Exact and approximate unitary 2-designs and their application to fidelity estimation

Christoph Dankert,[1] Richard Cleve,[1,2] Joseph Emerson,[3] and Etera Livine[2]
[1]*David R. Cheriton School of Computer Science and Institute for Quantum Computing, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1*
[2]*Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario, Canada, N2L 2Y5*
[3]*Department of Applied Mathematics and Institute for Quantum Computing, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1*

(Received 27 April 2009; published 6 July 2009)

We develop the concept of a *unitary t*-design as a means of expressing operationally useful subsets of the stochastic properties of the uniform (Haar) measure on the unitary group $U(2^n)$ on n qubits. In particular, sets of unitaries forming 2-designs have wide applicability to quantum information protocols. We devise an $O(n)$-size in-place circuit construction for an *approximate* unitary 2-design. We then show that this can be used to construct an efficient protocol for experimentally characterizing the fidelity of a quantum process on n qubits with quantum circuits of size $O(n)$ without requiring any ancilla qubits, thereby improving upon previous approaches.

DOI: 10.1103/PhysRevA.80.012304 PACS number(s): 03.67.Lx, 03.65.Fd

I. INTRODUCTION

The importance of generating random states and random unitary operators in quantum information processors has become increasingly clear from the growing number of algorithms and protocols that presume such a resource [1–8]. For many algorithms and protocols the invariant (Haar) measure on the unitary group $U(D)$ is the natural randomization measure [3,5,7,8]. It is well known that generating Haar-random unitary operators on a quantum information processor is inefficient: the number of gates grows exponentially with the number of qubits. Consequently it is useful to identify subsets of the unitary group that can adequately simulate the Haar measure for a given class of operational tasks and to identify efficient gate decompositions for these subsets.

Quantum data hiding [2], based on a process known as bilateral twirling [9], is an example where such a subset has been identified. Sampling over the discrete subset of $U(2^n)$ known as the Clifford group is sufficient for this task [2], and the number of gates required to implement such operations is $O(n^2)$ for n-qubit systems. It has been shown recently in [8] that a large family of protocols can be described in terms of bilateral twirling with Haar-random unitary operators, and noted that the Clifford group is sufficient to implement them. A related protocol, twirling a quantum channel with Haar-random unitary operators, has been studied recently for the purpose of reducing a generic gate to a standard form, which is among the results of Ref. [10], where it was also shown that the Clifford group is also an adequate substitute for that task. Moreover, the task of generating typical generic entanglement has been studied in Ref. [11], where it was shown that a discrete subset of elements of the unitary group that are constructible with $O(n^3)$ gates are sufficient for that purpose. Each of the above tasks satisfies the unitary 2-design condition and is therefore subsumed by our work.

In this paper we propose the concept of a *unitary t*-design as a generalization of the concept of quantum state t-design [12], and, in the nontrivial case of $t=2$, we give explicit efficient constructions for them as quantum circuits. Our circuits have size $O(n^2)$ for exact implementations and $O(n \log 1/\varepsilon)$ for approximations within ε (where ε is the degree of closeness to the exact case, which is formalized in the next section). We then discuss how our results are useful in the context of experimentally estimating the average fidelity of a quantum channel. In particular, we show how the (exponentially inefficient) protocol of Ref. [7] can be achieved efficiently with $O(n)$ operations for n-qubit systems.

II. DEFINITIONS AND SUMMARY OF RESULTS

We define a *unitary t*-design for D dimensions as a finite set $\{U_k\}_{k=1}^K \subset \mathcal{U}(D)$ of unitary operators on $\mathbb{C}^D$ such that, for every polynomial $P_{(t,t)}(U)$ of degree at most t in the matrix elements of U and at most t in the complex conjugates of those matrix elements [i.e., a polynomial of degree at most (t,t)],

$$\frac{1}{K}\sum_{k=1}^{K} P_{(t,t)}(U_k) = \int_{\mathcal{U}(D)} dU P_{(t,t)}(U), \qquad (1)$$

where, unless otherwise specified, integrals over $\mathcal{U}(D)$ are with respect to the unitarily invariant Haar measure (for an equivalent definition in terms of representation theory, see [13]). This definition is a natural extension of the definition of t-designs for quantum states.

The focus of the current paper is on the $t=2$ case; however, there are specific applications for other values of t (see [14]). The connection with operational tasks in the case $t=2$ can be seen as follows. Consider a general quantum channel Λ acting on D-dimensional quantum states. Such a (super)operator is a completely positive trace preserving linear map acting on $L(\mathbb{C}^D)$, the algebra of linear operators on $\mathbb{C}^D$. Suppose that Λ is conjugated by a randomly chosen unitary operation with respect to a probability measure μ on $\mathcal{U}(D)$. That is, U is chosen according to μ and then the channel is modified to $\hat{U}^\dagger \circ \Lambda \circ \hat{U}$, where

$$\hat{U}(\rho) = U\rho U^\dagger,$$

$$\hat{U}^{\dagger}(\rho) = U^{\dagger}\rho U. \tag{2}$$

Denoting the resulting superoperator by $\mathbb{E}_{\mu}(\Lambda)$, we have

$$\mathbb{E}_{\mu}(\Lambda){:}X \mapsto \int_{\mathcal{U}(D)} d\mu(U)U^{\dagger}\Lambda(UXU^{\dagger})U, \tag{3}$$

for all $X \in L(\mathbb{C}^D)$ (including the operationally significant case where the input to the channel X is a density operator). The process that transforms any superoperator Λ to the superoperator $\mathbb{E}_{\mu}(\Lambda)$ is called a μ-*twirl*. A particular case of interest is if the measure μ is taken as the unitarily invariant Haar measure. A unitary 2-design has the property that sampling uniformly from $\{U_1,\ldots,U_K\}$ is operationally equivalent to sampling from the Haar measure. In other words, if μ is set to the uniform probability measure on $\{U_1,\ldots,U_K\}$ then, for any quantum channel Λ,

$$E_{\mu}(\Lambda) = E_{\text{Haar}}(\Lambda). \tag{4}$$

This can be seen by considering the linear mappings Λ on $L(\mathbb{C}^D)$ of the form

$$\Lambda(X) = AXB, \tag{5}$$

where $A,B \in L(\mathbb{C}^D)$. Then the condition in Eq. (1) is equivalent to the condition

$$\frac{1}{K}\sum_{k=1}^{K} U_k^{\dagger}AU_kXU_k^{\dagger}BU_k = \int_{\mathcal{U}(D)} dUU^{\dagger}AUXU^{\dagger}BU, \tag{6}$$

for all $A,X,B \in L(\mathbb{C}^D)$; this equivalence can be seen explicitly by considering A,X,B of the form $|i\rangle\langle j|$. Although not all quantum channels are of the form of Eq. (5), they are convex combinations of mappings of this form. Therefore, Eq. (4) follows by linearity. Similarly, one can show that bilateral twirling [2,9] and generating typical subsystem purity [11] correspond to particular instances of Eq. (1) with $t=2$.

Our first contribution is a direct proof of the following theorem (which Leung has pointed out to us, can also be obtained as a corollary of Ref. [2], Sec. VI).

Theorem 1. The uniform distribution over the Clifford group on n qubits is a unitary 2-design with $D=2^n$.

As is known, the Clifford group $\mathcal{C}_n$ on n qubits can be implemented by quantum circuits of size $O(n^2)$ [2,15]. This general result allows us to immediately deduce that, for example, the Clifford group gives a more efficient solution to the protocol for generating generic entanglement given in [11]. Moreover, as described at the end of this paper, this result leads to an efficient solution to the experimental problem of estimating the average fidelity of a quantum process or a quantum channel [7].

Given the wide class of protocols that require unitary 2-designs, we show furthermore that a more efficient implementation on n qubits is possible if we consider *approximate* unitary 2-designs. We define these with respect to arbitrary linear superoperators $\Lambda{:}L(\mathbb{C}^D)\to L(\mathbb{C}^D)$ (that is, Λ need not be completely positive and trace preserving for this to make sense). Any such Λ can be expressed in the form $\Lambda(X) = \mathrm{Tr}_E[A(X\otimes \mathbb{1}_E)B]$, where A and B act on an extended Hilbert space $\mathcal{C}^D\otimes\mathcal{H}_E$. A μ twirl of the superoperator Λ with respect to a measure μ on a subset of $\mathcal{U}(D)$ is a mapping of the form $\Lambda\mapsto\mathbb{E}_{\mu}(\Lambda)$, where $\mathbb{E}_{\mu}(\Lambda)$ is the superoperator

$$\mathbb{E}_{\mu}(\Lambda){:}X \mapsto \int_{\mathcal{U}(D)} d\mu(U)U^{\dagger}\Lambda(UXU^{\dagger})U. \tag{7}$$

Note that, by linearity, unitary 2-designs satisfy Eq. (4) for all linear superoperators (i.e., not just for quantum channels).

Defining an ε-approximate unitary 2-design in terms of the diamond norm [16] as a measure on a finite subset of $\mathcal{U}(D)$ satisfying the property

$$\|\mathbb{E}_{\mu}(\Lambda) - \mathbb{E}_{\text{Haar}}(\Lambda)\|_{\diamond} \le \varepsilon\|\Lambda\|_{\diamond}. \tag{8}$$

Note that, in the interesting case where Λ is a quantum channel, $\|\Lambda\|_{\diamond}=1$; hence the channel $\mathbb{E}_{\mu}(\Lambda)$ is a good approximation of the channel $\mathbb{E}_{\text{Haar}}(\Lambda)$.

Our second contribution is to show that:

Theorem 2. For all $\varepsilon>0$, an ε-approximate unitary 2-design on n qubits can be implemented by in-place circuits of size $O(n \log 1/\varepsilon)$ and depth $O(\log n \log 1/\varepsilon)$.

Our third contribution is an application of this toward fidelity estimation:

Theorem 3. The average fidelity of a quantum channel Λ acting on n qubits, can be estimated to within $\delta>0$ with error probability $\varepsilon>0$ at a cost of $O(\log 1/\varepsilon)$ evaluations of the channel conjugated by in-place circuits of size $O(n \log 1/\varepsilon)$ and depth $O(\log n \log 1/\varepsilon)$.

III. EXACT CONSTRUCTION

We prove Theorem 1, which implies that a unitary 2-design on n qubits (dimension $D=2^n$) can be explicitly constructed by in-place circuits of size $O(n^2)$. Our approach is to construct a uniform probability distribution on a subset of the Clifford group $\mathcal{C}_n$, which defines a $\mathcal{C}_n$-twirl. It is sufficient to consider linear mappings of the form $\Lambda(X)=AXB$, where A, $B\in L(\mathbb{C}^D)$, and the results can be extended to arbitrary linear superoperators by linearity. Specifically, we prove that, for all X,

$$\frac{1}{|\mathcal{C}_n|}\sum_{U\in\mathcal{C}_n} U^{\dagger}AUXU^{\dagger}BU = \int_{U(D)} dUU^{\dagger}AUXU^{\dagger}BU.$$

As shown in Ref. [7], the right-hand side can be expressed in the form,

$$\int_{U(D)} dUU^{\dagger}AUXU^{\dagger}BU = \frac{\mathrm{Tr}(AB)\mathrm{Tr}(X)}{D}\frac{\mathbb{1}}{D} + \frac{D\mathrm{Tr}(A)\mathrm{Tr}(B) - \mathrm{Tr}(AB)}{D(D^2-1)} \times\left(X - \mathrm{Tr}(X)\frac{\mathbb{1}}{D}\right). \tag{9}$$

To evaluate the left-hand side, we will make use of the fact that $\mathcal{C}_n$ is the normalizer of the generalized Pauli group $\mathcal{P}_n$ consisting of all n-fold tensor products of the one-qubit Pauli operators $\{\mathbb{1},X,Y,Z\}$. We denote the elements of $\mathcal{P}_n$ as $\{P_j\}_{j=1}^{D^2}$, where P_1 is the n-fold tensor product of $\mathbb{1}$. Applying a $\mathcal{P}_n$-twirl to the mapping $\Lambda(X)=AXB$ results in a mapping of the form $X \mapsto \sum_{k=1}^{D^2} r_k P_k X P_k$, where $r_1 = \mathrm{Tr}(A)\mathrm{Tr}(B)/D^2$ and $\sum_{k=1}^{D^2} r_k = \mathrm{Tr}(AB)/D$. This follows from noting that we can express $A = \sum_{a=1}^{D^2} \alpha_a P_a$ and $B = \sum_{b=1}^{D^2} \beta_b P_b$. The resulting operation maps X to

$$
\begin{aligned}
1/D^2 \sum_{k=1}^{D^2} P_k A P_k X P_k B P_k
&= 1/D^2 \sum_{a=1}^{D^2} \sum_{b=1}^{D^2} \alpha_a \beta_b \left(\sum_{k=1}^{D^2} (-1)^{(k, a \oplus b)_{S_P}} \right) P_a X P_b \\
&= \sum_{a=1}^{D^2} \alpha_a \beta_a P_a X P_a, \qquad (10)
\end{aligned}
$$

with the symplectic inner product S_P on the index space (see [17] for further details; these techniques are discussed also in [10]). Therefore, setting $r_k = \alpha_k \beta_k$ leads to the above form.

We can express each $U \in \mathcal{C}_n$ as $U = Q_j P_k$, where $\{P_1, \ldots, P_{D^2}\} = \mathcal{P}_n$ and $\{Q_1, \ldots, Q_{|\mathcal{P}_n|/|\mathcal{C}_n|}\}$ contains a representative from each coset in $\mathcal{C}_n/\mathcal{P}_n$ (how these representatives are chosen does not matter). Hence, after twirling Λ by $\mathcal{P}_n$, we then twirl with $\{Q_1, \ldots, Q_{|\mathcal{P}_n|/|\mathcal{C}_n|}\}$, where we henceforth refer to the latter operation as a twirl by $\mathcal{C}_n/\mathcal{P}_n$. The $\mathcal{C}_n/\mathcal{P}_n$-twirl yields

$$
\frac{|\mathcal{P}_n|}{|\mathcal{C}_n|} \sum_{j=1}^{|\mathcal{C}_n|/|\mathcal{P}_n|} \sum_{k=1}^{D^2} r_k Q_j^\dagger P_k Q_j X Q_j^\dagger P_k Q_j. \qquad (11)
$$

Next we distinguish the identity element $P_1 = \mathbb{1}$ and make use of the fact that conjugation under the Clifford group maps each nonidentity Pauli element to every other nonidentity Pauli element with equal frequency. It follows that the final state is

$$
\begin{aligned}
r_1 X + \frac{|\mathcal{P}_n|}{|\mathcal{C}_n|} \sum_{k=2}^{D^2} r_k \sum_{j=1}^{|\mathcal{C}_n|/|\mathcal{P}_n|} Q_j^\dagger P_k Q_j X Q_j^\dagger P_k Q_j
&= r_1 X + \frac{1}{D^2 - 1} \left(\sum_{k=2}^{D^2} r_k \right) \sum_{l=2}^{D^2} P_l X P_l. \qquad (12)
\end{aligned}
$$

Using the identity $\sum_{j=1}^{D^2} P_j X P_j = D\mathrm{Tr}(X)\mathbb{1}$, it is straightforward to show that the right sides of Eqs. (12) and (9) are equal.

IV. APPROXIMATE CONSTRUCTION

We now prove Theorem 2, in which an ε-approximate unitary 2-design can be explicitly constructed in terms of circuits that are in-place, of size $O(n \log 1/\varepsilon)$, and of depth $O(\log n \log 1/\varepsilon)$. More precisely, we describe a probabilistic construction that produces an n-qubit quantum circuit, generated according to a probability distribution $(p_1, \ldots, p_m)$ on a sequence of circuits $(C_1, \ldots, C_m)$ with the following property. For any linear superoperator Λ on n qubits, the mapping,

$$
\rho \mapsto \sum_{i=1}^{m} p_i C_i^\dagger \Lambda(C_i \rho C_i^\dagger) C_i, \qquad (13)
$$

is ε close (with respect to $\|\cdot\|_\diamond$) to the mapping

$$
\rho \mapsto \int_{\mathcal{U}(2^n)} dU U^\dagger \Lambda(U \rho U^\dagger) U. \qquad (14)
$$

Since we are converting Λ to the superoperator

$$
\sum_{i=1}^{m} p_i \hat{C}_i^\dagger \circ \Lambda \circ \hat{C}_i, \qquad (15)
$$

we describe the probabilistic circuit construction as a series of simple operations that are each conjugations performed on the channel.

Uniformization procedure:

1. $\mathcal{C}_1/\mathcal{P}_1$-twirl qubit k for all $k \in \{1, \ldots, n\}$.
2. Conjugate the first qubit by a random XOR. (This operation is defined in Figure 2 below.)
3. H-conjugate the first qubit, and $\mathcal{C}_1/\mathcal{P}_1$-twirl qubit k for all $k \in \{2, \ldots, n\}$.
4. Conjugate the first qubit by a random XOR.
5. H-conjugate the first qubit, and $\mathcal{C}_1/\mathcal{P}_1$-twirl qubit k for all $k \in \{2, \ldots, n\}$.
6. With probability 1/2, S-conjugate the first qubit.
7. Conjugate the first qubit by a random XOR.
8. $\mathcal{C}_1/\mathcal{P}_1$-twirl the first qubit.

FIG. 1. The uniformization procedure

Our construction first applies a Pauli twirl to the superoperator, which consists of $O(n)$ gates and results in a superoperator that is a linear combination of Pauli channels of the form $\rho \mapsto P_a \rho P_a$. In order to convert an arbitrary Pauli channel into a good approximation of a depolarizing channel, we shall add slightly more than $O(n)$ further twirling operations to approximately uniformize the probabilities associated with each P_a for all $a \neq 1$. The process consists of a series of repetitions of the procedure in Fig. 1 (where the operation *conjugating the first qubit by a random* XOR is defined in Fig. 2).

A $\mathcal{C}_1/\mathcal{P}_1$-twirl of a qubit can be analyzed as follows. Let $R = SH$, where $S = |0\rangle\langle 0| + i|1\rangle\langle 1|$, and H is the Hadamard transform. Select $i \in \{0, 1, 2\}$ uniformly and conjugate the register by R^i. This operation has the property that, if it is applied to the identity channel $\mathbb{1}$, it has no net effect; however, for a Pauli channel of the form X, Y, or Z, this operation causes the channel to become a uniform mixture of X, Y, and Z.

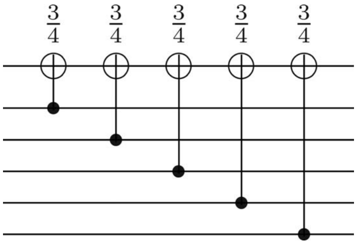

FIG. 2. Conjugating the first qubit by a random XOR is the conjugation by a randomly generated circuit of the above form, where a numerical label associated with a gate indicates that it occurs with probability of 3/4 (with probability of 1/4 there is no gate). That is, for each $k \in \{2,\dots,n\}$, with independent probability of 3/4, there is a CNOT gate with the first qubit as target and qubit k as control.

By Eq. (10), the initial Pauli twirl results in a linear combination of mappings of the form $\rho \mapsto P_a \rho P_a$. We consider each term separately: start with a channel of the form $\rho \mapsto P_a \rho P_a$, for some fixed $a \neq 1$, and apply the above procedure. To analyze the result, we trace through the effect of each of the eight steps of the uniformization procedure:

(1) For each k, if component k of P_a is $\mathbb{1}$ then it remains $\mathbb{1}$, and if component k is X, Y, or Z, then it becomes a uniform mixture of X, Y, and Z.

(2) Call an execution of this procedure *good* if, after step (2), the first component of the channel is X or Y. This happens with probability of at least 1/2, which can be seen by considering these cases:

Case 1: for all $k \in \{2,\dots,n\}$, component k of P_a is $\mathbb{1}$. In this case, the controlled-NOT (CNOT) gates have no effect but since $a \neq 1$, component 1 of P_a is *not* $\mathbb{1}$. Therefore, after the previous step, the first component of P_a is uniformly distributed over X, Y, and Z. Hence the first component of the channel is X or Y with probability of 2/3.

Case 2: for some $k \in \{2,\dots,n\}$, component k of P_a is not $\mathbb{1}$. With probability $(2/3)(3/4)=1/2$, component k has X or Y *and* the CNOT gate is present. This causes the first component to evolve as follows. If it is X or $\mathbb{1}$ then it becomes an equal mixture of $\mathbb{1}$ and X. Also, if it is Y or Z then it becomes an equal mixture of Y and Z.

In both of the above cases, the first component is X or Y with probability of 1/2.

(3) If the execution is good then the first component is Y or Z. For each $k \in \{2,\dots,n\}$, component k is either $\mathbb{1}$ or a uniform mixture of X, Y, and Z.

(4) If the execution is good then for each $k \in \{2,\dots,n\}$, component k is $\mathbb{1}$ with independent probability of 1/4, and some mixture of X, Y, Z with probability of 3/4. To see why this is so, for each k, consider the effect of the back action of the CNOT gates in the following two cases separately.

Case 1: after the previous step, component k is $\mathbb{1}$. In this case, it remains $\mathbb{1}$ with probability of 1/4, and it becomes Z with probability of 3/4.

Case 2: after the previous step, component k is a uniform mixture of X, Y, and Z. In this case, with probability of 3/4, the channel becomes a uniform mixture of Y, X, and $\mathbb{1}$. Hence the component becomes $\mathbb{1}$ with probability $(3/4)(1/3)=1/4$.

(5) If the execution is good then, after this step, the first component of the channel is X or Y, and, for each $k \in \{2,\dots,n\}$, component k is independently a uniform mixture of $\mathbb{1}$, X, Y, and Z.

(6) If the execution is good then, after this step, the first component of the channel is a uniform mixture of X and Y.

(7) Call a good execution *typical* if, after step (6), there is at least one component $k \in \{2,\dots,n\}$ that is not $\mathbb{1}$. The probability that a good execution is also typical is $1-(1/4)^{n-1}$. If the execution is good and typical, the first component of the channel is a uniform mixture of $\mathbb{1}$, X, Y, and Z (independent of the other components of the channel).

To see why this is so, consider the effect of any non-$\mathbb{1}$ component $k \in \{2,\dots,n\}$. Prior to the potential conjugation by CNOT, the first component is uniformly distributed among X and Y, and component k is uniformly distributed among X, Y, and Z. Therefore, with probability $(2/3)(3/4)=1/2$, the first component becomes a uniform mixture of I and Z.

(8) If the execution is good then, after this step, the first component of the channel is: in a uniform mixture of X, Y, and Z if the execution is not typical, and a uniform mixture of $\mathbb{1}$, X, Y, and Z if it is typical.

For executions that are both good and typical, there are $4(4^{n-1}-1)$ possible Pauli channels that can result (namely, those that are not $\mathbb{1}$ in at least one of the components from 2 through n). Conditional on the execution being good, each of these cases arises with probability of $1/4^n$. For executions that are good but not typical, there are three possible outcomes (namely, all channels that are X, Y, or Z in the first component and $\mathbb{1}$ in components 2 through n). Conditional on the execution being good, each of these three cases arises with probability of $1/(3\times 4^{n-1})$. The resulting probability distribution on Pauli channels can be expressed as a convex combination of these two distributions: (a) the uniform distribution of all nontrivial Pauli channels, and (b) the uniform distribution on the three nontypical Pauli channels. Distribution (a) occurs with probability $(1/2)[1-(1/4)^n]$, and distribution (b) occurs with probability $(1/2)[1+(1/4)^n]$. Note that distribution (a) corresponds to a perfect 2-design. Repeating the procedure $O(\log 1/\varepsilon)$ times, we can increase probability weighting associated with (a) from $(1/2)[1-(1/4)^n]$ to $1-\varepsilon/2$ (the perfect Pauli channel need only arise in one of the repetitions). Each execution of the uniformization procedure consists of $O(n)$ gates, which can be implemented in $O(\log n)$ depth—the only nontrivial part is the conjugations by a random XOR, whose log-depth implementation is based on the construction in Fig. 3.

The net result of this construction can be viewed as a mixture of two mixed-unitary operations, one of which is a perfect 2-design. The perfect 2-design occurs with probability at least $1-\varepsilon/2$ and the other operation occurs with probability at most $\varepsilon/2$. Therefore the construction yields a linear superoperator of the form

$$(1-\varepsilon/2)\mathbb{E}_{\mathrm{Haar}}(\Lambda)+(\varepsilon/2)\mathbb{E}_{\nu}(\Lambda), \tag{16}$$

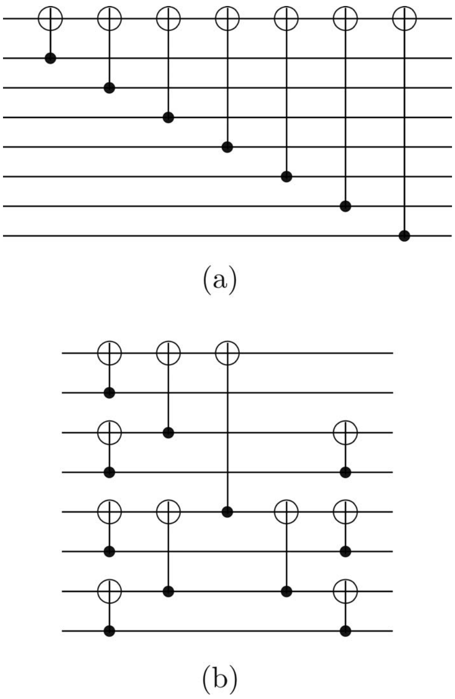

FIG. 3. Part (a) shows a circuit consisting of $n=7$ CNOT gates with common target; part (b) shows an equivalent circuit (based on a binary tree of addition modulo 2) whose depth is bounded by 2 log n and size is bounded $2n$.

for some probability measure ν on $\mathcal{U}(2^n)$. The diamond norm distance between this operation and $\mathbb{E}_{\text{Haar}}(\Lambda)$ is bounded by

$$\|(1-\varepsilon/2)\mathbb{E}_{\text{Haar}}(\Lambda)+(\varepsilon/2)\mathbb{E}_{\nu}(\Lambda)-\mathbb{E}_{\text{Haar}}(\Lambda)\|_{\diamond}\leq(\varepsilon/2) \times\|\mathbb{E}_{\text{Haar}}(\Lambda)\|_{\diamond}+(\varepsilon/2)\|\mathbb{E}_{\nu}(\Lambda)\|_{\diamond}=\varepsilon\|\Lambda\|_{\diamond}, \tag{17}$$

where we have made use of the fact that $\|\cdot\|_{\diamond}$ is invariant under twirling: $\|\mathbb{E}_{\mu}(\Lambda)\|_{\diamond}=\|\Lambda\|_{\diamond}$. It follows that the construction produces an ε-approximate 2-design.

V. APPLICATION TO FIDELITY ESTIMATION

We now turn to a discussion of the experimental problem of fidelity estimation for which the above unitary 2-design constructions lead to an efficient scalable protocol that is accessible with current experimental techniques on systems of a few qubits. Consider the Haar-averaged fidelity [7,18]

$$\langle F\rangle \equiv \int_{\mathcal{U}(D)} dU \mathrm{Tr}[U|0\rangle\langle 0|U^{\dagger}\Lambda(U|0\rangle\langle 0|U^{\dagger})] = \sum_{k}\frac{|\mathrm{Tr}(A_k)|^2+D}{D^2+D}. \tag{18}$$

of a quantum operation $\Lambda(\rho)=\Sigma_k A_k\rho A_k^{\dagger}$. The Haar-averaged fidelity can be related to two standard fidelity benchmarks: the entanglement fidelity F_e, which has been proposed as means of characterizing the noise strength in a physical quantum channel Λ [19], and the gate fidelity F_g, which has been used to characterize the quality of quantum memory [20] or of an implementation of a target unitary U_g on a noisy quantum processing device [21,22]. In the latter scenario we imagine the implementation of a gate sequence U_g followed immediately by its inverse $U_g^{\dagger}$, and make the identification $\Lambda(\rho)=U_g^{\dagger}\mathcal{E}(U_g\rho U_g^{\dagger})U_g$, where the map $\mathcal{E}(\rho)$ represents the noise accumulated over the course of implementing $U_g^{\dagger}U_g$. Then, using the results of Refs. [7,18,19,21], we find the following relationship between the Haar-average fidelity and the previously proposed gate fidelity and entanglement fidelity,

$$\langle F\rangle = \frac{DF_g+1}{D+1} = \frac{DF_e+1}{D+1}. \tag{19}$$

We emphasize that this relationship holds in the context where F_g and F_e are understood to characterize errors under the composed sequence $U_g^{\dagger}U_g$ rather than errors under U_g itself.

There are two experimental approaches to estimating F_e and F_g for a given quantum channel. The first is based on ancilla-assisted process tomography, or some variant such as direct characterization, both of which require creating an entangled state of $2n$ qubits: the first n of which are subjected to the unknown transformation and the remaining n of which are ancilla qubits that are subject to the identity channel [23], followed by joint measurements on the final $2n$ qubit state. A significant disadvantage of this approach is the requirement of n noise-free ancilla qubits, as well as the requirement of joint operations on $2n$, rather than n, qubits. The second approach is standard process tomography, which suffers from the requirement of a number of experiments that grows exponentially with $n=\log_2 D$ [18,23].

However, as described in Ref. [7], we can estimate $\langle F\rangle$ directly by the following protocol: apply a random unitary operator U to the initial state $|0\rangle$, followed by the quantum operation Λ, and then apply $U^{\dagger}$ to the output state. Then from Eq. (18) we see that $\langle F\rangle$ can be estimated by repeating this procedure with U sampled randomly from the Haar measure in each experiment. For an arbitrary but fixed average fidelity $0\leq\langle F\rangle\leq 1$, the number of experiments required to estimate $\langle F\rangle$ to precision $\delta>1/4^n$ is independent of the dimension D. A serious limitation of the approach of Ref. [7] is that the implementation of a random unitary requires exponential resources. However, given that F is a polynomial function of homogeneous degree (2,2), it follows that we can estimate $\langle F\rangle$ by sampling from any unitary 2-design instead of the Haar-random unitary operators presumed in Ref. [7]. Hence the results of this paper, and in particular the ε-approximate unitary 2-design described above, imply that each experiment requires only $O(n\log(1/\varepsilon))$ gates. Hence the fidelity $\langle F\rangle$, and equivalently F_g and F_e, may be estimated by an efficient experimental protocol that can be applied with existing levels of quantum control in systems of a few qubits. We remark that, after the original submission of this work, a randomization approach has been developed [24] that offers an improvement over the resource requirements discussed above for the task of fidelity estimation. However, the randomization approach of this paper is strictly stronger than that of Ref. [24].

It remains an interesting open question whether an arbitrary quantum randomization algorithm can be reduced to a t-design condition, and hence classified within this framework. This would provide further motivation to generalize the methods of this paper to obtain unitary and state t-designs for $t>2$. The alternate definition proposed in [13] might be a good starting point for research in this direction.

ACKNOWLEDGMENTS

We thank Robin Blume-Kohout, David Cory, Daniel Gottesman, Debbie Leung, Pranab Sen, and John Watrous for helpful discussions, and Markus Grassl for pointing out an omission in an earlier version of this work. This research was supported by Canada's NSERC, MITACS, and CIFAR, and the U.S. ARO.

[1] A. Ambainis, M. Mosca, A. Tapp, and R. de Wolf, Proceedings of the 41st FOCS (IEEE Computer Society Press, Los Alamitos, CA, 2000), pp. 547–553.

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

[3] J. Radishkran, M. Roetteler, and P. Sen, Lect. Notes Comput. Sci. **3580**, 1399 (2005); P. Sen, in Proceedings of the 21st Annual IEEE Conference of Computational Complexity (IEEE Computer Society Press, Los Alamitos, CA, 2006), pp. 274–287.

[4] A. Ambainis and A. Smith, Lect. Notes Comput. Sci. **3122**, 249 (2004).

[5] P. Hayden, D. Leung, P. Shor, and A. Winter, Commun. Math. Phys. **250**, 371 (2004); C. H. Bennett, P. Hayden, D. Leung, P. Shor, and A. Winter, IEEE Trans. Inf. Theory **51**, 56 (2005); A. Harrow, P. Hayden, and D. Leung, Phys. Rev. Lett. **92**, 187901 (2004).

[6] J. Emerson, Y. Weinstein, M. Saraceno, S. Lloyd, and D. Cory, Science **302**, 2098 (2003).

[7] J. Emerson, R. Alicki, and K. Zyczkowski, J. Opt. B: Quantum Semiclassical Opt. **7**, S347 (2005).

[8] A. Abeyesinghe, I. Devetak, P. Hayden, and A. Winter, e-print arXiv:quant-ph/0606225 Proc. R. Soc. A (to be published).

[9] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).

[10] W. Dür, M. Hein, J. I. Cirac, and H.-J. Briegel, Phys. Rev. A **72**, 052326 (2005).

[11] R. Oliveira, O. Dahlsten, and M. Plenio, Phys. Rev. Lett. **98**, 130502 (2007).

[12] J. Renes, R. Blume-Kohout, A. Scott, and C. Caves, J. Math. Phys. **45**, 2171 (2004); A. Klappenecker and M. Roetteler, in Proceedings of the 2005 IEEE International Symposium on Information Theory (IEEE, Adelaide, Australia, 2005), p. 1740–1744.

[13] An equivalent definition of a unitary t-design is such that $\Sigma_{k=1}^{K}D^{\mathcal{I}}(U_k)=0$ for all nontrivial irreducible representations $D^{\mathcal{I}}$ contained in the tensor power $V^{\otimes t}\otimes\bar{V}^{\otimes t}$ of the fundamental representation V and its conjugate. This characterization is especially relevant when analyzing the efficiency of "random circuit" constructions for generating pseudorandom sets of unitaries [25].

[14] For example, $t=1$ corresponds to the conditions of a private quantum channel [1], and $t=4$ corresponds to the case of the state-distinction problem [26,27].

[15] D. Gottesman, Ph.D. thesis, California Institue of Technology, 1997.

[16] A. Yu. Kitaev, A. H. Shen, and M. N. Vyali, *Classical and Quantum Computation* (American Mathematical Society, Providence, RI, 2002), Vol. 47.

[17] C. Dankert, MMath thesis, University of Waterloo, 2005.

[18] M. A. Nielsen, Phys. Lett. A **303**, 249 (2002).

[19] B. Schumacher, Phys. Rev. A **54**, 2614 (1996).

[20] N. Boulant, T. F. Havel, M. A. Pravia, and D. G. Cory, Phys. Rev. A **67**, 042322 (2003).

[21] E. M. Fortunato, M. A. Pravia, N. Boulant, G. Teklemariam, T. F. Havel, and D. G. Cory, J. Chem. Phys. **116**, 7599 (2002).

[22] Y. Weinstein, T. F. Havel, J. Emerson, N. Boulant, M. Saraceno, and D. Cory, J. Chem. Phys. **121**, 6117 (2004).

[23] I. Chuang and M. Nielsen, J. Mod. Opt. **44**, 2455 (1997); J. B. Altepeter, D. Branning, E. Jeffrey, T. C. Wei, P. G. Kwiat, R. T. Thew, J. L. OBrien, M. A. Nielsen, and A. G. White, Phys. Rev. Lett. **90**, 193601 (2003).

[24] J. Emerson *et al.*, Science **317**, 1893 (2007).

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

[26] P. Sen, Proceedings of the 21st Annual IEEE Conference on Computational Complexity (IEEE Computer Society Press, Los Alamitos, CA, 2006), pp. 274–287.

[27] A. Ambainis and J. Emerson, Proceedings of the 22nd Annual IEEE Conference on Computational Complexity(IEEE Computer Society Press, Los Alamitos, CA, 2007), pp. 129–140.