Statistical benchmark for BosonSampling

New Journal of Physics

The open access journal at the forefront of physics

CrossMark

FAST TRACK COMMUNICATION

Statistical benchmark for BosonSampling

Mattia Walschaers[1,2], **Jack Kuipers**[3,4], **Juan-Diego Urbina**[3], **Klaus Mayer**[1], **Malte Christopher Tichy**[5], **Klaus Richter**[3] and **Andreas Buchleitner**[1]

[1] Physikalische Institut, Albert-Ludwigs Universität Freiburg, Freiburg, D-79104 Germany
[2] Instituut voor Theoretische Fysica, KU Leuven, Celestijnenlaan 200D, 3001 Leuven, Belgium
[3] Institut für Theoretische Physik, Universität Regensburg, D-93040 Regensburg, Germany
[4] D-BSSE, ETH Zürich, Mattenstrasse 26, 4058 Basel, Switzerland
[5] Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, DK-8000 Aarhus, Denmark

E-mail: mattia.walschaers@physik.uni-freiburg.de

Abstract

Boson samplers—set-ups that generate complex many-particle output states through the transmission of elementary many-particle input states across a multitude of mutually coupled modes—promise the efficient quantum simulation of a classically intractable computational task, and challenge the extended Church–Turing thesis, one of the fundamental dogmas of computer science. However, as in all experimental quantum simulations of truly complex systems, one crucial problem remains: how to certify that a given experimental measurement record unambiguously results from enforcing the claimed dynamics, on bosons, fermions or distinguishable particles? Here we offer a statistical solution to the certification problem, identifying an unambiguous statistical signature of many-body quantum interference upon transmission across a multimode, random scattering device. We show that statistical analysis of only partial information on the output state allows to characterise the imparted dynamics through particle type-specific features of the emerging interference patterns. The relevant statistical quantifiers are classically computable, define a falsifiable benchmark for BosonSampling, and reveal distinctive features of many-particle quantum dynamics, which go much beyond mere bunching or anti-bunching effects.

1. Introduction

As the world waits for the first universal and fully operational quantum computer [1], quantum information scientists are eager to already show the power of quantum physics to perform computational tasks which are out of reach for a classical computer. Rather than designing devices that can perform a wide range of calculations, machines which are specialised in specific tasks have joined the scope [2, 3]. Here, we focus on one type of such devices, the BosonSamplers, which may hold the key to falsifying the extended Church–Turing thesis [1, 4, 5]. This conjecture, rooted in the early days of computer science, states that any efficient calculation performed by a physical device can also be performed in polynomial time on a classical computer. It is now proposed that all that is necessary to falsify this foundational dogma of computer science is a set of m photonic input modes, which are connected to m output modes by a random photonic circuit [2]. This immediately indicates why BosonSampling attracts such attention, as these systems are experimentally in reach [6–11].

Since such a device is operating on bosons, indistinguishable particles, interesting physics arises when multiple particles are simultaneously injected into the system [12–16]. A schematic overview, indicating the essential ingredients of a sampling device such as considered here, is provided in figure 1. BosonSampling essentially consists in sampling an occupation vector $\vec{y} = (y_1,...,y_m)$ for the output modes, given that the initial mode occupation was $\vec{x} = (x_1,...,x_m)$, where x_i and y_i are the number of photons found in the ith input and output mode, respectively. In the case with at most one photon per input mode, this probability is given by

Figure 1. Sketch of the system under consideration. From a set of input modes (here $m = 12$), four particles are injected (depicted in red). The particles traverse the system via the depicted channels. At the crossings, different paths are interconnected such that particles can travel in each direction, thus leading to single-particle *and* many-body interference. The yellow zone of the setup is described by a unitary matrix U, which connects the input modes to the output modes. The output signal is probabilistic in nature (hence different intensities of red), with its statistics governed by the many-body quantum state. The key object of this work, the C-dataset (see main text), is obtained by calculating the two-point correlations between different modes as depicted in green.

$$p_{\vec{x} \to \vec{y}} = \frac{|\mathrm{perm}\, U_{\vec{x},\vec{y}}|^2}{\prod_{i=1}^{m} y_i!}, \tag{1}$$

where $U_{\vec{x},\vec{y}}$ is the $n \times n$ matrix, constructed from U, which describes how the n occupied input modes are connected to the selected output modes [13], and 'perm' denotes the *permanent* [17]. To evaluate the many-body output wave function, and thus the sampling statistics, we must deal with these permanents, which are 'hard' to compute, i.e. there is no algorithm that can do so in polynomial time [17, 18]. *A priori*, there may be algorithms that *sample* from the probability distribution $p_{\vec{x} \to \vec{y}}$ without explicit computation of the probabilities. This possibility was recently refuted [2] and it is argued that it is in general not possible[6] to efficiently simulate such a BosonSampling procedure on a *classical* computer. However, it turns out that single photons, linear optics and photon counters are all the required building blocks to build a *quantum* device that performs the task. Thus, the BosonSampler is a physical system that efficiently samples bosons according to the bosonic many-body wave function, even though this many-body state *cannot* be calculated by a classical computer. Its realisation would in this perspective invalidate the extended Church-Turing thesis in the sense that it is a concrete example of a quantum device that performs a task efficiently where a classical computer could not.

However, this strength of BosonSampling apparently also implies a profound weakness: since it is impossible to compute the many-body wave function, one cannot certify that a given experimental output unambiguously stems from sampling bosons–how can we be sure that an alleged device samples events from the correct probability distribution? Note that different scientific disciplines impose very different requirements on certification protocols: from the perspective of theoretical computer science, which is motivated by the scrupulous applications of cryptography, a proof for the correct functionality of a BosonSampler can only be accepted if the possibility of a fraudulent classical device that generates a false-positive result is provably ruled out. That is, generating a sample that passes the certification must be computationally much harder than the very act of certifying. Setting aside this cryptographic perspective and disregarding the possibility of the wilful use of fraudulent devices, we focus on unambiguous physical signatures of the functionality of the device. Bunching of bosons was proposed [19] as a decisive criterion, but dismissed in [20], where a mean-field ansatz incarnated by 'simulated bosons' (to be defined below) was shown to precisely reproduce coarse-grained bunching. Reference [20] furthermore offers a highly symmetric, analytically solvable test-case (recently implemented experimentally [21, 22]), which however remains, by construction, outside the realm of a true BosonSampler. A reliable physical benchmark for the functionality of BosonSamplers therefore is unachieved to date, yet highly desirable if one wishes to fully exploit the computational power brought about by many-particle quantum interference. To fill this gap, we conceive a solution based on a statistical certifier that excludes simulated bosons—even for the original formulation [2] of the BosonSampling problem with random scattering matrices. Borrowing from the vast set of techniques offered by statistical mechanics, we formulate the problem in terms of a *transport process* in a scattering system (e.g. a photonic network), described by a random unitary matrix U. We show that, by

---

[6] To be precise: it was shown [2] that the realisation of an efficient classical algorithm that simulates BosonSampling implies a collapse of the polynomial hierarchy [5] to the third order. As it is not our goal to enter the mathematical subtleties related to the theory of computational complexity, let us stress that this is simply the complexity theorist's way of saying that it is highly unlikely to be true.

combining quantum optics with methods found in statistical physics and random matrix theory, it is indeed possible to identify signatures of genuine bosonic many-particle interference with manageable overhead.

2. Statistical signatures of many-particle interference

It must be stressed that the resulting many-body wave function of a scattering process, such as manifested in a BosonSampler, is determined by several factors. One can separate them into those which are of quantum statistical origin and those which are dynamical in nature. The former concern all effects related to the structure of state space (e.g., bunching for bosons, and the Pauli exclusion principle for fermions), whereas the latter involve all sorts of interference effects. One can divide these interferences into the well-known single-particle interference [23], encoded within the matrix U, which arises due to the wave-like nature of quantum transport [24–26], and the far more intricate many-body interference [15, 16, 27]. Below, we shall scrutinise the interference exhibited by various particle types—bosons, fermions, distinguishable particles and simulated bosons.

   Distinguishable particles are the simplest of the considered species, as their signals are governed only by single-particle interference. Transport processes with bosons or fermions, in contrast, exhibit the entire range of statistical *and* interference effects. As indistinguishable particles, they obey quantum statistics (as dictated by the relevant algebra), which for two particles in two modes either leads to *bunching* (bosons) [12] or to *anti-bunching* (fermions) [28]. Similar effects have also been observed in larger setups, with more particles, and have hence been proposed as hallmarks for many-boson [19] or many-fermion [29] behaviour.

   However, for larger setups one expects a much richer phenomenology due to many-body interference [15], resulting from the coherent superposition of distinct many-particle transmission amplitudes [16]. For BosonSampling, bunching or clouding behaviour [19] as such is therefore not a sufficient tool for certification, as it can easily be achieved by the so-called mean-field sampler [20] which was inspired by semiclassical models [30]. These devices are designed to replicate the bunching behaviour by approximating the many-body output state by a macroscopically populated single-particle state with random phases added. Averaging over the random phases successfully mimics boson-like bunching. However, all relative phases between the initially populated many-particle transmission amplitudes are scrambled by the averaging procedure, and, consequently, all many-body interference effects are deleted. This type of sampling [19] is easily simulated by Monte Carlo methods [20]—hence we here refer to such sampled particles as *simulated bosons*—and is unable to harness quantum granularity on the many-particle level. It is therefore conceptually decisive to set mean field sampling apart from actual BosonSampling. We now introduce a general method which is specifically conceived to detect quantum interference structures as induced by the coherent superposition of many-particle transmission amplitudes. This approach thus distinguishes simulated from true bosons, and identifies many-particle interference features which are characteristic of both—particle type *and* the specific scattering dynamics.

3. Random matrix methods

The scattering matrix U that describes the photonic circuit in the BosonSampling setup is randomly sampled from the *Haar measure* [31–33]. Therefore it is only natural to treat this problem in a framework of statistics and random matrix theory (RMT) [31, 34–37]. Often the lack of grasp on the statistical distribution of the full many-body wave function is put forth as the core of the certification problem. Exhaustive statistical characterisation of the many-body state would require the full distribution of permanents over the set of unitary matrices. To date, only the first moment of this distribution is known [38] and it is not enough to provide certification, while sufficiently precise higher order moments are out of reach [2]. The reason is that, in terms of the quantum state, permanents depend on high-order correlation functions. In contrast, we will emphasise below that a gigantic amount of information on the many-body state *is* within reach in the form of distributions of low-order correlation functions.

   In probability theory, the knowledge of all possible correlation functions implies full knowledge of the (joint) probability distribution itself. Therefore, correlation functions play a central role in many probabilistic theories, from RMT [31] to quantum statistical mechanics [39]. In practical applications of RMT, such as occur in quantum chaos, a study of the statistics related to two-point correlations is often sufficient to certify the RMT ensemble. Similarly, here, we do not (and, for sufficiently large systems, cannot) know *all* correlation functions of the output many-body quantum state. There is, however, a large set of correlation functions which *are* accessible (both theoretically and experimentally) [14, 38, 40]. Hence, the relevant question is whether this set offers a sufficient amount of information to distinguish the many-particle dynamics undergone by simulated or true bosons, fermions and distinguishable particles.

4. Statistical benchmarking

As the arguably simplest accessible correlation function we propose the mode correlator

$$C_{ij} = \langle \hat{n}_i \hat{n}_j \rangle - \langle \hat{n}_i \rangle \langle \hat{n}_j \rangle \tag{2}$$

(with $\hat{n}_i = a_i^\dagger a_i$ the bosonic number operator) [14], which quantifies how the number of outgoing photons in modes i and j are correlated (as indicated in figure 1). Scattering dynamics and particle type are encrypted in the many-body output quantum state $|\varphi_{\text{out}}\rangle$, which enters C_{ij} via the expectation value $\langle . \rangle = \langle \varphi_{\text{out}} | . | \varphi_{\text{out}} \rangle$.

Closer inspection of the expression for $\langle \hat{n}_i \hat{n}_j \rangle$ in [14], for particles initially prepared in mode $\{q_1, ..., q_n\}$, shows that

$$\langle \hat{n}_i \hat{n}_j \rangle = \sum_{k \neq l=1}^{n} |U_{q_k,i}|^2 |U_{q_l,j}|^2 + t \, U_{q_k,i} U_{q_l,j} U_{q_l,i}^* U_{q_k,j}^*, \tag{3}$$

where $t = -1, 0, +1$ for fermions, distinguishable particles, and bosons, respectively (the more intricate result for simulated bosons can be found in appendix A). With respect to $p_{\vec{x} \to \vec{y}}$ in (1), expression (3) can be interpreted as a sum over all two-particle processes which connect a pair of input particles to the two selected output modes. *In concreto*, $|U_{q_k,i}|^2 |U_{q_l,j}|^2$ gives the classical transmission probability, independent of the particle type. The second term on the right hand side of (3) provides the interference contribution for the two particles selected from the input state. In the definition (2) of the correlators we suppress the classical contribution to the transmission probability by subtracting $\langle \hat{n}_i \rangle \langle \hat{n}_j \rangle$, and thus focus on the interference terms (see also equations (A.4) and (A.5)). Therefore, C_{ij} probes all possible two-particle interference contributions to the population of modes i and j.
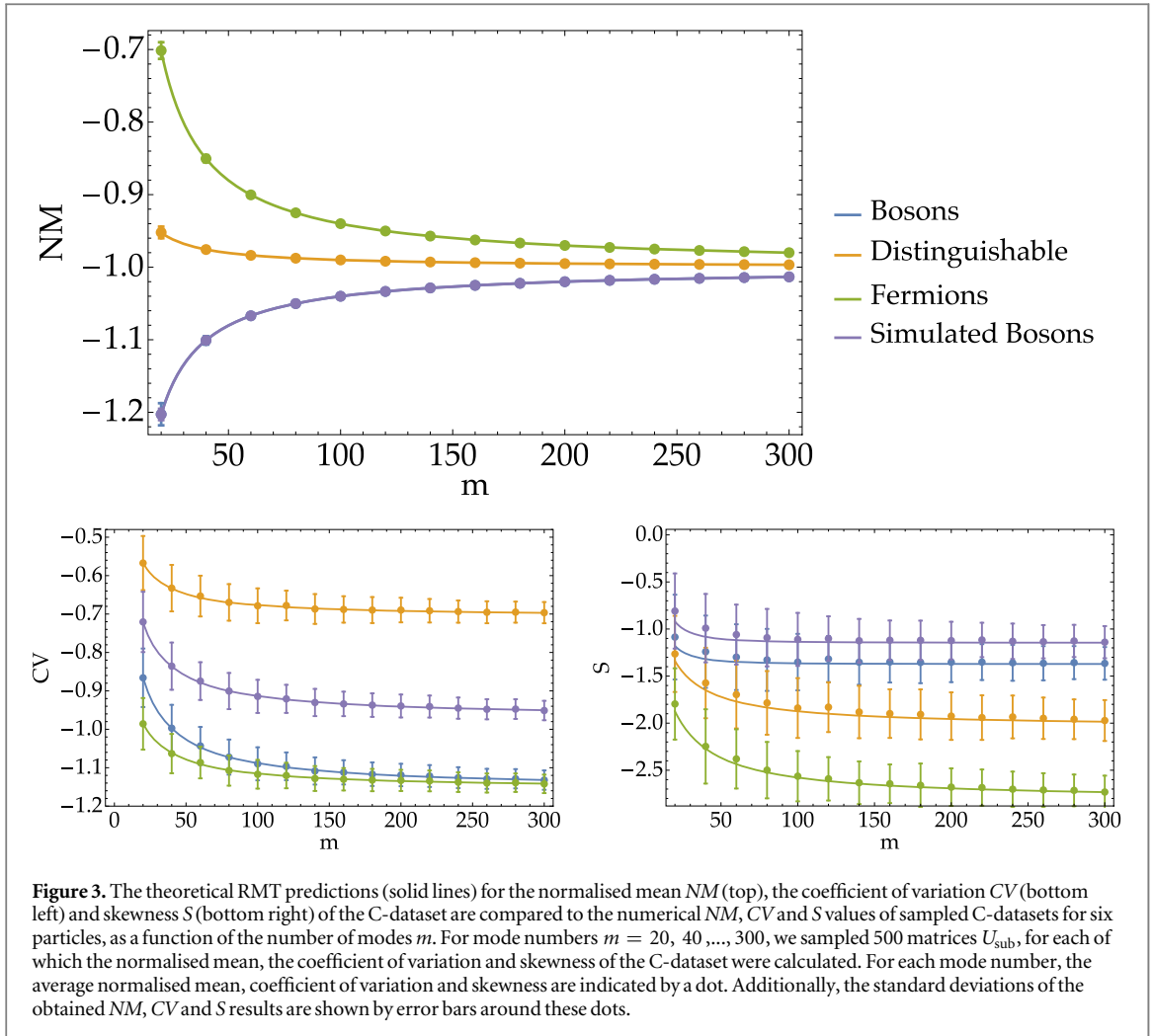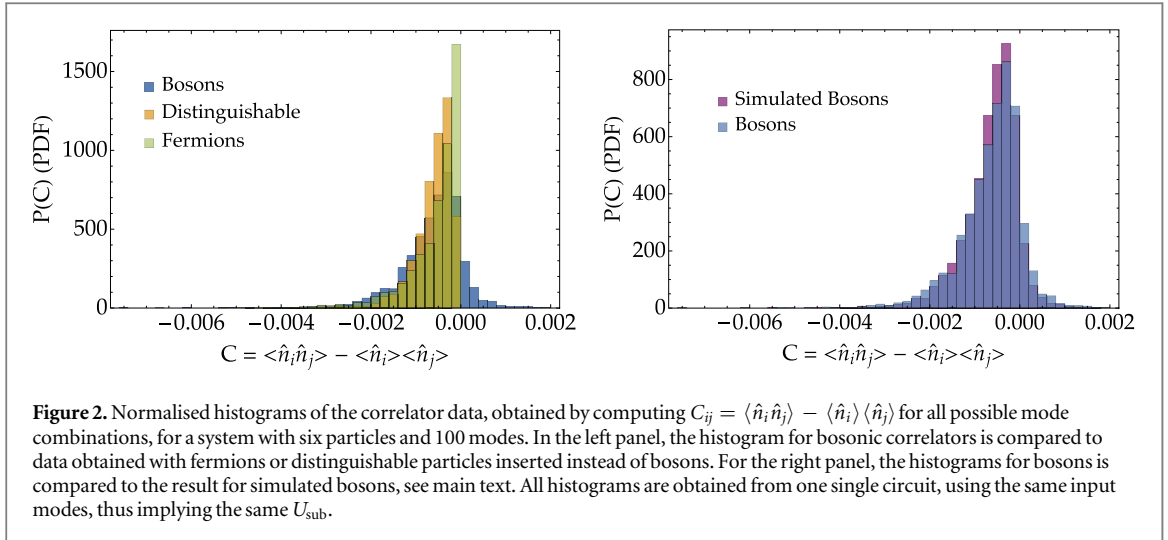
A single such correlator (typically) cannot probe in sufficient detail the structure of the full output wave function, but some characteristic feature of the interference pattern should become apparent once a sufficiently large set of correlators is considered by variation of i and j. The resulting dataset, which we from now on refer to as the *C-dataset*, is easily obtained in the experimental setup: One must consider sufficiently many choices for output modes, $i, j \in \{1, ..., m\}$ such that $i < j$, and for each choice compute the correlation C_{ij} between the number of particles sampled in the two modes. Now, for a *single* choice of input modes and hence a *single* $n \times m$ unitary matrix U_{sub}, the submatrix of U that describes how the input modes are coupled to all possible output modes, we obtain a set of data on which we can do statistics.

Given U_{sub}, it is possible to calculate the C-dataset for each particle type numerically exactly [14] (see (3) and appendix A). Although we can explore this via histograms, moments and other statistical properties, it is far from straightforward to get an analytical grasp of its statistics. Ideally, we would like to predict the exact shape of the distribution, given the number of modes m and the number of incoming particles n, but this appears to be an unrealistic goal. Nevertheless, after longwinded RMT calculations, we obtained analytical predictions for the first three moments of the set of possible outcomes when varying U_{sub} for fixed i and j (rather than fixing U_{sub} and varying i and j). Although the distributions are mathematically not exactly equivalent, we find good agreement with numerics (for a more elaborate discussion, see appendix B).
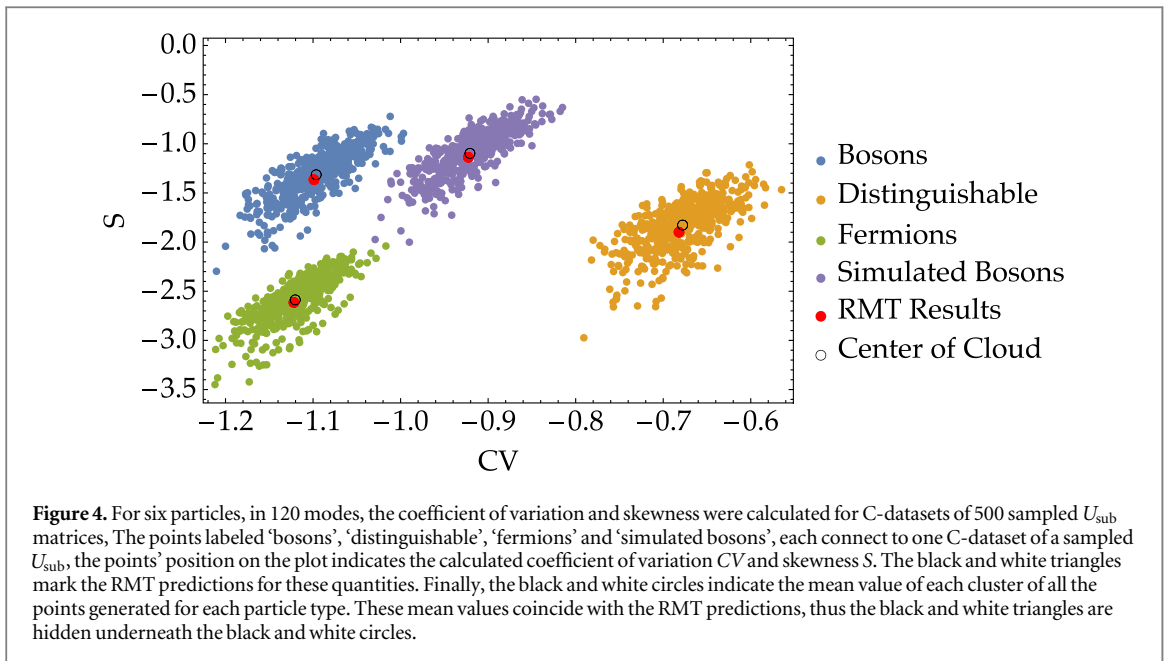
For a generated C-dataset for the different particle types, considering one fixed U_{sub} for 120 output modes and six particles, the left panel in figure 2 clearly shows a qualitative difference in the histograms for different particle types. In contrast, the right panel indicates that the histograms of the true bosons (where quantum statistical bunching and multi-particle interference do both contribute) and their simulated counterparts (which only exhibit bunching, but no multi-particle interference) bear a strong resemblance. Consequently, a quantitative understanding is essential to clearly distinguish the signature of bosonic many-particle interference from the quantum dynamics of other species. The second and third moment of the obtained correlator dataset can exactly provide us with such insight. To obtain these quantities involves averaging products of components of unitary matrices, for which straightforward (but tedious) combinatorics are used [36, 37].

5. Particle type-specific features of interference patterns

To acquire the clearest distinction between the many-particle interference patterns generated by different particle types, we propose the *normalised mean* (NM)—the first moment divided by n/m^2, the *coefficient of variation* (CV)—the standard deviation divided by the mean—and the *skewness* (S) [41] of the C-dataset as benchmarks. For these quantities, we have obtained an analytical RMT prediction in terms of mode and particle number, but as these expressions are rather longwinded, we present them in the appendix. Since the dataset is generated for a single U_{sub}, as explained before, we do expect slight deviations from these RMT results. In figure 3, we show the theoretical predictions (solid lines) for NM, CV and S, for a sampler in which six particles were injected, as a function of the number of modes. In order to quantify the deviations from the RMT prediction, we sampled, for various numbers of modes, 500 different U_{sub} matrices, calculated the respective C-dataset and its moments, and indicated the resulting average normalised mean, coefficient of variation, and

Figure 2. Normalised histograms of the correlator data, obtained by computing $C_{ij} = \langle \hat{n}_i \hat{n}_j \rangle - \langle \hat{n}_i \rangle \langle \hat{n}_j \rangle$ for all possible mode combinations, for a system with six particles and 100 modes. In the left panel, the histogram for bosonic correlators is compared to data obtained with fermions or distinguishable particles inserted instead of bosons. For the right panel, the histograms for bosons is compared to the result for simulated bosons, see main text. All histograms are obtained from one single circuit, using the same input modes, thus implying the same U_{sub}.

Figure 3. The theoretical RMT predictions (solid lines) for the normalised mean *NM* (top), the coefficient of variation *CV* (bottom left) and skewness *S* (bottom right) of the C-dataset are compared to the numerical *NM*, *CV* and *S* values of sampled C-datasets for six particles, as a function of the number of modes *m*. For mode numbers *m* = 20, 40 ,..., 300, we sampled 500 matrices U_{sub}, for each of which the normalised mean, the coefficient of variation and skewness of the C-dataset were calculated. For each mode number, the average normalised mean, coefficient of variation and skewness are indicated by a dot. Additionally, the standard deviations of the obtained *NM*, *CV* and *S* results are shown by error bars around these dots.

skewness by a point. The error bars indicate the standard deviation from this mean value and thus quantify the typical spread of possible outcomes. We show here that for these parameters, even NM is fit to effectively distinguish the interference of bosons, fermions and distinguishable particles, the curves for simulated and true bosons, however, collapse. *CV*, on the other hand, is a trustworthy quantity to distinguish truly bosonic interference from that of distinguishable particles and even of simulated bosons. The skewness *S* completes the certification that the observed many-particle interference pattern is actually generated by bosons. Similar plots where *n* is varied such that $m \sim \mathcal{O}(n)$ and $m \sim \mathcal{O}(n^2)$ are shown in appendix C.

Figure 4. For six particles, in 120 modes, the coefficient of variation and skewness were calculated for C-datasets of 500 sampled U_{sub} matrices, The points labeled 'bosons', 'distinguishable', 'fermions' and 'simulated bosons', each connect to one C-dataset of a sampled U_{sub}, the points' position on the plot indicates the calculated coefficient of variation CV and skewness S. The black and white triangles mark the RMT predictions for these quantities. Finally, the black and white circles indicate the mean value of each cluster of all the points generated for each particle type. These mean values coincide with the RMT predictions, thus the black and white triangles are hidden underneath the black and white circles.

We must emphasise that the curves of figure 3 represent the *typical* values for *NM*, *CV* and *S*, and that it might be possible to encounter a large deviation from such quantities. Moreover, one might dwell into a parameter regime where standard deviation bars overlap and hence it is unrealistic to certify the sampler with a single U_{sub} measurement. Luckily, a simple change of input modes implies a change in U_{sub} and hence it is feasible to generate several C-datasets from one circuit. Figure 4 shows the outcomes for various such U_{sub} matrices as points, where the x-coordinate indicates the coefficient of variation and the *y*-coordinate shows the skewness. The colour coded sets of points for different particle types are all separated from each other, showing clearly that the associated interference structures can be distinguished. As is indicated in figure 4, upon averaging over all the points in each cloud, one finds values (indicated by the black circles) which are very well estimated by the RMT predictions (indicated by the red points), thus providing a strong quantitative tool for such certification. This quality of the certification is further enhanced in large systems by noticing that the cloud is expected to shrink with the effective number of scattering events inside the array, namely the typical number of crossings between optical paths in figure 1 (state of the art experiments have $\simeq 20$).

In figure 3 we have indicated that bosonic, fermionic and distinguishable particles' interference patterns can be distinguished by studying the averages of the lowest-order statistical moments of the C-dataset. Furthermore, figure 4 shows that for a rather large number of modes and a large set of sampled U_{sub} matrices, we can classify all types of many-particle scattering dynamics. The method presented, however, does not require such an abundance of modes and samples since we can perform additional statistical analyses on the obtained cluster of data points. We emphasise this in figure 5, where data points for only 20 samples of U_{sub} matrices for $m = 20$ are shown. We focus specifically on bosons (indicated by blue points), where for each sample the average is calculated (red dot) and the red ellipses indicate two and four standard errors of the sample mean. The RMT prediction for bosons (a blue filled circle), with a slight bias, falls within the four standard errors, whereas the RMT prediction for simulated bosons (purple square) is well outside this region. Thus, we can successfully differentiate bona fide bosonic many-particle interference from the quantum dynamics of simulated bosons, using the RMT-based techniques described here.

6. Experimental overhead to measure correlators

In a realistic setup, repeated measurements are required to evaluate the correlators which constitute the C-dataset with sufficient accuracy. A fundamental question to be answered is therefore whether our protocol can be performed efficiently: Does the number of required experimental runs scale polynomially in the number of modes, m, and particles, n? The answer is positive.

Measurement of quantum observables is probabilistic by its very nature, hence one initially sets forth a target accuracy ϵ to which to determine the correlator. The RMT results of the previous section can be used to determine the required ϵ, since it needs to be significantly smaller than the distance between the points in figure 4. Focusing now on one correlator C_{ij}, the central limit theorem dictates that the number M of
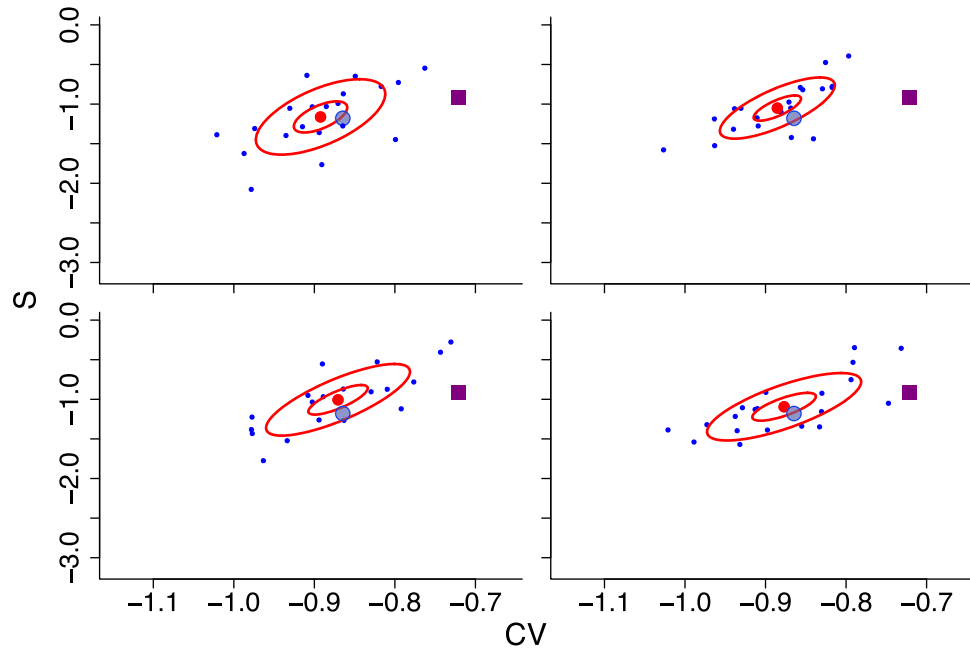
Figure 5. Four scatter plots, each containing 20 randomly sampled U_{sub} matrices for six particles in $m = 20$ output modes, from which the coefficient of variation CV and the skewness S of the bosonic C-dataset were calculated. The average of the cloud is indicated (red dot) as are the two- and four standard error regions (small and large ellipsoid, respectively). The RMT prediction for bosons is shown (large blue dot) to be contained within the ellipses for each of the four samples. The RMT prediction for simulated bosons (purple square) falls well outside the four standard errors.

measurements necessary to achieve this accuracy is given by $M = \mathrm{Var}(C_{ij})/\epsilon^2$, where

$$\mathrm{Var}(C_{ij}) = \langle (\hat{n}_i - \langle \hat{n}_i \rangle)^2 (\hat{n}_j - \langle \hat{n}_j \rangle)^2 \rangle - C_{ij}^2. \tag{4}$$

Rather than analysing the scaling behaviour thereof in full rigour, we provide a simple argument for polynomial scaling.

For increasing numbers of particles, the difficulty of BosonSampling is hidden in the increasing size of the matrix $U_{\vec{x},\vec{y}}$ of which to calculate the permanent. However, when calculating correlation functions, the problem simplifies considerably. The highest order term, in creation and annihilation operators, which enters $\mathrm{Var}(C_{ij})$ is given by $\langle (a_i^\dagger)^2 (a_j^\dagger)^2 (a_j)^2 (a_i)^2 \rangle$, and from [14] it is known that the calculation of this term involves the calculation of the permanents of 4×4 matrices, constructed with components of U. The dimensions of the required matrices are independent of the number of particles n and of the number of modes m. As shown in [14], the parameter n only governs the number of permanents of different 4×4 matrices that need to be summed to obtain the final variance. From the general expression in [14] one sees that this number of terms scales polynomially in n and is independent of the number of modes. Thus the variance and hence the final number of required measurements scale polynomially with the number of particles n.

The above argument holds for determining the correlator of a single choice of output modes i and j, but to obtain the full C-dataset one considers all combinations of output modes. Even if we assume the most inefficient scenario where one measures one correlator at a time, this still only implies a low order polynomial scaling in m.

7. Discussion

As evident from our above discussion, our certification method can be applied in a broad variety of experimental setups, for a wide range of particle numbers n and mode numbers m. Of course, smaller m lead to smaller C-datasets, making it more difficult to achieve statistical significance. In our numerical studies, however, we successfully distinguish particle type-specific interference structures for 20 modes or fewer. A nice advantage of our method as compared to [42] is that we can certainly treat regimes where $m < n^{5.1}$, we even can explore regimes in which $m \sim \mathcal{O}(n)$. However, as n and m grow, the curves for true and simulated bosons in figure 3 will approach one another to a distance of the order $\mathcal{O}(1/n)$ (for an extended discussion, see appendix C). As the limit $1/n \to 0$ is the semi-classical limit, where mean-field theory is exact, this is as such not surprising. The similarity of the two curves essentially implies that the statistics of the C-dataset is still strongly affected by the quantum statistics of the particles and thus by bosonic bunching. In this sense, we see a lot of potential in

elaborating the methods employed here, to further distill specific signatures of many-body interference, e.g. by filtering out the undesired bunching contributions to the statistics.

Let us emphasise yet another important advantage of our method, especially in the regime of large m and n: every experimental measurement outcome contributes to the C-dataset. Alternative methods which only seek to tell bosons apart from other particles, via book-keeping of bunching events [19], are far less favourable in this respect, since bunching events are rare events in the statistical sense. Therefore, not only does our method detect genuine many-particle inference as induced by the actual dynamics, much beyond mere quantum statistical effects such as bunching or anti-bunching—in the regime where BosonSampling can be considered really hard it is also expected to be more efficient, even when only employed to identify the particle species.

What concerns a possible confrontation of our theoretical analysis with real experiments, we have here focussed on the conceptual ideas behind statistical certification using correlation functions, thus describing an ideal scenario. For an actual experimental implementation, a wide range of imperfections kick in and eventually need to be accounted for: most prominent are losses [10], but to a lesser extent also decoherence, *i.e.* mode-mismatch leading to partial distinguishability and phase-fluctuations due to alignment instability [16, 43–45]. If losses occur before the particles enter the scattering system, this effect can be modelled by a statistical mixture of input states with different numbers of particles. Alternatively, particles lost inside the scattering system can be dealt with by increasing the number of output modes, while assuming that these additional modes are not observed [46]. In both cases, our RMT results have straightforward generalisations. Decoherence phenomena are more difficult to incorporate, since they require an expansion of our formal framework by an additional (environmental) degree of freedom [16]. This does fundamentally not affect the here proposed certification strategy, but imposes some non-trivial technical overhead. Both, losses and decoherence effects, ultimately require a careful technical treatment, which we set forth as an objective for future work. Likewise, generalisations of our RMT approach for the scenario of multiboson correlation sampling [47, 48] appear feasible.

In summary, we demonstrated that, by measuring the mode correlators $C_{i,j} = \langle \hat{n}_i \hat{n}_j \rangle - \langle \hat{n}_i \rangle \langle \hat{n}_j \rangle$ for all possible combinations of outgoing modes, one holds the key to certifying BosonSampling. The *mean*, the *variance*, and the *skewness* of such a C-dataset are sufficient to identify the multi-particle interference patterns resulting from transmission through a multimode random scatterer, as generated by bosons, fermions or distinguishable particles beyond reasonable doubt. By varying the chosen input channels, and thereby generating multiple such datasets one can efficiently distinguish true bosons from simulated bosons—which only exhibit bunching, yet no many-particle interference—through comparison of the first moments of their respective C-datasets. From a more general perspective, since our statistical quantifiers are constructed to distill bona fide many-particle quantum interference, beyond quantum statistical bunching or anti-bunching effects, these results improve our understanding of the fundamental differences between the quantum many-body dynamics of distinct particle types.

What is the consequence of our results for the extended Church–Turing thesis? Clearly, the capacity of any classical computer will be quickly exhausted when confronted with the task of evaluating a many-body wave function represented by a permanent, as soon as the number of bosonic constituents and modes is large enough. However, much as in the classical theory of gases or of chaotic (classical or quantum) systems, where, e.g., classical single particle trajectories are computationally unaccessible, too [49], we have shown that there are robust and selective statistical quantifiers which indeed can be handled, and which are accessible in state of the art experiments. In this sense, we suggest that a 'thermodynamic' or 'statistical' re-interpretation of the extended Church–Turing thesis will prevail.

Acknowledgments

Appendix A. Correlators

Initially, let us present a short and slightly more technical introduction to the central objects that constitute the C-dataset, the two-mode correlators. Given some pure quantum state $|\varphi\rangle$, these objects are defined as $C_{ij} = \langle \varphi | \hat{n}_i \hat{n}_j | \varphi \rangle - \langle \varphi | \hat{n}_i | \varphi \rangle \langle \varphi | \hat{n}_j | \varphi \rangle$, and the main goal of studying these object is to gain insight in the structure of states $|\varphi\rangle$, which results from the scattering of a many-particle Fock state in a system (one might think of a complicated network) which is described by a single-particle scattering matrix U (which we numerically generate following the algorithm described in [32]). As we initially start from a Fock state for which

modes $q_1, ..., q_n$ are populated by a single particle, we can describe the initial state $|\varphi_{\text{in}}\rangle$ in terms of creation operators $a_q^\dagger$ (for creation in the qth mode) that act on the vacuum state $|\Omega\rangle$, as

$$|\varphi_{\text{in}}\rangle = a_{q_1}^\dagger ... a_{q_n}^\dagger |\Omega\rangle. \tag{A.1}$$

Now, by traversing the system, the matrix U acts by connecting an input mode $a_q^\dagger$ to all possible output modes $a_i^\dagger$

$$a_q^\dagger \rightarrow \sum_{i=1}^m U_{q,i} a_i^\dagger \tag{A.2}$$

and thus we obtain that

$$|\varphi\rangle = \sum_{i_1,...,i_n=1}^m U_{q_1,i_1} a_{i_1}^\dagger ... U_{q_n,i_n} a_{i_n}^\dagger |\Omega\rangle. \tag{A.3}$$

For bosons (B) and fermions (F), an application of the (anti)commutation relations, $[a_i, a_j^\dagger]_\pm = \delta_{ij}\mathbf{1}$, and a long but straightforward computation leads to expressions for C_{ij}:

$$C_{ij}^B = -\sum_{k=1}^n U_{q_k,i} U_{q_k,j} U_{q_k,i}^* U_{q_k,j}^* + \sum_{k \neq l=1}^n U_{q_k,i} U_{q_l,j} U_{q_l,i}^* U_{q_k,j}^*, \tag{A.4}$$

$$C_{ij}^F = -\sum_{k=1}^n U_{q_k,i} U_{q_k,j} U_{q_k,i}^* U_{q_k,j}^* - \sum_{k \neq l=1}^n U_{q_k,i} U_{q_l,j} U_{q_l,i}^* U_{q_k,j}^*. \tag{A.5}$$

In the case of distinguishable particles, one can in principle treat the particles in an independent fashion, and thus a particle starting in input mode q will be found in output mode i with a probability $p_{q \rightarrow i} = |U_{i,q}|^2$. As the particles are distinguishable, these probabilities are not influenced by the presence of other particles, and via simple probability theory we now find that

$$C_{ij}^D = \sum_{k<l=1}^n (p_{q_k \rightarrow i} p_{q_l \rightarrow j} + p_{q_k \rightarrow j} p_{q_l \rightarrow i}) - \sum_{k,l=1}^n p_{q_k \rightarrow i} p_{q_l \rightarrow j}$$

$$= -\sum_{k=1}^n U_{q_k,i} U_{q_k,j} U_{q_k,i}^* U_{q_k,j}^*. \tag{A.6}$$

Finally, simulated bosons behave similarly to distinguishable particles, with the sole exception that the initial state is different and that (uniformly distributed) random phases are included over which one needs to average [20]. We essentially sample distinguishable particles, which are inserted in the form of a single-particle state that superposes all input modes with the same amplitude, but with random phases, implying a probability $p_i = \frac{1}{n}\left|\sum_{r=1}^n e^{i\theta_{q_r}} U_{q_r,i}\right|^2$ to find a particle in output mode i. Since each time we consider n such indistinguishable particles, a simple calculation yields

$$C_{ij}^S = \mathbb{E}(n(n-1)p_i p_j) - \mathbb{E}(np_i)\mathbb{E}(np_j)$$

$$= \frac{n-1}{n} \frac{1}{(2\pi)^n} \int_0^{2\pi} d\theta_{q_1} ... d\theta_{q_n} \left|\sum_{r=1}^n e^{i\theta_{q_r}} U_{q_r,i}\right|^2 \left|\sum_{r=1}^n e^{i\theta_{q_r}} U_{q_r,j}\right|^2$$

$$- \left(\frac{1}{(2\pi)^n} \int_0^{2\pi} d\theta_{q_1} ... d\theta_{q_n} |\sum_{r=1}^n e^{i\theta_{q_r}} U_{q_r,i}|^2\right)\left(\frac{1}{(2\pi)^n} \int_0^{2\pi} d\theta_{q_1} ... d\theta_{q_n} |\sum_{r=1}^n e^{i\theta_{q_r}} U_{q_r,j}|^2\right)$$

$$= \frac{(n-1)}{n} \sum_{r_1,r_2,s_1,s_2=1}^n U_{q_{r_1},i} U_{q_{r_2},j} U_{q_{s_1},i}^* U_{q_{s_2},j}^* \frac{1}{(2\pi)^n} \int_0^{2\pi} d\theta_{q_1} ... d\theta_{q_n} \exp[i(\theta_{q_{r_1}} + \theta_{q_{r_2}} - \theta_{q_{s_1}} - \theta_{q_{s_2}})] \tag{A.7}$$

$$- \sum_{r,s=1}^n U_{q_r,i} U_{q_s,i}^* \frac{1}{(2\pi)^n} \int_0^{2\pi} d\theta_{q_1} ... d\theta_{q_n} \exp(i\theta_{q_r} - i\theta_{q_s})$$

$$\times \sum_{r,s=1}^n U_{q_r,j} U_{q_s,j}^* \frac{1}{(2\pi)^n} \int_0^{2\pi} d\theta_{q_1} ... d\theta_{q_n} \exp(i\theta_{q_r} - i\theta_{q_s}).$$

Evaluating the integrals, we eventually obtain

$$C_{ij}^S = \left(1 - \frac{1}{n}\right) \sum_{r \neq s=1}^n U_{q_s,i} U_{q_r,j} U_{q_r,i}^* U_{q_s,j}^* - \frac{1}{n} \sum_{r,s=1}^n U_{q_r,i} U_{q_s,j} U_{q_r,i}^* U_{q_s,j}^*. \tag{A.8}$$

Appendix B. Random matrix theory

From expressions for the correlators of different particle types, we are able to construct the C-dataset by varying i and j (with $i < j$) to obtain all different mode combinations. In order to make theoretical predictions (or at least up to very good approximation), we use RMT methods. Rather than varying i and j, these methods keep the two output modes under consideration fixed and formally average over all possible matrices U in the unitary group with the *Haar measure* imposed on it. One might understand this as an analogue to the Bohigas-Giannoni-Schmit conjecture [50] for unitary matrices. The averaging contains one fundamental identity for an $N \times N$ random unitary matrix U:

$$\mathbb{E}_U \left(U_{a_1, b_1} \dots U_{a_n, b_n} U^*_{\alpha_1, \beta_1} \dots U^*_{\alpha_n, \beta_n} \right) = \sum_{\sigma, \pi \in S_n} V_N(\sigma^{-1}\pi) \prod_{k=1}^{n} \delta(a_k - \alpha_{\sigma(k)}) \delta(b_k - \beta_{\pi(k)}), \tag{B.1}$$

where $\mathbb{E}_U(.)$ denotes the average over the unitary group and V are class coefficients also known as Weingarten functions, which are determined recursively. The details of this method can be found in [34–37]. With this formula, we efficiently average long products of coefficients of unitary matrices to compute $\mathbb{E}_U(C_{ij})$, $\mathbb{E}_U(C_{ij}^2)$ and $\mathbb{E}_U(C_{ij}^3)$ for each particle type. Combining these quantities we can find the coefficient of variation CV and the skewness S for each particle type. We find, with n particles in m modes, for bosons:

$$\mathbb{E}_U(C_B) = -\frac{n(m + n - 2)}{m(m^2 - 1)}, \tag{B.2}$$

$$\mathbb{E}_U(C_B^2) = \frac{2n(m^2 n + m^2 + 9mn - 11m + n^3 - 2n^2 + 5n - 4)}{m^2(m + 2)(m + 3)(m^2 - 1)}, \tag{B.3}$$

$$\mathbb{E}_U(C_B^3) = -2n\left(\frac{m^3 n^2 + 15 m^3 n + 2m^3 + 3 m^2 n^3 + 6 m^2 n^2 + 213 m^2 n - 222 m^2 - 3 mn^4}{m^2(m + 1)(m + 2)(m + 3)(m + 4)(m + 5)(m^2 - 1)} \right.$$
$$\left. + \frac{45mn^3 + 32mn^2 + 372mn - 464m + 3n^5 - 6n^4 + 45n^3 + 78n^2 + 168n - 288}{m^2(m + 1)(m + 2)(m + 3)(m + 4)(m + 5)(m^2 - 1)} \right), \tag{B.4}$$

for fermions

$$\mathbb{E}_U(C_F) = \frac{n(n - m)}{m(m^2 - 1)}, \tag{B.5}$$

$$\mathbb{E}_U(C_F^2) = \frac{2n(n + 1)(m - n)(m - n + 1)}{m^2(m + 2)(m + 3)(m^2 - 1)}, \tag{B.6}$$

$$\mathbb{E}_U(C_F^3) = -\frac{6n(n + 1)(n + 2)(m - n)(m - n + 1)(m - n + 2)}{m^2(m + 1)(m + 2)(m + 3)(m + 4)(m + 5)(m^2 - 1)}, \tag{B.7}$$

for distinguishable particles

$$\mathbb{E}_U(C_D) = -\frac{n}{m(m + 1)}, \tag{B.8}$$

$$\mathbb{E}_U(C_D^2) = \frac{n(m^2 n + 3m^2 + mn - 5m + 2n - 2)}{m^2(m + 2)(m + 3)(m^2 - 1)}, \tag{B.9}$$

$$\mathbb{E}_U(C_D^3) = -\frac{n(m^2 n^2 + 9m^2 n + 26m^2 + 5mn^2 + 21mn - 62m + 12n^2 + 60n - 72)}{m^2(m + 2)(m + 3)(m + 4)(m + 5)(m^2 - 1)}, \tag{B.10}$$

and finally for the simulated bosons

$$\mathbb{E}_U(C_S) = -\frac{n(m + n - 2)}{m(m^2 - 1)}, \tag{B.11}$$

$$\mathbb{E}_U(C_S^2) = \frac{4mn - m - 14n^2 + 8n - 2}{m^2(m + 2)(m + 3)(m^2 - 1)n}$$
$$+ \frac{2m^2 n^3 - m^2 n^2 + 4m^2 n - m^2 + 18mn^3 - 25mn^2 + 2n^5 - 4n^4 + 10n^3}{m^2(m + 2)(m + 3)(m^2 - 1)n}, \tag{B.12}$$

$$\mathbb{E}_U(C_S{}^3) = \left(\frac{-2m^3n^5 - 21m^3n^4 + 30m^3n^3 - 41m^3n^2 - 10m^3n + 8m^3 - 6m^2n^6 - 3m^2n^5}{(m-1)m^2(m+1)^2(m+2)(m+3)(m+4)(m+5)n^2} \right.$$

$$+ \frac{-285m^2n^4 + 261m^2n^3 + 75m^2n^2 - 66m^2n + 24m^2 + 6mn^7 - 90mn^6 - 55mn^5}{(m-1)m^2(m+1)^2(m+2)(m+3)(m+4)(m+5)n^2}$$

$$+ \frac{-360mn^4 + 591mn^3 + 8mn^2 - 128mn + 64m}{(m-1)m^2(m+1)^2(m+2)(m+3)(m+4)(m+5)n^2}$$

$$+ \left. \frac{-6n^8 + 12n^7 - 90n^6 - 120n^5 - 24n^4 + 396n^3 - 168n^2 - 48(n-1)}{(m-1)m^2(m+1)^2(m+2)(m+3)(m+4)(m+5)n^2} \right). \tag{B.13}$$

Although these formulas do not appear remarkably elegant due the lack of any form of assumption on m and n (apart from $m > n$), they are necessary to obtain sufficiently accurate results. Once these moments are defined, we can use them to find *NM*, *CV* and *S* by the following definitions:

$$NM = \frac{\mathbb{E}_U(C)m^2}{n} \tag{B.14}$$

$$CV = \frac{\sqrt{\mathbb{E}_U(C^2) - \mathbb{E}_U(C)^2}}{\mathbb{E}_U(C)}, \tag{B.15}$$

$$S = \frac{\mathbb{E}_U(C^3) - 3\mathbb{E}_U(C)\mathbb{E}_U(C^2) + 2\mathbb{E}_U(C)^3}{(\mathbb{E}_U(C^2) - \mathbb{E}_U(C)^2)^{3/2}}. \tag{B.16}$$

With these results, one can now calculate the expected coefficient of variation and the expected skewness for each of the samplers we described, with an arbitrary number of modes and particles.

Appendix C. RMT results and the number of particles

In this appendix, we assess the lowest-order scaling properties of the RMT results of appendix B as a function of the particle number. We consider large sets of particles which are injected into $m \sim \mathcal{O}(n)$ or $m \sim \mathcal{O}(n^2)$ modes.

First, let us set $m = \lambda n$, with $\lambda \in \mathbb{N}$ and $\lambda > 1$, and consider the limit $n \to \infty$. In this regime, we can evaluate the leading order scaling behaviour for the normalised mean *MN*, the coefficient of variation *CV* and the skewness *S* for different particle types. We now find that for bosons

$$NM \approx \left(-\frac{1}{\lambda} - 1 \right) + \frac{2}{\lambda n}, \tag{C.1}$$

$$CV \approx \left(\frac{2}{\lambda + 1} - 1 \right) + \frac{5 - \lambda^4 + 9\lambda^3 + 2\lambda^2 - 3\lambda}{(\lambda - 1)\lambda(\lambda + 1)^2 n}, \tag{C.2}$$

$$S \approx 2 - \frac{6(5\lambda^3 + 2\lambda^2 - 10\lambda + 5)}{((\lambda - 1)^2 \lambda)n}, \tag{C.3}$$

for fermions,

$$NM \approx \left(\frac{1}{\lambda} - 1 \right) + \frac{1 - \lambda}{\lambda^3 n^2}, \tag{C.4}$$

$$CV \approx -\frac{\lambda}{(\lambda - 1)n} + \frac{5}{\lambda n} - 1, \tag{C.5}$$

$$S \approx \frac{\frac{30}{\lambda} - \frac{6\lambda}{\lambda - 1}}{n} - 2, \tag{C.6}$$

for distinguishable particles,
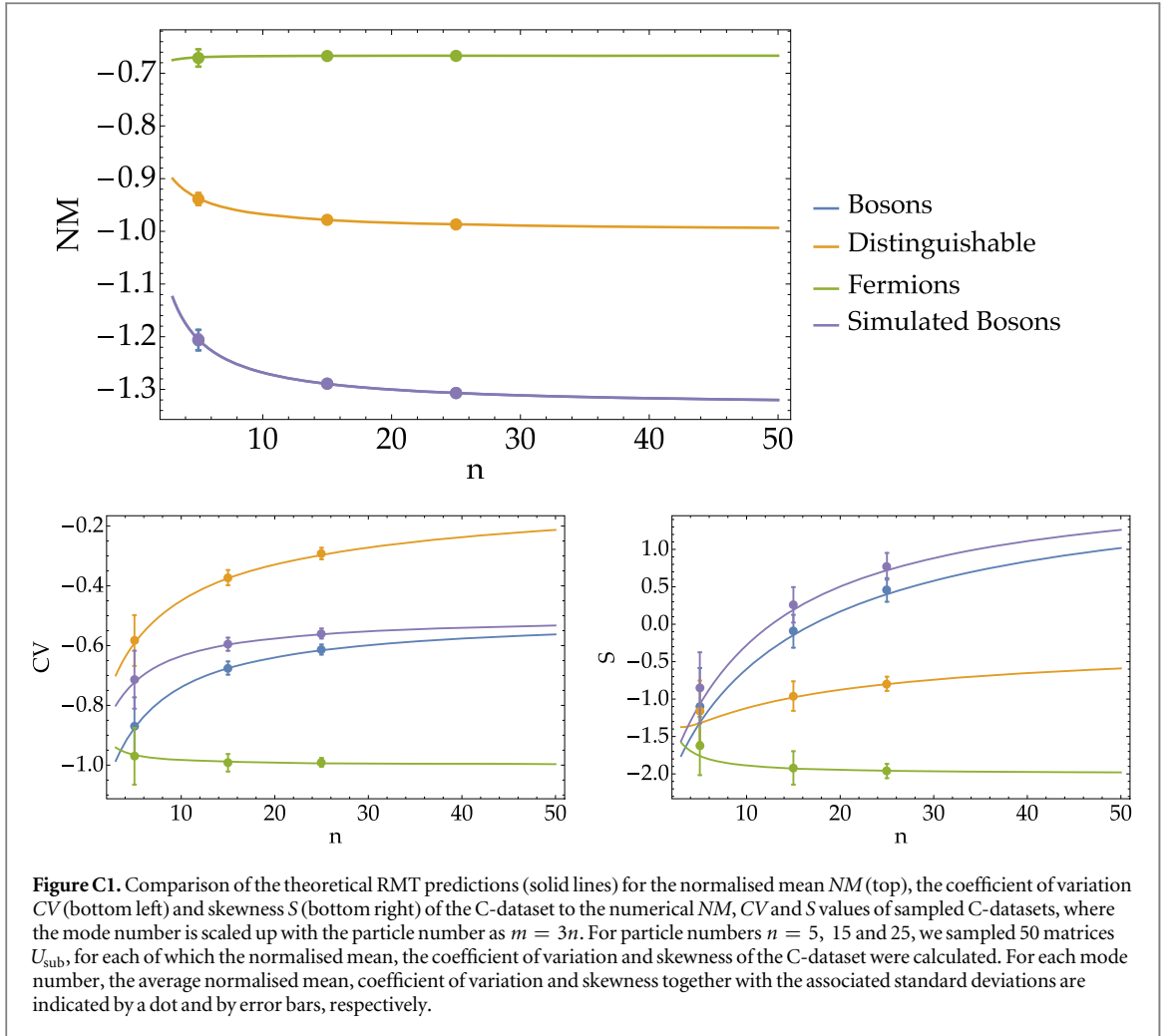
$$NM \approx \frac{1}{\lambda n} - 1, \tag{C.7}$$

$$CV \approx -\left(\sqrt{3 - \frac{2}{\lambda}} \sqrt{\frac{1}{n}} \right), \tag{C.8}$$

$$S \approx -\frac{2(\lambda(13\lambda - 18) + 7)}{\sqrt{\lambda}(3\lambda - 2)^{3/2}} \sqrt{\frac{1}{n}}, \tag{C.9}$$

and, finally, for simulated bosons

$$NM \approx \left(-\frac{1}{\lambda} - 1 \right) + \frac{2}{\lambda n}, \tag{C.10}$$

Figure C1. Comparison of the theoretical RMT predictions (solid lines) for the normalised mean *NM* (top), the coefficient of variation *CV* (bottom left) and skewness *S* (bottom right) of the C-dataset to the numerical *NM*, *CV* and *S* values of sampled C-datasets, where the mode number is scaled up with the particle number as $m = 3n$. For particle numbers $n = 5$, 15 and 25, we sampled 50 matrices U_{sub}, for each of which the normalised mean, the coefficient of variation and skewness of the C-dataset were calculated. For each mode number, the average normalised mean, coefficient of variation and skewness together with the associated standard deviations are indicated by a dot and by error bars, respectively.

$$CV \approx \left(\frac{2}{\lambda + 1} - 1 \right) + \frac{\lambda^4 - 15\lambda^3 - 4\lambda^2 + 6\lambda + 10}{2(\lambda - 1)\lambda(\lambda + 1)^2 n}, \tag{C.11}$$

$$S \approx 2 - \frac{3(\lambda(\lambda + 2)(7\lambda - 10) + 10)}{((\lambda - 1)^2 \lambda) n}. \tag{C.12}$$

Observe that even the leading *n*-dependence is different when comparing particle types. This implies that, while the RMT results for bosons and simulated bosons converge to the same point for *NM*, *CV*, and *S*, they do so in a quantitatively distinct manner.

We now repeat the computation for a different type of scaling, where we set $m = \lambda n^2$. Again, when we evaluate the scaling behaviour as $n \to \infty$, we find for bosons:

$$NM \approx -\frac{1}{\lambda n} - 1, \tag{C.13}$$

$$CV \approx \frac{2 - \lambda}{\lambda n} - 1, \tag{C.14}$$

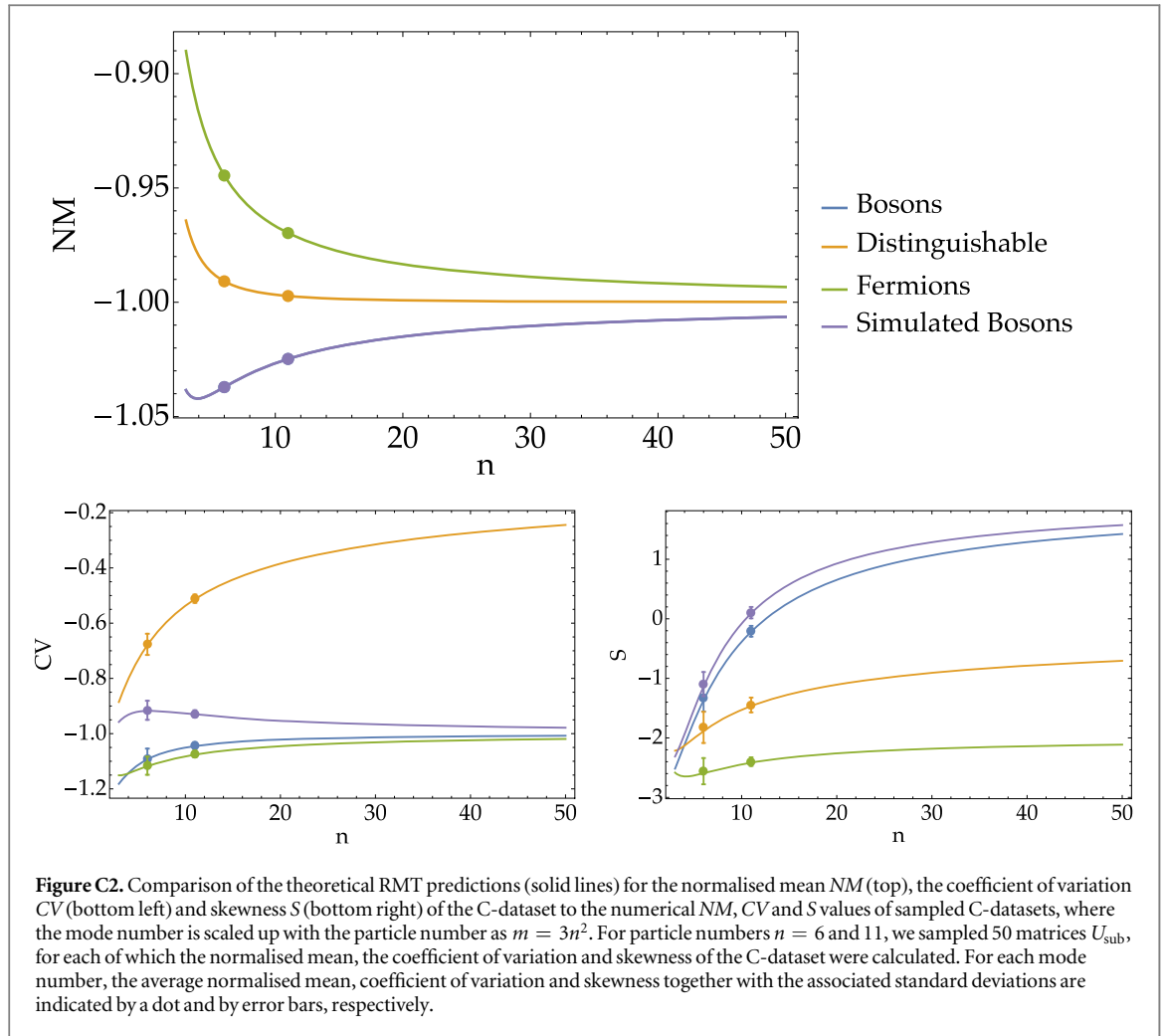$$S \approx 2 - \frac{30}{n}, \tag{C.15}$$

for fermions:

$$NM \approx \frac{1}{\lambda n} - 1, \tag{C.16}$$

$$CV \approx -\frac{1}{n} - 1, \tag{C.17}$$

$$S \approx -\frac{6}{n} - 2, \tag{C.18}$$

Figure C2. Comparison of the theoretical RMT predictions (solid lines) for the normalised mean *NM* (top), the coefficient of variation *CV* (bottom left) and skewness *S* (bottom right) of the C-dataset to the numerical *NM*, *CV* and *S* values of sampled C-datasets, where the mode number is scaled up with the particle number as $m = 3n^2$. For particle numbers $n = 6$ and 11, we sampled 50 matrices U_{sub}, for each of which the normalised mean, the coefficient of variation and skewness of the C-dataset were calculated. For each mode number, the average normalised mean, coefficient of variation and skewness together with the associated standard deviations are indicated by a dot and by error bars, respectively.

for distinguishable particles:

$$NM \approx \frac{1}{\lambda n^2} + 1, \tag{C.19}$$

$$CV \approx -\left(\sqrt{3}\sqrt{\frac{1}{n}} \right), \tag{C.20}$$

$$S \approx -\frac{26}{3\sqrt{3}}\sqrt{\frac{1}{n}}, \tag{C.21}$$

and for simulated bosons:

$$NM \approx -\frac{1}{\lambda n} - 1, \tag{C.22}$$

$$CV \approx \frac{\lambda + 4}{2\lambda n} - 1, \tag{C.23}$$

$$S \approx 2 - \frac{21}{n}. \tag{C.24}$$

The first notable observation is that, again, the bosons and simulated bosons converge to the same point for *NM*, *CV*, and *S*, but approach this limit in a different fashion. Moreover, it is an interesting observation that the leading order term obtained for *S* are independent of λ.

The physical interpretation of these results remains unclear to us at the present stage. One can understand the similarity between bosons and simulated bosons in the thermodynamic limit by realising that this is also the limit where mean-field theory is expected to be exact. One may even verify this through (A.4) and (A.8): the terms which distinguish the two expressions vanish in the limit $n \to \infty$.

Finally, to visualise the behaviour of the RMT results with $m \sim \mathcal{O}(n)$ or $m \sim \mathcal{O}(n^2)$, we present additional plots in figures C1 and C2 , which compare to figure 3. With the arbitrary choice $\lambda = 3$, straightforward numerical evaluation rapidly becomes demanding, therefore we limit ourselves to just a few data points.

References

[1] Nielsen M A and Chuang I L 2010 *Quantum Computation and Quantum Information* (Cambridge: Cambridge University Press)

[2] Aaronson S and Arkhipov A 2013 The computational complexity of linear optics *Theory of Computing* **9** 143–252

[3] Lanting T *et al* 2014 Entanglement in a quantum annealing processor *Phys. Rev.* X **4** 021041

[4] Deutsch D 1985 Quantum theory, the church-turing principle and the universal quantum computer *Proc. R. Soc. Lond.* A **400** 97–117

[5] Moore C and Mertens S 2011 *The Nature of Computation* (Oxford University Press)

[6] Broome M A, Fedrizzi A, Rahimi-Keshari S, Dove J, Aaronson S, Ralph T C and White A G 2013 Photonic boson sampling in a tunable circuit *Science* **339** 794–8

[7] Crespi A, Osellame R, Ramponi R, Brod D J, Galvão E F, Spagnolo N, Vitelli C, Maiorino E, Mataloni P and Sciarrino F 2013 Integrated multimode interferometers with arbitrary designs for photonic boson sampling *Nat. Photon.* **7** 545–9

[8] Ralph T C 2013 Quantum computation: Boson sampling on a chip *Nat. Photon.* **7** 514–5

[9] Spagnolo N *et al* 2014 Experimental validation of photonic boson sampling *Nat. Photon.* **8** 615–20

[10] Spring J B *et al* 2013 Boson sampling on a photonic chip *Science* **339** 798–801

[11] Tillmann M, Dakić B, Heilmann R, Nolte S, Szameit A and Walther P 2013 Experimental boson sampling *Nat. Photon.* **7** 540–4

[12] Hong C K, Ou Z Y and Mandel L 1987 Measurement of subpicosecond time intervals between two photons by interference *Phys. Rev. Lett.* **59** 2044–6

[13] Tichy M C, Tiersch M, de Melo F, Mintert F and Buchleitner A 2010 Zero-transmission law for multiport beam splitters *Phys. Rev. Lett.* **104** 220405

[14] Mayer K, Tichy M C, Mintert F, Konrad T and Buchleitner A 2011 Counting statistics of many-particle quantum walks *Phys. Rev.* A **83** 062307

[15] Tichy M C, Tiersch M, Mintert F and Buchleitner A 2012 Many-particle interference beyond many-boson and many-fermion statistics *New J. Phys.* **14** 093015

[16] Ra Y-S, Tichy M C, Lim H-T, Kwon O, Mintert F, Buchleitner A and Kim Y-H 2013 Nonmonotonic quantum-to-classical transition in multiparticle interference *Proc. Natl Acad. Sci. USA* **110** 1227–31

[17] Minc H 1978 *Permanents, Encyclopedia of Mathematics and its Applications.* vol 16 (Boston, MA: Addison-Wesley)

[18] Troyansky L and Tishby N 1996 On the quantum evaluation of the determinant and the permanent of a matrix *Proc. Fourth Workshop on Physics and Computation* (*PhysComp 96*) (*Boston, 22–24 November 1996*) **96**

[19] Carolan J *et al* 2014 On the experimental verification of quantum complexity in linear optics *Nat. Photon.* **8** 621–6

[20] Tichy M C, Mayer K, Buchleitner A and Mølmer K 2014 Stringent and efficient assessment of boson-sampling devices *Phys. Rev. Lett.* **113** 020502

[21] Carolan J *et al* 2015 universal linear optics *Science* **349** 711–6

[22] Crespi A, Osellame R, Ramponi R, Bentivegna M, Flamini F, Spagnolo N, Viggianiello N, Innocenti L, Mataloni P and Sciarrino F 2015 Quantum suppression law in a 3-D photonic chip implementing the fast fourier transform (arXiv: 1508.00782)

[23] Akkermans E and Montambaux G 2007 *Mesoscopic Physics of Electrons and Photons* (Cambridge: Cambridge University Press)

[24] Bloch F 1929 Über die quantenmechanik der elektronen in kristallgittern *Z. Physik.* **52** 555–600

[25] Aharonov Y, Davidovich L and Zagury N 1993 Quantum random walks *Phys. Rev.* A **48** 1687–90

[26] Scholak T, Wellens T and Buchleitner A 2014 Spectral backbone of excitation transport in ultracold rydberg gases *Phys. Rev.* A **90** 063415

[27] Engl T, Dujardin J, Argüelles A, Schlagheck P, Richter K and Urbina J D 2014 Coherent backscattering in fock space: a signature of quantum many-body interference in interacting bosonic systems *Phys. Rev. Lett.* **112** 140403

[28] Jeltes T *et al* 2007 Comparison of the hanbury brown-twiss effect for bosons and fermions *Nature* **445** 402–5 01

[29] Rom T, Best Th, van Oosten D, Schneider U, Folling S, Paredes B and Bloch I 2006 Free fermion antibunching in a degenerate atomic fermi gas released from an optical lattice *Nature* **444** 733–6 12

[30] Chuchem M, Smith-Mannschott K, Hiller M, Kottos T, Vardi A and Cohen D 2010 Quantum dynamics in the bosonic josephson junction *Phys. Rev.* A **82** 053617

[31] Mehta M L 2004 *Random Matrices* (Amsterdam: Elsevier)

[32] Mezzadri F 2007 How to generate random matrices from the classical compact groups *Notices AMS* **54** 592–604 (www.ams.org/notices/200705/fea-mezzadri-web.pdf)

[33] Zyczkowski K and Kus M 1994 Random unitary matrices *J. Phys. A: Math. Gen.* **27** 4235

[34] Samuel S 1980 U(N) integrals, 1/N, and the De Wit-'t Hooft anomalies *J. Math. Phys.* **21** 2695–703

[35] Mello P A 1990 Averages on the unitary group and applications to the problem of disordered conductors *J. Phys.* A **23** 4061–80

[36] Brouwer P W and Beenakker C W J 1996 Diagrammatic method of integration over the unitary group, with applications to quantum transport in mesoscopic systems *J. Math. Phys.* **37** 4904–34

[37] Berkolaiko G and Kuipers J 2013 Combinatorial theory of the semiclassical evaluation of transport moments. I. equivalence with the random matrix approach *J. Math. Phys.* **54** 112103

[38] Urbina J-D, Kuipers J, Hummel Q and Richter K 2014 Multiparticle correlations in complex scattering and the mesoscopic boson sampling problem (arXiv:1409.1558)

[39] Bratteli O and Robinson D W 1997 *Operator Algebras and Quantum Statistical Mechanics: Equilibrium States. Models in Quantum Statistical Mechanics* (Berlin: Springer)

[40] Peruzzo A *et al* 2010 Quantum walks of correlated photons *Science* **329** 1500–3

[41] MacGillivray H L 1986 Skewness and asymmetry: Measures and orderings *Ann. Stat.* **14** 994–1011

[42] Aaronson S and Arkhipov A 2014 Bosonsampling is far from uniform *Quantum Info. Comput.* **14** 1383–423 (www.rintonpress.com/journals/qiconline.htm)

[43] Shchesnovich V S 2014 Sufficient condition for the mode mismatch of single photons for scalability of the boson-sampling computer *Phys. Rev.* A **89** 022333

[44] Tichy M C 2015 Sampling of partially distinguishable bosons and the relation to the multidimensional permanent *Phys. Rev.* A **91** 022316

[45] Shchesnovich V S 2015 Partial indistinguishability theory for multiphoton experiments in multiport devices *Phys. Rev.* A **91** 013844

[46] Mayer K 2012 Many-particle quantum walks *Master's Thesis* Albert-Ludwigs Universität Freiburg, Freiburg

[47] Tamma V and Laibacher S 2015 Multiboson correlation interferometry with arbitrary single-photon pure states *Phys. Rev. Lett.* **114** 243601

[48] Laibacher S and Tamma V 2015 From the physics to the computational complexity of multiboson correlation interference *Phys. Rev. Lett.* **115** 243605

[49] Schuster H G 1989 *Deterministic Chaos: An Introduction* (Weinheim: VCH) 1st reprint of 2nd ed

[50] Bohigas O, Giannoni M J and Schmit C 1984 Characterization of chaotic quantum spectra and universality of level fluctuation laws *Phys. Rev. Lett.* **52** 1–4