $\hat{C}_2$  is



**Fig. 2.** Characterization of photonic integrated circuits using an iterative machine learning-assisted process and harnessing by an imperfection mitigation. (a) In our photonic integrated circuit (PIC) characterization process, the virtual replica model is initialized with parameter values given in Section 3.A that are optimized subsequently. (b) The first step in the characterization process is a voltage interference fringe measurement (V-IFM), detailed in Section 3.B. Each on-chip phase shifter is individually swept in voltage. Fitting each recorded interference fringe allows us to populate the passive phases vector  $\hat{c}_0$  and diagonal elements of the matrix  $\hat{C}_2$  in the phase-voltage relation. (c) The model is subsequently fine-tuned by a machine learning (ML) step. The ML step requires a dataset of the form  $\{(\vec{V}, \text{input port}, \vec{p})\}$  acquired as described in Section 3.C. ML consists of a gradient-descent algorithm that updates the  $\hat{C}_2$  matrix, the beamsplitter reflectivity vector  $\hat{R}$ , and output transmissions  $\hat{T}_{out}$ . The minimized cost function is the mean square error (MSE) between the data sample output light intensities  $\vec{p}$  and corresponding model predictions  $\hat{p}$ . (d) Phase interference fringe measurement ( $\phi$ -IFM): the learned phase-voltage relation is solved to sweep the phase of the individual phase shifters. The offset between the measured data points and the expected curve is used to update the passive phases  $\hat{c}_0$ . The process does multiple ML +  $\phi$ -IFM iterations until the MSE stops improving compared to the previous iteration (see Section 3.D). (e) The last step is an input transmission measurement (ITM). Light intensities are measured without normalization on the physical PIC from each used input. Differential output losses are compensated using the estimated model parameters, yielding a measurement of  $\hat{T}_{in}$ . Further information is in Section 3.E. (f) The parameters of the fully trained model are harnessed in our imperfection mitigation (see Section 4). To implement a target unitary matrix  $U_{target}$  with the PIC, a compilation step first relabels the detector outputs and computes a set of phase shifts  $\hat{\phi}$  that faithfully recreates the target  $U_{target}$  taking into account the learned beamsplitter reflectivity values  $\hat{R}$  (see Section 4.A). The phase shifts are then converted into voltages  $\vec{V}$  by solving the learned phase-voltage relation  $\hat{\phi} = \hat{C}_2 \cdot \vec{V}^{\otimes 2} + \hat{c}_0$  (see Section 4.B).  $\vec{V}$  is afterwards communicated to the physical PIC.

the zero matrix,  $\vec{c}_0 = \vec{0}$ , beamsplitters have their target reflectivity value, and  $\hat{T}_{in}^{(i)} = \hat{T}_{out}^{(i)} = 1$ .

## B. Voltage Interference Fringe Measurement

The first stage denoted V-IFM in our characterization protocol illustrated on Fig. 2(b) is an interference fringe measurement. V-IFM contributes to establishing the phase-voltage relation  $\vec{\phi} = \hat{C}_2 \cdot \vec{V}^{\otimes 2} + \vec{c}_0$  of the PIC. For each phase shifter  $PS_i$ , light is injected into the PIC via a single input chosen such that light is routed to  $PS_i$  using only already characterized phase shifters and then exits through a monitored output. The voltage applied on  $PS_i$  is swept, starting from 0 V to up to a designated safe maximum.

The voltage sweep produces an oscillating output light intensity that is recorded and fitted to recover initial values for  $PS_i$ 's self-heating coefficient  $C_2^{(i,i)}$  and passive phase  $c_0^{(i)}$ . To accommodate various PIC mesh designs, we algorithmically generate our interference fringe measurement protocol. This protocol provides the phase shifter characterization sequence, the routing of light from input to output for measurement, and phase offsets that need to be accounted for (see Supplement 1 C.3). Most experimental PIC characterizations [37–39] only carry out the V-IFM, but in the general case, the retrieved passive phase  $c_0^{(i)}$  is then flawed because of the imperfect directional coupler reflectivities and uncompensated crosstalk between phase shifters (see Supplement 1 C.4).

### C. Fine Tuning the Model Parameters Via Machine Learning

Estimating physical PIC parameters like crosstalk between phase shifters is a challenging task due to the substantial number of parameters  $n_{\text{PS}}^2$  that need to be determined, with  $n_{\text{PS}}$  the number of phase shifters. Therefore, we employ machine learning (ML) to find optimized values for the model parameters  $\hat{C}_2$ ,  $\hat{R}$ , and  $\hat{T}_{\text{out}}$  as shown on Fig. 2(c).

Our method requires a number of training samples  $n_{\text{train}}$  of the order of the number of parameters to optimize ( $n_{\text{PS}}^2$  + number of beamsplitters + number of modes). This is discussed in Section 3.F. For each data sample, a set of random voltages  $\vec{V}$  is generated and applied on the phase shifters. Light is then injected into a randomly chosen input port, and the corresponding output light intensity distribution  $\vec{p}$  is measured and normalized to sum to 1. Depending on the PIC layout, gathering data samples only from a restricted number of input ports might be sufficient. The data samples are a collection of the form  $\{(\vec{V}, \text{port}, \vec{p})\}$ .

We also acquire a set of test  $n_{\text{test}}$  samples with a ratio  $n_{\text{train}}/n_{\text{test}} = 80/20$ . The gradient-descent algorithm is implemented with the PyTorch package in Python [40]. The Adam optimizer [41] updates the model parameters via stochastic gradient descent to minimize the mean squared error (MSE) between the training set  $\{\vec{p}\}$  and the model predictions  $\{\hat{\vec{p}}\}$  for output light intensity distributions. The model is evaluated on the set of test samples to monitor the progress of the learning process. Each model parameter type is tuned with a different learning rate.

### D. Phase Interference Fringe Measurements and Iterating the Process

The estimated passive phases  $\hat{c}_0$  remain fixed during the ML stage because the model usually converges to a local minimum for the passive phases. We introduce a step [ $\phi$ -IFM on Fig. 2(d)] that updates  $\hat{c}_0$  while leveraging the knowledge gained by the virtual replica in the ML stage. The so-far learned phase-voltage relation is solved to sweep the phase of each phase shifter from 0 to  $2\pi$  while compensating crosstalk. The passive phases  $\vec{c}_0$  are then updated by estimating the offset between the measured and expected interference fringe, generated by taking into account the learned beamsplitter reflectivity values and output transmissions (details in Supplement 1 C.2).

The automated characterization protocol performs multiple (ML +  $\phi$ -IFM) iterations, gradually acquiring more precise information about the physical device. The gradient descent learning rates of the virtual replica are all divided by a constant factor after each iteration to enable faster convergence to the optimum. The protocol exits the iteration loop when the minimum MSE measured on the test dataset during the ML stage gradient descent exceeds that of the previous iteration.

### E. Input Transmission Measurement

The last stage in our characterization protocol measures the optical input transmissions of the PIC [ITM on Fig. 2(e)]. For each addressed input port  $i$ , we select among the training dataset the voltage configuration  $\vec{V}$  yielding the best digital replica prediction. The voltage configuration is applied on the physical PIC, and the output light intensity  $\vec{p}$  is measured without normalization. Differential output transmissions are compensated by computing

the vector  $\vec{p} \oslash \hat{T}_{\text{out}}$ , where  $\oslash$  is element-wise division and  $\hat{T}_{\text{out}}$  represents the output transmissions estimated from the ML stages [see Fig. 2(c) and Section 3.C]. The sum of the components of this vector yields  $T_{\text{in}}^{(i)} \times P$ , where  $T_{\text{in}}^{(i)}$  is the transmission of input port  $i$  and  $P$  is the input light intensity.

### F. Simulation Benchmarking the PIC Characterization Method

We benchmark our characterization protocol on simulated universal-scheme PICs with a Clements mesh and 6 to 12 modes. [see Fig. 3(a)]. The evaluation metric  $\text{TVD}_{\text{test}}$  for the virtual replica prediction accuracy is the average total variation distance (TVD, see Supplement 1 D) evaluated on the test dataset.  $\text{TVD}_{\text{test}}$  is the average distance between the measured output light intensity distributions and the corresponding model predictions.

- Reduction of the characterization duration.* In practice, the time needed to characterize a PIC is predominantly consumed by the  $\phi$ -IFM stages, which are constrained in terms of speed by the PIC reconfiguration and light intensity integration times. Hence, reducing the number of (ML +  $\phi$ -IFM) iterations results in a significantly more time-efficient PIC characterization. We show on Fig. 3(b) that the number of (ML +  $\phi$ -IFM) iterations required to characterize a PIC can be substantially reduced by increasing the number of gradient descent epochs processed in the ML stage. For the following, we use for each PIC size the value indicated by the corresponding arrow on the plot. We note that the total number of epochs to process increases with the number of modes, as the precision required on model parameters increases with PIC complexity in order to minimize error propagation across an increasing number of component layers.
- Sample efficiency.* We define the data-to-parameter ratio as the number of training samples divided by the number of parameters trained during the ML stage. Figure 3(c) demonstrates that a data-to-parameter ratio of 1 is a good compromise between low iteration count, convergence guarantees, and data acquisition time. This demonstrates that our clear-box PIC characterization is more sample efficient than black-box [18] and gray-box [19] alternatives (see Supplement 1 A). We simulate in Supplement 1 E the characterization of a 24-mode Clements interferometer and demonstrate that the sample efficiency also holds for significantly larger interferometer sizes.
- Impact of photon-counting noise.* In practice, measurement noise naturally sets a limit on the minimally attainable  $\text{TVD}_{\text{test}}$  denoted  $\text{TVD}_{\text{test, limit}}$ , due to the inherent difference between the noisy test dataset and the noiseless replica predictions. Knowing the order of magnitude of  $\text{TVD}_{\text{test, limit}}$  is crucial from an experimental point of view to evaluate if the characterization protocol has converged to the physically allowed limit. We examine the case where single photons are injected into the PIC and detected by single-photon detectors. Simulated characterizations yield values  $\text{TVD}_{\text{test, limit}}$  graphed as a function of the input photon count rate on Fig. 3(d), assuming a detection integration time of 1 s. The reached  $\text{TVD}_{\text{test, limit}}$  values agree with the theoretical threshold plotted as continuous lines. Estimating  $\text{TVD}_{\text{test, limit}}$  in the case of laser light and powermeter arrays necessitates a case-by-case approach, as powermeter arrays commonly exhibit dominant Gaussian noise originating from internal components.

## 4. IMPERFECTION MITIGATION

To mitigate imperfections, universal-scheme PICs benefit from a compilation process that translates target unitary matrices into phases to implement on the integrated phase shifters. To compute voltages from these phases, the PIC phase-voltage relation is retrieved from the trained virtual replica and solved. Imperfection mitigation is necessary to demonstrate enhanced PIC control [see Fig. 2(f)], leveraging the estimated imperfections by our PIC characterization process.

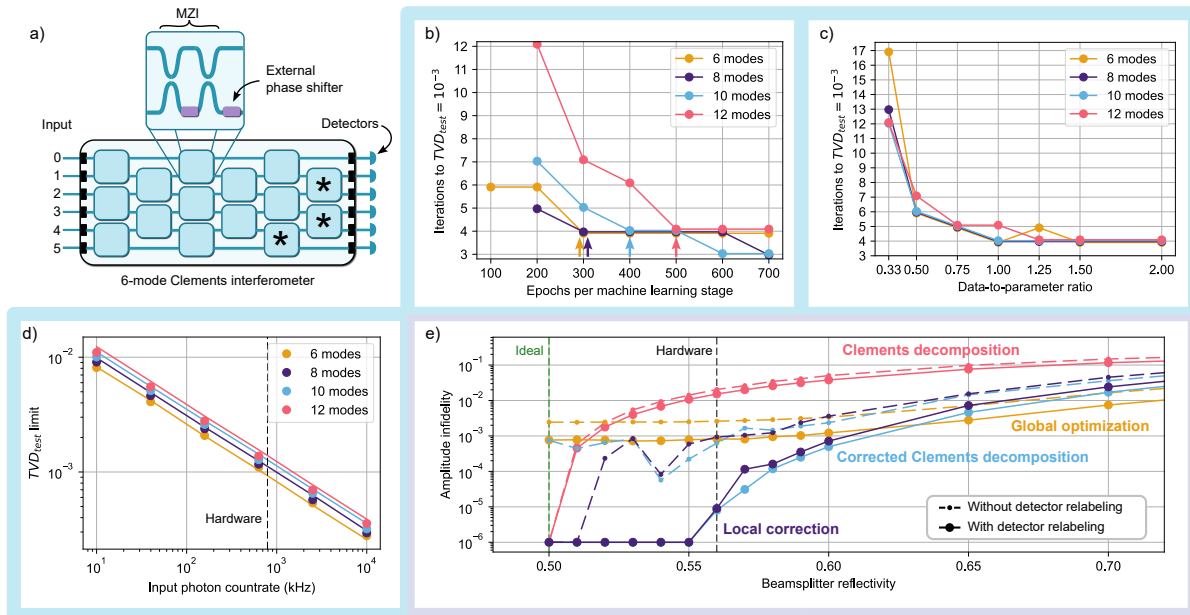
### A. Compilation from Unitary Matrices to Phases and Comparison of Methods

Universal-scheme PICs, like Clements interferometers [see Fig. 3(a)] with a tiled rectangular mesh of Mach–Zehnder interferometers (MZI), can by definition implement any unitary matrix acting on the spatial input modes. The compilation process computes the phase shifts that implement a target unitary matrix. Ideally, the compilation should deliver adequate phases leading to a high-fidelity reconstitution of the target matrix with minimal computation times to guarantee fast PIC reconfiguration times. While deterministic compilation methods achieve fast computation with a minimal overhead, another point to consider is that deterministic compilations are generally tied to a specific mesh, e.g., Clements interferometers. In contrast, gradient descent based-methods

work on any interferometer mesh. The optimized cost function of gradient-based methods may be in addition modified to compute phase shifts while adhering to additional specific constraints, for instance, to reduce the overall power consumption.

- a. *Comparison of compilation methods.* The “Clements decomposition” is the canonical deterministic compilation method for Clements interferometers [26], but assumes a uniform beamsplitter reflectivity value of 0.5. To cope with beamsplitter reflectivity errors, the “corrected Clements decomposition” [22] takes into account reflectivity errors while following the algorithm of the standard Clements decomposition. The “local correction” compiler [21] adjusts deterministically the phases yielded by the Clements decomposition by looking at each Clements interferometer unit cell individually. The gradient based-compilation method, denoted “global optimization,” optimizes all the phases globally to maximize the fidelity of the implemented matrix with respect to the target (see Supplement 1 F).

We compare the different methods in simulations on 100 Haar-random target unitary matrices as shown on Fig. 3(e) (dashed lines). The simulation scenario is a 12-mode Clements interferometer with a uniform beamsplitter reflectivity error, in agreement with the features of our hardware (see Section 5). This choice is also relevant for PICs in general because beamsplitters tend to exhibit spatially correlated



**Fig. 3.** Evaluating the photonic integrated circuit characterization method and the imperfection mitigation through simulation benchmarking. (a) The characterization process and imperfection mitigation are simulation benchmarked on universal-scheme PICs with a Clements mesh with 6 (drawn here) up to 12 modes. Blue lines: waveguides. Blue square: mesh unit cell. The unit cell consists of a Mach–Zehnder interferometer (MZI) followed by an external phase shifter (see inset). Purple rectangles: reconfigurable phase shifter. Waveguides closely spaced form a beamsplitter. Unit cells marked with a star (\*) do not feature an external phase shifters. (b) to (d): each point is a single simulation run. (b) Impact of the number of gradient descent epochs per ML stage (see Section 3.C) on the number of required (ML +  $\phi$ -IFM) iterations to characterize a PIC.  $\text{TVD}_{\text{test}}$  evaluates the accuracy of the model predictions (see Supplement 1 D). (c) Number of (ML +  $\phi$ -IFM) iterations against the data-to-parameter ratio. (d) Effect of photon counting noise on the learning process. Lowest reached  $\text{TVD}_{\text{test}}$  as a function of the input single-photon count rate, assuming a detection integration time of 1 s with single-photon detectors. Continuous lines indicate the theoretical threshold. The black dotted line indicates the photon count rate for the experimental validation of our characterization process (see Section 5). Details on the simulation benchmarking of the characterization protocol are in Section 3.F. (e) Simulated comparison of compilation methods (see Section 4.A). Average amplitude infidelity ( $= 1 - \text{amplitude fidelity}$ , see Supplement 1 D.3) as a function of uniform beamsplitter reflectivity for a 12-mode Clements interferometer evaluated on 100 Haar-random unitary matrices. Dashed lines with small dots: standard method. Continuous line with big dots: method with prior detector relabeling. Red: Clements decomposition [26]. Light blue: Clements decomposition with corrected unit cell [22]. Purple: Local deterministic phase correction [21]. Yellow: Global phase optimization. Green dashed line: ideal reflectivity value for a Clements interferometer. Black dashed line: average reflectivity on our hardware (see Section 5). Value zero is clipped to  $10^{-6}$ .

reflectivity values [42]. Consequently the effects of deterministic errors prevail over random fabrication errors in practice [9,21]. Each target unitary matrix is compiled into phases taking into account the uniform reflectivity value. The computed phases are then applied on the simulated imperfect PIC. All simulations were carried out with the `Perceval` photonic circuit simulator [43]. For each Haar-random matrix, the amplitude fidelity to the target (see Supplement 1 D.3) of the implemented matrix is computed. We observe that the Clements decomposition is strongly affected by beamsplitter reflectivity deviating from the ideal 0.5 value. Deterministic corrected methods “local correction” and “corrected Clements decomposition” achieve very similar good results, while “global optimization” shows an advantage for high reflectivity error.

- b. *Introducing prior detector relabeling.* We introduce in our compilation process a detector relabeling step before the phases computation, which significantly increases the fidelity of all of the presented methods [continuous lines on Fig. 3(e)]. Intuitively, relabeling detectors for a 2-mode circuit is an easy imperfection mitigation technique. Indeed, when considering, e.g., an MZI with uniform beamsplitter reflectivity values differing from 0.5, the *cross* configuration is not feasible, but the *bar* configuration is (see Supplement 1 C.1 for definitions). Permuting the labels of the two output detectors and setting the MZI to the *bar* configuration yield a perfect measured *cross* configuration. On the full circuit level, obtaining the detector relabeling which maximizes the fidelity is less obvious. Iterating on all  $m!$  possible permutations is unpractical; hence, we randomly sample 32 random permutations, which yields a significant improvement as illustrated in Fig. 3(e). Detector relabelling has also been used in [22] to decrease PIC power consumption.

We discuss in Supplement 1 G a method for mitigating inhomogeneous input and output transmissions.

## B. Thermal Crosstalk Compensation

Knowing the phases  $\vec{\phi}$  to apply on the PIC phase shifters, the phase-voltage matrix equation [Eq. (2)] has to be solved for the voltages. Non-linear matrix equations are notoriously hard to solve analytically. We implement an iteration-based solver detailed in Supplement 1 H to approximate a solution. The returned voltage configuration readily compensates thermal crosstalk because of the matrix formulation of the equation and implements the target phases within 0.1 mrad precision as set by our convergence threshold.

## 5. APPLYING THE PROCESS ON A 12-MODE UNIVERSAL-SCHEME PIC

We validate our process experimentally on a physical 12-mode universal-scheme PIC with a Clements mesh [44]. It features 126 reconfigurable thermo-optic phase shifters and 132 directional couplers. This is one of the biggest available PICs in terms of number of on-chip components (see table in Supplement 1 A). This showcases that our method scales beyond the few-phase-shifters case. We use a single-photon source, addressing only even input ports, and single-photon detectors on all output ports. It would be equivalent to use a continuous-wave laser and powermeters, but

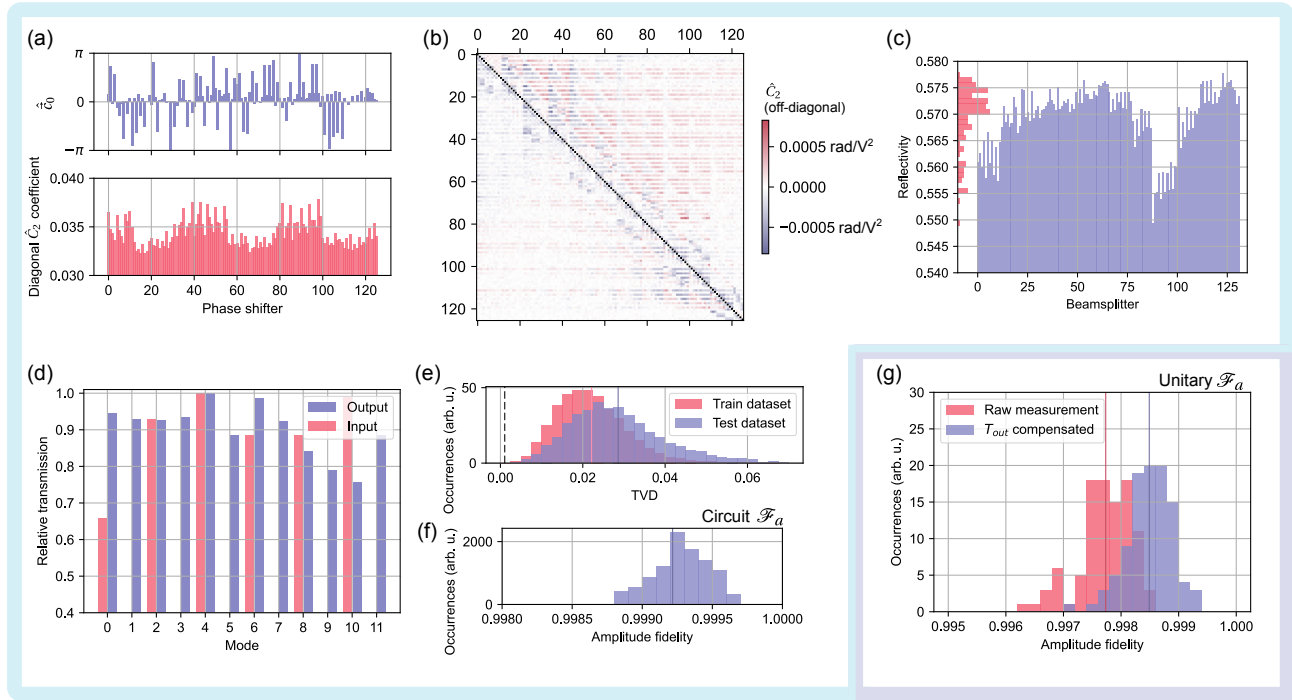
the measurement noise may be greater. We characterize the phase-voltage relation assumed to be of the form  $\vec{\phi} = C_2 \cdot \vec{V}^{\odot 2} + \vec{c}_0$ , beamsplitter reflectivity errors, and input/output transmissions. The entire characterization process, including the data sample acquisition, took approximately 18 h. See Figs. 4(a)–4(d) for characterization results. In particular, the estimated  $\hat{C}_2$  matrix on Fig. 4(b) has predominant values around its diagonal, meaning the estimated crosstalk is as expected weaker between distant components. The beamsplitter reflectivity values on Fig. 4(c) are strongly deviating from the ideal 0.5 value, because the single-photon source emission wavelength and the PIC fabrication wavelength are 11 nm apart. The digital replica stabilizes at an average difference between its predictions and the test data samples  $\text{TVD}_{\text{test}} \approx 2.9\%$ , which is one order of magnitude higher than the expected value of 0.1% [see Fig. 3(d)]. This discrepancy hints towards additional PIC imperfections not taken into account by the virtual replica. We experimentally assess the accuracy of the characterization by implementing 100 random phase configurations  $\vec{\phi}$  on the PIC phase shifters. For each configuration, the amplitude fidelity (see Supplement 1 D.3) is measured between the implemented matrix and the expected matrix, generated from the learned beamsplitter errors and output transmissions. We call this metric the circuit amplitude fidelity and measure  $\mathcal{F}_a = 99.92(2)\%$  [see Fig. 4(f)], where the error bar is one standard deviation. In general, we empirically observe the relation  $\mathcal{F}_a \approx 1 - \text{TVD}_{\text{test}}^2$  between the measured amplitude fidelity and  $\text{TVD}_{\text{test}}$ .

We then measure the average amplitude fidelity with respect to 100  $12 \times 12$  Haar-random unitary matrices [see Fig. 4(g)]. We compile the matrices into phases using the “local correction” method (see Section 4.A) with detector relabeling based on the estimated beamsplitter reflectivity values  $\hat{R}$ . We obtain a unitary amplitude fidelity  $\mathcal{F}_a = (99.77 \pm 0.04)\%$ , which represents the highest value recorded for Clements interferometers to this date (see table in Supplement 1 A). Correcting for the estimated output transmissions  $\hat{T}_{\text{out}}$ , we arrive at a more precise evaluation of the amplitude fidelity of the actual implemented matrix on the PIC, with value  $\mathcal{F}_a = (99.85 \pm 0.04)\%$ . Both values are to our knowledge the highest reported in the literature, obtained on the most complex PIC characterized with machine learning so far (see table in Supplement 1 A).

## 6. DISCUSSION AND OUTLOOK

Our characterization method combines machine learning with a clear-box approach, modeling both the physical PIC to characterize and its imperfections. We thus overcome accuracy limitations imposed by PIC characterization processes based solely on interference fringe measurements. Our method requires significantly lower amounts of training samples and computational power than neural network-based black- and grey-box methods (see table in Supplement 1 A) to train the model. Finally, our method does not depend on the interferometer structure and can therefore be used to handle complex interferometer meshes and large amounts of parameters, which ensures scalability.

We validate our characterization and modeling method on a 12-mode Clements interferometer with a record value of 0.08% amplitude infidelity between the measured PIC behavior and virtual replica model predictions.



**Fig. 4.** Experimental validation of our photonic chip characterization protocol on a 12-mode universal-scheme interferometer. The phase-voltage relation of the PIC is of the form  $\vec{\phi} = C_2 \cdot \vec{V}^{\text{O}2} + \vec{c}_0$  [see Eq. (2)]. The top plot in (a) depicts the estimated values of the passive phases vector  $\hat{c}_0$ , and the bottom plot the diagonal elements of the matrix  $\hat{C}_2$ . The off-diagonal elements of  $\hat{C}_2$  are displayed in (b). They represent thermal crosstalk between phase shifters. (c) Reflectivity of each individual on-chip beamsplitter. Histogram of the values on the plot y-axis. (d) The bar plot indicates the relative optical input and output transmissions of the PIC, normalized such that the maximum is 1 for each set. Notice that we only use 6 inputs of the PIC. (e) Histogram of total variation distance (TVD) that is the difference between the fully trained virtual replica predictions and the training/test dataset output light intensity distributions (see Supplement 1 D for definition). Vertical lines indicate the average for the train dataset (2.2%) and for the test dataset (2.9%). (f) Histogram of measured circuit amplitude fidelity values for 100 random phase configurations implemented on the physical PIC (not Haar-random). For each phase configuration, the amplitude fidelity between the expected and the measured amplitude matrix is acquired. The expected matrix is generated by using the beamsplitter reflectivity values  $\hat{R}$  and output transmissions  $\hat{T}_{\text{out}}$  estimated by the virtual replica. The voltages to apply are computed by solving the learned phase-voltage relation. The vertical line indicates the average amplitude fidelity of 99.93%. (g) Histogram of measured unitary amplitude fidelity values for 100  $12 \times 12$  Haar-random target unitary matrices. We post-process the measurement to compensate for output transmissions. Vertical lines indicate averages of 99.77% and 99.85%. See Section 5.

Furthermore, we have compared and enhanced imperfection mitigation methods. We demonstrate optimal chip control by combining imperfection mitigation with our PIC characterization protocol, allowing us to implement unitary matrices with 0.23% amplitude infidelity, which is to our knowledge the best value in the literature. We emphasize that in addition we obtained it on one of the most complex available PICs.

Our process is straightforward to implement experimentally and can be performed with only a continuous-wave laser and powermeters. In addition, our strategy can be tailored to model other on-chip components and imperfections, extending its applicability to various photonic circuits and systems. To reach even better PIC control, one route is to further investigate the modeling and characterization of other types of PIC imperfections. The development of faster compilation methods that account for and effectively mitigate these additional imperfections will also push the boundaries of PIC optimal control in the context of increasingly intricate and large interferometer meshes.

The clear-box approach finally guarantees transparency of the characterization process and holds the promise for improved fabrication processes by accurately probing individual characteristics of photonic devices that are typically hard to reach, like thermal crosstalk or individual beamsplitter reflectivity values.

The increased reliability of photonic devices presents a transformative opportunity to rise to the current technological challenges in telecommunications, data processing, sensing, and quantum information processing. In particular, photonic quantum computing stands to reap substantial benefits from increased PIC accuracy, by achieving higher qubit fidelities. These advances open the door to efficient near-term quantum processors with demonstrations of boson sampling with reconfigurable circuits and graph problem solvers and lay the foundations for fault-tolerant quantum computing harnessing integrated photonic components.

## 7. METHODS

### A. Simulation Benchmarking the Characterization Process

We simulate Clements interferometers with 6, 8, 10, and 12 modes featuring a phase-voltage relation of the form  $\vec{\phi} = C_2 \cdot \vec{V}^{\text{O}2} + \vec{c}_0$ . The parameters of the simulated physical device that are to be estimated by the virtual replica are set as follows. The matrix  $C_2$  is generated deterministically: the diagonal coefficients are set to 0.034, the off-diagonal elements are computed following  $a/d^2$  where  $d$  is the distance between components on a PIC schematic

as shown in Fig. 3, and  $a$  is set such that the coefficient relating two phase shifters belonging to the same unit cell is 1% of the diagonal coefficient. These are realistic crosstalk values. For phase shifters in Mach–Zehnder interferometers, the coefficients describing received crosstalk is either positive or negative depending on the position of the heat emitter. The passive phase vector  $\vec{c}_0$  is generated following a Gaussian distribution of mean 0 and standard deviation 0.7 rad. The reflectivity of the beamsplitters is generated according to a Gaussian distribution of mean 0.56 and standard deviation 0.007, closely agreeing with our experimental hardware. Output transmissions are chosen uniformly between 0.7 and 1. The maximum voltage threshold is set to 14 V, allowing all the phase shifters to achieve a  $2\pi$  phase sweep. The initial values before training of the virtual replica are set to start with a zero vector for the passive phases  $\hat{c}_0$  and a zero matrix for the crosstalk matrix  $\hat{C}_2$ . The estimated output transmissions are configured to be 1, and the estimated beamsplitter reflectivity values are set to 0.5. For each voltage and phase sweep during the characterization, 15 data points are acquired. The learning rates are the following:  $\hat{C}_2$ :  $10^{-5}$ ,  $\hat{R}$ :  $10^{-3}$ , and  $\hat{T}_{\text{out}}$ :  $10^{-3}$ . The learning rate scheduling factor is set to 0.7.

## B. Simulation Benchmarking the Imperfection Mitigation

For the global optimization compilation, 32 simultaneous phase computations via gradient descent are performed simultaneously following Supplement 1 F. Optimization stops for  $10^{-6}$  relative variation of both loss function and optimization parameters. For all methods, 32 random relabelings are tested.

## C. Experimental Validation

We use a single-photon source based on a quantum dot embedded in a cavity, emitting single-photons at 929 nm. A pump laser excites the source with an 80 MHz pulse rate. The 2-photon emission probability is estimated at 1%. The stream of emitted single photons is divided into 6 by an active demultiplexer. Fiber delays synchronize the photons such that they enter the 12-mode PIC at the same time. Photons are injected in even input port numbers. An automated mechanical shutter system ensures that only one input port is addressed at a time. The PIC features 126 thermo-optic phase shifters and 132 beamsplitters with fabrication wavelength 940 nm. The PIC reconfiguration time is 2 s when reconfiguring all the phase shifters. To speed up the process, we update only voltages that differ by more than 0.1 mV (leading to phase errors below  $1 \times 10^{-9}$  rad). We acquire 15 points per phase shifter during voltage and phase sweeps. Photons are detected by 12 superconducting nanowire single-photon detectors, and countrates are measured by a correlator. We use a countrate integration time of 1 s. We use the same learning rates and scheduling factor as for the simulation benchmarking to train the virtual replica model. Electric crosstalk is not considered because of an already built-in compensation subroutine by the manufacturer. The characterization protocol carried out the following sequence of stages: V-IFM  $\rightarrow$  ML  $\rightarrow$   $\phi$ -IFM  $\rightarrow$  ML  $\rightarrow$  ITM. A total of 16,500 training and 4125 test samples have been acquired in 17 h (the data-to-parameter ratio is 1.03). The V-IFM and  $\phi$ -IFM stages take about 30 min. The protocol is run on a i7-12700 2.1 GHz processor, allowing the ML stages with 500 epochs each to

be completed in under 10 min. The characterization process took thus about 18 h. We measure the unitary amplitude fidelity by compiling 100 Haar-random unitary matrices into phases using the “local correction” method with detector relabeling (see Section 4.A). We post-process the acquired matrices by dividing each row by the estimated output transmission. The columns are then normalized such that their elements squared sum to 1.

**Funding.** HORIZON EUROPE European Innovation Council (190188855); Agence Nationale de la Recherche (ANR-23-PETQ-0013).

**Acknowledgment.** The authors would like to express their gratitude to Elianor Hoffmann, Thomas Liege, and Francesca Zanichelli for their valuable contributions. We would like to acknowledge Simone Piacentini and Alexia Salavrakos for providing valuable feedback on the manuscript. We sincerely thank Nicolas Heurtel, Alexandre Mercier, Rawad Mezher, and Mathias Pont for fruitful discussions. This work has been co-funded by the European Commission as part of the EIC accelerator program. We acknowledge funding from the Plan France 2030 through the project. Authors are listed in the main title order. Characterization protocol: A.F., O.F., & J.S. Protocol generation algorithm: A.F. & O.F. Imperfection mitigation: A.F. & J.S. Experimental investigation: A.F., N.M., J.S., & N.B. Experimental validation: A.F., N.M., & J.S. Manuscript writing and visualization: A.F. & N.B. Project supervision: N.M., J.S., & N.B.

**Disclosures.** A patent has been filed listing A.F., J.S., N.M., and N.B. as inventors.

**Data availability.** The codebase and data generated as part of this work are available to research groups upon reasonable request from the corresponding author.

**Supplemental document.** See Supplement 1 for supporting content.

## REFERENCES

1. J. Wang, F. Sciarrino, A. Laing, *et al.*, “Integrated photonic quantum technologies,” *Nat. Photonics* **14**, 273–284 (2020).
2. H.-S. Zhong, H. Wang, Y. H. Deng, *et al.*, “Quantum computational advantage using photons,” *Science* **370**, 1460–1463 (2020).
3. L. S. Madsen, F. Laudenbach, M. F. Askarani, *et al.*, “Quantum computational advantage with a programmable photonic processor,” *Nature* **606**, 75–81 (2022).
4. W. Luo, L. Cao, Y. Shi, *et al.*, “Recent progress in quantum photonic chips for quantum communication and internet,” *Light Sci. Appl.* **12**, 175 (2023).
5. A. Fyrillas, B. Bourdoncle, A. Maños, *et al.*, “Certified randomness in tight space,” *arXiv*, arXiv:2301.03536 (2023).
6. E. Polino, M. Riva, M. Valeri, *et al.*, “Experimental multiphase estimation on a chip,” *Optica* **6**, 288–295 (2019).
7. W. Bogaerts, D. Pérez, J. Capmany, *et al.*, “Programmable photonic circuits,” *Nature* **586**, 207–216 (2020).
8. H. Zhou, J. Dong, J. Cheng, *et al.*, “Photonic matrix multiplication lights up photonic accelerator and beyond,” *Light Sci. Appl.* **11**, 30 (2022).
9. S. Banerjee, M. Nikdast, and K. Chakrabarty, “Characterizing coherent integrated photonic neural networks under imperfections,” *J. Lightwave Technol.* **41**, 1464–1479 (2023).
10. J. Mower, N. C. Harris, G. R. Steinbrecher, *et al.*, “High-fidelity quantum state evolution in imperfect photonic integrated circuits,” *Phys. Rev. A* **92**, 032322 (2015).
11. D. Pérez-López, A. López, P. DasMahapatra, *et al.*, “Multipurpose self-configuration of programmable photonic circuits,” *Nat. Commun.* **11**, 6359 (2020).
12. I. V. Kondratyev, V. V. Ivanova, S. A. Zhuravitskii, *et al.*, “Large-scale error-tolerant programmable interferometer fabricated by femtosecond laser writing,” *arXiv*, arXiv:2308.13452 (2023).
13. R. Hamerly, S. Bandyopadhyay, and D. Englund, “Accurate self-configuration of rectangular multiport interferometers,” *Phys. Rev. Appl.* **18**, 024019 (2022).
14. A. Laing and J. L. O’Brien, “Super-stable tomography of any linear optical device,” *arXiv*, arXiv:1208.2868 (2012).
15. S. Rahimi-Keshari, M. A. Broome, R. Fickler, *et al.*, “Direct characterization of linear optical networks,” *Opt. Express* **21**, 13450–13458 (2013).

16. I. Dhand, A. Khalid, H. Lu, *et al.*, “Accurate and precise characterization of linear optical interferometers,” *J. Opt.* **18**, 035204 (2016).
17. N. Spagnolo, E. Maiorino, C. Vitelli, *et al.*, “Learning an unknown transformation via a genetic approach,” *Sci. Rep.* **7**, 14316 (2017).
18. V. Cimini, E. Polino, M. Valeri, *et al.*, “Calibration of multiparameter sensors via machine learning at the single-photon level,” *Phys. Rev. Appl.* **15**, 044003 (2021).
19. A. Youssry, Y. Yang, R. J. Chapman, *et al.*, “Experimental graybox quantum system identification and control,” *arXiv*, arXiv:2206.12201 (2023).
20. R. Burgwal, W. R. Clements, D. H. Smith, *et al.*, “Using an imperfect photonic network to implement random unitaries,” *Opt. Express* **25**, 28236–28245 (2017).
21. S. Bandyopadhyay, R. Hamerly, and D. Englund, “Hardware error correction for programmable photonics,” *Optica* **8**, 1247–1255 (2021).
22. S. P. Kumar, L. Neuhaus, L. G. Helt, *et al.*, “Mitigating linear optics imperfections via port allocation and compilation,” *arXiv*, arXiv:2103.03183 (2021).
23. Y. Zhu, G. L. Zhang, B. Li, *et al.*, “Countering variations and thermal effects for accurate optical neural networks,” in *Proceedings of the 39th International Conference on Computer-Aided Design (ACM, 2020)*, pp. 1–7.
24. B. J. Metcalf, J. B. Spring, P. C. Humphreys, *et al.*, “Quantum teleportation on a photonic chip,” *Nat. Photonics* **8**, 770–774 (2014).
25. M. Reck, A. Zeilinger, H. J. Bernstein, *et al.*, “Experimental realization of any discrete unitary operator,” *Phys. Rev. Lett.* **73**, 58–61 (1994).
26. W. R. Clements, P. C. Humphreys, B. J. Metcalf, *et al.*, “Optimal design for universal multiport interferometers,” *Optica* **3**, 1460–1465 (2016).
27. B. A. Bell and I. A. Walmsley, “Further compactifying linear optical unitaries,” *APL Photon.* **6**, 070804 (2021).
28. N. Maring, A. Fyrrillas, M. Pont, *et al.*, “A general-purpose single-photon-based quantum computing platform,” *arXiv*, arXiv:2306.00874 (2023).
29. S. Bandyopadhyay, A. Sludds, S. Krastanov, *et al.*, “Single chip photonic deep neural network with accelerated training,” *arXiv*, arXiv:2208.01623 (2022).
30. E. A. J. Marcatili, “Dielectric rectangular waveguide and directional coupler for integrated optics,” *Bell Syst. Tech. J.* **48**, 2071–2102 (1969).
31. N. J. Russell, L. Chakhmakhchyan, J. L. O’Brien, *et al.*, “Direct dialling of Haar random unitary matrices,” *New J. Phys.* **19**, 033007 (2017).
32. Y. Yang, Y. Ma, H. Guan, *et al.*, “Phase coherence length in silicon photonic platform,” *Opt. Express* **23**, 16890–16902 (2015).
33. N. C. Harris, Y. Ma, J. Mower, *et al.*, “Efficient, compact and low loss thermo-optic phase shifter in silicon,” *Opt. Express* **22**, 10487–10493 (2014).
34. M. Milanizadeh, D. Aguiar, A. Melloni, *et al.*, “Canceling thermal cross-talk effects in photonic integrated circuits,” *J. Lightwave Technol.* **37**, 1325–1332 (2019).
35. M. Dong, G. Clark, A. J. Leenheer, *et al.*, “High-speed programmable photonic circuits in a cryogenically compatible, visible–near-infrared 200 mm CMOS architecture,” *Nat. Photonics* **16**, 59–65 (2022).
36. M. Li, J. Ling, Y. He, *et al.*, “Lithium niobate photonic-crystal electro-optic modulator,” *Nat. Commun.* **11**, 4123 (2020).
37. X. Qiang, X. Zhou, J. Wang, *et al.*, “Large-scale silicon quantum photonics implementing arbitrary two-qubit processing,” *Nat. Photonics* **12**, 534–539 (2018).
38. J. C. Adcock, C. Vigliar, R. Santagati, *et al.*, “Programmable four-photon graph states on a silicon chip,” *Nat. Commun.* **10**, 3528 (2019).
39. J. M. Arrazola, V. Bergholm, K. Brádler, *et al.*, “Quantum circuits with many photons on a programmable nanophotonic chip,” *Nature* **591**, 54–60 (2021).
40. A. Paszke, S. Gross, F. Massa, *et al.*, “PyTorch: an imperative style, high performance deep learning library,” in *Advances in Neural Information Processing Systems (Curran Associates, Inc., 2019)*, Vol. **32**, pp. 8024–8035.
41. D. P. Kingma and J. Ba, “Adam: a method for stochastic optimization,” *arXiv*, arXiv:1412.6980 (2017).
42. Z. Lu, J. Jhoja, J. Klein, *et al.*, “Performance prediction for silicon photonics integrated circuits with layout-dependent correlated manufacturing variability,” *Opt. Express* **25**, 9712–9733 (2017).
43. N. Heurtel, A. Fyrrillas, G. de Gliniasty, *et al.*, “Perceval: a software platform for discrete variable photonic quantum computing,” *Quantum* **7**, 931 (2023).
44. C. Taballione, R. van der Meer, H. J. Snijders, *et al.*, “A universal fully reconfigurable 12-mode quantum photonic processor,” *Mater. Quantum Technol.* **1**, 035002 (2021).