BosonSampling is robust against small errors in the network matrix

Alex Arkhipov*
Ray and Maria Stata Center, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
(Received 25 June 2015; published 14 December 2015)

We demonstrate the robustness of BosonSampling against imperfections in the linear optical network that cause a small deviation in the matrix it implements. We show that applying a noisy matrix $\tilde{U}$ that is within ϵ of the desired matrix U in operator norm leads to an output distribution that is within ϵn of the desired distribution in variation distance, where n is the number of photons. This lets us derive a sufficient tolerance for beam splitters and phase shifters in the network. This result only concerns errors that result from the network encoding a different unitary than desired and not from other sources of noise such as photon loss and partial distinguishability.

DOI: 10.1103/PhysRevA.92.062326 PACS number(s): 03.67.Lx, 42.50.−p

I. BACKGROUND

A. BosonSampling

BosonSampling [1] is a computational problem inspired by linear optics and closely related to the matrix permanent. It models a one-step linear-optical experiment where n identical photons are produced in distinct modes and passed simultaneously through a linear network that encodes an $m\times m$ unitary matrix U. The initial state $|1_n\rangle$ consists of one boson in each of the first n modes, with the rest empty.[1] A photon-counting measurement is performed on each output mode and we consider the outcome to be the list of photon counts $S=(s_1,\ldots,s_m)$, where the s_i are non-negative integers whose sum is S. This output is random and we define $\mathcal{D}_U$ to be the resulting probability distribution over outcomes.[2]

As a computational problem, we can define BosonSampling as follows.

Definition 1 (BosonSampling). Given an $m\times m$ matrix U and a parameter n, sample the distribution of photon counts $\mathcal{D}_U$ given by

$$\Pr_{D_U}[S]=\frac{|\mathrm{Per}(U_{[n],S})|^2}{s_1!\cdots s_m!}, \tag{1.1}$$

where Per is the matrix permanent and $U_{[n],S}$ is the submatrix of U consisting of the first n rows and the columns given by S with multiplicity.

We can think of the linear optical network as acting on the n-photon Hilbert space, each of whose basis elements is labeled by each photon count. Its dimension is $N=\binom{m}{n}$, the number of partitions of n unlabeled photons into m labeled modes. Let φ be the homomorphism that takes the $m\times m$ unitary matrix U defining the action of one photon to the $N\times N$ unitary matrix $\varphi(U)$ defining its action on n identical photons. See [1] for a precise definition of φ and a proof that it is indeed a homomorphism. The definition of BosonSampling is motivated by a surprising result about its computational complexity that gives evidence that it cannot be approximated by a classical computer and thus evidence against the Church-Turing thesis.

B. Experimental realization

The definition of BosonSampling is partially motivated by it modeling a linear optical experiment. Moreover, the prospect of a computation beyond what is possible in the efficient classical world asks for a such a computational device to be built and tested, if only to check that quantum mechanics works as we would expect.

Four independent groups [2–5] have built devices to implement the BosonSampling setup for small numbers of photons and modes and checked the results to be as statistically expected. These experiments were done with $n=3$ photons and a number of modes m ranging between 5 and 9. While these are modest parameters, the goal is to eventually scale up the experiments to the point that the problem it solves is intractable for the fastest classical computers of the time.

C. Experimental noise

The question of scaling naturally leads to the issue of noise. Real experiments have imperfections that cause them to deviate slightly from the ideal model and we would like to understand what level of error is tolerable that it creates only a small deviation in the output distribution.

There are four main sources of noise: (i) incorrect or correlated initial states, (ii) imperfect coding of the unitary U by the linear optical network, (iii) partial distinguishability of photons (caused by nonsimultaneous arrival), such as mode mismatch within the circuits, and (iv) photon loss (whether in the network or due to failure to measure). In this work we will consider (ii), the effect of imperfect coding of the unitary. In current experiments, although individual components are accurate, there is difficulty in either aligning a large number of components or fabricating precise integrated optics. As a result, inaccurate unitaries remain a significant source of output error in some experiments. The five-mode and seven-mode experiments in [6] achieved respective fidelities of 0.975 and 0.950, a minority but significant contribution to the variation distance in the output distribution.

D. Bounds on noise

Many results have proven upper and lower bounds on the amount of noise in various forms that a BosonSampling

*arkhipov@mit.edu

[1]In this work we will loosen this assumption to allow any pure n-boson state.

[2]We deviate slightly from the definitions in [1], which considers only the $m\times n$ submatrix A of U consisting of the first n rows, the ones relevant to the start state $|1_n\rangle$, and defines the distribution $\mathcal{D}_A$ in terms of this.

experiment can withstand in terms of the number of photons n, either in terms of the accuracy of the output distribution or in preserving the conjectured computational hardness of BosonSampling. Leverrier and García-Patrón [7] demonstrate that to obtain a nearly correct output distribution, each linear optical element must have fidelity $1 - O(1/n^2)$ under certain assumptions. The work of Kalai and Kindler [8] argues that a noise level of additive $\omega(1/n)$ Gaussian error applied to the overall unitary matrix leads to large deviations in the output distribution and therefore allows classical simulation. Shchesnovich [9] gives sufficient conditions for an experimental realization of BosonSampling to demonstrate a conflict with the extended Church-Turing thesis. He also proves that for a small distinguishability error, a state fidelity of $O(\frac{1}{n})$ is necessary and sufficient to obtain constant distance in the distribution. Rohde and Ralph [10] give evidence that linear optical systems remain out of reach of classical simulation even in the presence of photon loss and mode mismatch. Tichy [11] bounds the difference in outcomes between partially distinguishable and perfectly identical photons.

II. MAIN RESULT

We look at the effect caused by imperfections in the linear optical network that cause a deviation in the unitary matrix that it encodes. We assume that the actual network still applies a unitary matrix $\tilde{U}$ (in particular, it takes pure states to pure states), but one that is slightly different from the desired matrix U. We will give an upper bound for the error in the output distribution in terms of the error in U. In particular, we will show that for n photons, the operator distance of $o(1/n)$ suffices to give an error of $o(1)$ in the output distribution.

Our main result is a bound on the error in the BosonSampling distribution D_U (Definition 1) caused by inaccuracy in the single-particle unitary U that encodes the action of the beam splitters and phase shifters.

Theorem 1. For unitary matrices U and $\tilde{U}$, the L_1 distance between the corresponding n-photon BosonSampling distributions $\mathcal{D}_U$ and $\mathcal{D}_{\tilde{U}}$ is bounded as

$$\|\mathcal{D}_{\tilde{U}} - \mathcal{D}_U\|_1 \leqslant n\|\tilde{U} - U\|_{\text{op}}. \tag{2.1}$$

Note that there is no dependence on the number of modes m. As a result, the accuracy of the unitaries only needs to depend on the number of photons n, with $o(\frac{1}{n})$ error sufficing.

Corollary 1. To obtain vanishingly small error $\|\mathcal{D}_{\tilde{U}} - \mathcal{D}_U\|_1 = o(1)$, it suffices for the unitary representing the entire transformation to have $\|\tilde{U} - U\|_{\text{op}} = o(\frac{1}{n})$.

This can be achieved by having each beam splitter and phase shifter in the network be sufficiently accurate. Since such a network can be made with a depth of $O(n \log m)$ components (Theorem 45 of [1]), it suffices to divide the tolerable error by that amount.

Corollary 2. In order to have $\|\mathcal{D}_{\tilde{U}} - \mathcal{D}_U\|_1 = o(1)$, it suffices for every component in the network to have an accuracy of $\|\tilde{A} - A\|_{\text{op}} = o(\frac{1}{n^2 \log m})$.

III. COMPARISON TO PREVIOUS RESULTS

The result is comparable to the standard result for the qubit-based circuits of Bernstein and Vazirani [12]. To better parallel our main result, we state the result here with identical gates and in particle language. We also generalize qubits to m-mode qudits, which does not affect the bound.

Theorem 2. Suppose one applies a noisy unitary matrix $\tilde{U}$ to each of n distinguishable particles (qudits) and then measures each particle to sample an n-tuple of measurement outcomes from $\{1,2,\ldots,m\}$ Then the distance in the outcome distribution $\mathcal{D}_{\tilde{U}}$ from that with the error-free matrix U is bounded as

$$\|\mathcal{D}_{\tilde{U}} - \mathcal{D}_U\|_1 \leqslant n\|\tilde{U} - U\|_{\text{op}}. \tag{3.1}$$

Previous work on BosonSampling noise sensitivity has given necessary bounds for the required accuracy of the linear optical network. In other words, it has been shown that above certain thresholds of noise, one gets large inaccuracies in the distribution of outcomes. Thus, it proves a certain level of noise to be prohibitive for BosonSampling. This work, in contrast, shows a certain level of accuracy to be sufficient.

The work of Leverrier and García-Patrón [7] demonstrates that each linear optical element must have fidelity $1 - O(1/n^2)$ by considering a composite experiment in which the network is applied followed by its inverse, with independent noise in each part. As shown in the Appendix, this corresponds to a required single-operator distance of $O(1/n^2)$, which has a factor of $\log m$ gap from our sufficient bound of $o(1/(n^2 \log m))$ per operator being sufficient. From our methods in Sec. IV D, this implies an overall distance of $\|\tilde{U} - U\|_{\text{op}} = O(\log m/n)$, again a factor of $\log m$ off of our result.

The work of Kalai and Kindler [8] argues that a noise level of additive $\omega(1/n)$ Gaussian error applied to the overall unitary matrix leads to large deviations in the output distribution. Specifically, above such a threshold, one finds vanishingly little correlation between the original and noise permanent of a submatrix and thus between outcomes of a BosonSampling experiment. Translating to our error model of unitary noise in the Appendix, a typical such error corresponds to the operator distance $\omega(1/\sqrt{n})$, significantly above the $O(1/n)$ distance that we show.

In both cases, once we convert the error measures to a consistent scale, we find that the sufficient bound for noise shown in this work is consistent with the necessary bound shown in previous work. Moreover, a gap remains for potential improvement. Our resulting scaling is similar to that obtained in [13], where for a small distinguishability error, a state fidelity of $O(\frac{1}{n})$ is necessary and sufficient to obtain a constant distance in the distribution.

IV. PROOF OF RESULT

A. Outline of proof

We give an outline of the proof here and prove each part in the upcoming sections. Let Ψ_0 be the initial n-boson state and let φ be the homomorphism from a unitary acting on one boson to that acting on n identical bosons. Applying unitaries U and $\tilde{U}$ to the initial state Ψ_0 produces, respectively,

$$\Psi = \varphi(U)\Psi_0, \quad \tilde{\Psi} = \varphi(\tilde{U})\Psi_0$$

Measuring Ψ and $\tilde{\Psi}$ in the standard basis gives outcome distributions $\mathcal{D}_U$ and $\mathcal{D}_{\tilde{U}}$, respectively.

The main step is Theorem 3, which states that the distance between the n-boson unitaries is at most a factor of n times that between the one-boson unitaries

$$\|\varphi(\tilde{U}) - \varphi(U)\|_{\text{op}} \leqslant n\|\tilde{U} - U\|_{\text{op}}. \tag{4.1}$$

We then conclude with a standard argument (Lemma 3) that the distance between the output distributions is at most the operator distance between the matrices that produced them

$$\|\mathcal{D}_{\tilde{U}} - \mathcal{D}_U\|_1 \leqslant \|\varphi(\tilde{U}) - \varphi(U)\|_{\text{op}}. \tag{4.2}$$

B. Effect of the homomorphism

We first show that close unitaries U and $\tilde{U}$ induce nearby n-boson unitaries $\varphi(U)$ and $\varphi(\tilde{U})$. Thus, if two operations act similarly on single bosons, then they also act similarly on n identical bosons. The blowup is simply a factor of n, the number of bosons.

Theorem 3. Let φ be the homomorphism that takes an $m\times m$ unitary matrix U acting on a single boson and produces an $N\times N$ unitary matrix acting on n identical bosons with $N = \binom{m}{n}$. Then

$$\|\varphi(\tilde{U}) - \varphi(U)\|_{\text{op}} \leqslant n\|\tilde{U} - U\|_{\text{op}}. \tag{4.3}$$

In order to prove this, it will be useful to have two lemmas. Lemma 1 expresses the operator distance between two unitary matrices A and B in terms of the eigenvalues of AB^{-1}. Lemma 2 relates the eigenvalues of $\varphi(M)$ to those of M.

Lemma 1. If A and B are unitary, their operator distance can be expressed in terms of the eigenvalues $\{\lambda_i\}$ of AB^{-1} as

$$\|A - B\|_{\text{op}} = \max_i |\lambda_i - 1|. \tag{4.4}$$

Proof. Since AB^{-1} is unitary, it diagonalizes via unitaries as $AB^{-1} = V\text{diag}(\lambda_i)V^*$. Using the operator norm's invariance to left multiplication or right multiplication by a unitary, we have

$$\begin{aligned}\|A - B\|_{\text{op}} &= \|AB^{-1} - I\|_{\text{op}}\\ &= \|V(\text{diag}(\lambda_i) - I)V^*\|_{\text{op}}\\ &= \|\text{diag}(\lambda_i - 1)\|_{\text{op}}\\ &= \max_i |\lambda_i - 1|.\end{aligned}$$

Lemma 2. If M has eigenvalues $(\lambda_1,\ldots,\lambda_m)$, then the eigenvalues of $\varphi(M)$ are $\lambda_1^{s_1}\cdots\lambda_m^{s_m}$ for each ordered partition S of n into m parts with sizes $s_1,\ldots,s_m$.

Proof. Let v_i be the eigenvector corresponding to λ_i. We will construct eigenvectors of $\varphi(M)$ in terms of the v_i and note that they have the desired eigenvalues.

For each eigenvector v_i, let $v_i(x)$ be the formal polynomial $(v_i)_1x_1 + \cdots + (v_i)_nx_n$. For each S, let p_S be the degree-n polynomial

$$p_S(x) = v_1^{s_1}(x)\cdots v_m^{s_m}(x). \tag{4.5}$$

If we consider $\varphi(M)$ as it acts on the Fock basis, we see that each $p_S(x)$ is an eigenvector with eigenvalue $\lambda_1^{s_1}\cdots\lambda_m^{s_m}$:

$$\begin{aligned}\varphi(M)[p_S(x)] &= (Mv_1)^{s_1}(x)\cdots(Mv_m)^{s_m}(x)\\ &= (\lambda_1 v)^{s_1}(x)\cdots(M\lambda_m v)^{s_m}(x)\\ &= \lambda_1^{s_1}\cdots\lambda_m^{s_m}[p_S(x)].\end{aligned}$$

Since we have one eigenvalue for each S, the number of which equal the dimension $\binom{m}{n}$ of $\varphi(M)$, this is the full set of eigenvalues.

Now we are ready to prove Theorem 3.

Proof. Let $(\lambda_1,\ldots,\lambda_m)$ be the eigenvalues of $\tilde{U}U^{-1}$. From Lemma 2, the eigenvalues of $\varphi(\tilde{U})\varphi(U)^{-1}$, which equal $\varphi(\tilde{U}U^{-1})$ because φ is a homomorphism, are $\lambda_1^{s_1}\cdots\lambda_m^{s_m}$ for each ordered partition S of n into m parts, which we write as λ^S for brevity.

We now bound the distance of λ^S from 1 in terms of the distances of the λ_i from 1. As eigenvalues of a unitary matrix, the λ_i are complex phases with norm 1, we can inductively apply

$$\begin{aligned}|ab - 1| &= |ab - a + a - 1|\\ &\leqslant |a||b - 1| + |a - 1|\end{aligned}$$

to get

$$|\lambda^S - 1| \leqslant \sum_i s_i|\lambda_i - 1| \leqslant n\max_i|\lambda_i - 1|. \tag{4.6}$$

From Lemma 1 we have

$$\max_i|\lambda_i - 1| = \|\tilde{U} - U\|_{\text{op}} \tag{4.7}$$

and

$$\max_S|\lambda^S - 1| = \|\varphi(\tilde{U}) - \varphi(U)\|_{\text{op}}, \tag{4.8}$$

so Eq. (4.6) gives the desired result

$$\|\varphi(\tilde{U}) - \varphi(U)\|_{\text{op}} \leqslant n\|\tilde{U} - U\|_{\text{op}}. \tag{4.9}$$

C. Bounding distance between the output distributions

In Sec. IV B we showed that $\tilde{U}$ being close to U implies that the corresponding n-boson transition matrices $\varphi(U)$ and $\varphi(\tilde{U})$ are close. We now finish with a standard argument that applying close transition matrices to the same input produces close measurement distributions.

Let Ψ_0 be the initial n-boson state. For BosonSampling, this is a Fock basis state $|1_n\rangle$, but this is not necessary for this result. Applying unitaries U and $\tilde{U}$ to Ψ_0 produces states

$$\Psi = \varphi(U)\Psi_0, \quad \tilde{\Psi} = \varphi(\tilde{U})\Psi_0.$$

The distributions $\mathcal{D}_U$ and $\mathcal{D}_{\tilde{U}}$ are produced by measuring Ψ and $\tilde{\Psi}$, respectively, in the standard basis.

We show that the distance between the distributions is bounded by the operator distance between the respective operators that produced them.

Lemma 3. $\|\mathcal{D}_{\tilde{U}} - \mathcal{D}_U\| \leqslant \|\varphi(\tilde{U}) - \varphi(U)\|_{\text{op}}$.

Proof. We first bound the Euclidian distance of the resulting states from the definition of the operator norm

$$\begin{aligned}\|\tilde{\Psi} - \Psi\| &= \|[\varphi(U) - \varphi(\tilde{U})]\Psi_0\|\\ &\leqslant \|\varphi(\tilde{U}) - \varphi(U)\|_{\text{op}}\|\Psi_0\|\\ &= \|\varphi(\tilde{U}) - \varphi(U)\|_{\text{op}}.\end{aligned} \tag{4.10}$$

Now we show that the variation distance between $\mathcal{D}_U$ and $\mathcal{D}_{\tilde{U}}$ is bounded by this distance $\|\tilde{\Psi} - \Psi\|$.

The variation distance $\|\mathcal{D}_{\tilde{U}} - \mathcal{D}_U\|_1$ corresponding to the distributions obtained from a standard basis measurement is

bounded by the trace distance, the maximum such variation over all projective measurements

$$\|\mathcal{D}_{\tilde{U}} - \mathcal{D}_U\|_1 \leqslant \|\tilde{\Psi} - \Psi\|_{\mathrm{tr}}. \tag{4.11}$$

We use the expression for trace distance between pure states and bound this expression in terms of $\|\tilde{\Psi} - \Psi\|$,

$$\begin{aligned}\|\mathcal{D}_{\tilde{U}} - \mathcal{D}_U\|_1 &\leqslant \|\tilde{\Psi} - \Psi\|_{\mathrm{tr}} \\ &= \sqrt{1 - |\langle\tilde{\Psi}|\Psi\rangle|^2} \\ &\leqslant \sqrt{1 - (\mathrm{Re}\langle\tilde{\Psi}|\Psi\rangle)^2} \\ &= \sqrt{1 - (1 - \tfrac{1}{2}\|\tilde{\Psi} - \Psi\|)^2} \\ &\leqslant \|\tilde{\Psi} - \Psi\|.\end{aligned}$$

Combining this with Eq. (4.10) gives the bound

$$\|\mathcal{D}_{\tilde{U}} - \mathcal{D}_U\|_1 \leqslant \|\varphi(\tilde{U}) - \varphi(U)\|_{\mathrm{op}}, \tag{4.12}$$

which, along with Theorem 3, gives the main result.

D. Error tolerance of components of the linear optical network

We now investigate the maximum error on components of the linear optical network that still guarantees that the output distribution is vanishingly close to the ideal one. This requires bounding the error of the unitary produced by a linear optical network in terms of that of its components.

Proposition 1. If each component $\tilde{A}$ of a linear optical network is within operator distance ϵ of the ideal component A,

$$\|\tilde{A} - A\|_{\mathrm{op}} \leqslant \epsilon, \tag{4.13}$$

then the produced unitary U acting on the first n modes has accuracy

$$\|\tilde{U} - U\|_{\mathrm{op}} = O(n\epsilon \log m) \tag{4.14}$$

and the measured output has

$$\|\mathcal{D}_{\tilde{U}} - \mathcal{D}_U\|_1 = O(n^2\epsilon \log m). \tag{4.15}$$

Proof. We wish to bound the operator distance error of the network in terms of that of its components. We use two familiar facts about operator distance.

(i) For components applied in parallel, the overall operator distance error is at most that of each component. So if each component has some maximum error, so does each layer in the network.

(ii) For components applied in series, the total operator distance error is at most the sum of the operator distance errors of the components.

A linear optical network for n fixed input modes and m output modes can be implemented using $O(mn)$ beam splitters and phase shifters in a network of depth $O(n \log m)$ (Theorem 45 of [1]). So if each optimal element is within ϵ of the ideal in operator norm, we are guaranteed the following accuracy for a linear optical network:

$$\|\tilde{U} - U\|_{\mathrm{op}} = O(n \log m)\|\tilde{A} - A\|_{\mathrm{op}} = O(n\epsilon \log m). \tag{4.16}$$

Applying the main theorem then gives an overall error of

$$\|\mathcal{D}_{\tilde{U}} - \mathcal{D}_U\|_1 = O(n^2\epsilon \log m). \tag{4.17}$$

Corollary 3. In order to have $\|\mathcal{D}_{\tilde{U}} - \mathcal{D}_U\|_1 = o(1)$, it suffices for every component in the network to have an accuracy of $\|\tilde{A} - A\|_{\mathrm{op}} = o(\frac{1}{n^2 \log m})$.

V. INTERPRETATION

Note that we do not obtain that constant error suffices. In fact, a constant error does not suffice, as shown in [7,8], suggesting that fault tolerance is necessary to perform scalable quantum computing. This is not surprising: We expect that more photons require higher accuracy for the unitary because each photon interaction with the unitary introduces error. Similarly, as the network requires more and more components, each component must have better accuracy to maintain the same overall accuracy.

We conjecture that the requirement we obtain, that $\|\tilde{U} - U\|_{\mathrm{op}} = o(\frac{1}{n})$, is the best possible. It parallels the Bernstein-Vazirani result for qubit-based circuits [12]. Because each photon passes through the network and experiences its imperfections, it is natural to conjecture that the acceptable error in the network falls inversely with the number of photons. Likewise, since each photon passes through a depth $O(n \log m)$ network, one might guess that the acceptable error of each component is $O(\frac{1}{n \log m})$ times that of the full network, corresponding to the sufficient bound in Corollary 3.

Future work

This work solely addresses one type of noise: errors in the beam splitters and phase shifters that cause them to implement a slightly erroneous unitary matrix. We would like to extend these results to other sources of noise. The more plausible potential extensions of this approach are those dealing with continuous errors rather than discrete ones such as photon losses. One such source is the partial distinguishability of the photons as they pass through the network, a phenomenon that has been mathematically modeled by Tichy [14] and Xu [15].

The gaps between the sufficient bound proven here and the necessary bounds proven in [7,8] mean that an improvement must be possible to at least one of the sides. Moreover, all the results are fine-tuned for models of noise, so it would be ideal to bound the error under each of the noise models.

ACKNOWLEDGMENT

A.A. gratefully acknowledges financial support from an NSF Alan T. Waterman award given to Scott Aaronson.

APPENDIX: TRANSLATION BETWEEN NOISE MODELS

Previous work on BosonSampling noise [7,8] used measures of error different from the one we used. In order to put these results on the same scale as ours, we find the amount of operator distance error that corresponds to the errors they prove prohibitive. Note that because these results are optimized for their specific model of error, the converted results are not necessarily the strongest possible.

The work of Leverrier and García-Patrón [7] demonstrates that each linear optical element must have fidelity $1 - O(1/n^2)$. This corresponds to the operator distance $O(1/n^2)$ for each element. From the observation in Sec. IV D

that the operator distance of the whole network is at most its depth times that of each component and the result that an $O(n \log m)$ depth suffices (Theorem 45 of [1]), this corresponds to the necessary error $O(\log m/n)$.

The work of Kalai and Kindler [8] argues that a noise level of additive $\epsilon = \omega(1/n)$ Gaussian error is prohibitive for BosonSampling. We show that this corresponds to the operator distance

$$\|\tilde{U} - U\|_{\text{op}} = \omega(1/\sqrt{n}), \tag{A1}$$

so we may put it on the same scale as out result.

Consider an ϵ-noise of a matrix X. In order to match the operator distance, we consider X to be the entire $m \times m$ unitary matrix, rather than an $n \times n$ submatrix, since we expect the error to affect entries in the whole matrix just as it does in the submatrix. Since each entry of a unitary matrix has a norm of $1/\sqrt{m}$ in rms average, the error should be $\epsilon/\sqrt{m}$.

So an ϵ-noise of a unitary matrix U is given by

$$\tilde{U} = \sqrt{1-\epsilon}U + \sqrt{\epsilon}G/\sqrt{m}, \tag{A2}$$

where G is a matrix of independent and identically distributed complex Gaussian random variables. To first order in ϵ, the difference $\tilde{U} - U$ is given by

$$\tilde{U} - U = -\epsilon U/2 + \sqrt{\epsilon}G/\sqrt{m} + O(\epsilon^2). \tag{A3}$$

Since U and $G/\sqrt{m}$ have entries of the same rms norm, for small ϵ, the term with coefficient $\sqrt{\epsilon}$ dominates the remaining terms:

$$\tilde{U} - U = \sqrt{\epsilon}G/\sqrt{m} + O(\epsilon). \tag{A4}$$

Then the prohibitive amount of noise $\epsilon = \omega(1/n)$ corresponds to

$$\tilde{U} - U = \omega(1/\sqrt{m})G/\sqrt{m}. \tag{A5}$$

Finally, with the result from [16] that a random $m \times m$ Gaussian matrix has operator norm $\Theta(\sqrt{m})$ with high probability $\|G/\sqrt{m}\|_{\text{op}} = \Theta(1)$, the corresponding operator distance is

$$\|\tilde{U} - U\|_{\text{op}} = \omega(1/\sqrt{n}). \tag{A6}$$

[1] S. Aaronson and A. Arkhipov, The computational complexity of linear optics, Theory Comput. **9**, 143 (2013).

[2] M. A. Broome, A. Fedrizzi, S. Rahimi-Keshari, J. Dove, S. Aaronson, T. C. Ralph, and A. G. White, Photonic BosonSampling in a tunable circuit, Science **339**, 794 (2013).

[3] J. B. Spring, B. J. Metcalf, P. C. Humphreys, W. S. Kolthammer, X.-M. Jin, M. Barbieri, A. Datta, N. Thomas-Peter, N. K. Langford, D. Kundys, J. C. Gates, B. J. Smith, P. G. R. Smith, and I. A. Walmsley, BosonSampling on a photonic chip, Science **339**, 798 (2013).

[4] M. Tillmann, B. Dakić, R. Heilmann, S. Nolte, A. Szameit, and P. Walther, Experimental BosonSampling, Nat. Photon. **7**, 540 (2013).

[5] A. Crespi, R. Osellame, R. Ramponi, D. J. Brod, E. F. Galvão, N. Spagnolo, C. Vitelli, E. Maiorino, P. Mataloni, and F. Sciarrino, Experimental BosonSampling in arbitrary integrated photonic circuits, Nat. Photon. **7**, 545 (2013).

[6] N. Spagnolo *et al.*, Experimental validation of photonic BosonSampling, Nat. Photon. **8**, 615 (2014).

[7] A. Leverrier and R. García-Patrón, Does BosonSampling Need Fault Tolerance? Quantum Inf. Comput. **15**, 0489 (2015).

[8] G. Kalai and G. Kindler, Gaussian noise sensitivity and BosonSampling, arXiv:1409.3093.

[9] V. S. Shchesnovich, Conditions for an experimental BosonSampling computer to disprove the extended Church-Turing thesis. arXiv:1403.4459.

[10] P. P. Rohde and T. C. Ralph, Error tolerance of the BosonSampling model for linear optics quantum computing, Phys. Rev. A **85**, 022332 (2012).

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

[12] E. Bernstein and U. V. Vazirani, Quantum complexity theory, SIAM J. Comput. **26**, 1411 (1997).

[13] V. S. Shchesnovich, Tight bound on the trace distance between a realistic device with partially indistinguishable bosons and the ideal BosonSampling, Phys. Rev. A **91**, 063842 (2015).

[14] M. Tichy, Interference of identical particles from entanglement to BosonSampling, J. Phys. B **47**, 103001 (2014).

[15] C. Xu (private communication).

[16] R. Latala, Some estimates of norms of random matrices, Proc. Am. Math. Soc. **133**, 1273 (2005).