BosonSampling is robust against small errors in the network matrix

Alex Arkhipov

Phys. Rev. A **92**, 062326 — Published 14 December 2015

DOI: 10.1103/PhysRevA.92.062326

BOSON SAMPLING IS ROBUST TO SMALL ERRORS IN THE NETWORK MATRIX

ALEX ARKHIPOV

Abstract. We demonstrate the robustness of BosonSampling to 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 beamsplitters and phaseshifters 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.

1. Background

1.1. **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 nonnegative 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

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

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 labelled by each photon count. Its dimension is $N = \binom{m}{n}$, the number of partitions of n unlabelled photons into m labelled modes. Let φ be the homomorphism from the unitary $m \times m$ unitary U defining the action of one photon the $N \times N$ unitary matrix $\varphi(U)$ defining the 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.

1.2. **Experimental realization.** The definition of BosonSampling is partially motivated by it modelling 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'd expect.

Four independent groups (based in Brisbane [2], Oxford [3], Vienna [4], and Rome [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 are a number of modes

This research was supported by Scott Aaronson's NSF Waterman Award.

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

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.

1.3. **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:

(1) Incorrect or correlated initial states
(2) Imperfect coding of the unitary U by the linear optical network
(3) Partial distinguishability of photons (caused by non-simultaneous arrival), such as mode mismatch within the circuits
(4) Photon loss (whether in the network or due to failure to measure)

In this work, we will consider (2), 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 in fabricating precise integrated optics. As a result, inaccurate unitaries remain a significant source of output error in some experiments. The 5-mode and 7-mode experiments in [12] achieved respective fidelities of 0.975 and 0.950, a minority but significant contribution to the variation distance in the output distribution.

1.4. **Bounds on noise.** Many results have proven upper and lower bounds on the amount of noise in various forms that a BosonSampling experiment can withstand in terms of the number of photons n, either in terms of accuracy of the output distribution or in preserving the conjectured computational hardness of BosonSampling.

Leverrier and Patrón [6] demonstrate that to obtain a nearly-correct output distribution, each linear optical element must have fidelity $1 - O\left(1/n^2\right)$ under certain assumptions.

The work of Kalai and Kindler [7] argues that a noise level of additive $\omega\left(1/n\right)$ Gaussian error applied to the overall unitary matrix leads to large deviations in the output distribution, and therefore allows classical simulation.

Shchesnovich [14] 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 $O\left(\frac{1}{n}\right)$ is necessary and sufficient to obtain constant distance in the distribution.

Rohde and Ralph [15] give evidence that linear optical systems remain out of reach of classical simulation even in the presence of photons loss and mode mismatch.

Tichy [16] bounds the difference in outcomes between partial-distinguishable and perfectly identical photons.

2. 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, an operator distance of $o\left(1/n\right)$ suffices to give $o\left(1\right)$ error 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 beamsplitters and phaseshifters.

Theorem 2. *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*

$$\left\|\mathcal{D}_{\tilde{U}} - \mathcal{D}_U\right\|_1 \leq n \left\|\tilde{U} - U\right\|_{\text{op}} \tag{2.1}$$

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 $o\left(\frac{1}{n}\right)$ error sufficing.

Corollary 3. *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* $\left\|\tilde{U} - U\right\|_{\text{op}} = o\left(\frac{1}{n}\right)$.

This can be achieved by having each beamsplitter and phaseshifter 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 4. *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* $\left\|\tilde{A} - A\right\|_{\text{op}} = o\left(\frac{1}{n^2 \log m}\right)$.

3. Comparison to Previous Results

3.1. **Comparison to previous work.** The result is comparable to the standard result for qubit-based circuits of Bernstein and Vazirani [10]. To better parallel our main result, we state here with identical gates and in particle language. We also generalized qubits to m-mode qudits, which does not affect the bound.

Theorem 5. *Bernstein-Vazirani, adapted Suppose one applies a noisy unitary matrix* $\tilde{U}$ *to each of* n *distinguishable particles (qudits), 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 error-free matrix* U *is bounded as*

$$\|\mathcal{D}_{\tilde{U}} - \mathcal{D}_U\|_1 \leq n \left\|\tilde{U} - U\right\|_{\text{op}} \tag{3.1}$$

Previous work on BosonSampling noise sensitivity has given necessary bounds for the required accuracy of the linear optical network. In other words, it's 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 Patrón [6] demonstrates that each linear optical element must have fidelity $1 - O\left(1/n^2\right)$ by considering a composite experiment in which the network is applied followed by its inverse, with independent noise in each part. As shown in Appendix A, this corresponds to a required single-operator distance of $O\left(1/n^2\right)$, which has a factor of $\log m$ gap from our sufficient bound of $o\left(1/\left(n^2 \log m\right)\right)$ per operator being sufficient. From our methods in Section 4.4, this implies an overall distance of $\left\|\tilde{U} - U\right\|_{op} = O(\log m/n)$, again a factor of $\log m$ off of our result.

The work of Kalai and Kindler [7] 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 Appendix A, a typical such error corresponds to 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 the sufficient bound for noise shown in this work is consistent with the necessary bound shown in the previous. 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 $O\left(\frac{1}{n}\right)$ is necessary and sufficient to obtain constant distance in the distribution.

4. Proof of Result

4.1. **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 $\tilde{U}$ respectively to the initial state Ψ_0 produce:

$$\begin{aligned} \Psi &= \varphi(U)\,\Psi_0 \\ \tilde{\Psi} &= \varphi\left(\tilde{U}\right)\Psi_0 \end{aligned}$$

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

The main step is Theorem 6, which states that the distance between the n-boson unitaries is at most a factor of n times that between the 1-boson unitaries

$$\left\|\varphi\left(\tilde{U}\right)-\varphi\left(U\right)\right\|_{\mathrm{op}}\leq n\left\|\tilde{U}-U\right\|_{\mathrm{op}} \tag{4.1}$$

We then conclude with a standard argument (Lemma 10) that the distance between the output distributions is at most the operator distance between the matrices that produced them

$$\left\|\mathcal{D}_{\tilde{U}}-\mathcal{D}_{U}\right\|_{1}\leq\left\|\varphi\left(\tilde{U}\right)-\varphi\left(U\right)\right\|_{\mathrm{op}} \tag{4.2}$$

4.2. **Effect of the homomorphism.** We first show that close unitaries U and $\tilde{U}$ induce nearby n-boson unitaries $\varphi\left(U\right)$ and $\varphi\left(\tilde{U}\right)$. 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 6. *Let φ be the homomorphism that takes a $m\times m$ unitary matrix U acting on a single boson and produces a $N\times N$ unitary matrix acting on n identical bosons with $N=\binom{m}{n}$. Then,*

$$\left\|\varphi\left(\tilde{U}\right)-\varphi\left(U\right)\right\|_{\mathrm{op}}\leq n\left\|\tilde{U}-U\right\|_{\mathrm{op}} \tag{4.3}$$

In order to prove this, it will be useful to have two lemmas. Lemma 7 expresses the operator distance between two unitary matrices A and B in terms of the eigenvalues of AB^{-1}. Lemma 8 relates the eigenvalues of $\varphi\left(M\right)$ to those of M.

Lemma 7. *If A and B are unitary, their operator distance can be expressed in terms of the eigenvalues $\{\lambda_i\}$ of AB^{-1} as*

$$\left\|A-B\right\|_{\mathrm{op}}=\max_{i}\left|\lambda_i-1\right|. \tag{4.4}$$

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

$$\begin{aligned}\left\|A-B\right\|_{\mathrm{op}} &= \left\|AB^{-1}-I\right\|_{op}\\ &= \left\|V\left(\mathrm{diag}\left(\lambda_i\right)-I\right)V^{*}\right\|_{\mathrm{op}}\\ &= \left\|\mathrm{diag}\left(\lambda_i-1\right)\right\|_{\mathrm{op}}\\ &= \max_{i}\left|\lambda_i-1\right|.\end{aligned}$$

□

Lemma 8. *If M has eigenvalues $(\lambda_1,\ldots,\lambda_m)$, then the eigenvalues of $\varphi\left(M\right)$ 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 $\varphi\left(M\right)$ in terms of the v_i and note that they have the desired eigenvalues.

For each eigenvector v_i, let $v_i\left(x\right)$ be the formal polynomial $\left(v_i\right)_1 x_1+\cdots+\left(v_i\right)_n x_n$. For each S, let p_S be the degree-n polynomial

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

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

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

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

Now, we're ready to prove Theorem 6, which we restate here.

Theorem 9. *Let φ be the homomorphism that takes a $m \times m$ unitary matrix U acting on a single boson and produces a $N \times N$ unitary matrix acting on n identical bosons with $N = \binom{m}{n}$. Then,*

$$\left\|\varphi\left(\tilde{U}\right) - \varphi\left(U\right)\right\|_{\text{op}} \leq n \left\|\tilde{U} - U\right\|_{\text{op}} \tag{4.6}$$

Proof. Let $(\lambda_1, \ldots, \lambda_m)$ be the eigenvalues of $\tilde{U}U^{-1}$. From Lemma 8, the eigenvalues of $\varphi\left(\tilde{U}\right)\varphi\left(U\right)^{-1}$, which equals $\varphi\left(\tilde{U}U^{-1}\right)$ 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

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

to get

$$\left|\lambda^S - 1\right| \leq \sum_i s_i \left|\lambda_i - 1\right| \leq n \max_i \left|\lambda_i - 1\right| \tag{4.7}$$

From Lemma 7, we have

$$\max_i \left|\lambda_i - 1\right| = \left\|\tilde{U} - U\right\|_{\text{op}} \tag{4.8}$$

and

$$\max_S \left|\lambda^S - 1\right| = \left\|\varphi\left(\tilde{U}\right) - \varphi\left(U\right)\right\|_{\text{op}}, \tag{4.9}$$

so equation 4.7 gives the desired result

$$\left\|\varphi\left(\tilde{U}\right) - \varphi\left(U\right)\right\|_{\text{op}} \leq n \left\|\tilde{U} - U\right\|_{\text{op}} \tag{4.10}$$

□

4.3. **Bounding distance between the output distributions.** In Section 4.2, we showed that $\tilde{U}$ being close to U implies that the corresponding n-boson transition matrices $\varphi\left(U\right)$ and $\varphi\left(\tilde{U}\right)$ are close. We now argue 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 necesarry for this result. Applying unitaries U and $\tilde{U}$ to Ψ_0 produces states that we call

$$\begin{aligned} \Psi &= \varphi\left(U\right)\Psi_0 \\ \tilde{\Psi} &= \varphi\left(\tilde{U}\right)\Psi_0 \end{aligned}$$

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 10. $\left\|\mathcal{D}_{\tilde{U}} - \mathcal{D}_U\right\| \leq \left\|\varphi\left(\tilde{U}\right) - \varphi\left(U\right)\right\|_{\text{op}}$

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

$$\begin{aligned} \left\|\tilde{\Psi} - \Psi\right\| &= \left\|\left(\varphi\left(U\right) - \varphi\left(\tilde{U}\right)\right)\Psi_0\right\| \\ &\leq \left\|\varphi\left(\tilde{U}\right) - \varphi\left(U\right)\right\|_{\text{op}} \left\|\Psi_0\right\| \\ &= \left\|\varphi\left(\tilde{U}\right) - \varphi\left(U\right)\right\|_{\text{op}} \end{aligned} \tag{4.11}$$

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

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.

$$\|\mathcal{D}_{\tilde{U}} - \mathcal{D}_U\|_1 \leq \left\|\tilde{\Psi} - \Psi\right\|_{\mathrm{tr}} \tag{4.12}$$

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

$$\begin{aligned}
\|\mathcal{D}_{\tilde{U}} - \mathcal{D}_U\|_1 &\leq \left\|\tilde{\Psi} - \Psi\right\|_{\mathrm{tr}} \\
&= \sqrt{1 - \left|\left\langle \tilde{\Psi} \mid \Psi \right\rangle\right|^2} \\
&\leq \sqrt{1 - \left(\mathrm{Re}\left\langle \tilde{\Psi} \mid \Psi \right\rangle\right)^2} \\
&= \sqrt{1 - \left(1 - \frac{1}{2}\left\|\tilde{\Psi} - \Psi\right\|\right)^2} \\
&\leq \left\|\tilde{\Psi} - \Psi\right\|
\end{aligned}$$

□

Combining this with Equation 4.11 gives the bound

$$\|\mathcal{D}_{\tilde{U}} - \mathcal{D}_U\|_1 \leq \left\|\varphi\left(\tilde{U}\right) - \varphi\left(U\right)\right\|_{\mathrm{op}}, \tag{4.13}$$

which, along with Theorem 6

$$\left\|\varphi\left(\tilde{U}\right) - \varphi\left(U\right)\right\|_{\mathrm{op}} \leq n\left\|\tilde{U} - U\right\|_{\mathrm{op}} \tag{4.14}$$

gives the main result.

4.4. **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 a 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 11. *If each component $\tilde{A}$ of a linear optical network is within operator distance ϵ of the ideal component A*

$$\left\|\tilde{A} - A\right\|_{\mathrm{op}} \leq \epsilon, \tag{4.15}$$

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

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

and the measured output has

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

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:

- For components are 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.
- For components applied in series, the total operator distance error is at most the sum of the operator distance error of the components.

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

$$\left\|\tilde{U} - U\right\|_{op} = O(n \log m) \left\|\tilde{A} - A\right\|_{\text{op}} = O(n\epsilon \log m) \tag{4.18}$$

Applying the main theorem then gives an overall error of □

$$\left\|\mathcal{D}_{\tilde{U}} - \mathcal{D}_U\right\|_1 = O\left(n^2 \epsilon \log m\right). \tag{4.19}$$

Corollary 12. *In order to have* $\left\|\mathcal{D}_{\tilde{U}} - \mathcal{D}_U\right\|_1 = o(1)$*, it suffices for every component in the network to have an accuracy of* $\left\|\tilde{A} - A\right\|_{\text{op}} = o\left(\frac{1}{n^2 \log m}\right)$.

5. Interpretation

Note that we do not obtain that a constant error suffices. In fact, constant error does not suffice, as shown in [7, 6], 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 $\left\|\tilde{U} - U\right\|_{op} = o\left(\frac{1}{n}\right)$ is the best possible. It parallels the Bernstein-Vazirani result for qubit-based circuits [10]. 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 has guessed that the acceptable error of each component is $O\left(\frac{1}{n \log m}\right)$ time that of the full network, as corresponding to the sufficient bound in 12.

5.1. **Future work.** This work solely addresses one type of noise: errors in the beamsplitters and phaseshifters 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 like photon losses. One such source is the partial distinguishability of the photons as they pass through the network, a phenomenon that has been mathematically modelled by Tichy [9] and Xu [8].

The gaps between the sufficient bound proven here and the necessary bounds proven in [6, 7] 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.

6. appendix A: Translation Between Noise Models

Previous work on BosonSampling noise [6, 7] used different measures of error than we did. In order to put these results on the same scale as ours, we will 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 Patrón [6] demonstrates that each linear optical element must have fidelity $1 - O\left(1/n^2\right)$. This corresponds to operator distance $O\left(1/n^2\right)$ for each element. From the observation in Section 4.4 that the operator distance of the whole network is at most it depth times that of each component, and the result that $O(n \log m)$ depth suffices (Theorem 45 of [1]), this corresponds to necessary error $O(\log m/n)$.

The work of Kalai and Kindler [7] argues that a noise level of additive $\epsilon = \omega(1/n)$ Gaussian error is prohibitive for Boson Sampling. We show that this corresponds to operator distance

$$\left\|\tilde{U} - U\right\|_{op} = \omega(1/\sqrt{n}) \tag{6.1}$$

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

Consider an ϵ-noise of a matrix X. In order to match with 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 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

$$\tilde{U} = \sqrt{1-\epsilon}U + \sqrt{\epsilon}G/\sqrt{m}, \tag{6.2}$$

where G is a matrix of i.i.d. complex Gaussians. To first order in ϵ, the difference $\tilde{U} - U$ is given by

$$\tilde{U} - U = -\epsilon U/2 + \sqrt{\epsilon}G/\sqrt{m} + O(\epsilon^2) \tag{6.3}$$

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:

$$\tilde{U} - U = \sqrt{\epsilon}G/\sqrt{m} + O(\epsilon) \tag{6.4}$$

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

$$\tilde{U} - U = \omega(1/\sqrt{m})G/\sqrt{m} \tag{6.5}$$

Finally, with the result from [11] that a random $m \times m$ Gaussian matrix has operator norm $\Theta(\sqrt{m})$ with high probability, $\|G/\sqrt{m}\|_{op} = \Theta(1)$, and so the corresponding operator distance is

$$\left\|\tilde{U} - U\right\|_{op} = \omega(1/\sqrt{n}) \tag{6.6}$$

References

[1] S. Aaronson and A. Arkhipov. The computational complexity of linear optics. Theory of Computing, 9(4):143–252, 2013. Conference version in Proceedings of ACM STOC'2011. ECCC TR10-170, arXiv:1011.3245.

[2] M. A. Broome, A. Fedrizzi, S. Rahimi-Keshari, J. Dove, S. Aaronson, T. C. Ralph, and A. G. White. Photonic boson sampling in a tunable circuit. Science, (10.1126/science.1231440), 2012. arXiv:1212.2234.

[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. Boson sampling on a photonic chip. Science, (10.1126/sci- ence.1231692), 2012. arXiv:1212.2622.

[4] M. Tillmann, B. Daki´c, R. Heilmann, S. Nolte, A. Szameit, and P. Walther. Experimental boson sampling. Nature Photonics, 7:540–544, 2013. arXiv:1212.2240.

[5] A. Crespi, R. Osellame, R. Ramponi, D. J. Brod, E. F. Galv ao, N. Spagnolo, C. Vitelli, E. Maiorino, P. Mataloni, and F. Sciarrino. Experimental boson sampling in arbitrary integrated photonic circuits. Nature Photonics, 7:545–549, 2013. arXiv:1212.2783.

[6] A. Leverrier and R. Garcia-Patrón. Does Boson Sampling Need Fault Tolerance? Quant Inf & Comp 15, 0489-0512 (2015)

[7] G. Kalai and G. Kindler. Gaussian Noise Sensitivity and BosonSampling. preprint arXiv: 1409.3093

[8] C. Xu, Physically Realistic Formulations of BosonSampling under Photon Loss or Partial Distinguishability, Manuscript by private communication, 2013.

[9] M. Tichy. Interference of Identical Particles from Entanglement to Boson-Sampling. J. Phys. B 47, 103001 (2014).

[10] E. Bernstein and U.V. Vazirani. Quantum complexity theory - SIAM. J. on Comput Vol. 26, No. 5, pp. 1411-1473, 1997.

[11] R. Latala. Some estimates of norms of random matrices. Proceedings of the AMS, 133(5): 1273–1282, 2005.

[12] Spagnolo, Nicolò, et al. Experimental validation of photonic boson sampling. Nature Photonics 8.8 (2014): 615-620.

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

[14] V. S. Shchesnovich. Conditions for an experimental Boson-sampling computer to disprove the Extended Church-Turing thesis. arXiv:1403.4459.

[15] P. P. Rohde, T. C. Ralph, Error tolerance of the boson-sampling model for linear optics quantum computing. Phys. Rev. A 85, 022332 (2012).

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