

Marginal probabilities in boson samplers with arbitrary input states

Jelmer Renema

January 1, 2021

Abstract

With the recent claim of a quantum advantage demonstration in photonics by Zhong *et al*, the question of the computation of lower-order approximations of boson sampling with arbitrary quantum states at arbitrary distinguishability has come to the fore. In this work, we present results in this direction, building on the results of Clifford and Clifford. In particular, we show:

- 1) How to compute marginal detection probabilities (i.e. probabilities of the detection of some but not all photons) for arbitrary quantum states.
- 2) Using the first result, how to generalize the sampling algorithm of Clifford and Clifford to arbitrary photon distinguishabilities and arbitrary input quantum states.
- 3) How to incorporate truncations of the quantum interference into a sampling algorithm.
- 4) A remark considering maximum likelihood verification of the recent photonic quantum advantage experiment.

1 Introduction

In a recent landmark result, Zhong *et al* [1] claimed the first demonstration of a quantum advantage in photonics. This result immediately led to discussion [2] focussing on the extent to which the samples which this experiment produced truly manifest large-scale quantum interference, and whether the sampling carried out by the photonic device could be efficiently classically simulated.

This discussion focussed on *marginal probabilities*, i.e. the probability to observe some $k \leq n$ photons in a particular output configuration ϕ_k , summed over all possible outcomes for the remaining photons. Informally, these probability distributions answer the question ‘what is the probability to see a photon *here*, *here* and *here*, irrespective of where the other photons go?’

The reason for this interest is that unlike in random circuit sampling [3], marginal distributions carry some information about the overall large-scale interference process, and can therefore be used as a diagnostic tool, to build up evidence of the success of a quantum advantage experiment. Specifically, k -th

order marginals (i.e. marginal distributions involving k particles) carry information about the interference processes of k and fewer photons.

Applying these ideas, the authors of [1] showed that their experiment has first and second order marginal distributions roughly consistent with the theoretical predictions. Furthermore, the authors generated samples which have the first order marginals correct (i.e. in agreement with the noiseless case) but not the higher order marginals. The authors of [1] introduced a statistical test (which has since received the name ‘CHOG’) which essentially tests which of two sets of samples is most likely to come from the ideal distribution. They showed that the first-order marginal samples lose this test against the experimental samples. Kalai suggested [2], based on an earlier paper of his and Kindler, [4] that a series of samples which in addition to the first-order marginals have up to k order marginals correct would win against the experimental samples. There was a subsequent counterclaim [5] by the authors of [1] that for the case of $k = 2$, these samples also fail to beat the experimental samples at the CHOG test. The issue of what would happen when comparing higher marginals to the experiment is currently stuck at the question of how to generate samples which have the correct third and higher order marginal distributions.

In this work, we resolve this issue by showing how to generate samples which have the correct k -th order marginals for any k , for an arbitrary quantum state, with arbitrary properties of the photons in that state. This discussion builds on a rich literature of how to approximate boson sampling with imperfections [4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23].

2 The setting

Consider a linear optical system of N modes described by a transmission matrix U . At the input of this optical system, an arbitrary quantum state of light $|\Psi\rangle$ is impinging¹. Furthermore, we assume that there is photodetection in the Fock basis, i.e. at the detector the set of measurement operators $\Pi_i = |i\rangle\langle i|$ is applied to each output mode of the interferometer.

We wish to accomplish the task of computing marginal probabilities for arbitrary input quantum states $|\Psi\rangle$ and arbitrary mutual distinguishabilities (wavefunction overlaps) of the input photons, and ultimately generating samples which follow these distributions. To simplify notation, we will assume that the photons in this quantum state have equal mutual wave function overlaps i.e. $x_{ij} = \langle\psi_i|\psi_j\rangle = x + (1-x)\delta_{ij}$, but our results can be trivially generalized to the case of unequal wavefunction overlaps².

As a special case, for $|\Psi\rangle = |1\rangle^{\otimes n}|0\rangle^{\otimes N-n}$, this is the boson sampling problem as proposed by Aaronson and Arkhipov [24]. In that case, the probability of observing a given outcome ϕ is given by $P(\phi) = |\text{Perm}(M_{\xi,\phi})|^2$, where $M_{\xi,\phi}$ is

¹ $|\Psi\rangle$ may be entangled, and need not be a state of definite overall photon number.

²In this work, we will consider mutual distinguishability as the main source of imperfections. The relation between this noise source and noise on the interferometer as treated in [4] is discussed by Shchesnovich in [20]

the submatrix of U connecting the first n input modes of U to the set of output modes ϕ of interest, and Perm is the permanent function³. $\xi = (1, 2, \dots, n)$ is the vector of input modes containing a photon, which is introduced for consistency with later notation. For this case, Clifford & Clifford [25] solved the question of how to compute marginal probabilities, which they used to construct a sampling algorithm which generates samples at close to the computational cost of evaluating a single permanent. We will import their results to the broader setting of arbitrary distinguishabilities and input quantum states.

3 Preliminaries

We follow the approach of [15] and project the input state $|\Psi\rangle$ onto the subspace Ξ of the Fock basis at the input of the interferometer with exactly n incident photons, i.e. for all $|\xi\rangle \in \Xi, \langle \xi | \hat{N} | \xi \rangle = n$, where $\hat{N} = \sum \hat{n}$ is the multimode photon number operator⁴.

In the rest of this work, we will assume the set $\{|\xi\rangle\}$ forms an orthonormal basis of Ξ , and that they are chosen such that each $|\xi\rangle$ is a product state over the input modes, i.e. $|\xi\rangle = \prod_{i=1}^N |m_i\rangle$, where $|m_i\rangle$ is a Fock state in mode i with m photons, and the condition that $|\xi\rangle$ lies in Ξ enforces $\sum_{i=0}^N m_i = n$. Note that this can always be done, since different product states of this form are automatically orthonormal, and a simple counting argument guarantees they form a basis. Furthermore, we will use the notation that ξ is the vector of Fock-state inputs corresponding to $|\xi\rangle$. We will use Ξ both for the subspace and for the set of all valid assignments of ξ corresponding to that subspace.

Note that since the photon number is conserved by the interferometer, conditional on observing n photons and barring photon loss⁵, the dynamics of an n -photon detection event are entirely governed by interference in the subspace Ξ . The probability to observe some outcome ϕ is then given by [15]:

$$P(\phi) = \sum_{\xi \in \Xi} \sum_{\chi \in \Xi} c_\xi c_\chi^\dagger \text{Perm}(M_{\xi, \phi}) \text{Perm}(M_{\chi, \phi})^\dagger, \quad (1)$$

where $c_\xi = \langle \Psi | \xi \rangle / \sqrt{\mu(\xi)\mu(\phi)}$, μ is the multiplicity function, Perm is the permanent function, and $M_{\xi, \phi}$ is the submatrix that connects the modes in ξ which contain a nonzero number of photons with the output modes of interest. In what follows, for simplicity and following the constraints of the experiment [1], we will assume $\mu(\phi) = 1$, i.e. all n photons emerge from distinct modes.

As an example of how to construct the set of ξ , consider Gaussian boson sampling (GBS). In that case, the input state is given by a product state of pairs of modes over a given subset modes, with the remaining ones empty, i.e.

³We may renumber the input without loss of generality so that the first n modes contain a photon

⁴See [15] for a more formal derivation

⁵In the presence of photon loss, the probability is given by a classical sum over various Ξ_j for $j > n$, so *mutatis mutandis* the discussion below holds in the presence of photon loss as well.

$|\Psi\rangle = |\psi\rangle^{\otimes n/2}|0\rangle^{N-n}$, with $|\psi\rangle = \frac{1}{\cosh(r)} \sum_{n=0}^{\infty} (\exp(-i\phi) \tanh(r))^n |n, n\rangle$ the *two-mode squeezed state*, where r and ϕ form the *squeezing parameter* via $\zeta = re^{i\phi}$. As an example, for a three-squeezer boson sampling experiment detecting four photons, the complete set of ξ would be $(1, 2, 1, 2)$, $(1, 2, 3, 4)$, $(1, 2, 5, 6)$, $(3, 4, 3, 4)$, $(3, 4, 5, 6)$, $(5, 6, 5, 6)$, i.e. all 6 ways of choosing 2 pairs of 3 possible sources with repetition (i.e. $\binom{3+2-1}{2}$ ways). Interference between these different emission processes has been observed in experiments. In the case of GBS, for equal squeezing parameters on all sources, these ξ are all equiprobable, which follows from the exponential distribution of pairs.

3.1 Results from Clifford & Clifford

We introduce a few results and techniques from [25].

First, to compute the marginal output distributions, we switch to the expanded sample space notation. In this formalism, we remove the notational restriction that the vector specifying the location of the output photons at ϕ needs to be ordered. Normally, if we give a list of the output modes of a given output configuration ϕ of a boson sampling experiment, we require that that list is ordered, i.e. that $\phi(i) \leq \phi(i+1)$. Removing this restriction lifts the spurious correlation between elements of ϕ arising from this ordering. In the expanded sample space formalism, we have that [25]:

$$P(\phi_u) = \frac{\mu(\phi)}{n!} P(\phi), \quad (2)$$

where ϕ_u stands for an unordered version of ϕ ,

Next, we introduce two key expressions from [25] (rephrased in our notation).

First:

$$\text{Perm}(M_{\xi, \phi}) \text{Perm}(M_{\chi, \phi}^\dagger) = \sum_{\sigma \in S} \text{Perm}(M_{\xi, \phi} \circ M_{\sigma(\chi), \phi}^\dagger), \quad (3)$$

Where the sum runs over all permutations and \circ is an elementwise product. Secondly:

$$\sum_{\phi \in \Phi} \text{Perm}(M_{a, \phi} \circ M_{b, \phi}^\dagger) = m! \delta(a, b), \quad (4)$$

where m is the size of M , and Φ is the set of all valid assignment of the output modes (i.e. all possible outcomes), and δ is an elementwise Kronecker delta function. This result follows from the orthonormality of the rows and columns of U . Note that combining equations 1-4 guarantees that $P(\phi)$ is normalized for any choice of $|\Psi\rangle$.

4 Results

We obtain marginal distributions by applying equations 2-4 to 1. First, we apply eqn 2 and 3.

$$P(\phi_u) = \frac{1}{n!} \sum_{\xi \in \Xi} \sum_{\chi \in \Xi} c_\xi c_\chi^\dagger \sum_{\sigma} \text{Perm}(M_{\xi, \phi} \circ M_{\sigma(\chi), \phi}), \quad (5)$$

then, we apply Laplace expansion along the output modes, to split up the permanent into those photons over which we will marginalize and those which will remain. Note that without loss of generality, we can pick the unmarginalized photons to be the first k ones. This results in:

$$P(\phi_u) = \frac{1}{n!} \sum_{\xi \in \Xi} \sum_{\chi \in \Xi} c_\xi c_\chi^\dagger \sum_{\sigma} \sum_{\rho} \text{Perm}(M_{\rho(\xi), \phi_{1:k}} \circ M_{\rho(\sigma(\chi)), \phi_{1:k}}) \text{Perm}(M_{\rho(\xi), \phi_{k+1:n}} \circ M_{\rho(\sigma(\chi)), \phi_{k+1:n}}), \quad (6)$$

where ρ is the set of all k combinations out of n , and $\bar{\rho}$ is the complement of ρ .

Next, we apply eqn 4, summing over the $k+1$ st to n th photon. Using the independence of the various sums, this gives:

$$P(\phi_k) = \frac{k!}{n!} \sum_{\xi \in \Xi} \sum_{\chi \in \Xi} c_\xi c_\chi^\dagger \sum_{\sigma} \sum_{\rho} \text{Perm}(M_{\rho(\xi), \phi_{1:k}} \circ M_{\rho(\sigma(\chi)), \phi_{1:k}}) \delta(\bar{\rho}(\xi), \bar{\rho}(\sigma(\xi))). \quad (7)$$

We will now introduce partial distinguishability to the problem. Interestingly, it is straightforward to introduce partial distinguishability to equation 7. The reason for this is that in the model of partial distinguishability outlined in Section 1, each term that doesn't correspond to a fixed point picks up a factor x , since it is sensitive to the wavefunction overlap between two different photons (fixed points pick up a factor 1 by construction). Since the delta function in eqn 7 tests whether all marginalized photons correspond to fixed points of the partial permutation $(\sigma_{(\rho)}^\xi)$, this means that all non-fixed points must be in the first $1 \dots k$ modes. Therefore, the expression at partial distinguishability reads:

$$P(\phi) = \frac{k!}{n!} \sum_{\xi \in \Xi} \sum_{\chi \in \Xi} c_\xi c_\chi^\dagger \sum_{j=0}^k x^j \sum_{\sigma_j} \sum_{\rho} \text{Perm}(M_{\rho(\xi), \phi_{1:k}} \circ M_{\rho(\sigma(\chi)), \phi_{1:k}}) \delta(\bar{\rho}(\xi), \bar{\rho}(\sigma(\xi))), \quad (8)$$

Which is the central result of this work. In this expression, σ_j is a permutation with $k-j$ fixed points.

Equation 8 shows the interplay between marginalizing over photons and partial distinguishability. It shows that there are only two possibilities: either all unfixed points lie ρ , in which case the effect of partial distinguishability doesn't change (j -th order interference remains at j -th order) or there are one or more unfixed points in $\bar{\rho}$, in which case the entire term sums to zero. This proves the claim in the introduction that marginal distributions can be used to test lower-order interference up to k -th order. This was noted previously by Walschaers *et al* for two-point correlators, which are closely related to second order marginal probabilities [26].

5 From computing output probabilities to sampling

Eqn 8 only allows us to compute marginal probabilities, however what we wish to do is sample from the distribution given by eqn 1, which has these marginal probabilities, or even sample from some approximation of that distribution. If we wish to generate samples from eqn 1, without any further approximations, things are fairly straightforward. Eqn 8 allows us to use the procedure of Clifford and Clifford, which relies on the repeated application of the chain rule for probabilities to place the photons sequentially. We place the first photon according to $P(\phi_1)$, which results in placement of that photon in the j -th mode. The second photon is placed according to $P(\phi_2)|_{k_1=j}$, i.e. $k = 2$ with the first photon fixed in the position where we placed it in the previous step, and so on. We repeat this until we have placed n photons. If we are sampling over a state $|\Psi\rangle$ of indeterminate photon number, we can place an outer loop around this procedure in which we first draw n according to the appropriate distribution given by $|\Psi\rangle$, and then proceed as before.

If we wish to approximate eqn 8 as only low-order interference, we can do this by truncating the sum over j to some value j_{\max} . From previous work [4, 13], we know that the distribution given by these probabilities is close in L_1 distance to the actual distribution, if the experiment has strong enough imperfections (i.e. low mutual indistinguishability or high photon loss). In that case, the j -th order marginal approximation $P'_{j_{\max}}(\phi_k)$ is given by:

$$P'_{j_{\max}}(\phi) = \frac{k!}{n!} \sum_{\xi \in \Xi} \sum_{\chi \in \Xi} c_{\xi} c_{\chi}^{\dagger} \sum_{j=0}^{j_{\max}} x^j \sum_{\sigma_j} \sum_{\rho} \text{Perm}(M_{\rho(\xi), \phi_{1:k}} \circ M_{\rho(\sigma(\chi)), \phi_{1:k}}) \delta(\bar{\rho}(\xi), \bar{\rho}(\sigma(\xi))). \quad (9)$$

Since there are n^j permanents with j unfixed points, the inner sum now sums over polynomially many terms.

By incorporating this approximation into the Clifford algorithm, we have introduced some additional error. The reason for this is that since equation 9 does not correspond to a physically realizable Gram matrix for the wave function overlap of all the photons, it is no longer guaranteed that the approximate probabilities will be positive. The solution to this is to set all negative probabilities to be zero. However, since the unmodified distribution is normalized, this influences the normalization. We leave the analysis of this effect to future work.

6 Closing remarks

We conclude with a few points:

6.1 Application of Gaussian boson sampling, first and second order marginals

As an illustration, we carry out some of the work of applying this result to Gaussian boson sampling. For cases with a high degree of symmetry, it is not necessary to explicitly enumerate all $\xi \in \Xi$ and $\chi \in \Xi$. For example, for Gaussian boson sampling, we can use the pairwise product structure of the modes to simplify our computation. Furthermore, we will assume equal squeezing in all modes, leading to $c_\xi = c_\chi$ for all ξ and χ .

For the first order marginal, the pairwise structure of the state imposes that $\xi = \chi$, $\sigma = I$, where I is the identity permutation, since if $\xi \neq \chi$ they must differ in at least two places. Hence by symmetry we have

$$P(\phi_1) = \frac{1}{m} \sum_{j=1}^m |U_{j,k}|^2,$$

where m is the number of modes connected to a squeezer and k enumerates the output modes.

For the second order marginals, we have two kinds of terms: those where $\xi = \chi$, and those where ξ and χ differ by one pair. The first case gives rise to terms analogous to boson sampling with Fock states. The second case gives rise to phase-dependent terms, which have no analog in regular boson sampling. The presence of these terms was noted as a special feature of Gaussian boson sampling previously, by Phillips *et al* [27]. We now see how these terms arise: they correspond to quantum interference of different histories of how the photons were emitted. Higher order marginals will have the corresponding higher order interference terms, i.e. at the fourth marginal ξ and χ can differ in two pairs, and so on.

6.2 Maximum likelihood estimation

In [28], the idea was floated to use maximum likelihood estimation to test the quality of a boson sampler. The present results show that this can also be done via the marginal distributions, since they contain information about the mutual distinguishability. How to interpret such an analysis would depend on the level of security assumptions which we are willing to make: if we consider these measurements a trusted characterization experiment, then the information about all $k > 2$ photon interference processes is contained in the second-order marginals.

However, with the procedure outlined above, it would be possible to generate outcomes which have the higher order marginals correct, so in an adversarial setting (as is commonly employed in quantum advantage analysis), such a result cannot be used to certify large-scale photonic interference. Interestingly, this leads to a hierarchy of both of verification and of spoofing, where an analysis looking at the k -th order marginals can test for k photon interference,

but not higher. The inadequacy of the lower-order marginals in proving large-scale quantum interference in an adversarial setting was noted previously by Shchesnovich [22].

6.3 Role of photon loss

It is interesting to note that if losses in the optical system are identical over all modes, they commute with the interferometer, and may therefore arbitrarily be moved to any point in the experiment. This means that Eqn 8 is also the expression for gaussian boson sampling with loss, since we may simply consider the marginalized photons to have disappeared into a uniformly present loss channel. This in turn means that the marginal distributions cannot be used to measure the effect of photon loss onto the effective distinguishability parameter of [28].

6.4 Arbitrary quantum states

Kalai [4] asked whether it is possible to construct quantum states which are resistant against the approximation method outlined in that paper. We give an example: the quantum state $|\Psi\rangle = \frac{1}{2}\{|\psi_1\rangle + |\psi_2\rangle\}$, with $|\psi_1\rangle$ and $|\psi_2\rangle$ both of the form of Fock state boson sampling (i.e. a product state of n single photon states and vacuum), but where the modes where these two states have photons are completely disjoint (e.g. $|\psi_1\rangle$ has photons in modes $1..n$ and $|\psi_2\rangle$ has them in $n + 1..2n$). In this case, the set of ξ consists of two elements, $|\xi_1\rangle = |\psi_1\rangle$ and $|\xi_2\rangle = |\psi_2\rangle$, and any permutation σ in the cross term consists of n -photon interference. Since there are n such terms, this means the n -photon interference is enhanced exponentially compared to all others, similar to the case of random circuit sampling.

Even more intriguingly, such a state cannot be made with a linear interferometer, but requires many photon-photon nonlinearities (e.g. n CNOT gates, which are known to not be realizable deterministically in linear optics). It is an interesting question whether it is the structure of linear optics itself that imposes the noise sensitivity noticed by Kalai, since (as the above example shows) it is possible to remove this noise sensitivity by allowing nonlinear resources.

References

- [1] H.-S. Zhong *et al.*, Science **370**, 1460 (2020), <https://science.sciencemag.org/content/370/6523/1460.full.pdf>.
- [2] G. Kalai, <https://www.scottaaronson.com/blog/?p=5122>.
- [3] S. Boixo, V. N. Smelyanskiy, and H. Neven, Fourier analysis of sampling from noisy chaotic quantum circuits, 2017, arXiv:1708.01875.
- [4] G. Kalai and G. Kindler, Gaussian noise sensitivity and boson sampling, 2014, arXiv:1409.3093.

- [5] C. Lu, Private communication.
- [6] S. Aaronson and D. J. Brod, *Phys. Rev. A* **93**, 012335 (2016).
- [7] A. Arkhipov, *Phys. Rev. A* **92**, 062326 (2015).
- [8] A. E. Moylett, R. García-Patrón, J. J. Renema, and P. S. Turner, *Quantum Science and Technology* **5**, 015001 (2019).
- [9] M. Oszmaniec and D. Brod, [arXiv:1801.06166](https://arxiv.org/abs/1801.06166) .
- [10] R. Garcia-Patron, J. J. Renema, and V. Shchesnovich, [arXiv:1712.10037](https://arxiv.org/abs/1712.10037) .
- [11] H. Qi, D. J. Brod, N. Quesada, and R. Garcia-Patron, Regimes of classical simulability for noisy gaussian boson sampling, 2019, [arXiv:1905.12075](https://arxiv.org/abs/1905.12075).
- [12] S. Rahimi-Keshari, T. C. Ralph, and C. M. Caves, *Phys. Rev. X* **6**, 021039 (2016).
- [13] J. J. Renema *et al.*, *Phys. Rev. Lett.* **120**, 220502 (2018).
- [14] J. Renema, V. Shchesnovich, and R. Garcia-Patron, [arXiv:1809.01953](https://arxiv.org/abs/1809.01953) (2018).
- [15] J. J. Renema, *Physical Review A* **101**, 063840 (2020).
- [16] V. S. Shchesnovich, *Phys. Rev. A* **89**, 022333 (2014).
- [17] V. S. Shchesnovich, *Phys. Rev. A* **91**, 063842 (2015).
- [18] V. S. Shchesnovich, *Phys. Rev. A* **91**, 013844 (2015).
- [19] V. Shchesnovich, Partial distinguishability and photon counting probabilities in linear multiport devices, 2017, [arXiv:1712.03191](https://arxiv.org/abs/1712.03191).
- [20] V. S. Shchesnovich, *Physical Review A* **100** (2019).
- [21] V. Shchesnovich, [arXiv:1904.02013](https://arxiv.org/abs/1904.02013) (2019).
- [22] V. Shchesnovich, Distinguishing noisy boson sampling from classical simulations, 2019, [arXiv:1905.11458](https://arxiv.org/abs/1905.11458).
- [23] R. van der Meer, J. J. Renema, B. Brecht, C. Silberhorn, and P. W. H. Pinkse, (2020), [arXiv:2001.03596](https://arxiv.org/abs/2001.03596).
- [24] S. Aaronson and A. Arkhipov, *Theory Comput.* **9**, 143 (2013).
- [25] P. Clifford and R. Clifford, The classical complexity of boson sampling, 2017, [arXiv:1706.01260](https://arxiv.org/abs/1706.01260).
- [26] M. Walschaers *et al.*, *New Journal of Physics* **18**, 032001 (2016).
- [27] D. S. Phillips *et al.*, *Physical Review A* **99** (2019).
- [28] J. J. Renema *et al.*, Sample-efficient benchmarking of multi-photon interference on a boson sampler in the sparse regime, 2020, [arXiv:2008.09077](https://arxiv.org/abs/2008.09077).