Concerns About Recent Claims of a Huge Quantum Computational Advantage via Gaussian Boson Sampling*

Gil Kalai† Guy Kindler‡

Abstract

A recent paper [28] claims to have achieved "quantum computational advantage" using photons. Specifically, the paper reports a Gaussian boson sampling experiment representing a quantum state in $\sim 10^{30}$-dimensional Hilbert space, and a sampling rate that is $\sim 10^{14}$ faster than that of digital supercomputers. This claim is based on certain statistical tests measuring the proximity of the empirical samples to the outcomes of noiseless simulations of the quantum experiment on a classical computer. We point out a polynomial-time algorithm from our 2014 paper [15] that may achieve similar or better sampling quality for the statistical methods of [28]. Our algorithm is based on taking a truncated Fourier–Hermite expansion on the Boson Sampling distribution.

1 Introduction

A recent paper [28] claims to have achieved a "quantum computational advantage" using photons. Specifically, the paper reports a a room-temperature photonic device implementing Gaussian boson sampling experiment with a sampling rate that is $\sim 10^{14}$ faster than that of digital supercomputers. This claim is based on certain statistical tests measuring the proximity of the empirical samples to the outcomes of noiseless simulations of the quantum experiment on a classical computer, and on comparing the empirical samples against a few other distributions.

We point out a polynomial-time algorithm from our 2014 paper [15] that may achieve similar or better sampling quality as tested by the statistical methods of [28]. Our algorithm is based on taking a truncated Fourier–Hermite expansion of the boson sampling distribution.

In view of the results of [15], the statistical reasoning of [28] is incorrect and therefore the conclusion of achieving huge quantum computational advantage is unfounded.

We note that apart from the matter of quantum advantage, the photonic experiments in [28] appear to represent remarkable progress in controlling photonic systems.

*First version: Dec 17, 2020; Current version March 25, 2021

†Efi Arazy School of Computer Science, IDC and Einstein Institute of Mathematics, the Hebrew University of Jerusalem.

‡School of Computer Science and Engineering, the Hebrew University of Jerusalem.

Boson Sampling. Boson sampling (Aaronson and Arkhipov [3]; see also Tishby and Troyansky [26]) is the following computational task.

1. The input is an $n \times m$ complex matrix whose rows are unit vectors.
2. The output is a sample from a probability distribution on all multisets of size n from $\{1, 2, \ldots, m\}$, where the probability of a multiset S is proportional to $\mu(S)$ times the square of the absolute value of the permanent of the associated n by n minor. Here, if the elements of the multiset occur with multiplicities $r_1, r_2, \ldots, r_k$, then $\mu(S) = 1/r_1!r_2! \ldots r_k!$.

This sampling task can be achieved by an (ideal) quantum computer [26, 3]. It is not hard to see (as noted in [15]) that noisy quantum computers with the full apparatus of quantum fault-tolerance can also achieve boson sampling with a negligible error. Boson sampling can be realized by linear systems of n noninteracting photons that describe a restricted regime of quantum algorithms. Aaronson and Arkhipov proved that a polynomial-time algorithm for boson sampling would imply that the polynomial hierarchy collapses to the third level [3]. (This result is related to a line of research initiated by Terhal and DiVincenzo [25], and others.)

Aaronson and Arkhipov proposed a way, based on boson sampling, to demonstrate a quantum computational advantage (also referred to as "quantum supremacy") without quantum fault-tolerance. This bold proposal is based on three conjectures. The first conjecture, on the computational complexity side, is that achieving an approximate version of boson sampling, even for a (complex) Gaussian random matrix, will be computationally hard for classical computers. The second conjecture is that this computational complexity asymptotic hardness will come to play for moderate-size systems. The third conjecture is that approximate versions of boson sampling will be able to be achieved experimentally with quantum devices when the number of bosons is not very large, but still large enough (by the second conjecture) to demonstrate quantum computational advantage.

The experiment reported in [28] indeed claims to have achieved a huge quantum computational advantage. It is based on a variant of boson sampling called "Gaussian boson sampling" [10], where the amplitudes are described in terms of hafnians rather than permanents. (We note that, in view of several later variants, the original version of boson sampling is referred to as "Fock state" boson sampling.)

2 Fourier–Hermite decomposition and noise

Fourier–Hermite decomposition and truncation Let $h_j(x)$ be the normalized Hermite polynomial of degree j. For $d = (d_1, \ldots, d_n)$ we can define a multivariate Hermite polynomial $h_d(X) = \prod_{i=1}^{n} h_{d_i}(x_i)$, and the set of such polynomials is an orthonormal basis for $L_2(\mathbb{R}^n)$.

Let f be a function from $\mathbb{R}^n$ to $\mathbb{R}$. Consider the expansion of f in terms of Hermite polynomials, i.e.,

$$f(x) = \sum_{\beta \in \mathbb{N}^d} \hat{f}(\beta) \prod_{i=1}^{d} h_{\beta_i}(x_i). \tag{1}$$

The values $\hat{f}(\beta)$ are called the Hermite coefficients of f. Let $|\beta| = \beta_1 + \cdots + \beta_n$. Let $k > 0$ be an integer; define

$$L_k(f(x)) = \sum_{\beta \in \mathbb{N}^d, |\beta| \leq k} \hat{f}(\beta) \prod_{i=1}^{d} h_{\beta_i}(x_i). \tag{2}$$

A Fourier–Hermite noise operator Let $\epsilon > 0$ be a noise parameter and let $\rho = \sqrt{1-\epsilon}$. We define $T_\rho(f)(x)$ to be the expected value of $f(y)$ where $y = \sqrt{1-\epsilon}x + \sqrt{\epsilon}u$, and u is a Gaussian random variable in $\mathbb{R}^n$ of variance 1. The description of this noise operator in terms of Hermite expansion is [16, 15]:

$$T_\rho(f) = \sum_{\beta \in N^d} \hat{f}(\beta)\rho^\beta \prod_{i=1}^{d} h_{\beta_i}(x_i). \tag{3}$$

For boson sampling the probabilities depend on complex variables and we need a slight variant of the Fourier–Hermite decomposition and the operator T_ρ that is described in [15] and reproduced below in the second Section. We also remark that the Fourier–Hermite decomposition is closely related to the ANOVA decomposition from statistics.

Various models for noise Kalai and Kindler conjectured in [15] that the effect of other types of noise on the Fourier–Hermite expansion will be similar to their (mathematically motivated) noise. This conjecture has now been supported by several works. For example, Renema, Shchesnovich, and Garcia-Patron [22], succeeded in combining many forms of noise into the formalism of the Fourier–Hermite expansion, including optical losses, photon distinguishability, dark counts in detectors, and noise on the interferometer. They also established "exchange rates" between the various forms of noise and are able to say that 'such and such an amount of noise in this component is equivalent to so much noise in that component.' Shchesnovich, [24] identified the Kalai–Kindler model as expressing "noise on the interferometer." (We would guess that a loss of 50% of the photons would correspond to $\rho \sim 1/2$ or so in our model.) For other related papers we refer the reader to items 4–23 in the bibliography of [20].

A simple Metropolis–Hastings approximate sampling algorithm Let $\mathcal{P}(x) = (P_1(x), P_2(x), \ldots, P_M(x))$ be a discrete probability distribution that depends on a real (or later complex) vector $x = (x_1, x_2, \ldots, x_N)$. Let $Q_i(x)$ be an approximate value of $P_i(x)$ that can be easily computed.

Now, apply the following sampling algorithm:

Let A be a real number, $A > 1$. The algorithm consists of repeating two steps:

(1) Sample uniformly at random j among $i = 1, 2, \ldots, M$.

(2) If $Q_j(x) \leq 0$ throw the sample away, if $Q_j(x) \geq A$ accept the sample, and if $0 < Q_j(x) \leq A$ accept this sample with probability $g_j(x)/A$.

Then, repeat.

Polynomial-time-sampling for noisy boson sampling Kalai and Kindler [15] used this algorithm with $Q_i(x) = L_k(T_\rho(P_i(x)))$ to get a polynomial-time algorithm for approximating noisy boson sampling when the noise is described by (the complex version of) (3) for a constant level ρ of noise. They went on to show that their algorithm can be implemented by bounded-depth (classical) circuits. Part of the analysis in [15] depends on the number of modes being larger than the square of the number of photons (which was a standard assumption of various theoretical studies at that time) and Kalai and Kindler conjectured [15] that their algorithmic conclusions would hold also in the range where the number of modes is sub-quadratic and even proportional to the number of photons. This is more relevant to current experiments. In this case, one needs to analyze the Fourier–Hermite description of permanents with some repeated columns which is more complex.

Polynomial-time "spoofing" For the sake of giving an easy-to-compute sampling algorithm that succeeds in the statistical tests of [28] ("spoofing") we let $Q_i(x) = L_k(P_i(x))$.

This sampling will give us approximately the same k-marginals as $\mathcal{P}$. Forcing the probabilities to be nonnegative may slightly distort the marginals but will, of course, improve the correlation with the noiseless distribution.

To understand theoretically the correlation between samples obtained by this algorithm and the ideal probability distribution we need to estimate the L_2 weight of Fourier–Hermite coefficients for degrees at most k. When there are no repeated columns the correlation behaves like $(k-1)/n$ and a similar behavior can be expected for typical cases with repeated columns as well. Of course we can get a fairly good picture from simulations.

Noise sensitivity A further result from [15] is that for a wide range of sub-constant levels of noise the correlation between the noisy distribution and the ideal distribution tends to zero. (This property is referred to as *noise sensitivity*.) This suggests that empirical distributions for boson sampling experiments will be non-stationary and even chaotic and it will be interesting to test this empirically also on the data of [28].

Interpretation It is argued in [15] that the two results, polynomial-time approximation for noisy boson sampling for a constant rate of noise, and noise sensitivity even for sub-constant rate of noise, weaken the possibility of demonstrating a quantum advantage via boson sampling without quantum fault-tolerance. Kalai [13, 14] later extended this interpretation to general NISQ systems.

Judging the new claims of a huge quantum computational advantage We can expect that our algorithm will give an easy polynomial-time algorithm that may pass the statistical tests applied in the Gaussian boson sampling experiments from [27, 28], as well

as or better than the reported results. This is in contrast to the fantastic claims of the huge quantum computational advantage reported in [28]. In any case, the statistical methodology used in [28] is flawed since a correct argument for quantum advantage needs to exclude (or give good evidence against) the possibility of efficient classical sampling methods leading to samples of similar quality.

Further Remarks

The k-marginals of boson sampling distributions. The fact that any k-marginal distribution of the boson sampling distribution can be computed by a classical deterministic algorithm of complexity $O(n^k)$ is proved in [11], see also [8] and Gurvits's pioneering paper [9]. The fact that the k-marginal distributions are not flat was proved in [4].

Gaussian boson sampling, hafnians, and torontians For full analysis of the new photonic experiment that is based on Gaussian boson sampling we would need to explore the Fourier–Hermite expansion for hafnians [10, 19]. The hafniam of a symmetric $2n \times 2n$ matrix z is the sum of $\prod_{i=1}^{n} z_{a(i)b(i)}$ where $\{a(1), b(1)\}, \{a(2), b(2)\}, \ldots \{a(n), b(n)\}$ goes over all partition of $[2n]$ into n sets with two elements. While permanents correspond to weighted counting of perfect matchings in bipartite graphs, hafnians correspond to counting perfect matchings in general graphs. A basic reference on Gaussian boson sampling and hafnians is by Hamilton, Kruse, Sansoni, Barkhofen, Silberhorn, and Jex [10].

Torontians refers to a situation where we can only identify if a mode is occupied but cannot identify its multiplicity. This is relevant both to Fock state boson sampling and to Gaussian boson sampling and may actually simplify the combinatorics.

Noise-stable and purely noise-sensitive bosonic states Kalai and Kindler posed the question of identifying "noise-stable" boson sampling states, namely, states that are well approximated by their low-degree expansions. Another interesting question is whether there are pseudo-random boson sampling states, namely, states where all the Fourier–Hermite coefficients are very small in absolute value. Renema [20] provides an interesting example giving an affirmative answer to the second question.

Random circuit sampling and the Google experiment Another notable approach to demonstrating huge quantum computational advantage is via random (quantum) circuit sampling (RCS) [5]. In this case, the relevant expansion is the Fourier–Walsh expansion and the behavior is different since the Fourier coefficients are flat. See, Boixo, Smelyanskiy, and Neven ,[7] and [5] (supplement, IX.D). Therefore, with random quantum circuits we cannot expect truncation to give polynomial-time approximation, but we can still expect it to reduce the number of required steps proportionally to the value of the target fidelity.

Statistical aspects of the Google experiment and NISQ systems are studied in [23]. This study is relevant also to some aspects of the new photonic experiment. In both cases, we can model the noisy distribution as a mixture of the noiseless distribution and other distributions that describe various error events (Formula (2) in [23]). A notable difference is that in the Google experiment these other distributions are largely uncorrelated with the

noiseless distribution (and also among themselves), and this is not the case in the boson sampling experiment.

Learning The approximation from [15] allows efficient (approximate) learning of the kth-level approximation (or the noisy approximation) of $|\mathsf{permanent}(z)|^2$. This is much stronger than the ability of efficient approximation and suggests using learning packages for learning robust quantum distributions coming from NISQ systems. It could be conjectured that, in practice, robust probability distributions (or robust parameters of such distributions) in NISQ systems are learnable. Namely, that it is possible to efficiently learn from examples to sample based on the parameters of the NISQ device as to achieve similar quality to empirical noisy sampling. (This is, of course, in very stark contrast to any claim of a huge quantum advantage.) For learnability and complexity of quantum samples, see also [17].

Related critique Related concerns regarding [28] were raised by Jelmer Renema, Sergio Boixo, John Martinis, and others; see [2, 1], and [20].

3 Complex variable Fourier–Hermite expansion and a formula for permanents

We equip $\mathbb{C}^n$ with the product measure where in each coordinate we have a Gaussian normal distribution with mean 0 and variance 1. We call a random vector $z \in \mathbb{C}^n$ which is distributed according to this measure a (complex) Gaussian vector. The measure also defines a natural inner-product structure in the space of complex-valued functions on $\mathbb{C}^n$.

An orthonormal set. It is useful to use the following set of orthonormal functions, related to the real Hermite basis.

Proposition 3.1. *The functions 1, z, $\bar{z}$, and $h_2(z) = z\bar{z} - 1$ form an orthonormal set of functions. Moreover, these functions are all eigenvectors of T_ρ, with eigenvalues $1, \rho, \rho$, and ρ^2, respectively.*

For the simple proof see [15]. (**Remark:** For the present application we did not need to consider a full orthonormal system of eigenvalues of T_ρ.)

The recipe If the amplitude is described by a function $a(z_{ij})$ of a complex matrix $z = z_{ij}$, let $f(z) = |a(z)|^2$ be the associated probability, and expand f as a polynomial in products of z_{ij}, $\bar{z}_{ij}$, and $h_2(z_ij)$. $L_k(f)$ is the degree-k truncation of this expansion.

Permanents. Let $\mathbf{z} = \{z_{i,j}\}_{i,j=1,\ldots,n}$ be an $n \times n$ complex matrix of, and let $\mathsf{permanent}(\mathbf{z}) = \sum_{\sigma \in S_n} \prod_{i=1}^n z_{i,\sigma(i)}$ be its permanent. We also let

$$f(\mathbf{z}) = |\mathsf{permanent}(\mathbf{z})|^2 = \sum_{\sigma,\tau \in S_n} \prod_{i=1}^n z_{i,\sigma(i)} \bar{z}_{i,\tau(i)}. \tag{4}$$

Observation: In order to study $L_k(f)$ and $T_\rho(f)$, consider one term in the above formula that corresponds to the permutations σ and τ, and let T be the indices i on which they agree, and $T^c = [n] \setminus T$ be its complement. We can write such a term as

$$\prod_{i=1}^{n} z_{i,\sigma(i)}\bar{z}_{i,\tau(i)} = \prod_{i\in T}(z_{i,\sigma(i)}\bar{z}_{i,\sigma(i)}) \cdot \prod_{i\in T^c} z_{i,\sigma(i)}\bar{z}_{i,\tau(i)} = \prod_{i\in T}(1+h_2(z_{i,\sigma(i)})) \prod_{i\in T^c} z_{i,\sigma(i)}\bar{z}_{i,\tau(i)}$$

$$= \sum_{R\subseteq T}\left[\prod_{i\in T\setminus R} h_2(z_{i,\sigma(i)}) \prod_{i\in T^c} z_{i,\sigma(i)}\bar{z}_{i,\tau(i)}\right]$$

The degree of a term For each product in the above sum we assign a degree: we add 1 to the degree for each multiplicand of the form $z_{i,j}$ or $\bar{z}_{i,j}$, and 2 for each multiplicand of the form $h_2(z_{i,j})$. The degree of a term $\prod_{i\in T\setminus R} h_2(z_{i,\sigma(i)}) \prod_{i\in\bar{T}} z_{i,\sigma(i)}\bar{z}_{i,\tau(i)}$ is thus $2(|T|-|R|) + 2(n-|T|) = 2(n-|R|)$.

Now, $L_k(f)$ is the sum of all terms of degrees at most k.

Repeated columns, and torontians The above formula extends to permanents of matrices with repeated columns. For those, the expression will still be multilinear, so the expansion involves only $z, \bar{z}$ and h_2. As far as we can see, it extends to Hafnians as well.

For permanents of complex Gaussian matrices without repeated columns the L_2-contributions of the degree-k terms are all the same. Therefore the degree-k truncation yields a correlation of $(k-1)/n$ with the noiseless distribution. These terms are dominant when m, the number of modes, is more than quadratic in n the number of photons. When m is proportional to n typical terms will have repeated columns and the combinatorics is more complicated. (I would still expect that for a typical term $< L_k(f), f >$ is rather large.)

For the case where there are repeated columns we did not find closed formulas and it is less clear how the L_2 norm is distributed between levels. However, note that in the case at hand it is only possible to measure whether a mode is occupied or not and this leads to the notion of *torontians*, which means that we sample according to the set of occupied modes. In the case of torontians based on "Fock state boson sampling," we need to sum up the probabilities corresponding to all minors with the same support. In this case the L_2 computations are different and perhaps simpler compared to the case where we record the multiplicities. It is possible that there will be closed formulas for the combined L_2 contributions of degree-k terms of $\sum |\mathsf{permanent}(\mathbf{z})|^2$ when we sum up permanents with a given support. This is an interesting topic for further study, that may extend to the hafnian case.

3.1 Low-degree expansions

For convenience we present the first few terms in the expansion of $|\mathsf{permanent}(\mathbf{z})^2|$, where $(\mathbf{z}) = (z_{ij})$ is an $n \times n$ complex Gaussian matrix. (To move to boson sampling probabilities we need of course to normalize.)

$k = 0$ **approximation** $S(0) = n!$

$k = 1$ **approximation**

$$S(1) = S(0) + (n-1)! \sum_{i,j} [(|z(i,j)|^2 - 1)] + (n-1)! \sum_{i,j,j';j' \neq j} z(i,j)\bar{z}(i,j')$$

$k = 2$ **approximation**

$$S(2) = S(1) +$$

$$+(n-2)! \sum_{i_1,\ j_1,\ i_2,\ j_2;\ i_1 \neq j_1,\ i_2 \neq j_2} (|z(i_1,j_1)|^2 - 1)[(|z(i_2,j_2)|^2 - 1)$$

$$+(n-2)! \sum \{(|z(i_1,j_1)|^2 - 1) z(i_2,j_2)\bar{z}(i_2,j_2') :$$

$$i_1,\ i_2,\ j_1,\ j_2,\ j_2',\ j_2',\ \text{where all } i\text{'s and all } j\text{'s are distinct}\}$$

$$+(n-2)! \sum \{(z(i_1,j_1)\bar{z}(i_1,j_1'))z(i_2,j_2)\bar{z}(i_2,j_2')) :$$

$$i_1,\ i_2,\ j_1,\ j_1',\ j_2,\ j_2',\ j_2',\ \text{where the } i\text{'s are all distinct, and the } j\text{'s are all distinct}\}$$

$k = 3$ **approximation**

$$S(3) = S(2) +$$

$$+(n-3)! \sum \{(|z(i_1,j_1)|^2 - 1)(|z(i_2,j_2)|^2 - 1)(|z(i_3,j_3)|^2 - 1) :$$

$$i_1,\ i_2,\ i_3,\ j_1,\ j_2,\ j_3,\ \text{where all } i\text{'s and all } j\text{'s are distinct}\} +$$

$$+(n-3)! \sum \{(|z(i_1,j_1)|^2 - 1)(|z(i_2,j_2)|^2 - 1) z(i_3,j_3)\bar{z}(i_3,j_3') :$$

$$i_1,\ i_2,\ i_3,\ j_1,\ j_2,\ j_3,\ j_3',\ \text{where all } i\text{'s and all } j\text{'s are distinct}\} +$$

$$+(n-3)! \sum \{(|z(i_1,j_1)|^2 - 1) z(i_2,j_2)\bar{z}(i_2,j_2') z(i_3,j_3)\bar{z}(i_3,j_3') :$$

$$i_1,\ i_2,\ i_3,\ j_1,\ j_2,\ j_2',\ j_3,\ j_3',\ \text{where all } i\text{'s and all } j\text{'s are distinct}\} +$$

$$+(n-3)! \sum \{z(i_1,j_1)\bar{z}(i_1,j_1') z(i_2,j_2)\bar{z}(i_2,j_2') z(i_3,j_3)\bar{z}(i_3,j_3') :$$

$$i_1,\ i_2,\ i_3,\ j_1,\ j_1',\ j_2,\ j_2',\ j_3,\ j_3',\ \text{where all } i\text{'s and all } j\text{'s are distinct}\}.$$

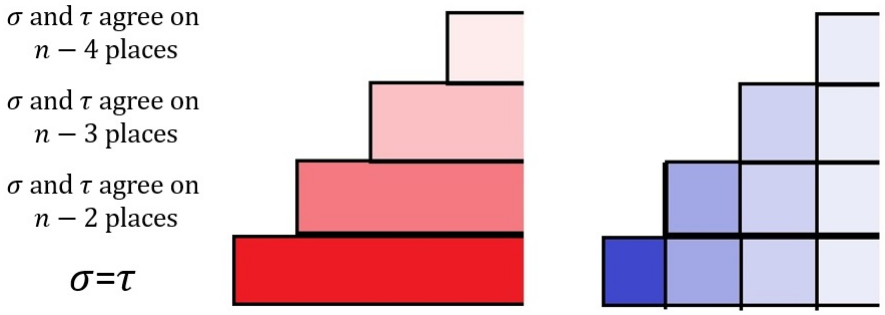

Figure 1: A schematic picture of the two hierarchies of approximations

4 A tale of two hierarchies and related noise models

Renema's [20] decomposition Renema [19, 20] considered an interesting hierarchy of approximations related to the analysis of a specific noise model for "distinguishability," where the noise pushes the system toward a classical behavior. This hierarchy is based on splitting the interference pattern into a series of j-photon interference terms, where j runs from 0 to the number of photons. It relies on several earlier works, such as Renema, Shchesnovich, and Garcia-Patron [22].

Mathematically speaking, Renema's kth approximation consists of taking all the contributions (in equation (4) coming from pairs of permutations π, τ that agree in at least $n-k$ places. Our hierarchy consists of further truncating each term in Renema's hierarchy. (In a companion manuscript [21], Renema also presented our Fourier-Hermite hierarchy in his language.) Renema's hierarchy for $k = 1$ is of special interest: the weight corresponding to an $n \times n$ matrix z_{ij} is the permanent of the real nonnegative matrix $(|z_{ij}|^2)$. See Figure 1 for a schematic picture of the two hierarchies.

Computational matters Renema [20] proposed to use an algorithm by Clifford and Clifford [8] to efficiently sample based on the kth term in his approximation. For $k = 1$ we consider "distinguishable" boson sampling where the probability corresponding to an $n \times n$ matrix z_{ij} is the permanent of the real nonnegative matrix $(|z_{ij}|^2)$. It is interesting to note that computing precisely the probabilities is **#P** complete, but classical computers are able to sample efficiently on the nose by a simple row after row procedure. The Clifford and Clifford algorithm extends this observation for larger values of k. Earlier Renema noted that when k is fixed, polynomial-time algorithms based on Jerrum, Sinclair, and Vigoda's [12] (JSV) algorithm for approximating permanents of positive matrices can be used for approximate sampling. It follows immediately that JSV allows for polynomial-time algorithm to approximate the probabilities for the case $k = 1$ and what is more remarkable and requires some work is that for a fixed $k > 1$ one only needs polynomially many calls to JSV.

We could expect that our hierarchy would allow better approximation of the probabilities for the same amount of computation. (The JSV algorithm is polynomial time but, practically, not very quick.) On the other hand, when it comes to approximate sampling, it is not clear what is the way to get better approximation for the same computational effort. It would be

interesting to test these matters empirically.

Modeling noise, and noise in practice Renema's hierarchy is motivated by his study of a specific noise model related to "distinguishability" Indeed, it allows to give a complete description of this type of noise. Our noise model was mathematically motivated but, as it turns out, it expresses "noise on the interferometer" (see, Shchesnovich [24]). We note also that in [15] (Appendix 1) we considered modeling mode-mismatches noise, which seems relevant to "distinguishability," and that it reduced high Fourier–Hermite terms as in our original suggestion.

Of course, an interesting question is which of the truncations (or rather heat-kernel operators based on them) is more relevant to realistic forms of noise. We expect that our truncation is more relevant to realistic noise. This is because while both hierarchies correspond to two different types of realistic noise we can expect that in reality *both* of them would occur. For realistic noise modeling we may consider heat operators based on different exponents for terms according to both hierarchies. In any case, this is an interesting question for further theoretical and empirical study, and, in particular, on the data from [28]. We note that it follows from the analysis in [15] that NISQ sampling exhibits non-stationary and even chaotic noise ingredients. This is supported by statistical studies of the Google data, and could be checked for the new photonic data.

Connection with computations based on Feynman diagrams Jelmer Renema kindly pointed out to us a connection that he himself learned from Valery Shchesnovich between his hierarchy of approximations and computations based on Feynman diagrams. Renema's hierarchy is based on taking the first k terms in an expansion based on Feynman diagrams, and we (apparently) take them and truncate them as well, and this can be of interest in some wider contexts. In general, computations based on Feynman diagrams are notoriously difficult even for the first few terms, and indeed this difficulty gave one of the early motivations for quantum computers. Efficient approximation methods based on Fourier-like truncations could be of great experimental and theoretical interest. (It is possible that this will shed a new point of view on known heuristic computational methods.)

Acknowledgements

We would like to thank Chaoyang Lu and Jelmer Renema, for insightful discussion and other colleagues including Scott Aaronson, Alex Arkhipov, Sergio Boxio, John Martinis, Yosi Rinott, and Gali Weinstein for helpful comments.

Gil Kalai was supported by an ERC grant 320924. Guy Kindler was supported by ISF Grant No. 2635/19.

References

[1] S. Aaronson, Quantum supremacy, now with BosonSampling, Blog post, December 3, 2020.

[2] S. Aaronson, Subsequent blog post, December 16, 2020.

[3] S. Aaronson and A. Arkhipov, The computational complexity of linear optics, *Theory of Computing* 4 (2013), 143–252. Conference version: STOC 2011.

[4] S. Aaronson and A. Arkhipov, BosonSampling is far from uniform, preprint, 2013, arXiv:1309.7460.

[5] F. Arute et al., Quantum supremacy using a programmable superconducting processor, *Nature,* 574 (2019), 505–510.

[6] I. Benjamini, G. Kalai, and O. Schramm, Noise sensitivity of Boolean functions and applications to percolation, *Publications Mathématiques de l'Institut des Hautes Études Scientifiques* 90 (1999), 5–43. arXive:9811.5157.

[7] S. Boixo, V. N. Smelyanskiy, and H. Neven, Fourier analysis of sampling from noisy chaotic quantum circuits. arXiv:1708.01875 (2017).

[8] P. Clifford and R. Clifford, The classical complexity of boson sampling, 2017, arXiv:1706.01260.

[9] L. Gurvits, On the complexity of mixed discriminants and related problems, *Mathematical Foundations of Computer Science 2005*, Lecture Notes in Computer Science, vol. 3618, Springer, Berlin, 2005, pp. 447–458.

[10] C. S. Hamilton, R. Kruse, L. Sansoni, S. Barkhofen, C. Silberhorn, I. Jex Gaussian Boson Sampling, *Physical Review Letters* 119 (2017), 170501. arXiv:1612.01199.

[11] D. A. Ivanov, L. Gurvits, Complexity of full counting statistics of free quantum particles in product states, arXiv:1904.06069.

[12] M. Jerrum, A. Sinclair and E. Vigoda, A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries, *Journal of the ACM* 51 (2004), 671–697

[13] G. Kalai, The quantum computer puzzle, *Notices of the American Mathematical Society* 63 (2016), 508–516. (Expanded version), arXiv:1605.00992.

[14] G. Kalai, The argument against quantum computers, in: M. Hemmo, and O. Shenker (eds.) *Quantum, Probability, Logic: Itamar Pitowsky's Work and Influence*, Springer, 2020, pp. 399–422. arXiv:1908.02499.

[15] G. Kalai and G. Kindler, Gaussian noise sensitivity and BosonSampling, 2014, arXiv:1409.3093.

[16] G. Kindler and R. O'Donnell, Gaussian noise sensitivity and Fourier tails, 2011, *CCC '12*.

[17] M. Niu et al. Learnability and complexity of quantum samples, 2020 arXiv:2010.11983.

[18] R. O'Donnell, *Analysis of Boolean Functions*, Cambridge University Press, 2014.

[19] J. Renema, Simulability of imperfect Gaussian and superposition boson sampling, 2019 arXiv:1911.10112.

[20] J. Renema, Marginal probabilities in boson samplers with arbitrary input states, 2020 arXiv:2012.14917

[21] J. Renema, Kalai rephrased, preprint 2020.

[22] J. Renema, V. Shchesnovich, R. Garcia-Patron, Classical simulability of noisy boson sampling, 2018 arXiv:1809.01953.

[23] Y. Rinott, T. Shoham, and G. Kalai, Statistical aspects of the quantum supremacy demonstration, 2020 arXiv:2008.05177.

[24] V. Shchesnovich, Noise in BosonSampling and the threshold of efficient classical simulatability, Physical Review A 100 (2019), 012340. arXiv:1902.02258.

[25] B. Terhal and D. DiVincenzo, Adaptive quantum computation, constant depth quantum circuits and Arthur–Merlin games, *Quantum Information and Computation* 4 (2004), 134–145. arXiv:quantph/0205133.

[26] L. Troyansky and N. Tishby, Permanent uncertainty: On the quantum evaluation of the determinant and the permanent of a matrix, In *Proc. 4th Workshop on Physics and Computation*, 1996.

[27] H.-S. Zhong et al., 12-Photon entanglement and scalable scattershot boson sampling with optimal entangled-photon pairs from parametric down-conversion. *Physical Review Letters* 121 (2018), 250505.

[28] H.-S. Zhong, et al., Quantum computational advantage using photons, *Science* 370 (2020).