

Matrix Phase-Space Representations for Gaussian Boson Sampling

Peter D. Drummond^{✉,*}, Alexander S. Dellios[✉], and Margaret D. Reid

Centre for Quantum Science and Technology Theory, Swinburne University of Technology, Melbourne 3122, Australia



(Received 2 June 2025; accepted 10 December 2025; published 6 January 2026)

We introduce coherent matrix phase-space distributions. These use conservation laws and symmetries to improve the accuracy and speed of quantum phase-space representations. As an example, this is applied to the validation of low-loss Gaussian boson sampling (GBS) quantum computational advantage experiments, where classical generation of the random photon-number counts is exponentially hard. Large improvements in sampling errors are demonstrated compared to previous methods. Matrix phase-space representations also provide a large numerical speedup due to their (at worst) quadratic scaling, compared to other methods for validating total count probabilities of large-scale, low-loss GBS networks.

DOI: [10.1103/htk1-rn8v](https://doi.org/10.1103/htk1-rn8v)

Introduction—Phase-space representations have a long history in quantum theory [1,2] and were applied to quantum optics using the normally ordered Glauber-Sudarshan P representation [3,4]. Here, we develop matrix phase-space representations that include symmetries in the basis, thereby leveraging symmetry theory to make phase-space methods more efficient. The resulting representation is complete, and it has greatly reduced sampling errors when symmetries exist. This uses superpositions of coherent Schrödinger cat states [5], allowing one to include conserved quantities. We treat a phase symmetry in this Letter, but the method holds for more general cases. Our approach unifies normally ordered positive-P(+P) [6] and gauge-P representations [7] with stochastic wave-function methods [8].

To illustrate this, a simulation of moments and grouped probabilities for low-loss Gaussian boson sampling (GBS) quantum computers (QCs) with photon-number-resolving (PNR) detectors [9–11] is demonstrated for up to 8000 modes. GBS is a nonuniversal optical QC that sends squeezed photons into a large interferometer. Output photons are measured, making a physical random number generator (RNG) whose probabilities would require computing the #P-hard Hafnian matrix function [11]. For large systems, exact simulation of the counts is not feasible due to exponentially growing computation times and rounding errors [12,13].

An essential task is to validate the output data of GBS in an efficient and scalable way. Algorithms that replicate the

GBS task [14] are not scalable to large sized networks, and checking large-scale validity is necessary for quantum advantage (QA) claims [15]. Computers must be both fast and accurate. We show that matrix-P representations projecting conserved parity can validate large-scale GBS experiments in the low-loss limit, where earlier phase-space methods are impractical. Our approach complements studies of unitary ensembles [16–19] and anticoncentration [18,19] by computing necessary, but not sufficient, tests of validity similar to other RNG validation tests [20].

The example treated is a calculation of binned count probabilities. Other methods using loop Hafnian generator functions [21] are dominated by up to quartic scaling for the 10^3 – 10^4 mode networks of current interest. For these, matrix-P run times scale nearly linearly with mode number, allowing much faster computation of count probabilities. The matrix-P representation can also be used for other tests [15,16,22] in parallel. We investigate scalability up to 8000 modes. This size was reached in recent experiments [23], although not using PNR, and is used in related QA proposals [18].

We compare our results to the +P method, which can represent any nonclassical state [6]. It has found many applications [24–32], including GBS experiments in current intermediate loss regimes with threshold [15,33,34] and PNR [35] detectors. The End Matter shows +P converges too slowly for practical application to large, low-loss PNR experiments.

Matrix-P representations—A phase-space representation [36,37] with a basis $\hat{\Lambda}(\vec{\alpha})$ on can be extended to a matrix representation by including a complete set of $\mathcal{M}$ symmetry projectors, $\hat{\Pi}_p$, such that $\sum \hat{\Pi}_p = \hat{1}$, where $p = 0, \dots, \mathcal{M} - 1$ indexes the global eigenvalues. The quantum density matrix $\hat{\rho}$ is expanded in probabilities $P_{\mathcal{M}}(\lambda)$ of coherent matrix projectors $\hat{\Lambda}_{pq}(\vec{\alpha}) = \hat{\Pi}_p \hat{\Lambda}(\vec{\alpha}) \hat{\Pi}_q$, so that

*Contact author: peterddrummond@protonmail.com

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

$$\hat{\rho} = \int P_{\mathcal{M}}(\lambda) \text{tr}_{\mathcal{M}}[\hat{\Lambda}(\vec{\alpha})\Omega] d\lambda. \quad (1)$$

Here, $\lambda = (\vec{\alpha}, \Omega)$, $d\lambda \equiv d\vec{\alpha}d\Omega$, while Ω is an $\mathcal{M} \times \mathcal{M}$ effective stochastic density matrix, and $\text{tr}_{\mathcal{M}}$ is the matrix trace. The method treated here starts with the +P distribution, $P_+(\alpha, \beta)$, with $\vec{\alpha} = (\alpha, \beta)$, which is positive and always exists. The matrix-P representation expands in projected coherent states $|\alpha\rangle_p = \hat{\Pi}_p|\alpha\rangle$ [38], so that the matrix kernel $\hat{\Lambda}$ is

$$\hat{\Lambda}_{pq}(\vec{\alpha}) = |\alpha\rangle_p \langle \beta^*|_q e^{-w_{pq}(\vec{\alpha})}, \quad (2)$$

where $|\alpha\rangle = \exp(\alpha \cdot a^\dagger)|0\rangle$ is an unnormalized coherent state [3], with M bosonic creation operators $a_k^\dagger$ and an amplitude vector α . We choose that $w_{pq}(\vec{\alpha}) = \log\left(\sqrt{g_p(\vec{\alpha})g_q(\vec{\alpha})}\right)$, where $g_p(\vec{\alpha}) \equiv \langle \beta^*|\hat{\Pi}_p|\alpha\rangle$, so that $\hat{\Lambda}$ has unit quantum trace and $\langle \Omega \rangle$ is a density matrix. One can also use other normalizations [8], and the phase variables α, β can depend on the eigenvalues.

The expansion of Eq. (1) unifies and extends previous phase-space representations [3,6,8,39]. To prove that our expansion exists for all density matrices, we use both completeness of the positive-P expansion and the projectors $\hat{\Pi}_p$. Defining $n = \alpha \cdot \beta$, one can show that at least one positive matrix distribution exists such that

$$P_{\mathcal{M}}(\vec{\alpha}) = P_+(\alpha, \beta) \prod_{pq} \delta(\Omega_{qp} - e^{w_{pq}-n}). \quad (3)$$

This is not unique, because coherent states are not orthogonal. The advantage of this approach is to distinguish the $\mathcal{M}$ global eigenvalues in each symmetry, which are finite and scalable, from local, mode-dependent fluctuations in the exponentially large Hilbert space basis for large mode numbers. Such local quantum effects are described best by the coherent amplitudes in phase space that can then be randomly sampled efficiently.

Coherent state projections—We utilize projectors that generate Schrödinger cat states [38,40–43]. We define them to have $\mathcal{M}$ coherent amplitudes $U^q\alpha$, where $q = 0, \dots, \mathcal{M}-1$, and U is a unitary matrix that is an $\mathcal{M}$ th root of unity and defines the symmetry. We treat phase symmetry with $U^q = \exp(iq\phi)$ and $\phi = 2\pi/\mathcal{M}$ in the results below. Other, more general cases with different symmetries will be treated elsewhere.

The basis vector, $|\alpha\rangle \equiv [|\alpha\rangle_0, \dots, |\alpha\rangle_{\mathcal{M}-1}]^T$, is

$$|\alpha\rangle_p = \frac{1}{\mathcal{M}} \sum_{q=0}^{\mathcal{M}-1} e^{-ipq\phi} |U^q\alpha\rangle. \quad (4)$$

Each cat state has a different set of photon numbers. If the maximum photon number is $N_{\max} < \mathcal{M}$, the state $|\alpha\rangle_p$ has a fixed global photon number p .

For dynamics, we define derivatives $\partial_j = \partial/\partial\alpha_j$, together with global annihilation and creation matrices, σ and $\sigma^\dagger$, where $\sigma_{p,q} = \delta_{p,q+1}^{(\mathcal{M})}$ such that $\delta_{p,q}^{(\mathcal{M})}$ is a cyclic Kronecker delta (modulo $\mathcal{M}$). This leads to the general operator identities for any $\mathcal{M}$:

$$\begin{aligned} a_j|\alpha\rangle &= \alpha_j\sigma|\alpha\rangle \\ a_j^\dagger|\alpha\rangle &= \partial_j\sigma^\dagger|\alpha\rangle. \end{aligned} \quad (5)$$

In the simplest case of $\mathcal{M} = 2$, both matrices are a Pauli matrix, $\sigma = \sigma^\dagger = \sigma^x$. Operator products give similar identities, for example, $a_i^\dagger a_j|\alpha\rangle = \alpha_j\partial_i|\alpha\rangle$.

Differential identities for the kernel $\hat{\Lambda}$ require the introduction of a diagonal matrix T , where

$$T_{pp'} = \delta_{pp'} T_p = \delta_{pp'} g_{p-1}/g_p. \quad (6)$$

The indices are cyclic, so $g_{-1} = g_{\mathcal{M}-1}$. A super-matrix $\mathcal{T}\hat{\Lambda} \equiv (T\hat{\Lambda} + \hat{\Lambda}T)/2$ is also needed for the identities. Applying Eq. (5), one obtains

$$\begin{aligned} a_i^\dagger a_j \hat{\Lambda} &= \alpha_j [\partial_i + \beta_i T] \hat{\Lambda} \\ \hat{\Lambda} a_i^\dagger a_j &= \beta_i [\partial_j + \alpha_j T] \hat{\Lambda}, \end{aligned} \quad (7)$$

where we have defined $\tilde{\partial}_j \equiv \partial/\partial\beta_j$. The above identities can also be used to calculate expectation values of arbitrary observables $\hat{O}$ corresponding to c-number matrix functions $O(\vec{\alpha})$, such that

$$\langle \hat{O} \rangle = \int P_{\mathcal{M}}(\vec{\alpha}) \text{tr}_{\mathcal{M}}[O(\vec{\alpha})\Omega] d\vec{\alpha}. \quad (8)$$

In the simplest parity representation with $\mathcal{M} = 2$, $T_0 = \tanh(n)$, $T_1 = \coth(n)$.

Applications to GBS—To demonstrate the advantages of matrix-P in a topical case, we analyze the validation of lossless and low-loss GBS with PNR detectors using the approach of multiple statistical tests [20]. In these regimes, it is classically hard to generate the output samples for large mode numbers [12,14,44]. We show that in both lossless and lossy cases, very large-scale validation is feasible due to greatly reduced sampling errors. We treat one- and two-dimensional count binning, but other, low-order tests should be used as well [35]. The resulting distributions in Fig. 1 give excellent agreement with exact results in soluble cases.

Since squeezed states are always generated in correlated photon pairs, the GBS examples treat a parity symmetry with $\mathcal{M} = 2$, using two-component cat states, $|\alpha\rangle_p = \frac{1}{2} [|\alpha\rangle + (-1)^p |-\alpha\rangle]$, so $|U^q\alpha\rangle = |(-1)^q\alpha\rangle = |\alpha\rangle_0 + (-1)^q |\alpha\rangle_1$. Opposite parity states are orthogonal, because even parity ($p = 0$) states have only even photon

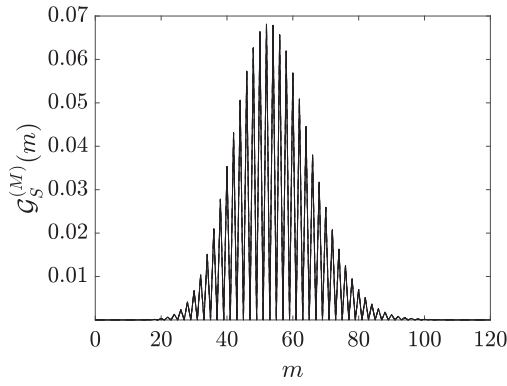


FIG. 1. Lossless matrix-P simulations of $\mathcal{G}$, the total photon-count probability in GBS, versus counts m . These are indistinguishable from the exact distribution for uniform pure squeezed states with $M = 200$, $r = 0.5$, and a Haar random unitary network. Errors are $< 6 \times 10^{-5}$ and are not visible. An even larger size of $M = 10^4$ gives an error of $\sim 6 \times 10^{-6}$, showing excellent scalability.

numbers and odd parity ($p = 1$) states have odd photon numbers.

To calculate GBS photonic dynamics, we treat a linear network with a Hamiltonian $\hat{H} = \hbar \omega_{ij} a_i^\dagger a_j$. The equations are identical to those in the positive P representation, with

$$\partial_t P_{\mathcal{M}} = i \omega_{ij} [\partial_i \alpha_j - \tilde{\partial}_i \beta_j] P_{\mathcal{M}}. \quad (9)$$

This equation is solved using characteristics so that

$$[\alpha(t), \beta(t)] = [e^{-i\omega t} \alpha(0), e^{i\omega t} \beta(0)]. \quad (10)$$

For a linear network, amplitude vectors are multiplied by a Haar random unitary matrix U_H , just as in the +P solutions [33]. Each coherent component evolves independently of the others, as first pointed out by Schrödinger [45]. Damping is treated in detail below.

Next, take the GBS case, with squeezed vacuum input states $|\mathbf{r}\rangle$. These are an integral over coherent states [46], using an even function of α defined as $\xi(\alpha) = \prod_j C_j e^{-\alpha_j^2 \coth(r_j)/2}$, with $C_j = [2\pi \sinh(r_j)]^{-1/2}$, for real $r_j > 0$. A squeezed state, $|\mathbf{r}\rangle$, is an integral over a real line, $|\mathbf{r}\rangle = \int \xi(\alpha) |\alpha\rangle d^M \alpha$. The +P distribution for pure squeezed states is defined on a $2M$ -dimensional real subspace [15]. For multimode squeezed states $|\mathbf{r}\rangle \langle \mathbf{r}|$ with an $M = 2$ parity symmetric matrix-P method, one can improve convergence enormously by noting that $\xi(\alpha)$ is even, so that one can replace $|\alpha\rangle$ by the cat state $|\alpha\rangle_0$:

$$\hat{\rho} = \int \xi(\alpha) \xi(\beta) |\alpha\rangle_0 \langle \beta|_0 \delta[\text{Im}(\vec{\alpha})] d\vec{\alpha}. \quad (11)$$

The resulting matrix-P distribution, $P_2(\vec{\alpha})$, has $\Omega_{00} = 1$ and is even in n . The +P sampling procedure [33] for initial stochastic trajectories can still be used. The resulting

output photon number sample is obtained from matrix transformations $\mathbf{t}$ of the input amplitudes, where $n' = \sum_i n'_i = \sum_i \alpha'_i \beta'_i$ such that $\alpha' = \mathbf{t}\alpha$, $\beta' = \mathbf{t}^* \beta$ are the transformed amplitudes of a linear network, as one would expect classically.

Photon-counting observables—Using Eq. (7) and taking the density matrix trace with the number operator, $\hat{n}_j = a_j^\dagger a_j$, gives the result, on defining $p_m = [1 - (-1)^m]/2 = [0, 1]$, that photon-counting moments are given by

$$\langle : \hat{n}_j^m : \rangle = \int (\alpha_j \beta_j)^m \text{tr}_{\mathcal{M}} [T^{p_m} \Omega] P_{\mathcal{M}}(\vec{\lambda}) d\vec{\lambda}. \quad (12)$$

Apart from the weights, the even powers give expressions identical to the +P distribution. The observables of interest for PNR detectors are the d th dimensional probability of binned number patterns [35],

$$\mathcal{G}_S^{(\tilde{n})}(\mathbf{m}) = \left\langle \prod_{j=1}^d \left[\sum_{c_j=m_j} \hat{G}(\mathbf{c}) \right] \right\rangle, \quad (13)$$

where the projection operator for an M -mode output photon number pattern $\mathbf{c} = [c_1, \dots, c_M]$, where $c_i = 0, 1, 2, \dots$ are observed photon counts in the set S , is

$$\hat{G}(\mathbf{c}) = \bigotimes_{i \in S} \frac{1}{c_i!} : \hat{n}_i^{c_i} e^{-\hat{n}_i} :. \quad (14)$$

Here, $m_j = \sum_{i \in S_j} c_i$ is the $j = 1, \dots, d$ th photon count bin, $\tilde{n} = \sum_{j=1}^d M_j \leq M$ is the correlation order, and $S = (S_1, \dots, S_d)$ is a vector of subsets of modes.

We first treat the total count probability, which arises when $d = 1$, $\tilde{n} = M$, and $S = \{1, \dots, M\}$, such that

$$\mathcal{G}_S^{(M)}(m) = \left\langle \sum_{c_i=m} \hat{G}(\mathbf{c}) \right\rangle. \quad (15)$$

In the case of lossless GBS with uniform pure squeezed state inputs, an exact M -mode total count probability distribution is known [47,48]:

$$\mathcal{G}_S^{(M)}(m) = \begin{cases} \left(\frac{M+m}{2} - 1 \right) p_s^{M/2} (1-p_s)^{m/2} & \text{even } m \\ 0 & \text{odd } m \end{cases}, \quad (16)$$

where $p_s = 1/(1 + \bar{n})$ is the success probability for detecting a photon with $\bar{n} = \sinh^2(r)$ mean photons per mode. With uniform photon loss, an exact M -mode probability distribution for uniform input squeezing also exists [12,35], which provides a useful analytic test case. This shows the validity of the method, which can readily model more realistic, experimental networks.

The projection operator, Eq. (14), is rewritten in terms of odd and even powers,

$$\hat{G}(\mathbf{c}) = [C(\hat{n}_S) - S(\hat{n}_S)] \bigotimes_{i \in S} \frac{\hat{n}_i^{c_i}}{c_i!}, \quad (17)$$

where $\hat{n}_S = \sum_{i \in S} \hat{n}_i$, and we have abbreviated $C \equiv \cosh$, $S \equiv \sinh$. This result also holds if the counts $\hat{n}_i$ come from binned groups of channels. Using the matrix-P method, we find that the c-number function corresponding to $\hat{G}(\mathbf{c})$ for parity-symmetric inputs is

$$G(\mathbf{m}) = g_0(n_S, \mathbf{m}) \prod_{i \in S} \frac{n_i^{c_i}}{c_i!} e^{-n_i}, \quad (18)$$

where, defining parity as $\pi_m \equiv 1 - 2 * \text{rem}(m, 2) = \pm 1$,

$$g_0(n_S, m) = \left(\frac{1 + \pi_m e^{2(n_S - m)}}{1 + e^{-2n_S}} \right). \quad (19)$$

For the case that all channels are counted in the set S , so that $n = n_S$, the c-number pattern projector for odd counts becomes $g_0(2m+1) = 0$, as expected, while the c-number pattern projector for even counts is $g_0(2m) = n^{2m}/[C(n)(2m)!]$.

Computational results and losses—Calculations using the matrix-P distribution are given in Fig. 1 for a 200-mode lossless GBS and compared to the exact distribution [Eq. (16)]. Results with $E_S = 1.2 \times 10^6$ samples agree with the exact distribution with errors of $< 5 \times 10^{-5}$, in agreement with central limit theorem sampling error estimates [49]. Extensions to an 8,000-mode network, as in current threshold detection experiments [23], give similar relative errors. All simulations were carried out with a publicly available MATLAB software package, XQSIM [50].

We now compare this approach with the positive-P method, which is useful for lossy cases and click detection [15,22,34]. To simulate total counts with the +P distribution, one uses Eq. (18), except with a unit prefactor, $g_0 = 1$. For a sample number of $E_S = 1.2 \times 10^6$, the +P results in the lossless PNR case have extremely large sampling errors due to a tail-heavy distribution of trajectories, as shown in the End Matter.

Losses are modeled by adding reservoir channels that are not excited or counted, so that the network matrix becomes a lossy transmission matrix $\mathbf{t}$. We compare simulations for this lossy case with exact results in Fig. 2 for the previous 200-mode network, with an intensity loss of 1%. The even-odd count oscillations are retained, but with reduced amplitude. These simulations use 10^3 subensembles of 10^4 trajectories each. Typical probability errors are 5×10^{-6} , indicating an excellent fit to the exact, lossy distribution. For these ensemble sizes, +P simulations have

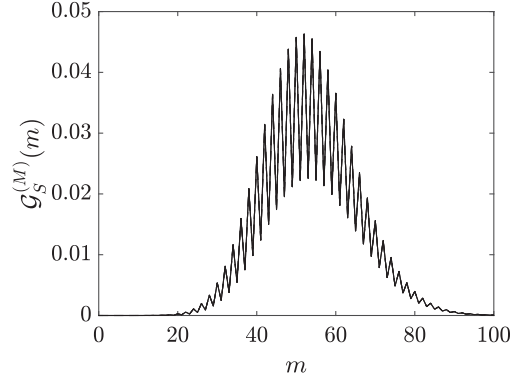


FIG. 2. Simulations with 1% loss; otherwise as in Fig. 1.

large sampling errors whenever the even-odd oscillations are present, becoming accurate if parity conservation is violated with large losses [35] or if there are only small mode numbers. The matrix-P method also gives correct results with large losses, but there is less computational benefit in this regime.

Another alternative, the loop Hafnian generator method [21], agrees with our results for total counts but has at least cubic or quartic scaling. We show in Appendix B that matrix-P timing scales nearly linearly with mode number up to $M = 8000$, making it faster by orders of magnitude for experiments with $M > 1000$.

Higher-dimensional binning—Matrix-P methods can also model higher-dimensional binning ($d \geq 2$), where the outputs are divided into multiple groups, denoted by the vector $\mathbf{S}$ in Eq. (13), giving many independent validation tests. When $\tilde{n} = M$, the photon-count distribution becomes a joint probability distribution over d count bins $G_S^{(M)}(m_1, \dots, m_d)$. Unlike the $d = 1$ case, if the network matrix is a Haar random unitary, no elementary higher-dimensional binning result is known, but the matrix-P method is still applicable.

In Fig. 3, we simulate a $d = 2$ dimensional binning of a lossless 200-mode network using the matrix-P distribution with $E_S = 1.2 \times 10^6$. The even-odd oscillations are still present, with nonzero probabilities occurring only if $m_1 + m_2$ is even. The sampling errors remain small at $< 3 \times 10^{-5}$. Loop Hafnian numerical results are not yet extended to higher-dimensional binning for timing comparisons.

Conclusion—Rapid convergence is obtained for matrix-P simulations of GBS photon-count distributions at large mode numbers. This will allow precise validation of photon-counting data in future large-scale, low-loss GBS quantum computing experiments, which are predicted to have quantum advantage [18]. Improvements occur because the method projects the physically relevant part of the Hilbert space, allowing one to include any global symmetries. This method allows other symmetries to be included, and it has potential applicability in many

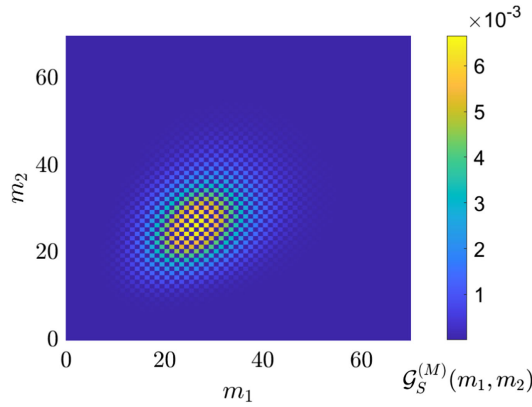


FIG. 3. Surface plot of matrix-P simulated $d = 2$ dimensional binned photon-count probability $\mathcal{G}_{S_1, S_2}^{(M)}(m_1, m_2)$, with $S_1 = \{1, \dots, 100\}$ and $S_2 = \{101, \dots, 200\}$, for pure squeezed states into a lossless GBS network. All parameters as in Fig. 1. Sampling errors are $< 3 \times 10^{-5}$.

disparate areas, such as Bell violations [29], ultracold atoms in optical lattices [8], soliton propagation [51–53], and others.

Acknowledgments—We thank Javier Martínez-Cifuentes and coauthors for their Hafnian code. This publication was made possible through an NTT Phi Laboratories grant and support of Grant No. 62843 from the John Templeton Foundation.

The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

Data availability—The data that support the findings of this article are openly available [50].

[1] E. Wigner, *Phys. Rev.* **40**, 749 (1932).
 [2] K. Husimi, *Proc. Phys. Math. Soc. Jpn.* **22**, 264 (1940).
 [3] R. J. Glauber, *Phys. Rev.* **131**, 2766 (1963).
 [4] E. C. G. Sudarshan, *Phys. Rev. Lett.* **10**, 277 (1963).
 [5] E. Schrödinger, *Naturwissenschaften* **23**, 823 (1935).
 [6] P. D. Drummond and C. W. Gardiner, *J. Phys. A* **13**, 2353 (1980).
 [7] P. Deuar and P. D. Drummond, *Phys. Rev. A* **66**, 033812 (2002).
 [8] I. Carusotto, Y. Castin, and J. Dalibard, *Phys. Rev. A* **63**, 023606 (2001).
 [9] S. Aaronson, *Proc. R. Soc. A* **467**, 3393 (2011).
 [10] S. Aaronson and A. Arkhipov, *Theory Comput.* **9**, 143 (2013).
 [11] C. S. Hamilton, R. Kruse, L. Sansoni, S. Barkhofen, C. Silberhorn, and I. Jex, *Phys. Rev. Lett.* **119**, 170501 (2017).
 [12] A. Deshpande, A. Mehta, T. Vincent, N. Quesada, M. Hinsche, M. Ioannou, L. Madsen, J. Lavoie, H. Qi, J.

Eisert, D. Hangleiter, B. Fefferman, and I. Dhand, *Sci. Adv.* **8**, eabi7894 (2022).
 [13] A. Björklund, B. Gupt, and N. Quesada, *J. Exp. Algorithmics (JEA)* **24**, 1 (2019).
 [14] J. F. F. Bulmer, B. A. Bell, R. S. Chadwick, A. E. Jones, D. Moise, A. Rigazzi, J. Thorbecke, U.-U. Haus, T. Van Vaerenbergh, R. B. Patel, I. A. Walmsley, and A. Laing, *Sci. Adv.* **8**, eabl9236 (2022).
 [15] P. D. Drummond, B. Opanchuk, A. Dellios, and M. D. Reid, *Phys. Rev. A* **105**, 012427 (2022).
 [16] P. D. Drummond, B. Opanchuk, L. Rosales-Zárate, M. D. Reid, and P. J. Forrester, *Phys. Rev. A* **94**, 042339 (2016).
 [17] J. Martínez-Cifuentes, H. de Guise, and N. Quesada, *PRX Quantum* **5**, 040312 (2024).
 [18] A. Ehrenberg, J. T. Iosue, A. Deshpande, D. Hangleiter, and A. V. Gorshkov, *Phys. Rev. Lett.* **134**, 140601 (2025).
 [19] A. Ehrenberg, J. T. Iosue, A. Deshpande, D. Hangleiter, and A. V. Gorshkov, *Phys. Rev. A* **111**, 042412 (2025).
 [20] A. Rukhin, J. Soto, J. Nechvatal, M. Smid, E. Barker, S. Leigh, M. Levenson, M. Vangel, D. Banks, A. Heckert *et al.*, NIST Special Publication 800-22rev1a (2010), <http://csrc.nist.gov/publications/nistpubs/800-22-rev1a/SP800-22rev1a.pdf>.
 [21] J. F. Bulmer, J. Martínez-Cifuentes, B. A. Bell, and N. Quesada, *arXiv:2412.17742*.
 [22] P. D. Drummond and B. Opanchuk, *Phys. Rev. Res.* **2**, 033304 (2020).
 [23] H.-L. Liu, H. Su, S.-Q. Gong, Y.-C. Gu, H.-Y. Tang, M.-H. Jia, Q. Wei, Y. Song, D. Wang, M. Zheng *et al.*, *arXiv:2508.09092*.
 [24] S. J. Carter, P. D. Drummond, M. D. Reid, and R. M. Shelby, *Phys. Rev. Lett.* **58**, 1841 (1987).
 [25] W. J. Munro and M. D. Reid, *Phys. Rev. A* **47**, 4412 (1993).
 [26] P. Deuar and P. D. Drummond, *Phys. Rev. Lett.* **98**, 120402 (2007).
 [27] S. Kiesewetter, R. Y. Teh, P. D. Drummond, and M. D. Reid, *Phys. Rev. Lett.* **119**, 023601 (2017).
 [28] S. Kiesewetter, Q. Y. He, P. D. Drummond, and M. D. Reid, *Phys. Rev. A* **90**, 043805 (2014).
 [29] L. Rosales-Zárate, B. Opanchuk, P. D. Drummond, and M. D. Reid, *Phys. Rev. A* **90**, 022109 (2014).
 [30] K. Takata, A. Marandi, and Y. Yamamoto, *Phys. Rev. A* **92**, 043821 (2015).
 [31] P. D. Drummond and S. Chaturvedi, *Phys. Scr.* **91**, 073007 (2016).
 [32] R. Y. Teh, S. Kiesewetter, P. D. Drummond, and M. D. Reid, *Phys. Rev. A* **98**, 063814 (2018).
 [33] A. Dellios, P. D. Drummond, B. Opanchuk, R. Y. Teh, and M. D. Reid, *Phys. Lett. A* **429**, 127911 (2022).
 [34] A. S. Dellios, B. Opanchuk, N. Goodman, M. D. Reid, and P. D. Drummond, *Phys. Lett. A* **549**, 130529 (2025).
 [35] A. S. Dellios, M. D. Reid, and P. D. Drummond, *Quantum Sci. Technol.* **10**, 045030 (2025).
 [36] J. F. Corney and P. D. Drummond, *Phys. Rev. A* **68**, 063822 (2003).
 [37] J. F. Corney and P. D. Drummond, *Phys. Rev. Lett.* **93**, 260401 (2004).
 [38] P. D. Drummond and M. D. Reid, *Phys. Rev. A* **94**, 063851 (2016).

- [39] P. Deuar and P. D. Drummond, *J. Phys. A* **39**, 2723 (2006).
 [40] B. Yurke and D. Stoler, *Phys. Rev. Lett.* **57**, 13 (1986).
 [41] M. Wolinsky and H. J. Carmichael, *Phys. Rev. Lett.* **60**, 1836 (1988).
 [42] L. Krippner, W. J. Munro, and M. D. Reid, *Phys. Rev. A* **50**, 4330 (1994).
 [43] L. Gravina, F. Minganti, and V. Savona, *PRX Quantum* **4**, 020337 (2023).
 [44] S. Aaronson and A. Arkhipov, *Quantum Inf. Comput.* **14**, 1383 (2014).
 [45] E. Schrödinger, *Naturwissenschaften* **14**, 664 (1926).
 [46] P. Adam, I. Földesi, and J. Janszky, *Phys. Rev. A* **49**, 1281 (1994).
 [47] J. Huang and P. Kumar, *Phys. Rev. A* **40**, 1670 (1989).
 [48] C. Zhu and C. M. Caves, *Phys. Rev. A* **42**, 6794 (1990).
 [49] B. Opanchuk, L. Rosales-Zárate, M. D. Reid, and P. D. Drummond, *Phys. Rev. A* **97**, 042304 (2018).
 [50] All phase-space simulations were performed using the publicly available software package, XQSIM, which is written in the MATLAB programming language, P. D. Drummond, A. S. Dellios, and N. Goodman, XQSIM3, <https://github.com/peterddrummond/xqsim> (2025).
 [51] S. J. Carter, P. D. Drummond, M. D. Reid, and R. M. Shelby, *Phys. Rev. Lett.* **58**, 1841 (1987).
 [52] P. D. Drummond, R. M. Shelby, S. R. Friberg, and Y. Yamamoto, *Nature (London)* **365**, 307 (1993).
 [53] K. L. Ng, B. Opanchuk, M. D. Reid, and P. D. Drummond, *Phys. Rev. Lett.* **122**, 203604 (2019).

End Matter

Appendix A: Probability densities of observables— Here, we examine the distribution of stochastic observables for the +P and matrix-P phase-space methods.

The effect that conservation laws have on convergence of any observable can be obtained from the probability density of the stochastic trajectories for a specific observable. Trajectory amplitudes for any phase-space observable $O(\alpha, \beta)$, with corresponding operator $\hat{O}$, are binned into N_b equally spaced bins with spacing $\Delta = (O_{\max} - O_{\min})/N_b$ to estimate the probability density as

$$P(O_j) = \frac{1}{N_S \Delta_b} \sum_{i=1}^{N_R} \sum_{k=1}^{N_S} N_j^{(i,k)}. \quad (\text{A1})$$

The sum is over the number of stochastic trajectories $N_j^{(i,k)}$ contained in the j th bin. To improve sampling error estimates [15], we use N_R subensembles, each with N_S trajectories, giving an ensemble size of $E_S = N_S N_R$.

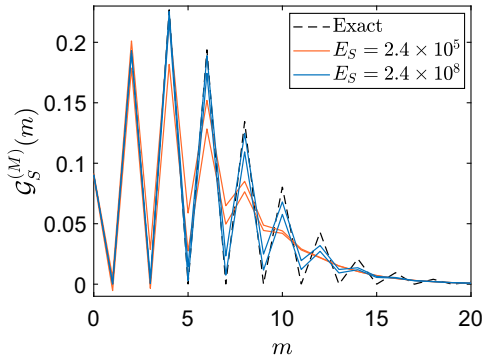


FIG. 4. Comparisons of +P simulations of the total count distribution versus the exact multimode photon-counting distribution (dashed black line) for lossless GBS with uniform $r = 0.5$ and $M = 20$. +P moments are averages over $E_S = 2.4 \times 10^5$ (solid orange line) and 2.4×10^8 (solid blue line) samples. Upper and lower lines correspond to $\pm 1\sigma_{T,j}$ estimated sampling errors.

Our simulations in the main text used $N_R = 120$ subensembles of $N_S = 10^4$ trajectories each, giving maximum errors of $< 6 \times 10^{-5}$ for the matrix-P photon-number probabilities, which are stochastic observables. The +P method gave errors up to ~ 1000 times larger or more. To show the discrepancy, simulation results with $r = 0.5$ and $M = 20$ are given in Fig. 4. In the lossless case with a Haar unitary, +P simulations have not yet converged to the exact distribution, even with $E_S = 2.4 \times 10^8$ samples.

To explain this, we computed the probability densities of stochastic observables at different counts m of the +P phase-space observable $G_+(m)$ for a photon-number probability in a PNR type GBS experiment, using

$$G_+(m) = \frac{1}{m!} (n)^m e^{-n}. \quad (\text{A2})$$

This demonstrates that the results in Fig. 4 occur because in the lossless +P case, the probability densities are highly skewed with long tails, as shown in Fig. 5, where the $G_+(4)$ amplitudes have positive tails, while $G_+(3)$ amplitudes have negative tails. The largest probabilities in both cases are within the range $G_+ \in [0, 0.5]$ (see inset of Fig. 5), giving an overlap between the odd and even densities.

The projector Eq. (A2) uses in the +P simulations is unbounded due to the exponential in Eq. (A2). From the inset of the distribution in Fig. 5, the $G_+(4)$ values with the highest probability correspond to observables with small amplitudes. However, large positive trajectories also occur with small probabilities. To obtain correct results, these tails must be sampled, requiring enormous numbers of samples. The number of samples required for good accuracy grows rapidly as $M \rightarrow \infty$, and even for the $M = 200$ simulations in the main text, it is far too large to be practical.

By comparison, projecting onto the smaller Hilbert space of even photon numbers makes matrix-P computation far

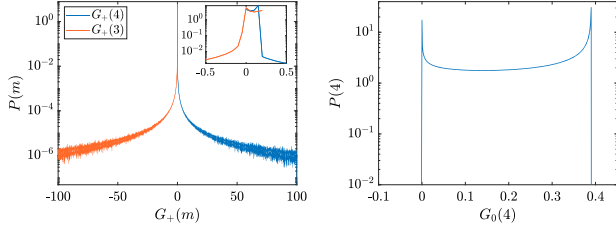


FIG. 5. Left plot: Logarithmic plot of binned +P probability densities $P(m)$ versus observable $G_+(m)$ for $m = 4$ (solid blue line) and $m = 3$ (solid orange line). Trajectories are from simulations of a lossless $M = 20$ mode GBS network with $E_S = 2.4 \times 10^8$, using bins of size $\Delta_b = 0.05$. The inset shows the range $[-0.5, 0.5]$, where the odd and even probability densities overlap. Right plot: Logarithmic plot of binned matrix-P probability densities $P(4)$ versus observable $G_0(4)$ for the same trajectory and mode count. Bins of size $\Delta_b = 0.001$ are used to estimate probabilities, with no probability tails.

more efficient. Due to the parity representation expansion, odd counts are always zero in lossless cases. In the right-hand panel of Fig. 5, we bin matrix-P amplitudes of the observable $G_0(4)$ for the total photon count distribution of the $M = 20$ mode network. Although accurate convergence is obtained for smaller ensembles, we use $E_S = 2.4 \times 10^8$ for comparison with the +P density estimates.

The matrix-P distribution does not have the large tails present in the +P case. Instead, the distribution is highly localized and bimodal. In these calculations, the matrix-P projector is within the range $(0, 0.4)$ for $m = 4$ and is bounded for any mode number. This bimodal matrix-P distribution has its largest probabilities at $G_0(4) \approx 0.39$ and $G_0(4) \approx 0$, with no extended tails.

In summary, once projected onto a smaller Hilbert space, the number of samples required is reduced enormously, because the range of values each trajectory can take is much smaller. This improvement is not just for lossless cases, and the method is applicable to general states. One must check that there are no other relevant conservation laws, since these may also cause enhanced sampling errors if they lead to heavy tailed distributions, but these too can be eliminated with projections.

Appendix B: Numerical scaling properties—Another method of computing binned count probabilities proposed by Bulmer *et al.* [21] utilizes the loop Hafnian generator function, the f_m function. This useful development gives results in agreement with ours. Here, we compare run times (T) versus mode numbers, up to $M = 8192$, of matrix-P and the f_m function. Scaling properties of the maximum matrix-P sampling errors, $\max(\sigma_T)$, are also investigated when losses and nonuniform squeezing is present. Results are shown in Fig. 6.

Matrix phase-space methods scale generally as $T \sim C_1 E_S M + C_2 E_S M^2 + C_3 M^2$, for an overall additive error $\sim 1/\sqrt{E_S}$. For the mode numbers investigated here,

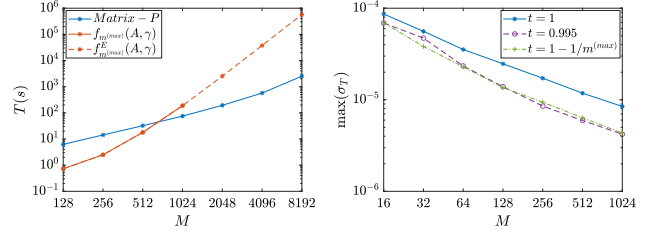


FIG. 6. Scaling properties of matrix-P total count distribution simulations using $E_S = 1.2 \times 10^6$. Left plot: Run-time T (seconds) versus mode number M of matrix-P (blue) and loop Hafnian f_m -function algorithm (orange). Lossless GBS with uniform squeezing $r = 0.89$ and a Haar unitary U_H is simulated for mode numbers $M = 128, 256, \dots, 8192$. The dashed orange line is the extrapolated f_m function run-time, denoted f_m^E . Right plot: Scaling of the maximum matrix-P sampling error $\max(\sigma_T)$ versus $M = 16, 32, \dots, 1024$ for uniform amplitude loss rates $t = 1$ (solid blue), $t = 0.995$ (dashed purple), and $t = 1 - 1/m^{(\max)}$ (dash-dot green), applied as tU_H , with nonuniform squeezing $r_i = 0.89 + 0.1w_i$ for the i th input with Gaussian random noise w_i . Here, $m^{(\max)} = 85, 118, 174, 273, 450, 776, 1388$ for each network.

the linear phase-space memory access term $\mathcal{C}_1$ dominates (see Fig. 6). As $E_S, M \rightarrow \infty$, the times should converge asymptotically to the quadratic matrix multiply time contribution, $\mathcal{C}_2$. Unlike the Hafnian approach, the same phase-space ensemble can be used efficiently to compute a large number of distinct moment or marginal tests, which are necessary for full validation.

The f_m function has run times estimated to scale as $T \sim c_1 m M^3 + c_2 m^2 \log(m)$, where m denotes the photon count bin. At the time of writing, the f_m function algorithm has been applied to total count distributions only, with one test at a time. Therefore, we treat this case for comparative timing, where the number of matrix-P samples is constant at $E_S = 1.2 \times 10^6$.

The largest observable count bin $m^{(\max)}$ is chosen to occur when $\mathcal{G}_S^{(M)}(m^{(\max)}) = 10^{-7}$. This cutoff is implemented in both the matrix-P and f_m -function algorithms. With inputs $m^{(\max)} \propto M$, as used here, this gives $T \propto M^4$, while if $m^{(\max)} \propto \sqrt{M}$ as in anticoncentration proposals [18], one has $T \propto M^{3.5}$. To optimize the matrix-P algorithm, we also define a minimum count bin $m^{(\min)}$, such that $\mathcal{G}_S^{(M)}(m^{(\min)}) = 10^{-7}$.

The left-hand panel of Fig. 6 graphs matrix-P timing for lossless GBS networks of size $M = 128, 256, \dots, 8192$ with $r = 0.89$. Both algorithms use $m^{(\max)} = 273, 450, 776, 1388, 2562, 4831, 9267$, while minimums of $m^{(\min)} = 37, 122, 318, 755, 1678, 3605, 7561$ are used for matrix-P. Simulations were performed on a 12-core 3.7 GHz AMD CPU computer with 32 GB of RAM.

The f_m -function algorithm is faster than matrix-P for networks in the range $16 \leq M \leq 512$ [21]; however, it becomes slower once $M > 512$ due to its quartic scaling in

M . Matrix-P times scale nearly linearly in M for this range, making it much more efficient for large networks.

For networks with $M > 1024$, the publicly available f_m -function algorithm gave no useful results due to floating point overflows and underflows. For this reason, some of the run times in Fig. 6 were estimated (dashed orange line) by extrapolation. Using the $M \leq 1024$ times, least squares minimization found the c_1, c_2 coefficients in the f_m scaling law, which were used to estimate times for $M = 2048, 4096, 8192$.

We now investigate the scaling of matrix-P sampling errors $\sigma_T \propto 1/\sqrt{E_S}$ in experimentally realistic networks with losses and nonuniform squeezing. Using $E_S = 1.2 \times 10^6$ samples, networks of size $M = 16, 32, \dots, 1024$ are simulated for randomly generated nonuniform squeezing parameters $r_i = 0.89 + 0.1w_i$ for the i th input mode with

Gaussian random noise w_i . Results of $\max(\sigma_T)$ are presented in the right-hand panel of Fig. 6. Uniform amplitude transmissions of $t = 1, 0.995$ are used, as well as $t = 1 - (m^{(\max)})^{-1}$, where $m^{(\max)}$ is the same as the timing simulations for each M . This loss scaling is chosen so that the even-odd oscillations are still present in the total count distributions. Matrix-P sampling errors decrease for larger networks.

In summary, quartic scaling causes the loop Hafnian f_m -function run times to increase rapidly. Matrix-P simulations of total count probabilities for $M = 8192$ are ~ 220 times faster than the estimated f_m -function run-time, and they can treat multiple tests at once. The ability to rapidly simulate large networks is vital to validation of QC outputs in future experiments [18].