ULB Faculté des Sciences

Distinguishability in Quantum Multiphoton Interference

From Bunching Phenomena to the Validation of Boson Sampling

Thesis presented by Benoît SERON

in fulfilment of the requirements of the Doctorate in Physics ("Doctorat en Physique")

Année académique 2022-2023

Supervisor : Nicolas CERF

Co-supervisor : Petr TINYAKOV

Centre for Quantum Information and Communication (QuIC)

Thesis jury :

Nathalie VAECK (Université libre de Bruxelles, Chair)
Magdalena STOBIŃSKA (University of Warsaw)
Andreas BUCHLEITNER (Albert-Ludwigs-Universität Freiburg)
Leonardo NOVO (International Iberian Nanotechnology Laboratory)
Petr TINYAKOV (Université libre de Bruxelles)
Nicolas CERF (Université libre de Bruxelles)

fnrs FREEDOM TO RESEARCH

Abstract

This thesis investigates multiphoton quantum interference with a particular emphasis on photonic partial distinguishability stemming from factors such as timelags or polarization mismatches between the photons. We explore the gray area between the behaviour of distinguishable classical particles and ideally indistinguishable bosons in large linear interferometers, with practical implications for photonic quantum computers – especially boson samplers – as well as fundamental consequences on the many-body physics of bosonic systems.

Our research leverages recent findings in the mathematical theory of matrix permanents to unravel new aspects of interference phenomena. In particular, we exhibit a surprising counterexample to the rule of thumb that "identical bosons bunch most", which can be observed with 7 photons in a 7-mode linear interferometer.

We develop robust techniques to validate boson samplers – that is, to verify the incompatibility with classical samplers – by analyzing the photon distribution among partitions of the output modes. This innovative method extends existing validation tests, providing a comprehensive approach to assessing boson sampling experiments.

An extra contribution of this thesis is the development of BosonSampling.jl, an open-source, high-performance software package for multiphoton interferometry simulations, featuring versatile tools for boson sampling tasks and a suite of validation tools in order to rigorously evaluate boson samplers in the presence of partial distinguishability and other noise sources.

Keywords: Multiphoton interference, Photonic partial distinguishability, Boson sampling, Validation tests, Quantum computing, Matrix permanents

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my advisor, Prof. Nicolas Cerf, for his unwavering support, guidance, and encouragement throughout this journey. I am extremely fortunate to have had the opportunity to learn from such an accomplished and inspiring mentor.

I am also profoundly grateful to my co-advisor, Prof. Petr Tinyakov, for his invaluable insights and for helping me become a better physicist through our thought-provoking discussions. His expertise and knowledge have truly enriched my doctoral experience.

I owe a debt of gratitude to Dr. Leonardo Novo, who, as a postdoc in our group, provided me with daily coaching and support. I could not have completed this thesis without his assistance and dedication.

My heartfelt appreciation goes to my partner, Ursula, for her love, understanding, and unwavering support. Thank you for standing by my side throughout this journey.

I would like to extend my deepest thanks to my parents for their love, guidance, and constant encouragement throughout my life. Their support has been a pillar of strength for me.

My friends have also played a vital role in this journey, providing me with support and insights. In particular, I would like to thank Romain Ruzziconi, Sacha Ferrari, and Manon Gortz for their friendship and contributions to my academic and personal growth.

I am grateful to all the members of our research group, as well as the colleagues I have had the pleasure of meeting during my travels. Their camaraderie and collaboration have made this experience truly rewarding.

My sincere appreciation goes to my thesis follow-up committee, for their valuable feedback and guidance. I would like to give a special mention to Prof. François Léo for his support during the challenging first year of my thesis.

Lastly, I would like to extend my gratitude to Prof. Mustapha Tlidi for his presence and invaluable assistance in my development as a researcher.

I would like to express my gratitude to those whom I fondly refer to as Lucy and her sisters, as they have guided me on a series of enlightening adventures, providing me with the inspiration and motivation to explore uncharted territories both in my personal and academic life. The impact of these experiences cannot be understated, and they have undoubtedly contributed to my growth as an individual and a researcher.

I am truly grateful for the opportunity to have worked with and learned from such incredible individuals, and I dedicate this thesis to them.

Scientific acknowledgments and funding

In this thesis, I would like to extend my gratitude to the following individuals for their valuable contributions to each of the three articles included:

For the article "BosonSampling.jl: A Julia package from quantum multiphoton interference," I thank Dr. Fulvio Flamini for critical reading of the manuscript, Dr. Christoph Hotter and Dr. David Plankensteiner for discussions and providing the LaTeX template to display Julia code, Leonardo Banchi for discussions and his contribution to the Permanents.jl package, and Dr. Leonardo Novo and Prof. Nicolas Cerf for discussions.

For the article "Efficient validation of Boson Sampling from binned photon-number distributions," I thank Dr. Fulvio Flamini, Prof. Paul Valiant, and Prof. Gregory Valiant for valuable discussions.

For the article "Boson bunching is not maximized by indistinguishable particles," I thank Prof. Stephen Drury for useful correspondence, Dr. Fulvio Flamini, and Prof. Valery Shchesnovich for valuable discussions.

I am grateful for the financial support that has made this research possible. I am a Research Fellow of the Fonds National de la Recherche Scientifique – FNRS. Prof. Nicolas Cerf acknowledges support from the Fonds de la Recherche Scientifique – FNRS under Grant No T.0224.18 and from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 956071. Dr. Leonardo Novo acknowledges support from the Fonds de la Recherche Scientifique – FNRS, from FCT-Fundação para a Ciência e a Tecnologia (Portugal) via the Project No. CEECINST/00062/2018, and from the European Union's Horizon 2020 research and innovation program through the FET project PHOQUSING ("PHOtonic QUantum SamplING machine" - Grant Agreement No. 899544).

During my research, I had the opportunity to travel thanks to the AppQInfo MSCA ITN support from Antoine Restivo. This has allowed me to meet various colleagues and establish collaborations.

I would like to express my appreciation to my colleagues, and in particular, Antoine Restivo and Léo Pioge, for helping me with code tasks.

Once again, I extend my deepest gratitude to all who have contributed to my research journey and the completion of my thesis.

Contents

Part I
Introduction

This thesis is an exploration of the complex and intriguing world of multiphoton interference, with a particular focus on the role of photonic partial distinguishability - the incomplete overlap between the wavefunctions of the photons which may arise from timelags between them or mismatched polarization.

Intriguingly, the concept of partial distinguishability introduces a certain 'gray area' into our understanding of quantum interference, straddling the boundary between the behavior of classical particles and indistinguishable bosons. Indistinguishable bosons behave in ways that can only be explained through quantum mechanical effects. The paradigmatic Hong-Ou-Mandel experiment - where two identical bosons will always bunch together at the output of a 50:50 linear coupler - is the simplest example. As particles become distinguishable, this bunching effect declines, which can be seen as a quantum-to-classical transition. However, this transition becomes much more complex when considering larger systems. This territory is mostly unexplored and turns out to be even more challenging and surprising than standard multiphoton interference, which already is a complex area in itself.

At the heart of this analysis of quantum interference lies the fruitful interplay between mathematical theory and physical properties. The theory of matrix permanents, in particular, has emerged as a powerful tool for examining and understanding the behavior of multiphoton interference. The research presented in this thesis exploits this knowledge to get new insights into interference phenomena with partially distinguishable photons, while simultaneously offering intriguing counterexamples that push the boundaries of our existing understanding. Conversely, we use tools and intuition stemming from quantum optics to extend mathematical knowledge on matrix permanents.

In the context of quantum computation, a major effect of partial distinguishability is to reduce the potential computational advantage that is expected from photonic quantum computers, such as boson samplers. In the time span of a decade, these boson samplers went from a theoretical idea to large-scale optical devices with claims of quantum computational advantage by multiple teams. However, as an intrinsic side effect of the intractability of the task performed, these claims are hard to justify without proper validation techniques. Indeed, even modest levels of distinguishability between particles can make a boson sampler easy to simulate on a classical computer, defeating their purpose.

In this regard, our work in this thesis has developed robust techniques to substantiate the correct functioning of boson samplers. We present a validation method, grounded on how photons distribute among partitions of the output modes. The versatility of our approach allows it to encompass and extend upon previous validation tests based on bunching phenomena, marginal distributions, and some suppression laws. We develop a new formalism that allows us to approximate the binned distributions classically in a computationally efficient manner. Through this, we are able to analyze and assess

the performance of boson samplers under various noise models, including those with partially distinguishable photons. We support our findings with rigorous analytical evidence and extensive numerical simulations. Importantly, our validation method is also able to process experimental data with photon losses, making the validation process more practical and efficient, especially in the early stages of "debugging" a boson sampler.

In the process of this exploration, an original computational tool was conceived — BosonSampling.jl, an open-source software package designed for high-performance simulations of multiphoton interferometry. Conceived in Julia, this versatile software provides a range of features for tasks related to boson sampling and interferometry, delivering an adaptive and extendable framework that caters to the evolving demands of this very active field. The package also encompasses a robust suite of validation tools, enabling researchers to rigorously assess the performance of boson samplers and ensure accurate results in the presence of partial distinguishability and other sources of noise.

As we progress through this thesis, readers will be introduced to the complexities of photonic interference, offering a detailed examination of partial distinguishability and its implications on experiments. We will provide insight into the development and functionality of the BosonSampling.jl package, focusing on its utility for high-performance simulations and the integral validation tools it provides. Finally, we will consider the broader implications of these findings for the field of quantum computing, particularly in relation to photon-based systems such as boson samplers. This work, we hope, contributes to the ongoing dialogue within this challenging and rapidly evolving field.

Works covered in this dissertation

This dissertation addresses a subset of the articles published during this thesis:

- *Boson bunching is not maximized by indistinguishable particles*, with L. Novo and N. J. Cerf, accepted in Nature Photonics, arXiv 2203.01306 [quant-ph]
- *Efficient validation of Boson Sampling from binned photon-number distributions*, with L. Novo, A. Arkhipov, and N. J. Cerf, submitted to Phys. Rev. X, arXiv 2212.09643 [quant-ph]
- *BosonSampling.jl: a Julia package for multiphoton interferometry*, with A. Restivo, submitted to Quantum, arXiv 2212.09537 [quant-ph]

A detailed introduction to each of these works will be given after presenting the essential theoretical ideas that are needed to understand the context and questions that these works answer.

Other works realised during this thesis

The following work was published on the arXiv preprint server but will not be discussed in this dissertation:

- *Programmable multi-photon quantum interference in a single spatial mode*, with L. Carosini, V. Oddi, F. Giorgino, L. M. Hansen, J. S. Piacentini, T. Guggemos, I. Agresti, J. C. Loredo, and P. Walther, submitted to Science Advances, arXiv 2305.11157 [quant-ph]

The following work was completed during this thesis but is not discussed in the present dissertation as it concerns a different field:

- *Dynamical elastic contact of a rope with the ground*, with G. Kozyreff, published in Physical Review Research 3, L022026 (2021), https://doi.org/10.1103/PhysRevResearch.3.L022026.

The following works also concern multiphoton interferometry and have led to significant progress, but have not yet been submitted to the arXiv preprint server.

- *Validation of Gaussian Boson Sampling*, with G. Bressanini, L. Novo, N. J. Cerf, and M. Kim. This works by applying similar techniques as those found in "Efficient validation of Boson Sampling from binned photon-number distributions".
- *Enhanced bunching of nearly indistinguishable bosons* with L. Pioge, L. Novo, N. J. Cerf.
- Experimental demonstration of the effects shown in "Efficient validation scheme for multiphoton interference" with L. Novo, N. J. Cerf, J. Renema, M. Anguita.

Part II
Theoretical grounds

1 Multiphoton interference

This section provides a concise overview of the fundamental theoretical concepts of multiphoton interference, focusing on the essential aspects required to comprehend the content of this thesis. It is important to note that this only touches the basic notions required to understand the papers presented in the next part of this manuscript. Readers seeking a more comprehensive understanding of the subject matter are encouraged to consult additional references for a thorough discussion ([1, 2, 3, 4, 5, 6, 7, 8, 9]).

The structure of this section is as follows: first, we give a brief reminder of the quantization of the electromagnetic field, underlying the formalism that will be used in the rest of this manuscript. Next, we discuss as an example model the interference of two particles. This allows us to present notions that translate to the more complicated case of multiparticle interference. After that, we study in generality the mathematics of generic interference. Special emphasis is put on the notion of partial distinguishability, which will play an important role in the key publication resulting from this thesis, "*Boson bunching is not maximized by indistinguishable particles*", presented in Section 4. We also show how the probabilities of different events are computed. This allows us to introduce the important matrix function known as the matrix permanent, a direct analog of the matrix determinant (and the Slater determinant in quantum physics). We will discuss this object in great detail through this thesis. Indeed, all behaviour studied is understood from the intricate properties of the permanent. These properties and their implications will be studied in more depth in the following sections of this introductory part of the thesis.

1.1 Quantization of the electromagnetic field

Quantum optics is an important field of study in quantum physics, focusing on the interaction of light with matter at the quantum level. One of the key aspects of quantum optics is the quantization of the electromagnetic field, which allows for the description of light as a collection of discrete quanta or photons. We introduce the quantization of the electromagnetic field through the colloquial example of a one dimensional cavity with perfectly conducting boundaries in the planes $z = 0$ and $z = L$. In vacuum, the Maxwell equations give

$$E_x(z,t) = \sqrt{\frac{2\omega^2}{V\epsilon_0}} q(t) \sin(kz) \tag{1}$$

$$B_y(z,t) = \frac{\mu_0\epsilon_0}{k} \sqrt{\frac{2\omega^2}{V\epsilon_0}} \dot{q}(t) \cos(kz) \tag{2}$$

setting the $x-$axis along the polarization of the electric field and defining ω the pulsation and $k = \omega/c$ the wave number and with V a normalization. Only the frequencies given by $\omega_m = cm\pi/L$ with $m = 1, 2, \ldots$ are allowed in this example from the boundary conditions. Let us focus on a single allowed frequency. The Hamiltonian of the electromagnetic field becomes in this case

$$H = \frac{1}{2}(p^2 + \omega^2 q^2) \tag{3}$$

where we identified $p = \dot{q}$. It is quantized by applying the correspondence rules of replacing the canonical variables q, p by their operators. We identify

$$\hat{a} = (2\hbar\omega)^{-1/2}(\omega\hat{q} + i\hat{p}) \tag{4}$$

$$\hat{a}^\dagger = (2\hbar\omega)^{-1/2}(\omega\hat{q} - i\hat{p}) \tag{5}$$

In turn, we recover the Hamiltonian of a harmonic oscillator

$$\hat{H} = \hbar\omega(\hat{a}^\dagger\hat{a} + \frac{1}{2}) \tag{6}$$

The product $\hat{n} = \hat{a}^\dagger\hat{a}$ is called the number operator. We write

$$|n\rangle = \frac{(\hat{a}^\dagger)^n}{\sqrt{n!}}|0\rangle \tag{7}$$

This state is an eigenvector of $\hat{H}$, with energy $\hbar\omega(n + 1/2)$. The action of $\hat{a}^\dagger$ can therefore be understood as adding a quanta $\hbar\omega$ of energy, while $\hat{a}$ removes one. While many other states are used in quantum optics, we will be mostly interested in this thesis in the states of type $|n\rangle$, colloquially referred to as *number states*. They contain a fix number of photons, n.

1.2 Interference of two particles

We begin by studying the interference of two particles through a *beam splitter*. We follow the presentation of [1, 2, 3]. When going through a beam splitter, each particle has a probability R of being reflected and $T = 1 - R$ of being transmitted. We can find the two following configurations in the detectors: either a *coincidence*, where each detector sees one particle, or a bunched event, where the two particles are observed by the same detector.

1.2.1 Distinguishable particles

Consider first two classical particles ("balls"). The following is also valid for elementary particles (behaving quantumly) as long as they can be deterministically distinguished, for instance, if their polarization is orthogonal. We consider these two particles to be identical. Although they can be traced throughout the entirety of their motion, we will act as if our detectors are blind to the origin of the particles. That is, these particles are

fundamentally distinguishable but we consider them *observationally* indistinguishable. Given their distinguishability, the particles behave independently to the presence of the other. Therefore, overall probabilities are computed by summing the probabilities of each particle.

The probability to find a coincidence is given by $p_D(1,1) = T^2 + R^2$ while a bunched event happens with probability $p_D(0,2) = p_D(2,0) = TR$.

1.2.2 Indistinguishable particles

Consider now identical particles, such as photons or electrons. These particles cannot be distinguished: unlike in the classical case, one cannot attach a label to the particles. They are *fundamentally indistinguishable*. Let us compute their evolution through the beam splitter. Their initial state is given by

$$|\Psi_{in}\rangle = \hat{a}_1^\dagger \hat{a}_2^\dagger |0\rangle \tag{8}$$

where $\hat{a}_i^\dagger$ is the creation operator in mode i and $|0\rangle$ the vacuum. The evolution operators evolve as

$$\hat{a}_j^\dagger \to \hat{U}\hat{a}_j^\dagger \hat{U}^\dagger = \begin{cases} \sqrt{T}\hat{b}_1^\dagger + i\sqrt{R}\hat{b}_2^\dagger \text{ if } j = 1 \\ i\sqrt{R}\hat{b}_1^\dagger + \sqrt{T}\hat{b}_2^\dagger \text{ if } j = 2 \end{cases} \tag{9}$$

by the action of the beam splitter. Restricting to a balanced beam splitter ($R = T = 1/2$), the state at the output of the beam splitter is thus

$$|\Psi_{out}\rangle = \frac{1}{2}\left(i((\hat{b}_1^\dagger)^2 + (\hat{b}_2^\dagger)^2) + [\hat{b}_1^\dagger, \hat{b}_2^\dagger]\right)|0\rangle \tag{10}$$

with $[a,b] = ab - ba$ the commutator. For bosons, whose creation operators commute, we are left with

$$|\Psi_{out,B}\rangle = \frac{i}{2}\left((\hat{b}_1^\dagger)^2 + (\hat{b}_2^\dagger)^2\right)|0\rangle \tag{11}$$

while for fermions, whose creation operators anticommute,

$$|\Psi_{out,D}\rangle = \hat{b}_1^\dagger \hat{b}_2^\dagger |0\rangle \tag{12}$$

The possible output patterns can be read out from these states: bosons will always bunch into a single mode, and will never lead to coincidences. The opposite stands true for fermions (as is expected by Pauli's exclusion principle). The bunching and absence of coincidences was first demonstrated by Hong-Ou-Mandel [10].

1.2.3 Partial distinguishability

Photons described above were assumed to be indistinguishable. In practice, this is hard to realise experimentally: photons sent in different spatial modes may have slightly different polarizations, or arrival times. This leads to an imperfect interference, as we will describe now. These photons are said to be *partially distinguishable*. Photons are described by their input mode i, as well as an *internal wave function* $|\phi_i\rangle$. The latter

regroups all information not contained in the spatial mode occupation. Therefore we can write

$$|\Psi_{in}\rangle = \hat{a}^\dagger_{1,\phi_1}\hat{a}^\dagger_{2,\phi_2}|0\rangle \tag{13}$$

A core assumption is that both the network and detectors are blind to the internal degrees of freedom. This means that creation operators evolve as

$$\hat{a}^\dagger_{i,\phi} \to \hat{U}\hat{a}^\dagger_{i,\phi}\hat{U}^\dagger = \begin{cases} \sqrt{T}\hat{b}^\dagger_{1,\phi} + i\sqrt{R}\hat{b}^\dagger_{2,\phi} \text{ if } j=1 \\ i\sqrt{R}\hat{b}^\dagger_{1,\phi} + \sqrt{T}\hat{b}^\dagger_{2,\phi} \text{ if } j=2 \end{cases} \tag{14}$$

regardless of the internal wave function $|\phi\rangle$. Likewise, detectors trace out the information contained in $|\phi\rangle$.

Repeating the same calculation as in the previous section, we find

$$p(\boldsymbol{s}) = |\langle\phi_1|\phi_2\rangle|^2 p_B(\boldsymbol{s}) + (1 - |\langle\phi_1|\phi_2\rangle|^2)p_D(\boldsymbol{s}) \tag{15}$$

for $\boldsymbol{s} = (1,1),(2,0),(0,2)$. The probability of a given outcome is now a convex combination of the ideal probabilities, weighted by the overlap between the internal states of the photons. This equation can be read as a quantum-to-classical transition, and follows straightforwardly the intuition that as photons are more distinguishable, the more they behave as classical particles.

In the case of a varying time delay, this gives rise to the typical HOM-curve, as displayed in Fig.1

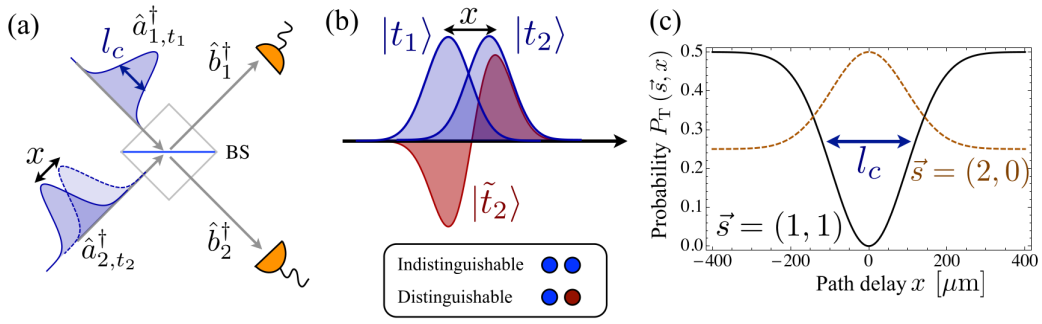


Figure 1: **The Hong-Ou-Mandel effect.** (a) Two independently prepared photons are sent through a beam splitter with a possible time delay, or equivalently, a path delay x. (b) This path delay leads to imperfect overlap of the wave functions of the photons, leading to partial distinguishability. (c) Exact cancellation of amplitudes leads to the suppression of the event $\boldsymbol{s} = (1,1)$ if $x = 0$. Partial distinguishability monotically reduces this cancellation. Image from [1].

1.2.4 Bunching and anti-bunching

The bunching of two photons in a single detector observed by HOM is prototypical of bosons. While one could imagine that the particles "stick" together, we above considered no interaction between them: this effect is entirely due to the statistics of the particles,

and the symmetries implied in their wavefunction.

We will study an extension of this phenomenon in one of the papers of this thesis.

1.3 Many-particle interference

We studied above the case of two particles interfering in a simple beam-splitter in order to introduce basic concepts of photonic interferometry. We will now consider a more generic case, where n particles are allowed to interfere in a m-mode network described by a unitary matrix U. The particles occupy the input modes according to $\boldsymbol{r} = (r_j)$ where $0 \leq r_j \leq n$ is the number of photons occupying input mode j. Naturally $\sum_{j=1}^{m} r_j = n$. We will denote output occupation numbers in the same way by $\boldsymbol{s}$. We define the mode assignment list $\boldsymbol{d}(\boldsymbol{r}) = \oplus_{j=1}^{m} \oplus_{k=1}^{r_j} (j)$ which allows us to define an effective scattering matrix for $(\boldsymbol{r}, \boldsymbol{s})$

$$\mathcal{U} = U_{\boldsymbol{d}(\boldsymbol{r}),\boldsymbol{d}(\boldsymbol{s})} \tag{16}$$

with as convention the jth row corresponding to the jth input, and likewise for columns and outputs. That is, $\mathcal{U}$ is the submatrix constructed by selecting the rows from U corresponding to the occupied input ports, and the columns for the output ports.

Particles entering mode j have a corresponding *internal wave function* $|\phi_j\rangle$. In particular, it is implicitly assumed that photons in a given mode have the same internal wave function : in practice, we are only interested in situations where there is at most one photon input mode.

We define the distinguishability matrix, or *Gram matrix*

$$\mathcal{S}_{ij} = \langle \phi_i | \phi_j \rangle \tag{17}$$

which contains the set of overlaps of the internal wave functions of the input photons. Note that this is relative to a given input configuration $\boldsymbol{d}(\boldsymbol{r})$, but in practice we will only be interested in the $\boldsymbol{d}(\boldsymbol{r}) = (1, ..., n)$ input, that is, a single photon in the first n modes. In particular, $S = \mathbb{1}$ if all photons are distinguishable, while $\mathcal{S} = \mathbb{E}$ ($\mathbb{E}_{j,k} = 1$ for all j, k) in the case where they are fully indistinguishable. We define a basis Φ_j for the internal states, with

$$\sum_j \langle \phi_k | \Phi_j \rangle \langle \Phi_j | \phi_k \rangle = 1 \tag{18}$$

for all k.

1.4 Time evolution

The initial state is

$$|\Psi\rangle_{in} = \prod_{j=1}^{m} \frac{1}{\sqrt{r_j!}} \left(\hat{a}^{\dagger}_{j,\phi_j} \right)^{r_j} |0\rangle \tag{19}$$

where $|0\rangle$ is the vacuum state. Defining $\mu(r) \equiv \prod_{j=1}^{m} r_j!$, we can rewrite the input state as

$$|\Psi\rangle_{in} = \frac{1}{\sqrt{\mu(\boldsymbol{r})}} \prod_{j=1}^{n} \hat{a}^{\dagger}_{d_j(\boldsymbol{r}),\phi_{d_j(\boldsymbol{r})}} |0\rangle \tag{20}$$

The interferometer U acts leaving the internal modes untouched but distributing the photons in the various output modes, with the amplitude U_{jk} for the photon j to end up in mode k

$$\hat{a}^\dagger_{j,\phi_j} \to \hat{U}\hat{a}^\dagger_{j,\phi_j}\hat{U}^\dagger = \sum_{k=1}^{m} U_{jk}\hat{b}^\dagger_{k,\phi_j} \tag{21}$$

(Note that an opposite convention is sometimes used in the literature, leading to easy confusions. For instance, [1] and [3] have U and $U^\dagger$ interchanged.) The output state is thus

$$\begin{aligned} |\Psi\rangle_{out} &= \hat{U}\,|\Psi\rangle_{in} \\ &= \frac{1}{\sqrt{\mu(\boldsymbol{r})}} \prod_{j=1}^{n}\sum_{k=1}^{m} U_{d_j(\boldsymbol{r}),k}\hat{b}^\dagger_{k,\phi_{d_j(\boldsymbol{r})}}\,|0\rangle \end{aligned} \tag{22}$$

1.5 Output pattern probabilities

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:

$$|\boldsymbol{d}(\boldsymbol{s}), \Phi_{\boldsymbol{a}}\rangle = \frac{1}{\sqrt{\mu(\boldsymbol{s})}} \prod_{j=1}^{n} \hat{b}^\dagger_{d_j(\boldsymbol{s}),\Phi_{a_j}}\,|0\rangle\,. \tag{23}$$

Here, $\boldsymbol{a}$ is a vector of dimension n, whose indices a_j define that the internal state of the jth particle is $\left|\Phi_{a_j}\right\rangle$.

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

$$|\Psi_{out}\rangle = \sum_{\boldsymbol{d}(\boldsymbol{s}),\boldsymbol{a}} \alpha(\boldsymbol{d}(\boldsymbol{s}), \Phi_{\boldsymbol{a}})\,|\boldsymbol{d}(\boldsymbol{s}), \Phi_{\boldsymbol{a}}\rangle \tag{24}$$

where

$$\alpha(\boldsymbol{d}(\boldsymbol{s}), \Phi_{\boldsymbol{a}}) = \left\langle \prod_{j=1}^{n} \hat{b}_{d_j(\boldsymbol{s}),\Phi_{a_j}} \middle| \Psi_{out} \right\rangle \tag{25}$$

Tichy defines a set of projectors sending a state on an output mode configuration $\boldsymbol{s}$ disregarding the internal state

$$\hat{P}_{\boldsymbol{s}} = \sum_i \prod_{j=1}^{n} \hat{b}^\dagger_{d_j(\boldsymbol{s}),\Phi_i}\,|0\rangle\,\langle 0| \prod_{j=1}^{n} \hat{b}_{d_j(\boldsymbol{s}),\Phi_i}$$

such that the probability to find the output state $\boldsymbol{s}$ is [2]

$$P(\boldsymbol{s}) = \frac{1}{\mu(\boldsymbol{r})\mu(\boldsymbol{s})} \left\langle \Psi_{out} \middle| \hat{P}_{\boldsymbol{s}} \middle| \Psi_{out} \right\rangle \tag{26}$$

Tichy shows that this probability can be expanded as a multi-dimensional *tensor*

permanent

$$P(\boldsymbol{s}) = \frac{1}{\mu(\boldsymbol{r})\mu(\boldsymbol{s})} \sum_{\sigma,\rho\in S_n} \prod_{j=1}^{n} \left(\mathcal{U}_{\sigma_j,j} \mathcal{U}^*_{\rho_j,j} \mathcal{S}_{\rho_j,\sigma_j} \right) \tag{27}$$

or alternatively

$$P(\boldsymbol{s}) = \frac{1}{\mu(\boldsymbol{r})\mu(\boldsymbol{s})} \sum_{\rho\in S_n} \left[\prod_{j=1}^{n} \mathcal{S}_{j,\rho_j} \right] \mathrm{perm}\left(\mathcal{U} \odot \mathcal{U}^*_{\rho,\mathbb{E}} \right) \tag{28}$$

This approach makes it clear that computing probabilities in the partial distinguishability regime is exponentially more costly than in the perfectly bosonic regime, since there is one more permutation to sum over. In particular, the fully indistinguishable case (bosonic) $\mathcal{S} = \mathbb{E}$ as

$$P(\mathbb{E}, U) = \frac{1}{\mu(\boldsymbol{r})\mu(\boldsymbol{s})} \left(\sum_{\sigma\in S_n} \prod_{j=1}^{n} \mathcal{U}_{\sigma_j,j} \right) \left(\sum_{\rho\in S_n} \prod_{j=1}^{n} \mathcal{U}^*_{\rho,j} \right) = \frac{|\mathrm{perm}\left(\mathcal{U}\right)|^2}{\mu(\boldsymbol{r})\mu(\boldsymbol{s})} \tag{29}$$

as well as the fully distinguishable case $\mathcal{S} = \mathbb{1}$ as

$$P(\mathbb{E}, U) = \frac{1}{\mu(\boldsymbol{r})\mu(\boldsymbol{s})} \sum_{\sigma\in S_n} \prod_{j=1}^{n} \mathcal{U}_{\sigma_j,j} \mathcal{U}^*_{\sigma_j,j} = \frac{\mathrm{perm}\left(|\mathcal{U}|^2\right)}{\mu(\boldsymbol{r})\mu(\boldsymbol{s})} \tag{30}$$

1.6 Matrix permanents

The core of this work and the general interest of the field rely on the properties of the matrix permanent, as introduced in the computation of the event amplitudes and probabilities. We will now explore some of the properties of this function. The permanent belongs to a class of *generalized matrix functions*. Calling S_n the group of permutations (also called symmetric group), taking H as a subgroup of S_n and χ are character of H, we define

$$d^H_\chi(A) = \sum_{\sigma\in H} \chi(\sigma) \prod_{i=1}^{n} A_{i\sigma_i} \tag{31}$$

the generalized matrix function of a matrix A with respect to H, χ. We can now see that this definition encompasses multiple useful matrix functions. If $H = S_n$ and $\chi(g) = \mathrm{sign}(g)$, we recover the *matrix determinant*

$$\det(A) = \sum_{\sigma\in S_n} \mathrm{sign}(\sigma) \prod_{i=1}^{n} A_{i\sigma_i} \tag{32}$$

while if $\chi(g) = 1$ we have the *matrix permanent*

$$\mathrm{perm}\left(A\right) = \sum_{\sigma\in S_n} \prod_{i=1}^{n} A_{i\sigma_i} \tag{33}$$

More generally, when $H = S_n$ and χ is an irreducible character, we refer to the generalized matrix function as the *immanant*. Another noteworthy case is if $H = 1$, from which we recover the product of the diagonal elements of A.

While at first sight expressions Eq. (32) and Eq. (33) might seem very similar, the properties of the determinant and the permanent are very different. Indeed, while the expression for the permanent looks more simple, it lacks many of the symmetries that induce the properties of the determinant. For instance, $\det(AB) = \det(A)\det(B)$ while $\mathrm{perm}\,(AB) \neq \mathrm{perm}\,(A)\,\mathrm{perm}\,(B)$. However, the Laplace rule holds for both the determinant and permanent, with the difference of the sign of the permutation

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \det(A[i,j]) \tag{34}$$

$$\mathrm{perm}(A) = \sum_{j=1}^{n} A_{ij}\,\mathrm{perm}(A[i,j]) \tag{35}$$

where $A[i,j]$ is the matrix constructed by removing row i and column j of A. It is then easy to understand that the permanent is "just the determinant without the signs".

A particularly striking difference between the permanent and the determinant lies in the efficiency with which each can be computed. Fast algorithms exist for the determinant—the most naive being Gaussian elimination—needing only $\mathcal{O}(n^3)$ operations for a matrix of size n [11]. In contrast, the best known algorithm to compute the permanent requires exactly $n2^n$ operations, as described by Ryser in [12] (and improved by a factor n compared to original paper using Gray ordering). This places the two functions in different complexity classes, as we will describe in Section 3.2. In fact, is it known that the permanent of matrices with entries of zeros and ones is #P complete [13].

Another, more physically relevant question is that of approximating matrix permanents. Indeed, experiments consist of a limited number of runs, and for this reason it is often sufficient to approximate matrix permanents at least as accurately as the experiment allows. We consider two type of approximations: given U, estimating $\mathrm{perm}\,(U)$ is said to be with an *additive error* if an algorithm outputs a value $\mathrm{perm}\,(U) + \epsilon$ while it is said to be of *multiplicative error* if it outputs $(1+\epsilon)\mathrm{perm}\,(U)$, where ϵ can be made arbitrarily small (through more computation steps).

The difference between a multiplicative and additive error is especially important. Indeed, the probabilities of most patterns appearing at the output of linear interferometer are exponentially small. That is, if n is the number of input photons and m is the size of the network, the *average probability* $p(\boldsymbol{s})$ scales as

$$\log(p(\boldsymbol{s})) \sim n\epsilon(\rho) + \mathcal{O}(\log(n)) \tag{36}$$

with $\epsilon(\rho) = \rho\log\rho - (1+\rho)\log(1+\rho)$ and $\rho = n/m$ the boson density [14]. Therefore, a algorithm running with an additive error $\epsilon = \mathcal{O}(\mathrm{poly}(n))$ would not bring a relevant approximation of outcome probabilities (because the additive error would need to be

kept exponentially small), while a multiplicative error algorithm would. An additive error algorithm exists by Gurvits in $\mathcal{O}(n^2/\epsilon^2)$ [15] and was later improved by Aaronson [16] for events with more than one photon per mode.

General matrix permanents are thought as hard to approximate with a multiplicative error, but some special cases are noteworthy and have deep ties with the physical processes studied in this thesis. In general, the U matrix describing the network will be composed of complex entries. However, if U is a matrix of non negative (real) numbers, the *JSV algorithm* allows for efficient multiplicative error approximation [17]. This result has the following physical implication. When looking at Eq. (29), we can see that the permanent determining event probabilities is that of a generic complex matrix in the case of indistinguishable photons, while for distinguishable photons in Eq. (30), the matrix is made of non negative numbers. This implies that bosonic probabilities are hard to approximate meaningfully, while distinguishable ones are easy. An important distinction must be made between computing the probabilities of a given event and sampling. This will be discussed in the Section 3.2, with similar conclusions.

2 Bosonic behaviours

We saw above how the statistics (bosonic, fermionic, distinguishable) of the particles can dramatically change their behaviour. A natural question arises: what happens between these cases? Indeed, it is hard to generate perfectly indistinguishable particles in experiments, and one can reasonable expect that a non negligible degree of distinguishability will always be present between particles. Generally, fermions are simple to treat, because the event probabilities are linked to matrix determinants, which are much easier than the matrix permanents of the bosons. Therefore, we take interest in what happens in the grey zone of partially distinguishable bosons. One of the paradigmatic bosonic effects is that of bunching, as emphasized by the HOM experiment, where bosons are more probable to be grouped together in a single output mode. Any degree of partial distinguishability reduces this effect, monotonically. It is known that when considering more photons and modes, the event probabilities may not transition monotically [18] from their bosonic to distinguishable values. In general, little formal results exist concerning the qualitative behaviour of partial distinguishable particles, while some rules of thumb are known by most physicists, such as that bosons tend to bunch. Shchesnovich formalized this important aspect for more than two modes and photons, in direct translation the HOM effect, and with more general parameters [19]. We will first study his theory, and then his conjecture that the HOM effect extends to more than two modes and particles: ideal bosons bunch most. One of the findings from this thesis shows that this is not the case, and that the general rule of thumb held by many physicists is not, in fact, a generic behaviour. This is particularly surprising and has piqued the interest of the community. Indeed, we will study in this thesis the numerous reasons why this conjecture appears plausible, and the various important cases in which it does hold.

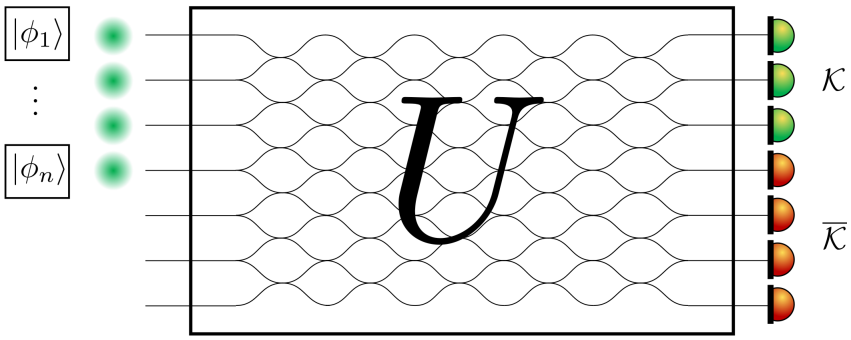


Figure 2: **Bunching in a subset.** Single photons are sent through the first n input modes of a m-mode linear interferometer U. Each photon at input j carries internal degrees of freedom (polarization, arrival time, etc.) described by an (internal) wave function $|\phi_j\rangle$. Not perfectly overlapping wave functions (measured via the Gram matrix S) gives rise to partial distinguishability amongst the photons, reducing the degree of quantum interference and bosonic effects such as bunching (see Conjecture 1). We focus in particular on the probability that all n photons bunch in a subset K (corresponding to the green detectors) of the output modes, while the red detectors in $\overline{K}$ do not click. This is a natural extension of the HOM experiment for more than two modes.

2.1 Generalized bunching

Shchesnovich studies *generalized bunching*, defined as finding all photons in a subset K of the output modes. This subset could consist of none or all the modes, in which case we call it trivial (as it brings no information). Generally, people refer to (strict) bunching as finding all photons in a single output mode, hence the term generalized bunching. Below, we will not make this distinction and assume that we are studying generalized bunching.

Regarding strict bunching, analytical results extend the conclusion of the HOM experiment to more than 2 photons/modes. If we choose $\mathcal{K} = \{j\}$, the *single-mode bunching* probability is given by

$$P_n(S) = \prod_{i=1}^{n} |U_{ij}|^2 \operatorname{perm}(S) = P_n^{(\text{dist.})} \operatorname{perm}(S), \tag{37}$$

where $P_n^{(\text{dist.})} = P_n(\mathbb{E})$ corresponds to the multimode bunching probability when the photons are fully distinguishable. The HOM effect can be recovered: for a two-mode 50:50 beam-splitter, the maximum probability of bunching in a single output mode is attained for fully indistinguishable photons ($S = \mathbb{E}$) and given by $1/4 \times 2! = 1/2$. The conclusion is straightforwardly generalized for larger interferometric setups as $P_n(S)$ is indeed maximum for fully indistinguishable photons since the maximum value of $\operatorname{perm}(S)$ is attained when $S = \mathbb{E}$, with $\operatorname{perm}(\mathbb{E}) = n!$.

One could hope that this still holds for a subset K with more than one mode. Shchesnovich conjectures that

Conjecture 1 (Generalized Bunching)**.** Consider any input state of n classically

correlated photons. For any linear interferometer and any nontrivial subset K of output modes, the probability that all photons are found in K is maximal if the photons are (perfectly) indistinguishable.

There is, in fact, plenty of evidence to support this conjecture, beyond the one present in Eq. (37). Indeed, when considering random configurations of interferometers and Gram matrices, we never found a counter example in tens of millions of trials in various dimensions. The procedure goes as follows: a unitary U is sampled randomly according to the Haar measure and used to compute the matrix H. The matrix S is constructed by taking random normalized vectors in a space of dimension r (for instance, when considering photons sent simultaneously but with potentially different polarization states, $r = 2$). Note that r determines the rank of the matrix S.

However, note that with *educated* numerical choices, counterexamples can be found in less than a second. This relies on setting a dimension $n \geq 7$, and taking an matrix H of rank for the matrix S of $r = 2, 3, \cdots \ll n$. In this case, we set the interferometer as fixed (through the choice of matrix H) and then we use a classical gradient-descent type optimisation algorithm over the coefficients of the n, $r-$dimensional generating vectors of the Gram matrix by optimising on the *violation ratio*

$$R = \frac{P_n(S)}{P_n(\mathbb{E})} \tag{38}$$

With this definition, a counter example to Shchesnovich's conjecture consists of finding $R > 1$. We did not investigate in depth the many numerical counter examples we found. A few of them can be found in the generated with the repository on github and the data found at OSF. While we can easily find counter examples with these tailored routines, this does not diminish their scarcity when randomly sampled according to the Haar measure.

2.2 Mathematical description

Let us see the main mathematical tools needed to formalize Shchesnovich's statement (a careful analysis is found in the associated paper). Consider the probability that all n photons are found in a subset $\mathcal{K}$ of the output modes, which we refer to as the *multimode bunching probability* $P_n(S)$. We define the matrix

$$H_{a,b} = \sum_{l \in \mathcal{K}} U^*_{l,a} U_{l,b}, \tag{39}$$

where $a, b \in \{1, ..., n\}$ and where $U, \mathcal{K}$ are the interferometer and the subset, respectively. Shchesnovich shows how to compute $P_n(S)$ in terms of matrix permanents

$$P_n(S) = \text{perm}\left(H \odot S^T\right) \tag{40}$$

where the special product is called the Hadamard (or element-wise) product

$$(A \odot B)_{ij} \equiv A_{ij} B_{ij} \tag{41}$$

unlike the standard matrix product [19] (see also Methods 4.5 of the associated paper). The attentive reader will be surprised at the simplicity of this formula. Indeed, when we are speaking of bunching in subset that contains more than a single mode (or conversely, which is not composed of all modes but one), finding n photons in the subset can be done in a great number of different output arrangements. As an example, if $n = 2$ and the subset K consists of the first two modes, all three configurations $(2, 0, ...), (1, 1, ...), (0, 2, ...)$ equally qualify as bunching. In general, one may consider subsets with poly(n) number of modes, making the number of such bunched configurations exponentially large. The probability of each of these configurations requires the computation of a matrix permanent. Therefore, it is surprising that when combining exponentially many of such probabilities, each one being computationally hard, we obtain this single matrix permanent. In fact, this specific permanent is itself easy to compute. Indeed, note that H, S and $H \odot S^T$ are all positive semidefinite matrices. Shchesnovich shows that these specific permanents are easy to compute in [19], while another result from Chakhmakhchyan, Cerf, and Garcia-Patron also provides a quantum-inspired algorithm to approximate such permanents [20] up to quasi-multiplicative error.

Note too that H, S and $H \odot S^T$ being Hermitian and positive semidefinite ensures that their permanent is positive [21]. This is, of course, expected given their physical interpretation as probabilities. Given that, for indistinguishable photons, $P_n^{(\mathrm{bos})} = \mathrm{perm}\,(H)$, the Conjecture 1, takes the following mathematical form:

$$\mathrm{perm}\left(H \odot S^T\right) \overset{?}{\leq} \mathrm{perm}\,(H) \tag{42}$$

with equality holding iff S corresponds to indistinguishable photons.

2.3 Bapat-Sunder conjecture

This inequality was known to mathematicians in the context of matrix permanent properties (with no apparent ties to quantum optics [22]). In the field of mathematics, it is referred to as the Bapat-Sunder conjecture [23]:

Conjecture 2 (Bapat-Sunder)**.** For any two positive semi-definite $n \times n$ matrices $A = (A_{ij})$ and $B = (B_{ij})$, we have

$$\mathrm{perm}\,(A \odot B) \leq \mathrm{perm}\,(A) \prod_{i=1}^{n} B_{ii}. \tag{43}$$

Although Conjecture 2 takes a slightly different form to Eq. (42), they are in fact equivalent as shown by Zhang in [24]. Only a few analytical results are known regarding this equation. Zhang [25] studied properties of matrices maximizing the function

$$f_A(X) = \mathrm{perm}\,(A \odot X) \tag{44}$$

for a given choice of matrix A. We call M_A a *maximizing matrix* of A if

$$f_A(M_A) = \max_{X \in \mathcal{C}_n} \operatorname{perm}(A \odot X) \tag{45}$$

with $\mathcal{C}_n$ the space of Gram matrices of dimension n. A violation to Conjecture 2 happens when

$$f_A(X) > f_A(\mathbb{E}) \tag{46}$$

(and in particular if $f_A(M_A) > f_A(\mathbb{E})$). Physically, maximizing matrices correspond to the partial distinguishability encoding that maximize bunching, given a choice of interferometer and subset. Zhang argues that such matrices always exist from the continuity of f and the compactness of $\mathcal{C}_n$.

These may, however, not be unique. Consider for instance

Proposition 1 (Rank 1 matrices)**.** Let $B = (b_{ij}) \in \mathcal{C}_n$. The following statements are equivalent.

1. $|B_{ij}| = 1$ for all i, j.
2. B is of rank 1 .
3. There exist complex numbers $u_1, \ldots, u_n$ all having modulus 1 such that $B = u^* u$, where $u = (u_1, \ldots, u_n)$, $u^* = \bar{u}^t$ is the conjugate transpose of u
4. $B = D^* \mathbb{E} D$, where D is some diagonal unitary matrix.
5. $\prod_{i=1}^n B_{i\sigma(i)} = 1$, for every $\sigma \in S_n$

As Zhang argues, statement 4 implies that the rank one correlation matrix is unique up to diagonally unitary similarity. Equivalently, by scaling through unit complex numbers, there is only one correlation matrix of rank 1, that is, $\mathbb{E}$.

In his article, Zhang takes the position that Bapat-Sunder's conjecture holds. In this case, $\mathbb{E}$ is clearly a maximizing matrix. While we now know that the conjecture is false, the results presented by Zhang are still highly interesting as they have a direct translation into physical properties.

Zhang points out that for the Hadamard product by a matrix whose elements are of modulus one does not change the value of the matrix permanent. Indeed, if $B \geq 0$ and $|b_{ij}| = 1$ for all i, j, we write $B = u^* u$ according to (3). Then for any $A \geq 0$

$$\operatorname{perm}(A \circ B) = \operatorname{perm}(A \circ (u^* u)) = \operatorname{perm}(\overline{u_i} A_{ij} u_j) = \prod_{i=1}^{n} |u_i|^2 \operatorname{perm} A = \operatorname{perm} A \tag{47}$$

This calculation recovers, in fact, something that is to be expected trivially from physics: the elements of u are likened to a phase held by each particle's internal wave function. These phases multiply into a global phase in front of the input state. What the above calculation confirms is that such phases do not change the bunching probabilities. We can intepret this result as some gauge freedom in the choice of Gram matrices.

Zhang then proved the following theorem.

Theorem 1 (Singularity of maximizing matrices)**.** If $A \in \mathcal{C}_n$ and $A \neq \mathbb{E}$, then M_A is singular.

Therefore, maximizing matrices are necessarily rank-deficient. In physical terms, this means the following:

Corollary 1 (Linear dependance of maximally bunching internal states)**.** Gram matrices leading to maximal bunching are generated from internal wave functions $\{|\phi_1\rangle, \ldots, |\phi_n\rangle\}$. These functions must be linearly dependant. In particular, they can be described by $r < n$ basis elements. In turn, the internal wave functions can be seen as belonging to a space of dimension r.

We remind that this holds true in any interferometric situation, regardless of the existence of a violation of Conjecture 1. Note that this is stated for maximizing matrices. Interestingly, this singularity can be destroyed by a minor perturbation. As all the functions treated are continuous (the permanent is a sum of products), if a violation is found, full rank counterexamples must exist.

While we saw that maximizing matrices may not be unique, Zhang shows that they are all linked by the following invariance:

Theorem 2 (Invariance row/column deletion)**.** Let A be an $n \times n$ Gram matrix. There exist real constants $\lambda_1, \ldots, \lambda_n$ depending solely on A such that for any maximizing matrix X of A

$$\operatorname{perm}(A \circ X) = \operatorname{perm}((A \circ X)[i,i]) + \lambda_i, \quad i = 1, \ldots, n. \tag{48}$$

where $A[i,j]$ is the matrix whose row i and column j are removed.

Corollary 2. Let X and Y be maximizing matrices of A. Then,

$$\operatorname{perm}((A \circ X)[i,i]) = \operatorname{perm}((A \circ Y)[i,i]) \quad \forall i = 1, \ldots, n \tag{49}$$

While deleting a row and column might seem like a strange operation, it is interesting on multiple levels. In his paper, Zhang compares this property of maxizing matrices to the properties of matrices relating to the most famous conjecture on matrix permanents: the Van Der Waerden conjecture, later proven by Egorycev [26, 27]. It concerns doubly stochastic matrices, i.e. matrices with nonnegative real entries and where each row and column sums to one. These matrices can be likened to the unitary U, for classical mixing processes. The simplest case of such a matrix is maximally mixing matrix $1/n * \mathbb{E}$, whose permanent is trivially

$$\operatorname{perm}\left(\frac{1}{n}\mathbb{E}\right) = \frac{n!}{n^n} \tag{50}$$

The Van der Waerden's conjecture states that:

Theorem 3 (Van der Waerden)**.** A doubly stochastic matrix A obeys

$$\operatorname{perm}(A) \geq \frac{n!}{n^n} \tag{51}$$

with equality achieved iff A is the maximally mixing matrix.

The existence of a similar theorem for quantum processes was one of the initial questions we pondered in this thesis. In the proof of this theorem, Egorycev makes use of minimizing matrices, which directly compare to the maximizing matrices of Zhang. The proof aims to show that the only minimizing matrix is the maximally mixing matrix, by constraining its structure. These matrices are shown to be invariant by row and column deletion [28]:

$$\operatorname{perm}(A[i,j]) = \operatorname{perm}(A)\,\forall i,j \tag{52}$$

making Theorem 2 a very interesting parallel.

Physically, row and column deletion is also relevant. To understand it, let's study the formulae of Renema to approximate bosonic processes. In [29], he shows that partially distinguishable particles have event probabilities that can be expressed as a sum over r perfectly interfering particles and $n-r$ perfectly distinguishable, with $r = 0, \ldots, n$. The simplest model of partial distinguishability sets all particles as equally distinguishable and refer to this colloquial toy-model as the "x-model". We assume that the internal states are given by

$$|\phi_j\rangle = \sqrt{x}\,|\Phi_0\rangle + \sqrt{1-x}\,|\Phi_j\rangle\,, \tag{53}$$

where $\langle\Phi_k|\Phi_l\rangle = \delta_{kl}$ for $k,l \in \{0,1,...n\}$. In this case the S matrix becomes

$$S_{ii} = 1 \tag{54}$$

$$S_{ij} = x \text{ if } i \neq j, \tag{55}$$

interpolating between the two extreme cases ($x = 0$ for the distinguishable case, $x = 1$ for the bosonic one), i.e. $S = (1-x)S_D + xS_B$ is a convex interpolation between both extremes. Let us call σ^j the permutations in S_n which have $n-j$ elements as fixed points (a fixed point being defined as $\sigma(l) = l$ for some value of $l \in 1, \ldots, n$). Renema then obtains a new form for Eq. (27) as

$$P(\boldsymbol{s}) = \sum_{j=0}^{n} \sum_{\sigma^j} x^j \sum_{\rho} \operatorname{Perm}\left(M_{\rho,1} * M^*_{\rho,\sigma_p}\right) \operatorname{Perm}\left(|M_{\bar{\rho},\sigma_u}|^2\right) \tag{56}$$

where ρ denotes the $\begin{pmatrix} n \\ j \end{pmatrix}$ possible combinations of j columns from the matrix M, $\bar{\rho}$ denotes the complemenary rows, and σ_p and σ_u denote the permuted and unpermuted elements of σ, respectively. In each term of the summation, the first matrix permanent corresponds to fully interfering particles, while the second corresponds to fully distinguishable particles. A way of approximating event probabilities is to truncate the sum in j.

In the case of Corollary 2, the situation is slightly different than that of the formalism developed by Renema. He considers the approximation of event probabilities, while we consider photonic bunching. Nevertheless, the removal of row and column i physically

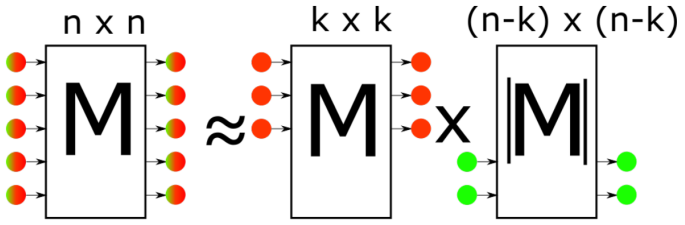


Figure 3: The interference process between n partially distinguishable photons splits into a sum over terms containing k perfectly interfering photons with $n-k$ perfectly distinguishable ones. The k-th order terms are weighted by the degree of partial distinguishability x^k, making high order interference terms less apparent as photons become distinguishable. Image from [30].

consists in removing the photon i, as can be readily understood from the definition of the Gram matrix and the H matrix, and owing to the presence of a Hadamard product between the two.

Corollary 3. Any two Gram matrices leading to maximal bunching have the same bunching probability upon removal of a single photon i, for any i.

Again, this holds in any bunching scenario, without the necessity of a violation of the bunching conjecture. Note also the parallel of this statement with the singularity requirement. It is, as far as we know, an open question to show that maximizing matrices are unique modulo gauge transformations.

2.4 General realization

The link between the conjecture of Bapat Sunder and a physical experiment is not entirely direct: if we compare Eq. (43) and Eq. (64), it is clear that the matrix B corresponds to the transpose of S. To obtain the internal wave functions corresponding to a Gram matrix, we need to perform the *Cholesky decomposition* (note that special care must be taken when dealing with incomplete rank). Regarding the matrix A, it directly corresponds to H. However, while using an interferometer U and subset K to build up a matrix H from Eq. (39) is trivial, going from a matrix H to an interferometric setup seems less evident. Let's see how this is performed. The matrix A admits a Cholesky decomposition

$$A = M^\dagger M \tag{57}$$

as it is hermitian positive semi-definite. The matrix M is of size $r * n$ where r is the rank of A. Thus, it is tempting to identify M with U with a subset K of the output modes being the first r modes as

$$A_{a,b} = \sum_{k=1}^{r} M^\dagger_{a,k} M_{k,b} = \sum_{k=1}^{r} M^\dagger_{a,k} \left(M^\dagger_{b,k}\right)^* \tag{58}$$

closely resembles the definition of the matrix H in Eq. (39). The matrix $M^\dagger$ however, may not be unitary as is U in Eq. (39). It is always possible to incorporate a matrix, up to renormalization, into a bigger one that is unitary following Lemma 29 of the seminal paper of Aaronson-Arkhipov defining the Boson Sampling paradigm [31]. It states :

Lemma 1 (Incorportation in a unitary)**.** Let $X \in \mathbb{C}^{n\times n}$. Then for all $m \geq 2n$ and $\varepsilon \leq 1/\|X\|$, there exists an $m \times m$ unitary matrix U that contains εX as a submatrix. Furthermore, U can be computed in polynomial time given X.

It is easy to understand the need for renormalization as the columns (and rows) of a unitary matrix need to be of unit norm. This does not impact the Bapat-Sunder conjecture as we have noted above. The proof of the lemma is constructive and starts by finding a $2n * n$ matrix whose columns are orthonormal, and then by filling the n remaining columns of the described $2n * 2n$ matrix so that it forms a unitary matrix. This later step can be implemented, for instance, by filling the remaining n columns with random entries, and then orthnormalizing them through the Gram-Schmidt procedure. Note that when A is of incomplete rank ($r < n$) we can restrict the U to the uppermost $(n + r) * (n + r)$ submatrix found through the procedure of the lemma.

2.5 Other conjectures on permanents

We studied above (generalized) bosonic bunching in its most commonly admitted form, with independently prepared particles. Shchesnovich also studied another permanental conjecture, the *permanent on top* conjecture [32], which relates to bunching of a generic (possibly entangled) input [19].

In this more generic context, we need to study properties of the *Schur power matrix* built from the matrix H of above as

$$\Pi_{\sigma,\rho}(H) = \prod_{\alpha=1}^{n} H_{\sigma(\alpha),\rho(\alpha)} \tag{59}$$

where $\sigma, \rho \in S_n$ are permutations. Shchesnovich shows that the maximum (minimum) eigenvalues of this matrix directly related with maximally (minimally) bunching configurations. Schur showed that the minimum eigenvalue is always the determinant, corresponding to an input made of indistinguishable fermions.

Conjecture 3 (Permanent On Top)**.** The largest eigenvalue of the Schur matrix $\Pi(H)$ is perm (H).

This conjecture was disproven by direct numerical search by Shchesnovich in dimension $n = 5$. This result implies the existence of a generic 5 particles input that bunches more than ideal photons. This counter example does not display any apparent structure. We are of the opinion that speaking of bunching in this context is less relevant to the common understanding of physicists, given that the particles are not independently prepared. Therefore, on an intuitive level, this type of enhanced bunching compared to ideal bosons is less unexpected, given the surprising features found with the presence of entanglement in quantum physics.

Other conjectures on generalized matrices exist in the literature, and are summarised in [33]. It is interesting to see if more of these purely mathematical conjectures can be translated into physical behaviours.

3 Sampling problems

Intrinsic to the probabilistic nature of quantum mechanics is the notion of sampling. In the particular context of the evolution of a state through a linear network, such as the case presented above of single photons evolving through passive linear optics, one can define the probability of finding the output state in a particular mode configuration $\boldsymbol{s}$. A problem of interest is to obtain samples from the distribution $\mathcal{D} = \{\boldsymbol{s}_i\}$ of possible output configurations, given a particular input state $|\Psi_{in}\rangle$. This sampling problem is referred to as *Boson Sampling*, as introduced by Aaronson and Arkhipov in [31]. We will speak of boson sampling whether the particles are indistinguishable or have some degree of distinguishability. A variant of boson sampling where the input particles are fermions is called Fermion Sampling.

3.1 Boson Sampling

Boson sampling is the quantum version of the famous Galton board tabletop experiment, in which pebbles fall in a network and distribute in bins as shown in Fig. 4. In the Galton board experiment, the distribution of balls in bins $\mathcal{D}'$ tends towards a Gaussian [34]. Because classical balls do not interfere, the problem of sampling from Galton's board is easy. Indeed, while probabilities are given by matrix permanents as in Eq. (30), sampling can be done by considering each particle independently. Let us remind that if the network is described by the matrix U_{ij}, the index i refers the the input mode i and j to the output mode. Restricting to at most 1 particle per input mode, we will consider, without loss of generality, that they are sent in modes $1, \ldots, n$. The particle in mode i will end up in mode j with probability $|U_{ij}|^2$. Therefore, one can obtain a sample of the distinguishable distribution by sampling mode output indexes according to the weights $(|U_{i1}|^2, \ldots, |U_{im}|^2)$ for each particle i independently.

The situation is different for indistinguishable particles. As we saw when studying the HOM effect, particles do not behave independently: while there is no interaction term in the Hamiltonian, an effective interaction exists from the symmetry of the wave function describing the multiparticle state. In practice, the best known algorithm to sample from $\mathcal{D}$ takes a time equal to the computation of a single matrix permanent of size n, i.e. $n2^n$ [35] (with some improvements if the boson density is high [36], which is not the main case of interest as will be discussed shortly).

Boson sampling can be seen as a primitive quantum computer. A core difference with most of the schemes found in the quantum computing literature is that is it not universal: while the choice of input state and interferometer may be altered, one cannot hope to run a generic algorithm (such as Shor's prime factorization [37, 38]) on a boson sampler [31]. It is important to note that if the network that may be adapted given

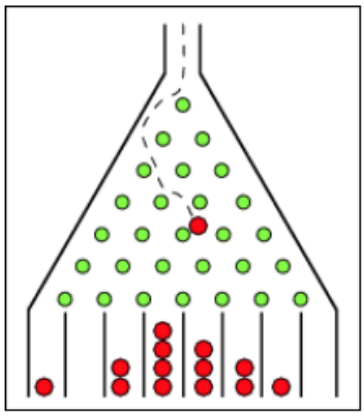


Figure 4: **The Galton board** table-top experiment is the inspiration for boson sampling. Particles are sent in a network (here, of poles on which they can bounce) and then collected into bins. After sufficiently many particles are sent, an overall pattern appears: the distribution of pepples in bins becomes a gaussian. Image from [31].

the result of measurements, then generic quantum computing can be done, as was demonstrated by Knill, Laflamme and Milburn in [39].

Therefore, it is natural to ask: why one should take interest in boson sampling? While opinions differ on the possibility of running practical tasks on a boson sampler, the initial interest comes from the detailed complexity analysis performed by AA. Under strongly held conjectures of computational complexity theory, boson sampling is a problem that is hard to simulate on a classical computer. However, as we described above, the physical implementation of boson sampling is direct in principle. If one was able to demonstrate an experimental boson sampler, this would invalidate the famous Church-Turing thesis. In informal terms, it states that problems are computable if and only if they can be translated to run on a Turing machine, which is a direct equivalent to classical computers. An extended version of the thesis adds the ability to draw random numbers, in which case we speak of probabilistic algorithms.

3.2 Computational complexity of boson sampling

Let us present here a few elements on complexity theory in order to better understand the motivations behind boson sampling and the precise claims of AA. Multiple metrics can be used to categorize algorithms, such as the number of elementary computation steps (such as an addition) necessary for execution or the size of the memory array required. We will be interested only in the number of steps required to execute an algorithm. Complexity theory is interested in the asymptotic behaviour of the metrics. If a problem has a size n, (such as the size of two $n \times n$ matrices to be multiplied), how many steps are required to perform the algorithm, at the limit of large n? The answer to this question allows to classify tasks in different classes. Algorithms that require only a polynomial number of steps to be executed, often written as poly(n), are said to

belong to the class P. For instance, naive matrix multiplication requires n^3 operations. We say that these algorithms are efficient. Practically, they may still be very slow: the polynomial could be dominated by a large power, such as n^{2048}, or simply have large numerical prefactors.

Another interesting complexity class, NP, is that of problems which may be difficult to solve, but whose solution can be checked efficiently. A famous example is the task of factorizing one integer of n digits in a product of prime numbers. This problem is conjectured to be hard to solve. In fact, most of modern cryptography relies on this assumption, such as RSA [40]. Best known algorithms take a (quasi-)exponential time to complete. However, if given a tentative solution, it is trivial to multiply the different primes to see if we obtain the initial number. One of the most important unsolved question in computer science is $P \stackrel{?}{=} NP$. It is strongly believed that the answer is negative: there are problems whose solution can be verified efficiently but not computed efficiently.

A pivotal moment of quantum computing theory was the publication of an algorithm by Shor, which efficiently solved the task of prime factorization on a quantum computer [37, 38]. This naturally lead to the definition of a class of problems that can be solved efficiently with bounded error on quantum hardware, BQP.

Whether quantum hardware allows us to gain a theoretical advantage compared to classical hardware relies on the relationship between the various complexity classes. While little is known with certainty, under the commonly held conjectures by computer scientists, a boson sampler would not be possible to simulate efficiently on a classical device [31]. This would, in turn, invalidate the Church-Turing thesis.

An important constraint of the proof of hardness of boson sampling is that the number of modes must scale roughly as $m \equiv n^6$. It is strongly conjectured that a scaling of $m \equiv n^2$ is sufficient. In this case, we speak of the *no-collision regime*: it is unlikely to find more than one photon per detector [41].

3.3 Experimental realizations of boson sampling

While very simple when stated conceptually - just send photons through an optical network and detect them - boson sampling revealed challenging experimentally. Nonetheless, experiments went from $n = 3$ to 20 single photons in the time span of a decade [42]. Even more impressively, the alternative scheme of Gaussian Boson Sampling, being easier experimentally, led to multiple experiments with claims of quantum supremacy [43, 44]. The most recent experiment displays a mean of 125 photons, up to 219 photon detected from 216 squeezed modes at the input.

Let us study here how boson sampling is realised experimentally, and what challenges are faced by experimentalists. This will allow us to better understand the sources of errors that arise in experiments.

3.3.1 Sources

Generating single photons can be done in different ways. Historically, *Spontaneous Parametric Down Conversion* sources are used. There, a nonlinear (χ^2) crystal is pumped by a laser and splits one photon into a pair of two photons of lower energies, called signal and idler. The signal photons will be sent to the experiment and interfere, while the idler signal will be detected as an auxiliary photon that asserts the presence of the signal photon. The problem with SPDC is its low efficiency, making it challenging to generate many single photons simulatenously when combining different sources.

Another, more recent technique is to use *Quantum Dots*. They are artificial two-level systems realised by creating defects in semiconductors. The whole system is usually placed inside of an optical microcavity. The dot can be excited by a laser and will spontaneously descend in energy level by sending a single photon. Quantum dots are overall vastly superior sources [45] and are now favored in boson sampling experiments.

Typically, a single quantum dot is used and its output is divided (demultiplexed) into different spatial modes. The laser excitation of the dot is periodic, and therefore a train of regularly spaced photons is produced at the output of the dot. This predictability can be used to separate photons deterministically in different modes using optical switches.

One of the key requirements for multiphoton interference is the indistinguishability between pairs of photons. In this regard, quantum dots provide high levels of indistinguishability, with pairwise overlaps above $x > 0.90$ (see Eq. (53)).

3.3.2 Network

One typically distinguishes two types of optical networks: bulk optics, where optical elements are macroscopic, and integrated optics, where the network is realised on a semiconductor of much smaller size. One of the most stringent requirements in the proof of hardness of boson sampling is the ability to pick a network distributed randomly in the Haar-measure. In general, any $m * m$ unitary can be realised by a collection of beam-splitter and phase-shifters, with depth m [46] (see also [47, 48]).

It is desirable to have an experimental setup that allows for reconfigurability of the network [49], in particular for potential applications of boson sampling, as they are encoded in the choice of unitary.

One of the biggest limitations of boson sampling, with photon distinguishability, is photon loss. It is known that boson sampling remains hard with a few lost photons [50], however, in realistic scenarios with universal networks, loss cannot be circumvented when considering asymptotic scalings. Indeed, the circuit depth needs to be of order $D = \text{poly}(n)$ and the loss probability, for each photon, goes as e^{-D}.

3.3.3 Detectors

Detecting the presence or absence of photons is relatively easy, and is most commonly done by *threshold detectors*. These are blind to the number of single photons. In general, a more advanced type of detector, called photon number resolving (PNR) is preferred. The latter can distinguish between $0, 1, 2, ...$ up to a couple of photons with some level

of confidence. While genuine PNR detectors are still an advanced piece of equipment, pseudo photon number resolution can be done by demultiplexing a mode in a few extra modes using beam splitters. For instance, one mode can be divided into 4 by using one balanced beam splitter, then a layer of two other beam splitters on the output modes of the first. Photons present in the initial mode will distribute with some probability in the 4 exit modes, where they can be detected by more readily available threshold detectors. This commonly used technique allows for some statistical information on the photon number content, but does not operate strict, deterministic photon number resolution [51]. Typical problems with detectors are false positive results, called *dark counts*. Conversely, missing the presence of light is a false negative, referred to as detectors inefficiency.

3.3.4 Time-binned boson sampling

A specific type of architecture, called time-binned boson sampling, has attracted attention recently [52, 44, 53]. It uses an array of impinging photons coming from a quantum dot excited periodically every τ units of time. These photons are sent through a time-dependant beam splitter, whose reflectivity can be changed within a time lesser than τ. One of the output branches of the beam-splitter is sent back to the entrance of the beam splitter, making for its second input mode. The link is done with an optical fibre, which we call a delay line. It is configured so that photons need exactly τ units of time to travel through it. The remaining output port of the beam splitter can either be sent directly to detectors, or through another network of loops. To achieve a universal interferometer, it is sufficient to have a second, larger loop of travel time $> n\tau$, see Fig. 5.

In practice, the first impinging photon goes through the beam splitter and part of its amplitude stays in the loop, while part goes to the detector. The part remaining in the loop meets the second impinging photon τ later, and interferes with it. In this way, the array of photons can be decomposed in time bins which are likened to spatial modes in the original boson sampling picture.

This architecture is especially interesting for experimentalists as it minimizes the number of devices necessary [54]. Indeed, general boson sampling can performed with one source (the quantum dot), one time-dependant beam splitter, two optical switches, and a single detector on top of the optical fibres for the delay lines. In contrast, standard boson sampling is generally realised with one reconfigurable linear optical network (either in bulk optics or integrated) with one detector per mode.

3.3.5 Alternative sampling schemes

One of the main difficulties of standard boson sampling is to generate a large number indistinguishable single photons. Another type of input that is more readily accessible are single mode squeezed states. We can use this type of Gaussian state at the input of a linear network and use similar detectors to the standard boson sampling case. In this case, it is shown that event probabilities depend on the matrix function called the

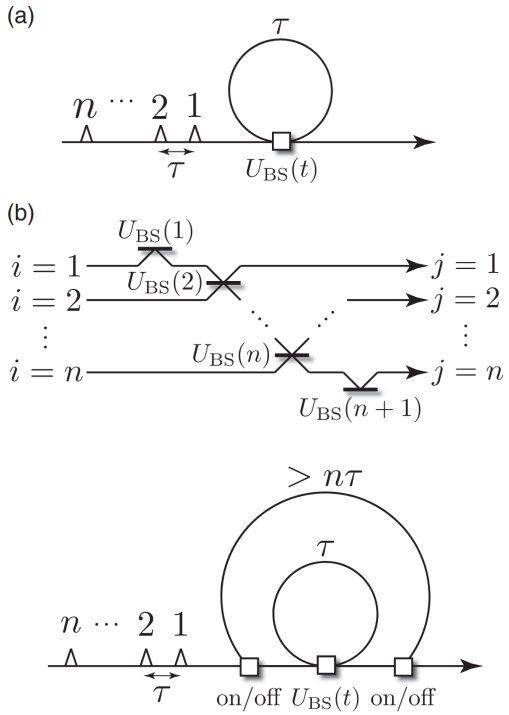


Figure 5: **Time binned boson sampling**. (a) Time binned sampling with one loop. Photons are sent through a time dependant beam splitter, implementing the matrix element $U_{BS}(t)$. The loop takes a time τ to travel, equal to the spacing between the photons. (b) The equivalent picture with spatial modes shows that this scheme is incomplete. Below is shown one way to create a universal boson sampler, by the addition of a longer loop, allowing to send back the photon train through the varying beam splitter multiple times. Figure from [52].

Hafnian. It is shown that this matrix function is also #P hard, and likewise, that it is hard to sample from the GBS distribution [55]. Many of the recent developments around boson sampling have revolved around the GBS paradigm, including major experiments claiming quantum supremacy [43, 44] as well as practical applications [56, 57, 58, 59].

3.4 Validation

By construction of the boson sampling problem, it is hard to verify whether an experimental boson sampler is working properly. Indeed, the experiment outputs a series of samples. Calculating the probability of each sample requires the calculation of a matrix permanent, which is exponentially hard. On top of that, in the regimes where AA proved hardness of boson sampling there are exponentially many samples that can be observed. An experiment can therefore not hope to reconstruct the entire boson sampling distribution when working with high photon numbers. Other ways have to be devised to infer the correct working of the experiment, which we refer to as *validation*. An even harder task is that of *certification*, which differentiates between a working experiment and possibly adversarial actors, actively trying to cheat the certification method (i.e. samples do not have to come from a physical device). In contrast, a validation method provides evidence that an experiment, operating under reasonable physical assumptions, is running as desired.

We expect from a good validation method to apply to any interferometer U. This *generality* expectation is not satisfied by all validation methods, such as those relying on specific symmetries of a limited class of networks (such as the Fourier or Hadamard interferometer). In this case, a large quantity of events may be forbidden. Observing them at the output rules out a working boson sampler. In this case we speak of *suppression laws*.

Given that the greatest initial motivation in doing boson sampling was reaching a regime of quantum supremacy, validating the quality of the photonic interference is particularly high on the list of desiderata. Indeed, if photons are sufficiently distinguishable, then the experiment becomes easy to simulate on a classical computer. The threshold of indistinguishability is quite high, in realistic contexts [29] estimates it at $x > 0.95$. Many methods rely on considering quantities that only relate to low order photon interference. For instance, on can study low order correlation functions of the number of detected photons in a couple of detectors. While practical, these methods do not probe the interesting, high order interference terms, whose role is crucial in boson sampling [60, 61, 62, 63].

We saw above that generating the entire boson sampling distribution $\mathcal{D}$ quickly becomes prohibitively expensive as n increases, and, in fact, even computing a single event probability. We ask that a validation method be efficient in classical computations. Likewise, the experiment only has access to a limited number of samples, and we ask that the method be efficient in the number of samples needed.

In the publication "Efficient validation of Boson Sampling from binned photon-number distributions", we study more in depth the question of verifying boson sampling experiments. We then propose a new formalism, unifying under a common frame many

of the previously existing methods of validation. We then show that it ticks all desired boxes mentionned above.

Part III
Publications

This part contains the three articles that form the core of this thesis in their published form. These articles aim at advancing our understanding of multiphoton interference, with special emphasis on the specific case of boson sampling — a key topic in quantum computing that has been the subject of intense investigation over the past decade. The articles address various aspects of boson sampling, including the role of indistinguishability in photon bunching, the development of efficient validation methods, and the introduction of cutting-edge simulation tools.

The first article, "*Boson bunching is not maximized by indistinguishable particles*," investigates the relationship between indistinguishability and photon bunching, a fundamental aspect of quantum optics. By exploiting a recent finding in the theory of matrix permanents, this study challenges the widely-held belief that indistinguishable photons bunch most. We exhibit a family of optical circuits that show increased photon bunching when the photons are made partially distinguishable through a specific pattern. This unexpected behavior calls into question our understanding of multiparticle interference and has implications for the validation of photonic devices.

In the second article, "*Efficient validation of Boson Sampling from binned photon-number distributions*," we address the critical issue of validating experimental boson samplers. We propose a versatile test based on the distribution of photons among partitions of output modes. This method encompasses previous validation tests and efficiently distinguishes ideal boson samplers from those affected by realistic imperfections, such as partial distinguishability. The development of reliable validation techniques is crucial for establishing the validity of experimental data and for substantiating claims of quantum computational advantage.

The third article, "*BosonSampling.jl: A Julia package for quantum multiphoton interference*," presents an open-source package designed for high-performance simulation and numerical investigation of boson samplers and multi-photon interferometry in general. Written in Julia, the package combines the benefits of high-level coding with C-like performance, enabling researchers to quickly develop and modify code without requiring advanced programming skills. BosonSampling.jl provides a range of routines for boson sampling tasks, such as statistical tools, optimization methods, classical samplers, and validation tools, addressing the growing need for optimized and adaptable numerical simulations in this rapidly evolving field. The extendable framework allows researchers to explore new paradigms in multi-photon interferometry and contribute to the ongoing competition for demonstrating a computational advantage of quantum devices over classical computers.

4 Boson bunching is not maximized by indistinguishable particles

Abstract

Boson bunching is amongst the most remarkable features of quantum physics. A celebrated example in optics is the Hong-Ou-Mandel effect, where the bunching of two photons arises from a destructive quantum interference between the trajectories where they both either cross a beam splitter or are reflected. This effect takes its roots in the indistinguishability of identical photons. Hence, it is generally admitted – and experimentally verified – that bunching vanishes as soon as photons can be distinguished, e.g., when they occupy distinct time bins or have different polarizations. Here we disproof this alleged straightforward link between indistinguishability and bunching by exploiting a recent finding in the theory of matrix permanents. We exhibit a family of optical circuits where the bunching of photons into two modes can be significantly boosted by making them partially distinguishable via an appropriate polarization pattern. This boosting effect is already visible in a 7-photon interferometric process, making the observation of this phenomenon within reach of current photonic technology. This unexpected behavior questions our understanding of multiparticle interference in the grey zone between indistinguishable bosons and classical particles.

In quantum physics, it is common knowledge that a gain of information results in the extinction of quantum interference. In the iconic double-slit experiment, the interference fringes originate from the absence of which-path information [64, 65, 66]. As emphasized for instance by Feynman [67], the fringes necessarily disappear as soon as the experiment allows us to learn that the particles have taken one or the other path. In quantum optics, photon bunching is another distinctive feature that follows from quantum interference. Specifically, the indistinguishability of photons makes it such that one cannot know which photon has followed a given trajectory in a linear interferometer (see Fig. 6). The Hong-Ou-Mandel (HOM) effect [10], for example, arises from the fact that one cannot distinguish the trajectory where two photons have crossed a 50:50 beam splitter from the trajectory where they have both been reflected. The net result is a tendency of indistinguishable photons to occupy the same mode – that is, to bunch – as a consequence of this lack of information. In accordance with Feynman's rule of thumb, quantum interference effects become less pronounced as soon as the photons become distinguishable, for example if they occupy distinct temporal or polarization modes so that one gains information about their individual trajectories [68]. As a consequence, the HOM dip disappears for two photons with orthogonal polarization since their distinct trajectories can then be fully traced back. Hence, even in more general scenarios involving multiple independently prepared photons and larger interferometers, it is commonly admitted that bunching effects are maximum for fully indistinguishable photons and gradually decline when photons are made increasingly distinguishable [69, 19, 70].

Here, we find a quantum interferometric scenario that goes against this intuition and contradicts the very idea that distinguishability undermines photon bunching. We

consider the probability of multimode bunching, i.e., the probability that all photons entering a linear interferometer end up in a certain subset of the output modes. In accordance with a longstanding mathematical conjecture on matrix permanents due to Bapat and Sunder [23], this bunching probability must indeed be maximum if the input state consists of fully indistinguishable photons [19]. However, inspired by a counter-example to this conjecture recently discovered by Drury [71] we have found that multimode bunching may, against all odds, be enhanced if photons are partially distinguishable. To do so, we not only convert Drury's counterexample into a linear interferometric experiment but also find a natural generalization, leading to a family of physical setups where partial distinguishability enhances multimode bunching. Incidentally, on the mathematical side, this finding even implies new counterexamples to the Bapat-Sunder conjecture.

More specifically, we construct a family of interferometers such that a higher two-mode bunching probability is attained if the photons are prepared in a well-chosen polarization state (making them partially distinguishable) rather than in the same state (in which case they would all be indistinguishable). In other words, gaining partial information about the photons paradoxically results in a higher probability for them to coalesce on two output modes, contradicting common knowledge. We give an interpretation of the physical process behind this counterintuitive effect and prove that, in our setup, the enhancement ratio of the two-mode bunching probability actually grows (at least linearly) with the system size. In the simplest case, an enhancement of 7% is already visible for 7 photons in 7 modes with a specific polarization pattern, which makes the observation of this remarkable phenomenon within reach of today's photonic technology.

4.1 Multimode boson bunching

Consider a general interferometric experiment in which n bosons are sent through a linear interferometer of m modes, described by the $m \times m$ unitary matrix U. Although our discussion is valid for any bosonic particle, we focus on photons here since, in practice, controlled linear interferometric experiments are easier to carry out in photonics. In Ref. [19], Shchesnovich provides compelling evidence for the following conjecture:

Conjecture 1 (Generalized Bunching) Consider any input state of n classically correlated photons. For any linear interferometer and any nontrivial subset $\mathcal{K}$ of output modes, the probability that all photons are found in $\mathcal{K}$ is maximal if the photons are (perfectly) indistinguishable.

Note that the considered class of input states must exclude entangled photons, so that it keeps a closer resemblance with the original Hong-Ou-Mandel setting (indeed, a related conjecture with entangled input states was proven to be false in Ref. [72]). In fact, it is enough for our purposes to consider a further simplification of the statement of Conjecture 1 and assume that photons are prepared independently of each other, i.e., they are uncorrelated, and that there is only one photon in each of the first n modes. Moreover, we may also assume that the state of each photon is pure. This setting, depicted in Fig. 6, is enough to demonstrate that Conjecture 1 is false. We refer

to Ref. [19] for a mathematical treatment of bunching probabilities in more general scenarios.

The state of each photon is not only described by the spatial mode it occupies but also by other degrees of freedom, such as its polarization and its spectral distribution, which we will refer to as internal degrees of freedom. Partial distinguishability can then be modelled by considering that the internal state of the photon entering mode j is described by an (internal) wavefunction $|\phi_j\rangle$. Let us denote the creation operator associated to this state as $\hat{a}^\dagger_{j,\phi_j}$. We make the common assumption that the interferometer acts only on spatial modes, leaving the internal state of the photon invariant [2, 3, 73]. More precisely, the interferometer is described by an operator $\hat{U}$ which acts as

$$\hat{U}\hat{a}^\dagger_{j,\phi_j}\hat{U}^\dagger = \sum_k U_{jk}\hat{a}^\dagger_{k,\phi_j}, \quad \forall j. \tag{60}$$

Following Refs. [2, 3], it can be shown that the probabilities of the different outcomes of the linear interferometric process not only depend on the unitary U but also on the *distinguishability matrix*, defined as

$$S_{ij} = \langle\phi_i|\phi_j\rangle\,. \tag{61}$$

This is a $n \times n$ Gram matrix constructed from all possible overlaps of the internal wave functions of the input photons. In particular, $S = \mathbb{1}$ if all photons are fully distinguishable, while $S = \mathbb{E}$ (with $\mathbb{E}_{i,j} = 1$ for all i, j) in the case where they are fully indistinguishable [1]. Intermediate situations between these two extreme cases are called *partially distinguishable.*

In order to compute the probability that all n photons are found in a subset $\mathcal{K}$ of the output modes, which we refer to as the *multimode bunching probability* $P_n(S)$, it is useful to define the matrix

$$H_{a,b} = \sum_{l\in\mathcal{K}} U^*_{l,a}U_{l,b}, \tag{62}$$

where $a, b \in \{1, ..., n\}$. For a fixed interferometer U and subset $\mathcal{K}$, the multimode bunching probability is a function of the distinguishability matrix S and can be expressed as

$$P_n(S) = \mathrm{perm}\left(H \odot S^T\right) \tag{63}$$

i.e., the permanent of the Hadamard (or element-wise) product $(H \odot S^T)_{ij} \equiv H_{ij}S_{ji}$ [19] (see also Methods). It is important to remark that H, S and $H \odot S^T$ are all positive semidefinite matrices, which ensures their permanent is positive [21]. Moreover, note that for indistinguishable photons, $P_n^{(\mathrm{bos})} = \mathrm{perm}\,(H)$. Hence, Conjecture 1, when restricted to the setting that we consider (see Fig. 6), takes the following mathematical

[1] The internal wavefunction of each photon may be multiplied by a phase $|\phi_i\rangle \to e^{i\theta_i}|\phi_i\rangle$, which does not affect event probabilities. Hence, there is an equivalence class of distinguishability matrices S for each physical situation as $S_{ij} \equiv e^{i(\theta_j-\theta_i)}S_{ij}$. Here, we represent the equivalence class of fully indistinguishable particles with a single matrix $S = \mathbb{E}$.

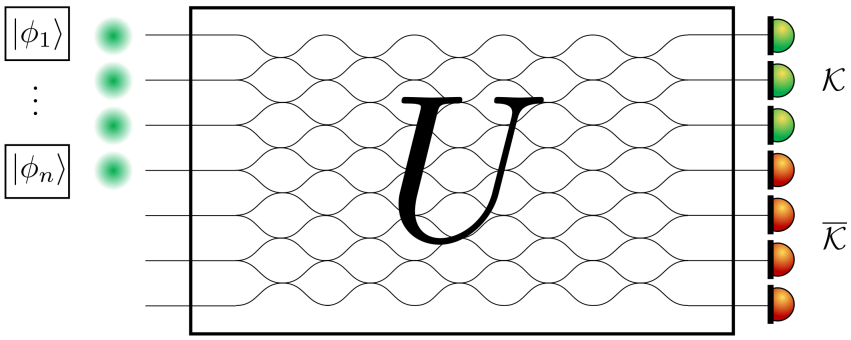


Figure 6: **Interferometric setup.** Single photons are sent through the first n input modes of a m-mode linear interferometer U, which can always be decomposed into a network of two-mode couplers [46]. Each photon at input j carries internal degrees of freedom (polarization, arrival time, etc.) described by an (internal) wave function $|\phi_j\rangle$. Not perfectly overlapping wave functions (measured via the Gram matrix S) give rise to partial distinguishability amongst the photons, reducing the degree of quantum interference and bosonic effects such as bunching (see Conjecture 1). We focus in particular on the probability that all n photons bunch in a subset $\mathcal{K}$ (corresponding to the green detectors) of the output modes, while the red detectors in $\overline{\mathcal{K}}$ do not click. This is a natural extension of the HOM experiment for more than two modes.

form:

$$\operatorname{perm}\left(H \odot S^T\right) \overset{?}{\leq} \operatorname{perm}\left(H\right), \tag{64}$$

with equality holding if S corresponds to indistinguishable photons. The reasons to presume that this conjecture might be an actual physical law governing multiparticle interferences are manifold and, in what follows, we detail several evidences supporting this hypothesis.

4.1.1 Single-mode bunching.

If the subset $\mathcal{K}$ is a single output mode, then the conjecture holds. In this case, if we choose $\mathcal{K} = \{1\}$, the single-mode bunching probability is given by

$$P_n(S) = \prod_{j=1}^{n} |U_{1j}|^2 \operatorname{perm}\left(S\right) = P_n^{(\text{dist.})} \operatorname{perm}\left(S\right), \tag{65}$$

where $P_n^{(\text{dist.})} = P_n(\mathbb{1})$ corresponds to the multimode bunching probability when the photons are fully distinguishable. Then, $P_n(S)$ is indeed maximum for fully indistinguishable photons since the maximum value of $\operatorname{perm}\left(S\right)$ is attained when $S = \mathbb{E}$, with $\operatorname{perm}\left(\mathbb{E}\right) = n!$. This is also a reason why $\operatorname{perm}\left(S\right)$ can be seen as a measure of indistinguishability of the input photons [2, 74]. Note that the celebrated HOM effect can be simply recovered from Eq. (65): for a two-mode 50:50 beam-splitter, the maximum probability of bunching in a single output mode is attained for fully indistinguishable photons and given by $1/4 \times 2! = 1/2$.

In addition, even if we take $|\mathcal{K}| > 1$ but restrict to compare fully indistinguishable

with fully distinguishable photons, it appears that the multimode bunching probability satisfies $P_n(\mathbb{E}) \geq P_n(\mathbb{1})$, see [19], further suggesting that any partial distinguishability is bound to decrease bunching effects.

4.1.2 Fermion antibunching.

Another physical motivation for Conjecture 1 is the fact that an analogous statement on fermion antibunching can be proved. Indeed, an input state of n fully indistinguishable fermions *minimizes* the probability that all of the n fermions are bunched in any subset $\mathcal{K}$ of the output modes. This can be shown using a famous result by Schur [75, 19]. In the setting we consider, the fermionic multimode bunching probability obeys

$$P_n^{(\mathrm{ferm})}(S) = \det(H \odot S^T) \geq \det H, \tag{66}$$

which follows from the Oppenheim inequality for determinants [76] stating that for any two positive semidefinite matrices A and B we have

$$\det(A \odot B) \geq \det A \det B. \tag{67}$$

4.1.3 Bapat-Sunder conjecture.

In 1985, Bapat and Sunder have questioned whether an analogue to the Oppenheim inequality holds for permanents [23]. They conjectured the following:

Conjecture 2 (Bapat-Sunder) For any two positive semi-definite $n \times n$ matrices $A = (a_{ij})$ and $B = (b_{ij})$, we have

$$\mathrm{perm}\,(A \odot B) \leq \mathrm{perm}\,(A) \prod_{i=1}^{n} b_{ii}.$$

It is easy to see that if this statement was valid, it would imply Eq. (64), thus confirming the validity of Conjecture 1 in the specific setting of Fig. 6 [19].

Counterexample 1 (Drury)**.** The Bapat-Sunder conjecture was recently disproved by Drury, who found a 7-dimensional counterexample [71]. It consists in a positive semidefinite matrix A of dimension 7, whose diagonals are $a_{ii} = 1$, which is such that

$$\frac{\mathrm{perm}\left(A \odot A^T\right)}{\mathrm{perm}\,(A)} = \frac{1237}{1152} \approx 1.07, \tag{68}$$

thus implying a violation of Conjecture 2. Instead of presenting the matrix A, it will be useful to consider its Cholesky decomposition, which can always be found for any positive semi-definite matrix [77]. Thus, we write $A = M^\dagger M$, with

$$M = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{2} & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & \sqrt{2} & 1 & \omega^1 & \omega^2 & \omega^3 & \omega^4 \end{pmatrix} \tag{69}$$

where $\omega = \exp(2i\pi/5)$ is the fifth root of unity. As we shall see, the existence of

this counterexample implies that there are 7-dimensional matrices H and S for which Eq. (64) is false, hence contradicting Conjecture 1.

Consequently, in spite of the compelling evidences listed above suggesting that multimode bunching should be maximum for indistinguishable bosons, Drury's counterexample allows us to predict the existence of *boosted boson bunching* with partially distinguishable bosons. Before turning to the optical realization and physical mechanism behind this counterintuitive phenomenon, we stress that brute-force numerical trials really seem to support the (now proven wrong) Conjectures 1 and 2. In particular, we generated 10^7 samples of dimension $n = 7$ and of rank $r = 2$ with the following physically-inspired procedure: a unitary U is sampled randomly according to the Haar measure and used to compute the matrix H. The matrix S is constructed by taking random normalized vectors in a space of dimension r. We did not encounter a single counterexample, which seems to imply that Conjecture 1 holds in practically all cases, even when restricting to a dimension and rank where we know that a counterexample actually exists. Numerical trials of Conjecture 2 with two random Gram matrices also gave similar results. These observations are corroborated by the numerical searches reported in Refs. [19, 72] and thus make the finding of enhanced boson bunching via partial distinguishability even more surprising. Arguably, the violation of the extended Conjecture 1 for entangled input photons observed in Ref. [72] could be viewed as an instance of enhanced boson bunching, but these counterexamples do not break any common assumption on multi-photon interference and it is unclear whether such states can be prepared experimentally using linear interferometry and currently available photon sources.

4.2 Optical realization of boosted bunching

As we now prove, Drury's counterexample entails the existence of a physical experiment that violates the generalized bunching conjecture. In Fig. 7 (left panel), we present a possible optical setup realizing this violation and involving 7-photon interferometry. The details on how to construct the internal states of the photons $|\varphi_j\rangle$ and the unitary matrix U to obtain the desired H and S matrices are explained in Methods. In a nutshell, the states $|\varphi_j\rangle$ can be read from the columns of the matrix M in Eq. (69), while U is chosen in order to contain a rescaled version of $M^\dagger$ as a submatrix.

The resulting setup is surprisingly simple: the 7-mode interferometer U is constructed with a 5-mode Discrete-Fourier Transform (DFT) supplemented with two additional beam splitters (with the same transmittance $\eta = 2/7$). Since S is of rank 2, the internal states live in a two-dimensional Hilbert space, hence it is most natural to use photon polarization, which can easily be manipulated via waveplates (other encodings would also be possible, such as time-bin encoding).

A single photon is sent in each of the 7 input modes with an appropriately chosen polarization state, making the 7 photons partially distinguishable. The polarization pattern is shown in Fig. 7 (right panel) and can be viewed as a 5-star polarization state denoted as $\star$ (for the 5 photons entering the DFT), supplemented with a horizontally-

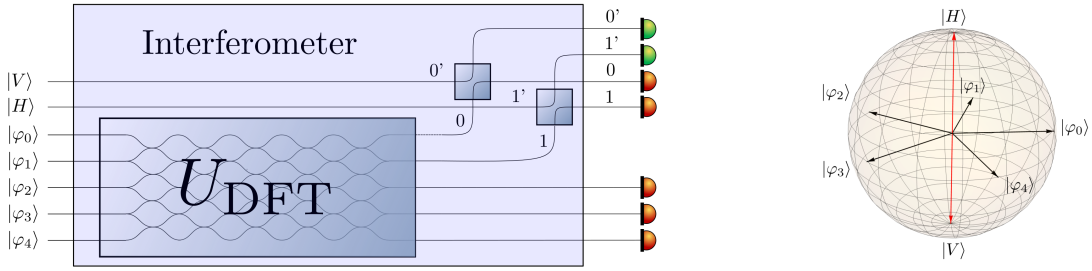


Figure 7: **Boosted two-mode bunching.** (Left) A 7-mode interferometer which violates the generalized bunching conjecture for an appropriate input polarization pattern (depicted in the right panel). Here, 7 photons are sent into the 7 input modes and the probability of detecting them all in the two output modes indicated with green detectors exceeds its value for indistinguishable photons (all with the same polarization). We assume the action of the interferometer is polarization independent, see Eq. (60). This setup can be generalized to n modes as follows. Defining $q = n - 2$, we first send q photons in a q-mode Discrete Fourier Transform (DFT) interferometer $U_{jk} = \frac{1}{\sqrt{q}}\omega^{jk}$ with polarization states $|\varphi_j\rangle = \frac{1}{\sqrt{2}}(|H\rangle + \omega^j |V\rangle)$, where $j, k = 0, \ldots, q-1$ and $\omega = \exp(2i\pi/q)$. The upper two output modes (labelled 0 and 1) of the DFT are then sent to two beam splitters of equal transmittance $\eta = 2/n$, achieving interference respectively with a vertically-polarized photon (in mode 0') and horizontally-polarized photon (in mode 1'). We measure the bunching of all n photons in the subset $\mathcal{K}$ corresponding to the output modes $0'$ and $1'$ indicated with green detectors, thus all red detectors do not click. For $n \geq 7$, we observe a boosted 2-mode bunching probability by comparison with indistinguishable photons. (Right) Bloch-sphere representation of the input polarization pattern for $n = 7$. The polarization states of the 5 input photons of the DFT (indicated as black arrows) are equally spaced along the equator of the Bloch sphere. We call this special state a 5-star polarization state (or q-star polarization state for general q) and denote it as $\star$. The two extra photons (indicated as red arrows) have antipodal – horizontal and vertical – polarization states.

polarized and a vertically-polarized state. The probability of detecting all 7 photons in the output bin (modes labelled by $0'$ and $1'$) is then given by $P_7^{(\star)} \approx 7.5 \times 10^{-3}$. By comparison, for fully indistinguishable photons (all with the same polarization), this probability is only $P_7^{(\mathrm{bos})} \approx 7 \times 10^{-3}$. This simple experimental setup thus exhibits a bunching violation greater than 7%, in accordance with Eq. (68). The photon number distribution in modes $0'$ and $1'$ is depicted in Fig. 8 for different scenarios (fully indistinguishable, partially distinguishable, and fully distinguishable photons).

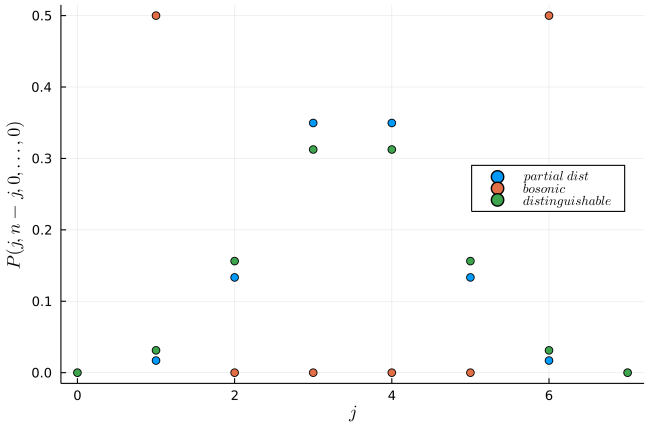


Figure 8: Photon-number probability distributions in mode $0'$ at the output of the 7-mode circuit of Fig. 7 for different scenarios. The distributions are normalized, i.e., the probabilities are conditioned on events where all 7 photons end up in modes 0' and 1' (two-mode bunching events). Due to the symmetry of the circuit, the probability of event $(j, 7-j)$ is the same as event $(7-j, j)$. For fully indistinguishable bosons (red points), most events are interferometrically suppressed, originating from the fact that the output of the 5-mode DFT is a NOON state, see Eq. (71). The only surviving events are (1,6) and (6,1). Using partially distinguishable bosons (blue points) erases these destructive interferences and, as proven in this work, enhances the overall two-mode bunching probability. This also leads to a qualitatively very different photon-number distribution, which has a Bell-like shape. Interestingly, fully distinguishable particles (green points) lead to a photon-number distribution of similar shape but a significantly smaller two-mode bunching probability. The two-mode bunching probabilities in the three cases are $P_7^{(\mathrm{bos})} \approx 7 \times 10^{-3}$, $P_7^{(\star)} \approx 7.5 \times 10^{-3}$ and $P_7^{(\mathrm{dist.})} \approx 1.5 \times 10^{-4}$, respectively.

The outstanding recent progress in boson sampling experiments indicates that the experimental observation of such a boosted bunching due to partial distinguishability should be possible with present-day technology [78, 79, 80, 42, 81, 82]. For example, 5-photon coincidence rates in the hundreds of Hz range have already been demonstrated in optical circuits of a much bigger size that the one represented in Fig. 7 [42]. Moreover, optical implementations of the DFT of dimension 6 and 8 have already been reported [83, 84]. The required photon number resolution (up to 7 photons) could be achieved with single-photon resolving detectors by first multiplexing the output modes $0'$ and $1'$ into several spatial or temporal modes [85, 51]. In fact, very recently a photon number resolution of up to a hundred photons was achieved [86]. As a further feasibility argument, we note that the experimental scheme realizing this boosted bunching is

stable under small perturbations to the matrix elements of the unitary U as well as to the distinguishability matrix S, see Fig. 9. This is easy to understand intuitively as the permanent is a sum of products of matrix elements, thus smooth and well behaved under Taylor expansion (see Supplementary Information for a formal proof). A fully realistic treatment of how perturbations affect the bunching violation would depend on the particular details of the physical implementation of the scheme and is out of the scope of the current work.

4.3 Physical mechanism and asymptotically large violations

The interferometer shown in Fig. 7 arguably provides only a small relative violation of the generalized bunching conjecture, as seen from Eq. (68). It is natural to ask whether larger relative violations can be obtained and what would be the corresponding physical set-up. Moreover, from a physics perspective, it is important to pinpoint the underlying mechanism that explains the violation, at least in some particular setting. We answer these questions and find a way to generalize the 7-mode circuit into an n-mode circuit, as described in the caption of Fig. 7. For this family of circuits, we show that the ratio of the bunching probabilities, which we refer to simply as the *bunching violation ratio* R_n, obeys the following bound

$$R_n = \frac{P_n^{(\star)}}{P_n^{(\mathrm{bos})}} \geq \frac{n}{8} + \frac{1}{32}\frac{(n-2)^2}{n-1}, \tag{70}$$

for any $n \geq 4$. Hence, partial distinguishability can lead to an asymptotically larger multimode bunching probability with respect to fully indistinguishable bosons. Incidentally, this family of circuits also yields some previously unreported family of H and S matrices that violate the Bapat-Sunder conjecture. While a detailed derivation of this bound is given in Methods 4.5.3, we present here the main arguments by comparing the physical mechanism of bunching for fully indistinguishable and partially distinguishable photons.

4.3.1 Fully indistinguishable photons.

The first step of the argument is to note that the only possibility for all of the n photons to be observed in the subset $\mathcal{K}$ is if there are $q = n - 2$ photons in the first two output modes of the DFT interferometer of dimension q and vacuum on the rest (we assume $q \geq 2$). The corresponding conditional (subnormalized) state of these two output modes is given by the NOON state [87]

$$\begin{aligned}\left|\psi_{\mathrm{out}}^{(\mathrm{bos})}\right\rangle &= \frac{1}{q^{q/2}}\prod_{j=0}^{q-1}\left(\hat{a}_0^\dagger + \omega^j\,\hat{a}_1^\dagger\right)|0\rangle \\ &= \frac{1}{q^{q/2}}\left((\hat{a}_0^\dagger)^q + (-1)^q(\hat{a}_1^\dagger)^q\right)|0\rangle\,,\end{aligned} \tag{71}$$

where $\omega = \exp(2i\pi/q)$ is the qth root of unity. The probability of having q photons in these two modes is simply given by the square norm of this state, namely $2q!/q^q$. The next part of the circuit realizes the interference between modes 0 and $0'$ via beam splitter $\hat{U}_{\mathrm{BS}}^{0,0'}$, as well as between modes 1 and $1'$ via beam-splitter $\hat{U}_{\mathrm{BS}}^{1,1'}$. The action of these beam-splitters on state $\hat{a}_{0'}^\dagger \hat{a}_{1'}^\dagger \left|\psi_{\mathrm{out}}^{(\mathrm{bos})}\right\rangle$ followed by postselection on vacuum in both output modes 0 and 1 is analyzed in Methods. The resulting probability is governed by the bunching mechanism sketched in the upper part of Fig. 10, where η denotes the transmittance of the two beam splitters. The first term of state (71) describing q photons in mode 0 undergoes bunching with the extra photon in mode 0' with probability $(q+1)\,\eta(1-\eta)^q$, while the extra photon in mode 1' is simply transmitted with probability η. The second term of state (71) behaves similarly. Consequently, the multimode bunching probability (in output modes $0'$ and $1'$) is given by

$$P_n^{(\mathrm{bos})} = \frac{2\,(q+1)!}{q^q}\,\eta^2(1-\eta)^q \tag{72}$$

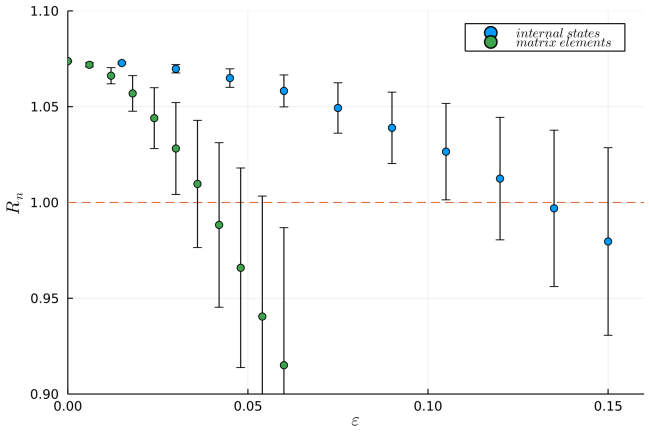


Figure 9: Perturbation effects on the bunching violation ratio R_n (vertical axis) for the setup of Fig. 7 with $n = 7$. (Blue) Perturbation of the internal wave functions defined in the caption of Fig. 7 by a random amount drawn from a Gaussian distribution with zero mean and standard deviation ϵ (horizontal axis). For each ϵ value, 10^4 samples are taken. We observe violations of Conjecture 1 (on average) up to $\epsilon_{max} \approx 0.135$, which corresponds to the components of the internal wave functions being perturbed by about $\epsilon_{max}/(1+\epsilon_{max}^2) \approx 13.3\%$. The vertical bars represent the standard deviation for each ϵ value. (Green) Perturbation of the matrix elements U_{ij} of the 7-mode interferometer. The same random Gaussian perturbations are added to the columns of the matrix U, which are then Gram-orthonormalized. In this case, the matrix elements can be perturbed by about $\epsilon_{max} \approx 0.039$ to still exhibit a violation (on average). The optical scheme is thus resilient to perturbations both to the internal states of the photons and to the interferometer, making a good case for its experimental feasibility.

4.3.2 Partially distinguishable photons.

We consider the following q-star polarization pattern

$$\left|\psi_{\text{in}}^{(\star)}\right\rangle = \frac{1}{2^{q/2}} \prod_{j=0}^{q-1} \left(\hat{a}_{h,j}^{\dagger} + \omega^{j}\, \hat{a}_{v,j}^{\dagger}\right) |0\rangle\,, \tag{73}$$

as the input state sent to the DFT interferometer. Here, $\hat{a}_{h,j}^{\dagger}$ ($\hat{a}_{v,j}^{\dagger}$) are creation operators of a photon in spatial mode j and horizontal (vertical) polarization. This state is a generalization of the 5-star polarization pattern shown in Fig. 7 (right). In order to understand why multimode bunching is boosted with this special input state, it is convenient to define the spatio-polarization modes

$$\hat{c}_{\pm}^{\dagger} = \frac{\hat{a}_{h,1}^{\dagger} \pm \hat{a}_{v,0}^{\dagger}}{\sqrt{2}}. \tag{74}$$

Similarly as for indistinguishable photons, we compute the conditional output state of the DFT interferometer that contains q photons in the first two output modes (and vacuum for the rest) for input state (73), namely

$$\left|\psi_{\text{out}}^{(\star)}\right\rangle = \frac{1}{q^{q/2}} \prod_{j=0}^{q-1} \left(\frac{\hat{a}_{h,0}^{\dagger}}{\sqrt{2}} + \omega^{j}\, \hat{c}_{+}^{\dagger} + \omega^{2j}\, \frac{\hat{a}_{v,1}^{\dagger}}{\sqrt{2}}\right) |0\rangle\,. \tag{75}$$

This polarization multi-mode state is the counterpart of the NOON state, Eq. (71). It comprises many terms, most importantly the one containing q photons in mode $\hat{c}_{+}^{\dagger}$, namely

$$\left|\psi_{\text{out}}^{(\star)}\right\rangle = \frac{(-1)^{q+1}}{q^{q/2}} (\hat{c}_{+}^{\dagger})^{q} |0\rangle + \dots. \tag{76}$$

In what follows, we show that this term alone is enough to prove that the bunching violation ratio in Eq. (70) grows at least linearly (see Methods for a more detailed derivation). The other components of state (76) are orthogonal to the state $(c_{+}^{\dagger})^{q} |0\rangle$ and thus can only contribute with additional positive terms to the bunching probability. The probability to have q photons (regardless of their polarization) in the first two modes is thus lower bounded by $q!/q^{q}$. The subsequent part of the interferometer is fed with the state

$$\hat{a}_{h,1'}^{\dagger}\, \hat{a}_{v,0'}^{\dagger} \left|\psi_{\text{out}}^{(\star)}\right\rangle = \frac{(\hat{c'}_{+}^{\dagger})^{2} - (\hat{c'}_{-}^{\dagger})^{2}}{2} \left|\psi_{\text{out}}^{(\star)}\right\rangle, \tag{77}$$

where we have described the two extra photons with antipodal polarization (H and V) using the spatio-polarization modes

$$\hat{c'}_{\pm}^{\dagger} = \frac{\hat{a}_{h,1'}^{\dagger} \pm \hat{a}_{v,0'}^{\dagger}}{\sqrt{2}}, \tag{78}$$

defined in analogy with Eq. (74). Thus, the leading term of the final output state of the interferometer is

$$\frac{(-1)^{q+1}}{2\,q^{q/2}}\,\hat{V}\left((\hat{c'}_+^\dagger)^2-(\hat{c'}_-^\dagger)^2\right)(\hat{c}_+^\dagger)^q\,|0\rangle+\ldots, \tag{79}$$

where we have defined $\hat{V}=\hat{U}_{BS}^{0,0'}\,\hat{U}_{BS}^{1,1'}$ as the operator describing the joint operation of the two beam-splitters in both polarization (see Methods). Using the fact that the beam splitters have the same transmittance η, it appears that $\hat{V}$ also acts as a beam splitter of transmittance η that couples modes $\hat{c}_+^\dagger$ and $\hat{c'}_+^\dagger$ [2]. We thus have interference between q photons in mode $\hat{c}_+^\dagger$ (occurring with probability $q!/q^q$) and two photons in mode $\hat{c'}_+^\dagger$ (occurring with probability $1/2$). Hence, the resulting probability of obtaining $n=q+2$ photons in output mode $\hat{c'}_+^\dagger$ (which leads to the detection of n photons in the output bin $\mathcal{K}=\{0',1'\}$) is governed by the double-bunching mechanism sketched in the lower part of Fig. 10, associated with probability $\binom{q+2}{2}\eta^2(1-\eta)^q$. As a result, we obtain the following bound on the 2-mode bunching probability in the case of partially distinguishable photons

$$P_n^{(\star)}\geq\frac{(q+2)!}{4\,q^q}\,\eta^2(1-\eta)^q. \tag{80}$$

This probability is asymptotically larger than its counterpart for fully indistinguishable photons, Eq. (72). Using Eqs. (72) and (80), we indeed obtain a lower bound on the bunching violation ratio

$$R_n=\frac{P_n^{(\star)}}{P_n^{(\mathrm{bos})}}\geq\frac{q+2}{8}=\frac{n}{8}\,, \tag{81}$$

which confirms the dominant term in Eq. (70) and shows that it grows at least linearly with n. A more detailed calculation given in Methods leads to the second term in Eq. (70). As seen in Fig. 11, this bound seems to describe well enough the behavior of R_n up to $n=30$. In the special case where $n=7$, we may compute exactly all terms in Eq. (76), which gives $R_7=1237/1152$, in perfect agreement with Eq. (68). Note also that Eq. (70) shows no dependence on the transmittance η. However, the absolute probability of bunching events can be maximized by maximizing the term $\eta^2(1-\eta)^q$ over η, which yields $\eta=2/n$ as mentioned in the caption of Fig. 7.

4.4 Discussion and outlook

The complex behavior of interferometric experiments with multiple partially distinguishable photons has been explored in several theoretical and experimental works [88, 18, 1, 89, 90, 91, 92, 93, 94], revealing that many-body interference does not reduce to a simple dichotomy between distinguishable and indistinguishable photons. This is evident, for example, from the fact that certain outcome probabilities do not behave monotonically as one makes photons more distinguishable [88, 18]. However, the scheme

[2] The beam splitter $\hat{V}$ also couples $\hat{c}_-^\dagger$ and $\hat{c'}_-^\dagger$ but we disregard the corresponding term here as its contribution is much smaller and we seek a lower bound on the probability, see Methods.

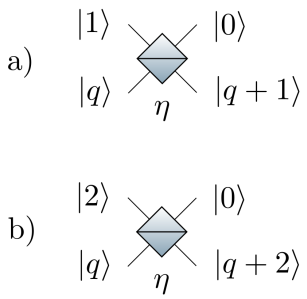


Figure 10: Mechanism at the origin of boosted bunching. (a) For indistinguishable photons, the extra photon in mode $\hat{a}^\dagger_{0'}$ (or $\hat{a}^\dagger_{1'}$) bunches with q photons in mode $\hat{a}^\dagger_0$ (or $\hat{a}^\dagger_1$) coming from the NOON state (71). (b) For partially distinguishable photons, the largest term contributing to the bunching probability (80) comes from the *double bunching* of two photons in the delocalized mode $\hat{c'}^\dagger_+$ [see Eq. (78)] with q photons in the delocalized mode $\hat{c}^\dagger_+$ [see Eq. (74)]. The asymptotics of the bunching violation ratio R_n as shown in Eq. (70) originates from the probabilities of the processes depicted here.

of Fig. 7 is the first explicit setup showing that boson bunching can be boosted by partial distinguishability to the point where it actually beats ideal (fully indistinguishable) bosons. This disproves the common belief that bunching effects are necessarily maximized in this ideal scenario.

It is intriguing to observe that the state of partially distinguishable photons we have found to exhibit boosted bunching is, in a sense, *far* from the state of fully indistinguishable photons. This can be seen by computing the relative contribution of the fully (permutation-) symmetric component of the internal wavefunction $|\Phi\rangle = |\phi_1\rangle\,|\phi_2\rangle \ldots |\phi_n\rangle$ [74]

$$d(S) = \left\langle \Phi \middle| \hat{\mathcal{S}}_n \middle| \Phi \right\rangle = \frac{\mathrm{perm}\,(S)}{n!}, \tag{82}$$

where $\hat{\mathcal{S}}_n = (1/n!)\sum_{\sigma\in S_n}\hat{P}_\sigma$ is the symmetrizer in n dimensions. This measure is 1 for fully indistinguishable photons ($S = \mathbb{E}$) and $1/n!$ for fully distinguishable ones ($S = \mathbb{1}$). For the simplest case of seven photons, we have that

$$\begin{aligned} d(\mathbb{1}) &= \frac{1}{7!} \approx 1.98 \times 10^{-4} \\ d(S^{(\star)}) &= \frac{45}{7!} \approx 8.93 \times 10^{-3} \ll 1 \end{aligned} \tag{83}$$

where $S^{(\star)}$ is the Gram matrix of the partially distinguishable polarization state shown in Fig. 7. This state is thus somehow *closer* from fully distinguishable photons. It is also natural to ask whether there may exist states violating the generalized bunching conjecture already *in the vicinity* of a fully indistinguishable state. We show in Supplementary Information that first-order perturbations around $S = \mathbb{E}$ leave the multimode bunching probability constant, which suggests a negative answer. But the question remains open whether this probability always decreases near $S = \mathbb{E}$ if second-order terms are taken into account, in which case it would be a general feature.

Furthermore, it must be noted that we have modeled distinguishable photons with a two-dimensional internal degree of freedom, namely polarization. In contrast, in the fully distinguishable setting, each photon occupies a different internal state, forming an orthonormal basis of an n-dimensional space. Is it possible to find counterexamples where the internal states live in a larger space, going beyond small perturbations around our conterexamples? This could model realistic situations with photons occupying partly overlapping time-bins, possibly leading to larger than linear asymptotic scalings of the bunching violation ratio. We leave this question for future work.

On a final note, we stress that our findings corroborate the deep connections between bosonic interferences in quantum physics on the one hand, and the algebra of matrix permanents on the other hand. The transposition of Drury's 7-dimensional matrix counterexample into a quantum interferometric experiment has inspired us to find a new family of n-dimensional matrices that not only violate the Bapat-Sunder conjecture but also exhibit a relative violation increasing with n. We anticipate that other mathematical conjectures on permanents may be addressed by exploiting this fruitful interplay with physics-inspired mechanisms such as those shown in Fig. 10. This may even help solve other questions on the Bapat-Sunder conjecture [33]. For example, the smallest known counterexample is 7-dimensional, so it would be interesting to find a simpler counterexample if it exists, or show that this is not possible. Another open question in matrix theory is whether there is a counterexample involving a real matrix of dimension smaller than 16 [95], which may be resolved by considering interferometry within real quantum mechanics.

Overall, we hope this work will open new paths to explore the connection between distinguishability and boson bunching, leading not only to a better understanding of multiparticle quantum interference but perhaps also to novel applications of partially distinguishable photons to quantum technology.

4.5 Methods

4.5.1 Bunching probability

In this section we summarize the main steps needed to derive Eq. (63). Following the colloquial conventions of [96], we consider n photons sent through a (m, m) linear interferometer described by the unitary matrix U. We limit ourselves to at most one photon per input mode. Without loss of generality, we consider that the photons occupy the first n input modes. We denote the vector of occupation number of the output modes as $\boldsymbol{s} = (s_i)$ where $0 \leq s_i \leq n$ is the number of photons in output mode i. Naturally $\sum s_i = n$. We define the mode assignment list $\boldsymbol{d} = \boldsymbol{d}(\boldsymbol{s}) = \oplus_{j=1}^{m} \oplus_{k=1}^{s_j} (j)$. For instance if $\boldsymbol{s} = (2, 0, 1)$ then $\boldsymbol{d} = (1, 1, 3)$.

Consider the probability $P(\boldsymbol{d})$ that the photons give an output configuration $\boldsymbol{s}$ with a mode assignment list $\boldsymbol{d} = \boldsymbol{d}(\boldsymbol{s})$. Tichy shows that this probability can be expanded as

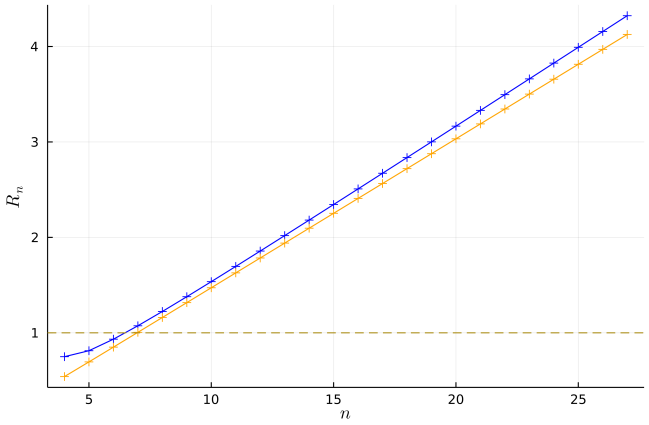


Figure 11: Bunching violation ratio R_n for the family of optical schemes shown in Fig. 7 as a function of the size n (blue). For $n \geq 7$, it appears that partially distinguishable particles outperform indistinguishable bosons as witnessed by $R_n > 1$. The lower bound given by Eq. (70) is also shown (orange).

a multi-dimensional tensor permanent [2]

$$P(\boldsymbol{d}) = \frac{1}{\mu(\boldsymbol{s})} \sum_{\sigma,\rho \in S_n} \prod_{j=1}^{n} \left(U_{\sigma_j,d_j} U^*_{\rho_j,d_j} S_{\rho_j,\sigma_j} \right) \tag{84}$$

with $\mu(\boldsymbol{s}) = \prod_{j=1}^{m} s_j!$.

Let us now compute the probability $P_n(S)$ that all n photons bunch in a subset $\mathcal{K}$ of the output modes, for a set interferometer and a Gram matrix S, following a derivation of Shchesnovich [19]. Without loss of generality, consider that the subset $\mathcal{K}$ is the first $K = |\mathcal{K}|$ output modes. The bunching probability is the sum over all event probabilities $P(\boldsymbol{d})$ with $s_i = 0$ for all $i > K$

$$P_n(S) = \frac{1}{n!} \sum_{d_1=1}^{K} \cdots \sum_{d_n=1}^{K} \sum_{\sigma,\rho \in S_n} \prod_{j=1}^{n} \left(U_{\sigma_j,d_j} U^*_{\rho_j,d_j} S_{\rho_j,\sigma_j} \right) \tag{85}$$

Now, calling

$$H_{a,b} = \sum_{i=1}^{K} U^*_{i,a} U_{i,b} \tag{86}$$

we can rewrite

$$P_n(S) = \sum_{\sigma' \in S_n} \prod_{j=1}^{n} \left(H_{j,\sigma'_j} S^T_{j,\sigma'_j} \right) = \mathrm{perm}\left(H \odot S^T \right) \tag{87}$$

where the last quantity is the permanent of the Hadamard (or elementwise) product $(H \odot S^T)_{ij} \equiv H_{ij} S_{ji}$. Note that for indistinguishable particles $S_{ij} = 1$, $\forall i,j$, so that $P_n(\mathbb{E}) = \mathrm{perm}\,(H)$.

4.5.2 Physical realization of violating matrices

It is possible to show that any counterexample to the Bapat-Sunder conjecture can be used to construct a physical interferometer U and a set of internal states of the photons $\{|\phi_i\rangle\}$ that provide a counterexample to the generalized bunching conjecture. We assume, without loss of generality [97], a simplified form of the Bapat-Sunder conjecture, where A and B are Gram matrices and so $a_{ii} = b_{ii} = 1, \forall i \in \{1, ..., n\}$. In this case, Conjecture 2 takes the form $\mathrm{perm}\,(A \odot B) \leq \mathrm{perm}\,(A)$. Since the distinguishability matrix S is a Gram matrix, we can choose $S^T = B$. The set of quantum states realizing any given distinguishability matrix can be obtained from its Cholesky decomposition

$$B = M'^{\dagger} M'. \tag{88}$$

The matrix M' is of size $r' \times n$ where r' is the rank of B. The n internal photon states $\{|\phi_i\rangle\}$ that realize this Gram matrix can be read out from the columns of M'. Thus, the rank of B determines the dimension of the Hilbert space spanned by the states $\{|\phi_i\rangle\}$. For the physical realization of Drury's counterexample we chose $S = A = M^{\dagger} M$, with M given in Eq. (69). This implies that the 7 internal states of the bosons live in a two-dimensional space, where each state is obtained from each column of M .

In addition, it is always possible to construct an interferometer U such that $H = \alpha A$, where α is a positive rescaling factor such that $\alpha \leq 1$. Note that this rescaling is not important when it comes to showing a violation of the generalized bunching conjecture [3]. We can write the Cholesky decomposition of αA as

$$H_{a,b} = \alpha \sum_{k=1}^{r} M^{\dagger}_{a,k} M_{k,b} = \alpha \sum_{k=1}^{r} M^{\dagger}_{a,k} \left(M^{\dagger}_{b,k} \right)^{*}, \tag{89}$$

where r is the rank of A. By comparing with the definition of the matrix H in Eq. (214), it is possible to see that we obtain $H = \alpha A$ if we appropriately incorporate the matrix $\sqrt{\alpha} M^{\dagger}$ as a submatrix of U, for example, in the top left corner. This choice determines that the subset $\mathcal{K}$ is given by the first r output modes. Note also that it is always possible to incorporate an arbitrary complex matrix, up to renormalization, into a bigger unitary matrix, using arguments similar to Lemma 29 of Ref. [31].

In the case discussed in the main text, the aim is to construct an interferometer U which contains a rescaled version of the 7×2 matrix $\sqrt{\alpha} M^{\dagger}$, where M is given in Eq. (69). Here, the procedure to construct U is simplified by the fact that the columns of $M^{\dagger}$ are already orthogonal vectors. Hence, we can choose $\alpha = 2/7$ to normalize these columns and find 5 other orthonormal vectors to construct a 7×7 unitary matrix. The unitary built from the circuit presented in Fig. 7 gives one possibility to construct such a unitary, which was chosen for its simplicity.

[3] If $\mathrm{perm}\,(A \odot B) \geq \mathrm{perm}\,(A)$ then $\mathrm{perm}\,(\alpha A \odot B) \geq \mathrm{perm}\,(\alpha A)$

4.5.3 Bound on the bunching violation ratio

We consider the circuit described in the caption Fig. 7, which is a generalization of the 7-mode optical circuit depicted in this figure to a circuit of n modes. The circuit is composed by a discrete Fourier transform (DFT) circuit of size $q = n - 2$ applied to the input modes $\{0, 1, ..., q-1\}$ followed by two beam splitters of equal transmittance applied between the modes $0'$ and 0 as well as between the modes $1'$ and 1.

Our quantity of interest is the probability of observing all the n photons in the output modes $0'$ and $1'$. Let us first compute this quantity when the input photons are fully indistinguishable. The quantum state at the output of the DFT is given by

$$\hat{U}_{dft} \left|\psi_{in}\right\rangle = \hat{U}_{dft} \prod_{j=0}^{q-1} \hat{a}_j^\dagger \left|0\right\rangle \tag{90}$$

$$= \frac{1}{q^{q/2}} \prod_{j=0}^{q-1} \left(\sum_{k=0}^{q-1} e^{2\pi i jk/q} \hat{a}_k^\dagger \right) \left|0\right\rangle . \tag{91}$$

The only possibility for all of the n photons to be observed in modes $0'$ or $1'$ at the output of the full circuit is if at the output of the DFT interferometer there are $q = n-2$ photons in modes 0 or 1 and vacuum on the rest. Hence, we only consider the subnormalized component of the wavefunction in these output modes, given by

$$\left|\psi_{\text{out}}^{(\text{bos})}\right\rangle = \frac{1}{q^{q/2}} \prod_{j=0}^{q-1} \left(\hat{a}_0^\dagger + e^{2\pi i j/q} \hat{a}_1^\dagger \right) \left|0\right\rangle \tag{92}$$

To expand this expression, one can think of $\hat{a}_0^\dagger$ and $\hat{a}_1^\dagger$ as complex numbers since these two operators commute. In this sense, following [87], we can consider that each term $\hat{a}_0^\dagger + e^{2\pi i j/q} \hat{a}_1^\dagger$ is an eigenvalue of the circulant matrix of dimension q given by $\text{circ}(\hat{a}_0^\dagger, \hat{a}_1^\dagger, 0, \ldots, 0)$. Hence, the previous equation can be rewritten as

$$\left|\psi_{\text{out}}^{(\text{bos})}\right\rangle = \frac{1}{q^{q/2}} \det(\text{circ}(\hat{a}_0^\dagger, \hat{a}_1^\dagger, 0, \ldots, 0)) \left|0\right\rangle \tag{93}$$

$$= \frac{1}{q^{q/2}} \left((\hat{a}_0^\dagger)^q + (-1)^q (\hat{a}_1^\dagger)^q \right) \left|0\right\rangle , \tag{94}$$

which is a NOON state. The subsequent part of the interferometer couples this state with the ancillary modes $0'$ and $1'$, each containing a single photon. We denote these beam-splitters by $\hat{U}_{BS}^{0,0'}$ and $\hat{U}_{BS}^{1,1'}$ and their transmittance by η. We use the following convention for the unitary representing the action of the beam-splitter

$$U_{BS} = \begin{pmatrix} \sqrt{\eta} & \sqrt{1-\eta} \\ -\sqrt{1-\eta} & \sqrt{\eta} \end{pmatrix} . \tag{95}$$

The joint application of $\hat{U}_{BS}^{0,0'}$ and $\hat{U}_{BS}^{1,1'}$ results in the state

$$\hat{U}_{BS}^{0,0'}\hat{U}_{BS}^{1,1'}\hat{a}_{0'}^{\dagger}\hat{a}_{1'}^{\dagger}\left|\psi_{\text{out}}^{(\text{bos})}\right\rangle \tag{96}$$

$$=\frac{\hat{U}_{BS}^{0,0'}(\hat{a}_{0}^{\dagger})^{q}\hat{a}_{0'}^{\dagger}\hat{U}_{BS}^{1,1'}\hat{a}_{1'}^{\dagger}+(-1)^{q}\hat{U}_{BS}^{1,1'}(\hat{a}_{1}^{\dagger})^{q}\hat{a}_{1'}^{\dagger}\hat{U}_{BS}^{0,0'}\hat{a}_{0'}^{\dagger}}{q^{q/2}}\left|0\right\rangle . \tag{97}$$

The postselection on the component where all photons occupy output modes $\{0',1'\}$ yields

$$\left|\psi_{\text{post}}^{(\text{bos})}\right\rangle=\frac{(1-\eta)^{q/2}\eta}{q^{q/2}}\left((\hat{a}_{0'}^{\dagger})^{q+1}\hat{a}_{1'}^{\dagger}+(-1)^{q}(\hat{a}_{1'}^{\dagger})^{q+1}\hat{a}_{0'}^{\dagger}\right)\left|0\right\rangle \tag{98}$$

Finally, the bunching probability in modes $0'$ and $1'$ is given by

$$P_{n}^{(\text{bos})}=\left\langle\psi_{\text{post}}^{(\text{bos})}\middle|\psi_{\text{post}}^{(\text{bos})}\right\rangle=\frac{2\,(q+1)!}{q^{q}}\,(1-\eta)^{q}\eta^{2}. \tag{99}$$

Consider now the analogous calculation for the specially chosen state of partially distinguishable photons, described in the caption of Fig. 7. In this case, as discussed in the main text, the counterpart of the NOON state obtained in Eq. (94) is given by

$$\left|\psi_{\text{out}}^{(\star)}\right\rangle=\frac{1}{q^{q/2}}\prod_{j=0}^{q-1}\left(\frac{\hat{a}_{h,0}^{\dagger}}{\sqrt{2}}+\omega^{j}\,\hat{c}_{+}^{\dagger}+\omega^{2j}\,\frac{\hat{a}_{v,1}^{\dagger}}{\sqrt{2}}\right)\left|0\right\rangle . \tag{100}$$

After the Fourier interferometer, one ancillary photon is introduced in mode $0'$ with vertical polarization and another one in mode $1'$ with horizontal polarization. At this point, the state of the system is given by

$$\hat{a}_{h,1'}^{\dagger}\,\hat{a}_{v,0'}^{\dagger}\left|\psi_{\text{out}}^{(\star)}\right\rangle=\frac{(\hat{c'}_{+}^{\dagger})^{2}-(\hat{c'}_{-}^{\dagger})^{2}}{2}\left|\psi_{\text{out}}^{(\star)}\right\rangle , \tag{101}$$

with $\hat{c'}_{\pm}^{\dagger}$ defined in Eq. (78). To analyse the action of the subsequent part of the interferometer, it is useful to define the joint action of the beam-splitter operators $\hat{U}_{BS}^{0,0'}$ and $\hat{U}_{BS}^{1,1'}$ as

$$\hat{V}=\hat{U}_{BS}^{0,0'}\hat{U}_{BS}^{1,1'}. \tag{102}$$

Using the fact that both beam-splitters have equal transmittance, it can be seen that the action of $\hat{V}$ on the delocalized modes $\hat{c}_{\pm}^{\dagger}$ and $\hat{c'}_{\pm}^{\dagger}$ is given by

$$\hat{V}\hat{c'}_{\pm}^{\dagger}\hat{V}^{\dagger}=\sqrt{\eta}\hat{c'}_{\pm}^{\dagger}+\sqrt{1-\eta}\hat{c}_{\pm}^{\dagger}, \tag{103}$$

$$\hat{V}\hat{c}_{\pm}^{\dagger}\hat{V}^{\dagger}=\sqrt{\eta}\hat{c}_{\pm}^{\dagger}-\sqrt{1-\eta}\hat{c'}_{\pm}^{\dagger}. \tag{104}$$

We will see that the interference between the modes $\hat{c}_{+}^{\dagger}$, which are occupied in the state $\left|\psi_{\text{out}}^{(\star)}\right\rangle$, with the two bosons in mode $\hat{c'}_{+}^{\dagger}$ from the ancillary photon state will lead to bosonic bunching effects that are responsible for the largest asymptotic contributions to the bunching probability in modes $\{0',1'\}$. In contrast, the other modes occupied in state $\left|\psi_{\text{out}}^{(\star)}\right\rangle$, i.e. $\hat{a}_{h,0}^{\dagger}$ and $\hat{a}_{v,1}^{\dagger}$, do not undergo any enhanced bunching effects since

they do not couple neither to $\hat{c'}_+^\dagger$ nor to $\hat{c'}_-^\dagger$. For completeness, we also write the action of $\hat{V}$ on these modes

$$\hat{V}\hat{a}_{h,0}^\dagger\hat{V}^\dagger = \sqrt{\eta}\hat{a}_{h,0}^\dagger + \sqrt{1-\eta}\hat{a}_{h,0'}^\dagger, \tag{105}$$

$$\hat{V}\hat{a}_{v,1}^\dagger\hat{V}^\dagger = \sqrt{\eta}\hat{a}_{v,1}^\dagger + \sqrt{1-\eta}\hat{a}_{v,1'}^\dagger. \tag{106}$$

We are now ready to analyse the action of $\hat{V}$ on the state given in Eq. (101), with the aim of computing a bound for the bunching probability. For the aforementioned reasons, it will be useful to expand the state $\left|\psi_{\text{out}}^{(\star)}\right\rangle$ as a superposition of states with different occupation numbers in mode $\hat{c}_+^\dagger$. To do so, it is useful to define

$$\hat{B}_j^\dagger = \frac{1}{\sqrt{2}}(\omega^{-j}\hat{a}_{h,0}^\dagger + \omega^j\hat{a}_{v,1}^\dagger). \tag{107}$$

With this definition, we can write

$$\left|\psi_{\text{out}}^{(\star)}\right\rangle = \frac{(-1)^{q-1}}{q^{q/2}}\prod_{j=0}^{q-1}\left(\hat{c}_+^\dagger + \hat{B}_j^\dagger\right)|0\rangle\,. \tag{108}$$

Since all operators involved in this expression commute with each other, we can expand it as if $\hat{c}_+^\dagger$ and $\hat{B}_j^\dagger$ were complex numbers. Precisely, we have the following expansion

$$\prod_{j=0}^{q-1}\left(\hat{c}_+^\dagger + \hat{B}_j^\dagger\right) = \sum_{k=0}^{q}(\hat{c}_+^\dagger)^{q-k}\hat{e}_k(\hat{B}_0^\dagger, ..., \hat{B}_{q-1}^\dagger), \tag{109}$$

where we have defined

$$\hat{e}_0(\hat{B}_0^\dagger, ..., \hat{B}_{q-1}^\dagger) = 1, \tag{110}$$

$$\hat{e}_k(\hat{B}_0^\dagger, ..., \hat{B}_{q-1}^\dagger) = \sum_{0\le j_1<...<j_k\le q-1}\hat{B}_{j_1}^\dagger ... \hat{B}_{j_k}^\dagger\,. \tag{111}$$

To simplify the notation we denote $\hat{e}_k(\hat{B}_0^\dagger, ..., \hat{B}_{q-1}^\dagger)$ simply as $\hat{e}_k$. Newton's identities give us the following recursion relation

$$k\hat{e}_k = \sum_{i=1}^{k}(-1)^{i-1}\hat{e}_{k-i}\;\hat{p}_i, \tag{112}$$

where $\hat{p}_k$ is the k-th power sum

$$\hat{p}_k = \sum_{l=0}^{q-1}(\hat{B}_l^\dagger)^k. \tag{113}$$

These sums take a simple form for any $k \geq 1$, with

$$\hat{p}_k = \begin{cases} 0 \quad \text{for } k \text{ odd}, \\ \frac{q}{2^{k/2}} \binom{k}{k/2} (\hat{a}^\dagger_{h,0} \hat{a}^\dagger_{v,1})^{k/2}, \quad \text{for } k \text{ even}. \end{cases} \tag{114}$$

Using this expression for $\hat{p}_k$ together with Eq. (112), it can be seen that all the terms with odd k in the expansion given in Eq. (109) are suppressed. Moreover, these equations provide a simple way to calculate the first few terms of the expansion in Eq. (109) and obtain

$$\left|\psi^{(\star)}_{\text{out}}\right\rangle = \frac{(-1)^{q-1}}{q^{q/2}} \left((\hat{c}^\dagger_+)^q - \frac{1}{2} q (\hat{c}^\dagger_+)^{q-2} \hat{a}^\dagger_{h,0} \hat{a}^\dagger_{v,1} + ... \right) |0\rangle \,. \tag{115}$$

The other terms of the expansion are orthogonal to the first two terms and can only contribute with additional positive terms to the bunching probability. In fact, these two terms are enough to obtain the lower bound for the bunching violation ratio presented in the main text (Eq. (70)). After the action of the interferometer $\hat{V}$ it can be shown that component of the wavefunction containing all the n photons in modes $0'$ and $1'$ is given by

$$\begin{aligned} &\left|\psi^{(\star)}_{\text{post}}\right\rangle = \\ &= \frac{(1-\eta)^{q/2}\eta}{2q^{q/2}} \left((\hat{c'}^\dagger_+)^{q+2} - \frac{1}{2} q (\hat{c'}^\dagger_+)^q \hat{a}^\dagger_{h,0'} \hat{a}^\dagger_{v,1'} + ... \right) |0\rangle \,. \end{aligned} \tag{116}$$

The omitted terms in the previous equation are orthogonal to the first two terms. Hence, we obtain the following lower bound for the bunching probability

$$P^{(\star)}_n = \left\langle \psi^{(\star)}_{\text{post}} \middle| \psi^{(\star)}_{\text{post}} \right\rangle \geq \frac{\eta^2 (1-\eta)^q}{4\, q^q} \left((q+2)! + \frac{1}{4} q^2 q! \right) . \tag{117}$$

Finally, we can use Eqs. (117) and Eq. (99) to obtain

$$R_n = \frac{P^{(\star)}_n}{P^{(\text{bos})}_n} \geq \frac{q+2}{8} + \frac{1}{32} \frac{q^2}{q+1} \tag{118}$$

$$\geq \frac{n}{8} + \frac{1}{32} \frac{(n-2)^2}{n-1}, \tag{119}$$

demonstrating the bound on the bunching violation ratio given in Eq. (70).

Code availability

The following project relies on the packages Permanents.jl and BosonSampling.jl [98], and the source code generating the figures and data of this paper is available on github and the data at OSF.

Appendix

4.6 Resilience to perturbations

It is natural to ask whether small perturbations of the interferometer U or the partial distinguishability matrix S would immediately negate the violation of the generalized bunching conjecture, or, equivalently of the Bapat-Sunder's conjecture, in which case the discussion would be physically insignificant. However it is easy to prove that this is not the case, as we now show.

We consider for instance a perturbation of the matrix A in $\mathrm{perm}\,(A \odot B) > \mathrm{perm}\,(A)$, the reasoning being similar if we also consider perturbations in the matrix B. We consider the perturbation $A \to (1-\epsilon)A + \epsilon\Delta$ with $\Delta_{i,i} = 1, \Delta_{i,j} = \mathcal{O}(1)$, with $\epsilon \ll 1$. Note that

$$\begin{aligned}
&\mathrm{perm}\,((1-\epsilon)A + \epsilon\Delta) \\
&= \mathrm{perm}\left((1-\epsilon)\left(A + \frac{\epsilon}{(1-\epsilon)}\Delta\right)\right) \\
&= (1-\epsilon)^n \mathrm{perm}\,(A + \delta\Delta)
\end{aligned} \tag{120}$$

setting $\delta = \frac{\epsilon}{(1-\epsilon)} = \mathcal{O}(\epsilon)$. The same factorization may be operated on both sides of the Bapat-Sunder inequality.

Next, we present a formula from Minc [99] for the permanent of the a sum of two matrices

$$\begin{aligned}
&\mathrm{perm}\,(A+B) = \\
&\sum_{r=0}^{n} \sum_{\alpha,\beta \in Q_{r,n}} \mathrm{perm}\,(A[\alpha,\beta])\,\mathrm{perm}\,(B(\alpha,\beta))
\end{aligned} \tag{121}$$

where $Q_{r,n}$ is the set of increasing sequences. More precisely, if we denote $\Gamma_{r,n}$ as the set of all n^r sequences $\omega = (\omega_1, \ldots, \omega_r)$ of integers, $1 \le \omega_i \le n, i = 1, \ldots, n$, we define

$$Q_{r,n} = \{(\omega_1, \ldots, \omega_r) \in \Gamma_{r,n} \mid 1 \le \omega_1 < \cdots < \omega_r \le n\}\,. \tag{122}$$

Moreover, $A[\alpha,\beta]$ denotes the $r \times r$ matrix constructed by choosing the rows and columns of A according to the sets α and β. In contrast, $A(\alpha,\beta)$ is the $(n-r)\times(n-r)$ matrix where those rows and columns have been excluded. This way, we can expand to

first order

$$\begin{aligned}
&\mathrm{perm}\,(A+\delta\Delta) = \\
&= \mathrm{perm}\,(A) + \delta \sum_{i,j=1}^{n} \Delta_{i,j}\mathrm{perm}\,(A(i,j)) + \mathcal{O}(\delta^2) \\
&\mathrm{perm}\,((A+\delta\Delta)\odot B) = \\
&= \mathrm{perm}\,(A\odot B) + \delta \sum_{i,j=1}^{n} \Delta_{i,j}B_{i,j}\mathrm{perm}\,((A\odot B)(i,j)) \\
&\quad + \mathcal{O}(\delta^2)
\end{aligned} \tag{123}$$

so that if $\mathrm{perm}\,(A\odot B) > \mathrm{perm}\,(A)$ then, for small enough ϵ, we have that

$$\mathrm{perm}\,(((1-\epsilon)A+\epsilon\Delta)\odot B) > \mathrm{perm}\,((1-\epsilon)A+\epsilon\Delta)\,, \tag{124}$$

meaning that small enough perturbations to the matrix A still lead to a violation of the Bapat-Sunder inequality. A similar reasoning can be used to argue about robustness to perturbations to matrix B. For the particular counterexample of the generalized bunching conjecture considered in the main text, the robustness to perturbations can be seen in Fig. 9.

In order to further visualize how the choice of distinguishability matrix can affect the bunching violation ratio, we consider the following two-parameter family of S matrices of dimension 7

$$S(x,y) = (1-x-y)S^{(\star)} + xS^{(\mathrm{bos.})} + y, S^{(\mathrm{dist.})}. \tag{125}$$

where $x, y \geq 0$ and $x+y \leq 1$. Here, $S_\star$ corresponds to the S matrix of the partially distinguishable input state from Fig. 7, whereas $S_{\mathrm{bos}} = \mathbb{E}$ and $S_{\mathrm{dist}} = \mathbb{1}$ correspond to the fully indistinguishable and fully distinguishable cases, respectively. The bunching violation ratio for these different S matrices is plotted in Fig. 12. As one gets gets closer to the case of distinguishable particles, the bunching decreases significantly as expected. However, when we interpolate between the $S^{(\mathrm{bos.})}$ and $S_\star$, the bunching probability behaves non-monotonically and the bunching violation ratio $P_7(S(x,y))/P_7^{(\mathrm{bos})}$ attains values larger than 1 in a small region around $S_\star$.

4.7 Stability around the bosonic case

We now turn to the question of whether the violation of the generalized bunching conjecture could possibly be in the neighborhood of the fully indistinguishable bosonic case. We demonstrate that first order perturbations to the internal wave functions of the photons around the fully indistinguishable case leave multimode bunching probabilities invariant, which corroborates the claim that the violation can only be "far from" the bosonic case. We start with the internal functions of the photons

$$|\phi_i\rangle = |\phi_0\rangle \tag{126}$$

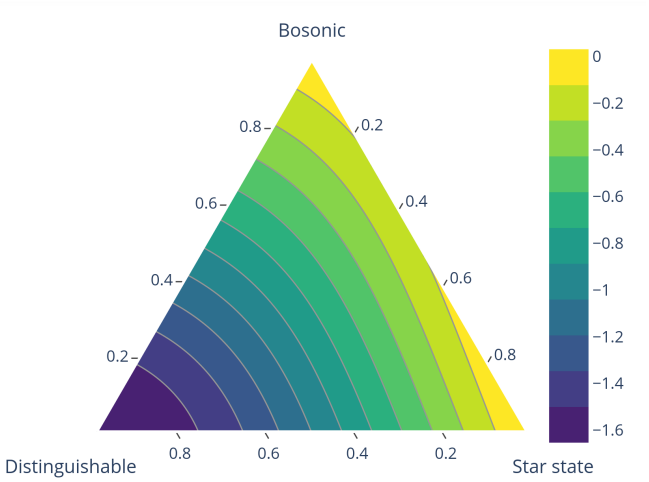


Figure 12: Ternary plot of the $\log_{10}$ of the bunching ratio $P_7(S(x,y))/P_7^{(\mathrm{bos})}$ for the optical scheme from Fig. 7, and the two-parameter family of distinguishability matrices $S(x,y)$ defined in Eq. (125).

for all $i = 1, ..., n$. Let us first consider a perturbation only to the first photon's internal wavefunction

$$\left|\phi_1'\right\rangle = \left|\phi_1 + \delta\phi_1\right\rangle \tag{127}$$

In order for this state to be normalized, we need

$$\langle\phi_1|\delta\phi_1\rangle = ix \tag{128}$$

with $x \in \mathbb{R}$ as our (small) perturbation parameter. Consider a perturbation around the bosonic case: $S' = \mathbb{E} + \delta S$ with

$$\delta S = \begin{pmatrix} 0 & -ix & \cdots & -ix \\ ix & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ ix & 0 & \cdots & 0 \end{pmatrix} \tag{129}$$

Note that the perturbation part of the S matrix is not a Gram matrix. Then

$$\begin{aligned} &\mathrm{perm}\left(A \odot S^T\right) = \mathrm{perm}\left(A + A \odot \delta S^T\right) \\ &= \mathrm{perm}\left(A\right) + \sum_{i,j=1}^{n} \delta S_{j,i} A_{i,j} \mathrm{perm}\left(A(i,j)\right) + \mathcal{O}(x^2) \end{aligned} \tag{130}$$

To alleviate the notations, let us define the matrix F such that

$$F_{i,j} = H_{i,j}\,\mathrm{perm}\left(H(i,j)\right), \tag{131}$$

and note that F can be proven to be Hermitian (since H is an Hermitian matrix). Then, the first-order perturbation to the bunching probability is

$$\sum_{i,j=1}^{n} (\delta S^T \odot F)_{i,j} = -2x \sum_{j=2}^{n} \mathcal{I}(F_{1,j}) = 0 \tag{132}$$

where $\mathcal{I}$ stands for the imaginary part. The second equality holds because the Laplace expansion of the permanent of H gives $\mathrm{perm}\,(H) = \sum_{j=1}^{n} F_{1,j} = F_{1,1} + \sum_{j=2}^{n} F_{1,j}$, which is real (in fact ≥ 0) since H is Hermitian (positive semidefinite), and $F_{1,1}$ is real (≥ 0) since F is Hermitian, hence $\sum_{j=2}^{n} F_{1,j}$ is real too. Thus, we have

$$\mathrm{perm}\left(H \odot (\mathbb{E} + \delta S^T)\right) = \mathrm{perm}\,(H) + \mathcal{O}(x^2). \tag{133}$$

A similar procedure can be applied when every photon's internal wave function is modified by a small quantity such that $\langle \phi_0 | \delta\phi_j \rangle = i x_j = \mathcal{O}(x)$ with $j = 1, \ldots, n$, giving the same result to first order. This means that a small physical perturbation of the internal wavefunctions of the photons around the bosonic case (fully indistinguishable particles) leads to a bunching probability that is unchanged to first order. We leave open the question of whether the bosonic case corresponds to a local maximum of the multimode bunching probability, as suggested by the scheme considered in the main text. This would be true only if second-order perturbations always give a negative contribution to this probability.

5 Efficient validation of Boson Sampling from binned photon-number distributions

Abstract

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.

5.1 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 [100, 101]. 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 [102], 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 [103, 55, 104, 105]. These efforts culminated in multiple claims of quantum computational advantage with Gaussian Boson sampling [106, 43, 44], while standard boson sampling saw experimental implementations with $n = 20$ photons in $m = 60$ modes [42]. Other experimental platforms than photonics were also considered [107].

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 [108, 109, 110, 50, 111, 30, 60]. 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 non-trivial. 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 [112]. Efficient verification schemes that guarantee closeness to the ideal distribution exist, but they require an active control over the experiment via the ability to do Gaussian measurements on the ouput states [113]. Upon reasonable physical assumptions about the nature of noise [114, 115, 116, 117, 118, 119, 120, 121], 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 question we consider in this paper.

5.2 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 [6]. Techniques such as pattern recognition or machine learning [122, 123], Bayesian testing [124], coarse-grained measurements [125, 83], Heavy Output Generation [106], or the analysis of marginal distributions have also been applied [63]. 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 [126]. 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 [60, 61]. 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 [109, 127]. 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.

5.2.1 Our contribution

We propose a validation scheme for Boson Sampling which aims at fulfilling the list of requirements stated above. Our scheme (Sec. 5.5) 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. 13). 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 $1/\mathrm{poly}(n)$), 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

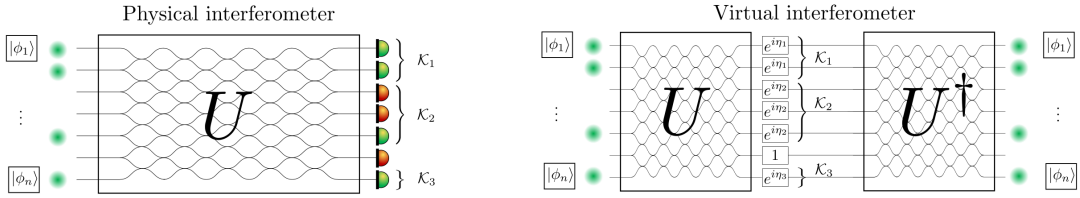


Figure 13: (Left) **Interferometric setup.** Single photons are sent through the first n input modes of a m-mode linear interferometer U. Each photon at input j carries internal degrees of freedom (polarization, arrival time, etc.) described by an (internal) wavefunction $|\phi_j\rangle$. To validate 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. (155) (through Eq. (148)) 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. (150) to (152)).

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.

5.3 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.

5.3.1 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 $\hat{n}_i$ [128, 129] such as

$$c_i = \langle n_i \rangle \tag{134}$$

$$c_{ij} = \langle n_i n_j \rangle - \langle n_i \rangle \langle n_j \rangle \tag{135}$$

Correlations of order 3 and 4 have also been considered [130, 43].

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 [63]. 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 [62]. In fact, it is possible to construct efficient classical mock-up samplers that are consistent with all marginals of order k [63]. 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.

5.3.2 Full bunching

Shchesnovich presents an interesting scheme that aims at fulfilling all the requirements stated in Sec. 5.2 [19, 61]. 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 [61].

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 [131], 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. 5.2, 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 photon counting distribution in one or more subsets. This allows for better distinguishing power, requiring fewer experimental samples for the validation task.

5.3.3 Suppression laws

Certain validation tests rely on a specific choice of optical network [132]. Symmetries in the unitary matrix lead to suppression laws where many outputs are prohibited , as initially noticed in [96] and later applied to boson sampling in [133, 134, 135]. 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 [102]. If one can precisely tune a reconfigurable network [44], 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. 5.6 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.

5.3.4 Validation from coarse-grained measurements

Other proposals exist to validate boson samplers using coarse-grained data. In [125], 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, distinguishable 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 [122].

5.4 Structure of the paper

The core of our work is divided into three sections followed by a discussion section. In Sec. 5.5, 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. 5.6, 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. 6.3.4). We conclude with a discussion section containing open questions and perspectives for future work.

5.5 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

$$k_z = \sum_{j \in \mathcal{K}_z} s_j. \tag{136}$$

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, efficiently estimates the probabilities $P(\boldsymbol{k})$ up to $1/\text{poly}(n)$ additive errors, 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 [136]. 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. 13, and thus it can be approximated via Gurvits randomized algorithm for permanent approximation [102].

5.5.1 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 [2]), 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

$$x(\boldsymbol{\eta}) = \mathop{\mathbb{E}}_{\boldsymbol{k}} \left[\exp\left(i\boldsymbol{\eta} \cdot \boldsymbol{k}\right)\right] \tag{137}$$

$$= \sum_{\boldsymbol{k} \in \Omega^K} P_{\boldsymbol{k}} \exp\left(i\boldsymbol{\eta} \cdot \boldsymbol{k}\right), \tag{138}$$

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

$$\Omega^K = \{(k_1, k_2, \ldots, k_K) \mid k_z \in \Omega, \forall z \in \{1, \ldots, K\}\}, \tag{139}$$

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,

$$\boldsymbol{\nu_l} = \frac{2\pi \boldsymbol{l}}{n+1}, \quad \text{with } l_z \in \Omega, \forall z \in \{1, \ldots, K\} \tag{140}$$

and taking the multidimensional Fourier transform, i.e.

$$P(\boldsymbol{k}) = \frac{1}{(n+1)^K} \sum_{\boldsymbol{l}\in\Omega^K} x(\boldsymbol{\nu_l}) \exp\left(-i\boldsymbol{\nu_l}\cdot\boldsymbol{k}\right). \tag{141}$$

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 5.10). 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

$$|\Psi\rangle_{in} = \prod_{j=1}^{n} \left(\hat{a}^\dagger_{j,\phi_j}\right) |0\rangle \tag{142}$$

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

$$\sum_j \langle\phi_k|\Phi_j\rangle \langle\Phi_j|\phi_k\rangle = 1, \quad \forall k. \tag{143}$$

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

$$\hat{n}_i = \sum_j \hat{a}^\dagger_{i,\Phi_j} \hat{a}_{i,\Phi_j}. \tag{144}$$

Following the formalism of [2, 3], we assume 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

$$\begin{gathered} \hat{a}^\dagger_{j,\phi_j} \to \hat{b}^\dagger_{k,\phi_j} \quad \text{with} \\ \hat{a}^\dagger_{j,\phi_j} = \hat{U}\hat{b}^\dagger_{j,\phi_j}\hat{U}^\dagger = \sum_{k=1}^{m} U_{jk}\, \hat{b}^\dagger_{k,\phi_j}. \end{gathered} \tag{145}$$

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

$$\hat{N}_{\mathcal{K}_z} = \sum_{j\in\mathcal{K}_z} \hat{n}_j = \sum_{j\in\mathcal{K}_z} \sum_k \hat{b}^\dagger_{j,\Phi_k} \hat{b}_{j,\Phi_k}. \tag{146}$$

We demonstrate in Appendix 5.10 that the characteristic function $x(\boldsymbol{\eta})$ can be computed as the quantum expectation value

$$x(\boldsymbol{\eta}) = \langle \Psi_{\text{out}} | \, e^{i\boldsymbol{\eta}\cdot\hat{\boldsymbol{N}}_{\boldsymbol{\mathcal{K}}}} \, | \Psi_{\text{out}} \rangle \tag{147}$$

$$= \langle \Psi_{\text{in}} | \, \hat{U}^\dagger e^{i\boldsymbol{\eta}\cdot\hat{\boldsymbol{N}}_{\boldsymbol{\mathcal{K}}}} \hat{U} \, | \Psi_{\text{in}} \rangle \, , \tag{148}$$

where we have used the notation

$$\boldsymbol{\eta}\cdot\hat{\boldsymbol{N}}_{\boldsymbol{\mathcal{K}}} = \sum_{z=1}^{K} \eta_z \hat{N}_{\mathcal{K}_z}. \tag{149}$$

At this point, it is useful to note that $\hat{V}(\boldsymbol{\eta}) = \hat{U}^\dagger e^{i\boldsymbol{\eta}\cdot\hat{\boldsymbol{N}}_{\boldsymbol{\mathcal{K}}}}\hat{U}$ is a linear interferometer characterized by an $m \times m$ unitary matrix $V(\boldsymbol{\eta})$ (see right panel of Fig. 13). This matrix is constructed as

$$V(\boldsymbol{\eta}) = U^\dagger \Lambda(\boldsymbol{\eta}) U, \tag{150}$$

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

$$\Lambda(\boldsymbol{\eta}) = \prod_{z=1}^{K} D^{(z)}(\eta_z), \tag{151}$$

such that

$$D^{(z)}_{ab}(\eta_z) = \begin{cases} e^{i\eta_z}, & \text{if } a = b \text{ and } a \in \mathcal{K}_z, \\ 1, & \text{if } a = b \text{ and } a \notin \mathcal{K}_z, \\ 0, & \text{if } a \neq b. \end{cases} \tag{152}$$

Using the results of Ref. [102], 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

$$x(\boldsymbol{\eta}) = \text{perm}\left(V_n(\boldsymbol{\eta})\right), \tag{153}$$

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. (142)), this expression is modified in a simple way. By defining the Gram matrix

$$S_{ij} = \langle \phi_i | \phi_j \rangle \tag{154}$$

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

$$x(\boldsymbol{\eta}) = \text{perm}\left(S \odot V_n(\boldsymbol{\eta})\right), \tag{155}$$

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 5.10.

5.5.2 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 [50]. 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. 14).

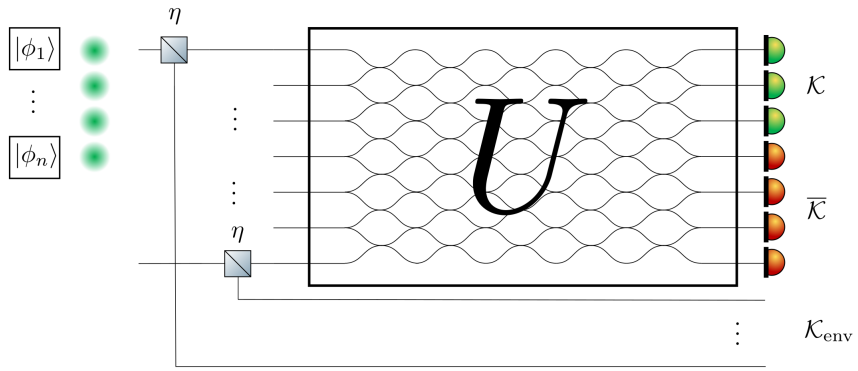


Figure 14: **Lossy interferometer model.** We use a simple model of uniform loss: each photon has a goes through a beam-splitter of transmissivity η before entering the interferometer. If the photon is reflected, it is sent into an environment mode, which simulates a lost photon. The photon number distribution in binned modes of a lossy boson sampler can be computed by considering this larger lossless interferometer.

Taking this into consideration, it is easy to adapt the formalism from Sec. 5.5.1 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. (155) 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. (148) and (155))), 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

$$q(k_d) = \binom{K}{k_d} p_d^{k_d} (1 - p_d)^{K - k_d} \tag{156}$$

The overall probability $P_d(k)$ of observing k photons is thus the convolution of the original probability distribution $P(k)$ with (156)

$$P_d(k) = (P * q)(k) = \sum_{k_d=0}^{k} q(k_d) P(k - k_d) \tag{157}$$

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

5.5.3 Complexity analysis

As shown in Eq. (141), 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

$$T = O(K(n+1)^K \log(n+1) C_x), \tag{158}$$

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 [15]. 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.

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 $\tilde{P}(\boldsymbol{k})$

$$\tilde{P}(\boldsymbol{k}) = \frac{1}{(n+1)^K} \sum_{\boldsymbol{l}\in\Omega^K} \tilde{x}(\nu_{\boldsymbol{l}}) \exp\left(-i\nu_{\boldsymbol{l}}\cdot\boldsymbol{k}\right), \tag{159}$$

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

$$\tilde{x}(\nu_{\boldsymbol{l}}) = x(\nu_{\boldsymbol{l}}) + \epsilon_{\boldsymbol{l}}, \tag{160}$$

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

$$\begin{aligned}\sum_{\boldsymbol{k}} |\tilde{P}(\boldsymbol{k}) - P(\boldsymbol{k})|^2 &= \sum_{\boldsymbol{k}} \left| \sum_{\boldsymbol{l}} \frac{\epsilon_{\boldsymbol{l}}}{(n+1)^K} e^{-i\nu_{\boldsymbol{l}}\cdot\boldsymbol{k}} \right|^2 \\ &= \frac{\sum_{\boldsymbol{l}} |\epsilon_{\boldsymbol{l}}|^2}{(n+1)^K},\end{aligned} \tag{161}$$

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

$$\sqrt{\sum_{\boldsymbol{k}} |\tilde{P}(\boldsymbol{k}) - P(\boldsymbol{k})|^2} \leq \epsilon. \tag{162}$$

This implies that the ℓ_1-norm is bounded by

$$\sum_{\boldsymbol{k}} |\tilde{P}(\boldsymbol{k}) - P(\boldsymbol{k})| \leq (n+1)^{K/2}\epsilon. \tag{163}$$

Therefore, in order to obtain an ℓ_1-norm bounded by some constant value β, we need to compute the approximate values $\tilde{x}(\nu_{\boldsymbol{l}})$ up to an error $\epsilon_{\boldsymbol{l}} \leq \epsilon \leq \beta(n+1)^{-K/2}$. 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 $O(n^2/\epsilon^2)$ [102]. Due to a theorem by Schur (see section of Hadamard product from Ref. [137]), we have that

$$||S \odot V_n(\boldsymbol{\eta})|| \leq (\max_i S_{ii})\ ||V_n(\boldsymbol{\eta})|| \leq 1. \tag{164}$$

Finally, the complexity of obtaining the full approximate probability distribution $\tilde{P}(\boldsymbol{k})$ can be bounded using Eq. (158) by

$$T = O(n^{2K+2}\log(n)\beta^{-2}), \tag{165}$$

where we considered K to be a constant independent of n. □

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 [4]. While this fact is practically irrelevant to efficiently estimate usual boson sampling

[4]We ignore the complexity of computing the matrix $V(\boldsymbol{\eta})$, which is polynomial in m, the number of

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.

5.5.4 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 [63]. 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 $z \in \{1, \ldots, K\}$. It is known that marginal distributions of ideal boson sampler (with fully indistinguishable photons) can be computed *exactly* in polynomial time [102, 138]. 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 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. (154)), which we denote as r_S.

To show this result, we follow very similar lines to Ref. [138] 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

$$x(\boldsymbol{\eta}) = \langle \Psi_{in} | \hat{U}^\dagger e^{i \sum_{j=1}^{r} \eta_j \hat{n}_j} \hat{U} | \Psi_{in} \rangle \tag{166}$$

$$= \mathrm{perm}\left(V_n(\boldsymbol{\eta}) \odot S\right). \tag{167}$$

In this case, we can write

$$V_n(\boldsymbol{\eta} \odot S) = \mathbb{1}_n + W(\boldsymbol{\eta}) \odot S, \tag{168}$$

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

$$W(\boldsymbol{\eta}) = U^\dagger(\Lambda(\boldsymbol{\eta}) - \mathbb{1}_n)U. \tag{169}$$

It is possible to see that $W(\boldsymbol{\eta})$ is a matrix of rank r and thus $\mathrm{rank}(W(\boldsymbol{\eta}) \odot S) \leq K r_S$. 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

$$C_x = O(n^{2Kr_S+1}). \tag{170}$$

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

modes, itself assumed to be polynomial in the number of photons n in most use cases.

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. [109, 139], we may obtain approximations where the error scales as $\log(1/x)$, where x represents a distinguishability parameter.

5.6 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 [96, 132]. 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. 5.5 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.

5.6.1 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. [140, 141], 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 5.11.2 that the formalism of Sec. 5.5 allows us to obtain this distribution analytically. The probability of observing

k photons, if the input photons are fully indistinguishable, is given by

$$P_k^B = \sum_{a=k}^{n} (-1)^{k+a} \binom{a}{k} \binom{n}{a} \frac{a!}{n^a}. \tag{171}$$

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 [96]. 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

$$P_k^D = \binom{n}{k} \frac{1}{n^k} \left(1 - \frac{1}{n}\right)^{n-k}. \tag{172}$$

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.

5.6.2 Photon number distribution in larger subsets

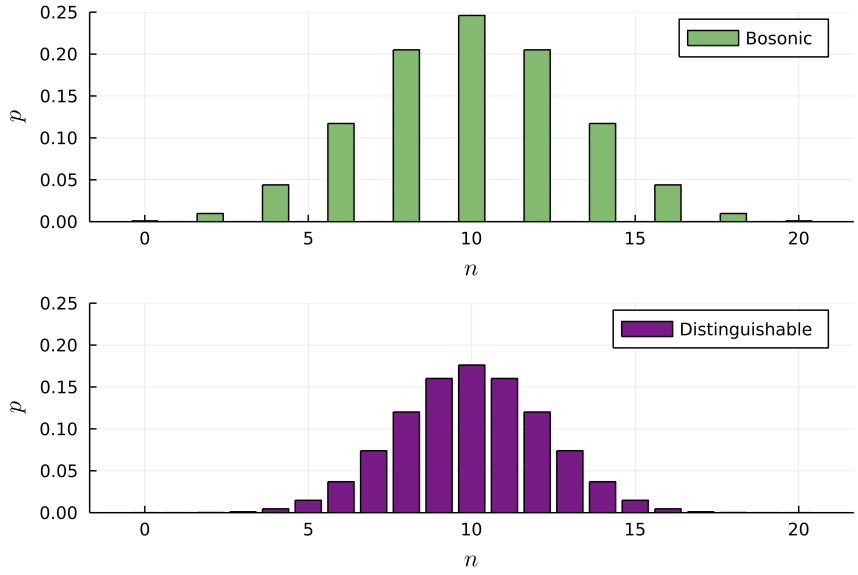


Figure 15: Comparison of the probability of seeing k photons in the odd output modes of the Fourier interferometer with 1 particle per input mode. For bosons we see a suppression of events with odd k, whereas the events with even k follow a binomial distribution. For distinguishable particles the probabilities follow a simple binomial distribution.

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 [34]. 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. 5.5, 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. 15 (see Appendix 5.11.3 for a detailed derivation). For ideal bosons, we obtain

$$P_k^B = \begin{cases} 0 & \text{if } k \text{ is odd,} \\ \frac{1}{2^{n/2}}\binom{n/2}{k} & \text{if } k \text{ is even.} \end{cases} \tag{173}$$

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

$$P_k^D = \frac{1}{2^n}\binom{n}{j}. \tag{174}$$

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. (173) may be used to bound the degree of genuine multiphoton indistinguishability of the input state [142].

5.7 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.

5.8 Noise models

Although the formalism of Sec. 5.5 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 $0 \leq x \leq 1$. The off-diagonal elements of the distinguishability matrix S from Eq. (154) 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. [2, 3, 61, 109], 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. 21). Each photon has a probability $0 \leq \eta \leq 1$ 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.

5.8.1 Comparison to asymptotic formulas

The results in Sec. 5.6 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. [143, 34], 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

$$p^D(\boldsymbol{k}) = \frac{n!}{\prod_{z=1}^K k_z!} \prod_{z=1}^K q_z^{k_z} \tag{175}$$

$$p^B(\boldsymbol{k}) = p^D(\boldsymbol{k}) \frac{\prod_{z=1}^K (\prod_{l=0}^{k_z-1} [1 + l/K_z])}{\prod_{l=0}^{n-1} [1 + l/m]} \tag{176}$$

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:

$$P^{\sigma}(\vec{k}|K) = \frac{\exp\{-n\sum_{z=1}^{K}\frac{(x_z-q_z)^2}{2(1+\sigma\rho)q_z}\}}{(2\pi(1+\sigma\rho)n)^{(K-1)/2}\prod_{z=1}^{K}\sqrt{q_z}} \times\left(1+\mathcal{O}\left(\frac{\rho\delta_{\sigma,+}}{n}\right)\right), \tag{177}$$

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. [34].

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. 5.5. For simplicity, we consider a bipartition of the output modes into two sets of equal size and the partially distinguishability model introduced in Sec. 5.8, 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. 16. 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. (176) and (175) 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.

5.8.2 Distance between distributions

A standard quantity used to quantify the distance between probability distributions is the total variation distance (TVD), defined as

$$\mathrm{tvd}(p,q) = \sum_{j}|p_j - q_j|. \tag{178}$$

This is an especially pertinent metric regarding the problem of distinguishing two distributions [144, 145] via sampling, since the number of samples n_s needed to distinguish a distribution p from another q scales as

$$n_s = \mathcal{O}(\mathrm{tvd}(p,q)^{-2}). \tag{179}$$

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

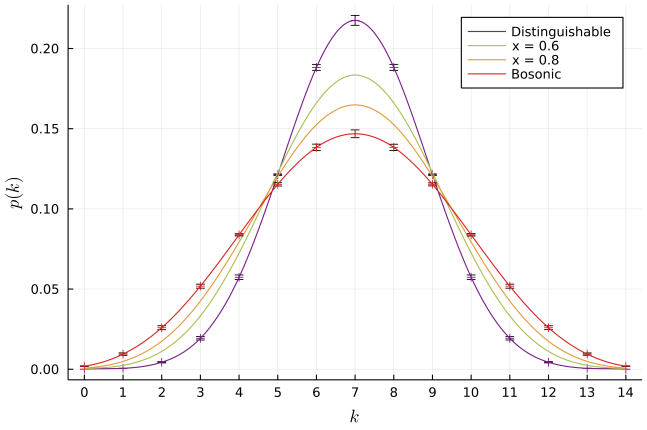


Figure 16: **Effect of partial distinguishability on a subset distribution.** We consider the photon number distribution in a subset consisting of the first half of the $m = 14$ output modes with $n = 14$. The probabilities are averaged on 1000 Haar-random unitaries. Error bars show one standard deviation. We see that they are well approximated by a Gaussian distribution, with the main difference between each case being the width of the curve.

For the rest of this paper, and unless specified otherwise, we will always select the bins to form an equipartition, as defined in Appendix 5.12.1.

Bosons vs. distinguishable particles First, let us compare the two extreme cases of indistinguishable vs distinguishable particles. In Fig. 17, 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 [102]. 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. 17 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. (179).

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 the experiment easy to simulate classically [109, 127]. 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

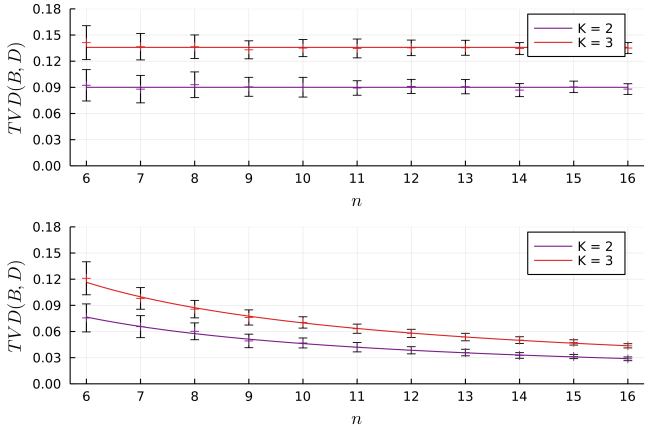


Figure 17: **Evolution of the TVD with the system size.** Haar-averaged TVD between the binned output distributions of bosonic and distinguishable inputs with varying system size. We consider two equipartions of size $K = 2$ and $K = 3$. In the top subplot, we consider a sparse but constant density $m = 5n$, 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 5.12). We also see a significant increase of the TVD when we choose a larger partition size.

validation test to partial distinguishability we again make use of the the one-parameter model interpolating between distinguishable and indistinguishable photons (see Sec. 5.8). In Fig. 18 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. 17: for constant densities the TVD remains constant whereas in the "collision free" regime it suggests a polynomial decay. A specific example for $x = 0.9$, can be found in Fig. 23 of Appendix 5.12. 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

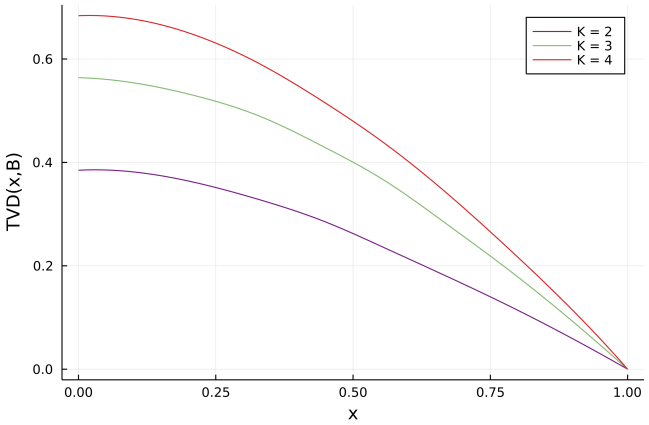


Figure 18: **Effect of partial distinguishability on TVD.** The TVD between the binned distributions corresponding to ideal photons ($x = 1$) and partially distinguishable ones is displayed as a function of the distinguishability parameter x. In this figure we took $m = n = 10$ and equipartitions of size $K = 2, 3, 4$.

described by the following behavior

$$\mathrm{tvd}(B, x) \simeq c(K, x)\rho. \tag{180}$$

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 $\mathrm{tvd}(B, x) \simeq c(K, x)\rho^r$, 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 5.12. 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.

5.8.3 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. [124, 146, 147]. 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

$$\chi = \prod_{i=1}^{n_s} \frac{p(\boldsymbol{s}^{(i)}|H_0)}{p(\boldsymbol{s}^{(i)}|H_a)}, \tag{181}$$

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

$$p_{null} = \chi/(\chi + 1). \tag{182}$$

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 [124, 146, 147], 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.

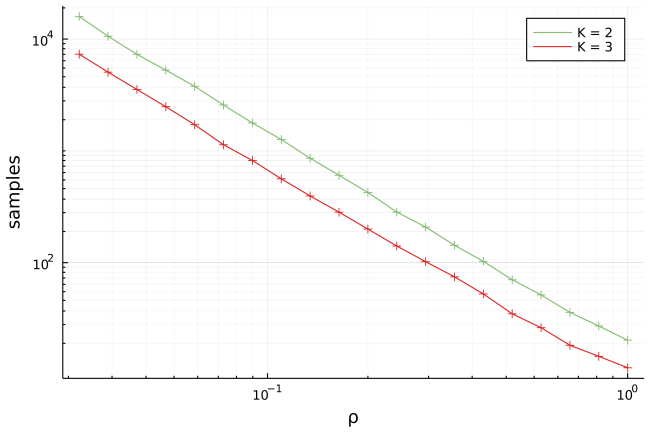


Figure 19: **Number of samples** required to distinguish the indistinguishable and distinguishable inputs. This number of samples is determined by taking the average number of samples required to obtain a certainty of $p_{null} = 95\%$ using the Bayesian approach laid out in the main text, see (181). 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.

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. 5.5. 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. 5.8. In Fig. 20, 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 increases 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. 5.8.2. 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

$$n_s \approx \frac{d(K)}{\rho^{5/2}}, \tag{183}$$

which we extract from the data of Fig. 19. 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 5.12). Such extrapolations are useful to predict the necessary sampling rates for validate experiments when scaling up the system size.

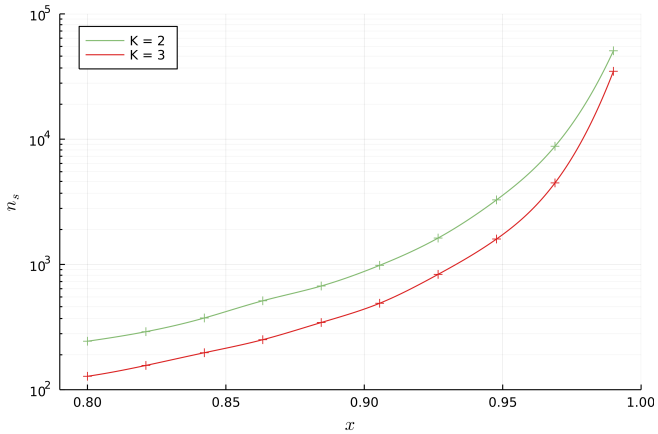


Figure 20: **Number of samples** required to reject the null hypothesis of having indistinguishable bosons when the input is actually the one parameter interpolation with indistinguishability $0.8 \leq x \leq 0.99$. This number of samples is determined by taking the average number of samples required to obtain a certainty less than $p_{null} = 5\%$ using the Bayesian approach laid out in the main text, see (181). This specific figure was generated using parameters $n = 10$, $m = 10$ and each point is an average over 1000 Haar-random unitaries.

5.8.4 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 optical circuit decreases exponentially with

the circuit depth [50], 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 [120].

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 speed-up 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. 21, 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. 21 also reveals that, as expected, 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.

5.9 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 [102].

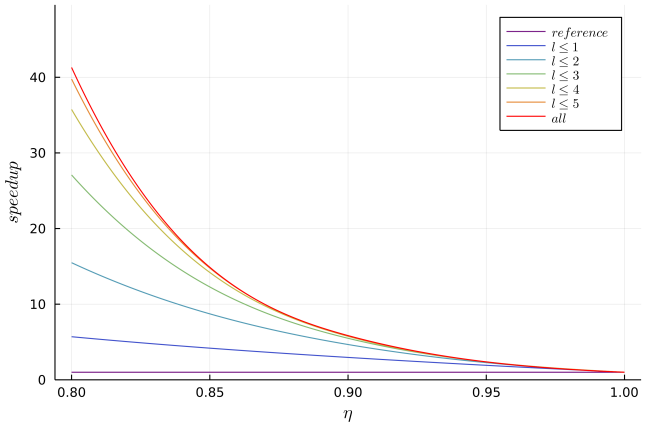


Figure 21: **Validation speed up using lost photons.** In this figure, we take a single subset of half the $m = 10$ output modes with $n = 10$ photons. We validate that the observations are indistinguishable bosons versus the one-parameter model with $x = 0.9$. We consider a model of uniform loss, with 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.

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. (176) and Refs. [34, 143]) may be used to spoof the test. We leave this question for future consideration.

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. [19],

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. [148, 149], 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. [148, 149] may be approximated via a similar formalism to that presented in Sec. 5.5.

During the completion of this work, we became aware of the recent works from Refs. [150, 151, 152]. Ref. [150] develops an efficient classical algorithm to approximate molecular vibronic spectra, using a Formalism similar to Sec. 5.5 based on approximating Fourier components of the target probability distribution using Gurvits algorithms. In turn, Refs. [151, 152] 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. [151, 152] 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 numerical analysis exposed in this paper to simple noise models such as uniform partial distinguishability and loss. In Ref. [153], 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 [98] regarding BosonSampling.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.

Appendix

5.10 Transition amplitudes in the partially distinguishable case

5.10.1 Expression of the characteristic function as an amplitude

Let us first show that the characteristic function introduced in Eq. 137

$$x(\boldsymbol{\eta}) = \mathop{\mathbb{E}}_{\boldsymbol{k}} \left[\exp\left(i\boldsymbol{\eta}\cdot\boldsymbol{k}\right)\right] \tag{184}$$

$$= \sum_{\boldsymbol{k}\in\Omega^K} P(\boldsymbol{k}) \exp\left(i\boldsymbol{\eta}\cdot\boldsymbol{k}\right) \tag{185}$$

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

$$x(\boldsymbol{\eta}) = \left\langle \Psi_{out} \middle| e^{i\sum_{z=1}^{K}\eta_z \hat{N}_{\mathcal{K}_z}} \middle| \Psi_{out} \right\rangle . \tag{186}$$

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:

$$|\boldsymbol{d}(\boldsymbol{s}), \Phi_{\boldsymbol{a}}\rangle = \frac{1}{\sqrt{\mu(\boldsymbol{s})}} \prod_{j=1}^{n} \hat{b}^{\dagger}_{d_j(\boldsymbol{s}),\Phi_{a_j}} |0\rangle . \tag{187}$$

Here, $\boldsymbol{d}(\boldsymbol{s})$ is the mode assignment list, a vector of dimension n constructed as $\boldsymbol{d}(\boldsymbol{s}) = \oplus_{j=1}^{n} \oplus_{k=1}^{s_j} (j)$ [1]. 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

$$|\Psi_{out}\rangle = \sum_{\boldsymbol{d}(\boldsymbol{s}),\boldsymbol{a}} \alpha(\boldsymbol{d}(\boldsymbol{s}), \Phi_{\boldsymbol{a}}) \, |\boldsymbol{d}(\boldsymbol{s}), \Phi_{\boldsymbol{a}}\rangle \tag{188}$$

where

$$\alpha(\boldsymbol{d}(\boldsymbol{s}), \Phi_{\boldsymbol{a}}) = \left\langle \prod_{j=1}^{n} \hat{b}_{d_j(\boldsymbol{s}),\Phi_{a_j}} \middle| \Psi_{out} \right\rangle \tag{189}$$

We can now expand part of the right side of Eq. (186) as

$$e^{i\sum_{z=1}^{K}\eta_z\hat{N}_{\mathcal{K}_z}}\left|\Psi_{out}\right\rangle \tag{190}$$

$$= e^{i\sum_{z=1}^{K}\eta_z\sum_{j\in\mathcal{K}_z}\hat{n}_j}\left|\Psi_{out}\right\rangle \tag{191}$$

$$= \sum_{\boldsymbol{d}(\boldsymbol{s}),\boldsymbol{a}} e^{i\sum_{z=1}^{K}\eta_z\sum_{j\in\mathcal{K}_z}\hat{n}_j}\alpha(\boldsymbol{d}(\boldsymbol{s}),\Phi_{\boldsymbol{a}})\left|\boldsymbol{d}(\boldsymbol{s}),\Phi_{\boldsymbol{a}}\right\rangle \tag{192}$$

$$= \sum_{k_1=0}^{n}\cdots\sum_{k_K=0}^{n} e^{i\sum_{z=1}^{K}\eta_z k_z}\sum_{\boldsymbol{d}(\boldsymbol{s})|\boldsymbol{k},\boldsymbol{a}}\alpha(\boldsymbol{d}(\boldsymbol{s}),\Phi_{\boldsymbol{a}})\left|\boldsymbol{d}(\boldsymbol{s}),\Phi_{\boldsymbol{a}}\right\rangle \tag{193}$$

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

$$x(\boldsymbol{\eta}) = \sum_{k_1=0}^{n}\cdots\sum_{k_K=0}^{n} e^{i\sum_{z=1}^{K}\eta_z k_z}\sum_{\boldsymbol{d}(\boldsymbol{s})|\boldsymbol{k},\boldsymbol{a}}\left|\alpha(\boldsymbol{d}(\boldsymbol{s}),\Phi_{\boldsymbol{a}})\right|^2 \tag{194}$$

Remark that

$$P(\boldsymbol{k}) = \sum_{\boldsymbol{d}(\boldsymbol{s})|\boldsymbol{k},\boldsymbol{a}}\left|\alpha(\boldsymbol{d}(\boldsymbol{s}),\Phi_{\boldsymbol{a}})\right|^2 \tag{195}$$

is the probability to find the photon count $\boldsymbol{k}$, by construction, which proves that we recover Eq. (185). 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:

$$P(\boldsymbol{k}) = \frac{1}{(n+1)^K}\sum_{l_1=0}^{n}\cdots\sum_{l_K=0}^{n} x\left(\frac{2\pi\boldsymbol{l}}{n+1}\right)e^{-2\pi i\boldsymbol{l}\cdot\boldsymbol{k}/(n+1)} \tag{196}$$

5.10.2 Computation of the amplitudes as permanents

Let us derive the expression for $x(\boldsymbol{\eta})$ from Eq. (155). We recall that

$$x(\boldsymbol{\eta}) = \left\langle\Psi_{\text{out}}\right| e^{i\boldsymbol{\eta}\cdot\hat{\boldsymbol{N}}_{\boldsymbol{\mathcal{K}}}}\left|\Psi_{\text{out}}\right\rangle \tag{197}$$

$$= \left\langle\Psi_{\text{in}}\right|\hat{U}^\dagger e^{i\boldsymbol{\eta}\cdot\hat{\boldsymbol{N}}_{\boldsymbol{\mathcal{K}}}}\hat{U}\left|\Psi_{\text{in}}\right\rangle \tag{198}$$

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

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

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

$$x(\boldsymbol{\eta}) = \frac{1}{\mu(\boldsymbol{r})} \left\langle 0 \left| \prod_{j=1}^{n} \hat{a}_{d_j(\boldsymbol{r}),\phi_{d_j(\boldsymbol{r})}} \prod_{j=1}^{n} \sum_{k=1}^{m} V_{d_j(\boldsymbol{r}),k} \hat{a}^\dagger_{k,\phi_{d_j(\boldsymbol{r})}} \right| 0 \right\rangle \tag{199}$$

$$= \frac{1}{\mu(\boldsymbol{r})} \sum_{k_1,\ldots,k_n=1}^{m} \left(\prod_{j=1}^{n} V_{d_i(\boldsymbol{r}),k_i} \right) \left\langle 0 \left| \prod_{j=1}^{n} \hat{a}_{d_j(\boldsymbol{r}),\phi_{d_j(\boldsymbol{r})}} \hat{a}^\dagger_{k_j,\phi_{d_j(\boldsymbol{r})}} \right| 0 \right\rangle . \tag{200}$$

$$\tag{201}$$

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

$$|\phi_j\rangle = \sum_{\alpha}^{m} c_{j\alpha} |\Phi_\alpha\rangle . \tag{202}$$

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

$$\left\langle 0 \left| \prod_{j=1}^{n} \hat{a}_{d_j(\boldsymbol{r}),\phi_{d_j(\boldsymbol{r})}} \hat{a}^\dagger_{k_j,\phi_{d_j(\boldsymbol{r})}} \right| 0 \right\rangle = \sum_{\alpha_1,\ldots,\alpha_n,\beta_1,\ldots,\beta_n=1}^{m} c^*_{d_1,\alpha_1} \ldots c^*_{d_n,\alpha_n} c_{d_1,\beta_1} \ldots c_{d_n,\beta_n} \tag{203}$$

$$\times \left\langle 0 \left| \prod_{j=1}^{n} \hat{a}_{d_j(\boldsymbol{r}),\Phi_{\alpha_j}} \hat{a}^\dagger_{k_j,\Phi_{\beta_j}} \right| 0 \right\rangle \tag{204}$$

This last quantity is given by

$$\left\langle 0 \left| \prod_{j=1}^{n} \hat{a}_{d_j(\boldsymbol{r}),\Phi_{\alpha_j}} \hat{a}^\dagger_{k_j,\Phi_{\beta_j}} \right| 0 \right\rangle = \sum_{\sigma \in S_n} \prod_{i=1}^{n} \delta_{d_i(\boldsymbol{r}),k_{\sigma_i}} \delta_{\alpha_i,\beta_{\sigma_i}} . \tag{205}$$

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

$$x(\boldsymbol{\eta}) = \frac{1}{\mu(\boldsymbol{r})} \sum_{\alpha_1,\ldots,\alpha_n}^{m} \sum_{\sigma \in S_n} \prod_{j=1}^{n} V_{d_i(\boldsymbol{r}),d_{\sigma_i^{-1}}(\boldsymbol{r})} c^*_{d_i,\alpha_i} c_{d_i,\alpha_{\sigma_i^{-1}}} \tag{206}$$

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

$$\langle \phi_i | \phi_j \rangle = \sum_{\alpha} c^*_{i,\alpha} c_{j,\alpha} \equiv S_{ij}. \tag{207}$$

This leads to the expression

$$x(\boldsymbol{\eta}) = \frac{1}{\mu(\boldsymbol{r})} \sum_{\sigma \in S_n} \prod_{j=1}^{n} V_{d_i(\boldsymbol{r}),d_{\sigma_i^{-1}}(\boldsymbol{r})} S_{d_i(\boldsymbol{r}),d_{\sigma_i^{-1}}(\boldsymbol{r})} \tag{208}$$

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

$$\tilde{M}_{ij} = V_{d_i(\boldsymbol{r}),d_j(\boldsymbol{r})} \tag{209}$$

$$\tilde{S}_{ij} = S_{d_i(\boldsymbol{r}),d_j(\boldsymbol{r})} \tag{210}$$

we recover a compact expression

$$x(\boldsymbol{\eta}) = \frac{1}{\mu(\boldsymbol{r})} \sum_{\sigma \in S_n} \prod_{j=1}^{n} \tilde{S}_{i,\sigma_i} \tilde{M}_{i,\sigma_i} \tag{211}$$

$$= \frac{1}{\mu(\boldsymbol{r})} \mathrm{perm}\left(\tilde{S} \odot \tilde{M}\right) \tag{212}$$

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.

5.11 Binned distributions of Fourier interferometers

5.11.1 Explicit probabilities for a single subset

In order to derive the expressions for the single-mode distributions obtained in Sec. 5.6.1, 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. (150) as

$$V(\eta) = \left(\mathbb{1} + (e^{i\eta} - 1)H\right), \tag{213}$$

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

$$H_{a,b} = \sum_{l \in \mathcal{K}} U^*_{l,a} U_{l,b}, \tag{214}$$

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

$$S \odot V_n(\eta) = S \odot \left(\mathbb{1}_n + (e^{i\eta} - 1)H_n\right) \tag{215}$$

$$= \mathbb{1}_n + (e^{i\eta} - 1)H'_n \tag{216}$$

Using an identity from Minc [154] (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

$$x(\nu_l) = \mathrm{perm}\left(\mathbb{1}_n + (e^{i\nu_l} - 1)H'_n\right) \tag{217}$$

$$= 1 + \sum_{a=1}^{n} c_a (1 - e^{i\nu_l})^a. \tag{218}$$

The coefficients c_a are given by

$$c_a = (-1)^a \sum_{\bar{w}\in Q_{a,n}} \mathrm{perm}\left(H'_n[\bar{w}]\right), \tag{219}$$

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. (218) into Eq. (141) we can obtain, after some manipulations, an explicit expression for the probabilities of observing k photons in the subset $\mathcal{K}$

$$P(k) = (-1)^k \sum_{a=k}^{n} \binom{a}{k} c_a. \tag{220}$$

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. 5.6.1) 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

$$P(n) = \mathrm{perm}\left(H'_n\right), \tag{221}$$

retrieving the result derived in [19]. It is also possible to see from the probability of not observing any photons in subset $\mathcal{K}$ is given by

$$P(0) = \mathrm{perm}\left(\mathbb{1}_n - H'_n\right). \tag{222}$$

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 [61]. 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 $\rho = n/m$.

5.11.2 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

$$H_{i,j} = F^*_{1i} F_{1j} = \frac{1}{m}. \tag{223}$$

For indistinguishable photons we have that $H'_n = H_n$ and the coefficients c_a from Eq. (219) are given by

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

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. (220), we obtain

$$P_k^B = \sum_{a=k}^{n} (-1)^{k+a} \binom{a}{k} \binom{n}{a} \frac{a!}{m^a}. \tag{224}$$

Although to our knowledge this expression has not been derived before, it is known from the general results of Ref. [140] 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

$$P_k^D = \binom{n}{k} \frac{1}{m^k} \left(1 - \frac{1}{m}\right)^{n-k}. \tag{225}$$

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 $m = n^2$.

5.11.3 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

$$x(\boldsymbol{\eta}) = \text{perm}\left(F^\dagger \Lambda(\eta) F\right). \tag{226}$$

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

$$\begin{aligned}
(F^\dagger \Lambda(\eta) F)_{ab} &= \delta_{ab} + \left(e^{i\eta} - 1\right) \sum_{j\in\mathcal{K}} F^\dagger_{aj} F_{jb} \\
&= \delta_{ab} + \left(e^{i\eta} - 1\right) \sum_{j\in\mathcal{K}} \frac{1}{n} \exp\left(\frac{2\pi i}{n}(j-1)(a-b)\right) \\
&= \delta_{ab} + \left(e^{i\eta} - 1\right) \frac{1}{n} \sum_{k=0}^{n/2-1} \exp\left(\frac{4\pi i}{n} k(a-b)\right)
\end{aligned}$$

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

$$\begin{aligned}
&\sum_{k=0}^{n/2-1} \exp\left(\frac{4\pi i}{n} k(a-b)\right) = \\
&= \begin{cases} \frac{n}{2}, & \text{if } |a-b| = 0 \text{ or } \frac{n}{2} \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

given that a, b are integers. Thus

$$(F^\dagger \Lambda(\eta) F)_{ab} = \begin{cases} \frac{e^{i\eta}+1}{2} & \text{if } a = b \\ \frac{e^{i\eta}-1}{2} & \text{if } |a-b| = \frac{n}{2} \\ 0 & \text{otherwise} \end{cases} \tag{227}$$

In other words,

$$V(\eta) = F^\dagger \Lambda(\eta) F = \tag{228}$$

$$= \text{circ}\left(\frac{e^{i\eta}+1}{2}, 0, \ldots, 0, \frac{e^{i\eta}-1}{2}, 0, \ldots, 0\right) \tag{229}$$

where $\text{circ}(\vec{v})$ 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

$$\text{perm}\left(V(\eta)\right) = \sum_{\sigma\in S_n} \prod_i V(\eta)_{i\sigma(i)}. \tag{230}$$

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

$$\sigma = \gamma_1 \gamma_2 \dots \gamma_{n/2} \tag{231}$$

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

$$\begin{aligned}
\operatorname{perm}(V(\phi)) &= \sum_{\gamma_1}\sum_{\gamma_2}\cdots\sum_{\gamma_{n/2}}\prod_{i=1}^{n/2} V(\eta)_{i,\gamma_i(i)} V(\eta)_{\frac{n}{2}+i,\gamma_i(\frac{n}{2}+i)} && (232)\\
&= \prod_{i=1}^{n/2}\sum_{\gamma_i} V(\eta)_{i,\gamma_i(i)} V(\eta)_{\frac{n}{2}+i,\gamma_i(\frac{n}{2}+i)} && (233)\\
&= \prod_{i=1}^{n/2}\left(\left(\frac{e^{i\eta}+1}{2}\right)^2 + \left(\frac{e^{i\eta}-1}{2}\right)^2\right) && (234)\\
&= \left(\left(\frac{e^{i\eta}+1}{2}\right)^2 + \left(\frac{e^{i\eta}-1}{2}\right)^2\right)^{n/2} && (235)\\
&= \frac{1}{2^{n/2}}(e^{2i\eta}+1)^{n/2} && (236)
\end{aligned}$$

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. (141) we obtain

$$P_k^B = \begin{cases} 0 & \text{if } k \text{ is odd} \\ \frac{1}{2^{n/2}}\binom{n/2}{k} & \text{if } k \text{ is even} \end{cases} \tag{237}$$

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. 15. It is possible to see that the TVD between these two probability distributions tends to a constant as n goes to infinity.

5.12 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.

5.12.1 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.

5.12.2 Variability of the TVD

In Fig. 22, we show the behaviour of the variance bars of Fig. 17, showing how the TVD converges to its mean value.

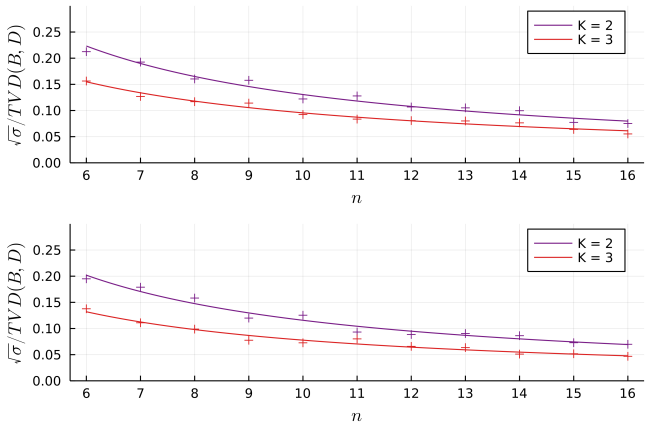


Figure 22: **Variability of the TVD with system size.** We plot the same parameters as in Fig. 17 but instead of looking at the TVD we display its coefficient of variation (CV), defined as the standard deviation divided by the mean. To a good approximation, it follows a power law of form $CV \propto n^{-1}$ in all plots with a slightly varying prefactor of order one.

5.12.3 TVD with partial distinguishability

In Fig. 23, we show here that the TVD decreases in a similar fashion than described in Fig. 17 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.

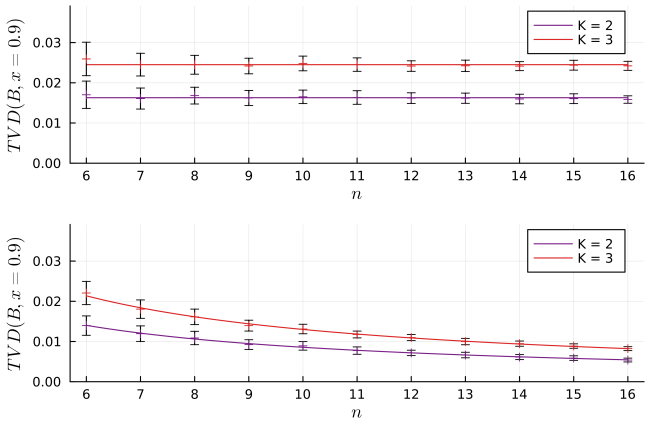


Figure 23: **TVD between bosonic and nearly indistinguishable particles.** This figure shows that the results shown in Fig. 17 hold even when comparing indistinguishable photons to nearly indistinguishable ones. All parameters are identical to those from Fig. 17, except from the fact that we compare 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.

5.12.4 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. 24 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

$$\mathrm{tvd}(B, x) \approx c(K, x)\rho \tag{238}$$

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. 25. 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. 26.

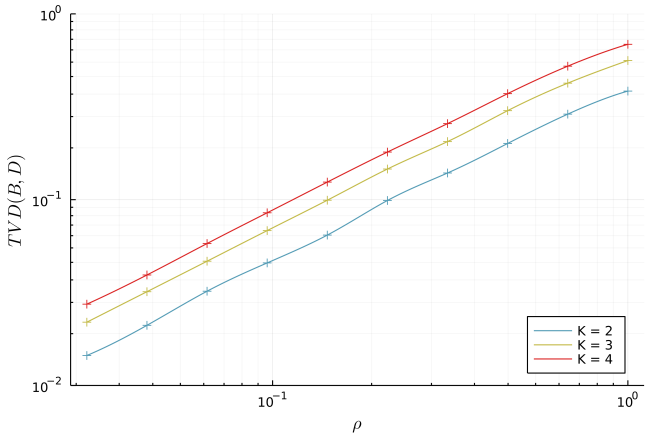


Figure 24: **Effect of boson density.** The Haar-averaged TVD between bosonic and distinguishable particles (asymptotically) depends only on the boson density ρ, see (177). We show how it evolves in the example of $n = 10$ and $m = 10, \ldots, 300$ with a number of subsets $K = 2, 3, 4$. Real-world experiments typically have a high boson density, for instance [42] 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 $m \propto n^2$ [102].

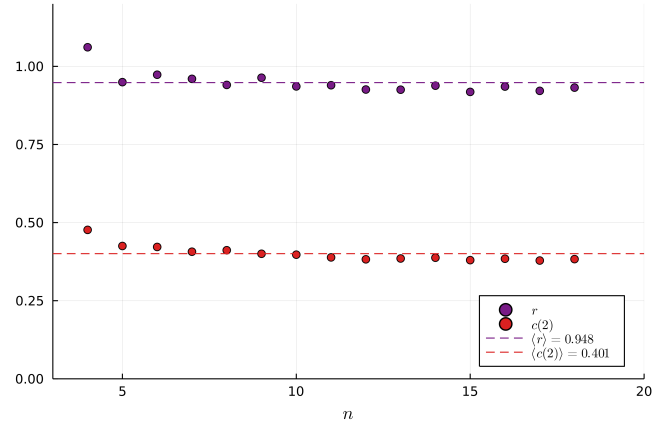


Figure 25: **Validity of the power law.** In Fig 24 we claim that the density is the prime factor modifying the TVD. Here we provide a justification for this by looking at the power-law fit $\mathrm{tvd}(B, D) = c(K, 0)\rho^r$ with K the number of subsets in the partition. We plot the coefficients for 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. 24). For each point, 100 iterations are performed.

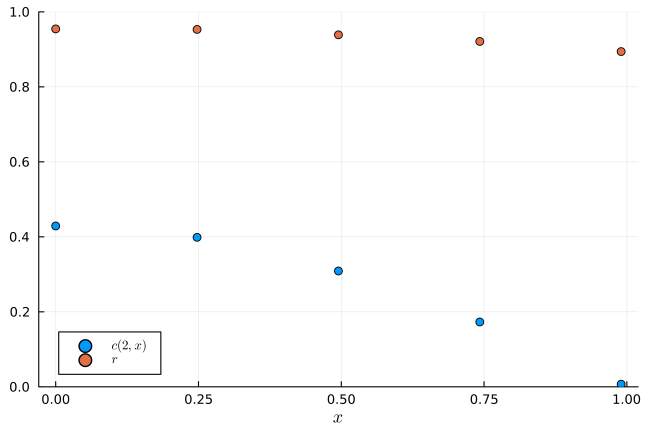


Figure 26: **Validity of the power law for various values of x.**. We show the evolution of the coefficient $c(2, x)$ as well as the power r for $n = 8$ over various values of x. We see small variations in the r but the overall behaviour seems to hold well. For each point, 100 iterations are performed. The last point on the right has value $x = 0.99$.

5.12.5 Dependency of number of samples on density

In this subsection, we give some details about Eq. 183. 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:

$$n_s \approx \frac{d(K, x)}{\rho^r} \tag{239}$$

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).

6 BosonSampling.jl: A Julia package from quantum multiphoton interference

Abstract

We present a free open source package for high performance simulation and numerical investigation of boson samplers and, more generally, multi-photon interferometry. Our package is written in Julia, allowing C-like performance with easy notations and fast, high-level coding. Underlying building blocks can easily be modified without complicated low-level language modifications. We present a great variety of routines for tasks related to boson sampling, such as statistical tools, optimization methods, classical samplers and validation tools.

6.1 Introduction

Quantum computing holds strong hopes for a profound change of computational paradigm and power in the coming years [155, 37, 156]. A large number of experimental and theoretical efforts took place in the last decades, leading to the fast growth in experiments' complexity. While a general purpose, universal quantum computer seems so far out of reach, non-universal models already claim a quantum advantage for specialized tasks on a quantum system that cannot be simulated in a reasonable amount of time with the most advanced classical super-computers [100, 101]. One of these restricted models is that of boson sampling. In their seminal paper [31], Aaronson and Arkhipov describe a quantum version of Galton's board where n bosons - generally photons - are sent through a m-modes linear interferometer implementing a unitary transformation U. The task consists in sampling output mode patterns of particles, such as finding 2 photons in the first mode, none in the second, etc. Depending on the importance of the noise sources from the experimental components, the output distribution $\mathcal{D}$ can be hard or easy to sample from. By hard we understand that it takes an exponential amount of steps $\mathcal{O}(e^n)$ to generate a sample from $\mathcal{D}$ on a classical computer. Indeed, for modest experimental noise, AA prove that this task remains hard, under widely believed conjectures in complexity theory. However, when sufficiently strong noise is present, e.g. due to partial distinguishability or particle loss, then classical algorithms can sample from $\mathcal{D}$ efficiently [108, 29, 110, 50, 111, 30, 60].

Boson sampling sparked a great interest both from theoreticians and experimentalists. Various alternative schemes were constructed, such as Scattershot and Gaussian boson sampling (GBS) [157, 55] to alleviate the technical difficulties in generating single photons. These efforts culminated in multiple claims of quantum advantage in GBS, while standard boson sampling saw experimental implementations with $n = 20$ photons in $m = 60$ modes [42]. Other experimental platforms than photonics were also considered [107].

These factors imply a great need for optimized, high performance numerical simulations of multi-photon interferometry. As the fight for quantum supremacy is played between quantum devices and classical computers, scalable, versatile and fast implementations of the boson sampling algorithms become even more important. In particular,

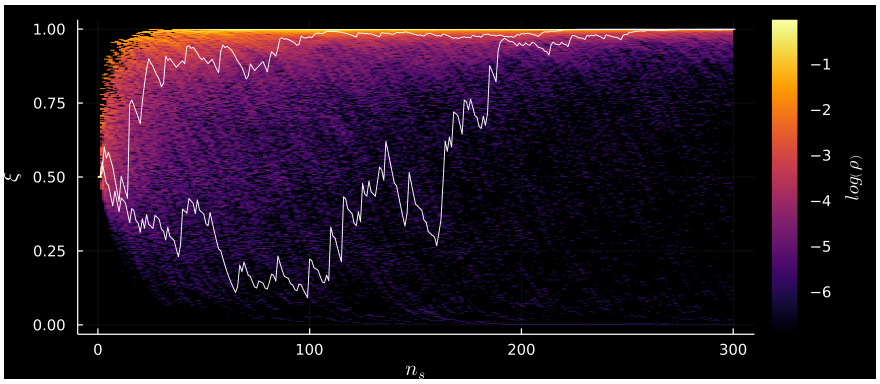


Figure 27: Validating boson sampling using Bayesian methods (see Sec. 6.3.4). Is displayed the confidence ξ that the input is made of 10 indistinguishable photons, averaged over $n_{trials} = 500$, as more samples n_s are provided to the validation protocol. The color scale represents the density of ξ curves. White curves are examples of validation runs.

an adaptive framework for the boson sampling community is of great interest given the speed of growth of this field of research. Indeed, many new boson sampling paradigms were introduced theoretically, and a great variety of experimental techniques were refined, enhancing the demand for evolutive code bases.

In this paper we expose our package, BosonSampling.jl, written in the Julia programming language to tackle classical tasks related to boson sampling. It intends to solve the challenges discussed above by providing an adaptable and extendable framework for multi-photon interferometry designed to study new boson sampling paradigms. The choice of Julia solves the two-language problem: instead of using a high-level wrapper (e.g. Python) with time-critical functions written in a low-level languages (e.g. C, Fortran), Julia is fast out-of-the-box while being easy to write. As a byproduct, all functions are fast, and not only time-critical ones. This is especially relevant for time-crunched researchers who are not professional-programmers but still require a fast execution time for this type of computationally intensive research. The package is evolutive and obeys the Open Source Software standards. It welcomes contributions from the community through its GitHub portal [158]. A complete documentation as well as a pedagogical tutorial are provided [159].

The paper is structured as follows. We first give an overview of the package content. We describe what kind of problems it aims to solve, as well as novel functionalities. We then explain how the specific choice of the Julia programming language is, in the eyes of the authors, optimal for both theoreticians and experimentalists and how it affects the structure of our package and benefits users. In the second section of this paper, we expose how our package is globally structured, including more specific descriptions and examples in the third section. It contains worked-through examples of well-known physical phenomena such as the Hong-Ou-Mandel effect. We also show how one can make use of the package architecture to take into account new sources of noise for instance. Finally, we compare our package to other existing software and discuss possible future features.

6.1.1 Package overview

Our package, BosonSampling.jl together with its paired package Permanents.jl, provides tools to classically simulate multi-photon interferometry experiments. This includes the well-known cases of standard boson sampling, scattershot and Gaussian boson sampling.

Our platform is aimed at high-performance computations for scientists. This is reflected in the choice of the language, Julia, and the modular architecture. Our goal is to make it both easy and fast to implement new models while maintaining all the provided methods and tools adapted to the new features with fast execution time.

The package is evolutive, and welcomes outside contributions from the community. It is intended to be a toolkit for both experimentalists and theoreticians, allowing to easily bridge between new theoretical models and their realistic implementation.

We provide state-of-the-art tools addressing relevant boson sampling problems. Let us now give an overview of these tools made available.

1. **Boson-samplers**, which allow to get samples as if performing an experiment. These samplers are based on the best known algorithms from the literature. Those samplers, unlike that found in other packages, include experimental noise such as photons' partial distinguishability and loss. This noise is critical as it will affect the classical hardness of simulating a given experiment.

2. **Event probability** computation routines, allow to observe specific input/output patterns. Special interest is given to noise, and in particular to a fast implementation in the case of partial distinguishability.

3. **Bunching** tools, aimed at investigating the tendency of bosons to bunch and of fermions to anti-bunch, as exemplified by the Hong-Ou-Mandel effect. We provide functions to investigate the link between bunching and matrix permanent conjectures [131, 19].

4. **Validation** tools for boson-samplers. These allow, given experimental as well as black-box samples, to discuss the degree of multi-photon interference observed in a boson sampler device, and to estimate the level of noise. This gives indications as to whether the device outputs samples that are classically hard to obtain, i.e. if quantum advantage can be claimed. Standard tools from the literature are unified in a recent validation protocol [160] along with the common method of correlators [128, 129], suppression laws [132, 133, 134, 135], and full bunching [19, 61] and also allowing to recover marginal probabilities [139].

5. **Optical circuits** built from optical elements. While in the standard boson sampling problem, networks, described by a unitary matrix U, are chosen from the Haar measure, specific interferometers are also implemented such as the Fourier and Hadamard transformations. In addition, a suite of linear-optical elements is provided, so that networks can be constructed as in an experimental setting.

6. **Optimization** routines to maximize cost functions over unitary matrices. Maximizing certain properties of interferometers, such as photonic bunching, can be convenient for the design of experiments. We include algorithms from Ref. [161] that can optimize cost-functions on the Haar measure.

7. **Time-bin interferometry** types and circuits to simulate the experimental boson sampling proposal of [52], where photons are separated in time bins and sent through a time-dependant beam-splitter and delay lines.

6.1.2 Julia

Julia is a high-level and dynamic programming language originally designed for technical computing, in particular, the simulation of physical systems. It is catching more and more attention from the scientific community and has been adopted in several high class projects such as working with dynamical systems [162], satellites simulations [163], the Celeste project (1.54 petaFLOPS) [164], the Climate Modelling Alliance [165], as well as NASA who saw a 15.000 times speed improvement over their previous Simulink/MATLAB code [166]. The field of quantum information and computation is already well covered, with several packages such as Yao.jl and Quantum Optics.jl rivalling in scope and efficiency with the leaders of the market [167, 168, 169, 170].

The main strength of Julia, in particular for scientists, is its goal to solve the "two-languages problem". It is very common that code is first prototyped in a high-level, easy to write, language such as Python and then made fast to execute in low-level, and hard to write languages such as C++. Julia allows the coding to be convenient and efficient, while being as fast as lower level languages (such as C and Fortran) out of the box. This performance comes from its just-in-time (JIT) compilation and specific type architecture, more akin to functional programming than Oriented Object Programming (OOP) found in most languages (C++, Java, Python,...). This bottom-up approach is especially suited for scientists, whose code evolve organically as results are found, without a set roadmap ("proofs then theorems") as would be preferred for OOP programming [171].

Another strength of Julia's type structure is multiple dispatch: functions are written without strong restrictions to the type of their arguments, and can be reused with custom types with no modifications [172] (provided that the function makes sense for these new types). For instance, a function computing the square root would be restricted to specific types (say, `Float64`) in strongly typed languages. A new type, say `BigFloat`, would require a rewriting of the function, while in Julia the same square root algorithm will work automatically. This design philosophy, is, once again, a blessing for scientists who may not know in advance what the future of their code looks like.

6.2 Package architecture

The two main procedures achievable with this package are the simulation of photonic experiments and validation for multi-photon quantum experiments. We will first describe the simulation architecture and then the validation procedure.

6.2.1 Simulation

The present package takes advantage of Julia's type system through a hierarchy used to classify the building blocks of a quantum interferometry experiment, namely: a photonic `Input` state sent through an `Interferometer` and measured according to an `OutputMeasurement`. This last category also includes (classical) sampling. Those three elements are represented by abstract types, that is, they cannot be instantiated and form the backbone of the package structure. They can be seen as containers of sets of related concrete types, which can be instantiated. Concrete types will represent the specific building blocks of a given interferometric setup. For instance, an interferometer defined from a Haar distributed unitary, `RandHaar`, is a concrete type of the abstract type `Interferometer` as displayed in Fig. 28.

The latter structure is at the core of the code efficiency and tunability. Indeed, the generic programming allowed by Julia combined to the inherent type hierarchy makes the package easy to adapt depending on the user's needs. As a new model can simply be embedded as a concrete type of an existing abstract type, implementing new elements can be done easily without modifying the overall package structure, or any function that acts in the same conceptual fashion on the new type. For instance, a function adding noise could work similarly on any type of `Interferometer`, be it `RandHaar` or `Hadamard`. We will now describe the three main abtract building blocks used in simulating an experiment.

Input When considering photons in Fock states at the input, BosonSampling.jl offers three families of input types depending on the partial distinguishability of the particles. We refer to indistinguishable photons (identical arrival time, polarization) using `Bosonic`. When they are partially distinguishable, they fall under `PartDist`. Finally, we use `Distinguishable` for completely distinguishable input photons. The type `PartDist` allows for a general description of the partial distinguishability. Common models used in the literature are also implemented as subtypes [2].

The input state can also be `Gaussian`. In order to take into account partial distinguishability, we adopt the model introduced in [173]. More precisely, any Gaussian input will be model of fully indistinguishable states while, partial distinguishability can be introduced at the level of the interferometer. The input state is decomposed into interfering and non-interfering modes which evolve independently through the interferometer.

Interferometers The second building block of a quantum multi-photon experiment is the interferometer. A common practice when simulating boson sampling is sample a unitary matrix according to the Haar random measure. Specific interferometers, such as the discrete Fourier and the Hadamard transforms are built-in, together with elementary linear optical circuit elements such as beamsplitters or phase shifts. This allows the user to build networks from scratch. Generic interferometers are supported. Loss can be accounted for in generality.

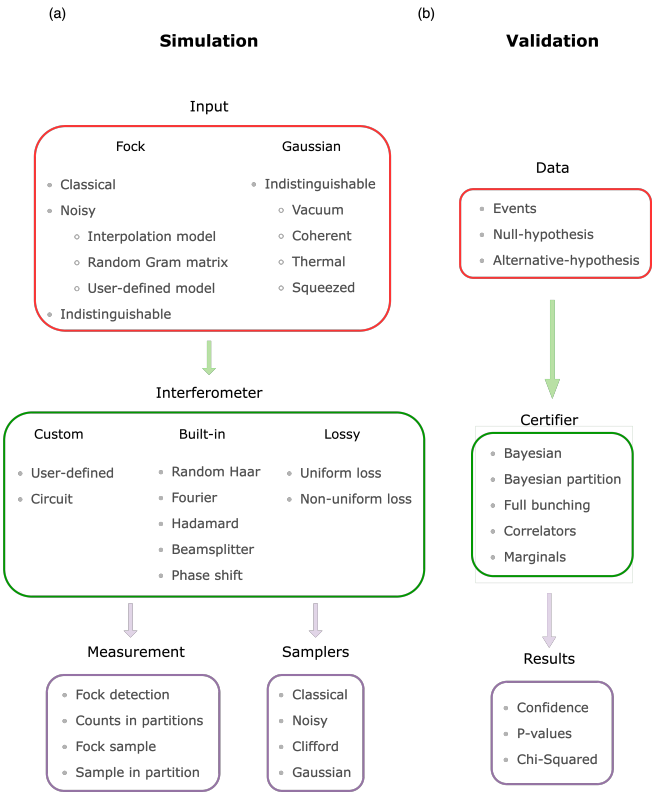


Figure 28: Architecture of BosonSampling.jl (a) Simulation framework: the abstract types representing the experimental setup are displayed as the colored boxes, containing they related concrete types. (b) Validation framework.

Measurement/Sample This object represents the output of the experiment, which we classify between a measurement that can be the probability of a specific output pattern, property or a randomly generated sample. For Fock states, a possibility is to observe a given output mode occupation $\boldsymbol{s} = (s_1, s_2, ...)$, where $0 \leq s_i \leq n$ is the number of photons in mode i. Likewise, we can observe a specific partition photon count $\boldsymbol{k} = (k_1, k_2, ...)$, where $0 \leq k_i \leq n$ is the number photons in a bin K_i gathering multiple output modes.

Different samplers are available depending on the level of distinguishability of the input particles or the use of a lossy interferometer. For exact sampling, we provide Clifford-Clifford's algorithm [35]. For imperfect sampling, we rely on the works of Moylett et al. [174]. Their detailed implementations are discussed in the package documentation.

6.2.2 Validation

Experimental data can be analysed within this package to validate boson sampling devices. The procedure is straightforward for the user: data is held in form of `Events`, comprising all known experimental parameters, such as the detector outputs. Hypotheses are provided, such as that the input is made of indistinguishable particles. An alternative hypothesis could be that they are distinguishable. Another task that can be addressed with this package is to estimate the amount of noise, e.g due to partial distinguishability.

Common validation protocols are implemented, with a comprehensive statistical analysis, providing p-values, etc. A detailed description of these protocols can be found in [6, 160, 60, 147].

6.3 Examples

This section presents some basic experiments that can be simulated by BosonSampling.jl. Further details about the package, as well as a complete, step-by-step tutorial, can be found in the documentation, hosted on the GitHub page of the package. As for any registered Julia package, BosonSampling.jl can be installed by simply executing the following commands in the Julia REPL:

```
using Pkg
Pkg.add("BosonSampling")
```

Code sample 1: *Installing* BosonSampling.jl.

6.3.1 Hong-Ou-Mandel

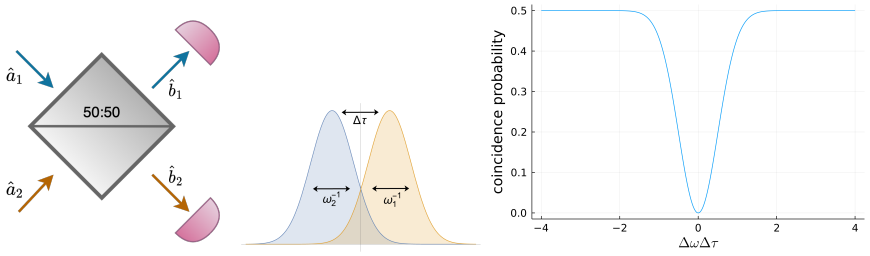
It is well known since the experimental work of Hong, Ou and Mandel [175] that two indistinguishable photons impinging on the two input modes of a balanced beamsplitter will bunch in one of the two output modes, leaving a zero probability to observe one photon at each output mode (the coincidence probability). As an introductory example, we review the effect of partial distinguishability in the HOM effect. In this context, it is a common practice to use the time delay $\Delta\tau$ between the two incoming beams as a

source of partial distinguishability, while the distinguishability parameter itself is $\Delta\omega\Delta\tau$ with $\Delta\omega$ the uncertainty of the frequency distribution. In order to make the parallel with the built-in interpolating model, we substitute its linear parameter x by $e^{-\Delta\omega^2\Delta\tau^2}$. In this way, a distinguishable input is recovered for $\Delta\omega\Delta\tau \to \infty$ and a bosonic input for $\Delta\tau = 0$.

We reproduce in the code sample 2 the HOM experiment as described in Fig. 29(a), with a variable overlap between the two initial wave functions (Fig. 29(b)) and compute the coincidence probability.

```
# Set experimental parameters
delta_omega = 1
# Set the model of partial distinguishability
T = OneParameterInterpolation
# Define the unbalanced beams-plitter
B = BeamSplitter(1/sqrt(2))
# Set each particle in a different mode
r_i = ModeOccupation([1,1])

# Define the output as detecting a coincidence
r_f = ModeOccupation([1,1])
o = FockDetection(r_f)

# Will store the events probability
events = []

for delta_t in -4:0.01:4
    # distinguishability
    dist = exp(-(delta_omega * delta_t)^2)
    i = Input{T}(r_i,dist)

    # Create the event
    ev = Event(i,o,B)
    # Compute its probability to occur
    compute_probability!(ev)

    # Store the event and its probability
    push!(events, ev)
end
```

Code sample 2: *Effect of Partial distinguishability in the Hong-Ou-Mandel effect.*



(a) (b) (c)

Figure 29: The Hong-Ou-Mandel effect: (a) experimental setup (b) time delay between the incoming wave packets (c) HOM dip translating the bosonic interference appears when the overlap between the beams is perfect.

6.3.2 Sampling

As described in Sec. 6.1.1, several classical algorithms for boson sampling are implemented, such as the Clifford and Clifford's algorithm [35] for exact sampling and algorithms taking into account typical sources of noise when considering realistic schemes [174]. To simulate a boson sampling experiment, one defines an input, the interferometer and a measurement. As in the previous example, an `Event` type is used as a container to hold all interferometric parameters.

```
n = 20
m = 400

# Define an input of 20 photons among 400 modes
i = Input{Bosonic}(first_modes(n,m))

# Define the interferometer
interf = RandHaar(m)

# Set the output measurement
o = FockSample()

# Create the event
ev = Event(i, o, interf)

# Simulate
sample!(ev)
# output:
# state = [0,1,0,...]
```

Code sample 3: *Simulate a perfect boson sampling experiment involving 20 single-photon Fock states in 400 modes via Clifford and Clifford's algorithm.*

Given n photons and m modes, the package proposes methods to compute the entire photon-counting distribution exactly or approximately when different sources of noise are included [30] through the `noisy_distribution` function. It returns by default the exact distribution, an approximated distribution in which the computation of the probabilities are truncated and a distribution reconstructed from a Metropolis Independent Sampler (MIS). Below we compute those three distributions for $n = 3$ photons in $m = 5$ modes with a distinguishability parameter $x = 0.74$ and a loss $l = 0.63$ (see Fig. 30).

```
# Set the model of partial distinguishability
T = OneParameterInterpolation

# Define an input of 3 photons placed
# among 5 modes
x = 0.74
i = Input{T}(first_modes(3,5), x)

# Interferometer
l = 0.63
U = RandHaar(i.m)

# Compute the full output statistics
p_exact, p_truncated, p_sampled =
noisy_distribution(input=i,loss=l,interf=U)
```

Code sample 4: *Computing the output mode occupation statistics with partial distinguishability and loss for 3 photons at the input of a (5,5) Haar distributed unitary.*

Both computing `p_truncated` and `p_sampled` have by default a probability of failure and a maximal error equal to 10^{-4}. It is still possible to refine their computations by setting the keywords `error` and `failure_probability` when calling `noisy_distribution`.

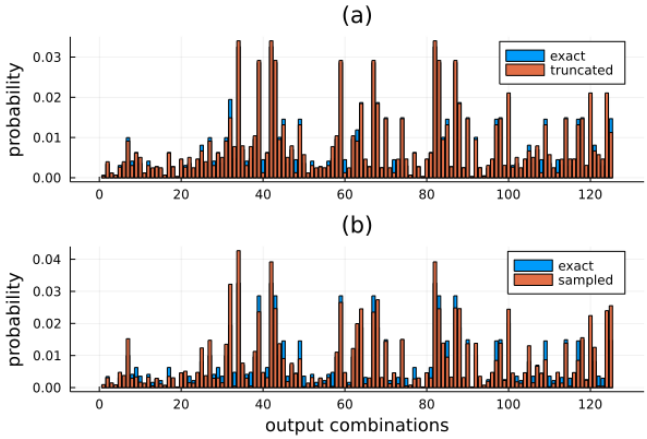


Figure 30: Photon-counting statistics at the output of a 5-mode Haar distributed unitary matrix. (a) Comparison between the ideal distribution and the approximated one. (b) Comparison between the theoretical distribution and a sampled distribution from 10^5 samples.

6.3.3 Photon counting in partitions

One of the novel theoretical tools implemented in this package is the ability to efficiently compute photon counting probabilities in partitions of the output modes of the interferometer. A detailed theoretical and numerical investigation can be found in Ref. [160].

Let us now show how we can obtain the probability of finding $k = 0, ..., n$ photons in a subset consisting of the first half of the modes of a Haar-randomly chosen an interferometer with $n = 25$ photons and $m = 400$ modes. With these parameters, a brute force calculation would be pratically impossible, while the new methods used in this package allow for a fast computation.

```
n = 25
m = 400

# Experiment parameters
input_state = first_modes(n,m)
interf = RandHaar(m)
i = Input{Bosonic}(input_state)

# Subset selection
s = Subset(first_modes(Int(m/2),m))
part = Partition(s)

# Want to find all photon counting probabilities
o = PartitionCountsAll(part)

# Define the event and compute probabilities
ev = Event(i,o,interf)
compute_probability!(ev)
# About 30s execution time on a single core
#
# output:
#
# 0 in subset = [1, 2,..., 200]
# p = 4.650035467008141e-8
# ...
```

Code sample 5: This snippet retrieves the probability of finding $k = 0, ..., n$ photons in the top half of the output modes, with $n = 25$ and $m = 400$. Note that this calculation would be completely out of reach using a brute-force approach.

6.3.4 Validation

We now show how experimental data can be validated. In this example, we compare the null hypothesis $H_0 =$ (bosonic input) to the alternative $H_a =$ (distinguishable input). Multiple validation protocols are implemented in this package, as shown for instance in Fig. 28. We choose to use the new formalism developed by some of the authors of photon counting in binned output modes.

In Fig. 27, we display an example of validation protocol. Data is assessed using a Bayesian procedure on the event probabilities (instead of partition photon counting in the code sample). As explained in detail in [147, 160], each sample updates the confidence level ξ. To be satisfied that the null hypothesis is true, we could require $\xi = 95\%$ for instance. We plot this curve averaged over $n_{trials} = 500$ with $n_{samples} = 300$ events for each run, in a randomly chosen interferometer with $m = 30$ modes and $n = 10$ photons. The input photons are set to be indistinguishable. The heatmap shows the density of the ξ-curves, on a logarithmic scale.

6.3.5 Incorporating new detectors: dark counts

We now show one of the main advantages of this package and Julia: the ability to create new models while keeping most of the structure untouched.

Consider a simple model of threshold detector. With a probability p the detector clicks even though no photon was present, so-called dark count.

We start by defining a new type of output measurement that we label `DarkCountFockSample`, which is a sub-type of `OutputMeasurement`. It takes as arguments the probability p to observe a dark count. A variable holds the sampled `ModeOccupation`, but is first initialised empty at creation of `DarkCountFockSample`. It will be filled later. This is an interesting feature of Julia: the absence of an argument is integrated efficiently by the compiler while a typical, variable-size array would not.

```
mutable struct DarkCountFockSample
    <: OutputMeasurement

    s::Union{ModeOccupation, Nothing}
    p::Real

    function DarkCountFockSample(p::Real)
        if isa_probability(p)
            new(nothing, p)
        else
            error("invalid probability")
        end
    end
end
```

Code sample 6: Implementation of dark counts in the detectors in the BosonSampling framework.

The second step is to define a new method `sample!` that implements a sampling algorithm for this specific type of detector.

As mentioned before, Julia enables multiple dispatch, which means that the new method will not overwrite the already existing ones but will extend the method to the new detectors. That is, when calling `sample!`, Julia will automatically apply the relevant method for a given type (`FockSample`, `DarkCountFockSample`,...). The new algorithm first extracts a sample without dark counts, then adds their influence on the photon counts:

```
function sample!(ev::Event{TIn,TOut})
    where {TIn <: InputType,
           TOut <: DarkCountFockSample}

    # sample without dark counts
    ev_no_dark = Event(ev.input_state,
                       FockSample(),
                       ev.interferometer)

    sample!(ev_no_dark)
    sample_no_dark =
    ev_no_dark.output_measurement.s

    # add count with probablity p
    observe_dark_count(p) =
    Int(do_with_probability(p))
    dark_counts =
    [observe_dark_count(ev.output_measurement.p)
    for i in 1: ev.input_state.m]

    ev.output_measurement.s =
    sample_no_dark + dark_counts
end
```

Code sample 7: Defining a new sampling algorithm.

Now, our new measurement can be used like any other in the previous examples

```
n = 10
m = 10
p_dark = 0.1
input_state = first_modes(n,m)
interf = RandHaar(m)
i = Input{Bosonic}(input_state)
o = DarkCountFockSample(p_dark)
ev = Event(i,o,interf)

sample!(ev)
# output:
# state = [3, 1, 0, 3, 0, 1, 2, 0, 0, 1]
```

Code sample 8: Usage of the newly added detectors.

6.4 Benchmarks

6.4.1 Julia is fast!

When simulating a boson sampling experiment, the most time consuming part comes from the computation of the transition probabilities, since they are related to the evaluation of the permanent. The best known exact algorithm for its computation is due to Ryser [176] and runs in $\mathcal{O}\left(2^{n-1}n\right)$ time for a matrix of size n (using Gray ordering). Having an intensive usage of Ryser's algorithm, we compare the running time of its implementation in Julia with two other interpreted languages: Matlab and Python. Fig. 31 shows how Julia outperforms the two other implementations by two orders of magnitude in computation time.

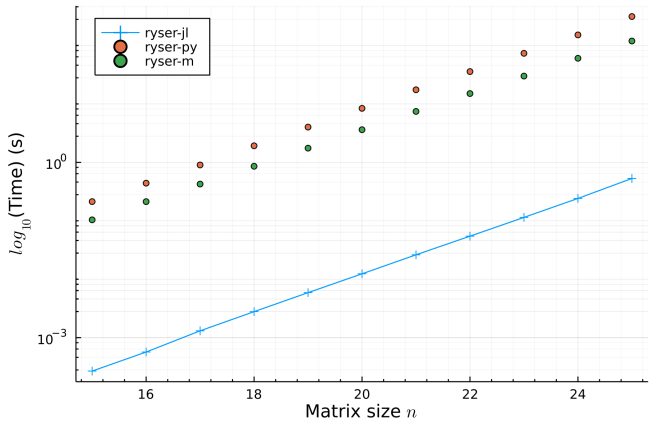


Figure 31: Benchmarks for the single-core running time of Ryser's algorithm in Julia, Matlab and Python on a 3.2GHz Apple M1 processor.

6.4.2 Comparison to existing softwares

The present package is designed for multi-photon interference and boson sampling experiments, taking into account sources of noise and providing a wide variety of validation tools while remaining highly flexible to new paradigms. Other packages for the numerical evaluations of matrix functions and simulation bosonic systems were

published in the last few years. Among the most popular softwares are The Walrus [177] and Perceval [178], both written in Python with a C++ backend.

We illustrate how the Julia programming language is efficient and how, indeed, it can reach C-like performance while keeping the coding as easy, fast and convenient as in Python. We benchmark common functionalities in the three softwares. We first compute permanents of Haar distributed unitary matrices[5] (Fig. 32.a) showing that our implementation of Ryser's algorithm remains faster even on single-core. Our simulations of perfect boson sampling through Clifford and Clifford's algorithm achieves comparable performance to what is possible on Perceval, and is more performant only via multithreading (Fig. 32.b).

[5] Benchmarks inspired from https://the-walrus.readthedocs.io/en/latest/gallery/permanent_tutorial.html and https://github.com/Quandela/Perceval/blob/main/scripts/performance.py

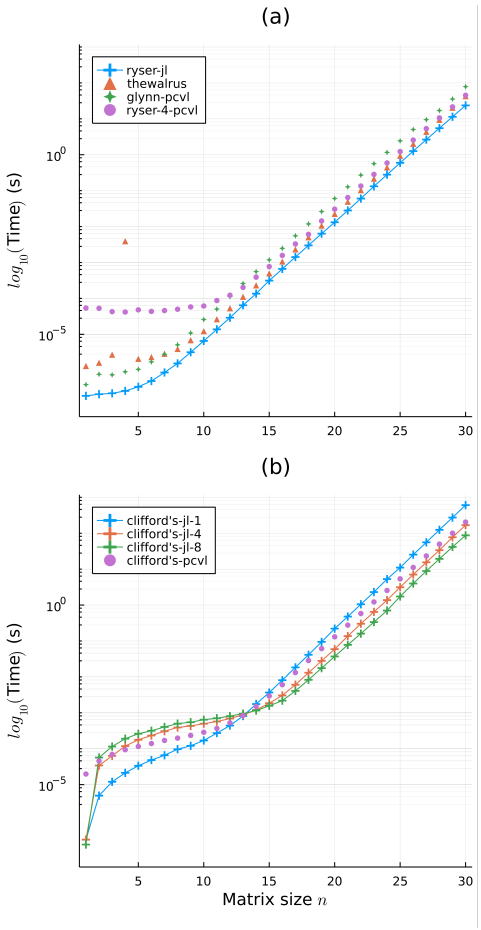


Figure 32: (a) Benchmarks measuring the average time elapsed when computing the permanent of matrices with increasing size n. Are displayed the running times of (blue dots) The Walrus, version 0.19, on a single-core, (orange curve) Ryser implementation of Permanent.jl, version 0.1.1, single-core, (green dots) Glynn implementation of Perceval, version 0.6.2 (Quandelibc version 0.5.3), on a single-core, (purple dots) Ryser implementation of Perceval, 4 threads. (b) Average time elapsed when simulating perfectly indistinguishable bosons via Clifford and Clifford's algorithm in Julia on 1, 4 and 8 threads compared to Perceval's implementation. Both benchmarks are realized with the same configuration as in Fig. 31.

In the final stages of writing this paper, SOQCS [179] was released, a C++ library centered around performance at cost of handiness and ease of modification.

6.5 Development path

This package is a continuously expanding, and many more features are expected to be implemented in the coming years. Current development areas are Gaussian boson sampling related functionalities, in particular simulating GBS with partial distinguishability in a lossy channel. Many of the functions used in this package will be made faster, such as computing permanents and hafnians via GPU computing, and HPC-friendly deployment through multithreading and distributed computing.

Potential users are welcome to contact the authors for more details or specific

development paths, either through the emails listed above or directly through the Github interface.

Outside contributions are welcome, and multiple improvement paths are outlined on the Github discussion page.

6.6 Conclusion

In this paper we introduced a new, free open source Julia package for the high-performance simulation of multi-photon interference. We motivated how the choice of language is particularly relevant for scientists, and allows for a package structure that is easy to modify while keeping execution time as fast as in low-level languages. This package is therefore aimed at experimentalists, who need to implement specifics of their hardware and at theoreticians, who want to develop new models and boson sampling paradigms. We showed how the performance positively compares to similar packages written in other languages.

Part IV
Conclusion

In conclusion, this thesis has explored the complex behavior of quantum interferences with multiple partially distinguishable photons, as well as the development of computational tools (BosonSampling.jl) for the simulation of such experiments. We also proposed an extensive study on the validation of photonic interferometers. Through this work, we have gained a deeper understanding of the intricacies of multiphoton interference and the impact of photon distinguishability on the observed phenomena.

We have presented the first explicit setup that demonstrates the boosting of photon bunching into several modes due to partial distinguishability, disproving the common belief that bunching effects are maximized in the ideal scenario of fully indistinguishable photons. We have also investigated the stability of the bunching violation ratio with respect to experimental imperfections and the potential existence of other counterexamples, raising questions about the generality of this phenomenon and the role of partial distinguishability in practical applications of quantum technology.

We hope to see this setup being realised in a photonic experiment within the near future. While the technology is available and progressing rapidly, our proposal is at the edge of what is feasible today as it combines different technical challenges. It is therefore an interesting open question as to whether one could find even simpler counter-examples with a better experimental prospect - either with a larger bunching probability or a smaller number of photons.

Furthermore, it is interesting to analyze whether other conjectures on permanents or related matrix functions could lead to similar physical interpretations. We recently found out that another conjecture on matrix permanents (from the same authors, Bapat and Sunder [180] and, funnily enough, disproven by the same author, Drury [181]) implies the existence of violations to the multimode bunching conjecture arbitrarily close to the bosonic case. While preliminary evidence indicated that this was a local extrema, as reported in this thesis, we now know that is possible to observe this strange behaviour even for nearly indistinguishable particles. This is true despite the fact that all numerically found counter-examples are far from indistinguishable particles.

From a more mathematical perspective, it is interesting to see if physics can help prove (or disprove) similar conjectures that are still open. Further, even if the validity of the conjecture is settled, it is relevant to make a link with boson bunching. One such conjecture that is also related to bunching is the Permanent-on-Top (POT) conjecture, which was disproved by Shchesnovich [72]. It can be associated with a boosted bunching with a generic input state, with photons having potentially entangled internal wave functions. Speaking of bunching is less justified here since the input photons are not independent, so that weird behaviours of the output probabilities are expected. Yet, Shchesnovich provides a counterexample to the POT conjecture that can be translated into physics, so understanding the reason for this violation would be interesting. Moreover, if a linear optics procedure could be found in order to prepare a state violating

the POT conjecture, then a direct correspondence could be asserted between the POT conjecture and the Bapat-Sunder conjecture.

Overall, a great deal of mathematical features related to matrix permanents are left essentially unexplored, with various matrix functions potentially playing a role in the physics of interfering quantum particles. Of particular interest to mathematicians is Lieb's permanental dominance conjecture, which has been actively discussed for more than half a century [182, 183]. Relating it to physics could lead to a deeper understanding of its meaning, and perhaps even to a solution. In turn, this type of exploration would strengthen our knowledge on the opposite interferometric behaviour of bosons and fermions: in the words of Lieb, "the few inequalities that are known for the permanent are suspiciously similar to certain special cases of classical inequalities for the determinants of such matrices – the only difference being that the direction of the inequality is reverse" [remember that determinants directly relate to fermions]. We even believe that studying the interference of anyons along the same lines could lead to interesting results. Indeed, so far we know that sampling bosons is computationally hard, while fermions and distinguishable particles are easy. A natural extension would be to study Anyon Sampling. It is unknown if any deviation of bosonic statistics would lead to classically efficient sampling. Interesting work in this direction was done by Tichy and Mølmer on hypothetical immanons [184].

Given the computationally demanding aspects of the physics and mathematics involved in this thesis, extensive numerical simulations were undertaken. In order to help advance the work of colleagues and future researchers interested in the same type of questions, we chose to publish this numerical development in full details. Given the size of the codebase, we have introduced a standalone free open-source Julia package for the high-performance simulation of multi-photon interference, providing a versatile and efficient tool for both experimentalists and theoreticians to study and develop new models and boson sampling paradigms. We have demonstrated the performance and advantages of this Julia package compared to similar tools written in other programming languages.

This package has recently been of use to experimentalists of the group of Philip Walter in Vienna, with whom a paper was published [54]. Of particular interest to this experiment, realising time-binned boson sampling, was the validation of the data. A large portion of the work from this thesis pushes forward the known techniques to assert the correct functioning of multiphoton interferometry experiments. We have proposed a validation test based on a coarse-graining of the output photon-number distribution, allowing for an efficient classical approximation of the binned distribution and hence offering a sensitive method to probe photon distinguishability and effectively use experimental data with photon losses.

This validation test raises questions about the classical spoofability of the method and the potential use of binned distributions to quantify the degree of indistinguishability of input photons. Indeed, we have so far limited ourselves to assert that this method can verify the quality of an experiment known to have physical sources of errors. At the opposite of the spectrum, we could ask if the binned distributions contain sufficient

information to discriminate between a valid boson sampler and a black-box adversary. We believe that this is the case, at least up to some level of precision. Indeed, the partition choice is completely arbitrary, and, in fact, multiple binned distributions can be classically computed and compared to the single set of data that we wish to validate. As the number of subsets grows, an exponential number of partitions can be chosen. Our numerical evidence suggests that the information from each partition is different enough to provide additional certification power (more precisely, polynomial decrease of the standard deviation). Therefore, it seems unlikely that an adversary could efficiently output a classical distribution compatible with such a large, *a priori* unknown, choice of partitions.

On another note, it is interesting to ask whether the same technique of grouped counts could be applied to the Gaussian Boson Sampling (GBS) paradigm. While this is yet ongoing work, we can answer this question positively. In particular, we have found interesting asymptotic laws echoing that of Shchesnovich [Eq. (177)] but for GBS. However, the notion of partial distinguishability with Gaussian states is more intricate and deserves further analysis. Only the works of Shchesnovich [185] and Renema [186] provide a starting point for formalizing partial distinguishability in this case, while the account of partial distinguishability is well established in the case of single photons. A noteworthy toy-model is that of Shi and Byrnes [187].

The formalism for binned photon number distributions should also allow us to question the computational hardness of some applications of boson sampling, such as some decision problems. A recent example is the cryptographic one-way function introduced by Nikopoulos [188, 189] and recently implemented experimentally by Wang [190]. We believe that this function could actually be simulated on a classical computer, exactly as the binned distribution used to validate it.

A broader question follows: Can any practical task be sped up using a boson sampler? Unlike for a generic-purpose quantum computer, it is indeed much less clear whether these primitive devices can hold any practical computational advantage. As a simple argument raising some doubts, let us remember that, while event probabilities are linked to matrix permanents (which are computationally hard), a boson sampler does not allow us to approximate these permanents efficiently [191]. In the past, certain tasks were thought to be efficient on a boson sampler, such as calculations of vibronic spectra of molecules [192] but were later shown to be classically simulatable [150]. In this regard, Gaussian boson sampling seems to hold the highest promises (although this might be due to skewed research efforts as GBS is a prime candidate for some private companies in the race for developing a quantum computer). Various tasks have been proposed using GBS, such as finding dense subgraphs [193, 194] or studying molecular docking [195] and drug discovery [59]. A recent proposal to achieve tasks related to graphs was also proposed for single photons [196]. A notable mention on quantum advantage and sampling problems is the work of Novo, Bermejo-Vega and García-Patrón [197].

We conclude this dissertation with a sense of optimism for the future of this field, anticipating not only a more profound understanding of nature's laws but also the

emergence of potent, novel quantum technologies that could greatly benefit society.

References

[1] Tichy, Malte C. 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.

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

[3] Shchesnovich, VS. Partial indistinguishability theory for multiphoton experiments in multiport devices. *Physical Review A*, 91(1):013844, 2015.

[4] Brunner, Eric. *Many-body interference, partial distinguishability and entanglement*. PhD thesis, Masterarbeit, Universitat Freiburg, 2019, 2020.

[5] Tichy, Malte Christopher. Entanglement and interference of identical particles. *Verhandlungen der Deutschen Physikalischen Gesellschaft*, 2013.

[6] Walschaers, Mattia. Signatures of many-particle interference. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 53(4):043001, 2020.

[7] Cerf, Nicolas J and Jabbour, Michael G. Two-boson quantum interference in time. *Proceedings of the National Academy of Sciences*, 117(52):33107–33116, 2020.

[8] Jabbour, Michael G and Cerf, Nicolas J. Multiparticle quantum interference in bogoliubov bosonic transformations. *Physical Review Research*, 3(4):043065, 2021.

[9] Chakhmakhchyan, Levon and Cerf, Nicolas J. Simulating arbitrary gaussian circuits with linear optics. *Physical Review A*, 98(6):062314, 2018.

[10] Hong, Chong-Ki and Ou, Zhe-Yu and Mandel, Leonard. Measurement of subpicosecond time intervals between two photons by interference. *Physical review letters*, 59(18):2044, 1987.

[11] Teukolsky, Saul A and Flannery, Brian P and Press, WH and Vetterling, WT. Numerical recipes in C. *SMR*, 693(1):59–70, 1992.

[12] Ryser, Herbert John. *Combinatorial mathematics*, volume 14. American Mathematical Soc., 1963.

[13] Valiant, Leslie G. The complexity of computing the permanent. *Theoretical computer science*, 8(2):189–201, 1979.

[14] Drummond, Peter D and Opanchuk, Bogdan and Rosales-Zárate, Laura and Reid, Margaret D and Forrester, Peter J. Scaling of boson sampling experiments. *Physical Review A*, 94(4):042339, 2016.

[15] Gurvits, Leonid and Samorodnitsky, Alex. 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.

[16] Scott Aaronson and Travis Hance. Generalizing and Derandomizing Gurvits's Approximation Algorithm for the Permanent, 2012.

[17] Jerrum, Mark and Sinclair, Alistair and Vigoda, Eric. A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries. *Journal of the ACM (JACM)*, 51(4):671–697, 2004.

[18] Ra, Young-Sik and Tichy, Malte C and Lim, Hyang-Tag and Kwon, Osung and Mintert, Florian and Buchleitner, Andreas and Kim, Yoon-Ho. Nonmonotonic quantum-to-classical transition in multiparticle interference. *Proceedings of the National Academy of Sciences*, 110(4):1227–1231, 2013.

[19] Shchesnovich, VS. Universality of generalized bunching and efficient assessment of boson sampling. *Physical review letters*, 116(12):123601, 2016.

[20] Chakhmakhchyan, Levon and Cerf, Nicolas J and Garcia-Patron, Raul. Quantum-inspired algorithm for estimating the permanent of positive semidefinite matrices. *Physical Review A*, 96(2):022329, 2017.

[21] Zhang, Fuzhen. *Matrix theory: basic results and techniques*. Springer Science & Business Media, 2011.

[22] Drury, Steven. Private Communication. 2021-2022.

[23] Bapat, Ravindra B and Sunder, Vaikalathur S. On majorization and Schur products. *Linear algebra and its applications*, 72:107–117, 1985.

[24] Zhang, Fuzhen. Notes on hadamard products of matrices. *Linear and Multilinear Algebra 1989-oct vol. 25 iss. 3*, 25, oct 1989. doi: 10.1080/03081088908817946.

[25] Zhang, Fuzhen. An analytic approach to a permanent conjecture. *Linear Algebra and its Applications*, 438(4):1570–1579, 2013.

[26] Minc, Henryk. The van der Waerden permanent conjecture. In *General Inequalities 3: 3rd International Conference on General Inequalities, Oberwolfach, April 26–May 2, 1981*, pages 23–40. Springer, 1983.

[27] Egorycev, GP. A solution of van der Waerden's permanent problem. In *Dokl. Akad. Nauk SSSR*, volume 258, pages 1041–1044, 1981.

[28] Minc, Henryk. Theory of permanents 1978–1981. *Linear and Multilinear Algebra 1983-jan vol. 12 iss. 4*, 12, jan 1983. doi: 10.1080/03081088308817488. URL libgen.li/file.php?md5=144685e1a4f24be7a35a86b0e08c81b7.

[29] Renema, Jelmer J and Menssen, Adrian and Clements, William R and Triginer, Gil and Kolthammer, William S and Walmsley, Ian A. Efficient classical algorithm for boson sampling with partially distinguishable photons. *Physical review letters*, 120(22):220502, 2018.

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

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

[32] Soules, George William. *Matrix functions and the Laplace expansion theorem.* PhD thesis, University of California, Santa Barbara, 1966.

[33] Fuzhen Zhang. An update on a few permanent conjectures. *Special Matrices*, 4 (1), 2016. doi: doi:10.1515/spma-2016-0030. URL https://doi.org/10.1515/spma-2016-0030.

[34] Shchesnovich, Valery S. Asymptotic Gaussian law for noninteracting indistinguishable particles in random networks. *Scientific reports*, 7(1):1–11, 2017.

[35] Clifford, Peter and Clifford, Raphaël. The Classical Complexity of Boson Sampling, 2017. URL https://arxiv.org/abs/1706.01260.

[36] Clifford, Peter and Clifford, Raphael. Faster classical boson sampling. *arXiv preprint arXiv:2005.04214*, 2020.

[37] Shor, Peter W. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th annual symposium on foundations of computer science*, pages 124–134. Ieee, 1994.

[38] Shor, Peter W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM review*, 41(2):303–332, 1999.

[39] Knill, Emanuel and Laflamme, Raymond and Milburn, Gerald J. A scheme for efficient quantum computation with linear optics. *nature*, 409(6816):46–52, 2001.

[40] Nielsen, Michael A and Chuang, Isaac L. *Quantum Computation and Quantum Information.* Cambridge University Press, 2010.

[41] Arkhipov, Alex and Kuperberg, Greg. The bosonic birthday paradox. *Geometry & Topology Monographs*, 18(1):10–2140, 2012.

[42] Wang, Hui and Qin, Jian and Ding, Xing and Chen, Ming-Cheng and Chen, Si and You, Xiang and He, Yu-Ming and Jiang, Xiao and You, L and Wang, Z and others. 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.

[43] Zhong, Han-Sen and Deng, Yu-Hao and Qin, Jian and Wang, Hui and Chen, Ming-Cheng and Peng, Li-Chao and Luo, Yi-Han and Wu, Dian and Gong, Si-Qiu and Su, Hao and others. Phase-programmable gaussian boson sampling using stimulated squeezed light. *Physical review letters*, 127(18):180502, 2021.

[44] Madsen, Lars S and Laudenbach, Fabian and Askarani, Mohsen Falamarzi and Rortais, Fabien and Vincent, Trevor and Bulmer, Jacob FF and Miatto, Filippo M and Neuhaus, Leonhard and Helt, Lukas G and Collins, Matthew J and others. Quantum computational advantage with a programmable photonic processor. *Nature*, 606(7912):75–81, 2022.

[45] Somaschi, Niccolo and Giesz, Valerian and De Santis, Lorenzo and Loredo, JC and Almeida, Marcelo P and Hornecker, Gaston and Portalupi, S Luca and Grange, Thomas and Anton, Carlos and Demory, Justin and others. Near-optimal single-photon sources in the solid state. *Nature Photonics*, 10(5):340–345, 2016.

[46] William R. Clements and Peter C. Humphreys and Benjamin J. Metcalf and W. Steven Kolthammer and Ian A. Walmsley. Optimal design for universal multiport interferometers. *Optica*, 3(12):1460–1465, Dec 2016. doi: 10.1364/OPTICA.3.001460. URL http://opg.optica.org/optica/abstract.cfm?URI=optica-3-12-1460.

[47] Cerf, Nicolas J and Adami, Christoph and Kwiat, Paul G. Optical simulation of quantum logic. *Physical Review A*, 57(3):R1477, 1998.

[48] Reck, Michael and Zeilinger, Anton and Bernstein, Herbert J and Bertani, Philip. Experimental realization of any discrete unitary operator. *Physical review letters*, 73(1):58, 1994.

[49] Hoch, Francesco and Piacentini, Simone and Giordani, Taira and Tian, Zhen-Nan and Iuliano, Mariagrazia and Esposito, Chiara and Camillini, Anita and Carvacho, Gonzalo and Ceccarelli, Francesco and Spagnolo, Nicolo and others. Reconfigurable continuously-coupled 3D photonic circuit for Boson Sampling experiments. *npj Quantum Information*, 8(1):55, 2022.

[50] Garcia-Patron, Raul and Renema, Jelmer J and Shchesnovich, Valery. Simulating boson sampling in lossy architectures. *Quantum*, 3:169, 2019.

[51] Jan Provaznik and Lukas Lachman and Radim Filip and Petr Marek. Benchmarking photon number resolving detectors. *Opt. Express*, 28(10):14839–14849, May 2020. doi: 10.1364/OE.389619. URL http://opg.optica.org/oe/abstract.cfm?URI=oe-28-10-14839.

[52] Motes, Keith R and Gilchrist, Alexei and Dowling, Jonathan P and Rohde, Peter P. Scalable boson sampling with time-bin encoding using a loop-based architecture. *Physical review letters*, 113(12):120501, 2014.

[53] Lubasch, Michael and Valido, Antonio A and Renema, Jelmer J and Kolthammer, W Steven and Jaksch, Dieter and Kim, Myungshik S and Walmsley, Ian and Garcia-Patron, Raul. Tensor network states in time-bin quantum optics. *Physical Review A*, 97(6):062304, 2018.

[54] Carosini, Lorenzo and Oddi, Virginia and Giorgino, Francesco and Hansen, Lena M and Seron, Benoit and Piacentini, Simone and Guggemos, Tobias and Agresti,

Iris and Loredo, Juan Carlos and Walther, Philip. Programmable multi-photon quantum interference in a single spatial mode. *arXiv preprint arXiv:2305.11157*, 2023.

[55] Craig S. Hamilton and Regina Kruse and Linda Sansoni and Sonja Barkhofen and Christine Silberhorn and Igor Jex. Gaussian Boson Sampling. *Physical Review Letters*, 119(17), 2017.

[56] Deng, Yu-Hao and Gong, Si-Qiu and Gu, Yi-Chao and Zhang, Zhi-Jiong and Liu, Hua-Liang and Su, Hao and Tang, Hao-Yang and Xu, Jia-Min and Jia, Meng-Hao and Chen, Ming-Cheng and others. Solving Graph Problems Using Gaussian Boson Sampling. *arXiv preprint arXiv:2302.00936*, 2023.

[57] Bradler, Kamil and Friedland, Shmuel and Izaac, Josh and Killoran, Nathan and Su, Daiqin. Graph isomorphism and Gaussian boson sampling. *Special Matrices*, 9(1):166–196, 2021.

[58] Banchi, Leonardo and Fingerhuth, Mark and Babej, Tomas and Ing, Christopher and Arrazola, Juan Miguel. Molecular docking with Gaussian boson sampling. *Science advances*, 6(23):eaax1950, 2020.

[59] Yu, Shang and Zhong, Zhi-Peng and Fang, Yuhua and Patel, Raj B and Li, Qing-Peng and Liu, Wei and Li, Zhenghao and Xu, Liang and Sagona-Stophel, Steven and Mer, Ewan and others. A universal programmable Gaussian Boson Sampler for drug discovery. *arXiv preprint arXiv:2210.14877*, 2022.

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

[61] Shchesnovich, Valery. Distinguishing noisy boson sampling from classical simulations. *Quantum*, March 2021.

[62] Shchesnovich, Valery. Boson sampling cannot be faithfully simulated by only the lower-order multi-boson interferences. *arXiv preprint arXiv:2204.07792*, 2022.

[63] Villalonga, Benjamin and Niu, Murphy Yuezhen and Li, Li and Neven, Hartmut and Platt, John C and Smelyanskiy, Vadim N and Boixo, Sergio. Efficient approximation of experimental Gaussian boson sampling. *arXiv:2109.11525*, 2021. doi: 10.48550/ARXIV.2109.11525. URL https://arxiv.org/abs/2109.11525.

[64] Wootters, William K and Zurek, Wojciech H. Complementarity in the double-slit experiment: Quantum nonseparability and a quantitative statement of Bohr's principle. *Physical Review D*, 19(2):473, 1979.

[65] Greenberger, Daniel M and Yasin, Allaine. Simultaneous wave and particle knowledge in a neutron interferometer. *Physics Letters A*, 128(8):391–394, 1988.

[66] Mandel, Leonard. Coherence and indistinguishability. *Optics letters*, 16(23): 1882–1883, 1991.

[67] Feynman, Richard P and Leighton, Robert B and Sands, Matthew. The Feynman lectures on physics. *Reading: Addison-Wesley*, 1963.

[68] Gerry, Christopher and Knight, Peter and Knight, Peter L. *Introductory quantum optics*. Cambridge university press, 2005.

[69] Spagnolo, Nicolo and Vitelli, Chiara and Sansoni, Linda and Maiorino, Enrico and Mataloni, Paolo and Sciarrino, Fabio and Brod, Daniel J. and Galvao, Ernesto F. and Crespi, Andrea and Ramponi, Roberta and Osellame, Roberto. General Rules for Bosonic Bunching in Multimode Interferometers. *Phys. Rev. Lett.*, 111:130503, Sep 2013. doi: 10.1103/PhysRevLett.111.130503. URL https://link.aps.org/doi/10.1103/PhysRevLett.111.130503.

[70] Carolan, Jacques and Meinecke, Jasmin DA and Shadbolt, Peter J and Russell, Nicholas J and Ismail, Nur and Worhoff, Kerstin and Rudolph, Terry and Thompson, Mark G and O'brien, Jeremy L and Matthews, Jonathan CF and others. On the experimental verification of quantum complexity in linear optics. *Nature Photonics*, 8(8):621–626, 2014.

[71] Drury, Stephen. A counterexample to a question of Bapat and Sunder. *The Electronic Journal of Linear Algebra*, 31:69–70, 2016.

[72] Shchesnovich, Valery S. The permanent-on-top conjecture is false. *Linear Algebra and its Applications*, 490:196–201, 2016.

[73] Dittel, Christoph and Dufour, Gabriel and Walschaers, Mattia and Weihs, Gregor and Buchleitner, Andreas and Keil, Robert. Totally destructive interference for permutation-symmetric many-particle states. *Phys. Rev. A*, 97:062116, Jun 2018. doi: 10.1103/PhysRevA.97.062116. URL https://link.aps.org/doi/10.1103/PhysRevA.97.062116.

[74] Shchesnovich, VS. Tight bound on the trace distance between a realistic device with partially indistinguishable bosons and the ideal BosonSampling. *Physical Review A*, 91(6):063842, 2015.

[75] Schur, Isaac. Uber endliche gruppen und hermitesche formen. *Mathematische Zeitschrift*, 1(2):184–207, 1918.

[76] Oppenheim, Alexander. Inequalities connected with definite Hermitian forms. *Journal of the London Mathematical Society*, 1(2):114–119, 1930.

[77] Gloub, Gene H and Van Loan, Charles F. Matrix computations. *Johns Hopkins Universtiy Press, 3rd edtion*, 1996.

[78] Flamini, Fulvio and Spagnolo, Nicolo and Sciarrino, Fabio. Photonic quantum information processing: a review. *Reports on Progress in Physics*, 82(1):016001, 2018.

[79] Pelucchi, Emanuele and Fagas, Giorgos and Aharonovich, Igor and Englund, Dirk and Figueroa, Eden and Gong, Qihuang and Hannes, Hubel and Liu, Jin and Lu, Chao-Yang and Matsuda, Nobuyuki and others. The potential and global outlook of integrated photonics for quantum technologies. *Nature Reviews Physics*, pages 1–15, 2021.

[80] Wang, Hui and Li, Wei and Jiang, Xiao and He, Y-M and Li, Y-H and Ding, Xing and Chen, M-C and Qin, Jian and Peng, C-Z and Schneider, Christian and others. Toward scalable boson sampling with photon loss. *Physical review letters*, 120(23):230502, 2018.

[81] Wang, Jianwei and Sciarrino, Fabio and Laing, Anthony and Thompson, Mark G. Integrated photonic quantum technologies. *Nature Photonics*, 14(5):273–284, 2020.

[82] Hoch, Francesco and Piacentini, Simone and Giordani, Taira and Tian, Zhen-Nan and Iuliano, Mariagrazia and Esposito, Chiara and Camillini, Anita and Carvacho, Gonzalo and Ceccarelli, Francesco and Spagnolo, Nicolo and others. Boson Sampling in a reconfigurable continuously-coupled 3D photonic circuit. *arXiv preprint arXiv:2106.08260*, 2021.

[83] Jacques Carolan, Christopher Harrold, Chris Sparrow, Enrique Martin-Lopez, 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.

[84] Crespi, Andrea and Osellame, Roberto and Ramponi, Roberta and Bentivegna, Marco and Flamini, Fulvio and Spagnolo, Nicolo and Viggianiello, Niko and Innocenti, Luca and Mataloni, Paolo and Sciarrino, Fabio. Suppression law of quantum states in a 3D photonic fast Fourier transform chip. *Nature communications*, 7(1): 1–8, 2016.

[85] Jönsson, Mattias and Björk, Gunnar. Evaluating the performance of photon-number-resolving detectors. *Phys. Rev. A*, 99:043822, Apr 2019. doi: 10.1103/PhysRevA.99.043822. URL https://link.aps.org/doi/10.1103/PhysRevA.99.043822.

[86] Cheng, Risheng and Zhou, Yiyu and Wang, Sihao and Shen, Mohan and Taher, Towsif and Tang, Hong X. Unveiling photon statistics with a 100-pixel photon-number-resolving detector. *arXiv:2206.13753*, 2022.

[87] Pryde, Geoff J and White, Andrew G. Creation of maximally entangled photon-number states using optical fiber multiports. *Physical Review A*, 68(5):052315, 2003.

[88] Tichy, Malte C and Lim, Hyang-Tag and Ra, Young-Sik and Mintert, Florian and Kim, Yoon-Ho and Buchleitner, Andreas. Four-photon indistinguishability transition. *Physical Review A*, 83(6):062111, 2011.

[89] Tillmann, Max and Tan, Si-Hui and Stoeckl, Sarah E and Sanders, Barry C and De Guise, Hubert and Heilmann, Rene and Nolte, Stefan and Szameit, Alexander and Walther, Philip. Generalized multiphoton quantum interference. *Physical Review X*, 5(4):041015, 2015.

[90] Peter S. Turner. Postselective quantum interference of distinguishable particles. *arXiv: Quantum Physics*, 2016.

[91] Jones, Alex E and Menssen, Adrian J and Chrzanowski, Helen M and Wolterink, Tom AW and Shchesnovich, Valery S and Walmsley, Ian A. Multiparticle interference of pairwise distinguishable photons. *Physical Review Letters*, 125(12):123603, 2020.

[92] Menssen, Adrian J and Jones, Alex E and Metcalf, Benjamin J and Tichy, Malte C and Barz, Stefanie and Kolthammer, W Steven and Walmsley, Ian A. Distinguishability and many-particle interference. *Physical review letters*, 118(15): 153603, 2017.

[93] Shchesnovich, VS and Bezerra, MEO. Collective phases of identical particles interfering on linear multiports. *Physical Review A*, 98(3):033805, 2018.

[94] Jones, Alexander Edward. Distinguishability in quantum interference. 2019.

[95] Drury, Stephen. A real counterexample to two inequalties involving permanents. *Mathematical Inequalities & Applications*, 2017.

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

[97] Zhang, Fuzhen. Notes on Hadamard products of matrices. *Linear and Multilinear Algebra*, 25(3):237–242, 1989.

[98] Seron, Benoit and Restivo, Antoine. BosonSampling.jl: A Julia package for quantum multi-photon interferometry. 2022.

[99] Minc, Henryk. *Permanents*. Cambridge University Press, 1984.

[100] Preskill, John. Quantum computing in the NISQ era and beyond. *Quantum*, 2018.

[101] Preskill, John. Quantum computing 40 years later. *arXiv:2106.10522*, 2021.

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

[103] 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.

[104] Chakhmakhchyan, Levon and Cerf, Nicolas J. Boson sampling with Gaussian measurements. *Physical Review A*, 96(3):032326, 2017.

[105] 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.

[106] 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.

[107] Robens, Carsten and Arrazola, Inigo and Alt, Wolfgang and Meschede, Dieter and Lamata, Lucas and Solano, Enrique and Alberti, Andrea. Boson Sampling with Ultracold Atoms. *arXiv preprint arXiv:2208.12253*, 2022.

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

[109] Renema, J. J. and Menssen, A. and Clements, W. R. and Triginer, G. and Kolthammer, W. S. and Walmsley, I. A. Efficient Classical Algorithm for Boson Sampling with Partially Distinguishable Photons. *Phys. Rev. Lett.*, 120:220502, May 2018.

[110] Oszmaniec, Michal and Brod, Daniel J. Classical simulation of photonic linear optics with lost particles. *New Journal of Physics*, 20(9):092002, 2018.

[111] Brod, Daniel Jost and Oszmaniec, Michal. Classical simulation of linear optics subject to nonuniform losses. *Quantum*, 4:267, 2020.

[112] 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.

[113] Chabaud, Ulysse and Grosshans, Frederic and Kashefi, Elham and Markham, Damian. Efficient verification of Boson sampling. *Quantum*, 5:578, 2021.

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

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

[116] Kalai, Gil and Kindler, Guy. Gaussian noise sensitivity and BosonSampling. *arXiv:1409.3093*, 2014.

[117] Leverrier, Anthony and Garcia-Patron, Raul. Analysis of circuit imperfections in BosonSampling. *Quantum Information & Computation 15, 0489-0512*, 2015.

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

[119] Arkhipov, Alex. BosonSampling is robust against small errors in the network matrix. *Physical Review A*, 92(6):062326, 2015.

[120] Aaronson, Scott and Brod, Daniel J. BosonSampling with lost photons. *Physical Review A*, 93(1):012335, 2016.

[121] Latmiral, Ludovico and Spagnolo, Nicolo and Sciarrino, Fabio. Towards quantum supremacy with lossy scattershot boson sampling. *New Journal of Physics*, 18 (11):113008, 2016.

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

[123] Flamini, Fulvio and Spagnolo, Nicolo and Sciarrino, Fabio. Visual assessment of multi-photon interference. *Quantum Science and Technology*, 4(2):024008, 2019.

[124] 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.

[125] Wang, Sheng-Tao and Duan, Lu-Ming. Certification of boson sampling devices with coarse-grained measurements. *arXiv:1601.02627*, 2016.

[126] Mezher, Rawad and Mansfield, Shane. Assessing the quality of near-term photonic quantum devices. *arXiv preprint arXiv:2202.04735*, 2022.

[127] Moylett, Alexandra E and Garcia-Patron, Raul and Renema, Jelmer J and Turner, Peter S. Classically simulating near-term partially-distinguishable and lossy boson sampling. *Quantum Science and Technology*, 5(1):015001, 2019.

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

[129] Walschaers, Mattia. *Statistical Benchmarks for Quantum Transport in Complex Systems: From Characterisation to Design*. Springer, 2018.

[130] Giordani, Taira and Flamini, Fulvio and Pompili, Matteo and Viggianiello, Niko and Spagnolo, Nicolo and Crespi, Andrea and Osellame, Roberto and Wiebe, Nathan and Walschaers, Mattia and Buchleitner, Andreas and others. Experimental statistical signature of many-body quantum interference. *Nature Photonics*, 12(3):173–178, 2018.

[131] Seron, Benoit and Novo, Leonardo and Cerf, Nicolas J. Boson bunching is not maximized by indistinguishable particles. *arXiv:2203.01306*, 2022.

[132] Tichy, Malte C and Mayer, Klaus and Buchleitner, Andreas and Molmer, Klaus. Stringent and efficient assessment of boson-sampling devices. *Physical review letters*, 113(2):020502, 2014.

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

[134] Viggianiello, Niko and Flamini, Fulvio and Innocenti, Luca and Cozzolino, Daniele and Bentivegna, Marco and Spagnolo, Nicolo and Crespi, Andrea and Brod, Daniel J and Galvao, Ernesto F and Osellame, Roberto and others. Experimental generalized quantum suppression law in Sylvester interferometers. *New Journal of Physics*, 20(3):033017, 2018.

[135] Crespi, Andrea. Suppression laws for multiparticle interference in Sylvester interferometers. *Physical Review A*, 91(1):013811, 2015.

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

[137] Johnson, Charles R. *Matrix theory and applications*, volume 40. American Mathematical Soc., 1990.

[138] Ivanov, Dmitri A. and Gurvits, Leonid. 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.

[139] Renema, Jelmer J. Marginal probabilities in boson samplers with arbitrary input states. *arXiv:2012.14917*, 2020.

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

[141] 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, 2021.

[142] 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.

[143] Shchesnovich, VS. Quantum de Moivre–Laplace theorem for noninteracting indistinguishable particles in random networks. *Journal of Physics A: Mathematical and Theoretical*, 50(50):505301, 2017.

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

[145] 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.

[146] Dai, Zhe and Liu, Yong and Xu, Ping and Xu, WeiXia and Yang, XueJun and Wu, JunJie. A Bayesian validation approach to practical boson sampling. *Science China Physics, Mechanics & Astronomy*, 63(5):1–8, 2020.

[147] Flamini, Fulvio and Walschaers, Mattia and Spagnolo, Nicolo and Wiebe, Nathan and Buchleitner, Andreas and Sciarrino, Fabio. Validating multi-photon quantum interference with finite data. *Quantum Science and Technology*, 5(4):045005, 2020.

[148] Nikolopoulos, Georgios M and Brougham, Thomas. Decision and function problems based on boson sampling. *Physical Review A*, 94(1):012315, 2016.

[149] Nikolopoulos, Georgios M. Cryptographic one-way function based on boson sampling. *Quantum Information Processing*, 18(8):1–25, 2019.

[150] Oh, Changhun and Lim, Youngrong and Wong, Yat and Fefferman, Bill and Jiang, Liang. Quantum-inspired classical algorithm for molecular vibronic spectra. *arXiv preprint arXiv:2202.01861*, 2022.

[151] 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.

[152] 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.

[153] Seron, Benoit and Restivo, Antoine. BosonSampling.jl, . URL https://juliahub.com/ui/Packages/BosonSampling/olGSq/1.0.1.

[154] Minc, Henryk. *Permanents*, volume 6. Cambridge University Press, 1984.

[155] Feynman, Richard P. Simulating physics with computers. In *Feynman and computation*, pages 133–153. CRC Press, 2018.

[156] Max Tillmann and Borivoje Dakic and Rene Heilmann and Stefan Nolte and Alexander Szameit and Philip Walther. Experimental boson sampling. *Nature Photonics*, 7(7):540–544, may 2013. doi: 10.1038/nphoton.2013.102. URL https://doi.org/10.1038%2Fnphoton.2013.102.

[157] Marco Bentivegna and Nicolo Spagnolo and Chiara Vitelli and Fulvio Flamini and Niko Viggianiello and Ludovico Latmiral and Paolo Mataloni and Daniel J. Brod and Ernesto F. Galvao and Andrea Crespi and Roberta Ramponi and Roberto Osellame and Fabio Sciarrino. Experimental scattershot boson sampling. *Science Advances*, 1(3), apr 2015. doi: 10.1126/sciadv.1400255. URL https://doi.org/10.1126%2Fsciadv.1400255.

[158] Seron, Benoit and Restivo, Antoine. BosonSampling.jl, . URL https://github.com/benoitseron/BosonSampling.jl.git.

[159] Seron, Benoit and Restivo, Antoine. BosonSampling.jl Documentation, . URL https://docs.juliahub.com/BosonSampling/olGSq/1.0.1/.

[160] Seron, Benoit and Novo, Leonardo and Arkhipov, Alex and Cerf, Nicolas J. Efficient validation of boson sampling from binned photon-number distributions. 2022.

[161] Abrudan, Traian and Eriksson, Jan and Koivunen, Visa. Conjugate gradient algorithm for optimization under unitary matrix constraint. *Signal Processing*, 89 (9):1704–1714, 2009.

[162] George Datseris. DynamicalSystems.jl: A Julia software library for chaos and nonlinear dynamics. *Journal of Open Source Software*, 3(23):598, 2018. doi: 10.21105/joss.00598. URL https://doi.org/10.21105/joss.00598.

[163] SatelliteToolBox. URL https://juliaspace.github.io/SatelliteToolbox.jl/dev/.

[164] Regier, Jeffrey and Pamnany, Kiran and Giordano, Ryan and Thomas, Rollin and Schlegel, David and McAuliffe, Jon and others. Learning an astronomical catalog of the visible universe through scalable Bayesian inference. *arXiv preprint arXiv:1611.03404*, 2016.

[165] Ali Ramadhan and Gregory LeClaire Wagner and Chris Hill and Jean-Michel Campin and Valentin Churavy and Tim Besard and Andre Souza and Alan Edelman and Raffaele Ferrari and John Marshall. Oceananigans.jl: Fast and friendly geophysical fluid dynamics on GPUs. *Journal of Open Source Software*, 5 (53):2018, 2020. doi: 10.21105/joss.02018. URL https://doi.org/10.21105/joss.02018.

[166] Ranjan Anantharaman and Kimberly Hall and Viral B. Shah and Alan Edelman. Circuitscape in Julia: High Performance Connectivity Modelling to Support Conservation Decisions. *Proceedings of the JuliaCon Conferences*, 1(1):58, 2020. doi: 10.21105/jcon.00058. URL https://doi.org/10.21105/jcon.00058.

[167] Piotr Gawron, Dariusz Kurzyk, and Lukasz Pawela. Quantuminformation. jl—a julia package for numerical computation in quantum information theory. *PLoS One*, 13(12):e0209358, 2018.

[168] Xiu-Zhe Luo and Jin-Guo Liu and Pan Zhang and Lei Wang. Yao.jl: Extensible, Efficient Framework for Quantum Algorithm Design. *Quantum*, 4:341, oct 2020. doi: 10.22331/q-20720-10-11-341.

[169] Sebastian Krämer and David Plankensteiner and Laurin Ostermann and Helmut Ritsch. QuantumOptics.jl: A Julia framework for simulating open quantum systems. *Computer Physics Communications*, 227:109–116, jun 2018. doi: 10.1016/j.cpc.2018.02.004.

[170] David Plankensteiner and Christoph Hotter and Helmut Ritsch. QuantumCumulants.jl: A Julia framework for generalized mean-field equations in open quantum systems. *Quantum*, 6:617, jan 2022. doi: 10.22331/q-2022-01-04-617.

[171] Bezanson, Jeff and Karpinski, Stefan and Shah, Viral B. and Edelman, Alan. Julia: A Fast Dynamic Language for Technical Computing, 2012. URL https://arxiv.org/abs/1209.5145.

[172] Avik Sengupta. *Julia High Performance*. Packt, 2019. ISBN 978-1-78829-811-7.

[173] Shi, Junheng and Byrnes, Tim. Gaussian boson sampling with partial distinguishability, 2021. URL https://arxiv.org/abs/2105.09583.

[174] Alexandra E Moylett and Rau l Garcia-Patron and 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/10.10882/2058-95652/ab5555.

[175] Hong, C. K. and Ou, Z. Y. and Mandel, L. Measurement of subpicosecond time intervals between two photons by interference. *Phys. Rev. Lett.*, 59:2044–2046, Nov 1987. doi: 10.1103/PhysRevLett.59.2044. URL https://link.aps.org/doi/10.1103/PhysRevLett.59.2044.

[176] Leech, John. H. J. Ryser, Combinatorial Mathematics (Carus Mathematical Monographs, No. 14; published by The Mathematical Association of America, distributed by John Wiley and Sons, 1963), xiv 154 pp., 30s. *Proceedings of the Edinburgh Mathematical Society*, 14(1):82–83, 1964. doi: 10.1017/S0013091500011299.

[177] Brajesh Gupt and Josh Izaac and Nicolás Quesada. The Walrus: a library for the calculation of hafnians, Hermite polynomials and Gaussian boson sampling. *Journal of Open Source Software*, 4(44):1705, 2019. doi: 10.21105/joss.01705. URL https://doi.org/10.21105/joss.01705.

[178] Heurtel, Nicolas and Fyrillas, Andreas and de Gliniasty, Gregoire and Bihan, Raphael Le and Malherbe, Sebastien and Pailhas, Marceau and Bourdoncle,

Boris and Emeriau, Pierre-Emmanuel and Mezher, Rawad and Music, Luka and Belabas, Nadia and Valiron, Benoit and Senellart, Pascale and Mansfield, Shane and Senellart, Jean. Perceval: A Software Platform for Discrete Variable Photonic Quantum Computing, 2022. URL https://arxiv.org/abs/2204.00602.

[179] Osca, Javier and Vala, Jiri. Implementation of photon partial distinguishability in a quantum optical circuit simulation, 2022. URL https://arxiv.org/abs/2208.03250.

[180] R.B. Bapat and V.S. Sunder. An extremal property of the permanent and the determinant. *Linear Algebra and its Applications*, 76:153–163, 1986. ISSN 0024-3795. doi: https://doi.org/10.1016/0024-3795(86)90220-X. URL https://www.sciencedirect.com/science/article/pii/002437958690220X.

[181] Drury, Stephen. A counterexample to a question of bapat sunder. *Mathematical Inequalities Applications*, 21:517–520, 04 2018. doi: 10.7153/mia-2018-21-37.

[182] Lieb, Elliott H. Proofs of some conjectures on permanents. *Inequalities: Selecta of Elliott H. Lieb*, pages 101–108, 2002.

[183] Wanless, Ian M. Lieb's permanental dominance conjecture. *arXiv preprint arXiv:2202.01867*, 2022.

[184] Tichy, Malte and Molmer, Klaus. Extending exchange symmetry beyond bosons and fermions. *Physical Review A*, 96(2), aug 2017. doi: 10.1103/physreva.96.022119. URL https://doi.org/10.1103/physreva.96.022119.

[185] Valery Shchesnovich. Distinguishability in quantum interference with multimode squeezed states. *Physical Review A*, 105(6), jun 2022. doi: 10.1103/physreva.105.063703. URL https://doi.org/10.11032/physreva.105.063703.

[186] Renema, Jelmer J. Simulability of partially distinguishable superposition and gaussian boson sampling. *Phys. Rev. A*, 101:063840, Jun 2020. doi: 10.1103/PhysRevA.101.063840. URL https://link.aps.org/doi/10.1103/PhysRevA.101.063840.

[187] Shi, Junheng and Byrnes, Tim. Effect of partial distinguishability on quantum supremacy in gaussian boson sampling. *npj Quantum Information*, 8(1):54, 2022.

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

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

[190] Wang, Xiao-Wei and Zhou, Wen-Hao and Fu, Yu-Xuan and Gao, Jun and Lu, Yong-Heng and Chang, Yi-Jun and Qiao, Lu-Feng and Ren, Ruo-Jing and Jiang,

Ze-Kun and Jiao, Zhi-Qiang and Nikolopoulos, Georgios M. and Jin, Xian-Min. Experimental boson sampling enabling cryptographic one-way function. *Phys. Rev. Lett.*, 130:060802, Feb 2023. doi: 10.1103/PhysRevLett.130.060802. URL https://link.aps.org/doi/10.1103/PhysRevLett.130.060802.

[191] Troyansky, Lidror and Tishby, Naftali. On the quantum evaluation of the determinant and the permanent of a matrix. *Proc. Phys. Comput*, page 96, 1996.

[192] Huh, Joonsuk and Guerreschi, Gian Giacomo and Peropadre, Borja and McClean, Jarrod R. and Aspuru-Guzik, Alan. Boson sampling for molecular vibronic spectra. *Nature Photonics*, 9(9):615–620, aug 2015. doi: 10.1038/nphoton.2015.153. URL https://doi.org/10.10382/nphoton.2015.153.

[193] Arrazola, Juan Miguel and Bromley, Thomas R. Using gaussian boson sampling to find dense subgraphs. *Phys. Rev. Lett.*, 121:030503, Jul 2018. doi: 10.1103/PhysRevLett.121.030503. URL https://link.aps.org/doi/10.1103/PhysRevLett.121.030503.

[194] Sempere-Llagostera, S. and Patel, R. B. and Walmsley, I. A. and Kolthammer, W. S. Experimentally finding dense subgraphs using a time-bin encoded gaussian boson sampling device. *Phys. Rev. X*, 12:031045, Sep 2022. doi: 10.1103/PhysRevX.12.031045. URL https://link.aps.org/doi/10.1103/PhysRevX.12.031045.

[195] Banchi, Leonardo and Fingerhuth, Mark and Babej, Tomas and Ing, Christopher and Arrazola, Juan Miguel. Molecular docking with gaussian boson sampling. *Science Advances*, 6, 2019.

[196] Mezher, Rawad and Carvalho, Ana Filipa and Mansfield, Shane. Solving graph problems with single-photons and linear optics. *arXiv preprint arXiv:2301.09594*, 2023.

[197] Novo, Leonardo and Bermejo-Vega, Juani and Garcia-Patron, Raul. Quantum advantage from energy measurements of many-body quantum systems. *Quantum*, 5:465, 2021.