


Effect of photonic errors on quantum enhanced dense-subgraph finding

Naomi R. Solomons^{1,2,*}, Oliver F. Thomas,³ and Dara P. S. McCutcheon^{2,3}

¹Quantum Engineering Centre for Doctoral Training, Centre for Nanoscience and Quantum Information, University of Bristol, Bristol BS8 1FD, United Kingdom

²Quantum Engineering Technology Labs, H. H. Wills Physics Laboratory and Department of Electrical and Electronic Engineering, University of Bristol, Bristol BS8 1UB, United Kingdom

³Duality Quantum Photonics, 6 Lower Park Row, Bristol BS1 5BJ, United Kingdom

 (Received 15 February 2023; revised 16 June 2023; accepted 11 October 2023; published 21 November 2023)

We investigate the effects of photon loss and spectrally impure sources on Gaussian boson sampling (GBS) when used in dense-subgraph-finding algorithms. We find that the effectiveness of these algorithms is remarkably robust to such errors, to such an extent that there exist classical algorithms that can efficiently simulate the underlying GBS. These results suggest that, unlike the GBS problem itself, the speed-up of GBS-based algorithms over classical approaches when applied to the dense-subgraph problem is not exponential. They do suggest, however, that any advantage offered could be achieved on a quantum device with far less stringent requirements on photon loss and purity than general GBS.

DOI: [10.1103/PhysRevApplied.20.054043](https://doi.org/10.1103/PhysRevApplied.20.054043)

I. INTRODUCTION

Gaussian boson sampling (GBS) is a nonuniversal model of quantum computation that samples from the probability distribution of photon-counting measurements on Gaussian states [1]. Unlike universal quantum computers, GBS can be realized by currently available devices at a scale that is not efficiently simulable by a classical computer and hence has been the subject of early claims of quantum advantage [2–4]. Recent proposals have suggested using GBS for a number of different applications [5], including dense-subgraph identification [6,7], a problem that occurs in computing [8], computational biology [9,10], and finance [11], as well as predicting molecular-docking configurations [12,13]. Given the strong evidence for the exponential complexity of simulating GBS on classical devices [14,15], it has been suggested that GBS may enable genuine useful quantum computing applications offering exponential speed-up.

However, current and near-future GBS experiments are likely to be significantly affected by error, impacting the effectiveness of GBS-based algorithms and the corresponding extent of any speed-up over classical approaches.

For example, in the presence of sufficient photon distinguishability and loss, GBS experiments can be efficiently classically simulated (sampling from a sufficiently similar distribution in polynomial time) [16,17].

This work uses methods for simulating GBS as described in Ref. [18], producing samples that are used in the densest- k -subgraph (DkS) finding algorithms in Refs. [5,19]. These “quantum enhanced” algorithms use samples from GBS to improve the outcome of stochastic classical algorithms. We examine how effective GBS experiments are for this application when including the effects of spectrally impure sources and photon loss, using numerical studies on particular example graphs. The DkS problem is known to be NP hard in general and, as such, an efficient quantum algorithm is unlikely (as it is widely thought that NP is not in bounded-error quantum polynomial time (BQP), the class of problems efficiently solvable by a quantum computer). Nevertheless, it is important to know how errors will affect quantum approaches and to elucidate the origin of any speed-ups offered. We show that with every graph studied, the DkS algorithms are extremely robust to these sources of errors, even in regimes that may be efficient to simulate classically.

Insofar as our numerical results can be generalized, they suggest that GBS applied to the DkS problem is unlikely to offer an exponential speed-up over classical computing approaches. This is supported by recent advances in stochastic solutions to the DkS problem, as well as the ongoing speculation that hafnians of positive-valued matrices are classically efficient to calculate. For example, the work in Ref. [20] demonstrates a method to efficiently

*naomi.solomons@bristol.ac.uk

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.

calculate hafnians of matrices with positive entries to within additive error, although our work suggests that GBS is effective for DkS within error regimes that can be efficiently simulated with multiplicative error, for which general algorithms calculating the hafnian of matrices with all positive values are not currently known.

However, we also see that this advantage does not seem to impose challenging hardware requirements on loss and photon purity. Given the level of errors that can be tolerated for these algorithms, we speculate whether any advantage offered on quantum devices should really be considered “quantum,” or if “analog” or “optical” may be more fitting terminology.

II. GAUSSIAN BOSON SAMPLING

An m -mode quantum state ρ can be defined in terms of the Wigner function $\mathcal{W}(\vec{q}, \vec{p}) = (1/\pi^m) \int_{\mathbb{R}^m} \langle \vec{q} + \vec{x} | \rho | \vec{q} - \vec{x} \rangle d^m \vec{x}$. A Gaussian state is defined as a quantum state with a Gaussian Wigner function $\mathcal{W}(\vec{r}) = (2^m/\pi^m \sqrt{\det(\sigma)}) \exp(-(\vec{r} - \vec{D})^T \sigma^{-1} (\vec{r} - \vec{D}))$, with $\vec{r} = (\vec{q}, \vec{p})^T$ and thus an m -mode Gaussian state is completely defined by a $2m \times 2m$ covariance matrix σ and a $2m$ displacement vector \vec{D} [21]. In the following, we assume that $\vec{D} = \vec{0}$. GBS consists of measurements in the Fock basis applied to a Gaussian squeezed state passed through an interferometer. The measurement probabilities for photon-number-resolving (PNR) detectors are given by [1]

$$P(S; \{s_i\}) = \frac{2^N}{s_1! s_2! \cdots s_n! \sqrt{|\sigma_Q|}} \text{Haf}(\mathcal{A}_S), \quad (1)$$

in which S is the subset of modes involved in the detection outcome, s_i is the number of photons measured in mode i , $N = \sum_i s_i$, and $\sigma_Q = \sigma + \mathbb{1}$. Given the matrix $\mathcal{A} = (X \otimes \mathbb{1})(\mathbb{1} - 2\sigma_Q^{-1})$, we construct \mathcal{A}_S by repeating the i th row and column of \mathcal{A} according to s_i .

The difficulty of classically simulating GBS is due to the complexity of calculating the matrix hafnian: $\text{Haf}(A) = \sum_{\mu \in M} \left(\prod_{k=1}^{|S|} A_{\mu_{2k-1}, \mu_{2k}} \right)$, where M is the set of perfect matchings of S , the different ways of “pairing” the indices (every permutation in which $\mu_{2k} < \mu_{2(k+1)}$ and $\mu_{2k} < \mu_{2k+1}$). Each measurement carried out in a GBS experiment draws a sample from the distribution described by Eq. (1). Generating samples according to this distribution by direct calculation is #P hard [22]. In the ideal case, the best algorithm for simulating GBS has complexity $O(mN^3 2^{N/2})$, depending on the number of modes m and the measured number of photons N [23]. The complexity with which an experimental apparatus produces samples is linear in the number of samples but alongside the difficulties of physically constructing the device, there is the initial computational overhead of calculating the correct interferometer settings to produce the desired output. Matrix

decompositions, e.g., Williamson and Bloch-Messiah, are needed to find the transformations to generate the correct state; these generally require $O(m^3)$ time in practice [24].

III. MAPPING GRAPHS TO GAUSSIAN STATES

A graph G is defined by a collection of vertices V and connecting edges E , characterized by an adjacency matrix A . We consider only equal-weight undirected graphs, although all graphs can be represented as Gaussian states [25]. To find this state, we use the procedure defined in Ref. [5]. First, the adjacency matrix A is diagonalized to find the eigenvalues $\{\lambda\}$. Next, we pick the scaling parameter c such that $c < \lambda_{\max}^{-1}$. We then construct $\mathcal{A} = c(A \oplus A)$ and the covariance matrix is found using

$$\sigma = 2(\mathbb{1} - X\mathcal{A})^{-1} - \mathbb{1}, \quad \text{with } X = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}. \quad (2)$$

This means that the probability of sampled outcomes will be proportional to $|\text{Haf}(A)|^2$ (as $\text{Haf}(A \oplus A) = |\text{Haf}(A)|^2$). The scaling parameter (c) ensures that σ corresponds to physical squeezing values and can be chosen to optimize the photon-number distribution [12]. In principle, the expected photon number can be arbitrarily high but in practice this is limited by the amount of squeezing possible. The largest eigenvalue of the adjacency matrix is bounded above by the largest vertex degree in the graph; hence more well-connected graphs are likely to require greater squeezing levels [26].

The algorithms described in Sec. IV are intended for use in the collision-free regime, which means that outcomes containing multiple photons in one mode have negligible probability and can be ignored. For this reason, we consider the use of threshold detectors that do not distinguish between photon numbers greater than zero. However, higher values of c , which result in more collisions, tend to favor more samples being drawn in the preferred photon number or click subspace for DkS and hence improve the efficiency of the algorithm. As the size of the graphs increases, the probability of collisions decreases, hence allowing data collected in this work to use higher values of c .

IV. DENSE-SUBGRAPH IDENTIFICATION

We consider the problem of finding the subgraph of k vertices with the largest density within the input graph G . The density of G is given by $\rho(G) = 2|E(G)|/|V(G)|(|V(G)| - 1)$, in which $|V(G)|$ and $|E(G)|$ represent the number of vertices and edges in the graph, respectively. The maximum possible density is therefore 1, for a fully connected graph (a clique). The DkS problem is NP hard [27], although solutions to variations on this problem can be found in polynomial time [28,29], as can approximate solutions [30].

Consider a GBS experiment where the Gaussian state leaving the interferometer has a covariance matrix as defined in Eq. (2) for some graph G . We can then associate subsets S of output modes, corresponding to measurement outcomes, to subgraphs of G , where the modes in S correspond to the vertices in the subgraph. As shown in Ref. [19], the number of perfect matchings in the graph, and hence the size of the hafnian of the adjacency matrix, is strongly correlated with the number of edges in a graph and hence more dense subgraphs are the most likely sampling results. GBS can therefore be used to seed algorithms for DkS.

Different classical and quantum enhanced algorithms exist for DkS. In this work, we focus on the simplest approach, where the classical case samples from the uniform distribution of k -vertex subgraphs and the quantum enhanced case takes samples using GBS and retains only those of the correct size (the random-search algorithm in Ref. [19]), using these to select subgraphs. We tune the scaling parameter c to maximize the probability of click patterns of the correct size. In the appendixes, we explore the performance of the base (photon-number-varying) GBS algorithm and also compare with the performance of simulated annealing algorithms [19]. We use as a benchmark a deterministic classical algorithm, consisting of iteratively removing the lowest-degree vertex [31], which does not always find the densest subgraph but is guaranteed to find one of a density within a reasonable approximation ratio.

Our main text focuses on one particular randomly generated graph, with further examples (changing the initial density and considering a graph with an embedded clique) considered in the appendixes, which show qualitatively similar results. To create the graphs in this work, we use the Erdős-Rényi form [32] for generating a random graph $G(M, \rho)$, of size M with density ρ , as described in Appendix B. The quantum Gaussian-optics emulator used to model the outcome statistics is available online [33].

V. MODELING SPECTRAL IMPURITY

The spectral impurity of the sources in a GBS setup can be modeled by replacing each interfered mode with an ensemble of spectral modes, across which the squeezed states are generated. When measured using broadband detectors that do not distinguish spectral modes, this gives rise to mixing of the quantum state [18]. In the limit of maximal mixing, no quantum interference occurs and the sampling can be efficiently classically simulated [34].

We now introduce a method for including ensembles of spectral modes in a Gaussian state representing a graph, which is used to model spectral impurity in our emulator. Our method is general enough to work for any physical state and can be thought of as replacing the idealized single (spectral-)mode squeezers used to generate the state with

more physical sources with imperfections. The procedure to include imperfect sources is as follows:

- (1) Construct a graph G with adjacency matrix A .
- (2) Choose a scaling value, c , and construct a quantum state, σ , according to Eq. (2).
- (3) Perform the Williamson decomposition to find the symplectic transformation that generates the state $\sigma = \mathbf{M}\mathbf{1}\mathbf{M}^\dagger$, $\mathbf{M} = \sigma^{1/2}$ (for a pure state).
- (4) Perform the Bloch-Messiah decomposition on the symplectic matrix to obtain the unitaries and single-mode squeezers, $\mathbf{M} = \mathbf{U}\mathbf{M}^D\mathbf{V}^\dagger$.
- (5) Replace the unitaries \mathbf{U} and \mathbf{V} with the full-sized versions including spectral modes, $\mathcal{U} = \mathbf{U} \otimes \mathbb{1}_{N_F}$, $\mathcal{V}^\dagger = \mathbf{V}^\dagger \otimes \mathbb{1}_{N_F}$ (in which N_F is the number of spectral modes).
- (6) Replace the set of single-mode squeezers with a set of single-spatial-mode multiple-spectral-mode squeezers $\mathbf{M}^D \mapsto \mathcal{M}^D$, which are normalized to the same amount of squeezing as the single spectral mode. Detectors are used that trace over the spectral modes, i.e., nonfrequency-resolving detectors.
- (7) The multimode state can then be constructed using $\sigma = \mathcal{M}\mathcal{M}^\dagger$, where $\mathcal{M} = \mathcal{U}\mathcal{M}^D\mathcal{V}^\dagger$.

Having constructed these multimode states, we can use the method described in Ref. [18] to measure photons in the multimode formalism.

Step (6) above requires characterizing the spectral properties of the sources by finding the nonzero Schmidt coefficients. To do this, we pick a set of basis functions— for integrated optics, these are typically the Hermite polynomials [35], although any complete set of normalized functions is suitable. We find that for purity P , at least $\lceil 1/P \rceil$ nonzero coefficients are required. These are constructed using a geometric scaling, i.e., we have the set $S = \{s_1\} \cup \{s_i^2 = k_{l,b}(i)(1 - s_1^2), 2 \leq i \leq l\}$, with $k_{l,b}(i) = b^{l-i} / (\sum_{j=1}^{l-1} b^j)$. Here, l is the number of nonzero Schmidt coefficients and b is the base that sets the factor for the scaling between the elements x_i . More information on this procedure is given in Appendix C.

VI. RESULTS

We consider a graph with $n = 24$ vertices and density 0.362. Appendix E also considers a graph with density 0.196 and a modified version containing a clique. We search for the densest subgraph of $k = 8$ vertices, which has density 0.714. Figure 1 compares the performance of the classical algorithm with the quantum enhanced algorithm that uses GBS samples produced in our simulation. The gray line shows the result of the deterministic classical algorithm. The number of steps on the x axis refers to the number of random subgraphs drawn from the target k -click subspace, with the final density being the maximum of these samples. The results are averaged over

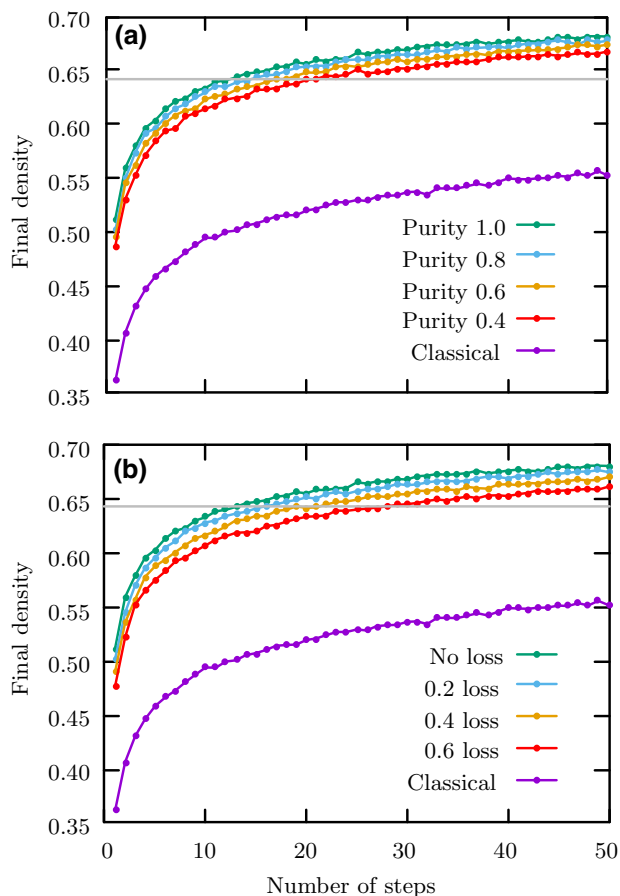


FIG. 1. The performance of the quantum enhanced and classical random-search algorithms. The quantum enhanced algorithm considerably outperforms the classical algorithm, even in the presence of high levels of noise. We vary (a) the spectral impurity and (b) the loss as indicated. The gray line shows the performance of the deterministic classical algorithm [31]. The final maximum density of sampled subgraphs is shown as a function of the sample batch size.

1000 iterations. We can see that the quantum enhanced algorithms outperform the classical algorithms, similarly to Ref. [19]. The deterministic algorithm performs well but is quickly overtaken by the quantum enhanced algorithms.

Figure 1(b) also shows the GBS algorithm including different photon loss rates. This is modeled by applying uniform loss to all modes after constructing the multiple-spectral-mode state. We also vary the spectral purity of the sources used in the GBS emulation using the method described above, shown in Fig. 1(a). Interestingly, the quantum advantage offered by the GBS algorithm appears to be remarkably robust to errors from photon loss and spectral impurity; in both cases, the simulated quantum enhanced routines retain a clear advantage over the classical routine, as well as approaching the best graph within the number of samples considered.

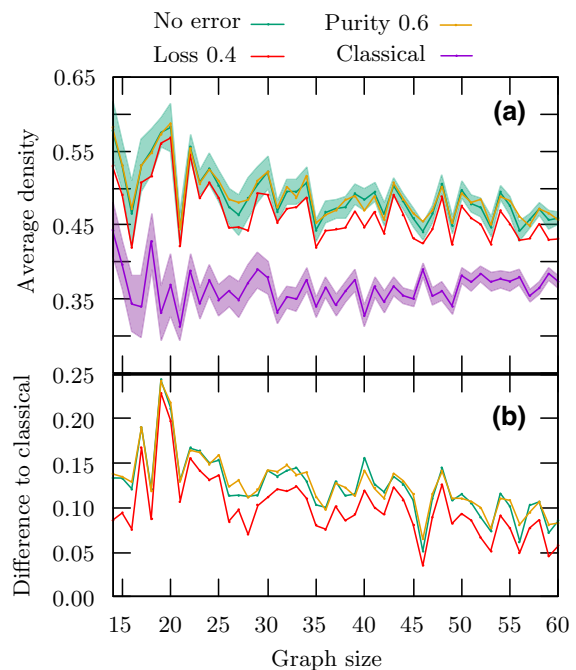


FIG. 2. The density of graphs chosen from GBS samples of different sizes. (a) The average density of sampled \sqrt{n} -subgraphs as a function of the graph size n with initial density 0.4. The variance is shown (shaded region) for the classical and quantum case with no error. (b) The difference between the average density of sampled subgraphs, in the quantum case, compared to the classical case. The pattern of noisy GBS samples having high density, compared to sampling from the uniform distribution, persists, although with decreasing advantage at larger graph sizes.

In order to see how the speed-up from the quantum enhanced algorithm scales with size, we also investigate graphs of different sizes, all generated using the Erdős-Rényi form [32] with $\rho = 0.4$. We draw samples corresponding to k -vertex subgraphs, where k is the closest integer to \sqrt{n} , either from the uniform distribution (for the classical case) or from a simulated GBS device. Figure 2 compares the average densities of these subgraphs after drawing 1000 samples, including with photon loss and spectral impurity. In Fig. 2(b), which shows the difference between the quantum samples drawn and the classical samples, we see that the extent of the quantum advantage here appears to decrease with increasing graph size.

VII. DISCUSSION

It is clear that the quantum enhanced DkS algorithms are resilient against the forms of error considered. A considerable speed-up (compared to the classical case) is shown by the examples with high levels of error (up to 60% loss and sources with purity 0.4), with little difference between the effectiveness of these and the simulations with no error. This can in part be understood from Fig. 3, which shows

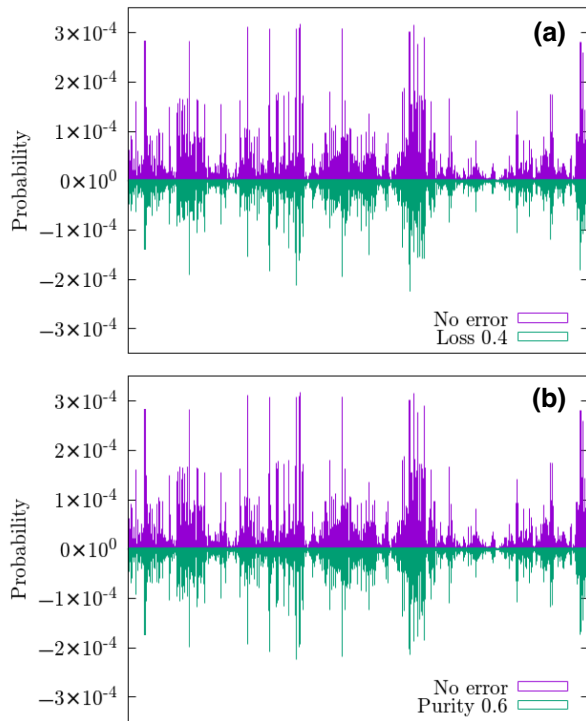


FIG. 3. The similarity of the probability distributions. We show the probability of different outcomes within the eight-click subspace, (a) in the ideal case (purple) and with loss (green) and (b) with spectrally impure sources (green).

the outcome-probability photon loss and spectral impurity. Although some subtle differences can be seen in the distributions, the overall structure appears unchanged—namely, samples with the highest probabilities retain high probabilities. This gives some intuition as to why the algorithms continue to select dense subgraphs in the presence of errors. This argument applies due to the nature of the DkS problem, as more likely outcomes are better solutions, but there is no need to estimate any probabilities; we would expect to continue to see good performance as long as the peaks remain in the same place.

It is worth noting that we have considered the performance of the quantum enhanced algorithm with samples selected from the appropriate photon-number subspace. In realistic implementations, samples drawn outside of this subspace can be manipulated into graphs of the correct size, with minimal overheads (e.g., by removing low-degree vertices or combining graphs from different samples). Without this postselection, the quantum speed-up remains but is reduced, as more samples need to be drawn. This is considered further in Appendix G.

When varying the graph size n , we see that simulations with photon loss and spectrally impure sources continue to perform well. This also indicates that the speed-up of the quantum algorithm is smaller with larger n . In addition, it can be seen that low spectral impurity has less of an impact

than increased levels of loss. The variance of subgraph density in these samples is, in general, lower than that for the case with loss. This suggests that the experiment in this case may occasionally not be able to sample the densest subgraph but also avoids sampling very low-density subgraphs.

GBS with a high level of photon distinguishability is known to be efficient to simulate classically [16]. Similarly, increasing the spectral impurity is less well studied but approaches the limit of simulating thermal states, which are also efficient to classically sample from [34] and hence reduces the usefulness of quantum resources. For a sufficient level of photon loss, GBS permits an efficient classical simulation [17]. Following the analysis of Ref. [17], with loss of at least 43.7%, the GBS considered here is efficiently classically (approximately) simulable. As loss is increased, we have increased the squeezing to compensate in order to maximize the likelihood of k -click events (as described in Appendix A). This increases the loss threshold above which GBS is classically simulable, up to a maximum of 47.3% for these parameter s . As such, although the algorithms appear robust to errors, the level of errors means that efficient classical algorithms exist that perform the same underlying sampling task with minimal error. For the cases studied in this work at least, we conclude that the apparent quantum advantage of GBS applied to the DkS problem may be achieved using efficient classical algorithms instead.

The levels of error that are tolerable suggest that these algorithms are not fully exploiting truly quantum mechanical effects. It is not straightforward to engineer a graph problem requiring such phenomena, since not all Gaussian states correspond to graphs in the manner described, i.e., adjacency matrices with real entries. Future work could examine whether our findings are due to the structure of Gaussian states described by graphs if, e.g., this is due to the limited tree width of randomly produced graphs [36]. It may be possible to find Gaussian states corresponding to DkS cases that are not robust to loss and impurity errors, although we were not able to construct any in our investigations.

The code used here is open source and can be run with a variety of different graph types [33].

VIII. CONCLUSIONS

We have used methods of simulating GBS to show that quantum enhanced dense-subgraph finding is particularly robust to spectral error and photon loss. This suggests that efficient classical methods exist that can implement the same algorithm as the quantum device, albeit with a potential polynomial overhead associated with drawing samples. Recently, the advantage of using GBS for DkS has inspired classical algorithms for the same problem [20]. The implication is that a quantum device with modest requirements

on photon loss and spectral purity could be used for the DKS problem with little loss in performance, though this in itself raises questions as to whether the advantage gained can be said to be “quantum” in the sense that it is generally understood when applied to algorithms. This work has focused on studying Erdős-Rényi random graphs and further work could be done to generalize these results to different graphs with potentially more elaborate structures. Furthermore, more study is needed into the Gaussian states generated from graphs and whether the conclusions drawn here apply to other applications of GBS. Finally, these results support ongoing speculation that GBS problem itself can be efficiently simulated classically within certain regimes. In particular, it has been suggested that it is possible to efficiently calculate hafnians of matrices with all positive values to within multiplicative error, as is possible in the case of permanents [37].

ACKNOWLEDGMENTS

We would like to thank Patrick Yard and Anthony Laing for feedback on the manuscript and Ryan Mann for useful discussions. N.R.S. is supported by the Quantum Engineering Centre for Doctoral Training, through Engineering and Physical Sciences Research Council (EPSRC) Grant No. EP/SO23607/1.

APPENDIX A: CALCULATING THE SCALING PARAMETER

Here, we discuss how to calculate the optimal scaling parameter, c , in order to maximize the probability of drawing samples in the k -click subspace. We first eigendecompose the adjacency matrix $A = U\Lambda U^T$, with $\Lambda = \text{diag}\{|\lambda_i|\}$, in which λ_i are the eigenvalues of A . We then rescale to give the diagonal matrix $t_D = c\Lambda$ with $t_D = \text{diag}\{\tanh r_i\}$ with r_i the squeezing in mode i . When using threshold detectors, the expected number of “clicks” (detectors registering at least one photon during a measurement) is given by

$$\langle \hat{C} \rangle = M - \sum_{i=1}^M P_{\text{vac}}(i), \quad (\text{A1})$$

where M is the number of modes, and the vacuum probability in mode i , $P_{\text{vac}}(i)$ is given by

$$P_{\text{vac}}(i) = \left(\left| \sum_j |U_{ij}|^2 \left(\frac{1 - t_D^2}{1 - t_D^2} \right)_j \right|^2 - (1 - t_D^2)^2 \left| \sum_j U_{ij}^2 \left(\frac{t_D}{1 - t_D^2} \right)_j \right|^2 \right)^{-1/2}. \quad (\text{A2})$$

Increasing c means increasing the squeezing parameter and hence this increases the likelihood of collision events,

which makes the GBS experiment more simulable [38]. The effect of squeezing on dense-subgraph finding has been investigated experimentally in Ref. [6].

APPENDIX B: GENERATING RANDOM GRAPHS

Our implementation of the Erdős-Rényi model [39] is as follows:

- (1) Create a set of vertices V of size M .
- (2) Start at the first vertex $v_{i=1}$ in V .
- (3) For every other vertex $\{v_j \in V \setminus v_i\}$, pick $r_j \in [0, 1]$ from a uniform distribution.
- (4) If $r_j < \rho/2$, then add an edge between v_i and v_j .
- (5) At the last vertex, $v_{i=M}$, stop, else move to the next vertex, $v_i \leftarrow v_{i+1}$, and go to step (3).

APPENDIX C: CONSTRUCTING ARBITRARY PURITY SOURCES

We wish to find the Schmidt coefficients of the multiple-spectral-mode sources that we will use in our model. We first construct the vector of singular values $\{S\}$, such that they obey the following normalization and purity conditions (in which P is purity):

$$\sum_i S_i^2 = 1, \quad \sum_i S_i^4 = P := \text{Tr}[\hat{\rho}^2]. \quad (\text{C1})$$

In order to simplify the two conditions in Eqs. (C1), we construct a new set of coefficients, $X = \{x_i = S_i^2\}$, which reduces the two conditions to

$$\sum_i x_i = 1, \quad \sum_i x_i^2 = P. \quad (\text{C2})$$

We define the set X (which are the Schmidt coefficients squared) as

$$X_{l,b} = \{x_1\} \cup \{x_i = k_{l,b}(i)(1 - x_1) | 2 \leq i \leq l\}, \quad (\text{C3})$$

so that, by construction, the normalization condition is satisfied, where the condition on the values $k_{l,b}(i)$ ensures that they are normalized to 1, $\sum_i k_{l,b}(i) = 1$. One useful form for the set $k_{l,b}$ is the geometric scaling

$$k_{l,b}(i) = \frac{b^{l-i}}{\sum_{j=1}^{l-1} b^{j-1}}. \quad (\text{C4})$$

Here, l is the number of nonzero Schmidt coefficients and b is the base that sets the factor for the geometric scaling between the elements x_i . i is the index of the element of the set of $X_{l,b}$ for the coefficient value.

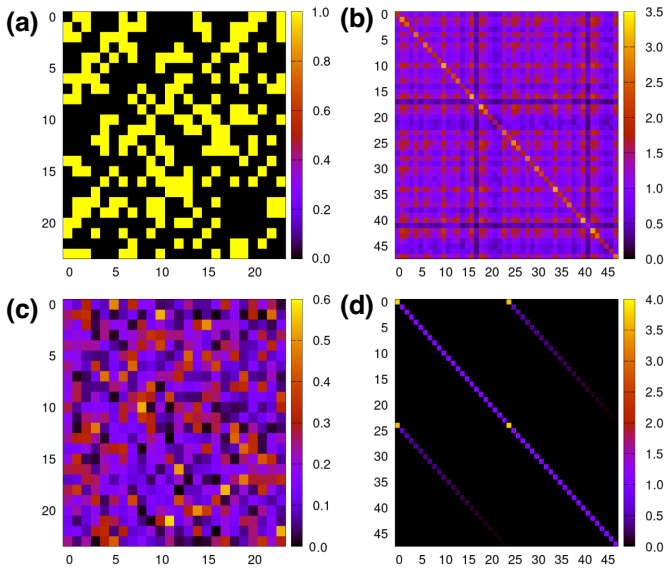


FIG. 4. The absolute values of the Gaussian transformation corresponding to our 24-vertex graph with a density 0.362. We show (a) the adjacency matrix of the graph, (b) the covariance matrix describing the associated Gaussian state, (c) the linear interferometers U to create the state, and (d) the single-mode squeezer matrix M^D .

Using these definitions, the purity condition given in Eq. (C2) is then

$$\sum_i x_i^2 = x_1^2 + (k_{l,b}(2))^2(1-x_1)^2 + \dots + (k_{l,b}(l))^2(1-x_1)^2 = P, \quad (\text{C5})$$

and grouping the terms by the order of x_1 , we see that it produces a quadratic equation in x_1 :

$$\left(1 + \sum_{i=2}^l (k_{l,b}(i))^2\right) x_1^2 - 2 \left(\sum_{i=2}^l (k_{l,b}(i))^2\right) x_1 + \left(\sum_{i=2}^l (k_{l,b}(i))^2\right) - P = 0, \quad (\text{C6})$$

which is then straightforward to solve. Since all the elements of S are positive, all the elements of X are positive and, as such, the positive root can be taken. Lastly, we calculate the remaining x_i by substituting x_1 back in:

$$x_i = k_{l,b}(i)(1-x_1), \quad (\text{C7})$$

and then the set S is the (positive) square root of X ,

$$S = \{\sqrt{x_i} | x_i \in X\}. \quad (\text{C8})$$

Due to the choice of the two parameters l and b , for a particular set of S with the desired purity there are many distributions of Schmidt coefficients that generate a source with the same purity. Also, we note that our construction does not generate all Schmidt distributions but it can be used to generate a large class of distributions.

One interesting consequence of this particular form for X is that for any purity with $\frac{1}{2} \leq P \leq 1$, only two terms are needed. For quantum information processing applications, we are typically interested in sources with purities close to 1 and, as such, the majority of physically relevant or interesting sources can be described with two Schmidt modes. This offers a significant simplification when including a large number of spectral degrees of freedom into a

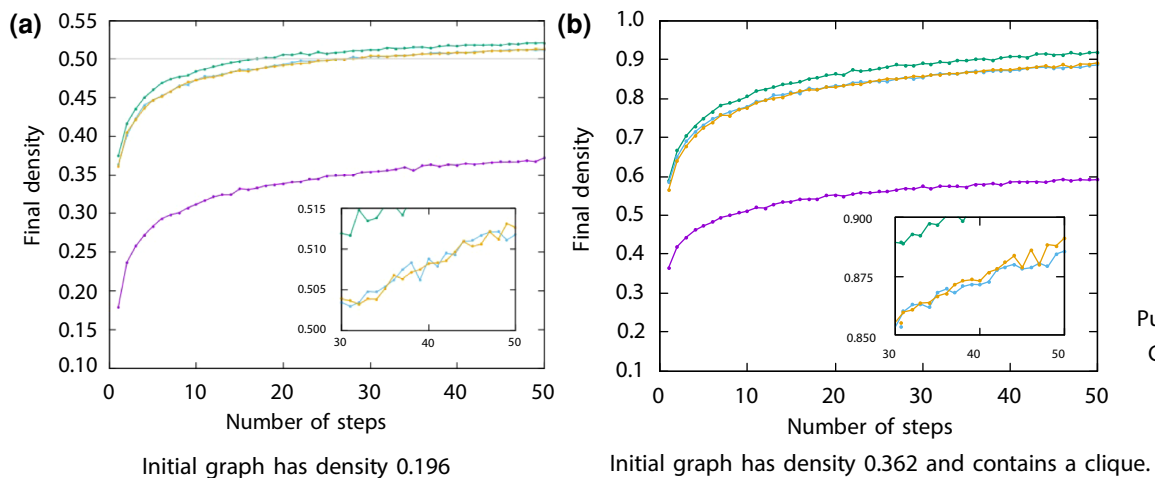


FIG. 5. A comparison of the performance of the classical (purple) and quantum (green) enhanced random-sampling algorithms, with different sources of error. The gray line shows the performance of the deterministic classical algorithm [31] (it is able to identify the clique). The inset focuses on the region between 30 and 50 steps. (a) The initial graph has density 0.196. (b) The initial graph has density 0.362 and contains a clique.

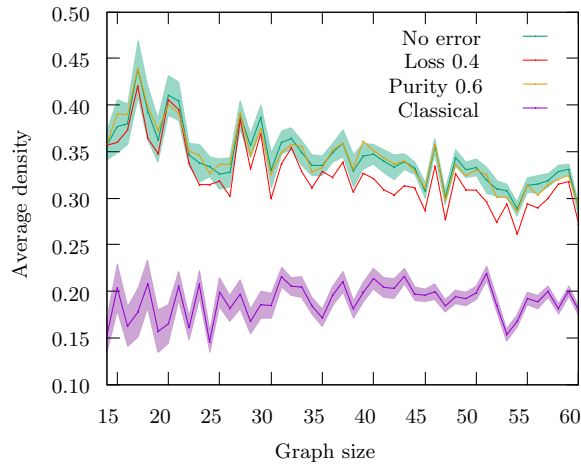


FIG. 6. The average density of subgraph found over 1000 samples, for different graph sizes, with an initial density of approximately 0.2. The variance is shown for the classical and quantum case with no error.

model. Although the inclusion of spectral modes that are traced over in the detection calculations only introduces a cubic overhead [18], these factors are still large. Using this relation we can, when required for speed, drastically reduce the number of spectral modes modeled from some large l to only 2.

APPENDIX D: VISUALIZING THE QUANTUM STATES

In the main text, we have described how to find a Gaussian state, described by covariance matrix σ from a particular graph. In Fig. 4, we plot the adjacency matrix [Fig. 4(a)] of our 24-mode example graph and the covariance matrix [Fig. 4(b)] of the associated Gaussian state.

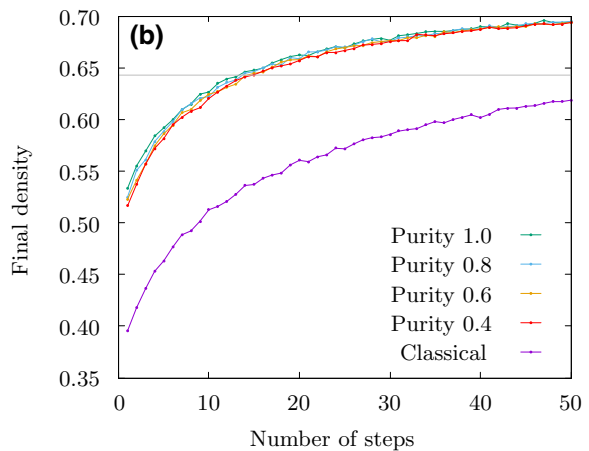
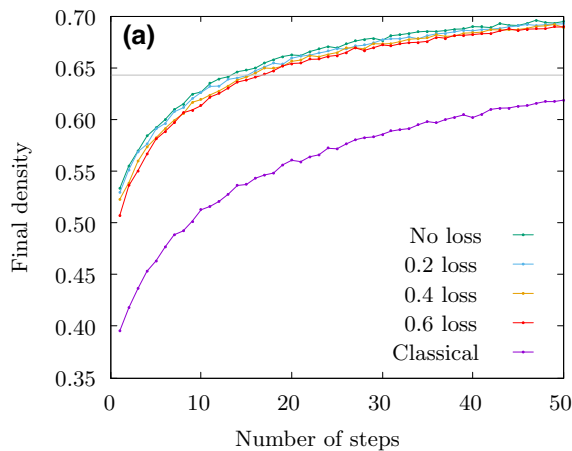


FIG. 7. A comparison of the performance of the classical and quantum enhanced simulated-annealing algorithms, (a) in the presence of varying amounts of loss and (b) with spectrally impure sources. The gray line shows the performance of the deterministic classical algorithm [31].

We then use the Williamson decomposition, followed by the Bloch-Messiah decomposition, to find the unitary matrix and matrix of single-mode squeezers that produce the state, also shown in Figs. 4(c) and Figs. 4(d), respectively).

In particular, it is interesting to note that the sources show one bright squeezer and many low-power single-mode squeezers.

APPENDIX E: FURTHER GRAPH EXAMPLES

To check the generalizability of our results, we carry out the same analysis as in the main text but with different graphs. We show the performance of random sampling on two initial subgraphs, one of which is an Erdős-Rényi graph with density 0.196 and one that is a modified version of the main graph from our paper, with the same initial density but containing a clique (this is identified by the deterministic algorithm). Figure 5 shows the impact of loss and spectrally impure sources on dense-subgraph finding applied to these graphs. Figure 6 shows the average density of graphs chosen by drawing samples from a GBS distribution or by randomly selecting, in the classical case, over 1000 repetitions. As in Fig. 2, this is done for a range of graph sizes, searching for subgraphs with size equal to the closest integer to \sqrt{n} , but with the initial graph created using the Erdős-Rényi method, with initial density 0.2.

These results suggest similar behavior to the example discussed in the main text, being very robust to error and showing a significant speed-up in the quantum case, even with high levels of loss and spectral impurity.

APPENDIX F: SIMULATED ANNEALING

We now employ the quantum enhanced simulated-annealing algorithm, as described in Ref. [19], for finding dense subgraphs in our main example graphs (as

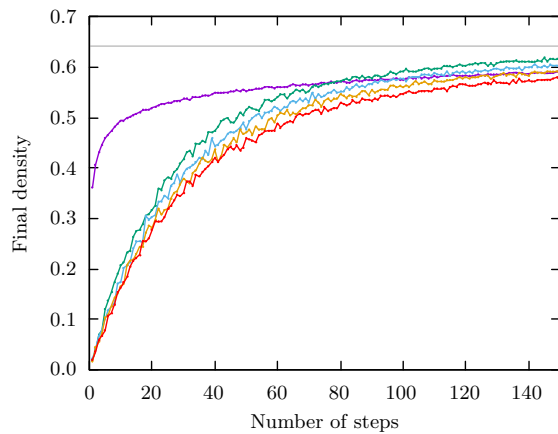


FIG. 8. The performance of the classical and quantum enhanced random-sampling algorithms for dense-subgraph finding, without postselecting GBS samples from the correct photon-number subspace. The gray line shows the performance of the deterministic classical algorithm.

introduced in the main text). Figure 7 shows the impact of loss [Fig. 7(a)] and impure sources [Fig. 7(b)]. As in Ref. [19], this algorithm shows a smaller difference between the effectiveness of the classical and quantum enhanced algorithms. Nonetheless, as with our previous results, this speed-up remains even in the presence of significant levels of error.

Note that our implementation of the simulated-annealing algorithm differs slightly from that described in Ref. [19]. During the GBS-Tweak subroutine, described in Ref. [40], a sample T (in this case, corresponding to a subgraph) is drawn that facilitates an exploration of the search space, after drawing an initial sample S . We have drawn the sample T from the modes that are not included in S , as opposed to the full set of modes. This has been done to reduce the run time while minimizing the impact on the performance of the algorithm. Here, “Number of steps” (as indicated on the x axis) refers to the number of steps in the GBS-Tweak subroutine.

APPENDIX G: PERFORMANCE WITHOUT POSTSELECTION

Previous results have considered the number of steps to only include detection events with k click events. However, unlike the Aaronson-Arhipov boson sampling scheme, which uses single photons as input [41], GBS does not have a fixed photon number at the detection stage. By choosing the scaling parameter c appropriately, it is possible to bias the output toward a particular photon number (or number of clicks when using threshold detectors) but a spread of photon numbers will be observed in practice.

In Fig. 8, we show the performance of the DkS random-sampling algorithm, where here “Number of steps” refers to the total number of samples drawn. We assume that

any samples with the incorrect click number are discarded (realistically, these may be useful for local searches). We vary the loss parameter and use the optimal scaling parameter c (using the method previously described) for each different loss value (thus, we assume that the error in the device is well understood and that the squeezing can be increased to compensate). For the classical comparison, we repeat the previous method and sample randomly across only the k -vertex subgraphs. As before, we can see that increased levels of loss do not have a significant impact on the performance of the quantum enhanced algorithm.

- [1] C. S. Hamilton, R. Kruse, L. Sansoni, S. Barkhofen, C. Silberhorn, and I. Jex, Gaussian boson sampling, *Phys. Rev. Lett.* **119**, 170501 (2017).
- [2] A. Lund, M. J. Bremner, and T. Ralph, Quantum sampling problems, BosonSampling and quantum supremacy, *npj Quantum Inf.* **3**, 1 (2017).
- [3] H.-S. Zhong, H. Wang, Y.-H. Deng, M.-C. Chen, L.-C. Peng, Y.-H. Luo, J. Qin, D. Wu, X. Ding, and Y. Hu, *et al.*, Quantum computational advantage using photons, *Science* **370**, 1460 (2020).
- [4] L. S. Madsen, F. Laudenbach, M. F. Askarani, F. Rortais, T. Vincent, J. F. Bulmer, F. M. Miatto, L. Neuhaus, L. G. Helt, and M. J. Collins, *et al.*, Quantum computational advantage with a programmable photonic processor, *Nature* **606**, 75 (2022).
- [5] T. R. Bromley, J. M. Arrazola, S. Jahangiri, J. Izaac, N. Quesada, A. D. Gran, M. Schuld, J. Swinarton, Z. Zabaneh, and N. Killoran, Applications of near-term photonic quantum computers: Software and algorithms, *Quantum Sci. Technol.* **5**, 034010 (2020).
- [6] S. Sempere-Llagostera, R. Patel, I. Walmsley, and W. Kolthammer, Experimentally finding dense subgraphs using a time-bin encoded Gaussian boson sampling device, *Phys. Rev. X* **12**, 031045 (2022).
- [7] Y.-H. Deng, S.-Q. Gong, Y.-C. Gu, Z.-J. Zhang, H.-L. Liu, H. Su, H.-Y. Tang, J.-M. Xu, M.-H. Jia, and M.-C. Chen, *et al.*, Solving graph problems using Gaussian boson sampling, *Phys. Rev. Lett.* **130**, 190601 (2023).
- [8] R. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins, Trawling the web for emerging cyber-communities, *Comput. Networks* **31**, 1481 (1999).
- [9] H. Hu, X. Yan, Y. Huang, J. Han, and X. J. Zhou, Mining coherent dense subgraphs across massive biological networks for functional discovery, *Bioinformatics* **21**, i213 (2005).
- [10] M. Habibi, C. Eslahchi, and L. Wong, Protein complex prediction based on k -connected subgraphs in protein interaction network, *BMC Syst. Biol.* **4**, 1 (2010).
- [11] S. Arora, B. Barak, M. Brunnermeier, and R. Ge, in *ICCS* (Association for Computing Machinery, New York, 2010), p. 49.
- [12] L. Banchi, M. Fingerhuth, T. Babej, C. Ing, and J. M. Arrazola, Molecular docking with Gaussian boson sampling, *Sci. Adv.* **6**, eaax1950 (2020).
- [13] S. Yu, Z.-P. Zhong, Y. Fang, R. B. Patel, Q.-P. Li, W. Liu, Z. Li, L. Xu, S. Sagona-Stophel, and E. Mer, *et al.*, A universal

- programmable Gaussian Boson Sampler for drug discovery, Preprint [ArXiv:2210.14877](https://arxiv.org/abs/2210.14877) (2022).
- [14] R. Kruse, C. S. Hamilton, L. Sansoni, S. Barkhofen, C. Silberhorn, and I. Jex, Detailed study of Gaussian boson sampling, *Phys. Rev. A* **100**, 032326 (2019).
- [15] N. Quesada, J. M. Arrazola, and N. Killoran, Gaussian boson sampling using threshold detectors, *Phys. Rev. A* **98**, 062322 (2018).
- [16] J. J. Renema, Simulability of partially distinguishable superposition and Gaussian boson sampling, *Phys. Rev. A* **101**, 063840 (2020).
- [17] H. Qi, D. J. Brod, N. Quesada, and R. García-Patrón, Regimes of classical simulability for noisy Gaussian boson sampling, *Phys. Rev. Lett.* **124**, 100502 (2020).
- [18] O. F. Thomas, W. McCutcheon, and D. P. S. McCutcheon, A general framework for multimode Gaussian quantum optics and photo-detection: Application to Hong-Ou-Mandel interference with filtered heralded single photon sources, *APL Photonics* **6**, 040801 (2021).
- [19] J. M. Arrazola and T. R. Bromley, Using Gaussian boson sampling to find dense subgraphs, *Phys. Rev. Lett.* **121**, 030503 (2018).
- [20] C. Oh, L. Jiang, and N. Quesada, Quantum-inspired classical algorithm for graph problems by Gaussian boson sampling, Preprint [ArXiv:2302.00536](https://arxiv.org/abs/2302.00536) (2023).
- [21] G. Adesso, S. Ragy, and A. R. Lee, Continuous variable quantum information: Gaussian states and beyond, *Open Syst. Inf. Dyn.* **21**, 1440001 (2014).
- [22] A. Deshpande, A. Mehta, T. Vincent, N. Quesada, M. Hinsche, M. Ioannou, L. Madsen, J. Lavoie, H. Qi, and J. Eisert, *et al.*, Quantum computational advantage via high-dimensional Gaussian boson sampling, *Sci. Adv.* **8**, eabi7894 (2022).
- [23] N. Quesada, R. S. Chadwick, B. A. Bell, J. M. Arrazola, T. Vincent, H. Qi, and R. García-Patrón, Quadratic speed-up for simulating Gaussian boson sampling, *PRX Quantum* **3**, 010306 (2022).
- [24] V. Vasudevan and M. Ramakrishna, A hierarchical singular value decomposition algorithm for low rank matrices, Preprint [ArXiv:1710.02812](https://arxiv.org/abs/1710.02812) (2017).
- [25] M. Walschaers, S. Sarkar, V. Parigi, and N. Treps, Tailoring non-Gaussian continuous-variable graph states, *Phys. Rev. Lett.* **121**, 220501 (2018).
- [26] H. Minc, *Nonnegative Matrices* (Technion—Israel Institute of Technology, Department of Mathematics, New York, 1974).
- [27] A. Bhaskara, M. Charikar, E. Chlamtac, U. Feige, and A. Vijayaraghavan, in *Proceedings of the Forty-Second ACM Symposium on Theory of Computing* (Association for Computing Machinery, New York, 2010), p. 201.
- [28] G. Gallo, M. D. Grigoriadis, and R. E. Tarjan, A fast parametric maximum flow algorithm and applications, *SIAM J. Comput.* **18**, 30 (1989).
- [29] C. Tsourakakis, in *Proceedings of the 24th International Conference on World Wide Web (IW3C2)*, Florence, 2015), p. 1122.
- [30] U. Feige, D. Peleg, and G. Kortsarz, The dense k -subgraph problem, *Algorithmica* **29**, 410 (2001).
- [31] Y. Asahiro, K. Iwama, H. Tamaki, and T. Tokuyama, Greedily finding a dense subgraph, *J. Algorithms* **34**, 203 (2000).
- [32] P. Erdős and A. Rényi, *et al.*, in *On Random Graphs*, Vol. 1 (Institute of Mathematics, University of Debrecen, Hungary, 1959), p. 290.
- [33] Duality Quantum Photonics, Quantum Gaussian optics toolkit, https://gitlab.com/dualityqp/qgot_public (accessed 2022-11-30).
- [34] S. Rahimi-Keshari, A. P. Lund, and T. C. Ralph, What can quantum optics say about computational complexity theory?, *Phys. Rev. Lett.* **114**, 060501 (2015).
- [35] P. P. Rohde, W. Mauerer, and C. Silberhorn, Spectral structure and decompositions of optical states, and their applications, *New J. Phys.* **9**, 91 (2007).
- [36] C. Oh, Y. Lim, B. Fefferman, and L. Jiang, Classical simulation of boson sampling based on graph structure, *Phys. Rev. Lett.* **128**, 190501 (2022).
- [37] M. Rudelson, A. Samorodnitsky, and O. Zeitouni, Hafnians, perfect matchings and Gaussian matrices, *Ann. Probab.* **44**, 2858 (2016).
- [38] J. Martínez-Cifuentes, K. Fonseca-Romero, and N. Quesada, Classical models may be a better explanation of the Jiuzhang 1.0 Gaussian Boson Sampler than its targeted squeezed light model, *Quantum* **7**, 1076 (2023).
- [39] P. Erdős and A. Rényi, On the evolution of random graphs, *Publ. Math. Inst. Hung. Acad. Sci.* **5**, 17 (1960).
- [40] J. M. Arrazola, T. R. Bromley, and P. Reberstrost, Quantum approximate optimization with Gaussian boson sampling, *Phys. Rev. A* **98**, 012322 (2018).
- [41] S. Aaronson and A. Arkhipov, BosonSampling is far from uniform, *Quantum Inf. Comput.* **14**, 1383 (2014).