Complexity of sampling as an order parameter

Abhinav Deshpande,[1,2] Bill Fefferman,[1] Michael Foss-Feig,[3,1,2] and Alexey V. Gorshkov[1,2]

[1] *Joint Center for Quantum Information and Computer Science, NIST/University of Maryland, College Park, MD 20742, USA*
[2] *Joint Quantum Institute, NIST/University of Maryland, College Park, MD 20742, USA*
[3] *United States Army Research Laboratory, Adelphi, MD 20783, USA*

We consider the classical complexity of approximately simulating time evolution under spatially local quadratic bosonic Hamiltonians for time t. We obtain upper bounds on the scaling of t with the number of bosons, n, for which simulation is classically efficient. We also obtain a lower bound on the scaling of t with n for which this problem reduces to a general instance of the boson sampling problem and is hence hard, assuming the conjectures of Aaronson and Arkhipov [Proc. 43rd Annu. ACM Symp. Theory Comput. STOC '11]. We view these results in the light of classifying phases of physical systems based on parameters in the Hamiltonian and conjecture a link to dynamical phase transitions. In doing so, we combine ideas from mathematical physics and computational complexity to gain insight into the behavior of condensed matter systems.

Usual approaches to studying the computational complexity of quantum many-body Hamiltonians have considered the problem of finding ground state energies; this research has spawned the rich area of Hamiltonian complexity [1–3] and also led to progress in other areas of physics [4, 5]. However, Hamiltonians are not only associated with energy eigenstates but also are generators of time evolution. There could be Hamiltonians that an experimentalist could simulate with polynomial resources in a lab, but still not be able to cool down to the ground state in polynomial time [6] because the corresponding problem of finding the ground state energy is QMA-hard [7–9].

This motivates us to consider a complementary problem of the classical complexity of simulating time evolution generated by a Hamiltonian, which could give evidence to refute the extended Church-Turing thesis [10–20], an area of research termed "quantum supremacy" [21] (or more recently, "quantum ascendancy"). In an early paper on this topic, Aaronson and Arkhipov [13] (AA) posed the boson sampling problem, where they showed that the task of efficient approximate sampling from a distribution produced by a system of bosons going through a linear optical circuit is classically impossible, under the assumption that the polynomial hierarchy (PH) does not collapse (and a few other conjectures about the hardness of permanents). This has in turn led to a spate of activity aimed at understanding easiness/hardness in modified settings [22–27] and an examination of resource requirements and validation [28, 29]. Inspired by this and other work on complexity classification of time evolution due to two-qubit Hamiltonians [19, 30, 31], we study the complexity of continuous time evolution of quantum systems.

Specifically, we consider the task of sampling from a distribution close in total variation distance to that obtained by evolving the initial state under a Hamiltonian and measuring in the computational basis [32]. This is a natural way of characterizing the classical complexity of simulating a Hamiltonian, since an experimentalist with access to a quantum system can only perform measurements and sample from an output state, but does not have access to the full description of the state. Henceforth, we mean simulation to be in the sense of approximate sampling [12, 33].

In this Letter, we study how the complexity of sampling from the state of a quantum system of size n evolved for a certain time t depends on the scaling of t with n. On one hand, intuitively, for very small times, Hamiltonian evolution does not change the state too much and sampling should still be easy; on the other hand, for longer times, the system can evolve into an arbitrarily complex state that would presumably be hard to sample from, by analogy with the AA result. Our primary accomplishment here is that we demonstrate and understand this transition from easy cases of sampling to hard ones. This behavior is reminiscent of systems where there is an Ehrenfest timescale that marks the transition between classical and quantum dynamics as measured by the out-of-time-ordered correlator [34]. Furthermore, we explore the link between this transition and dynamical phase transitions of quantum many-body systems.

Set-Up.— The model consists of free bosons hopping from one site (vertex) to another on a graph with m vertices (each representing a bosonic mode), and is described by the Hamiltonian $H = \sum_{i,j} J_{ij}(t) a_i^\dagger a_j$, where $a_i^\dagger$ is the creation operator of a boson at the i'th site. $J(t)$, which can be time-dependent in general, is an $m \times m$ Hermitian matrix that encodes the connectivity of the graph on which the bosons hop. Any linear optical unitary U acting on the bosonic modes can be generated through a free boson Hamiltonian by taking $H = i \log U$ and evolving it for unit time. However, this Hamiltonian can require arbitrarily long-range hops on the graph in general.

Since such arbitrarily long-ranged Hamiltonians may not always be realistic, we consider the complexity of a Hamiltonian implementing boson sampling with short-range hops. In particular, we consider the case where the

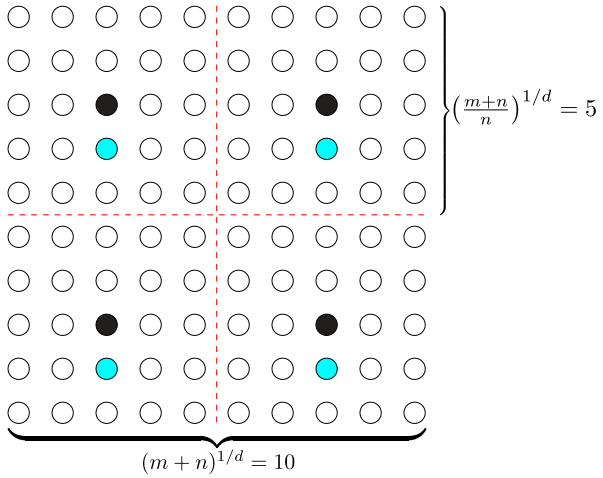

FIG. 1. (Color online). An example of the initial state in $d = 2$ dimensions. Here $m = 96$, $n = 4$, $c_1 = 3/2$ and $\beta = 3$. The black-filled circles represent modes with a single boson in them. For brevity, the mode indices are not shown here. The cyan-filled circles represent the ancilla modes.

graph is a hypercubic lattice in d dimensions (denoted d-D), where J_{ij} is nonzero only if $i = j$ or i and j label adjacent sites, thereby allowing only nearest-neighbor hops. We further restrict J_{ij} to satisfy $|J_{ij}| \leq 1$ in order to set an energy scale.

One can efficiently solve the equations of motion $i\dot{a_i}^\dagger(t) = [a_i^\dagger(t), H(t)]$ on a classical computer to obtain $a_i^\dagger(t) = \sum_k a_k^\dagger(0) R_{ki}(t)$ for some mode transformation matrix R. From here onward, we shall take $R(t)$ to be the input to the problem, since it can be determined from the input Hamiltonian H and time t in time $\text{poly}(m, \log t)$. In Ref. [35], the authors exploit the fact that the mode unitary for fermions can be found efficiently even for exponentially large t in order to implement exponentially precise measurements.

The m modes in the problem are numbered from 1 to m, and together with n ancilla modes, are arranged in a hypercubic lattice of side length $(m+n)^{1/d}$ in d dimensions. The initial state has n bosons equally spaced in the lattice as shown in Fig. 1. We take $m = c_1 n^\beta$, where β controls the sparsity of the occupied bosons in the lattice. The minimum spacing between any two bosons in the initial state is $L_{\min} = \left(\frac{m+n}{n}\right)^{1/d} > c_1^{1/d} n^{\frac{\beta-1}{d}}$. The quantity $L_{\min}$ is an important length scale in the problem. The ancilla modes (marked in cyan) in the lattice are not counted as part of the m modes, and are present in order to accelerate the time required to construct an arbitrary unitary. The presence of the ancillas does not change the scaling of quantities like $L_{\min}$ with n in our arguments.

The basis states are described by strings of the form $r = (r_1, \ldots, r_m)$ specifying the number of bosons in each mode, so that $r_1 + \ldots r_m = n$. In this Letter, r denotes the input state and s an output state. An alternative way of specifying an input/output basis state is by listing the indices of those modes occupied by one or more bosons, arranged in nondecreasing order. For example, in Fig. 1, the input state according to this notation is $\mathsf{in} = (23, 28, 71, 76)$. Modes with multiple bosons have repeated indices. Output states numbered similarly are denoted out and subscripts index these quantities, e.g. here, $\mathsf{in}_3 = 71$.

Measurement in the boson number basis gives a set of outcomes drawn from a distribution $\mathcal{D}_U$, which we aim to sample from. The probability of finding an output state $s = (s_1, s_2, \ldots, s_m)$ is given by

$$\Pr_{\mathcal{D}_U}[s] = \frac{1}{r!s!}|\text{Per}(A)|^2, \tag{1}$$

where $r! := r_1! \ldots r_m!$ (with $s!$ defined similarly) and $A_{n\times n}$ is the matrix formed by taking s_i copies of the i'th column and r_j copies of the j'th row of R, and $\text{Per}(A)$ denotes the permanent of A. The permanent has $n!$ terms, which come from the possible permutations of the n bosons in the output state, each of which is associated with a different amplitude [36, 37].

For the particular choice of initial states described in Fig. 1, the task is to sample from a distribution that is close to $\mathcal{D}_U$ in variation distance when given a description of the mode unitary $R(t)$. We now formalize the notion of efficient sampling.

Definition 1. *Efficient sampler: An efficient sampler is a classical randomized algorithm that takes as input the mode unitary R_{ij} and tolerance ϵ, and outputs a sample s from the distribution $\mathcal{D}_\mathcal{O}$ such that the variation distance between the distributions $\|\mathcal{D}_\mathcal{O} - \mathcal{D}_U\| \leq \epsilon$ in runtime* $\text{poly}(n, \frac{1}{\epsilon})$.

We will call the sampling problem *easy* if there exists an efficient sampler for the problem in the stated regime. Similarly, we call the problem *hard* in a particular regime if any hard instance of boson sampling can be realized in that regime. We restrict our attention to two special cases where we can show the existence of an efficient sampling algorithm: i) when the system evolves for a limited time so that the evolution is sufficiently close to trivial, and ii) when there is Anderson localization in the system [38]. These two cases correspond to promises on the input mode transformation unitary R. We now state our main results.

Theorem 1.A. *When $d < \beta - 1$, there exists a constant c such that the sampling problem is easy $\forall t \in o(n^c)$.*

Theorem 1.B. *(based on Theorem 3 of Ref. [13]) When $t \in \Omega(n^c)$ (where c is the same as in the statement of theorem 1.A), the sampling problem is hard in general, assuming that the* PH *does not collapse and assuming AA's conjectures on the hardness of Gaussian permanents.*

The value of c is constrained by our proof to lie in the interval $[\frac{\beta-1}{d}-1, \frac{\beta+d}{d}]$. The result for the case of Anderson localization comes out as a corollary from theorem 1.A.

Corollary 2. *For Anderson-localized systems, the sampling problem is easy for all times.*

Easiness at short times.— In this section, we prove theorem 1.A and corollary 2. First, let us examine the promise we have on the mode unitary in both cases.

At short times, we use the Lieb-Robinson bound [39] on the speed of information propagation in a system to achieve our result on easiness. Applying the bound to the free bosonic Hamiltonian, we get

$$|[a_i(t), a_j^\dagger(0)]| = |R_{ij}(t)| \leq c \exp\left(\frac{vt - \ell_{ij}}{\xi}\right), \quad (2)$$

where ℓ_{ij} is the distance between the two sites i and j, v is the upper bound to the velocity of information propagation called the Lieb-Robinson velocity, c is a constant and ξ is called the localization length scale. This bound applies even when the hopping matrix J is time-dependent. Note that the arguments from Ref. [40] that state the absence of a bound for interacting bosons do not apply, since Eq. (2) can be derived within the single-particle subspace of the Hilbert space.

When the Hamiltonian is Anderson localized, the mode unitary R satisfies the following promise at all times [41]:

$$|R_{ij}| \leq c \exp\left(\frac{-\ell_{ij}}{\xi}\right). \quad (3)$$

Here, ξ is the maximum localization length among all eigenvectors. Eq. (3) can be viewed as a consequence of Lieb-Robinson bounds with zero velocity [42]. On account of the zero-velocity Lieb-Robinson bound, all results for the time-dependent case can be ported to the Anderson localized case, setting $vt = 0$.

We now give an algorithm that efficiently samples from the output distribution for short times $t \in o(n^{\frac{\beta-1}{d}-1})$, given the promise in Eq. (2). The algorithm uses a subroutine we call the "classical-particle subroutine" that outputs a sample from the distribution $\mathcal{D}_{CP}$ obtained by assuming that the bosons are classical particles. Based on the upper bound δ to the variation distance $\|\mathcal{D}_{CP} - \mathcal{D}_U\|$ between this distribution and the actual distribution $\mathcal{D}_U$, the algorithm decides whether to use the classical-particle subroutine (if $\delta < \epsilon$) or to use a different subroutine, the brute-force subroutine that samples exactly from $\mathcal{D}_U$ (if $\delta \geq \epsilon$).

Algorithm 1: Sampling algorithm

Input: Mode unitary $R(t)$, tolerance ϵ
Output: Sample s drawn either from $\mathcal{D}_U$ or from $\mathcal{D}_{CP}$, a distribution that is close to $\mathcal{D}_U$

1 Calculate the upper bound δ to the variation distance $\|\mathcal{D}_{CP} - \mathcal{D}_U\|$ (described in the Supplemental Material [37]).
2 **if** *allowed* $\epsilon > \delta$ **then**
3 | Run `classical-particle subroutine(2)`
4 **else**
5 | Run `brute-force subroutine(3)`
6 **end**

Subroutine 2: `classical-particle subroutine`

1 $\mathcal{P}_{kl} := |R(t)|^2_{kl}$;
2 **for** i *in* $\{1, 2, \ldots n\}$, **do**
3 | Select mode l from the distribution $\mathcal{P}_{\mathsf{in}_i, l}$;
4 | Update $s_l \to s_l + 1$ (or equivalently, assign $\mathsf{out}_i = l$);
5 **end**
6 **return** configuration s (or out), a sample from $\mathcal{D}_{CP}$.

Subroutine 3: `brute-force subroutine`

1 Assign a lexicographic numbering to all the $\binom{m+n-1}{n}$ output states.
2 Draw a random number x in $[0, 1)$.;
3 $c = 0$;
4 **while** $x < 1$, **do**
5 | Set $c \to c + 1$;
6 | Set $x \to x + \Pr_{\mathcal{D}_U}(c)$ (here c is a number that refers to a particular output state) ;
7 **end**
8 **return** the configuration represented by c, a sample from $\mathcal{D}_U$.

Analysis.— We analyze the algorithm 1 by examining the correctness and runtime for both subroutines. It can be seen that the brute-force subroutine 3 outputs a state c with probability equal to $\Pr_{\mathcal{D}_U}(c)$, which is the measure of the points $x \in [0, 1]$ that lead to output c.

In the classical-particle subroutine 2, note that $\mathcal{P}$ from line 1 is a doubly stochastic matrix. We take $\mathcal{P}$ to describe the classical Markov process of particles undergoing a random walk. The probability of getting an outcome s is given by

$$\begin{aligned}\Pr_{\mathcal{D}_{CP}}[s] &= \sum_\sigma \frac{1}{s!} |R_{\mathsf{in}_1, \mathsf{out}_{\sigma(1)}}|^2 |R_{\mathsf{in}_2, \mathsf{out}_{\sigma(2)}}|^2 \ldots |R_{\mathsf{in}_n, \mathsf{out}_{\sigma(n)}}|^2 \\ &= \sum_\sigma \frac{1}{s!} \mathcal{P}_{\mathsf{in}_1, \mathsf{out}_{\sigma(1)}} \mathcal{P}_{\mathsf{in}_2, \mathsf{out}_{\sigma(2)}} \ldots \mathcal{P}_{\mathsf{in}_n, \mathsf{out}_{\sigma(n)}}, \quad (4)\end{aligned}$$

where the sum is over all permutations σ mapping the initial bosons to the output ones. We now state a result on how close the distribution $\mathcal{D}_{CP}$ is to the true distribution $\mathcal{D}_U$ (see the Supplemental Material [37] for a proof).

Lemma 3. *When $d < \beta - 1$, there exist constants*

c_2, c_3 *and* $N_0 > 0$ *such that* $\|\mathcal{D}_{CP} - \mathcal{D}_U\| \leq \delta = c_2 \exp(-c_3 n^{\frac{\beta-1}{d}})\ \forall n > N_0$ *when* $t \in o(n^{\frac{\beta-1}{d}-1})$.

Assuming this lemma, the classical-particle subroutine is sufficient for all n, ϵ such that $\delta < \epsilon$, that is, when $n > N := \max\left(\left[\frac{1}{c_3}\log\left(\frac{c_2}{\epsilon}\right)\right]^{\frac{d}{\beta-1}}, N_0\right)$. When $n < N \in O\left([\log(\frac{1}{\epsilon})]^{d/(\beta-1)}\right)$, we run the brute force subroutine that samples exactly from $\mathcal{D}_U$. The exponentially small error stated in the lemma is what enables us to prove theorem 1.A.

Proof of theorem 1.A. We have seen the correctness of algorithm 1. Showing that the runtime of both subroutines is polynomial suffices to prove the theorem. The classical-particle subroutine has runtime polynomial in n. This is because the loop runs n times, and in each run of the loop, we need to evaluate $m = \text{poly}(n)$ elements of the matrix $\mathcal{P}$ and select one among them, which takes $\text{poly}(n)$ time.

When $n < N$, we use the brute-force subroutine, whose runtime is $O(\binom{m+N-1}{N} \times N2^N)$, where the first factor is the dimension of the Hilbert space, the number of configurations of N bosons in m modes. The latter factor comes from Ryser's or Glynn's algorithm to find the permanent [43, 44]. The runtime is therefore $O(N2^N m^N) = O((2c_1)^N N^{\beta N+1})$. Writing this in terms of ϵ, the runtime is $\exp[O([\log(\frac{1}{\epsilon})]^{d/(\beta-1)} \times (\frac{\beta d}{\beta-1}\log\log(\frac{1}{\epsilon}) + \log 2c_1))]$. For $d < \beta - 1$, the exponent of the $\log(\frac{1}{\epsilon})$ factor appearing in the argument of the exp() function is less than 1. Therefore, the runtime of the algorithm is $(\frac{1}{\epsilon})^{O(1)}$, which is polynomial in $\frac{1}{\epsilon}$.

Therefore, in both cases, the overall runtime is upper bounded by a polynomial in both n and $\frac{1}{\epsilon}$: $O(\text{poly}(n, \frac{1}{\epsilon}))$, showing that we have an efficient sampler when $d < \beta - 1$ and $t \in o(n^{\frac{\beta-1}{d}-1})$. □

Hardness at longer times.— If we allow the system to evolve for a longer amount of time, we can use the time-dependent control to effect any arbitrary unitary and recover AA's results on hardness. We can perform phase gates on a mode by waiting for a particular time with an appropriate value of a diagonal element of J, with the hopping terms and other diagonal terms turned off (set to zero). We can apply a nontrivial two-mode gate between adjacent modes, for example the balanced beamsplitter unitary on the modes 1 and 2, $U = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$, by setting $H = -i(a_1^\dagger a_2 - a_1 a_2^\dagger)$ and letting it evolve for time $t = \frac{\pi}{4}$. One can also apply arbitrary unitaries using arbitrary on-site control J_{ii} and fixed, time-independent nearest-neighbor hopping $J_{ij}(t) = 1$.

Using the constructions in Refs. [45, 46], we can effect any arbitrary $m \times m$ unitary on the m optical modes from a nontrivial two-mode gate and arbitrary single-mode phase gates. This construction is possible with $O(m^2)$ gates, and consequently $O(m^2)$ depth without parallelisation. Instead, using the construction of AA that employs ancillas to obtain the desired final state rather than applying the full unitary on all the modes, we can reduce the required depth of the linear optical circuit to $O(nm^{1/d})$. Each of the n columns of the mode unitary are implemented in time $O(m^{1/d})$ [47], which corresponds to the timescale set by the Lieb-Robinson velocity and the distance between the furthest two modes in the system [48].

Proof of theorem 1.B. From the above, when $t \in \Omega(n^{1+\frac{\beta}{d}})$, we see that we can effect any arbitrary unitary, and in particular, any random unitary drawn from the Haar measure that hides a Gaussian matrix. Assuming AA's conjectures about the hardness of Gaussian permanents, we see that sampling from the output distribution is hard unless PH collapses to its third level. □

Conclusions.— We have defined the sampling problem for local Hamiltonian dynamics and given upper and lower bounds for the scaling of time $t(n)$ with the number of bosons n for which the problem is efficiently simulable or hard to classically simulate, respectively. From our results in the previous section, we discuss a few cases. We set $\beta = 5$, since this is the regime where AA's results have been shown to hold.

The Lieb-Robinson case shows two regimes of the scaling of t with n where sampling is provably easy/hard. We have shown that when $t \in o(n^{\frac{\beta-1}{d}-1}) = o(n^{\frac{4}{d}-1})$, sampling is provably easy, whereas when $t \in \Omega(n^{1+\frac{\beta}{d}}) = \Omega(n^{\frac{d+5}{d}})$, sampling is hard. Since our definitions of easiness and hardness are exhaustive and allow for no intermediate regimes, we argue that there must exist a constant c such that sampling is efficient for $t \in o(n^c)$ and hard for $t \in \Omega(n^c)$, which is illustrative of a phase transition in time. This can be viewed as a transition between two regimes, one for short times in which the system's dynamics is essentially indistinguishable from classical dynamics; and the other in which quantum mechanical effects dominate to such an extent as to forbid an efficient classical simulation. We leave as an open question whether this classical-quantum transition is associated with an observable order parameter that shows a non-analyticity at the transition time. We note that the system in Ref. [34] exhibits a similar classical-quantum transition in the out-of-time-ordered correlator with an Ehrenfest timescale. The transitions also closely mirror phase transitions in average case complexity seen in problems like k-SAT [49].

In the Anderson localized case, we observe that boson sampling is classically easy for all times when $d < \beta - 1 = 4$, i.e in 1-D, 2-D and 3-D. An open problem is whether there is a class of static, local Hamiltonians that generate a hard-to-sample output distribution at *some* (possibly exponentially large) time. An affirmative answer to this

question would show that complexity of sampling distinguishes Anderson-localized and delocalized systems, justifying the name "order parameter". This is akin to recent work that identified phase transitions based on features seen in the time-dependence of the out-of-time-ordered correlator [50, 51].

The above link to the out-of-time-ordered correlator raises the question of the connection between complexity of sampling and scrambling time [52] in quantum many-body systems [53] and fast scramblers like black holes [54–57]. Further, it would be interesting to study the classical complexity of simulating time evolution in various other physical settings, like systems with topological or many-body-localized phases, systems in their ground state after imaginary time evolution to link back with Hamiltonian complexity, systems with fast scrambling dynamics like the Sachdev-Ye [58] and Sachdev-Ye-Kitaev [50, 59–61] models and black holes, where one can explore the connection to recent conjectures on complexity in the dual CFT [57, 62, 63].

We are grateful to Mohammad Hafezi, Zhe-Xuan Gong, Sergey Syzranov, Ignacio Cirac and Emmanuel Abbe for discussions. AD and AVG acknowledge funding from ARL CDQI, ARO MURI, NSF QIS, ARO, NSF PFC at JQI, and AFOSR. After this Letter was completed, we learned of a related proposal [64].

[1] A. Kitaev, A. Shen, and M. Vyalyi, *Classical and Quantum Computation*, Graduate Studies in Mathematics, Vol. 47 (American Mathematical Society, Providence, Rhode Island, 2002).
[2] J. Kempe, A. Kitaev, and O. Regev, SIAM J. Comput. **35**, 1070 (2006).
[3] S. Gharibian, Y. Huang, Z. Landau, and S. W. Shin, Found. Trends® Theor. Comput. Sci. **10**, 159 (2015).
[4] M. Troyer and U.-J. Wiese, Phys. Rev. Lett. **94**, 170201 (2005).
[5] N. Schuch and F. Verstraete, Nat. Phys. **5**, 732 (2007).
[6] D. Janzing, P. Wocjan, and T. Beth, arXiv:quant-ph/0303186.
[7] A. M. Childs, D. Gosset, and Z. Webb, in *Autom. Lang. Program.* (Springer Berlin Heidelberg, 2013) pp. 308–319.
[8] A. M. Childs, D. Gosset, and Z. Webb, Quantum Inf. Comput. **16**, 1 (2015).
[9] T. Cubitt and A. Montanaro, SIAM J. Comput. **45**, 268 (2013).
[10] A. P. Lund, M. J. Bremner, and T. C. Ralph, arXiv:1702.03061.
[11] B. M. Terhal and D. P. DiVincenzo, Phys. Rev. A **65**, 032325 (2002).
[12] B. M. Terhal and D. P. DiVincenzo, Quantum Inf. Comput. **4**, 1 (2002).
[13] S. Aaronson and A. Arkhipov, in *Proc. 43rd Annu. ACM Symp. Theory Comput. - STOC '11* (ACM Press, New York, New York, USA, 2011) p. 333.
[14] M. J. Bremner, R. Jozsa, and D. J. Shepherd, Proc. R. Soc. A **467**, 459 (2010).
[15] B. Peropadre, A. Aspuru-Guzik, and J. J. Garcia-Ripoll, arXiv:1509.02703.
[16] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, J. M. Martinis, and H. Neven, arXiv:1608.00263.
[17] S. Aaronson and L. Chen, arXiv:1612.05903.
[18] E. Farhi and A. W. Harrow, arXiv:1602.07674.
[19] B. Fefferman, M. Foss-Feig, and A. V. Gorshkov, arXiv:1701.03167.
[20] X. Gao, S.-T. Wang, and L.-M. Duan, Phys. Rev. Lett. **118**, 040502 (2017).
[21] J. Preskill, arXiv:1203.5813.
[22] A. Arkhipov, Phys. Rev. A **92**, 062326 (2015).
[23] S. Aaronson and D. J. Brod, Phys. Rev. A **93**, 012335 (2016).
[24] B. Fefferman and C. Umans, in *11th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2016)*, Leibniz International Proceedings in Informatics (LIPIcs), Vol. 61 (2016) pp. 1:1–1:19.
[25] S. Rahimi-Keshari, A. P. Lund, and T. C. Ralph, Phys. Rev. Lett. **114**, 1 (2015).
[26] S. Rahimi-Keshari, T. C. Ralph, and C. M. Caves, Phys. Rev. X **6**, 021039 (2016).
[27] D. J. Brod, Phys. Rev. A **91**, 042316 (2015).
[28] S. Aaronson and A. Arkhipov, Quantum Inf. Comput. **14**, 1383 (2014).
[29] P. P. Rohde, Phys. Rev. A **86**, 052321 (2012); Phys. Rev. A **91**, 012307 (2015).
[30] A. M. Childs, D. Leung, L. Mančinska, and M. Ozols, Quantum Inf. Comput. **11**, 19 (2011).
[31] A. Bouland, L. Mańcinska, and X. Zhang, in *31st Conference on Computational Complexity (CCC 2016)*, Leibniz International Proceedings in Informatics (LIPIcs) (2016) pp. 28:1–28:33.
[32] We also assume that the initial state has a probability distribution that is easy to sample from.
[33] R. Jozsa, arXiv:quant-ph/0603163.
[34] E. B. Rozenbaum, S. Ganeshan, and V. Galitski, Phys. Rev. Lett. **118**, 086801 (2017).
[35] Y. Atia and D. Aharonov, arXiv:1610.09619.
[36] S. Aaronson, Proc. R. Soc. A **467**, 3393 (2011).
[37] See Supplemental Material at [URL] for further details.
[38] P. W. Anderson, Phys. Rev. **109**, 1492 (1958).
[39] E. H. Lieb and D. W. Robinson, Commun. Math. Phys. **28**, 251 (1972).
[40] J. Eisert and D. Gross, Phys. Rev. Lett. **102**, 240501 (2009).
[41] J. Fröhlich, F. Martinelli, E. Scoppola, and T. Spencer, Commun. Math. Phys. **101**, 21 (1985).
[42] E. Hamza, R. Sims, and G. Stolz, Commun. Math. Phys. **315**, 215 (2012).
[43] R. A. Brualdi and H. J. Ryser, *Combinatorial Matrix Theory*, Encyclopedia of mathematics and its applications No. 39 (Cambridge University Press, Cambridge [England] ; New York, 1991).
[44] D. G. Glynn, European J. Combin. **31**, 1887 (2010); Des. Codes Cryptogr. **68**, 39 (2013).
[45] M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani, Phys. Rev. Lett. **73**, 58 (1994).
[46] W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walsmley, Optica **3**, 1460 (2016).
[47] We can talk about circuit depth and the time for which the Hamiltonian is applied interchangeably, since in each

step of the circuit, we can apply a beamsplitter unitary in $O(1)$ time.

[48] AA's construction applied each column of the unitary in $O(\log m)$-depth, whereas we can only apply it in $O(m^{1/d})$ depth because of the spatial locality of the Hamiltonian and the Lieb-Robinson bounds that follow therein.

[49] S. Kirkpatrick and B. Selman, Science (80-.). **264**, 1297 (1994).

[50] S. Banerjee and E. Altman, arXiv:1610.04619.

[51] R.-Q. He and Z.-Y. Lu, Phys. Rev. B **95**, 054201 (2017).

[52] P. Hosur, X.-L. Qi, D. A. Roberts, and B. Yoshida, J. High Energ. Phys. **2016**, 4 (2016); D. A. Roberts and B. Yoshida, arXiv:1610.04903.

[53] Y. Huang, Y.-L. Zhang, and X. Chen, Ann. Phys. (Berlin), 1 (2016); R. Fan, P. Zhang, H. Shen, and H. Zhai, arXiv:1608.01914; Y. Chen, arXiv:1608.02765; B. Swingle and D. Chowdhury, Phys. Rev. B **95**, 060201 (2017).

[54] Y. Sekino and L. Susskind, J. High Energy Phys. **2008**, 065 (2008).

[55] N. Lashkari, D. Stanford, M. Hastings, T. Osborne, and P. Hayden, Journal of High Energy Physics **2013**, 22 (2013).

[56] B. Swingle, XRDS Crossroads, ACM Mag. Students **23**, 52 (2016).

[57] A. R. Brown, D. A. Roberts, L. Susskind, B. Swingle, and Y. Zhao, Phys. Rev. Lett. **116**, 191301 (2016); Phys. Rev. D **93**, 086006 (2016).

[58] S. Sachdev and J. Ye, Phys. Rev. Lett. **70**, 3339 (1993).

[59] A. Kitaev, "A simple model of quantum holography," (2015).

[60] J. Maldacena and D. Stanford, Phys. Rev. D **94**, 106002 (2016).

[61] J. Maldacena, S. H. Shenker, and D. Stanford, J. High Energy Phys. **2016**, 106 (2016).

[62] S. Aaronson, A. Bouland, and L. Schaeffer, arXiv:1607.05256.

[63] A. R. Brown and L. Susskind, arXiv:1701.01107.

[64] G. Muraleedharan, A. Miyake, and I. Deutsch, unpublished.