


## Efficient Classical Algorithm for Boson Sampling with Partially Distinguishable Photons

J. J. Renema,<sup>\*</sup> A. Menssen, W. R. Clements, G. Triginer, W. S. Kolthammer, and I. A. Walmsley  
*Clarendon Labs, Department of Physics, Oxford University, Parks Road OX1 3PU Oxford, United Kingdom*

 (Received 10 July 2017; revised manuscript received 8 October 2017; published 30 May 2018)

We demonstrate how boson sampling with photons of partial distinguishability can be expressed in terms of interference of fewer photons. We use this observation to propose a classical algorithm to simulate the output of a boson sampler fed with photons of partial distinguishability. We find conditions for which this algorithm is efficient, which gives a lower limit on the required indistinguishability to demonstrate a quantum advantage. Under these conditions, adding more photons only polynomially increases the computational cost to simulate a boson sampling experiment.

DOI: [10.1103/PhysRevLett.120.220502](https://doi.org/10.1103/PhysRevLett.120.220502)

Boson sampling [1] provides a promising route towards demonstrating a quantum advantage, i.e., a computation by a quantum system that exceeds what is possible with a classical one. In boson sampling, the task is to provide a sample from the output of a linear transformation of optical modes, some of which are fed with single photons. For a sufficient number of photons and modes, a suitable quantum machine directly implementing this problem will outperform a realistic classical computer simulating the experiment. This result has spurred a range of experimental efforts [2–8].

A crucial challenge for computational problems based on boson sampling is to accommodate imperfections that arise in real-world devices. An essential aspect of the original proposal [1] was to show that for small deviations from the ideal machine, the achieved sampling problem retains computational hardness. However, when this bound is translated into concrete values for photon distinguishability [9] or beam-splitter deviations [10], it turns out that this results in scaling which is extremely unfavorable compared to what is achievable experimentally. There is therefore a need to devise photon sampling problems with improved error tolerance.

There exist two approaches to demarcating the line between viable and nonviable extensions of boson sampling. In the top-down approach, the original complexity proof is extended “downwards,” showing that the resulting output distribution for systems of increasing imperfection is still hard to sample from. This approach has had some success: for example, it has been shown that the issue of unreliable single photon generation can be overcome, and that the hardness of the resulting “scattershot” boson sampling problem is equivalent to that of the original boson sampling problem [11,12].

The alternate bottom-up approach is to construct new efficient classical algorithms for boson samplers with particular imperfections. This approach leads to constructive proofs that rule out a computational advantage, thereby

showing performance limits that must be exceeded by quantum machines. For example, Rahimi-Keshari *et al.* [13] used generalizations of the Wigner function as a way to construct a classical algorithm which can efficiently simulate certain lossy boson samplers.

In this work, we consider which-way information of the interfering photons as an imperfection that compromises the hardness of a boson sampler. We show a classical algorithm that efficiently approximates detection probabilities in the limit of many photons, given the photons have some partial distinguishability. We then use this algorithm to consider problems of finite size and estimate a level of indistinguishability that must be surpassed to demonstrate a quantum advantage. The basis of our algorithm, schematically depicted in Fig. 1, is that for partially distinguishable photons, the probability of a given outcome can be approximated by terms that involve fewer interfering photons, where the remaining ones do not interfere at all. For a given error tolerance and indistinguishability, we determine a number of photons above which this approach succeeds, while requiring only a polynomial increase in the computational steps as the number of photons increases further. We use this result to estimate a lower bound on the

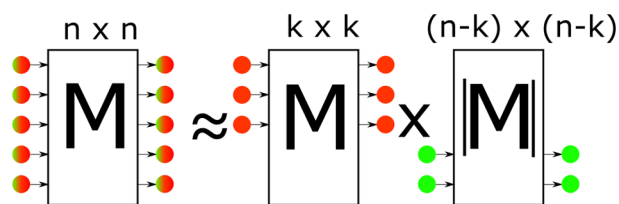


FIG. 1. A pictorial representation of our result. We show that boson sampling with  $n$  photons of partial distinguishability (represented by the mixed red-green balls) can be approximated as computing the outcome of a series of smaller permanents of size  $k$ , combined with probabilistic transmission of the remaining  $n - k$  photons. The value of  $k$  at which this approximation works is set by the value of the distinguishability.

photon quality required to demonstrate a quantum advantage: for 50 photons and an error threshold of 10%, the degree of indistinguishability of the interfering photons must be higher than 94.7%.

Multiphoton interference at partial distinguishability has been studied extensively [9,14–25]. For boson sampling with fully indistinguishable photons in separate modes, the probability of a particular detection outcome is given by  $P = |\text{Perm}(M)|^2$ , where  $M$  is a submatrix of the unitary  $U$ , where the rows and columns of  $M$  are chosen to correspond to the occupied input and output modes, respectively. In this work, we will use the formalism of Tichy [16], where the probability of a particular detection outcome (i.e., photons emerging at particular outputs) is

$$P = \sum_{\sigma \in \Sigma} \left( \prod_{j=1}^n S_{\sigma_j j} \right) \text{Perm}(M * M_{1,\sigma}^*), \quad (1)$$

where  $M$  is a submatrix constructed in the same way as for fully indistinguishable boson sampling,  $n$  the number of photons, and where  $*$  denotes the elementwise product, and  $M^*$  is the elementwise complex conjugation. Throughout this work, we will assume that  $M$  is selected from a Haar-random unitary  $U$  of size  $N$ , with  $N \gg n^2$ , so that the elements of  $M$  are independent and identically distributed (i.i.d.) complex Gaussians [1]. The notation  $M_{1,\sigma}$  indicates that the rows of  $M$  are unpermuted, and that the columns are permuted according to  $\sigma$ , and we will use this notational convention throughout. The matrix of mutual distinguishabilities  $S$  is given by  $S_{ij} = \langle \Psi_i | \Psi_j \rangle$ , where  $\Psi_i$  is the  $i$ th single-photon wave function. The set of permutations of size  $n$  is denoted  $\Sigma$ .

There are two extreme cases to note. First, if  $S_{ij} = 1$ , Eq. (1) reduces to the standard expression for boson sampling with fully indistinguishable particles. For  $S_{ij} = \delta_{ij}$ , Eq. (1) reduces to  $P = \text{Perm}(|M|^2)$ , which is the expression for boson sampling with distinguishable photons in separate modes. In this latter case, multiphoton interference is absent and the total probability is expressed in terms of single-photon transmission probabilities instead of transmission amplitudes. Since this matrix contains only positive elements, it can be evaluated to within a multiplicative error in polynomial time [26]. In our work, we will interpolate between these cases, and parametrize the mutual distinguishabilities by a single real parameter  $x$ , with

$$S_{ij}(x) = x + (1-x)\delta_{ij}. \quad (2)$$

We will argue at the end of our work and in the Supplemental Material [27] that our results apply to more general forms of  $S$ , and that  $x$  can be taken to be real without loss of generality.

The observation that underlies our work is that the degree of quantum interference in each term in Eq. (1) is determined by the number of fixed points (invariant

elements) in the corresponding permutation. Each fixed point in  $\sigma$  causes the corresponding row from  $M$  to enter into the permanent as the modulus squared, meaning that in that term, the corresponding photon does not exhibit interference. Therefore, the size of the matrix of complex elements which must be computed to evaluate each term is set by the number of noninvariant elements in the permutation [16,17]. Furthermore, terms with many fixed points have a larger weight in the sum: for each element in the permutation which is not a fixed point, the product in Eq. (1) will pick up a factor  $x < 1$  from the off-diagonal elements of  $S$  [9,15]. One can therefore construct a series of successively more accurate approximations by grouping the terms in Eq. (1) by number of fixed points, and then truncating the sum at some value, which we designate  $k$ . The resulting approximation  $P_k$  is given by

$$P_k = \sum_{j=0}^k \sum_{\sigma^j} x^j \text{Perm}(M * M_{1,\sigma}^*), \quad (3)$$

where we have introduced the notation  $\sigma^j$  to denote those permutations which have  $n-j$  elements as fixed points. To simplify Eq. (3), we note that in this permanent,  $n-j$  columns will be left unpermuted, and will therefore end up as the modulus squared of the elements. Expanding Eq. (3) in all the permuted columns and combining terms, we can separate the permanents of permuted and unpermuted columns as

$$P_k = \sum_{j=0}^k \sum_{\sigma^j} x^j \sum_{\rho} \text{Perm}(M_{\rho, I_\rho} * M_{\rho, \sigma_\rho}^*) \text{Perm}(|M_{\bar{\rho}, \sigma_u}|^2), \quad (4)$$

where  $\rho$  denotes the  $\binom{n}{j}$  possible combinations of  $j$  rows from the matrix  $M$ ,  $I_\rho$  is the identity permutation of the elements of  $\rho$ ,  $\bar{\rho}$  denotes the complementary rows, and  $\sigma_\rho$  and  $\sigma_u$  denote the permuted and unpermuted elements of  $\sigma$ , respectively.

We note that the lower the value of  $j$ , the easier the terms are to compute, since the second of these two permanents can be efficiently evaluated, and the first permanent is of a matrix of size  $j$ . The term with  $j=0$  represents the case where the photons are treated as fully distinguishable particles. The next term ( $j=2$ ) represents the first-order correction, where interference between each pair of photons is considered, and similarly for higher values of  $j$ . The sequence of  $P_k$  is therefore ordered by computational cost as well as by degree of quantum interference.

The rest of this Letter is dedicated to investigating the properties of these successive approximations. We start by investigating these approximations numerically, which will lead to some conjectures regarding the scaling behavior of these approximations with  $n$  and  $k$ , which we will then confirm through more rigorous analysis.

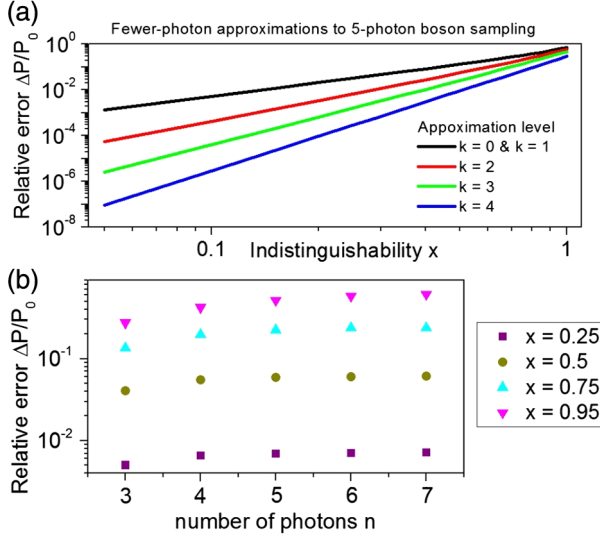


FIG. 2. Top: relative error  $\Delta P_k/P(x=0)$ , where  $P$  is the probability of observing some outcome of the boson sampler, as a function of indistinguishability, for different values of  $k$ . Bottom: scaling of the relative error with overall photon number for  $x = 0.25, 0.5, 0.75$ , and  $0.95$  (from bottom to top), at fixed  $k = 2$ .

In Fig. 2(a), we numerically investigate the quality of this approximation for a five-photon boson sampling experiment. We simulate  $10^4$  Haar-random unitary matrices of size  $N = 100$ , from which we take the first  $n = 5$  modes as input and output without loss of generality, and computed  $P_k$  for values of  $k$  from 1 to 4. Note that the case  $k = 1$  only encompasses the identity permutation ( $j = 0$ ), which corresponds to sampling with distinguishable photons. We plot the relative error of our approximation, defined as  $\Delta P_k/P_0$ , where  $\Delta P = |P_k - P|$  and  $P_0 = n!/N^n$  is the characteristic scaling in the number of photons.

Figure 2(b) shows how the relative error scales with the photon number  $n$ . We plot  $\Delta P_2/P_0$  as a function of the photon number  $n$ . We find that the relative error saturates at moderate values of  $n$ . This numerical result suggests that in the limit of large photon numbers, the accuracy of our approximation does not depend on the number of photons; i.e.,  $\Delta P_k/P_0$  is a function of  $k$  and  $x$ , but not of  $n$  if  $n$  is sufficiently large. This means that for large enough  $n$ , if our approximation  $P_k$  satisfies a given error threshold for a particular value of  $k$ , that approximation will also work for larger photon numbers. Since the size of the complex-valued permanents is given by  $k$ , this suggests that adding more photons does not significantly add to the cost of simulating the problem.

We now proceed to prove that this is indeed the case. To do this, we show two things. First, we need to show that our approximation is efficient, in the sense that adding more photons only induces a polynomial increase in the number of computations required. Second, we need to show that the intuition we obtained from the numerics above is correct,

and that the relative error on our approximation does not increase when  $n$  is increased.

We begin with the first task: counting the number of terms in Eq. (4). The number of terms in the middle of the three sums of Eq. (4) is given by the rencontres number [28,29]  $R_{n,n-j} = \binom{n}{j}!(j)$ , where  $!(j) = \lfloor j!/e \rfloor$  is the subfactorial.  $R_{n,n-j}$  is defined as counting the number of permutations which leave  $n - j$  elements invariant; it scales as  $n^j$ . The inner sum is over  $\rho$ , which as noted above scales as  $\binom{n}{j}$ , which is a quantity which also grows polynomially in  $n$ . The total number of complex-valued permanents which we are required to compute for particular value of  $j$  in the outer sum is therefore  $\binom{n}{j}^2!(j)$ . The problem therefore scales as  $n^{2j}$ , and when we truncate the terms at  $k$ , the number of terms is therefore of the order of  $n^{2k}$ . Using Ryser's algorithm [30] to evaluate each complex-valued permanent takes  $2^k k$  steps, and the whole algorithm scales as  $n^{2k} 2^k k$ , which scales polynomially in  $n$  as required, provided the choice of  $k$  needed to satisfy some error bound does not depend on  $n$ .

Therefore, we now consider the accuracy of the scheme. We write  $P$  as a polynomial in  $x$ :  $P(x) = \sum c_j x^j$ , where the coefficients are given by Eq. (4). We can therefore estimate the error  $\Delta P_k/P_0$  by computing the error term  $\sum_{j=k+1}^n c_j x^j$ . We find that if  $M$  is a matrix of independent Gaussian elements, the increase in the number of terms with  $j$  is precisely balanced out by the decrease in the magnitude of each term (see Supplemental Material [27]), and that these coefficients are given by

$$|c_j| \approx \sqrt{\left( \sum_{k=0}^{n-j} (1/k!) \right) R_{n,n-j} \binom{n}{j} j!(n-j)! / 2N^n} \approx n! / 2N^n,$$

where in the final step we have taken the limit of large  $n$ . It should be noted that this latter expression does not depend on  $j$ , and differs by a factor  $1/2$  from the expectation value of fully distinguishable boson sampling. As an illustration, Fig. 3 shows estimates of the absolute values of the coefficients  $|c_j|$ , for  $n = 8$ . These are compared against a numerical simulation on 500 submatrices of Haar-random matrices. The precise behavior of this function is discussed in the Supplemental Material [27]. Already for  $n = 8$ , some points reach the value obtained in the limit of large  $n$ .

We can now perform the required error estimate. Noting that the  $c_j$  are uncorrelated (see Supplemental Material [27]), we add the terms as a sum of squares:  $\Delta P_k/P_0 = \sqrt{\sum_{j=k+1}^n (|c_j| x^j)^2}$ . If we take the limit  $n \rightarrow \infty$ , this has the finite value

$$\Delta P_k/P_0 = x^{k+1} / 2(1-x^2)^{1/2}. \quad (5)$$

Our algorithm for simulating boson sampling with partially distinguishable photons is now as follows: given

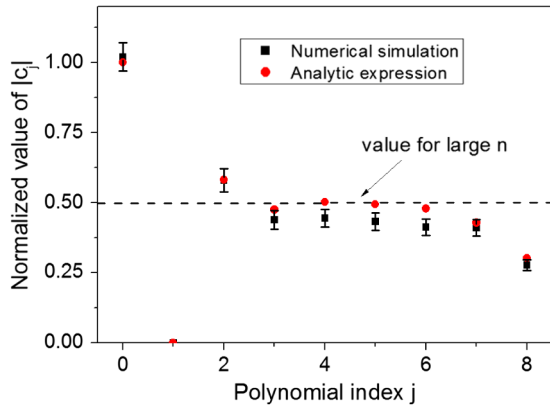


FIG. 3. Coefficients of the polynomial  $P = \sum_{j=0}^n c_j x^j$ , for  $n = 8$ , normalized to  $c_0$ . The black squares correspond to a numerical simulation of 500 random unitaries. The red circles correspond to the prediction from Eq. (5).

the desired accuracy of the simulation and the level of indistinguishability with which the experiment is performed, use Eq. (5) to evaluate the required value of  $k$ . Next, compute Eq. (3) up to the  $k$ th term, and feed the computed value of  $P_k$  into a classical sampling algorithm, such as the Metropolis algorithm [31]. Such algorithms can generate a sample from a probability distribution, even if the number of possible outcomes is large [32], which means that the number of modes does not enter into the hardness of boson sampling in general [33].

Finally, we determine the regions of the parameter space where our algorithm is efficient. The solid lines in Fig. 4 show the values of  $k$  and  $x$  for which Eq. (5) has solutions.

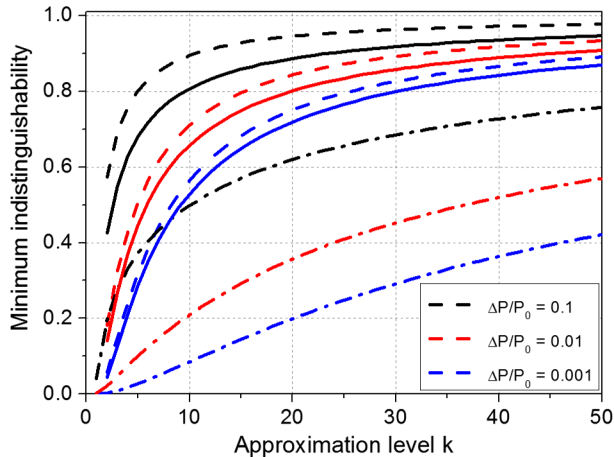


FIG. 4. Lower bound on the quality of photons required to achieve exponential scaling of computational hardness in boson sampling. The solid lines indicate solutions of Eq. (5), which demarcate the region of polynomial scaling in  $n$ . The dashed lines indicate the value of  $x$  below which  $n$ -photon interference can be expressed as interference of  $n - 1$  photons. The dash-dotted lines indicate the values of indistinguishability where  $n$ -photon interference can be described using fewer resources than required for the computation of an  $n$ -by- $n$  complex permanent.

These lines define a bound which must be exceeded for the complexity of the problem not to scale polynomially with the total number of photons. This bound holds for all  $n > k$ . State-of-the-art supercomputers can compute permanents of size 50 in approximately an hour [34], which might therefore be taken as an estimate of the number of photons required to obtain a quantum advantage. Using Eq. (5), we find that we require  $x = 0.870$ ,  $x = 0.908$ , and  $x = 0.947$  for  $P_{49}$  to be accurate to within  $\Delta P/P_0 = 0.1\%$ ,  $1\%$ , or  $10\%$ , respectively. Note that as the error tolerance increases, the degree of distinguishability that our algorithm can handle increases as well. We stress again that this result is a lower bound: achieving it experimentally is no guarantee that the experiment is not classically simulable through a more advanced algorithm.

If we relax the requirement that our algorithm must work for all  $n$ , we obtain a tighter requirement on the distinguishability. The dashed lines in Fig. 4 show a tighter bound, which holds only for  $k = n - 1$ . This line demarcates where  $n$ -photon interference can be expressed as  $n - 1$  photon interference. The area in between the dashed and solid lines is the region of parameter space where our approximation works for  $n \approx k$ , but where the approximation fails for  $n \gg k$ .

Finally, we consider the prefactor of our algorithm. The original expression from Ref. [16] requires the computation of roughly  $2^n$  permanents of size  $n$ , whereas the original boson sampling proposal requires the computation of only one such permanent. If we limit the number of computations steps we allow ourselves for computing  $P_k$  to be equal to the number of steps required to compute an  $n$ -by- $n$  permanent, we arrive at the dash-dotted curves in Fig. 4. The area in between the solid and dash-dotted curves represents the region of the parameter space where the computational cost scales polynomially in the number of photons, but where our approximation might be impractical for small values of  $n$ . Since Tichy's expression requires the computation of many very similar permanents, we expect that there is significant scope for improvement on the classical algorithm at this point, which we leave as an open problem.

Finally, we note that our results do not depend on the original parametrization of  $S$ . In particular, if one has  $S_{ij} = x_{ij} + \delta_{ij}(1 - x_{ij})$ , with  $|x_{ij}| < 1$ , and complex  $x$ , one can apply the error bound by setting  $x = \max(|x_{ij}|)$  in Eq. (4). Therefore, our results apply in the experimentally relevant case where all photons are not of equal distinguishability.

In summary, we have shown how the limited indistinguishability of photons affects the hardness of the boson sampling problem. We have presented a scheme that can express the probability of an outcome in boson sampling as a sum of smaller permanents when the photons are sufficiently distinguishable. We have demonstrated that this scheme scales polynomially in the overall photon number, while its accuracy does not depend on photon

number. We have used this scheme to estimate a lower bound on the indistinguishability required to achieve a quantum advantage.

We thank Raul Garcia-Patron and Nathan Walk for useful discussions. J.J.R. is supported by NWO Rubicon. A. M. is supported by the Buckee Scholarship from Merton College, Oxford. G. T. is supported by the Merton Scholarship Fund. W. R. C. and I. A. W. acknowledge an ERC Advanced Grant (MOQUACINO). W. S. K. is supported by EPSRC EP/M013243/1. I. A. W. acknowledges the H2020-FETPROACT-2014 project QUCHIP (G. A. 641039).

---

\*Present address: Complex Photonic Systems (COPS), University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands.

jelmer.renema@physics.ox.ac.uk

- [1] S. Aaronson and A. Arkhipov, *Theory Comput.* **9**, 143 (2013).
- [2] J. B. Spring *et al.*, *Science* **339**, 798 (2013).
- [3] M. A. Broome, A. Fedrizzi, S. Rahimi-Keshari, J. Dove, S. Aaronson, T. C. Ralph, and A. G. White, *Science* **339**, 794 (2013).
- [4] A. Crespi, R. Osellame, R. Ramponi, D. J. Brod, E. F. Galvão, N. Spagnolo, C. Vitelli, E. Maiorino, P. Mataloni, and F. Sciarrino, *Nat. Photonics* **7**, 545 (2013).
- [5] M. Tillmann, B. Dakić, R. Heilmann, S. Nolte, A. Szameit, and P. Walther, *Nat. Photonics* **7**, 540 (2013).
- [6] M. Bentivegna *et al.*, *Sci. Adv.* **1**, e1400255 (2015).
- [7] J. Carolan *et al.*, *Science* **349**, 711 (2015).
- [8] H. Wang *et al.*, *Nat. Photonics* **11**, 361 (2017).
- [9] V. S. Shchesnovich, *Phys. Rev. A* **91**, 063842 (2015).
- [10] A. Arkhipov, *Phys. Rev. A* **92**, 062326 (2015).
- [11] S. Aaronson, Scattershot boson sampling: a new approach to scalable boson sampling experiments, <https://www.scottaaronson.com/blog/>.
- [12] A. P. Lund, A. Laing, S. Rahimi-Keshari, T. Rudolph, J. L. O'Brien, and T. C. Ralph, *Phys. Rev. Lett.* **113**, 100502 (2014).
- [13] S. Rahimi-Keshari, T. C. Ralph, and C. M. Caves, *Phys. Rev. X* **6**, 021039 (2016).
- [14] V. S. Shchesnovich, *Phys. Rev. A* **89**, 022333 (2014).
- [15] V. S. Shchesnovich, *Phys. Rev. A* **91**, 013844 (2015).
- [16] M. C. Tichy, *Phys. Rev. A* **91**, 022316 (2015).
- [17] P. P. Rohde, *Phys. Rev. A* **91**, 012307 (2015).
- [18] V. Tamma and S. Laibacher, *Quantum Inf. Process.* **15**, 1241 (2016).
- [19] M. Walschaers, J. Kuipers, and A. Buchleitner, *Phys. Rev. A* **94**, 020104(R) (2016).
- [20] V. Tamma and S. Laibacher, *Phys. Rev. Lett.* **114**, 243601 (2015).
- [21] S. Laibacher and V. Tamma, [arXiv:1706.05578](https://arxiv.org/abs/1706.05578).
- [22] M. Tillmann, S.-H. Tan, S. E. Stoeckl, B. C. Sanders, H. de Guise, R. Heilmann, S. Nolte, A. Szameit, and P. Walther, *Phys. Rev. X* **5**, 041015(2015).
- [23] S.-H. Tan, Y. Y. Gao, H. deGuise, and B. C. Sanders, *Phys. Rev. Lett.* **110**, 113603 (2013).
- [24] H. de Guise, S.-H. Tan, I. P. Poulin, and B. C. Sanders, *Phys. Rev. A* **89**, 063819 (2014).
- [25] A. Khalid *et al.*, [arXiv:1709.06615](https://arxiv.org/abs/1709.06615).
- [26] M. Jerrum, A. Sinclair, and E. Vigoda, A polynomial-time approximation algorithm for the permanent of a matrix with non-negative entries, in *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing - STOC '01* (ACM Press, New York, NY, 2001).
- [27] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.120.220502>, for details on the derivations, the case of correlated matrices, and detailed simulability limits.
- [28] L. Euler, *Memoires de l'academie de Berlin* **7**, 1753 (1751).
- [29] G. Del Fiacco, Partial Derangement, in *From MathWorld—A Wolfram Web Resource* created by E. W. Weisstein, <http://mathworld.wolfram.com/PartialDerangement.html>
- [30] H. J. Ryser, *Combinatorial Mathematics* (Mathematical Association of America, Oberlin, OH, 1963).
- [31] A. Neville, C. Sparrow, R. Clifford, E. Johnston, P. M. Birchall, A. Montanaro, and A. Laing *et al.*, *Nat. Phys.* **13**, 1153 (2017).
- [32] L. Martino and V. Elvira, [arXiv:1704.04629](https://arxiv.org/abs/1704.04629).
- [33] P. Clifford and R. Clifford, [arXiv:1706.01260](https://arxiv.org/abs/1706.01260).
- [34] J. Wu *et al.*, [arXiv:1606.05836](https://arxiv.org/abs/1606.05836).