nature photonics

Article

https://doi.org/10.1038/s41566-024-01403-4

A versatile single-photon-based quantum computing platform

In the format provided by the authors and unedited

S-I. SINGLE-PHOTON SOURCE

The single-photon source is based on a gated InGaAs quantum dot (QD) embedded in a monolithic micropillar cavity [1] cooled down to 5K. It is optically excited using the near-resonant LA-phonon-assisted excitation scheme [2] at an 80 MHz rate. The neutral QD emits linearly-polarized single photons at 928 nm with a lifetime of 92 ps and a 55% first lens brightness. After spectrally filtering the remaining excitation laser using free-space spectral bandpass filters (FWHM=800 pm) and coupling into a single mode fiber, the single-photon source device shows a low multiphoton component with a single-photon purity $\mathcal{P} = 1 - g^{(2)}(0) > 99\,\%$ and a 2-photon indistinguishability $M_s > 94\,\%$.

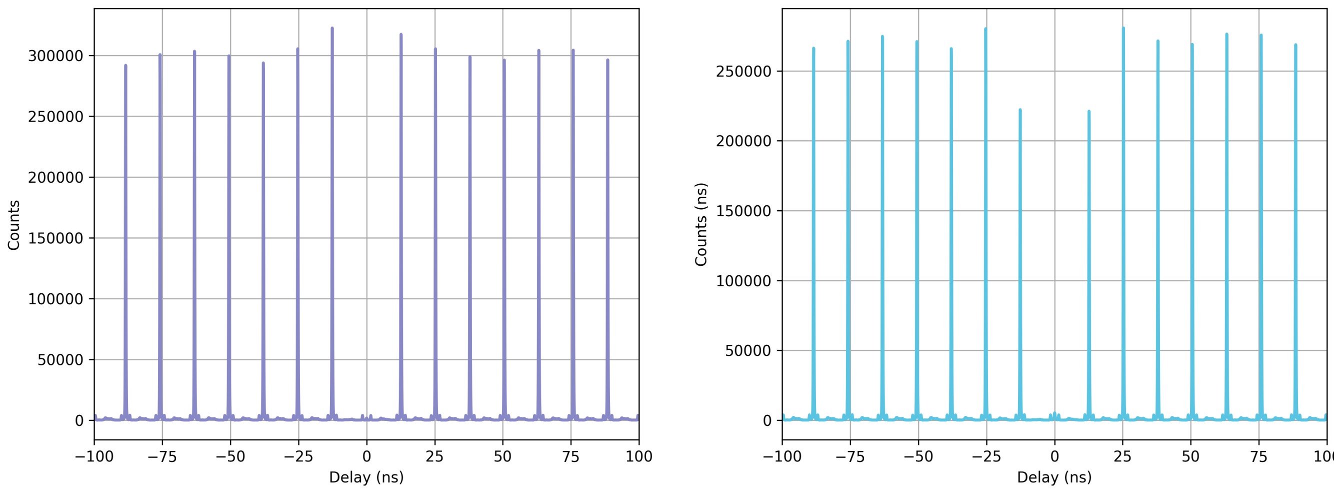

FIG. S1: **The single-photon source delivers pure and indistinguishable single photons.** Single-photon purity is quantified by the normalized second-order correlation function $g^{(2)}(0) = (7.32 \pm 0.07) \times 10^{-3}$ (left) and the Hong-Ou-Mandel visibility $V_{\mathrm{HOM}} = 0.9296 \pm 0.0003$ (right) yielding a corrected 2-photon indistinguishability of $M_s = 0.9438 \pm 0.0003$ [3]. The histograms were integrated for 5 s and the peak integration window is 1 ns. The integrated values are obtained with no background subtraction.

S-II. OPTICAL SETUP

The overall transmission of the optical setup (see Fig. 1.a in main text) is characterized by the transmission or efficiency of each module. A precise loss budget of the optical setup is provided in Tab. S1.

Module	Transmission/Efficiency	Near-term targets
First lens brightness	55 %	80% [4]
Single-mode fiber coupling	70 %	85% [5]
Spectral Filtering module	75 %	>82%[*]
Demultiplexer	70 %	>80%[*]
PIC insertion and transmission	45 %	70% [6]
SNSPDs	92 %	>95%[**]
Total	$8.4 \pm 0.2\,\%$	27%
Pump laser repetition rate	80 MHz	320 MHz [7]
6-photon countrate	4 Hz	∼35 kHz (computed)
12-photon countrate	200 nHz (computed)	∼10 Hz (computed)

TABLE S1: Current loss budget of the optical setup (Transmission/Efficiency) and near-term targets. [*] Optical module in development at Quandela. [**] Commercially available products

Experimentally, the input state of the PIC is fixed to $|101010101010\rangle$. To apply a unitary matrix U acting on an arbitraty input state $|\mathrm{IN}\rangle$ using the universal-scheme 12-mode photonic chip, we first compute the permutation matrix U_{perm} so that $U_{\mathrm{perm}}\,|101010101010\rangle = |\mathrm{IN}\rangle$. The matrix $U' = U \cdot U_{\mathrm{perm}}$ is also a 12×12 unitary and thus the universal photonic chip can implement U'. For input states with less than 6 photons, mechanical shutters are used to effectively reduce the number of photons at the input of the PIC.

Ascella's detection module is composed of superconducting nanowire single-photon detectors showing an average detection efficiency of 92 % and dark countrates under 20 Hz. Detection events are digitalized using a time-to-digital converter (Swabian instruments) and post-processed for sampling N-photon ($1 \leq N \leq 6$) coincidences between all 12 detectors within a 1 ns coincidence window.

The polarization in the delay fibers is optimized automatically to ensure a maximal transmission to the polarization-selective photonic circuit. The excitation laser power is similarly stabilized, ensuring optimal brightness and photon purity.

S-III. MULTIPHOTON INTERFERENCE CHARACTERIZATION

In this section, we measure all pairwise 2-photon indistinguishabilities, and the *genuine* 4- and 6-photon indistinguishability of the input state of Ascella, i.e. the probability that the n ($n = 4$ or $n = 6$) photons are identical [8, 9]. This initial characterization of our input state sets the basis for future practical applications on our platform. It also allows us to fine-tune our simulator to be able to reproduce experimental results with a good agreement.

We first measure the 2-photon pairwise indistinguishability M_{ij} $(i, j = 1, ..., 6)$ between all $C_6^2 = 15$ photon pairs. As shown in Fig S4, the reconfigurable chip is set to successively connect each pair with a Mach-Zehnder interferometer (MZI). We vary the internal phase of the MZI to measure the correlated $\phi = \pi/2$ (uncorrelated $\phi = \pi$) 2-photon coincidences at zero time delay, which gives access to the visibility of the 2-photon interference fringe. The imperfect single-photon purity [3] of our QD-source ($g^{(2)}(0) = 0.0075$), is taken into account to compute M_{ij} for all 15 pairs. The values of M_{ij} for each of the 15 pairs is reported in the *indistinguishability matrix* $\mathcal{M}$ where $\mathcal{M}_{ij} = M_{ij}$ and $\mathcal{M}_{ii} = 1$.

$$\mathcal{M} = \begin{bmatrix} 1 & 0.942 & 0.921 & 0.924 & 0.917 & 0.914 \\ & \ddots & 0.935 & 0.925 & 0.924 & 0.919 \\ & & \ddots & 0.932 & 0.911 & 0.925 \\ & & & \ddots & 0.943 & 0.941 \\ & & & & \ddots & 0.942 \\ & & & & & 1 \end{bmatrix}$$

The genuine N-photon indistinguishability is the probability p_N that all N photons are identical. To quantify experimentally the genuine 4- and 6-photon indistinguishability of our input multiphoton state we implement on the reconfigurable QPU 8-mode and 12-mode versions of the cyclic interferometer (first introduced in [9]) whose general layout is presented in Fig. S2.a. First, the single internal phase of the interferometer α is set to 0 (2π). Each odd input port of the interferometer is fed with single photons. We detect the output states corresponding to one photon per pair of output ports ($2k$, $2k+1$) (see Fig. S2.a). In Fig. S2.b-c we present the experimental output distribution for all outputs corresponding to constructive (orange) and destructive (blue) n-photon interference ($N = 4$ for b. and $N = 6$ for c.). The visibility of the interference fringe is the genuine N-photon indistinguishability. We experimentally measure $p_4 = 0.85 \pm 0.002$ 4-photon indistinguishability for photons $\{1, 2, 3, 4\}$ and $p_6 = 0.76 \pm 0.02$ 6-photon indistinguishability. This work constitutes the first experimental realization of this protocol for 6 photons, and sets a new state-of-the-art for genuine 4- and 6-photon indistinguishability.

To further study our ability to drive the internal phase of the interferometer shown in Fig. S2.a, we scan the single internal phase α to measure the full interference fringe visibility for 4-photon interferences with photons $\{1, 2, 3, 4\}$. Note that for each value of the internal phase α we compute the associated unitary matrix using *Perceval* and transpile the circuit to be implemented on the QPU. After normalization of the 4-photon counts, we fit the theoretical interference fringe [9] $p_4 = 1 \pm c_1 \cos(a \cdot \alpha + b)$. In the ideal case we expect $a_{\text{ideal}} = 1$ and $b_{\text{ideal}} = 0$. The experimental data is well fitted with $a = 1.00 \pm 0.01$ and $b = -0.06 \pm 0.08$ rad, which shows that the transpilation can very accurately implement an 8×8 unitary matrix.

S-IV. PHOTONIC CHIP CHARACTERIZATION

We use a machine learning-based process to characterize the photonic chip. The 126 on-chip thermo-optic phase shifters generate heat via the Joule effect, thus they can be modelled by a relation of the form $\vec{\phi} = A\vec{V}^{\odot 2} + \vec{b}$ between

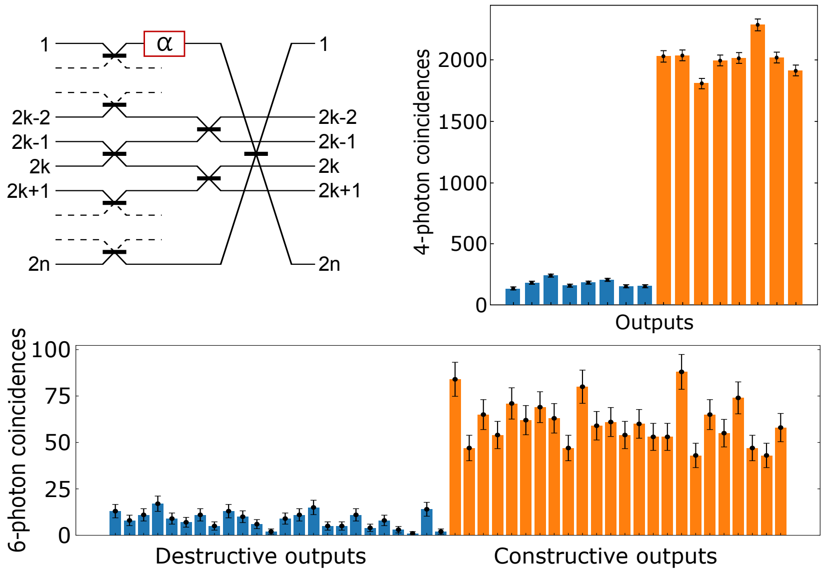

FIG. S2: Genuine indistinguishability of the input multiphoton state. (a) General layout of the multiport interferometer used to measure the probability that the N photons are identical. (b-c) Histogram of 4- (6-) photon outputs that undergo destructive (blue) and constructive (orange) interferences in an 8- (12-) mode version of the interferometer presented in (a) fed with 4 (6) photons. In (b) the genuine 4-photon indistinguishability is $p_4 = 0.85 \pm 0.02$. In (c) the genuine 6-photon indistinguishability is $p_6 = 0.76 \pm 0.02$. Error bars correspond to ± 1 SD of the photon counting statistics.

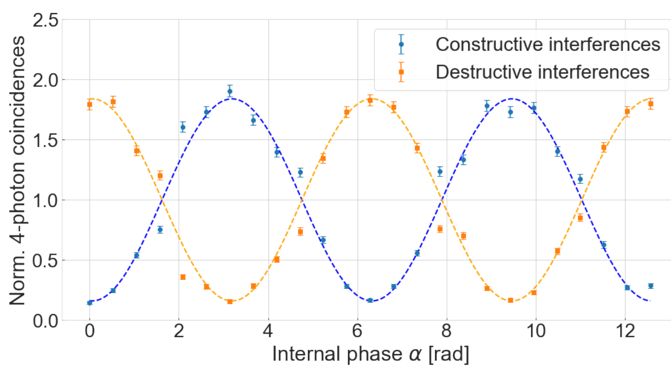

FIG. S3: Total normalized four-photon coincidence rate (sum of all eight output states) for the constructive and destructive outputs, as a function of the internal phase α. Error bars correspond to ± 1 SD of the photon counting statistics.

the vector $\vec{\phi}$ containing all 126 physically implemented phases and the vector $\vec{V}$ corresponding to the 126 applied voltages squared. $^{\odot 2}$ represents element-wise squaring. Off-diagonal elements of the 126×126 matrix A represent thermal crosstalk between phase shifters. A process based on machine-learning techniques optimizes the values of A and $\vec{b}$, which represent ≈ 16000 free parameters to determine. The same process also estimates individual directional coupler reflectivities and relative output losses. We show on Fig. S9 the estimated values. The elements of $\vec{b}$ have values $(0.1 \pm 1.2)\,\mathrm{rad}$, and the diagonal elements of A have values $(0.034 \pm 0.001)\,\mathrm{rad/V^2}$, ensuring that on average a π-phase shift can be achieved by applying around 10 V on a phase shifter. The matrix A seems to show long-range interactions between phase shifters, but these are, in reality, artefacts arising from certain transformations on A that leave the output quantum state unchanged. The directional coupler reflectivities have values $(56.8 \pm 0.6)\,\%$. One can observe regions of low and high reflectivities on Fig. S7, which is a signature of the photonic chip's folding; that is the interferometer is not laid out in a straight manner, but is folded to increase compactity. For the output losses, we notice that mode 10 has a a relative transmission of 75% compared to mode 6. This result was confirmed on two separate detection systems (power meter array and single-photon detectors), hinting that the defect lies in the photonic chip and not in the photon detectors.

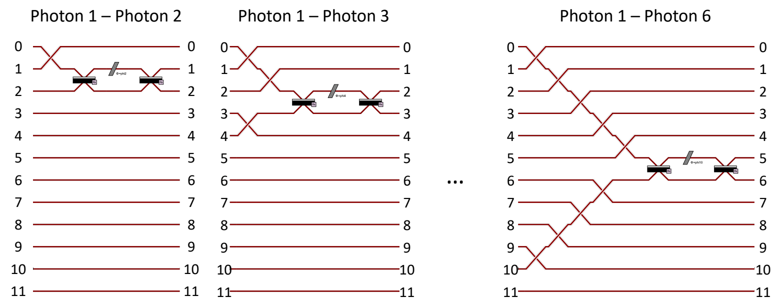

FIG. S4: Examples of the layouts of the photonic circuits used to successively connect each pair of photon inputs to an MZI to measure the pairwise 2-photon indistinguishability.

S-V. BENCHMARKING LOGIC GATES

In this supplementary section, we outline the method of Ref. [10] applied to benchmark 1-, 2-, and 3- qubit gates implemented by Ascella. We then give an explicit example of the method by deriving F_{avg} for the T-gate. Finally, we describe the general approach taken to obtain F_{avg} for multi-qubit gates and discuss the application to the 2- and 3-qubit gates benchmarked in the main text.

A. Symmetry-based benchmarking

Let $|\boldsymbol{i}\rangle := |i_1, \ldots, i_n\rangle$ with $i_j \in \{0,1\}$ for $j \in \{1,\ldots,n\}$ denote an n-qubit computational basis state. A noisy implementation of the gate unitary U is given by $\Lambda \circ \mathbf{U}(.)$, which is a completely positive trace preserving (CPTP) map [11] acting on n-qubit density matrices ρ as $\Lambda \circ \mathbf{U}(\rho) = \Lambda(U\rho U^\dagger)$ where Λ is the noise channel. The average fidelity F_{avg}, over all possible n-qubit states, given Λ corresponding to the noisy application of U is shown in [10] to be

$$F_{\text{avg}}(U) = \frac{1}{2^n(2^n+1)} \sum_{\boldsymbol{i},\boldsymbol{j},\boldsymbol{i'},\boldsymbol{j'}} \left(\alpha^{U^\dagger}_{|\boldsymbol{i'}\rangle\langle\boldsymbol{j'}|;|\boldsymbol{i}\rangle\langle\boldsymbol{j}|} + \alpha^{U^\dagger}_{|\boldsymbol{i'}\rangle\langle\boldsymbol{j'}|;|\boldsymbol{j}\rangle\langle\boldsymbol{i}|}\right) \mathsf{Trace}\left[|\boldsymbol{i}\rangle\langle\boldsymbol{j}|\,\Lambda\circ\mathbf{U}(|\boldsymbol{i'}\rangle\langle\boldsymbol{j'}|)\right], \tag{S1}$$

where

the coefficients $\alpha^{U^\dagger}_{|\boldsymbol{i'}\rangle\langle\boldsymbol{j'}|;|\boldsymbol{i}\rangle\langle\boldsymbol{j}|} := \langle\boldsymbol{i'}|\,U^\dagger\,|\boldsymbol{i}\rangle\,\langle\boldsymbol{j}|\,U\,|\boldsymbol{j'}\rangle \in \mathbb{C}$ depend only on the unitary U. This expression generally includes 2^{4n} terms, corresponding to 2^{4n} measurements. In the worst case, evaluating $F_{\text{avg}}(U)$ as above requires as many measurements as a full process tomography. However, for most gates of interest, a significant number of α coefficients will vanish, hence requiring many fewer measurements to evaluate the sum in Eq. (S1). A more formal proof of Eq. (S1) will appear in [10], but we will now outline the key technical steps needed to arrive at Eq. (S1).

We start by expanding $U|\psi\rangle = \sum_{\boldsymbol{i}} \beta_{\boldsymbol{i}}|\boldsymbol{i}\rangle$ in the basis $\{|\boldsymbol{i}\rangle\}$ for any initial state $|\psi\rangle$, with $\beta_{\boldsymbol{i}} \in \mathbb{C}$. Then, by plugging this into the expression of the final state fidelity $F_\psi(U)$ given in the main text, we obtain

$$F_\psi(U) = \sum_{\boldsymbol{i},\boldsymbol{i'},\boldsymbol{j},\boldsymbol{j'}} \beta^*_{\boldsymbol{i}}\beta^*_{\boldsymbol{i'}}\beta_{\boldsymbol{j}}\beta_{\boldsymbol{j'}}\langle\boldsymbol{i}|\Lambda(|\boldsymbol{j'}\rangle\langle\boldsymbol{i'}|)|\boldsymbol{j}\rangle,$$

and therefore

$$F_{\text{avg}}(U) = \sum_{\boldsymbol{i},\boldsymbol{i'},\boldsymbol{j},\boldsymbol{j'}} \mathbb{E}(\beta^*_{\boldsymbol{i}}\beta^*_{\boldsymbol{i'}}\beta_{\boldsymbol{j}}\beta_{\boldsymbol{j'}})\langle\boldsymbol{i}|\Lambda(|\boldsymbol{j'}\rangle\langle\boldsymbol{i'}|)|\boldsymbol{j}\rangle,$$

where the expectation values of the product of β coefficients is

$$\mathbb{E}(\beta^*_{\boldsymbol{i}}\beta^*_{\boldsymbol{i'}}\beta_{\boldsymbol{j}}\beta_{\boldsymbol{j'}}) := \int d\psi \beta^*_{\boldsymbol{i}}\beta^*_{\boldsymbol{i'}}\beta_{\boldsymbol{j}}\beta_{\boldsymbol{j'}}(\psi).$$

The quantities $\mathbb{E}(\beta^*_{\boldsymbol{i}}\beta^*_{\boldsymbol{i'}}\beta_{\boldsymbol{j}}\beta_{\boldsymbol{j'}})$ are the same as those in the expansion in the operator basis $\{|\boldsymbol{j}\rangle|\boldsymbol{j}^{'}\rangle\langle\boldsymbol{i}^{'}|\langle\boldsymbol{i}|\}$ of

$$\Pi := \int d\psi(|\psi\rangle\langle\psi|)^{\otimes 2}.$$

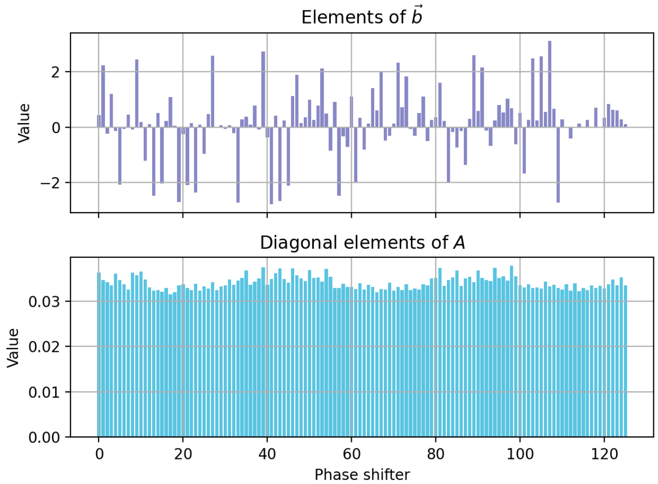

FIG. S5

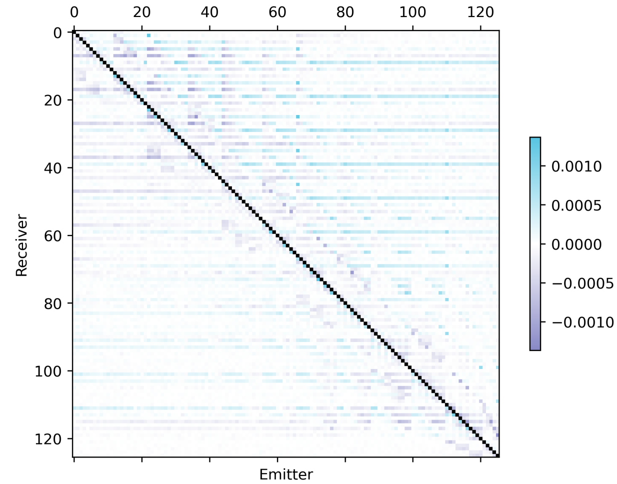

FIG. S6

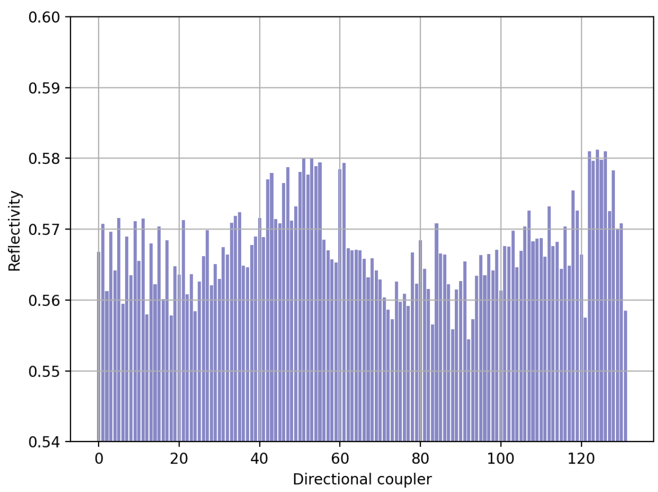

FIG. S7

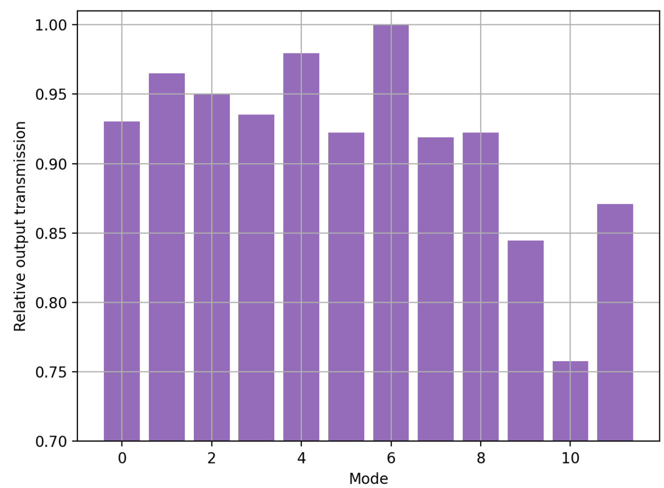

FIG. S8

FIG. S9: **Large-scale photonic chip imperfections estimated by a machine-learning process.** The phase shifter phase-voltage relation is modelled by a matrix relation of the form $\vec{\phi} = A\vec{V}^{\odot 2} + \vec{b}$, where $\vec{\phi}$ and $\vec{V}^{\odot 2}$ are vectors containing respectively the applied phase shifts and squared voltages. We show on S5 the diagonal elements of A and the elements of $\vec{b}$, and on S6 we show the off-diagonal elements of A, which account for thermal crosstalk. S7 represents the values of the reflectivity of the on-chip directional couplers. S8 displays the relative output losses per mode, scaled such that the maximum value is equal to 1.

In addition, Π is proportional to the projector onto the symmetric subspace of $(\mathbb{C}^d)^{\otimes 2}$ with $d = 2^n$ [12], and this projector has a known expansion in the basis $\{|\boldsymbol{j}\rangle|\boldsymbol{j}^{'}\rangle\langle\boldsymbol{i}^{'}|\langle\boldsymbol{i}|\}$ [13]. These key insights allow us to compute the coefficients $\mathbb{E}(\beta^*_{\boldsymbol{i}}\beta^*_{\boldsymbol{i'}}\beta_{\boldsymbol{j}}\beta_{\boldsymbol{j'}})$, to obtain

$$F_{\text{avg}}(U) = \frac{1}{2^n(2^n+1)} \sum_{\boldsymbol{i},\boldsymbol{j}} \big(\langle \boldsymbol{i}|\Lambda(|\boldsymbol{j}\rangle\langle\boldsymbol{i}|)|\boldsymbol{j}\rangle + \langle\boldsymbol{i}|\Lambda(|\boldsymbol{i}\rangle\langle\boldsymbol{j}|)|\boldsymbol{j}\rangle \big).$$

The last steps needed to arrive at Eq. (S1) are to note that $\langle\boldsymbol{i}|\Lambda(|\boldsymbol{j}\rangle\langle\boldsymbol{i}|)|\boldsymbol{j}\rangle = \mathsf{Trace}(|\boldsymbol{j}\rangle\langle\boldsymbol{i}|\Lambda(|\boldsymbol{j}\rangle\langle\boldsymbol{i}|))$, and then to rewrite $\Lambda = \Lambda \circ \mathbf{U} \circ \mathbf{U}^\dagger$, and finally expand $\mathbf{U}^\dagger(|\boldsymbol{j}\rangle\langle\boldsymbol{i}|)$ in the basis $\{|\boldsymbol{i}^{'}\rangle\langle\boldsymbol{j}^{'}|\}$ to identify the α coefficients of Eq. (S1).

Following this method, we can derive exact expressions of F_{avg} for any gate unitary as a discrete sum of a finite number of terms. As outlined in the following sections, the exact form of F_{avg} and the number of non-zero terms will depend on U, and these terms can be evaluated using a set of state preparation and measurement settings.

We note that there is in the literature another technique to evaluate F_{avg} using fewer measurements than full process tomography. The approach experimentally implemented in [14] obtains F_{avg} by first measuring an intermediate quantity called the entanglement fidelity [15–17]. In contrast, the symmetry-based method outlined here and used in the main text bypasses computing this intermediate quantity and directly evaluates F_{avg} as a weighted summation of measurements. As discussed below, we notably find that the symmetry-based benchmarking approach requires half as many measurements to benchmark a CNOT gate as were required in [14].

B. Average fidelity of a T-gate

As an illustration of the method in action, we explicitly compute $F_{\text{avg}}(U)$ for the case of the T-gate, which is a very important gate in the context of magic state distillation protocols for fault-tolerant quantum computing [18]. This gate is defined by the unitary transformation

$$T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix}. \tag{S2}$$

In this case, the following relations can directly be verified by taking $n = 1$ and replacing U by T to compute the α coefficients:

$$\begin{aligned} \alpha^{T^\dagger}_{|0\rangle\langle 0|;|0\rangle\langle 0|} &= \alpha^{T^\dagger}_{|1\rangle\langle 1|;|1\rangle\langle 1|} = 1, \\ \alpha^{T^\dagger}_{|0\rangle\langle 1|;|0\rangle\langle 1|} &= e^{i\frac{\pi}{4}},\ \alpha^{T^\dagger}_{|1\rangle\langle 0|;|1\rangle\langle 0|} = e^{-\frac{i\pi}{4}}, \\ \alpha^{T^\dagger}_{|i\rangle\langle j|;|i'\rangle\langle j'|} &= 0 \text{ when } i \neq i' \text{ or } j \neq j'. \end{aligned} \tag{S3}$$

Plugging these values into Eq. (S1) for $n = 1$ gives

$$\begin{aligned} F_{\text{avg}}(U) = &\frac{1}{3}\mathsf{Trace}\left[|0\rangle\langle 0|\,\Lambda\circ\mathbf{T}(|0\rangle\langle 0|)\right] + \frac{1}{3}\mathsf{Trace}\left[|1\rangle\langle 1|\,\Lambda\circ\mathbf{T}(|1\rangle\langle 1|)\right] \\ &+ \frac{e^{i\frac{\pi}{4}}}{6}\left(\mathsf{Trace}\left[|0\rangle\langle 1|\,\Lambda\circ\mathbf{T}(|0\rangle\langle 1|)\right] + \mathsf{Trace}\left[|0\rangle\langle 1|\,\Lambda\circ\mathbf{T}(|1\rangle\langle 0|)\right]\right) \\ &+ \frac{e^{-i\frac{\pi}{4}}}{6}\left(\mathsf{Trace}\left[|1\rangle\langle 0|\,\Lambda\circ\mathbf{T}(|0\rangle\langle 1|)\right] + \mathsf{Trace}\left[|1\rangle\langle 0|\,\Lambda\circ\mathbf{T}(|1\rangle\langle 0|)\right]\right). \end{aligned} \tag{S4}$$

Then, we can note that

$$\begin{aligned} |0\rangle\langle 1| &= |+\rangle\langle +| + i\,|+_i\rangle\langle +_i| - \frac{1+i}{2}I \\ |1\rangle\langle 0| &= |+\rangle\langle +| - i\,|+_i\rangle\langle +_i| - \frac{1-i}{2}I, \end{aligned} \tag{S5}$$

where $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, $|+_i\rangle = (|0\rangle + i\,|1\rangle)/\sqrt{2}$, and $I = |0\rangle\langle 0| + |1\rangle\langle 1|$. In addition, since $\Lambda\circ\mathbf{T}$ is a CPTP map, then $\Lambda\circ\mathbf{T}(I) = I$, and consequently $\mathsf{Trace}\left[|1\rangle\langle 0|\,\Lambda\circ\mathbf{T}(I)\right] = \mathsf{Trace}\left[|0\rangle\langle 1|\,\Lambda\circ\mathbf{T}(I)\right] = 0$. Thus, we can write the average fidelity in terms of four combinations of state preparations $\mathcal{I} = \{|0\rangle, |1\rangle, |+\rangle, |+_i\rangle\}$ and measurements $\mathcal{M} = \{|0\rangle\langle 0|, |1\rangle\langle 1|, |+\rangle\langle +|\}$ that can be implemented on Ascella:

$$\begin{aligned} F_{\text{avg}}(U) = &\frac{1}{3}\mathsf{Trace}\left[|0\rangle\langle 0|\,\Lambda\circ\mathbf{T}(|0\rangle\langle 0|)\right] + \frac{1}{3}\mathsf{Trace}\left[|1\rangle\langle 1|\,\Lambda\circ\mathbf{T}(|1\rangle\langle 1|)\right] \\ &- \frac{2}{3\sqrt{2}}\mathsf{Trace}\left[|+\rangle\langle +|\,\Lambda\circ\mathbf{T}(|+_i\rangle\langle +_i|)\right] \\ &+ \frac{2}{3\sqrt{2}}\mathsf{Trace}\left[|+\rangle\langle +|\,\Lambda\circ\mathbf{T}(|+\rangle\langle +|)\right]. \end{aligned} \tag{S6}$$

C. Multi-qubit gates

To evaluate Eq. (S1) for any given gate unitary U, we must generalize the approach applied above. This can be done by first pre-computing $\Lambda \circ \mathbf{U}$ in terms of single-qubit Pauli operators $S_j \in \{I, X, Y, Z\}$, and then explicitly evaluating $F_{\mathrm{avg}}(U)$ to significantly reduce the required set of measurements that will ultimately be performed by Ascella. This approach provides a solution for $F_{\mathrm{avg}}(U)$, but it is not necessarily the optimal way to determine the required measurement settings.

To evaluate the process map in terms of measurements that can be performed by Ascella, we can choose to prepare each qubit j independently following $|\psi_j\rangle \in \mathcal{I}$ as for the T-gate and measure each qubit independently following $S_j \in \{I, X, Y, Z\}$. This choice of state preparations and measurements also may not be optimal to minimize the number of measurements for any given gate, but it guarantees a system of 2^{4n} linearly independent equations describing an n-qubit noisy gate:

$$\langle S\rangle_\Psi = \mathsf{Trace}\left[S\ \Lambda \circ \mathbf{U}(|\Psi\rangle\langle\Psi|)\right], \tag{S7}$$

where $S = S_1 \otimes S_2 \cdots S_n$ and $|\Psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \ldots |\psi_n\rangle$.

For $n \leq 3$, this system can be solved directly using symbolic packages or software such as Mathematica. Solving 4-qubit gates or larger would require more advanced numerical methods or analytic simplification to circumvent memory limitations. Hence, although the approach to obtain $F_{\mathrm{avg}}(U)$ is general for any gate and results in equal or fewer required measurements compared to a full process tomography, pre-computing these required measurements is still a hard problem.

Evaluating Eq. (S7) for the CNOT and Toffoli gates provides all elements of the error process Λ in terms of Pauli correlations. By evaluating $F_{\mathrm{avg}}(U)$ using the methods of [10], we obtain a set of 58 (593) correlations for the CNOT (Toffoli) gate. The set of correlations m_i and their weight w_i required for the evaluation of $F_{\mathrm{avg}}(U)$ for a CNOT gate are given in Table S3. These solutions represent a significant reduction over the number of correlations that would be necessary to perform full arbitrary process tomography, which is $2^{4n} = 256$ (4096) for the CNOT (Toffoli) gate.

Since some of the correlations are measured among fewer than n qubits, it is possible to further reduce the number of measurement settings by tracing out some qubits from the measurements obtained from higher-order correlations. For example, $\langle IX\rangle$ can be evaluated from the same data used to compute $\langle ZX\rangle$ by tracing out the first qubit. In this case, assuming all I measurements can be evaluated by recycling the Z measurement data, the number of measurements is reduced to 36 (340) for the CNOT (Toffoli), corresponding to 86% (92%) fewer measurements than for a full process tomography. Recycling the X measurement further reduces this to 34 settings for the CNOT gate. Notably, 34 measurement settings are fewer than half of the 71 settings previously used to benchmark a linear-optical CNOT gate in [14].

D. Benchmarking results

We applied the symmetry-based benchmarking method to gates implemented by Ascella and also gates implemented on other online quantum computing platforms.

1. *Ascella*

For each individual state preparation and measurement (SPAM) configuration, the transpilation process converges to a high-fidelity implementation of the desired unitary. However, one of the assumptions needed to apply symmetry-based benchmarking is that the gate unitary remains unchanged for each SPAM configuration. As a result, re-transpiling the unitary for each setting can introduce a small systematic, but random, error in the estimate of the average gate fidelity.

When benchmarking the T-gate naively, we find that the re-transpilation error can occasionally cause the measured average gate fidelity to exceed 1 by up to 0.1%, implying that the re-transpilation error is on the same order of magnitude as the T-gate error for Ascella. To remove this systematic error, we fix the voltages applied to the specific part of the chip implementing the T-gate so that transpilation process only optimizes the SPAM procedure. As a result, the measured average gate fidelity given in the main text is less than 1 to within the measurement precision, but it also no longer fully benefits from the advantage of the machine-learned transpilation process.

We also observe that implementing multiple T-gates in a row, up to 4 T-gates each implemented by a separate part of the chip, does not significantly decrease the measured average single-qubit gate fidelity. This suggests that the dominant contribution to the remaining T-gate error of 0.4% is likely SPAM error.

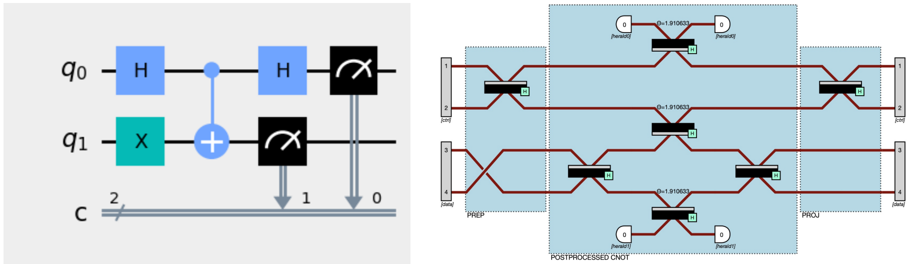

FIG. S10: One of the 36 circuits (`+1:XZ`), Qiskit quantum circuit (left) created using IBM Quantum and linear optics circuit (right) created using Perceval equivalent contributing to the 2-qubit CNOT fidelity measurement

Platform (Device)	Gate	F_{avg} (%)	Date - Benchmark Details
Quandela (Ascella)	T-gate	99.6 ± 0.1	2023/05/31 – average and standard deviation on 5× 1M-sample measurements, for 14 different gate locations on chip
	CNOT	93.8 ± 0.6	2023/03/20-2023/05/07 — average and standard deviation of 114 consecutive 100k-sample measurements over 46 days
	Toffoli	86 ± 1.2	2023/01/06 – calculated on 100000-sample tasks
IonQ (`AWS ionq.qpu`)	T-gate	99.6 ± 1	2022/12/16 – calculated on 4096-sample tasks
	CNOT	91.7 ± 1.5	2022/12/17 – calculated on 4096-sample tasks
	Toffoli	90 ± 3.1	2023/01/18 – calculated on 256-sample tasks
Rigetti (`AWS rigetti.aspen-11`)	T-gate	88.7 ± 1	2022/12/16 – calculated on 4096-sample tasks
	CNOT	71.2 ± 1.5	2022/12/17 – calculated on 4096-sample tasks
IBM (`Quito` or `Belem` depending on availability)	T-gate	96 ± 1.5	2022/12/16 – calculated on 4096-sample tasks
	CNOT	86.4 ± 1.5	2022/12/17 – calculated on 4096-sample tasks

TABLE S2: Average gate fidelity F_{avg} obtained from symmetry-based benchmarking applied to devices on other online quantum computing platforms. Margin of uncertainty is based on standard deviation assuming Poisson counting statistics unless otherwise stated. Note that this benchmarking procedure is not robust to SPAM errors, which may be a significant factor for some devices. In addition, the values returned by other online platforms may be post-processed using error mitigation techniques or subject to systematic errors analogous to the re-transpilation errors on Ascella.

Since the systematic error caused by re-transpilation is much smaller than the CNOT and the Toffoli gate errors and on the same order of magnitude as the measurement precision we apply the transpilation process to the entire chip for those cases. It is worth noting that implementing the Toffoli gate already saturates all 12 modes of the chip, meaning that it is not possible to implement SPAM without re-transpiling the part of the chip implementing the Toffoli gate.

2. *Other online platforms*

To place Ascella in context, we also apply the same symmetry-based benchmarking for gates implemented by several other online quantum computing platforms (see Table S2). We show in Figure S10 an example of a circuit for running benchmarking on these platforms compared to the equivalent circuit on Ascella.

w_i	m_i	w_i	m_i	w_i	m_i	w_i	m_i	w_i	m_i	w_i	m_i
1	00:II	1	01:II	2	0+:IX	−1	10:YY	1	11:YY	2	+0:YY
−1	00:IX	−1	01:IX	2	0+:XI	−1	10:ZI	−1	11:ZI	2	+1:XI
1	00:IZ	−1	01:IZ	2	0+:ZX	1	10:ZX	1	11:ZX	−2	+1:XX
−1	00:XI	−1	01:XI	−2	0i:XZ	1	10:ZZ	−1	11:ZZ	−2	+1:YY
1	00:XX	1	01:XX	1	10:II	1	11:II	2	1+:IX	−4	++:XI
1	00:XZ	1	01:XZ	−1	10:IX	−1	11:IX	2	1+:XI	−2	i0:XZ
−1	00:YY	1	01:YY	−1	10:IZ	1	11:IZ	−2	1+:ZX	−2	i1:XZ
1	00:ZI	1	01:ZI	−1	10:XI	−1	11:XI	−2	1i:XZ	4	ii:XZ
−1	00:ZX	−1	01:ZX	1	10:XX	1	11:XX	2	+0:XI		
1	00:ZZ	−1	01:ZZ	1	10:XZ	1	11:XZ	−2	+0:XX		

TABLE S3: List of 58 weights w_i and correlations m_i used to evaluate the average fidelity $F_{\text{avg}}(U) = (1/40)\sum_i w_i m_i$ of a CNOT gate. Correlations are labelled by ab:xy where $a, b \in \{0, 1, +, \mathrm{i}\}$ represent state preparation ($\{|0\rangle, |1\rangle, |+\rangle, |+_i\rangle\}$ respectively), and $x, y \in \{\mathrm{I,X,Y,Z}\}$ represent measurements ($\{I, X, Y, Z\}$ respectively).

S-VI. BOSON SAMPLING

Experiment	# of photons	Source	Modes	Platform	Reconfigurability	Distance	Fidelity
J. B. Spring *et. al.* [19]	3	SPDC	6	Integrated optics	No	0.094±0.014	∅
	4					0.097±0.004	∅
H. Wang *et. al.* [20]	3	QD	36	Bulk optics	No	0.125±0.001	0.984±0.001
	4					0.141±0.003	0.979±0.005
	5					0.178±0.005	0.973±0.009
Y. He *et. al.* [21]	3	QD	8	Fiber loop	Yes	∅	0.993±0.002
	4					∅	0.973±0.006
H.-S. Zhong *et. al.* [22]	3	SPDC	12	Integrated optics	No	0.122±0.018	0.982±0.004
	4					∅	∅
	5					∅	∅
H. Wang *et. al.* [23]	$4^{\dagger}$	QD	16	Bulk optics	No	0.085±0.001	0.994±0.001
	$5^{\dagger}$					0.106±0.002	0.989±0.002
	$6^{\dagger}$					0.201±0.003	0.960±0.004
	7					∅	∅
S. Paesani *et. al.* [24]	3	SPDC	12	Integrated optics	No	∅	0.92
	4					∅	0.87
H. Wang *et. al.* [25]	2	QD	60	Bulk optics	No	0.043±0.005	0.995±0.003
	3					0.095±0.001	0.988±0.001
	4					0.143±0.001	0.984±0.001
	...					∅	∅
	$20^{\dagger\dagger}$					∅	∅
J. Gao *et. al.* [26]	3	SPDC	35	Integrated optics	No	0.35	0.89
F. Hoch *et. al.* [27]	3	SPDC	32	Integrated optics	Yes	∅	∅
This work	6	QD	12	Integrated optics	Yes	0.16±0.02	0.97±0.03

TABLE S4: Comparison with previous demonstrations of Boson Sampling using single photons. We tabulate here only the works that report a metric with measured fidelity or total variation distances. †Measuring output states with 1 photon loss. ††Measuring output states with 6 photon loss.

In this section we study boson-sampling with photon loss. With a 6-photon input state, we acquire respectively $295 \cdot 10^9$, $41.5 \cdot 10^9$, $3.07 \cdot 10^9$, $110 \cdot 10^6$, and $1.78 \cdot 10^6$ samples for 1-, 2-, 3-, 4-, and 5-photon coincidences. We compute the total variation distance $D = \frac{1}{2}\sum_i |p_i - q_i|$ where $\{p_i\}$ and $\{q_i\}$ are the ideal and experimental output probability distributions respectively for output states with 2, 3, 4 and 5 photons. The results are reported in Tab. S5. For > 2 photon loss the experimental output probability is dominated by loss, and no longer describe the ideal distribution.

To validate that our boson-sampler device is functioning correctly, we use a phenomenological approach that models all sources of noise in the experimental apparatus and demonstrate that we reach a remarkable overlap between all threshold statistics and our simulations. Using *Perceval*, we develop a realistic simulator for our boson-sampler. We use the phenomenological model first introduced in [28] of our single-photon source to account for the partial distinguishability of the multiphoton state, the imperfect single-photon purity and the optical losses. Then, we account for the error related to the transpilation of the unitary matrix and the imperfect implementation of the physical phases

N-photon loss	Fidelity	Distance
0-	0.97±0.03	0.16±0.02
1-	0.989±0.002	0.118
2-	0.9950±0.0002	0.143
3-	0.99625±3e-05	0.225
4-	0.997333±6e-06	0.40

TABLE S5: Fidelity and total variation distance between the experimental data and the ideal output probability for all N-photon outcomes.

with the thermo-optic phase shifters. The simulations gives a total variation distance of $D_{\rm simu} = 0.132$ and a fidelity of $F_{\rm simu} = 0.978$. The simulations are compatible with the experimental data, which shows that our realistic simulator truthfully describes the boson-sampling device.

S-VII. CLASSIFICATION

Quantum algorithms for classification on near-term devices have been explored through a variety of approaches [29], although most results are supported by numerical simulations and not implemented in the lab. This is the case for instance of Ref. [30], where the authors present an ansatz for a variational quantum algorithm that is native to photonics. They study the expressivity of the resulting model through theory and numeric simulations. The circuit is made of two trainable blocks with a data encoding block in the middle. This data encoding block consists of phase shifters. For a k-dimensional data point $\vec{x} = (x_i, \ldots, x_k)$, each feature x_i is encoded into the phase of a phase shifter. The two trainable blocks are beamsplitter meshes that implement unitary operations, for instance through the encoding of [31] or [32]. The model is studied within the Fock space, considering input states $|n_1^{in}, \ldots, n_m^{in}\rangle$ and output states $|n_1^{out}, \ldots, n_m^{out}\rangle$. The total number of photons is denoted n and the number of modes is denoted m.

The authors of Ref. [30] show that the output of the circuit, i.e. the model, can be expressed as $f_{\boldsymbol{\theta}}(x) = \sum_\omega c_\omega(\boldsymbol{\theta}, \boldsymbol{\lambda}) e^{ix\omega}$, where the frequencies ω depend on the number of photons n input in the circuit, and the Fourier coefficients $c_\omega(\boldsymbol{\theta}, \boldsymbol{\lambda})$ depend on the measurement parameters $\boldsymbol{\lambda}$ and the chip parameters $\boldsymbol{\theta}$.

We can see how interesting it is to use a photonic encoding and in particular to exploit the Fock space from the perspective of expressivity: by adding more photons, more terms will be added to the Fourier series and the resulting model will be more expressive, without increasing the complexity of the circuit or the number of modes. However, it is important to note that this is only possible if PNR detectors are available, so that the output state $|n_1^{out}, \ldots, n_m^{out}\rangle$ can be resolved beyond $n_i^{out} = 1$.

Taking inspiration from the results of [30], we design an ansatz for the variational quantum classifier on Ascella containing two parameterizable blocks with a data-encoding block in between. We select modes 3 to 7 on Ascella and construct the first parameterized block using 16 of the reconfigurable thermo-optic phase shifters. The data-encoding block follows, where 4 phase shifters acting on modes 4 to 7 encode the 4 features of each IRIS data point. We then implement another parameterized block with 16 phase shifters. The remaining phases on the chip are either set to 0, when we do not wish to add any extra trainable or encoding phases, or to π, when an interferometer needs to be fully reflective so that no photon escapes to the other modes of the chip. Our ansatz is shown in Fig. S11.

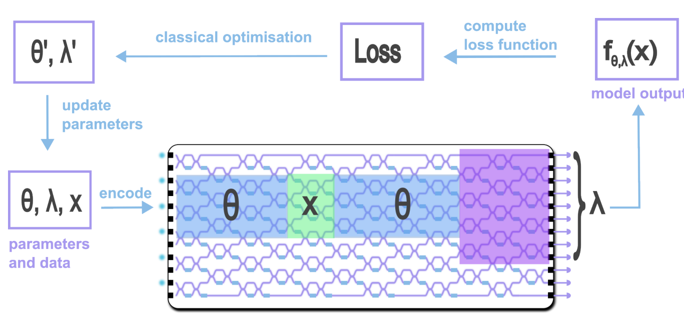

FIG. S11: Representation of our variational quantum classifier on Ascella for the classification algorithm. The trainable blocks parameterized by $\boldsymbol{\theta}$ are depicted in blue. The data encoding block is in green. The purple block represents the pseudo-PNR layer, which incorporates four extra modes into the circuit. The model is computed by assigning $\boldsymbol{\lambda}$ parameters to the outputs at the detectors. All parameters are updated via classical optimization.

We input into the circuit the 3-photon Fock state $|\psi_{in}\rangle = |0010101000000\rangle$ defined over the 12 modes of Ascella: one photon enters modes 3, 5 and 7 of the chip. We observe different output states $|\psi_{out}\rangle$ depending on the photon counts observed in the detectors. The resulting model takes the form $f_{\boldsymbol{\theta},\boldsymbol{\lambda}}(\boldsymbol{x}) = \langle\psi_{in}|\mathcal{U}^\dagger(\boldsymbol{x},\boldsymbol{\theta})\mathcal{M}(\boldsymbol{\lambda})\mathcal{U}(\boldsymbol{x},\boldsymbol{\theta})\,|\psi_{in}\rangle$. The weights $\boldsymbol{\lambda}$ are assigned to each possible output state observed at the detection step and thus define the observable $\mathcal{M}$ that we are effectively measuring. The operator $\mathcal{U}(\boldsymbol{x},\boldsymbol{\theta})$ describes the parameterized and data-encoding blocks. Following the variational approach, we train the model by optimizing classically over the phases $\boldsymbol{\theta}$ from the parameterized blocks, as well as over the weights $\boldsymbol{\lambda}$.

We demonstrate pseudo photon-number resolution (PNR) partially, on the first and on the last detectors. To this end, we set four phase shifters to $\pi/2$ in the final layer of the chip in order to redirect a portion of the photons from modes 3 and 7 into modes 1, 2 and 8, 9 respectively. We can then reinterpret the detection counts: for instance, observing photons in modes 1, 2 and 3 in our scheme corresponds to observing a three-photon count in mode 3 if we had PNR detectors. Note that implementing this partial pseudo-PNR adds expressivity to our model, as it increases the space of possible outcome states and thus the amount of $\boldsymbol{\lambda}$ parameters over which we optimize.

For the optimization, we found that performing a see-saw was the most efficient option. This means separating the chip parameters $\boldsymbol{\theta}$ from the observable parameters $\boldsymbol{\lambda}$ into two loops, and finding the best $\boldsymbol{\lambda}$ for each set of values of $\boldsymbol{\theta}$. The efficiency comes from the fact that tuning one or other set of parameters is not equally costly: changing the $\boldsymbol{\theta}$ requires re-configuring the chip, while modifying the $\boldsymbol{\lambda}$ only involves classical post-processing. We used Gaussian processes to optimize the $\boldsymbol{\theta}$ and for each iteration we optimized over the $\boldsymbol{\lambda}$ using the Nelder-Mead method.

A note about optimizing over the $\boldsymbol{\lambda}$ parameters: if we were to choose a fixed observable for the variational algorithm and dismiss the $\boldsymbol{\lambda}$ parameters completely, we may not obtain the same performance for the classifier. Indeed, we would not only have fewer parameters for the optimization but the remaining parameters would be the ones most sensitive to the noise of the experiment. Nevertheless, it is reasonable to grant degrees of freedom to the choice of the observable, so we choose to optimize over the $\boldsymbol{\lambda}$ as in [30].

In the main text, we summarized the performance of the model using confusion matrices. Fig. S14 shows an alternative display of the results, where the classification estimator is included for each prediction in the dataset along with an error. We evaluated this error knowing that we used 50000 samples for each run of the experiment. Note also that we adapted the definition of the classification estimator of [30] to our case of multi-class classification.

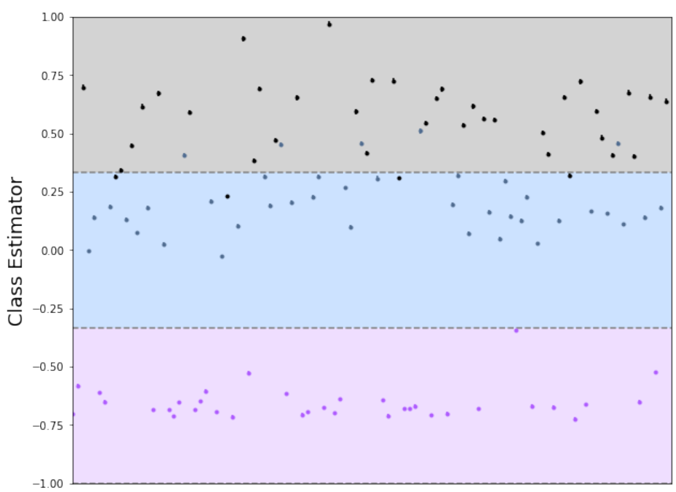

FIG. S12: Training dataset

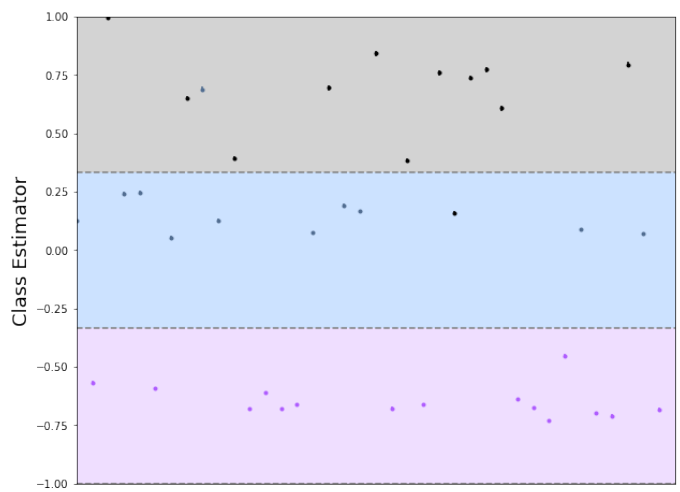

FIG. S13: Test dataset

FIG. S14: Value of the classification estimator, i.e. the model output, displayed for each data point in the train and the test set. The color of the data point indicates its true label while the background color indicates the predicted label. The error bars are obtained via Poissonian statistics from the number of samples used in each run to compute the value of the estimator. The points are simply ordered in the same way as in the dataset.

S-VIII. VQE

The Variational Quantum Eigensolver (VQE) [34, 35] for finding ground state energies for a target Hamiltonian can be broken into several steps as follows:

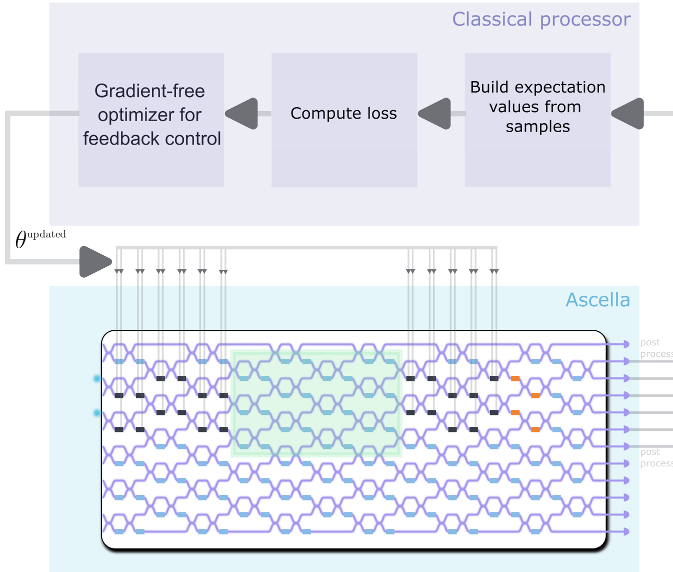

FIG. S15: Procedure for the hybrid photonic VQE algorithm we implement on Ascella. It comprises a QPU block on the bottom that enables the energy of the chosen system to be estimated. The ansatz is parameterized via thermo-electric phase shifters highlighted in black. The green block is a postselected Ralph CNOT [33]. The phase shifter in orange allows selection of the measurement bases required to reconstruct the correct Pauli terms from the qubit Hamiltonian (S8). The outputs from Ascella are fed into a classical block (at the top) which reconstructs the energy. Then it implements a feedback loop back into Ascella via a gradient-free optimizer in order to optimize the angles of the phase shifters to obtain an ansatz closer to the ground state.

1. Find a Hamiltonian formulation suitable for the problem at hand. To do so, we use the symmetry-conserving Bravyi-Kitaev transformation [36] (available through the *OpenFermion* [37] python package). This provides two very useful pieces of information:
   – The number of qubits necessary to run the VQE.
   – A description of the Hamiltonian in terms of Pauli words which then define coefficients to associate to each measurement basis. We can then construct the expectation value based on this description.
2. Prepare the ansatz. We need a parametrizable circuit expressive enough to be able to reach (a good approximation of) the ground state.
3. Rotate the ansatz into the basis for each Pauli word operator present in the molecular Hamiltonian, then take a number of samples. Restore the ideal sample counts through error mitigation.
4. Optimize the parameters via a classical procedure (COBYLA algorithm) based on the expectation value computed from the error-mitigated output of the QPU. A classical feedforward loop updates the parameters of the ansatz to reach a lower energy. The evolution of the ground state energy of H_2 with respect to the iterations is represented in Fig. S16.

A. Hamiltonian description

In order to calculate expectation values of our H_2 Hamiltonian we need to transform the second-quantized version of the Hamiltonian into a qubit basis. This can be done with a number of different transformations such as Jordan-Wigner, but here we use the Symmetry-Conserving Bravyi-Kitaev transform as described in [36, 38]. This involves first reordering the electronic orbitals and then using the Bravyi-Kitaev binary tree mapping. This gives a Hamiltonian which consists of four states that can be assigned in the following way: qubit 1 corresponds to spin-up on the first site, qubit 2 to spin-up on the second site, qubit 3 to spin-down on the first site, and qubit 4 to spin-down on the second site. The Hamiltonian can be further reduced with the symmetries first derived in [39]. There it was noted that the Hamiltonian acts off-diagonally on only two qubits, those indexed 0 and 2. The simulation is begun in the

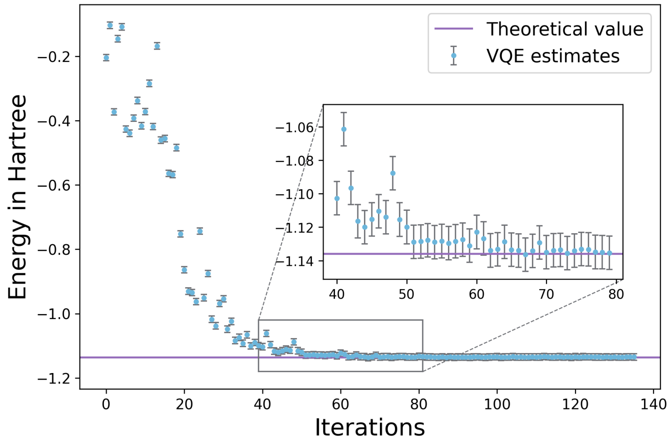

FIG. S16: Evolution of the ground state energy of H_2 obtained by the VQE procedure on Ascella with the number of iterations. Each data point corresponds to 10000 post-processed 2-photon coincidence samples. Error bars correspond to ±1 SD of the photon counting statistics.

Hartree-Fock state which stabilizes qubits 1 and 3 so that they are never flipped throughout the simulation. This symmetry can be used to reduce the Hamiltonian of interest to the following effective Hamiltonian which acts only on two qubits:

$$\hat{\mathcal{H}}_{\text{qubit}}(r) = \alpha\mathbb{I}\mathbb{I} + \beta Z\mathbb{I} + \gamma\mathbb{I}Z + \delta ZZ + \mu XX \tag{S8}$$

The constants vary with the choice of bond length and can be found in Table S6. The procedure described above is implemented in *OpenFermion* [37]. Thus finding the ground state energy for H_2 for varying bound length can be performed

- with 2 qubits (providing we have an expressive enough ansatz – this is what we tackle in the next subsection),
- by building the expectation values for the measurements ZZ and XX. This is done in a single job on the QPU by tuning the phase shifter in orange in Fig. S15 with micro-increments which allows only a few phases on the Ascella chip to be changed and avoids losing time with full reconfiguration of the chip.

B. Ansatz preparation

In the present context, the idea behind using the VQE algorithm is to produce a parameterizable ansatz expressive enough to get very close to a 2-qubit ground state of the desired Hamiltonian in order to obtain the ground state energy. The gate-based circuit below (see Fig. S17) can generate any 2-qubit state by Schmidt decomposition (one R_X rotation at the end can be removed in principle since it amounts to removing the global phase).

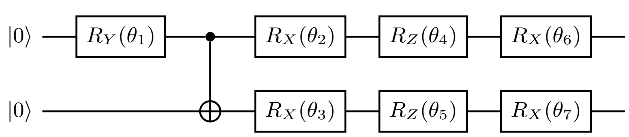

FIG. S17: Gate-based version of an ansatz circuit with 7 parameters controlling 7 parameterized rotations capable of generating any 2-qubit state.

This wave ansatz can be implemented photonically. To deal with noise and because using more parameters can be helpful for converging faster to the ground state energy, we use the ansatz represented in Fig. S15 where we path encode the 2 qubits with one photon per pair of modes $2-3$ and $4-5$ (the first mode being the 0^{th} one). This comprises 20 tunable phase shifters. Modes 1 and 6 are used as ancillary mode for the postselected Ralph CNOT. The CNOT is constructed from our transpilation algorithm. We use parametrizable thermo-electric phase shifter (represented in black on Fig. S15) to control the optical index in the waveguide and thus tune the phase of the photons so that we can achieve any 2-qubit state.

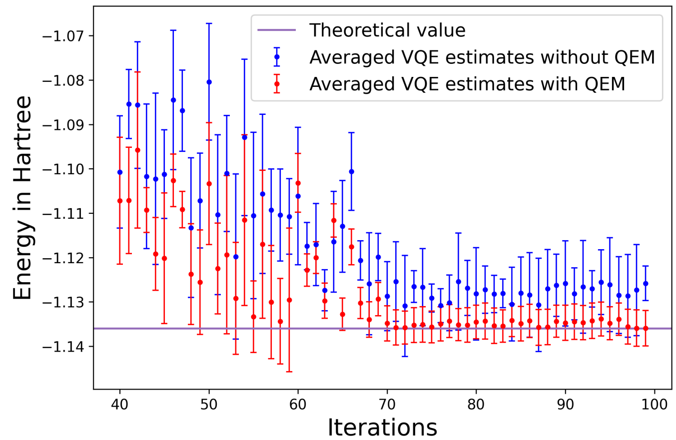

FIG. S18: Comparison with and without SPAM error mitigation by averaging loss values over 45 runs. Each data point corresponds to 10000 post-processed 2-photon coincidence samples. Error bars correspond to ±1 SD of the photon counting statistics.

C. Error Mitigation

We use the quantum error mitigation (QEM) technique proposed in [40] and first experimentally demonstrated in [41] to more consistently converge to the correct ground state energy by mitigating state preparation and measurement (SPAM) errors. The idea is to mitigate the errors arising from noisy evolution due to thermal noise from the heated phase shifters as well as from the error caused by changing the measurement basis. As detailed in the Methods section, we want to compute left-stochastic matrices Γ_b (for each measurement basis b) such that

$$q = \Gamma_b p, \tag{S9}$$

with q the noisy output probability and p the ideal noiseless output probability. In our case, we have two such Γ_b, which corresponds to a measurement basis in Eq. (S8): XX and ZZ (as $\mathbb{II}$, $\mathbb{I}Z$ and $Z\mathbb{I}$ can be obtained from ZZ by classical post-processing).

We construct Γ_b experimentally as follows: $(\Gamma_b)_{ij} = |\langle\psi|_i^b\, b\,|\psi\rangle_j^b|^2$ is the probability of obtaining the i^{th} eigenvector of $b\,|\psi\rangle_i^b$ when the j^{th} eigenvector $|\psi\rangle_j^b$ is prepared and measurement observable b is performed. For low enough SPAM errors, each Γ_b is a diagonally dominant matrix and we can retrieve the idealized probability distribution by inverting Γ_b in Eq. (S9). From this procedure we get the two desired matrices as:

$$\begin{aligned}
\Gamma_{ZZ} &= \begin{bmatrix} 9.99999952e-01 & 3.09568451e-02 & 3.09568451e-02 & 1.54929555e-09 \\ 2.34741773e-08 & 9.38086308e-01 & 1.45337301e-09 & 2.34741773e-08 \\ 2.34741773e-08 & 1.45337301e-09 & 9.38086308e-01 & 2.34741773e-08 \\ 1.54929555e-09 & 3.09568451e-02 & 3.09568451e-02 & 9.99999952e-01 \end{bmatrix} \\
\Gamma_{XX} &= \begin{bmatrix} 9.99999951e-01 & 2.47148265e-02 & 2.47148265e-02 & 1.24580719e-09 \\ 2.39578331e-08 & 9.50570344e-01 & 1.18422748e-09 & 2.39578331e-08 \\ 2.39578331e-08 & 1.18422748e-09 & 9.50570344e-01 & 2.39578331e-08 \\ 1.24580731e-09 & 2.47148287e-02 & 2.47148287e-02 & 9.99999951e-01 \end{bmatrix}
\end{aligned} \tag{S10}$$

Note that for the best performance, the Γ_b matrices should be experimentally evaluated immediately prior to the VQE experiment.

When we obtain our error-mitigated measurement probabilities, we can construct an eigenvalue estimate which is sent to the classical optimizer. We can note the difference between the expectation values compared to simulated values with and without error mitigation in Fig. S18 and see a noticeable improvement, particularly as the energy nears the ground state. We note that the simulated and error-mitigated values have a close to perfect agreement. This mainly comes from the correction of the basis rotation gates.

D. Classical optimization

The classical part of the VQE algorithm is performed on the dark blue classical processing unit (CPU) box in Fig. S15. The expectation values for the terms of $\hat{\mathcal{H}}_{\text{qubit}}$ are constructed from error-mitigated samples from the QPU.

Hamiltonian Coefficients					
Radius (Å)	α ($\mathbb{II}$)	β ($Z\mathbb{I}$)	γ ($\mathbb{I}Z$)	δ (ZZ)	μ (XX)
0.2	2.0115282039582	0.9304885285175	0.9304885285175	0.013623865138623	0.157972708628
0.25	1.4228278945358	0.8706459577114	0.8706459577114	0.013463487127669	0.15927658478468
0.3	1.0101820841922	0.8086489099089	0.8086489099089	0.013287977089941	0.16081851920392
0.35	1.0101820841922	0.8086489099089	0.8086489099089	0.013287977089941	0.16081851920392
0.4	0.4603634956295	0.688819429564	0.688819429564	0.01291396933589	0.1645154240225
0.45	0.2675472248053	0.6338897827590	0.6338897827590	0.012719203005418	0.16662140112466
0.50	0.11064654485357	0.5830796254889	0.5830796254889	0.01251643158428	0.16887022768973
0.55	−0.0183735206558	0.5364887845888	0.5364887845888	0.012300353656101	0.17124451736495
0.65	−0.2139316272136	0.45543342027862	0.4554334202786	0.011801922101754	0.17631845161020
0.75	−0.3498334175179	0.38874758809160	0.38874758809160	0.011177144762525	0.18177153657730
0.85	−0.4454236322275	0.3337464949796	0.33374649497965	0.01040606826223	0.18756184791877
0.95	−0.5135484185550	0.2877959899385	0.2877959899385	0.009503470221825	0.19365031698524
1.05	−0.5626001130028	0.24878328975518	0.24878328975518	0.008509936866414	0.19998426653596
1.15	−0.5979734705198	0.2152339371429	0.2152339371429	0.007477201225847	0.2064946748241
1.25	−0.6232232011799	0.1861731031999	0.18617310319995	0.006455593489009	0.2131024013141
1.35	−0.6408366121165	0.1609263900897	0.16092639008976	0.005486217221390	0.21972703573593
1.45	−0.6526612024877	0.13897677941251	0.13897677941251	0.004597585574594	0.22629425934361
1.55	−0.6601174612872	0.11989353736336	0.11989353736336	0.0038055776890	0.23274029161766
1.65	−0.6643091838424	0.10330532972950	0.10330532972950	0.003115459300252	0.23901364608341
1.75	−0.6660923101667	0.08889055166239	0.08889055166239	0.00252481503720	0.24507502046287
1.85	−0.6661263822710	0.07637119958665	0.07637119958665	0.002026489193459	0.2508961512677
1.95	−0.6649159578993	0.06550649596835	0.06550649596835	0.001610984321150	0.2564582470193
2.05	−0.6628441004621	0.05608661275595	0.05608661275595	0.001268117568204	0.26175037476834

TABLE S6: List of Hamiltonian coefficients from Eq. (1) for varying bond length.

Then the loss function $\langle \hat{\mathcal{H}}_{\text{qubit}} \rangle$ is constructed by summing the 5 terms comprising Eq. (1) (from main letter). The CPU calls the COBYLA optimizer to find the new set of angles to feed the QPU. Convergence of the procedure is shown in Fig. S16.

At each step, for each bond length, we compute the energy from 10000 processed samples. This amounts to a 0.013 probability that our sampling error is greater than 0.01 Hartree. From this we get the value of r for which the energy is minimum which we call r^{opt}. We make an additional run at r^{opt} with 400000 processed samples to reduce the sampling error to 0.00158 Hartree, below the error for chemical accuracy. It is worth noting here that the optimizer can occasionally converge to a local minimum of the objective function, due to barren plateaus problem which is avoided in most experiments.

S-IX. GHZ-FACTORY

The layout of the GHZ-factory photonic circuit adapted from Ref. [42, 43] is presented in Fig. S19.

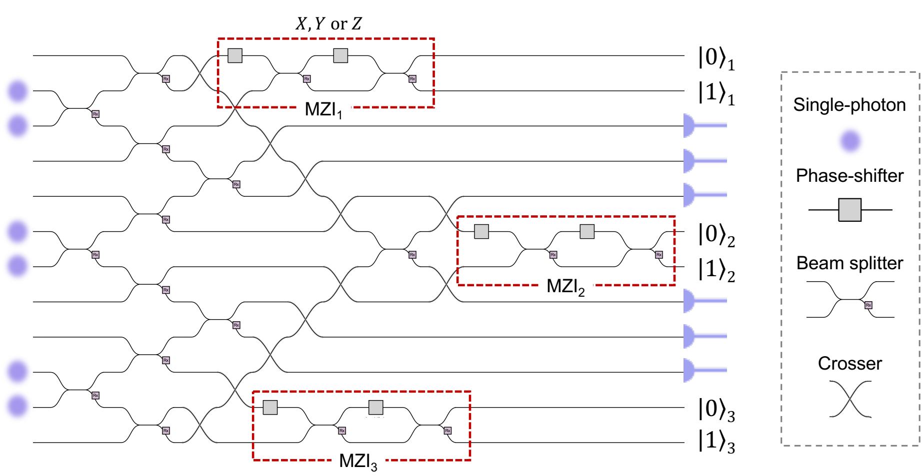

FIG. S19: GHZ-factory photonic circuit from *Perceval*. Six single photons (purple dots) are sent to the input optical modes $(2, 3, 6, 7, 10, 11)$. Three qubits are path encoded in the output optical modes $(1, 2)$, $(6, 7)$ and $(11, 12)$. To account for the losses in the optical system we postselect on the detection of one and only one photon per qubit. We use reconfigurable Mach-Zehnder interferometers to project the generated state in the Pauli matrix basis $(\mathbb{I}, X, Y, Z)$.

[1] Niccolo Somaschi, Valerian Giesz, Lorenzo De Santis, JC Loredo, Marcelo P Almeida, Gaston Hornecker, S Luca Portalupi, Thomas Grange, Carlos Anton, Justin Demory, et al. Near-optimal single-photon sources in the solid state. Nature Photonics, 10(5):340–345, 2016.

[2] SE Thomas, M Billard, N Coste, SC Wein, H Ollivier, O Krebs, L Tazaïrt, A Harouri, A Lemaitre, I Sagnes, et al. Bright polarized single-photon source based on a linear dipole. Physical review letters, 126(23):233601, 2021.

[3] H Ollivier, SE Thomas, SC Wein, I Maillette de Buy Wenniger, N Coste, JC Loredo, N Somaschi, A Harouri, A Lemaitre, I Sagnes, et al. Hong-ou-mandel interference with imperfect single photon sources. Physical Review Letters, 126(6):063602, 2021.

[4] O Gazzano, S Michaelis de Vasconcellos, C Arnold, A Nowak, E Galopin, I Sagnes, L Lanco, A Lemaître, and P Senellart. Bright solid-state sources of indistinguishable single photons. Nature communications, 4(1):1425, 2013.

[5] H Snijders, JA Frey, J Norman, VP Post, AC Gossard, JE Bowers, MP Van Exter, W Löffler, and D Bouwmeester. Fiber-coupled cavity-qed source of identical single photons. Physical Review Applied, 9(3):031002, 2018.

[6] Giuseppe Della Valle, Roberto Osellame, and Paolo Laporta. Micromachining of photonic devices by femtosecond laser pulses. Journal of Optics A: Pure and Applied Optics, 11(1):013001, 2008.

[7] C Antón, Juan Carlos Loredo, Guillaume Coppola, Helene Ollivier, Niko Viggianiello, Abdelmounaim Harouri, Niccolo Somaschi, Andrea Crespi, Isabelle Sagnes, Aristide Lemaitre, et al. Interfacing scalable photonic platforms: solid-state based multi-photon interference in a reconfigurable glass chip. Optica, 6(12):1471–1477, 2019.

[8] Daniel J Brod, Ernesto F Galvão, Andrea Crespi, Roberto Osellame, Nicolò Spagnolo, and Fabio Sciarrino. Photonic implementation of boson sampling: a review. Advanced Photonics, 1(3):034001–034001, 2019.

[9] Mathias Pont, Riccardo Albiero, Sarah E Thomas, Nicolò Spagnolo, Francesco Ceccarelli, Giacomo Corrielli, Alexandre Brieussel, Niccolo Somaschi, Hêlio Huet, Abdelmounaim Harouri, et al. Quantifying n-photon indistinguishability with a cyclic integrated interferometer. Physical Review X, 12(3):031033, 2022.

[10] Stephen Wein Rawad Mezher, Grégoire De Gliniasty. Direct computing of the average gate fidelity of any gate. in prep, 2024.

[11] Michael A Nielsen and Isaac Chuang. Quantum computation and quantum information, 2002.

[12] Joseph M Renes, Robin Blume-Kohout, Andrew J Scott, and Carlton M Caves. Symmetric informationally complete quantum measurements. Journal of Mathematical Physics, 45(6):2171–2180, 2004.

[13] Aram W Harrow. The church of the symmetric subspace. arXiv preprint arXiv:1308.6595, 2013.

[14] Jeremy L O'Brien, Geoff J Pryde, Alexei Gilchrist, Daniel FV James, Nathan K Langford, Timothy C Ralph, and Andrew G White. Quantum process tomography of a controlled-not gate. Physical review letters, 93(8):080502, 2004.

[15] Michael A Nielsen. A simple formula for the average gate fidelity of a quantum dynamical operation. Physics Letters A, 303(4):249–252, 2002.

[16] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. General teleportation channel, singlet fraction, and quasidistillation. Physical Review A, 60(3):1888, 1999.

[17] Steven T Flammia and Yi-Kai Liu. Direct fidelity estimation from few pauli measurements. Physical review letters, 106 (23):230501, 2011.

[18] Sergey Bravyi and Alexei Kitaev. Universal quantum computation with ideal clifford gates and noisy ancillas. Physical Review A, 71(2):022316, 2005.

[19] Justin B Spring, Benjamin J Metcalf, Peter C Humphreys, W Steven Kolthammer, Xian-Min Jin, Marco Barbieri, Animesh Datta, Nicholas Thomas-Peter, Nathan K Langford, Dmytro Kundys, et al. Boson sampling on a photonic chip. Science, 339(6121):798–801, 2013.

[20] Hui Wang, Yu He, Yu-Huai Li, Zu-En Su, Bo Li, He-Liang Huang, Xing Ding, Ming-Cheng Chen, Chang Liu, Jian Qin, et al. High-efficiency multiphoton boson sampling. Nature Photonics, 11(6):361–365, 2017.

[21] Yu He, X Ding, Z-E Su, H-L Huang, J Qin, C Wang, S Unsleber, C Chen, H Wang, Y-M He, et al. Time-bin-encoded boson sampling with a single-photon device. Physical review letters, 118(19):190501, 2017.

[22] Han-Sen Zhong, Yuan Li, Wei Li, Li-Chao Peng, Zu-En Su, Yi Hu, Yu-Ming He, Xing Ding, Weijun Zhang, Hao Li, et al. 12-photon entanglement and scalable scattershot boson sampling with optimal entangled-photon pairs from parametric down-conversion. Physical review letters, 121(25):250505, 2018.

[23] Hui Wang, Wei Li, Xiao Jiang, Y-M He, Y-H Li, Xing Ding, M-C Chen, Jian Qin, C-Z Peng, Christian Schneider, et al. Toward scalable boson sampling with photon loss. Physical review letters, 120(23):230502, 2018.

[24] Stefano Paesani, Yunhong Ding, Raffaele Santagati, Levon Chakhmakhchyan, Caterina Vigliar, Karsten Rottwitt, Leif K Oxenløwe, Jianwei Wang, Mark G Thompson, and Anthony Laing. Generation and sampling of quantum states of light in a silicon chip. Nature Physics, 15(9):925–929, 2019.

[25] Hui Wang, Jian Qin, Xing Ding, Ming-Cheng Chen, Si Chen, Xiang You, Yu-Ming He, Xiao Jiang, L You, Z Wang, et al. Boson sampling with 20 input photons and a 60-mode interferometer in a 1 0 14-dimensional hilbert space. Physical review letters, 123(25):250503, 2019.

[26] Jun Gao, Zhi-Qiang Jiao, Ruo-Jing Ren, Xiao-Wei Wang, Xiao-Yun Xu, Wen-Hao Zhou, Lu-Feng Qiao, and Xian-Min Jin. Experimental collision-free dominant boson sampling. arXiv preprint arXiv:1910.11320, 2019.

[27] Francesco Hoch, Simone Piacentini, Taira Giordani, Zhen-Nan Tian, Mariagrazia Iuliano, Chiara Esposito, Anita Camillini, Gonzalo Carvacho, Francesco Ceccarelli, Nicolò Spagnolo, et al. Reconfigurable continuously-coupled 3d photonic circuit for boson sampling experiments. npj Quantum Information, 8(1):55, 2022.

[28] Mathias Pont, Giacomo Corrielli, Andreas Fyrillas, Iris Agresti, Gonzalo Carvacho, Nicolas Maring, Pierre-Emmanuel Emeriau, Francesco Ceccarelli, Ricardo Albiero, Paulo HD Ferreira, et al. High-fidelity generation of four-photon ghz states on-chip. arXiv preprint arXiv:2211.15626, 2022.
[29] Weikang Li and Dong-Ling Deng. Recent advances for quantum classifiers. Science China Physics, Mechanics & Astronomy, 65(2), dec 2021. doi:10.1007/s11433-021-1793-6. URL https://doi.org/10.1007%2Fs11433-021-1793-6.
[30] Beng Yee Gan, Daniel Leykam, and Dimitris G. Angelakis. Fock state-enhanced expressivity of quantum machine learning models. EPJ Quantum Technology, 9(1), 2022. doi:10.1140/epjqt/s40507-022-00135-0.
[31] Michael Reck, Anton Zeilinger, Herbert J. Bernstein, and Philip Bertani. Experimental realization of any discrete unitary operator. Phys. Rev. Lett., 73:58–61, Jul 1994. doi:10.1103/PhysRevLett.73.58. URL https://link.aps.org/doi/10.1103/PhysRevLett.73.58.
[32] William R. Clements, Peter C. Humphreys, Benjamin J. Metcalf, W. Steven Kolthammer, and Ian A. Walmsley. Optimal design for universal multiport interferometers. Optica, 3(12):1460–1465, Dec 2016. doi:10.1364/OPTICA.3.001460. URL https://opg.optica.org/optica/abstract.cfm?URI=optica-3-12-1460.
[33] Timothy C Ralph, Nathan K Langford, TB Bell, and AG White. Linear optical controlled-not gate in the coincidence basis. Physical Review A, 65(6):062324, 2002.
[34] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J Love, Alán Aspuru-Guzik, and Jeremy L O'brien. A variational eigenvalue solver on a photonic quantum processor. Nature communications, 5(1): 4213, 2014.
[35] Jules Tilly, Hongxiang Chen, Shuxiang Cao, Dario Picozzi, Kanav Setia, Ying Li, Edward Grant, Leonard Wossnig, Ivan Rungger, George H Booth, et al. The variational quantum eigensolver: a review of methods and best practices. Physics Reports, 986:1–128, 2022.
[36] Sergey Bravyi, Jay M Gambetta, Antonio Mezzacapo, and Kristan Temme. Tapering off qubits to simulate fermionic hamiltonians. arXiv preprint arXiv:1701.08213, 2017.
[37] Jarrod R McClean, Nicholas C Rubin, Kevin J Sung, Ian D Kivlichan, Xavier Bonet-Monroig, Yudong Cao, Chengyu Dai, E Schuyler Fried, Craig Gidney, Brendan Gimby, et al. Openfermion: the electronic structure package for quantum computers. Quantum Science and Technology, 5(3):034014, 2020.
[38] Jacob T Seeley, Martin J Richard, and Peter J Love. The bravyi-kitaev transformation for quantum computation of electronic structure. The Journal of chemical physics, 137(22):224109, 2012.
[39] Peter JJ O'Malley, Ryan Babbush, Ian D Kivlichan, Jonathan Romero, Jarrod R McClean, Rami Barends, Julian Kelly, Pedram Roushan, Andrew Tranter, Nan Ding, et al. Scalable quantum simulation of molecular energies. Physical Review X, 6(3):031007, 2016.
[40] Junaid ur ur Rehman and Hyundong Shin. Entanglement-free parameter estimation of generalized pauli channels. Quantum, 5:490, 2021.
[41] Donghwa Lee, Jinil Lee, Seongjin Hong, Hyang-Tag Lim, Young-Wook Cho, Sang-Wook Han, Hyundong Shin, Junaid ur Rehman, and Yong-Su Kim. Error-mitigated photonic variational quantum eigensolver using a single-photon ququart. Optica, 9(1):88–95, 2022.
[42] Ying Li, Peter C Humphreys, Gabriel J Mendoza, and Simon C Benjamin. Resource costs for fault-tolerant linear optical quantum computing. Physical Review X, 5(4):041007, 2015.
[43] Clément Gouriou. Design and fabrication of an integrated photonic circuit for producing a maximally-entangled three-photon state. Politecnico di Milano, 2019. URL http://hdl.handle.net/10589/151271.