Generic Entanglement Can Be Generated Efficiently

R. Oliveira[*]

IBM Watson Research Center, Yorktown Heights, New York, 10598, USA

O. C. O. Dahlsten[†] and M. B. Plenio[‡]

Institute for Mathematical Sciences, Imperial College London, 53 Exhibition Road, London SW7 2PG, United Kingdom
and QOLS, Blackett Laboratory, Imperial College London, Prince Consort Road, London SW7 2BW, United Kingdom

(Received 19 June 2006; published 30 March 2007)

We find that generic entanglement is physical, in the sense that it can be generated in polynomial time from two-qubit gates picked at random. We prove as the main result that such a process generates the average entanglement of the uniform (unitarily invariant) measure in at most $O(N^3)$ steps for N qubits. This is despite an exponentially growing number of such gates being necessary for generating that measure fully on the state space. Numerics furthermore show a variation cutoff allowing one to associate a specific time with the achievement of the uniform measure entanglement distribution. Various extensions of this work are discussed. The results are relevant to entanglement theory and to protocols that assume generic entanglement can be achieved efficiently.

DOI: 10.1103/PhysRevLett.98.130502 PACS numbers: 03.67.Mn, 03.65.Ud, 05.70.−a

Introduction.—Entanglement has traditionally been viewed as a fundamental tool for studies of the foundations of quantum mechanics [1]. More recently, the viewpoint of using entanglement as a resource has also gained prominence; see [2] for a recent review. While a great deal of insight into the structure of two-particle entanglement has been gained, it has become equally clear that the complexity and diversity of multiparticle entanglement grows exponentially with the number of particles. It is thus difficult to imagine a structurally simple theory that characterizes and quantifies all details of multiparticle entangled states. On the other hand one may expect that large numbers of particles admit a notion of typical entanglement properties for which a structurally simple theory may be developed. This intuition gives hope that significant progress can be made by restricting attention to entanglement properties that are typical (generic) relative to the uniform (unitarily invariant) measure, the unbiased distribution of pure states. In this setting it was demonstrated that typically pure states of large numbers of spins exhibit maximal bipartite [3–7] and multipartite entanglement [3]. This suggests that the exploration of the entanglement properties of generic states is a promising approach.

But a big question mark exists as to whether statements about generic states relative to the uniform measure are physically relevant. This is because the generation of a typical unitary requires a sequence of 2-qubit unitaries whose length grows exponentially in the number of qubits [8], even if one allows for a finite fixed fidelity. Thus achieving the uniform distribution to a fixed accuracy requires sequences of random 2-qubit unitaries that grow exponentially with the size of the system, and quickly becomes unphysical, see, e.g., [9–12]. One could then argue that the entanglement properties of generic quantum states are mathematically sound and interesting but physically irrelevant, as a system undergoing a randomization of its state through two-party interactions would only get close to the uniform measure in an unfeasibly long time. On the other hand, entanglement properties represent a restricted class of physical properties of a quantum state. Accordingly the faithful reproduction of generic entanglement properties may be possible with far fewer physical resources, i.e., 2-qubit gates, than those required for the generation of the expectation value for an arbitrary observable.

It is thus crucial to explore whether generic entanglement properties can be obtained efficiently, i.e., polynomially in the number of qubits, using only one- and two-qubit gates. The present work answers this question positively (cf. Fig. 1).

Our results support the physical relevance of the exploration of generic entanglement towards a structurally simple entanglement theory and have direct practical relevance since certain quantum information processing protocols such as [3,13–15] assume that generic entanglement can be generated efficiently.

The presentation proceeds as follows. We first define the key process that is used throughout this work: random two-qubit interactions, modeled as random circuits on a quantum computer. Then we prove that the generic entanglement average as well as the purity of a subsystem are achieved efficiently and that the so generated states are typically very close to maximally entangled. This is followed by numerical evidence that the achievement of generic entanglement can be associated with a specific time, the variation cutoff, for large systems. We finish with a discussion and conclusion.

The setting.—We consider a set of N qubits split into two subsets A (with N_A qubits) and B (with N_B qubits). Let $|\psi_0\rangle$ be an initial state in AB and consider a random circuit C_n consisting of n randomly chosen two-qubit quantum gates. Define $|\psi_n\rangle = C_n|\psi_0\rangle$ and the reduced density matrix $\rho_{A,n} = \mathrm{Tr}_B(|\psi_n\rangle\langle\psi_n|)$ of system A. Then the entangle-

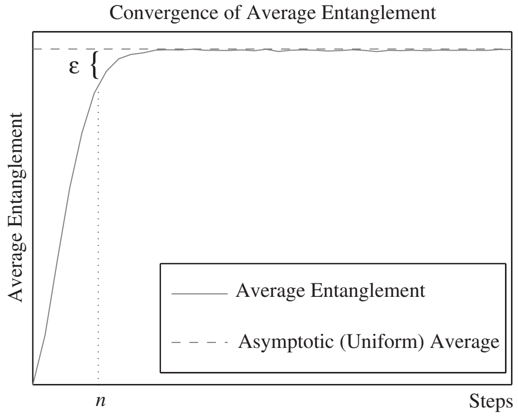

FIG. 1. Typical numerical simulation using the random circuit. The entanglement average of the uniform measure is reached to an accuracy ε in n steps. We prove it suffices with $n = O(N^3)$ to achieve a fixed arbitrary ε accuracy when increasing N.

ment in the state is given by $E(\psi_n) = S(\rho_{A,n})$ and its purity by $\mathrm{Tr}(\rho_{A,n}^2)$.

Definition of the random circuit.—The random circuit C_n is a product $W_n \ldots W_1$ of two-qubit gates where each W_i is independently chosen in the following way: A pair of distinct integers $c \neq t$ is chosen uniformly at random from $\{1, \ldots, N\}$. Next, single-qubit unitaries $U[c]$ and $V[t]$ acting on qubit c and t, respectively, are drawn independently from the uniform measure on $U(2)$. Then $W = \mathrm{CNOT}[c, t]U[c]V[t]$ where $\mathrm{CNOT}[c, t]$ is the controlled-NOT gate with control c and target t [16].

Asymptotics of random circuit.—The circuit, acting on N qubits, will asymptotically induce the uniform measure on states (see, e.g., [9]). In the above setting for states distributed according to the uniform measure the average bipartite entanglement can be found exactly [6] and is bounded from below such that $\mathbb{E}[E(\psi_n)] \geq N_A - \frac{1}{\ln 2}2^{-t}$ where $N_B - N_A = t \geq 0$ [3]. We find that the average purity of the subsystem A is $(2^{N_A} + 2^{N_B})/(2^N + 1)$ consistent with [11]. Furthermore, the distributions for entanglement and purity concentrate around their average with increasing N [3–6]. Thus one is overwhelmingly likely to find near-maximal entanglement for large systems.

Main theorem.—We will now be concerned with the approach to the asymptotic regime. For the above setting we prove that, independently of the initial state $|\psi_0\rangle$, convergence of the expected entanglement to its asymptotic value to an arbitrary fixed accuracy ε is achieved after a number of random two-qubit gates that is polynomial in the number of qubits. More precisely we find:

Theorem 1.—Suppose that $N_B - N_A = t \geq 0$ and that some arbitrary $\varepsilon \in (0, 1)$ is given. Then for a number n of gates in C_n satisfying

$$n \geq 9N(N-1)[(3\ln 2)N + \ln\varepsilon^{-1}]/4,$$

we have

$$\mathbb{E}[E(\psi_n)] \geq N_A - (2^{-t} + \varepsilon)/\ln 2 \tag{1}$$

and

$$\mathbb{E}[\max_{|\Psi\rangle_{AB}=\mathrm{maxent}} |\langle\psi_n|\Psi\rangle|] \geq 1 - \sqrt{\frac{2^{-t} + \epsilon}{2\ln 2}}. \tag{2}$$

Equation (2), follows from Eq. (1) employing $\sqrt{2S(\sigma\|\rho)} \geq \mathrm{Tr}|\sigma - \rho|_1$, where S is the relative entropy, and $\frac{1}{2}\mathrm{Tr}|\sigma - \rho|_1 \geq 1 - \mathrm{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}$ as well as Uhlmann's theorem [8]. To prove Eq. (1) we prove a lemma that considers the quantity $\mathbb{E}[\mathrm{Tr}(\rho_{A,n}^2)]$.

Lemma 1.—For arbitrary N, N_A, N_B and all n we have

$$\left|\mathbb{E}[\mathrm{Tr}(\rho_{A,n}^2)] - \frac{2^{N_A} + 2^{N_B}}{2^N + 1}\right| \leq 4^N e^{-(4n/9N(N-1))}.$$

To see that Lemma 1 implies Theorem 1 note first that $E(\psi_n) = S(\rho_{A,n}) \geq -\log_2 \mathrm{Tr}(\rho_{A,n}^2)$. By convexity we then find $\mathbb{E}[-\log_2 \mathrm{Tr}(\rho_{A,n}^2)] \geq -\log_2(\mathbb{E}[\mathrm{Tr}(\rho_{A,n}^2)])$ and a direct computation using $\ln(1 + x) \leq x$ for $x \geq 0$ completes the argument.

Proof of Lemma 1.—We proceed to outline the proof of Lemma 1 below, omitting some tedious but straightforward calculations to improve clarity. We begin with a useful representation of quantum states in terms of Pauli-operators. Indeed, $|\psi_n\rangle\langle\psi_n| = \sum_{p\in\{0,x,y,z\}^N} \xi_n(p) 2^{-N/2} \otimes_{i=1}^N \sigma^{p_i}[i]$, where each $\xi_n(p) = 2^{-N/2}\mathrm{Tr}(\otimes_{i=1}^N \sigma^{p_i}[i]|\psi_n\rangle\langle\psi_n|)$ and $\sigma^{p_i}[i]$ is a Pauli operator acting on qubit i [16]. Then for the reduced density operator $\rho_{A,n} = \mathrm{Tr}_B(|\psi_n\rangle\langle\psi_n|)$ we find

$$\mathbb{E}[\mathrm{Tr}(\rho_{A,n}^2)] = 2^{N_B} \sum_{\{p:\,\forall\, i\notin A, p_i=0\}} \mathbb{E}[\xi_n^2(p)]. \tag{3}$$

The main purpose will now be to analyze the evolution of the expected values of the squared coefficients, $\mathbb{E}[\xi_n^2(p)]$.

Evolution of the coefficients.—The key idea of the proof relies on the observation that the $\mathbb{E}[\xi_n^2(p)]$ form a probability distribution on $\{0, x, y, z\}^N$ for all n and that these probabilities evolve as a Markov chain with transition matrix P which takes q distributed according to $(\mathbb{E}[\xi_n^2(q)])_q$ in one step to p distributed according to $(\mathbb{E}[\xi_{n+1}^2(p)])_p$. To determine P we consider the action of a random unitary W_n at time n that acts on qubits c, t in state $|\psi_n\rangle$. This results in

$$\mathbb{E}[\xi_{n+1}^2(p)|\psi_n, c, t] = \frac{1}{16} \sum_{q,q'\in\{0,x,y,z\}^N:\,\forall\, i\notin\{c,t\}, q_i\neq q_i'} \xi_n(q)\xi_n(q')\mathbb{E}[\mathrm{Tr}[U\sigma^{q_c}[c]U^\dagger\sigma^{\hat{p}_c}[c]]\mathrm{Tr}[U\sigma^{q_c'}[c]U^\dagger\sigma^{\hat{p}_c}[c]]]$$
$$\times\, \mathbb{E}[\mathrm{Tr}[V\sigma^{q_t}[t]V^\dagger\sigma^{\hat{p}_t}[t]]\mathrm{Tr}[V\sigma^{q_t'}[t]V^\dagger\sigma^{\hat{p}_t}[t]]],$$

where the $(\hat{p}_c, \hat{p}_t)$ are uniquely determined by $\mathrm{CNOT}[c, t]\sigma^{p_c}[c]\sigma^{p_t}[t]\mathrm{CNOT}[c, t] = \pm\sigma^{\hat{p}_c}[c]\sigma^{\hat{p}_t}[t]$. Direct calculation with the uniform measure on $U(2)$ shows that the products of expectations in the sum vanish unless $q = q'$. Then with the

Kronecker symbol $\delta_{i,j}$ we find $\mathbb{E}[\xi^2_{n+1}(p)|\psi_n, c, t] = \sum_{q\in\{0,x,y,z\}^N} P^{(c,t)}(q, p)\prod_{i\notin\{c,t\}}\delta_{q_i,p_i}\xi^2_n(q)$ where $P^{(c,t)}(q, p) = 1$ if $\hat{p}_c = \hat{p}_t = q_c = q_t = 0$; $P^{(c,t)}(q, p) = 1/3$ if $\hat{p}_c = q_c = 0$ and $\hat{p}_t, q_t \neq 0$ or if $\hat{p}_t = q_t = 0$ and $\hat{p}_c, q_c \neq 0$; and $P^{(c,t)}(q, p) = 1/9$ otherwise. Averaging $P^{(c,t)}(q, p)\prod_{i\notin\{c,t\}}\delta_{q_i,p_i}$ over the $N(N-1)$ choices of c, t produces the entry $P(q, p)$ of the transition matrix of the desired Markov chain for $(\mathbb{E}[\xi^2_n(q)])_q$.

Simplifying the Markov chain.—Our aim is the evaluation of Eq. (3) and it turns out that this can be done via a simplified Markov chain. Consider $\{q(n) = (q(n)_1 q(n)_2 \ldots q(n)_N)\}_{n\geq 1}$ as an n-step evolution of our Markov chain P. Then the sets $S(n) = \{i \in \{1, \ldots, N\}: q(n)_i \neq 0\}$, identifying the nonzero elements of $q(n)$ also form a Markov chain. Using $S(n)$ in Eq. (3) we find

$$\mathbb{E}[\mathrm{Tr}(\rho^2_{A,n})] = 2^{N_B}\mathbb{P}(S(n) \subset A). \tag{4}$$

Thus we need only to consider the chain $\{S(n)\}_n$.

Convergence rate of the Markov chain.—As it turns out, our chain is almost ergodic: removing the isolated state $S(n) = \varnothing$, we obtain an ergodic chain on $\Omega = 2^{\{1,\ldots,N\}}\backslash\{\varnothing\}$. Since $\mathbb{P}(S(n) = \varnothing) = \mathbb{P}(S(0) = \varnothing) = 2^{-N}$, determining the convergence rate to the equilibrium of $S(n)$ on Ω given by $\mathcal{M}(S) = 3^{|S|}/(4^N - 1)$, $S \in \Omega$ is sufficient for our purposes. Let $Q = (Q(S, S'))_{S,S'\in\Omega}$ be the transition matrix of the restricted $S(n)$ chain. It has largest eigenvalue 1 whose eigenvector determines the steady state solution $\mathcal{M}$. The difference to the second largest eigenvalue, the spectral gap λ_Q, bounds the convergence rate to the steady state: for any initial distribution vector v, the component of $Q^n v$ orthogonal to $\mathcal{M}$ shrinks exponentially fast with $\lambda_Q n$. A quantitative result is provided in Chapter 2 of [17] (see Corollary 2.15): since our Q is a reversible chain with $Q(S, S) \geq \frac{1}{2}$ for all $S \in \Omega$, we obtain $|\mathbb{P}(S(n) \subset A) - \sum_{\varnothing\neq S\subset A}\mathcal{M}(S)| \leq e^{-\lambda_Q n}/\sqrt{\min_T \mathcal{M}(T)} \leq 2^N e^{-\lambda_Q n}$. We have $\sum_{\varnothing\neq S\subset A}\mathcal{M}(S) = (4^{N_A} - 1)/(4^N - 1)$. Putting back the isolated state $\varnothing$ into the calculations, applying (4) and noting that $2^{N_B} \leq 2^N$ yields $|\mathbb{E}[\mathrm{Tr}(\rho^2_{A,n})] - \frac{2^{N_A}+2^{N_B}}{2^N+1}| \leq 4^N e^{-\lambda_Q n}$ All that remains is to show that $\lambda_Q \geq 4/9N(N-1)$. We use a well-known variational principle for λ_Q [18]:

$$\lambda_Q = \inf\frac{\sum_{S,S'\in\Omega}\mathcal{M}(S)Q(S, S')(f(S) - f(S'))^2}{\sum_{T,T'\in\Omega}\mathcal{M}(T)\mathcal{M}(T')(f(T) - f(T'))^2}, \tag{5}$$

where the inf is taken over nonconstant $f: \Omega \to \mathbb{R}$. This is an application of Raleigh's principle to the second smallest eigenvalue of $I - Q$, which is precisely λ_Q. Equation (5) implies that if R is the transition matrix of a Markov chain on Ω with same stationary distribution $\mathcal{M}$ and $\alpha R(S, S') \leq Q(S, S')$ for all S, $S' \in \Omega$, then the gap λ_R of R satisfies $\lambda_Q \geq \alpha\lambda_R$. This allows us to estimate λ_Q by comparison with a simpler chain [19]. Indeed, our R will be the transition matrix of chain $\{B(n)\}$ on Ω defined as follows: Assume $B(n) = B$ and choose a $1 \leq j \leq N$ uniformly at random. If $j \in B$ and $|B| \geq 2$, set $B(n+1) = B\backslash\{j\}$ with probability $\frac{1}{3}$ and $B_{n+1} = B$ with probability $\frac{2}{3}$. If $j \in B$ and $|B| = 1$, do nothing. If $j \notin B$, set $B(n+1) = B \cup \{j\}$. This is a biased random walk on the hypercube $2^{\{1,\ldots,N\}}$ where transitions to state $\varnothing$ are suppressed. A coupling argument following, e.g., Chap. 4 of [20] shows that R has a spectral gap $\geq \frac{1}{3}N$. Moreover, one can check that $\alpha R(S, S') \leq Q(S, S')$ with $\alpha = \frac{4}{3}(N-1)$. It follows that $\lambda_Q \geq \alpha\lambda_R \geq \frac{4}{9}N(N-1)$, as desired, and the proof is finished. Numerics indicate convergence in approximately $N\ln N$ steps, so our bound is not tight.

Observe cutoff.—Many Markov chains exhibit the so called "cutoff effect"[21]. The cutoff refers to an abrupt approach to the stationary distribution occurring at a certain number of steps taken in the chain. Say we have a Markov chain defined by its transition matrix P, and that it converges to a stationary distribution π. Initially the total variation distance $\mathrm{TV} = ||P - \pi|| = \sup|P(E) - \pi(E)|$ between the corresponding probability distributions is given by $\mathrm{TV} = 1$. After k steps, $\mathrm{TV}(k) = |P^k - \pi|$. A cutoff occurs, basically, if $\mathrm{TV}(k) \simeq 1$ for $k = 0, 1, 2, \ldots a$ and thereafter falls quickly such that after a few steps $\mathrm{TV}(k) \simeq 0$. As we increase the size of the state space, the ratio of the number of steps during which the abrupt approach takes place and a should vanish asymptotically. Then we can say that the randomization occurs at a steps. Rigorously, this may be stated as follows [21]. Let P_n, π_n be Markov chains on sets χ_n. Let a_n, b_n be functions tending to infinity, with b_n/a_n tending to zero. Say the chains satisfy an a_n, b_n cutoff if for some starting states x_n and all fixed real θ with $k_n = \lfloor a_n + \theta b_n\rfloor$, then $||P_n^{k_n} - \pi_n|| \to c(\theta)$ with $c(\theta)$ a function tending to zero for θ tending to infinity and to 1 for θ tending to minus infinity. Here we observe this behavior in the entanglement distribution, a functional of the Markov chain on unitaries given by the random circuit, and we accordingly term this a cutoff.

Numerical observation.—Numerical simulations indicate a cutoff effect in the entanglement probability distribution under the random circuit on $|0\rangle^{\otimes N}$ may be observed both for single-qubit gates drawn from the uniform measure on $U(2)$ and for stabilizer gates; see Fig. 2.

The simulations using stabilizer gates allow us to consider far larger systems sizes. Here we choose the single-qubit gates U and V from the set $\{\sigma^x, \sigma^y, \sigma^z, S, H\}$ [16] with equal probability. It should be noted that the proof of Lemma 1 still holds (see [22] for details) and the entanglement behavior will remain similar [23,26]. The restriction allows us to use the efficient stabilizer formalism [24] and the tools developed in [25] which in turn allows for an efficient evaluation of state properties.

Extensions of the present result.—Similar methods can be used to address the mixed state setting through tracing out part of the system on which the random circuit is applied. Multipartite entanglement measures based on average purities [2] can be considered with the results established here. We also anticipate that one can use similar

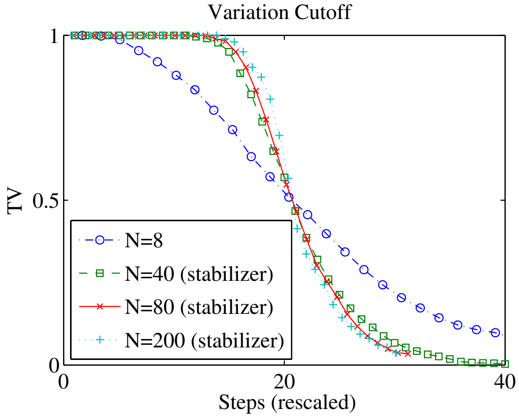

FIG. 2 (color online). Observe a variation cutoff of the entanglement probability distribution compared with that of the uniform measure as determined numerically. The state space has been discretised by rounding off entanglement values to the nearest integer. We observe that TV $\simeq 1$ for a while and then falls. Finally there is a stage where TV $\simeq 0$. The effect becomes more pronounced with increasing N. The results for $N > 8$ are done using the stabilizer random circuit.

techniques to obtain rigorous statements about the convergence rates of finite temperature Markov process quantum Monte Carlo simulations.

Our results may be applied to the protocols for superdense coding of quantum states presented in [3,13,15] to replace the inefficient process of creating random unitaries distributed according to the Haar measure by our efficient random circuits. After that replacement, Theorem 1 may be applied directly to verify that the main Lemma 1 of [15] still holds. It is an open question whether the performance of the protocols in [3,13,14] is adversely affected by this substitution. This cannot be decided on the basis of Theorem 1 alone but we expect that similar techniques as described here and in [22] will be able to decide this. The results of this work as well as the above extensions will be presented in detail in forthcoming publications.

Conclusion.—In this Letter we have proved that the average entanglement over the unitarily invariant measure is reached in a time that is polynomial in the size of the system by a quantum random process that is restricted to random two-qubit interactions. We also provided numerical evidence that for large systems the entanglement distribution of the uniform measure is achieved at a specific point in time, the variation cutoff. Our results demonstrate that the entanglement properties of generic entanglement are physical in the sense that they can be generated efficiently from random sequences of two-qubit gates. We have described extensions, including how this knowledge can be applied to render certain protocols efficient.

We gratefully acknowledge early discussions with J. Oppenheim and discussions with T. Rudolph, G. Smith, and J. Smolin. M. B. P. was funded by the EPSRC QIP-IRC, The Leverhulme Trust, EU Integrated Project QAP, and the Royal Society; O. D. by the IMS of Imperial College, and R. O. by NSA and ARDA through ARO Contract No. W911NF-04-C-0098.

*Electronic address: rob.oliv@gmail.com
†Electronic address: oscar.dahlsten@imperial.ac.uk
‡Electronic address: m.plenio@imperial.ac.uk

[1] M. Genovese, Phys. Rep. **413**, 319 (2005).
[2] M. B. Plenio and S. Virmani, Quantum Inf. Comput. **7**, 1 (2007).
[3] P. Hayden, D. W. Leung, and A. Winter, Commun. Math. Phys. **265**, 95 (2006).
[4] E. Lubkin, J. Math. Phys. (N.Y.) **19**, 1028 (1978).
[5] S. Lloyd and H. Pagels, Ann. Phys. (N.Y.) **188**, 186 (1988).
[6] D. N. Page, Phys. Rev. Lett. **71**, 1291 (1993); S. K. Foong and S. Kanno, Phys. Rev. Lett. **72**, 1148 (1994).
[7] O. C. O. Dahlsten and M. B. Plenio, Quantum Inf. Comput. **6**, 527 (2006).
[8] M. A. Nielsen and I. Chuang, *Quantum Information and Computation* (Cambridge University Press, Cambridge, 2000).
[9] J. Emerson, E. Livine, and S. Lloyd, Phys. Rev. A **72**, 060302 (2005).
[10] J. Emerson, Y. S. Weinstein, M. Saraceno, S. Lloyd, and D. G. Cory, Science **302**, 2098 (2003).
[11] G. Smith and D. W. Leung, Phys. Rev. A **74**, 062314 (2006).
[12] J. Emerson, AIP Conf. Proc. **734**, 139 (2004).
[13] A. Abeyesinghe, P. Hayden, G. Smith, and A. Winter, IEEE Trans. Inf. Theory **52**, 3635 (2006).
[14] H. Buhrman, M. Christandl, P. Hayden, H. K. Lo, and S. Wehner, Phys. Rev. Lett. **97**, 250501 (2006).
[15] A. Harrow, P. Hayden, and D. Leung, Phys. Rev. Lett. **92**, 187901 (2004).
[16] In the computational basis: $\sigma_x = [0,1;1,0]$, $\sigma_y = [0,-i;i,0]$, $\sigma_z = [1,0;0,-1]$, $S = [1,0;0,i]$, $H = [1,1;1,-1]/\sqrt{2}$, and CNOT $= [1,0,0,0;0,1,0,0;0,0,0,1;0,0,1,0]$.
[17] R. Montenegro and P. Tetali, *Series Foundations and Trends in Theoretical Computer Science* (NOW Publ., Boston, Delft, 2006), Vol. 1:3.
[18] See Lemma 2.21 in [17]. Our numerator is twice their Dirichlet form (Definition 2.1) and our denominator is twice the variance.
[19] P. Diaconis and L. Saloff-Coste, Ann. Appl. Prob. **3**, 696 (1993).
[20] D. A. Levin, Y. Peres, and E. L. Wilmer, "Markov Chains and Mixing Times" (to be published); see http://www.oberlin.edu/markov/.
[21] P. Diaconis, Proc. Natl. Acad. Sci. U.S.A. **93**, 1659 (1996).
[22] O. C. O. Dahlsten, R. Oliveira, and M. B. Plenio, quant-ph/0701125.
[23] C. Dankert, R. Cleve, J. Emerson, and E. Livine, quant-ph/0606161 which has since appeared gives further support for this restriction.
[24] D. Gottesman, quant-ph/9807006.
[25] K. M. R. Audenaert and M. B. Plenio, New J. Phys. **7**, 170 (2005); computer codes downloadable at www.imperial.ac.uk/quantuminformation.
[26] D. P. DiVincenzo, D. W. Leung, and B. M. Terhal, IEEE Trans. Inf. Theory **48**, 580 (2002).