# Scaling of boson sampling experiments

P. D. Drummond, B. Opanchuk, L. Rosales-Zárate, and M. D. Reid
*Centre for Quantum and Optical Science, Swinburne University of Technology, Hawthorn, Victoria 3122, Australia*

P. J. Forrester
*Department of Mathematics and Statistics, ARC Centre of Excellence for Mathematical & Statistical Frontiers, The University of Melbourne, Victoria 3010, Australia*

(Received 2 May 2016; published 24 October 2016)

Boson sampling is the problem of generating a multiphoton state whose counting probability is the permanent of an $n \times n$ matrix. This is created as the output $n$-photon coincidence rate of a prototype quantum computing device with $n$ input photons. It is a fundamental challenge to verify boson sampling, and therefore the question of how output count rates scale with matrix size $n$ is crucial. Here we apply results from random matrix theory as well as the characteristic function approach from quantum optics to establish analytical scaling laws for average count rates. We treat boson sampling experiments with arbitrary inputs, outputs, and losses. Using the scaling laws we analyze grouping of channel outputs and the count rates for this case.

DOI: 10.1103/PhysRevA.94.042339

## I. INTRODUCTION

Much recent attention has been given to the application of multichannel linear photonic networks to solving computational tasks thought to be inaccessible to any classical computer. Such devices are in the vanguard of a new generation of problem-specific quantum computers [1,2] and novel metrology devices [3,4]. As the prototypes of novel forms of quantum computer, these have many potential applications [5].

In particular, boson sampling is the problem of generating photon counts with a probability distribution equal to the modulus squared of the permanent of unitary submatrix. The result is created as the output $n$-photon coincidence rate from a single photon input to each of $n$ distinct channels. Generating a random sequence of this type is conjectured to be exponentially hard at large $n$, while being relatively simple to implement physically. This could become the first example of a quantum solution to a classically inaccessible problem, and new applications are already appearing.

It is widely appreciated that to verify the solution is correct is a significant challenge [6–13]. The task is to measure the coincidence rate of counting one photon in each of $n$ output channels and to confirm that this corresponds to the modulus squared of an $n \times n$ permanent of a unitary submatrix. The matrices are under experimental control [14–17], and a random ensemble of them is tested to verify the experiments. A number of strategies have been suggested [6–8,10–12,18,19].

Since permanents are known to be exponentially hard to compute [20], it is nontrivial to verify the output is correct [2,6] at large $n$ values. A verification strategy requires that both calculating the criterion and making the measurement take a finite time, and the criterion should not be easily forgeable with a classical computer. As a result, understanding scaling is essential because count rates can decline exponentially fast with $n$. Yet, the computational hardness of permanents makes it extremely difficult to estimate count rates for $n > 50$ [21]. These fundamental scaling relations are the central issue that this paper addresses.

Even though individual permanents are effectively uncomputable, the tools of random matrix theory can be used to tell us analytically how the average count rates scale on averaging over all unitary matrices. It is known that large random unitaries are similar to their unitary averages, apart from a small measure of extreme cases. This gives us a quantitative prediction of what scaling to expect at large matrix size, which cannot be obtained through numerics, owing to computational complexity.

Here we solve the average scaling problem, including losses, for arbitrary inputs and outputs. Previous work has considered specific input states, for example Fock states [7,11,14,15] or Gaussian states [22,23]. Importantly, we obtain the maximum scaling of the average improvements that are possible with recent channel grouping strategies [7]. While the average count rate on its own is not a complete verification—it is too nonspecific for this—we show in another work how these limits can be applied to give analytic tests for boson sampling for single unitary transformations [24].

To obtain scaling laws, we combine the characteristic function of the generalized $P$ representation with methods from random matrix theory to obtain averages over unitary transformations. Random matrix theory methods have also been recently used in order to study mesoscopic chaotic scattering [25]. The present approach allow us to describe the scaling of realistic bosonic experiments, with arbitrary inputs, outputs, and losses. Losses in boson sampling have also been investigated elsewhere [26,27].

Scaling issues like this arising in random matrix theory are widespread [28], as averages over unitaries are fundamental to quantum physics. We note an unexpected analogy with the statistics of a well-known classical device for generating random counts. On averaging over all unitaries, the probability of $n$ single-photon counts in $n$ preselected channels—a quantum Galton's board [29]—is identical to a type of classical photonic Galton's board.

The only difference is that there are now $n - 1$ additional virtual channels, which describe multiphoton events in an output mode. These extra channels can be thought of as nonclassical communication channels. Such channel capacity improvements [30] are closely related to Arkhipov and Kuperberg's "birthday paradox" for bosons [31].

We show, in particular, that while typical individual permanents give extremely low count rates, the scaling improvement with channel grouping strategies for verification depends on the channel occupation ratio, $k = m/n$, reaching well over a hundred orders of magnitude at $k = 10$, $n = 100$. This matrix size is beyond any current exact computation of matrix permanents. As a result, our calculations indicate that boson sampling verification with such large $n$ values is not impossible, provided one has fast, high-efficiency detectors.

## II. CHARACTERISTIC FUNCTION FORMALISM

We start with a result [32,33] from quantum optics: any bosonic correlation function is obtainable from the normally ordered quantum characteristic function,

$$\chi(\boldsymbol{\xi}) = \langle : e^{\boldsymbol{\xi}\cdot\hat{\boldsymbol{a}}^\dagger - \boldsymbol{\xi}^*\cdot\hat{\boldsymbol{a}}} : \rangle_Q. \tag{1}$$

Here, $\langle \hat{O} \rangle_Q \equiv \mathrm{Tr}[\hat{\rho}\hat{O}]$ is a quantum average, which we calculate using a generalized $P$ representation [34]. This approach extends the Glauber $P$ function [35], giving a distribution $P(\boldsymbol{\alpha},\boldsymbol{\beta})$ over are two $m$-component complex vectors, which exists for any $m$-mode bosonic state $\hat{\rho}$. The quantum characteristic function is then obtained from [36]

$$\begin{aligned}\chi(\boldsymbol{\xi}) &= \langle \chi(\boldsymbol{\xi}|\boldsymbol{\alpha},\boldsymbol{\beta})\rangle_P \\ &= \int P(\boldsymbol{\alpha},\boldsymbol{\beta})\chi(\boldsymbol{\xi}|\boldsymbol{\alpha},\boldsymbol{\beta})d\mu(\boldsymbol{\alpha},\boldsymbol{\beta}),\end{aligned} \tag{2}$$

where $d\mu(\boldsymbol{\alpha},\boldsymbol{\beta})$ is the integration measure, $\chi(\boldsymbol{\xi}|\boldsymbol{\alpha},\boldsymbol{\beta}) \equiv \exp(\boldsymbol{\xi}\cdot\boldsymbol{\beta} - \boldsymbol{\xi}^*\cdot\boldsymbol{\alpha})$ is the conditional characteristic function for $\boldsymbol{\xi}$ given a particular quantum phase-space trajectory $\boldsymbol{\alpha},\boldsymbol{\beta}$, and $\langle\rangle_P$ is a phase-space average.

Transmission through a linear network changes the input density matrix $\hat{\rho}^{(\mathrm{in})}$ to an output density matrix $\hat{\rho}^{(\mathrm{out})}$. An amplitude transmission matrix $T$ transforms the coherent amplitudes [37], so that $\boldsymbol{\alpha}^{(\mathrm{out})}, \boldsymbol{\beta}^{(\mathrm{out})} = T\boldsymbol{\alpha}, T^*\boldsymbol{\beta}$. The output characteristic $\chi^{(\mathrm{out})}$ now depends on the input phase-space amplitude $\boldsymbol{\alpha},\boldsymbol{\beta}$ as:

$$\chi^{(\mathrm{out})}(\boldsymbol{\xi}|\boldsymbol{\alpha},\boldsymbol{\beta}) = e^{\boldsymbol{\xi}\cdot\boldsymbol{T}^*\boldsymbol{\beta} - \boldsymbol{\xi}^*\cdot\boldsymbol{T}\boldsymbol{\alpha}}. \tag{3}$$

To calculate the average scaling behavior, we consider the case of $T = \sqrt{t}U$, in which the unitary mode transformation $U$ of the photonic network is combined with an absorptive transmission coefficient $t$, representing losses and detector inefficiencies.

### Unitary average of the characteristic function

We compute the average output correlations over all possible unitaries, indicated by $\langle\rangle_U$, from random matrix theory [38]. This allows one to evaluate averages over the unitary matrices in the conditional characteristic function, $\chi^{(\mathrm{out})}$ of Eq. (3). This result is obtained from considering the identity (5.6) of Ref. [38]. For clarity we write the Fyodorov identity using the notational conventions of the present paper, where $\boldsymbol{A}^*$ is a complex conjugate and $\boldsymbol{A}^\dagger \equiv \boldsymbol{A}^{T*}$ is a Hermitian (transposed) conjugate:

$$\begin{aligned}&\langle e^{-\mathrm{Tr}[U(\mathbf{z}_A\otimes\boldsymbol{\xi}_A^\dagger)+U^\dagger(\mathbf{z}_B\otimes\boldsymbol{\xi}_B^\dagger)]}\rangle_U \\ &\quad= (m-1)!\sum_{j=0}^{\infty}\frac{[(\boldsymbol{\xi}_A^*\cdot\mathbf{z}_B)(\boldsymbol{\xi}_B^*\cdot\mathbf{z}_A)]^j}{j!(m-1+j)!},\end{aligned} \tag{4}$$

The above matrix integral identity allows us to evaluate the average of the exponentials of the unitary matrices in the conditional characteristic function, which only involves complex matrices and vectors.

Making the substitutions of: $\mathbf{z}_A \equiv -\sqrt{t}\boldsymbol{\alpha}$, $\boldsymbol{\xi}_A \equiv \boldsymbol{\xi}$, we find that

$$\mathrm{Tr}[U(\mathbf{z}_A\otimes\boldsymbol{\xi}_A^\dagger)] = \boldsymbol{\xi}_A^*\cdot U\mathbf{z}_A = -\boldsymbol{\xi}^*\cdot\boldsymbol{T}\boldsymbol{\alpha}. \tag{5}$$

Similarly, with $\boldsymbol{\xi}_B^* \equiv \sqrt{t}\boldsymbol{\beta}$, $\mathbf{z}_B \equiv \boldsymbol{\xi}$, we have

$$\mathrm{Tr}[U^\dagger(\mathbf{z}_B\otimes\boldsymbol{\xi}_B^\dagger)] = \mathbf{z}_B\cdot U^*\boldsymbol{\xi}_B^* = \boldsymbol{\xi}\cdot\boldsymbol{T}^*\boldsymbol{\beta}. \tag{6}$$

Hence, we obtain the following conditional characteristic function averaged over the unitary matrices:

$$\begin{aligned}\langle\chi^{(\mathrm{out})}(\boldsymbol{\xi}|\boldsymbol{\alpha},\boldsymbol{\beta})\rangle_U &= \langle e^{\boldsymbol{\xi}\cdot\boldsymbol{T}^*\boldsymbol{\beta}-\boldsymbol{\xi}^*\cdot\boldsymbol{T}\boldsymbol{\alpha}}\rangle_U \\ &= (m-1)!\sum_{j=0}^{\infty}\frac{[-t|\boldsymbol{\xi}|^2\boldsymbol{\beta}\cdot\boldsymbol{\alpha}]^j}{j!(m-1+j)!}.\end{aligned} \tag{7}$$

The output photon statistics depend on phase-space averages of the inner product $\boldsymbol{\beta}\cdot\boldsymbol{\alpha}$, which is the phase-space equivalent of the total input photon number $\hat{N}$. In the complex-$P$ representation moments in phase-space correspond to expectations of normally ordered operators. Therefore, on phase-space averaging one obtains

$$\langle(\boldsymbol{\beta}\cdot\boldsymbol{\alpha})^j\rangle_P = \langle : \hat{N}^j : \rangle_Q^{(\mathrm{in})}, \tag{8}$$

As a result, the characteristic function after unitary and quantum averaging is

$$\langle\chi^{(\mathrm{out})}(\boldsymbol{\xi})\rangle_U = (m-1)!\sum_{j=0}^{\infty}\frac{(-t|\boldsymbol{\xi}|^2)^j\langle : \hat{N}^j : \rangle_Q^{(\mathrm{in})}}{j!(m-1+j)!}. \tag{9}$$

Simplifying this expression gives

$$\langle\chi^{(\mathrm{out})}(\boldsymbol{\xi})\rangle_U = \sum_{j=0}^{\infty} f_j|\boldsymbol{\xi}|^{2j}, \tag{10}$$

where we have defined

$$f_j = \frac{(m-1)!\langle : (-t\hat{N})^j : \rangle_Q^{(\mathrm{in})}}{j!(m-1+j)!}. \tag{11}$$

The effect of unitary averaging is such that all output averages are obtained solely from the total input photon number moments, $\langle : \hat{N}^j : \rangle_Q^{(\mathrm{in})}$, regardless of which input channels are used.

## III. UNITARY AVERAGE OF PHOTON-NUMBER COUNTS

Photon-number correlations, as commonly measured experimentally, are obtained from taking derivatives of the photon-number generating function [39],

$$G(\boldsymbol{\gamma}) \equiv Tr\left(\hat{\rho}\prod_i (1-\gamma_i)^{\hat{n}_i}\right). \tag{12}$$

We wish to obtain the expression for $P_{n|m}$, which is the unitary average probability of observing one photon in each of $n$ specified output channels, given an $n$-photon input and an $m$-mode network. This probability is obtained using photon-number correlations, which can be calculated from taking derivatives of the photon-number generating function $G(\boldsymbol{\gamma})$ in the output modes. We first obtain the general expression for any output photon-number correlation, using the result of Sec. II. Next, we will compare this with unitary averages of permanents, which is the well-known result for photon correlations in a lossless, unitary photonic network.

### A. Average photon-number-generating function

The general relationship between photon-number generator and characteristic function is given by [40]:

$$G(\boldsymbol{\gamma}) = \int\int \prod_i \left[\frac{1}{\pi\gamma_i}\exp\left(\frac{-|\xi_i|^2}{\gamma_i}\right)\right]\chi(\boldsymbol{\xi})d^{2m}\boldsymbol{\xi}. \tag{13}$$

Next, substituting in the above expression with Eq. (9) we get for the $m$-mode case:

$$\langle G(\boldsymbol{\gamma})\rangle_U = \int \langle\chi^{(\mathrm{out})}(\boldsymbol{\xi})\rangle_U \exp\left(\sum_{i=1}^{m}\frac{-|\xi_i|^2}{\gamma_i}\right)\prod_{i=1}^{m}\frac{d^2\xi_i}{\pi\gamma_i}. \tag{14}$$

Changing variables to polar coordinates with $R_i = |\xi_i|^2/\gamma_i$ gives us:

$$\langle G(\boldsymbol{\gamma})\rangle_U = \int \langle\chi^{(\mathrm{out})}(\boldsymbol{\xi})\rangle_U \exp\left(-\sum_{i=1}^{m}R_i\right)\prod_{i=1}^{m}dR_i. \tag{15}$$

We now expand the unitary average expression obtained above for $\chi(\boldsymbol{\xi})$ as a function of $R_i$, with the substitution:

$$\langle\chi^{(\mathrm{out})}(\boldsymbol{\xi})\rangle_U = \sum_{j=0}^{\infty} f_j|\boldsymbol{\xi}|^{2j} = \sum_{j=0}^{\infty} f_j\left[\sum_{i=1}^{m}\gamma_i R_i\right]^j, \tag{16}$$

and hence obtain the unitary average:

$$\langle G(\boldsymbol{\gamma})\rangle_U = \sum_{j=0}^{\infty} f_j \int\left[\sum_{i=1}^{m}\gamma_i R_i\right]^j \exp\left(-\sum_{i=1}^{m}R_i\right)\prod_{i=1}^{m}dR_i. \tag{17}$$

This is integrated using the multinomial expansion of the term in brackets,

$$\left[\sum_{i=1}^{m}\gamma_i R_i\right]^j = j!\sum_{\mathbf{p},\ \sum p_k=j}\frac{\gamma_1^{p_1}\cdots\gamma_m^{p_m}}{(p_1!)\cdots(p_m!)}, \tag{18}$$

followed by a product of integrals over each radial coordinate $R_j$, so that:

$$\langle G(\boldsymbol{\gamma})\rangle_U = \sum_{j=0}^{\infty} f_j j! \sum_{\mathbf{p},\ \sum p_k=j}\gamma_1^{p_1}\ldots\gamma_m^{p_m}. \tag{19}$$

Therefore we finally obtain:

$$\langle G(\boldsymbol{\gamma})\rangle_U = (m-1)!\sum_{j=0}^{\infty}\frac{(-t^2)^j\langle:\hat{N}^j:\rangle}{(m-1+j)!} \times \sum_{\mathbf{p},\ \sum p_k=j}\gamma_1^{p_1}\ldots\gamma_m^{p_m}. \tag{20}$$

Next we consider a specific case: an $n$-photon input number state with

$$\langle:\hat{N}^j:\rangle_Q^{(\mathrm{in})} = n!/(n-j)!, \tag{21}$$

so the sum in Eq. (9) vanishes for $j > n$. This is a consequence of photon-number conservation and the purely absorptive loss reservoirs, which we assume here for simplicity. In this case we get the result:

$$\langle G(\boldsymbol{\gamma})\rangle_U = (m-1)!\sum_{j=0}^{n}\frac{(-t^2)^j n!}{(n-j)!(m-1+n)!} \times \sum_{\mathbf{p},\ \sum p_k=j}\gamma_1^{p_1}\ldots\gamma_n^{p_n}. \tag{22}$$

This is a general expression for unitary averages given an $n$-photon input number state. This central result is completely general, describing any unitarily averaged photon correlation measured in lossy photonic networks with arbitrary inputs and outputs.

In the case of most interest for boson sampling, we obtain the average probability of observing one photon in each of a set of $n$ specified output modes with an $n$-photon input and an $m$-mode network, $P_{n|m}$. This is found on taking $n$ first derivatives of $\langle G(\boldsymbol{\gamma})\rangle_U$, so that:

$$P_{n|m} = \frac{t^n(m-1)!n!}{(m-1+n)!} = t^n\left[C_n^{m+n-1}\right]^{-1}. \tag{23}$$

### B. Averages using the permanent

We can also calculate these results following the route of calculating unitary averages over permanents. The permanent is a sum over all permutations $\sigma$ of the matrix indices of the product of $n$ terms, in which neither row nor column indices are repeated. It is an exponentially hard object to compute, and is one of the fundamental quantities addressed in boson sampling theory and in linear optical networks [41,42]. We note that the results in this section are somewhat less general, as they are restricted to number state inputs and outputs. For a pure, unitary state evolution, the photon counting probability is known to be the averaged permanent of a submatrix, $\langle|\mathrm{perm}(U_{n|m})|^2\rangle$, where $U_{n|m}$ is any $n\times n$ submatrix of $U$ [43].

More generally, we can see from the above results that one can simply replace $U_{n|m}\to T_{n|m}$, so as to include losses.

We now show that Eq. (23) obtained above can also be obtained from permanents of $T$. The permanent of a submatrix of $T$ is obtainable [44] from the permanental polynomial, which has similarities with moment-generating functions. This is given by:

$$p(x) = \text{perm}(xI - T) \equiv \sum_{n=0}^{m} b_n x^{m-n}. \tag{24}$$

Next, we consider how to compute the unitary average of products of the permanents. This is achieved through another elegant result from random matrix theory [38]. The unitary average of permanental polynomials in Eq. (24) is

$$\langle p(x)p(y)^*\rangle_U = m!(m—1)!\sum_{j=0}^{m}\frac{t^j(xy^*)^{m-j}}{(m-j)!(m-1+j)!}. \tag{25}$$

If $\omega^{(n)} = (\omega_1, \ldots \omega_n)$ where $\omega_1 < \omega_2 \ldots < \omega_n$, we can define an $n \times n$ submatrix $T_{\boldsymbol{\omega}} \equiv T_{\omega_i\omega_j}$. The coefficients of this polynomial are simply the sums over the permanents of all possible distinct submatrices:

$$b_n = (-1)^n \sum_{\boldsymbol{\omega}^{(n)}} \text{perm}(T_{\boldsymbol{\omega}^{(n)}}). \tag{26}$$

The number of submatrices in the sum has a multiplicity given by the binomial coefficient $C_n^m = m!/[n!(m-n)!]$, corresponding to the different ways to choose the distinct indices $\omega_j$. These indices have a straightforward physical interpretation: they are the channel numbers of the input or output modes of the photonic device.

The unitary averaged sum over all possible subpermanents of this size is a ratio of two binomial coefficients, which we define as $R_{n|m}$:

$$\sum_{\boldsymbol{\omega}^{(n)}}\langle|\text{perm}(T_{\boldsymbol{\omega}^{(n)}})|^2\rangle_U = R_{n|m} = \frac{t^n m!(m—1)!}{(m-n)!(m+n-1)!}. \tag{27}$$

Taking account of the number of different permanents, the final unitary averaged result is the same as in Eq. (23), except so that $P_{n|m} = \langle|\text{perm}(T_{n|m})|^2\rangle_U$. In the lossless limit of $t = 1$, this result agrees with the bosonic birthday paradox of Arkhipov and Kuperberg [31], derived using different techniques. We note that these unitary averaged scaling results for the permanents are analytic. This is important since evaluating permanents of individual large matrices is an exponentially hard problem; and even if it were feasible, there are exponentially many submatrices to check for each unitary.

## IV. SCALING RESULTS

We turn next to some limiting cases for large $n$, where $\ln P_{n|m} \approx n\epsilon$ for a scaling exponent $\epsilon$. Details of the following results are given in Appendix A. In all cases the limit is taken at a fixed ratio of $k = m/n$, so both $m$ and $n$ are taken as large integers.

### A. Entire matrix

If the matrix is the entire transmission matrix, then $n = m$. The scaling exponent is $\epsilon = \ln(t/4)$, and

$$\ln P_{n|m} \underset{n\to\infty}{\sim} n\epsilon + \tfrac{1}{2}\ln[4\pi n]. \tag{28}$$

This result generalizes one of Fyodorov [38].

### B. Gaussian limit

Next, take $n \ll m$, so that $k \gg 1$. Standard methods for approximating a binomial coefficient in this limit give the scaling exponent $\epsilon = \ln[t/(k+1/2)] - 1$, where $k = m/n$, so that:

$$\ln P_{n|m} \underset{n\to\infty}{\sim} n\epsilon + \tfrac{1}{2}\ln[2\pi n]. \tag{29}$$

This is consistent with the fact that for large $k$, unitary submatrices reduce to matrices with complex Gaussian random entries [2,45].

### C. General submatrix

For the general case, the scaling exponent is $\epsilon = \ln t + k\ln k - (1+k)\ln(1+k)$, and the asymptotic result is

$$\ln P_{n|m} \underset{n\to\infty}{\sim} n\epsilon + \tfrac{1}{2}\ln[2\pi n(1+1/k)]. \tag{30}$$

The exact result is plotted in Fig. 1, with different values of $k = m/n$. The power law is so close that it cannot be told apart from the exact result on this scale. For all numerical results, we choose $t = 1$, since results in more realistic cases with losses are readily obtained by adding $\ln t$ to the scaling exponents.

At large $k$ values, one obtains the Gaussian limit of Eq. (29). This gives an increasingly negative exponent, with exponentially small count rates.

### D. Numerical averages

To confirm and illustrate these scaling laws, in Fig. 2 we show the result of a numerical average over a finite, random ensemble of unitary matrices, with $S = 40000$ unitary samples. For $k = m/n = 2$ and $n \leqslant 25$ we plot relative errors in the asymptotic approximation and a numerical average, compared to the exact solution. To estimate statistical error bars, we take an ensemble $S$, and divide it into $\sqrt{S}$ subensembles, giving subensemble means that are approximately Gaussian from the central limit theorem.

These are averaged, and the error in the mean $\sigma_m$ is obtained using standard techniques. The error bars are given in the plots as $\pm\sigma_m$. The results agree with the exact equation with

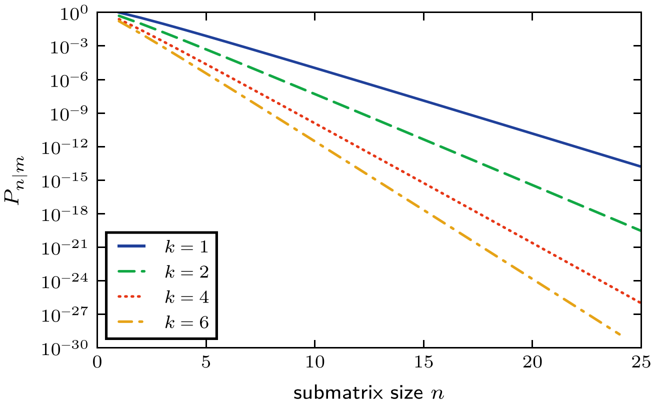

FIG. 1. Average subunitary permanent squared $P_{n|m}$ with $t = 1$ for $k = 1,2,4,6$, with $k = 1$ at the top and $k = 6$ at the bottom.

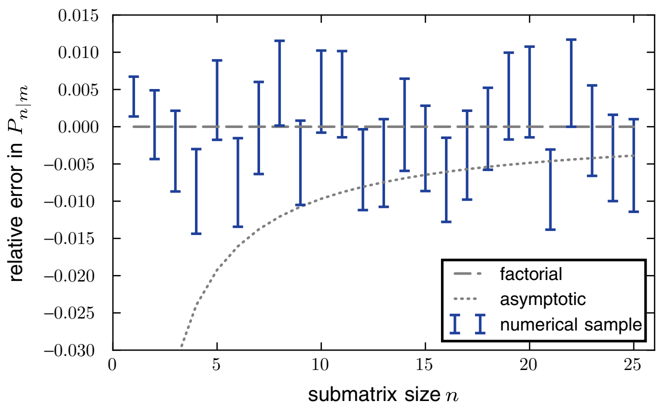

FIG. 2. Relative sampling errors in the subunitary permanent squared $P_{n|m}$ with $t = 1$ for $k = 2$. Numerical results for 40000 random unitaries (solid lines with error bars) are compared to exact results (dashed lines) and the asymptotic power law form (dotted line) for up to $n = 25$.

a relative error comparable to the sampling error bars. For the plotted ratio of $k = 2$, the errors are around ±1% for over 20 orders of magnitude range of values, and we see that the relative sampling error over unitaries is independent of matrix size for $n \geqslant 10$.

While the scaling treated here—up to $k = 6$—is better than in the Gaussian limit of $k \gg 1$, it is a problem for boson sampling verification. Even in the case of perfect efficiency, the average permanent for a photon number of $n = 25$ is $\sim 10^{-30}$ with a submatrix ratio of $k = 6$. This is the probability of a coincidence count. Hence, one needs $10^{30}$ samples to obtain one count for a typical unitary. At a repetition rate of $10^9$ Hz, one would require $10^{21}$ s of measurement time for each count: this is still not yet in the domain where classical computers fail.

The cause here is many-body complexity. There are too many quantum states possible. Monitoring the coincidence channels for just one many-body state takes too long, even though it is these counts that are of interest. This property, although making verification hard, is also the most interesting feature of these experiments. They give a uniquely controllable access to a laboratory system in which one can unravel the exponential complexity of a many-body system.

## V. CHANNEL GROUPING

We now consider what happens when one groups multiple output channels together, by using logic gate operations on the detector circuits, as in a recent, pioneering experiment [7]. A large number of randomized sets of channels can be combined, to obtain a unique distinguishing signature for each unitary. We analyze particular verification tests using channel combinations in another work [24]. Here we focus on the maximum scaling improvement possible.

This has important advantages over many previous proposals. It increases count rates by exponentially large factors, and may allow a test of the unitary output bitstream for permanents larger than $n = 50$, beyond the present classical limits with the fastest known supercomputers [21]. Yet, it is not restricted to any particular unitary, reducing the chance that the test may only work in special cases. One can also use the strategy in edge cases such as Fourier matrices [12].

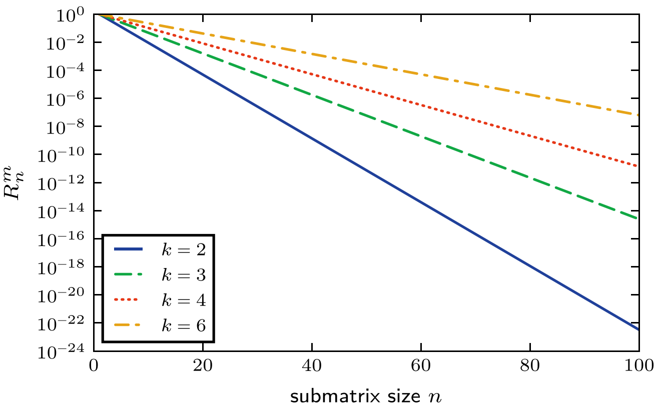

FIG. 3. Upper bound on count rates $R_n^m$ for $n$ photons occurring, in any $n$ output channels, for $k = 2, \dots 6$, with $k = 6$ at the top and $k = 2$ at the bottom.

### A. Grouped scaling laws

Our scaling laws predict the upper bound of the count-rate gain that can be achieved through submatrix multiplicity. The upper bound from channel grouping is given by $R_{n|m}$, in Eq. (27). This has an exponent of $\lambda = \ln t + 2k \ln k - (k - 1)\ln(k - 1) - (k + 1)\ln(k + 1)$, so that the scaling is (see Appendix B for details):

$$\ln R_{n|m} \underset{n\to\infty}{\sim} n\lambda + \frac{1}{2}\ln\left[\frac{k+1}{k-1}\right]. \tag{31}$$

For the Gaussian limit of $n,k \gg 1$, one finds that $\lambda \to \ln t - 1/k$. Unlike the single coincidence case, the grouped channel count rate is maximized for large $k$, rather than minimized as before. The corresponding upper-bound result is plotted in Fig. 3, for different values of $k = m/n$, again taking $t = 1$ for simplicity. The improvement is greatest for large $k$ values, which are of most interest.

### B. Physical estimates

What kind of count rates can be obtained if we extrapolate current technologies? Recently developed superconducting nanowire single-photon detectors have an upper efficiency limit of 90%, and a bandwidth limit of $10^9\,\mathrm{s}^{-1}$ [46]. These parameters give a best case scaling exponent of $\lambda = -0.2$ at $k = 10$, and an upper count rate limit of around $1s^{-1}$ at $n = 100$. This count rate would be enough to gather good statistics in a reasonable time, for a device much too large to classically emulate.

Thus, while boson sampling experiments still require much progress in state preparation and optical fabrication, we can see that verification of the quantum properties of these devices at large $n$, using channel grouping, does appear achievable as the technology improves.

At a more fundamental level, it is an intriguing result, mathematically and physically, that the quantum Galton's board has a close relationship with the binomial coefficients normally found in classical combinatorics. How can we interpret this result?

Suppose that we replaced the photonic network, equivalent to a unitary transformation, by an updated Galton's board device, which simply switched the photons from the $n$ input channels to the $n$ output channels in a random way. This would not involve interference, and would have a similar behavior to a mechanical board, apart from an increased number of inputs.

Under these conditions, the average probability of all counts occurring in a preselected set of $n$ output channels is an inverse binomial $[C_n^m]^{-1}$. We now see a truly remarkable result. Apart from losses, the quantum Galton's board has, on averaging over all unitaries, identical output coincidence probabilities to a classical Galton's board with a number of channels given by $\tilde{m} = m + n - 1$.

In other words, the fact that photons can bunch—one channel may carry up to $n$ photons—has a similar effect on the output statistics as if the device were classical, but with $n - 1$ additional channels available. These virtual channels represent, on average, the additional output possibilities available owing to the fact that several bosons can occupy the same mode, and hence occur in the same channel. This is the large-scale consequence of the Hong-Ou-Mandel effect in quantum optics [47,48].

It is this additional, virtual channel capacity that allows more quantum information to be transmitted in a quantum photonic network than is feasible if each channel was used separately, with one bit per channel. Such extra capacity is a fundamental and important property of quantum photonic networks [30].

## VI. CONCLUSIONS

In summary, the unitary average of count-rates in photonic networks has some interesting properties. Each computation of a coincidence rate requires knowledge of an exponentially complex permanent with no simple closed-form expression. One might imagine that taking an average over all possible unitaries would only make things harder. Yet the average over the unitary ensemble is just as simple as the closed-form expression applicable to a type of classical Galton's board.

This of course is only true on average. While guiding us to an understanding of scaling behavior, the knowledge of unitary averages does not change the underlying exponential hardness of boson sampling. We use these scaling laws to show that channel grouping strategies can lead to count rates that are within the range of current photon-counting technology, even for photon numbers as large as $n = 100$, far beyond current computational limits for exact permanents. We show in another work [24] how these results can be extended to a full verification of boson sampling. Our results are applicable to all linear quantum photonic networks, for any input state.

## ACKNOWLEDGMENT

This work was funded by Australian Research Council Grants No. DP140102613 and No. DP1401045484.

## APPENDIX A: SCALING RESULTS

In this section we provide the derivation of Eqs. (28)–(30) of the main text, which give the scaling exponents. Here we consider limiting cases for large $n$, of the expression for $P_{n|m}$, where we recall that

$$P_{n|m} = \frac{t^n(m-1)!n!}{(m-1+n)!} = t^n\left[C_n^{m+n-1}\right]^{-1}. \tag{A1}$$

The limit is taken at a fixed ratio of $k = m/n$, so both $m$ and $n$ are large integers.

### 1. Entire matrix

This case corresponds to $n = m$. Hence, from Eq. (A1) and using Stirling's approximation $n! \approx \sqrt{2\pi n}(n/e)^n$ as well as the simple manipulations $(n-1)! = n!/n$ and $(2n-1)! = (2n)!/2n$ we obtain:

$$P_{n|m} = \frac{t^n(n-1)!n!}{(2n-1)!} = \frac{2t^n(n!)^2}{(2n)!} \approx \frac{2t^n\sqrt{\pi n}}{4^n}. \tag{A2}$$

On taking natural logarithms and defining $\epsilon = \ln(t/4)$:

$$\ln P_{n|m} \underset{n\to\infty}{\sim} n\epsilon + \tfrac{1}{2}\ln[4\pi n]. \tag{A3}$$

This is Eq. (28) of the main text. In the lossless case this result generalizes Eq. (5.14) of Ref. [38], by giving the logarithmic correction to the leading exponent.

### 2. Gaussian limit

This case corresponds to $n \ll m, k \gg 1$, and $m = kn$. Here we use the following identity for approximating the binomial coefficient in the limit of $p \gg n$:

$$\binom{p}{n} \approx \frac{(p/n - 0.5)^n e^n}{\sqrt{2\pi n}}. \tag{A4}$$

On simplifying and taking natural logarithms we get Eq. (29) of the main text:

$$\ln P_{n|m} \underset{n\to\infty}{\sim} n\left(\ln\left[\frac{t}{k+0.5}\right] - 1\right) + \frac{1}{2}\ln[2\pi n]. \tag{A5}$$

### 3. General submatrix

This case corresponds to $m = kn$ and $n \gg 1$, so that we get:

$$\begin{aligned} P_{n|m} &= \frac{t^n(m-1)!n!}{(m-1+n)!} \\ &\underset{n\to\infty}{\sim} t^n\sqrt{\frac{2\pi n(m-1)}{m-1+n}}\,\frac{n^n(m-1)^{m-1}}{(m+n-1)^{m+n-1}}. \end{aligned} \tag{A6}$$

On taking natural logarithms we get:

$$\begin{aligned} \ln P_{n|m} \underset{n\to\infty}{\sim}\; & n\ln t + \frac{1}{2}\ln\left[\frac{2\pi n(m-1)}{m-1+n}\right] \\ & + n\ln n + (m-1)\ln[m-1] \\ & - (m+n-1)\ln[m+n-1]. \end{aligned} \tag{A7}$$

Next we use the following identity $\ln[x \pm 1] = \ln x + \ln(1 \pm 1/x)$. On dropping terms of the form $1/n$ and $1/m$ we get:

$$\begin{aligned} \ln P_{n|m} \underset{n\to\infty}{\sim}\; & n\ln t + \tfrac{1}{2}\ln[2\pi nm(m+n)] + n\ln n \\ & + (m-1)\ln[m] - (m+n)\ln[m+n]. \end{aligned} \tag{A8}$$

Simplifying this, and using that $m = kn$:

$$\begin{aligned}\ln P_{n|m} \underset{n\to\infty}{\sim} &= n[\ln t + k\ln k - (1+k)\ln(1+k)] \\ &+ \tfrac{1}{2}\ln[2\pi n(1+1/k)]. \qquad \text{(A9)}\end{aligned}$$

This is Eq. (30) of the main text.

## APPENDIX B: CHANNEL GROUPING SCALING

In this section we provide the derivation of Eq. (31) of the main text, in the limiting case for large $n$, of $R_{n|m}$. For reference, this expression is given below:

$$R_{n|m} = \frac{t^n m!(m\text{—}1)!}{(m-n)!(m+n-1)!}. \qquad \text{(B1)}$$

We wish to obtain an expression in the limit of $n \gg 1$ and $m = kn$. We start by using Stirling's approximation $n! \approx \sqrt{2\pi n}(n/e)^n$, which gives:

$$\begin{aligned}R_{n|m} &= \frac{t^n m!(m\text{—}1)!}{(m-n)!(m+n-1)!} \\ &\underset{n\to\infty}{\sim} t^n\sqrt{\frac{m(m\text{—}1)}{(m-n)(m+n-1)}} \\ &\times \frac{m^m(m\text{—}1)^{(m-1)}}{(m-n)^{(m\text{—}n)}(m+n-1)^{(m+n-1)}}. \qquad \text{(B2)}\end{aligned}$$

On taking natural logarithms we obtain:

$$\begin{aligned}\ln R_{n|m} \underset{n\to\infty}{\sim} & n\ln t + \frac{1}{2}\ln\left[\frac{m(m\text{—}1)}{(m-n)(m+n-1)}\right] \\ &+ m\ln m + (m-1)\ln[m-1] \\ &-(m-n)\ln[m-n] \\ &- (m+n-1)\ln[m+n-1]. \qquad \text{(B3)}\end{aligned}$$

Using that $\ln[x \pm 1] = \ln x + \ln(1 \pm 1/x)$ and simplifying terms, gives:

$$\begin{aligned}\ln R_{n|m} \underset{n\to\infty}{\sim} & n\ln t + \frac{1}{2}\ln\left[\frac{(m\text{—}1)(m+n-1)}{(m-n)m}\right] \\ &+ 2m\ln m + (m-1)\ln\left[1 - \frac{1}{m}\right] \\ &- (m-n)\ln[m-n] - (m+n) \\ &\times \left(\ln[m+n] + \ln\left[1 - \frac{1}{m+n}\right]\right). \qquad \text{(B4)}\end{aligned}$$

On dropping terms of order $1/m$ and $1/(m+n)$ and using that $m = kn$ we get:

$$\begin{aligned}\ln R_{n|m} \underset{n\to\infty}{\sim} & n\ln t + \frac{1}{2}\ln\left[\frac{(m\text{—}1)(m+n-1)}{(m-n)m}\right] \\ &+ 2m\ln m - (m-n)\ln[m-n] \\ &- (m+n)\ln[m+n] \\ = & \frac{1}{2}\ln\left[\frac{(kn\text{—}1)(kn+n-1)}{(kn-n)kn}\right] \\ &+ n\ln t + n[2k\ln k - (k-1)\ln[k-1] \\ &- (k+1)\ln[k+1]]. \qquad \text{(B5)}\end{aligned}$$

The term $\frac{1}{2}\ln\left[\frac{(kn\text{—}1)(kn+n-1)}{(kn-n)kn}\right]$ can be simplified by using the identity $\ln[x \pm 1] = \ln x + \ln(1 \pm 1/x)$ and dropping terms of the form $1/kn$ and $1/(k+1)n$; hence, we now obtain

$$\frac{1}{2}\ln\left[\frac{(kn\text{—}1)(kn+n-1)}{(kn-n)kn}\right] \underset{n\to\infty}{\sim} \frac{1}{2}\ln\left[\frac{k+1}{k-1}\right]. \qquad \text{(B6)}$$

Substituting the above expression in Eq. (B5) gives the final result:

$$\begin{aligned}\ln R_{n|m} \underset{n\to\infty}{\sim} & n\ln t + \frac{1}{2}\ln\left[\frac{k+1}{k-1}\right] + n[2k\ln k \\ &- (k-1)\ln[k-1] - (k+1)\ln[k+1]]. \\ & \qquad \text{(B7)}\end{aligned}$$

This corresponds to Eq. (31) of the main text.

[1] S. Aaronson and A. Arkhipov, in *Proceedings of the 43rd Annual ACM Symposium on Theory of Computing* (ACM Press, New York, 2011), pp. 333–342.

[2] S. Aaronson and A. Arkhipov, Theor. Comput. **9**, 143 (2013).

[3] J. L. O'Brien, A. Furusawa, and J. Vuckovic, Nat. Photon. **3**, 687 (2009).

[4] K. R. Motes, J. P. Olson, E. J. Rabeaux, J. P. Dowling, S. J. Olson, and P. P. Rohde, Phys. Rev. Lett. **114**, 170802 (2015).

[5] C. H. Papadimitriou, in *Encyclopedia of Computer Science* (Wiley, Chichester, 2003), pp. 260–265.

[6] S. Aaronson and A. Arkhipov, Quantum Inf. Comput. **14**, 1383 (2014).

[7] J. Carolan *et al.*, Nat. Photon. **8**, 621 (2014).

[8] N. Spagnolo *et al.*, Nat. Photon. **8**, 615 (2014).

[9] C. Gogolin, M. Kliesch, L. Aolita, and J. Eisert, arXiv:1306.3995.

[10] M. Bentivegna *et al.*, Sci. Adv. **1**, e1400255 (2015).

[11] L. Aolita, C. Gogolin, M. Kliesch, and J. Eisert, Nat. Commun. **6**, 8498 (2015).

[12] M. C. Tichy, K. Mayer, A. Buchleitner, and K. Mølmer, Phys. Rev. Lett. **113**, 020502 (2014).

[13] M. Bentivegna, N. Spagnolo, and F. Sciarrino, New J. Phys. **18**, 041001 (2016).

[14] M. A. Broome, A. Fedrizzi, S. Rahimi-Keshari, J. Dove, S. Aaronson, T. C. Ralph, and A. G. White, Science **339**, 794 (2013).

[15] A. Crespi, R. Osellame, R. Ramponi, D. J. Brod, E. F. Galvao, N. Spagnolo, C. Vitelli, E. Maiorino, P. Mataloni, and F. Sciarrino, Nat. Photon. **7**, 545 (2013).

[16] M. Tillmann, B. Dakic, R. Heilmann, S. Nolte, A. Szameit, and P. Walther, Nat. Photon. **7**, 540 (2013).

[17] J. B. Spring *et al.*, Science **339**, 798 (2013).

[18] M. Walschaers, J. Kuipers, J.-D. Urbina, K. Mayer, M. C. Tichy, K. Richter, and A. Buchleitner, New J. Phys. **18**, 032001 (2016).

[19] M. Walschaers, J. Kuipers, and A. Buchleitner, Phys. Rev. A **94**, 020104 (2016).

[20] L. Valiant, Theor. Comput. Sci. **8**, 189 (1979).

[21] J. Wu, Y. Liu, B. Zhang, X. Jin, Y. Wang, H. Wang, and X. Yang, arXiv:1606.05836.

[22] A. P. Lund, A. Laing, S. Rahimi-Keshari, T. Rudolph, J. L. O'Brien, and T. C. Ralph, Phys. Rev. Lett. **113**, 100502 (2014).

[23] S. Rahimi-Keshari, T. C. Ralph, and C. M. Caves, Phys. Rev. X **6**, 021039 (2016).

[24] B. Opanchuk, L. E. C. Rosales-Zárate, M. D. Reid, and P. D. Drummond, arXiv:1609.05614.

[25] J.-D. Urbina, J. Kuipers, S. Matsumoto, Q. Hummel, and K. Richter, Phys. Rev. Lett. **116**, 100401 (2016).

[26] K. R. Motes, J. P. Dowling, A. Gilchrist, and P. P. Rohde, Phys. Rev. A **92**, 052319 (2015).

[27] S. Aaronson and D. J. Brod, Phys. Rev. A **93**, 012335 (2016).

[28] P. J. Forrester, *Log-gases and random matrices (LMS-34)* (Princeton University Press, Princeton, 2010).

[29] B. T. Gard, J. P. Olson, R. M. Cross, M. B. Kim, H. Lee, and J. P. Dowling, Phys. Rev. A **89**, 022328 (2014).

[30] C. M. Caves and P. D. Drummond, Rev. Mod. Phys. **66**, 481 (1994).

[31] A. Arkhipov and G. Kuperberg, Geom. Topol. Monogr. **18**, 1 (2012).

[32] R. J. Glauber, Phys. Rev. **130**, 2529 (1963).

[33] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, Cambridge, 1995).

[34] P. D. Drummond and C. W. Gardiner, J. Phys. A **13**, 2353 (1980).

[35] R. J. Glauber, Phys. Rev. **131**, 2766 (1963).

[36] C. W. Gardiner and P. Zoller, *Quantum Noise*, 2nd ed. (Springer, Berlin, 2000).

[37] P. D. Drummond and M. Hillery, *The Quantum Theory of Nonlinear Optics* (Cambridge University Press, Cambridge, 2014).

[38] Y. V. Fyodorov, Int. Math. Res. Notices (2006) 61570.

[39] C. D. Cantrell, J. Math. Phys. **12**, 1005 (1971).

[40] E. B. Rockower, Phys. Rev. A **37**, 4309 (1988).

[41] S. Scheel, in *Quantum Information Processing*, edited by T. Beth and G. Leuchs (Wiley-VCH, Weinheim, 2005), Chap. 28, pp. 382–392.

[42] S. Scheel, arXiv:quant-ph/0406127.

[43] S. Aaronson, Proc. R. Soc. A **467**, 3393 (2011).

[44] W. Li and H. Zhang, Bull. Malaysian Math. Sci. Soc. **38**, 1361 (2015).

[45] T. Jiang, Probab. Theory Relat. Fields **144**, 221 (2009).

[46] H. Takesue, S. D. Dyer, M. J. Stevens, V. Verma, R. P. Mirin, and S. W. Nam, Optica **2**, 832 (2015).

[47] C. K. Hong, Z. Y. Ou, and L. Mandel, Phys. Rev. Lett. **59**, 2044 (1987).

[48] S. P. Walborn, A. N. de Oliveira, S. Pádua, and C. H. Monken, Phys. Rev. Lett. **90**, 143601 (2003).