Efficiency of producing random unitary matrices with quantum circuits

Ludovic Arnaud and Daniel Braun
Laboratoire de Physique Théorique, IRSAMC, UPS, CNRS, Université de Toulouse, F-31062 Toulouse, France
(Received 4 July 2008; published 17 December 2008)

We study the scaling of the convergence of several statistical properties of a recently introduced random unitary circuit ensemble towards their limits given by the circular unitary ensemble. Our study includes the full distribution of the absolute square of a matrix element, moments of that distribution up to order eight, as well as correlators containing up to 16 matrix elements in a given column of the unitary matrices. Our numerical scaling analysis shows that all of these quantities can be reproduced efficiently, with a number of random gates which scales at most as $n_q[\ln(n_q/\epsilon)]^\nu$ with the number of qubits n_q for a given fixed precision ϵ and $\nu>0$.

DOI: 10.1103/PhysRevA.78.062329 PACS number(s): 03.67.Ac, 03.67.Pp

I. INTRODUCTION

Random unitary matrices play an important role in many tasks of quantum-information processing, including quantum data hiding and quantum encryption [1,2], quantum state distinction [3], superdense coding of quantum states [4], and noise estimation [5]. In these applications, a random ensemble of $N\times N$ matrices U drawn uniformly from the Haar measure of the unitary group, the so-called circular unitary ensemble (CUE), is required [6]. In principle, any unitary matrix acting on vectors in the Hilbert space of dimension $N=2^{n_q}$ of n_q qubits can be approximated with arbitrary precision using a computationally universal set of quantum gates that act on one or two qubits at a time [7–11]. However, as simple parameter counting quickly confirms, the required number of quantum gates n_g grows typically exponentially with the number of qubits. Indeed, $O(N^2(\ln N)^3)$ gates are required to approximate all matrix elements of U using a fixed universal gate set [12]. This suggests that the construction of sets of random unitary matrices which are evenly distributed according to the Haar measure of the unitary group are highly inefficient. One explicit but inefficient procedure of constructing matrices drawn from CUE is based on the Hurwitz parametrization (see [13]).

In a seminal paper, Emerson *et al.* introduced the concept of pseudorandom unitary operators, i.e., random unitary operators which are drawn from a distribution that mimics a uniform distribution with respect to the Haar measure of the unitary group [14]. The construction of these operators was motivated by ideas from quantum chaos, and used a random quantum circuit consisting of random U(2) rotations on each qubit followed by two qubit gates that implement an Ising spin interaction between nearest neighbors. They showed that this circuit produced unitary matrices with a distribution of matrix elements which converges exponentially with the number of quantum gates to the well-known distribution of matrix elements of CUE [15]. Later, Emerson, Livine, and Lloyd showed that the joint distribution function of matrix elements of a product of unitary operators created by a random quantum circuit composed of a continuous or discrete universal gate set converges uniformly and exponentially with the number of quantum gates to the Haar measure on the unitary group, albeit with a rate which itself decreases exponentially with the number of qubits [5]. This left open the question of the efficiency of the creation of the pseudorandom unitary operators in the sense of the scaling of the number of gate operations with the number of qubits. Furthermore, the distribution $P_{ij}(U_{ij})$ of matrix elements U_{ij} contains only a small amount of information compared to the full joint distribution of matrix elements. Notably it is unclear how fast correlators of matrix elements would converge to the CUE values.

Quite different statistics have been studied so far for different random circuit ensembles. In [16] the question of the efficient generation of typical bipartite entanglement between two subsystems was addressed numerically for a quantum circuit composed of U(4) gates, each of which was a product of a fixed 2-qubit gate and two random single-qubit gates drawn uniformly from the Haar measure of U(2). Exponential convergence to the CUE value with a rate that depends on $n_q \ln n_q$ as the number of qubits was found. Oliveira *et al.* introduced the technique of Markov chain analysis to study the same question and were able to prove an upper bound of $O(n_q^3)$ quantum gates necessary to reach a given (absolute) precision ϵ for the average amount of bipartite entanglement [17,18]. Average gate fidelity was studied in [19]. The distributions of differences between nearest-neighbor eigenphases as well as the distribution of the amount of interference was studied in [20] for the same random unitary circuit ensemble as the one we will use here (see below). Exponential or even Gaussian convergence was observed, but the question of the efficiency remained open. Quantum pseudorandomness from cluster-state quantum computation was introduced in [21].

The study of pseudorandom unitaries is closely related to the theory of unitary k-designs. Dankert *et al.* defined a unitary k-design as a discrete set of unitary matrices such that the average of any polynomial of degree equal or smaller than k in the complex matrix elements of U over the set equals the average of that polynomial over the unitary group [22]. Harrow *et al.* showed that a random circuit of length polynomial in n_q yields an ϵ-approximate 2-design. Depending on the gate set used, the number of gates n_g needed to achieve a given precision ϵ scales as $O(n_q(n_q+\ln 1/\epsilon))$ or as $O(n_q \ln(n_q/\epsilon))$. They also conjectured that a random circuit on n_q qubits composed of poly (n_q,k) random 2-qubit gates chosen from a universal gate set is an ϵ-approximate k-design [23]. In [22,24] approximate k-designs were studied

in the context of twirling. Originally, unitary designs were defined for a fixed set of unitary matrices, each of which comes with the same weight (see [25] for an insightful discussion of their mathematical structure). The definition in [23] naturally extends the concept of unitary designs to probability distributions over sets of unitary matrices, such that each random unitary matrix corresponds to a realization of the random quantum circuit, and averaging over the random circuits realizes the average over the unitary design.

The results from [23] imply that random quantum circuits can efficiently [with $O(n_q \ln(n_q/\epsilon))$ gates] reproduce the CUE averages of $|U_{ij}|^2$, $|U_{ij}|^4$, and $|U_{ij}U_{kl}|^2$. Here ϵ was taken as bound on the completely bound norm of the differences between the matrices averaged over the circuit ensemble or over CUE. Note that CUE averages of unpaired matrix elements vanish [26]. Since entanglement fidelity and gate fidelity can be expressed as averages over polynomials of order (2,2) in U_{ij} and U^*_{kl}, this result confirms the numerical finding in [16].

In this paper we study numerically the efficiency with which the random unitary circuit ensemble (UCE) introduced in [20] (see below for its definition) reproduces various statistical properties of the matrix elements of CUE matrices. Our study includes the full distribution of the (absolute square of) matrix elements, moments of that distribution up to order $|U_{ij}|^{16}$, as well as correlators containing up to 16 matrix elements of a given column of the unitary matrix. Within the range of numerically accessible sizes of the quantum circuits (up to 28 qubits for the distribution of a single matrix element, down to 15 qubits for the correlators), our results show that, surprisingly, the number of gates required to reach a given precision ϵ for all of these quantities grows no faster than $n_q[\ln(n_q/\epsilon)]^\nu$ with $\nu>0$, indicating that the statistical properties of CUE which require an exponential number of quantum gates in order to be well approximated by a random quantum circuit, must be of a more sophisticated nature.

II. CONVERGENCE OF UCE TO CUE

The unitary circuit ensemble (UCE) introduced in [20] consists of quantum algorithms which use two kinds of quantum gates: U(2) gates which act on single qubits, and the controlled-NOT (CNOT) gate which acts on two qubits at the time. Each algorithm is built from a random sequence of these gates, where the probability that a given gate is a 1-qubit gate is p_g and the probability that it is a 2-qubit gate is $1-p_g$. We set $p_g=0.5$ throughout this paper, as $p_g=0.5$ appears to lead to the most rapid convergence at least for the distribution of interference [20]. The choice of the qubit(s) on which a gate acts, is made uniformly and independently for different gates over all the qubits. Figure 1 shows an example of this kind of algorithm for 3 qubits and 5 gates.

The U(2) gates are chosen uniformly with respect to the invariant Haar measure of the U(2) group. They can be parametrized with four angles α, ψ, χ chosen randomly and uniformly from $[0,2\pi[$, and $\varphi=\arcsin(\xi^{1/2})$ with ξ picked randomly and uniformly from [0,1],

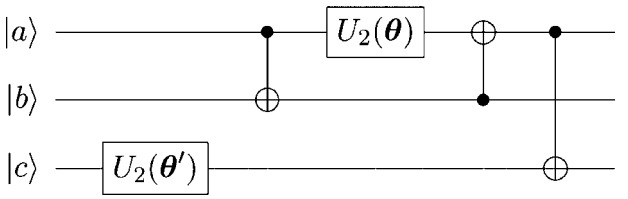

FIG. 1. A random UCE circuit. The two different angles $\boldsymbol{\theta}$ and $\boldsymbol{\theta}'$ mean two different random U(2) gates.

$$U_2(\boldsymbol{\theta}) = e^{i\alpha}\begin{pmatrix} \cos\varphi\, e^{i\psi} & \sin\varphi\, e^{i\chi} \\ -\sin\varphi\, e^{-i\chi} & \cos\varphi\, e^{-i\psi} \end{pmatrix} \equiv \begin{pmatrix} c & s \\ -\bar{s} & \bar{c} \end{pmatrix}, \quad (1)$$

where we have abbreviated $\boldsymbol{\theta}=(\alpha,\psi,\chi,\varphi)$ [13]. The phases α only modify the global phase of the algorithm and are irrelevant for the statistical properties that we are going to study. From the results of [5,14] it is clear that in the limit of the number of gates $n_g\to\infty$ and fixed n_q, UCE converges to CUE.

The UCE gate set might be summarized as $\Gamma=\{\{\frac{d\mu(U_2)}{4}, U_2\otimes 1_2\}, \{\frac{d\mu(U_2)}{4}, 1_2\otimes U_2\}, \{\frac{1}{4}, U_{\mathrm{CNOT1,2}}\}, \{\frac{1}{4}, U_{\mathrm{CNOT2,1}}\}\}$, where the first number in each pair in the list is the probability that the second member of the pair will be selected in any step of the algorithm, $\mu(U_2)$ means the Haar measure of U(2), and $U_{\mathrm{CNOTi,j}}$ is a controlled-NOT gate with control qubit i and target qubit j. It is easily checked that Γ is a "2-copy gapped gate set" in the terminology of [23]. This means that the operator $G=\int_{\mathrm{U}(4)} U\otimes U\otimes U^*\otimes U^* d\mu_\Gamma(U)$, defined for a general gate set distributed continuously over U(4) with measure $\mu_\Gamma(U)$, has only two eigenvalues with absolute value equal to 1. The difference between this largest degenerate eigenvalue 1 and the next smaller eigenvalue (in terms of its absolute value) is called the spectral gap Δ. Our gate set Γ has spectral gap $\Delta\simeq 0.232\,703$, if the gates are represented as 4×4 matrices. The gap is expected to decay as $1/n_q$ if the gates are represented as matrices of size 2^{n_q}, but will be finite for any finite n_q [27].

A. Distribution of matrix elements

The uniform distribution of CUE matrices of size N with respect to the Haar measure of the unitary group U(N) yields a specific joint probability distribution $P(U) \equiv P(U_{11},U_{12},\ldots,U_{NN})$ of the matrix elements U_{ij} which entirely defines this ensemble. Convergence of UCE to CUE means that the joint probability distribution $\tilde{P}(U)$ associated with UCE converges to $P(U)$. However, a direct numerical study of the joint distribution is impractical since the number of necessary realizations grows exponentially with the number of independent arguments of $\tilde{P}(U)$. More practical quantities can be obtained from the joint probability distribution by integrating out several variables. A natural quantity to consider is the distribution of matrix elements, which depends on one complex variable, and which is obtained by integrating out the other N^2-1 complex parameters,

$$P(U_{ij}) = \int \cdots \int \prod_{(k,l)\neq(i,j)} dU_{kl} P(U). \quad (2)$$

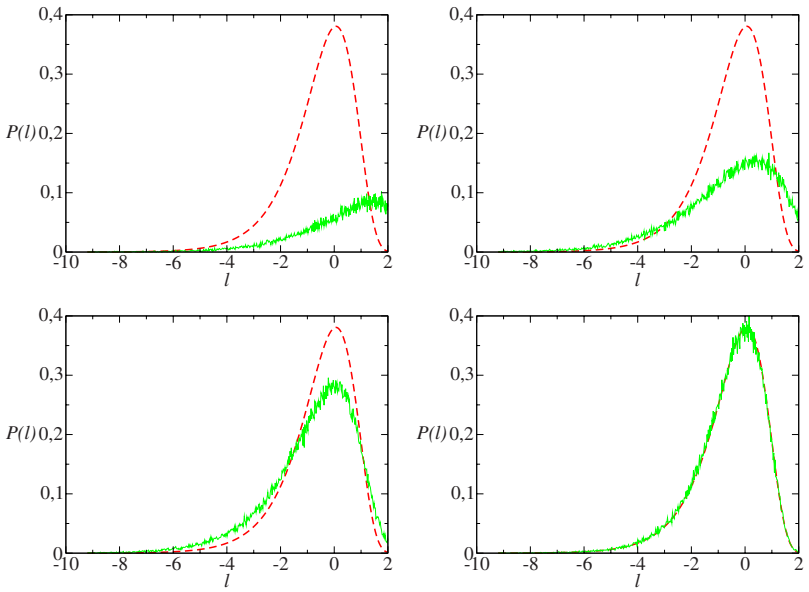

FIG. 2. (Color online) Convergence of $\tilde{P}(l)$ (green continuous line) to $P(l)$ (red dashed line) for 4 qubits with $n_g=5$, 10, 20, and 50 for an ensemble of 10^4 matrices.

The first quantity we study in this paper, closely related to $P(U_{ij})$, is the distribution of quantities defined by $l_{ij}=\ln(N|U_{ij}|^2)$. For CUE, one shows in random matrix theory (RMT) that all l_{ij} are distributed according to the normalized distribution

$$P(l)=\frac{(N-1)}{N}e^l\left(1-\frac{e^l}{N}\right)^{N-2}, \tag{3}$$

independently of the choice of the index of l_{ij} [15].

For UCE, the distribution of matrix elements is not independent of the elements chosen as long as the number of gates n_g is small, but becomes uniform over the matrix in the limit $n_g\to\infty$. For numerical efficiency, we have made two simplifications.

First, we produce and propagate only the first column of the matrix. This obviously reduces drastically the memory requirement, and moreover, the action of a CNOT gate on this vector requires only the manipulation of a subset of the matrix elements. With the binary notation of the row index i of a matrix element U_{i1}, $i=1+\sum_{\alpha=1}^{n_q}\sigma_\alpha 2^\alpha$, a CNOT between qubits k (control) and l (target) $\in[1,n_q]$ requires only the exchange of the 2^{n_q-2} elements in positions where $(\sigma_k=0,\ \sigma_l=1)$ with the 2^{n_q-2} elements in positions where $(\sigma_k=1,\ \sigma_l=1)$. For the U(2) gates, each element in the new column is a linear combination of two old elements, with c, s, $-\bar{s}$, or $\bar{c}$ as coefficients.

Second, we define $\tilde{P}(l)$ by averaging both over the realizations ($\langle\cdots\rangle_R$) and the elements in the first column ($\langle\cdots\rangle_C$)

$$\tilde{P}(l)=\frac{1}{n_rN}\sum_{r=1}^{n_r}\sum_{i=1}^{N}\tilde{h}(l_{i1}^{(r)})\equiv\langle\langle\tilde{h}(l)\rangle_C\rangle_R, \tag{4}$$

where $\tilde{h}(l_{i1}^{(r)})$ is the histogram for the ith component in the first column of the rth matrix. In order to obtain good statistics, it is important to produce a large enough number of matrices n_r. However, for a given number of matrices, when one adds one qubit, the calculation time roughly doubles, because the size of the Hilbert space is doubled. For this reason, as long as $n_q\leqslant 20$, we considered an ensemble of $n_r=a2^{b-n_q}$ matrices (with a and b integers). In other words, when adding a qubit, the increase of the size of the Hilbert space by a factor 2 is compensated by reducing the size of the ensemble by a factor 2, without a loss of statistics. We choose $a=10$ and $b=20$ leading to a total of about 10^7 matrix elements. Due to the correlations between the matrix elements (see below), averaging over a column is not quite as effective as averaging over realizations, and therefore the noise of the numerical data increases with growing n_q. For more than 20 qubits, numerical run time limitations forced us to fix the number of realizations to $n_r=10$.

Figure 2 shows, for $n_q=4$, the convergence of $\tilde{P}(l)$ to $P(l)$ with increasing n_g. To quantify the scaling of the convergence with the number of qubits, we use the statistical overlap function

$$D_P=\int_0^\infty[\sqrt{\tilde{P}(l)}-\sqrt{P(l)}]^2dl=2\left(1-\int_0^\infty\sqrt{\tilde{P}(l)P(l)}dl\right)\leqslant 2$$

which represents a distance between the square roots of the UCE and CUE distributions. This distance goes to zero as UCE converges toward CUE for $n_g\to\infty$. Using the square roots rather than the distributions themselves is motivated by the fact that D_P is bounded from above by the value 2, which simplifies the scaling analysis. Figure 3 shows the behavior of D_P as a function of n_g for $n_q=2,3,\ldots,28$. As expected, this quantity decays rapidly when the number of gates n_g grows, but the decay slows down with increasing n_q. Since our numerical calculations use a finite number of realizations, $\tilde{P}(l)$ fluctuates about $P(l)$. The distance D_P can therefore never vanish exactly, and we observe that it saturates for large n_g at a finite level $d_{\min}$ which depends on n_q. The level of saturation can be reduced by increasing n_r. When D_P saturates, our ensemble becomes indistinguishable from CUE within the precision of the numerics. We have fitted D_P for each value of n_q and we observed that when n_q is small ($n_q\lesssim 12$), D_P is well fitted by $2e^{-\alpha n_q}$ whereas for larger val-

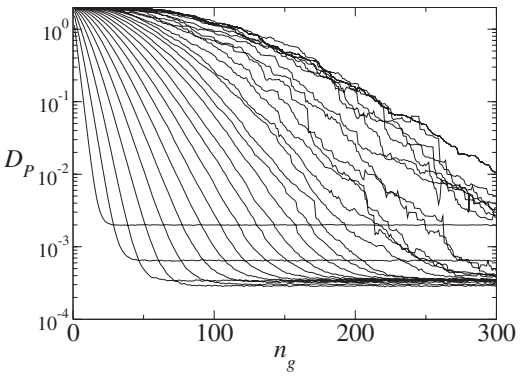

FIG. 3. The distance $D_P(n_g)$ between the distributions $P(l)$ and $\tilde{P}(l)$ as function of the number of gates n_g for $n_q=2,3,\ldots,28$ qubits (from left-hand to right-hand sides).

ues of n_q, D_P has a pronounced quadratic component in its exponent ($D_P \simeq 2e^{-\alpha n_q-\beta n_q^2}$). The change in the functional dependence of D_P on n_q makes it difficult to determine the scaling of the rate of convergence with n_q. We have therefore preferred to base our analysis on the number of gates n^* needed to achieve a fixed small value ϵ of D_P for a given number of qubits. Figure 4 shows the behavior of n^* as function of n_q for six different values of ϵ [$\ln(\epsilon)$ between -5 and 0]. We have fitted $n^*(n_q)$ with three different 2-parameters functions.

$$f_1 = a_1 n_q + b_1, \tag{5}$$

$$f_2 = a_2 n_q \ln(n_q/\epsilon) + b_2, \tag{6}$$

$$f_3 = a_3 n_q[n_q + \ln(1/\epsilon)] + b_3. \tag{7}$$

When fitting to f_i $(i=1,2,3)$, we permit a different value of a_i and b_i for different values of ϵ, and later study the ϵ-dependence of a_i and b_i. The linear function f_1 is an obvious choice given the appearance of the numerical data. The functions f_2 and f_3 are motivated by the results in [23] on 2-designs. These authors defined the convergence of unitary k-designs by the action on a test density matrix ρ of dimension $k2^{n_q}$. As measure of distance, they consider the completely bounded ("diamond") norm of the difference between the state $\mathcal{G}_W(\rho)=\Sigma_i p_i U_i^{\otimes k}\rho(U_i^\dagger)^{\otimes k}$ propagated by the k-design and $\mathcal{G}_H(\rho)=\int_U U^{\otimes k}\rho(U^\dagger)^{\otimes k}$ resulting from the propagator averaged over the unitary group. The gate set $\Gamma=\{\{p_i,U_i\}\}$ of unitary matrices U_i together with their probabilities p_i need to form a 2-copy gapped gate set. They show that a random quantum circuit of length n_g drawn from a 2-copy gapped gate set is an ϵ-approximate unitary 2-design if $n_g \geqslant C\{n_q[n_q+\ln(1/\epsilon)]\}$ with some positive constant C which may depend on the gate set. In the special case of a gate set drawn uniformly from U(4), which has maximum spectral gap $\Delta=1$ (i.e., $\mathcal{G}_W$ is a projector), it was found that an ϵ-approximate unitary 2-design is already reached for $n_g \geqslant Cn_q \ln(n_q/\epsilon)$.

As mentioned, our gate set Γ is indeed 2-copy gapped, with spectral gap $\Delta \simeq 0.232\,703$, and the results of [23] do therefore apply to the convergence of UCE to CUE. While one should be cautious in directly comparing these results, which constitute an upper bound, and are based on the propagation of a trial state and the use of the diamond norm with our results which use D_P as a measure of distance, it seems plausible that the convergence of the distribution of matrix elements of the propagator should be related to the convergence of a propagated test state (see also [24], where an efficient quantum algorithm for twirling was introduced). Based on the above cited results of [23] the function f_3 would therefore be the most natural candidate for a fit of $n^*(n_q)$. However, it turns out that the function f_2, even though relevant *a priori* for spectral gap $\Delta=1$, fits our numerical data much better, i.e., in what concerns the distribution of matrix elements, UCE converges to CUE much more rapidly than expected from the upper bound on n_g mentioned. Note that a similar result was found very recently in [28].

The quality of the fits is measured by χ^2, the sum of squares of deviations (see Fig. 4). We see that the simple

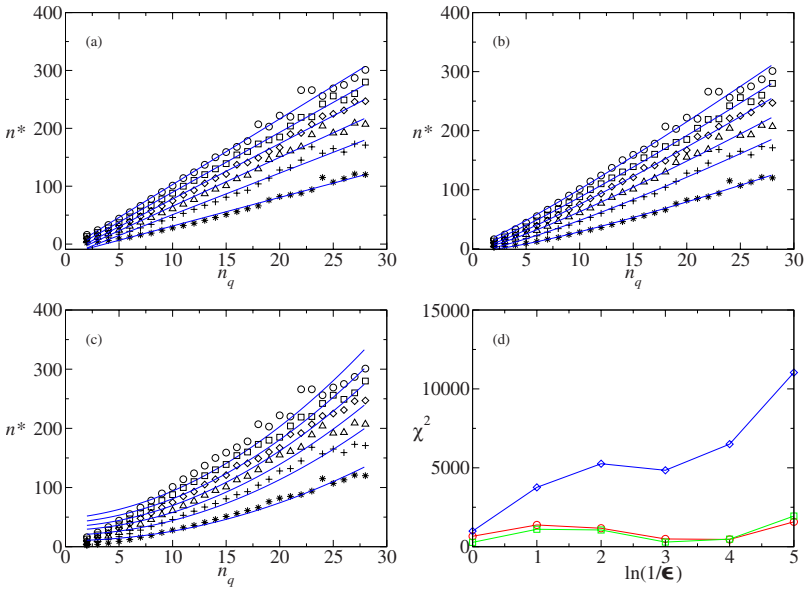

FIG. 4. (Color online) The number of gates n^* needed to achieve $D_P \leqslant \epsilon$ for $\ln(\epsilon)=0$, -1, -2, -3, -4 and -5 (*, +, △, ◇, □, ○, respectively) and $n_q=2,\ldots,28$. Straight lines are fits to the functions f_1, f_2, and f_3 [(a), (b), and (c) plot, respectively]. The plot (d) shows χ^2 for these fits [f_1 (○), f_2 (□), and f_3 (◇)].

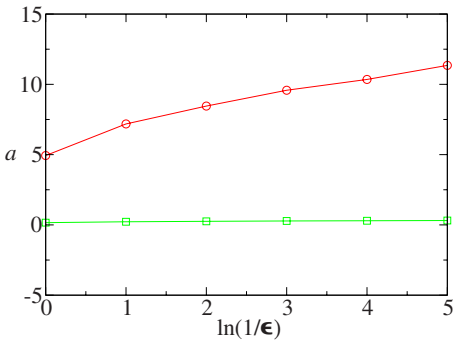

FIG. 5. (Color online) The coefficients a_1 (red circles) and a_2 (green squares) as a function of $\ln(1/\epsilon)$.

linear behavior f_1 fits better than the quadratic form despite a slight upwards curvature of the curves $n^*(n_q)$. That curvature is well captured by the $n_q \ln n_q$ behavior of f_2, whereas the quadratic behavior of f_3 fits much worse. The function f_2 gives in addition the correct ϵ-dependence for a constant coefficient $a_2 \simeq 0.2$ (see Fig. 5).

A clear distinction between f_1 and f_2 is not possible based on the numerical data, as both fit very well in the limited range of n_q available. Similarly we cannot exclude from the numerical data a $n_q[\ln(n_q/\epsilon)]^\nu$ behavior with $\nu > 0$. Nevertheless, our numerical results clearly indicate that, concerning the distribution of matrix elements, CUE can be efficiently simulated by UCE, in the sense that the number of gates used to achieve a given level of accuracy ϵ grows only like $n_q \ln(n_q/\epsilon)$ (or possibly like $n_q[\ln(n_q/\epsilon)]^\nu$, $\nu > 0$) with the number of qubits, and in any case more slowly than n_q^2. This is rather surprising, as $\tilde{P}(l)$ contains information about all moments, and one is therefore led to the conclusion that no moment of appreciable weight in the reconstruction of the distribution should need more than $O(n_q[\ln(n_q/\epsilon)]^\nu)$ gates before coming within distance ϵ relative to the CUE value. In order to confirm this hypothesis, we have studied several kth moments directly.

B. Moments of the distribution of matrix elements

The kth moment μ_k of the distribution of matrix elements is defined as $\mu_k = \langle y^k \rangle = N^k \langle |U_{ij}|^{2k} \rangle$. Invariant integration [26] leads for CUE to

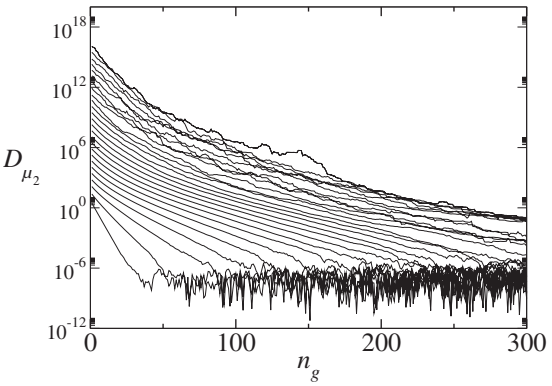

FIG. 6. $D_{\mu_2}(n_g)$ for a number of qubits n_q varying between 2 and 18 (from left-hand to right-hand sides).

$$\mu_k = \frac{k! N^k (N-1)!}{(N+k-1)!}, \tag{8}$$

which tends to $k!$ for $N \to \infty$ and k fixed. For UCE we average again over both random realizations and elements in the first column of U, analogously to (4), and define the kth moment as

$$\tilde{\mu}_k = \frac{1}{n_r} \sum_{r=1}^{n_r} \frac{1}{N} \sum_{i=1}^{N} (y_{i1}^{(r)})^k = \langle \langle y^k \rangle_C \rangle_R, \tag{9}$$

where $y_{i1}^{(r)}$ is equal to $N|U(l_{i1}^{(r)})|^2$ for the ith component in the first column of the rth matrix. As a measure of the deviation from the CUE result we use the relative deviations

$$D_{\mu_k} = \frac{|\tilde{\mu_k} - \mu_k|}{\mu_k}.$$

We have calculated D_{μ_k} for $k=2$, 4, and 8. For the latter two cases, we used $n_r = 10^5$ for $n_q = 2, \ldots, 14$ and $n_r = 5 \times 10^4$ for 15 qubits. Figure 6 shows the behavior of $D_{\mu_2}(n_g)$ for $n_q = 2, \ldots, 18$. The curves for D_{μ_4} and D_{μ_8} look very similar, with the exception of a higher saturation level, and are not shown.

In Fig. 7 we plot the number n^* needed to achieve a fixed small value $D_{\mu_k} < \epsilon$ for all three moments studied, $k = 2, 4, 8$, and for different values of $\ln(\epsilon)$, together with fits to

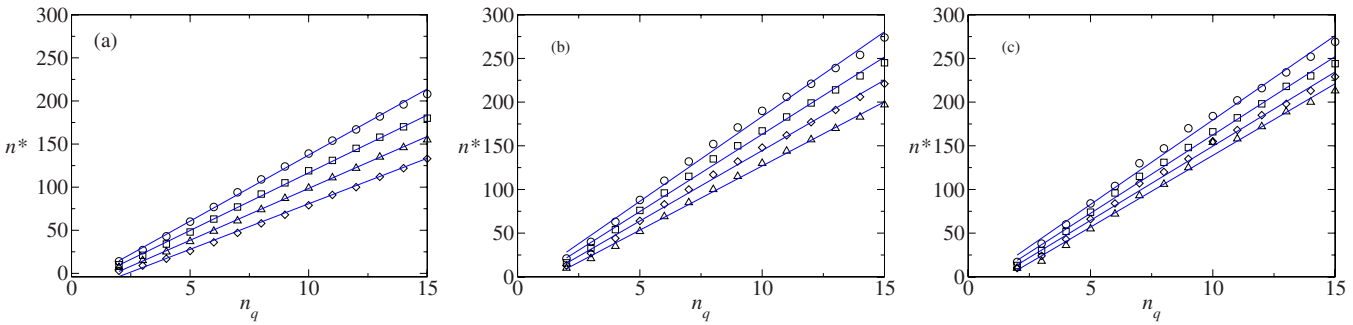

FIG. 7. (Color online) The number of gates n^* needed to achieve $D_{\mu_k} \leq \epsilon$ as a function of the number of qubits together with fits to the function f_1. Plots (a), (b), and (c) correspond to $k=2,4,8$, respectively. The different symbols in a plot (△, ◇, □, ○) represent different values of ϵ, with $\ln(\epsilon) = 0.5, -1.5, -2.5$, and -3.5 for $k=2$, $\ln(\epsilon) = -1, -2, -3$ and -4 for $k=4$, and $\ln(\epsilon) = 1, 0, -1$, and -2, respectively, for $k=8$.

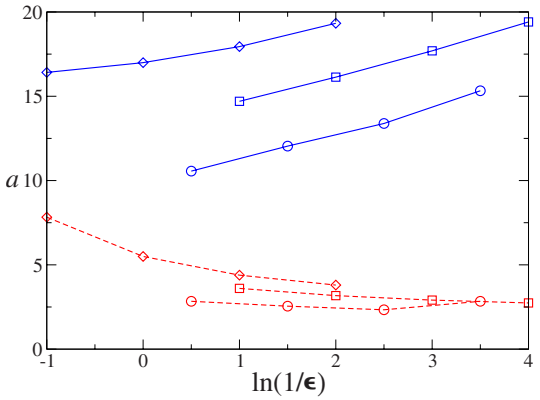

FIG. 8. (Color online) The coefficients a_1 (blue full lines) and a_2 (red dashed lines) for the convergence of D_{μ_2} (circles), D_{μ_4} (squares), and D_{μ_8} (diamonds) as a function of the available values of $\ln(1/\epsilon)$.

the function f_1 introduced above. The similarity between the three curves is striking. The figure demonstrates that $n^*(n_q)$ is very well described by a linear behavior for all k, and that furthermore even the slopes of that linear behavior are very similar for all moments.

We have also fitted the data to f_2, but our range of n_q is too small to decide which of the two functions f_1 and f_2 describes the scaling with n_q better. In fact, the numerical data $n^*(n_q)$ for D_{μ_4} and D_{μ_8} show a slight negative curvature, which makes f_2 nominally fit worse than f_1. However, we believe that the slight negative curvature is a numerical artifact explained below, and second, f_2 also represents the dependence on ϵ very well. This is shown in Fig. 8, where we have collected the coefficients a_1 and a_2 for all moments considered. Figure 8 shows that the ϵ dependence is correctly captured by the function $n_q \ln(n_q/\epsilon)$: a_2 becomes basically independent of ϵ for small enough ϵ. The prefactors a_2 of all moments considered converge to practically the same value once ϵ is small enough, underlining once more the very similar convergence behavior of the three moments with $k=2,4,8$ [a_2 takes on the values $a_2 \simeq 2.44$ for $k=2$, $a_2 \simeq 3.175$ for $k=4$, and $a_2 \simeq 3.80$ for $k=8$ at $\ln(1/\epsilon)=2$, but the differences between the values of a_2 appear to diminish further for decreasing ϵ].

The apparent slight sublinear behavior of $n^*(n_q)$ for D_{μ_4} and D_{μ_8} finds its explanation in the fact that the saturation levels of our numerical data for D_{μ_4} and D_{μ_8} are higher than for D_{μ_2} and D_P, such that the possible values we can choose for ϵ are closer to saturation than in D_{μ_2}. This slightly overestimates $n^*(n_q)$, but less so for large n_q, where the approach to saturation is slower, such that the curve $n^*(n_q)$ appears to curve downwards.

One might wonder if the slight negative curvature may not result from averaging over the column of the matrix. Indeed, while in CUE all matrix elements are equivalent in the sense that $\langle |U_{ij}|^{2k} \rangle$ is independent of i,j, and additionally averaging over a column would therefore give exactly the same result, this is not the case in a UCE circuit of given finite length n_g. For example, after one gate, the first element U_{11} (where the index 1 signifies the state $|0\cdots0\rangle$ in the computational basis) is never affected by a CNOT, whereas others are. One effect of the convergence of UCE to CUE is that these inhomogeneities decay. One might suspect that averaging over the first column effectively reduces the inhomogeneities and could therefore provide a mechanism of accelerated convergence compared to a moment that has not been averaged over a column of U. As the sample size (in the sense of the number of elements in a column used to average) increases exponentially with n_q, small differences in the $\langle |U_{ij}|^{2k} \rangle$ are rapidly averaged out, and this might suggest a more rapid convergence than for a single matrix element. Moreover, the effect is expected to become more pronounced for higher moments, which amplify small initial differences.

To test this hypothesis we calculated for restricted sample sizes ($n_q \leqslant 10$ and $n_r = 10^4$) the fourth and eighth moments μ_4' and μ_8' for a fixed matrix element (we chose U_{11} and U_{31}), defined as in (9) but without averaging over the first column. For larger values of n_q, a calculation that does not use averaging over a column is unfortunately beyond our numerical capacities. The corresponding signals D'_{μ_4} and D'_{μ_8} for the element U_{11} start off at a larger value than D_{μ_4} and D_{μ_8}, and decay more rapidly, until the latter are reached. However, this happens at rather small values of n_g, whereas for larger n_g, the two curves D'_{μ_k} and D_{μ_k} for the same k are basically indistinguishable within the precision of the data. Thus, averaging over a column does not significantly change $n^*(n_q)$.

On the other hand, we verified that also in D_{μ_2} and in D_P a slight negative curvature of $n^*(n_q)$ can be produced by pushing ϵ close to the saturation level. Furthermore, the quality of the fits to f_1 and f_2 deteriorates for decreasing ϵ. From a physical perspective a sublinear behavior seems impossible, as it would mean that the global state of a large enough quantum circuit equilibrates before even every qubit is touched by a quantum gate. All of these elements confirm the explanation of the slight negative curvature as numerical artifact as discussed above.

The main messages from Figs. 7 and 8 is that (1) all moments considered converge at basically the same rates; (2) the number of gates needed to achieve a given precision increases in good approximation linearly with the number of qubits, and (3) the additional ϵ dependence is well accounted for by a $n_q \ln(n_q/\epsilon)$ behavior of $n^*(n_q)$.

C. Correlations between matrix elements

Even in CUE, different matrix elements are not independently distributed (in contrast to CUE's Hermitian cousin GUE). One obvious reason for the correlations is the orthonormalization of columns and lines of a unitary matrix. We define correlations between k different y_{ij} for a same column j as $c_k = \langle \Pi_{i=1}^k y_{ij} \rangle = N^k \langle \Pi_{i=1}^k |U_{ij}|^2 \rangle$, where the average is over the considered ensemble. In the CUE case, one finds through invariant integration [26]

$$c_k = \frac{N^k (N-1)!}{(N+k-1)!}, \tag{10}$$

which differs from μ_k by a factor $\frac{1}{k!}$. Thus, for small k, the correlations are important, and comparable to the moments

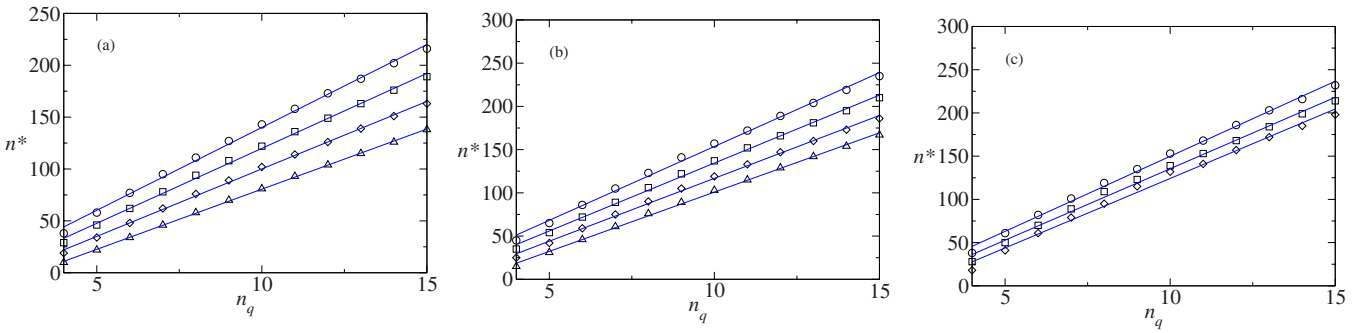

FIG. 9. (Color online) The number of gates n^* needed to achieve convergence of the correlators c_k, $D_{c_k} \leqslant \epsilon$ for $k=2,4,8$ together with fits to the functions f_1 [plots (a), (b), and (c), respectively]. For $k=2$, $\ln(\epsilon)=-1$, -2, -3, and -4 (△, ◇, □, ○); for $k=4$, $\ln(\epsilon)=0$, -1, -2, and -3 (△, ◇, □, ○); and for $k=8$, $\ln(\epsilon)=1$, 0, and -1 (◇, □, ○, respectively).

of the same order. For UCE we again average over the column, but include each element in at most one product in order not to create additional artificial correlations between the products,

$$\tilde{c_k} = \frac{1}{n_r\left(\frac{N}{k}\right)} \sum_{r=1}^{n_r} \sum_{i=1}^{N/k} \prod_{j=1}^{k} y^{(r)}_{(ki-k+j)1}. \tag{11}$$

Here, (x) means the integer part of x. We measure the distance to CUE through the relative deviations of $\tilde{c_k}$ from the CUE values

$$D_{c_k} = \frac{|\tilde{c_k} - c_k|}{c_k}. \tag{12}$$

Figure 9 shows results for the evolution of $n^*(n_q)$ in the cases $k=2$, $k=4$, and $k=8$, as well as fits to f_1. The behavior is predominantly linear, and very similar for all correlators considered. The numerical data can also be fitted very well to f_2, and again it is difficult within the limited range of n_q values available to us to clearly distinguish between one or the other. The function f_2 fits in general somewhat worse (plot not shown) than f_1, but this is due to the same numerical artifact of slight negative curvatures of $n^*(n_q)$ discussed above. Nevertheless, from Fig. 10 which shows the fitted coefficients a_1 and a_2 for all three correlators, it is clear that the ϵ dependence is correctly described by $n_q \ln(n_q/\epsilon)$, and that the prefactor a_2 is largely independent of the order of the correlator for sufficiently small ϵ.

Moreover, comparing Figs. 10 and 8, we see that the correlations c_k converge basically with the same rates as the moments of the same order, μ_k—a result to be expected from the theory of k-designs [22,23]. In fact, an alternative definition of a unitary k-design is that any polynomial in the complex matrix elements of degree (m,l) with $m,l \leqslant k$ has the same average over the unitary design as over the full unitary group [22].

III. CONCLUSION

We have studied the convergence of the distribution of matrix elements, moments of that distribution up to $\langle |U_{ik}|^{16} \rangle$, as well as correlations between matrix elements with up to 16 factors $|U_{ik}|$ within the same column, for random quantum algorithms drawn from the unitary circuit ensemble (UCE) to their counterparts in CUE. Simulating quantum circuits with up to 28 qubits (for the distribution of matrix elements), and up to 18 (15) qubits for the moments (correlations), we have shown that all of these quantities can be efficiently reproduced with a precision ϵ using quantum circuits from UCE containing a number of gates that scales at most as $n^* \leqslant C n_q [\ln(n_q/\epsilon)]^\nu$ with the number of qubits n_q, where C and ν are positive constants. Such fast convergence comes somewhat as a surprise, as for general 2-copy gapped gate sets with a gap $1 > \Delta > 0$ a quadratic upper bound, $n^* \leqslant C n_q [n_q + \ln(1/\epsilon)]$, has been shown [23] (UCE has spectral gap $\Delta \simeq 0.232\,703$). While our numerical results are restricted to moments and correlators of order $k \leqslant 16$, it has been conjectured that all moments can be reproduced efficiently with a number of gates that scales at most as poly(n_q,k) [23]. The obvious question arises then how the convergence of the full joint-probability distribution to the CUE counterpart can be inefficient, as the latter can presumably be reconstructed from a sufficiently large number of moments and correlators? Several answers seem possible: The number of moments and correlators needed increases exponentially with the number

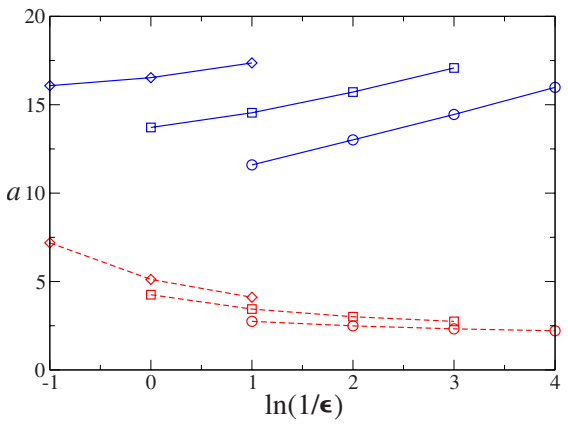

FIG. 10. (Color online) The coefficients a_1 (blue full lines) and a_2 (red dashed lines) for the convergence of D_{c_2} (circles), D_{c_4} (squares), and D_{c_8} (diamonds) as a function of the available values of $\ln(1/\epsilon)$.

of qubits, and the conjecture might breakdown for these, while it might be valid for moments and correlators of polynomial order. Alternatively, the reconstruction of the full joint distribution from the moments and the correlators may not be possible, even if an exponentially large number of moments and correlators are known. In any case, if the common wisdom that in order to faithfully reproduce the full joint probability distribution of CUE using UCE circuits one needs a number of gates which scales exponentially with the number of qubits is correct, our results suggest that the inefficiently reproduced quantities must be of more complex nature than the low moments of the distribution of absolute values of matrix elements and their low-order correlation functions.

ACKNOWLEDGMENTS

We thank Aram Harrow for useful correspondence, Lorenza Viola, Winton Brown, and Géza Tóth for useful discussions, and CALMIP (Toulouse) and IDRIS (Orsay) for the use of their computers. This work was supported by the Agence National de la Recherche (ANR), project INFOSYSQQ, and the EC IST-FET project EUROSQIP.

[1] A. Ambainis and A. Smith, Proceedings of RANDOM, Cambridge, MA, 2004.
[2] P. Hayden, D. Leung, P. W. Shor, and A. Winter, Commun. Math. Phys. **250**, 371 (2004).
[3] P. Sen, e-print arXiv:quant-ph/0512085v1.
[4] A. Harrow, P. Hayden, and D. Leung, Phys. Rev. Lett. **92**, 187901 (2004).
[5] J. Emerson, E. Livine, and S. Lloyd, Phys. Rev. A **72**, 060302 (2005).
[6] M. L. Mehta, *Random Matrices*, 2nd ed. (Academic, New York, 1991).
[7] D. Deutsch, Proc. R. Soc. London, Ser. A **400**, 97 (1985).
[8] D. P. DiVincenzo, Phys. Rev. A **51**, 1015 (1995).
[9] T. Sleator and H. Weinfurter, Phys. Rev. Lett. **74**, 4087 (1995).
[10] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter, Phys. Rev. A **52**, 3457 (1995).
[11] A. Barenco, Proc. R. Soc. London, Ser. A **449**, 679 (1995).
[12] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[13] M. Pozniak, K. Zyczkowski, and M. Kus, J. Phys. A **31**, 1059 (1998).
[14] J. Emerson, Y. Weinstein, M. Saraceno, S. Lloyd, and D. Cory, Science **302**, 2098 (2003).
[15] F. Haake, *Quantum Signatures of Chaos*, 2nd ed. (Springer, Berlin, Heidelberg, 2000).
[16] M. Žnidarič, Phys. Rev. A **76**, 012318 (2007).
[17] R. Oliveira, O. C. O. Dahlsten, and M. B. Plenio, Phys. Rev. Lett. **98**, 130502 (2007).
[18] O. Dahlsten, R. Oliveira, and M. Plenio, J. Phys. A **40**, 8081 (2007).
[19] E. M. Fortunato, M. A. Pravia, N. Boulant, G. Teklemariam, T. F. Havel, and D. G. Cory, J. Chem. Phys. **116**, 7599 (2002).
[20] L. Arnaud and D. Braun, Phys. Rev. A **75**, 062314 (2007).
[21] W. G. Brown, Y. S. Weinstein, and L. Viola, Phys. Rev. A **77**, 040303 (2008).
[22] C. Dankert, R. Cleve, J. Emerson, and E. Livine, e-print arXiv:quant-ph/0606161.
[23] A. Harrow and R. Low, e-print arXiv:quant-ph/0802.191v1.
[24] G. Toth and J. J. Garcia-Ripoll, Phys. Rev. A **75**, 042311 (2007).
[25] D. Gross, K. Audenaert, and J. Eisert, J. Math. Phys. **48**, 052104 (2007).
[26] S. Aubert and C. Lam, J. Math. Phys. **44**, 6112 (2003).
[27] A. Harrow (private communication).
[28] M. Znidaric, Phys. Rev. A **78**, 032324 (2008).