THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Simulating boson sampling in lossy architectures

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Quantum

OPEN ACCESS

Simulating boson sampling in lossy architectures

Raúl García-Patrón[1], Jelmer J. Renema[2,3], and Valery Shchesnovich[4]

[1]Centre for Quantum Information and Communication, Ecole Polytechnique de Bruxelles, CP 165, Université Libre de Bruxelles, 1050 Brussels, Belgium

[2]Clarendon Laboratory, Department of Physics, University of Oxford, Oxford OX1 3PU, United Kingdom

[3]University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands

[4]Centro de Ciências Naturais e Humanas, Universidade Federal do ABC, Santo André, SP, 09210-170 Brazil.

July 25, 2019

Photon losses are among the strongest imperfections affecting multi-photon interference. Despite their importance, little is known about their effect on boson sampling experiments. In this work we show that using classical computers, one can efficiently simulate multi-photon interference in all architectures that suffer from an exponential decay of the transmission with the depth of the circuit, such as integrated photonic circuits or optical fibers. We prove that either the depth of the circuit is large enough that it can be simulated by thermal noise with an algorithm running in polynomial time, or it is shallow enough that a tensor network simulation runs in quasi-polynomial time. This result suggests that in order to implement a quantum advantage experiment with single-photons and linear optics new experimental platforms may be needed.

1 Introduction

In 2003 Knill, Laflamme and Milburn showed that single-photon sources and linear optics are sufficient to achieve universal quantum computation [16]. A single-photon and linear optics version of measurement based quantum computation has also been thoroughly studied [25, 46]. In both proposals, a key component to reach universality is the capability for a measurement outcome to change the gates implemented later in time, i.e., using active feed-forward, a challenging experimental requirement [39]. In 2010, Aaronson and Arkhipov demonstrated that removing the feed-forward condition provides a framework, called

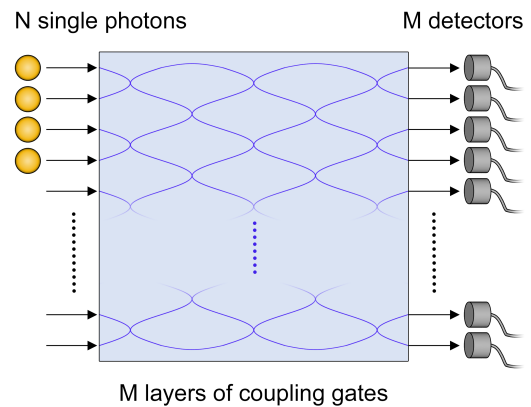Raúl García-Patrón: raulgarciapatron@gmail.com

Figure 1: In a boson sampling device N single photons are sent over an M mode linear optical circuit composed of M layers of two-mode coupling gates and detected at the output with photon counting detectors. The circuit being selected randomly from the Haar measure, we can place the N photons over the first N modes without lost of generality.

boson sampling, that does not seem to be sufficient for universal quantum computation but is hard to simulate on a classical computer [2]. The key idea behind their proof is the connection between the output statistics of non-interacting indistinguishable photons and the permanent [8], a quantity know to be #P-hard to compute [1, 28]. As shown in Figure 1, a boson sampling device implements an interference of N single photons over a randomly-selected M-mode linear optical interferometer and measures each output mode with a photon counting detector. An M-mode linear optical interferometer can be build out of two-mode couplers (beamsplitters) acting on neighboring modes and single-mode phases gates. In order to generate an arbitrary linear optical circuit, a depth of M is needed [12, 43].

Since 2013, a variety of experimental quan-

tum optics groups have implemented proof-of-principle implementations based on different architectures, such as reconfigurable integrated photonic circuits [7, 47, 51], fiber-loops [20], 3D waveguides [14] or multimode fibers [15], that have the potential of being scalable and therefore are candidates for a quantum advantage demonstration. The motivation to further simplify the experimental scheme led to the proposal of scattershot boson sampling, a sampling problem as hard to simulate as the initial boson sampling proposal. Scattershot boson sampling solves the problem of obtaining N single photons from state of the art probabilistic single-photon sources by using M heralded two-mode squeezed vacuum state sources, one per input mode of the boson sampling circuit [30].

The lack of fault-tolerant error correction in quantum advantage architectures, such as in boson sampling experiments, implies that increasing the size and depth of the circuit would ultimately lead to a system that is equivalent to sampling random noise. Therefore, the existence of an opportunity window where noise has not yet destroyed the quantum advantage but a classical algorithm, such as [13, 33], can no longer simulate the system is fundamental for a conclusive quantum advantage experiment. It is therefore of paramount importance to have a good understanding of when noise makes a quantum advantage architecture classically simulatable. Rahimi-Keshari et al. provided in [42] a first rigorous bound, which required both losses and dark counts of the detectors. Unfortunately, this bound is independent of the size of the system and can not provide an answer in terms of losses independently of additional noise (dark-counts of detectors).

Together with the indistinguishably of photons, for which an algorithm to simulate partially distinguishable photons was demonstrated recently in [45], losses are the most damaging imperfections challenging boson sampling. Despite its importance, little is know about the effect of losses. Firstly, it was proven in [3] that boson sampling with losses remains hard in the regime of a constant number of photons lost, a rather limiting assumption. Secondly, modeling a lossy circuit as a larger lossless one with additional environmental modes, it is rather straightforward to see that the algorithm of Clifford and Clifford [13]

can be extended to lossy architectures with only a constant overhead. This has two implications: (i) if an ideal boson sampling circuit with N input photons can be simulated (with an exponential-time algorithm), the same can be achieved for arbitrary losses; (ii) lossy multi-photon interference becomes classically simulatable when the number of photons left scales as $O(\log N)$. In section 3 of this work we will considerably improve over this trivial corollary of [13] by proving that if the numbers of photons that reach the detectors are less than $O(\sqrt{N})$, a boson sampling experiments becomes classically efficiently simulatable.

Most experimental boson sampling architectures, and all for which interference has been shown for more than 2 photons, are based on a planar geometry of depth proportional to the number of input systems, where the loss per coupler in the circuit is constant, leading to a law of exponential decay of the transmission with the depth of the circuit. In section 3 of this work we show that for those platforms, and platforms which have similar exponential decay, boson sampling experiments can be efficiently simulated classically. Therefore, we believe that for single-photons and linear optics to remain competitive in the race for a quantum advantage demonstration new ideas are needed. More precisely, we show that for those platforms either the depth of the circuit (D) is large enough $(D \geq O(\log M))$ that it can be simulated by thermal noise with an algorithm running in polynomial time, or the depth of the circuit is short enough $(D \leq O(\log M))$ that a tensor network simulation, similar in spirit to [49], runs in quasi-polynomial time.

Not all optical architectures suffer from an exponential decay of the transmission, for example free-space optics has a quadratic decay of transmission. In section 4 we extend the validity of the thermal noise simulation to this family of architectures, to scattershot boson sampling and boson sampling architectures where the photon-counting detectors are replaced by Gaussian measurements [11, 31].

In sections 5, 6 of the manuscript we provide the detailed proofs stated in section 3 and in section 7 we generalize the result to non-uniform losses. Finally, we conclude in section 8.

2 Preliminaries

In this section, we review the concept of boson sampling, the most established model of losses used in the quantum optics literature, and the trace distance and its properties, needed in the understanding of our main result in section 3. In this work we will use the notation $|\overline{n}\rangle = |n_1\rangle \otimes |n_2\rangle \otimes \ldots \otimes |n_M\rangle$, for the M-mode Fock basis, where $\overline{n}$ correspond to a vector of M integers and $|n_i\rangle$ is the Fock state corresponding to n_i photons in mode i.

2.1 The ideal boson sampling model

The boson sampling proposal concerns the interference of a multi-photon input state $|\overline{1}_N\rangle = |1\rangle^{\otimes N} \otimes |0\rangle^{\otimes M-N}$, over an M-mode linear optics interferometer modeled as a linear transformation of the annihilation operators

$$\hat{b}_i = \sum_j U_{ij}\hat{a}_j, \tag{1}$$

or $\hat{\mathbf{b}} = U\hat{\mathbf{a}}$ for an equivalent compact notation. We remark that the unitarity of U guarantees the preservation of the total photon number and that U is an $M \times M$ matrix acting on the creation operators, to which corresponds a homomorphism $\varphi(U)$ of dimension $\binom{N+M-1}{N}$ acting on the M-mode Fock space [2]. At the output of the interferometer we implement a measurement in the photon number basis on each mode, where the probability of obtaining an outcome $\overline{z}$ reads $|\langle\overline{z}|\varphi(U)|\overline{n}\rangle|^2$. When $\overline{z}$ corresponds to a string of bits, the probability outcome is connected to the permanent of a submatrix of U, a crucial tool in the hardness proof of boson sampling [2].

A necessary condition in the proof of the hardness of boson sampling is the fact the U needs to be a Haar random unitary. The proposal by Reck et al. [43] showed that any general linear-optics transformation can be achieved with a planar circuit composed of two-mode gates, using $M(M-1)/2$ gates distributed over $2M-3$ layers and M parallel modes. A recent improvement in [12] remarkably brought this result to depth $M+1$, which is one unit close to the lower-bound M obtained from a simple counting argument based on the degrees of freedom of a unitary matrix. In order to make our result as general as possible we will consider the depth of the circuit D as an additional free parameter.

In the initial boson sampling proposal, the proof necessitates a polynomial relation between the number of photons and modes ($M = O(N^5 \log^2 N)$). In this work we consider the generalized relation

$$N = kM^\gamma, \tag{2}$$

where $0 < k < 1$ and $0 < \gamma \leq 1$. It is easy to see that $\gamma = 1/2$ corresponds to the bosonic birthday paradox ratio [4]. This ratio ensures that for input states $|\overline{n}\rangle$ composed of single photons, the probability of two or more boson bunching at the output is negligible (on average over the set of Haar random unitaries and on the asymptotic limit of a large system). The case $\gamma = 1/6$ corresponds to $M = O(N^6)$, which guarantees the condition in the hardness proof in [2] to hold.

Finally, $\gamma = 1$ corresponds to the regime where the density of photons (with k satisfying $k \leq 1$ for single photons at the input) remains constant while the size of the system increases, as opposed to the original boson sampling proposal where it decreases with the size of the system.

2.2 Modeling losses

A lossy linear optics circuit, as in Figure 2 a), can be mathematically modeled by a complex matrix A satisfying $AA^\dagger \leq I$ and transforming the annihilation operators of M input modes $\hat{\mathbf{a}}$ and M environment modes $\hat{\mathbf{e}}$ as

$$\hat{\mathbf{a}}_{\text{out}} = A\hat{\mathbf{a}}_{\text{in}} + \sqrt{I - AA^\dagger}\hat{\mathbf{e}}. \tag{3}$$

A has a singular value decomposition $A = V\hat{\mu}W$, where V and W are unitary matrices and $\hat{\mu} = \text{diag}(\sqrt{\mu_1}, \sqrt{\mu_2}, ..., \sqrt{\mu_M})$ is a diagonal matrix of real values satisfying $\mu_i \in [0,1]$. The singular value decomposition has a very natural interpretation, see Figure 2 b), which is that the real interferometer with losses characterized by A has an equivalent circuit composed by a lossless linear optics transformation V, followed by M parallel set of pure-loss channels of transmission μ_i each, and a final lossless linear optics transformation W. A pure-loss channel is equivalent to a coupling interaction of transmission μ_i between our physical mode i and an environmental mode, see [17] for details. The matrix A can be efficiently inferred using a simple tomographic technique that only needs two-mode interferences of classical laser light [40].
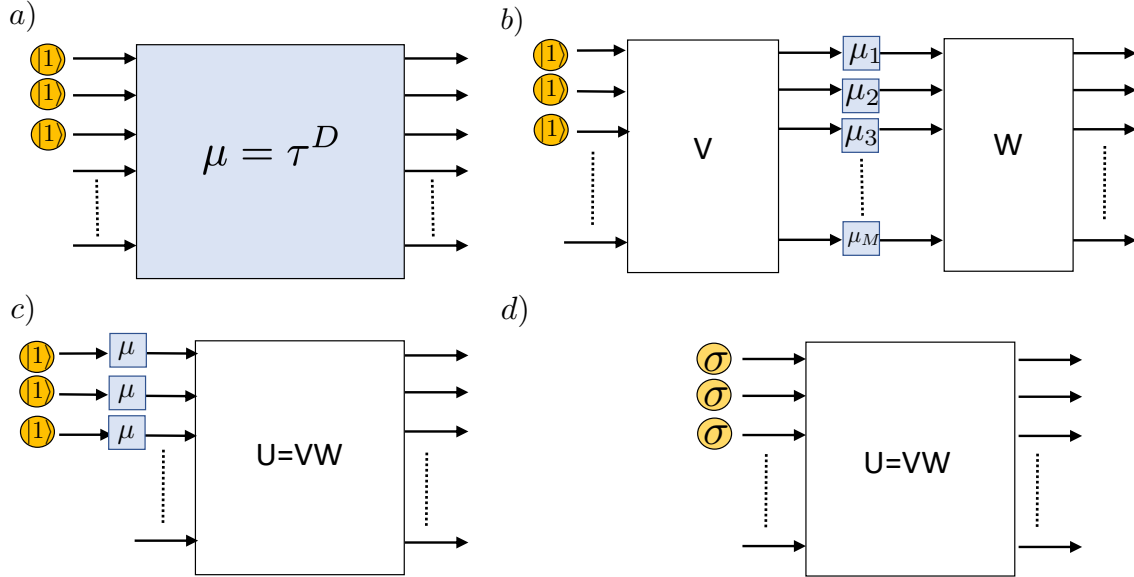
Figure 2: a) N single photons are sent over a lossy M mode linear optical circuit composed of D layers of two-mode couplers. An uniform transmission per coupler of τ leads to an exponentially decay of the transmission $\mu = \tau^D$; b) the real scheme in a) is indistinguishable from a circuit composed by a lossless linear optics transformation V, followed by M parallel set of pure-loss channels of transmission μ_i each, and a final lossless linear optics transformation W; c) In the case where all the μ_i are equal, the virtual representation can be simplified to a layer of M identical pure-loss channels of transmission μ followed by a virtual lossless linear optics transformation $U = VW$; d) The action of a pure-loss channel of transmission μ on a single-photon state $|1\rangle$ is equivalent to an erasure channel of probability μ, which outputs the mixed state $\sigma = (1 - \mu)|0\rangle\langle 0| + \mu|1\rangle\langle 1|$.

In practice, a linear optics circuit is composed of a network of two-mode couplers and single-mode phase gates, where each layer of gates of an M-mode linear optics circuit is given by a direct product of local 2×2 linear transformations and complex scalars, resulting in a $M \times M$ complex banded matrix A_i of bandwidth 1. The total linear optics circuit transformation results from the multiplication of D matrices A_i, i.e., $A = A_1 A_2 ... A_D$.

All currently existing architecture proposals to implement a boson sampling experiment, integrated photonic circuits, fiber-optical links and 3D-waveguides, suffer from exponential decay of the transmission with the length of the circuit. An intuitive explanation is that every photon has a constant probability of being lost per unit of length of the circuit or per layer of coupling gates. For a planar circuit composed of D layers of gates, where every gate has a transmission coefficient τ, we obtain that all the μ_i are equal and the transmission follows an exponential decay rule $\mu = \tau^D$. Because $\hat{\mu} = \sqrt{\tau^D} I$ commutes with any matrix, we can simplify the virtual representation of A to

a layer of M identical pure-loss channels of transmission μ followed by a virtual lossless linear-optics transformation $U = VW$ (see Figure 2 c)).

The action of a pure-loss channel of transmission μ into a single-photon state $|1\rangle$ is equivalent to an erasure channel of probability μ, resulting into a mixed state $\sigma = (1 - \mu)|0\rangle\langle 0| + \mu|1\rangle\langle 1|$, see Figure 2 d). Therefore, boson sampling over a realistic interferometer with uniform losses is equivalent to an ideal boson sampler over its virtual circuit $U = VW$ where we replace each of the N single-photons of the input state, located in the first N modes, by the state σ, leading to a global input state

$$\rho_{\text{in}} = \sigma^{\otimes N} \otimes |0\rangle\langle 0|^{\otimes(M-N)},$$
$$\text{with} \quad \sigma = (1 - \mu)|0\rangle\langle 0| + \mu|1\rangle\langle 1|. \quad (4)$$

2.3 Trace distance and its properties

Before presenting our result we need to provide some definitions and properties of the trace distance [34]. The trace distance between two quantum states ρ and σ reads (here and below $\|...\|$

stands for the 1-norm)

$$D(\rho, \sigma) = \frac{1}{2}\text{Tr}\left(\sqrt{(\rho - \sigma)^2}\right) = \frac{1}{2}||\rho - \sigma||. \quad (5)$$

When both density matrices are diagonal in the same basis, i.e., $\rho = \sum_i p_i|i\rangle\langle i|$ and $\rho = \sum_i q_i|i\rangle\langle i|$, the trace distance is equivalent to the total variation distance of its corresponding eigenvalues

$$D(\rho, \sigma) = \frac{1}{2}||\overline{p} - \overline{q}|| = D(\overline{p}, \overline{q}). \quad (6)$$

For simplicity, and following [34], we will use the abuse of notation $D(\overline{p}, \overline{q})$ for the total variation distance between two probability distributions. In what follows we will use three important properties of the trace distance:

1. Its invariance under unitary transformation $D(U\rho U^\dagger, U\sigma U^\dagger) = D(\rho, \sigma)$.

2. Its contractivity under a trace-preserving CP map $\mathcal{M}$, i.e., $D(\mathcal{M}(\rho), \mathcal{M}(\sigma)) \leq D(\rho, \sigma)$.

3. Let $\{E_x\}$ be a POVM, with $p(x) = \text{Tr}[\rho E_x]$ and $q(x) = \text{Tr}[\sigma E_x]$ the probabilities of obtaining a measurement outcome labeled by x. Then $D(p(x), q(x)) \leq D(\rho, \sigma)$. Where there always exist a POVM that saturates the bound.

Using the triangle inequality one can prove the following lemma.

Lemma 1. *The trace distance between N copies of ρ and σ satisfies the bound*

$$D(\rho^{\otimes N}, \sigma^{\otimes N}) \leq ND(\rho, \sigma). \quad (7)$$

For simplicity we show the proof for $n = 2$, where we make use of the triangle inequality and $||\rho \otimes \sigma|| = ||\rho||||\sigma||$:

$$\begin{aligned}
||\rho^{\otimes 2} - \sigma^{\otimes 2}|| &= ||\rho^{\otimes 2} - \sigma \otimes \rho + \sigma \otimes \rho - \sigma^{\otimes 2}|| \\
&\leq ||\rho^{\otimes 2} - \sigma \otimes \rho|| + ||\sigma \otimes \rho - \sigma^{\otimes 2}|| \\
&= ||(\rho - \sigma) \otimes \rho|| + ||\sigma \otimes (\rho - \sigma)|| \\
&= 2||\rho - \sigma||
\end{aligned}$$

Its generalization to $n > 2$ is straightforward.

3 Main result

In this section we summarize the results of this manuscript, the detailed derivation of the lemmas 3, 5 in the proofs are presented in sections 5 and 6. To make our proofs more accessible, we first restrict to the scenario of uniform loss and later generalize the results to arbitrary circuits in Section 7, where we show that any result that holds for uniform μ can be generalized to $\mu_{max} = \max_i \mu_i$.

3.1 Lossy boson sampling as thermal noise

The key element of our proof is to approximate the input state ρ_{in} in equation (4) with the state $\rho_T = \sigma_{\text{th}}^{\otimes N} \otimes |0\rangle\langle 0|^{\otimes(M-N)}$, composed of N thermal states in the first N modes and the remaining $M-N$ input modes in a vacuum state. A thermal state is given by the Bose-Einstein distribution

$$\sigma_{\text{th}} = (1 - \lambda) \sum_{x=0}^{\infty} \lambda^x |x\rangle\langle x|, \quad (8)$$

where $\lambda = \frac{z}{1+z}$, with z being the average number of photons. This allows us to proof the following theorem.

Theorem 2. *For any multi-photon interferometer circuit of uniform transmission μ satisfying*

$$\mu \leq \sqrt{\frac{\epsilon}{N}}, \quad (9)$$

we have $D(\rho_{in}, \rho_T) \leq \epsilon$.

Proof. One can use Lemma 1 to obtain the bound $D(\rho_T, \rho_{\text{in}}) \leq ND(\sigma_{\text{th}}, \sigma)$. To have $D(\rho_T, \rho_{\text{in}}) \leq \epsilon$ we require that $D(\sigma_{\text{th}}, \sigma) \leq \epsilon/N$. A simple calculation gives $D(\sigma_{\text{th}}, \sigma) = \frac{1}{2}\left(\lambda^2 + |\mu - \lambda| + |\lambda(1 - \lambda) - \mu|\right)$. There are three cases: (i) $\mu \leq \lambda(1-\lambda)$, (ii) $\lambda(1-\lambda) \leq \mu \leq \lambda$ and (iii) $\lambda \leq \mu$. All of them lead to the condition $\lambda \leq \sqrt{\epsilon/N}$ and three intervals for μ given λ. The union of the latter reads $\lambda - \epsilon/N \leq \mu \leq \lambda(1 - \lambda) + \epsilon/N$. We are looking for any λ and the maximal possible value of μ satisfying $D(\sigma_{\text{th}}, \sigma) \leq \epsilon/N$. Notice that the upper bound $\mu \leq \lambda(1 - \lambda) + \epsilon/N$ grows with λ (for $\lambda \leq 1/2$) and reaches its maximal value for $\lambda_{\max} = \sqrt{\epsilon/N}$ resulting in equation (9). Without loss of generality, we can set $\lambda = \mu$. □

In what follows we use the notation $p_{\text{in}}(\overline{n})$ and $p_T(\overline{n})$ for the outcome probabilities resulting from

applying a linear interferometer to ρ_{in} and ρ_{T} followed by a photon counting measurement. Because quantum operations and measurement can only decrease the trace distance, a corollary of Theorem 2 is the bound $D\left(p_{\text{in}}(\overline{n}), p_{\text{T}}(\overline{n})\right) \leq \epsilon$. Therefore, any classical algorithm efficiently simulating a boson sampling experiment with input state ρ_{T} will be a good ϵ-approximation of a lossy boson sampling experiment satisfying equation (9). Theorem 2 formalizes, for multi-photon interference, the commonly held belief that any quantum supremacy experiment without access to error correction is equivalent to random noise after some noise threshold.

Let's analyze the prospect of a finite size boson sampling experiment under realistic conditions. A simple calculation shows that for a multi-photon interference of N photons and transmission per coupler of $\tau = 1 - x$, the depth $\tilde{D}$ above which the output becomes ϵ-close to a thermal state reads

$$\tilde{D} = \frac{\log \frac{N}{\epsilon}}{2 \log \frac{1}{1-x}}. \tag{10}$$

Setting $\epsilon = 10^{-6}$, a boson sampling experiment with $N = 100$ photons, selected in order to discard a brute-force simulation via permanent calculations using [13], and realistic values for the loss of $x = 10^{-3}$ per coupler in the circuit [10, 19, 21, 48], we obtain $\tilde{D} = 9.205$. This value is smaller than the birthday paradox constraint $D = M \geq N^2$ and many orders of magnitude bellow the true boson sampling condition $O(N^5 \log^2 N)$. This is sufficient to discard a quantum supremacy experiment within the boson sampling paradigm with current experimental capabilities and theoretical know-how.

In order to satisfy the classical simulatability condition in equation (9) the transmission μ needs to decrease for fixed ϵ and increasing N. This shows that the bound is relatively loose, and improvements are certainly possible. Nonetheless, it is sufficient to prove that platforms suffering from exponential decay of transmission, such as integrated photonics and fiber optics, can be efficiently simulatable on a classical computer, as shown in the next subsection.

3.2 Exponential decaying transmission

In this section we provide an efficient algorithm for simulating multi-photon interference over ar-

chitectures with exponential decay of transmission.

3.2.1 Sufficient condition for efficient classical simulation of multi-photon interference

The following lemma on the efficient classical simulation of photo-counting of interfering thermal states, which was implicit in [41], is proven in section 5,

Lemma 3. *There exists a polynomial time classical algorithm that simulates the evolution of a thermal state ρ_T over an ideal or lossy interferometer followed by measurement in the photon number basis, where the output distribution is ϵ-close to the ideal one and the computational running time scales as $O\left(\frac{MN^2}{\epsilon}\left[\log N + \log(1/\epsilon)\right]^2\right)$ for transmission μ satisfying equation (9).*

As detailed in section 5 the algorithm combines the three following well-know facts in quantum optics. Firstly, any thermal state ρ_{T} has a Glauber-Sudarshan P-representation as a mixture of an N-mode tensor product of coherent states $|\boldsymbol{\alpha}\rangle = \bigotimes_{i=1}^{M}|\alpha_i\rangle$ according to a Gaussian distribution $P(\boldsymbol{\alpha})$ [18]. Secondly, a linear-optical circuit characterized by a unitary matrix U transforms a tensor product of coherent states $\boldsymbol{\alpha}$ into another tensor product of coherent states $|\boldsymbol{\beta}\rangle$ satisfying $\boldsymbol{\beta} = U\boldsymbol{\alpha}$. Thirdly, coherent states follow a Poisson photon number distribution, which can be sampled efficiently. Finally, we exploit the fact that there exist efficient constellations of coherent states that are arbitrarily close to the ideal thermal state with probability distribution $P(\boldsymbol{\alpha})$ [27]. The combination of Theorem 2 and Lemma 3 will allow us to simulate a lossy boson sampling architecture composed of D layers of gates with exponentially decaying transmission $\mu = \tau^D$, as stated in the following theorem.

Theorem 4. *The statistics $p_{in}(\overline{n})$ of N photons interfering over an M-mode linear optics planar circuit of depth D, transmission τ per layer of gates, and a relation between photons and modes given by equation (2) can be approximated with trace distance error ϵ in polynomial time for $D \geq D^*$, where*

$$D^* = \frac{1}{2\log\left(\frac{1}{\tau}\right)}\left[\gamma \log M + \log\left(\frac{k}{\epsilon}\right) + \log 2\right]. \tag{11}$$

Proof. Any classical algorithm that generates a distribution $\tilde{p}_{\mathrm{T}}(\overline{n})$ approximating the sampling from an ideal thermal state distribution $p_{\mathrm{T}}(\overline{n})$, where $\lambda = \mu$, satisfies the bound

$$D(p_{\mathrm{in}}(\overline{n}), \tilde{p}_{\mathrm{T}}(\overline{n})) \leq D(p_{\mathrm{in}}(\overline{n}), p_{\mathrm{T}}(\overline{n})) + D(\tilde{p}_{\mathrm{T}}(\overline{n}), p_{\mathrm{T}}(\overline{n})) \leq D(\rho_{\mathrm{in}}, \rho_{\mathrm{T}}) + D(\tilde{p}_{\mathrm{T}}(\overline{n}), p_{\mathrm{T}}(\overline{n})), \qquad (12)$$

where we use the triangle inequality in the first inequality and the fact that a measurement over a quantum state can only decrease its trace norm in the second. We can now apply Theorem 2 to set the bound $D(\rho_{\mathrm{in}}, \rho_{\mathrm{T}}) < \epsilon/2$ and Lemma 3 to bound $D(\tilde{p}_{\mathrm{T}}(\overline{n}), p_{\mathrm{T}}(\overline{n})) < \epsilon/2$. Then, the classical simulability condition $D \geq D^*$ can be trivially derived starting from the condition $\mu \leq \sqrt{\epsilon/(2N)}$ adapted from Theorem 4, replacing the transmission by $\mu = \tau^D$, taking the logarithm and replacing N by eq. (2). Therefore, when the condition $D \geq D^*$ is satisfied, by properly selecting a thermal state ρ_T satisfying $\lambda = \mu$ and running the algorithm sketched in the discussion after Lemma 3 (see section 5 for details), $\tilde{p}_{\mathrm{T}}(\overline{n})$ provides an ϵ-approximation of $p_{\mathrm{in}}(\overline{n})$ in polynomial time. $\qquad\square$

3.2.2 Simulation of shallow boson sampling circuits using tensor networks

In section 6 we show how one can simulate an ideal boson sampling circuit using tensor network techniques, which can be summarized in the following lemma.

Lemma 5. *An ideal boson sampling circuit with N interfering photons over an M-mode linear interferometer of depth D can be simulated exactly using tensor networks with a computational running time $O(M^2(N+1)^{8D})$.*

Tensor networks are a way of encoding quantum states and operating with them that have proven to be very successful in many-body physics [35, 36, 53, 56]. Our tensor network proof is a quantum optics version, adapted from [49], of the well-know result that logarithmic-depth planar circuit of M qudits can be simulated on polynomial time [22]. It is easy to see that when the depth of the circuit scales logarithmically with the number of modes M and N satisfies equation (2), our algorithm runs in quasipolynomial time. This is due to the unbounded nature of the Hilbert space of optical modes. In order to have an exact simulation we need to fix the local dimension on each mode to be as large as the total

number of photons in the circuit, which results into a quasipolynomial-time algorithm.

3.2.3 Efficient simulation of architectures with exponential decaying transmission

Now, combining theorem 4 and lemma 5 we can classically simulate any multi-photon interference architecture that has an exponential decaying transmission, as stated by the following theorem.

Theorem 6. *The statistics of N photons interfering over an M-mode planar linear-optical circuit of depth D, transmission τ per layer of gates, and a relation between photons and modes given by equation (2) can be approximated with trace distance error ϵ in polynomial time for $D \geq D^*$ and in quasi-polynomial time for $D \leq D^*$, where D^* is defined in equation (11).*

Proof. The classical simulability under condition $D \geq D^*$ is a direct corollary of Theorem 4. For a uniform losses circuit satisfying $D \leq D^*$ we exploit the equivalence between a lossy multi-photon interference and an ideal interference over the circuit U with input state $\rho_{\mathrm{in}} = \sigma^N \otimes |0\rangle\langle 0|^{M-N}$, as explained in subsection 2.2. We model the initial state ρ_{in} by starting with N single-photons and keeping each one with probability $\mu = \tau^D$, while the rest are transformed to vacuum inputs. We then proceed with the tensor network simulation of U using lemma 5, where we just need to change the input tensor accordingly to the random sequence of (surviving) input single-photons. This leads to a quasi-polynomial time algorithm with a running time $\mathrm{O}\left(M^{\frac{\gamma^2}{2\log\frac{1}{\epsilon}}\log M}\right).$ $\qquad\square$

4 Additional results

In this section we extend the previous results to architectures suffering from algebraic decay of transmission and to other quantum optics alternatives to boson sampling.

4.1 Algebraic decay of transmission

Not all optical architectures suffer from an exponential decay of the transmission, for example free-space optics suffers from a decay of transmission scaling as $1/D^2$. Suppose that a given architecture follows the following algebraic decay of losses

$$\mu = \left(1 + \frac{D}{d}\right)^{-\beta}, \qquad (13)$$

where d is a length scale that together with the parameter β model the algebraic decrease of transmission. Then theorem 4 can be adapted to the following weaker form.

Corollary 7. *The statistics of N photons interfering over an M-mode linear optics planar circuit of depth D, with algebraic losses given by eq. (13), and a relation between photons and modes given by eq. (2) can be approximated with trace distance error ϵ in polynomial time when $D \geq D^*$, where*

$$D^* = d\left[\left(\frac{2k}{\epsilon}\right)^{\frac{1}{2\beta}} M^{\frac{\gamma}{2\beta}} - 1\right]. \qquad (14)$$

It is not difficult to check that when the condition $\gamma/\beta < 2$ is satisfied, there always exist an M^* such that the condition $D^* \leq M - 1$ is satisfied for all $M \geq M^*$. This shows that any boson sampling experiment, which needs a depth $D = M$, on an architecture causing algebraic decay of optical transmission satisfying $\gamma/\beta < 2$ will be classically simulatable by an approximation by thermal noise sampling for all $M \geq M^*$.

4.2 Generalization to alternative boson sampling proposals

Scattershot boson sampling was presented in [30] to circumvent the main problem of the non-deterministic nature of state-of-the-art photon sources, where the probability of firing N photon at the same time decays exponentially. The protocol starts by generating M two-mode squeezed vacuum states,

$$|\psi\rangle = \sqrt{1 - \lambda} \sum_{n=0}^{\infty} \lambda^{n/2} |n\rangle |n\rangle, \qquad (15)$$

where λ is the same parameter as in the definition of a thermal state, as a two-mode squeezed vacuum state is its purification. Then we send

half of the two-mode squeezed vacuum states through a boson sampling circuit, while the remaining modes are used to herald the presence of photons. By properly tuning the squeezing parameter λ one can guarantee that most of the heralded sequences are collision-free, i.e., satisfy the birthday-paradox condition. The price to pay is that the modes where the photons enter the circuit are completely randomized, which is not a problem for boson sampling as the circuit is anyway randomly generated according to the Haar random distribution. Because right after the heralding process the setup is strictly equivalent to a traditional boson sampling device, up to the randomization of the modes where the single-photons enter the interferometer, both of our simulation algorithms (thermal state sampling and tensor network simulation) can be trivially adapted. We only need to randomly generate valid heralding sequences following the distribution given by eq. (15) and depending on the obtained heralded value we run the boson sampling simulations presented in subsection 3.2.1 and detailed in section 5. The only difference is that now the input photons enter the interferometer on a random selection of N input modes.

More recently, a variant of boson sampling, where photon detectors are replaced by a Gaussian measurement, has been proposed [11, 31]. Because quantum operations and any measurement can only decrease the trace distance, the outcome statistics of this alternative proposals will also remain ϵ-close. The evolved thermal state being Gaussian, we can extend our result to this scenario by using well-know techniques of simulating Gaussian measurement over Gaussian states [32, 52].

5 Proof of Lemma 3

An idealized algorithm for simulating the photo-counting of a set of interfering thermal states is composed of three steps. Firstly, any thermal states ρ_{T} has a Glauber-Sudarshan P-representation as a mixture of N-mode coherent states $|\boldsymbol{\alpha}\rangle \equiv |\alpha_1, ..., \alpha_N\rangle = \bigotimes_{i=1}^{N} |\alpha_i\rangle$ according to a Gaussian distribution

$$\rho_{\mathrm{T}} = \int_{\mathbb{C}^N} d\boldsymbol{\alpha} \, p(\boldsymbol{\alpha}) |\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}| \otimes |\mathbf{0}\rangle\langle\mathbf{0}|^{M-N}, \qquad (16)$$

where

$$p(\boldsymbol{\alpha}) = \prod_{i=1}^{N} \left[\frac{d^2\alpha_i}{\pi z} \exp\left(-\frac{|\alpha_i|^2}{z} \right) \right]. \qquad (17)$$

Secondly, a linear-optical circuit characterized by a unitary matrix U transforms a tensor product of coherent states $\bigotimes_{i=1}^{M} |\alpha_i\rangle$, into another tensor product of coherent states $\bigotimes_{i=1}^{M} |\beta_i\rangle$ with amplitudes

$$\beta_i = \sum_{j=1}^{M} U_{ij}\alpha_j. \qquad (18)$$

In other words, coherent states remain in a tensor product form while evolved through a linear optical circuit. Thirdly, coherent states follow a Poisson photon number distribution

$$P(n_i, \beta_i) = e^{-|\beta_i|^2} \frac{|\beta_i|^{2n_i}}{n_i!}. \qquad (19)$$

Therefore, a concatenation of three stochastic processes simulates the photo-counting of a set of interfering thermal states. The first process generates a complex vector $\boldsymbol{\alpha}$ following the probability distribution $p(\boldsymbol{\alpha})$. The second one applies the map $\mathcal{U}$ to the vector $\boldsymbol{\alpha}$ generating the output $\boldsymbol{\beta} = U\boldsymbol{\alpha}$. The third process $\mathcal{P}$ generates an M-dimensional vector $\overline{n}$ from $\boldsymbol{\beta}$ by sampling from the M-dimensional Poisson distribution $p(\overline{n}, \boldsymbol{\beta})$, where averaging over $\boldsymbol{\beta}$ gives

$$p_{\mathrm{T}}(\overline{n}) = \int d\boldsymbol{\beta}\, p(\boldsymbol{\beta}) p(\overline{n}, \boldsymbol{\beta}) \qquad (20)$$

This algorithm is an idealized one, as it assumes access to oracles that sample exactly from Gaussian and Poisson distributions.

In order to build a realistic algorithm we define a new three step process, where the sampling oracles are replaced by efficient approximation algorithms. The first step, detailed in subsection 5.2, consist of sampling from a N-mode constellation $\mathcal{C}_{N,\lambda,m^2}(\boldsymbol{\alpha})$ composed of $m^2 N$ coherent states of distribution $\tilde{p}(\boldsymbol{\alpha})$, satisfying

$$\sum_{i=1}^{m^2 N} \tilde{p}(\boldsymbol{\alpha}_i)|\boldsymbol{\alpha}_i\rangle\langle\boldsymbol{\alpha}_i| = \tilde{\rho}_{\mathrm{T}}, \qquad (21)$$

which efficiently approximates the N-mode thermal state ρ_{T} [27]. The N-mode constellation is composed of N identical single-mode constellations $\mathcal{C}_{\lambda,m^2}(\alpha)$ of size m^2 providing each a good approximation of the single-mode thermal state σ_{th}. The second step $\tilde{\mathcal{U}}$ implements an approximation of matrix multiplication, discussed in subsection 5.3, mapping $\boldsymbol{\alpha}$ to $\boldsymbol{\beta}$, transforming the constellation $\mathcal{C}_{N,\lambda,m^2}(\boldsymbol{\alpha})$ into $\mathcal{C}_{N,\lambda,m^2}(\boldsymbol{\beta})$. Finally the third step $\tilde{\mathcal{P}}$ generates $\tilde{p}(\overline{n}, \boldsymbol{\beta})$, an approximation of an M-dimensional Poisson distributions (19) satisfying

$$\sum_{j}^{m^2 N} \tilde{p}(\overline{n}, \boldsymbol{\beta}_j) = \tilde{p}_T(\overline{n}). \qquad (22)$$

To approximate the Poisson distribution we use a scalable number of Bernouilli trials, as detailed in subsection 5.4 following [5].

5.1 Error analysis

We want to show that the trace distance between the ideal and approximate algorithms, above described, satisfies $D(p_{\mathrm{T}}(\overline{n}), \tilde{p}_{\mathrm{T}}(\overline{n}) \leq \epsilon$ while the algorithm remaining polynomial-time. It is easy to see from the definition of $\mathcal{P}, \tilde{\mathcal{P}}, \mathcal{U}, \tilde{\mathcal{U}}$ that

$$\|p_{\mathrm{T}}(\overline{n}) - \tilde{p}_{\mathrm{T}}(\overline{n})\| = \|\mathcal{P} \circ \mathcal{U}(\rho_{\mathrm{T}}) - \tilde{\mathcal{P}} \circ \tilde{\mathcal{U}}(\tilde{\rho}_{\mathrm{T}})\|. \qquad (23)$$

One can rewrite it as the following norm of a linear combination of terms,

$$\|p_{\mathrm{T}}(\overline{n}) - \tilde{p}_{\mathrm{T}}(\overline{n})\| = \|\mathcal{P} \circ \mathcal{U}(\rho_{\mathrm{T}}) - \mathcal{P} \circ \mathcal{U}(\tilde{\rho}_{\mathrm{T}}) + \mathcal{P} \circ \mathcal{U}(\tilde{\rho}_{\mathrm{T}}) - \mathcal{P} \circ \tilde{\mathcal{U}}(\tilde{\rho}_{\mathrm{T}}) + \mathcal{P} \circ \tilde{\mathcal{U}}(\tilde{\rho}_{\mathrm{T}}) - \tilde{\mathcal{P}} \circ \tilde{\mathcal{U}}(\tilde{\rho}_{\mathrm{T}})\| \qquad (24)$$

which using the triangle inequality and the fact that a trace preserving map, such as $\mathcal{U}, \mathcal{P}, \tilde{\mathcal{U}}, \tilde{\mathcal{P}}$ can only decrease the trace distance, we obtain the upper-bound

$$\|p_{\mathrm{T}}(\overline{n}) - \tilde{p}_{\mathrm{T}}(\overline{n})\| \leq \|\rho_{\mathrm{T}} - \tilde{\rho}_{\mathrm{T}}\| + \|\mathcal{U}(\tilde{\rho}_{\mathrm{T}}) - \tilde{\mathcal{U}}(\tilde{\rho}_{\mathrm{T}})\| + \|\mathcal{P} \circ \tilde{\mathcal{U}}(\tilde{\rho}_{\mathrm{T}}) - \tilde{\mathcal{P}} \circ \tilde{\mathcal{U}}(\tilde{\rho}_{\mathrm{T}})\|, \qquad (25)$$

that can be simplified using the definition of the constellation $\mathcal{C}_{N,\lambda,m}(\boldsymbol{\alpha})$ and $\mathcal{C}_{N,\lambda,m}(\boldsymbol{\beta})$, resulting in

$$\|p_{\mathrm{T}}(\overline{n}) - \tilde{p}_{\mathrm{T}}(\overline{n})\| \leq \|\rho_{\mathrm{T}} - \tilde{\rho}_{\mathrm{T}}\| + \max_{\mathcal{C}(\boldsymbol{\alpha})} \|\mathcal{U}(|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}|) - \tilde{\mathcal{U}}(|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}|)\| + \max_{\mathcal{C}(\boldsymbol{\beta})} \|\mathcal{P}(|\boldsymbol{\beta}\rangle\langle\boldsymbol{\beta}|) - \tilde{\mathcal{P}}(|\boldsymbol{\beta}\rangle\langle\boldsymbol{\beta}|)\|, \quad (26)$$

where $\mathcal{C}(\boldsymbol{\beta}) = \mathcal{U}(\mathcal{C}(\boldsymbol{\alpha}))$. In what follows we will show how one can build efficient algorithm for each step such that the last three terms on the r.h.s. of (26) are bounded by $2\epsilon/3$, leading to $D(p_{\mathrm{T}}(\overline{n}), \tilde{p}_{\mathrm{T}}(\overline{n})) \leq \epsilon$.

5.2 Efficient coherent states constellations

We follow the derivation in [27] to bound the trace norm between a single-mode thermal state σ_{th} and the single-mode coherent state constellation

$$\mathcal{C}_{\lambda,m}(\alpha) = \{P_{\lambda,m}(\alpha), |\alpha\rangle\langle\alpha|\}_{\alpha\in\mathbb{C}}, \quad (27)$$

which needs to have at least the same first and second moments. The constellation is described by the P-distribution $P_{\lambda,m}(\alpha)$ supported on m^2 points, such that

$$\frac{V}{2} P_{\lambda,m}\left[\sqrt{\frac{V}{2}}(x+iy)\right] = P_{X_m}(x) P_{X_m}(y), \quad (28)$$

where $P_{X_m}(x)$ is one constellations of the normal distribution $X \sim N(0,1)$ [57]. In terms of random variables, $\alpha_m = \sqrt{\frac{V}{2}}(X_m + iX'_m)$, where X_m and X'_m are independent realizations of the given constellation. The factor $\sqrt{V}$ ensures that the resulting P function has the same variance

$$V = \frac{\lambda}{1-\lambda} \quad (29)$$

as the one corresponding to the thermal state σ_{th}, where the factor $1/\sqrt{2}$ takes care of the conversion from two real variables to one complex variable. As shown in [50] the χ^2-divergence bounds the trace norm as

$$||\rho - \sigma||^2 \leq \chi^2(\rho, \sigma) = \mathrm{Tr}\left[\left(\rho\sigma^{-1/2}\right)^2\right]. \quad (30)$$

As proven in the Appendix of [27], the quantum χ^2-divergence $\chi^2(\sigma_{\mathrm{th}}, \mathcal{C}_{\lambda,m}(\alpha))$ satisfies the upper-bound

$$1 + \chi^2(\sigma_{\mathrm{th}}, \mathcal{C}_{\lambda,m}(\alpha)) = (1 + \chi^2(P_{X_m}, P_X))^2 \quad (31)$$

The Gauss-Hermite constellation X_m of size m of the normal distribution $X \sim N(0,1)$ is given by the m roots of the mth Hermite polynomial, with weight P_{X_m} selected to provide exact integration with respect to P_{X_m} for all polynomials up to degree $2m - 1$ [57]. Then one can relate the χ^2-divergence between the normal distribution and the Gauss-Hermite constellation $\chi^2(P_{X_m}, P_X)$ to the moments of the Hermite polynomials of X_m [27, 57]

$$1 + \chi^2(P_{X_m}, P_X) = \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{s}{1+s}\right)^k |\mathbb{E}[H_k(X_m)]|^2 \quad (32)$$

where

$$s = \frac{V}{\sqrt{V(V+1)} - V}. \quad (33)$$

Because $\mathbb{E}[H_k(X_m)] = 0$ for odd k and by definition of the Gauss-Hermite quadrature $\mathbb{E}[H_k(X_m)] = 0$ for all $k \leq 2m - 1$, together with using equation (33) and definition (29) to obtain the equality

$$\frac{s}{1+s} = \sqrt{\lambda}, \quad (34)$$

we reach the simplification

$$\chi^2(P_{X_m}, P_X) = \sum_{k\geq m}^{\infty} \frac{\lambda^k}{(2k)!} |\mathbb{E}[H_{2k}(X_m)]|^2. \quad (35)$$

Finally, following the proof of Theorem 8 in [57] we can show

$$\chi^2(P_{X_m}, P_X) \leq 2\kappa^2 \sum_{k\geq m}^{\infty} \lambda^k = 2\kappa^2 \frac{\lambda^m}{1-\lambda}, \quad (36)$$

where κ is a constant such that $2\kappa^2 \approx 2.36$. This provides us with an upper-bound to the trace distance between the single-mode constellation and the single-mode thermal state

$$D(\sigma_{\mathrm{th}}, \mathcal{C}_{\lambda,m}(\alpha)) \leq \frac{1}{2}\left(\left[1 + 2\kappa^2 \frac{\lambda^m}{1-\lambda}\right]^2 - 1\right)$$
$$\leq 3\kappa^2 \frac{\lambda^m}{1-\lambda}. \quad (37)$$

Using Lemma 1 one can trivially obtain an upper-bound between the N-mode constellation and thermal state from its single-mode version

$$D(\rho_{\mathrm{T}}, \mathcal{C}_{N,\lambda,m}(\boldsymbol{\alpha})) \leq 3\kappa^2 N \frac{\lambda^m}{1-\lambda}. \quad (38)$$

5.3 Matrix multiplication

The transformation of $\boldsymbol{\beta} = U\boldsymbol{\alpha}$ can be approximated using standard numerical linear algebra within error ϵ in $O(MN)$ operations [24]. It is

$$||\mathcal{U}\left(|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}|\right) - \tilde{\mathcal{U}}\left(|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}|\right)|| \leq 2\sqrt{1 - F\left(\mathcal{U}\left(|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}|\right), \tilde{\mathcal{U}}\left(|\boldsymbol{\alpha}\rangle\langle\boldsymbol{\alpha}|\right)\right)} \leq 2\sqrt{1 - e^{-|\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}|^2}} \leq 2|\tilde{\boldsymbol{\beta}} - \boldsymbol{\beta}| \leq 2\epsilon, \tag{39}$$

which shows that the error on the coherent sates is upper bounded by the error of the matrix multiplication, which can be made arbitrary small with a polynomial-time overhead.

between the probability distribution of the sum of t independent Bernoulli trials, $S_t = \xi_1 + \ldots + \xi_t$, with $p_B(\xi = 1) = x/t$, $p_B(\xi = 0) = 1 - x/t$ and the Poisson distribution $P(n, x)$. This implies that the error of the M-dimensional output $\overline{n}$ will be bounded as

$$D\left(\mathcal{P}\left(|\boldsymbol{\beta}\rangle\langle\boldsymbol{\beta}|\right), \tilde{\mathcal{P}}\left(|\boldsymbol{\beta}\rangle\langle\boldsymbol{\beta}|\right)\right) \leq \frac{M}{t} \max_{\boldsymbol{\beta}} |\boldsymbol{\beta}|^4. \tag{41}$$

Because the unitary matrix U preserves the norm of a vector, it is easy to show

$$\max_{\boldsymbol{\beta}} |\boldsymbol{\beta}| \leq N \max|\alpha| \leq N\sqrt{2m}, \tag{42}$$

where we used the fact that $\boldsymbol{\alpha}$ results from Nidentical single-mode constellations of coherent states and Krasikov's upper bound on the roots of Hermite polynomials [26], which define the location of the coherent states of the constellation. Therefore, the number of necessary Bernoulli trials t for simulation of a Poisson distribution with a trace distance error $2\epsilon/3$ reads

$$t = 6\frac{MN^2}{\epsilon}m^2. \tag{43}$$

then easy, using the fidelity upper-bound of the trace norm, the well-know exponential decrease of the overlap between two coherent states, and the inequality $1 - e^{-x} \leq x$ to derive the bound

5.4 Approximate Poisson sampling with Bernouilli trials

Starting from an M-dimensional vector $\boldsymbol{\beta}$ the map $\mathcal{P}(\boldsymbol{\beta})$ outputs samples $\overline{n}$ from M Poisson distributions, i.e., for each β_l with $l = 1, \ldots, M$ the Poisson distributions $P(n_l, |\beta_l|^2)$ (19) is sampled via independent Bernoulli trials. To determine the number of Bernoulli trials t to achieve a given error we can use the trace distance bound [5]

$$\frac{1}{2} \sum_{k=0}^{\infty} \left| \binom{t}{k} p_B^k (1 - p_B)^{t-k} - P(k, x) \right| \leq (1 - e^{-x})\frac{x}{t} \leq \frac{x^2}{t} \tag{40}$$

5.5 Algorithm scaling

The algorithm is composed of three steps. First, the random generation of a N-mode coherent state, a task with a computational cost scaling as $O(N)$. Secondly, the matrix multiplication $\boldsymbol{\beta} = U\boldsymbol{\alpha}$, with a computational cost scaling as $O(MN)$. Finally, we need to approximate a Poisson distribution by sampling from t Bernoulli trials, as specified in equation (43). Using equation (38) and that $\lambda = \mu$ we obtain

$$m = \frac{\log N + \log(1/\epsilon) + \log(1/(1-\mu)) + 2\log(3\kappa)}{\log(1/\mu)}. \tag{44}$$

It is easy to see that a constant μ guarantees a computational cost scaling as $O\left(\frac{MN^2}{\epsilon}[\log N + \log(1/\epsilon)]^2\right)$ in terms of the number of input photons N, the transmission M and the quality of the approximation ϵ. Observe that by theorem 2, $\mu \leq \sqrt{\epsilon/N}$, hence the right hand side on equation (44) is regular (for $\epsilon/N \leq 1 - \delta$ for all $\delta > 0$).

6 Proof of Lemma 5

A quantum state of M bosonic modes with at most N total number of photons reads

$$|\psi\rangle = \sum_{\{\overline{n}:|\overline{n}|=N\}} C_{n_1,n_2,...,n_M} |n_1, n_2, ..., n_M\rangle. \quad (45)$$

The memory and computational cost of a brute-force simulation is associated with the number of degrees of freedom of the coefficient $C_{n_1,n_2,...,n_M}$, which correspond to the size of the Hilbert space. In our case it is given by the binomial $\binom{N+M-1}{N}$ that grows exponentially if both N and M increase proportionally to each other. The idea behind tensor networks is to interpret $C_{n_1,n_2,...,n_M}$ as a big tensor with M free indices. A tensor that can be recovered from the contraction of a network of tensors of smaller size, where the virtual degrees of freedom contract each other leaving M free parameters corresponding to the physical indices n_i. This provides a very intuitive representation of quantum states and allows for a very efficient encoding and manipulation when the states have a high degree of locality [35, 36, 53, 56].

6.1 Matrix product states

In our case we are interested in the evolution of a particular example of a tensor network called matrix product states,

$$|\psi\rangle = \sum_{n_1=0}^{d_1} \sum_{n_2=0}^{d_2} \cdots \sum_{n_M=0}^{d_M} B_{n_1}^{[1]} B_{n_2}^{[2]} ... B_{n_M}^{[M]} |n_1, ..., n_M\rangle, \quad (46)$$

where $B_{n_1}^{[1]}$ is a transposed vector of dimension $\tilde{\chi}_1$, $B_{n_M}^{[M]}$ is a vector of dimension $\tilde{\chi}_M$, and $B_{n_i}^{[i]}$ for $1 < i < M$ is a matrix of dimension $\tilde{\chi}_i \times \tilde{\chi}_{i+1}$. The physical indexes n_i take values $\{0, 1, ..., d\}$.

As shown in Figure 3 (a), one can associate to each matrix product state a 1-dimensional graph where each vertex is associated to a three index tensor $B_{n_i,\alpha,\beta}^{[i]}$ (Figure 3 (b)) and the edges determine the contraction rule of the tensor indices.

6.1.1 Canonical form

It is well known that any bipartite quantum state can be rewritten as

$$|\psi\rangle = \sum_{i=0}^{d_1} \sum_{j=0}^{d_2} c_{ij} |ij\rangle = \sum_{\alpha=0}^{\min d_1, d_2} \lambda_\alpha |\varphi_\alpha\rangle |\psi_\alpha\rangle, \quad (47)$$

where the Schmidt coefficients λ_α result from the singular value decomposition $c_{ij} = \sum_\alpha U_{i,\alpha} \lambda_\alpha \overline{V}_{j,\alpha}$. Every matrix product state can be also transformed into a canonical form

$$|\psi\rangle = \sum_{n_1,n_2,...,n_M} \left(\Gamma_{n_1}^{[1]} \lambda^{[1]} \Gamma_{n_2}^{[2]} \lambda^{[2]} \ldots \Gamma_{n_M}^{[N]} \right) |n_1, n_2, \ldots, n_M\rangle \quad (48)$$

by iteratively applying the singular value decomposition [38, 54], with its graph representation shown in Figure 3 (c). The matrices $\lambda^{[i]}$ are diagonal and contain the Schmidt coefficients corresponding to the bipartition of modes $(1, ..., i)$ versus $(i+1, .., M)$. The Schmidt rank of each link reads χ_k, and $\chi = \max_k\{\chi_k\}$ is called the bond dimension. The total number of parameters and thus the storage cost for such a matrix product state scales as $O(M(d+1)\chi^2)$.

6.2 Simulating ideal linear optics circuits

A linear optics circuit is composed of one-mode phase gates and two-mode couplers implementing an interaction between two adjacent modes.

In what follows we first explain how to update a matrix product state that goes under the evolution of linear optics gates and later discuss how to sample from the final output state.

6.2.1 One mode gates (phase rotation)

As shown in Figure 4, a single-mode gate acting on mode i is modeled by a matrix $G_{n_i',n_i}^{[i]}$ that transform the input physical indices n_i to the output physical indices n_i'. The evolution corresponds to the contraction of the physical indices n_i of $\Gamma_{\alpha,n_i,\beta}^{[i]}$ and $G_{n_i',n_i}^{[i]}$ as

$$\tilde{\Gamma}_{\alpha,\beta,n_i'}^{'i} = \sum_{n_i} G_{(n_i'),(n_i)}^{[i]} \Gamma_{(n_i),(\alpha,\beta)}^{[i]}. \quad (49)$$
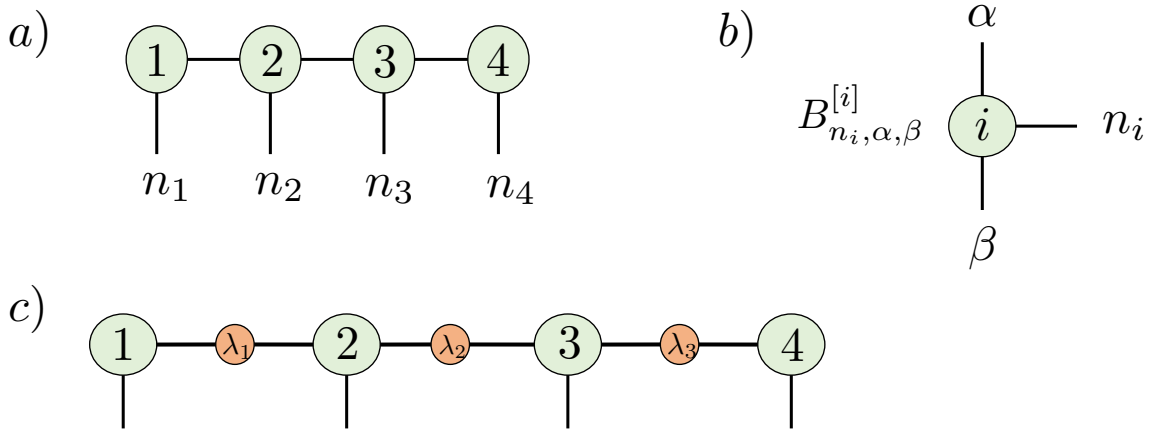
Figure 3: (a) Graphical representation of a quantum state $|\psi\rangle$ corresponding to a matrix product state of four tensors, as defined in Eq. (46). (b) $B^{[i]}_{n_i,\alpha,\beta}$ is a tensor of rank three: represented by a vertex i with three edges, one corresponding to the physical index n_i and the two other to the virtual indexes α, and β. (c) All matrix product states can be transformed into a canonical form where a diagonal matrix λ of Schmidt coefficients is assigned to every edge between two vertices of (a).
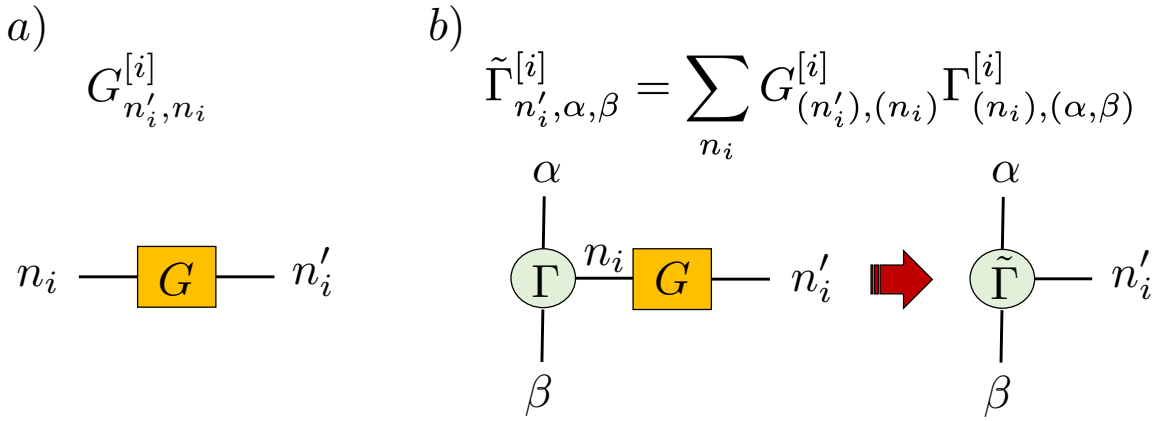
Figure 4: a)A single-mode phase-rotation acting on mode i is modeled by a matrix (tensor) $G^{[i]}_{n'_i,n_i}$ that transform the input physical indexes n_i to the output physical indexes n'_i; b) The tensor $\Gamma^{[i]}$ of virtual indexes α,β and physical index n_i is transformed to the tensor $\tilde{\Gamma}^{[i]}$ by implementing a tensor contraction between the tensor $\Gamma^{[i]}$ and the phase-rotation $G^{[i]}$

.

A phase rotation θ has a matrix $G^{[i]}$ that is diagonal with coefficients $G^{[i]} = \exp(i\theta n_i)$. Therefore, the computational cost of the update of a single local gate scales as $O\left((d+1)\chi^2\right)$. Notice that applying a single-mode gate will not change the Schmidt coefficients of the matrix product state, as it acts only on the physical indexes of one vertex of the graph.

6.2.2 Two-mode couplers

A two-mode coupler $B^{[k,k+1]}$ acting on modes k and $k+1$ is modeled by a 4 legs tensor, i.e., a matrix product operator, with physical indexes n_k and n_{k+1} for the input and n'_k and n'_{k+1} for the output,

$$B^{[k,k+1]} = \sum_{n_k,n_{k+1},n'_k,n'_{k+1}} C^{n'_k,n'_{k+1}}_{n_k,n_{k+1}} |n'_k,n'_{k+1}\rangle\langle n_k,n_{k+1}|, \tag{50}$$

where the coefficients $C^{n'_k,n'_{k+1}}_{n_k,n_{k+1}}$ are the well-known input-output amplitudes of a beamsplitter [9, 23]
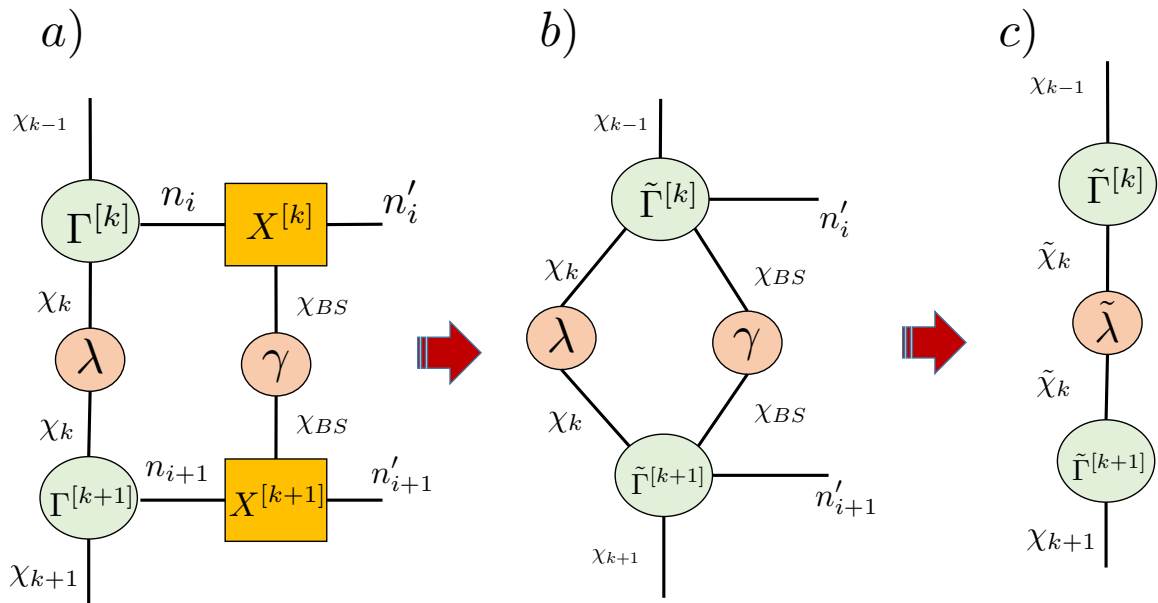
Figure 5: a) The action of a two-mode coupler on modes k and $k+1$ of a matrix product state written on its canonical form is first obtained by doing a singular-value decomposition of the matrix product operator of the coupler; b) Secondly, we contract the tensors $\Gamma^{[k]}$ and $X^{[k]}$ of mode k and $\Gamma^{[k+1]}$ and $X^{[k+1]}$ of mode $k+1$, giving $\tilde{\Gamma}^{[k]}$ and $\tilde{\Gamma}^{[k+1]}$ respectively; c) Finally, we relabel the two singular values λ and γ into a new label $\tilde{\lambda}$ of the resulting matrix product state.

(see also equation (3.9) in [2]).

In [49] an algorithm was constructed based on directly applying the unitary $B^{[k,k+1]}$ to the matrix product state followed by a singular-value decomposition to rebuild the canonical form of the output state, reaching a computational cost scaling as $\mathrm{O}(\chi^3 d^3)$. In what follows we present an alternative algorithm that provides a better scaling when the bond dimension χ is higher than the physical dimension d, which is generally the case in most realistic simulations.

As shown in Figure 5 a), in order to model the evolution of modes k and $k+1$ under a two-mode coupler operation, we first implement a singular-value decomposition of the matrix product operator $B^{[k,k+1]}$ with respect to the separation between (n_k, n_k') and (n_{k+1}, n_{k+1}') indexes, i.e.,

$$B^{[k,k+1]}_{(n_k,n_k'),(n_{k+1},n_{k+1}')} = \sum_{\gamma=1}^{\chi_{BS}} X^{[k]}_{n_k,n_k',\gamma} \sigma^{[k]}_\gamma X^{[k+1]}_{n_{k+1},n_{k+1}',\gamma} \tag{51}$$

where χ_{BS} is the Schmidt rank of the matrix product operator. The Schmidt rank of a singular-value decomposition of a matrix being upper-bounded by the largest of the two local dimensions provides the bound

$$\chi_{BS} \leq (d+1)^2, \tag{52}$$

where the running time of the matrix product operator decomposition scales as $\mathrm{O}\left((d+1)^6\right)$.

As shown in Figure 5 b), the next step is to contract the tensors $\Gamma^{[k]}$ and $X^{[k]}$ of mode k and $\Gamma^{[k+1]}$ and $X^{[k+1]}$ of mode $k+1$ in order to generate the tensors $\tilde{\Gamma}^{[k]}$ and $\tilde{\Gamma}^{[k+1]}$ of the state resulting after the beamsplitter transformation.

The running time of the contraction leading to the tensor $\tilde{\Gamma}^{[k]}$ scales as $\chi_{k-1}\chi_k\chi_{BS}(d+1)^2$ where for $\tilde{\Gamma}^{[k+1]}$ scales as $\chi_{k+1}\chi_k\chi_{BS}(d+1)^2$ which leads to a scaling of the contraction running time

$$T_C = \mathrm{O}\left(\chi^2\chi_{BS}(d+1)^2\right) \leq \mathrm{O}\left(\chi^2(d+1)^4\right). \tag{53}$$

We remark that, as shown in Figure 5 c), the tensors $\tilde{\Gamma}^{[k]}$ and $\tilde{\Gamma}^{[k+1]}$ are connected by two pairs of singular values, χ_k from the initial state and χ_{BS} from the beamsplitter matrix product operator, which can be merged into a single $\tilde{\chi}$ satisfying $\tilde{\chi}_k = \chi_k\chi_{BS}$, which combined with equation (52) provides the bound

$$\tilde{\chi} \leq \chi(d+1)^2, \tag{54}$$

which is the equivalent of Lemma 4 (i) in [22].

6.2.3 Circuit simulation

The ideal boson sampling input state corresponds to a trivial matrix product state of

bond dimension $\chi = 1$, composed of N tensors $\Gamma_1^{[i]} = \delta_{n_i,1}\delta_{\alpha_{i-1},0}\delta_{\alpha_i,0}$ encoding single-photon inputs and $M - N$ tensors $\Gamma_0^{[i]} = \delta_{n_i,0}\delta_{\alpha_{i-1},0}\delta_{\alpha_i,0}$ encoding vacuum inputs. For every layer of couplers we apply in parallel the matrix product update detailed in subsections 6.2.1 and 6.2.2. The bond dimension scales with the depth of the circuit D as $\mathrm{O}\left((d+1)^{2D}\right)$, the storage space for the tensors as $\mathrm{O}\left(M(d+1)^{4D+1}\right)$, and the computational cost of the contraction of the matrix product state scales as $\mathrm{O}\left(M(d+1)^{4(D+1)}\right)$, where the leading order corresponds to the contractions of the last layer of gates.

6.2.4 Sampling from a matrix product state

Once the matrix product state resulting from D layers of gates has been calculated, it is well known that one can exploit the chain rule

$$p(n_1, \ldots, n_M) = p(n_M | n_{M-1}, \ldots, n_1) \ldots p(n_1) \tag{55}$$

to generates samples of $p(\overline{n})$. For completeness, we reproduce the explanation in [29] bellow.

First calculate for each of the $d+1$ outcomes n_1 the probability $\mathrm{Tr}\left[|\psi\rangle\langle\psi||n_1\rangle\langle n_1| \otimes I_{2\ldots M}\right]$, where $|n_1\rangle\langle n_1|$ is the projector onto the photon number state n_1 of mode 1 and $I_{2\ldots M}$ is the identity operator on modes 2 to M. This is done by contraction of the matrix product state with itself, interleaved with a matrix product operator representing the measurement projector $|n_1\rangle\langle n_1|$. Then we randomly select one of the $d+1$ potential outcomes n_1 and update our state by generating $|\psi_{\tilde{n}_1}\rangle := \langle\tilde{n}_1|\psi\rangle$, where the bra $\langle\tilde{n}_1|$ acts only on mode 1. The result of this contraction is a new, unnormalized matrix product state $|\psi_{\tilde{n}_1}\rangle$ of size $N-1$. Note that this new matrix product state satisfies the condition $\langle\psi_{\tilde{n}_1}|\psi_{\tilde{n}_1}\rangle = p(\tilde{n}_1)$. The second step now uses the state $|\psi_{\tilde{n}_1}\rangle$. Firstly, we calculate the $d+1$ outcome probabilities $p(n_2, \tilde{n}_1) := \langle\psi_{\tilde{n}_1}|\left(|n_2\rangle\langle n_2| \otimes \mathcal{I}_{3,\ldots,N}\right)|\psi_{\tilde{n}_1}\rangle$ and randomly select a $\tilde{n}_2$ from the probability distribution $p(n_2|\tilde{n}_1) := p(n_2, \tilde{n}_1)/p(\tilde{n}_1)$. Secondly, we generate a new, unnormalized matrix product state $|\psi_{\tilde{n}_1,\tilde{n}_2}\rangle := \langle\tilde{n}_2|\psi_{\tilde{n}_1}\rangle$ of size $N-2$. Continuing this procedure for the remaining $M-3$ output modes, we end up with one sample drawn according to the probability distribution $p(n_1, n_2, \ldots, n_N)$. The highest computational cost corresponds to the con-

traction leading to $p(n_1, n_2, \ldots, n_N)$. A trivial contraction algorithm provides a running time of $\mathrm{O}\left(M\chi^4(d+1)^2\right)$, which for matrix product state resulting from D layers of couplers becomes $\mathrm{O}(M^2(d+1)^{8D+2})$.

6.3 Simulating ideal logarithmic depth circuits

It is important to notice that the bond dimension scales exponentially with the depth of the circuit. If d was a constant, such as in spin systems simulations, a shallow circuit satisfying a logarithmic depth constraint as in eq. (11) would lead to a polynomial time algorithm. In a tensor network simulation of quantum optics, the potential bunching of photons, which can all potentially accumulate in a given mode, makes the simulation harder. In order to obtain an exact simulation of the evolution and the sampling of N input single photons over a circuit of depth D we fixed the physical dimension over the whole evolution to $d = N$, the total number of photons. Because N scales with the number of modes M, see eq. (2), the computational cost of contraction, storage and sampling becomes quasipolynomial in the size of the system.

7 Generalization to non-uniform losses

In subsection 2.2 we have shown that a lossy linear optics interferometer is mathematically modeled by a complex linear transformation A with singular value decomposition $A = V\hat{\mu}W$. As shown in Figure 2, this is equivalent to a circuit composed by a lossless linear optics transformation V, followed by M parallel set of different pure-loss channels of transmission μ_i, and a final lossless linear optics transformation W.

As shown in Figure 6, each of the M parallel pure-loss channels of transmission μ_i can be decomposed into a concatenation a pure-loss channels of transmission $\mu = \max \mu_i$ followed by one of transmission $\tilde{\mu}_i = \mu_i/\mu$. Because M parallel pure-loss channels of transmission μ commute with the unitary V, we obtain a scheme where M parallel set of pure-loss channels of transmission μ are followed by the ideal circuit V, M parallel different pure-loss channels ($\tilde{\mu}_i$) and a final ideal circuit W. The state ρ_{in} now results from applying M parallel set of pure-loss channels of transmission $\mu = \max_i \mu_i$ and the thermal state
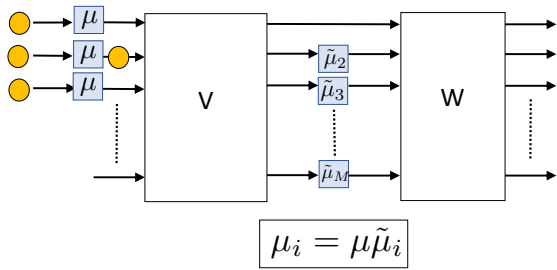
Figure 6: A non-uniform lossy interferometer is equivalent to M parallel set of pure-loss channels of transmission $\mu = \max \mu_i$ (where we have chosen $\mu = \mu_1$ without loss of generality) followed by the losseless circuit V, M parallel set of pure-loss channels ($\tilde{\mu}_i$) and a final losseless circuit W.

ρ_{T} is chosen such that $\lambda = \mu$, in the same way we did in subsection 3. The only difference is that now the set of operations after the pure-loss channels of transmission μ is not longer a lossless interferometer U but a M-mode quantum channel $\mathcal{L}$ resulting from concatenating V, M parallel pure-loss channels ($\tilde{\mu}_i$) and W.

7.1 Generalizing Theorem 2

Because applying a quantum operation $\mathcal{L}$ can only make the trace-distance decrease, similarly as for a measurement, it is trivial to see that one can generalize Theorem 2 by replacing the uniform losses μ by $\mu = \max \mu_i$.

7.2 Generalizing lemma 3

It is a well-known fact in quantum optics that the action of a pure loss channel of transmission μ_i on a coherent state $|\alpha\rangle$ outputs a weaker coherent state $|\mu_i \alpha\rangle$. Therefore, the evolution of an input multimode coherent state $|\boldsymbol{\alpha}\rangle$ can be easily computed, by implementing the matrix multiplication $\boldsymbol{\beta} = A\boldsymbol{\alpha}$. Once the output coherent states have been determined, the sampling from their respective Poisson distributions proceeds as before.

7.3 Generalizing lemma 5

The adaptation of the tensor network simulation is slightly more involved. Let's use the notation $A_{i,j}$ for the coupler acting on modes i and $i{+}1$ at the layer of couplers j. Every $A_{i,j}$ has a decomposition into a unitary $V_{i,j}$ followed by two independent pure-loss channels and a final unitary $W_{i,j}$.

Because every pure-loss channel can be seen as a lossless coupling interaction with an environmental mode, it is easy to see that a circuit with losses can be transformed into an ideal lossless circuit by doubling the number of couplers and adding two ancillary modes per coupler with losses. We can then place all the ancillary modes interacting with mode i between input modes i and $i{+}1$, i.e., D of them bellow each input mode for a circuit of depth D. For a lossy circuit of depth D there is at most $3D$ lossless gates acting on each mode with a range of at most D. As detailed in [22], one can transform a D range gate into $2D$ nearest-neighbor gates. Therefore, our initial circuit with losses of depth $D = M$ becomes a lossless circuit with M^2 modes and $6M^2$ nearest-neighbor gates. This leads to a less favorable scaling of the computational cost of contraction, storage and sampling, but which remains quasipolynomial in M. This last algorithm is certainly not optimal and we are convinced that more elaborate choices can certainly improve the simulation of multi-photon interference with non-uniform losses.

8 Conclusion

The vast majority of currently proposed boson sampling architectures suffer from exponential decay of the transmission with the length of the circuit. We have shown that multi-photon interference over an M modes interferometer of depth D can be efficiently simulated classically. More precisely, we have show that either the depth of the circuit is large enough ($D \geq \mathrm{O}(\log M)$) that it can be simulated by thermal noise with an algorithm running in polynomial time, or the depth of the circuit is shallow enough ($D \leq \mathrm{O}(\log M)$) that a tensor network simulation runs in quasipolynomial time.

We also showed that for even very optimistic experimental parameters, a quantum supremacy experiment using the current boson sampling hardness proof is out of reach. We believe that our result suggests that in order to implement a quantum advantage experiment with single-photons and linear optics we need novel theoretical ideas or radically new technological developments. One possibility would be to shift to platforms with very low algebraic transmission where our result would not be applicable. Another option would be prove the hardness of novel boson

sampling architectures beyond the planar circuit architecture. A promising route would be to reduce the depth of the circuit to the shallow regime while maintaining the complexity by moving to a lattice structure. A potential candidate would be an adaptation to quantum optics of the recent proposal of quenching a spin lattice [6].

We discussed that the potential bunching of photons makes the tensor network simulation of quantum optical systems more involved than its finite spin counterparts. One could potentially restore the polynomial scaling of logarithmic depth circuits observed for finite systems by designing an ϵ-approximate algorithm that truncates every mode to a finite size. To our knowledge, this is a non-trivial result that is certainly worth pursuing in future research.

One of the motivations of our work was to show that simulating boson sampling with imperfections is indeed easier than ideal boson sampling. We achieved this goal for a restricted regime of losses where the system becomes classically simulatable. We conjecture that there should exist a family of algorithms that optimally interpolate between the Clifford and Clifford algorithm for ideal devices and fully efficient algorithms for noisy devices. This work is only a first step in this direction and we believe that further results will improve even further the classical simulation of imperfect boson sampling devices.

An interesting open question is whether our proof can be adapted to other technological platforms candidates to a quantum advantage test.

It has been recently proposed to use postselection to improve boson sampling rates while maintaining its hardness [55]. The proposal, inspired by the fact that the hardness of boson sampling is preserved when only a constant number of photons is lost [3], proposes to postselect outcome events with a constant number of lost photons. If the probability of a successful post-selection decreases exponentially with the increasing size of the system, it was argued that it could be a useful tool for experiment of the limited size needed for a practical quantum advantage demonstrations. If the proof presented in this manuscript can not be used to discard this novel approach to boson sampling, a recent work by the same authors [44] imposes strong constrains on its viability.

After the completion of this article we learned about Ref. [37] that obtains a similar result to our Theorem 2 and its generalization to non-uniform losses using very different techniques.

9 Acknowledgements

References

[1] S. Aaronson. A linear-optical proof that the permanent is #P-hard. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 467:3393–3405, 2011. DOI: 10.1098/rspa.2011.0232.

[2] S. Aaronson and A. Arkhipov. The computational complexity of linear optics. *Theory of Computing*, 9:143–252, 2013. DOI: 10.4086/toc.2013.v009a004.

[3] S. Aaronson and D. J. Brod. Boson sampling with lost photons. *Phys. Rev. A*, 93:012335, 2016. DOI: 10.1103/PhysRevA.93.012335.

[4] A. Arkhipov and G. Kuperberg. The bosonic birthday paradox. *Geometry & Topology Monographs*, 18:1, 2012. DOI: 10.2140/gtm.2012.18.1.

[5] A. D. Barbour and P. Hall. On the rate of poisson convergence. *Mathematical Proceedings of the Cambridge Philosophical Society*, 95:473–480, 1984. DOI: 10.1017/S0305004100061806.

[6] J. Bermejo-Vega, D. Hangleiter, M. Schwarz, R. Raussendorf, and J. Eisert. Architectures for quantum simulation showing a quantum speedup. *Phys. Rev. X*, 8:021010, 2018. DOI: 10.1103/PhysRevX.8.021010.

[7] M. A. Broome, A. Fedrizzi, S. Rahimi-Keshari, J. Dove, S. Aaronson, T. C. Ralph,

and A. G. White. Photonic boson sampling in a tunable circuit. *Science*, 339:794–798, 2013. DOI: 10.1126/science.1231440.

[8] E. R. Caianiello. On quantum field theory i: explicit solution of dyson's equation in electrodynamics without use of feynman graphs. *Il Nuovo Cimento*, 10:1634–1652, 1953. DOI: 10.1007/BF02781659.

[9] R. A. Campos, B. E. A. Saleh, and M. C. Teich. Quantum-mechanical lossless beam splitter: SU(2) symmetry and photon statistics. *Phys. Rev. A*, 40:1371–1384, 1989. DOI: 10.1103/PhysRevA.40.1371.

[10] J. Carolan, C. Harrold, C. Sparrow, E. Martín-López, N. J. Russell, J. W. Silverstone, P. J. Shadbolt, N. Matsuda, M. Oguma, M. Itoh, G. D. Marshall, M. G. Thompson, J. C. F. Matthews, T. Hashimoto, J. L. O'Brien, and A. Laing. Universal linear optics. *Science*, 349:711–716, 2015. DOI: 10.1126/science.aab3642.

[11] L. Chakhmakhchyan and N. J. Cerf. Boson sampling with gaussian measurements. *Phys. Rev. A*, 96:032326, 2017. DOI: 10.1103/PhysRevA.96.032326.

[12] W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley. Optimal design for universal multiport interferometers. *Optica*, 3:1460–1465, 2016. DOI: 10.1364/OPTICA.3.001460.

[13] P. Clifford and R. Clifford. The classical complexity of boson sampling. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '18, pages 146–155, Philadelphia, PA, USA, 2018. Society for Industrial and Applied Mathematics. ISBN 978-1-6119-7503-1. URL http://dl.acm.org/citation.cfm?id=3174304.3175276.

[14] A. Crespi, R. Osellame, R. Ramponi, D. J. Brod, E. F. Galvão, N. Spagnolo, C. Vitelli, E. Maiorino, P. Mataloni, and F. Sciarrino. Experimental boson sampling. *Nat. Photonics*, 7:540, 2013. DOI: 10.1038/nphoton.2013.112.

[15] H. Defienne, M. Barbieri, I. A. Walmsley, B. J. Smith, and S. Gigan. Two-photon quantum walk in a multimode fiber. *Science Advances*, 2, 2016. DOI: 10.1126/sciadv.1501054.

[16] R. Laflamme E. Knill and G. J. Millburn.

A scheme for efficient quantum computation with linear optics. *Nature*, 409:46, 2001. DOI: 10.1038/35051009.

[17] V. Giovannetti, A. S. Holevo, and R. Garcia-Patron. A solution of gaussian optimizer conjecture for quantum channels. *Communications in Mathematical Physics*, 334:1553–1571, 2015. DOI: 10.1007/s00220-014-2150-6.

[18] R. J. Glauber. *Quantum Theory of Optical Coherence: Selected Papers and Lectures*. Wiley-VCH Verlag, Weinheim, Germany, 10th edition, 2007. ISBN 9783527610075. DOI: 10.1002/9783527610075.

[19] N. C. Harris, D. Bunandar, M. Pant, G. R. Steinbrecher J. Mower, M. Prabhu, T. Baehr-Jones, M. Hochberg, and D. Englund. Large-scale quantum photonic circuits in silicon. *Nanophotonics*, 5:456, 2016. DOI: 10.1515/nanoph-2015-0146.

[20] Y. He, X. Ding, Z.-E. Su, H.-L. Huang, J. Qin, C. Wang, S. Unsleber, C. Chen, H. Wang, Y.-M. He, X.-L. Wang, W.-J. Zhang, S.-J. Chen, C. Schneider, M. Kamp, L.-X. You, Z. Wang, S. Höfling, Chao-Yang Lu, and J.-W. Pan. Time-bin-encoded boson sampling with a single-photon device. *Phys. Rev. Lett.*, 118:190501, 2017. DOI: 10.1103/PhysRevLett.118.190501.

[21] M. J. R. Heck, J. F. Bauters, M. L. Davenport, D. T. Spencer, and J. E. Bowers. Ultra-low loss waveguide platform and its integration with silicon photonics. *Laser & Photonics Reviews*, 8:667–686, 2014. DOI: 10.1002/lpor.201300183.

[22] R. Jozsa. On the simulation of quantum circuits. *arXiv*, page 0603163, 2006. URL https://arxiv.org/abs/quant-ph/0603163.

[23] M. S. Kim, W. Son, V. Bužek, and P. L. Knight. Entanglement by a beam splitter: Nonclassicality as a prerequisite for entanglement. *Phys. Rev. A*, 65:032323, 2002. DOI: 10.1103/PhysRevA.65.032323.

[24] D. E. Knuth. *The Art of Computer Programming, Volume 1 (3rd Ed.): Fundamental Algorithms*. Addison Wesley Longman Publishing Co., Inc., Redwood City, CA, USA, 1997. ISBN 0-201-89683-4.

[25] P. Kok, W. J. Munro, K. Nemoto, T. C. Ralph, J. P. Dowling, and G. J. Mil-

burn. Linear optical quantum computing with photonic qubits. *Rev. Mod. Phys.*, 79:135–174, 2007. DOI: 10.1103/RevModPhys.79.135.

[26] I. Krasikov. Nonnegative quadratic forms and bounds on orthogonal polynomials. *J. Approx. Theory*, 111:31–49, 2001. DOI: 10.1006/jath.2001.3570.

[27] F. Lacerda, J. M. Renes, and V. B. Scholz. Coherent-state constellations and polar codes for thermal gaussian channels. *Phys. Rev. A*, 95:062343, 2017. DOI: 10.1103/PhysRevA.95.062343.

[28] L.G.Valiant. The complexity of computing the permanent. *Theoretical Computer Science*, 8:189–201, 1979. DOI: 10.1016/0304-3975(79)90044-6.

[29] M. Lubasch, A. Valido, J. Renema, W. S. Kolthammer, D. Jaksch, M. S. Kim, I. Walmsley, and R. García-Patrón. Tensor network states in time-bin quantum optics. *Phys. Rev. A*, 97:062304, 2018. DOI: 10.1103/PhysRevA.97.062304.

[30] A. P. Lund, A. Laing, S. Rahimi-Keshari, T. Rudolph, J. L. O'Brien, and T. C. Ralph. Boson sampling from a gaussian state. *Phys. Rev. Lett.*, 113:100502, 2014. DOI: 10.1103/PhysRevLett.113.100502.

[31] A. P. Lund, S. Rahimi-Keshari, and T. C. Ralph. Exact boson sampling using gaussian continuous-variable measurements. *Phys. Rev. A*, 96:022301, 2017. DOI: 10.1103/PhysRevA.96.022301.

[32] A. Mari and J. Eisert. Positive wigner functions render classical simulation of quantum computation efficient. *Phys. Rev. Lett.*, 109:230503, 2012. DOI: 10.1103/PhysRevLett.109.230503.

[33] A. Neville, C. Sparrow, R. Clifford, E. Johnston, P. M. Birchall, A. Montanaro, and A. Laing. Experimental boson sampling. *Nature Physics*, 13:1153, 2017. DOI: 10.1038/nphys4270.

[34] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition.* Cambridge University Press, New York, NY, USA, 10th edition, 2011. ISBN 1107002176, 9781107002173. DOI: 10.1017/CBO9780511976667.

[35] R. Orús. A practical introduction to tensor networks: Matrix product states and projected entangled pair states. *Annals of Physics*, 349:117 – 158, 2014. DOI: 10.1016/j.aop.2014.06.013.

[36] S. Ostlund and S. Rommer. Thermodynamic limit of density matrix renormalization. *Phys. Rev. Lett.*, 75:3537–3540, 1995. DOI: 10.1103/PhysRevLett.75.3537.

[37] Michał Oszmaniec and Daniel J Brod. Classical simulation of photonic linear optics with lost particles. *New Journal of Physics*, 20:092002, 2018. DOI: 10.1088/1367-2630/aadfa8.

[38] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac. Matrix product state representations. *Quantum Info. Comput.*, 7:401–430, 2007.

[39] R. Prevedel, P. Walther, F. Tiefenbacher, P. Böhi, R. Kaltenbaek, T. Jennewein, and A. Zeilinger. High-speed linear optics quantum computing using active feed-forward. *Nature*, 445:65–69, 2007. DOI: 10.1038/nature05346.

[40] S. Rahimi-Keshari, M. A. Broome, R. Fickler, A. Fedrizzi, T. C. Ralph, and A. G. White. Direct characterization of linear-optical networks. *Opt. Express*, 21:13450–13458, 2013. DOI: 10.1364/OE.21.013450.

[41] S. Rahimi-Keshari, A. P. Lund, and T. C. Ralph. What can quantum optics say about computational complexity theory? *Phys. Rev. Lett.*, 114:060501, 2015. DOI: 10.1103/PhysRevLett.114.060501.

[42] S. Rahimi-Keshari, T. C. Ralph, and C. M. Caves. Sufficient conditions for efficient classical simulation of quantum optics. *Phys. Rev. X*, 6:021039, 2016. DOI: 10.1103/PhysRevX.6.021039.

[43] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani. Experimental realization of any discrete unitary operator. *Phys. Rev. Lett.*, 73:58–61, 1994. DOI: 10.1103/PhysRevLett.73.58.

[44] J. Renema, V. Shchesnovich, and R. Garcia-Patron. Classical simulability of noisy boson sampling. *arXiv*, page 809.01953, 2018. URL https://arxiv.org/abs/1809.01953.

[45] J. J. Renema, A. Menssen, W. R. Clements, G. Triginer, W. S. Kolthammer, and I. A. Walmsley. Efficient classical algorithm for boson sampling with partially dis-

tinguishable photons. *Phys. Rev. Lett.*, 120:220502, 2018. DOI: 10.1103/PhysRevLett.120.220502.

[46] T. Rudolph. Why I am optimistic about the silicon-photonic route to quantum computing. *APL Photonics*, 2:030901, 2017. DOI: 10.1063/1.4976737.

[47] J. B. Spring, B. J. Metcalf, P. C. Humphreys, W. S. Kolthammer, X.-M. Jin, Ma. Barbieri, A. Datta, N. Thomas-Peter, N. K. Langford, D. Kundys, J. C. Gates, B. J. Smith, P. G. R. Smith, and I. A. Walmsley. Boson sampling on a photonic chip. *Science*, 339:798–801, 2013. DOI: 10.1126/science.1231692.

[48] C. Taballione, T. A. W. Wolterink, J. Lugani, A. Eckstein, B. A. Bell, R. Grootjans, I. Visscher, J. J. Renema, D. Geskus, C. G. H. Roeloffzen, I. A. Walmsley, P. W. H. Pinkse, and K.-J. Boller. 8x8 programmable quantum photonic processor based on silicon nitride waveguides. In *Frontiers in Optics / Laser Science*. Optical Society of America, 2018. DOI: 10.1364/FIO.2018.JTu3A.58.

[49] K. Temme and P. Wocjan. Efficient computation of the permanent of block factorizable matrices. *arXiv*, page 1208.6589, 2012. URL https://arxiv.org/abs/1208.6589.

[50] K. Temme, M. J. Kastoryano, M. B. Ruskai, M. M. Wolf, and F. Verstraete. The χ^2-divergence and mixing times of quantum markov processes. *Journal of Mathematical Physics*, 51:122201, 2010. DOI: 10.1063/1.3511335.

[51] M. Tillmann, B. Dakić, R. Heilmann, S. Nolte, A. Szameit, and P. Walther. Experimental boson sampling. *Nat. Photonics*, 7:540, 2013. DOI: 10.1038/nphoton.2013.102.

[52] V. Veitch, C. Ferrie, D. Gross, and J. Emerson. Negative quasi-probability as a resource for quantum computation. *New Journal of Physics*, 14:113011, 2012. DOI: 10.1088/1367-2630/14/11/113011.

[53] F. Verstraete, D. Porras, and J. I. Cirac. Density matrix renormalization group and periodic boundary conditions: A quantum information perspective. *Phys. Rev. Lett.*, 93:227205, 2004. DOI: 10.1103/PhysRevLett.93.227205.

[54] G. Vidal. Efficient classical simulation of slightly entangled quantum computations. *Phys. Rev. Lett.*, 91:147902, 2003. DOI: 10.1103/PhysRevLett.91.147902.

[55] H. Wang, W. Li, X. Jiang, Y.-M. He, Y.-H. Li, X. Ding, M.-C. Chen, J. Qin, C.-Z. Peng, C. Schneider, M. Kamp, W.-J. Zhang, H. Li, L.-X. You, Z. Wang, J. P. Dowling, S. Höfling, C.-Y. Lu, and J.-W. Pan. Toward scalable boson sampling with photon loss. *Phys. Rev. Lett.*, 120:230502, 2018. DOI: 10.1103/PhysRevLett.120.230502.

[56] S. R. White. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, 69:2863–2866, 1992. DOI: 10.1103/PhysRevLett.69.2863.

[57] Y. Wu and S. Verdú. The impact of constellation cardinality on gaussian channel capacity. In *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 620–628, 2010. DOI: 10.1109/ALLERTON.2010.5706965.