Boson Sampling on a Photonic Chip

Justin B. Spring[1], Benjamin J. Metcalf[1], Peter C. Humphreys[1], W. Steven Kolthammer[1], Xian-Min Jin[1,2], Marco Barbieri[1], Animesh Datta[1], Nicholas Thomas-Peter[1], Nathan K. Langford[1,3], Dmytro Kundys[4], James C. Gates[4], Brian J. Smith[1], Peter G.R. Smith[4], and Ian A. Walmsley[1*]

[1]Clarendon Laboratory, Department of Physics, University of Oxford, OX1 3PU, UK,
[2]Department of Physics, Shanghai Jiao Tong University, Shanghai 200240, PR China
[3] Department of Physics, Royal Holloway, University of London, TW20 0EX, UK
[4]Optoelectronics Research Centre, University of Southampton, Southampton, SO17 1BJ, UK

*To whom correspondence should be addressed; E-mail: i.walmsley1@physics.ox.ac.uk

While universal quantum computers ideally solve problems such as factoring integers exponentially more efficiently than classical machines, the formidable challenges in building such devices motivate the demonstration of simpler, problem-specific algorithms that still promise a quantum speedup. We construct a quantum boson sampling machine (QBSM) to sample the output distribution resulting from the nonclassical interference of photons in an integrated photonic circuit, a problem thought to be exponentially hard to solve classically. Unlike universal quantum computation, boson sampling merely requires indistinguishable photons, linear state evolution, and detectors. We benchmark our QBSM with three and four photons and analyze sources of sampling inaccuracy. Scaling up to larger devices could offer the first definitive quantum-enhanced computation.

Universal quantum computers require physical systems that are well-isolated from the decohering effects of their environment, while at the same time allowing precise manipulation during computation. They also require qubit-specific state initialization, measurement, and the generation of quantum correlations across the system (*1–4*). Although there has been substantial progress in proof-of-principle demonstrations of quantum computation (*5–8*), simultaneously meeting these demands has proven difficult. This motivates the search for schemes that can demonstrate quantum-enhanced computation under more favorable experimental conditions. Investigating the space between classical and universal quantum computers has attracted broad interest (*9–11*).

Boson sampling has recently been proposed as a specific quantum computation that is more efficient than its classical counterpart but only requires indistinguishable bosons, low decoherence linear evolution, and measurement (*12*). The distribution of bosons that have undergone a unitary transformation U is thought to be exponentially hard to sample from classically (*12*). The probability amplitude of obtaining a certain output is directly proportional to the permanent of a corresponding submatrix of U (*13*). The permanent expresses the wavefunction of identical bosons, which are symmetric under exchange (*14, 15*); in contrast, the Slater determinant expresses the wavefunction of identical fermions, which are antisymmetric under exchange. While determinants can be evaluated efficiently, permanents have long been believed to be hard to compute (*16*); the best known algorithm scales exponentially with the size of the matrix.

One can envision a race between a classical and a quantum machine to sample the boson distribution given an input state and U. The classical machine would evaluate at least part of the probability distribution, which requires the computation of matrix permanents. An ideal QBSM instead creates indistinguishable bosons, physically implements U, and records the outputs. While the QBSM is not believed to efficiently estimate any individual matrix permanent, for a sufficiently large system it is expected to beat the classical computer in sampling over the entire

distribution (*12*).

Photonics is a natural platform to implement boson sampling since sources of indistinguishable photons are well-developed (*17*), and integrated optics offers a scalable route to low decoherence linear transformations over many modes (*18*). Such circuits can be rapidly reconfigured to sample from a user-defined operation (*19, 20*). Importantly, boson sampling requires neither nonlinearities nor on-demand entanglement, unlike photonic approaches to universal quantum computation (*21*). This clears the way for experimental boson sampling with existing photonic technology, building on the extensively studied two-photon Hong-Ou-Mandel (HOM) interference effect (*22*).

A QBSM (Fig. 1) samples the output distribution of a multi-particle bosonic quantum state $|\Psi_{\rm out}\rangle$, prepared from a specified initial state $|\mathbf{T}\rangle$ and linear transformation $\mathbf{\Lambda}$. Unavoidable losses in the system imply $\mathbf{\Lambda}$ will not be unitary, though lossy QBSMs can still surpass classical computation (*12, 23*). A trial begins with the input state $|\mathbf{T}\rangle{=}|T_1...T_M\rangle \propto \prod_{i=1}^{M}(\hat{a}_i^\dagger)^{T_i}|0\rangle$, which describes $N{=}\sum_{i=1}^{M} T_i$ particles distributed in M input modes in the occupation-number representation. The output state $|\Psi_{\rm out}\rangle$ is generated according to the linear map between input and output mode creation operators $\hat{a}_i^\dagger{=}\sum_{j=1}^{M}\Lambda_{ij}\hat{b}_j^\dagger$, where $\mathbf{\Lambda}$ is an $M{\times}M$ matrix. Finally, the particles in each of the M output modes are counted. The probability of a particular measurement outcome $|\mathbf{S}\rangle{=}|S_1...S_M\rangle$ is given by

$$P(\mathbf{S}|\mathbf{T}) = |\langle\mathbf{S}|\Psi_{\rm out}\rangle|^2 = \frac{\left|\mathrm{Per}(\mathbf{\Lambda}^{(\mathbf{S},\mathbf{T})})\right|^2}{\prod_{j=1}^{M} S_j! \prod_{i=1}^{M} T_i!} \tag{1}$$

where the $N{\times}N$ submatrix $\mathbf{\Lambda}^{(\mathbf{S},\mathbf{T})}$ is obtained by keeping S_j (T_i) copies of the j^{th} column (i^{th} row) of $\mathbf{\Lambda}$ (*13*).

Our QBSM consists of sources of indistinguishable single photons, a multiport linear optical circuit, and single-photon counting detectors. Two parametric down-conversion (PDC) pair sources (*24*) are used to inject up to four photons into a silica-on-silicon integrated photonic

circuit, fabricated by UV writing (*19, 25*). The circuit is shown in Fig. 2A and consists of M=6 input and output spatial modes coupled by a network of ten beam splitters (*18*). The output state is measured with single-photon avalanche photodiodes on each mode. We only consider outcomes in which the number of detections equals the intended number of input photons (*13*).

Our central result of three- and four-boson sampling is shown in Fig. 3. In the first case, we repeatedly inject three photons in the input state $|\mathbf{T}\rangle = |011010\rangle$, monitor all outputs, and collect all three-fold coincident events. In the four-photon experiment, we use the input $|\mathbf{T}\rangle = |202000\rangle$ and record all four-fold events (*26*). For each experiment, the measured relative frequencies P_S^{exp} for every allowed outcome $|\mathbf{S}\rangle$ are shown along with their observed statistical variation. The corresponding theoretical P_S^{th}, calculated using the right-hand side of Eq. 1, are shown along with their uncertainties arising from the characterization of $\mathbf{\Lambda}$, described below.

We reconstruct $\mathbf{\Lambda}$ with a series of one- and two-photon transmission measurements to determine its complex-valued elements $\Lambda_{ij} = \tau_{ij} e^{i\phi_{ij}}$ (*27*). The characterization results for the circuit used in the three-photon experiment are shown in Fig. 2, B and C. To obtain the magnitude τ_{ij}, single photons are injected in mode i. The probability of a subsequent detection in mode j is given by $P_1(j,i) = |\Lambda_{ij}|^2 = \tau_{ij}^2$. The phases ϕ_{ij} are determined from two-photon quantum interference measurements. The probability that a photon is detected in each of modes j_1 and j_2 when they are injected in modes i_1 and i_2 is given by $P_2(j_1, j_2, i_1, i_2) = |\Lambda_{i_1 j_1}\Lambda_{i_2 j_2} + \Lambda_{i_2 j_1}\Lambda_{i_1 j_2}|^2$. This expression is used to find the relevant phases ϕ_{ij} given the previously determined magnitudes τ_{ij} (*13*).

To analyze the performance of our QBSM we compare our results to an ideal machine. We quantify the match of two sets of relative frequencies $\mathbf{P}^{(1)}$ and $\mathbf{P}^{(2)}$ by calculating the L_1 distance $d^{(N)}(\mathbf{P}^{(1)}, \mathbf{P}^{(2)}) = \frac{1}{2}\sum_S |P_S^{(1)} - P_S^{(2)}|$, where N denotes the number of photons in a sample (*28*). Identical and maximally dissimilar distributions correspond to d=0 and d=1, respectively. For our experiments we calculate $d^{(N)}(\mathbf{P}^{\text{exp}}, \mathbf{P}^{\text{th}})$ to give $d^{(3)}$=0.094 $\pm$ 0.014

and $d^{(4)}$=0.097 $\pm$ 0.004 (Fig. 3). Even in an ideal QBSM with perfect state preparation and detection, the statistical variations result in nonzero d. If we substitute for our experimental data a Monte Carlo sampling of $\mathbf{P}^{\mathrm{th}}$ with sample size equivalent to our experiments, we instead calculate $d^{(3)}$=0.043 $\pm$ 0.012 and $d^{(4)}$=0.059 $\pm$ 0.022. This suggests there are appreciable contributions to $d(\mathbf{P}^{\mathrm{exp}}, \mathbf{P}^{\mathrm{th}})$ beyond statistical deviation.

Due to experimental limitations, our QBSM occasionally samples distributions other than $\mathbf{P}^{\mathrm{th}}$. The dominant sources of this sampling inaccuracy in our experiment are multi-photon emission and partial distinguishability amongst the photons. In practice, all single-photon sources produce multiple photons with a finite probability (*17*). For our PDC sources, the output state is approximately $|00\rangle + \lambda|11\rangle + \lambda^2|22\rangle$, with $\lambda \ll 1$. Both single-photon and undesired multiphoton terms increase with λ. In our three-photon experiments, for example, multiphoton emission from the two PDC sources can lead to input states $|\mathbf{T}\rangle = |021010\rangle$ or $|012020\rangle$, which contribute to three-fold coincident events if photons are lost or emerge in the same output mode. In addition, partial distinguishability of the photons contaminates the distribution sampled by the QBSM by mixing in one- and two-photon interference effects (*29*).

We form a new distribution $\mathbf{P}^{\mathrm{mod}}$ that accounts for the effects of multiphoton emission and photon distinguishability (*13*). The distance $d(\mathbf{P}^{\mathrm{exp}}, \mathbf{P}^{\mathrm{mod}})$ shown by the green point (insets of Fig. 4, A and B), is found to be consistent with the statistical variation due to a finite sample size, for both the three- and four-photon experiments. This suggests we have correctly identified and modeled the sources of inaccuracy. To investigate how the performance of our QBSM depends on λ and photon distinguishability, we calculate $d(\mathbf{P}^{\mathrm{mod}}, \mathbf{P}^{\mathrm{th}})$ for a range of operating parameters (Fig. 4, C and D). In terms of λ, a clear tradeoff is presented between data rate and inaccuracy due to multiphoton emission, which is an intrinsic consequence of using PDC sources. Improvement in photon indistinguishability increases the fidelity to the ideal machine, and additionally is thought to enhance the computational power of a QBSM (*29*).

Our results demonstrate that boson sampling is related to the computation of matrix permanents, a problem believed to be classically hard. Our successful diagnosis of the source and magnitude of the principal sampling errors, as validated by a reduction in $d(\mathbf{P}^{\mathrm{exp}}, \mathbf{P}^{\mathrm{mod}})$ to within the statistical variation of a perfect QBSM, will inform the design of next-generation devices. While investigations into quantum-enhanced computation in the presence of errors is ongoing, it already appears that the boson sampling model makes less stringent demands on device performance than universal photonic quantum computers (*12, 23, 29*). There is thus reason for optimism that ongoing advances in integrated photonics such as reduced transmission loss, efficient number-resolving detectors (*30*), and multiplexed (*31, 32*) or single-emitter (*17*) photon sources, will enable larger QBSMs that outperform classical computers. Beyond the specific boson sampling problem, such a device would provide clear evidence for the computational power of quantum mechanics.

References and Notes

1. D. P. DiVincenzo, *Fortschritte der Physik* **48**, 771 (2000).

2. R. Raussendorf, H. J. Briegel, *Phys. Rev. Lett.* **86**, 5188 (2001).

3. M. A. Nielsen, *Phys. Lett. A* **308**, 96 (2003).

4. A. M. Childs, *Phys. Rev. Lett.* **102**, 180501 (2009).

5. P. Walther, *et al.*, *Nature* **434**, 169 (2005).

6. C.-Y. Lu, D. E. Browne, T. Yang, J.-W. Pan, *Phys. Rev. Lett.* **99**, 250504 (2007).

7. B. P. Lanyon, *et al.*, *Science* **334**, 57 (2011).

8. E. Lucero, *et al.*, *Nature Phys.* **8**, 719 (2012).

9. E. Knill, R. Laflamme, *Phys. Rev. Lett.* **81**, 5672 (1998).

10. S. P. Jordan, *Quant. Inf. Comput.* **10**, 470 (2010).

11. D. Shepherd, M. J. Bremner, *Proc. R. Soc. A* **465**, 1413 (2009).

12. S. Aaronson, A. Arkhipov, *Proceedings of ACM Symposium on the Theory of Computing*, STOC (2011).

13. *Materials and methods are available as supplementary material on Science Online* .

14. E. Caianiello, *Il Nuovo Cimento (1943-1954)* **10**, 1634 (1953).

15. L. Troyansky, N. Tishby, *Proceedings of PhysComp* (1996).

16. L. G. Valiant, *Theoretical Computer Science* **8**, 189 (1979).

17. M. D. Eisaman, J. Fan, A. Migdall, S. V. Polyakov, *Rev. Sci. Instrum.* **82**, 071101 (2011).

18. B. J. Metcalf, *et al.*, *arXiv:1208.4575v2* (2012).

19. B. J. Smith, D. Kundys, N. Thomas-Peter, P. G. R. Smith, I. A. Walmsley, *Opt. Express* **17**, 264 (2009).

20. P. J. Shadbolt, *et al.*, *Nature Photon.* **6**, 45 (2011).

21. E. Knill, R. Laflamme, G. J. Milburn, *Nature* **409**, 46 (2001).

22. C. K. Hong, Z. Y. Ou, L. Mandel, *Phys. Rev. Lett.* **59**, 2044 (1987).

23. P. P. Rohde, T. C. Ralph, *Phys. Rev. A.* **85**, 022332 (2012).

24. P. J. Mosley, *et al.*, *Phys. Rev. Lett.* **100**, 133601 (2008).

25. D. O. Kundys, J. C. Gates, S. Dasgupta, C. Gawith, P. G. R. Smith, *IEEE Photon. Technol. Lett.* **21**, 947 (2009).

26. Eq. 1 is expected to hold for any $|\mathbf{S}\rangle$ and $|\mathbf{T}\rangle$, however the classical hardness of sampling $P(\mathbf{S}|\mathbf{T})$ is maximized when $S_j, T_i \in \{0, 1\}$ for a given N and $\mathbf{\Lambda}$.

27. A. Laing, J. L. O'Brien, *arXiv:1208.2868v1* (2012).

28. A. Gilchrist, N. K. Langford, M. A. Nielsen, *Phys. Rev. A* **71**, 062310 (2005).

29. P. P. Rohde, *Phys. Rev. A* **86**, 052321 (2012).

30. T. Gerrits, *et al.*, *Phys. Rev. A* **84**, 060301 (2011).

31. A. L. Migdall, D. Branning, S. Castelletto, *Phys. Rev. A* **66**, 053805 (2002).

32. J. Nunn, *et al.*, *arXiv:1208.1534v1* (2012).

Acknowledgements: We thank Josh Nunn for valuable insights. This work was supported by the EPSRC (EP/C013840/1), the EC project Q-ESSENCE (248095), the Royal Society, and the AFOSR EOARD. XMJ is supported by an EU Marie-Curie Fellowship (PIIF-GA-2011-300820). JS acknowledges support from the United States Air Force Institute of Technology. The views expressed in this article are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.

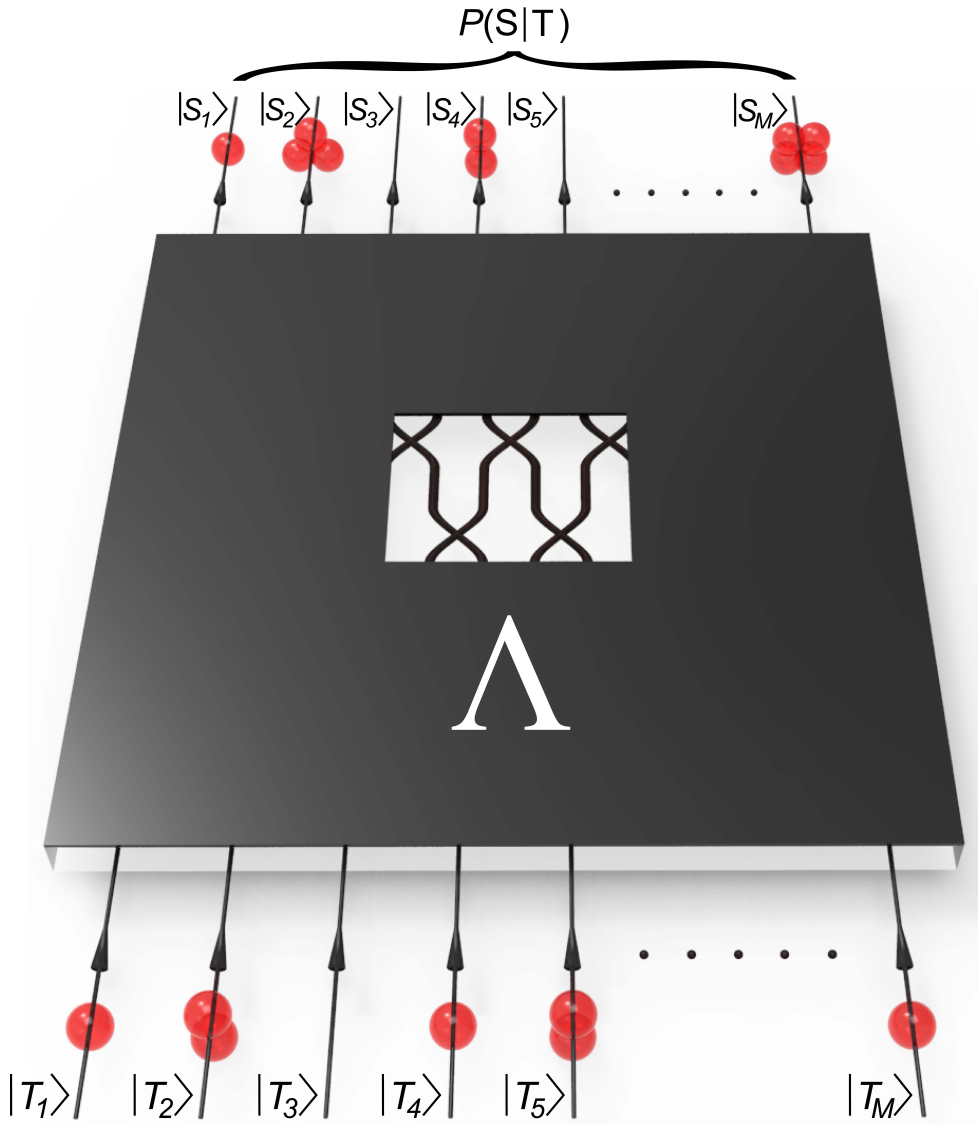

Figure 1: Model of quantum boson sampling. Given a specified initial number state $|\mathbf{T}\rangle=|T_1...T_M\rangle$ and linear transformation $\mathbf{\Lambda}$, a quantum boson sampling machine efficiently samples from the distribution $P(\mathbf{S}|\mathbf{T})$ of possible outcomes $|\mathbf{S}\rangle=|S_1...S_M\rangle$.

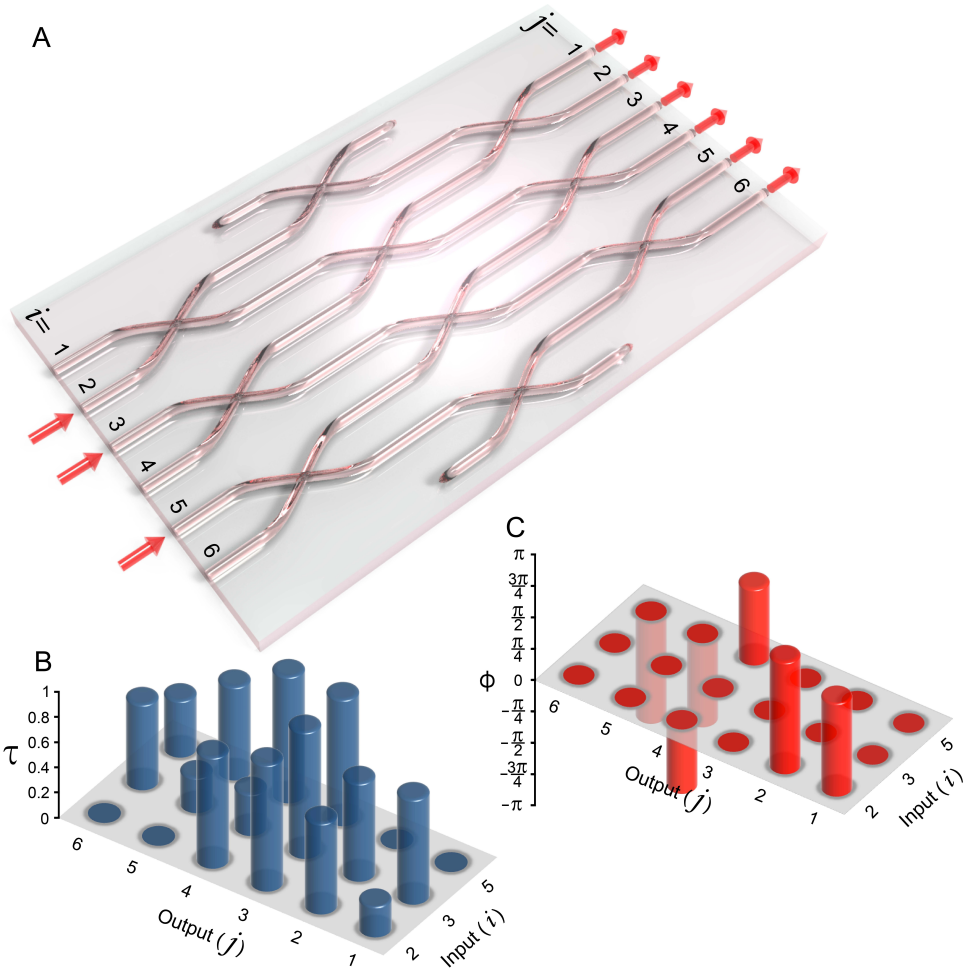

Figure 2: Schematic and characterization of the photonic circuit. (**A**) The silica-on-silicon waveguide circuits consist of $M{=}6$ accessible spatial modes (labeled $1{-}6$). For the three-photon experiment, we launch photons into inputs $i{=}2$, 3 and 5 from two parametric down-conversion sources which produce near-single photons and postselect outcomes in which three detections are registered amongst the output modes j. For the four-photon experiment, which is implemented on a different chip of identical geometry, we inject a double photon pair from a single source into the modes $i{=}1$, 3 and postselect on four detection events. (**B**-**C**) Measured elements of the linear transformation $\Lambda_{ij}{=}\tau_{ij}e^{i\phi_{ij}}$ linking the input mode i to the output mode j of our three-photon apparatus. The circuit geometry dictates that several τ_{ij} are zero, and our phase-insensitive input states and detection methods imply only six non-zero ϕ_{ij}. Since only relative values are needed due to post-selection, we rescale each row of τ so that its maximum value is unity.

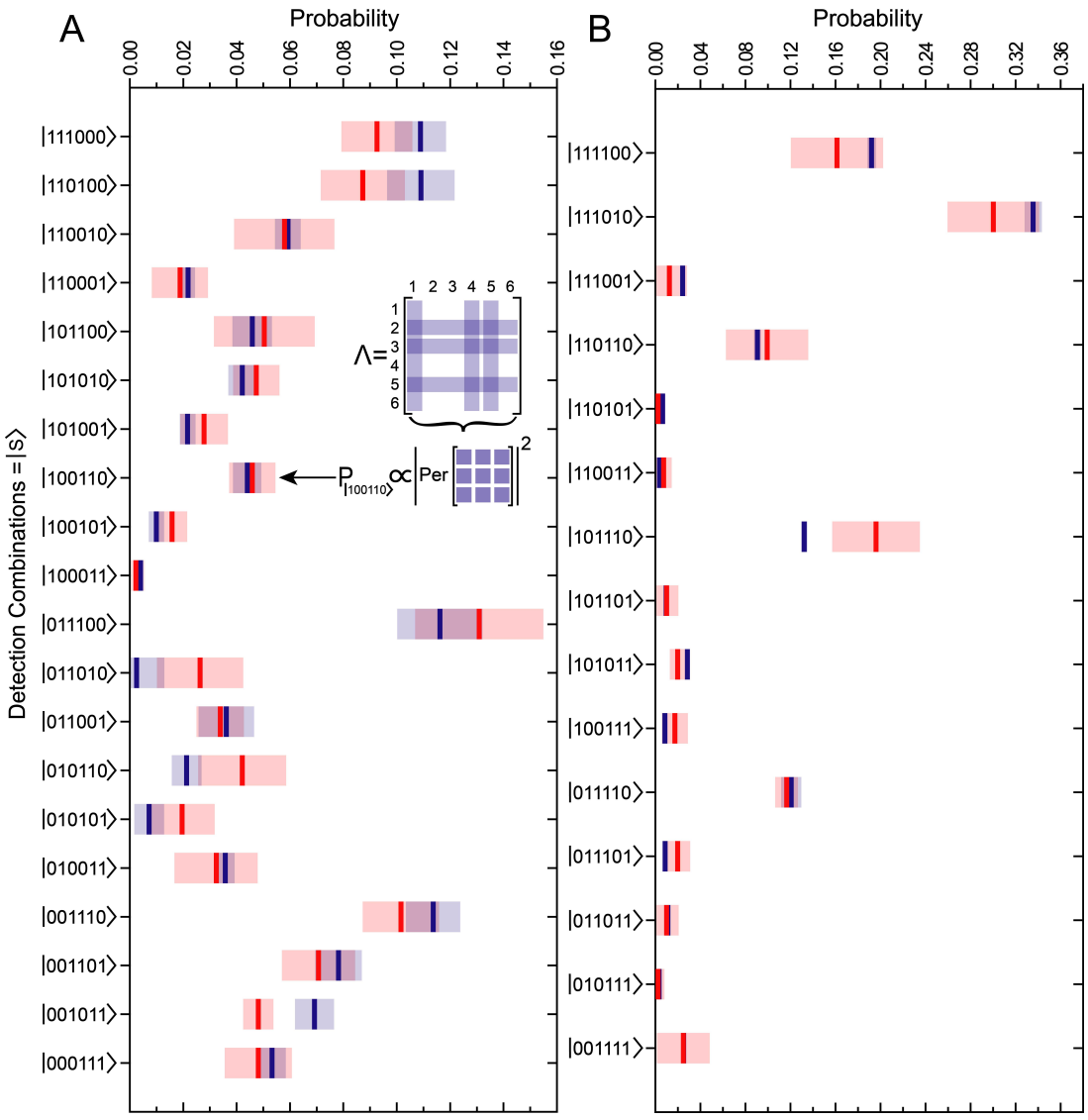

Figure 3: Boson sampling results. The measured relative frequencies $\mathbf{P}^{\text{exp}}$ of outcomes in which the photons are detected in distinct modes are shown in red for (**A**) three- and (**B**) four-photon experiments. Each data set is collected over 160 hours, and statistical variations in counts are shown by the red shaded bars. The theoretical distributions $\mathbf{P}^{\text{th}}$ (blue) are obtained from the permanents of submatrices constructed from the full transformation $\mathbf{\Lambda}$, as depicted in the inset. The blue error bars arise from uncertainties in the characterization of $\mathbf{\Lambda}$.

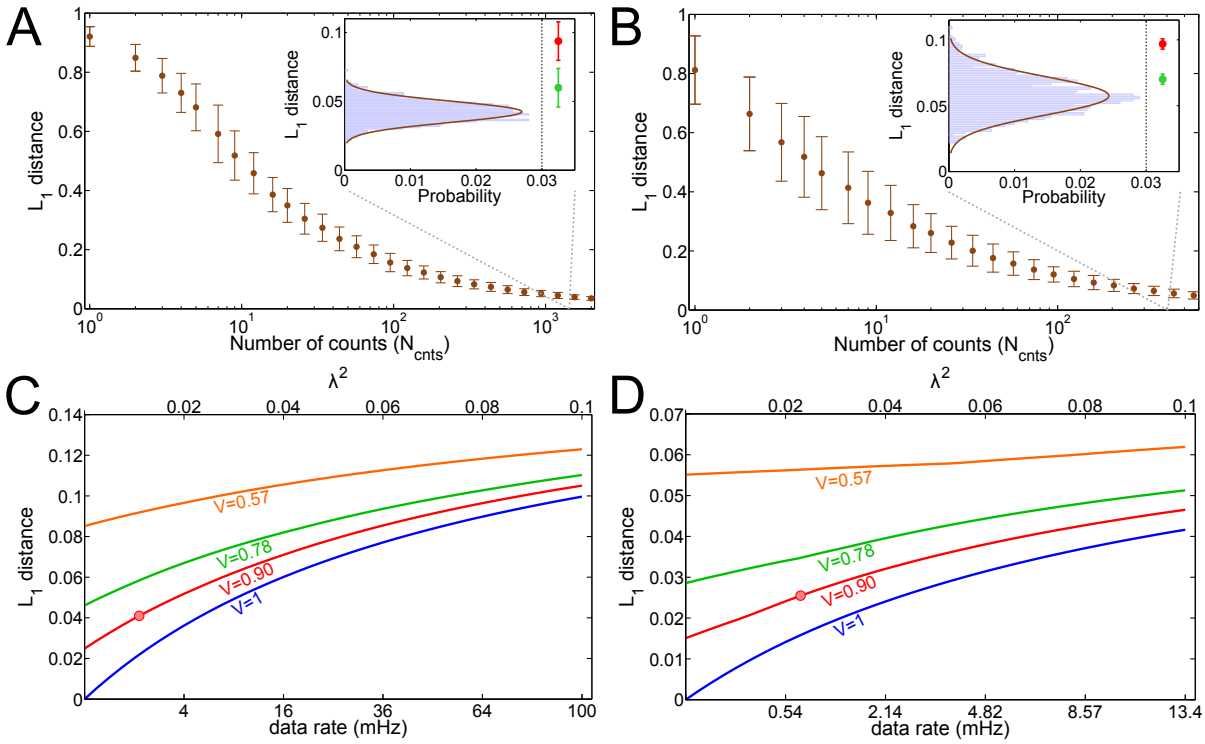

Figure 4: Sampling accuracy. We consider several boson distributions: the experimental samples $\mathbf{P}^{\text{exp}}$, the ideal predictions of the matrix permanent $\mathbf{P}^{\text{th}}$, and the predictions of the full model $\mathbf{P}^{\text{mod}}$ that includes higher-order emission and photon distinguishability. **(A)** The L_1 distance d between $\mathbf{P}^{\text{th}}$ and a Monte Carlo simulation of an ideal machine that samples $\mathbf{P}^{\text{th}}$ a finite number of times for our three- and **(B)** four-photon cases. The errors in this case are solely a result of the finite number of samples collected by the ideal machine. The inset histograms show the variation in d expected for a sample size corresponding to the 1421 and 405 counts collected in our three- and four-photon experiments. The distance $d(\mathbf{P}^{\text{exp}}, \mathbf{P}^{\text{th}})$ (red) suggests an underlying systematic inaccuracy as it falls outside the range of outputs of an ideal machine indicated in the histogram. Our full model is validated by the distance $d(\mathbf{P}^{\text{exp}}, \mathbf{P}^{\text{mod}})$ (green) which is consistent with statistical variation. The red and green dot positions correspond to the L_1 axis only. **(C)** The predicted variation in $d(\mathbf{P}^{\text{th}}, \mathbf{P}^{\text{mod}})$ is shown as a function of λ and the photon distinguishabilites, represented by the reduction in two-photon interference visibility V, for the three- and **(D)** four-photon cases. Our experimental conditions are marked (red dot).

Supplementary Materials for Boson Sampling on a Photonic Chip

Justin B. Spring[1], Benjamin J. Metcalf[1], Peter C. Humphreys[1], W. Steven Kolthammer[1], Xian-Min Jin[1,2], Marco Barbieri[1], Animesh Datta[1], Nicholas Thomas-Peter[1], Nathan K. Langford[1,3], Dmytro Kundys[4], James C. Gates[4], Brian J. Smith[1], Peter G.R. Smith[4], and Ian A. Walmsley[1*]

[1]Clarendon Laboratory, Department of Physics, University of Oxford, OX1 3PU, UK,

[2]Department of Physics, Shanghai Jiao Tong University, Shanghai 200240, PR China

[3] Department of Physics, Royal Holloway, University of London, TW20 0EX, UK

[4]Optoelectronics Research Centre, University of Southampton, Southampton, SO17 1BJ, UK

[*]To whom correspondence should be addressed; E-mail: i.walmsley1@physics.ox.ac.uk

1 Materials and Methods

Multiphoton states generation. An 80 MHz Ti-Sapphire oscillator outputs 100 fs pulses at 830 nm, which undergo type-I second harmonic generation in a 700 μm β-BaB_2O_4 (BBO) crystal. This 415 nm light is then used used to pump type-II collinear parametric downconversion (PDC) in 8 mm-long AR-coated Potassium Dihydrogen Phosphate (KDP) crystals (*23*). Both modes of the resulting squeezed state are passed through spectral filters (Semrock, $\Delta\lambda$=3 nm), to maximize photon indistinguishability. In the three-photon experiment, two such PDC sources are coupled to four polarization-maintaining (PM) fibers (*17*). One of the photons is sent directly to a heralding detector, while the other three are coupled into the photonic circuit via a PM v-groove array (VGA) on a 6-axis alignment stage at the chip input. Programmable optical delay is provided by motor-controlled translation stages on each single-photon mode preceding the circuit, allowing the photons to arrive temporally coincident at the interferometric network in Fig 2A. Another PM VGA at the chip output couples the output modes into avalanche photodiode (APD) single photon counting modules (PerkinElmer SPCM-AQ4C), which are monitored by a home-built coincidence counting program loaded onto a commercially available FPGA development board (Xilinx SP605) operating with a 5 ns coincidence window. In the four-photon experiment, the higher order emission ($|22\rangle$) from a single PDC crystal was launched into two

spatial modes of a different chip, though of identical geometry, which had VGAs glued to input and output. This reduced the coupling efficiency uncertainty for the four-photon QBSM, which reduced the size of the error bars in Fig. 3B. In both cases, to minimize undesired PDC emission terms the sources were pumped as lightly as possible while maintaining reasonable count rates, yielding $\lambda^2{=}0.011$ and 0.023 in the three- and four-photon sampling experiments respectively.

Photonic circuit fabrication. The boson sampling is performed on a UV-written silica-on-silicon integrated chip (*17*). The chip was fabricated by focusing a continuous wave UV laser (244 nm) onto a photosensitive layer to create a local increase in the index of refraction. The chip is then moved, via computer-controlled precision translation stages, transversely to the incident UV beam to trace out the desired waveguide network geometry. One can fabricate cross-coupling beamsplitters of a certain reflectivity by crossing waveguides at a specific angle. Such beamsplitters take up less space than traditional evanescent couplers in such circuits, allowing reduced chip size and lower loss (*32*). While there are eight spatial modes in the middle of the circuit (Fig. 2A), the outer modes are not fabricated to the edge of the chip and are thus not accessible to the experimenter. They can then be accurately treated as losses on the neighboring modes, leaving a six (accessible) mode interferometric network.

Boson sampling data collection. For the four-photon QBSM, the FPGA simultaneously counts all possible combinations of coincidence detection events (one detection event, two coincident detections etc) amongst the six APDs monitoring the output modes of the integrated circuit. For the three-photon QBSM, a similar set of coincidences is taken, but conditioned on detection of the herald photon by a separate APD. There is a nonzero background count rate on each detector which contributes to erroneous N-fold detection events when $N{-}1$ APDs detect photons and another APD erroneously fires. The $N{-}1$ coincidence rate is included in the above set of statistics collected by the FPGA and, by temporarily blocking all input modes to the circuit, one can estimate the background rate on each detector. The resulting background contribution to N-fold coincidences, comprising approximately 5% of the total counts, are then subtracted from the raw data, with the results plotted in Fig. 3.

Circuit characterization. Our photonic chip performs a non-unitary, complex-valued linear mapping of input to output modes $\Lambda_{ij}{=}\tau_{ij}e^{i\phi_{ij}}$. We follow the method outlined in (*25*) to use a set of one- and two-photon data to reconstruct $\mathbf{\Lambda}$.

We can find τ_{i_1,j_1} by coupling light (single photons) into mode i_1 and monitoring the probability of output in j_1 yielding

$$\begin{aligned}\left|\langle 0|\hat{b}_{j_1}\hat{a}^\dagger_{i_1}|0\rangle\right|^2 &= \left|\langle 0|\hat{b}_{j_1}\sum_{j=1}^{M}\Lambda_{i_1 j}\hat{b}^\dagger_j|0\rangle\right|^2 \\ &= |\Lambda_{i_1 j_1}|^2 \\ &= |\tau_{i_1 j_1}|^2 \end{aligned} \quad (1)$$

It is sufficient to measure the ratio of values between output ports, $(\tau_{i_1,j_1}/\tau_{i_1,j_2})^2$, as this describes the transformation induced by our circuit up to a constant factor for each input, x_i. If $\mathbf{\Lambda}$ is the true transformation, then this procedure gives us $\mathbf{X\Lambda}$ where $\mathbf{X}$ is a diagonal matrix with entries $X_{ii}{=}x_i$. However, the properties of the matrix permanent give $\text{Per}(\mathbf{X}\,\mathbf{\Lambda}){=}(\prod_i x_i)\text{Per}(\mathbf{\Lambda})$. Therefore, Eq. 1 of the main text shows, since we always launch the same input state, every portion of our boson distribution is multiplied by a constant factor that cancels in the normalization. This data is collected by periodically pausing boson sampling data collection, blocking two of the three photon inputs, and measuring the relative power in each output mode. We thus obtain accurate values for τ as well as a variance that is important for calculating the error bars in Fig. 3.

One can then use two photon interference to find $\phi_{i,j}$ (*25*). If one inputs two photons into modes i_1, i_2 and detect in modes j_1, j_2, then we have $|S\rangle{=}\hat{a}^\dagger_{j_1}\hat{a}^\dagger_{j_2}|0\rangle$ and $|T\rangle{=}\hat{a}^\dagger_{i_1}\hat{a}^\dagger_{i_2}|0\rangle$. If single, indistinguishable photons are used, then the probability of post-selecting this output is

$$\begin{aligned}
P_{\text{indist}} &= \left|\text{Per}(\mathbf{\Lambda}^{(\mathbf{S},\mathbf{T})})\right|^2 \\
&= \left|\Lambda^{(\mathbf{S},\mathbf{T})}_{11}\Lambda^{(\mathbf{S},\mathbf{T})}_{22} + \Lambda^{(\mathbf{S},\mathbf{T})}_{12}\Lambda^{(\mathbf{S},\mathbf{T})}_{21}\right|^2 \\
&= (\tau_{i_1,j_1}\tau_{i_2,j_2})^2 + (\tau_{i_1,j_2}\tau_{i_2,j_1})^2 \\
&+ 2\tau_{i_1,j_1}\tau_{i_1,j_2}\tau_{i_2,j_1}\tau_{i_2,j_2} \\
&\times \cos(\phi_{i_1,j_1} - \phi_{i_1,j_2} - \phi_{i_2,j_1} + \phi_{i_2,j_2})
\end{aligned} \tag{2}$$

If the photons launched into the individual modes are distinguishable, then we get the incoherent sum of their individual statistics. We again take a matrix permanent, but as this is an incoherent process one finds

$$P_{\text{dist}} = (\tau_{i_1,j_1}\tau_{i_2,j_2})^2 + (\tau_{i_1,j_2}\tau_{i_2,j_1})^2 \tag{3}$$

One can then perform a Hong-Ou-Mandel experiment and find that the resulting interference visibility is

$$\begin{aligned}
V &= \frac{P_{dist} - P_{indist}}{P_{dist}} \\
&= \frac{2\tau_{i_1,j_1}\tau_{i_1,j_2}\tau_{i_2,j_1}\tau_{i_2,j_2}}{(\tau_{i_1,j_1}\tau_{i_2,j_2})^2 + (\tau_{i_1,j_2}\tau_{i_2,j_1})^2} \\
&\times \cos(\phi_{i_1,j_1} + \phi_{i_2,j_2} - \phi_{i_1,j_2} - \phi_{i_2,j_1})
\end{aligned} \tag{4}$$

We perform this two-photon interference experiment and fit a Gaussian to the resulting data to determine the visibility. Using the known $\tau_{i,j}$, one can then find $|\phi_{i_1,j_1} + \phi_{i_2,j_2} - \phi_{i_1,j_2} - \phi_{i_2,j_1}|$. Repeating this procedure for all accessible two photon dips, and applying additional constraints one can determine $\phi_{i,j}$ (*25*).

For our circuit geometry, we encounter an overconstrained problem as we measure more interference visibilities than unknown ϕ. Therefore, we run a least squares minimization to find

the set of ϕ that best fits our two photon interference data. To find the error bars in Fig. 3, we use a Monte Carlo method where the elements of $\mathbf{\Lambda}$ are selected from a normal distribution with an appropriate variance for each element. The one photon measurement was repeated periodically during the 160 hour long boson sampling data collection, thus yielding the variance in the τ_{ij} characterization, which was determined to dominate the uncertainty in the predicted boson distribution in Fig. 3. This explains why the four-photon QBSM, which had VGAs glued to the ends and thus was much less susceptible to changes in the coupling, has significantly smaller error bars in the predicted distribution shown in Fig. 3B. The uncertainty in the Gaussian fit to the two-photon interference patterns and in τ_{ij} was then used to determine the variance in ϕ_{ij}.

This characterization process is efficient, as a general linear transformation over M modes can be described by $\mathcal{O}(M^2)$ parameters, requiring $\mathcal{O}(M^2)$ measurements with this technique. While photons from the PDC sources were used for the circuit characterization here, one can also use classical coherent states (*25, 33*). However, the single-photon based technique outlined here benefits from not requiring phase-stable path length matching and, because we use the same sources for boson sampling and characterization, the photonic degrees of freedom (polarization, spectrum etc) for the characterization match that used in the experiment.

With $\Lambda_{i,j}$ experimentally determined over the accessible modes, one can predict the post-selected boson distribution for any input/output $|S\rangle$, $|T\rangle$ by constructing $\mathbf{\Lambda}^{(\mathbf{S},\mathbf{T})}$ from $\mathbf{\Lambda}$ and taking the permanent according to Eq. (1) in the main text. Thus, we effectively use the one and two photon boson distributions to characterize our non-unitary operation over the accessible modes. One can then take various matrix permanents of this non-unitary matrix to predict the boson distribution for *any* N.

2 Supplementary Text

In this section we first outline how the boson distribution is given by a set of matrix permanents. We then show how the boson distribution can be accurately predicted by the permanents associated with a non-unitary matrix describing a lossy channel, $\mathbf{\Lambda}$, if one post-selects on trials where no photons are lost. Finally, the principal sources of error in this experiment, namely the photon distinguishability and higher order terms from our PDC photon sources, are discussed.

2.1 The boson distribution is given by a set of matrix permanents

We assume N bosons are injected into a network that performs a unitary transformation over M modes. We consider the special case, appropriate to our experiment, where the input (and output) states contain no more than one boson per mode, though the general case is treated elsewhere (*34*). Without loss of generality, let modes 1 to N contain an input boson, while

modes $N+1$ to M have vacuum inputs. The input state can then be described by

$$|\Psi_{\text{in}}\rangle = |T\rangle = \prod_{i=1}^{N} \hat{a}_i^\dagger |0\rangle \tag{5}$$

The unitary transformation allows one to evolve the operators according to

$$\hat{a}_i^\dagger = \sum_{j=1}^{M} U_{ij} \hat{b}_j^\dagger \tag{6}$$

where $\hat{a}_i^\dagger$ and $\hat{b}_j^\dagger$ are creation operators on the i-th input and j-th output mode respectively. We then obtain the output state

$$|\Psi_{\text{out}}\rangle = \prod_{i=1}^{N} \left(\sum_{j=1}^{M} U_{ij} \hat{b}_j^\dagger \right) |0\rangle \tag{7}$$

To find the boson distribution, we project our output onto a state $|\mathbf{S}\rangle$ which, in the number state basis, we describe by an N element vector $\mathbf{S}$, where S_j gives the mode of the j-th boson. The probability of measuring this state is then

$$\begin{aligned} P_S &= |\langle \mathbf{S} | \psi_{\text{out}} \rangle|^2 \\ &= \left| \langle 0 | \prod_{i=1}^{N} \hat{b}_{S_i} \left[\prod_{j=1}^{N} \left(\sum_{k=1}^{M} U_{jk} \hat{b}_k^\dagger \right) \right] |0\rangle \right|^2 \end{aligned} \tag{8}$$

The term in square brackets can be expanded and includes M^N terms, as one is selecting N bosons from M modes where repetitions are allowed (> 1 boson in a mode). One can rewrite this term in square brackets to give

$$P_S = \left| \langle 0 | \prod_{i=1}^{N} \hat{b}_{S_i} \left[\sum_{j=1}^{M^N} \left(\prod_{k=1}^{N} U_{k,\widetilde{V}_k^j} \hat{b}_{\widetilde{V}_k^j}^\dagger \right) \right] |0\rangle \right|^2 \tag{9}$$

where $\widetilde{\mathbf{V}}$ is the set of M^N permutations of N photons amongst M modes, repetitions allowed. The tilde notation will be used throughout this paper for a set of permutations. Then, $\widetilde{V}_k^j$ indicates the mode of the k-th boson in the j-th permutation. As an example, consider the case with $M = 3$ modes and $N = 2$ input bosons, then

$$\begin{array}{lll} \widetilde{V}^1 = [1,1] & \widetilde{V}^2 = [1,2] & \widetilde{V}^3 = [1,3] \\ \widetilde{V}^4 = [2,1] & \widetilde{V}^5 = [2,2] & \widetilde{V}^6 = [2,3] \\ \widetilde{V}^7 = [3,1] & \widetilde{V}^8 = [3,2] & \widetilde{V}^9 = [3,3] \end{array} \tag{10}$$

Let us denote all $N!$ permutations of $\mathbf{S}$ by $\widetilde{\mathbf{S}}$ where $\widetilde{S}_k^j$ indicates the mode of the k-th boson in the j-th permutation. For example, if we project onto the state $|S\rangle$=$|011\rangle$ then $\mathbf{S}$=$[2,3]$ and

$$\widetilde{S}^1 = [2,3] \quad \widetilde{S}^2 = [3,2] \tag{11}$$

It is clear we only retain terms from the summation in Eq. 9 where $\widetilde{V}^j \in \widetilde{S}$, otherwise at least one annihilation operator will act on vacuum and give P_S=0. This then leaves us with

$$P_S = \left| \sum_{j=1}^{N!} \prod_{k=1}^{N} U_{k,\widetilde{S}_k^j} \right|^2 \tag{12}$$

The formula for the permanent of an $n \times n$ matrix A with elements a_{ij} is

$$\text{Per}(A) = \sum_{i=1}^{n!} \prod_{j=1}^{n} a_{j,\widetilde{\sigma}_j^i} \tag{13}$$

where $\widetilde{\sigma}_j^i$ gives the j-th element of the i-th permutation of the numbers $1, 2, ..., n$. The term inside the modulus of Eq. 12 has the same form as the matrix permanent in Eq. 13. Our original unitary, U, can be described by an $M \times M$ matrix. However, it is obvious from Eq. 12 that, in general, we take the permanent of a subsection of U. Specifically, we only keep rows $1 \rightarrow N$, those rows corresponding to modes with input photons. In addition, we only keep columns corresponding to the elements of S. Let us call this modified subsection of our original unitary $U^{(\mathbf{S},\mathbf{T})}$. Then, using the definition of the permanent we can rewrite Eq. 12

$$P(\mathbf{S}|\mathbf{T}) = \left|\text{Per}(U^{(\mathbf{S},\mathbf{T})})\right|^2 \tag{14}$$

If one allows the possibility of more than one photon per input and output mode, then a similar analysis yields Eq. 1 in the main text (*34*).

We also note that the above treatment assumes the bosonic commutation relation $[\hat{b}_i^\dagger, \hat{b}_j^\dagger] = 0 \, \forall \, i,j$. If the system consisted of indistinguishable fermions, then the corresponding anticommutation relation $\{\hat{b}_i^\dagger, \hat{b}_j^\dagger\} = 0 \, \forall \, i,j$ would be used, leading to alternating plus and minus signs introduced in the summation in Eq. 12, yielding the easily classically computable determinant.

2.2 Effects of loss

In any experimental implementation of boson sampling there will be losses. These losses, regardless of where they occur, can be modeled as beam splitters that link accessible to inaccessible modes (*35*). When these losses are considered, it is important to ask whether the boson distribution is still given by a set of matrix permanents and, if so, what linear transformation does that matrix describe.

Let us adopt the convention that modes 1 to M are accessible modes while inaccessible loss modes are given the labels $M+1$ to L. There is an $L \times L$ unitary operation describing the

evolution of our pure input state, though we must trace over these loss modes at the output, yielding a mixed state over the accessible modes. Let us assume that photons are input into modes 1 to N where $N \leq M$ and we post-select on cases where N photons are detected, by definition, in the accessible modes 1 to M. Then Eq. 9 becomes

$$P_S = \left| \langle 0| \prod_{i=1}^{N} \hat{b}_{S_i} \left[\sum_{j=1}^{L^N} \left(\prod_{k=1}^{N} U_{k,\widetilde{V}_k^j} \hat{b}^\dagger_{\widetilde{V}_k^j} \right) \right] |0\rangle \right|^2 \tag{15}$$

but $S_i \leq M$ as we can only project on accessible modes. Therefore, even though U is an $L \times L$ matrix and the elements of $\widetilde{V}$ range from $1 \rightarrow L$, when we project onto $\widetilde{S}$ (all the permutations of S) we are left with

$$\begin{aligned} P_S &= \left| \sum_{j=1}^{N!} \prod_{k=1}^{N} U_{k,\widetilde{S}_k^j} \right|^2 \\ &= \left| \text{Per}(\mathbf{U}^{(\mathbf{S},\mathbf{T})}) \right|^2 = \left| \text{Per}(\mathbf{\Lambda}^{(\mathbf{S},\mathbf{T})}) \right|^2 \end{aligned} \tag{16}$$

where $U^{(\mathbf{S},\mathbf{T})}$ is again a modified version of the original unitary but only keeping rows $1 \rightarrow N$ and columns in $S'_i \leq M$. Since these elements always describe the accessible modes, then we can equivalently work in terms of Λ, where $\Lambda_{i,j}$=$U_{i,j}$ but $i, j \leq M$. In summary, when post-selecting on no bosons being lost, one can work in terms of Λ, a non-unitary linear transformation that is simply the subsection of U over the accessible modes. The matrix permanents of such a non-unitary linear transformation lead to the theoretical predictions in Fig. 3 of the main text.

Equivalently, we can describe our system as a noisy (lossy) quantum channel in the operator sum representation (*36*). This formalism will be useful later when we discuss sources of error from higher order PDC terms. In this picture, the accessible modes are in the system Q, while all inaccessible loss modes form the environment system E. We can describe the transformation induced by our circuit over the full space $Q \otimes E$ by a unitary operation U. Let ρ and σ be the inputs to Q and E respectively, then the output in Q after a projective measurement P_m and tracing over the environment is described by

$$\rho_{\text{out}} = \text{tr}_E(P_m U(\rho \otimes \sigma) U^\dagger P_m) \tag{17}$$

Let the basis for E be described by $|e_k\rangle$ and the initial state of the environment be $\sigma = \sum_j q_j |j\rangle\langle j|$, then we can express Eq. 17 as

$$\rho_{\text{out}} = \sum_{jk} E_{jk} \rho E_{jk}^\dagger \tag{18}$$

where $E_{jk} = \sqrt{q_j}\langle e_k | P_m U | j \rangle$ are the Kraus operators. We do not directly characterize U, as it extends over the environment which is inaccessible to the experimenter. However, with photons

we can assume $\sigma = |0\rangle\langle 0|$, and all of our boson sampling results post-select on the case where no photons are lost to the environment. Therefore, the summation in Eq. 18 reduces to only one term with postselection

$$\rho_{\rm out} = E_{00}\rho E_{00}^{\dagger} \tag{19}$$

Experimental boson sampling efforts will sample a non-unitary transformation that is equivalent, when post-selecting on no bosons being lost, to the E_{00} Kraus operator.

2.3 Sources of error

The computational difficulty for a classical machine to sample a boson distribution increases as the maximum error threshold is lowered. Therefore, while a QBSM need not sample the true boson distribution perfectly (*12*), it will be easier to beat a classical machine if future QBSMs designs minimize their sampling errors. In this paper, we have benchmarked the accuracy of our QBSMs by inferring the probability distribution from our data, labeled $\mathbf{P}^{\rm exp}$, and comparing it to the distribution obtained from Eq. 1, labeled $\mathbf{P}^{\rm th}$. Throughout the text, we quantify the distance between two probability distributions via the L_1 distance, $d(\mathbf{P}^{(1)}, \mathbf{P}^{(2)}) = \frac{1}{2}\sum_{\rm i} |{\rm P}_{\rm i}^{(1)} - {\rm P}_{\rm i}^{(2)}|$.

Our method of benchmarking QBSM accuracy will always yield a nonzero d due to the finite number of collected samples. We perform a Monte Carlo simulation of $\mathbf{P}^{\rm exp}$ from a QBSM that perfectly samples $\mathbf{P}^{\rm th}$ as a function of the number of counts collected (Fig. 4, A and B), to show the rate at which d asymptotically approaches zero as the number of samples collected increases. In the inset histograms, we show the range of outputs for this ideal QBSM for the actual number of experimental counts collected in the three and four photon cases, while the red dots indicate $d(\mathbf{P}^{\rm th}, \mathbf{P}^{\rm exp})$. These two probability distributions show close agreement in Fig. 3, however a comparison of the histogram and red dots in Fig. 4, A and B indicates there is an additional source of error beyond the finite number of samples.

Due to experimental limitations, occasionally we sample distributions other than $\mathbf{P}^{\rm th}$. We model the effect of two such imperfections, photon distinguishability and higher order terms from our PDC sources, and form a new distribution $\mathbf{P}^{\rm mod}$ that accounts for these effects. We ignore the effects of photon impurity, as our post-selected data collection and use of nearly-spectrally factorable photon sources (*23*), minimizes this contribution. We find that the new $d(\mathbf{P}^{\rm exp}, \mathbf{P}^{\rm mod})$, indicated with the green dot in Fig. 4, A and B, comes well within the output variance of an ideal machine. This indicates we have correctly diagnosed and modeled the principal sources of experimental error, which will be important in guiding designs of future, larger N, QBSMs.

2.3.1 Effect of using heralded single photon sources

The parametric downconversion sources we use to generate our photons actually generate a two-mode squeezed state that is given by

$$|\Psi_{\mathrm{PDC}}\rangle = \sqrt{1-\lambda^2}\sum_{n=0}^{\infty}\lambda^n|nn\rangle \tag{20}$$

where $0 \leq \lambda < 1$ is the squeezing parameter whose magnitude is determined, in part, by the type of crystal and pump power used. For our experiment we wish to minimize higher order terms ($|22\rangle$ and $|33\rangle$ for the three- and four-photon QBSMs respectively) and so we deliberately lower our pump power as much as possible while maintaining a feasible count rate. For the three-photon experiment, $\lambda{=}\sqrt{0.011}$ and for the four-photon experiment $\lambda{=}\sqrt{0.023}$. However, even in this case we will sometimes inject more than N photons into our circuit which, due to losses, could be observed as an N-fold detection at the output.

Due to the circuit characterization method employed, we have no information about what such terms will be. In the operator sum representation, our characterization only determines the E_{00} Kraus operator, which describes the transformation of our input state when the environment, which includes all loss modes, starts and ends with zero photons. For example, if we instead inject five photons and lose two, then this process is described by the E_{20} operator, about which we have no information.

To model the effect of using such squeezed sources, we start with a circuit with the same geometry used in the experiment. Non-uniform losses throughout the circuit can then be modeled by adding beam splitters that link the depicted accessible modes shown in Fig. 2A to inaccessible loss modes, where the beam splitter reflectivity indicates the loss in that channel (*35*). Such 'loss beamsplitters' are added throughout the circuit. We cannot directly characterize these losses in our current circuit, however we can estimate them numerically. The characterized linear transformation $\mathbf{\Lambda}$ is a function of these losses as well as the fabricated interferometers shown in Fig. 2A. We have performed a loss-independent characterization of the interferometers (*17*), and then input this data into a genetic algorithm to find the relative losses between modes that best reproduces the characterized $\mathbf{\Lambda}$. We then apply three independent scaling factors to the relative losses at the sources, circuit and detectors. The source scaling factor is chosen to match the known source heralding efficiency, which is a measure of the loss in each source arm. The detector losses are scaled such that no detector has an efficiency greater than 50%, which is appropriate for APDs detecting photons at 830 nm. Finally, we scale the relative losses in the circuit to reproduce the known overall system transmission observed experimentally.

With knowledge of these losses, we reproduce an actual unitary linear transformation $\mathbf{U}$ that extends over both the $N{=}6$ modes as well as all loss modes and can be used to accurately predict the effect of higher order PDC terms. Taking the three-photon QBSM as an example, we use Eq. 1 of the main text to find the probability distributions when the input is $|\mathbf{T}\rangle{=}|011010\rangle$ (desired single-photon input), as well as $|011020\rangle$ or $|022010\rangle$ (the first higher-order terms from our two sources), which we label $\mathbf{P}^{111}$, $\mathbf{P}^{112}$ and $\mathbf{P}^{221}$ respectively. The boson distributions are then

found by summing the terms where three photons appear in the desired accessible modes and zero, one, or two (respectively) photons appear in any combination of loss modes. For example, the probability of obtaining $|\mathbf{S}\rangle$=$|111000\rangle$ given input $|\mathbf{T}\rangle = |022010\rangle$ is the summation of all terms where three photons appear in the first three accessible modes and two photons appear in any combination of loss modes. The higher order probability distributions, $\mathbf{P}^{221}$ and $\mathbf{P}^{112}$, are weighted by λ^2 which is obtained via a conditional second order correlation measurement, $g^{(2)}(0)$ (*37*). As λ^2 is small, we only consider the first higher order terms from each source.

2.3.2 Photon Distinguishability

Boson sampling assumes indistinguishable bosons, while experimental implementations will always have some distinguishability. Assuming pure inputs we follow the notation of (*28*) to write our input state as

$$|\psi_{\text{in}}\rangle = \prod_{i}^{N} \left(\alpha A^{\dagger}_{\xi_0,i} + \sqrt{1-\alpha^2} A^{\dagger}_{\xi_i,i}\right) |0\rangle \tag{21}$$

where N is again the number of photons, α is a distinguishability parameter, and $A^{\dagger}_{\xi_j,i}$ is the creation operator for photon i in mode ξ_j. Each photon is in a superposition of a desired mode ξ_0 and another mode ξ_i. By analyzing the reduction in HOM dip visibility at a beamsplitter inside our circuit we find $\alpha = 0.974$ on average in our experiment.

If one photon is distinguishable from the others, then the new probability distribution is given by the permanents of $N-1$ matrices which are incoherently summed. For example, assume an input state $|\mathbf{T}\rangle$=$|1\rangle^{\tau}|11000\rangle$ where τ labels a distinguishable photon, then for a unitary transformation U the probability of obtaining an output $|\mathbf{S}\rangle$ is

$$\begin{aligned} P(\mathbf{S}|\mathbf{T}) = &\left|U^{(\mathbf{S},\mathbf{T})}_{11}\right|^2 |U^{(\mathbf{S},\mathbf{T})}_{22}U^{(\mathbf{S},\mathbf{T})}_{33} + U^{(\mathbf{S},\mathbf{T})}_{23}U^{(\mathbf{S},\mathbf{T})}_{32}|^2 \\ &+ \left|U^{(\mathbf{S},\mathbf{T})}_{12}\right|^2 \left|U^{(\mathbf{S},\mathbf{T})}_{21}U^{(\mathbf{S},\mathbf{T})}_{33} + U^{(\mathbf{S},\mathbf{T})}_{23}U^{(\mathbf{S},\mathbf{T})}_{31}\right|^2 \\ &+ \left|U^{(\mathbf{S},\mathbf{T})}_{13}\right|^2 \left|U^{(\mathbf{S},\mathbf{T})}_{21}U^{(\mathbf{S},\mathbf{T})}_{32} + U^{(\mathbf{S},\mathbf{T})}_{22}U^{(\mathbf{S},\mathbf{T})}_{31}\right|^2 \end{aligned} \tag{22}$$

where the terms in parentheses are permanents of 2×2 matrices. We calculate these probability distributions when one photon is distinguishable and weight them by $|\alpha^2\sqrt{1-\alpha^2}|^2$ and $|\alpha^3\sqrt{1-\alpha^2}|^2$, the probability that one photon is distinguishable from the others for the three- and four-photon cases respectively. As α is large, we ignore the case when two photons are distinguishable.

References and Notes

1. D. P. DiVincenzo, *Fortschritte der Physik* **48**, 771 (2000).

2. R. Raussendorf, H. J. Briegel, *Phys. Rev. Lett.* **86**, 5188 (2001).

3. M. A. Nielsen, *Phys. Lett. A* **308**, 96 (2003).

4. A. M. Childs, *Phys. Rev. Lett.* **102**, 180501 (2009).

5. P. Walther, *et al.*, *Nature* **434**, 169 (2005).

6. C.-Y. Lu, D. E. Browne, T. Yang, J.-W. Pan, *Phys. Rev. Lett.* **99**, 250504 (2007).

7. B. P. Lanyon, *et al.*, *Science* **334**, 57 (2011).

8. E. Lucero, *et al.*, *Nature Phys.* **8**, 719 (2012).

9. E. Knill, R. Laflamme, *Phys. Rev. Lett.* **81**, 5672 (1998).

10. S. P. Jordan, *Quant. Inf. Comput.* **10**, 470 (2010).

11. D. Shepherd, M. J. Bremner, *Proc. R. Soc. A* **465**, 1413 (2009).

12. S. Aaronson, A. Arkhipov, *Proceedings of ACM Symposium on the Theory of Computing*, STOC (2011).

13. *Materials and methods are available as supplementary material on Science Online* .

14. E. Caianiello, *Il Nuovo Cimento (1943-1954)* **10**, 1634 (1953).

15. L. Troyansky, N. Tishby, *Proceedings of PhysComp* (1996).

16. L. G. Valiant, *Theoretical Computer Science* **8**, 189 (1979).

17. B. J. Metcalf, *et al.*, *arXiv:1208.4575v2* (2012).

18. B. J. Smith, D. Kundys, N. Thomas-Peter, P. G. R. Smith, I. A. Walmsley, *Opt. Express* **17**, 264 (2009).

19. P. J. Shadbolt, *et al.*, *Nature Photon.* **6**, 45 (2011).

20. E. Knill, R. Laflamme, G. J. Milburn, *Nature* **409**, 46 (2001).

21. C. K. Hong, Z. Y. Ou, L. Mandel, *Phys. Rev. Lett.* **59**, 2044 (1987).

22. P. P. Rohde, T. C. Ralph, *Phys. Rev. A.* **85**, 022332 (2012).

23. P. J. Mosley, *et al.*, *Phys. Rev. Lett.* **100**, 133601 (2008).

24. Eq. 1 is expected to hold for any $|\mathbf{S}\rangle$ and $|\mathbf{T}\rangle$, however the classical hardness of sampling $P(\mathbf{S}|\mathbf{T})$ is maximized when $S_j, T_i \in \{0, 1\}$ for a given N and $\mathbf{\Lambda}$.

25. A. Laing, J. L. O'Brien, *arXiv:1208.2868v1* (2012).

26. A. Gilchrist, N. K. Langford, M. A. Nielsen, *Phys. Rev. A* **71**, 062310 (2005).

27. M. D. Eisaman, J. Fan, A. Migdall, S. V. Polyakov, *Rev. Sci. Instrum.* **82**, 071101 (2011).

28. P. P. Rohde, *Phys. Rev. A* **86**, 052321 (2012).

29. T. Gerrits, *et al.*, *Phys. Rev. A* **84**, 060301 (2011).

30. A. L. Migdall, D. Branning, S. Castelletto, *Phys. Rev. A* **66**, 053805 (2002).

31. J. Nunn, *et al.*, *arXiv:1208.1534v1* (2012).

32. D. O. Kundys, J. C. Gates, S. Dasgupta, C. Gawith, P. G. R. Smith, *IEEE Photon. Technol. Lett.* **21**, 947 (2009).

33. S. Rahimi-Keshari, *et al.*, *arXiv:1210.6463v1* (2012).

34. S. Scheel, *arXiv:0406127v1* (2004).

35. N. Thomas-Peter, *et al.*, *New J. Phys.* **13**, 055024 (2011).

36. M. A. Nielsen, I. L. Chuang, *Quantum Computation and Quantum Information (Cambridge Series on Information and the Natural Sciences)* (Cambridge University Press, 2004).

37. B. J. Smith, P. Mahou, O. Cohen, J. S. Lundeen, I. A. Walmsley, *Opt. Express* **17**, 23589 (2009).