# Convergence Rates for Arbitrary Statistical Moments of Random Quantum Circuits

Winton G. Brown and Lorenza Viola

*Department of Physics and Astronomy, Dartmouth College, 6127 Wilder Laboratory, Hanover, New Hampshire 03755, USA*

(Received 11 September 2009; revised manuscript received 21 May 2010; published 21 June 2010)

We consider a class of random quantum circuits where at each step a gate from a universal set is applied to a random pair of qubits, and determine how quickly averages of arbitrary finite-degree polynomials in the matrix elements of the resulting unitary converge to Haar measure averages. This is accomplished by mapping the superoperator that describes $t$ order moments on $n$ qubits to a multilevel $SU(4^t)$ Lipkin-Meshkov-Glick Hamiltonian. We show that, for arbitrary fixed $t$, the ground-state manifold is exactly spanned by factorized eigenstates and, under the assumption that a mean-field ansatz accurately describes the low-lying excitations, the spectral gap scales as $1/n$ in the thermodynamic limit. Our results imply that random quantum circuits yield an efficient implementation of $\epsilon$ approximate unitary $t$ designs.

DOI: 10.1103/PhysRevLett.104.250501 PACS numbers: 03.67.Ac, 05.30.−d, 05.40.−a

Random quantum states and unitary operators are broadly useful across theoretical physics and applied mathematics. Within quantum information science [1], they play a key role in tasks ranging from quantum data hiding [2] and quantum cryptography [3] to noise estimation in open quantum systems [4–6]. Unfortunately, generating an ensemble of $N$-dimensional unitary matrices which are evenly distributed according to the invariant Haar measure on $U(N)$ is inefficient, in the sense that the number of required quantum gates grows exponentially with the number of qubits, $n = \log_2 N$ [1]. So-called unitary $t$ designs provide a powerful substitute for Haar-distributed ensembles. Building on the notion of a state $t$ design [7], a unitary $t$ design is an ensemble of unitaries whose statistical moments up to order $t$ equal (exactly or approximately) Haar-induced values [6]. That is, a unitary $t$ design faithfully simulates the Haar measure with respect to any test that uses at most $t$ copies of a selected $n$-qubit unitary. Ramifications of the theory of $t$ designs [8] are being uncovered in problems as different as black hole evaporation and fast "scrambling" of information [9], efficient quantum tomography and randomized gate benchmarking [10], quantum channel capacity [11], and the foundations of quantum statistical mechanics [12].

Prompted by the above advances, significant effort has been devoted recently to identifying efficient constructions of $t$ designs and characterizing their convergence properties [2,6,13–16]. Harrow and Low established, in particular, the equivalence between approximate 2 designs and random quantum circuits as introduced in [4], and conjectured that a random circuit consisting of $k = \text{poly}(n, t)$ gates from a two-qubit universal gate set yields an approximate $t$ design [16]. While supporting numerical evidence was gathered in [17] for low-order moments, and efficient constructions of $t$ designs were reported in [16] for any $t = \mathcal{O}(n/\log n)$, the extent to which random quantum circuits could be used to implement an approximate $t$ design for arbitrary, fixed $t$ remained open.

In this Letter, we address this question by determining the rate at which, for sufficiently large circuit depth, statistical moments of arbitrary order converge to their limiting Haar values. Our strategy involves two steps: first, for given $t$, we show that the asymptotic convergence rate is determined by the spectral gap of a certain superoperator, which encapsulates moments up to order $t$; next, we compute this gap by mapping the $t$-moment superoperator to a multilevel version of the Lipkin-Meshkov-Glick (LMG) model, whose low-energy spectrum is well understood in the thermodynamic limit $n \to \infty$ [18]. Remarkably, the ground-state manifold may be exactly characterized for arbitrary $n$ and $t$, whereas obtaining the first excitation energy relies on a mean-field ansatz whose validity has been extensively tested for the class of infinitely coordinated models of interest. Our approach ties together $t$ design theory with established mean-field techniques from many-body physics, extending earlier results by Znidaric [14] for $t = 2$. Furthermore, asymptotic convergence rates allow us to upper bound the convergence time (minimum circuit length, $k_c$) needed for a desired accuracy $\epsilon$ relative to the Haar measure to be reached. For any fixed $t$, we find that the scaling $k_c \sim n \log(1/\epsilon)$ holds for sufficiently large $n$ and small $\epsilon$.

*Moment superoperator.*—Let a random quantum circuit of length $k$ be a sequence $U_k \dots U_1$ of $k$ unitary operators on an $n$-qubit Hilbert space $\mathcal{H} = \otimes_j^n \mathcal{H}_{q_j}$, where each $U_i$ is selected from an ensemble $\{\mu(U), U\}$, for a probability distribution $\mu$ with support on a universal gate set. To analyze $t$-order moments, we introduce a Hilbert space $\mathcal{H}_{M_t} = \mathcal{H}^{\otimes 2t}$, which consists of $2t$ copies of $\mathcal{H}$ and we refer to as the moment space, with $\dim(\mathcal{H}_{M_t}) \equiv D = N^{2t}$, and a local moment space $\mathcal{H}_{l_t}$, which results from grouping factors corresponding to the same qubit in $\mathcal{H}_{M_t}$. That is, $\mathcal{H}_{M_t} = \otimes_j^n \mathcal{H}_{q_j}^{\otimes 2t} = \mathcal{H}_{l_t}^{\otimes n}$, with $\dim(\mathcal{H}_{l_t}) \equiv d = 4^t$. Moments of order $t$ may be described in terms of the following linear operator on $\mathcal{H}_{M_t}$:

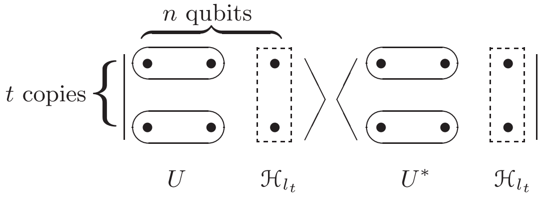

FIG. 1. The moment space $\mathcal{H}_{M_t}$ may be visualized as an array of $2nt$ qubits, such that $t$ copies support a ket in the state space on $nt$ qubits, and the remaining $t$ copies the corresponding bra. In this way, a unitary $U$ on $n$ qubits induces a transformation $U^{\otimes t,t}$ on density operators on $nt$ qubits. Dashed rectangles indicate the $2t$ qubits corresponding to a local moment space $\mathcal{H}_{l_t}$. Ovals correspond to a unitary $U$ acting only on the first two qubits.

$$M_t[\mu] = \int d\mu(U) U^{\otimes t} \otimes U^{*\otimes t} \equiv \int d\mu(U) U^{\otimes t,t}.$$

Physically, $M_t[\mu]$ may be viewed as the superoperator induced by the action of $t$ copies of the random circuit on $nt$-qubit density operators defined on $\mathcal{H}^{\otimes t}$ (see also Fig. 1). In line with standard practice in open-system theory [19], we shall introduce "operator kets" in $\mathcal{H}_{M_t}$, denoted $|A\rangle\rangle \equiv A$, and, correspondingly, $\langle\langle A| = A^\dagger$. Thus, a $D^2$-dimensional operator ket transforms according to $(U \otimes U^*)|A\rangle\rangle \equiv UAU^\dagger$, under $U \in U(D)$. Once a basis for $\mathcal{H}_{M_t}$ is chosen, the matrix representation of $M_t[\mu]$ specifies a complete set of $t$-order moments.

The probability distribution that describes a random circuit of length $k$, $\mu_k(U)$, is given by the $k$th fold convolution of $\mu$ with itself [5,16]. That is, $\mu_k(U) = \int \prod_{i=1}^k d\mu(U_i)\delta(U - \Pi_{i=1}^k U_i)$. It then follows that

$$M_t[\mu_k] = \int \Pi_{i=1}^k d\mu(U_i) \Pi_{i=1}^k U_i^{\otimes t,t} = \prod_{i=1}^k \int d\mu(U_i) U_i^{\otimes t,t}$$
$$= (M_t[\mu])^k \equiv M_t^k[\mu].$$

Note that, under the assumption that the ensemble $\{\mu(U), U\}$ is invariant under Hermitian conjugation, $\mu(U) = \mu(U^\dagger)$, $M_t[\mu]$ is an Hermitian operator on $\mathcal{H}_{M_t}$.

If $\mu(U)$ has support on a universal set of gates, then the measure over the random circuit converges to the Haar measure on $U(N)$ in the limit of infinite circuit length [5], $\lim_{k\to\infty}\mu_k(U) = \mu_H(U)$. We begin by characterizing how these convergence properties translate in terms of $t$-order moments. Let $M_t[\mu_H] = \int d\mu_H(U) U^{\otimes t,t}$, and let

$$\mathcal{V}_t = \text{span}\{|\phi\rangle\rangle \in \mathcal{H}_{M_t} | U^{\otimes t,t}|\phi\rangle\rangle = |\phi\rangle\rangle,\ \forall U \in U(N)\}$$

be the subspace of fixed points of $U^{\otimes t,t}$, $U \in U(N)$, with $\mathcal{P}_{\mathcal{V}_t}$ denoting the corresponding projector. We claim that

$$\lim_{k\to\infty}(M_t[\mu])^k = \mathcal{P}_{\mathcal{V}_t} = M_t[\mu_H], \quad \forall t. \tag{1}$$

While this is implied by the results in [16], a self-contained proof follows. Let $|\phi\rangle\rangle$ be an eigenoperator of $M_t[\mu]$ with eigenvalue $\lambda$, and $|\phi_U\rangle\rangle \equiv U^{\otimes t,t}|\phi\rangle\rangle$. Since $|\langle\langle M_t[\mu]\rangle\rangle| \leq \int d\mu(U)|\langle\langle\phi|\phi_U\rangle\rangle|$, it follows that $|\lambda| \leq 1$, with equality holding if and only if $U^{\otimes t,t}|\phi\rangle\rangle = |\phi\rangle\rangle$ for all $U$ with $\mu(U) \neq 0$. Any such operator ket $|\phi\rangle\rangle$ is also invariant under any unitary of the form $U^{\otimes t,t}$, where $U$ is generated by a random circuit of arbitrary length, that is, $U = \Pi_{i=1}^k U_i$ for any $k$, as long as $\mu(U_i) \neq 0$. Thus, if $\mu(U)$ has support on a universal gate set, the eigenspace of eigenvalue 1 is precisely $\mathcal{V}_t$. Since all other eigenvalues of $M_t[\mu]$ have magnitude less than 1, $M_t^k[\mu]$ converges to $\mathcal{P}_{\mathcal{V}_t}$. To establish the second equality in Eq. (1), we invoke the invariance of the Haar measure under $U(N)$, $\mu_H(U) = \mu_H(U'U)$. For $|\phi\rangle\rangle$ an eigenoperator of $M_t[\mu_H]$ with eigenvalue $\lambda$, it follows that $M_t[\mu_H]|\phi\rangle\rangle = \int d\mu_H(U) U^{\otimes t,t}|\phi\rangle\rangle = \lambda|\phi\rangle\rangle$. Thus,

$$\begin{aligned} U'^{\otimes t,t}\lambda|\phi\rangle\rangle &= \int d\mu_H(U)(U'U)^{\otimes t,t}|\phi\rangle\rangle \\ &= \int d\mu_H(U'^\dagger U) U^{\otimes t,t}|\phi\rangle\rangle \\ &= \int d\mu_H(U) U^{\otimes t,t}|\phi\rangle\rangle = \lambda|\phi\rangle\rangle. \end{aligned}$$

If $\lambda \neq 0$, it follows that $|\phi\rangle\rangle \in \mathcal{V}_t$, otherwise $\lambda = 0$, which establishes the desired result.

Our next goal is to obtain the rate at which $M_t^k[\mu]$ approaches $M_t[\mu_H]$. Since $M_t[\mu_H]$ projects onto the eigenspace of $M_t[\mu]$ of eigenvalue 1, the distance $\|M_t^k[\mu] - M_t[\mu_H]\|$ with respect to any norm depends only on the remaining eigenvalues $\{\lambda_i\}$ of $M_t[\mu]$ and the corresponding eigenprojectors $\{\Pi_i\}$. Specifically, if $k$ is sufficiently large, $\|M_t^k[\mu] - M_t[\mu_H]\| = \|\sum_{\lambda_i\neq 1}\lambda_i^k\Pi_i\| \approx |\lambda_1|^k\|\Pi_1\|$, where $\lambda_1 \equiv 1 - \Delta_t$ is the subdominant eigenvalue of $M_t[\mu]$. Thus, the asymptotic convergence rate is entirely determined by the spectral gap $\Delta_t$ of $M_t[\mu]$.

*Spectral gap determination.*—The starting point for mapping $M_t[\mu]$ to a multilevel LMG model is to ensure that the following conditions are obeyed: (i) The applied quantum gates consist only of single- and two-qubit gates selected according to a distribution $\tilde{\mu}(U)$ on $U(4)$, with $\tilde{\mu}(U) = \tilde{\mu}(U^\dagger)$; (ii) the target pair of qubits is picked uniformly at random. We shall refer to the class of circuits obeying (i)–(ii) as permutationally invariant random quantum circuits. Since, in each application of a random gate $U$ to a fixed pair of qubits, $U^{\otimes t,t}$ acts nontrivially only on the associated bilocal moment space $\mathcal{H}_{l_t} \otimes \mathcal{H}_{l_t}$, and each pair is equally likely to be chosen, the moment superoperator may be written as follows:

$$M_t[\mu] = \frac{2}{n(n-1)} \sum_{i<j=1}^{n} m_t^{ij}[\tilde{\mu}], \tag{2}$$

where for any pair $i$, $j$ the restriction $m_t[\tilde{\mu}]$ of $m_t^{ij}[\tilde{\mu}]$ to $\mathcal{H}_{l_t} \otimes \mathcal{H}_{l_t}$ acts as $m_t[\tilde{\mu}] = \int d\tilde{\mu}(U) U^{\otimes t,t}$. Recalling that $\dim(\mathcal{H}_{l_t}) = d$, $M_t[\mu]$ thus defines a qudit Hamiltonian, which is invariant under the symmetric group $\mathcal{S}_n$ of permutations of the $n$ local moment spaces. Explicitly, if $\{b_{\alpha\beta}^i = |\alpha\rangle\rangle\langle\langle\beta|\}$ denotes an outer-product basis for operators acting on any $\mathcal{H}_{l_t}$, we may expand $m_t^{ij} = \sum_{\alpha\beta\gamma\delta=1}^d \langle\langle\alpha\gamma|m_t|\beta\delta\rangle\rangle b_{\alpha\beta}^i b_{\gamma\delta}^j \equiv \sum_{\alpha\beta\gamma\delta=1}^d c_{\alpha\beta\gamma\delta} b_{\alpha\beta}^i b_{\gamma\delta}^j$, and rewrite $M_t[\mu]$ as a quadratic function of the collective operators $B_{\alpha\beta} = \sum_{i=1}^{n} b_{\alpha\beta}^{i}$, that is, $M_t[\mu] = \frac{1}{n(n-1)} \times \sum_{\alpha\beta\gamma\delta=1}^{d} c_{\alpha\beta\gamma\delta}(B_{\alpha\beta}B_{\gamma\delta} - \delta_{\beta\gamma}B_{\alpha\delta})$. Since the operators $B_{\alpha\beta}$ obey $SU(d)$ commutation rules, $[B_{\alpha\beta}, B_{\gamma\delta}] = B_{\alpha\delta}\delta_{\beta\gamma} - B_{\beta\gamma,}\delta_{\alpha\beta}$, $\mathcal{H}_{M_t}$ carries the (reducible) collective $n$-fold tensor product representation of $SU(d)$, and $M_t[\mu]$ provides a $d$-level extension of the standard, spin-1/2 LMG model [20].

Thanks to the invariance under $\mathcal{S}_n$, each of the eigenoperators of $M_t[\mu]$ belongs to an irreducible representation (irrep) of $SU(d)$. Our first step is to show that the eigenspace $\mathcal{V}_t$ of $M_t[\mu]$ corresponding to the ground-state (extremal) eigenvalue of 1 lies in the totally symmetric irrep, of dimension $d_S = \binom{4^t+n-1}{n}$ [21]. Recall that $\mathcal{V}_t$ consists of operators in $\mathcal{H}_{M_t}$ that commute with all $t$-fold tensor power unitaries $U^{\otimes t}$. By Schur-Weyl duality [16,21], every such operator is a linear combinations of elements of $\mathcal{S}_t$, under the natural representation in $\mathcal{H}^{\otimes t}$. Note that the operators spanning $\mathcal{V}_t$ are permutations of the $t$ copies of $\mathcal{H}$ rather than permutations of the $n$ local moment spaces $\mathcal{H}_{l_t}$. One may write any such permutation as $|\sigma^{(n)}\rangle\rangle = \sum_{i_1\ldots i_t=1}^{N} |i_1 \ldots i_t\rangle\langle i_{\sigma(1)} \ldots i_{\sigma(t)}|$, where $\sigma \in \mathcal{S}_t$. Furthermore, each such permutation may be viewed as a product ket relative to the factorization $\mathcal{H}_{M_t} = \mathcal{H}_{l_t}^{\otimes n}$. Explicitly, $|\sigma^{(n)}\rangle\rangle = (|\sigma\rangle\rangle)^{\otimes n}$, where $|\sigma\rangle\rangle = \sum_{i_1\ldots i_t=0,1} |i_1 \ldots i_t\rangle\langle i_{\sigma(1)} \ldots i_{\sigma(t)}| \in \mathcal{H}_{l_t}$. The fact that $\mathcal{V}_t$ is exactly spanned by product states is significant from the perspective of mean-field theory. For arbitrary $SU(d)$ quadratic Hamiltonians, it has been rigorously established that the exact ground-state energy is given in the thermodynamic limit by a mean-field ansatz equivalent to assuming that the ground-state is an $SU(d)$ coherent state [20]. Since, for the completely symmetric irrep, the manifold of coherent states consists precisely of all product states [22], the mean-field extremal eigenspace of $M_t[\mu]$ is, in fact, exact for any finite $n$.

The next step is to determine the lowest excitation energy in the large-$n$ limit, which is accomplished by expanding $M_t[\mu]$ around an arbitrary extremal mean-field state for each irrep. For the totally symmetric irrep, the required diagonalization procedure is most straightforwardly carried out by realizing the $U(d)$ algebra in terms of $d$ canonical Schwinger boson operators $\{a_\alpha, a_\beta^\dagger\}$ [23]. That is, we let $B_{\alpha\beta} = a_\alpha^\dagger a_\beta$ and rewrite $M_t[\mu] = \frac{1}{n(n-1)} \times \sum_{\alpha\beta\gamma\delta=1}^{d} c_{\alpha\beta\gamma\delta} a_\alpha^\dagger a_\gamma^\dagger a_\beta a_\delta$. Since the totally symmetric irrep of $\mathcal{H}_{M_t}$ contains exactly $n$ Schwinger bosons, it is possible to eliminate one boson mode by regarding it as "frozen" in the vacuum for a generalized Holstein-Primakoff transformation [23]. Specifically, let the local basis be chosen so that the frozen mode corresponds to $|\sigma\rangle\rangle$, and let $\theta(n) \equiv (n - \sum_{\alpha\neq\sigma} a_\alpha^\dagger a_\alpha)^{1/2}$, with $a_\sigma^\dagger \rightarrow \theta(n)$, $a_\sigma \rightarrow \theta(n)$. Two simplifications may now be invoked: first, the fact that $|\sigma^{(n)}\rangle\rangle$ is an exact ground-state causes any coefficient of the form $c_{\alpha\sigma\beta\sigma}$, $c_{\alpha\sigma\sigma\sigma}$ (and their complex conjugates) to vanish; second, only terms up to the leading order in $1/n$ need to be kept in $\theta(n)$. This finally yields $M_t[\mu] = 1 - \frac{1}{n}\sum_{\alpha\beta=1}^{d} E_{\alpha\beta} a_\alpha^\dagger a_\beta + \mathcal{O}(1/n^2)$, where $E_{\alpha\beta} = 2(\delta_{\alpha\beta} - \langle\langle\sigma\alpha|m_t|\sigma\beta\rangle\rangle - \langle\langle\sigma\alpha|m_t|\beta\sigma\rangle\rangle)$. To leading order, the desired gap is then determined by the smallest eigenvalue, $a_1$, of $E_{\alpha\beta}$. That the latter is nonzero may be shown by exploiting basic properties of the superoperator $m_t$ [24]. That no other excitation with a larger eigenvalue may exist, which is not captured by the $1/n$ expansion, has not been rigorously justified to the best of our knowledge—although, for LMG Hamiltonians, this is supported by an extensive body of theoretical and numerical investigations [18] (see also [24]). Subject to this conjecture, our main result follows: For any permutationally invariant random quantum circuit, and for any fixed $t > 0$, the spectral gap may be expanded as

$$\Delta_t = \sum_{p=1}^{\infty} a_p n^{-p} = \frac{a_1}{n} + \mathcal{O}\left(\frac{1}{n^2}\right), \qquad (3)$$

for coefficients $\{a_p\}$ that may in general depend on $t$. A stronger result may be obtained for a subclass of random quantum circuits which are, in addition, locally invariant, that is, $\tilde{\mu}(U)$ is invariant under the subgroup $U(2) \times U(2) \subset U(4)$ of local unitary transformations on the two target qubits. In this case, it is possible to choose a basis for each local moment space $\mathcal{H}_{l_t}$, which includes a maximal set of $t$-qubit operators $\{|\omega\rangle\rangle\}$ in the commutant of $U^{\otimes t}$, with $U \in U(2)$. Accordingly, every matrix element $\langle\langle\alpha\beta|m_t|\gamma\delta\rangle\rangle = 0$, unless each local basis element is itself an invariant, and the large-$n$ behavior of the gap is determined by matrix elements of the form $\langle\langle\sigma\omega|m_t|\sigma\omega\rangle\rangle$ and $\langle\langle\sigma\omega|m_t|\omega\sigma\rangle\rangle$, with $\sigma \in \mathcal{S}_t$ (without loss of generality, we may choose $|\sigma\rangle\rangle = |I\rangle\rangle$) and $|\omega\rangle\rangle$ an arbitrary $U(2)$ invariant with $\langle\langle\sigma|\omega\rangle\rangle = 0$. Since, for $t > 1$, the maximum value of any such matrix element is independent of $t$ (see [24] for full details), it follows that the leading order term $a_1$ does not depend on $t$ for locally invariant random quantum circuits.

*Example.*—Consider the simplest case where $t = 2$ and $\tilde{\mu}(U) = \mu_H(U)$ on $U(4)$. The invariant eigenspace $\mathcal{V}_2$ of $M_2$ is spanned by the identity $|I^{(n)}\rangle\rangle = (|I\rangle\rangle)^{\otimes n}$ and the permutation $|S^{(n)}\rangle\rangle = (|S\rangle\rangle)^{\otimes n}$ that swaps the $t = 2$ copies of $\mathcal{H} = \mathcal{H}_q^{\otimes n}$. Since $\tilde{\mu}(U)$ is the Haar measure, $m_2$ is the projector onto the subspace $\mathcal{V}_2$ for $n = 2$ qubits. To find the excitation energies, we choose one of the extremal local kets, $|I\rangle\rangle$, and minimize $E_{\min} = 2\min(1 - \langle\langle I\alpha|m_2|I\alpha\rangle\rangle - \langle\langle I\alpha|m_2|\alpha I\rangle\rangle)$ over all local operators $|\alpha\rangle\rangle \in \mathcal{H}_{l_t}$ orthogonal to $|I\rangle\rangle$. This yields $|\alpha\rangle\rangle = |S\rangle\rangle - \langle\langle S|I\rangle\rangle|I\rangle\rangle = \vec{\sigma}^1 \cdot \vec{\sigma}^2$, and $\Delta_t = 6/5n + \mathcal{O}(1/n^2)$. To determine how quickly the large-$n$ scaling sets in, the fully symmetric sector of $M_t[\mu]$ under $\mathcal{S}_n$ was numerically diagonalized. Since $\mu_H(U)$ is invariant under $U(2) \times U(2)$ transformations, $\mathcal{H}_{l_2}$ may be restricted to the subspace of $SU(2)$ invariants. From angular momentum theory [21], the number of such invariants is $\sum_J m_J^2 = \frac{(2t)!}{(t+1)!t!} = C_t$, where $m_J$ is

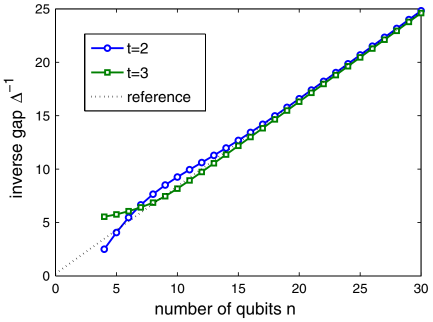

FIG. 2 (color online). Inverse spectral gap $\Delta_t^{-1}$ of $M_t[\mu_H]$ with $t = 2, 3$ for a random circuit consisting of two-qubit gates selected according to the Haar measure on $U(4)$. The line with slope 5/6 corresponds to the asymptotic result.

the multiplicity of the $SU(2)$ irrep with angular momentum $J$. This yields $d_S^{\rm loc} = \binom{C_t+n-1}{n} \ll d_S$, which makes numerical comparisons tractable for small $t$. Exact results for $t = 2$ and 3 (see Fig. 2) indicate that the scaling prediction for $\Delta_t$ becomes very accurate for $n \gtrsim 14$.

*Convergence time.*—In order to establish the usefulness of a random circuit as an $\epsilon$ approximate unitary $t$ design [6,16], we need to upper bound the circuit length required to achieve a specified accuracy $\epsilon$. Let the convergence time with respect to a given norm be defined by the minimum length $k_c$ for which $\|M_t^{k_c}[\mu] - M_t[\mu_H]\| \leq \epsilon$. That $M_t[\mu]$ be operationally indistinguishable from $M_t[\mu_H]$ requires that the supremum of $\|(M_t^k[\mu] - M_t[\mu_H])(\rho)\|_1$ be sufficiently small over all $nt$-qubit density operators $\rho$. We may bound the 1 norm starting from the 2 norm [16]. For any density matrix $\rho$, $\|(M_t^k[\mu] - M_t[\mu_H])(\rho)\|_2 \leq \lambda_1^k$. This follows from normalization of $\rho$ and the fact that $M_t[\mu_H]$ projects onto the eigenspace of eigenvalue 1 of $M_t[\mu]$. In conjunction with the Cauchy-Schwartz inequality, this implies $\|(M_t^k[\mu] - M_t[\mu_H])(\rho)\|_1 \leq 2^{nt}\lambda_1^k$. Requiring that $2^{nt}\lambda_1^{k_c} \leq \epsilon$ finally yields $k_c \leq \Delta_t^{-1}(\log(1/\epsilon) + nt\log(2))$. Since, using Eq. (3), $\Delta_t^{-1} = \sum_{p=1}^{\infty} a'_p n^{2-p} \sim a_1^{-1} n$ to leading order, $k_c = a_1^{-1} n \log(1/\epsilon)$ for sufficiently small $\epsilon$. It is important to stress that since the coefficients, $a_i$, for $i > 1$, may scale as polynomials in $d = 4^t$, the above result for $k_c$ is required to hold only for $t = \mathcal{O}(\log n)$. While improving this asymptotic bound is important for fully characterizing random circuits, our results are directly relevant to physical applications, where $t$ is fixed.

In summary, we have shown that, subject to a well-supported mean-field ansatz, a large class of random quantum circuits are efficient $\epsilon$ approximate unitary $t$ designs for arbitrary finite $t$. The fact that the extremal eigenoperators are separable suggests that similar results may apply to more general random circuits for which the Hermiticity and/or the $\mathcal{S}_n$-invariance assumptions of the moment superoperator need not hold [25]. A remaining open question is to determine how the circuit length scales as the limits of large $n$ and large $t$ are taken together. This may resolve the apparent paradox that while Haar random unitaries are inefficient, arbitrary $t$ designs are not, possibly with equal asymptotic rates.

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

[2] D. P. DiVincenzo, D. W. Leung, and B. Terhal, IEEE Trans. Inf. Theory **48**, 580 (2002).

[3] A. Harrow, P. Hayden, and D. W. Leung, Phys. Rev. Lett. **92**, 187901 (2004); P. Hayden *et al.*, Commun. Math. Phys. **250**, 371 (2004); A. Ambainis and A. Smith, Lect. Notes Comput. Sci. **3122**, 249 (2004).

[4] J. Emerson *et al.*, Science **302**, 2098 (2003).

[5] J. Emerson, E. Livine, and S. Lloyd, Phys. Rev. A **72**, 060302(R) (2005).

[6] C. Dankert *et al.*, Phys. Rev. A **80**, 012304 (2009).

[7] A. Ambainis and J. Emerson, in *Proceedings of the IEEE Conference on Computational Complexity* (IEEE, New York, 2007), p. 129.

[8] D. Gross, K. Audenaert, and J. Eisert, J. Math. Phys. (N.Y.) **48**, 052104 (2007).

[9] J. Preskill and P. Hayden, J. High Energy Phys. 09 (2007) 120; Y. Sekino and L. Susskind, *ibid.* 10 (2008) 065.

[10] A. Bendersky, F. Pastawski, and J. P. Paz, Phys. Rev. A **80**, 032116 (2009); E. Magesan, R. Blume-Kohout, and J. Emerson, arXiv:0910.1315.

[11] M. B. Hastings, Nature Phys. **5**, 255 (2009).

[12] R. A. Low, arXiv:0903.5236.

[13] R. Oliveira, O. C. Dahlsten, and M. B. Plenio, Phys. Rev. Lett. **98**, 130502 (2007); O. C. O. Dahlsten, R. Oliveira, and M. B. Plenio, J. Phys. A **40**, 8081 (2007).

[14] M. Znidaric, Phys. Rev. A **76**, 012318 (2007); **78**, 032324 (2008).

[15] W. G. Brown, Y. S. Weinstein, and L. Viola, Phys. Rev. A **77**, 040303(R) (2008); Y. S. Weinstein, W. G. Brown, and L. Viola, *ibid.* **78**, 052332 (2008).

[16] A. W. Harrow and R. Low, Commun. Math. Phys. **291**, 257 (2009); A. W. Harrow and R. Low, arXiv:0811.2597.

[17] L. Arnaud and D. Braun, Phys. Rev. A **78**, 062329 (2008).

[18] G. Ortiz *et al.*, Nucl. Phys. **B707**, 421 (2005); S. Dusuel and J. Vidal, Phys. Rev. B **71**, 224420 (2005); F. Leyvraz and W. D. Heiss, Phys. Rev. Lett. **95**, 050402 (2005); P. Ribeiro, J. Vidal, and R. Mosseri, *ibid.* **99**, 050402 (2007). While the standard two-level LMG model is addressed in these papers, none of the derivations depends specifically on the subsystems' dimension.

[19] R. Alicki and K. Lendi, *Quantum Dynamical Semigroups and Applications* (Springer, Berlin, 1987).

[20] R. Gilmore, J. Math. Phys. (N.Y.) **20**, 891 (1979).

[21] J. P. Elliott and P. G. Dawber, *Symmetry in Physics* (Oxford University Press, New York, 1979), Vol. 2.

[22] W. Zhang, D. Feng, and R. Gilmore, Rev. Mod. Phys. **62**, 867 (1990).

[23] S. Okubo, J. Math. Phys. (N.Y.) **16**, 528 (1975).

[24] See supplementary material at http://link.aps.org/supplemental/10.1103/PhysRevLett.104.250501 for additional technical details.

[25] W. G. Brown and L. Viola (to be published).