# Simulating and assessing boson sampling experiments with phase-space representations

Bogdan Opanchuk, Laura Rosales-Zárate, Margaret D. Reid, and Peter D. Drummond

Phys. Rev. A **97**, 042304 — Published 4 April 2018

DOI: 10.1103/PhysRevA.97.042304

# Simulating and assessing boson sampling experiments with phase-space representations

Bogdan Opanchuk[1], Laura Rosales-Zárate[1,2], Margaret D Reid[1] and Peter D Drummond[1,3]
[1] *Centre for Quantum and Optical Science, Swinburne University of Technology, Melbourne 3122, Australia*
[2] *Centro de Investigaciones en Óptica A.C., León, Guanajuato 37150, México and*
[3] *Kavli Institute for Theoretical Physics, UC Santa Barbara, USA*
(Dated:)

The search for new, application-specific quantum computers designed to outperform any classical computer is driven by the ending of Moore's law and the quantum advantages potentially obtainable. Photonic networks are promising examples, with experimental demonstrations and potential for obtaining a quantum computer to solve problems believed classically impossible. This introduces a challenge: how does one design or understand such photonic networks? One must be able to calculate observables using general methods capable of treating arbitrary inputs, dissipation and noise. We develop novel complex phase-space software for simulating these photonic networks, and apply this to boson sampling experiments. Our techniques give sampling errors orders of magnitude lower than experimental correlation measurements for the same number of samples. We show that these techniques remove systematic errors in previous algorithms for estimating correlations, with large improvements in errors in some cases. In addition, we obtain a scalable channel-combination strategy for assessment of boson sampling devices.

## I. INTRODUCTION

Bosonic quantum systems have an exponential complexity which has long been recognized as a fundamental theoretical challenge. At the same time, such complexity is also a potential resource in quantum technology. One solution to this challenge is in the application of coherence theory to these problems. Yet very little of the intuition and power of coherence theory[1] and quasi-probability theory[2] has been brought to bear on the problems of quantum technology. As an example, the quantum statistical properties of bosonic networks are almost exclusively discussed in terms of number state representations. Such treatments can be useful, but quickly run into the inevitable complexity limits of number state expansions, when applied to networks of large size.

Linear bosonic networks are now being used to implement novel quantum technologies, including boson sampling [3, 4] and high accuracy, quantum Fourier interferometers [5]. In this paper, we utilize complex P-representation methods [6] that can treat any quantum inputs and outputs to such networks, even at large sizes. The novelty of our approach is in a transformation from an exponentially hard number state problem, to a simpler one in which coherent states with equivalent moments and correlations is sampled. These techniques help analyze quantum hardware. As an application, we treat the open problem of how to *assess* that boson sampling experiments work correctly. We propose a strategy for assessment that is scalable to a large network size.

In such experiments one prepares an $M$-mode bosonic state $\hat{\rho}$, which is input into a passive linear optical multimode device, followed by a measurement on the output [7]. For $N$ single boson inputs into multiple channels, the generation of the random output counts is an exponentially hard computational problem [4, 8]. The most advanced known classical algorithm [9] takes a time exponential in $N$ to generate a single random sample of this type, which makes it impractical at large $N$. Quantum Fourier interferometers are an application of this quantum technology in high-precision metrology [10]. However, most of the development of the theory has taken place using orthogonal number states, which have inherent limits when treating large numbers of bosons or modes. Our methods are based on general quantum phase-space representation theorems, so that the underlying techniques are generally applicable to linear photonic networks.

Boson sampling experiments [5, 7, 11–17] have the goal of demonstrating computations thought to be impossible on classical computers, and lay the foundations for new quantum technologies. To fully realize this potential, one must have tools to analyze them. Ultimately it will be necessary to take account of imperfect inputs, as well as losses and other non-ideal behavior. Here, we derive a hybrid computational and analytic approach to allow the assessment of such networks. This algorithm is not intended to solve the BosonSampling problem of generating the random counts. However, it does make it feasible to analyze, design and assess photonic networks used for this and other quantum technology applications.

Linear photonic networks are defined by an $M \times M$ mode unitary matrix $\boldsymbol{U}$, or more generally by an input-output transformation matrix $\boldsymbol{T}$, which can include losses. The simplest boson sampling experiments [11–16, 18] have an initial $N$-photon state $|\boldsymbol{n}\rangle$. Here $\boldsymbol{n}$ is a vector such that $N = \sum n_j$, $n_j = 0, 1$, and $|n\rangle$ is the number state basis. More generally, one may anticipate that other nonclassical input states will be used. The output photon numbers, $\boldsymbol{n}'$, are the observables. The permanent-squared of the sub-matrices defined by the input and output modes [19, 20] gives the probability of measuring one photon in each preselected output mode [5, 13], as illustrated in Fig. 1. The optimal classi-

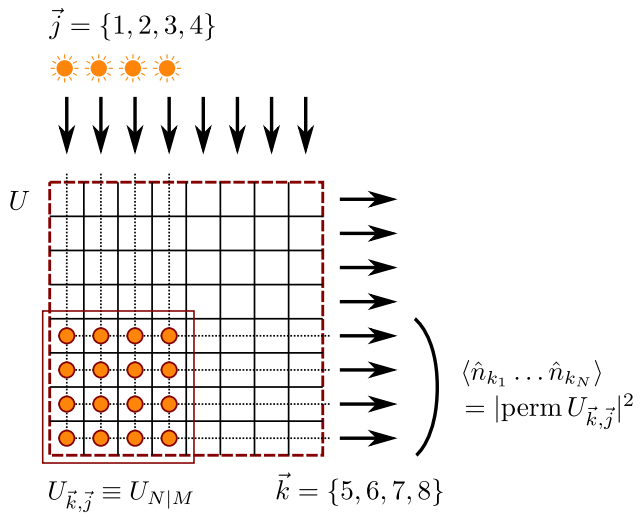


Figure 1. Channel diagram of a boson sampling photonic network gedanken experiment.

cal techniques for calculating this $\#P$ hard problem [21] scale exponentially as $N2^N$ [22] operations for $N$ inputs, and become rapidly infeasible above $N = 50$ [23].

To understand the reason for this scaling, we note that the permanent of a square matrix $\boldsymbol{U}$ is defined as a sum over all permutations $\boldsymbol{m}$ of the set of indices $\{1, \ldots, N\}$, $\mathrm{perm}(\boldsymbol{U}) = \sum_{\boldsymbol{m}} \prod_j U_{jm_j}$. The number of such permutations is $N!$. This complexity occurs naturally in a photonic network, driven by the enormous number of possible interference paths that a photon can take. Typically, the filled channels $N$ are only a small fraction of the open channels $M$, so that $N = M/k$ where $k \gg 1$. Thus, not only is the calculation of each permanent exponentially hard as $N$ increases, but in addition there are exponentially many possible sets of $N$ output channels. Each of these combinations corresponds to a different, unique permanent — all exponentially hard — which must be calculated to predict the full set of output probabilities. While it is known [9] that there exists a random bit generation algorithm that is only exponential in $N$ (rather than $M$), for generating single random samples, this does not solve the assessment problem.

We simulate such linear photonic experiments using "quantum simulation software", by transforming this problem into an exact expansion on non-classical phase space [6], combined with a random sampling method that generates single phase-space samples with an algorithm that is only polynomial in $N$ and $M$. The approach allows unbiased estimates of the experimentally measured probabilities, and scales *better* than an experiment, in terms of time taken for a correlation measurement with a given sampling error. This is an important point, since theories used to assess the experiments must be calculable in a time comparable to the measurements themselves, in order to be useful. In addition, it removes systematic errors that can occur when estimating correlations using previous algorithms. We expect that the general approach will be applicable to experiments using a range of nonclassical input states.

However, neither measurements nor calculations of correlations can be readily scaled to large sizes for assessment in a finite time with typical unitaries. Instead, we propose an analytic signature for assessing claims to solve Boson Sampling, using a hierarchy of measurable combinations of $N$-th order correlations, with successive channels deleted.

Assessment is an integral part of analyzing solutions to exponential complexity. Our proposal involves sums over exponentially large numbers of sub-permanents present in the experiments. This is essential to obtain significant experimental counts. Any assessment is limited by experimental and theoretical scaling limits, due to count rate issues. Our analytic theory itself involves a conjectured solution that is verified numerically. It is likely that it can be rigorously proved using random matrix theory.

Finally, our approach can calculate in principle any input state or output measurement. More details of these applications to quantum Fourier transform interferometry including dissipative effects from phase noise can be found elsewhere [24]. This is essential for understanding how this technology can be scaled to large sizes with imperfect sources, with recent progress in this direction for boson sampling using time-binning techniques [25] and in metrology applications [26].

The contents of the paper are as follows. Section II explains how the complex-P continuous phase-space representation can be used to calculate arbitrary observables given any quantum input state to a photonic network, while Section III gives a discrete version of this method. Section IV gives sampled results for both these approaches, using a randomized calculation of a boson sampling experiment. This section includes a treatment of scaling behavior and a comparison both to experimental sampling errors and to the Gurvits approximation for permanents. Section V explains how these results are generalized to obtain a scalable assessment method for boson sampling, with comparisons to some other proposals. Section VI gives our conclusions.

## II. COMPLEX P-REPRESENTATION

We first explain our phase-space method, which uses the complex P-distribution. This is a quantum phase-space expansion over a basis of coherent state projectors [6]. It is an extension of methods developed originally by Glauber [2] in quantum optics, and has the capability of treating any input quantum density matrix. We use this method to simulate linear photonic experiments, evaluating the measurable photonic moments by probabilistically sampling over specifically selected contours in the higher dimensional complex space $\mathcal{C}^M \times \mathcal{C}^M$. Our simulations *do not* generate a classical sequence of photon counts with a permanental distribution, which is known

to be exponentially hard. Instead we employ something more useful for assessment purposes: a sequence of numbers that generates equivalent correlations and moments to an experiment, but has much *lower* sampling errors for the same number of samples used.

The integration contour is illustrated schematically in Fig. 2, noting that only one complex variable of the $2N$ that are integrated is shown. We note that this simulation does not directly model photon-counting measurements, since there is a different representation for every operator ordering and measurement, as originally mentioned in Dirac's review [27]. The advantage of this approach is that the simulation is no longer exponentially hard to carry out and can still be used to evaluate correlations or moments of photon-counts, which has important applications for assessment protocols.

### A. Contour integrals and input states

There are several different generalized P-representations, including complex and positive valued representations [6]. Because it has a compact form, without large radius tails, the complex P-representation is very scalable when treating the high-order correlations measured in boson sampling experiments. With this representation, the input quantum density matrix $\hat{\rho}^{(\mathrm{in})}$ is represented by an integral over a closed contour $C$ enclosing the origin in a multi-dimensional complex plane as in Fig. 2:

$$\hat{\rho}^{(\mathrm{in})} = \oiint_C P(\boldsymbol{\alpha},\boldsymbol{\beta})\hat{\Lambda}\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right)\mathrm{d}\boldsymbol{\alpha}\mathrm{d}\boldsymbol{\beta}\,. \tag{2.1}$$

Here, $P(\boldsymbol{\alpha},\boldsymbol{\beta})$ is a complex distribution function that is a function of the input photon numbers, while $\boldsymbol{\alpha} = [\alpha_1,\cdots\alpha_M]$ is a vector of $M$ complex numbers, as is $\boldsymbol{\beta}$. The quantum operator basis $\hat{\Lambda}$ is a set of generalized coherent state projectors:

$$\hat{\Lambda}\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right) = \frac{\|\boldsymbol{\alpha}\rangle\left\langle\boldsymbol{\beta}^{*}\right\|}{\left\langle\boldsymbol{\beta}^{*}\right\|\boldsymbol{\alpha}\rangle}\,, \tag{2.2}$$

onto un-normalized Bargmann-Glauber [1] coherent states $\|\boldsymbol{\alpha}\rangle$. These are defined using photon number states $|n_k\rangle$ with $n_k$ photons in the $k$-th input mode,

$$\|\boldsymbol{\alpha}\rangle \equiv \prod_k \left[\sum_{n_k=0}^{\infty}\frac{\alpha_k^{n_k}}{\sqrt{n_k!}}\left|n_k\right\rangle\right]\,. \tag{2.3}$$

For the purposes of the sampling, it is only necessary to know that $\alpha_k$ and $\beta_k$ are complex numbers, and that the phase space method will generate the correct quantum moments. We integrate around a circular contour of radius $r$, by randomly sampling unit modulus complex numbers $z, z'$, where $\alpha = rz$ and $\beta = rz'$ , with a complex-valued weight $P\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right)$. The contour is illustrated in Fig. 2.

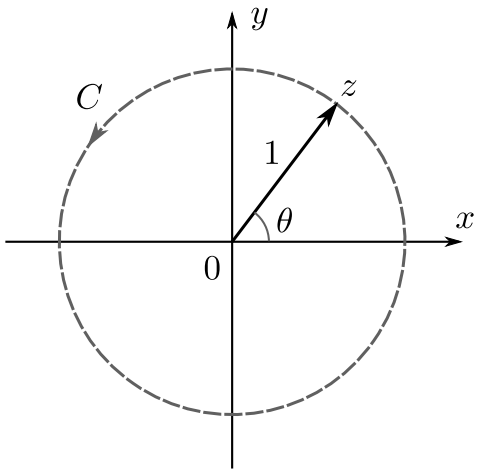


Figure 2. Contour integral for $z$, where $\alpha = rz$. Here $z$ is chosen randomly with unit modulus for sampling purposes, to give a Monte-Carlo sampled contour integral.

The over-completeness of the coherent states in quantum mechanics means that there is more than one way to choose the contour $C$, and in particular the circular radius $r$ can be varied. It is an essential feature of our method that this flexibility enables us to tailor the representation to optimize the sampling for different tasks.

For greater accuracy when developing sampling techniques to be used later, one can also use the fact that the density matrix is hermitian. In this refinement, we define the expansion in an explicitly hermitian form by taking the real part, thus imposing the constraint that the final result must be real:

$$\hat{\rho}^{(\mathrm{in})} = \Re\oiint_C P(\boldsymbol{\alpha},\boldsymbol{\beta})\hat{\Lambda}\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right)\mathrm{d}\boldsymbol{\alpha}\mathrm{d}\boldsymbol{\beta}\,. \tag{2.4}$$

To perform simulations using this approach it is necessary to have a representation, $P\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right)$, of the initial state. This distribution exists for an *arbitrary* initial density matrix, either pure or mixed, with any coherence properties and photon numbers [6]. In this paper, we focus on the simplest case: a pure state in which the $k$-th channel has an $n_k$-boson input. The usual boson sampling experiments send one photon into each of a set $\sigma$ of $N$ input modes, so that $n_k = 0, 1$, and, therefore:

$$N = \sum_k n_k = \sum_{k\in\sigma} n_k\,.$$

There is more than one way to represent these states, which are illustrated below. However, the same techniques can be readily adapted to treat other quantum inputs, for example squeezed or entangled states, which can be more easily generated in experiments.

### B. Output measurements

Any bosonic input density matrix $\hat{\rho}^{(\mathrm{in})}$ is changed by transmission through a linear network to an output density matrix $\hat{\rho}^{(\mathrm{out})}$. In general, to take into account inevitable losses, one can imagine a larger unitary that includes loss channels, but only consider the sub-matrix $\boldsymbol{T}$ for the accessible channels that are measured, with other channels just being ignored. An example is the transmission matrix $\boldsymbol{T} = \sqrt{t}\boldsymbol{U}$ in which all channels experience equal loss. This combines the $M \times M$ unitary mode transformation $\boldsymbol{U}$ of the entire network, with an absorptive transmission coefficient $t$.

The effect of the transmission matrix $\boldsymbol{T}$ on the phase-space distribution is straightforward, owing to the normal ordering property of the complex P-representation. It simply transforms the coherent amplitudes in a deterministic way [28], such that $\boldsymbol{\alpha}^{(\mathrm{out})} = \boldsymbol{T}\boldsymbol{\alpha}$, $\boldsymbol{\beta}^{(\mathrm{out})} = \boldsymbol{T}^{*}\boldsymbol{\beta}$. The resulting output density matrix is therefore a contour integral with the same weight, but a modified projector:

$$\hat{\rho}^{(\mathrm{out})} = \Re \oiint_{C} P(\boldsymbol{\alpha},\boldsymbol{\beta})\hat{\Lambda}\left(\boldsymbol{T}\boldsymbol{\alpha},\, \boldsymbol{T}^{*}\boldsymbol{\beta}\right) \mathrm{d}\boldsymbol{\alpha}\mathrm{d}\boldsymbol{\beta}\,. \tag{2.5}$$

Any output number correlation is given by computing moments of the known input $P$-function, using the equivalent output number variable,

$$\begin{aligned} n_{k}^{(\mathrm{out})}(\boldsymbol{\alpha},\boldsymbol{\beta}) &= \alpha_{k}^{(\mathrm{out})}\beta_{k}^{(\mathrm{out})} \\ &= \left(\sum_{j} T_{kj}\alpha_{j}\right)\left(\sum_{j} T_{kj}^{*}\beta_{j}\right)\,. \end{aligned} \tag{2.6}$$

A typical observable in these experiments [29–31] is an arbitrary normally ordered quantum correlation of $k$-th mode operators, $\hat{n}_{k} \equiv \hat{a}_{k}^{\dagger}\hat{a}_{k}$, belonging to a set of output modes $\sigma'$. In boson sampling we are interested in the correlator over $N$ output channels. Since it is assumed that no photons are generated inside the device — it has no gain — any $N$-fold coincidence count means that there is exactly one photon detected in each output channel. The correlator or coincidence count therefore has accessible eigenvalues of $1, 0$.

This property of binary counts means that the correlator is also the probability of an $N$-fold coincidence count, given an $M \times M$ transmission matrix,

$$P_{N|M} \equiv \left\langle \prod_{k\in\sigma'} \hat{n}_{k} \right\rangle_{\mathrm{Q}}\,, \tag{2.7}$$

where quantum expectation values are denoted $\langle\rangle_{Q}$. In other words, the correlation and the count probability are identical. Since these observables are hermitian, we can take the real part of the expansion as in Eq. (2.5). The quantity of interest here is therefore given by the following expression for $P_{N|M}$, where a contour integral over the complex P-function is denoted as $\langle\rangle_{P}$:

$$\begin{aligned} P_{N|M} &= \Re \oiint_{C} P(\boldsymbol{\alpha},\boldsymbol{\beta}) \prod_{k\in\sigma'} n_{k}^{(\mathrm{out})}(\boldsymbol{\alpha},\boldsymbol{\beta})\mathrm{d}\boldsymbol{\alpha}\mathrm{d}\boldsymbol{\beta} \\ &\equiv \Re\left\langle \prod_{k\in\sigma'} n_{k}^{(\mathrm{out})}(\boldsymbol{\alpha},\boldsymbol{\beta})\right\rangle_{\mathrm{P}}\,. \end{aligned} \tag{2.8}$$

As the representation $P$ is not unique (the choice of contour being flexible), one can choose different representations of the input state. The main point that we wish to emphasize is that these techniques comprise a *complete* set of efficient strategies for treating quantum photonic networks. Even though we focus here on number state inputs and outputs, the general approach is not limited to this case. Such methods have previously been used in nonlinear quantum problems as well, and give a general strategy for calculating any normal-ordered observable correlation [1].

The complex P-representation is known to have exact solutions for the non-equilibrium steady-state solutions in other quantum photonic devices. These include the driven anharmonic cavity [32–34], the nonlinear two-photon absorber [35], the degenerate parametric oscillator [36], the non-degenerate parametric oscillator [37] as well as approximate solutions for coupled nonlinear cavities [38] and Bose condensates [39]. Thus, it may not be impossible to extend these techniques to nonlinear photonic networks.

## III. DISCRETE COMPLEX P-REPRESENTATION

Just as the coherent states do not provide unique expansions of quantum states, neither do complex P-representations. This provides an opportunity to adapt the representation to a given task. In this section, we recall the most well-known type of complex P-representation, and also explain an alternative form using discrete sums, which is more suitable to sampling high-order correlations. For number state inputs, the complex P-distribution has a well-known solution that is readily proved to exist using Cauchy's theorem:

$$P\left(\boldsymbol{\alpha},\boldsymbol{\beta}\right) = \prod_{k}\left[\left(\frac{n_{k}!}{2\pi\mathrm{i}}\right)^{2}\frac{e^{\alpha_{k}\beta_{k}}}{(\alpha_{k}\beta_{k})^{n_{k}+1}}\right]\,, \tag{3.1}$$

where $n_{k}$ is the photon count in the mode $k$. For $n_{k} = 0$, the single pole at the origin means that one can replace the input variable by its vacuum value of $\alpha_{k} = \beta_{k} = 0$ in any average, so that $P\left(\alpha_{k},\beta_{k}\right) = \delta\left(\alpha_{k}\right)\delta\left(\beta_{k}\right)$, where for the case of a vacuum mode input into a single mode $k$.

### A. Discrete summation method

While the contour integral solution given above always exists, other forms are also possible. These give a differ-

ent strategy allowing efficient random sampling methods. Here we will use the limit of an infinitely small contour, which allows us to use a discrete phase summation method. Other choices can be utilized for more general states. Discrete phase sums have particular utility in cases where the initial photon number is bounded, for example with a fixed input boson number. We now introduce one of these discrete approaches, which we term the discrete or qudit complex P-representation (QCP) [40]. This construction is useful in the limit of the circular radius $r \to 0$, which leads to $d$ coherent phases distributed on an infinitesimal circle.

This alternative solution for $P(\boldsymbol{\alpha},\boldsymbol{\beta})$, is still in the form of Eq. (2.1), but with the distribution consisting of $d$ discrete delta-functions arranged on a circle of radius $r \to 0$. With appropriate choices of the complex amplitude at each point, one can represent an arbitrary quantum state of a $d$-dimensional qudit, with initial photon occupation numbers in each mode of up to $n = d-1$. We will show that this approach unifies quantum representation theory [6] with discrete sampling permanent approximation methods [41]. It has numerical properties that make it both accurate and efficient.

In this approach the input quantum state is expanded as a superposition of a discrete set of coherent amplitudes in each of the $N$ non-vacuum modes, defined as:

$$\begin{aligned} \alpha^{(q)} &= rz^{q} \quad q = 0,\dots d-1 \\ \beta^{(\tilde{q})} &= rz^{-\tilde{q}}, \quad \tilde{q} = 0,\dots d-1\,, \end{aligned} \tag{3.2}$$

where $z = \exp(i\phi)$ is the $d$-th root of unity and the phase interval is $\phi = 2\pi/d$. The density matrix is then expanded in coherent states with a discrete summation for the $N$ modes that are not initially in the vacuum state, using a discretized complex P-function, $P_Q(\boldsymbol{q},\tilde{\boldsymbol{q}})$ which is expanded in terms of a number-projected kernel, $\hat{\Lambda}^{(d)}$:

$$\hat{\rho} = \frac{1}{d^{2M}} \sum_{\boldsymbol{q},\tilde{\boldsymbol{q}}} P_Q(\boldsymbol{q},\tilde{\boldsymbol{q}}) \hat{\Lambda}^{(d)}(\boldsymbol{q},\tilde{\boldsymbol{q}})\,. \tag{3.3}$$

Here $\hat{\Lambda}^{(d)}(\boldsymbol{q},\tilde{\boldsymbol{q}}) = \mathcal{P}\left\|\boldsymbol{\alpha}^{(\boldsymbol{q})}\right\rangle\left\langle\boldsymbol{\beta}^{(\tilde{\boldsymbol{q}})*}\right\|$, where $\mathcal{P}$ is a projector onto a subspace of up to $d-1$ photons per input mode, and the coherent amplitudes are:

$$\begin{aligned} \boldsymbol{\alpha}^{(\boldsymbol{q})} &\equiv \left[\alpha^{(q_1)},\alpha^{(q_2)},\dots\right] \\ \boldsymbol{\beta}^{(\tilde{\boldsymbol{q}})} &\equiv \left[\beta^{(\tilde{q}_1)},\beta^{(\tilde{q}_2)},\dots\right]. \end{aligned} \tag{3.4}$$

This form of representation, the complex qudit P-distribution (QCP), can be shown to always exist for Hilbert spaces of bounded occupation numbers, using discrete Fourier transforms. While the full derivation [40] is given elsewhere, there is an important point to be noted. The discrete form is combined with a projection operator at finite $r$. However, these coherent states have unit norm in the limit of $r \to 0$, and in this limit no projection is needed. Input states of higher photon number are excluded automatically, either from Fourier orthogonality or because they have negligible weight in this limit.

Therefore, using the above expansion, a complex qudit P-function $P_{\rm Q}$ always exists for the input density matrix $\hat{\rho}$ used here, where:

$$P_{\rm Q}(\boldsymbol{q},\tilde{\boldsymbol{q}}) = \sum_{\boldsymbol{n},\boldsymbol{m}} \langle\boldsymbol{m}|\,\hat{\rho}\,|\boldsymbol{n}\rangle \prod_j \frac{\sqrt{n_j!m_j!}}{r^{n_j+m_j}} e^{[{\rm i}\phi(n_j\tilde{q}_j - m_j q_j)]}\,. \tag{3.5}$$

Here the expansion allows occupation numbers for mode $k$ up to $n_k = 0,\dots d-1$. We note that such discrete sampling gives continuous sampling in the limit of $d \to \infty$, and in this limit simply reduces to the earlier contour integral result of (3.1) at small radius. The discrete expansion can be verified as a solution, by inserting this distribution into the expansion of the density matrix, and noting that, from the properties of the discrete Fourier transform,

$$\frac{1}{d^N} \sum_{\boldsymbol{q}} {\rm e}^{{\rm i}\boldsymbol{q}\cdot(\boldsymbol{n}-\boldsymbol{m})\phi} = \delta_{\boldsymbol{n}-\boldsymbol{m}}. \tag{3.6}$$

In the simplest binary, or qubit, case where $d = 2$ and $n_k = 0,1$ we consider $q_k = \{0,1\}$. This implies that $\alpha_k = \pm r$.

For finite $d$, this method is only applicable to photon number inputs that are bounded, which is precisely the case in many boson sampling experiments, such as Quantum Fourier interferometry [5]. It is also the case in the assessment scheme described in Section V. Since this is a case of the complex P-representation, the input-output transformation used previously in (2.8) is still valid. The output coherent amplitude for a given discrete input $\boldsymbol{\alpha}^{(\boldsymbol{q})}$ is therefore $\boldsymbol{T}\boldsymbol{\alpha}^{(\boldsymbol{q})}$. This output is no longer restricted to the same input set of discrete phases, and hence can include other particle numbers, different to those in each input channel. Physically this means that interference effects can occur, leading to the coherent Hong-Ou-Mandel type phenomena that are responsible for the nonclassical boson sampling output statistics.

The output photon number phase-space variable, given an input of single bosons into the first $N$ modes, is:

$$n_k^{(\rm out)}(\boldsymbol{\alpha}^{(\boldsymbol{q})},\boldsymbol{\beta}^{(\tilde{\boldsymbol{q}})}) = r^2 \left(\sum_{j\in\sigma} T_{kj} z^{q_j}\right)\left(\sum_{j'\in\sigma} T_{kj'} z^{\tilde{q}_{j'}}\right)^*. \tag{3.7}$$

Here the notation $j \in \sigma$ is used to restrict the sum to the $N$ occupied input channels. This is then further summed over all the possible input values of $\boldsymbol{q},\tilde{\boldsymbol{q}}$, and weighted with $P_{\rm Q}(\boldsymbol{q},\tilde{\boldsymbol{q}})$. The quantum expectation value or probability of observing a simultaneous count in each of a particular set $\sigma'$ of $N$ output channels is then given by:

$$P_{N|M} = \frac{1}{d^{2N}} \Re\left[\sum_{\boldsymbol{q},\tilde{\boldsymbol{q}}} P_Q(\boldsymbol{q},\tilde{\boldsymbol{q}}) \prod_{k\in\sigma'} n_k^{(o)}(\boldsymbol{q},\tilde{\boldsymbol{q}})\right]. \tag{3.8}$$

### B. P-Representation for single-photon inputs

We have shown that there is a mapping between the phase-space expansion and a combinatoric sum. In the simplest one mode, one boson case, an inspection of the general discrete result, Eq. (3.5), shows that the P-function is simply the product of two complex variables together with a radial factor:

$$P_{\mathrm{Q}}(q,\tilde{q}) = \frac{1}{r^2} z^{\tilde{q}} z^{-q} \,. \tag{3.9}$$

The advantage of the small radius limit is that the known identities for the generalized P-representation are all valid, and the $r \to 0$ limit can be taken after the calculation. Thus, we can use the standard result that after transmission through a linear optical system with phase-shifts, beam-splitters and losses, the output coherent amplitudes are multiplied by the relevant linear transmission matrix. In calculating $N$-th order correlations of an $N$-photon input, all the factors proportional to the radius $r$ simply cancel.

As a result, after including the complex P-function weights, an $N$-th order output correlation is:

$$\begin{aligned} P_{N|M} &= \left| \frac{1}{d^N} \sum_{\boldsymbol{q}} \left\{ \prod_{\ell \in \sigma} z^{-q_\ell} \prod_{k \in \sigma'} \left( \sum_{j \in \sigma} T_{kj} z^{q_j} \right) \right\} \right|^2 . \\ &= \left| \mathrm{perm}\left[ T\left(\sigma', \sigma\right) \right] \right|^2 . \end{aligned} \tag{3.10}$$

As expected [8], this is the square of the permanent of the sub-matrix of $\boldsymbol{T}$ with rows in $\sigma'$ and columns in $\sigma$, which we call $\mathbf{T}\left(\sigma', \sigma\right) \equiv \mathbf{M}$. After summation on the $\boldsymbol{q}$ indices, the only terms that survive involve products of distinct permutations of the matrix indices, which is the permanent. However, there are exponentially many terms involved at large $N$, which requires a sampling technique that we explain in the next section.

We finally note that the complexity for exact calculations depends on the value of the discretization number, $d$. The minimum value of $d$ for a boson sampling experiment with 1 photon per input channel is $d = 2$. This gives the least complexity in an exact computation. It also directly corresponds to the most efficient and well-known computational technique for computing permanents.

However, this is only a lower bound. One can use any value of $d \geq 2$. In particular, we can also consider the exact continuum path-integral as well. This simply corresponds to replacing the sums in Eq. (3.10) by integrals over $\boldsymbol{\phi} = [\phi_1, \ldots \phi_N]$, so that:

$$P_{N|M} = \left| \int \frac{d^N \boldsymbol{\phi}}{(2\pi)^N} \prod_{\ell} e^{-i\phi_\ell} \prod_{k} \left( \sum_{j} M_{kj} e^{i\phi_j} \right) \right|^2 . \tag{3.11}$$

The last expression is also the $d \to \infty$ limit of the discrete complex P-representation, which is a complete representation. It can be generalized to treat arbitrary input states, in which case one should use the full expression of Eq. (3.5) to specify the P-distribution. While an integral is more complex numerically than a discrete sum, we show in the next section that it is extremely efficient when random sampling methods are employed.

## IV. METHODS FOR SAMPLING PHASE-SPACE

While the results of the previous section are exact, they are also exponentially complex. Next, we will explain the approximate randomized technique that is utilized to calculate the photon counting coincidence probabilities in polynomial time. This does not contradict complexity theory results, which only prohibit polynomial time methods for direct generation of random photon counts or the exact evaluation of matrix permanents. Verifying such coincidence probabilities is an important step in any assessment procedure for boson sampling devices.

For reasons of efficiency, the combinatoric sums can be evaluated approximately by taking pairs of randomly chosen integer vectors $\left(\boldsymbol{q}^{(j)}, \tilde{\boldsymbol{q}}^{(j)}\right)$, and averaging over samples of these random phases. In a Hilbert space of dimension $d$, there are $d$ possible random discrete phases. This can be taken to the limit of $d \to \infty$ for a continuously sampled Monte Carlo integral.

There are similarities between experimental measurements and the use of sampled quantum simulations. In both cases there is a sampling error, since one must calculate or measure results for correlations using a finite number of samples. The time taken is proportional to the number of samples used. It is therefore crucial to know how the average sampling error scales with the number of active channels $N$, which determines the computational time

We find some important results, as follows:

- Random sampling method is scalable even for the *most* complex exact method, that is, the $d \to \infty$ limit.
- Computed correlations have a lower sampling error than in an experimental measurement.
- These techniques eliminate a systematic error that occurs using previous approximate methods.

### A. Overview of calculations and sampling procedure

In order to explain the general procedure, we will summarize the steps involved in analyzing a photonic network.

1. Firstly, the input $M$-mode bosonic state $\hat{\rho}_0$ must be known, which can be any density matrix. For bounded photon number, a discrete complex-P expansion can be utilized, otherwise a continuous expansion is necessary. The existence theorems [6] are

used to choose a contour and obtain the P-function, $P(\boldsymbol{\alpha},\boldsymbol{\beta}) = |P|\, e^{i\phi}$, which is a distribution over a contour in $N$ complex dimensions.

2. The $M \times M$ transmission matrix $T$ must be known or measured. This can include any arbitrary loss in principle. In cases of low-frequency phase noise — for example, from $1/f$ noise [42, 43] — the refractive index and hence the transmission matrix is a random variable, since each repetition of the experiment has a different $T$. Dispersion and noise correlated on time-scales shorter than pulse durations require a more sophisticated theory [28, 44, 45].

3. The input contour is sampled with equal probability, $|P|$, to give coherent samples $\boldsymbol{\alpha}^{(j)}, \boldsymbol{\beta}^{(j)}$, together with the relevant complex phase. These are multiplied by the transmission matrix giving the output coherent samples whose moments, combined with the phase factor $\phi$, give the expected quantum correlations.

This gives a detailed procedure that would be followed in the most general case. The present paper treats the simplest case of considering correlated single-photon outputs in a predefined set of channels, without dissipation. More general examples are treated elsewhere.

### B. Sampled calculations of the permanent squared

In both the continuous phase and discrete phase approach, we use independent random samples of $\boldsymbol{z}^{(j)} = \exp\left(i\phi \boldsymbol{q}^{(j)}\right)$, $\tilde{\boldsymbol{z}}^{(j)} = \exp\left(i\phi \tilde{\boldsymbol{q}}^{(j)}\right)$. These amplitudes of unit modulus correspond to scaled versions of $\boldsymbol{\alpha}, \boldsymbol{\beta}^*$ in the original complex P-representation. For each channel $(i)$ and sample $(j)$, we have $z_i^{(j)} = z^{q_i^{(j)}}$. In order to obtain an unbiased estimate of the absolute value squared, we use an ensemble of size $L$ of *two independent* conjugate noise vectors, $\boldsymbol{z}, \tilde{\boldsymbol{z}}$.

To explain the notation, we define a polynomial function $p(\mathbf{M}, \boldsymbol{z})$ of the $N \times N$ sub-matrix of interest, as:

$$p(\mathbf{M}, \boldsymbol{z}) = \prod_{\ell=1}^{N} z_\ell^{-1} \prod_{k=1}^{N} \left( \sum_{j=1}^{N} M_{kj} z_j \right). \tag{4.1}$$

Each approximation of the permanent, which we denote as $p(\mathbf{M}, \boldsymbol{z})$, is a function of the sub-matrix $\mathbf{M}$ and a noise vector $\boldsymbol{z}$.

Regardless of the discretization parameter $d$, any of these sampling methods give an unbiased estimate of the modulus-squared of the permanent, and are applicable to any type of matrix permanent calculation, although the error scaling depends on the algorithm and type of matrix. We define $\langle\rangle_L$ as the stochastic expectation value over an ensemble of $L$ random samples of a stochastic vector. The sampled calculation of a permanent is then given by:

$$\langle p(\mathbf{M}, \boldsymbol{z}) \rangle_L \equiv \frac{1}{L} \sum_{j=1}^{L} p(\mathbf{M}, \boldsymbol{z}^{(j)}). \tag{4.2}$$

When applied to calculating a permanent, with the discretization value of $d = 2$, this method corresponds to a previously known permanent approximation method, known as the Gurvits approximation [41].

We do not use this more traditional approach. The methods described here lead to a much larger class of permanent approximations, as well as algorithms that are specifically optimized for calculating the *modulus squared* of the permanent — which is the quantity of interest in boson sampling. In fact, the optimal sampling method for the modulus squared is not simply using the permanent algorithm and squaring the result. The drawback with this method is that when taking the modulus squared of the stochastic estimate, an additional term is obtained equal to the sampling variance. Hence, the resulting estimate is not unbiased.

Instead, following the derivation given above, our coincidence rate — or modulus squared of the permanent — is approximated as the real part of a product of two *independent sets of* samples, $\mathbf{z}, \tilde{\mathbf{z}}$, each composed of $E$ random vectors:

$$P_{N|M} \approx \Re\left[ \langle p(\mathbf{M}, \boldsymbol{z}) \rangle_E \langle p(\mathbf{M}, \tilde{\boldsymbol{z}}) \rangle_E^* \right]. \tag{4.3}$$

The expression has two terms which are independent but conjugate on average. The factored terms give independent estimates of the permanent and its conjugate, and their product is an unbiased estimate of the modulus squared.

### C. Estimating the sampling error

In any calculation using random sampling, it is essential to have a statistical procedure for estimating the errors. A further modification is therefore used in our numerical procedure to obtain error estimates. Since we wish to understand how reliable the procedure is in terms of its sampling error, the random sampling process described above is repeated and averaged a large number of times for each sub-matrix, to obtain independent statistical estimates. This allows the standard deviation in the mean due to sampling error to be estimated. This gives a theoretical error-bar in our graphs, to indicate how accurate the calculation is.

To achieve this in our calculations, we divide the total number of ensembles $L$ into $L_2$ sub-ensembles, so that $L = L_1 L_2$. Each of the $L_2$ sub-ensembles has a large number $L_1$ of independent noise terms. The final estimate is then an average over the $L_2$ sub-ensemble estimates:

$$|\mathrm{perm}(\mathbf{M})|^2 \approx \frac{1}{L_2} \Re \sum_{i=1}^{L_2} \langle p(\mathbf{M}, \boldsymbol{z}) \rangle_{L_1}^{(i)} \langle p(\mathbf{M}, \tilde{\boldsymbol{z}}) \rangle_{L_1}^{(i)*}. \tag{4.4}$$

The notation $\langle\rangle_L^{(i)}$ indicates an average over the $i$-th subensemble, as given in Eq. (4.2). These individual averages have independent distributions. For large $L_1$, by the central limit theorem, they are approximately Gaussian since they involve sums of many independent random variables. In the limit of small variances, the products of independent Gaussian distributed quantities with a non-vanishing average remain nearly Gaussian distributed. Hence each term of the sum over sub-ensembles is approximately Gaussian distributed, as we are using independent estimates of the permanent and its conjugate, or at worst has a $\chi^2$ distribution.

Combining sub-ensemble estimates in 'quasi-conjugate' pairs that are complex conjugate on average, but are independent, allows for an unbiased estimate of the permanent squared. The use of sub-ensembles additionally gives an approximate sampling error, using statistical methods valid for Gaussian distributions. We have verified that these statistical estimates are in fact reliable, by also calculating the error from exact calculations of the permanent, where available. We find that there is a good agreement between calculated and estimated errors, and this agreement is maintained over a wide range of matrix sizes, as shown in Fig. 3. Results are shown here for $d = 2$ and $d \to \infty$, with similar results for the sampling errors, and good agreement with exact calculations. Since the errors depend on the unitary, results are obtained here by averaging over randomly chosen unitaries, as explained next.

### D. Unitary averages

Although one can simulate individual unitaries with this approach, it is more useful to know how the general performance scales with network size. This is only meaningful when expressed as an average over all possible unitary matrices. Serendipitously, infinite sums over exponentially complex permanents are analytically calculable, using results from the theory of matrix polar coordinates and random matrices. This can be usefully employed to obtain the expected scaling of the average count rate, for comparison purposes. We therefore can express our results on computing permanents as averages over the Haar measure of unitary matrices. These averaging techniques are well-known, and can be applied to numerical averages.

In these calculations, we therefore obtain a set of randomly chosen unitary matrices according to the Haar measure over the unitaries, to give a *third* level of random averaging. We repeat the above process for each unitary matrix, to get unitary averages over the sampling errors that occur in simulating typical random unitary matrices.

For comparison purposes, we note that a boson sampling experiment, the probability of finding $N$ photons in the $M$ output modes of a linear optical network is given by $P_{N|M}$. In the lossless case where $T$ is unitary, this is the permanent squared of an $N \times N$ sub-matrix of an $M \times M$ unitary matrix, perm $(\boldsymbol{U}_{N|M})$ [13]. That is, in terms of our earlier notation, we take the lossless case for simplicity ($t = 1$) and consider the transmission matrix of interest as a unitary sub-matrix: $\boldsymbol{M} = \boldsymbol{U}_{N|M}$. We denote the average over all unitary matrices of this permanent squared as:

$$\begin{aligned}\left\langle P_{N|M}\right\rangle_{\mathrm{U}} &= \left\langle \left|\mathrm{perm}\left(\boldsymbol{U}_{N|M}\right)\right|^2\right\rangle_{\mathrm{U}} \\ &\equiv \bar{P}_{N|M}. \end{aligned} \tag{4.5}$$

Using techniques from the theory of matrix polar coordinates and unitary averages, this has a known scaling law given by [46, 47]:

$$\ln \bar{P}_{N|M} = N\epsilon\left(k\right) + O(\ln N), \tag{4.6}$$

where $\epsilon\left(k\right) = k\ln k - (1+k)\ln(1+k)$ and $k = M/N$ is the channel ratio.

Unitary averages that are carried out numerically involve a different type of sampling error. Since the space of unitaries is extremely high dimensional, this average is carried out in a Monte Carlo way, by choosing random unitaries according to their Haar measure and averaging over a finite set. In most cases we find that this variance over the unitaries is rather small. Where it is significant, the unitary sampling errors are plotted. This does not exclude the possibility that there could be atypical unitaries where the count rates or estimation errors are very different from the unitary average. However, these fall into a set of very small measure, especially at large $N$ values.

### E. Sampling error comparisons

There is a close relationship between the QCP method with qubits for the permanent squared, and the Gurvits method for permanent approximation [48], in the case of an $N$-photon input and output. Both methods employ random sampling over binary numbers. We extend this to qudit and continuous cases, which has advantages in certain calculations. The Gurvits method provides an unbiased estimate of the permanent. However, it introduces systematic errors when computing the permanent squared, owing to effects of the finite distribution width. It also has large sampling errors when used for combinations of permanents, as described in the next section.

Figs. 4(a) and (b) show the systematic errors of the modulus-squared of the permanent using the Gurvits method and the QCP representation with $d \to \infty$ for a ratio $k = 4$, when averaged over 100 randomly chosen $40 \times 40$ and $80 \times 80$ unitary matrices respectively. These figures show that the Gurvits algorithm has a large systematic bias for the permanent squared. This is an order of magnitude larger than the standard deviation in the examples given here. This statistical bias is reduced as the number of samples is increased, but the samples required rapidly become impractical as the matrix size $N$ is increased.

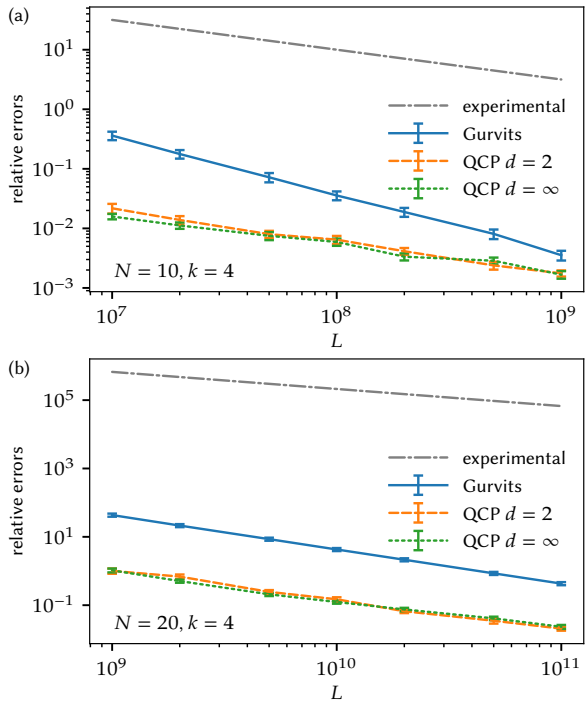


Figure 3. Actual error of the Gurvits method (solid blue lines) and the QCP representation with $d = 2$ (dashed orange lines) and $d \to \infty$ (dotted green lines) for the modulus-squared of the permanent, relative to the value of the permanent-squared, as a function of the total number of samples $L = L_1 L_2$. For each point we have used $N_{\rm m} = 100$ random unitary matrices and $L_2 = 1000$ sub-ensembles of size (a) $L_1 = 10^4 \ldots 10^6$ and (b) $L_1 = 10^6 \ldots 10^8$. Here we consider $k = 4$, (a) $N = 10$ and (b) $N = 20$. Dash-dotted grey lines denote the estimated experimental error $\sqrt{\langle P \rangle_{\rm e,m}/L}$.

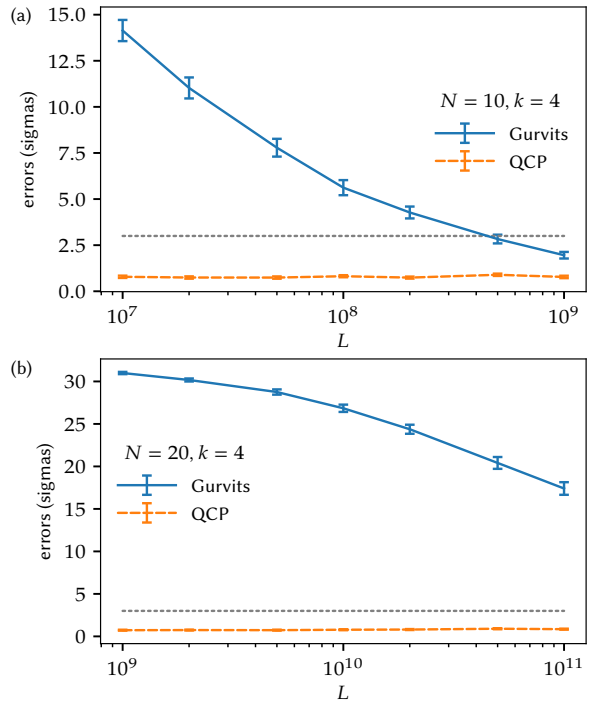


Figure 4. Actual error of the Gurvits method (solid blue lines) and the QCP representation with $d \to \infty$ (dashed orange lines) for the modulus-squared of the permanent, expressed in the units of the sampling error, as a function of the total number of samples $L$. Graphs have $N = 10$ and $N = 20$ respectively, with $k = 4$. The dotted line indicates a $3\sigma$ confidence interval. Other parameters are the same as in Fig. 3.

The data processing for Fig. 4 is performed as follows. Each data point in the plot corresponds to $L$ total samples, which is split in $L_2$ sub-ensembles of $L_1 = L/L_2$ samples. For the Gurvits method the effective number of sub-ensembles is $\tilde{L}_2^{(G)} = L_2 \equiv 10^3$, and for the QCP it is $\tilde{L}_2^{(\rm QCP)} = L_2/2$ (since two independent sub-ensembles are used to produce a single sub-ensemble value of the permanent squared). In this way we ensure that the total number of random samples used for the comparison is the same for both methods. For each data point we obtain a $N_{\rm m} \times \tilde{L}_2$ matrix $P_i^{(j)}$ of permanent squared values for each of $N_{\rm m}$ random matrices and each sub-ensemble. We also calculate the exact permanent squared values $\tilde{P}^{(j)}$ for each random matrix.

Per-matrix actual errors are:

$$E^{(j)} = \left| \frac{1}{\tilde{L}_2} \sum_{i=1}^{\tilde{L}_2} P_i^{(j)} - \tilde{P}^{(j)} \right| \equiv \left| \langle P^{(j)} \rangle_{\rm e} - \tilde{P}^{(j)} \right|, \quad (4.7)$$

estimated sampling errors are calculated as:

$$E_S^{(j)} = \frac{\sqrt{\langle \left(P^{(j)}\right)^2 \rangle_{\rm e} - \langle P^{(j)} \rangle_{\rm e}^2}}{\sqrt{L_2}}. \quad (4.8)$$

We then plot the ratio $\Delta$ of actual errors $E$ relative to the sampling errors $E_{\rm samp}$, that is, we calculate the values

$$\Delta^{(j)} = E^{(j)}/E_S^{(j)}, \quad (4.9)$$

and plot the mean of the relative error $\langle \Delta \rangle_{\rm m}$ averaged over a finite set of unitaries and the estimated error of the mean $\sqrt{\langle \Delta^2 \rangle_{\rm m} - \langle \Delta \rangle_{\rm m}^2}/\sqrt{N_{\rm m}}$ as error-bars.

In summary, our methods gives an unbiased estimate of the permanent squared, which is the relevant quantity in boson sampling experiments. Even more importantly, our methods can also be used in the practical case that the input is not a pure binary number state, or where the output measurements are not $n$-th order correlations.

From the mathematical viewpoint, these numerical methods unify the complex P-representation approach in quantum optics with the computational problem of efficient approximations of permanents and the permanent squared, which is the relevant calculation for boson sampling.

### F. Experimental vs simulated errors

To make useful *predictions* about experimental observables, perfect accuracy is not essential. It is only necessary to calculate the output correlations with better than experimental errors. Such errors can be accurately estimated from the count rates, which are determined by the permanents in the idealized case of number-state inputs. In the experimental case, the measured standard deviations are simply Poissonian errors in the counts, whose scaling is estimated in the figures.

To make sensible comparisons, one must choose timescales that are comparable. Here we note that both digital computers and photonic devices use similar counting and logic electronics, which limit the speed of any one simulation or measurement task. To obtain good statistics, both types of device require repeated measurement and averages over many samples, to reduce sampling errors.

The time taken is then just $T = LT_N$ for $L$ samples, where $T_N$ is the time taken per sample for $N$ channels. We assume that the factor $T_N$ is similar for digital and photonic devices. In both cases, there are other effects due to computational overheads and/or physical limits as $N$ increases, which are neglected here for simplicity. Given this assumption, we can compare the sampling errors over similar time-scales, that is, with the same number of samples. Our main purpose is to show that, at a given error, the computation of quantum correlations is as feasible as an experimental measurement.

Figs. 5(a)-(d) show the average modulus squared of the permanent of an $N \times N$ sub-matrix, denoted by $\bar{P}_{N|kN}$, as well as the mean error, or average deviation of the sampled result from the exact value. The mean error $E$ is defined as $E = \langle |P - \langle \tilde{P} \rangle_{\rm e}| \rangle_{\rm m}$, where $P$ is exact, and $\langle \tilde{P} \rangle_{\rm e}$ is the quantum simulation ensemble average. All results are averaged over a finite unitary ensemble of matrices $\langle \rangle_{\rm m}$. We have also plotted the Poissonian experimental error, which is asymptotically *larger* than the error of our simulations for an identical sample number.

While exact results for permanents are not available at large $N$, owing to their complexity, the average scaling behavior of $N$-channel coincidences over all possible unitaries can be calculated analytically. One could make comparisons of every possible unitary, but as they are uncountably infinite in number, this approach is not very practical. In each experiment with $E$ trials, the unitary average Poissonian measurement variance due to shot-noise is $\left\langle \sigma^2 \right\rangle_{\rm U} = \bar{P}/L$, which therefore scales as $\ln \left\langle \sigma^2 \right\rangle_{\rm U} = N\epsilon\left(k\right) - \ln L + O(\ln N)$.

To verify an experimental probability, one must have a theoretical error less than this experimental sampling error. Typically the channel ratio $k$ ranges from unity — its minimum value — up to a number of order $N$, which has theoretical advantages in computability theory [8]. Hence our numerical results given below are compared with sampling errors in the corresponding experiments, over a range of parameters. The results show that the computational error using our method has a *better* scaling than experiment, for the same size $L$ of the ensemble of measurements or samples.

The graphical comparisons demonstrate a necessary requirement for any assessment method. Without this favorable scaling, the calculation of expected correlations would take exponentially longer than the experiment itself. Instead, computational sampling errors reduce rapidly with $N$, faster than experimental error reductions. This means that the computational time is not prohibitive.

However, the graphs also show that the average permanent values reduce even faster. This makes direct permanent *measurements* problematic at large $N$, except for matrices with large permanents. In other words, at sufficiently large $N$ the average count rate for any individual set of channels goes rapidly to zero for most unitary submatrices, so that the measurement of an individual correlation or moment is impractical. We turn to a solution of this problem in the next section.

## V. CORRELATION BASED VERIFICATION STRATEGY

As boson sampling experiments improve, the problem of verification at large $N$ will become acute [16, 29, 31, 49–52]. Ideally, one needs a verification protocol that has the following properties. It should be:

1. Calculable at large $N$, in practical timescales.
2. Measurable at large $N$, in practical timescales.
3. Able to distinguish the required distribution from other ones.
4. Applicable to all possible transmission unitaries.
5. Unable to be readily mimicked.
6. Part of a well-defined progression of tests.

We emphasize that satisfying every possible desirable requirement for assessment may not be feasible. The requirement to differentiate the desired permanent squared distribution from all other distributions is still an open question, and will not be treated here in detail for brevity. Hence, the last condition is based on the assumption that no single test will be conclusive.

Since the exact permanents are exponentially hard to compute, a test that relies on knowing the exact permanent for any sub-unitary will not satisfy the first criterion. Verification methods that rely on calculating single permanents exactly have a limited applicability. They may work well at small $N$ values. Eventually, exponential complexity will make such methods impractical. Therefore, other methods must be found.

Even our quantum simulation techniques cannot reliably predict the permanent squared with high enough accuracy for all unitaries. At large $N$, it is really only

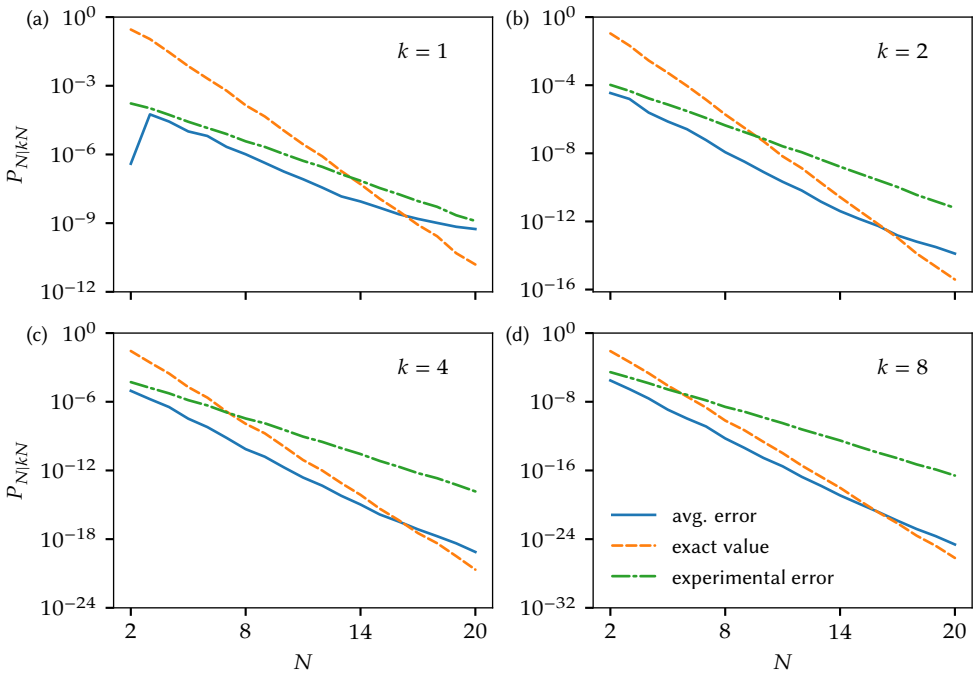


Figure 5. Estimation of the error for unitary-averaged coincidence rate $\bar{P}_{N|kN}$ as a function of $N$ using the QCP method with a random discrete phase $d \to \infty$. For each point we have used $N_{\rm m} = 100$ random unitary matrices and $L_2 = 100$ ensembles of $L_1 = 10^5$ samples each. Dashed orange line corresponds to the average for the exact value of $P_{N|kN}$. Solid blue line is the average error $\epsilon$, compared to the exact value. Dash-dotted green line denotes the estimated experimental error $\sqrt{\langle P\rangle_{\rm e,m}/L}$. **This is always greater than the simulation error, for the same number of samples.** Here we consider: (a) $k = 1$, (b) $k = 2$, (c) $k = 6$ and (d) $k = 10$.

useful for permanents that have large values. However, it is the low experimental count rate that is the main limiting factor. The methods described here can compute any permanent that is measurable with *better* than the experimental error. An ideal solution is to have an assessment signature that is both measurable and calculable at large $N$ values.

We turn to this challenging question in this section.

### A. Combined correlations

To obtain a quantity that is both measurable and calculable for typical unitaries at large $N$, we use the fact that collecting data from a single combination of $N$ output channels is extremely inefficient. Almost all the output information is lost with this procedure. Yet both the experimental data and the simulation data can provide parallel information about all channel combinations simultaneously, utilizing far more of the available information.

Hence, we first consider the *combined* correlation $C_{N|M} \equiv \langle \hat{C}_{N|M}\rangle$, defined as the sum over all different channel combinations $\sigma$ of length $N$:

$$\hat{C}_{N|M} = \sum_{\sigma} \prod_{j\in\sigma}^{N} \hat{n}_j. \tag{5.1}$$

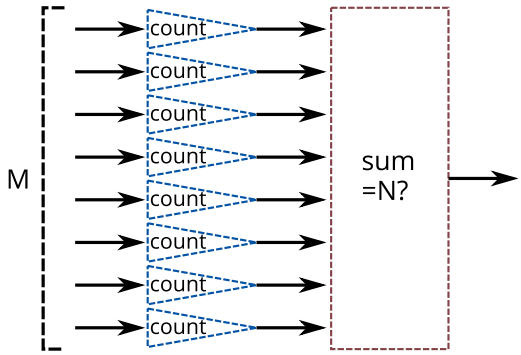


Figure 6. A schematic of the combined correlations strategy. Photo-detectors on output channels are indicated by triangles. Counter outputs must be binary or passed through a step function filter. Thus, channel counts are always 0 or 1, even for multiphoton events, which give total counts less than $N$. As a result, the final output counts all events with exactly $N$ photons in $N$ distinct channels. This sums over all possible channel combinations.

These sums combine an exponentially large number of permanents, each of which is exponentially hard to compute. Despite this, they can be evaluated efficiently with the QCP method using a modification of this technique

that employs a discrete Fourier transform (DFT); see the Appendix for details. While Gurvits type binary methods could also be used in combination with the DFT approach, the sampling error is much greater than with a qudit or continuous sampling, in addition to the systematic error problem already identified with such methods.

This combined channel, continuous sampling method provides an exceptionally efficient route to the randomized calculation of *all* the exponentially large number of $N$-th order correlations, each involving an exponentially large permanent calculation. The advantage of using $C_{N|M}$ as the assessment signature is that it has high count rates even at large $N$ values, and therefore is scalable. This can be measured experimentally with high average count rates, especially in the important large $k$ regime [46, 47]. A possible experimental realization would be to attach a detector to every output of the Boson Sampling device, triggered if it detects any photons. The outputs of these detectors are joined in a correlator which registers an event when exactly $N$ of the detectors are triggered. See Fig. 6 for an illustration of the strategy.

Naturally, this is still not exact for finite resources, due to sampling errors both in the experiment measurements and theoretical estimates. This leads to a problem: for large $N$, $C_{N|M}$ for any unitary is almost always given by its unitary average. This is known as the self-averaging property of a large random unitary.

Although satisfying the other criteria, $C_{N|M}$ therefore doesn't adequately distinguish between the different unitaries, due to sampling errors in the measurement process, combined with self-averaging. A fraudulent boson sampling device could be constructed to approximately replicate the required statistics, without having to process any information about the unitary.

### B. Channel-deletion verification

An improved strategy is needed in order to discriminate between different unitary matrices, and obtain a unique signature of the required permanent distribution. Therefore, to assess the boson sampling device, we now consider a hierarchy of experiments in which one or more channels is deleted from the channel combinations. This leads to channel-deleted, combined correlation $C^{(p)}_{N|M}$, which sum over all $N$-fold correlations that *don't* include a specific channel $p$, as illustrated in Fig. 7:

$$\hat{C}^{(p)}_{N|M} = \sum_{\sigma, p\notin\sigma} \prod_{j\in\sigma}^{N} \hat{n}_j. \tag{5.2}$$

In this approach we measure the combined permanents conditioned on channel $p$ having no counts. Similarly, one can have a hierarchy of measurements $C^{(\rho)}_{N|M}$, which are conditioned on channels in the set $\rho = \{p_1, p_2, \ldots\}$ having no counts. Eventually this exhaustively enumerates all measurable $N$-th order correlations, starting with the most readily measured combinations having the highest count rates and the lowest experimental errors.

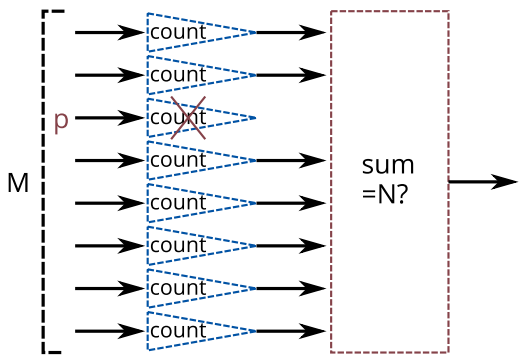


Figure 7. A schematic of the channel deleted combined correlations strategy. One or more channels are deleted randomly by switching off their counters. Therefore, only the events with $N$ counts in $N$ distinct channels, excluding deleted channels, are counted. This gives a high total count rate. At the same time, it provides a successively more detailed fingerprint of the unitary. This becomes effectively unique as the number of deleted channels is increased.

The advantage of this approach is that the goal of a boson sampling device is to generate samples with a permanent distribution. However, any probability distribution over a finite range has a unique fingerprint [53]: the set of all its moments. The hierarchy of channel-deleted combinations converges to this unique signature, in the limit in which all possible deletions are included.

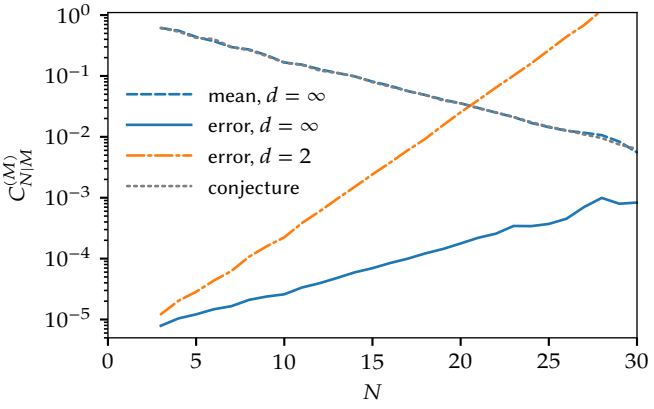


Figure 8. Combined correlations $C^{(M)}_{N|M}$ given in Eq. (5.3) evaluated using the QCP method. Here we have used $d \to \infty$, $k = 6$, $N_{\rm m} = 1$ and $L_2 = 10^4$ ensembles of $L_1 = 10^6$ samples. The dotted grey line is the analytical result given in Eq. (5.4), where $N = M/k$. The dashed blue line corresponds to our numerical results. The solid blue line is the average estimate of the error in the mean $E_S = \langle\sqrt{\langle C^2\rangle - \langle C\rangle^2}\rangle_{\rm m}/\sqrt{L_2}$. The dash-dotted orange line is the estimate of the error for the case of $d = 2$.

Calculation with complex-P distributions is straight-

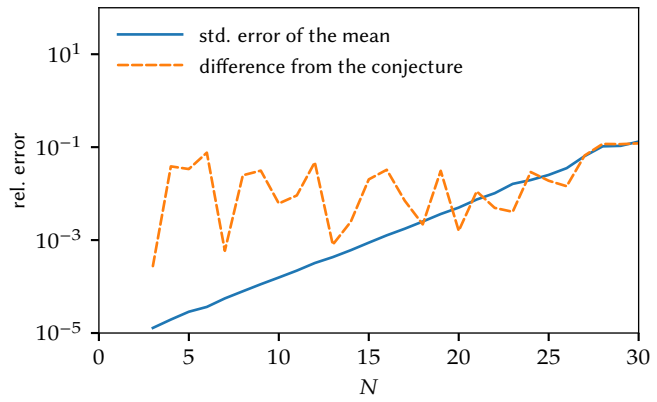


Figure 9. Modulus of the difference between the numerically obtained combined correlation, and the conjecture $C^{(M),\mathrm{conj}}_{N|M}$ (dashed orange line) plotted against the error estimate from Fig. 8, normalized on $C^{(M),\mathrm{conj}}_{N|M}$ (solid blue line). All other parameters as in Fig. 8.

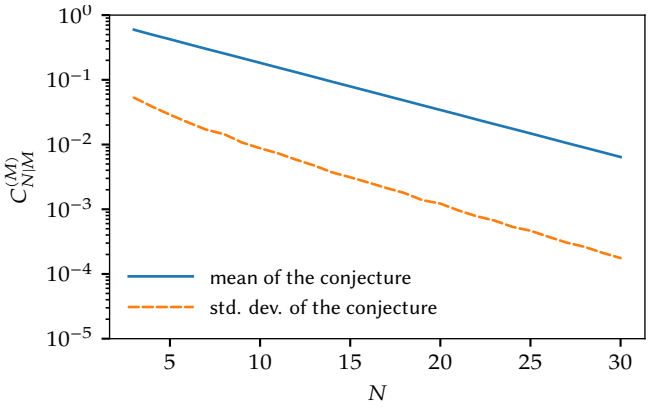


Figure 10. Standard deviation of the analytic conjecture $C^{(M),\mathrm{conj}}_{N|M}$ for $10^4$ random matrices (dashed orange line) plotted against its mean value (solid blue line), showing that there is a substantial variation between unitaries even for relatively large $N$. All other parameters as in Fig. 8.

forward. The specified channel becomes a loss reservoir for the remaining $M-1$ channels. This modified correlation can be readily evaluated using the QCP method, since its expectation value can be calculated probabilistically as:

$$C^{(p)}_{N|M} = \left\langle \frac{1}{M-1} \sum_j e^{-\mathrm{i}jk\delta} \prod_{k=1}^{M-1} \left[1 + \mathrm{e}^{\mathrm{i}j\delta} m_k\right] \right\rangle. \quad (5.3)$$

Here $m_k$ is a scaled boson number. We note, however, that even this technique has too large a sampling error for large enough $N$ values. Eventually, it will become essential to obtain an analytically calculable signature to verify boson sampling for the large $N$ case, which is important from the complexity theory viewpoint.

For this modified correlation with a single channel-deletion we give a physical argument that allows to conjecture an analytical form at large $N$. In this case the selected channel $p$ acts as a loss reservoir for the other $M-1$ channels. Therefore, we can write $\alpha^{(\mathrm{out})} = U\alpha^{(\mathrm{in})}$, $\beta^{(\mathrm{out})} = \beta^{(\mathrm{in})}U^\dagger$ so that

$$\left\langle n_j^{(\mathrm{out})} \right\rangle = \left\langle U\alpha^{(\mathrm{in})}\beta^{(\mathrm{in})}U^\dagger \right\rangle_{jj} = \sum_{i=1}^{N} |U_{ij}|^2 = T_j,$$

and $\sum_{j=1}^{M} T_j = N$. Next, we exclude the counts in channel $p$, since this acts as a loss reservoir. From unitarity, $\sum_{j=1}^{M} T_j = Nt$ , and we get:

$$\sum_{j\neq p}^{M} T_j = Nt - T_p = Nt\,(1 - T_p/Nt) = t_p N,$$

where the effective channel loss rate for channel $p$ is $t_p = t\left(1 - \sum_{i=1}^{N} |U_{pi}|^2 / N\right)$ .

For large $k \equiv M/N$, the modified correlation — which is the probability of $N$ coincidences in any of the remaining $M-p$ channels — is given by the sum of permanents of the $N \times N$ sub-matrices of the remaining $(M-1)\times(M-1)$ matrix. We conjecture that this is given asymptotically by taking an $M-1$ dimensional unitary average, together with an additional reduction of $t_p^N$ due to channel loss, in channel $p$, so that:

$$C^{(p),\mathrm{conj}}_{N|kN} \underset{k\to\infty}{\sim} \frac{t_p^N (kN-1)!(kN-2)!}{((k-1)N-1)!((k+1)N-2)!}, \quad (5.4)$$

We have tested this result using our simulation methods, that allow us to sum over exponentially large numbers — up to $10^{34}$ — of large permanents in parallel. In Figs. 8 and 9 we show numerical results for the channel combinations $C^{(p)}_{N|M}$ using the QCP method as well as for the conjecture $C^{(p),\mathrm{conj}}_{N|M}$. This remarkable task of summing over exponentially large numbers of quantities, each of which is exponentially difficult to compute, demonstrates the versatility of the simulation techniques used here.

Having established that the conjecture gives asymptotically correct results — with the exception of a set of unitaries of measure close to zero, such as unitaries near to the identity — one might inquire whether it is able to distinguish between the different random unitaries. This is investigated in Fig. 10, which plots the mean and standard deviation of the calculated last-channel-deleted count rate, $C^{(M),\mathrm{conj}}_{N|M}$, for 28 randomly chosen sets of $10^4$ unitaries with $k = 6$. The largest unitaries considered are $180\times180$. This graph shows that, while the standard deviation of the count-rate due to the randomness of the unitaries is less than the rate itself, the deviation from one unitary to the next is far from being negligible.

Our numerical simulations show that the conjecture is asymptotically valid at large $k$ for a single unitary, apart from anomalous matrices with zero asymptotic measure, such as the unit matrix. It is in excellent agreement with our quantum simulations of a random unitary, to a relative accuracy of order 1% for $k = 6$. This signature is

purely analytic, and calculable for large sizes. It is experimentally accessible, and can distinguish unitaries. It can be generalized recursively to allow increasing numbers of channel deletions, giving an increasingly unique fingerprint of the relevant permanent squared distribution, so that it will eventually distinguish any unitaries. Of course, such distinction is only possible if the unitaries are sufficiently distant relative to the sampling error.

There have been many different approaches to assess boson sampling experiments. We must start by ruling out techniques that require an exact calculation of the permanent for an arbitrary matrix. This is a #P hard computation and hence cannot be carried out for large $N$ with known algorithms. Measurement of the permanent within multiplicative error is similarly hard (on average), due to exponential suppression of count rates.

There are interesting methods that distinguish the boson sampling network from an uniform sampler [16, 49]. Other tests are based on considering either distinguishable or indistinguishable photons on the boson sampler [16, 49, 50] or distinguishing particles based on their statistics [31, 54].

In [54], an assessment protocol using the generalized bunching effect for bosons was proposed. In this case, for indistinguishable photons an absolute maximum of the probability of detecting all input particles in some output modes exists. This assessment protocol has the advantage that is detectable in a polynomial number of runs for an arbitrary network, and can distinguish a boson sampler from some other distributions. However, it is not proven to distinguish boson sampling from all other possible output distributions, and so does not satisfy criterion 3. This appears to be a fundamental problem shared by all tests that are scalable, and would imply that a wide variety of tests is desirable, rather than any single test.

Many-particle interference can also be used as an assessment for boson sampling [31] using a mode correlator, which contains the statistical properties of the particles. Using random matrix theory it is possible to obtain analytical result for the first three moments for certain conditions on the sub-matrix that contains the information of the coupling between the input and output modes. We note that such low order moments are not used in our assessment protocol, which instead relies on high-order moments that are directly relevant to the boson sampling problem. Recent experiments that consider if the input photons are distinguishable or indistinguishable photons have been performed based on machine learning [55] or Sylvester matrices [56]. Generally speaking, these methods also do not seem to distinguish different unitaries, although this is a nontrivial requirement.

In [29] a method based on symmetric sampling matrices was proposed. This distinguishes a boson sampler from another device as a mean field sampler. The test relies on verifying the suppression law for Fourier matrices [57]. It has the advantage that the proposal can be carried out for any $N$, which makes it extremely powerful. There is no doubt that this is a very sensitive test. Yet, as it does not allow assessment for any other unitary, it cannot satisfy criterion 4.

Our proposed method has the advantage that it can, in principle, treat any type of matrices, not just Fourier-transform unitary ones, since it does not rely on the zero-transmission law that uses the Fourier transform. Our proposal treats N-th order correlations, and does not rely on a generalized average bunching effect. In this regard, while still not necessarily able to eliminate all other possible distributions, it test for features in the distribution that were not addressed in previous proposals.

Finally, we note that counting rates are rather crucial at large $N$ values. Efficient nanowire detectors with $t = 0.9$ (using an optimistic estimate) and 1 GHz measurement rates [58] could allow one to reach high count rates. Coincidence rates as high as one per second even for $N = 100$ and $k = 10$, are not impossible, making these tests possible for large networks. Even very large boson sampling experiments beyond the limit of exact computability for permanents are potentially viable for our test. Such large $N$values would give an example of a quantum computation of a random bitstream that is inaccessible with digital computers. It is important to note that our complex phase-space methods do **not** replicate the random bitstream of coincidence counts, so they cannot replace the quantum network. Their role is to enable the design and assessment of the device.

## VI. CONCLUSION

In summary, we obtain novel quantum simulation methods based on complex phase-space P-distributions. These can be used for the efficient simulation of completely general linear photonic networks, using samples of integrals over coherent states. Our method is applicable to arbitrary inputs, losses and outputs. The generation of each sample is possible within polynomial time in $N$ and $M$. By contrast, classical generation of random photon number counts takes an exponentially large time in $N$ for each sample, as we expect from known fundamental complexity properties of boson sampling. Thus, our sampling technique is exponentially faster than generating photon counts.

We also calculate how the sampling error in an estimated counting probability scales with the number of samples. Here there is an unexpected result. Rather than giving an exponentially slower method than experiment, as one would obtain using classical random bit generation, the algorithm is exponentially *faster* than experiment. This is because of the properties of the coherent basis. Weighted contour integral sampling of moments has a much lower error at large $N$ values than estimating moments from numbers of photon counts. This makes possible the prediction of any photonic correlation, with less error than an experimental measurement taking a similar time.

In a very large network — the target scenario for boson sampling — the limitation of any assessment protocol is the low experimental count rate in any given set of channels. This is especially an issue when one is no longer able to efficiently compute permanents using standard methods. Accordingly, we propose a combined channel grouping protocol that allows one to assess large-scale boson sampling experiments. This uses a channel deletion protocol to distinguish different unitaries with relatively low experimental sampling error. The number of deleted channels can be recursively increased for more and more fine-grained distinction between the possible distributions. In such cases, one could then claim validation at successively more challenging levels of channel deletion.

We emphasize that our quantum simulation results are still limited by sampling errors. This is a strong limitation, even with the relatively efficient methods we use. However, these sampling errors are unbiased, and are generally much less than with other techniques carried out with the same computational time requirement. In addition, since the analytic test we propose is a conjecture, its limitations need to be investigated by further studies using random matrix theory. Application of the simulations to dissipation and noise in boson sampling interferometry will be given elsewhere. Finally, an interesting open question is the applicability of these methods to other input states.

## ACKNOWLEDGMENTS

This research was supported in part by the National Science Foundation under Grant No. NSF PHY-1125915, and by the Australian Research Council.

### Appendix A: Combined correlation calculations

We wish to use the QCP complex phase-space method in order to evaluate the sums of combined correlations $C_{N|M} \equiv \langle \hat{C}_{N|M} \rangle$ shown in Fig. 6, given by:

$$\hat{C}_{N|M} = \sum_{\sigma \in S_N} \prod_{j \in \sigma} \hat{n}_j,$$

where $S_N$ is the set of combinations of length $N$ of integers from 1 to $M$ ($\{1, 2, \ldots, N-1, N\}$, $\{1, 2, \ldots, N-1, N+1\}$, ..., $\{M-N+1, M-N+2, \ldots M-1, M\}$). This set corresponds to all different $N$-th order correlations of the output channels. The expectation value of this operator combines an exponentially large number of permanents of $N \times N$ sub-matrices of an $M \times M$ unitary matrix, *each* of which is exponentially hard to compute.

We first consider the following correlation polynomial, defined for $j = 0, \ldots M-1$:

$$\hat{D}_j = \prod_{k=1}^{M} \left[1 + \mathrm{e}^{\mathrm{i}j\delta} \hat{n}_k\right],$$

with $\delta \equiv 2\pi/M$. There are $M$ distinct correlation polynomials for $j = 0, \ldots M-1$, each including all possible combinations of $\hat{n}_k$. Their Fourier transform is given by:

$$\hat{F}_k = \frac{1}{M} \sum_{j=1}^{M} \mathrm{e}^{-\mathrm{i}jk\delta} \hat{D}_j.$$

Here $\hat{D}_j$ includes all possible multinomials in $\hat{n}_k$. For $N-$th order multinomial terms, the phase factors cancel, giving us the required correlation:

$$\hat{F}_N \equiv \hat{C}_{N|M}.$$

Using the results of Section III:

$$\begin{aligned}
\langle \hat{C}_{N|M} \rangle &= \frac{1}{M} \sum_{j=1}^{M} \mathrm{e}^{-\mathrm{i}jN\delta} \langle \hat{D}_j \rangle \\
&= \frac{1}{M} \sum_{j=1}^{M} \mathrm{e}^{-\mathrm{i}jN\delta} \langle \prod_{k=1}^{M} \left[1 + \mathrm{e}^{\mathrm{i}j\delta} \hat{n}_k\right] \rangle \\
&= \frac{1}{M} \sum_{j=1}^{M} \mathrm{e}^{-\mathrm{i}jN\delta} \langle \prod_{k=1}^{M} \left[1 + \mathrm{e}^{\mathrm{i}j\delta} n_k^{(o)}(\boldsymbol{q}, \tilde{\boldsymbol{q}})\right] \rangle_P
\end{aligned}$$

where $n_k^{(out)}$ is given by Eq (3.7) as:

$$n_k^{(out)} = r^2 \left( \sum_{s \in \sigma} T_{ks} z^{q_s} \right) \left( \sum_{t \in \sigma} T_{kt} z^{\tilde{q}_t} \right)^*$$

for an $M \times M$ transmission matrix $\boldsymbol{T}$, and single-photon inputs in a set of distinct channels $\sigma \in S_N$. Here $z = e^{2\pi i/d}$ and $q_s$ and $\tilde{q}_t$ are random integers uniformly distributed in the range $[0, d-1]$. The $P$-function used in the averaging $\langle \rangle_P$ is the product of the factors (3.9) for one-photon inputs and ones for zero-photon inputs:

$$P(\boldsymbol{q}, \tilde{\boldsymbol{q}}) = \prod_{m=1}^{N} \frac{1}{r^{2N}} z^{\tilde{q}_m} z^{-q_m}$$

Denoting $\alpha_k = \sum_{s \in \sigma} T_{ks} z^{q_s}$, $\beta_k^* = \sum_{j \in \sigma} T_{kt} z^{\tilde{q}_t}$, and introducing a scaled boson number $m_k = \alpha_k \beta_k$, and a combined unit random $y_m = z^{\tilde{q}_m} z^{-q_m}$, we get for the combined correlation:

$$\begin{aligned}
\langle \hat{C}_{N|M} \rangle = \frac{1}{M} \sum_{j=1}^{M} \mathrm{e}^{-\mathrm{i}jN\delta} \times \\
\langle \left( \prod_{m=1}^{N} \frac{1}{r^{2N}} y_m \right) \left( \prod_{k=1}^{M} \left[1 + \mathrm{e}^{\mathrm{i}j\delta} r^2 m_k\right] \right) \rangle.
\end{aligned}$$

Since all the terms with more or less than $N$ of $m_k$ factors are eliminated by the Fourier transform technique, the only ones that are left have the $r^{2N}$ multiplier that annihilates the $1/r^{2N}$ of the P-function:

$$\langle \hat{C}_{N|M} \rangle = \frac{1}{M} \sum_{j=1}^{M} \mathrm{e}^{-\mathrm{i}jN\delta} \langle \left( \prod_{m=1}^{N} y_m \right) \left( \prod_{k=1}^{M} \left[ 1 + \mathrm{e}^{\mathrm{i}j\delta} m_k \right] \right) \rangle.$$

## Appendix B: Combined correlations with excluded channels

We now consider how to distinguish unitaries. Define a count rate $C^{(\rho)}_{N|M}$, such that we definitely have no count in the set $\rho$ consisting of $Q$ different channels, and we don't care where the other counts are except that there are $N$ single counts in total. Define $S^{(\rho)}_N$ as the set of all $N$ channel combinations that exclude the channels from the set $\rho$.

Similarly to the previous section, we have

$$\hat{C}^{(\rho)}_{N|M} = \sum_{\sigma \in S^{(\rho)}_N} \prod_{j \in \sigma} \hat{n}_j.$$

Consider the modified correlation polynomial, defined for $j = 0, \ldots M - Q - 1$:

$$\hat{D}^{(\rho)}_j = \prod_{k=1 \ldots M,\ k \notin \rho} \left[ 1 + \mathrm{e}^{\mathrm{i}j\delta_Q} \hat{n}_k \right],$$

with $\delta_Q \equiv 2\pi / (M - Q)$. The Fourier transform is given by:

$$\hat{F}^{(\rho)}_k = \frac{1}{M - Q} \sum_{j=1}^{M-Q} \mathrm{e}^{-\mathrm{i}jk\delta_Q} \hat{D}^{(\rho)}_j.$$

And as before, for $N-$th order multinomial terms, the phase factors cancel, giving us the required correlation:

$$\hat{F}^{(\rho)}_N \equiv \hat{C}^{(\rho)}_{N|M}.$$

Similarly to the previous section, the probabilistic formula for the expectation value of $\hat{C}^{(\rho)}_{N|M}$ can be found to be

$$\langle \hat{C}^{(\rho)}_{N|M} \rangle = \frac{1}{M - Q} \sum_{j=1}^{M-Q} \mathrm{e}^{-\mathrm{i}jN\delta_Q} \times \langle \left( \prod_{m=1}^{N} y_m \right) \left( \prod_{k=1 \ldots M,\ k \notin \rho} \left[ 1 + \mathrm{e}^{\mathrm{i}j\delta_Q} m_k \right] \right) \rangle.$$


[1] R. J. Glauber, Phys. Rev. **130**, 2529 (1963).
[2] R. J. Glauber, Phys. Rev. **131**, 2766 (1963).
[3] S. Aaronson and A. Arkhipov, in *Proceedings of the 43rd Annual ACM Symposium on Theory of Computing* (ACM Press, 2011) pp. 333–342.
[4] S. Aaronson and A. Arkhipov, Theory of Computing **9**, 143 (2013).
[5] K. R. Motes *et al.*, Phys. Rev. Lett. **114**, 170802 (2015).
[6] P. D. Drummond and C. W. Gardiner, J. Phys. A **13**, 2353 (1980).
[7] J. Carolan *et al.*, Science **349**, 711 (2015).
[8] S. Aaronson, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences **467**, 3393 (2011).
[9] P. Clifford and R. Clifford, arXiv:1706.01260 (2017).
[10] Z.-E. Su, Y. Li, P. P. Rohde, H.-L. Huang, X.-L. Wang, L. Li, N.-L. Liu, J. P. Dowling, C.-Y. Lu, and J.-W. Pan, Phys. Rev. Lett. **119**, 080502 (2017).
[11] M. A. Broome *et al.*, Science **339**, 794 (2013).
[12] A. Crespi *et al.*, Nat. Photon. **7**, 545 (2013).
[13] M. Tillmann *et al.*, Nat. Photon. **7**, 540 (2013).
[14] J. B. Spring *et al.*, Science **339**, 798 (2013).
[15] A. Crespi *et al.*, Nat. Commun. **7**, 10469 (2016).
[16] N. Spagnolo *et al.*, Nat. Photon. **8**, 615 (2014).
[17] H. Wang *et al.*, Nat. Photon. **11**, 361 (2017).
[18] J. C. Loredo, M. A. Broome, P. Hilaire, O. Gazzano, I. Sagnes, A. Lemaitre, M. P. Almeida, P. Senellart, and A. G. White, Phys. Rev. Lett. **118**, 130503 (2017).
[19] S. Scheel, arXiv:quant-ph/0406127 (2004).
[20] S. Scheel, in *Quantum Information Processing*, edited by T. Beth and G. Leuchs (Wiley-VCH, Weinheim, 2005) Chap. 28, pp. 382–392.
[21] L. Valiant, Theoretical Computer Science **8**, 189 (1979).
[22] D. G. Glynn, European Journal of Combinatorics **31**, 1887 (2010).
[23] J. Wu, Y. Liu, B. Zhang, X. Jin, Y. Wang, H. Wang, and X. Yang, ArXiv e-prints (2016), arXiv:1606.05836.
[24] B. Opanchuk, L. Rosales-Zárate, M. D. Reid, and P. D. Drummond, arXiv:1711.07153 (2017).
[25] Y. He *et al.*, Phys. Rev. Lett. **118**, 190501 (2017).
[26] J. P. Olson, K. R. Motes, P. M. Birchall, N. M. Studer, M. LaBorde, T. Moulder, P. P. Rohde, and J. P. Dowling, Phys. Rev. A **96**, 013810 (2017).
[27] P. A. M. Dirac, Rev. Mod. Phys. **17**, 195 (1945).
[28] P. D. Drummond and M. Hillery, *The Quantum Theory of Nonlinear Optics* (Cambridge University Press, 2014).
[29] M. C. Tichy, K. Mayer, A. Buchleitner, and K. Mølmer, Phys. Rev. Lett. **113**, 020502 (2014).
[30] K. Mayer, M. C. Tichy, F. Mintert, T. Konrad, and A. Buchleitner, Phys. Rev. A **83**, 062307 (2011).

[31] M. Walschaers *et al.*, New Journal of Physics **18**, 032001 (2016).
[32] P. Drummond and D. Walls, Journal of Physics A: Mathematical and General **13**, 725 (1980).
[33] N. Bartolo, F. Minganti, W. Casteels, and C. Ciuti, Phys. Rev. A **94**, 033841 (2016).
[34] F. Minganti, N. Bartolo, J. Lolli, W. Casteels, and C. Ciuti, Scientific Reports **6**, 26987 (2016).
[35] P. D. Drummond, C. W. Gardiner, and D. F. Walls, Phys. Rev. A **24**, 914 (1981).
[36] P. Drummond, K. McNeil, and D. Walls, J. Mod. Opt. **28**, 211 (1981).
[37] K. McNeil and C. Gardiner, Physical Review A **28**, 1560 (1983).
[38] B. Cao, K. W. Mahmud, and M. Hafezi, Phys. Rev. A **94**, 063805 (2016).
[39] P. Drummond, J. Phys. A **50**, 45LT01 (2017).
[40] P. D. Drummond and M. D. Reid, Phys. Rev. A **94**, 063851 (2016).
[41] L. Gurvits, in *Mathematical Foundations of Computer Science 2005*, edited by J. Jędrzejowicz and A. Szepietowski (Springer, Berlin, Heidelberg, 2005) pp. 447–458.
[42] S. Perlmutter, M. Levenson, R. Shelby, and M. Weissman, Physical review letters **61**, 1388 (1988).
[43] R. Shelby, P. Drummond, and S. Carter, Physical Review A **42**, 2966 (1990).
[44] S. J. Carter and P. D. Drummond, Phys. Rev. Lett. **67**, 3757 (1991).
[45] P. D. Drummond and J. F. Corney, J. Opt. Soc. Am. B **18**, 139 (2001).
[46] A. Arkhipov and G. Kuperberg, Geometry & Topology Monographs **18**, 1 (2012).
[47] P. D. Drummond, B. Opanchuk, L. Rosales-Zárate, M. D. Reid, and P. J. Forrester, Phys. Rev. A **94**, 042339 (2016).
[48] L. Gurvits and A. Samorodnitsky, Discrete & Computational Geometry **27**, 531 (2002).
[49] J. Carolan *et al.*, Nat. Photon. **8**, 621 (2014).
[50] S. Aaronson and A. Arkhipov, Quantum Info. Comput. **14**, 1383 (2014).
[51] M. Bentivegna *et al.*, Science Advances **1**, e1400255 (2015).
[52] L. Aolita, C. Gogolin, M. Kliesch, and J. Eisert, Nat. Commun. **6**, 8498 (2015).
[53] P. A. P. Moran, *An introduction to Probability Theory* (Oxford, Clarendon Press, 1968).
[54] V. S. Shchesnovich, Phys. Rev. Lett. **116**, 123601 (2016).
[55] I. Agresti, N. Viggianiello, F. Flamini, N. Spagnolo, A. Crespi, R. Osellame, N. Wiebe, and F. Sciarrino, arXiv:1712.06863 (2017).
[56] N. Viggianiello, F. Flamini, M. Bentivegna, N. Spagnolo, A. Crespi, D. J. Brod, E. F. Galvão, R. Osellame, and F. Sciarrino, arXiv:1710.03578v2 (2017).
[57] M. C. Tichy, M. Tiersch, F. de Melo, F. Mintert, and A. Buchleitner, Phys. Rev. Lett. **104**, 220405 (2010).
[58] H. Takesue, S. D. Dyer, M. J. Stevens, V. Verma, R. P. Mirin, and S. W. Nam, Optica **2**, 832 (2015).