Efficient validation of Boson Sampling from binned photon-number distributions

Benoit Seron[1], Leonardo Novo[1,2], Alex Arkhipov[3], and Nicolas J. Cerf[1]

[1]Quantum Information and Communication, Ecole polytechnique de Bruxelles, CP 165/59, Université libre de Bruxelles (ULB), 1050 Brussels, Belgium

[2]International Iberian Nanotechnology Laboratory (INL), Av. Mestre José Veiga, 4715-330 Braga, Portugal

[3]Independent researcher

In order to substantiate claims of quantum computational advantage, it is crucial to develop efficient methods for validating the experimental data. We propose a test of the correct functioning of a boson sampler with single-photon inputs that is based on how photons distribute among partitions of the output modes. Our method is versatile and encompasses previous validation tests based on bunching phenomena, marginal distributions, and even some suppression laws. We show via theoretical arguments and numerical simulations that binned-mode photon number distributions can be used in practical scenarios to efficiently distinguish ideal boson samplers from those affected by realistic imperfections, especially partial distinguishability of the photons.

1 Introduction

An important milestone in the field of quantum computing is the construction of a quantum device that can surpass even the most advanced classical super-computers at a specific task [1, 2]. For the purposes of demonstrating quantum computational advantage with near-term quantum devices, one of the problems that has been intensely investigated is that of Boson sampling. In their seminal paper [3], Aaronson and Arkhipov presented strong complexity theoretic arguments showing that the task of sampling the output of a linear interferometry process involving many single photons is likely to be intractable for classical computers. Boson sampling sparked a great interest over the last decade and various alternative schemes were constructed, such as Scattershot boson sampling, Gaussian Boson sampling and others, in order to facilitate experimental implementations [4, 5, 6, 7]. These efforts culminated in multiple claims of quantum computational advantage with Gaussian Boson sampling [8, 9, 10], while standard boson sampling saw experimental implementations with $n = 20$ photons in $m = 60$ modes [11]. Other experimental platforms than photonics were also considered [12].

Crucially, experimental implementations are unavoidably subject to different noise sources, such as those induced by partial distinguishability or particle loss, which may compromise claims of quantum computational advantage. Indeed, if the amount of noise is too large, then classical algorithms can sample from the outcome distribution efficiently [13, 14, 15, 16, 17, 18, 19]. Therefore, a thin line exists between the regime of classical computational hardness and efficient classical simulability.

It is therefore of highest importance to develop efficient methods to discriminate an ideal boson sampler from a noisy one. However, the very formulation of the task at hand, which involves sampling from an exponentially large set of possibilities, makes the problem of verifying that the device is working properly highly nontrivial. Ideally, one would like to assert that the experiment is generating samples from a distribution that is close enough to the ideal one simply by post-processing the classical data generated by the experiment. However, due to the flatness of the boson sampling distribution, this requires exponentially many samples [20]. Efficient verification schemes that guarantee closeness to the

Benoit Seron: benoitseron@gmail.com

Leonardo Novo: leonardo.novo@inl.int

Alex Arkhipov: arkhipov@alum.mit.edu

Nicolas J. Cerf: ncerf@ulb.ac.be

ideal distribution exist, but they require an active control over the experiment via the ability to do Gaussian measurements on the ouput states [21]. Upon reasonable physical assumptions about the nature of noise [22, 23, 24, 25, 26, 27, 28, 29], and forgetting about adversarial scenarios, we can restrict ourselves to the easier task of validation. It consists of verifying that the experiment passes some easy-to-check tests, which a boson sampler working in the quantum supremacy regime is expected to pass. These tests should be sufficiently sensitive to noise, so we can efficiently discriminate an ideal boson sampler from a noisy one. This is the main question we consider in this paper.

In our work we propose a simple validation scheme of boson sampling based on jointly counting the number of photons in binned-together output modes. We give numerical evidence that these binned distributions are sensitive to typical sources of noise such as photon distinguishability, even for a small number of bins. At the same time, for a fixed choice of binning with a constant number of bins, we show that there is an efficient classical algorithm that approximates the corresponding binned distribution coming from an ideal boson sampler (Theorem 1 in Sec. 2). This way, it is possible to efficiently compare the data coming from an experiment or a mock-up sampler to that of an ideal sampler. This provides an advantage over other validation tests such as, for example, Bayesian approaches [30, 31, 32], which require exponentially hard classical computations. Additionally, while other commonly used validation tests based on correlators or marginal distributions, which require only polynomial-time classical computations, can be spoofed by classical algorithms [33], we argue that spoofing binned distribution tests is much harder, since there are exponentially many choices of bins and sufficient variability between the distributions corresponding to different possible choices. We leave as an open question whether a clever efficient classical algorithm exists that can spoof the validation method proposed in our work.

In what follows we introduce a list of desiderata for scalable validation tests for boson sampling, explaining how the validation scheme based on binned output modes fulfills them and detailing how the method stands in comparison with other methods in the literature.

1.1 Validation of Boson Sampling

A plethora of validation tests for boson samplers have been proposed which are able to discriminate between ideal boson samplers and other mock-up distributions, such as the uniform distribution, distributions generated by distinguishable input photon, or mean-field samplers [34]. Techniques such as pattern recognition or machine learning [35, 36], Bayesian testing [30], coarse-grained measurements [37, 38], Heavy Output Generation [8], or the analysis of marginal distributions have also been applied [33]. Each method offers its own advantages and disadvantages. A recent publication combined a variety of the tests cited the above applied sequentially to come up with a single metric describing the quality of the experiment, the Photonic Quality Factor [39]. In the context of our work, we put forward the following list of desiderata for a faithful validation test, focusing on sensitivity to realistic noise sources and computational efficiency. We would like a validation test to obey the following criteria:

1. *Generality:* The interferometer in a boson sampling experiment is drawn at random from the Haar measure, so an important requirement for a validation test is that it is applicable to an *arbitrary* linear interferometer.

2. *Sensitivity to multiphoton interference:* High-order multiphoton interference is at the core of the classical hardness of boson sampling [19, 40]. Partial distinguishability of the input photons is one of the most important noise sources in boson sampling, which may render the outcome probabilities easy to approximate. Experiments with a constant amount of photon distinguishability may be simulated by considering only k-photon interference terms, for some fixed value of k [14, 41]. Hence, a validation test should be sensitive to high-order multiphoton interference in order to discriminate an ideal boson sampler from one with partially distinguishable photons.

3. *Sampling efficiency:* Resources available for validation are limited since, compared to the exponentially large system size (domain of the sampled distribution), only a moderate

amount of samples are observed. For this reason, we request that a validation test is able to discriminate a noisy boson sampler (with some fixed noise parameters) from an ideal one by using a number of samples that scales only polynomially in the system size.

4. *Computational efficiency:* A last requirement is efficiency in post-processing the classical data. Many validation tests that aim at comparing the experiment to an ideal boson sampler require the computation of outcome probabilities, or the classical simulation of ideal boson sampling, which takes exponential time. This becomes prohibitive as experiments grow larger, which is why we restrict ourselves to polynomial-time computations in the system size.

Additionally, it is important to remark that, in a realistic setting, photon loss also plays a prominent role. Unlike partial distinguishability, the amount of loss is easy to estimate from data coming from the experiment (or even tests with classical light). We can assume that a lossy boson sampling experiment that aims at demonstrating quantum computational advantage is able to obtain high-enough output photon counts such that, if the rest of the experiment was ideal, it would still surpass the best classical simulation algorithms. As most of the outcomes would correspond to events with lost photons, it is important that a validation test is able to use this data in order to diagnose *other sources of noise*, such as partial distinguishability, which may render the experiment classically simulable. If validation required considering only postselected outcomes with no lost photons, this would sharply decrease the sampling rate of usable events, possibly making the validation unfeasible in a reasonable amount of time. We shall come back to this point later in this work.

1.2 Our contribution

We propose a validation scheme for Boson Sampling which aims at fulfilling the list of requirements stated above. Our scheme (Sec. 2) is based on a simple coarse-graining of the data coming from the boson sampling device: we group the output modes into different subsets and count how many photons end up in each of them. Our outcomes are thus given by a vector $\boldsymbol{k} =$

$(k_1, \ldots, k_K)$, where k_z is the number of photons observed in subset $\mathcal{K}_z$ (see left panel of Fig. 1). This *binning* of the output modes into different subsets allows us to deal with a space of events of much smaller size than the exponentially many outcomes of the boson sampler. For a fixed number K of subsets, the number of possible configurations of the photons in the bins is bounded by $(n+1)^K$, where n is the photon number. This implies that some of the probabilities are relatively large (of size $1/\text{poly}(n)$), and thus a meaningful estimation (i.e. up to relative error) of these large probabilities can be obtained from a polynomial number of experimental runs. Crucially, we show that a classical algorithm exists that can also estimate these probabilities efficiently, not only for ideal boson samplers but also noisy ones, involving partially distinguishable input photons as well as loss.

The validation test we consider is based on comparing these theoretically predicted probabilities to the experimentally observed ones. We provide analytical and numerical evidence that it is possible to use binned output distributions to efficiently discriminate between bosonic and classical input particles, as well as some models of partial distinguishability. We also argue that this way of validating boson samplers is sensitive to partial distinguishability, even in the presence of small amounts of loss.

1.3 Comparison to other existing methods

The versatility of the method we consider lies in the fact that it can be used for any interferometer and that the choice of the bins is completely arbitrary. It can even be done after the experiment – the same data can be tested using multiple choices of subsets, possibly ones chosen randomly. This versatility allows us to connect this method to some important validation tests for standard boson samplers that have been suggested in previous literature and even retrieve some of them as particular cases.

1.3.1 Correlators and marginal distributions

One of the most common validation protocol for boson samplers relies on low-order correlation functions of the output mode counts $\hat{n}_i$ [42, 43]
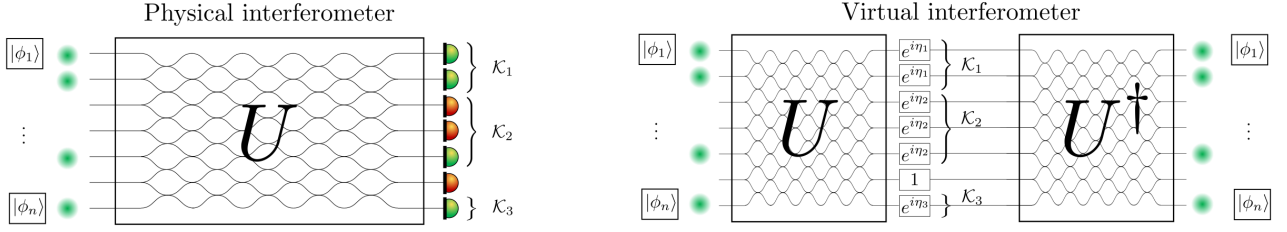
Figure 1: (Left) **Interferometric setup.** Single photons are sent through the first n input modes of a m-mode linear interferometer U. Each photon at input j carries internal degrees of freedom (polarization, arrival time, etc.) described by an (internal) wavefunction $|\phi_j\rangle$. To validate the device we use the probability distribution obtained by counting photons in binned-together output modes. In the figure, we represent in green the detectors that have detected one photon and in red the detectors that did not click. In this example, we have 3 bins and the outcome observed is $\vec{k} = (2, 1, 1)$ (the first subset $\mathcal{K}_1$ observes two photons, $\mathcal{K}_2$ observes a single photon, etc.). Such outcome probabilities are sensitive to partial distinguishability of the input and can be used to discriminate an ideal boson sampler from a noisy one. (Right) **Virtual interferometer.** The binned output mode probabilities can be obtained from the characteristic function of this distribution. The latter can be interpreted as a probability amplitude in a virtual interferometry process. More precisely, computing $x(\boldsymbol{\eta})$ as described in Eq. (22) (through Eq. (15)) is equivalent to computing the amplitude of the process where the input state of the boson sampler is left unchanged while going through a virtual interferometer V built by sandwiching the physical interferometer U and its hermitian conjugate $U^\dagger$ with a diagonal matrix of phases, encoding the choice of partition (as defined in Eq. (17) to (19)).

such as

$$c_i = \langle n_i \rangle \tag{1}$$
$$c_{ij} = \langle n_i n_j \rangle - \langle n_i \rangle \langle n_j \rangle \tag{2}$$

Correlations of order 3 and 4 have also been considered [44, 9].

A closely related validation method is that of computing k-marginals of the boson sampling distributions of an ideal experiment – which results from looking at the photon counts coming from a constant number k of detectors – and comparing them to the experiment itself [33]. This data can be used to estimate the correlators mentioned above. While marginals of fixed size and low-order correlators can be computed in polynomial time, they are known to be insensitive to higher order multiphoton interferences. As previously mentioned, the latter are crucial to reproduce the boson sampling distribution to sufficient precision [45]. In fact, it is possible to construct efficient classical mock-up samplers that are consistent with all marginals of order k [33]. Therefore, this scheme alone cannot be used to justify claims of quantum computational advantage.

We note that marginals distribution can be recovered as a particular case of our scheme, as it corresponds to choosing $K = k$ subsets with a single output mode in each. Our scheme however can be adapted to be sensitive to higher order interferences: if, for example, two equal-sized subsets are chosen, the way photons distribute in

this partition of the output modes cannot be well approximated by taking into account only few-photon interference terms.

1.3.2 Full bunching

Shchesnovich presents an interesting scheme that aims at fulfilling all the requirements stated in Sec. 1.1 [46, 40]. It relies on observing full bunching in a subset of the output modes, i.e. all the input photons are found in some chosen subset. Equivalently, one can focus on observing no photon in the complementary subset [40].

Full bunching probabilities can be approximated efficiently and numerical simulations predict that, for Haar random matrices, this quantity is maximized when photons are fully indistinguishable and decreases when they are partially distinguishable. While explicit counter-examples to this general rule of thumb were demonstrated in [47], going against general physical intuition, the method remains practical as it holds very well on average, independently of which subset is chosen.

While there is evidence that validation using bunching probabilities ticks all the boxes of desirable properties put forward in Sec. 1.1, we improve on this method quantitatively. In fact, the full bunching probability in a subset can be seen as a particular outcome probability of our more general scheme, which takes into account full pho-

ton counting distribution in one or more subsets. This allows for better distinguishing power, requiring fewer experimental samples for the validation task.

1.3.3 Suppression laws

Certain validation tests rely on a specific choice of optical network [48]. Symmetries in the unitary matrix lead to suppression laws where many outputs are prohibited [49, 50, 51]. These methods are interesting tools to diagnose noise in the network or input state, but are restricted to only a handful of networks, such as the Fourier or Sylvester matrices. However, the arguments for the hardness of boson sampling require the unitary to be chosen randomly according the the Haar measure [3]. If one can precisely tune a reconfigurable network [10], then validation can be executed first through suppression laws, and then the network can be changed to a random unitary to obtain samples. The main drawbacks are that this method leaves the door open for errors to appear in the samples when the circuit is reconfigured and provides no way to validate the final samples after circuit reconfiguration.

Nevertheless, interferometers possessing some symmetries such as the Fourier interferometer are interesting devices to test multiphoton interference due these suppression laws and their sensitivity to distinguishability of the photons. We show in Sec. 3 that our formalism allows us to compute analytically some binned output distributions for Fourier interferometers, revealing striking differences between the behavior of distinguishable particles and ideal bosons. We also observe that some characteristic suppressions observed for ideal bosons are inherited by the binned output distribution, for an appropriate choice of the bins.

1.3.4 Validation from coarse-grained measurements

Other proposals exist to validate boson samplers using coarse-grained data. In [37], the authors classify observed events in bubbles in the state space (unlike our binning of output spatial modes). Bubbles are constructed iteratively and are centered around high probability events. Numerical evidence is given to show that this is a good validation method against, for example, dis-

tinguishable input photons. However, with this validation method, the comparison of the experiment to an ideal boson sampler would still require a classical simulator of the latter which would not be efficient. A similar statement also applies to pattern recognition techniques [35].

1.4 Structure of the paper

The core of our work is divided into three sections followed by a discussion section. In Sec. 2, we explain in detail the mathematical techniques to compute the binned output distribution generated by ideal or noisy boson samplers. We focus on the analysis of the complexity of the method, showing that efficient approximations of the distribution can be obtained. In Sec. 3, we present analytical results for binned distributions in Fourier interferometers. This section may be skipped entirely by a reader who is only interested in our results regarding the validation of experiments with Haar-random interferometers. The latter are presented in our main results section (Sec. 4). We conclude with a discussion section containing open questions and perspectives for future work.

2 Formalism

As previously mentioned, the boson sampling validation method we consider in this work is based on how photons distribute into subsets of output modes, which we will also refer to as bins.

We consider the partition of $\mathcal{M} = \{1, 2, \ldots, m\}$ into K non-empty and mutually disjoint subsets $\mathcal{K}_z \subset \mathcal{M}$ with $z \in \{1, \ldots, K\}$. If the photon configuration at the output of a boson sampler is $\boldsymbol{s} = (s_1, s_2, \ldots, s_m)$, the way photons distribute in this partition denoted as $\mathcal{K} = \{\mathcal{K}_1, \ldots, \mathcal{K}_K\}$ is fully defined by a vector $\boldsymbol{k}$ of dimension K, whose components are given by

$$k_z = \sum_{j \in \mathcal{K}_z} s_j. \qquad (3)$$

We are interested in computing the probabilities $P(\boldsymbol{k})$ of observing the different possible photon number configurations in this partition. In this section, we show a classical algorithm that, for an experiment with n input photons and constant number of subsets K, estimates the probabilities

$P(\boldsymbol{k})$ up to total variation distance β in polynomial time $O(n^{2K+2}\log(n)\beta^{-2})$, given some theoretical model for the experiment which may include partial distinguishability between the photons as well as losses. Our derivation is based on the approximation of the characteristic function of the distribution $P(\boldsymbol{k})$ and is inspired by a result of Arkhipov for approximating linear statistics of ideal boson samplers [52]. The main idea we use is that this characteristic function can be interpreted as a probability amplitude of a virtual interferometric process, as depicted in Fig. 1, and thus it can be approximated via Gurvits randomized algorithm for permanent approximation [3].

2.1 Photon-counting probabilities in partitions

For a given partition $\mathcal{K} = \{\mathcal{K}_1, \ldots, \mathcal{K}_K\}$, the probability of observing a certain photon number configuration $\boldsymbol{k}$ can be obtained by summing the probabilities of all outcomes of the boson sampler that are consistent with it. However, this is impractical – each outcome probability is hard to compute (a tensor permanent if photons are partially distinguishable [53]), and there can be an exponentially large number of events that are consistent with a given photon number configuration in the partition.

A better way to compute these probabilities $P(\boldsymbol{k})$ is via the characteristic function associated with this distribution, defined as

$$x(\boldsymbol{\eta}) = \mathbb{E}_{\boldsymbol{k}}\left[\exp\left(i\boldsymbol{\eta} \cdot \boldsymbol{k}\right)\right] \tag{4}$$

$$= \sum_{\boldsymbol{k} \in \Omega^K} P(\boldsymbol{k})\exp\left(i\boldsymbol{\eta} \cdot \boldsymbol{k}\right), \tag{5}$$

with $\boldsymbol{\eta} \in \mathbb{R}^K$. Here, we have defined the set

$$\Omega^K = \{(k_1, k_2, \ldots, k_K) \mid k_z \in \Omega, \forall z \in \{1, \ldots, K\}\}, \tag{6}$$

with $\Omega = \{0, 1, \ldots, n\}$. It can be seen that the probabilities $P(\boldsymbol{k})$ can be retrieved by evaluating $x(\boldsymbol{\eta})$ at $(n+1)^K$ points on a K-dimensional grid, namely,

$$\boldsymbol{\nu_l} = \frac{2\pi \boldsymbol{l}}{n+1}, \quad \text{with } l_z \in \Omega, \forall z \in \{1, \ldots, K\} \tag{7}$$

and taking the multidimensional Fourier transform, i.e.

$$P(\boldsymbol{k}) = \frac{1}{(n+1)^K} \sum_{\boldsymbol{l} \in \Omega^K} x(\boldsymbol{\nu_l})\exp\left(-i\boldsymbol{\nu_l} \cdot \boldsymbol{k}\right). \tag{8}$$

To evaluate the characteristic function we consider the usual boson sampling setting where n photons are sent through a linear interferometer of m modes, with one photon occupying each of the first n input modes (a more general input, with more than one photon per mode, is considered in Appendix A). In order to model partial distinguishability between photons, we assume the internal degrees of freedom of the photon entering mode j, such as polarization or spectral distribution, are described by an internal state $|\phi_j\rangle$. The input state can then be written as

$$|\Psi\rangle_{in} = \prod_{j=1}^{n}\left(\hat{a}^{\dagger}_{j,\phi_j}\right)|0\rangle \tag{9}$$

where $|0\rangle$ is the vacuum state and $\hat{a}^{\dagger}_{j,\phi_j}$ is the creation operator corresponding to a photon in mode j and internal state $|\phi_j\rangle$. We also define a basis $\{|\Phi_j\rangle\}$ for the internal Hilbert space of the photons such that

$$\sum_j \langle\phi_k|\Phi_j\rangle\langle\Phi_j|\phi_k\rangle = 1, \quad \forall k. \tag{10}$$

Note that, even though the internal Hilbert space of the photons may be of infinite dimension, we only need at most n basis elements to span the Hilbert space generated by the n states $|\phi_j\rangle$. In addition, since the basis $\{|\Phi_j\rangle\}$ is orthonormal, the operators $\hat{a}_{i,\Phi_j}$ and $\hat{a}^{\dagger}_{i,\Phi_j}$ obey the usual commutation relations $[\hat{a}_{i,\Phi_j}, \hat{a}^{\dagger}_{k,\Phi_l}] = \delta_{ik}\delta_{jl}$. Therefore, we can define the number operator, which counts the number of photons in a spatial mode independently of their internal states, as

$$\hat{n}_i = \sum_j \hat{a}^{\dagger}_{i,\Phi_j}\hat{a}_{i,\Phi_j}. \tag{11}$$

Following the formalism of [53, 54], we use the standard assumption that the interferometer $\hat{U}$ acts only on the spatial modes, leaving the internal wavefunctions untouched. The relation between input and output modes is hence described by an $m \times m$ unitary matrix U via the equation

$$\hat{a}^{\dagger}_{j,\phi_j} \to \hat{b}^{\dagger}_{k,\phi_j} \quad \text{with}$$

$$\hat{a}^{\dagger}_{j,\phi_j} = \hat{U}\hat{b}^{\dagger}_{j,\phi_j}\hat{U}^{\dagger} = \sum_{k=1}^{m} U_{jk}\,\hat{b}^{\dagger}_{k,\phi_j}. \tag{12}$$

The operator which counts the number of photons in a given subset of output spatial modes is then

$$\hat{N}_{\mathcal{K}_z} = \sum_{j \in \mathcal{K}_z} \hat{n}_j = \sum_{j \in \mathcal{K}_z}\sum_k \hat{b}^{\dagger}_{j,\Phi_k}\hat{b}_{j,\Phi_k}. \tag{13}$$

We demonstrate in Appendix A that the characteristic function $x(\boldsymbol{\eta})$ can be computed as the quantum expectation value

$$x(\boldsymbol{\eta}) = \langle \Psi_{\text{out}} | e^{i\boldsymbol{\eta} \cdot \hat{\boldsymbol{N}}_{\mathcal{K}}} | \Psi_{\text{out}} \rangle \qquad (14)$$

$$= \langle \Psi_{\text{in}} | \hat{U}^\dagger e^{i\boldsymbol{\eta} \cdot \hat{\boldsymbol{N}}_{\mathcal{K}}} \hat{U} | \Psi_{\text{in}} \rangle, \qquad (15)$$

where we have used the notation

$$\boldsymbol{\eta} \cdot \hat{\boldsymbol{N}}_{\mathcal{K}} = \sum_{z=1}^{K} \eta_z \hat{N}_{\mathcal{K}_z}. \qquad (16)$$

At this point, it is useful to note that $\hat{V}(\boldsymbol{\eta}) = \hat{U}^\dagger e^{i\boldsymbol{\eta} \cdot \hat{\boldsymbol{N}}_{\mathcal{K}}} \hat{U}$ is a linear interferometer characterized by an $m \times m$ unitary matrix $V(\boldsymbol{\eta})$ (see right panel of Fig. 1). This matrix is constructed as

$$V(\boldsymbol{\eta}) = U^\dagger \Lambda(\boldsymbol{\eta}) U, \qquad (17)$$

where $\Lambda(\boldsymbol{\eta})$ is a diagonal matrix given by a product of diagonal matrices

$$\Lambda(\boldsymbol{\eta}) = \prod_{z=1}^{K} D^{(z)}(\eta_z), \qquad (18)$$

such that

$$D_{ab}^{(z)}(\eta_z) = \begin{cases} e^{i\eta_z}, & \text{if } a = b \text{ and } a \in \mathcal{K}_z, \\ 1, & \text{if } a = b \text{ and } a \notin \mathcal{K}_z, \\ 0, & \text{if } a \neq b. \end{cases} \qquad (19)$$

Using the results of Ref. [3], it can be shown that in the ideal boson sampling scenario where all photons are indistinguishable, the computation of the characteristic function is given by a matrix permanent

$$x(\boldsymbol{\eta}) = \text{perm}\left(V_n(\boldsymbol{\eta})\right), \qquad (20)$$

where V_n corresponds to the $n \times n$ upper left submatrix of the matrix V. In the more general case where the input photons can have different internal wavefunctions (see Eq. (9)), this expression is modified in a simple way. By defining the Gram matrix

$$S_{ij} = \langle \phi_i | \phi_j \rangle \qquad (21)$$

of the overlaps of the internal states of the photons, the expression takes the form

$$x(\boldsymbol{\eta}) = \text{perm}\left(S \odot V_n(\boldsymbol{\eta})\right), \qquad (22)$$

where $\odot$ is the Hadamard (elementwise) product: $(A \odot B)_{ij} = A_{ij}B_{ij}$. An explicit derivation of the expression is done in Appendix A.
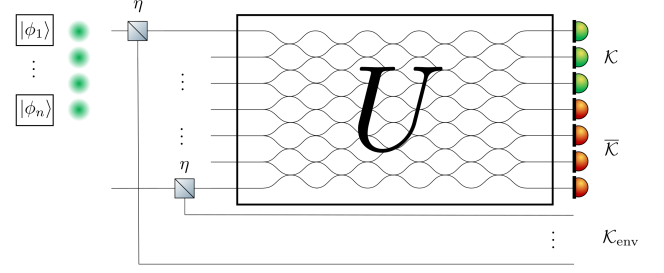

Figure 2: **Lossy interferometer model.** We use a simple model of uniform loss: each photon has a goes through a beam-splitter of transmissivity η before entering the interferometer. If the photon is reflected, it is sent into an environment mode, which simulates a lost photon. The photon number distribution in binned modes of a lossy boson sampler can be computed by considering this larger lossless interferometer.

2.2 Loss and dark counts

We can accommodate photon loss at little extra cost with this formalism. In general, a lossy linear optical circuit can be described by first applying a lossless linear interferometer W_1, followed by m parallel loss channels and a final lossless linear interferometer W_2 [16]. In turn, a loss channel acting on a given optical mode can be modelled in a *unitary* way, by introducing an ancillary environment mode in the vacuum state and applying a beam-splitter with a given transmissivity λ_i. This implies that the output statistics of a lossy boson sampler of m modes can be recovered by considering a larger lossless interferometer of $2m$ modes, described by a unitary matrix $\tilde{U}$, where only the first m modes are measured. In the case of *uniform loss*, the scheme can be simplified by considering an array of beam-splitters with the environment modes before the interferometer described by a unitary $U = W_1 W_2$ (see Fig. 2).

Taking this into consideration, it is easy to adapt the formalism from Sec. 2.1 to obtain the photon-counting probabilities in a partition of the output modes of a lossy interferometer. We simply consider the distribution in a partition of the larger interferometer $\{\mathcal{K}_1, \ldots, \mathcal{K}_K, \mathcal{K}_{env}\}$, where the last subset contains all the environment modes. Note that the size of the matrix whose permanent we need to compute in Eq. (22) depends only on the number of input photons. Hence, even for lossy interferometers, the characteristic function of the photon number distribution in the binned output modes is still given by a permanent of an $n \times n$ matrix (see Eqs. (15)

and (22))), which in this case is a submatrix of a $2m \times 2m$ unitary matrix.

Another source of experimental noise is dark counts from the detectors. Even though we do not consider explicitly this effect in our work we note that the effect of dark counts can be incorporated in the calculation of the probabilities $P(\boldsymbol{k})$. Indeed, dark counts are a default of detectors, and do not change the underlying physics of the experiment. It suffices to add their statistical contribution on top of the of the experiment with no dark counts. For example, consider the simple case of a uniform dark count probability generation $p_d \ll 1$, and a single subset of size K, the probability of observing k_d dark counts is given by a binomial

$$q(k_d) = \binom{K}{k_d} p_d^{k_d} (1 - p_d)^{K-k_d} \qquad (23)$$

The overall probability $P_d(k)$ of observing k photons is thus the convolution of the original probability distribution $P(k)$ with (23)

$$P_d(k) = (P * q)(k) = \sum_{k_d=0}^{k} q(k_d) P(k - k_d) \quad (24)$$

This has a moderate effect in the complexity of computing the probability distribution we are interested in.

2.3 Complexity analysis

As shown in Eq. (8), the probabilities $P(\boldsymbol{k})$ can be evaluated by taking a multidimensional DFT of the values of the characteristic function $x(\boldsymbol{\eta})$. Using fast methods to compute the multidimensional DFT, the full distribution can be computed in time

$$T = O(K(n+1)^K \log(n+1) C_x), \qquad (25)$$

where C_x is the cost of computing a single value of $x(\boldsymbol{\eta})$. Each quantity $x_{\boldsymbol{\eta}}$ requires the evaluation of a $n \times n$ permanent which can be computed exactly using Ryser's algorithm – the best known classical algorithm for the exact computation of permanents – in time $\mathcal{O}(n2^n)$. However, in a practical scenario, the exact computation of the probabilities is not necessary since the experimental estimation of these probabilities will always carry an error due to the finite number of samples. Precisely, we need $\mathcal{O}(1/\epsilon^2)$ samples to

estimate the probabilities $P_{\text{exp.}}(\boldsymbol{k})$ up to an additive error ϵ. Hence, if we assume we can run the experiment a polynomial number of times, we can only estimate the probabilities to a polynomially small error.

In what follows, we show that classical algorithms can also efficiently obtain such polynomially small additive error approximations due to Gurvits' permanent approximation algorithm [55]. This algorithm allows for the approximation of permanents of unitary matrices up to error ϵ in time $O(n^2/\epsilon^2)$. Our result regarding the computation of the approximate binned distribution up to a fixed *total variation distance* β is the following.

Theorem 1. *For a constant partition size K, there is a classical algorithm that computes an approximate distribution of probabilities $\tilde{P}(\boldsymbol{k})$ such that*

$$\sum_{\boldsymbol{k}} |\tilde{P}(\boldsymbol{k}) - P(\boldsymbol{k})| \leq \beta$$

in time $O(n^{2K+2} \log(n)\beta^{-2})$

Proof. Consider an approximate distribution of probabilities $\tilde{P}(\boldsymbol{k})$

$$\tilde{P}(\boldsymbol{k}) = \frac{1}{(n+1)^K} \sum_{\boldsymbol{l} \in \Omega^K} \tilde{x}(\nu_{\boldsymbol{l}}) \exp\left(-i\nu_{\boldsymbol{l}} \cdot \boldsymbol{k}\right), \tag{26}$$

obtained from the Fourier transform of approximate values of the characteristic function

$$\tilde{x}(\nu_{\boldsymbol{l}}) = x(\nu_{\boldsymbol{l}}) + \epsilon_{\boldsymbol{l}}, \tag{27}$$

where $\epsilon_{\boldsymbol{l}}$ is an error term. It can be shown that

$$\sum_{\boldsymbol{k}} |\tilde{P}(\boldsymbol{k}) - P(\boldsymbol{k})|^2 = \sum_{\boldsymbol{k}} \left| \sum_{\boldsymbol{l}} \frac{\epsilon_{\boldsymbol{l}}}{(n+1)^K} e^{-i\nu_{\boldsymbol{l}} \cdot \boldsymbol{k}} \right|^2 \tag{28}$$

$$= \frac{\sum_{\boldsymbol{l}} |\epsilon_{\boldsymbol{l}}|^2}{(n+1)^K},$$

where we have used Parseval's theorem. By defining $\epsilon = \max_{\boldsymbol{l}} \epsilon_{\boldsymbol{l}}$, we can bound the ℓ_2-norm between the approximate and exact distribution by

$$\sqrt{\sum_{\boldsymbol{k}} |\tilde{P}(\boldsymbol{k}) - P(\boldsymbol{k})|^2} \leq \epsilon. \tag{29}$$

This implies that the ℓ_1-norm is bounded by

$$\sum_{\boldsymbol{k}} |\tilde{P}(\boldsymbol{k}) - P(\boldsymbol{k})| \leq (n+1)^{K/2}\epsilon. \tag{30}$$

Therefore, in order to obtain an ℓ_1-norm bounded by some constant value β, we need to compute the approximate values $\tilde{x}(\nu_l)$ up to an error $\epsilon_l \leq \epsilon \leq \beta(n+1)^{-K/2}$. The cost of evaluating each of these values using Gurvits algorithm is $C_x = O(n^{K+2}\beta^{-2})$. This can be seen as follows. For any distinguishability matrix S, we can evaluate $x(\boldsymbol{\eta})$ up to error $\epsilon||S \odot V_n(\boldsymbol{\eta})||^n$ in time $O(n^2/\epsilon^2)$ [3]. Due to a theorem by Schur (see section of Hadamard product from Ref. [56]), we have that

$$||S \odot V_n(\boldsymbol{\eta})|| \leq (\max_i S_{ii}) \, ||V_n(\boldsymbol{\eta})|| \leq 1. \quad (31)$$

Finally, the complexity of obtaining the full approximate probability distribution $\tilde{P}(\boldsymbol{k})$ can be bounded using Eq. (25) by

$$T = O(n^{2K+2}\log(n)\beta^{-2}), \quad (32)$$

where we considered K to be a constant independent of n. $\qquad\square$

This shows that the photon counting probabilities in the binned output modes can be approximated efficiently (in polynomial time) for any polynomially small additive error [1]. While this fact is practically irrelevant to efficiently estimate usual boson sampling event probabilities (as they, on average, decrease exponentially, hence requiring ϵ to be exponentially small), it is relevant here as the probabilities $P(\boldsymbol{k})$ sum to one and that there are only $O((n+1)^K)$ of them.

2.3.1 Complexity of computing marginals

The formalism we presented also allows us to compute marginal boson sampling distributions, a commonly used validation test for boson sampling experiments [33]. For example, if we are interested in the distribution over the first K modes, we take each subset to be a single mode, i.e. $\mathcal{K}_z = \{z\}$, for $z \in \{1,\dots,K\}$. It is known that marginal distributions of ideal boson sampler (with fully indistinguishable photons) can be computed *exactly* in polynomial time [3, 57]. Here we show that the formalism we consider allows us to recover this result and extend it to any input (pure) state of partially distinguishable

---

[1]We ignore the complexity of computing the matrix $V(\boldsymbol{\eta})$, which is polynomial in m, the number of modes, itself assumed to be polynomial in the number of photons n in most use cases.

photons, as long as the internal states of the photons belong to a Hilbert space of constant size. The latter is given by the rank of the S matrix (see Eq. (21)), which we denote as r_S.

To show this result, we follow very similar lines to Ref. [57] and use on the existence of an efficient algorithm to exactly compute the permanent of n-dimensional square matrices of the form $\mathbb{1} + A$, where A has some constant rank r_A. Precisely, this takes time $O(n^{2r_A+1})$. In order to compute the marginal distribution we now consider the characteristic function (also called generating function) given by

$$x(\boldsymbol{\eta}) = \langle \Psi_{in}| \hat{U}^\dagger e^{i\sum_{j=1}^r \eta_j \hat{n}_j} \hat{U} |\Psi_{in}\rangle \quad (33)$$
$$= \mathrm{perm}\,(V_n(\boldsymbol{\eta}) \odot S). \quad (34)$$

In this case, we can write

$$V_n(\boldsymbol{\eta}) \odot S = \mathbb{1}_n + W(\boldsymbol{\eta}) \odot S, \quad (35)$$

where $\mathbb{1}_n$ is the identity matrix of dimension n and

$$W(\boldsymbol{\eta}) = U^\dagger(\Lambda(\boldsymbol{\eta}) - \mathbb{1}_n)U. \quad (36)$$

It is possible to see that $W(\boldsymbol{\eta})$ is a matrix of rank r and thus $\mathrm{rank}(W(\boldsymbol{\eta}) \odot S) \leq Kr_S$. This way, we can bound the cost of exactly calculating the generating function of the marginal distribution corresponding to a subsystem of K output modes by

$$C_x = O(n^{2Kr_S+1}). \quad (37)$$

This result can be of use to speed up computations of marginals, for example, when the main source of partial distinguishability are perturbations to the polarization state of the photons.

In more general scenarios though, the n internal states of the photons span a Hilbert space of dimension at most n and so this result is of limited use as $W(\eta)$ can have a rank which scales with the system size, implying that computing the marginals exactly with this method takes exponential time in the system size. In this case, we can may use different approaches allowing us to exploit partial distinguishability for more efficient approximations of the characteristic function. For example, using the techniques from Refs. [14, 58], we may obtain approximations where the error scales as $\log(1/x)$, where x represents a distinguishability parameter.

3 Signatures of multiphoton interference in Fourier interferometers

In this section, we give analytical evidence that the photon distribution in binned output modes contains important information about multiphoton interference. To do so, we focus on the Fourier interferometer, defined by the unitary transformation

$$F_{jk} = \frac{1}{\sqrt{m}} e^{-2\pi i (j-1)(k-1)/m}.$$

This interferometer that has been widely studied in the context of validating multiphoton interference. Due to its symmetries, it has been shown that most of the outcome probabilities are suppressed (that is, equal zero) if the inputs are fully indistinguishable [48]. A violation of these suppression laws can be used to test indistinguishability of the input photons.

Here we demonstrate that the formalism discussed in Sec. 2 can be used to obtain analytical results about how photons distribute in subsets of output modes of a Fourier interferometer. We show that even considering a single subset, the way indistinguishable photons behave is drastically different than distinguishable ones. We focus on two interesting examples, namely, the computation of the single-mode density matrix and on the photon-counting distribution on the odd output modes. Our results go beyond suppression laws as they allow us to predict the full distribution in these subsets and not only which events are suppressed.

3.1 Single-mode density matrix

One of the simplest ways of looking for signatures of multiphoton interference is by measuring subsystems of the output state of the linear interferometer, i.e. the reduced state of a few output modes. We focus here on the single-mode density matrix of a Fourier interferometer, with a single-photon in each of the input modes, i.e. $n = m$. This in an interesting setting as it falls within the scope of the results of Refs. [59, 60], which allows us to predict that the asymptotics of the single-mode density matrix is given by a thermal state with average photon number $\langle n \rangle = 1$.

To our knowledge, the exact form of the distribution for finite-sized systems has not been shown explicitly before. We show in Appendix B.2 that
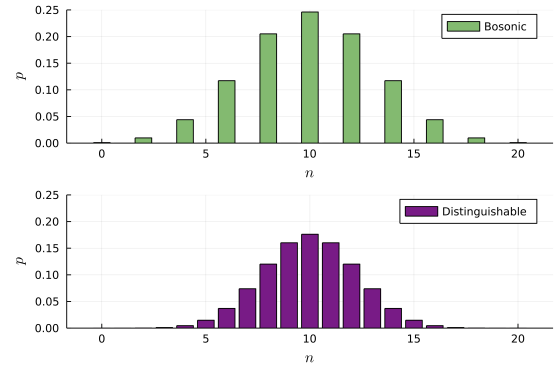

Figure 3: Comparison of the probability of seeing k photons in the odd output modes of the Fourier interferometer with 1 particle per input mode. For bosons we see a suppression of events with odd k, whereas the events with even k follow a binomial distribution. For distinguishable particles the probabilities follow a simple binomial distribution.

the formalism of Sec. 2 allows us to obtain this distribution analytically. The probability of observing k photons, if the input photons are fully indistinguishable, is given by

$$P_k^B = \sum_{a=k}^{n} (-1)^{k+a} \binom{a}{k} \binom{n}{a} \frac{a!}{n^a}. \qquad (38)$$

Although this distribution is well approximated by the geometric distribution with a mean equal to 1, it has some important differences. For example, the probability of observing $n-1$ photons is always 0, a fact that can also be predicted from suppression laws [61]. In contrast, interference of distinguishable photons results into a very different distribution. Using simple combinatorial arguments, one can see that the probability of observing k-photons in a single mode is given by the binomial distribution

$$P_k^D = \binom{n}{k} \frac{1}{n^k} \left(1 - \frac{1}{n}\right)^{n-k}. \qquad (39)$$

Asymptotically, this tends to a Poisson distribution with a mean equal to 1. This shows that photon distinguishability already plays a significant role in the photocounting statistics of a single detector.

3.2 Photon number distribution in larger subsets

In the previous section, we have chosen to look at a very simple subset of the output modes, given

by a single mode. Our formalism, however, allows us to look for signatures of multiphoton interference by considering more general subsets. If the interferometer has some particular symmetries, it is expected that choices of subsets that reflect these symmetries should reveal larger differences between the behavior of indistinguishable and distinguishable photons [62]. We give a specific example in what follows. Let us consider again the Fourier interferometer with a single photon in each of the input modes and analyse how many photons end up in the odd modes, i.e. the set $\mathcal{K} = \{1, 3, \ldots, n-1\}$, where we take the number of modes n to be even. Using the formalism of Sec. 2, we compute analytically the photon-counting probabilities in this subset in the two extreme cases of indistinguishable vs distinguishable photons, which we plot in Fig. 3 (see Appendix B.3 for a detailed derivation). For ideal bosons, we obtain

$$P_k^B = \begin{cases} 0 & \text{if } k \text{ is odd,} \\ \frac{1}{2^{n/2}} \binom{n/2}{k} & \text{if } k \text{ is even.} \end{cases} \quad (40)$$

Events with an odd number of photons are fully suppressed, whereas the events with an even photon number follow a simple binomial distribution. A sharp contrast is observed with respect to the behavior of distinguishable photons, which follows a simple binomial distribution

$$P_k^D = \frac{1}{2^n} \binom{n}{j}. \quad (41)$$

This suggests that, in an experimental setting, the analysis of photon distributions in properly chosen subsets may be used to diagnose partial distinguishability in the input photons. In particular, it would be interesting to investigate whether the measured statistical deviations to the ideal distribution of Eq. (40) may be used to bound the degree of genuine multiphoton indistinguishability of the input state [63].

4 Validation of boson samplers

The main premise of our work is that, from the way photons distribute in partitions of the output modes of a boson sampler, it is possible to tell whether we are in the presence of an ideal boson sampler from a noisy one. In this section we justify this claim for Haar-random interferometers.

First, we introduce the models of noise we analyse, namely partial distinguishability and photon loss. Subsequently we give analytical and numerical arguments showing the binned output distributions are sensitive to partial distinguishability and may be used for efficient validation tests. Finally, we stress that taking into account outcomes with a few lost photons may significantly speed-up validation tests as they still carry information about photon distinguihability.

4.1 Noise models

Although the formalism of Sec. 2 is able to encompass arbitrary inputs of partially distinguishable photons, here we consider a specific model of partial distinguishability, for the sake of performing numerical simulations about the validation method we propose in this work. We assume that the wave-functions describing the internal degrees of freedom of any pair of photons have an overlap given by a distinguishability parameter $0 \le x \le 1$. The off-diagonal elements of the distinguishability matrix S from Eq. (21) are thus $S_{ij} = x$ whereas the diagonal elements are one by definition. In this model, the distinguishability matrix is a convex interpolation of the distinguishability matrices of two extreme cases. At $x = 0$, the photons behave as fully distinguishable particles and the observed statistics corresponds to the classical case (no interference), when each one of the photons is sent at different times to the interferometer. At $x = 1$, we recover the ideal boson sampling case of linear interference between fully indistinguishable bosons. This interpolation model has been widely studied in works such as Refs. [53, 54, 40, 14], both from the perspective of understanding interference phenomena in the "quantum-to-classical" transition or with the aim of providing efficient classical simulation algorithms for noisy boson samplers.

Regarding photon loss, we restrict to the uniform loss model (Fig. 9). Each photon has a probability $0 \le \eta \le 1$ to go through the interferometer and thus lead to a detection. Mathematically, this is equivalent to adding a beam-splitter of transmissivity η at the end of each output source, with the reflected branch of the beam-splitter becoming an environment mode to which the photon can be sent (and thus represent a lost photon). We are interested in the photon configuration in a partition of the first m modes,

from which we can infer how many photons where lost to the environment modes.

4.2 Comparison to asymptotic formulas

The results in Sec. 3 indicate that the photon-counting statistics in output mode partitions can reveal striking signatures of multiparticle interference in certain symmetric interferometers. It is important to understand if this is also true in the usual boson sampling scenario, where the unitary characterizing the interferometer is drawn at random from the Haar measure. This question was addressed in the work of Shchesnovich in Refs. [64, 62], with the derivation of asymptotic laws characterizing how photons distribute in the binned output modes in large interferometers. Using combinatorial arguments, it was found that these photon-counting probabilities, when averaged over the Haar-random interferometers, are given by

$$p^D(\boldsymbol{k}) = \frac{n!}{\prod_{z=1}^{K} k_z!} \prod_{z=1}^{K} q_z^{k_z} \tag{42}$$

$$p^B(\boldsymbol{k}) = p^D(\boldsymbol{k}) \frac{\prod_{z=1}^{K}(\prod_{l=0}^{k_z-1}[1 + l/K_z])}{\prod_{l=0}^{n-1}[1 + l/m]} \tag{43}$$

where D signifies distinguishable particles, and B fully indistinguishable (bosonic) ones. We also define the bin size $K_z = |\mathcal{K}_z|$ as well as the relative bin size $q_z = K_z/m$. Here, it is assumed that the K bins span all the output modes, so from particle number conservation we have that $k_K = n - \sum_{z=1}^{K-1} k_z$. Assuming a constant relative bin size q_z, the asymptotic form of the previous expressions, as $n, m \to \infty$, is given by a multivariate Gaussian:

$$P^\sigma(\vec{k}|K) = \frac{\exp\{-n\sum_{z=1}^{K}\frac{(x_z-q_z)^2}{2(1+\sigma\rho)q_z}\}}{(2\pi(1+\sigma\rho)n)^{(K-1)/2}\prod_{z=1}^{K}\sqrt{q_z}}$$

$$\times \left(1 + \mathcal{O}\left(\frac{\rho\delta_{\sigma,+}}{n}\right)\right), \tag{44}$$

with $x_z = k_z/m$ and $\sigma = 1$ for indistinguishable particles and 0 for distinguishable ones. The difference between the behavior of these two extreme cases shows up via the particle density $\rho = n/m$ which influences the standard deviation of the Gaussian statistics when the input particles are ideal bosons. For the technical details about the validity regime of the asymptotic formula, as well as the error of this approximation, we refer to Ref. [62].
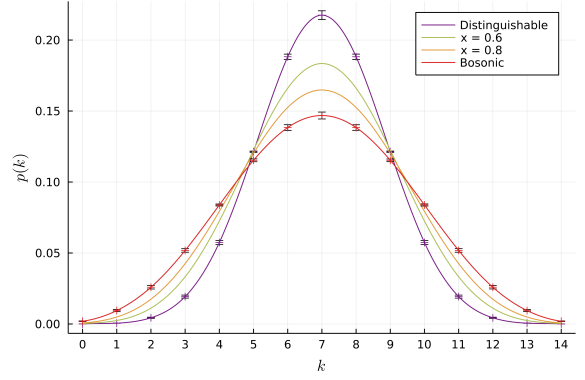

Figure 4: **Effect of partial distinguishability on a subset distribution.** We consider the photon number distribution in a subset consisting of the first half of the $m = 14$ output modes with $n = 14$. The probabilities are averaged on 1000 Haar-random unitaries. Error bars show one standard deviation. We see that they are well approximated by a Gaussian distribution, with the main difference between each case being the width of the curve.

These results suggest that the probability distribution in binned output modes is sensitive to partial distinguishability between the photons even for Haar-random unitaries. To our knowledge, there exist no explicit asymptotic formulae in this scenario and so we resort to numerical simulations to confirm this hypothesis, using the method detailed in Sec. 2. For simplicity, we consider a bipartition of the output modes into two sets of equal size and the partially distinguishability model introduced in Sec. 4.1, which interpolates between distinguishable and indistinguishable particles via a indistinguishability parameter x. The observed distribution for 14 photons in 14 output modes for several values of x is plotted in Fig. 4. The figure reveals significant differences in the probabilities as the indistinguishability parameter is varied. The width of the bell-shaped curve decreases as photons become more and more distinguishable, as suggested by the asymptotic formulas from Eqs. (42) and (43) which describe the extremes. This indicates that boson bunching effects play a role, since events where a large fraction of the photons are observed in the same bin are more likely as the photons become more indistinguishable.

4.3 Distance between distributions

A standard quantity used to quantify the distance between probability distributions is the total vari-

ation distance (TVD), defined as

$$\text{tvd}(p, q) = \sum_j |p_j - q_j|. \qquad (45)$$

This is an especially pertinent metric regarding the problem of distinguishing two distributions [65, 66] via sampling, since the number of samples n_s needed to distinguish a distribution p from another q scales as

$$n_s = \mathcal{O}(\text{tvd}(p, q)^{-2}). \qquad (46)$$

In what follows, we analyse how the TVD varies in different cases, depending on system size, partial distinguishability and loss.

For the rest of this paper, and unless specified otherwise, we will always select the bins to form an equipartition, as defined in Appendix C.1.

Bosons vs. distinguishable particles First, let us compare the two extreme cases of indistinguishable vs distinguishable particles. In Fig. 5, we analyse how the TVD, averaged over Haar-random unitaries, depends on the number of input photons. We do so in two different scenarios: when the density $\rho = n/m$ is constant as well as in the regime usually considered in boson sampling, where the number of modes $m = O(n^2)$ (and thus $\rho = 1/n$), which ensures that the probability of observing events with collisions is small [3]. As suggested by the asymptotic formulae, the density plays an important role. For constant density the TVD remains constant independently of the number of photons and consequently, the number of samples needed to distinguish the two distributions does not scale with the system size. In contrast, the bottom curve in Fig. 5 suggests an inverse polynomial decay for the TVD in the *collision-free* regime. This implies that the two distributions can still be distinguished efficiently, i.e. with a polynomial number of samples. We also remark the significant increase of the TVD if we take a larger partition size. For example, we observe that the TVD roughly doubles when we compare $K = 2$ with $K = 4$, which implies we need 4 times less samples to distinguish the two distributions, according to Eq. (46).

Partial distinguishability As previously mentioned, partial distinguishability between the input photons is one of the main sources of noise in boson samplers and may render
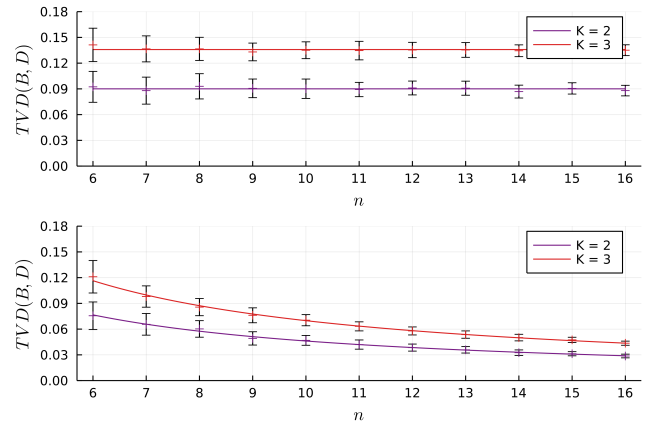
Figure 5: **Evolution of the TVD with the system size.** Haar-averaged TVD between the binned output distributions of bosonic and distinguishable inputs with varying system size. We consider two equipartions of size $K = 2$ and $K = 3$. In the top subplot, we consider a sparse but constant density $m = 5n$, while in the bottom subplot it is the no-collision regime $m = n^2$. The error bars represent the standard deviation when averaged over 100 trials. Fluctuations come from the intrinsic random Haar sampling, as well as the limited batch size used in the averaging. It can be seen that coefficient of variation decreases with increased system size (see Appendix C). We also see a significant increase of the TVD when we choose a larger partition size.

the experiment easy to simulate classically [14, 41]. Hence, a good validation test should be able to differentiate between an ideal boson sampler from one with partially distinguishable photons. To have a better understanding about the sensitivity of our validation test to partial distinguishability we again make use of the the one-parameter model interpolating between distinguishable and indistinguishable photons (see Sec. 4.1). In Fig. 6 we compare the distance between the photon-counting probabilities in the partitions when the input photons are indistinguishable ($x = 1$) and when they are partially distinguishable, as a function of the parameter x. We see a sharp increase in the TVD as we move away from the ideal case, suggesting that the probability distributions can be distinguished in practical scenarios.

Similarly, we have also analysed how the variation of the TVD between the distributions coming from ideal bosons or partially distinguishable ones (with some fixed x) varies as a function of the system size. Interestingly, the behavior follows the same trend as that that of Fig. 5: for constant densities the TVD remains con-
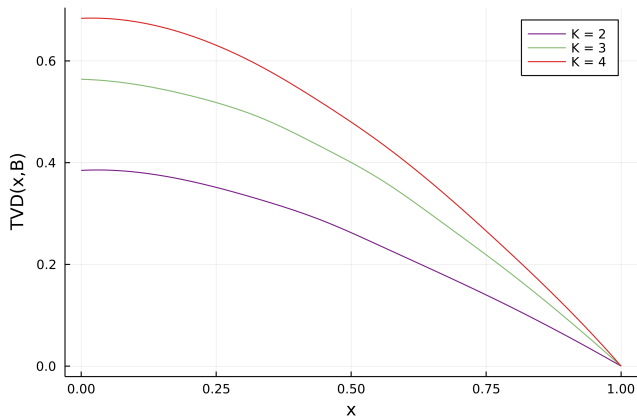
Figure 6: **Effect of partial distinguishability on TVD.** The TVD between the binned distributions corresponding to ideal photons ($x = 1$) and partially distinguishable ones is displayed as a function of the distinguishability parameter x. In this figure we took $m = n = 10$ and equipartitions of size $K = 2, 3, 4$.

stant whereas in the "collision free" regime it suggests a polynomial decay. A specific example for $x = 0.9$, can be found in Fig. 11 of Appendix C. This numerical evidence strongly suggests that the method of analysing photon-counting distributions in subsets can efficiently distinguish ideal boson samplers from ones with partially distinguishable inputs.

Dependency on photon density The previous results reaffirm the important role of photon density in the efficiency of discriminating ideal and noisy boson samplers. Although analytical results about this dependency may be difficult to obtain, we may use the numerical data to extract power laws that approximately govern this behavior. For different values of partial distinguishability, and considering equipartitions with a small number of subsets, the data suggests that TVD between the ideal distribution and that coming from partially distinguishable input photons with a fixed x is approximately described by the following behavior

$$\text{tvd}(B, x) \simeq c(K, x)\rho. \qquad (47)$$

Here, $\rho = n/m$ is the photon density and $c(K, x)$ is a numerical constant depending on the number of subsets K and the level of partial distinguishability x. More precisely, when fitting the numerical data with an ansatz model $\text{tvd}(B, x) \simeq c(K, x)\rho^r$, we obtain a value of $r \approx 0.95$, with some small variability depending on the value of

partial distinguishability chosen, which may be also due to the finite number of trials. Further plots and details regarding the quality of the approximation are given in Appendix C. While not formally proven, this approximate power law in the regimes we explored further suggests the efficiency of the validation scheme, with a polynomial decrease of the TVD between binned distributions when ρ decreases polynomially in the number of photons.

4.4 Hypothesis testing

Given a collection of experimental samples and two possible theoretical descriptions of the experiment, the formalism of Bayesian hypothesis testing allows us to predict how many samples are needed to decide which one is more likely to describe the observed data. This strategy has been exploited in the context of boson sampling in Refs. [30, 31, 32]. We may assume that one of the hypothesis to describe the experiment is an ideal boson sampler, with indistinguishable input photons. We call this the null hypothesis H_0, which we would like to test against an alternative description of the experiment H_a. The later could be for example a boson sampler with distinguishable or partially distinguishable input photons. Given an output sample $\boldsymbol{s} = (s_1, s_2, \ldots, s_m)$, we can compute the ratio between the probability of observing this sample assuming the null hypothesis $p(\boldsymbol{s}|H_0)$, and its counterpart assuming the alternative hypothesis $p(\boldsymbol{s}|H_a)$. The product of these ratios over the different samples give us the Bayesian factor

$$\chi = \prod_{i=1}^{n_s} \frac{p(\boldsymbol{s}^{(i)}|H_0)}{p(\boldsymbol{s}^{(i)}|H_a)}, \qquad (48)$$

where $\boldsymbol{s}^{(i)}$ refers to the i-th sample and n_s to the total number of samples. The confidence in the hypothesis H_0 can be computed from χ as

$$p_{null} = \chi/(\chi + 1). \qquad (49)$$

Although there is numerical evidence that this method requires only a modest number of samples to validate an ideal boson sampler against certain alternative hypothesis [30, 31, 32], the main drawback is that the computation of the confidence p_{null} is not efficient. Indeed, the output probabilities of ideal boson samplers $p(\boldsymbol{s}^{(i)}|H_0)$ are exponentially hard to approximate.
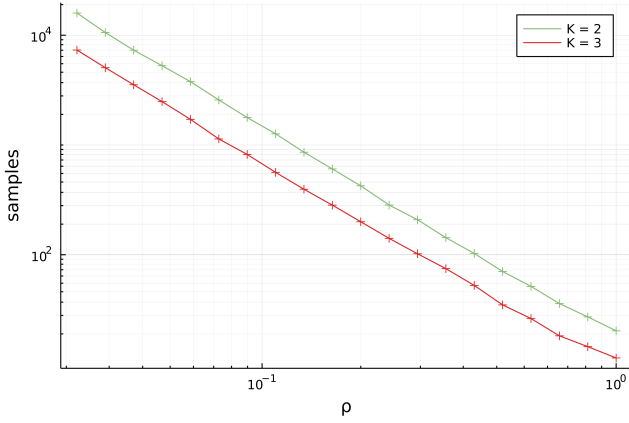
Figure 7: **Number of samples** required to distinguish the indistinguishable and distinguishable inputs. This number of samples is determined by taking the average number of samples required to obtain a certainty of $p_{null} = 95\%$ using the Bayesian approach laid out in the main text, see (48). We numerically confirmed that this quantity depends to a good approximation only from boson density ρ in numerical trials, as is discussed in the main text for the TVD. This specific figure was generated using parameters $n = 10$, $m = 10, \ldots, 300$ and each point is an average over 1000 Haar-random unitaries.

Figure 8: **Number of samples** required to reject the null hypothesis of having indistinguishable bosons when the input is actually the one parameter interpolation with indistinguishability $0.8 \leq x \leq 0.99$. This number of samples is determined by taking the average number of samples required to obtain a certainty less than $p_{null} = 5\%$ using the Bayesian approach laid out in the main text, see (48). This specific figure was generated using parameters $n = 10$, $m = 10$ and each point is an average over 1000 Haar-random unitaries.

In this section, we consider as output samples the events $\boldsymbol{k}$, corresponding to the photon number distribution in K bins as discussed in Sec. 2. We show that this simpler-to-compute probability distribution can be used to validate boson sampling experiments, instead of the full outcome distribution. In particular, we are interested in the following question: how many samples do we need to reject the hypothesis that we have an ideal boson sampler when we are in the presence of a noisy one? To give an example, we again focus on the interpolating model of partial distinguishability discussed in Sec. 4.1. In Fig. 8, we plot the number of samples needed to reject the null hypothesis with a confidence of 95% if the experiment is described by a boson sampler whose input has a distinguishability parameter x. We observe that for 10 photons in 10 modes and a distinguishability parameter $x = 0.8$, a few hundred samples are enough to reject the null hypothesis even in the simplest case where we choose two equal-sized bins. This is improved by a factor of about one half if we bin the output modes into three subsets, thus gaining more information about the full probability distribution. We remark also that, as expected, the number of samples to reject the null hypothesis sharply in-

creases as the noisy boson sampler becomes closer to ideal, i.e. as x tends to one.

Numerical evidence also suggests that it is possible to extract approximate power laws, that allow us to predict the number of samples needed as a function of the photon density, in analogy to what was done for the TVD in Sec. 4.3. In the case where the task is to differentiate between distinguishable vs. indistinguishable input photons, we verify numerically that this dependence is well described by the following power law

$$n_s \approx \frac{d(K)}{\rho^{5/2}}, \qquad (50)$$

which we extract from the data of Fig. 7. Here, $d(K)$ a constant depending on the choice of partition and level of partial distinguishability. Similar power laws may be extracted when the input photons are partially distinguishable (more details in Appendix C). Such extrapolations are useful to predict the necessary sampling rates for validate experiments when scaling up the system size.

4.5 Validation in the presence of loss

Thus far, we have not yet considered the role of loss in the validation task. The average number of photons that go through the linear opti-

cal circuit decreases exponentially with the circuit depth [16], which is usually linear with the number of modes. Hence, most of the experimental observations will be of lossy events and even though postselection on "lossless" events is possible, it would lead to an exponential decrease of the sampling rate. For this reason, it is interesting to consider events with a few lost photons for the task of validating an experiment trying to demonstrate a quantum computational advantage – this not only increases significantly the sampling rate but also these experiments may still be difficult to simulate with classical algorithms, provided that the photons are fully indistinguishable from each other [28].

The question we address in this section is whether such lossy events can be used to validate the experiment faster, i.e. whether they contain useful information about other sources of noise affecting the experiment, namely photon distinguishability, which may render the experiment easy to simulate classically. Let us consider again the hypothesis testing setting using the data from how photons distribute in subsets of output modes. Here we consider only a single subset with half the modes for simplicity. We define the validation time as the time needed to distinguish between the ideal hypothesis H_0 and an alternative one H_a with some predetermined confidence (say 95%), assuming a constant sampling rate of the lossy boson sampler. As we will see, using data with lost photons to test for photon distinguishability, may lead to a significant speedup in the validation time. To give a concrete example, we consider the task of validating a boson sampler with 10 photons in 10 modes for different loss parameters. We assume our boson sampler has a partially distinguishable input state with distinguishability parameter $x = 0.9$ (hypothesis H_a) and the task is to test if it is an ideal one with $x = 1$ (hypothesis H_0). We define T_l as the average validation time over Haar-random interferometers, if we take into account the data up to l lost photons. In Fig. 9, we plot the ratio T_0/T_l as a function of the loss rate which we assume to be uniform. This quantity reflects the average speed-up obtained by considering data with lost photons. For a loss rate of 0.2, a speed-up of around 40-fold is obtained when considering all the data, independently of how many photons were lost. Fig. 9 also reveals that, as expected,
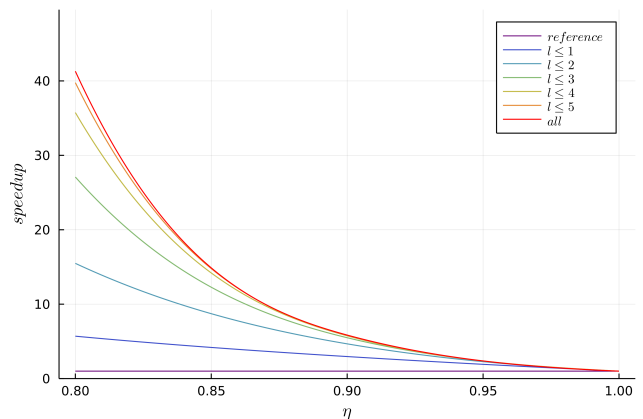

Figure 9: **Validation speed up using lost photons.** In this figure, we take a single subset of half the $m = 10$ output modes with $n = 10$ photons. We validate that the observations are indistinguishable bosons versus the one-parameter model with $x = 0.9$. We consider a model of uniform loss, with single photon transmissivity η. The reference curve shows validation with no lost photons (y-coordinate is 1). Curves with $l \leq 1, \ldots, 5$ show that the including output data with up to l photons lost gives significantly a faster validation scheme. This speed up becomes less and less important as l increases. The data is averaged over 100 Haar-random unitaries and 1000 validation runs for each unitary.

events where more photons were lost contain less information about the distinguishability of the input. This is visible, for example, from the fact that the speed-up obtained when taking the data with up to five lost photons is very similar to taking all the data. We have verified numerically that this speed up tends to increase in larger systems, even for a fixed η.

5 Discussion

In this work, we have showed that a coarse-graining of the boson sampling output distribution by grouping the output modes into bins, provides a simpler to analyse outcome distribution and a natural validation test for boson samplers. We demonstrate that, given a theoretical model of the experiment, the binned output distribution can be classically approximated as efficiently as if we run the experiment itself. The main technique we use to obtain this result is the computation of the (discrete) characteristic function of the binned distribution via Gurvits randomized algorithm for permanent approximation [3].

Even for a small number of output bins, the distribution reveals great sensitivity to photon

distinguishability. Our numerical simulations using this validation test suggest that a polynomial number of samples is sufficient to distinguish an ideal boson sampler from one with partially distinguishable input photons. We also showed how in realistic situations – where the experimental data includes a vast majority of events where photons are lost – the outcomes with lost photons contain useful information about partial distinguishability and that the effective use of this data can greatly speed-up validation tests.

Multiple interesting research questions arise related to this validation test. If we do not trust that the data is coming from an actual physical experiment, can we guarantee that no efficient classical algorithm exists that may spoof the test? An important property of the validation test we consider in our work is that the subsets we choose to test the experimental data can be chosen arbitrarily *a posteriori*. At first sight, this makes the test harder to spoof: a potential adversary trying to mimic the behavior of an ideal boson sampler would have to generate samples such that they are consistent with the correct coarse-grained photon-counting distributions for exponentially many possible subset choices. However, one may wonder whether the knowledge of the analytic form of the average over Haar-random unitaries of the binned-output distribution (see Eq. (43) and Refs. [62, 64]) may be used to spoof the test. We argue in Appendix C.6 that the knowledge of the mean of the binned distributions is not enough to spoof the test, because different partitions choices exhibit significant fluctuations around the mean which should be measurable with a polynomial number of samples.

Another interesting question is whether it is possible to obtain analytical results corroborating our numerical evidence about the sample efficiency of the method. Previous results on validation tests based on generalized bunching probabilities from Ref. [46], which can be seen as a particular outcome of a binned distribution, suggest that comparing bosonic and distinguishable particles can be done in a sample efficient way. However, the problem becomes more difficult when different models of partial distinguishability come into play.

Moreover, while we set our interest in using binned output probabilities as a validation

method, we also believe that it could be a useful tool for probing partial distinguishability. Deviations from the expected binned output distributions may possibly be used to quantify the degree of indistinguishability of the input photons, specially in highly symmetric interferometers such as the Fourier transform, where large differences between distinguishable and indistinguishable particles are observed.

Our work also opens up the question of whether certain decision or function problems that can be solved by boson samplers proposed in Refs. [67, 68], with potential cryptographic applications, may actually be solved by efficient classical algorithms. Some of these problems are also based on questions related to probability distributions obtained after certain binning of the boson sampling data. Even though the binning procedure is not directly equivalent to ours, it would be worth investigating if the binned distributions from Refs. [67, 68] may be approximated via a similar formalism to that presented in Sec. 2.

During the completion of this work, we became aware of the recent works from Refs. [69, 70, 71]. Ref. [69] develops an efficient classical algorithm to approximate molecular vibronic spectra, using a formalism similar to Sec. 2 based on approximating Fourier components of the target probability distribution using Gurvits algorithms. In turn, Refs. [70, 71] consider photon-number distributions in binned output modes as validation tests of Gaussian boson samplers. The authors use phase space methods to approximate these distributions, which are referred to as group-count probabilities. In contrast, we focus on validation of standard boson samplers and develop a different formalism to compute group-count probabilities which does not involve phase space averages. Another difference of our work is that, while the authors of Refs. [70, 71] focus on noise sources more likely to affect Gaussian boson samplers, we focus on testing the sensitivity of the method to photon distinguishability as this is one of the main noise sources affecting standard boson samplers. Overall, we believe our contribution, together with those previous works, suggests that analysing how photons distribute in binned output modes is a scalable and practical method to use for validation of near-future experiments.

Note that, for the sake of conciseness and ease of reading, we made the choice to limit the nu-

merical analysis exposed in this paper to simple noise models such as uniform partial distinguishability and loss. In Ref. [72], we provide tools that allow for general noise models, see below.

Code availability

A complete Julia package, BOSONSAMPLING.JL, and its related package, PERMANENTS.JL, includes all the tools presented in this paper and many more regarding boson sampling. They are written in a user-friendly way and are aimed at experimentalists wanting to use this work. This package is already being used in boson sampling experiments. The package is also focused on making it easy to write new models (such as noisy detectors, or new types of boson sampling) in an easy to write manner while being as fast as low-level languages such as C.

A related publication [72] regarding BOSON-SAMPLING.JL is available on the arxiv.

A complete tutorial and documentation are provided and interested users are welcome to contact the authors for possible extensions or specific needs.

All available Figures and data found in this article can be reproduced directly from the `/docs/publication/partition/` folder of the package.

Acknowledgments

References

[1] John Preskill. Quantum computing in the nisq era and beyond. *Quantum*, 2:79, 2018. DOI: 10.22331/q-2018-08-06-79.

[2] John Preskill. Quantum computing 40 years later. *arXiv preprint arXiv:2106.10522*, 2021. DOI: 10.1201/9781003358817-7.

[3] Scott Aaronson and Alex Arkhipov. The computational complexity of linear optics. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 333–342, 2011. DOI: 10.1145/1993636.1993682.

[4] Lund, Austin P and Laing, Anthony and Rahimi-Keshari, Saleh and Rudolph, Terry and O'Brien, Jeremy L and Ralph, Timothy C. Boson sampling from a Gaussian state. *Physical review letters*, 113 (10):100502, 2014. DOI: 10.1103/PhysRevLett.113.100502.

[5] Craig S. Hamilton, Regina Kruse, Linda Sansoni, Sonja Barkhofen, Christine Silberhorn, and Igor Jex. Gaussian boson sampling. *Physical Review Letters*, 119(17), oct 2017. DOI: 10.1103/physrevlett.119.170501. URL `https://doi.org/10.1103%2Fphysrevlett.119.170501`.

[6] Chakhmakhchyan, Levon and Cerf, Nicolas J. Boson sampling with Gaussian measurements. *Physical Review A*, 96(3):032326, 2017. DOI: 10.1103/PhysRevA.96.032326.

[7] Chabaud, Ulysse and Douce, Tom and Markham, Damian and Van Loock, Peter and Kashefi, Elham and Ferrini, Giulia. Continuous-variable sampling from photon-added or photon-subtracted squeezed states. *Physical Review A*, 96(6):062307, 2017. DOI: 10.1103/PhysRevA.96.062307.

[8] Zhong, Han-Sen and Wang, Hui and Deng, Yu-Hao and Chen, Ming-Cheng and Peng, Li-Chao and Luo, Yi-Han and Qin, Jian and Wu, Dian and Ding, Xing and Hu, Yi and others. Quantum computational advantage using photons. *Science*, 370(6523):1460–1463, 2020. DOI: 10.1126/science.abe8770.

[9] Han-Sen Zhong, Yu-Hao Deng, Jian Qin, Hui Wang, Ming-Cheng Chen, Li-Chao Peng, Yi-Han Luo, Dian Wu, Si-Qiu Gong, Hao Su, et al. Phase-programmable gaussian boson sampling using stimulated squeezed light. *Physical review letters*, 127(18):180502, 2021. DOI: 10.1364/IPRSN.2022.ITu3B.1.

[10] Lars S Madsen, Fabian Laudenbach, Mohsen Falamarzi Askarani, Fabien Rortais, Trevor Vincent, Jacob FF Bulmer, Filippo M Miatto, Leonhard Neuhaus, Lukas G Helt, Matthew J Collins, et al. Quantum computational advantage with a programmable photonic processor. *Nature*, 606(7912):75–81, 2022. DOI: 10.1038/s41586-022-04725-x.

[11] Hui Wang, Jian Qin, Xing Ding, Ming-Cheng Chen, Si Chen, Xiang You, Yu-Ming He, Xiao Jiang, L You, Z Wang, et al. Boson sampling with 20 input photons and a 60-mode interferometer in a 10^{14}-dimensional Hilbert space. *Physical review letters*, 123 (25):250503, 2019. DOI: 10.1103/PhysRevLett.123.250503.

[12] Carsten Robens, Iñigo Arrazola, Wolfgang Alt, Dieter Meschede, Lucas Lamata, Enrique Solano, and Andrea Alberti. Boson sampling with ultracold atoms. *arXiv preprint arXiv:2208.12253*, 2022. DOI: 10.48550/arXiv.2208.12253.

[13] Saleh Rahimi-Keshari, Timothy C Ralph, and Carlton M Caves. Sufficient conditions for efficient classical simulation of quantum optics. *Physical Review X*, 6(2):021039, 2016. DOI: 10.1103/PhysRevX.6.021039.

[14] 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 distinguishable photons. *Phys. Rev. Lett.*, 120:220502, May 2018. DOI: 10.1103/PhysRevLett.120.220502. URL `https://link.aps.org/doi/10.1103/PhysRevLett.120.220502`.

[15] Michał Oszmaniec and Daniel J Brod. Classical simulation of photonic linear optics with lost particles. *New Journal of Physics*, 20(9):092002, 2018. DOI: 10.1088/1367-2630/aadfa8.

[16] Raúl García-Patrón, Jelmer J Renema, and Valery Shchesnovich. Simulating boson sampling in lossy architectures. *Quantum*, 3:169, 2019. DOI: 10.22331/q-2019-08-05-169.

[17] Daniel Jost Brod and Michał Oszmaniec. Classical simulation of linear optics subject

Accepted in 〈 〉uantum 2024-08-15, click title to verify. Published under CC-BY 4.0.

19

to nonuniform losses. *Quantum*, 4:267, 2020. DOI: 10.22331/q-2020-05-14-267.

[18] Jelmer Renema, Valery Shchesnovich, and Raul Garcia-Patron. Classical simulability of noisy boson sampling. *arXiv preprint arXiv:1809.01953*, 2018. DOI: 10.48550/arXiv.1809.01953.

[19] Valery S Shchesnovich. Noise in boson sampling and the threshold of efficient classical simulatability. *Physical Review A*, 100(1):012340, 2019. DOI: 10.1103/PhysRevA.100.012340.

[20] Hangleiter, Dominik and Kliesch, Martin and Eisert, Jens and Gogolin, Christian. Sample Complexity of Device-Independently Certified "Quantum Supremacy". *Phys. Rev. Lett.*, 122:210502, May 2019. DOI: 10.1103/PhysRevLett.122.210502.

[21] Ulysse Chabaud, Frédéric Grosshans, Elham Kashefi, and Damian Markham. Efficient verification of boson sampling. *Quantum*, 5: 578, 2021. DOI: 10.22331/q-2021-11-15-578.

[22] Samuel Goldstein, Simcha Korenblit, Ydan Bendor, Hao You, Michael R Geller, and Nadav Katz. Decoherence and interferometric sensitivity of boson sampling in superconducting resonator networks. *Physical Review B*, 95(2):020502, 2017. DOI: 10.1103/PhysRevB.95.020502.

[23] Peter P Rohde and Timothy C Ralph. Error tolerance of the boson-sampling model for linear optics quantum computing. *Physical Review A*, 85(2):022332, 2012. DOI: 10.1103/PhysRevA.85.022332.

[24] Gil Kalai and Guy Kindler. Gaussian noise sensitivity and bosonsampling. *arXiv preprint arXiv:1409.3093*, 2014. DOI: 10.48550/ARXIV.1409.3093.

[25] Anthony Leverrier and Raúl García-Patrón. Analysis of circuit imperfections in bosonsampling. *arXiv preprint arXiv:1309.4687*, 2013. DOI: 10.26421/QIC15.5-6-8.

[26] VS Shchesnovich. Sufficient condition for the mode mismatch of single photons for scalability of the boson-sampling computer. *Physical Review A*, 89(2):022333, 2014. DOI: 10.1103/PhysRevA.89.022333.

[27] Alex Arkhipov. Bosonsampling is robust against small errors in the network matrix. *Physical Review A*, 92(6):062326, 2015. DOI: 10.1103/PhysRevA.92.062326.

[28] Scott Aaronson and Daniel J Brod. Bosonsampling with lost photons. *Physical Review A*, 93(1):012335, 2016. DOI: 10.1103/PhysRevA.93.012335.

[29] Ludovico Latmiral, Nicolò Spagnolo, and Fabio Sciarrino. Towards quantum supremacy with lossy scattershot boson sampling. *New Journal of Physics*, 18 (11):113008, 2016. DOI: 10.1088/1367-2630/18/11/113008.

[30] Bentivegna, Marco and Spagnolo, Nicolo and Vitelli, Chiara and Brod, Daniel J and Crespi, Andrea and Flamini, Fulvio and Ramponi, Roberta and Mataloni, Paolo and Osellame, Roberto and Galvao, Ernesto F and others. Bayesian approach to boson sampling validation. *International Journal of Quantum Information*, 12(07n08):1560028, 2015. DOI: 10.1142/S021974991560028X.

[31] Dai, Zhe and Liu, Yong and Xu, Ping and Xu, WeiXia and Yang, XueJun and Wu, Jun-Jie. A Bayesian validation approach to practical boson sampling. *Science China Physics, Mechanics & Astronomy*, 63(5):1–8, 2020. DOI: 10.1007/s11433-019-1440-y.

[32] Fulvio Flamini, Mattia Walschaers, Nicolò Spagnolo, Nathan Wiebe, Andreas Buchleitner, and Fabio Sciarrino. Validating multi-photon quantum interference with finite data. *Quantum Science and Technology*, 5(4):045005, 2020. DOI: 10.1088/2058-9565/aba03a.

[33] Benjamin Villalonga, Murphy Yuezhen Niu, Li Li, Hartmut Neven, John C Platt, Vadim N Smelyanskiy, and Sergio Boixo. Efficient approximation of experimental gaussian boson sampling. *arXiv:2109.11525*, 2021. DOI: 10.48550/arxiv.2109.11525.

[34] Mattia Walschaers. Signatures of many-particle interference. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 53(4):043001, 2020. DOI: 10.1088/1361-6455/ab5c30.

[35] Iris Agresti, Niko Viggianiello, Fulvio Flamini, Nicolò Spagnolo, Andrea Crespi, Roberto Osellame, Nathan Wiebe, and Fabio Sciarrino. Pattern recognition techniques for boson sampling validation. *Physical Review X*, 9(1):011013, 2019. DOI: 10.1103/PhysRevX.9.011013.

[36] Fulvio Flamini, Nicolò Spagnolo, and Fabio Sciarrino. Visual assessment of multi-photon interference. *Quantum Science and Technology*, 4(2):024008, 2019. DOI: 10.1088/2058-9565/ab04fc.

[37] Wang, Sheng-Tao and Duan, Lu-Ming. Certification of boson sampling devices with coarse-grained measurements. *arXiv:1601.02627*, 2016. DOI: https://doi.org/10.48550/arXiv.1601.02627.

[38] Jacques Carolan, Christopher Harrold, Chris Sparrow, Enrique Martín-López, Nicholas J Russell, Joshua W Silverstone, Peter J Shadbolt, Nobuyuki Matsuda, Manabu Oguma, Mikitaka Itoh, et al. Universal linear optics. *Science*, 349(6249):711–716, 2015. DOI: 10.1126/science.aab3642.

[39] Mezher, Rawad and Mansfield, Shane. Assessing the quality of near-term photonic quantum devices. *arXiv preprint arXiv:2202.04735*, 2022. DOI: https://doi.org/10.48550/arXiv.2202.04735.

[40] Valery Shchesnovich. Distinguishing noisy boson sampling from classical simulations. *Quantum*, 5:423, March 2021. ISSN 2521-327X. DOI: 10.22331/q-2021-03-29-423. URL https://doi.org/10.22331/q-2021-03-29-423.

[41] Alexandra E Moylett, Raúl García-Patrón, Jelmer J Renema, and Peter S Turner. Classically simulating near-term partially-distinguishable and lossy boson sampling. *Quantum Science and Technology*, 5(1):015001, nov 2019. DOI: 10.1088/2058-9565/ab5555. URL https://doi.org/10.1088%2F2058-9565%2Fab5555.

[42] Mattia Walschaers, Jack Kuipers, Juan-Diego Urbina, Klaus Mayer, Malte Christopher Tichy, Klaus Richter, and Andreas Buchleitner. Statistical benchmark for bosonsampling. *New Journal of Physics*, 18(3):032001, 2016. DOI: 10.1088/1367-2630/18/3/032001.

[43] Mattia Walschaers. *Statistical Benchmarks for Quantum Transport in Complex Systems: From Characterisation to Design*. Springer, 2018. DOI: 10.1007/978-3-319-93151-7.

[44] Taira Giordani, Fulvio Flamini, Matteo Pompili, Niko Viggianiello, Nicolò Spagnolo, Andrea Crespi, Roberto Osellame, Nathan Wiebe, Mattia Walschaers, Andreas Buchleitner, et al. Experimental statistical signature of many-body quantum interference. *Nature Photonics*, 12(3):173–178, 2018. DOI: 10.1038/s41566-018-0097-4.

[45] Valery Shchesnovich. Boson sampling cannot be faithfully simulated by only the lower-order multi-boson interferences. *arXiv preprint arXiv:2204.07792*, 2022. DOI: https://doi.org/10.48550/arXiv.2204.07792.

[46] VS Shchesnovich. Universality of generalized bunching and efficient assessment of boson sampling. *Physical review letters*, 116(12):123601, 2016. DOI: 10.1103/PhysRevLett.116.123601.

[47] Benoit Seron, Leonardo Novo, and Nicolas J Cerf. Boson bunching is not maximized by indistinguishable particles. *Nature Photonics*, 17(8):702–709, 2023. DOI: 10.1038/s41566-023-01213-0.

[48] Malte C Tichy, Klaus Mayer, Andreas Buchleitner, and Klaus Mølmer. Stringent and efficient assessment of boson-sampling devices. *Physical review letters*, 113(2):020502, 2014. DOI: 10.1103/PhysRevLett.113.020502.

[49] Christoph Dittel, Gabriel Dufour, Mattia Walschaers, Gregor Weihs, Andreas Buchleitner, and Robert Keil. Totally destructive many-particle interference. *Physical Review Letters*, 120(24):240404, 2018. DOI: 10.1103/PhysRevLett.120.240404.

[50] Niko Viggianiello, Fulvio Flamini, Luca Innocenti, Daniele Cozzolino, Marco Bentivegna, Nicolò Spagnolo, Andrea Crespi, Daniel J Brod, Ernesto F Galvão, Roberto Osellame, et al. Experimental generalized quantum suppression law in sylvester interferometers. *New Journal of Physics*, 20(3):033017, 2018. DOI: 10.1088/1367-2630/aaad92.

[51] Andrea Crespi. Suppression laws for multiparticle interference in sylvester interferometers. *Physical Review A*, 91(1):013811, 2015. DOI: 10.1103/PhysRevA.91.013811.

[52] Arkhipov, Alex. Computing the distribution of linear statistics of Boson Sampling. *Private notes*, 2014.

[53] Malte C Tichy. Sampling of partially distinguishable bosons and the relation to the multidimensional permanent. *Physical Review A*, 91(2):022316, 2015. DOI: 10.1103/PhysRevA.91.022316.

[54] VS Shchesnovich. Partial indistinguishability theory for multiphoton experiments in multiport devices. *Physical Review A*, 91(1):013844, 2015. DOI: 10.1103/PhysRevA.91.013844.

[55] Leonid Gurvits and Alex Samorodnitsky. A deterministic algorithm for approximating the mixed discriminant and mixed volume, and a combinatorial corollary. *Discrete & Computational Geometry*, 27:531–550, 06 2002. DOI: 10.1007/s00454-001-0083-2.

[56] Johnson, Charles R. *Matrix theory and applications*, volume 40. American Mathematical Soc., 1990. DOI: 10.1090/psapm/040.

[57] Dmitri A. Ivanov and Leonid Gurvits. Complexity of full counting statistics of free quantum particles in product states. *Phys. Rev. A*, 101:012303, Jan 2020. DOI: 10.1103/PhysRevA.101.012303. URL https://link.aps.org/doi/10.1103/PhysRevA.101.012303.

[58] Jelmer J Renema. Marginal probabilities in boson samplers with arbitrary input states. *arXiv preprint arXiv:2012.14917*, 2020. DOI: 10.48550/arXiv.2012.14917.

[59] Clive D Cushen and Robin L Hudson. A quantum-mechanical central limit theorem. *Journal of Applied Probability*, 8(3): 454–469, 1971. ISSN 00219002. DOI: 10.2307/3212170.

[60] Becker, Simon and Datta, Nilanjana and Lami, Ludovico and Rouze, Cambyse. Convergence rates for the quantum central limit theorem. *Communications in Mathematical Physics*, 383(1):223–279, feb 2021. DOI: 10.1007/s00220-021-03988-1. URL https://doi.org/10.1007%2Fs00220-021-03988-1.

[61] Malte Christopher Tichy, Markus Tiersch, Fernando de Melo, Florian Mintert, and Andreas Buchleitner. Zero-transmission law for multiport beam splitters. *Physical review letters*, 104(22):220405, 2010. DOI: 10.1103/PhysRevLett.104.220405.

[62] Valery S Shchesnovich. Asymptotic gaussian law for noninteracting indistinguishable particles in random networks. *Scientific reports*, 7(1):1–11, 2017. DOI: 10.1038/s41598-017-00044-8.

[63] Brod, Daniel J and Galvao, Ernesto F and Viggianiello, Niko and Flamini, Fulvio and Spagnolo, Nicolo and Sciarrino, Fabio. Witnessing genuine multiphoton indistinguishability. *Physical review letters*, 122(6):063602, 2019. DOI: 10.1103/PhysRevLett.122.063602.

[64] VS Shchesnovich. Quantum de moivre–laplace theorem for noninteracting indistinguishable particles in random networks. *Journal of Physics A: Mathematical and Theoretical*, 50(50):505301, 2017. DOI: 10.1088/1751-8121/aa8e8f.

[65] Gregory Valiant and Paul Valiant. An automatic inequality prover and instance optimal identity testing. *SIAM Journal on Computing*, 46(1):429–455, 2017. DOI: 10.1137/151002526.

[66] Blais, Eric and Canonne, Clement Louis and Gur, Tom. Alice and Bob Show Distribution Testing Lower Bounds (They don't talk to each other anymore.). In *Electron. Colloquium Comput. Complex.*, volume 23, page 168, 2016. DOI: https://doi.org/10.4230/LIPIcs.ITCS.2018.53.

[67] Georgios M Nikolopoulos and Thomas Brougham. Decision and function problems based on boson sampling. *Physical Review A*, 94(1):012315, jul 2016. DOI: 10.1103/PhysRevA.94.012315. URL https://doi.org/10.1103%2Fphysreva.94.012315.

[68] Nikolopoulos, Georgios M. Cryptographic one-way function based on boson sampling. *Quantum Information Processing*, 18 (8):1–25, jul 2019. DOI: 10.1007/s11128-019-2372-9. URL https://doi.org/10.10072Fs11128-019-2372-9.

[69] Changhun Oh, Youngrong Lim, Yat Wong, Bill Fefferman, and Liang Jiang. Quantum-inspired classical algorithms for molecular vibronic spectra. *Nature Physics*, 20(2):225–231, 2024. DOI: 10.1038/s41567-023-02308-9.

[70] Dellios, Alexander S and Reid, Margaret D and Opanchuk, Bogdan and Drummond, Peter D. Validation tests for GBS quantum computers using grouped count probabilities. *arXiv:2211.03480*, 2022. DOI: https://doi.org/10.48550/arXiv.2211.03480.

[71] Drummond, Peter D and Opanchuk, Bogdan and Dellios, Alexander and Reid, Margaret D. Simulating complex networks in phase

space: Gaussian boson sampling. *Physical Review A*, 105(1):012427, 2022. DOI: 10.1364/QIM.2021.F1C.5.

[72] Benoit Seron and Antoine Restivo. Bosonsampling. jl: A julia package for quantum multi-photon interferometry. *Quantum*, 8:1378, 2024. DOI: https://doi.org/10.22331/q-2024-06-18-1378.

[73] Malte C Tichy. Interference of identical particles from entanglement to boson-sampling. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 47(10):103001, May 2014. ISSN 1361-6455. DOI: 10.1088/0953-4075/47/10/103001. URL `http://dx.doi.org/10.1088/0953-4075/47/10/103001`.

[74] Henryk Minc. *Permanents*, volume 6. Cambridge University Press, 1984. DOI: 10.1017/CBO9781107340688.

A Transition amplitudes in the partially distinguishable case

A.1 Expression of the characteristic function as an amplitude

Let us first show that the characteristic function introduced in Eq. 4

$$x(\boldsymbol{\eta}) = \mathbb{E}_{\boldsymbol{k}} \left[\exp\left(i\boldsymbol{\eta} \cdot \boldsymbol{k}\right) \right] \tag{51}$$

$$= \sum_{\boldsymbol{k} \in \Omega^K} P(\boldsymbol{k}) \exp\left(i\boldsymbol{\eta} \cdot \boldsymbol{k}\right) \tag{52}$$

can be expressed as through computing expectation values (also referred to as amplitudes)

$$x(\boldsymbol{\eta}) = \left\langle \Psi_{out} \left| e^{i \sum_{z=1}^{K} \eta_z \hat{N}_{\mathcal{K}_z}} \right| \Psi_{out} \right\rangle. \tag{53}$$

Given an orthonormal basis for the internal states of the photons $\{|\Phi_j\rangle\}$, we can expand any n-photon state with mode occupation numbers $\boldsymbol{s} = (s_1, \ldots, s_m)$ into the following orthonormal basis:

$$|\boldsymbol{d}(\boldsymbol{s}), \Phi_{\boldsymbol{a}}\rangle = \frac{1}{\sqrt{\mu(\boldsymbol{s})}} \prod_{j=1}^{n} \hat{b}_{d_j(\boldsymbol{s}), \Phi_{a_j}}^{\dagger} |0\rangle. \tag{54}$$

Here, $\boldsymbol{d}(\boldsymbol{s})$ is the mode assignment list, a vector of dimension n constructed as $\boldsymbol{d}(\boldsymbol{s}) = \oplus_{j=1}^{n} \oplus_{k=1}^{s_j} (j)$ [73]. The component $d_j(\boldsymbol{s})$ reflects the spatial mode occupied by the jth particle and is formally constructed by repeating s_j times the mode number j. In turn, $\boldsymbol{a}$ is also a vector of dimension n, whose indices a_j define that the internal state of the jth particle is $\left|\Phi_{a_j}\right\rangle$. Moreover, we also define $\mu(\boldsymbol{s}) = s_1! \cdots s_m!$.

The output state of a boson sampler with partially distinguishable input photons can then be written as

$$|\Psi_{out}\rangle = \sum_{\boldsymbol{d}(\boldsymbol{s}), \boldsymbol{a}} \alpha(\boldsymbol{d}(\boldsymbol{s}), \Phi_{\boldsymbol{a}}) |\boldsymbol{d}(\boldsymbol{s}), \Phi_{\boldsymbol{a}}\rangle \tag{55}$$

where

$$\alpha(\boldsymbol{d}(\boldsymbol{s}), \Phi_{\boldsymbol{a}}) = \left\langle \prod_{j=1}^{n} \hat{b}_{d_j(\boldsymbol{s}), \Phi_{a_j}} \middle| \Psi_{out} \right\rangle \tag{56}$$

We can now expand part of the right side of Eq. (53) as

$$e^{i \sum_{z=1}^{K} \eta_z \hat{N}_{\mathcal{K}_z}} |\Psi_{out}\rangle \tag{57}$$

$$= e^{i \sum_{z=1}^{K} \eta_z \sum_{j \in \mathcal{K}_z} \hat{n}_j} |\Psi_{out}\rangle \tag{58}$$

$$= \sum_{\boldsymbol{d}(\boldsymbol{s}), \boldsymbol{a}} e^{i \sum_{z=1}^{K} \eta_z \sum_{j \in \mathcal{K}_z} \hat{n}_j} \alpha(\boldsymbol{d}(\boldsymbol{s}), \Phi_{\boldsymbol{a}}) |\boldsymbol{d}(\boldsymbol{s}), \Phi_{\boldsymbol{a}}\rangle \tag{59}$$

$$= \sum_{k_1=0}^{n} \cdots \sum_{k_K=0}^{n} e^{i \sum_{z=1}^{K} \eta_z k_z} \sum_{\boldsymbol{d}(\boldsymbol{s}) | \boldsymbol{k}, \boldsymbol{a}} \alpha(\boldsymbol{d}(\boldsymbol{s}), \Phi_{\boldsymbol{a}}) |\boldsymbol{d}(\boldsymbol{s}), \Phi_{\boldsymbol{a}}\rangle \tag{60}$$

where $\boldsymbol{d}(\boldsymbol{s}) | \boldsymbol{k}$ is understood as the vectors $\boldsymbol{s}$ compatible with finding the partition output count $\boldsymbol{k}$. By applying $\langle \Psi_{out} |$, we obtain

$$x(\boldsymbol{\eta}) = \sum_{k_1=0}^{n} \cdots \sum_{k_K=0}^{n} e^{i \sum_{z=1}^{K} \eta_z k_z} \sum_{\boldsymbol{d}(\boldsymbol{s}) | \boldsymbol{k}, \boldsymbol{a}} |\alpha(\boldsymbol{d}(\boldsymbol{s}), \Phi_{\boldsymbol{a}})|^2 \tag{61}$$

Remark that

$$P(\boldsymbol{k}) = \sum_{\boldsymbol{d}(\boldsymbol{s}) | \boldsymbol{k}, \boldsymbol{a}} |\alpha(\boldsymbol{d}(\boldsymbol{s}), \Phi_{\boldsymbol{a}})|^2 \tag{62}$$

is the probability to find the photon count $\boldsymbol{k}$, by construction, which proves that we recover Eq. (52). Therefore, by simply applying a K-dimensional Fourier transform, we can obtain the binned output probabilities from the values of the $x_{\boldsymbol{\eta}}$'s by computing:

$$P(\boldsymbol{k}) = \frac{1}{(n+1)^K} \sum_{l_1=0}^{n} \cdots \sum_{l_K=0}^{n} x\left(\frac{2\pi \boldsymbol{l}}{n+1}\right) e^{-2\pi i \boldsymbol{l} \cdot \boldsymbol{k}/(n+1)} \tag{63}$$

A.2 Computation of the amplitudes as permanents

Let us derive the expression for $x(\boldsymbol{\eta})$ from Eq. (22). We recall that

$$x(\boldsymbol{\eta}) = \langle \Psi_{\text{out}} | e^{i\boldsymbol{\eta} \cdot \hat{\boldsymbol{N}}_\kappa} | \Psi_{\text{out}} \rangle \tag{64}$$

$$= \langle \Psi_{\text{in}} | \hat{U}^\dagger e^{i\boldsymbol{\eta} \cdot \hat{\boldsymbol{N}}_\kappa} \hat{U} | \Psi_{\text{in}} \rangle \tag{65}$$

Written in this form, $x(\boldsymbol{\eta})$ can be interpreted as the amplitude of $|\Psi_{\text{in}}\rangle$ staying intact through a virtual interferometer

$$\hat{V}(\boldsymbol{\eta}) = \hat{U}^\dagger e^{i\boldsymbol{\eta} \cdot \hat{\boldsymbol{N}}_\kappa} \hat{U}$$

To ease notation, we will denote $V(\boldsymbol{\eta})$ as simply V. Even though in the main part of the paper we consider an input state with one photon per mode, here we do a slightly more general derivation encompassing states of more than one photon per input mode, with mode occupation numbers given by a vector $\boldsymbol{r}$. Using the standard trick of inserting $\hat{V}\hat{V}^\dagger$ in between each creation operator, we obtain

$$x(\boldsymbol{\eta}) = \frac{1}{\mu(\boldsymbol{r})} \left\langle 0 \left| \prod_{j=1}^{n} \hat{a}_{d_j(\boldsymbol{r}), \phi_{d_j(\boldsymbol{r})}} \prod_{j=1}^{n} \sum_{k=1}^{m} V_{d_j(\boldsymbol{r}),k} \hat{a}^\dagger_{k, \phi_{d_j(\boldsymbol{r})}} \right| 0 \right\rangle \tag{66}$$

$$= \frac{1}{\mu(\boldsymbol{r})} \sum_{k_1,\ldots,k_n=1}^{m} \left(\prod_{j=1}^{n} V_{d_i(\boldsymbol{r}),k_i} \right) \left\langle 0 \left| \prod_{j=1}^{n} \hat{a}_{d_j(\boldsymbol{r}), \phi_{d_j(\boldsymbol{r})}} \hat{a}^\dagger_{k_j, \phi_{d_j(\boldsymbol{r})}} \right| 0 \right\rangle. \tag{67}$$

$$\tag{68}$$

Let's now compute the quantity between brackets by expanding the internal degrees of freedom into a basis

$$|\phi_j\rangle = \sum_\alpha^m c_{j\alpha} |\Phi_\alpha\rangle. \tag{69}$$

As there are only n photons, we could stop the sum to n as the other coefficients will be zero, but to ease the notation we keep the sum running up to m with the understanding that some coefficients are zero by construction. Thus

$$\left\langle 0 \left| \prod_{j=1}^{n} \hat{a}_{d_j(\boldsymbol{r}), \phi_{d_j(\boldsymbol{r})}} \hat{a}^\dagger_{k_j, \phi_{d_j(\boldsymbol{r})}} \right| 0 \right\rangle = \sum_{\alpha_1,\ldots,\alpha_n,\beta_1,\ldots,\beta_n=1}^{m} c^*_{d_1,\alpha_1} \ldots c^*_{d_n,\alpha_n} c_{d_1,\beta_1} \ldots c_{d_n,\beta_n} \tag{70}$$

$$\times \left\langle 0 \left| \prod_{j=1}^{n} \hat{a}_{d_j(\boldsymbol{r}), \Phi_{\alpha_j}} \hat{a}^\dagger_{k_j, \Phi_{\beta_j}} \right| 0 \right\rangle \tag{71}$$

This last quantity is given by

$$\left\langle 0 \left| \prod_{j=1}^{n} \hat{a}_{d_j(\boldsymbol{r}), \Phi_{\alpha_j}} \hat{a}^\dagger_{k_j, \Phi_{\beta_j}} \right| 0 \right\rangle = \sum_{\sigma \in S_n} \prod_{i=1}^{n} \delta_{d_i(\boldsymbol{r}),k_{\sigma_i}} \delta_{\alpha_i,\beta_{\sigma_i}}. \tag{72}$$

Plugging into the expression for $x(\boldsymbol{\eta})$ we obtain

$$x(\boldsymbol{\eta}) = \frac{1}{\mu(\boldsymbol{r})} \sum_{\alpha_1,\ldots,\alpha_n} \sum_{\sigma \in S_n} \prod_{j=1}^{n} V_{d_i(\boldsymbol{r}), d_{\sigma_i^{-1}}(\boldsymbol{r})} c^*_{d_i,\alpha_i} c_{d_i,\alpha_{\sigma_i^{-1}}} \tag{73}$$

where we use the fact that if $\sigma \in S_n$ then $\prod_{i=1}^{n} A_{i,\sigma_i} = \prod_{i=1}^{n} A_{\sigma_i^{-1},i}$ to rearrange the product. Next, we eliminate the sums over $\alpha's$ by recombining the coefficients $c_{j,\alpha}$ into scalar products of the wavefunctions of internal degrees of freedom as

$$\langle \phi_i | \phi_j \rangle = \sum_{\alpha} c_{i,\alpha}^* c_{j,\alpha} \equiv S_{ij}. \tag{74}$$

This leads to the expression

$$x(\boldsymbol{\eta}) = \frac{1}{\mu(\boldsymbol{r})} \sum_{\sigma \in S_n} \prod_{j=1}^{n} V_{d_i(\boldsymbol{r}),d_{\sigma_i^{-1}}(\boldsymbol{r})} S_{d_i(\boldsymbol{r}),d_{\sigma_i^{-1}}(\boldsymbol{r})} \tag{75}$$

By defining, as is conventional, the following reduced matrices to lighten the notation

$$\tilde{M}_{ij} = V_{d_i(\boldsymbol{r}),d_j(\boldsymbol{r})} \tag{76}$$

$$\tilde{S}_{ij} = S_{d_i(\boldsymbol{r}),d_j(\boldsymbol{r})} \tag{77}$$

we recover a compact expression

$$x(\boldsymbol{\eta}) = \frac{1}{\mu(\boldsymbol{r})} \sum_{\sigma \in S_n} \prod_{j=1}^{n} \tilde{S}_{i,\sigma_i} \tilde{M}_{i,\sigma_i} \tag{78}$$

$$= \frac{1}{\mu(\boldsymbol{r})} \text{perm} \left(\tilde{S} \odot \tilde{M} \right) \tag{79}$$

involving a single $n \times n$ permanent of the elementwise (Hadamard) product between two matrices: the first constructed from entries of the Gram matrix of the internal degrees of freedom and the second from the interferometer V. When considering an input of photons occupying modes $(1, \ldots, n)$, this reduces to the expression presented in the main text.

B Binned distributions of Fourier interferometers

B.1 Explicit probabilities for a single subset

In order to derive the expressions for the single-mode distributions obtained in Sec. 3.1, we first derive a general expression for the photon-number distribution in a single subset of output modes which may be of independent interest. First we note that we can write Eq. (17) as

$$V(\eta) = \left(\mathbb{1} + (e^{i\eta} - 1)H \right), \tag{80}$$

where the matrix H is defined in terms of the interferometer U as

$$H_{a,b} = \sum_{l \in \mathcal{K}} U_{l,a}^* U_{l,b}, \tag{81}$$

for some subset of interest denoted as $\mathcal{K}$. Let us define H_n as the $n \times n$ submatrix obtained from the first n rows and columns of H_n and $H_n' = S \odot H_n$. Using the fact that the diagonal elements of the S matrix are $S_{ii} = 1$ we can write

$$S \odot V_n(\eta) = S \odot \left(\mathbb{1}_n + (e^{i\eta} - 1)H_n \right) \tag{82}$$

$$= \mathbb{1}_n + (e^{i\eta} - 1)H_n' \tag{83}$$

Using an identity from Minc [74] (Chapter 2.2 exercise 5), the expression for the amplitudes $x(\nu_l)$, with $\nu_l = 2\pi l/(n+1)$ can be expanded as follows

$$x(\nu_l) = \text{perm} \left(\mathbb{1}_n + (e^{i\nu_l} - 1)H_n' \right) \tag{84}$$

$$= 1 + \sum_{a=1}^{n} c_a (1 - e^{i\nu_l})^a. \tag{85}$$

The coefficients c_a are given by

$$c_a = (-1)^a \sum_{\bar{w} \in Q_{a,n}} \text{perm}\left(H'_n[\bar{w}]\right), \tag{86}$$

where $Q_{a,n}$ denotes the set of all strictly ordered subsets of $\bar{w} \subset \{1, 2, \ldots, n\}$ containing a elements. Furthermore, $H'_n[\bar{w}]$ denotes an $a \times a$ submatrix of H'_n whose rows and columns are picked according to $\bar{w}$. Plugging in Eq. (85) into Eq. (8) we can obtain, after some manipulations, an explicit expression for the probabilities of observing k photons in the subset $\mathcal{K}$

$$P(k) = (-1)^k \sum_{a=k}^{n} \binom{a}{k} c_a. \tag{87}$$

This expression is of limited use since in general it is given by a sum of exponentially many permanents. However, in some particular cases (such as the one in Sec. 3.1) it can be used to obtain analytical results for the probabilities. Moreover, it can be used to recover some results that were previously obtained. In particular, it can be seen from this expression that the probability that all photons end up in the chosen subset, which can be seen as a generalized bunching probability, is given by

$$P(n) = \text{perm}\left(H'_n\right), \tag{88}$$

retrieving the result derived in [46]. It is also possible to see from the probability of not observing any photons in subset $\mathcal{K}$ is given by

$$P(0) = \text{perm}\left(\mathbb{1}_n - H'_n\right). \tag{89}$$

The latter expression is consistent with the fact that this probability is the same as that of seeing n photons in the complement of subset $\mathcal{K}$ and was considered in [40]. Therein, it was also shown that this quantity can be used to distinguish certain efficient classical simulation algorithms from boson samplers with partially distinguishable inputs for constant density $\rho = n/m$.

B.2 Single mode output distribution

We now apply our previous derivation to the problem of obtaining the photon number distribution in a single detector of the Fourier interferometer. Without loss of generality we choose the subset $\mathcal{K}_1 = \{1\}$. In this case, for an m-mode Fourier interferometer the H matrix takes the simple form

$$H_{i,j} = F^*_{1i} F_{1j} = \frac{1}{m}. \tag{90}$$

For indistinguishable photons we have that $H'_n = H_n$ and the coefficients c_a from Eq. (86) are given by

$$\begin{aligned}
c_a &= (-1)^a \sum_{\bar{w} \in Q_{a,n}} \text{perm}\left(H_n[\bar{w}]\right) \\
&= (-1)^a \sum_{\bar{w} \in Q_{a,n}} \frac{a!}{m^a} \\
&= (-1)^a \binom{n}{a} \frac{a!}{m^a},
\end{aligned}$$

where we used the fact that the number of possible strictly order subsets from $Q_{a,n}$ is $|Q_{a,n}| = \binom{n}{a}$. Plugging in Eq. (87), we obtain

$$P^B_k = \sum_{a=k}^{n} (-1)^{k+a} \binom{a}{k} \binom{n}{a} \frac{a!}{m^a}. \tag{91}$$

Although to our knowledge this expression has not been derived before, it is known from the general results of Ref. [59] that the single mode density matrix converges to a thermal state with an average number of photons n/m. In contrast, it can be seen that if we have fully distinguishable particles at the input, the distribution P_k^D is a binomial distribution. In this case, the probability that each particle appears in the first output mode is $1/m$ and hence the probability of observing k particles in this mode is

$$P_k^D = \binom{n}{k} \frac{1}{m^k} \left(1 - \frac{1}{m}\right)^{n-k}. \tag{92}$$

In the regime of constant density $\rho = n/m = \text{const.}$, this probability distribution tends to a Poisson distribution with mean ρ. It can be seen that these two distributions are sufficiently far apart and that it should be possible to distinguish them efficiently, i.e. with a number of samples growing polynomially in n, both in the constant density regime and in the collision free regime $m = n^2$.

B.3 Photon number distribution in the odd modes

We will now restrain ourselves to the case where the number of modes m is equal to the number of input photons n and thus consider exactly one photon per input mode. We will show that for the choice of a specific partition of the output modes, a striking difference exists between the distinguishable and indistinguishable cases by computing analytically those probability density functions. Namely, we will consider the special subset of the output modes consisting of the modes with odd numbers $\mathcal{K} = \{1, 3, 5, \ldots, n-1\}$ and will restrict ourselves to n even. In this scenario, we show that if the input are bosons the probability of seeing an odd number of photons in $\mathcal{K}$ is 0, whereas the probability of seeing an even number of photons follows a simple binomial distribution. We start by computing

$$x(\boldsymbol{\eta}) = \text{perm}\left(F^\dagger \Lambda(\eta) F\right). \tag{93}$$

We will first simplify the matrix product inside the permanent. We start by writing

$$\Lambda(\eta) = \mathbb{1} + \sum_{j \in \mathcal{K}} \left(e^{i\eta} - 1\right) |j\rangle\langle j|$$

so that

$$(F^\dagger \Lambda(\eta) F)_{ab} = \delta_{ab} + \left(e^{i\eta} - 1\right) \sum_{j \in \mathcal{K}} F_{aj}^\dagger F_{jb}$$

$$= \delta_{ab} + \left(e^{i\eta} - 1\right) \sum_{j \in \mathcal{K}} \frac{1}{n} \exp\left(\frac{2\pi i}{n}(j-1)(a-b)\right)$$

$$= \delta_{ab} + \left(e^{i\eta} - 1\right) \frac{1}{n} \sum_{k=0}^{n/2-1} \exp\left(\frac{4\pi i}{n}k(a-b)\right)$$

Now, let's note that the sum is a geometric series

$$\sum_{k=0}^{n/2-1} \exp\left(\frac{4\pi i}{n}k(a-b)\right) =$$

$$= \begin{cases} \frac{n}{2}, & \text{if } |a-b| = 0 \text{ or } \frac{n}{2} \\ 0 & \text{otherwise} \end{cases}$$

given that a, b are integers. Thus

$$(F^\dagger \Lambda(\eta) F)_{ab} = \begin{cases} \frac{e^{i\eta}+1}{2} & \text{if } a = b \\ \frac{e^{i\eta}-1}{2} & \text{if } |a-b| = \frac{n}{2} \\ 0 & \text{otherwise} \end{cases} \tag{94}$$

In other words,

$$V(\eta) = F^{\dagger}\Lambda(\eta)F = \tag{95}$$

$$= \mathrm{circ}\left(\frac{e^{i\eta}+1}{2}, 0, \ldots, 0, \frac{e^{i\eta}-1}{2}, 0, \ldots, 0\right) \tag{96}$$

where $\mathrm{circ}(\vec{v})$ denotes the circulant matrix whose first row is $\vec{v}$. Now let's see that this form allows us to compute the permanent analytically. We have that

$$\mathrm{perm}\left(V(\eta)\right) = \sum_{\sigma \in S_n}\prod_i V(\eta)_{i\sigma(i)}. \tag{97}$$

Due to the large number of zeros in $V(\phi)$ the sum over all permutations can be greatly simplified. In fact, the only permutations that lead to a non-zero product $\prod_i V(\eta)_{i\sigma(i)}$ are those such that $\sigma(i) = i$ or $\sigma(i) = i + n/2 \pmod{n}$. Such permutations can be written as a product of disjoint permutations

$$\sigma = \gamma_1\gamma_2\ldots\gamma_{n/2} \tag{98}$$

where $\gamma_i \in \{e, \tau_i\}$, e is the identity permutation and τ_i is the transposition that switches elements i with element $i + n/2 \pmod{n}$. Hence we can write

$$\mathrm{perm}\left(V(\phi)\right) = \sum_{\gamma_1}\sum_{\gamma_2}\cdots\sum_{\gamma_{n/2}}\prod_{i=1}^{n/2} V(\eta)_{i,\gamma_i(i)}V(\eta)_{\frac{n}{2}+i,\gamma_i(\frac{n}{2}+i)} \tag{99}$$

$$= \prod_{i=1}^{n/2}\sum_{\gamma_i} V(\eta)_{i,\gamma_i(i)}V(\eta)_{\frac{n}{2}+i,\gamma_i(\frac{n}{2}+i)} \tag{100}$$

$$= \prod_{i=1}^{n/2}\left(\left(\frac{e^{i\eta}+1}{2}\right)^2 + \left(\frac{e^{i\eta}-1}{2}\right)^2\right) \tag{101}$$

$$= \left(\left(\frac{e^{i\eta}+1}{2}\right)^2 + \left(\frac{e^{i\eta}-1}{2}\right)^2\right)^{n/2} \tag{102}$$

$$= \frac{1}{2^{n/2}}(e^{2i\eta}+1)^{n/2} \tag{103}$$

In the second step we used the fact that the γ_i's are disjoint permutations to switch the product with the sum. Finally, by applying Eq. (8) we obtain

$$P_k^B = \begin{cases} 0 & \text{if } k \text{ is odd} \\ \frac{1}{2^{n/2}}\binom{n/2}{k} & \text{if } k \text{ is even} \end{cases} \tag{104}$$

Comparatively, we can find the counterpart of this expression for distinguishable particles following a simple statistical argument. For an unbiased interferometer (such as the Fourier interferometer), the probability that a given particle falls into the partition A is $q = |A|/m$ such that the probability to find j photons inside the partition is given by the binomial distribution

$$P_j^D = \binom{n}{j}q^j(1-q)^{n-j}$$

which, in the case discussed above, reduces to

$$p_j^D = \frac{1}{2^n}\binom{n}{j}.$$

The striking difference between the two cases is represented in Fig. 3. It is possible to see that the TVD between these two probability distributions tends to a constant as n goes to infinity.

C   Further numerical investigations

In this appendix, we present extra numerical evidence to support the claims made in the main text. All plots of this paper are reproducible by using the associated project, BOSONSAMPLING.JL which contains the exact code needed for their execution. The package also allows to explore different parameter ranges.

C.1   Default partition choice

Unless otherwise specified, the choice of partition is made as follows. We divide all physical output modes m in K bins. For a lossy boson sampler, all the environment modes are always grouped together in a single bin. If K divides m, each physical bin is contains m/K modes. Otherwise, we can decompose $m = p(K - 1) + q$. Then the first $K - 1$ bins are of size p while the remaining one of size q. We always take consecutive modes in building each bin. This choice is not very important as we approximately obtain the Haar average value of the probabilities by computing them for several Haar random unitaries. We remark that for the Haar averaged probabilities only the size of each bin matters and not the specific choice of partitionas. Arbitrary arrangements for the bins, which may be useful to test particular interferometers, can readily be simulated with the provided codes.

C.2   Variability of the TVD

In Fig. 10, we show the behaviour of the variance bars of Fig. 5, showing how the TVD converges to its mean value.
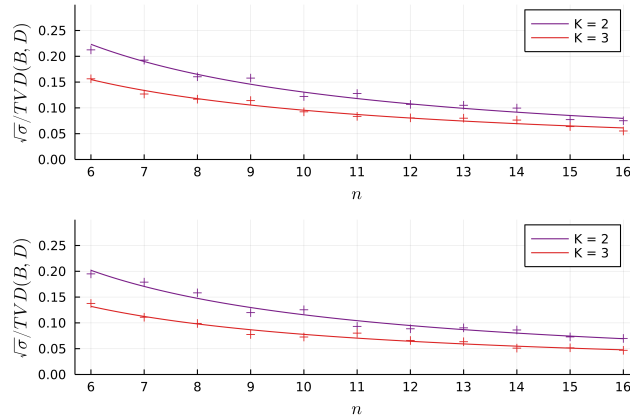

Figure 10: **Variability of the TVD with system size.** We plot the same parameters as in Fig. 5 but instead of looking at the TVD we display its coefficient of variation (CV), defined as the standard deviation divided by the mean. To a good approximation, it follows a power law of form $CV \propto n^{-1}$ in all plots with a slightly varying prefactor of order one.

C.3   TVD with partial distinguishability

In Fig. 11, we show here that the TVD decreases in a similar fashion than described in Fig. 5 when comparing indistinguishable particles to nearly indistinguishable ones. This is an important observation as this gives strong evidence for the scalability of our validation protocol even in real-world use cases. In addition, the fact that the distance between distributions may allow for the estimation of the distinguishability levels in the experiment, by finding the distinguishability parameter that gives the highest agreement with the experimental data.
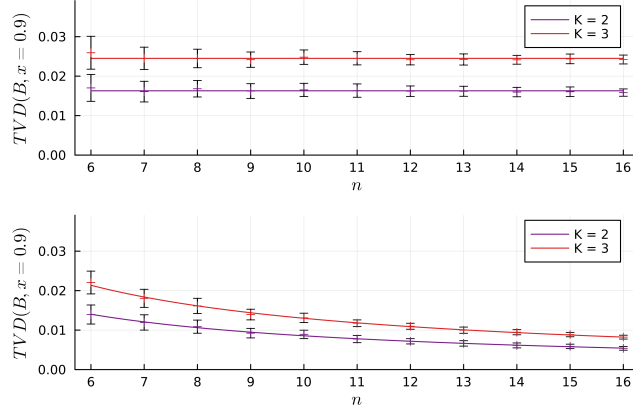
Figure 11: **TVD between bosonic and nearly indistinguishable particles.** This figure shows that the results shown in Fig. 5 hold even when comparing indistinguishable photons to nearly indistinguishable ones. All parameters are identical to those from Fig. 5, except from the fact that we compare ideal bosons to partially distinguishable ones with $x = 0.9$ (instead of $x = 0$). We confirmed this trend in multiple numerical trials with other values of distinguishability parameter x.

C.4 The role of boson density

Our numerical investigations highlight the preponderant role of boson density $\rho = n/m$ in the behavior of the TVD between binned distributions of different types of input particles. This dependency is shown in Fig. 12 for the TVD between distinguishable and indistinguishable photons. These power law fits at low number of photons are useful to understand the scalability of the validation technique we consider for larger experiments. Overall, we find that

$$\mathrm{tvd}(B, x) \approx c(K, x)\rho \tag{105}$$

with $c(2, 0) \approx 0.41$, $c(3, 0) \approx 0.67$. We see how the above equation holds when using different values of n in Fig. 13. While some slight variation is found in small sized systems, the power law seems to become nearly independent of n in larger systems. Likewise, we also check that it holds to a good approximation when comparing indistinguishable photons to some with partial distinguishability x in Fig. 14.
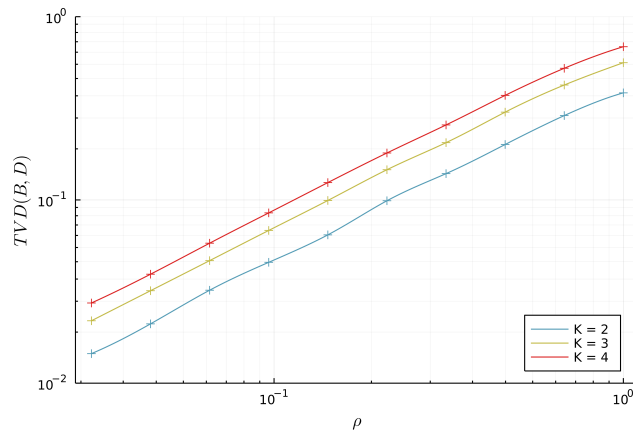

Figure 12: **Effect of boson density.** The Haar-averaged TVD between bosonic and distinguishable particles (asymptotically) depends only on the boson density ρ, see (44). We show how it evolves in the example of $n = 10$ and $m = 10, \ldots, 300$ with a number of subsets $K = 2, 3, 4$. Real-world experiments typically have a high boson density, for instance [11] has $n = 20$ and $m = 60$ thus $\rho = 1/3$. Note that the hardness of boson sampling is shown in a very dilute case of $m \propto n^6$ and conjectured in the "no collision" regime $m \propto n^2$ [3].
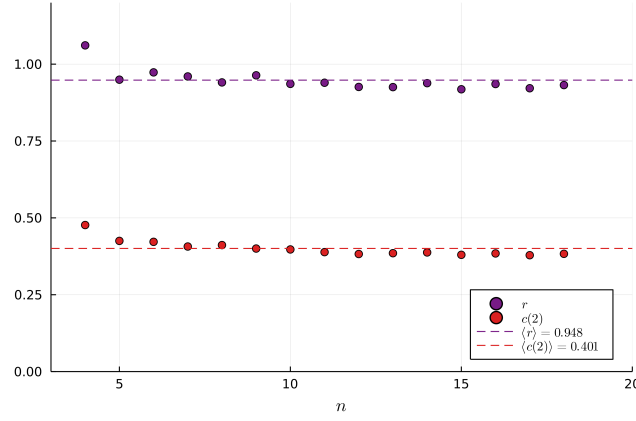
Figure 13: **Validity of the power law.** In Fig 12 we claim that the density is the prime factor modifying the TVD. Here we provide a justification for this by looking at the power-law fit $\mathrm{tvd}(B, D) = c(K, 0)\rho^r$ with K the number of subsets in the partition. We plot the coefficients for various values of n. The values of m are such that the density ranges from $1, \ldots, 0.03$. 10 values of m are taken and are equally distributed in logarithm (such as seen on Fig. 12). For each point, 100 iterations are performed.
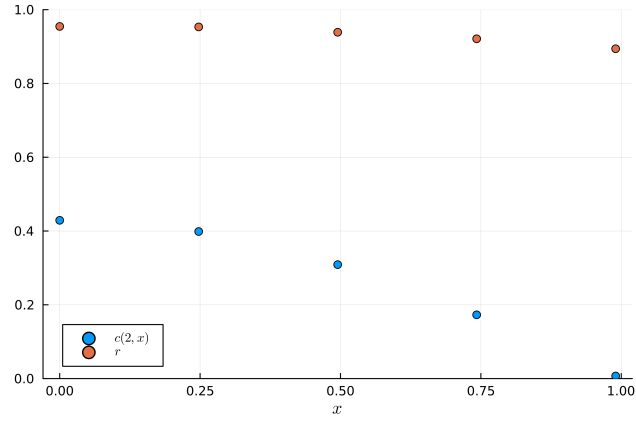

Figure 14: **Validity of the power law for various values of x..** We show the evolution of the coefficient $c(2, x)$ as well as the power r for $n = 8$ over various values of x. We see small variations in the r but the overall behaviour seems to hold well. For each point, 100 iterations are performed. The last point on the right has value $x = 0.99$.

C.5 Dependency of number of samples on density

In this subsection, we give some details about Eq. 50. We use as ansatz the following equation describing the dependence of the number of samples to validate the ideal boson sampler hypothesis against a model with partial distinguishable photons:

$$n_s \approx \frac{d(K, x)}{\rho^r} \tag{106}$$

The value of $d(K, x)$ depends on the partition size and the value of partial distinguishability. For $x = 0$ we obtain $d(2, 0) = 17.9$ and the power fit gives an exponent of $r = 2.53$ when using $n = 10$ input photons. Our numerical evidence suggests that the fit becomes better and better as the density decreases, as finite size/density effects may be at play. The same can be said for the number of input photons. We also obtained preliminary evidence suggesting that the power law holds for other values of partial distinguishability, although a complete analysis is out of the scope of the paper. More precise numbers can be obtained using the publicly available codes (see Code Availability).

C.6 Spoofing

An open question left by our work is whether an adversary can spoof the validation test based on binned distributions and successfully pretend to generate samples from an ideal boson sampler, for example, by constructing a clever mock-up sampler based on an efficient classical algorithm. In this subsection, we study whether an adversary can easily reproduce the binned distributions obtained from an ideal boson sampler, without the knowledge of the choice of partitions used to validate the sampler. To do so, we quantify numerically in two different ways how much the binned distribution changes over different random choices of the bins, showing there is a significant variability in the distributions between different choices.

First we consider a unitary drawn at random according to the Haar measure and compute the total variation distance between two random partitions with the same number of bins, but not necessarily the same sizes for each bin. An example for $m = 4$ could be to consider subsets of modes $\{1, 2\}, \{3, 4\}$ for the first partition while taking $\{1\}, \{2, 3, 4\}$ for the second. In Fig. 15 we plot the total variation distance between two randomly picked distributions, averaging over random partition choices and Haar random unitaries. The large values of the TVD observed are explained by the fact that there are large differences of the distributions with different bin sizes, as can be seen from the asymptotic expression in Eq. (43). For example, in a bin of size L we would expect to see on average Ln/m photons.

In our second example, we want to illustrate that even an adversary who is able to produce samples that correctly reproduce the Haar averaged properties of the binned distributions (e.g. by exploiting knowledge of Eq. (43)) would not be able to spoof the validation test. To do so, in Fig. 16, we compare only binned distributions from randomly chosen subsets of fixed sizes. More precisely, in this second case, we divide the m modes with K equally sized subsets (up to the remainder of m/K). In this case, for $m = 4$ we could for instance choose $\{1, 2\}, \{3, 4\}$ and $\{3, 2\}, \{1, 4\}$ (but not $\{1\}, \{2, 3, 4\}$ as the first subset does not have the same size as $\{1, 2\}$). In each case, we compute the TVD between the binned distributions resulting from these random bin choices, averaging over different random bin choices of equal sizes and also over Haar random unitaries.

Even in this latter scenario, we see that the TVD between distribution decreases slowly with the number of photons. This suggests that with a polynomial number of samples, one could identify differences between a mock-up sampler that is able to reproduce the right Haar averaged properties of binned distributions and an ideal boson sampler. It would be interesting to analytically prove a polynomial decrease of the TVD in this scenario, a question we leave for future work.
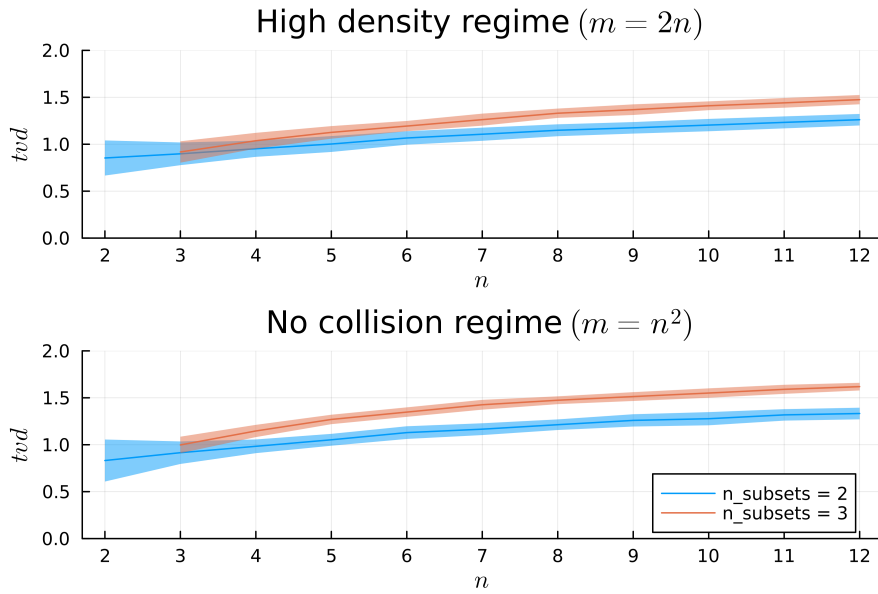
Figure 15: **Distance between randomly selected partitions with subsets of possibly different sizes.** The same procedure is repeated 100 random bin choices for each unitary. We then average, for each value of n, over 100 Haar random unitaries. Two scenarios are presented: a high density regime, where $m = 2n$ and a no collision regime with $m = n^2$. For both cases, we study the effect of choosing 2 or 3 bins.
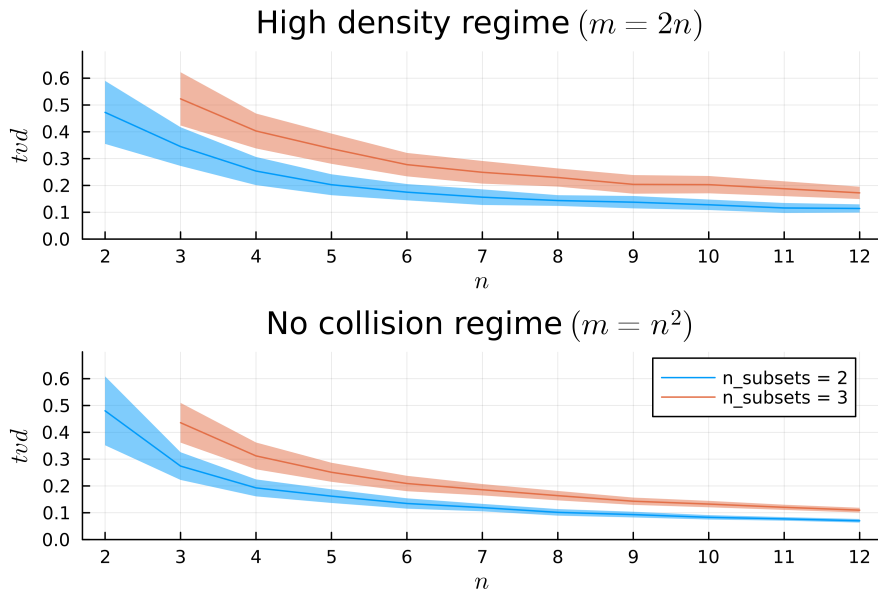

Figure 16: **Total variation distance between randomly selected partitions containing subsets of equal sizes.** The same procedure is repeated 100 random bin choices for each unitary. We then average, for each value of n, over 100 Haar random unitaries. Two scenarios are presented: a high density regime, where $m = 2n$ and a no collision regime with $m = n^2$. For both cases, we study the effect of choosing 2 or 3 bins.