Efficient Quantum Pseudorandomness

Fernando G. S. L. Brandão,[1,2] Aram W. Harrow,[3] and Michał Horodecki[4]
[1]*Quantum Architectures and Computation Group, Microsoft Research, Redmond 11728, Washington, USA*
[2]*Department of Computer Science, University College London, London, WC1E 6BT, United Kingdom*
[3]*Center for Theoretical Physics, MIT, Cambridge MA 02139, USA*
[4]*National Quantum Information Center of Gdansk, 81-824 Sopot, Poland*

(Received 14 March 2015; published 29 April 2016)

Randomness is both a useful way to model natural systems and a useful tool for engineered systems, e.g., in computation, communication, and control. Fully random transformations require exponential time for either classical or quantum systems, but in many cases pseudorandom operations can emulate certain properties of truly random ones. Indeed, in the classical realm there is by now a well-developed theory regarding such pseudorandom operations. However, the construction of such objects turns out to be much harder in the quantum case. Here, we show that random quantum unitary time evolutions ("circuits") are a powerful source of quantum pseudorandomness. This gives for the first time a polynomial-time construction of quantum unitary designs, which can replace fully random operations in most applications, and shows that generic quantum dynamics cannot be distinguished from truly random processes. We discuss applications of our result to quantum information science, cryptography, and understanding the self-equilibration of closed quantum dynamics.

DOI: 10.1103/PhysRevLett.116.170502

Random processes are ubiquitous in both natural and engineered systems. They are both an effective way to model many systems and a vital tool in algorithms, communication, control, cryptography, and elsewhere. However, a random function on n bits is known to require $\exp[\Omega(n)]$ elementary operations [1] to implement and a similar number of random bits to even specify [2], meaning that, in fact, such random functions can neither be found in nature nor designed on a computer. Instead, we now know many methods for engineering *pseudorandom* functions using far less randomness. These pseudorandom functions can be proven to be indistinguishable from truly random functions either by any test that examines their first k moments [3] (in which case they are called k designs) or by any computationally limited test [4] (for which case the term *pseudorandom function* is usually reserved). This Letter will focus on k designs.

While these constructions mean that a carefully designed algorithm can simulate a random function in many circumstances, they do not speak to the question of whether we should expect natural processes to also resemble random functions. However, it was proved in Refs. [5,6] that even reasonably short random reversible circuits yield approximate k designs, meaning that they approximate well the first k moments of a truly random function. These circuits are defined to be sequences of basic reversible operations, each involving three bits, which is the simplest type of reversible circuit that is computationally universal. As such, they form plausible toy models for the dynamics of actual systems and provide some rigorous justification for the intuition that generic dynamics cannot easily be distinguished from fully random functions.

In recent decades, quantum mechanics has been found to dramatically change the nature of information and information processing [7], implying, among other things, potential new applications such as quantum cryptography and computation. The above story needs then to be modified to account for quantum mechanics. The problem of quantum pseudorandomness was posed in Ref. [8], where it was asked to what extent random quantum circuits can mimic random quantum transformations. The simplest quantum systems are two-level systems, called *qubits*, and any larger quantum system can be decomposed into some number n of qubits [e.g., the state of m fermions in n modes can be expressed using $\lceil \binom{n}{m} \rceil$ qubits]. A quantum circuit is a sequence of gates, each acting on a constant number of qubits. Short quantum circuits are roughly equivalent in power to time evolution by local Hamiltonians for short times. The uniform distribution over unitary matrices is called the *Haar measure* and, as in the classical case, it has been extensively studied as a model of natural generic dynamics [9], with applications to black holes [10,11], quantum information processing [12–14], and elsewhere. In a further parallel to the classical case, Haar-random unitary matrices on n qubits cannot be implemented, even approximately, without $\Omega(4^n)$ elementary operations and $\Omega(4^n)$ random bits [15].

We thus have the same need for quantum pseudorandomness and an analogous notion of unitary k designs. Again, unitary k designs can be used in place of Haar-random unitaries in most applications (e.g., for encoding quantum information to protect from errors [12], or realizing more efficient quantum process tomography [8]), but it has been much harder to prove that efficient unitary k designs exist. It was long conjectured that random quantum circuits yield approximate k designs, but for a long time this was known only for $k = 1$; in other words, only the first moments of a quantum state were known to rapidly equilibrate under random dynamics. However, 1 designs can be realized even without entangling operations (e.g., a random product of Pauli matrices will suffice), while Haar-uniform unitaries create states with large amounts of entanglement [16], so 1 designs do not give a qualitatively good fit for the Haar measure. A better, but still imperfect, goal is to achieve a 2 design. It is known how to efficiently engineer a 2 design on a quantum computer by selecting a random element of the so-called Clifford group—a restricted class of quantum gates—which involves creating significant entanglement while still performing operations far simpler than those resulting from the Haar measure [17].

Initial numerical work suggested that random quantum circuits were indeed approximate 2 designs [8,18]. Later work was able to establish that random circuits matched the entanglement properties of 2 designs [19,20] and, finally, that they in fact were approximate 2 designs [18,21–23]. Since even the Clifford group, which is not universal for quantum computation, yields a 2 design, this too is a limited proxy for the Haar measure. Later work achieved 3 designs [22,24] (see also Ref. [25]). In this Letter we settle the question and achieve k designs for any k via circuits of length $\mathrm{poly}(n, k)$. Full details are given in Ref. [26], where it is also shown that this work cannot be substantially improved. We follow part of the framework of Ref. [27], which conjectured our result and gave a mean-field argument supporting it.

Definitions.—Let us give a more precise definition of approximate unitary designs. First, we say a probability measure μ on $\mathbb{U}(d)$ (the group of $d \times d$ unitary matrices) is a unitary k design if for every monomial $q(U) = U_{i_1 j_1} \ldots U_{i_k j_k} U^*_{m_1 n_1} \ldots U^*_{m_k n_k}$ of a degree that is, at most, k, in the entries of the unitary U_{ij} and their complex conjugate U^*_{nm}, the average of $q(U)$ over the Haar measure is the same as the average over μ.

In turn, we say a measure μ on $\mathbb{U}(d)$ forms an ε-approximate k design if

$$|\mathbb{E}_{\mathrm{Haar}} q(U) - \mathbb{E}_{\mu} q(U)| \leq \varepsilon, \tag{1}$$

with $\mathbb{E}_{\mathrm{Haar}}$, $\mathbb{E}_{\mu}$ representing the expectations over the Haar measure and μ, respectively. There are other definitions of ε-approximate k designs [26], but they turn out to be equivalent to the one above, up to a rescale of the approximation factor [28].

We model random quantum circuits as random walks on $\mathbb{U}(2^n)$ following Refs. [10,24]. In each step of the walk, an index i is chosen uniformly at random from $\{1, \ldots, n-1\}$ and a unitary $U_{i,i+1}$ drawn from the Haar measure on $\mathbb{U}(4)$ is applied to the two neighboring qubits i and $i+1$. This is illustrated in Fig. 1. Other choices of random circuits are possible (e.g., considering non-nearest-neighbor gates or gates from different universal sets as in Ref. [29]), and variants of our results will apply to them as well. However, for concreteness we focus here on the model above. A related model of random walk on the unitary (or orthogonal) group is Kac's random walk, extensively studied in connection to statistical mechanics (see, e.g., Ref. [30]).

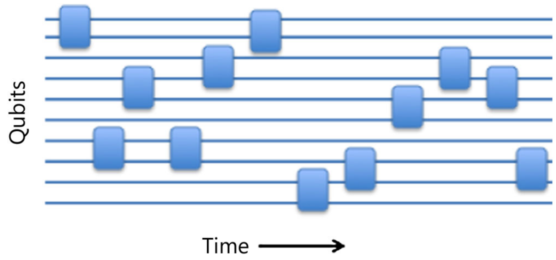

FIG. 1. In our model of a random quantum circuit, there are n qubits on a line, here arranged vertically. Time proceeds from left to right. In each time step, two adjacent qubits are chosen at random and a random two-qubit unitary operator is applied to them.

Our main result is the following.

Theorem 1.— Random circuits with $O(nk^9[nk + \log(1/\epsilon)])$ gates form ϵ-approximate k designs.

We comment briefly on these parameters. The n^2 dependence cannot be improved without changing the graph or using nonrandom circuits since it corresponds to $O(n)$ layers of $O(n)$ gates each, and with fewer layers the circuit would not be able to destroy correlations between the end points of the chain. The $\log(1/\epsilon)$ scaling also cannot be improved because of the possibility of an unlucky choice of the random gates. As for the k dependence, Proposition 8 of Ref. [26] shows that the circuit size should grow at least linearly with k. There is evidence for this scaling from a mean-field argument [27] and numerical studies [31]. However, no previous work could rule out the possibility that the scaling with k would be exponential or even worse. Our result is the first proof that the circuit size scales polynomially with k. While the degree of the polynomial is too large to use in practical applications, we expect that a more careful analysis would reduce its degree perhaps even to being linear in k.

Applications.—As k-wise independent distributions (or classical k designs) have widespread applications, so too do approximate unitary k designs [8,12,32].

Here, we briefly outline one application of our result to the problem of understanding equilibration in closed quantum dynamics. Consider the unitary time evolution of a system, initially in a fixed state—say, all spins up $|\uparrow\rangle^{\otimes n}$. The total state at any particular time is pure and hence does not appear to equilibrate in any sense. However, a long sequence of investigations, starting with von Neumann in 1929 [33], has elucidated that the state does equilibrate if one imposes constraints on the kind of observations possible [34–37]. For instance, suppose that one only has access to measurements on a few of the particles. Then it turns out that the building up of entanglement in the quantum state leads to local equilibration of every small subset of particles, for almost all times [36]. The limits of equilibration in closed quantum dynamics is an interesting problem. What is the largest class of observables for which equilibration holds? Our result on unitary designs allows us to advance this question significantly.

Define the circuit complexity of a measurement as the minimum size of any circuit of two-qubit gates that implements the measurement. Physical measurements (e.g., measurement of magnetization or heat capacity or even topological invariants) generally have low complexity. We show that in generic quantum dynamics given by random circuits (which model the case of generic evolutions under time-dependent Hamiltonians), the system equilibrates with respect to all measurements of low complexity. A general quantum measurement with two outcomes can be represented by operators M and $I - M$, with $0 \leq M \leq I$. We then have the following.

Corollary 2.— For every $k \geq 1$, for sufficiently large n's and almost all random circuits U of size $O(n^{11k+9})$ on n qubits,

$$\left| \langle \uparrow^{\otimes n} | U^{\dagger} M U | \uparrow^{\otimes n} \rangle - \frac{\operatorname{tr}(M)}{2^n} \right| \leq 2^{-n/4}, \tag{2}$$

for every measurement $\{M, I - M\}$ of circuit complexity less than n^k.*Proof sketch.*— If M is fixed and U is Haar uniform, then Eq. (2) holds with high probability; large deviations are suppressed with a probability exponential in the dimension, meaning $\exp(-2^{\Omega(n)})$. If instead U is drawn from a t design then this probability becomes $\exp(-t)$. We can approximate any low-complexity M with a measurement drawn from a set roughly of the size $\exp(n^k)$. Thus, Eq. (2) holds (approximately) for *all* low-complexity M's with a probability $\leq \exp(n^k - t)$. For this to be $\ll 1$ we need $t \gg n^k$, which, according to Theorem 1, can be achieved by a random circuit of length $O(n^{11k+9})$. The full details of this proof are in Ref. [26].

One interpretation of Corollary 2 is that long random quantum circuits appear to be Haar random when tested by significantly shorter quantum circuits, even when the tester has a full description of the random circuit. Thus it is crucial that the tester be weaker than the circuit U being tested; otherwise, the tester could apply the random circuit, then apply $U^{\dagger}$ and verify that the resulting dynamics are trivial (say, by applying them to half of a maximally entangled state).

Another interpretation of Corollary 2 is in the context of quantum cryptography. It gives a procedure for so-called quantum data hiding against a computationally bounded adversary, meaning that information is present in a state but cannot be measured without using a long quantum computation (cf. Ref. [17]). Indeed, Corollary 2 shows that all but a $2^{-\Omega(n)}$ fraction of states generated by circuits of size $O(n^{11k+9})$ cannot be distinguished from the maximally mixed state with a bias larger than $n^{-\Omega(1)}$ by any circuit of size n^k. So whether one has the particular pure state or the maximally mixed state is hidden from any adversary that is constrained to run in time n^k.

One situation where the assumption is satisfied is in the somewhat unrealistic setting when the adversary has less computational power than the honest parties. A more realistic situation is when the time it takes the adversary to decode the message is longer than the time is takes to send the message from one honest party to the other. In this case the honest parties can abort the protocol if the message is not delivered in time.*Proof overview.*— The remainder of the Letter will give a high-level description of the proof of Theorem 1. A full proof is in Ref. [26]. The proof is based on an interplay of techniques from quantum many-body theory [38], representation theory [39], and the theory of Markov chains [30], and we believe similar ideas might find further applications elsewhere.

Expressing the problem in terms of spectral properties of matrices: Classical case.—As a warm-up to understanding the properties of our random circuits, consider the classical analogue. If C is a t-gate reversible classical circuit acting on n bits, then we can think of it as a permutation matrix of the size 2^n. Since mixing over the set of all permutations requires exponentially long circuits, we instead examine the behavior of the moments of the circuit. To represent the circuit's kth moments, we can examine its action on sets of k inputs, each of which are n-bit strings. Using the tensor product, this action can be also be described as a matrix: $C^{\otimes k}$, which maps $|i_1\rangle \otimes \ldots \otimes |i_k\rangle$ to $C|i_1\rangle \otimes \cdots \otimes C|i_k\rangle$. The advantage of this representation is that the average over t-step circuits of $C^{\otimes k}$ (call it $A_{t,k}$) is simply the tth power of the average over one-step circuits $A_{1,k}$; i.e., $A_{t,k} = A_{1,k}^t$. Moreover, if the gate set is universal, then $A_{t,k}$ will approach the average over all permutations as $t \to \infty$.

Determining the rate of convergence now reduces to an eigenvalue problem. Since $A_{\infty,k} = A_{1,k}^{\infty}$, it must have only eigenvalues 0 or 1. The 1 eigenspace corresponds to the degrees of freedom that are preserved when the same circuit is applied to each element of $i_1, \ldots, i_k$, e.g., information about whether $i_1 = i_2$ or $i_1 \neq i_2$. When the set of gates is universal, the matrix $A_{1,k}$ has the same eigenspace with eigenvalue one and has all other eigenvalues smaller than one (i.e., there are no additional "constants of motion"). Thus, everything orthogonal to this subspace will decay to 0 as $t \to \infty$ at a rate controlled by the eigenvalues of $A_{1,k}$.

Our distance after t steps to the average over random permutations can be quantified by $\|A_{t,k} - A_{\infty,k}\|$. Because of the above arguments, this is given just by $(1 - \delta)^t$, where $1 - \delta$ is the second-largest eigenvalue of $A_{1,k}$ (disregarding multiplicity). This is the source of the exponential convergence typically exhibited by Markov chains on discrete state spaces. As a result, error ϵ is achieved by taking circuit length $t \geq \delta^{-1} \log \frac{1}{\epsilon}$. By proving [5,6] that $\delta \geq 1/\mathrm{poly}(k, n)$, it follows that n-bit circuits of length $\mathrm{poly}(k, n)$ have kth moments that approximate those of random permutations.

Expressing the problem in terms of spectral properties of matrices: Quantum case.—In the quantum case, the picture is similar [27] if we replace the action on n-bit strings with the action on 2^n-dimensional density matrices. The kth moments of this action can be expressed [27,40] in terms of the matrix

$$G_\mu = \int_{\mathbb{U}(2^n)} U^{\otimes k} \otimes (U^*)^{\otimes k} \mu(dU), \tag{3}$$

where μ is a distribution over unitary transformations of n qubits. This matrix can also be thought of as the matrix form of the map that sends ρ to $\int U^{\otimes k} \rho (U^\dagger)^{\otimes k} \mu(dU)$. If μ is taken to be of Haar measure, we obtain an analogue of $A_{\infty,k}$. To obtain an analogue of $A_{1,k}$, one sets $\mu = \nu \equiv \nu_n$, with ν_n representing an average over $n - 1$ choices of pairs of neighboring qubits and over a random choice of a two-qubit gate applied to those qubits. As in the classical case, G_{Haar} is the projector onto the 1 eigenspace of G_ν, which, we will argue below, has the dimension $k!$. Let $1 - \delta$ denote the next largest eigenvalue of G_ν. Then we again have that

$$\|(G_\nu)^t - G_{\mathrm{Haar}}\| = (1 - \delta)^t, \tag{4}$$

so the length of the circuit ensuring $\|(G_\nu)^t - G_{\mathrm{Haar}}\| \le \epsilon$ is given by

$$t = \delta(n, k)^{-1} \log \frac{1}{\epsilon}, \tag{5}$$

where we have made explicit the dependence of δ on n, k. Now, our main result is the following estimate:

$$\delta(n, k) \ge \Omega\left(\frac{1}{nk^{8.1} \log^2(k)}\right), \tag{6}$$

which implies that a random circuit of length $t = O(nk^{8.1} \log^2(k)[nk \log(d) + \log(1/\epsilon)])$ approximates up to ϵ the kth moment of random unitary, thus proving Theorem 1.

Connection to many-body theory.—The matrix G_ν can be expressed in terms of a *local Hamiltonian*, bringing the problem within the scope of quantum many-body theory. The quantity $\delta(n, k)$ that determines circuit length in our case was shown in Refs. [23,24,27] to be directly related to the spectral gap of some Hamiltonian consisting of nearest-neighbor interactions between n $(D = 4^k)$-dimensional systems on a line. This Hamiltonian does not correspond to a physical system, but tools from many-body theory still apply and can help evaluate the gap.

To construct the Hamiltonian, note that our random circuit ν consists of picking with probability $(1/n - 1)$ a random gate on two adjacent qubits. Thus, $G_\nu = (1/n- 1) \sum_{i=1}^{n-1} P_{i,i+1}$, where $P_{i,i+1} {:=} \int_{\mathbb{U}(4)} (U_{i,i+1})^{\otimes k,k} \mu_{\mathrm{Haar}}(dU)$ and $U_{i,i+1}$ acts on the ith and $(i + 1)$th qubit. Define

$$H_{n,k} {:=} (n - 1)(I - G_\nu) = \sum_{i=1}^{n-1} h_{i,i+1} \tag{7}$$

with the local terms $h_{i,i+1} {:=} I - P_{i,i+1}$, where I is the identity operator. It can then be shown that the "energy" of a ground state of the Hamiltonian is zero (corresponding to the 1 eigenspace of G_ν), while that of the first excited level is $(n - 1)\delta(n, k)$. In other words, the spectral gap of $H_{n,k}$ [denoted by $\Delta(H_{n,k})$] is directly related to the difference of the largest and second-largest eigenvalue of G_ν, according to $\delta(n, k) = (\Delta(H_{n,k})/n - 1)$. Thus, to determine the length of the random circuit, which approximates the k design, it suffices to lower bound the spectral gap of $H_{n,k}$.

Bounding the spectral gap.—Despite decades of research on many-body systems, the evaluation of spectral gaps of local Hamiltonians is often a formidable task. Fortunately, our Hamiltonian has a nice feature of being *nonfrustrated*: ground states of the total Hamiltonian are at the same time ground states of its local constituents. For such Hamiltonians, Nachtergaele [38] provided a sufficient condition for existence of a constant gap (in the number of qubits n), together with an estimate for the gap. Nachtergaele's criterion is given in terms of ground subspaces of Hamiltonians consisting of various numbers of systems $m \le n$. One finds that the ground space of our Hamiltonian is spanned by the $k!$ product vectors $|\psi_\sigma\rangle^{\otimes n}$, labeled by k-element permutations σ. This originates from the fact that the only operators which commute with $U^{\otimes k}$, where U is an arbitrary unitary transformation, are linear combinations of operators that permute systems (the vectors are actually isomorphic to those operators). Were the $|\psi_\sigma\rangle^{\otimes n}$ strictly orthogonal, the Nachtergaele criterion would apply immediately, but this does not hold here. Yet, by use of group representation theory, we obtain that the vectors $|\psi_\sigma\rangle^{\otimes n}$ have sufficiently small overlaps (see the Supplemental Material [41]) to enable us to apply the criterion and obtain a tight gap.

As a result, we obtain that the spectral gap of a Hamiltonian over n qubits for any n can be bounded by the gap of Hamiltonian of a *fixed* number of qubits, depending only on the moment k:

Lemma 1.— For all integers n, k with $n \ge \lceil 2.5 \log(4k) \rceil$,

$$\Delta(H_{n,k}) \ge \frac{\Delta(H_{\lceil 2.5 \log(4k) \rceil, k})}{4 \lceil 2.5 \log(4k) \rceil}, \tag{8}$$

where $\lceil x \rceil$ denotes the least integer no smaller than x.

It remains only to control the gap of the Hamiltonian for systems with $O(\log(k))$ qubits. Here, since we are concerned with a relatively small number of qubits, it suffices to establish that random circuits on m qubits mix after a number of steps that is exponential in m. We prove this in Ref. [26] using the path-coupling method in Ref. [30]. If $m = O(\log(k))$, then this means a number of steps that is polynomial in k. Translating these mixing bounds back into a statement about the gap of H, we find that $\Delta(H_{O(\log k),k}) \geq 1/\mathrm{poly}(k)$. This completes the proof of Theorem 1.

Discussion.—Our main result shows that $O(n^2 k^{10})$ random nearest-neighbor unitary interactions yield a distribution over unitaries that approximately matches the first k moments of the Haar measure. The dependence on n is approximately optimal for one dimension; indeed, we can think of it (for a fixed k) as $O(n)$ rounds of parallel nearest-neighbor interactions. Since it takes at least this much time for information to propagate along a line of n qubits, this n dependence cannot be improved. On the other hand, we did not try hard to reduce the degree of the polynomial in k and it may even be that the number of gates required is independent of k, as in Ref. [42]. Another open question is whether the n dependence could be improved for better connected geometries, e.g., nearest-neighbor gates in higher-dimensional lattices or gates with arbitrary connectivity.

Theorem 1 shows that k designs can be implemented efficiently on a quantum computer and also describes a plausible natural process that can give rise to them. Other work has studied open systems with nonunitary dynamics [43] and closed systems with Hamiltonian dynamics [44]. Our model differs from both of these but could plausibly arise from a randomly time-varying Hamiltonian. Still, a major challenge left open is to understand the conditions under which realistic physical systems thermalize.

We would like to thank Dorit Aharonov, Itai Arad, Winton Brown, Daniel Jonathan, Bruno Nachtergaele, Alex Russell, Tomasz Szarek, and Andreas Winter. F. G. S. L. B. acknowledges support from the Swiss National Science Foundation, via the National Centre of Competence in Research QSIT, and the Ministry of Education and the National Research Foundation, Singapore. A. W. H. was funded by NSF Grants No. CCF-1111382 and No. CCF-1452616 and ARO Contract No. W911NF-12-1-0486. M. H. is supported by the EU QESSENCE grant, by Polish Ministry of Science and Higher Education Grant No. N N202 231937. M. H. also acknowledges the QUASAR grant of the National Centre for Research and Development of Poland for preparing the final version of the Letter. Part of this work was done in the National Quantum Information Center of Gdansk. We thank the Institute Mittag Leffler for its hospitality within the program "Quantum Information Science," where (another) part of this work was done.

[1] $\Omega(f(n))$ refers to a function that is $\geq cf(n)$ for some cases of $c > 0$ and for a sufficiently large n.

[2] C. E. Shannon, The synthesis of two-terminal switching circuits, Bell Syst. Tech. J. **28**, 459 (1949).

[3] N. Alon, L. Babai, and A. Itai, A fast and simple randomized parallel algorithm for the maximal independent set problem, J. Algorithms **7**, 567 (1986).

[4] O. Goldreich, S. Goldwasser, and S. Micali, How to construct random functions, J. Assoc. Comput. Mach. **33**, 792 (1986).

[5] A. Brodsky and S. Hoory, Simple permutations mix even better, Random Struct. Algorithms **32**, 274 (2008).

[6] S. Hoory and A. Brodsky, Simple permutations mix even better, arXiv:math/0411098.

[7] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).

[8] J. Emerson, Y. S. Weinstein, M. Saraceno, S. Lloyd, and D. G. Cory, Pseudo-random unitary operators for quantum information processing, Science **302**, 2098 (2003).

[9] F. Haake, *Quantum Signatures of Chaos* (Springer, Berlin, 1991).

[10] P. Hayden and J. Preskill, Black holes as mirrors: Quantum information in random subsystems, J. High Energy Phys. 09 (2007) 120.

[11] L. Susskind, Computational complexity and black hole horizons, arXiv:1402.5674.

[12] A. Abeyesinghe, I. Devetak, P. Hayden, and A. Winter, The mother of all protocols: Restructuring quantum information's family tree, Proc. R. Soc. A **465**, 2537 (2009).

[13] P. Hayden, D. W. Leung, P. W. Shor, and A. Winter, Randomizing quantum states: Constructions and applications, Commun. Math. Phys. **250**, 371 (2004).

[14] P. Sen, in *Proceedings of the 21st Annual IEEE Conference on Computational Complexity (CCC 2006), Prague, 2006* (IEEE, New York, 2007), p. 274.

[15] E. Knill, Approximation by quantum circuits, arXiv:quant-ph/9508006.

[16] P. Hayden, D. W. Leung, and A. Winter, Aspects of generic entanglement, Commun. Math. Phys. **265**, 95 (2006).

[17] D. P. DiVincenzo, D. W. Leung, and B. M. Terhal, Quantum data hiding, IEEE Trans. Inf. Theory **48**, 580 (2002).

[18] L. Arnaud and D. Braun, Efficiency of producing random unitary matrices with quantum circuits, Phys. Rev. A **78**, 062329 (2008).

[19] O. C. O. Dahlsten, R. Oliveira, and M. B. Plenio, The emergence of typical entanglement in two-party random processes, J. Phys. A **40**, 8081 (2007).

[20] R. Oliveira, O. C. O. Dahlsten, and M. B. Plenio, Efficient Generation of Generic Entanglement, Phys. Rev. Lett. **98**, 130502 (2007).

[21] I. Diniz and D. Jonathan, Comment on the paper Random quantum circuits are approximate 2-designs, Commun. Math. Phys. **304**, 281 (2011).

[22] A. W. Harrow and R. A. Low, Random quantum circuits are approximate 2-designs, Commun. Math. Phys. **291**, 257 (2009).

[23] M. Žnidarič, Exact convergence times for generation of random bipartite entanglement, Phys. Rev. A **78**, 032324 (2008).

[24] F. G. S. L. Brandao and M. Horodecki, Exponential quantum speed-ups are generic, Quantum Inf. Comput. **13**, 0901 (2013).

[25] P. Cwiklinski, M. Horodecki, M. Mozrzymas, L. Pankowski, and M. Studzinski, Efficient algorithm for multiqudit twirling for ensemble quantum computation, J. Phys. A **46**, 305301 (2013).

[26] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, Local random quantum circuits are approximate polynomial-designs, arXiv:1208.0692.

[27] W. G. Brown and L. Viola, Convergence Rates for Arbitrary Statistical Moments of Random Quantum Circuits, Phys. Rev. Lett. **104**, 250501 (2010).

[28] R. A. Low, Ph.D. thesis, University of Bristol, 2010, arXiv:1006.5227.

[29] J. Emerson, Y. S. Weinstein, M. Saraceno, S. Lloyd, and D. G. Cory, Pseudo-random unitary operators for quantum information processing, Science **302**, 2098 (2003).

[30] R. I. Oliveira, On the convergence to equilibrium of Kac's random walk on matrices, Ann. Appl. Probab. **19**, 1200 (2009).

[31] P. Ćwikliński, M. Horodecki, M. Mozrzymas, Ł. Pankowski, and M. Studziński, Local random quantum circuits are approximate polynomial-designs: Numerical results, J. Phys. A **46**, 305301 (2013).

[32] R. A. Low, Large deviation bounds for k-designs, Proc. R. Soc. A **465**, 3289 (2009).

[33] J. von Neumann, Proof of the ergodic theorem and the H-theorem in quantum mechanics, Eur. Phys. J. H **35**, 201 (2010).

[34] M. C. Bañuls, J. I. Cirac, and M. B. Hastings, Strong and Weak Thermalization of Infinite Nonintegrable Quantum Systems, Phys. Rev. Lett. **106**, 050405 (2011).

[35] S. Goldstein, J. L. Lebowitz, R. Tumulka, and N. Zanghı, Canonical Typicality, Phys. Rev. Lett. **96**, 050403 (2006).

[36] N. Linden, S. Popescu, A. J. Short, and A. Winter, Quantum mechanical evolution towards thermal equilibrium, Phys. Rev. E **79**, 061103 (2009).

[37] S. Trotzky, Y.-A. Chen, A. Flesch, I. P. McCulloch, U. Schollwöck, J. Eisert, and I. Bloch, Probing the relaxation towards equilibrium in an isolated strongly correlated one-dimensional bose gas, Nat. Phys. **8**, 325 (2012).

[38] B. Nachtergaele, The spectral gap for some spin chains with discrete symmetry breaking, Commun. Math. Phys. **175**, 565 (1996).

[39] A. W. Harrow, Approximate orthogonality of permutations (unpublished).

[40] A. W. Harrow and R. A. Low, in *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, edited by I. Dinur, K. Jansen, J. Naor, and J. Rolim, Lecture Notes in Computer Science Vol. 5687 (Springer, New York, 2009), p. 548.

[41] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.116.170502 for [brief description].

[42] J. Bourgain and A. Gamburd, A spectral gap theorem in $SU(d)$, arXiv:1108.6264.

[43] H. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, New York, 2002).

[44] R. Nandkishore and D. A. Huse, Many-body localization and thermalization in quantum statistical mechanics, Annu. Rev. Condens. Matter Phys. **6**, 15 (2015).