Research Article **Open Access**

Kamil Brádler*, Shmuel Friedland, Josh Izaac, Nathan Killoran, and Daiqin Su

Graph isomorphism and Gaussian boson sampling

https://doi.org/10.1515/spma-2020-0132
Received November 6, 2020; accepted March 14, 2021

Abstract: We introduce a connection between a near-term quantum computing device, specifically a Gaussian boson sampler, and the graph isomorphism problem. We propose a scheme where graphs are encoded into quantum states of light, whose properties are then probed with photon-number-resolving detectors. We prove that the probabilities of different photon-detection events in this setup can be combined to give a complete set of graph invariants. Two graphs are isomorphic if and only if their detection probabilities are equivalent. We present additional ways that the measurement probabilities can be combined or coarse-grained to make experimental tests more amenable. We benchmark these methods with numerical simulations on the Titan supercomputer for several graph families: pairs of isospectral nonisomorphic graphs, isospectral regular graphs, and strongly regular graphs.

Keywords: Gaussian boson sampling, graph isomorphism, hafnian, quantum GI algorithm, strongly regular graph

MSC: 05C50, 05C60, 15A15, 68Q12, 81P68

1 Introduction

The problem of graph isomorphism (GI) lies at an interesting point in the landscape of computational complexity theory. Though algorithms have been recently proposed which run in 'quasi-polynomial' time [3, 27], it is still an open question in theoretical computer science whether there exists a polynomial-time algorithm that can determine whether two graphs are isomorphic; indeed, graph isomorphism is likely to belong to the class of NP-intermediate computational problems. Two other well-known problems which have similar status in the complexity landscape are integer factoring and the discrete logarithm problem. Famously, while no classically efficient algorithm for these two problems is known, they can be solved in polynomial time on quantum computers [36, 37]. Quantum algorithms with a superpolynomial runtime advantage have also been proposed for linear systems [16, 26], semidefinite programming [11, 12], knot invariants [19, 20, 42], and partition functions [2, 23, 40], among many others. Boson sampling is a strong candidate to demonstrate the quantum computational advantage [1]. Zhong et al. measured a sampling rate that is about 1014-fold faster than using state-of-the-art classical simulation strategies and supercomputers [43]

Given these other success cases, it is natural to hypothesize that may also be useful for the graph isomorphism problem.

Over the last several years, several works have explored this problem, with quantum algorithms for tackling graph isomorphism proposed based on quantum annealing [15, 21, 44] and quantum graph states [32].

***Corresponding Author: Kamil Brádler:** ORCA Computing, 84 Wood Lane, London, W12 0BZ, UK, formerly Xanadu, E-mail: kamilbradler@gmail.com
Josh Izaac, Nathan Killoran, Daiqin Su: Xanadu, 777 Bay Street, Suite 2902, Toronto, ON M5G 2C8, Canada, E-mail: josh@xanadu.ai, nathan@xanadu.ai, sudaiqin@gmail.com
Shmuel Friedland: Department of Mathematics, Statistics and Computer Science, University of Illinois, 851 South Morgan Street, Chicago, Illinois 60607-7045, USA, E-mail: friedlan@uic.edu

However, the bulk of quantum algorithm proposals to distinguish non-isomorphic graphs have utilized the time-evolution of a quantum walker to calculate 'graph invariants' or 'graph certificates' which, ideally, produce identical results for two graphs if and only if they are isomorphic. Of the algorithms proposed, they differ mainly in the number of particles involved, the presence of interactions, localised perturbations, and construction of the GI certificate [7, 17, 18, 33, 41]. It has subsequently been proven using this approach that conventional quantum walk algorithms, both discrete-time and continuous-time, with interactions and perturbations, cannot distinguish arbitrary non-isomorphic graphs [31, 33, 34].

To test the distinguishing ability of proposed quantum GI algorithms, a common benchmark has become their capacity to distinguish nonisomorphic strongly regular graphs (SRGs) with the same graph parameters. This provides an analytic approach to investigate the effectiveness of graph isomorphism proposals; if a particular certificate will always fail to distinguish two non-isomorphic SRGs, this can be shown to be because all elements of a certificate, as well as their multiplicities, are functions of SRG family parameters [22].

In this work, we present an approach to graph isomorphism which uses a near-term quantum computational device, namely a photonics-based Gaussian boson sampling apparatus [25, 29]. For this method, graphs are encoded into quantum-optical states of light – specifically Gaussian states – which are then subjected to photon-number-resolving measurements. Mathematically, we show that the resulting measurement outcome probabilities can be combined to give a complete set of graph invariants. Two graphs are isomorphic if and only if these graph invariants are equal. We also present several ways that these measurement probabilities can be combined and coarse-grained to obtain new quantities which can be used to distinguish nonisomorphic graphs. Finally, we perform classical numerical simulations of our proposed method on the Titan supercomputer. Using these results, we are able to distinguish 3852 out of 3854 nonisomorphic graphs using only a subset of measurement events. The remaining two graphs were distinguished by failing to satisfy a necessary condition introduced here as well.

2 Main results summarized

Our main result is a necessary and sufficient condition to distinguish isospectral nonisomorphic graphs by virtue of comparing the probabilities of the measurement patterns of the graphs encoded in a Gaussian boson sampling (GBS) apparatus. We discovered the vital role played by a matrix function called the hafnian [14], applied to an adjacency matrix, for the GI problem. It leads to a complete set of graph invariants. The hafnian belongs to the family of matrix functions such as the determinant, permanent and pfaffian [6]. It has been established that photon detection probabilities can be expressed in terms of the hafnians of a collection of graphs related to the original graph [9]. Multiphoton detection probabilities are handled by introducing a new matrix product related to the Kronecker product and by showing how the output probabilities depend on the hafnian of the graph adjacency matrix as well. We further strengthen our graph invariant results by introducing the so-called symmetrized graphs invariants and showing that they correspond to coarse-grained measurement events in GBS. The measurement events are given by the stratification according to the total photon number and partitioned into the orbits of the permutation group. Their hafnian-based coarse-grained probabilities are again sufficient to distinguish isospectral nonisomorphic graphs. We extend these insights by deriving necessary conditions for isospectral graphs to be isomorphic by comparing the coarse-grained partition-averaged photon distribution from the Gaussian boson sampler.

Our method differs from previous quantum GI algorithm proposals. A great majority have utilized quantum walks, either using discrete-time quantum walks (DTQWs) [7, 18] or continuous-time quantum walks (CTQWs) [22, 31, 33, 34]. Although the graph invariants constructed via quantum walk propagation on graph structures have shown success in distinguishing various families of SRGs, it has been proven that this distinguishing power is not universal — there will always exist graphs which (current) quantum walk-based algorithms cannot distinguish [38]. In order to execute a quantum walk-based algorithm in a universal quantum photonic platform, it is necessary to implement a non-Gaussian operation as a vertex-dependent shift or via multiple interacting walkers. This is a major obstacle with the current and near-term technology our

proposal does not suffer from. GBS is a Gaussian circuit followed by an array of photon-number-resolving detector (PNR) representing a non-Gaussian element. Unlike non-Gaussian unitary transformations, the PNRs are available in the state-of-the-art laboratories.

The most comprehensive simulations of quantum methods for GI were performed in [17] and [33]. We successfully tested three types of isospectral graphs: pairs of isospectral nonisomorphic graphs (PINGs) as the first examples of such graphs [5], isospectral regular graphs [30] and mainly SRGs. There are numerous resources available detailing the SRG families containing more than one non-isomorphic graph [13, 39]; as a result, SRGs have become a common benchmark in studying the distinguishing powers of the GI algorithms. Note that there may be other graph classes (such as k-equivalent graphs) which have been proven to be harder to distinguish than strongly regular graphs for particular quantum GI algorithms [38] — however, SRGs remain an ideal testing set, simply due to the large number of relatively small non-isomorphic graphs present in specific families [13, 39]. The largest tested and distinguished family by our approach was SRG(35,18,9,9) containing 3854 isospectral graphs. This family is supposedly tested in [33]. However, the size of the family is mistakenly taken to be only 227 graphs (see Table I.). The same error appears in [17]. Ironically, another SRG family considered there (SRG(35,16,16,8)), that happens to be complementary to SRG(35,18,9,9) and thus containing 3854 graphs as well, is counted properly and analyzed (see Table 1.).

Section 3 contains all necessary definitions and previous results used in the paper including a detailed GBS description and a formal introduction of SRGs. Section 4 contains the main result and is split into four subsections: In 4.1 we gather several supporting results followed by the main results in Sections 4.2, 4.3 and 4.4. Section 5 contains the simulation results and Section 6 concludes with a scalability discussion and other open questions. In Appendix A we informally introduce the hardware setup (Gaussian boson sampler) where studied graphs are encoded. In Appendix B we present the GBS quantum GI algorithm applied to various SRG families and other isospectral graphs. In Appendix C we summarize with a table the most important symbols and their meaning.

3 Notation and preliminaries

In the following text the symbol $\mathbb{J}_{k,\ell}$ denotes an all-ones rectangular matrix of size $k \times \ell$ and $\mathbb{J}_k \equiv \mathbb{J}_{k,k}$. The following notation is extensively used: $\partial_x^n \equiv \partial_{x,\dots,x} = \frac{\partial^n}{\partial x^n}$ and $\partial_{x_i,\bar{x}_i}^{n_i} = \frac{\partial^{n_i}}{\partial x_i^{n_i}} \frac{\partial^{n_i}}{\partial \bar{x}_i^{n_i}}$. Letting $\boldsymbol{n} = (n_1, \dots, n_M)$, $\boldsymbol{x} = (x_1, \dots, x_M)$, we occasionally write $\prod_{i=1}^M \partial_{x_i}^{n_i} = \partial_{\boldsymbol{x}}^{|\boldsymbol{n}|}$ and $\prod_{i=1}^M \partial_{x_i,\bar{x}_i}^{n_i} = \partial_{\boldsymbol{x},\bar{\boldsymbol{x}}}^{|\boldsymbol{n}|}$. The symbol $\stackrel{\text{df}}{=}$ stands for 'defined' and a positive-definite matrix A will be denoted by $A \succ 0$. Recall that any Gaussian n-dimensional real distribution with zero mean, denoted as G_Σ, is given by

$$\frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{\det \Sigma}} \exp\left[-\frac{1}{2}\boldsymbol{x}^\top \Sigma^{-1} \boldsymbol{x}\right].$$

Here, Σ is a positive definite matrix which is the covariance matrix of the Gaussian variables $X_1, \dots, X_n$.

Since Σ is positive definite, there exists a unique positive definite matrix A such that $\Sigma = A^2$. Let us change the variables $\boldsymbol{y} = A^{-1}\boldsymbol{x}$. That is, $\boldsymbol{x} = A\boldsymbol{y}$. Hence the determinant of the Jacobian is $\det A$. As $\det \Sigma = (\det A)^2$ we get that the density distribution for $(Y_1, \dots, Y_n)$ is the standard normal density distribution $\frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left[-\frac{1}{2}\boldsymbol{y}^\top \boldsymbol{y}\right]$. Therefore $Y_1, \dots, Y_n$ are independent standard random variables. Assume that $A = [a_{ij}]$ is a positive definite symmetric matrix. Then

$$X_i = \sum_{j=1}^n a_{ij} Y_j, \quad i \in [n]$$

and

$$\mathbb{E}[X_i X_j] = \mathbb{E}\Big[\Big(\sum_{p=1}^n a_{ip} Y_p\Big)\Big(\sum_{q=1}^n a_{jq} Y_q\Big)\Big] = \sum_{p,q=1}^n a_{ip} a_{iq} \mathbb{E}[Y_p Y_q] = \sum_{p=1}^n a_{ip} a_{jp} = \Sigma_{ij}.$$

Observe the well known fact that the odd moments $\mathbb{E}[\prod_{i=1}^{n} X^{m_i}]$, where $(m_1, \ldots, m_n) \in \mathbb{Z}_+^n$ and $\sum_{i=1}^{n} m_i$ is odd, are zero. A polynomial $p(\boldsymbol{x})$, $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ is called symmetric if for each permutation $\sigma : [n] \to [n] = (1, \ldots, n)$ we have the equality $p(\boldsymbol{x}) = p(\boldsymbol{x}_\sigma)$, where $\boldsymbol{x}_\sigma = (x_{\sigma(1)}, \ldots, x_{\sigma(n)})$.

Denote by $\mathfrak{S}_n$ the symmetric group of bijections $\sigma : [n] \to [n]$. Denote by $\mathcal{P}_n \subset \mathbb{R}^{n\times n}$ the group of $n \times n$ permutation matrices. So $P(\sigma)\boldsymbol{x} = \boldsymbol{x}_\sigma$.

Recall that two square matrices A, B are permutationally similar, if $B = PAP^\top$, where P is a permutation matrix. In this case $P^{-1} = P^\top$. Two Gaussian distributions corresponding to positive definite covariance matrices $\Sigma, \Sigma' \in \mathbb{R}^{n\times n}$ are called isomorphic, if $\Sigma' = P^\top \Sigma P$ for some permutation $P = P(\sigma)$. That is $\boldsymbol{x}^\top (\Sigma')^{-1}\boldsymbol{x} = \boldsymbol{x}_\sigma^\top (\Sigma)^{-1}\boldsymbol{x}_\sigma$, where $\sigma \in \mathfrak{S}_n$, for all $\boldsymbol{x} \in \mathbb{R}^n$.

Denote by $\mathcal{H}_N \supset \mathcal{H}_{+,N} \supset \mathcal{H}_{++,N}$ the real space of $N \times N$ hermitian matrices, the closed cone of positive semidefinite hermitian matrices, and the open set of positive definite hermitian matrices. For $F \in \mathcal{H}_N$ denote by $\lambda_1(F) \geq \cdots \geq \lambda_N(F)$ the N eigenvalues of F. Recall that the spectral norm of F is given by $\|F\|_2 = \max(\lambda_1(F), -\lambda_N(F))$. For $X, Y \in \mathcal{H}_N$ we denote $X \preceq Y$ and $X \prec Y$ if $Y - X \in \mathcal{H}_{+,N}$ or $Y - X \in \mathcal{H}_{++,N}$, respectively.

3.1 Gaussian Boson Sampling

Definition 1. Let $C = [c_{ij}] \in \mathbb{R}^{2M\times 2M}$ be a symmetric matrix and let $\mathcal{M}_{2M}$ denote all partitions ς to unordered disjoint pairs. Then

$$\operatorname{haf} C \stackrel{\text{df}}{=} \sum_{\varsigma\in\mathcal{M}_{2M}} \prod_{(uv)\in\varsigma} c_{uv} \tag{1}$$

is the hafnian of C [14].

Given a detection event $\boldsymbol{n}$, its measurement probability was shown in [25] to be

$$p(\boldsymbol{n}) = \frac{1}{\boldsymbol{n}!\sqrt{\det \sigma_Q}} \partial_{\boldsymbol{\beta},\overline{\boldsymbol{\beta}}}^{|\boldsymbol{n}|} e^{\frac{1}{2}\gamma^\top C\gamma}\Big|_{\gamma=0}, \tag{2}$$

where $\boldsymbol{n}! = n_1! \times \cdots \times n_M!$ and $\gamma \stackrel{\text{df}}{=} (\boldsymbol{\beta}, \overline{\boldsymbol{\beta}}) = (\beta_1, \ldots, \beta_M, \overline{\beta}_1, \ldots, \overline{\beta}_M) \in \mathbb{C}^{2M}$ which we view as a column vector (even though $\overline{\boldsymbol{\beta}}$ is a complex conjugate of a complex $\boldsymbol{\beta}$ (entrywise), we consider $\overline{\beta}_i$ as a new variable). We denote

$$X_{2M} = \begin{bmatrix} 0 & \mathbb{I}_M \\ \mathbb{I}_M & 0 \end{bmatrix}. \tag{3}$$

Then

$$\sigma_Q = (\mathbb{I}_{2M} - X_{2M}C)^{-1} \tag{4}$$

and $\sigma = \sigma_Q - \mathbb{I}_{2M}/2$ is the covariance matrix in Eq. (62).

Note that in order for σ to be a covariance matrix, C has to satisfy certain restrictions which will be elaborated on later. We call a GBS detection event corresponding to the measurement pattern $\boldsymbol{n} \stackrel{\text{df}}{=} (n_1, \ldots, n_M)$ of an M-mode matrix C of size $2M$ *pure* if $n_i = n_j$, $\forall i, j$ and *mixed* otherwise. For $\boldsymbol{n} = (1, \ldots, 1)$ Eq. (2) reduces to (63) [25].

3.2 Graphs: eigenvalues, isospectral graphs, strongly regular graphs and graph isomorphism

We now recall briefly some well known results on graphs that we use in this paper. An undirected simple graph consist of a set of n-vertices $V = \{v_1, \ldots, v_n\}$ which we identify with $[n]$. The set of edges $E(G) = E = \{e_1, \ldots, e_m\}$ is the set of unordered pairs $e = (u, v)$ where $u \neq v \in V$. We say that e connects u and v, or e is adjacent to u and v. A simple path in G is an ordered subset of E: $(u_1, u_2), \ldots, (u_k, u_{k+1})$. A graph G is called connected if for each pair of vertices $u \neq v$ there is simple path such that $u_1 = u$ and $u_{k+1} = v$.

A compete graph on n-vertices is denoted by K_n, so the cardinality of its edges is $|E(K_n)| = n(n-1)/2$. Given a simple graph G on n-vertices then $E(G)$ is a subset of $E(K_n)$. Hence $m \leq n(n-1)/2$. The complement of G, denoted as G^c, is a graph on $[n]$ vertices with the edges $E(G^c) = E(K_n) \setminus E(G)$. Thus the complement of K_n is a graph with no edges, the null graph. Given a subset $W \subset V$ the induced subgraph $G(W) = (W, E(G(W))$, where $E(G(W))$ is the subset of edges in V that connect two vertices in W. A clique is a subgraph $G(W)$ which is a complete graph on W. A graph G is called regular (or k-regular) if each vertex $v \in V$ is adjacent to exactly k-edges. A graph G is called bipartite if V is a union of disjoints subsets of vertices V_1, V_2 where the edges E connect vertices in V_1 to vertices in V_2.

The adjacency matrix $A(G) = A = [a_{ij}]$ is an $n \times n$ symmetric matrix with zero diagonal whose off diagonal entries are zero or one. Then $a_{ij} = 1$ if and only if $(i, j) \in E(G)$. If we relabel the vertices of G, i.e., apply a bijection $\sigma : [n] \to [n]$, then the new adjacency matrix $\tilde{A}$ is $PAP^\top$ for some permutation matrix $P \in \mathcal{P}_n$. As A is real symmetric it has n real eigenvalues, counted with their multiplicities: $\lambda_1(G) = \lambda_1 \geq \cdots \geq \lambda_n(G) = \lambda_n$. As A is a nonnegative matrix the Perron-Frobenius theorem yields that $\lambda_1 \geq |\lambda_n|$. If G is connected equality holds if and only if G is bipartite. If G is connected then $\lambda_1 > \lambda_2$, that is, λ_1 is a simple eigenvalue of G. Furthermore, if G is a k-regular graph then $\lambda_1 = k$. Since $\mathcal{P}_n$ is a subgroup of the group of orthogonal matrices, it follows that the eigenvalues of G do not depend on the labeling of the vertices of G.

A graph H on n-vertices is called isomorphic to G, if H is a relabeling G. That is, if $A(H) = PA(G)P^\top$ for some $P \in \mathcal{P}_n$. Thus a necessary conditions for two graph to be isomorphic is to be isospectral, i.e., to have the same sequence of eigenvalues. That is, $A(G)$ and $A(H)$ have the same characteristic polynomial. Hence we can check in polynomial time if G and H are isospectral. As we pointed out in Introduction, the problem of graph isomorphism (GI) lies at an interesting point in the landscape of computational complexity theory.

In studying the graph isomorphism problem, it is convenient to consider a class of graphs known to be classically intractable to distinguish. An important tractable feature is the graph eigenvalues and the first examples of isospectral graphs were pairs of isospectral nonisomorphic graphs (PINGs) [5]. The smallest connected example of a PING is on six vertices, see Fig. 1. PINGs may have some tractable features enabling one

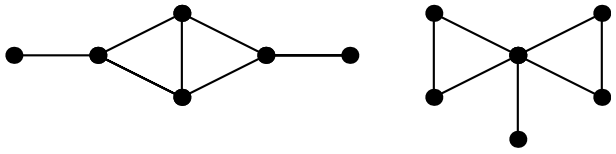

Figure 1: PING on six vertices.

to easily decide that they are not isomorphic.

One of best known positive results in graph isomorphism is the following result of Babai-Grigoryev-Mount [4]: Let k be a fixed integer. Assume that G and H are isospectral, and the multiplicity of each eigenvalue is at most k. Then there exists a polynomial time algorithm that decides if G and H are isomorphic.

The situation gets complicated for graphs with symmetries such as strongly regular graphs (SRGs), defined as follows [24].

Definition 2. Let $G(V, E)$ be a regular graph of degree k consisting of N vertices and adjacency matrix A, that is neither a complete graph ($A \neq \mathbb{J}_N - \mathbb{I}_N$) nor a null graph. G is then said to be strongly regular with parameters SRG(N, k, λ, μ) if every pair of adjacent vertices have exactly λ common neighbours, and every pair of *non*-adjacent vertices have exactly μ common neighbours.

Recall also that an SRG graph has three distinct eigenvalues: simple eigenvalue $\lambda_1 = k$, and other two eigenvalues with high multiplicity: at least one eigenvalue of multiplicity at least $(n-1)/2$.

Lemma 1. *Let $v_i \in V$ be a vertex in SRG(N, k, λ, μ). Then $k(k-\lambda-1) = \mu(N-k-1)$. Thus the SRG parameters are not independent.*

Proof. Consider a vertex v in a graph with parameters $SRG(N, k, \lambda, \mu)$, and count in two ways the number of edges from vertices adjacent to v to vertices non-adjacent to v. □

If multiple non-isomorphic strongly regular graphs share the same set of SRG parameters, we refer to this graph set as an *SRG family*, often simply denoted by the SRG parameters (N, k, λ, μ). Graphs within the same SRG family share various properties that are dependent only on the SRG parameters. The spectrum is rather special, consisting of just three eigenvalues with known multiplicities. SRG families contain non-isomorphic graphs which are isospectral, and difficult to distinguish using common classical measures [24].

4 Graph isomorphism via Gaussian boson sampling

4.1 Multiphoton contributions in GBS – supporting results

We start with the exploration of how to interpret Eq. (2) for the detection events $\boldsymbol{n}$ where $n_i > 1$ for some i. This corresponds to a multiphoton contribution of the output probability function. The multiphoton contributions play a vital role in our analysis.

Definition 3. Assume that $A = [a_{ij}] \in \mathbb{R}^{m_1 \times m_2}$ and in total $m_1 \times m_2$ matrices $B = [B_{ij}]$ where each B_{ij} is an $n_i \times n_j$ matrix. Then the *reduced Kronecker product* $C = A \not\otimes B$ will denote a block matrix C partitioned as B and the blocks are $C = [a_{ij}B_{ij}]$.

Remark. The dimension of $A \not\otimes B$ is the dimension of the matrix B. Then $A \not\otimes B$ is a submatrix of the Kronecker tensor product of two matrices $A \otimes B = [a_{ij}B]$. The matrix B in this paper will always be assembled from $B_{ij} = \mathbb{J}_{n_i,n_j} \in \mathbb{R}^{n_i \times n_j}$ where $\boldsymbol{n}$ is a measurement pattern. In this case we will write $B = \mathbb{J}_{|\boldsymbol{n}|} \in \mathbb{R}^{|\boldsymbol{n}| \times |\boldsymbol{n}|}$ where $|\boldsymbol{n}| = \sum_{i=1}^{M} n_i$. Note that if $\boldsymbol{n}$ is a pure event then the reduced Kronecker product $\not\otimes$ becomes the ordinary Kronecker product $A \otimes \mathbb{J}_{n_1}$. Also note that if $\dim B_{ij} = 1, \forall i, j$ the reduced Kronecker product becomes the Hadamard (Schur) product.

Example. Let $\boldsymbol{n} = (3, 2, 1, 4)$ and A an adjacency matrix of a simple weighted graph (without loops). Then

$$A \not\otimes \mathbb{J}_{|\boldsymbol{n}|} = \left[\begin{array}{ccc|cc|c|cccc} & & & a_{12} & a_{12} & a_{13} & a_{14} & a_{14} & a_{14} & a_{14} \\ & 0 & & a_{12} & a_{12} & a_{13} & a_{14} & a_{14} & a_{14} & a_{14} \\ & & & a_{12} & a_{12} & a_{13} & a_{14} & a_{14} & a_{14} & a_{14} \\ \hline a_{12} & a_{12} & a_{12} & & & a_{23} & a_{24} & a_{24} & a_{24} & a_{24} \\ a_{12} & a_{12} & a_{12} & & 0 & a_{23} & a_{24} & a_{24} & a_{24} & a_{24} \\ \hline a_{13} & a_{13} & a_{13} & a_{23} & a_{23} & 0 & a_{34} & a_{34} & a_{34} & a_{34} \\ \hline a_{14} & a_{14} & a_{14} & a_{24} & a_{24} & a_{34} & & & & \\ a_{14} & a_{14} & a_{14} & a_{24} & a_{24} & a_{34} & & 0 & & \\ a_{14} & a_{14} & a_{14} & a_{24} & a_{24} & a_{34} & & & & \\ a_{14} & a_{14} & a_{14} & a_{24} & a_{24} & a_{34} & & & & \end{array}\right].$$

The reason for introducing a new kind of structure is a compact expression for the probability of measurement of a mixed multiphoton event $\boldsymbol{n}$ as a hafnian function not unlike Eq. (63) for $\boldsymbol{n} = (1, \ldots, 1)$.

Definition 4. A $2M \times 2M$-dimensional real symmetric matrix R will be called *GBS encodable* if we can find a covariance matrix σ_Q such that

$$R = X_{2M}\left(\mathbb{I}_{2M} - \sigma_Q^{-1}\right). \tag{5}$$

Ref. [9] introduced a necessary criterion for R to be GBS encodable. For some real symmetric $\tilde{R}$ not satisfying the conditions a general procedure was created to produce a matrix related to $\tilde{R}$ that is GBS-encodable. It consists of taking $\tilde{R} \mapsto R \stackrel{\text{df}}{=} c(\tilde{R} \oplus \tilde{R})$ where $0 < c < 1/\|\tilde{R}\|_2$.

Even though this procedure is always guaranteed to succeed in creating a Gaussian covariance matrix, it is not a necessary condition. Here we strengthen this previous result by loosening the requirements on R.

Lemma 2. *Let $G \in \mathcal{H}_N$ and assume that $G = (\mathbb{I}_N - F)^{-1} - \frac{1}{2}\mathbb{I}_N$. Then*

1. *$G \succ 0$ if and only if $\|F\|_2 < 1$.*
2. *Suppose that $N = 2M$, $F = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix}$. Then the following conditions hold*

$$G \succ 0, \quad G + \frac{1}{2}\begin{bmatrix} \mathbb{I}_M & 0 \\ 0 & -\mathbb{I}_M \end{bmatrix} \succeq 0 \tag{6}$$

if and only if $\|F\|_2 < 1$ and $F_{22} \succeq 0$.

Proof.

1. Clearly we need to have that $(\mathbb{I}_N - F)^{-1} \succ 0$. This equivalent to $\lambda_1(F) < 1$. The assumption that $(\mathbb{I}_N - F)^{-1} \succ \frac{1}{2}\mathbb{I}_N$ is equivalent to $\lambda_N((\mathbb{I}_N - F)^{-1}) > \frac{1}{2}$. This is equivalent to $\lambda_N(F) > -1$. Therefore the claim follows.
2. Use (1) to get the assumption that $G \succ 0$ is equivalent to $\|F\|_2 < 1$. The assumption that $G + \frac{1}{2}\begin{bmatrix} \mathbb{I}_M & 0 \\ 0 & -\mathbb{I}_M \end{bmatrix} \succeq 0$ is equivalent to $(\mathbb{I}_N - F)^{-1} \succeq \begin{bmatrix} 0 & 0 \\ 0 & \mathbb{I}_M \end{bmatrix}$. This is equivalent to

$$\mathbb{I}_N \succeq (\mathbb{I}_N - F)^{\frac{1}{2}} \begin{bmatrix} 0 & 0 \\ 0 & \mathbb{I}_M \end{bmatrix} (\mathbb{I}_N - F)^{\frac{1}{2}} \succeq 0.$$

This inequality is equivalent to

$$1 \geq \lambda_1\left((\mathbb{I}_N - F)^{\frac{1}{2}} \begin{bmatrix} 0 & 0 \\ 0 & \mathbb{I}_M \end{bmatrix} (\mathbb{I}_N - F)^{\frac{1}{2}}\right) = \lambda_1\left((\mathbb{I}_N - F) \begin{bmatrix} 0 & 0 \\ 0 & \mathbb{I}_M \end{bmatrix}\right).$$

Observe next that

$$(\mathbb{I}_N - F) \begin{bmatrix} 0 & 0 \\ 0 & \mathbb{I}_M \end{bmatrix} = \begin{bmatrix} 0 & -F_{12} \\ 0 & \mathbb{I}_M - F_{22} \end{bmatrix}.$$

Hence

$$1 \geq \lambda_1\left((\mathbb{I}_N - F) \begin{bmatrix} 0 & 0 \\ 0 & \mathbb{I}_M \end{bmatrix}\right) = 1 - \lambda_M(F_{22})$$

is equivalent to $\lambda_M(F_{22}) \geq 0$, that is $F_{22} \succeq 0$.

□

Corollary 3. *Let $R \in \mathbb{R}^{2M \times 2M}$ be a nonzero real symmetric matrix with the following partition to $M \times M$ blocks: $R = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$. Then there exists a Gaussian covariance matrix σ such that $cR = X_{2M}[\mathbb{I}_{2M} - (\sigma + \frac{1}{2}\mathbb{I}_{2M})]$ if and only if:*

1. *$R_{11} = R_{22}$ and $R_{12} = R_{21}$.*
2. *$R_{12} \succeq 0$.*
3. *$c \in (0, \frac{1}{\|R\|_2})$.*

Part *(3)* follows from the observation the singular values of R and $X_{2M}R$ are the same. Hence $\|R\|_2 = \|X_{2M}R\|_2$.

Remark. The previous lemma was presented for the sake of completeness. In the rest of the paper we use the construction of GBS-encodable matrices introduced already in [9].

For future reference we recall the following straightforward result [9].

Lemma 4. *Let M be even and $C = A \oplus A$ a real symmetric matrix of dimension $2M \times 2M$. Then*

$$\operatorname{haf}[c(C + k\mathbb{I}_{2M})] = c^M \operatorname{haf}^2 A, \tag{7}$$

where $c > 0$ and $k \in \mathbb{R}$.

Lemma 5. *Assume $C = [c_{ij}] \in \mathbb{R}^{2M\times 2M}$ is a symmetric and GBS-encodable matrix. Then the probability of sampling in the GBS event, Eq.* (2)*, can be expressed as*

$$p(\boldsymbol{n}) = \frac{1}{\boldsymbol{n}!\sqrt{\det \sigma_Q}} \operatorname{haf}[C\not{\oplus}\mathbb{J}_{2|\boldsymbol{n}|}]. \tag{8}$$

Proof. To prove this equality we assume that $C = A \oplus A$ where A is an arbitrary symmetric matrix of order M. Let $\tilde{p}(n)$ be defined by the right-hand side of (2). We will show that $\tilde{p}(\boldsymbol{n}) = p(\boldsymbol{n})$. Consider $\mathbb{J}_{2|\boldsymbol{n}|}$ as a $2M \times 2M$ block matrix $[\mathbb{J}_{n_i,n_j}]$ where $n_{i+M} = n_i$. Hence $C\not{\oplus}\mathbb{J}_{2|\boldsymbol{n}|} = A\not{\oplus}\mathbb{J}_{|\boldsymbol{n}|} \oplus A\not{\oplus}\mathbb{J}_{|\boldsymbol{n}|}$. We now consider the quadratic form $\gamma^\top(C\not{\oplus}\mathbb{J}_{2|\boldsymbol{n}|})\gamma$ and focus on $\boldsymbol{\beta}^\top(A\not{\oplus}\mathbb{J}_{|\boldsymbol{n}|})\boldsymbol{\beta}$ (the 'second half' is treated in exactly the same way). We substitute

$$\beta_i \mapsto (\alpha_{i1}, \ldots, \alpha_{in_i}). \tag{9}$$

in (2). We define its action for the quadratic form to be

$$\boldsymbol{\beta}^\top A\boldsymbol{\beta} = \sum_{ij} \beta_i\beta_j a_{ij} \mapsto \sum_{ij}(\alpha_{i1} + \cdots + \alpha_{in_i})(\alpha_{j1} + \cdots + \alpha_{jn_j})a_{ij} = \boldsymbol{\alpha}^\top[a_{ij}\mathbb{J}_{n_in_j}]\boldsymbol{\alpha} \equiv \boldsymbol{\alpha}^\top(A\not{\oplus}\mathbb{J}_{|\boldsymbol{n}|})\boldsymbol{\alpha}, \tag{10}$$

where we used Def. 3 by setting $B_{ij} = \mathbb{J}_{n_i,n_j}$. We further set $\partial_{\beta_i}^{n_i} = \partial_{\alpha_{i1},\ldots,\alpha_{in_i}}$, $\forall i$ (and similarly for the conjugated variables $\overline{\beta}_i$ and the corresponding $\overline{\alpha}_{in_i}$) and write

$$\tilde{p}(\boldsymbol{n}) = \frac{1}{\boldsymbol{n}!\sqrt{\det \sigma_Q}} \prod_{k=1}^{M} \partial_{\alpha_{k1},\ldots,\alpha_{kn_k}} e^{\frac{1}{2}\boldsymbol{\alpha}^\top(A\not{\oplus}\mathbb{J}_{|\boldsymbol{n}|})\boldsymbol{\alpha}}\Big|_{\boldsymbol{\alpha}=0}. \tag{11}$$

It remains to show that $p(\boldsymbol{n}) = \tilde{p}(\boldsymbol{n})$. Indeed, the higher-order partial derivatives in (2) result in the same expression as the first-order ones in (11) whenever we set $\boldsymbol{\alpha} = \boldsymbol{\beta} = 0$ at the end of the calculation. This is a consequence of the elementary properties of the differential operator, namely,

$$\partial_{x_{k1},\ldots,x_{k1}} f(\textstyle\sum_{\ell=1}^{n_k} x_{k\ell}) = \partial_{x_{k1},\ldots,x_{kn_k}} f(\sum_{\ell=1}^{n_k} x_{k\ell}),\ \forall k \tag{12}$$

and the chain rule for the n-th derivative given by Faà di Bruno's formula for $(f \circ g)^{(n)}(x)$ in the special case of $g(x) = x + K$ where K is a constant:

$$f^{(n)}(x + K) = h(x + K) \tag{13}$$

whenever $f^{(n)}(x) \overset{\text{df}}{=} h(x)$. Now we put all the pieces together. The RHS of (12) is identified with (11) through $(x_{k1}, \ldots, x_{kn_k}) \mapsto (\alpha_{k1}, \ldots, \alpha_{kn_k})$ (or its conjugate) for a given $1 \le k \le M$ and so $f(\sum_{k=1}^{M}\sum_{\ell=1}^{n_k}\alpha_{k\ell}) = e^{\frac{1}{2}\boldsymbol{\alpha}^\top(A\not{\oplus}\mathbb{J}_{|\boldsymbol{n}|})\boldsymbol{\alpha}}$. But then, according to the LHS of (12) we may write (11) as

$$\tilde{p}(\boldsymbol{n}) = \frac{1}{\boldsymbol{n}!\sqrt{\det \sigma_Q}} \prod_{k=1}^{M} \partial_{\alpha_{k1},\ldots,\alpha_{k1}} e^{\frac{1}{2}\boldsymbol{\alpha}^\top(A\not{\oplus}\mathbb{J}_{|\boldsymbol{n}|})\boldsymbol{\alpha}}\Big|_{\boldsymbol{\alpha}=0}. \tag{14}$$

The RHS of (14) is identified with the LHS of (13) by setting $x = \alpha_{k1}$, $n = n_k$ and $K = \sum_{\ell=2}^{n_k}\alpha_{k\ell}$ for a given k. Since

$$h(\beta_1) = \frac{\mathrm{d}^n}{\mathrm{d}\beta_1^n} e^{\frac{1}{2}\boldsymbol{\beta}^\top A\boldsymbol{\beta}}$$

forms $p(\boldsymbol{n})$ and from (11) we get

$$\prod_{k=1}^{M} \partial_{\alpha_{k1},\ldots,\alpha_{kn_k}} e^{\frac{1}{2}\boldsymbol{\alpha}^\top(A\not{\oplus}\mathbb{J}_{|\boldsymbol{n}|})\boldsymbol{\alpha}}\Big|_{\boldsymbol{\alpha}=0} = \operatorname{haf}[A\not{\oplus}\mathbb{J}_{|\boldsymbol{n}|}]$$

we may conclude that $\tilde{p}(\boldsymbol{n}) = p(\boldsymbol{n})$ due to $f^{(n)}(\beta_1)|_{\beta_1=0} = f^{(n)}(\alpha_{k1} + K)|_{\alpha_{k1}=K=0}$. This follows from (13) and the definition of $h(x)$. □

Interestingly, many of the detection events have probability zero:

Lemma 6. *Let $C = c(A \oplus A) \in \mathbb{R}^{2M\times 2M}$ and let $\boldsymbol{n}$ be a detection event where $|\boldsymbol{n}| = \sum_{i=1}^{M} n_i$. If there is any $n_i > |\boldsymbol{n}|/2$ then $p(\boldsymbol{n}) = 0$.*

Proof. Here we may assume $c = 1$ and so $C = A \oplus A$. The probability expression (Eq. (8)) contains haf $[A\not\oplus\mathbb{J}_{|\boldsymbol{n}|}]$ and if $n_i > |\boldsymbol{n}|/2$ then $A\not\oplus\mathbb{J}_{|\boldsymbol{n}|}$ contains a zero matrix of size greater than $\lfloor M/2 \rfloor$ (placed in the lower right corner of $A\not\oplus\mathbb{J}_{|\boldsymbol{n}|}$), see Definition 3. The hafnian of such a matrix is zero. This follows from the hafnian definition (Def. 1) where the hafnian of a $2M \times 2M$ matrix is a sum of products of M entries a_{ij}. Since $i < j$ and none of the indices repeats for any summand then inevitably at least one of the a_{ij}'s in every summand equals zero. □

We prove several useful properties of the reduced Kronecker product $\not\oplus$.

Lemma 7. *Let $C = [c_{ij}] \in \mathbb{R}^{2M\times 2M}$ and $c \in \mathbb{R}$. Then*

$$\text{haf}\,[(cC)\not\oplus\mathbb{J}_{2|\boldsymbol{n}|}] = c^{|\boldsymbol{n}|}\,\text{haf}\,[C\not\oplus\mathbb{J}_{2|\boldsymbol{n}|}]$$

where $|\boldsymbol{n}| = \sum_{i=1}^{M} n_i$.

Proof. From Def. 3 we find $(cC)\not\oplus\mathbb{J}_{2|\boldsymbol{n}|} = c(C\not\oplus\mathbb{J}_{2|\boldsymbol{n}|})$. This trivially follows from $[(cc_{ij})B_{ij}] = [c(c_{ij}B_{ij})]$. We then observe that $\dim\,[C\not\oplus\mathbb{J}_{2|\boldsymbol{n}|}] = 2|\boldsymbol{n}|$ and the claim follows. □

Remark. For $n_i = 1, \forall i$ we recover haf $[cC] = c^M$ haf C since $|\boldsymbol{n}| = M$.

Lemma 8. *Let A, B, C be matrices and assume that $A\not\oplus C$ and $B\not\oplus C$ are defined. Then $(A \oplus B)\not\oplus(C \oplus C) = (A\not\oplus C) \oplus (B\not\oplus C)$.*

Proof. The ordinary Kronecker product satisfies $(A \oplus B) \otimes C' = A \otimes C' \oplus B \otimes C'$. By removing columns and the corresponding rows from A, B on both sides of the expression we arrive at the claimed result. □

Lemma 9. *Let P_π be a permutation matrix. Then the following diagram commutes for any matrix A and a detection event $\boldsymbol{n}$*

$$\begin{array}{ccc} A & \xrightarrow{\not\oplus\mathbb{J}_{|\boldsymbol{n}|}} & A\not\oplus\mathbb{J}_{|\boldsymbol{n}|} \\ {\scriptstyle P_\pi}\downarrow & & \downarrow{\scriptstyle \hat{P}_\pi} \\ \tilde{A} & \xrightarrow{\not\oplus\mathbb{J}_{|\boldsymbol{m}|}} & \tilde{A}\not\oplus\mathbb{J}_{|\boldsymbol{m}|} \end{array}$$

where $\hat{P}_\pi$ is another permutation matrix and $\boldsymbol{m} = \pi(\boldsymbol{n})$.

Proof. Following the lower route, the spanning basis $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_k)$ of A becomes $\pi(\boldsymbol{\beta}) = (\beta_{\pi(1)}, \ldots, \beta_{\pi(k)})$ for $\tilde{A}$. The new basis is then expanded by considering

$$\beta_{\pi(i)} \mapsto (\alpha_{\pi(i)1}, \ldots, \alpha_{\pi(i)m_{\pi(i)}}) \tag{15}$$

and so $\pi(\boldsymbol{\alpha})$ is a spanning basis of $\tilde{A}\not\oplus\mathbb{J}_{|\boldsymbol{m}|}$. Going through the upper route, we observe that $\boldsymbol{\beta} \mapsto \boldsymbol{\alpha}$ by the action of

$$\beta_i \mapsto (\alpha_{i1}, \ldots, \alpha_{in_i}). \tag{16}$$

Then, a permutation matrix exists transforming (16) into (15). Its construction is straightforward. The reordering (permutation) $(\alpha_{i1}, \ldots, \alpha_{in_i}) \mapsto (\alpha_{\pi(i)1}, \ldots, \alpha_{\pi(i)n_{\pi(i)}})$ is followed by setting $n_{\pi(i)} = m_{\pi(i)}$. Naturally, the overall transformation is an action of a permutation matrix we denoted by $\hat{P}_\pi$. □

We use another result from [9] to prove the following lemma.

Lemma 10. *The matrices σ_{Q,G_i} of two isospectral graphs G_1, G_2 encoded as adjacency matrices A_1, A_2 of dimension $2M$ satisfy*

$$\det \sigma_{Q,G_1} = \det \sigma_{Q,G_2}.$$

Proof. In order to encode an arbitrary graph we take two copies of a graph's adjacency matrix [9]. Also, it is advantageous to rewrite (4) as

$$\sigma_{A_i} = \frac{1}{2}(\mathbb{I}_{2M} + X_{2M}A_i)(\mathbb{I}_{2M} - X_{2M}A_i)^{-1} = \frac{1}{2}(\mathbb{I}_{2M} - X_{2M}A_i)^{-1}(\mathbb{I}_{2M} + X_{2M}A_i). \tag{17}$$

The matrix A_i commutes with X_{2M} [9] and so the eigenvalues of $X_{2M}A_i$ are products of eigenvalues of the constituents. Furthermore, using the second equality in (17) and $\sigma_{Q,G_i} = \sigma_{A_i} + \mathbb{I}_{2M}/2$ we conclude by a direct calculation that the eigenvalues of σ_{Q,G_1} and σ_{Q,G_2} coincide. The claim follows from the fact that the determinant is a product of eigenvalues. □

4.2 GBS and a complete set of graph invariants

Remark. We will use the transformation $A \mapsto C = A \oplus A$ for the application of GBS to the graph isomorphism problem. We recall $C \in \mathbb{R}^{2M\times 2M}$. By 'doubling' A, one copy is conveniently spanned by $(\beta_1, \ldots, \beta_M)$ whereas the second one by $(\overline{\beta}_1, \ldots, \overline{\beta}_M)$. Moreover, to make the matrix C GBS-encodable (see Def. 4) we simply take $R = c(C + k\mathbb{I}_{2M}) = c(A \oplus A + k\mathbb{I}_{2M})$ where $0 < c < 1/(\|A\|_2 + k)$ for $k \geq 0$. The additional multiple of an identity on the diagonal does not affect the hafnian of $A \oplus A$ as follows from Lemma 4 but it will become useful in the next sections.

GBS and Moments of Multivariate Gaussians

The moments $\mu_{n_1,\ldots,n_{2M}}(\Sigma)$ of a (zero-mean) $2M$-dimensional multivariate real normal distribution $\mathcal{N}(0,\Sigma)$ are given by the following formula:

$$\mu_{n_1,\ldots,n_{2M}}(\Sigma) = \partial_{\boldsymbol{x}}^{|\boldsymbol{n}|} e^{\frac{1}{2}\boldsymbol{x}^\top \Sigma \boldsymbol{x}}\big|_{\boldsymbol{x}=0}, \tag{18}$$

where Σ is the covariance matrix. This follows from the fact that $\exp[\frac{1}{2}\boldsymbol{x}^\top \Sigma \boldsymbol{x}]$ is the moment-generating function of the multivariate normal.

Let $R = c(A \oplus A + k\mathbb{I}_{2M})$ be GBS encodable for k sufficiently high so that $R \succ 0$. Then

$$p(\boldsymbol{n}) = \frac{1}{\mathbf{n}!\sqrt{\det \sigma_Q}} \partial_{\boldsymbol{\beta},\overline{\boldsymbol{\beta}}}^{|\boldsymbol{n}|} e^{\frac{1}{2}\gamma^\top R\gamma}\big|_{\gamma=0}. \tag{19}$$

is exactly the moment $\mu_{\boldsymbol{n}}(R)$ of the $2M$-dimensional (zero-mean) multivariate normal distribution $\mathcal{N}(0,R)$ if we ignore the prefactor $(\boldsymbol{n}!\sqrt{\det \sigma_Q})^{-1}$. For clarity, we have changed variables so that $x_j = \beta_j$ and $x_{M+j} = \overline{\beta}_j$ for $j = 0, \ldots, M$.

From the above equations, it is clear that the different photon-counting probabilities of a GBS setup are directly related to various moments of a multivariate normal distribution. Importantly, however, they do not give us *all moments* $(\mu_{n_1,\ldots,n_m,n_{m+1},\ldots,n_{2M}}(R))$, but rather the smaller set $(\mu_{n_1,\ldots,n_M,n_1,\ldots,n_M}(R))$. This is something we need to be careful of. In [10] the authors relate these moments to the matching of the prism graph induced by the graph G.

The moment-generating function factorizes:

$$\exp[\tfrac{1}{2}\boldsymbol{x}^\top R\boldsymbol{x}] = \exp[\tfrac{c}{2}(\boldsymbol{x}^{(M)})^\top (A + k\mathbb{I}_M)\boldsymbol{x}^{(M)}] \times \exp[\tfrac{c}{2}(\boldsymbol{x}^{(2M)})^\top (A + k\mathbb{I}_M)\boldsymbol{x}^{(2M)}], \tag{20}$$

where we have used the notation $\boldsymbol{x}^{(M)} \overset{\text{df}}{=} (x_1, \ldots, x_M)$ and $\boldsymbol{x}^{(2M)} \overset{\text{df}}{=} (x_{M+1}, \ldots, x_{2M})$. We set $c = 1$ (since we omit the determinant prefactor where it otherwise plays a role) and also $k = 0$. This step will cost us the positive-definiteness of A but at the moment this is just a formality to properly define the moment generating function.

We would have set $k = 0$ afterwards anyway to recover the correct probability expression. The moments of this factorized distribution are then

$$\mu_{n_1,\ldots,n_M,n_{M+1},\ldots,n_{2M}}(A) = \partial_{\boldsymbol{x}}^{|\boldsymbol{n}|}\left[\exp\left[\tfrac{1}{2}(\boldsymbol{x}^{(M)})^\top A\,\boldsymbol{x}^{(M)}\right]\times\exp\left[\tfrac{1}{2}(\boldsymbol{x}^{(2M)})^\top A\,\boldsymbol{x}^{(2M)}\right]\right]\Big|_{\boldsymbol{x}=0} \tag{21}$$

Rewriting, we find

$$\begin{aligned}\mu_{n_1,\ldots,n_M,n_{M+1},\ldots,m_{2M}}(A) &= \left[\partial_{\boldsymbol{x}^{(M)}}^{|\boldsymbol{n}|}\exp\left[\tfrac{1}{2}(\boldsymbol{x}^{(M)})^\top A\,\boldsymbol{x}^{(M)}\right]\right]\Big|_{\boldsymbol{x}^{(M)}=0}\left[\partial_{\boldsymbol{x}^{(2M)}}^{|\boldsymbol{n}|}\exp\left[\tfrac{1}{2}(\boldsymbol{x}^{(2M)})^\top A\,\boldsymbol{x}^{(2M)}\right]\right]\Big|_{\boldsymbol{x}^{(2M)}=0}\\ &= \mu_{n_1,\ldots,n_N}(A)\times\mu_{n_{N+1},\ldots,n_{2N}}(A),\end{aligned} \tag{22}$$

where $\mu_{n_1,\ldots,n_M}(A)$ and $\mu_{n_{M+1},\ldots,n_{2M}}(A)$ are moments of the M-dimensional normal distributions $\mathcal{N}(0,A)$.

Connecting back to photon-counting probabilities, we recover c and conclude that, for the considered case of block-diagonal R,

$$p(\boldsymbol{n}) = \frac{c^{|\boldsymbol{n}|}}{\boldsymbol{n}!\sqrt{\det\sigma_Q}}\mu^2_{n_1,\ldots,n_M}(A). \tag{23}$$

Finally, we note that the moments are exactly the hafnian of some appropriate matrix, so

$$p(\boldsymbol{n}) = \frac{c^{|\boldsymbol{n}|}}{\boldsymbol{n}!\sqrt{\det\sigma_Q}}\,\mathrm{haf}\,[A^{\oplus 2}\not\oplus\mathbb{J}_{2|\boldsymbol{n}|}] = \frac{c^{|\boldsymbol{n}|}}{\boldsymbol{n}!\sqrt{\det\sigma_Q}}\,\mathrm{haf}^{\,2}\,[A\not\oplus\mathbb{J}_{|\boldsymbol{n}|}], \tag{24}$$

where the second equality also follows from Lemma 4 and 8.

Proposition 11. *Suppose we have two isospectral graphs G_1 and G_2. Assume we can encode the adjacency matrices A_i of either graph into a Gaussian boson sampling setup. Then these graphs are isomorphic iff the hafnians are related by a permutation,* $\mathrm{haf}\,[A_1\not\oplus\mathbb{J}_{|\boldsymbol{n}|}] = \mathrm{haf}\,[A_2\not\oplus\mathbb{J}_{|\pi(\boldsymbol{n})|}]$*, for all $\boldsymbol{n}$. Furthermore, the permutation π must be the same for all $\boldsymbol{n}$.*

Proof of $\Rightarrow$:. Suppose G_1 and G_2 are isomorphic. Equivalently, their adjacency matrices are related by a permutation

$$A_1 = P^\top A_2 P, \tag{25}$$

where $P_{\pi(i)i} = \delta_{i,\pi(i)}$ for some permutation π. If we encode these adjacency matrices directly into the covariance matrices of two Gaussian states, then graph isomorphism is equivalent to the multivariate normal distributions corresponding to these two Gaussian states being related by a permutation of coordinates:

$$\mathcal{N}(0, A_1\oplus A_1) = \mathcal{N}(0,(P^\top A_2 P)\oplus(P^\top A_2 P)) = \mathcal{N}(0,(P^{\oplus 2})^\top(A_2\oplus A_2)(P^{\oplus 2})). \tag{26}$$

All moments of these $2M$-dimensional distributions must correspondingly be related by the permutation $\pi\oplus\pi$,

$$\mu_{n_1,\ldots,n_M,n_{M+1},\ldots,n_{2M}}(R_1) = \mu_{(\pi\oplus\pi)(n_1,\ldots,n_M,n_{M+1},\ldots,n_{2M})}(R_2),\quad \forall\,(n_1,\ldots,n_{2M}), \tag{27}$$

where $R_i = A_i\oplus A_i$ are made into GBS encodable matrices (we keep on omitting the c factors). Looking back to Eq. (22), we have

$$\mu_{n_1,\ldots,n_M}(A_1)\mu_{n_{M+1},\ldots,n_{2M}}(A_2) = \mu_{\pi(n_1,\ldots,n_M)}(A_2)\mu_{\pi(n_{M+1},\ldots,n_{2M})}(A_2),\quad \forall\,(n_1,\ldots,n_M),(n_{M+1},\ldots,n_{2M}). \tag{28}$$

These moments must be equal for any choices $\boldsymbol{p} = (n_1,\ldots,n_M)$ and $\boldsymbol{q} = (n_{M+1},\ldots,n_{2M})$. Thus, we conclude that

$$\mathrm{haf}\,[A_1\not\oplus\mathbb{J}_{|\boldsymbol{p}|}]\,\mathrm{haf}\,[A_1\not\oplus\mathbb{J}_{|\boldsymbol{q}|}] = \mathrm{haf}\,[A_2\not\oplus\mathbb{J}_{|\pi(\boldsymbol{p})|}]\,\mathrm{haf}\,[A_2\not\oplus\mathbb{J}_{|\pi(\boldsymbol{q})|}]. \tag{29}$$

In particular, for $\boldsymbol{p} = \boldsymbol{q} = \boldsymbol{n}$, where $\boldsymbol{n}$ is arbitrary, we get $\mathrm{haf}^{\,2}\,[A_1\not\oplus\mathbb{J}_{|\boldsymbol{n}|}] = \mathrm{haf}^{\,2}\,[A_2\not\oplus\mathbb{J}_{|\pi(\boldsymbol{n})|}]$. We now use the fact that adjacency matrices A contain only 0s or 1s, so $\mathrm{haf}\,[A\not\oplus\mathbb{J}_{|\boldsymbol{n}|}]\geq 0$ for any possible A or $\boldsymbol{n}$. This leads to

$$\mathrm{haf}\,[A_1\not\oplus\mathbb{J}_{|\boldsymbol{n}|}] = \mathrm{haf}\,[A_2\not\oplus\mathbb{J}_{|\pi(\boldsymbol{n})|}],\ \forall\,\boldsymbol{n}, \tag{30}$$

which proves the statement.

4.2.0.1 Proof of ⇐:

Eq. (30) immediately implies Eq. (29), even when the hafnians are not positive. Furthermore, Eqns. (25)-(29) are all equivalent. Hence, the graphs having adjacency matrices A_1 and A_2 are isomorphic. □

GBS and Symmetrized Moments of Multivariate Gaussians

In this part we find a new criterion for isomorphism of two isospectral graphs by showing that symmetrized moments are also complete invariants for graph isomorphism:

Theorem 12. *Let G_1 and G_2 be two isospectral graphs on an even number of vertices M. Denote by $p_1(\boldsymbol{n})$ and $p_2(\boldsymbol{n})$, the probabilities corresponding to G_1 and G_2, given in Eq.* (2)*. Then G_1 and G_2 are isomorphic if and only if the symmetrized sums*

$$\sum_{\sigma\in\mathfrak{S}_n}\sqrt{p(\boldsymbol{n}_\sigma)}$$

are the same for the two graphs for all possible $\boldsymbol{n}$.

To avoid a notational clash in this section we use $\boldsymbol{n}_\sigma$ instead of $\sigma(\boldsymbol{n})$ used in the previous sections. We start with the following result.

Theorem 13. *The following statements are equivalent for two Gaussian distributions with zero mean and positive definite covariance matrices $\Sigma, \Sigma' \in \mathbb{R}^{n\times n}$:*

1. *The two Gaussian distributions are isomorphic.*
2. *The matrices Σ and Σ' are permutationally similar.*
3. *The matrices Σ^{-1} and $(\Sigma')^{-1}$ are permutationally similar.*
4. *For each homogeneous symmetric polynomial $p(\boldsymbol{x})$ of even degree the expected value of $p(\boldsymbol{x})$ is the same for the two Gaussian distributions.*
5. *The symmetrized moments of the two Gaussian distributions are the same.*

Proof. For a given Gaussian distribution with zero mean and the covariance matrix Σ let us denote by H_Σ the distribution with the following density

$$h_\Sigma(\boldsymbol{x}) = \frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\exp\left[-\boldsymbol{x}^\top P^\top(\sigma)\Sigma^{-1}P(\sigma)\boldsymbol{x}\right]. \tag{31}$$

By G_Σ we denote the density function of the normal distribution $\exp\left[-\boldsymbol{x}^\top\Sigma^{-1}\boldsymbol{x}\right]$.

Clearly if Σ and Σ' are permutationally similar then $h_\Sigma(\boldsymbol{x}) = h_{\Sigma'}(\boldsymbol{x})$ for each $\boldsymbol{x}$. So we need to prove the other direction. Let $p(\boldsymbol{x})$ be a monomial $x_1^{m_1}\cdots x_n^{m_n}$ of even degree. We denote $\boldsymbol{m} = (m_1, \ldots, m_n)$. Consider the moments

$$\mu_{G_\Sigma}(\boldsymbol{m}) = \mathbb{E}_{G_\Sigma}[X_1^{m_1}\cdots X_n^{m_n}],$$
$$\mu_{H_\Sigma}(\boldsymbol{m}) = \mathbb{E}_{H_\Sigma}[X_1^{m_1}\cdots X_n^{m_n}].$$

Then

$$\mu_{H_\Sigma}(\boldsymbol{m}) = \frac{1}{n!}\sum_{\sigma\in\mathfrak{S}_n}\mu_{G_\Sigma}(\boldsymbol{m}_\sigma).$$

We call $\mu_{H_\Sigma}(\boldsymbol{m})$ the symmetrized moments of G_Σ. Clearly, $\mu_{H_\Sigma}(\boldsymbol{m}) = \mu_{H_\Sigma}(\boldsymbol{m}_\sigma)$ for each $\sigma \in \mathfrak{S}_n$. Thus if two Gaussian distributions are isomorphic, then $H_\Sigma = H_{\Sigma'}$ and the corresponding density functions are the same. In particular, the moments of H_Σ and $H_{\Sigma'}$ are the same.

Our first main result is the claim that if the moments of H_Σ and $H_{\Sigma'}$ are the same then $H_\Sigma = H_{\Sigma'}$. It is not true that the equality of the moments yield that the distribution are the same. However, it is true for all

distributions that have the form of H_Σ. Indeed a sufficient condition is $M(\boldsymbol{u}) = \mathbb{E}[\exp[\langle \boldsymbol{u}, \boldsymbol{X}\rangle]]$ is a well-defined vector for all $\boldsymbol{u} \in \mathbb{R}^n$ [28]. This condition is satisfied for H_Σ, where Σ is a positive definite matrix.

It is left to show that if $H_\Sigma = H_{\Sigma'}$ then Σ and Σ' are permutationally similar. We first analyze the behavior of the $n!$ numbers

$$\exp[-\boldsymbol{x}^\top P^\top(\sigma)\Sigma^{-1}P(\sigma)\boldsymbol{x}], \tag{32}$$

for a fixed but arbitrary vector $\boldsymbol{x} \in \mathbb{R}^n$ and $\sigma \in \mathfrak{S}_n$.

Let $\operatorname{orb}\Sigma$ be all pairwise distinct matrices of the form $P^\top(\sigma)\Sigma P(\sigma)$ for $\sigma \in \mathfrak{S}_n$. Recall that $|\operatorname{orb}\Sigma|$ divides $n!$ and $n!/|\operatorname{orb}\Sigma|$ is the cardinality of the automorphism group, i.e., all $\sigma \in \mathfrak{S}_n$ such that $P^\top(\sigma)\Sigma P(\sigma) = \Sigma$. Let A_1, A_2 be two real symmetric matrices of order n. Define two corresponding quadratic forms $f_1(\boldsymbol{x}) = \boldsymbol{x}^T A_1 \boldsymbol{x}, f_2(\boldsymbol{x}) = \boldsymbol{x}^\top A_2 \boldsymbol{x}$. Assume that $A_1 \neq A_2$. Then $f_1(\boldsymbol{x}) = f_2(\boldsymbol{x})$ if and only if $h(\boldsymbol{x}) = \boldsymbol{x}^\top(A_1 - A_2)\boldsymbol{x} = 0$. Hence for generic, (randomly selected $\boldsymbol{x}$) we have that $f_1(\boldsymbol{x}) \neq f_2(\boldsymbol{x})$. Similarly: let $A_1, \ldots, A_k$ be k pairwise distinct symmetric matrices. Set $f_i(\boldsymbol{x}) = \boldsymbol{x}^\top A_i \boldsymbol{x}$ for $i \in [k]$. Then for generic $\boldsymbol{x}$, $f_i(\boldsymbol{x}) \neq f_j(\boldsymbol{x})$ for $i \neq j$. Since any two pairs of matrices in $\operatorname{orb}\Sigma$ are pairwise distinct it follows that for a generic $\boldsymbol{x}$ we will have exactly $|\operatorname{orb}\Sigma|$ distinct values in (31) and each value is repeated $n!/|\operatorname{orb}\Sigma|$ times.

Assume now that for each $\boldsymbol{x} \in \mathbb{R}^n$ in (31) we have the equality $h_\Sigma(t\boldsymbol{x}) = h_{\Sigma'}(t\boldsymbol{x})$:

$$\sum_{\sigma\in\mathfrak{S}_n} \left(\exp[-\boldsymbol{x}^\top P^\top(\sigma)\Sigma^{-1}P(\sigma)\boldsymbol{x}]\right)^{t^2} = \sum_{\sigma\in\mathfrak{S}_n} \left(\exp[-\boldsymbol{x}^\top P(\sigma)^\top(\Sigma')^{-1}P(\sigma)\boldsymbol{x}]\right)^{t^2} \tag{33}$$

for some fixed $t \geq 0$. Fix $\boldsymbol{x}$ in general position. Then for $t = 1$ we get that the number of distinct values in (31) is $|\operatorname{orb}\Sigma|$ for Σ and $|\operatorname{orb}\Sigma'|$ for Σ', respectively. Let

$$a(\sigma) = \exp[-\boldsymbol{x}^\top P(\sigma)^\top(\Sigma)^{-1}P(\sigma)\boldsymbol{x}], \tag{34}$$

$$a'(\sigma) = \exp[-\boldsymbol{x}^\top P(\sigma)^\top(\Sigma')^{-1}P(\sigma)\boldsymbol{x}] \tag{35}$$

for $\sigma \in \mathfrak{S}_n$. In the equality (33) set $t = \sqrt{k}$ for $k = 0, 1, \ldots, n!$. Thus we have the equalities:

$$\sum_{\sigma\in\mathfrak{S}_n} a(\sigma)^k = \sum_{\sigma\in\mathfrak{S}_n} a'(\sigma)^k$$

for $k = 1, \ldots, n!$. These equalities yield that the two multisets $\{a(\sigma), \sigma \in \mathfrak{S}_n\}$ and $\{a'(\sigma), \sigma \in \mathfrak{S}_n\}$ are the same. Hence the $n!$ moments of discrete distributions equally distributed on $n!$ points given in (32) for Σ and Σ' are the same. Hence these two multisets are the same. First it yields that $|\operatorname{orb}\Sigma| = |\operatorname{orb}\Sigma'|$. Moreover there exists $P(\sigma)$ such that $\boldsymbol{x}^\top P^\top(\sigma)\Sigma^{-1}P(\sigma)\boldsymbol{x} = \boldsymbol{x}^\top(\Sigma')^{-1}\boldsymbol{x}$. Moreover, for each $\Sigma_i = P_i^\top \Sigma P_i$ in the orbit of Σ (under the action of the group of permutations) we have a permutation Q_i such that $\boldsymbol{x}^\top Q_i^\top(\Sigma_i')^{-1}Q_i\boldsymbol{x} = \boldsymbol{x}^\top\Sigma_i^{-1}\boldsymbol{x}$. Now if we change $\boldsymbol{x}$ to $\boldsymbol{y}$ we still have the same equality $\boldsymbol{y}^\top P^\top(\sigma)\Sigma^{-1}P(\sigma)\boldsymbol{y} = \boldsymbol{y}^\top(\Sigma')^{-1}\boldsymbol{y}$. This finally shows that $P^\top(\sigma)\Sigma^{-1}P(\sigma) = \Sigma'$. So indeed the covariance matrices are permutationally similar. □

Let $\boldsymbol{e}_i = (\delta_{i,1}, \ldots, \delta_{i,M})$, where $\delta_{i,j}$ is the Kronecker delta function. Assume that $\boldsymbol{n} = (n_1, \ldots, n_M) \in \mathbb{Z}_+^M$. Let $B = [b_{i,j}]$ be a real symmetric matrix and recall the "$\boldsymbol{n}$-th moment corresponding to B" from the beginning of this section as

$$\mu(\boldsymbol{n}, B) = \frac{\partial^{|\boldsymbol{n}|}}{\partial\boldsymbol{x}^{\boldsymbol{n}}} \exp\left[\frac{1}{2}\boldsymbol{x}^\top B\boldsymbol{x}\right]\Big|_{\boldsymbol{x}=0} = \frac{1}{(|\boldsymbol{n}|/2)!}\frac{\partial^{|\boldsymbol{n}|}}{\partial\boldsymbol{x}^{\boldsymbol{n}}}\left(\frac{1}{2}\boldsymbol{x}^\top B\boldsymbol{x}\right)^{|\boldsymbol{n}|/2}\Big|_{\boldsymbol{x}=0}. \tag{36}$$

(If $B \succ 0$ then $E(X_1^{n_1}\cdots X_M^{n_M})$, the $\boldsymbol{n}$-th moment of the Gaussian distribution given by the covariance B, is equal to $\mu(\boldsymbol{n}, B)$ up to a multiplicative constant.)

To proceed, we also recall the generalization of the classical Leibniz's formula of the derivative of the product of m functions in one variable:

$$\left(\prod_{i=1}^m f_i\right)^{(n)} = \sum_{a_1,\ldots,a_m\in\mathbb{Z}_+,\sum_{i=1}^m a_i=n} \binom{n}{a_1, a_2, \ldots, a_m} \prod_{i=1}^m f_i^{(a_i)}, \quad \binom{n}{a_1, a_2, \ldots, a_m} = \frac{n!}{a_1! \times \ldots \times a_m!}.$$

Assume now that $f_1 = \cdots = f_m = f(\boldsymbol{x}) = f(x_1, \ldots, x_M)$. Then for $\boldsymbol{n} = (n_1, \ldots, n_M) \in \mathbb{Z}_+^M$ we denote by $\partial^{\boldsymbol{n}} = \partial_1^{n_1} \cdots \partial_M^{n_M}$. For $\boldsymbol{n}, \boldsymbol{a} \in \mathbb{Z}_+^M$ let $\binom{\boldsymbol{n}}{\boldsymbol{a}} = \prod_{i=1}^M \binom{n_i}{a_i}$. Then Leibniz's formula yields the multilinear Leibniz's formula:

$$\partial^{\boldsymbol{n}}(f^m) = \sum_{\sum_{i=1}^m \boldsymbol{a}_i = \boldsymbol{n}} \binom{\boldsymbol{n}}{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m} \prod_{i=1}^m (\partial^{\boldsymbol{a}_i} f), \tag{37}$$

where

$$\binom{\boldsymbol{n}}{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m} = \prod_{j=1}^M \binom{n_j}{a_{1,j}, a_{2,j}, \ldots, a_{m,j}} \tag{38}$$

and $\boldsymbol{a}_i = (a_{i,1}, \ldots, a_{i,M})$ and $i \in [m]$. We now apply this formula for $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^\top B\boldsymbol{x}$ and $m = |\boldsymbol{n}|/2$. Then in (37) we need to consider only the case where $|\boldsymbol{a}_i| = 2$ for each $i \in |\boldsymbol{n}|/2$. So

$$\mu(\boldsymbol{n}, B) = \frac{1}{(|\boldsymbol{n}|/2)!} \sum_{\sum_{i=1}^{|\boldsymbol{n}|/2} \boldsymbol{a}_i = \boldsymbol{n}, |\boldsymbol{a}_1| = \ldots = |\boldsymbol{a}_{|\boldsymbol{n}|/2}| = 2} \binom{\boldsymbol{n}}{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{|\boldsymbol{n}|/2}} \prod_{i=1}^{|\boldsymbol{n}|/2} \partial^{\boldsymbol{a}_i}(1/2(\boldsymbol{x}^\top B\boldsymbol{x})). \tag{39}$$

We have two kinds of $\boldsymbol{a}_i$. Namely, either $\boldsymbol{a}_i = 2\boldsymbol{e}_p$ or $\boldsymbol{a}_i = \boldsymbol{e}_p + \boldsymbol{e}_q$, where $1 \le p < q \le n$. Let us discuss briefly all possibilities for the decomposition of $\boldsymbol{n}$ as $\boldsymbol{n} = \sum_{i=1}^{|\boldsymbol{n}|/2} \boldsymbol{a}_i$. We claim that the set $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{|\boldsymbol{n}|/2}\}$ corresponds to the following multigraph $G = G(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{|\boldsymbol{n}|/2})$ with multiple edges and self loops. Each $\boldsymbol{a}_i = \boldsymbol{e}_p + \boldsymbol{e}_q$, where $1 \le p < q \le n$ corresponds to an edge $\{p, q\}$. Each $\boldsymbol{a}_i = 2\boldsymbol{e}_p$ corresponds to a self loop on vertex p (the degree of a self loop is 2). So $A(G) = [c_{pq}(G)]$, the adjacency matrix of G, is a symmetric matrix whose entries are nonnegative integers with the following properties. Each diagonal entry $c_{pp}(G)$ is an even integer. $c_{pp}(G)/2$ is the number of $\boldsymbol{a}_i$ of the form $2\boldsymbol{e}_p$. For $1 \le p < q \le M$ the integer $c_{pq}(G)$ is the number of $\boldsymbol{a}_i$ of the form $\boldsymbol{e}_p + \boldsymbol{e}_q$.

Let $2k = \sum_{p=1}^M c_{pp}$ be a nonnegative integer. That is, the set $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{|\boldsymbol{n}|/2}\}$ has k vectors of the form $2\boldsymbol{e}_p$ for all possible $p \in [n]$. Assume that $k = 0$. Then $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{|\boldsymbol{n}|/2}\}$ correspond to a given multigraph $G = G(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{|\boldsymbol{n}|/2})$ with no loops. If one permutes the vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{|\boldsymbol{n}|/2}$ one obtains the same loopless multigraph G, whose degree sequence is $\boldsymbol{n} = (n_1, \ldots, n_M)$, where $n_i = \sum_{p=1}^M c_{ip}$. Let $\mathcal{G}(\boldsymbol{n})$ be all loopless multigraphs G whose degree sequence is $\boldsymbol{n}$. We arrange the edges of G in a fixed (say lexicographic order): $(1, 2), \ldots, (1, M), (2, 3), \ldots, (M-1, M)$. For example the sequence of edges on 3 vertices $(2, 3), (1, 3), (1, 2), (1, 3)$ is arranged as $(1, 2), (1, 3), (1, 3), (2, 3)$. It corresponds to a degree sequence $\boldsymbol{n} = (3, 2, 3)$.

We denote by $\mathcal{S}_{M,0}$ all $M \times M$ symmetric matrices with zero diagonal. Assume that $A \in \mathcal{S}_{M,0}$. Note that for $f = 1/2(\boldsymbol{x}^\top A\boldsymbol{x})$ we get that $\partial_i^2 f = 0$ for each i.

Lemma 14. *Let $A \in \mathcal{S}_{M,0}$. Then*

$$\mu(\boldsymbol{n}, A) = \sum_{G(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{|\boldsymbol{n}|/2}) \in \mathcal{G}(\boldsymbol{n})} \frac{1}{\prod_{1 \le p < q \le n} c_{pq}(G(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{|\boldsymbol{n}|/2}))!} \binom{\boldsymbol{n}}{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{|\boldsymbol{n}|/2}} \prod_{i=1}^{|\boldsymbol{n}|/2} (\partial^{\boldsymbol{a}_i} f). \tag{40}$$

Proof. Given a decomposition $\sum_{i=1}^{|\boldsymbol{n}|/2} \boldsymbol{a}_i = \boldsymbol{n}$, corresponding to the graph $G(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{|\boldsymbol{n}|/2})$, how may different decompositions are there? Since the edge $\{p, q\}$, $1 \le p < q \le M$ appears $c_{pq}(G(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{|\boldsymbol{n}|/2}))$ times, the number of different decompositions is $\frac{(|\boldsymbol{n}|/2)!}{\prod_{1 \le p < q \le n} c_{pq}(G(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_{|\boldsymbol{n}|/2}))!}$. Use (39) to deduce (40). □

For a given $k \in [|\boldsymbol{n}|/2]$ denote by

$$\mathcal{F}_k(\boldsymbol{n}) = \Big\{(j_1, \ldots, j_k) \in \mathbb{N}^k, 1 \le j_1 \le \cdots \le j_k \le M, \ 2\sum_{l=1}^k \boldsymbol{e}_{j_l} \le \boldsymbol{n}\Big\}.$$

For each $(j_1, \ldots, j_k) \in \mathcal{F}_k(\boldsymbol{n})$ and $i \in [M]$ denote $m_i(j_1, \ldots, j_k)$ the number of j_l that are equal to i. So $\sum_{i=1}^M m_i(j_1, \ldots, j_k) = k$.

Lemma 15. *Let $A \in \mathbb{S}_{M,0}$ and $t \in \mathbb{R}$ are given. Assume that $|\boldsymbol{n}|$ is even. Then*

$$\mu(\boldsymbol{n}, t\mathbb{I}_M + A) = \mu(\boldsymbol{n}, A) \tag{41}$$
$$+ \sum_{k=1}^{|\boldsymbol{n}|/2} t^k \sum_{(j_1,\dots,j_k)\in\mathcal{F}_k(\boldsymbol{n})} \Big(\prod_{i=1}^{M} \frac{d_i!}{m_i(j_1,\dots,j_k)!2^{m_i(j_1,\dots,j_k)}(d_i - 2m_i(j_1,\dots,j_k))!} \Big) \mu(\boldsymbol{n} - 2\sum_{l=1}^{k} \boldsymbol{e}_{j_l}, A).$$

Here $\mu(0, A) = 1$.

Proof. We consider the formula (39). Suppose we have $\{\boldsymbol{a}_1, \dots, \boldsymbol{a}_{|\boldsymbol{n}|/2}\}$ satisfying: (i) $|\boldsymbol{a}_l| = 2$ for each l, and, (ii) $\sum_{l=1}^{|\boldsymbol{n}|/2} \boldsymbol{a}_l = \boldsymbol{n}$. These terms define $G = G(\boldsymbol{a}_1, \dots, \boldsymbol{a}_{|\boldsymbol{n}|/2})$. In how may different ways can we represent $\boldsymbol{n}$ corresponding to $G(\boldsymbol{a}_1, \dots, \boldsymbol{a}_{|\boldsymbol{n}|/2})$? We can do it by permuting the factors $\boldsymbol{a}_1, \dots, \boldsymbol{a}_{|\boldsymbol{n}|/2}$. As in the proof of Lemma 14 it is

$$\frac{(|\boldsymbol{n}|/2)!}{\prod_{1\le p<q\le M} c_{pq}(G(\boldsymbol{a}_1, \dots, \boldsymbol{a}_{|\boldsymbol{n}|/2}))! \prod_{p=1}^{M} (c_{pp}(G(\boldsymbol{a}_1, \dots, \boldsymbol{a}_{|\boldsymbol{n}|/2}))/2)!}.$$

The entry $c_{pq}(G(\boldsymbol{a}_1, \dots, \boldsymbol{a}_{|\boldsymbol{n}|/2}))$ stands for the number of times the edge $\{p, q\}$ appears. The number of self-loops (p, p) is $c_{pp}(G(\boldsymbol{a}_1, \dots, \boldsymbol{a}_{|\boldsymbol{n}|/2}))/2)$. Dividing by $(|\boldsymbol{n}|/2)!$ we see that the contribution of $G(\boldsymbol{a}_1, \dots, \boldsymbol{a}_{|\boldsymbol{n}|/2})$ is

$$\frac{1}{\prod_{1\le p<q\le M} c_{pq}(G(\boldsymbol{a}_1, \dots, \boldsymbol{a}_{|\boldsymbol{n}|/2}))! \prod_{p=1}^{M} (c_{pp}(G(\boldsymbol{a}_1, \dots, \boldsymbol{a}_{|\boldsymbol{n}|/2}))/2)!} \times \binom{\boldsymbol{n}}{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_{|\boldsymbol{n}|/2}} \prod_{i=1}^{|\boldsymbol{n}|/2} \partial^{\boldsymbol{a}_i}(1/2(\boldsymbol{x}^\top B\boldsymbol{x})).$$

Let k be the number of terms in $\{\boldsymbol{a}_1, \dots, \boldsymbol{a}_{|\boldsymbol{n}|/2}\}$ of the form $2\boldsymbol{e}_i$. The contribution of all terms $\{\boldsymbol{a}_1, \dots, \boldsymbol{a}_{|\boldsymbol{n}|/2}\}$ for which $k = 0$ is $\mu(\boldsymbol{n}, A)$. Let us assume that $k \in [|\boldsymbol{n}|/2]$. Without loss of generality we can assume that $\boldsymbol{a}_l = 2\boldsymbol{e}_{j_l}$ for $l \in [k]$. So $\boldsymbol{a}_l = \boldsymbol{e}_{p(l)} + \boldsymbol{e}_{q(l)}$, $p(l) < q(l)$ for $l > k$. Clearly, $\partial^{\boldsymbol{a}_l} f = t$ for $l \in [k]$. Hence $\prod_{l=1}^{k} \partial^{\boldsymbol{a}_l} f = t^k$. The sum of all

$$\frac{1}{\prod_{1\le p<q\le M} c_{pq}(G(\boldsymbol{a}_{k+1}, \dots, \boldsymbol{a}_{|\boldsymbol{n}|/2}))!} \binom{\boldsymbol{n} - 2\sum_{l=1}^{k} \boldsymbol{e}_{j_l}}{\boldsymbol{a}_{k+1}, \dots, \boldsymbol{a}_{|\boldsymbol{n}|/2}} \prod_{l=k+1}^{|\boldsymbol{n}|/2} \partial^{\boldsymbol{a}_l} f,$$

where $\boldsymbol{a}_l = \boldsymbol{e}_{p(l)} + \boldsymbol{e}_{q(l)}$ for $l > k$, and $\boldsymbol{n} - 2\sum_{l=1}^{k} \boldsymbol{e}_{j_l} = \sum_{l=k+1}^{|\boldsymbol{n}|/2} \boldsymbol{a}_l$ is exactly $\mu(\boldsymbol{n} - 2\sum_{l=1}^{k} \boldsymbol{e}_{j_l}, A)$. It is left to justify the coefficient $\prod_{i=1}^{M} \frac{n_i!}{m_i(j_1,\dots,j_k)!2^{m_i(j_1,\dots,j_k)}(n_i - 2m_i(j_1,\dots,j_k))!}$ in front of $\mu(\boldsymbol{n} - 2\sum_{l=1}^{k} \boldsymbol{e}_{j_l}, A)$. This comes from the equality

$$\frac{1}{\prod_{1\le p<q\le M} c_{pq}(G(\boldsymbol{a}_1, \dots, \boldsymbol{a}_{|\boldsymbol{n}|/2}))! \prod_{p=1}^{M} (c_{pp}(G(\boldsymbol{a}_1, \dots, \boldsymbol{a}_{|\boldsymbol{n}|/2}))/2)!} \binom{\boldsymbol{n}}{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_{|\boldsymbol{n}|/2}}$$
$$= \Big(\prod_{i=1}^{M} \frac{n_i!}{m_i(j_1,\dots,j_k)!2^{m_i(j_1,\dots,j_k)}(n_i - 2m_i(j_1,\dots,j_k))!} \Big) \frac{1}{\prod_{1\le p<q\le M} c_{pq}(G(\boldsymbol{a}_{k+1}, \dots, \boldsymbol{a}_{|\boldsymbol{n}|/2}))!}$$
$$\times \binom{\boldsymbol{n} - 2\sum_{l=1}^{k} \boldsymbol{e}_{j_l}}{\boldsymbol{a}_{k+1}, \dots, \boldsymbol{a}_{|\boldsymbol{n}|/2}}.$$

We need to see this equality on the level of the derivative with respect to the variable x_i. Let $m_i = m_i(j_1, \dots, j_k)$. If $m_i = 0$, then $(\boldsymbol{n} - 2\sum_{j=1}^{k} \boldsymbol{e}_{j_l})_i = n_i$ for the coordinate i we have obvious equality. Assume now that $m_i \ge 1$. Then on the left-hand side of the above equality we the factor $\frac{n_i!}{(2!)^{m_i} m_i!}$. The factor $m_i!$ is equal to $(c_{ii}(G)/2)!$. On the right-hand side we have the factors:

$$\frac{n_i!}{m_i!(2!)^{m_i}(n_i - 2m_i)!}(n_i - 2m_i)!.$$

□

We are now ready to give the proof for the main result of this section.

Proof of Theorem 12.. We write the symmetrized sums of $\mu(\boldsymbol{n}, A)$ as

$$\mu_{\mathrm{sym}}(\boldsymbol{n}, A) = \sum_{\sigma \in \mathfrak{S}_n} \mu(\boldsymbol{n}, P^\top(\sigma) A P(\sigma)). \tag{42}$$

Now assume $\mu_{\mathrm{sym}}(\boldsymbol{n}, A_1) = \mu_{\mathrm{sym}}(\boldsymbol{n}, A_2), \forall \boldsymbol{n}$ for two isospectral graphs A_1, A_2. Then from (41) we get $\mu_{\mathrm{sym}}(\boldsymbol{n}, B_1) = \mu_{\mathrm{sym}}(\boldsymbol{n}, B_2), \forall \boldsymbol{n}$. Finally, Theorem 13 yields that B_1 and B_2 are permutationally similar and so are A_1 and A_2. □

Example. Let us illustrate Lemma 15 and Theorem 12 on an example of a graph whose adjacency matrix is $A \oplus A$ of size $2M = 6$ for orbit of $\boldsymbol{n} = (2, 3, 3)$. Since $|\boldsymbol{n}| = 8$ it is clear that only the fourth power ($|n|/2 = 4$) of $(\boldsymbol{x}^{(M)})^\top A(t) \boldsymbol{x}^{(M)}$ survives an encounter with the partial derivatives. So, following Lemma 15, we write

$$\frac{1}{2^4 4!} \frac{\partial^8}{\partial x_1^2 \partial x_2^3 \partial x_3^3} ((\boldsymbol{x}^{(M)})^\top A(t) \boldsymbol{x}^{(M)})^4 = \frac{1}{2^4 4!} \frac{\partial^8}{\partial x_1^2 \partial x_2^3 \partial x_3^3} \sum_{k=0}^{4} \binom{4}{k} t^k |\boldsymbol{x}^{(M)}|^{2k} ((\boldsymbol{x}^{(M)})^\top A \boldsymbol{x}^{(M)})^{4-k} \tag{43a}$$

$$\overset{\boldsymbol{x}=0}{\mapsto} 36 a_{12} a_{13} a_{23}^2 + 6t a_{23} (3(a_{12}^2 + a_{13}^2) + a_{23}^2) + 18 t^2 a_{12} a_{13} + 9 t^3 a_{23}. \tag{43b}$$

Each t^k coefficient corresponds to a polynomial of the matrix entries in the exponential of some $\mu(\boldsymbol{n} - 2\sum_{l=1}^{k} \boldsymbol{e}_{j_l}, A) \equiv \mu(\boldsymbol{m}, A)$. In accordance with Eq. (41) we get, for example for t^2, $\boldsymbol{m} = (2, 1, 1)$ since

$$\frac{1}{16} \frac{\partial^4}{\partial x_1^2 \partial x_2 \partial x_3} \left((\boldsymbol{x}^{(M)})^\top A \boldsymbol{x}^{(M)} \right)^2 \Big|_{\boldsymbol{x}=0} = a_{12} a_{13}.$$

As the final step, we symmetrize the orbit represented by $\boldsymbol{n}$ (in this case the orbit size equals 3) which causes a permutation of indices in (43b).

It is advantageous to stratify the measurement events of an M-mode interferometer according to the total photon number $|\boldsymbol{n}| \geq 0$. Once M and $|\boldsymbol{n}|$ are fixed, all possible detection events can be split into the orbits O_i (equivalence classes under permutation) that partition the set of all events for a fixed M and $|\boldsymbol{n}|$. We choose the class representative to be a detection event $\boldsymbol{n} = (n_j)_{j=1}^M$ such that $n_i \leq n_j, \forall i, j$ and denote by $G_{\boldsymbol{n}}$ its stabilizer. Clearly $G_{\boldsymbol{n}} \subset G = \mathfrak{S}_M$ and the orbits are generated by the left action of the coset $G/G_{\boldsymbol{n}}$. In order to find the orbits with a great number of detection events (presumably the most likely ones) we count the orbit size according to $|O_{\boldsymbol{n}}| = |\mathfrak{S}_M|/|G_{\boldsymbol{n}}| = \binom{M}{k_0, k_1, \ldots, k_\ell}$, where k_j are the multiplicities of the j-th photon events satisfying $\sum_{j=0}^{\ell} j k_j = |\boldsymbol{n}|$ and $\ell \leq M$. The probability of measurement of a given pattern $(n_1, \ldots, n_M)$ is given by $p(\boldsymbol{n})$ in Eq. (8) (or, more precisely, by its doubled version, Eq. (24), see Lemma 4 and the remark on page 175), where $\sum_{i=1}^{M} n_i = |\boldsymbol{n}|$. Hence the probability of orbit $O_{\boldsymbol{n}}$ for a graph G reads

$$p_G(O_{\boldsymbol{n}}) = \frac{1}{\sqrt{\det \sigma_{Q,G}}} \frac{c^{|\boldsymbol{n}|}}{\boldsymbol{n}!} \sum_{\boldsymbol{n} \in O_{\boldsymbol{n}}}^{|O_{\boldsymbol{n}}|} \mathrm{haf}^2 [A \not\oplus \mathbb{J}_{|\boldsymbol{n}|}]. \tag{44}$$

How does the number of orbits increase with $|\boldsymbol{n}|$? This is equivalent to the question of integer $|\boldsymbol{n}|$ partition, that is, in how many ways one can write

$$\lambda_1 + \cdots + \lambda_m = |\boldsymbol{n}|, \tag{45}$$

where the order of the sum plays no role and the number of parts is $1 \leq m \leq M$. We naturally order the parts such that $\lambda_i \leq \lambda_{i+1}, \forall i$. Suppose $M \geq |\boldsymbol{n}|$ first. No closed formula is known but the generating function for integer partition provides the number of orbits for a given $|\boldsymbol{n}|$. Also, very precise estimates have been uncovered and the growth of the number of orbits is exponential in $|\boldsymbol{n}|$. For $M < |\boldsymbol{n}|$, not all number partitions can be realized and the counting is given by the generating function for the number of integer partitions into exactly M parts. Note that we only partition even numbers in this paper, since GBS assigns zero probability for odd $|\boldsymbol{n}|$.

Corollary 16 (of Lemma 6). $p_G(O_{\boldsymbol{n}}) = 0$ *whenever* $p(\boldsymbol{n}) = 0$ *for the orbit representative* $\boldsymbol{n}$.

Curiously, if we try to coarse-grain the probability distribution further and introduce the *partition probability*

$$p_G(|\boldsymbol{n}|) \stackrel{\text{df}}{=} \sum_{\boldsymbol{n}\text{ s.t. }|\boldsymbol{n}|\text{ fixed}} p(\boldsymbol{n}) = \frac{1}{\sqrt{\det \sigma_{Q,G}}} \sum_{n_1+\cdots+n_M=|\boldsymbol{n}|} \frac{1}{\boldsymbol{n}!} \sum_{\boldsymbol{n}\in O_{\boldsymbol{n}}}^{|O_{\boldsymbol{n}}|} \mathrm{haf}^2\,[A \not\oplus \mathbb{J}_{|\boldsymbol{n}|}], \tag{46}$$

where the first sum on the RHS is over the partitions of $|\boldsymbol{n}|$ and the second sum over the orbit elements. We find

Lemma 17. *$p_{G_1}(|\boldsymbol{n}|) = p_{G_2}(|\boldsymbol{n}|)$ for all $|\boldsymbol{n}|$ whenever the graphs G_1, G_2 are isospectral.*

Proof. Any undirected graph G on $2M$ vertices can be encoded as a pure covariance matrix whose circuit decomposition consists of an array of M single-mode squeezing transformation $S(r_k)$ $(0 \le k \le M)$ followed by an M-mode linear interferometer U [9]. For each $S(r_k)$ we find

$$S(r_k)\,|0\rangle = \frac{1}{\sqrt{\cosh r_k}} \sum_{n=0}^{\infty} \frac{\sqrt{(2n)!}}{2^n n!} \tanh^n r_k\,|2n\rangle \tag{47}$$

and so

$$\bigotimes_{k=1}^{M} S(r_k)\,|0\rangle = \sum_{\boldsymbol{n}/2=0}^{\infty} \sum_{i=1}^{\binom{|\boldsymbol{n}|/2+M-1}{|\boldsymbol{n}|/2}} \alpha_{i\boldsymbol{n}}(r_1,\ldots,r_M)\,|(\boldsymbol{n},M)\rangle_i, \tag{48}$$

where $|(\boldsymbol{n},M)\rangle$ carry all completely symmetric representations of $su(M)$ (each representation labeled by $\boldsymbol{n}/2$). Given $\lambda_k(A)$ and $0 < c < 1/\|A\|_2$ for G's adjacency matrix A ([9], see also Lemma 2) we can write $c\lambda_k = \tanh r_k$. Therefore $\alpha_{i\boldsymbol{n}}(r_1,\ldots,r_M) = \alpha_{i\boldsymbol{n}}(\lambda_1,\ldots,\lambda_M)$. Since the interferometer U preserves the number of particles $|\boldsymbol{n}|$ the partition probability $p_G(|\boldsymbol{n}|)$ is unaffected by it. Then

$$p_G(|\boldsymbol{n}|) = \sum_{i=1}^{\binom{|\boldsymbol{n}|/2+M-1}{|\boldsymbol{n}|/2}} |\alpha_{i\boldsymbol{n}}(r_1,\ldots,r_M)|^2 = \sum_{i=1}^{\binom{|\boldsymbol{n}|/2+M-1}{|\boldsymbol{n}|/2}} |\alpha_{i\boldsymbol{n}}(\lambda_1,\ldots,\lambda_M)|^2.$$

However, the RHS is independent on the graph (depends only on λ_k common for isospectral graphs) and the claim follows. □

Remark. Note that the similar argument does not hold for the less coarse-grained probability $p_G(O_{\boldsymbol{n}})$ in Eq. (44) since the interferometer 'mixes' the orbits.

Even though the coarse-grained probability, Eq. (46), cannot be used to distinguish nonisomorphic graphs, not all hope is lost. Possible strategies and the closely related problem of scalability is discussed in Sec. 6.

4.3 Modifying the results for $C = A \oplus A$ and beyond

Given A of even order consider $p(\boldsymbol{n}, C)$ in (2) and $\mu(\boldsymbol{n}, C)$. If for the two graphs A and B we have the equalities for the symmetrized sums

$$\sum_{\sigma\in\mathfrak{S}_n} p(\boldsymbol{n}_\sigma, A\oplus A) = \sum_{\sigma\in\mathfrak{S}_n} p(\boldsymbol{n}_\sigma, B\oplus B),$$

what can we say? If instead of considering just the matrix A we will consider the matrix $A \oplus A$ then we can conclude that our bigger graph is a disjoint union of two isomorphic graphs. So if the union of two isomorphic graphs is isomorphic to the union of another two isomorphic graphs, then the two graphs are also isomorphic. If we consider the functions $\mu(\boldsymbol{m}, A\oplus A)$. Note that $\boldsymbol{m} = (m_1,\ldots,m_{2M}) = (\boldsymbol{n},\boldsymbol{n}')$ where $\boldsymbol{n} = (n_1,\ldots,n_M)$, $\boldsymbol{n}' = (n_{M+1},\ldots,n_{2M})$. It now follows that

$$p((\boldsymbol{n},\boldsymbol{n}'), A\oplus A) = p(\boldsymbol{n}, A)p(\boldsymbol{n}', A).$$

Hence, if $\gamma = (\boldsymbol{x}, \boldsymbol{y}, \overline{\boldsymbol{x}}, \overline{\boldsymbol{y}})$ then

$$\exp\left[\frac{1}{2}\gamma^\top (C \oplus C)\gamma\right] = \exp\left[\frac{1}{2}(\boldsymbol{x}^{(M)})^\top A \boldsymbol{x}^{(M)}\right] \exp\left[\frac{1}{2}(\boldsymbol{y}^{(M)})^\top A \boldsymbol{y}^{(M)}\right] \exp\left[\frac{1}{2}(\overline{\boldsymbol{x}}^{(M)})^\top A \overline{\boldsymbol{x}}^{(M)}\right] \exp\left[\frac{1}{2}(\overline{\boldsymbol{y}}^{(M)})^\top A \overline{\boldsymbol{y}}^{(M)}\right].$$

Therefore, if we have equalities for the symmetrized sums, we get a whole hierarchy of necessary conditions by considering $A^{\oplus k}$ for $k > 2$.

4.4 Partition-averaged photon distribution as a necessary condition for graph isomorphism

The main result of this section will be a simpler necessary condition for isospectral graphs to be isomorphic.

Definition 5. Let A be the adjacency matrix of an M-vertex graph G and $1 \le k \le M$. Then we call the partition-averaged photon distribution of the k-th detector the function

$$\langle n_k \rangle_G \stackrel{\text{df}}{=} \sum_{\boldsymbol{n} \in O_{\boldsymbol{n}}}^{|O_{\boldsymbol{n}}|} n_k p(\boldsymbol{n}) = \frac{1}{\sqrt{\det \sigma_{Q,G}}} \frac{1}{\boldsymbol{n}!} \sum_{\boldsymbol{n} \in O_{\boldsymbol{n}}}^{|O_{\boldsymbol{n}}|} n_k \operatorname{haf}^2 [A \not\oplus \mathbb{J}_{|\boldsymbol{n}|}] \tag{49}$$

and its coarse-grained version reads

$$\langle\langle n_k \rangle\rangle_G = \frac{1}{\sqrt{\det \sigma_{Q,G}}} \sum_{n_1 + \cdots + n_M = |\boldsymbol{n}|} \frac{1}{\boldsymbol{n}!} \sum_{\boldsymbol{n} \in O_{\boldsymbol{n}}}^{|O_{\boldsymbol{n}}|} n_k \operatorname{haf}^2 [A \not\oplus \mathbb{J}_{|\boldsymbol{n}|}]. \tag{50}$$

Remark. The complexities of coarse-grained probability $p_G(O_{\boldsymbol{n}})$ in Eq. (44), of the partition-averaged photon distribution of the k-th detector (49), and its coarse-grained version (50) are NP-hard, as we need to sum on the number of elements in O_n, which may be of order $M!$.

Theorem 18. *The partition-averaged photon distributions introduced in Definition 5 of two isomorphic graphs are identical up to a permutation of output modes which can be verified in polynomial time in M.*

Proof. Given partition-averaged photon distribution of a graph G we replace it by a distribution $\langle \tilde{n}_k \rangle_G$, $k \in [n]$, where $0 \le \tilde{n}_1 \le \cdots \le \tilde{n}_M$. Clearly, this rearrangement can be done in $O(M^2)$, actually $O(M \log M)$, time. Two isomorphic graphs will have the same rearranged partition-averaged photon distribution. □

Lemma 19. *Let G_A and $G_{\tilde{A}}$ be isomorphic graphs. Then the output probability distribution from GBS with encoded graphs is related by a permutation.*

Proof. Consider pure events first where we present two proofs. A graph $\tilde{A}$ is isomorphic to A iff there exists a permutation π such that $\tilde{A} = P_\pi^\top A P_\pi$. Ignoring the prefactor $c^{|\boldsymbol{n}|}/(\boldsymbol{n}!\sqrt{\det \sigma_Q})$ in Eq. (24) (it is identical for A and $\tilde{A}$ – see Lemma (10)), it follows that $\tilde{A}^{\oplus 2} \otimes \mathbb{J}_{2|\boldsymbol{n}|}$ is also a permutation of $A^{\oplus 2} \otimes \mathbb{J}_{2|\boldsymbol{n}|}$ since $\not\oplus \equiv \otimes$ for pure detection events (see Remark below Def. (3)). Hence $\operatorname{haf}[\tilde{A}^{\oplus 2} \otimes \mathbb{J}_{2|\boldsymbol{n}|}] = \operatorname{haf}[A^{\oplus 2} \otimes \mathbb{J}_{2|\boldsymbol{n}|}]$ and the probability expressions are invariant.

We prove the same statement by using Eq. (2) where $C = c(A \oplus A)$, $\tilde{C} = c(\tilde{A} \oplus \tilde{A})$ and we can ignore c here by setting $c = 1$. We introduce $P \stackrel{\text{df}}{=} P_\pi \oplus P_\pi$ and write

$$\partial^{|\boldsymbol{n}|}_{\boldsymbol{\beta}, \overline{\boldsymbol{\beta}}} e^{\frac{1}{2}\gamma^\top \tilde{A}^{\oplus 2} \gamma} = \partial^{|\boldsymbol{n}|}_{\boldsymbol{\beta}, \overline{\boldsymbol{\beta}}} e^{\frac{1}{2}(P\gamma)^\top A^{\oplus 2} (P\gamma)} \tag{51}$$

But that implies that the probability of a pure event remains the same since $P\boldsymbol{\beta}$ by definition merely relabels the output modes and the partial derivatives do not care:

$$\partial^{|\boldsymbol{n}|}_{\boldsymbol{\beta}, \overline{\boldsymbol{\beta}}} e^{\frac{1}{2}(P\gamma)^\top A^{\oplus 2} (P\gamma)} = \partial^{|\boldsymbol{n}|}_{\boldsymbol{\beta}, \overline{\boldsymbol{\beta}}} e^{\frac{1}{2}\gamma^\top \tilde{A}^{\oplus 2} \gamma}. \tag{52}$$

For mixed detection events the situation is different. If one of the n_i's in

$$\partial^{|\boldsymbol{n}|}_{\boldsymbol{\beta},\overline{\boldsymbol{\beta}}} e^{\frac{1}{2}(P\gamma)^\top A^{\oplus 2}(P\gamma)}$$

is different from the rest, the corresponding partial derivative breaks the symmetry and unlike the pure case one concludes that

$$\partial^{|\boldsymbol{n}|}_{\boldsymbol{\beta},\overline{\boldsymbol{\beta}}} e^{\frac{1}{2}(P\gamma)^\top A^{\oplus 2}(P\gamma)} \neq \partial^{|\boldsymbol{n}|}_{\boldsymbol{\beta},\overline{\boldsymbol{\beta}}} e^{\frac{1}{2}\gamma^\top A^{\oplus 2}\gamma}. \tag{53}$$

However, if we permute the derivative variables (symbolically written as $\partial_{\beta_i,\overline{\beta}_i} \mapsto \partial_{(P_\pi\beta_i),(P_\pi\overline{\beta}_i)}$), we find the desired equality

$$\partial^{|\boldsymbol{n}|}_{P_\pi\boldsymbol{\beta},P_\pi\overline{\boldsymbol{\beta}}} e^{\frac{1}{2}(P\gamma)^\top A^{\oplus 2}(P\gamma)} = \partial^{|\boldsymbol{n}|}_{\boldsymbol{\beta},\overline{\boldsymbol{\beta}}} e^{\frac{1}{2}\gamma^\top A^{\oplus 2}\gamma}. \tag{54}$$

Next, using map (16), we rewrite the both sides of the last equation as

$$\prod_{i=1}^{M} \partial_{(\hat{P}_\pi\alpha_i),(\hat{P}_\pi\overline{\alpha}_i)} e^{\frac{1}{2}(\hat{P}\boldsymbol{\delta})^\top (A^{\oplus 2}\oplus\mathbb{J}_{2|\boldsymbol{n}|})(\hat{P}\boldsymbol{\delta})} = \prod_{i=1}^{M} \partial_{\alpha_i,\overline{\alpha}_i} e^{\frac{1}{2}\boldsymbol{\delta}^\top (A^{\oplus 2}\oplus\mathbb{J}_{2|\boldsymbol{n}|})\boldsymbol{\delta}}. \tag{55}$$

where $\boldsymbol{\delta} \overset{\text{df}}{=} (\boldsymbol{\alpha},\overline{\boldsymbol{\alpha}})$ and $\hat{P} \overset{\text{df}}{=} \hat{P}_\pi \oplus \hat{P}_\pi$ was introduced in Lemma 9. We used Eq. (8), Lemma 8 and Lemma 9 (the upper route in the commutative diagram to go from the LHS of (54) to the LHS of (55)). But since $(\hat{P}\boldsymbol{\delta})^\top (A^{\oplus 2}\oplus\mathbb{J}_{2|\boldsymbol{n}|})(\hat{P}\boldsymbol{\delta}) = \boldsymbol{\delta}^\top (\hat{P}^\top (A^{\oplus 2}\oplus\mathbb{J}_{2|\boldsymbol{n}|})\hat{P})\boldsymbol{\delta}$ and $\hat{P}$ is a permutation, the hafnian is preserved and the output probability distribution is merely permuted.

To conclude the proof we notice that the overall detection probability is a sum of invariant (for pure events) or permuted (for the mixed ones) probability distributions where the permutation is the same for all mixed $\boldsymbol{n}$'s. □

To simplify the notation in the rest of the section we write $\mathrm{haf}^2_G(\boldsymbol{n}) \equiv \mathrm{haf}^2[A\oplus\mathbb{J}_{|\boldsymbol{n}|}]$ in Eq. (50). Given the stratification into orbits, it is advantageous to collect n_k together with the factorial coefficients and the (squared) hafnians of a graph G to N and haf_G, respectively, and rewrite (50) as

$$\mathsf{n}_G = \frac{1}{\sqrt{\det\sigma_{Q,G}}} \mathsf{N}\,\mathsf{haf}_G, \tag{56}$$

where n_G is M-tuple of numbers.

Example. Let us illustrate (56) for a graph G on $M = 4$ vertices and for $|\boldsymbol{n}| = 2$. There are two orbits represented by $(0,0,0,2)$ and $(0,0,1,1)$. Since the graph is doubled, we have $M = 4$ detectors and then

$$\mathsf{n}_G = \frac{1}{\sqrt{\det\sigma_{Q,G}}} \begin{bmatrix} 2/2! & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 2/2! & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 2/2! & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2/2! & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \mathrm{haf}^2_G(2000) \\ \mathrm{haf}^2_G(0200) \\ \mathrm{haf}^2_G(0020) \\ \mathrm{haf}^2_G(0002) \\ \mathrm{haf}^2_G(1100) \\ \mathrm{haf}^2_G(1010) \\ \mathrm{haf}^2_G(0110) \\ \mathrm{haf}^2_G(1001) \\ \mathrm{haf}^2_G(0101) \\ \mathrm{haf}^2_G(0011) \end{bmatrix}. \tag{57}$$

We can clearly identify the sums on the RHS of (50). Note that due to Corollary 16 the hafnians of the $(0,0,0,2)$ orbit are zero and so are the corresponding contributions to n_G.

The following proof is best viewed together with the above example.

Proof of Theorem 18. Lemma 19 shows that, if graphs G_1 and G_2 are isomorphic, then the ordered set of hafnians for one graph is a permutation of the same ordered set for the other graph. This statement translates into a permutation of haf_G introduced earlier: $\mathsf{haf}_{G_2} = \pi(\mathsf{haf}_{G_1})$. Note that the pure orbit elements are fixed points of π. Now we observe that the i-th row of N by construction coincides with the sequence assembled from the i-th elements of all $\boldsymbol{n}$ (see (57)). So instead of swapping the rows of N we correspondingly swap these sequences in the argument of all haf_G^2 forming haf_G. But this transformation is a permutation of the set of all $\boldsymbol{n}$'s for a fixed $|\boldsymbol{n}|$ since it preserves the photon number. Hence

$$\mathsf{N}\,\pi(\mathsf{haf}_{G_1}) = P_\pi(\mathsf{N})\,\mathsf{haf}_{G_1}, \tag{58}$$

where P_π swaps the rows of N. Since $\mathsf{Nhaf}_{G_2} = \mathsf{N}\,\pi(\mathsf{haf}_{G_1})$ we get

$$\mathsf{Nhaf}_{G_2} = P_\pi(\mathsf{N})\,\mathsf{haf}_{G_1} \tag{59}$$

and thanks to Lemma 10 we can rewrite the equality as

$$\frac{1}{\sqrt{\det \sigma_{Q,G_2}}}\mathsf{Nhaf}_{G_2} = \frac{1}{\sqrt{\det \sigma_{Q,G_1}}}P_\pi(\mathsf{N})\,\mathsf{haf}_{G_1}. \tag{60}$$

But the LHS is n_{G_2} and the action of a permutation P_π on the RHS implies that it is equal to $\pi(\mathsf{n}_{G_1})$. Therefore

$$\mathsf{n}_{G_2} = \pi(\mathsf{n}_{G_1})$$

which is the main result.

To conclude the proof we observe that it takes only a polynomial number of steps to uncover how the partition-averaged photon distribution is permuted. We order n_{G_1} and n_{G_2} in an increasing order and if the two ordered sets differ the graphs cannot be isomorphic □

One could be tempted to argue that the opposite is true (that is, if the partition-averaged photon distributions the same then the graphs are isomorphic). The following counterexample shows that there is no hope for the converse of Theorem 18.

Example (Counterexample based on SRG(16,6,2,2))**.** SRG(16,6,2,2,2) is the smallest family of SRGs containing two isospectral graphs on 16 vertices. Let $|\boldsymbol{n}| = 4$ which can be partitioned in five different ways. Orbits represented by $\boldsymbol{n} = (1, 3)$ and $\boldsymbol{n} = (4)$ (zeros omitted) do not contribute in accord with Corollary 16. Calculating Eq. (50) we find $\langle\langle n_k\rangle\rangle_{G_1} = \langle\langle n_k\rangle\rangle_{G_2}$. What about the less coarse-grained version, Eq. (49). Let's check the orbit of $\boldsymbol{n} = (1, 1, 1, 1)$ where $|O_{\boldsymbol{n}}| = 1820$. Here the situations is quite interesting and generic for SRGs. The sets of hafnians differ: $\mathrm{hafs}[G_1] = (0_{992}, 1_{768}, 2_{60})$ and $\mathrm{hafs}[G_2] = (0_{984}, 1_{792}, 2_{36}, 3_8)$ where the subscripts count the hafnian. Yet, we find $\langle n_k\rangle_{G_1} = \langle n_k\rangle_{G_2}$.

Remark. Note that since the hafnian sets differ in the previous example we know that the graphs are not isomorphic. It just can't be concluded from comparing the partition-averaged photon distributions for $|\boldsymbol{n}| = 4$ and it can't even be concluded from (44) since $p_{G_1}(O_n) = p_{G_2}(O_n)$ for all orbits for $|\boldsymbol{n}| = 4$ (including $\boldsymbol{n} = (1, 1, 1, 1)$ again!). The first differences both in $\langle n_k\rangle_G$ and $p_G(O_n)$ appear for some orbits of $|\boldsymbol{n}| = 8$. Interestingly, $\langle n_k\rangle_G$ is always uniform for SRGs and when it differs for two nonisomorphic SRGs, it differs in a magnitude.

Remark. Similarly to the partition (46), the coarse-grained partition-averaged photon distribution $\langle\langle n_k\rangle\rangle_G$ is efficiently calculable.

5 Simulations for isospectral graphs

In the following section, we present the results of the GBS quantum GI algorithm applied to various SRG families and other isospectral graphs. The algorithm itself is presented in the Appendix B. Among other graphs,

we examine the SRG(35,18,9,9) family, and show that, using various detection patterns, the GBS fully distinguishes all 3854 graphs in this family. Due to the large number of photon event permutations required to calculate the probability of detection, and the classically intractable graph hafnian calculation, the results were computed in parallel using the Python Hafnian library [8] and the Titan supercomputer [1]. Recall our convention for GBS encodable graphs: $C = c(A \oplus A) \in \mathbb{R}^{2M\times 2M}$ where we set $c = 1$ whenever we are allowed to.

Example. Let us start with the smallest connected PING in Fig. 1. The hafnians of the adjacency matrices coincide so we have to look to all possible GBS-measurable submatrices. These correspond to all measurement patterns with at most one photon per mode (in this example we won't study the multiphoton contributions coming from $A\not\oplus\mathbb{J}_{\boldsymbol{n}}$). Hence, we can measure in total $\binom{6}{4} = 15$ graphs on 4 vertices as well as 2 vertices (the subgraphs with an odd number of vertices have zero perfect matchings and therefore zero hafnian). The hafnians of the latter (let's call them 2-hafnians) do not differ but the 4-hafnian sets do differ:

$$4\text{-haf}\, A_{G_1} = (0,0,0,0,0,0,0,1,1,1,1,1,1,1,2) \equiv (0_7, 1_7, 2_1), \tag{61a}$$

$$4\text{-haf}\, A_{G_2} = (0_8, 1_7). \tag{61b}$$

Example. Consider another example of a PING [5] in Fig. 2, this time on nine vertices. Their hafnians are

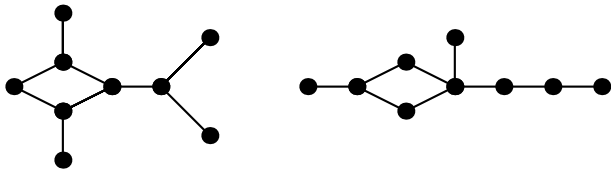

Figure 2: PING on nine vertices.

zero since the number of vertices is odd but also all (8,6,4,2)-hafnians are identical (that is, all subgraphs accessible to GBS have the same number of perfect matchings). We thus have to change the strategy and systematically investigate multiphoton detection events by using the stratification according to the overall photon number and analyze all detection events, both pure and mixed. This is the way we will proceed in the upcoming examples. Here it turns out that the first differences between the two graphs happen for $|\boldsymbol{n}| = 6$. Table 1 on page 187 summarizes the result. The leftmost column contains all partitions of $|\boldsymbol{n}| = 6$ (see (45)) where each partition is represented by a naturally ordered orbit representative. The orbit size is in the second column in accordance with the discussion preceding Eq. (44). In experimental terms, an orbit consists of a measurement pattern and all of its permutations. The two rightmost columns contain the hafnians haf $[A\not\oplus\mathbb{J}_{|\boldsymbol{n}|}]$. We notice a difference in three orbits (greyed): (1, 1, 1, 1, 2), (1, 1, 2, 2) and (1, 1, 1, 3) (the zeros omitted). Also note that the last four rows corresponding to events which do not occur as predicted by Lemma 6. For another graph G_3, isomorphic to G_2, we get the same hafnians for all orbits as an additional check.

Another interesting piece of information is the actual partition-averaged photon distribution for a given orbit provided by (49) or (50) (we omit the determinants in this example). For non-SRGs the partition-averaged photon distribution is typically non-flat. To illustrate (49) let's choose an orbit where no difference was found: $O_{\boldsymbol{n}}$ for $\boldsymbol{n} = (1, 2, 3)$. The first two plots in Fig. 3 are clearly different (that is, non-permutationally invariant). In accordance with the result of Section 4.4, namely Theorem 18, this is enough to decide that the two graphs

1 https://www.olcf.ornl.gov/olcf-resources/compute-systems/titan/

Table 1: All measurement patterns and their permutations $O_{\boldsymbol{n}}$ (orbits) for the PING on nine vertices in Fig. 2 for $|\boldsymbol{n}| = 6$. The second column is the orbit size and in the last two columns we list the hafnians whose total number (the sum of subscripts) equals $|O_{\boldsymbol{n}}|$. Note that the notation we are using for the hafnian sets is defined in (61).

Orbit representative of $O_{\boldsymbol{n}}$	$\lvert O_{\boldsymbol{n}}\rvert$	hafs$[G_1]$	hafs$[G_2]$
$(0,0,0,1,1,1,1,1,1)$	$\binom{9}{6}$	$(0_{69}, 1_{13}, 2_2)$	$(0_{69}, 1_{13}, 2_2)$
$(0,0,0,0,1,1,1,1,2)$	$\binom{9}{5}\binom{5}{1}$	$(0_{586}, 2_{41}, 4_3)$	$(0_{585}, 2_{42}, 4_3)$
$(0,0,0,0,0,1,1,2,2)$	$\binom{9}{4}\binom{4}{2}$	$(0_{698}, 2_{42}, 4_{12}, 6_4)$	$(0_{700}, 2_{42}, 4_{10}, 6_4)$
$(0,0,0,0,0,0,2,2,2)$	$\binom{9}{3}$	(0_{84})	(0_{84})
$(0,0,0,0,0,1,1,1,3)$	$\binom{9}{4}\binom{4}{1}$	$(0_{500}, 6_4)$	$(0_{499}, 6_5)$
$(0,0,0,0,0,0,1,2,3)$	$\binom{9}{3}3!$	$(0_{478}, 6_{26})$	$(0_{478}, 6_{26})$
$(0,0,0,0,0,0,0,3,3)$	$\binom{9}{2}$	$(0_{27}, 6_9)$	$(0_{27}, 6_9)$
$(0,0,0,0,0,0,1,1,4)$	$\binom{9}{3}\binom{3}{1}$	(0_{252})	(0_{252})
$(0,0,0,0,0,0,0,2,4)$	$\binom{9}{2}\binom{2}{1}$	(0_{72})	(0_{72})
$(0,0,0,0,0,0,0,1,5)$	$\binom{9}{2}\binom{2}{1}$	(0_{72})	(0_{72})
$(0,0,0,0,0,0,0,0,6)$	9	(0_9)	(0_9)

are not isomorphic. For a graph G_3 isomorphic to G_2 we notice a mere permutation of bars in the rightmost panel of Fig. 3 again in accordance with Theorem 18.

A similar conclusion follows from the analysis of (50) and the situation is depicted in Fig. 4 for $|\boldsymbol{n}| = 8$.

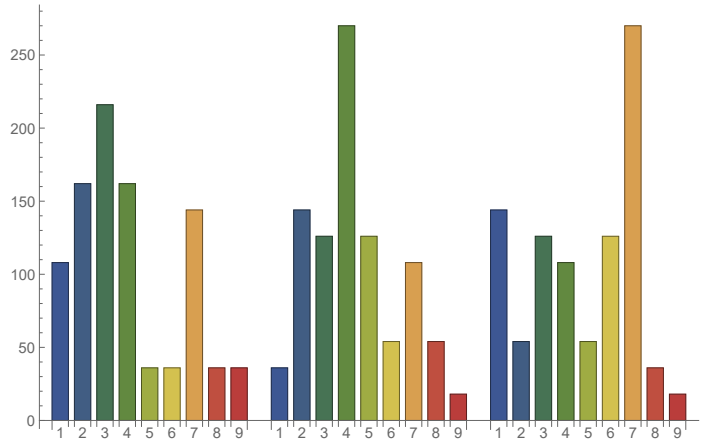

Figure 3: Partition-averaged photon distribution, Eq. 49, for G_1, G_2 and $G_3 \simeq G_2$ for the orbit of $\boldsymbol{n} = (1,2,3)$. The x axis labels the detectors. Note that since we omitted the determinantal prefactor and set $c = 1$ the distribution is not normalized.

Example. Regular isospectral nonisomorphic graphs appear to be somewhere between PINGs and SRGs in terms of the difficulty to distinguish them. We analyzed a pair of graphs on ten vertices introduced in [30, page 110, (a) and (b)]. Neither the coarse-grained photon distribution, Eq. (50), nor its probability equivalent differ for the two graphs for any tested orbit of $\boldsymbol{n}$. This is what we witnessed for all examined SRGs as well. But there is a difference, most likely related to the fact that regular graphs have less symmetry than SRGs. First, a difference in $\langle n_k \rangle_G$ but not in $p_G(O_{\boldsymbol{n}})$ appears for some orbits of $|\boldsymbol{n}| = 6$. For $|\boldsymbol{n}| = 8$ both quantities differ in an ever increasing number orbits. What makes regular graphs different from SRGs is that $\langle n_k \rangle_G$ is not uniform (c.f. with the example and remark at the end of Section 4.4). This is more similar to the PINGs we mentioned

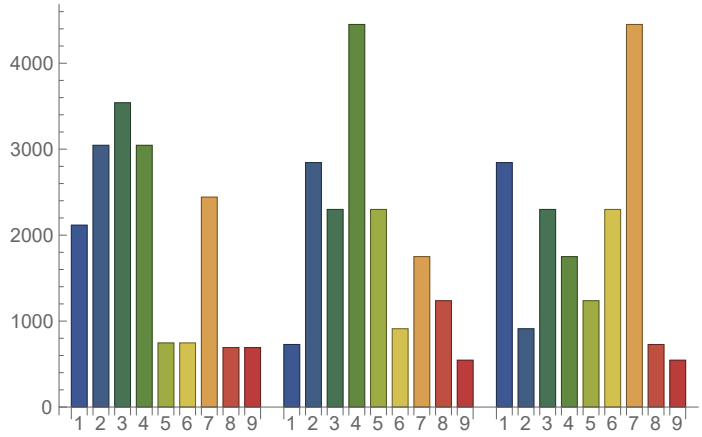

Figure 4: Coarse-grained partition-averaged photon distribution, Eq. 50, for G_1, G_2 and $G_3 \simeq G_2$ for all orbits contributing to $|\boldsymbol{n}| = 8$. The x axis labels the detectors. Note that since we omitted the determinantal prefactor and set $c = 1$ the distribution is not normalized.

previously. So we may get some information on the actual permutation operation from $\langle n_k \rangle_{G_{1,2}}$ if we cannot find any difference for any $\boldsymbol{n}$.

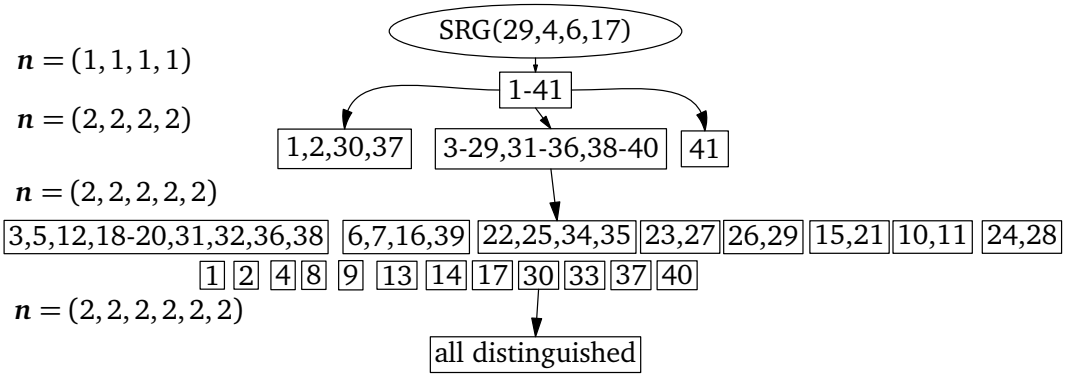

Figure 5: The SRG(29,14,6,7) family of 41 isospectral graphs can be fully distinguished by considering orbit (2, 2, 2, 2, 2, 2). The numbers in the rectangular boxes label the graphs according to [39]. Considering 'smaller' orbits (in terms of $|\boldsymbol{n}|$ or the number of nonzero n_i's) typically leads to a partial separation. Orbit (2, 2, 2, 2, 2, 2) may not be the smallest one to distinguish all the graphs.

Example (SRG(29,14,6,7))**.** Fig. 5 summarizes the path to distinguish all 41 isospectral graphs. As the starting point we took orbit $O_{\boldsymbol{n}}$ for $\boldsymbol{n} = (1, 1, 1, 1)$. This orbit has no distinguishing power. This is indicated by a single square box containing all graphs. One could start with a measurement event containing more single photons but the problem is, as the number of vertices increases, the orbit size grows rapidly making the simulations rather resource-expensive. Also, it is desirable to find the orbit with the smallest possible $|\boldsymbol{n}|$ distinguishing all graphs to (i) heuristically assess the performance of our algorithm and (ii) make sure that the physical resources needed to run the algorithm are not excessive. This is because the smaller $|\boldsymbol{n}|$ is the less squeezing we need in an actual experiment. Sometimes, however, a smaller $|\boldsymbol{n}|$ does not guarantee a faster simulation. In the current example the orbit of $\boldsymbol{n} = (2, 2, 2, 2, 2, 2)$ where $|\boldsymbol{n}| = 12$ is much more computationally feasible than $\boldsymbol{n} = (1, 2, 3, 4)$ where $|\boldsymbol{n}| = 10$. This is because $|O_{\boldsymbol{n}}| = 475020$ for the former but $|O_{\boldsymbol{n}}| = 11400480$ for the

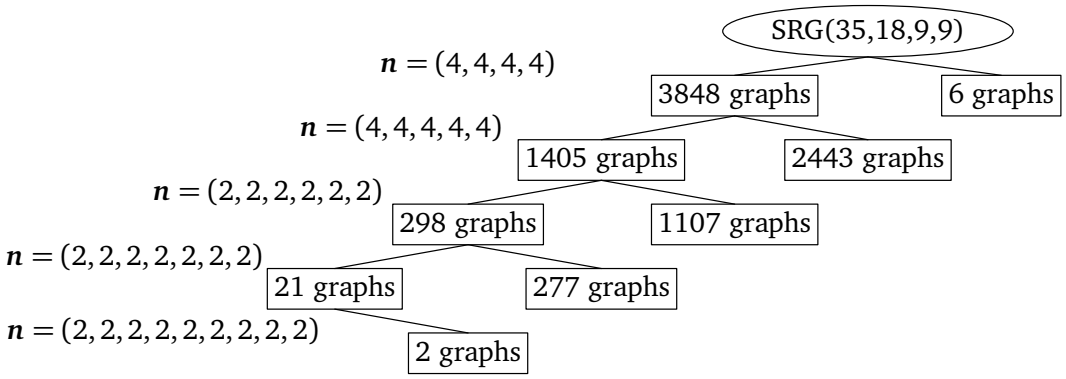

Figure 6: The SRG(35,18,9,9) family of 3854 isospectral graphs can be fully distinguished by starting with the orbit of $\boldsymbol{n}$ = (4, 4, 4, 4) up to $\boldsymbol{n}$ = (2, 2, 2, 2, 2, 2, 2, 2, 2).

latter. So perhaps the orbit of $\boldsymbol{n}$ = (1, 2, 3, 4) distinguishes them all. Hence without exploring all alternative routes we cannot claim optimality (here and in any other example).

Example (SRG(35,18,9,9))**.** The analysis of the biggest family of SRGs we studied is summarized in Fig. 6. Starting from the top, the orbit representatives on the left indicate how successful they are in distinguishing the graphs. The right box contains the number of *newly* distinguished graphs whereas the left box contains the number of remaining graphs. The effect is cumulative so for example the orbit of $\boldsymbol{n}$ = (4, 4, 4, 4, 4) together with $\boldsymbol{n}$ = (4, 4, 4, 4) distinguishes 2443 graphs. Our computational resources were not enough to distinguish the two remaining graphs using Theorem 12. The necessary condition developed in Sec. 4.3 was used instead.

Example (SRG(16,6,2,2))**.** The two graphs can be easily distinguished by our method but in this case we illustrate the probability function for all partitions and their orbits up to $|\boldsymbol{n}| = 14$. In Fig. 7 we plot Eq. (44) for orbits whose probability is nonzero (so their number is less than given by partitioning $|\boldsymbol{n}|$). The x axis labels these orbits and in the plot we indicate the actual partitioning by white and gray background. Even for a fixed $|\boldsymbol{n}|$ some orbits are more likely than others. We observed that the probability is correlated to the size of the orbit. This confirms our intuition from the paragraph before Eq. (44).

6 Open problems and Discussion

6.1 Open questions

Q1. Can we replace in Theorem 12 the quantity $\sum_{\sigma\in\mathfrak{S}_n}\sqrt{p(\boldsymbol{n}_\sigma)}$ with $\sum_{\sigma\in\mathfrak{S}_n} p(\boldsymbol{n}_\sigma)$?

Q2. The main result of this paper is a necessary and sufficient condition for two isospectral graphs to be isomorphic. We found a complete set of graph invariants. However, this is only half satisfactory because we don't know where the difference between two graphs 'kicks in'. Without this knowledge we can use the iff condition only in one direction. The ideal situation would be to have a deterministic or probabilistic criterion for the existence of such a *threshold* orbit as a polynomial function of the graph size. The numerical experiments are favorable as far as the polynomial growth goes – there is no indication that the threshold value grows fast. As the SRG example at the end of Section 4.4 shows, this is actually a more subtle problem: the set of hafnians may be different which is one sufficient condition as shown in Section 4.2 but their sum of squares is forming $p_{G_i}(O_{\boldsymbol{n}})$ (the coarse-grained probability as another sufficient condition) is the same. The latter often comes 'later', that is, for higher orbits than the former.

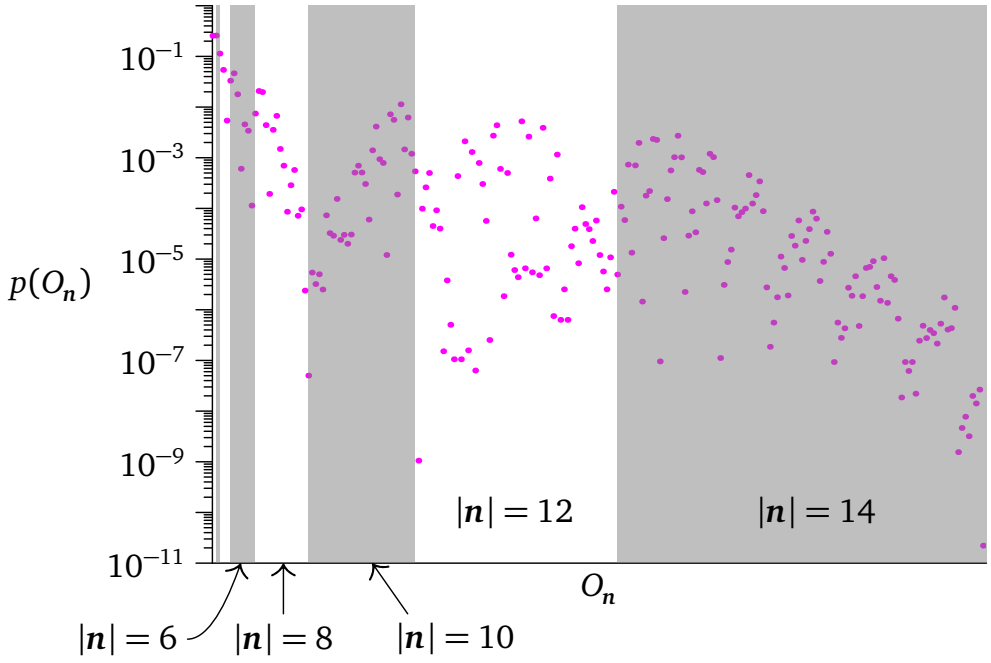

Figure 7: Probabilities of various orbits. Each point is a probability of an orbit $O_{\boldsymbol{n}}$ given by Eq. (44) for one of the graphs from the SRG(16,6,2,2) family. The prefactor in (44) is $1/\sqrt{\det \sigma_{Q,G}} = ((-1 + 4c^2)^{15}(-1 + 36c^2))^{1/2}$ where we set c=1/6.9. Orbits whose probability is zero are omitted.

Q3. When a threshold orbit $\boldsymbol{n}$ is reached we made the following observation: upon examining orbits $\boldsymbol{m}$ for $|\boldsymbol{m}| > |\boldsymbol{n}|$ satisfying $\boldsymbol{m} > \boldsymbol{n}$ ($m_i \geq n_i, \forall i$ and at least one $n_i \neq m_i$) the set of hafnians is again different. We call this a *branching effect*. It seems to be a probabilistic effect but from our experiments it holds overwhelmingly. This results in a dramatic increase of orbits with different hafnian sets as we increase $|\boldsymbol{m}|$ and greatly helps in the practical use of our algorithm. The intuition behind this behavior is that when the hafnian sets differ for an orbit of $\boldsymbol{n}$, then another orbit of $\boldsymbol{m} > \boldsymbol{n}$ (obtained by adding two new rows/columns or copies thereof to the adjacency matrix corresponding to $\boldsymbol{n}$) contains as its subgraphs all the graphs that already had different hafnian sets. The question to answer is how likely this effect is to occur.

Q4. A problem closely related to the previous question is how long it takes to approach a true probability distribution within a given precision for a chosen orbit of interest. This is typically answered by the methods that are standard in random graph and probability theory.

Q5. We did not address two related experimental points in this paper. The first one is the photon losses, which will make the statistics of two isomorphic graphs different. Hence one needs to come up with cutoff estimates for distinguishing nonisomorphic graphs. The second one is how easy to estimate probabilities of photon counting experiments.

6.2 Scalability discussion

As discussed in the previous subsection, we don't know at what point two non-isomorphic graphs start to differ when sampled using a GBS device. Presumably this critical $|\boldsymbol{n}|$ grows with the graph size and the numerical evidence suggests that the growth is not exponential or even fast in general in the graph size. Nonetheless, even a moderate rate of growth of the threshold orbit $\boldsymbol{n}$ with the graph size M could be fatal for large graphs. This is because the physical interpretation of $0 < c < \|A\|_2$ that directs the amount of squeezing – a quantity directly related to $|\boldsymbol{n}|$. As we see from Fig. 7, the class with a (nearly) maximal probability (a single orbit in fact) is $\boldsymbol{n} = (0)_M$ and this is a generic case. We can tune the c parameter by increasing squeezing such that the

probability of a different (desired) orbit $O_{\boldsymbol{n}}$ increases. This typically makes the vacuum contribution smaller but still dominating the probability landscape. What happens is that the desired orbit's probability $p_{G_A}(O_{\boldsymbol{n}})$ increases with respect to the vacuum and other lower contributions but the probability flattens and therefore its values are inevitably impractically low to gather enough statistics reasonably fast. To put it differently, there is no significant peak (or a concentration effect) for the desired orbit $O_{\boldsymbol{n}}$. The situation is a bit alleviated by a heuristic observation that beyond the threshold orbit $O_{\boldsymbol{n}}$, where the difference in the hafnian set is first observed, the orbits $O_{\boldsymbol{m}}$ where $\boldsymbol{n} < \boldsymbol{m}$ overwhelmingly add to distinguishability. But it is tempting to avoid this 'probability dilution' altogether. The rise of the number of orbits for a fixed $|\boldsymbol{n}|$ is equivalent to the integer partition problem briefly discussed towards the end of Section 4.2.

Let us look closer at this problem. Any graph G_A can be encoded as a pure covariance matrix whose circuit decomposition consists of an array of M squeezers S followed by an interferometer U on M modes. Since the eigenvalues (more precisely the singular values) of A are related to the squeezing parameters, two isospectral graphs, G_A, $G_{\tilde{A}}$, have the same S – the fact already used in Lemma 17. The input to the interferometer is given by (48). Suppose that we know or suspect that a given $|\boldsymbol{n}|$ contains orbits that are different if two graphs are not isomorphic. Since an interferometer is a passive unitary operator we know the input state responsible for it – it is the state whose coefficients are $\alpha_{i\boldsymbol{n}}(r_1, \ldots, r_M)$ from (48). So the task becomes to prepare such a state. This could be a computationally hard task. Even though $\bigotimes_{k=1}^{M} S(r_k)\,|0\rangle$ is factorized the states $\sum_{i=1}^{\binom{|\boldsymbol{n}|/2+M-1}{|\boldsymbol{n}|/2}} \alpha_{i\boldsymbol{n}}(r_1, \ldots, r_M)\,|(\boldsymbol{n}, M)\rangle_i$ living in the completely symmetric subspaces are, in general, entangled. There are two problems, though. First, generation of such states could potentially require a circuit of a great depth. Second, even if $\boldsymbol{n}$ scales favorably with M, the number of coefficients $\alpha_{i\boldsymbol{n}}(r_1, \ldots, r_M)$ can be overwhelming to work with for $M \gg 0$ and so we may run into the issue of tractability to describe the necessary unitary operation. If these issues were resolved we would gained a complete control over the output distribution behavior. But we would also switch from GBS to a generalized (multiphoton) boson sampling (BS) [1]. The orbit representative is a state with a given photon number per input mode. The difference compared to BS is that we do not require the input/output state to be in the 0,1 subspace per mode. Hence we arrived (by a detour) to the output probability function whose form is most likely governed by some permanent function [35] – a form which will most likely be a variation on the reduced Kronecker product we have introduced in Sec. 4.1.

The use of a fixed input photon distribution will dramatically change the odds of detecting the difference in the probability distribution. We can take a look at Fig. 7 for, say, $|\boldsymbol{n}| = 14$ and $p_{G_A}(O_{\boldsymbol{n}})$ for all orbits in this partition will be considerably higher with their mutual ratio preserved. We leave for a future exploration the question if the probability distribution is always skewed (or even concentrated) such that some events (preferably the ones where there is a difference) are overwhelmingly likely than the others. What can we do if this is not the case? The thing we cannot do is to coarse-grain the probability more than in Eq. (44) due to lemma 17. Then, one option would be to coarse-grain the probability more than in (44) but less than in (46). The reason for this effort is to have a favorable scaling. Recall that the number of partitions of $|\boldsymbol{n}|$ grows exponentially with $|\boldsymbol{n}|$. If we partially coarse-grain a given orbit we may keep the scaling polynomial and still detect a difference for non-isomorphic graphs. The question is how to split a given orbit. At this point we can offer only certain heuristic rules based on our simulations.

Acknowledgments: The authors greatly appreciate thorough reading and suggested modifications of the manuscript by Robert Israel. This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

Data Availability Statement: Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

References

[1] Scott Aaronson and Alex Arkhipov. The computational complexity of linear optics. In STOC'11-*Proceedings of the 43rd ACM Symposium on Theory of Computing*, 333-342, ACM, New York, 2011.

[2] Dorit Aharonov, Itai Arad, Elad Eban, and Zeph Landau. Polynomial quantum algorithms for additive approximations of the Potts model and other points of the Tutte plane. *arXiv preprint quant-ph/0702008*, 2007.

[3] László Babai. Graph isomorphism in quasipolynomial time. In STOC'16-*Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing*, 684-697, ACM, New York, 2016.

[4] L. Babai, D.Yu. Grigoryev and D.M. Mount. Isomorphism of graphs with bounded eigenvalue multiplicities. In STOC'82-*Proceedings of the 14th ACM symposium on Theory of Computing*, 310–324, ACM, New York,1982 .

[5] George A Baker Jr. Drum shapes and isospectral graphs. *J. Mathematical Phys.*, 7 (1966), no 12, 2238–2242.

[6] Alexander Barvinok. *Combinatorics and Complexity of Partition Functions*, Algorithms and Combinatorics, 30. Springer, Cham, 2016. vi+303 pp.

[7] Scott D Berry and Jingbo B Wang. Two-particle quantum walks: Entanglement and graph isomorphism testing. *Phys. Rev. A*, 83 (2011), no. 4, 042317, 12 pp.

[8] Andreas Björklund, Brajesh Gupt, and Nicolás Quesada. A faster hafnian formula for complex matrices and its benchmarking on a supercomputer. *ACM J. Exp. Algorithmics*, 24 (2019), Art. 1.11, 17 pp.

[9] Kamil Brádler, Pierre-Luc Dallaire-Demers, Patrick Rebentrost, Daiqin Su, and Christian Weedbrook. Gaussian boson sampling for perfect matchings of arbitrary graphs. *Physical Review A*, 98 (2018), no. 3, 032310, 15 pp. 2018.

[10] Kamil Brádler, Robert Israel, Maria Schuld, Daiqin Su. A duality at the heart of Gaussian boson sampling, *arXiv preprint arXiv:1910.04022*, 2019.

[11] Fernando GSL Brandão and Krysta M Svore. Quantum speed-ups for solving semidefinite programs. *58th Annual IEEE Symposium on Foundations of Computer Science-FOCS* 2017, 415-426, IEEE Computer Soc., Los Alamitos, CA, 2017

[12] Fernando GSL Brandão, Amir Kalev, Tongyang Li, Cedric Yen-Yu Lin, Krysta M Svore, and Xiaodi Wu. Exponential quantum speed-ups for semidefinite programming with applications to quantum learning. *arXiv preprint arXiv:1710.02581*, 2017.

[13] Andries E. Brouwer. Parameters of strongly regular graphs. https://www.win.tue.nl/~aeb/graphs/srg/srgtab.html, June 2017.

[14] Eduardo R Caianiello. On quantum field theory I: explicit solution of Dyson's equation in electrodynamics without use of Feynman graphs. *Nuovo Cimento* (9) 10 (1953), 1634-1652.

[15] Cristian S Calude, Michael J Dinneen, and Richard Hua. QUBO formulations for the graph isomorphism problem and related problems. *Theoret. Comput. Sci.* 701 (2017), 54-69.

[16] Andrew M Childs, Robin Kothari, and Rolando D Somma. Quantum algorithm for systems of linear equations with exponentially improved dependence on precision. *SIAM J. Comput.* 46 (2017), no. 6, 1920-1950.

[17] Brendan L Douglas and Jingbo B Wang. A classical approach to the graph isomorphism problem using quantum walks. *J. Phys. A* 41 (2008), no. 7, 075303, 15 pp.

[18] David Emms, Simone Severini, Richard C. Wilson, and Edwin R. Hancock. Coined quantum walks lift the cospectrality of graphs and trees. In: Rangarajan A., Vemuri B., Yuille A.L. (eds) *Energy Minimization Methods in Computer Vision and Pattern Recognition. EMMCVPR 2005. Lecture Notes in Computer Science*, vol 3757. Springer, Berlin, Heidelberg.

[19] Michael H Freedman, Alexei Kitaev, and Zhenghan Wang. Simulation of topological field theories by quantum computers. *Comm. Math. Phys.* 227 (2002), no. 3, 587-603.

[20] Michael H Freedman, Michael Larsen, and Zhenghan Wang. A modular functor which is universal for quantum computation. *Comm. Math. Phys*. 227 (2002), no. 3, 605–622.

[21] Frank Gaitan and Lane Clark. Graph isomorphism and adiabatic quantum computing. *Phys. Rev. A*, 89 (2014), no. 2, 022342, 20 pp.

[22] John King Gamble, Mark Friesen, Dong Zhou, Robert Joynt, and S. N. Coppersmith. Two-particle quantum walks applied to the graph isomorphism problem. *Phys. Rev. A*, 81 (2010), no. 5, 052313, 11 pp.

[23] Joseph Geraci and Daniel A Lidar. On the exact evaluation of certain instances of the Potts partition function by quantum computers. *Comm. Math. Phys.* 279 (2008), no. 3, 735-768.

[24] Chris Godsil and Gordon Royle. Strongly regular graphs. In *Algebraic graph theory*, pages 217–247. Springer, 2001.

[25] C. S. Hamilton, R. Kruse, L. Sansoni, S. Barkhofen, C. Silberhorn, and I. Jex. Gaussian boson sampling. *Phys. Rev. Lett.*, 119 (2017), no. 17, 170501, 5 pp.

[26] Aram W Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum algorithm for linear systems of equations. *Phys. Rev. Lett.*, 103 (2009), no. 15, 150502, 4 pp.

[27] Harald Andrés Helfgott, Jitendra Bajpai, and Daniele Dona. Graph isomorphisms in quasi-polynomial time. *arXiv preprint arXiv:1710.04574*, 2017.

[28] Christian Kleiber and Jordan Stoyanov. Multivariate distributions and the moment problem. *Journal of Multivariate Analysis*, 113 (2013), 7-18.

[29] Regina Kruse, Craig S Hamilton, Linda Sansoni, Sonja Barkhofen, Christine Silberhorn, and Igor Jex. Detailed study of gaussian boson sampling. *Phys. Rev. A*, 100 (2019), no. 3, 032326, 15 pp.

[30] C.H.C. Little. *Combinatorial Mathematics V.: Proceedings of the Fifth Australian Conference, Held at the Royal Melbourne Institute of Technology, August 24 - 26, 1976.* Lecture Notes in Mathematics. Springer Berlin Heidelberg, 2006.
[31] Anuradha Mahasinghe, Josh A Izaac, Jingbo B Wang, and Jagath K Wijerathna. Phase-modified CTQW unable to distinguish strongly regular graphs efficiently. *J. Phys. A*, 48 (2015), no. 26, 265301, 13 pp.
[32] PW Mills, RP Rundle, VM Dwyer, Todd Tilma, Simon J Devitt, JH Samson, and Mark J Everitt. A proposal for an efficient quantum algorithm solving the graph isomorphism problem. *arXiv preprint arXiv:1711.09842*, 2017.
[33] Kenneth Rudinger, John Gamble, Mark Wellons, Eric Bach, Mark Friesen, Robert Joynt, and S. Coppersmith. Noninteracting multiparticle quantum random walks applied to the graph isomorphism problem for strongly regular graphs. *J. Phys. A*, 86 (2012), no. 2, 022334, 10 pp.
[34] Kenneth Rudinger, John King Gamble, Eric Bach, Mark Friesen, Robert Joynt, and S. N. Coppersmith. Comparing algorithms for graph isomorphism using discrete- and Continuous-Time quantum random walks. *Journal of Computational and Theoretical Nanoscience*, 10 (2013), no. 7, 1653–1661.
[35] Stefan Scheel. Permanents in linear optical networks. *arXiv preprint quant-ph/0406127*, 2004.
[36] Peter W Shor. Algorithms for quantum computation: Discrete logarithms and factoring. In *35th Annual Symposium on Foundations of Computer Science* (Santa Fe, NM, 1994), 124-134, IEEE Comput. Soc. Press, Los Alamitos, CA, 1994.
[37] Peter W Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM review*, 41 (1999), no. 2, 303–332.
[38] Jamie Smith. *Algebraic Aspects of Multi-Particle Quantum Walks*. PhD, University of Waterloo, December 2012.
[39] Ted Spence. Strongly regular graphs. http://www.maths.gla.ac.uk/~es/srgraphs.php, August 2018.
[40] Maarten Van den Nest, Wolfgang Dür, Robert Raussendorf, and Hans J Briegel. Quantum algorithms for spin models and simulable gate sets for quantum computation. *Phys. Rev. A*, 80 (2009), no. 5, 052334, 5 pp.
[41] Huiquan Wang, Junjie Wu, Xuejun Yang, and Xun Yi. A graph isomorphism algorithm using signatures computed via quantum walk search model. *J. Phys. A* , 48 (2015), no. 11, 115302, 23 pp.
[42] Paweł Wocjan and Jon Yard. The Jones polynomial: quantum algorithms and applications in quantum complexity theory. *Quantum Inf. Comput.* 8 (2008), no. 1-2, 147–180.
[43] Han-Sen Zhong et al. Quantum computational advantage using photons. *Science* 370 (2020), 1460-1463.
[44] Kenneth M Zick, Omar Shehab, and Matthew French. Experimental quantum annealing: case study involving the graph isomorphism problem. *Scientific Reports*, 5 (2015), 11168, 11 pp.

A Basic hardware setup

The basis for our graph isomorphism method is a near-term photonic quantum processor, specifically a GBS apparatus. This apparatus consists of three main components. First, squeezed states are generated in M quantum-optical modes. These states are then sent through an M-port linear-optical interferometer. Finally, a photon-number-resolving measurement is performed on each of the M output modes. The first two steps lead to the preparation of a zero-mean quantum-optical Gaussian state, which can be described efficiently using a covariance matrix. For a single mode, the covariance matrix has dimension 2×2, and encodes the covariances of the canonical quadrature operators $(\hat{x}, \hat{p})$ of that mode:

$$\sigma_{ij} = \tfrac{1}{2}\langle \hat{\xi}_i \hat{\xi}_j + \hat{\xi}_j \hat{\xi}_i \rangle - \langle \hat{\xi}_i \rangle \langle \hat{\xi}_j \rangle, \tag{62}$$

with $\hat{\xi}_k \in \{\hat{x}_k, \hat{p}_k\}$. For M modes, we have M pairs of quadrature operators and an $2M \times 2M$ covariance matrix built from the set $\hat{\boldsymbol{\xi}} \in \{\hat{x}_1, \hat{p}_1, \ldots, \hat{x}_M, \hat{p}_M\}$.

Multimode Gaussian states themselves are of limited interest in quantum computing. While they can be prepared with quantum hardware and exhibit entanglement, the covariance matrix scales linearly in the number of modes, so they can be efficiently simulated classically. However, when we introduce the photon-number measurement, the story changes. A single photon-number measurement in mode k will return a nonnegative integer $n_k \in \mathbb{N}^+$, representing the number of photons which were detected. For measurement on M modes, we denote the collective photon-number output pattern by $\boldsymbol{n} = (n_1, \ldots, n_M)$ and call it a detection event. From [25], whenever $n_i = 1, \forall i$ the probability of this detection event is proportional to a function called the *hafnian* [14] (see Def. 1):

$$p(1, \ldots, 1) = \frac{1}{\sqrt{\det \sigma_Q}} \operatorname{haf} C, \tag{63}$$

where the matrix C is obtained from σ_Q by basic matrix transformations (see Eq. (4)).

Unlike the simulation of Gaussian states, computing the hafnian is a #P-hard problem. In addition, approximating the GBS photon-number distribution is believed to be computationally hard [25]. Thus, by combining Gaussian states with photon-number measurements – representing the wavelike and particle-like properties of light, respectively – we have a physical sampling apparatus whose behaviour is classically hard to replicate. This paper explores the question of how we can leverage this GBS device for the graph isomorphism problem, specifically, how we set the squeezing and interferometer parameters to represent the problem, and how to interpret the photon-number measurement outcomes to solve the problem.

B Algorithm

Algorithm 1 GBS graph isomorphism algorithm: returns the GI invariant of adjacency matrix A with respect to orbit o, considering hafnians' to the nth power.

```
function GBS_CERT(A, o, n)
    ▷ Generate a list of unique permutations of the orbit
    perms ← unique_permutations(o)
    result ← array[len(perms)]
    for p ∈ perms do
        ▷ Generate the reduced Kronecker product of matrix A
        Ap ← kron_reduced(A, p)
        ▷ Append the hafnian to the nth power
        result ← append(result, haf(Ap)^n)
    end for

    ▷ Calculate the sum of hafnians
    hafSum ← sum(results)

    ▷ Calculate the mean photon distribution for the orbit
    phDist ← array[len(o)]
    for i ∈ len(perms) do
        for j ∈ len(o) do
            phDist[j] ← phDist[j] + perms[i, j]*result[i]
        end for
    end for
end function
```

Algorithm 2 Function to return the reduced Kronecker product of matrix A, given a sequence of integers n indicating the multi-mode photon detection event.

```
function KRON_REDUCED(A, n)
    rows ← array()
    for i ∈ len(n) do
        for j = 0, n[i] do
            rows ← append(rows, i)
        end for
    end for
    return A[rows][rows]
end function
```

Algorithm 3 Function to return unique permutations of an orbit

```
function UNIQUE_PERMUTATIONS(orbit)
    if len(orbit) = 1 then
        ▷ If orbit is length 1, return the value
        return orbit[0]
    else
        ▷ Else, store the list of unique elements in the orbit
        elements ← drop_duplicates(orbit)
        for e ∈ elements do
            ▷ Unique elements except e
            remaining ← elements - e
            for p ∈ unique_permutations(remaining) do
                ▷ Use recurrence to concatenate element e with all remaining permutations of elements
                return e + p
            end for
        end for
    end if
end function
```

C List of symbols

In this appendix we summarize with a table the most important symbols and their meaning.

description	notation	defined at page
vectors	$\boldsymbol{n} = (n_1, \ldots, n_M), \boldsymbol{x} = (x_1, \ldots, x_M)$	page 3
partial derivative operator	$\partial^{n_i}_{x_i,\overline{x}_i} = \frac{\partial^{n_i}}{\partial x_i^{n_i}} \frac{\partial^{n_i}}{\partial \overline{x}_i^{n_i}}$	page 4
symmetric group of bijections $\sigma : [n] \to [n]$	$\mathfrak{S}_n$	page 4
$N \times N$ hermitian, psd and positive definite	$\mathcal{H}_N \supset \mathcal{H}_{+,N} \supset \mathcal{H}_{++,N}$	page 4
hafnian	$\mathrm{haf}\, C$	page 4
measurement probability of $\boldsymbol{n}$	$p(\boldsymbol{n})$	page 4
covariance matrix	σ_Q	page 5
graph	G	page 5
$i-th$ eigenvalue of G	$\lambda_i(G)$	page 5
strongly regular graph with parameters	$\mathrm{SRG}(N, k, \lambda, \mu)$	page 6
all-ones $\lvert\boldsymbol{n}\rvert \times \lvert\boldsymbol{n}\rvert$ matrix	$\mathbb{J}_{\lvert\boldsymbol{n}\rvert}$	page 7
reduced Kronecker product	$A \circledast \mathbb{J}_{\lvert\boldsymbol{n}\rvert}$	page 7
moments of Gaussian distribution induced by Σ	$\mu_{n_1,\ldots,n_{2M}}(\Sigma)$	page 11
multinomial coefficients	$\binom{\boldsymbol{n}}{\boldsymbol{a}_1,\boldsymbol{a}_2,\ldots,\boldsymbol{a}_m}$	page 15
symmetrized sums	$\mu(\boldsymbol{n}, A)$	page 16
probability of orbit $O_{\boldsymbol{n}}$	$p_G(O_{\boldsymbol{n}})$	page 18
coarse-grain probability distribution	$p_G(\lvert\boldsymbol{n}\rvert)$	page 19
k-mode partition-averaged photon distribution	$\langle n_k \rangle_G$	page 20
coarse-grained version	$\langle\langle n_k \rangle\rangle_G$	page 20