Optimal Strategies for Estimating the Average Fidelity of Quantum Gates

Daniel M. Reich,[1] Giulia Gualdi,[1,2,3] and Christiane P. Koch[1]

[1]*Theoretische Physik, Universität Kassel, Heinrich-Plett-Straße 40, D-34132 Kassel, Germany*
[2]*Dipartimento di Fisica e Astronomia, Università di Firenze, Via Sansone 1, 50019 Sesto Fiorentino, Italy*
[3]*QSTAR, Largo Enrico Fermi 2, 50125 Firenze, Italy*

(Received 6 June 2013; published 12 November 2013)

We show that the minimum experimental effort to estimate the average error of a quantum gate scales as 2^n for n qubits and requires classical computational resources $\sim n^2 2^{3n}$ when no specific assumptions on the gate can be made. This represents a reduction by 2^n compared to the best currently available protocol, Monte Carlo characterization. The reduction comes at the price of either having to prepare entangled input states or obtaining bounds rather than the average fidelity itself. It is achieved by applying Monte Carlo sampling to so-called 2-designs or two classical fidelities. For the specific case of Clifford gates, the original version of Monte Carlo characterization based on the channel-state isomorphism remains an optimal choice. We provide a classification of the available efficient strategies to determine the average gate error in terms of the number of required experimental settings, average number of actual measurements, and classical computational resources.

DOI: 10.1103/PhysRevLett.111.200401 PACS numbers: 03.67.Ac, 03.65.Wj

Introduction.—The development of quantum technologies is currently facing a number of obstacles. One of them is the difficulty to assess efficiently how well a quantum device implements a desired operation. The corresponding performance measure is the average fidelity or the average gate error which can be determined via quantum process tomography [1]. Full process tomography scales, however, strongly exponentially in the number of qubits and provides the full process matrix, i.e., much more information than just the gate error. For practical applications, a more targeted and less resource-intensive approach is required. Recent attempts at reducing the resources employ stochastic sampling [2–7]. The process matrix can be estimated efficiently if it is sparse in a convenient basis [4,5,8,9]. Randomized benchmarking utilizing so-called unitary t-designs [10] is scalable when estimating the gate error for Clifford gates [6,7]. For general unitary operations, Monte Carlo sampling combined with the channel-state isomorphism currently appears to be the most efficient approach [2,3]. It comes with the advantage of separable input states. A second promising approach for general unitaries utilizes state 2-designs [11]. They are given, for example, by the states of $d+1$ ($d=2^n$) mutually unbiased bases [12]. Since only three out of the $d+1$ mutually unbiased bases consist of separable states [13], entangled input states need to be prepared. The approaches of Refs. [2,3,11] yield the average fidelity for general unitaries with an arbitrary, prespecified accuracy. They have been tested experimentally, albeit so far only for two- and three-qubit operations, without taking advantage of the protocols' efficiency [14,15]. Alternatively to estimating the average fidelity directly, an upper and a lower bound can be obtained from two classical fidelities that are evaluated for the states of two mutually unbiased bases [16]. This approach has been employed in an experiment demonstrating quantum simulation with up to 6 qubits [17]. When designing a quantum device, one is thus faced with a number of options to determine the average gate errors, the optimality of which will depend on specific experimental constraints and the required accuracy.

Here, we provide a unified classification of the currently available approaches in terms of the number and type of input states and measurements that need to be available, the number of actual experiments that need to be carried out, and the classical computational resources that are required. We show that applying Monte Carlo estimation to the 2-design protocol and the classical fidelities yields a reduction by a factor of 2^n in resources for general unitary operations. For the specific task of characterizing Clifford gates, the two strategies are as efficient as Monte Carlo sampling combined with the channel-state isomorphism [2,3] or randomized benchmarking [6,7]. The reduction in resources that we report here for general unitary operations is made possible by avoiding the channel-state isomorphism: Instead of estimating the average fidelity in terms of a single state fidelity in Liouville space, it is determined by a sum over state fidelities in Hilbert space. We have recently shown that a minimal set of states in Hilbert space is sufficient for device characterization [18]. Therefore, the number of states that enter the sum is determined only by the desired bounds.

Monte Carlo sampling.—We first review Monte Carlo estimation of the average fidelity as introduced in Refs. [2,3] before applying it to 2-designs [5,11] and classical fidelities [16]. The fidelity of a quantum state or process is of the form $F=\mathrm{Tr}[\rho^{\mathrm{id}}\rho^{\mathrm{act}}]$ with $\rho^{\mathrm{id(act)}}$ the ideal (actual) state. Monte Carlo sampling estimates the quantity F from a small random sample of measurements. F thus needs to be

expressed in terms of measurement results. To this end, ρ^{id} and ρ^{act} are expanded in an orthonormal basis of Hermitian operators W_k, yielding $F \sim \sum_k \mathrm{Tr}[\rho^{\mathrm{id}} W_k]\mathrm{Tr}[\rho^{\mathrm{act}} W_k]$. The measurement results are treated as a random variable X taking values X_κ which occur with probability $\mathrm{Pr}(\kappa)$ [$\sum_{\kappa=1}^{T} \mathrm{Pr}(\kappa) = 1$ with T the size of the event space], i.e.,

$$F = \sum_{\kappa=1}^{T} \mathrm{Pr}(\kappa) X_\kappa. \tag{1}$$

Introducing $\chi_\rho(\kappa) = \mathrm{Tr}[\rho W_\kappa]$, X_κ is given by

$$X_\kappa = \frac{\chi_{\rho^{\mathrm{act}}}(\kappa)}{\chi_{\rho^{\mathrm{id}}}(\kappa)} = \frac{\mathrm{Tr}[\rho^{\mathrm{act}} W_\kappa]}{\mathrm{Tr}[\rho^{\mathrm{id}} W_\kappa]} \quad \text{and} \quad \mathrm{Pr}(\kappa) = \frac{\chi_{\rho^{\mathrm{id}}}(\kappa)^2}{\mathcal{N}},$$

with $\mathcal{N}$ ensuring proper normalization. $\mathrm{Pr}(\kappa)$ is also called the relevance distribution. Two levels of stochastic sampling are involved in the Monte Carlo estimation of a fidelity: According to Eq. (1), F is the expectation value of the random variable X taking values X_κ with known probability $\mathrm{Pr}(\kappa)$. However, the X_κ's cannot be accessed directly, since they depend on another random variable, the expectation value of W_κ for ρ^{act}. Because of the statistical nature of quantum measurements, as well as random errors in the experiment, it will be necessary to repeatedly measure W_κ in order to determine X_κ. Assuming that the X_κ's have been determined with sufficient accuracy, Monte Carlo sampling estimates their expectation value F from a finite number of realizations L,

$$F = \lim_{L\to\infty} F_L \quad \text{with} \quad F_L = \frac{1}{L}\sum_{l=1}^{L} X_{\kappa_l}. \tag{2}$$

with κ_l taking values between 1 and T. L is chosen such that the probability for F_L to differ from F by more than ϵ is less than δ. The key point of the Monte Carlo approach is that L depends only on the desired accuracy ϵ and confidence level δ and is independent of the system size. Note that ϵ is lower bounded by measurement and state preparation errors. Since, as quantum mechanical expectation values, the X_{κ_l}'s are known only approximately, also F can be obtained only approximately $\tilde{F}_L = (1/L)\sum_{l=1}^{L} \tilde{X}_{\kappa_l}$ (with $\tilde{X}_{\kappa_l}$ denoting the approximate values of X_{κ_l}). Therefore, in addition to ensuring that F_L approximates F with a statistical error of at most ϵ, one also has to guarantee that $\tilde{F}_L$ approximates F_L with the desired accuracy. This implies repeated measurements for a given element l ($l = 1, \ldots, L$) of the Monte Carlo sample. Denoting the number of respective measurements by N_l, the total number of experiments is given by $N_{\mathrm{exp}} = \sum_{l=1}^{L} N_l$. It can be shown that a proper choice of $\langle N_l\rangle$ and $\langle N_{\mathrm{exp}}\rangle$ guarantees the approximations of F_L by $\tilde{F}_L$ and of F by F_L to hold with the desired confidence level. While L is independent of system size, the choices of $\langle N_l\rangle$ and $\langle N_{\mathrm{exp}}\rangle$ in general depend on it.

For a quantum process, the average fidelity can be obtained by Monte Carlo estimation when combining it with the channel-state isomorphism [2,3]. F_{av} is then expressed in terms of the entanglement fidelity F_e via $F_{\mathrm{av}} = (dF_e + 1)/(d + 1)$ [19]. Since F_e is a state fidelity in Liouville space and Liouville space vectors correspond to Hilbert space operators, this implies evaluation of F_{av} with respect to an operator basis comparing input to output operators. Since a complete operator basis consists of 2^{2n} elements and the size of the event space is given by all possible combinations of input and output operators, $T = 2^{4n}$. The fact that only states not operators can be prepared as input is remedied by randomly selecting eigenstates of the input operators. There are six eigenstates for the three Pauli operators for each qubit. Therefore, the number of experimental settings, i.e., pairs of input state or output measurement operator, is given by $N_{\mathrm{setting}} = N_{\mathrm{input}} \times N_{\mathrm{meas}} = 6^n \times 2^{2n}$. The random selection of experimental settings requires classical computational resources $\mathcal{C}_{\mathrm{class}}$ that scale as $n^2 2^{4n}$ [2,3]. Although only some of the settings will be selected, the ability to implement all of them in the experiment is implied. Because of the statistical nature of measurements, all in all $\langle N_{\mathrm{exp}}\rangle$ runs of the experiment have to be carried out. For the experimental implementation, N_{setting}, $\langle N_{\mathrm{exp}}\rangle$, and $\mathcal{C}_{\mathrm{class}}$ thus characterize the procedure.

Monte Carlo estimation of classical fidelities and 2-designs.—State fidelities in Hilbert space as opposed to a state fidelity in Liouville space are sufficient to estimate the average fidelity of an arbitrary quantum gate [16,18]. We therefore distinguish in a Monte Carlo event κ_l between input states and measurement operators, $\kappa_l = (i_l, k_l)$. This allows for applying Monte Carlo sampling to the classical fidelities of Ref. [16] and the 2-design approach [11].

The two classical fidelities which yield an upper and a lower bound to the average fidelity [16] can be written as [18]

$$\begin{aligned} F_j &= \frac{1}{d}\sum_{i=1}^{d} \mathrm{Tr}[\rho_i^{j,\mathrm{id}} \rho_i^{j,\mathrm{act}}] \\ &= \frac{1}{d}\sum_{i=1}^{d} \mathrm{Tr}[U|\Psi_i^j\rangle\langle\Psi_i^j|U^+ \mathcal{D}(|\Psi_i^j\rangle\langle\Psi_i^j|)], \end{aligned} \tag{3}$$

with $|\Psi_i^j\rangle$ the states of two mutually unbiased bases in d-dimensional Hilbert space ($i = 1, \ldots, d$, $j = 1, 2$, $d = 2^n$), $\mathcal{D}$ the dynamical map describing the actual evolution, and U the desired unitary. Expanding the states $\rho_i^{j,\mathrm{id}}$, $\rho_i^{j,\mathrm{act}}$ in terms of Pauli operators, Eq. (3) becomes

$$F_j = \sum_{i=1}^{d}\sum_{k=1}^{d^2} \mathrm{Pr}^j(i,k) \frac{\chi_{\mathcal{D}}^j(i,k)}{\chi_U^j(i,k)}, \tag{4}$$

with characteristic function $\chi_U^j(i,k) = \mathrm{Tr}[W_k U|\Psi_i^j\rangle\langle\Psi_i^j|U^+]$ and relevance distribution $\mathrm{Pr}^j(i,k) = (1/d^2)[\chi_U^j(i,k)]^2$. Note that $\mathrm{Tr}[W_k W_{k'}] = d\delta_{k,k'}$. We show in the Supplemental Material [21] that $\mathrm{Pr}^j(i,k)$ is properly

normalized such that we can estimate the two classical fidelities F_j, and thus an upper and a lower bound to F_{av}, by Monte Carlo sampling.

The expression for the average fidelity when using 2-designs given in terms of $d+1$ mutually unbiased bases [11]

$$F_{\mathrm{av}}^{\mathrm{2des}} = \frac{1}{d(d+1)} \sum_{i=1}^{d(d+1)} \mathrm{Tr}[U|\Psi_i\rangle\langle\Psi_i|U^{+}\mathcal{D}(|\Psi_i\rangle\langle\Psi_i|)] \tag{5}$$

is formally similar to Eq. (3); i.e., it can be interpreted as the sum over $d+1$ classical fidelities. Equation (5) can thus be rewritten

$$F_{\mathrm{av}}^{\mathrm{2des}} = \sum_{i=1}^{d(d+1)} \sum_{k=1}^{d^2} \mathsf{Pr}^{\mathrm{2des}}(i,k) \frac{\chi_{\mathcal{D}}^{\mathrm{2des}}(i,k)}{\chi_{U}^{\mathrm{2des}}(i,k)}, \tag{6}$$

with characteristic function χ_U^{2des} analogous to χ_U^j, and the relevance distribution differing only in normalization, $\mathsf{Pr}^{\mathrm{2des}}(i,k) = (1/d^2(d+1))[\chi_U^{\mathrm{2des}}(i,k)]^2$. We show in the Supplemental Material [21] that also $\mathsf{Pr}^{\mathrm{2des}}(i,k)$ is properly normalized such that $F_{\mathrm{av}}^{\mathrm{2des}}$ can be estimated by Monte Carlo sampling.

Resources for estimating the gate error of general unitaries.—Evaluating Eqs. (4) or (6) by Monte Carlo estimation involves randomly selecting L times a pair (i_l, k_l) of input state-measurement operator. Compared to Refs. [2,3], the number of input states is significantly reduced for the two approaches based on state fidelities in Hilbert space. This yields a correspondingly smaller number of settings that an experimentalist needs to be able to implement, cf. Table I. Moreover, the smaller number of input states reduces the classical computational resources required for the random selection by a factor of 2^n for the classical fidelities. This is due to $\mathcal{C}_{\mathrm{class}} = N_{\mathrm{input}} \times \mathcal{C}_{\mathrm{single}}$ with $\mathcal{C}_{\mathrm{single}}$ the classical computational cost for sampling a single state fidelity in Hilbert space ($\mathcal{C}_{\mathrm{single}} \sim n^2 2^{2n}$ [3]). It reflects the fact that the relevance distribution for the classical fidelities depends on $\mathcal{O}(d^3)$ parameters, whereas the relevance distribution of Refs. [2,3] depends on $\mathcal{O}(d^4)$ parameters. The reduced number of parameters is sufficient to determine whether the actual evolution matches the desired unitary [18].

TABLE I. Resources required for determining the average gate error of a general unitary operation in terms of classical computational effort $\mathcal{C}_{\mathrm{class}}$ required for the random selection, number N_{input} of input states that need to be prepared, the number of experimental settings N_{setting} from which the actual experiments will be randomly chosen, and the average number $\langle N_{\mathrm{exp}}\rangle$ of experiments to be performed. $N_{\mathrm{setting}} = N_{\mathrm{input}} \times N_{\mathrm{meas}}$ with the number of measurement operators $N_{\mathrm{meas}} = 2^{2n}$ for all cases. A: Monte Carlo sampling based on the channel-state isomorphism [2,3]. B: Monte Carlo sampling for 2-designs. C: Monte Carlo sampling for classical fidelities.

Approach	$\mathcal{C}_{\mathrm{class}}$	N_{input}	N_{setting}	$\langle N_{\mathrm{exp}}\rangle$
A	$\mathcal{O}(n^2 2^{4n})$	6^n	$\mathcal{O}(6^n 2^{2n})$	$\mathcal{O}(2^{2n})$
B	$\mathcal{O}(n^2 2^{4n})$	$2^n(2^n+1)$	$\mathcal{O}(2^{4n})$	$\mathcal{O}(2^n)$
C	$\mathcal{O}(n^2 2^{3n})$	2×2^n	$\mathcal{O}(2^{3n})$	$\mathcal{O}(2^n)$

Analogously to Refs. [2,3], we determine the sample size L by Chebyshev's inequality. It provides an upper bound for the probability of a random variable Z with variance σ_Z to deviate from its mean,

$$\mathsf{Pr}[|Z - \langle Z\rangle| \geq \sigma_Z/\sqrt{\delta}] \leq \delta, \tag{7}$$

with $\delta > 0$. In our case, $\langle Z\rangle = F$ with $F = F_{\mathrm{av}}^{\mathrm{2des}}$ or F_j, $Z = F_L = 1/L\sum_{l=1}^{L} X_{\kappa_l}$, and $X_{\kappa_l} \equiv X_l = \chi_{\mathcal{D}}(i_l,k_l)/\chi_U(i_l,k_l)$, cf. Eqs. (4) and (6). We show in the Supplemental Material [21] that the variance of X_l is smaller than 1, and thus $\mathsf{var}(F_L) \leq 1/L$. Then the choice $L = 1/(\epsilon^2\delta)$ guarantees that the probability for the estimate F_L to differ from F by more than ϵ is smaller than δ. Specifying the experimental inaccuracy and choosing the confidence level thus determines the sample size.

In order to estimate the number of required experiments, we first determine the number of experiments for one setting, N_l. For each l, the observable W_{k_l} has to be measured N_l times to account for the statistical nature of the measurement. The corresponding approximation to X_l is given by

$$\tilde{X}_l = \frac{1}{\chi_U(i_l,k_l)} \frac{1}{N_l} \sum_{j=1}^{N_l} w_{lj}, \tag{8}$$

with w_{lj} the measurement result for the jth repetition of experimental setting l, equal to either $+1$ or -1 for Pauli operators. Since $\tilde{X}_l$ is given as the sum of independent random variables w_{lj}, N_l can be determined using Hoeffding's inequality. It provides an upper bound for the probability of a sum $S = \sum_{i=1}^{n} Y_i$ of independent variables Y_i with $a_i \leq Y_i \leq b_i$ to deviate from its expected value by more than ϵ,

$$\mathsf{Pr}(|S - \langle S\rangle| \geq \epsilon) \leq 2\exp\left(-\frac{2\epsilon^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right), \tag{9}$$

$\forall\, \epsilon > 0$. In our case, $S = \tilde{F}_L = (1/L)\sum_{l=1}^{L} \tilde{X}_l$ and, using Eq. (8), $\sum_{i=1}^{n}(b_i - a_i)^2 = \sum_{l=1}^{L} 4N_l[LN_l\chi_U(i_l,k_l)]^{-2}$. Inserting this into Eq. (9), it is obvious that the choice

$$N_l = \frac{2}{L\epsilon[\chi_U(i_l,k_l)]^2} \log\left(\frac{2}{\delta}\right) = N_l(i_l,k_l) \tag{10}$$

ensures the right-hand side of Eq. (9) to be $\leq \delta$. The setting l is chosen with probability $\mathsf{Pr}^{j/\mathrm{2des}}(i_l,k_l)$. The average number of times that this specific experiment (with input state i_l and measurement operator W_{k_l}) is carried out is therefore given by

$$\langle N_l\rangle = \sum_{i_l=1}^{d}\sum_{k_l=1}^{d^2} \mathsf{Pr}^j(i_l,k_l)N_l(i_l,k_l)$$
$$= \frac{1}{d^2}\sum_{i_l=1}^{d}\sum_{k_l=1}^{d^2}[\chi_U^j(i_l,k_l)]^2\frac{4}{[\chi_U^j(i_l,k_l)]^2 L\epsilon^2}\log\left(\frac{2}{\delta}\right)$$
$$\leq 1+\frac{2d}{L\epsilon^2}\log\left(\frac{2}{\delta}\right) \quad (11)$$

for the two classical fidelities ($j = 1, 2$). The same $\langle N_l\rangle$ is obtained for the 2-designs due to normalization of $\mathsf{Pr}^{2\text{des}}(i,k)$. The total number of experiments that need to be carried out is then estimated by

$$\langle N_{\text{exp}}\rangle = \sum_{l=1}^{L}\langle N_l\rangle \leq L\left[1+\frac{2d}{L\epsilon^2}\log\left(\frac{2}{\delta}\right)\right]$$
$$\leq 1+\frac{1}{\epsilon^2\delta}+\frac{2d}{\epsilon^2}\log\left(\frac{2}{\delta}\right). \quad (12)$$

This number is sufficient to account for both the sampling error due to finite L and statistical experimental errors in the measurement results. Notably, $\langle N_{\text{exp}}\rangle \sim 2^n$ only; i.e., the average number of experiments to estimate F_{av} scales like that required for characterizing a general pure quantum state [3]. This represents a reduction by a factor of 2^n compared to Refs. [2,3], cf. Table I. These savings come at the expense of (i) obtaining only bounds on the average fidelity when using two classical fidelities F_j or (ii) the necessity to prepare entangled input states when using 2-designs. The latter scales quadratically in n [11]. Even factoring this additional cost in, a Monte Carlo estimation of the average fidelity for a general unitary operation using 2-designs is significantly more efficient than that based on the channel-state isomorphism [2,3].

Resources for Monte Carlo estimation of Clifford gates.—The scaling of $\langle N_{\text{exp}}\rangle$ with the number of qubits changes dramatically for Clifford gates [2,3]. This is due to the property of Clifford gates to map eigenstates of a d-dimensional set of commuting Pauli operators into eigenstates from the same set. The mutually unbiased bases in Eqs. (3) and (5) can be chosen to be such eigenstates [20]. Given a generic eigenstate $|\Psi_i\rangle$ of a commuting set $\mathcal{W}$ of Pauli operators, the characteristic function of a Clifford gate U_{Cl} becomes

$$\chi_{U_{\text{Cl}}}(i,k) = \mathsf{Tr}[W_k U_{\text{Cl}}|\Psi_i\rangle\langle\Psi_i|U_{\text{Cl}}^\dagger] = \mathsf{Tr}[W_k|\Psi_j\rangle\langle\Psi_j|]$$
$$= \begin{cases}\pm 1 & \text{if } W_k \in \mathcal{W}\\ 0 & \text{otherwise}\end{cases}. \quad (13)$$

The relevance distribution for Clifford gates $\mathsf{Pr}(\mathsf{i},\mathsf{k}) \sim [\chi_{U_{\text{Cl}}}(i,k)]^2$ is thus zero for many settings and uniform otherwise. Since settings with $\mathsf{Pr}(\mathsf{i},\mathsf{k}) = 0$ will never be selected, the sampling complexity becomes independent of system size. Calculating $\langle N_l\rangle$ according to Eq. (11) for a uniform relevance distribution, and accounting for the correct normalizations of $\mathsf{Pr}(\mathsf{i},\mathsf{k})$, $\langle N_l\rangle$ is found to be independent of d, $\langle N_l\rangle \leq 1 + 2\log(2/\delta)/(L\epsilon^2)$, for all three approaches. Consequently, also $\langle N_{\text{exp}}\rangle$ does not scale with system size, $\langle N_{\text{exp}}\rangle \leq 1 + 1/(\epsilon^2\delta) + 2\log(2/\delta)/\epsilon^2$, cf. Table II. For Clifford gates, the three approaches require, therefore, a similar, size-independent number of measurements. A difference is found, however, for the number of possible experimental settings. For each input state i, there are only d (instead of d^2) measurement operators W_k with nonzero expectation value. This leads to $N_{\text{setting}} = N_{\text{input}} \times 2^n$ for Clifford gates, cf. Table II. The larger N_{setting} required for approaches A and B in Table II comes with a potentially higher accuracy of the estimate which is, however, limited by the experimental error of state preparation and measurement.

TABLE II. Resources required for determining the average gate error of a Clifford gate ($N_{\text{setting}} = N_{\text{input}} \times 2^n$). Symbols as in Table I.

Approach	$\mathcal{C}_{\text{class}}$	N_{input}	N_{setting}	$\langle N_{\text{exp}}\rangle$
A	$\mathcal{O}(1)$	6^n	$\mathcal{O}(6^n 2^n)$	$\mathcal{O}(1)$
B	$\mathcal{O}(1)$	$2^n(2^n+1)$	$\mathcal{O}(2^{3n})$	$\mathcal{O}(1)$
C	$\mathcal{O}(1)$	2×2^n	$\mathcal{O}(2^{2n})$	$\mathcal{O}(1)$

Conclusions.—We find the number of measurements required to estimate the gate error for a general unitary to scale as 2^n for n qubits. Our reduction by a factor of 2^n compared to the best currently available approach [2,3] comes at the expense of either determining bounds from two classical fidelities instead of the average fidelity itself or allowing for entangled input states. For the classical fidelities, the number of experimental settings that one needs to be able to prepare and the classical computational resources required for the sampling are also reduced by a factor of 2^n. All three approaches are significantly more efficient than traditional process tomography requiring of the order of 2^{4n} measurements and a computational cost scaling as 4^{6n}. For the special case of Clifford gates, we find the number of experiments to be independent of system size, just as in a Monte Carlo estimation based on the channel-state isomorphism [2,3] and randomized benchmarking [6,7].

We have shown earlier [18] that the minimum number of pure input states for device characterization is of the order 2^n. This corresponds to the number of states required by the classical fidelities. Monte Carlo sampling of two classical fidelities therefore realizes a strategy of minimal resources. Our comprehensive classification should allow an experimentalist to choose the most suitable procedure to determine the average fidelity defined in terms of the number of experimental settings, from which a Monte Carlo procedure randomly draws realizations and the actual number of experiments to be carried out.

We would like to thank Daniel Burgarth, Tommaso Calarco, and David Licht for helpful comments. Financial

support from the EC through the FP7-People IEF Marie Curie action Grant No. PIEF-GA-2009-254174 and the MIUR-PRIN Grant No. 2010LLKJBX is gratefully acknowledged. QSTAR is the MPQ, LENS, IIT, UniFi Joint Center for Quantum Science and Technology in Arcetri.

[1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).
[2] S. T. Flammia and Y.-K. Liu, Phys. Rev. Lett. **106**, 230501 (2011).
[3] M. P. da Silva, O. Landon-Cardinal, and D. Poulin, Phys. Rev. Lett. **107**, 210404 (2011).
[4] A. Shabani, R. L. Kosut, M. Mohseni, H. Rabitz, M. A. Broome, M. P. Almeida, A. Fedrizzi, and A. G. White, Phys. Rev. Lett. **106**, 100401 (2011).
[5] C. T. Schmiegelow, A. Bendersky, M. A. Larotonda, and J. P. Paz, Phys. Rev. Lett. **107**, 100502 (2011).
[6] E. Magesan, J. M. Gambetta, and J. Emerson, Phys. Rev. Lett. **106**, 180504 (2011).
[7] E. Magesan, J. M. Gambetta, and J. Emerson, Phys. Rev. A **85**, 042311 (2012).
[8] M. Mohseni and A. T. Rezakhani, Phys. Rev. A **80**, 010101 (2009).
[9] M. Cramer, M. B. Plenio, S. T. Flammia, R. Somma, D. Gross, S. D. Bartlett, O. Landon-Cardinal, D. Poulin, and Y.-K. Liu, Nat. Commun. **1**, 149 (2010).
[10] C. Dankert, R. Cleve, J. Emerson, and E. Livine, Phys. Rev. A **80**, 012304 (2009).
[11] A. Bendersky, F. Pastawski, and J. P. Paz, Phys. Rev. Lett. **100**, 190403 (2008).
[12] W. K. Wootters and B. D. Fields, Ann. Phys. (N.Y.) **191**, 363 (1989).
[13] J. Lawrence, Phys. Rev. A **84**, 022338 (2011).
[14] L. Steffen, M. P. da Silva, A. Fedorov, M. Baur, and A. Wallraff, Phys. Rev. Lett. **108**, 260506 (2012).
[15] C. T. Schmiegelow, M. A. Larotonda, and J. P. Paz, Phys. Rev. Lett. **104**, 123601 (2010).
[16] H. F. Hofmann, Phys. Rev. Lett. **94**, 160504 (2005).
[17] B. P. Lanyon, C. Hempel, D. Nigg, M. Müller, R. Gerritsma, F. Zähringer, P. Schindler, J. T. Barreiro, M. Rambach, G. Kirchmair *et al.*, Science **334**, 57 (2011).
[18] D. M. Reich, G. Gualdi, and C. P. Koch, Phys. Rev. A **88**, 042309 (2013).
[19] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. A **60**, 1888 (1999).
[20] J. Lawrence, C. Brukner, and A. Zeilinger, Phys. Rev. A **65**, 032320 (2002).
[21] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.111.200401 for a detailed proof of the proper normalization of the relevance distributions and the variance of X_κ being smaller than one.