Hybrid benchmarking of arbitrary quantum gates

Tobias Chasseur,[1] Daniel M. Reich,[2] Christiane P. Koch,[2] and Frank K. Wilhelm[1]
[1]*Theoretical Physics, Saarland University, 66123 Saarbrücken, Germany*
[2]*Theoretische Physik, Universität Kassel, Heinrich-Plett-Strasse 40, 34132 Kassel, Germany*
(Received 28 June 2016; published 27 June 2017)

We present a protocol for interleaved randomized benchmarking of arbitrary quantum gates using Monte Carlo sampling of quantum states. It is generally applicable, including non-Clifford gates while preserving key advantages of randomized benchmarking such as error amplification as well as independence from state preparation and measurement errors. This property is crucial for implementations in many contemporary systems. Although the protocol scales exponentially in the number of qubits, it is superior to direct Monte Carlo sampling of the average gate fidelity in both the total number of experiments by orders of magnitude and savings in classical preprocessing, that are exponential.

DOI: 10.1103/PhysRevA.95.062335

A central goal of quantum information science is to engineer a physical system capable of functioning as a scalable quantum computer that systematically outperforms classical computers in certain applications. To this end, it is imperative to drive arbitrary unitary evolution in a suitable quantum system consisting of $n \gg 1$ qubits and to benchmark the implementation of that operation.

Efficient benchmarking protocols, i.e., protocols that scale at most polynomially in n, are available for quantum operations in the Clifford group $\mathcal{C}$ [1–4], an important subset of quantum operations [5,6]. In particular, randomized benchmarking (RB) is a method to estimate the average error of the Clifford group based on the fidelity of random Clifford gate sequences [1,2]. RB has proven itself as a popular and experimentally viable approach not only because of its scaling properties but also due to its independence from state preparation and measurement (SPAM) errors [1,2]. For individual gates of the Clifford group, the fidelity can be estimated with interleaved randomized benchmarking (IRB) [7]. The remarkable RB construction hinges on the Clifford group elements being distributed sufficiently uniformly on the special unitary group $\mathrm{SU}(d = 2^n)$. A central prerequisite for the scalability of RB is that $\mathcal{C}$ can be simulated efficiently on a classical computer [8]. However, for the same reason, quantum algorithms based on those Clifford gates alone cannot outperform a classical computer. To realize the full potential of quantum computation, one has to access the full unitary group which is generated by $\mathcal{C}$ and one additional non-Clifford gate, e.g., a single qubit gate such as the $\pi/8$ gate. While IRB, for example, can be generalized to an arbitrary gate, the fidelity estimation becomes highly challenging [9]: Simulation and inversion of the sequence become increasingly inefficient as alternating Clifford and, e.g., $\pi/8$ gates generate the full $\mathrm{SU}(2^n)$ [8]. In order to sample all gates, including those not in $\mathcal{C}$, one needs to rely on strategies the experimental and classical resources of which scale exponentially.

One such generally applicable protocol is given by Monte Carlo sampling of the average gate fidelity which allows for the validation of arbitrary quantum gates [3,4,10]. It requires significantly less resources than the canonical approach, which is to extract this information from full quantum process tomography. However, Monte Carlo sampling is limited by SPAM errors, which for many physical systems can overshadow the gate error. Moreover its scaling in both experimental and classical resources, although favorable compared to process tomography, is still exponential in n. This poses the question whether it is possible to combine the generality of Monte Carlo sampling, i.e., going beyond the Clifford group, with the experimental advantages of RB.

Here, we answer this question and demonstrate the benefit of combining both methods. We show how arbitrary gates can be benchmarked by replacing the inverting gate at the end of each IRB sequence with Monte Carlo sampling of the resulting quantum state. Our approach outperforms direct Monte Carlo sampling of the average gate fidelity regarding the number of measurements and yields an exponential saving in classical computational resources while retaining the independence on SPAM errors. Therefore it enables the benchmarking of arbitrary gates in experimental settings.

We first briefly review the original RB protocol [1,2] as well as IRB [7,9]. RB provides an estimate for the average fidelity of a unitary two-design such as the Clifford group based on the idea that random sequences of Clifford gates also randomize the effect of error channels, turning them depolarizing. For every sequence of y Clifford gates $\hat{\mathrm{C}}_j$, $1 \leqslant j \leqslant y$, there is a unique Clifford gate $\hat{\mathrm{C}}_{y+1}$ inverting the sequence which can be efficiently found via the Gottesmann-Knill theorem. By applying the sequence and its inverse to an initial state $\hat{\rho}_0$ and measuring the survival probability of that state, the sequence fidelity is accessible experimentally. Averaging over all possible sequences and making the additional assumption of a gate independent error channel Λ result in an average sequence fidelity

$$\Phi_y = \frac{1}{\sharp\mathcal{C}^y} \sum_{\{C_j\}\in\mathcal{C}^y} \mathrm{Tr}\left[\hat{\rho}_0 \left(C_{y+1} \prod_{j=y}^{1} (\Lambda C_j)\right)(\hat{\rho}_0)\right], \qquad (1)$$

where $\sharp$ denotes cardinality and $C(\rho) = \hat{\mathrm{C}}\hat{\rho}\hat{\mathrm{C}}^\dagger$ is the operation of the Clifford gate. The reverse order of the product ensures the correct arrangement of the gates with the earlier operation applied to the state appearing on the right of the latter operation.

Equation (1) can be rewritten as $\Phi_y \equiv \mathrm{Tr}[\hat{\rho}_0 \Lambda^y_{\mathrm{twirl}}(\hat{\rho}_0)]$ where Λ_{twirl}—the twirl of the error channel Λ over the Clifford group [6]—is completely depolarizing, i.e., $\Lambda_{\mathrm{twirl}}(\hat{\rho}) = p\hat{\rho} +$

$\frac{1-p}{d}\mathbb{1}$ with decay parameter p [11]. The average fidelity associated with the channel Λ_{twirl} then is $\Phi = p + \frac{1-p}{d}$; the average sequence fidelity becomes

$$\Phi_y = \mathrm{Tr}\left[\hat{\rho}_0\left(p^y\hat{\rho}_0 + \frac{1-p^y}{d}\mathbb{1}\right)\right] = \frac{d-1}{d}p^y + \frac{1}{d}. \quad (2)$$

To access p, one has to estimate Φ_y for several sequence lengths y by sampling over a (in practice small [9,12]) subset of possible sequences for each y; p and hence Φ are derived by fitting the experimental data to an exponential decay. Incorporating the error channels for SPAM and C_{y+1} leads to $\Phi_y = Ap^y + B$, which leaves the exponential decay unchanged and therefore yields a protocol robust against imperfect SPAM [1,2]. An important extension to RB is the IRB protocol that sets limits to the fidelity of a single Clifford gate $\hat{\mathrm{V}}_{\mathcal{C}}$ using the fidelity of this gate interleaved with a random sequence, i.e., of the combined error channel $\Lambda_V\Lambda_{\mathcal{C}}$, in comparison with the fidelity obtained for the Clifford group [7]. This assessment of individual gates not only provides information of possible error sources but can be used directly for model free optimal control in an experiment [13,14].

Potential loopholes in RB and IRB such as gate dependent errors and leakage can be accounted for by considering linear maps acting on quantum channels instead of just quantum channels [9]. Specifically, the extension of the sequence length by one acts as a linear map T on the operator on $\hat{\rho}$ representing the shorter sequence (which depends on V for IRB). The average sequence fidelity is a linear functional of the y-fold product of T resulting in a multiexponential fidelity decay:

$$\Phi_y = \sum_i a_i\lambda_i^y\,. \quad (3)$$

The eigenvalues, λ_i, of the linear operator T are close to real and their absolute values are smaller than or equal to 1. The resulting fidelity decay can be typically fitted using just a few different exponential decays [9].

In this paper we go beyond the Clifford group. To do so, we rely on the technique of Ref. [9] that does not depend on $V_{\mathcal{C}}$ being an element of the Clifford group. If it is not, the unitary matrix representing an ideal implementation of the sequence can be quite general since $\mathcal{C}$ and V generate a dense subset of the whole special unitary group. The construction of the inverse C_{y+1} would be highly challenging and defeat the concept of performing quantum computation using a restricted set of gates.

Alternatively, one could be content to approximate the inverting gate using the Solovay-Kitaev theorem. However, this turns out to be inadequate for the following reason: The theorem states that any gate can be composed out of a small number l of gates depending only logarithmically on the permitted inaccuracy but exponentially on the number of qubits [15,16]. For RB to be reliable, the error rate ε_{y+1} associated with C_{y+1} should be much smaller than the error of the sequence. Since both sequence and inverting gate are composed of the same gate set, this is roughly equivalent to $l \ll y$. Satisfying this is possible only for ε_V and $\varepsilon_{\mathcal{C}}$ sufficiently small so that y is large while the Hilbert-space dimension must be kept small as it enters the sequence length exponentially in the Solovay-Kitaev algorithm. In other words, satisfying $l \ll y$ implies the ability to implement an arbitrary quantum gate to a relatively high precision, i.e., availability of a universal quantum computer as a starting point.

To overcome the limitations of these ideas, we present a pragmatic approach to the problem. Consider the fidelity of a specific sequence $\mathbf{y}$, written as a vector of gates:

$$\Phi_{\mathbf{y}} = \mathrm{Tr}\left[\rho_0 C_{y+1}\prod_{j=y}^{1}(V\Lambda_V\Lambda_j C_j)(\rho_0)\right] \equiv \mathrm{Tr}\left[\rho_{\text{id}}^{\mathbf{y}}\rho_{\text{act}}^{\mathbf{y}}\right], \quad (4)$$

where $\rho_{\text{id}}^{\mathbf{y}} = {C_{y+1}}^{-1}(\rho_0) = (\prod_{j=y}^{1} C_j)(\rho_0)$ is the state ideally generated by the sequence and determined on a classical computer, and $\rho_{\text{act}}^{\mathbf{y}}$ is the one actually realized by applying the gates V and C_j (including their errors Λ_V and Λ_j) in the experiment. Equation (4) is of the form used in Refs. [3,4,10] to estimate the overlap of two states via *Monte Carlo sampling*. Employing the notation of Ref. [10], the states are rewritten in the basis of the generalized Pauli matrices on n qubits normalized for the canonical scalar product defined by the (un-normalized) trace, $\mathcal{W} = \frac{1}{\sqrt{d}}\mathcal{P}^{\otimes n}$:

$$\begin{aligned}\Phi_{\mathbf{y}} &= \mathrm{Tr}\left[\rho_{\text{id}}^{\mathbf{y}}\rho_{\text{act}}^{\mathbf{y}}\right] = \sum_k^{d^2}\mathrm{Tr}[W_k\rho_{\text{id}}]\mathrm{Tr}[W_k\rho_{\text{act}}]\\ &\equiv \sum_k^{d^2}\chi_{\text{id}}(k)\chi_{\text{act}}(k) = \sum_k^{d^2}\chi_{\text{id}}(k)^2\frac{\chi_{\text{act}}(k)}{\chi_{\text{id}}(k)} \equiv \sum_k^{d^2}\Pr(k)X_k\,,\end{aligned} \quad (5)$$

where $\Pr(k) = \chi_{\text{id}}(k)^2$ and $X_k = \frac{\chi_{\text{act}}(k)}{\chi_{\text{id}}(k)}$. $\sum_k^{d^2}\Pr(k) = 1$ since $\sum_k\chi_{\text{id}}(k)^2 = \mathrm{Tr}[\rho_{\text{id}}^2]$ and ρ_{id} being a pure state. Therefore $\Pr(k)$ can be used as a sampling probability where the expectation value of the corresponding sampling is the desired fidelity $\Phi_{\mathbf{y}}$.

This is the core of our approach: Instead of actually implementing the gate that inverts the random sequence and measuring the error on identity, we treat $\Phi_{\mathbf{y}}$ as a state fidelity which is estimated with Monte Carlo sampling. Following Eq. (5), this consists in choosing a total of L Pauli measurement operators $W_{k_l} \in \mathcal{W}$, $1 \leqslant l \leqslant L$, according to the sampling probability $\Pr(k)$ and measuring W_{k_l} (and hence X_{k_l}) N_l times. We summarize the IRB protocol with Monte Carlo sampling of quantum states as follows.

(1) Perform standard IRB to estimate the average error of the Clifford group $\epsilon_{\mathcal{C}}$ as a reference point.

(2) Choose q different sequence lengths y such that the sequence fidelities Φ_y can be assumed to provide a reliable fit. This means the Φ_y shall be close neither to 1 nor to the fidelity limit for long sequences.

(3) For each selected sequence length, choose m different sequences $\mathbf{y}$ of random Clifford gates interleaved with the gate V. They are used to estimate the average fidelity Φ_y by comparing the actual and ideal state [see Eq. (4)] via Monte Carlo sampling.

(4) Determine the ideal state on a classical computer, i.e., apply $2y$ unitary matrices onto the pure initial state vector. This scales as $O(yd^2)$ as it cannot be done efficiently, since V is not necessarily a Clifford gate.

(5) Choose L measurement operators W_k at random, following the distribution $\Pr(k)$ defined in Eq. (5).

(6) For each chosen measurement operator apply the sequence $\mathbf{y}$ and measure W_k. This is repeated N_l times.

(7) Determine an estimate for the sequence fidelities Φ_y by averaging over all N_l measurements, the X_k for all L measurement operators W_{k_l}, and the m different sequences as given by Eq. (5).

(8) Fit Φ_y to the multiexponential decay, Eq. (3), analogously to the original IRB, and derive the combined average error as $\frac{\Phi_{y=1}}{\Phi_{y=0}}$.

(9) Calculate the average error of the arbitrary n qubit gate V as $\varepsilon_V = \varepsilon_{\mathcal{C}\times V} - \varepsilon_{\mathcal{C}}$ and estimate the lower and upper bounds as $\max(0,(\sqrt{\varepsilon_{\mathcal{C}\times V}} - \sqrt{\varepsilon_{\mathcal{C}}})^2$ and $(\sqrt{\varepsilon_{\mathcal{C}\times V}} + \sqrt{\varepsilon_{\mathcal{C}}})^2$ as in the original IRB.

The parameters of the protocol are chosen as follows: A valid fidelity estimate via RB requires sufficient experimental data for a fit to a (multi)exponential decay; hence q different values for y, all provided with a substantiated estimate for Φ_y. Because it is sufficient to fit to only a few decays, a rather small q suffices. The amount m of different sequences for each value of y can be upper bounded by a global constant using the leading order in gate errors (see Refs. [9,12]), yielding m not larger than 100. Higher-order corrections to the uncertainty in the error per gate originating from finite m can be bounded using the fact that Φ_y lies in the range [0,1] and invoking Hoeffding's inequality [17]. We choose sequence lengths y in a way that the error is neither too small to be measured efficiently nor so big that the decaying terms are already close to zero. This condition is satisfied for

$$\varepsilon y = O(1), \tag{6}$$

as can easily be seen using the simplified model of a single decay.

In Monte Carlo sampling, there are two sources for inaccurate fidelity assessment, namely, the sampling inaccuracy due to (a) the incomplete subset of the measurement operators and (b) the finite number of measurements. The inaccuracies can be bounded by Chebyshev's and Hoeffding's inequality, respectively, to be allowed to exceed $\frac{\alpha}{2}$ with a probability of at most $\frac{\delta}{2}$. For a given error bound, this leads to an estimate of the total number of experiments [3,4,10].

The sampling inaccuracy (a) is bounded by Chebyshev's inequality, which provides an upper limit to the probability of deviating from the mean value of a distribution, depending on its standard deviation:

$$\Pr\left(|Z - [Z]| \geqslant \frac{\sigma_Z}{\sqrt{\delta}}\right) \leqslant \delta. \tag{7}$$

Here, $Z \equiv \frac{1}{L}\sum_{l=1} X_{k_l}$ is the fidelity estimate obtained by the random choice of measurement operators W_{k_l} and $[Z]$ its classical expectation value, i.e., $\Phi_{\mathbf{y}}$. The variance can be estimated as

$$\begin{aligned}\sigma_Z^2 = [Z^2] - [Z]^2 &= \sum_{l=1}^{L}\sum_{k_l}\Pr\left(\frac{X_{k_l}}{L}\right)^2 - \Phi_{\mathbf{y}}^2 \\ &\leqslant \frac{1}{L}\sum_k \chi_{\mathrm{act}}(k)^2 = \frac{1}{L}\mathrm{Tr}\left[\rho_{\mathrm{act}}^2\right] \leqslant \frac{1}{L},\end{aligned} \tag{8}$$

using the fact that ρ_{act} is a convex sum of projectors. Thus

$$\Pr\left[|Z - \Phi_{\mathbf{y}}| \geqslant \sqrt{\frac{2}{L\delta}}\right] \leqslant \frac{\delta}{2} \tag{9}$$

and the choice $L = \lceil 8/(\alpha^2\delta)\rceil$ ensures the intended inequality, where the outer brackets denote the ceiling function. To limit the deviation (b) due to a finite number of measurements one relies on Hoeffding's inequality:

$$\Pr(|S - \langle S\rangle| \geqslant \alpha/2) \leqslant 2\exp\left(-\frac{\alpha^2}{2\sum_i (b_i - a_i)^2}\right). \tag{10}$$

S is the sum over random variables with outcomes in the range $[a_i,b_i]$, given by the adequately normalized sum of all $\sum_l N_l$ single shot measurements, and $\langle S\rangle = Z$. Since the measurement outcomes of Pauli matrices are bimodal, they are situated at the boundaries of the respective range $[a_i,b_i]$. Therefore, the range over the variance ratio is most suitable for Hoeffding's inequality. To ensure that the probability to exceed $\frac{\alpha}{2}$ is at most $\frac{\delta}{2}$, it suffices to demand

$$\frac{\delta}{2} \overset{!}{\geqslant} 2\exp\left(-\frac{\alpha^2}{2\sum_l 4N_l d^{-1}[LN_l\chi_{\mathrm{id}}(k)]^{-2}}\right), \tag{11}$$

which, with the natural choice $N_l \propto \chi_{\mathrm{id}}(k)^{-2}$, is satisfied for

$$N_l = \left\lceil \frac{8}{dL\alpha^2\chi_{\mathrm{id}}(k)^2}\ln\left(\frac{4}{\delta}\right)\right\rceil. \tag{12}$$

Compared to Refs. [3,4,10], the total inaccuracy α as well as the probability δ of exceeding it were chosen smaller by a factor of 2 to simplify the further treatment.

The classical average over the total number of experiments can be estimated as follows using Eq. (12):

$$\begin{aligned}[N_{\mathrm{exp}}] &= L\sum_{k=1}^{d^2}\Pr(k)N_k \\ &\leqslant L\left[1 + \frac{8d}{L\alpha^2}\ln\left(\frac{4}{\delta}\right)\right] \\ &\leqslant 1 + \frac{8}{\alpha^2\delta} + \frac{8d}{\alpha^2}\ln\left(\frac{4}{\delta}\right).\end{aligned} \tag{13}$$

Equation (13) is also valid for direct Monte Carlo sampling of the average gate fidelity and represents an exponential speedup in the number of qubits compared to full process tomography, which scales as $O(d^4)$ [10]. An important aspect is the scaling with $\frac{1}{\alpha^2}$. It is key to the advantageous scaling of IRB with Monte Carlo sampling of quantum states in comparison with direct Monte Carlo sampling of the average fidelity as shown below.

For the resource estimate, we aim for an inaccuracy of fidelity measurements that is one order of magnitude smaller than the error rate ε. Average gate fidelities are not fundamental quantities of physics but estimators on how well a quantum algorithm composed of a set of gates performs. Therefore any attempt at an overly precise characterization of gate errors does not yield a valuable gain in information. In addition, the systematic uncertainty α_{IRB} of IRB caused by Clifford gate errors limits the accuracy that can reasonably be achieved, even more so for other methods not robust against SPAM

errors. Based on Eq. (6), $\Phi_y \sim 1 - y\varepsilon$ such that uncertainties in its estimation affect the estimate of ε roughly with a factor of $\frac{1}{y}$. Therefore relative errors in Φ_y approximately translate to relative errors in ε. Using the above statement and the fact that the inaccuracy of an IRB based estimation α_{IRB} is aimed to be close to ε, one chooses an inaccuracy $\alpha_{\mathrm{MC}}(y)$ for the Monte Carlo sampling of sequence fidelities Φ_y that result in an estimation without unnecessary additional precision compared to α_{IRB}. It scales linearly with εy, which is on the order of 1. Therefore $\alpha_{\mathrm{MC}}(y)$ varies distinctly but not excessively over the q different sequence lengths y but depends on neither the error rate ε nor the Hilbert-space dimension $d = 2^n$. For the sake of simplicity, let α_{MC} be defined as an effective average value for $\alpha_{\mathrm{MC}}(y)$ setting an average as to what precision each sequence fidelity has to be assessed. α_{MC} as a system independent constant of the protocol can safely be assumed to not deceed $10^{-1.5}$.

The above derivation of $\alpha_{\mathrm{MC}}(y)$ ensures the required accuracy for each of the $q \times m$ single sequence fidelities rather than just for the resulting estimate for ε. This provides a reasonable fit to the decay function as each data point provides sufficient accuracy. Exploiting that in a more rigorous way may result in an improvement of prefactors but cannot improve the scaling since q and m are largely system independent [9,12].

The total number of experiments then adds up to

$$[N_{\mathrm{exp}}] \leqslant qm\left[1 + \frac{8}{\alpha_{\mathrm{MC}}^2\delta} + \frac{8d}{\alpha_{\mathrm{MC}}^2}\ln\left(\frac{4}{\delta}\right)\right], \tag{14}$$

which differs by a factor of $qm\frac{\alpha^2}{\alpha_{\mathrm{MC}}^2}$ compared to direct Monte Carlo sampling of the average fidelity [10]. Translating this factor into numbers relating to recent advances in the implementation of quantum gates as well as the error threshold for quantum computing highlights the advantage of our protocol. A specific set of values taking into account recent experimental results [18–20] corresponds to $q = 20$, $m = 50$, and $\varepsilon = 10^{-3}$ based on relatively high error rates of two qubit gates. These values yield $\alpha = 10^{-4}$ and two orders of magnitude of improvement in the total number of experiments via the above factor.

Another concern regarding scalability is the use of classical computational resources. Although more easily accessible, classical resources are not infinite and therefore become relevant eventually, especially for Monte Carlo sampling where classical resources scale exponentially with a higher exponent than the number of experiments. The sampling of measurement operators can be done using conditional probabilities, scaling with n^2d^2 for states and n^2d^4 for processes and hence outperforming the naive approach of calculating all $\Pr(k)$ [4,10]. Accounting also for the necessity to calculate ρ_{id} for each sequence, the classical resources needed for our protocol scale asa

$$N_{\mathrm{class}} = O\left(\frac{qm}{\alpha_{\mathrm{MC}}^2}\left(\frac{d^2}{\varepsilon} + n^2d^2\right)\right), \tag{15}$$

compared to $O(\frac{1}{\alpha^2}n^2d^4)$ for direct Monte Carlo sampling of the average gate fidelity. Hence, we obtain an exponential speedup of $O(d^2)$ in classical resources in addition to the reduction of the number of experiments.

Combining the currently best but individually restricted methods for estimating quantum fidelities (interleaved randomized benchmarking and Monte Carlo sampling), we have extended the former to arbitrary quantum operations, outside of the Clifford group, while reducing the enormous overheads and avoiding the SPAM dependence associated with the latter. The extension to non-Clifford gates is made possible by treating the RB sequence fidelity as a state fidelity that can be estimated with Monte Carlo sampling. This avoids the actual accurate physical implementation of the inverting gate in the RB sequence which, for a non-Clifford gate, would require availability of a universal quantum computer. Our protocol inherits from IRB robustness with respect to SPAM errors; for current experimental settings this can completely mask the actual error channel. As a conclusion the resulting hybrid algorithm is a viable tool for SPAM-independent, robust benchmarking of arbitrary quantum gates. While nonexponential scaling is still out of reach and might well be impossible, the proposed protocol reduces the total number of experiments compared to direct Monte Carlo sampling of the gate fidelity due to error amplification and yields exponential savings in the classical preprocessing resources.

We acknowledge travel support from the European Union through Coordination Action on Optimal Control of Quantum Systems (QUAINT) as well as funding through Scalable Superconducting Processors for Entangled Quantum Information Technology (SCALEQIT) and a Google Faculty Research Award.

[1] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, Phys. Rev. A **77**, 012307 (2008).

[2] E. Magesan, J. M. Gambetta, and J. Emerson, Phys. Rev. Lett. **106**, 180504 (2011).

[3] S. T. Flammia and Y. K. Liu, Phys. Rev. Lett. **106**, 230501 (2011).

[4] M. P. da Silva, O. Landon-Cardinal, and D. Poulin, Phys. Rev. Lett. **107**, 210404 (2011).

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

[6] C. Dankert, R. Cleve, J. Emerson, and E. Livine, Phys. Rev. A **80**, 012304 (2009).

[7] E. Magesan, J. M. Gambetta, B. R. Johnson, C. A. Ryan, J. M. Chow, S. T. Merkel, M. P. da Silva, G. A. Keefe, M. B. Rothwell, T. A. Ohki, M. B. Ketchen, and M. Steffen, Phys. Rev. Lett. **109**, 080505 (2012).

[8] D. Gottesman, in *Group 22: Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics*, edited by S. Corney, R. Delbourgo, and P. Jarvis (International, Cambridge, MA, 1999), pp. 32–43.

[9] T. Chasseur and F. K. Wilhelm, Phys. Rev. A **92**, 042333 (2015).

[10] D. M. Reich, G. Gualdi, and C. P. Koch, Phys. Rev. Lett. **111**, 200401 (2013).

[11] J. M. Gambetta, A. D. Córcoles, S. T. Merkel, B. R. Johnson, J. A. Smolin, J. M. Chow, C. A. Ryan, C. Rigetti, S. Poletto,

T. A. Ohki, M. B. Ketchen, and M. Steffen, Phys. Rev. Lett. **109**, 240504 (2012).
[12] J. Wallman and S. Flammia, New J. Phys. **16**, 103032 (2014).
[13] D. J. Egger and F. K. Wilhelm, Phys. Rev. Lett. **112**, 240503 (2014).
[14] J. Kelly, R. Barends, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, I. C. Hoi, E. Jeffrey, A. Megrant, J. Mutus, C. Neill, P. J. J. O'Malley, C. Quintana, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. N. Cleland, and J. M. Martinis, Phys. Rev. Lett. **112**, 240504 (2014).
[15] 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).
[16] E. Knill, arXiv:quant-ph/9508006.
[17] E. Magesan, J. M. Gambetta, and J. Emerson, Phys. Rev. A **85**, 042311 (2012).
[18] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. White, J. Mutus, A. Fowler, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, C. Neill, P. O'Malley, P. Roushan, A. Vainsencher, J. Wenner, A. Korotkov, A. Cleland, and J. Martinis, Nature (London) **508**, 500 (2014).
[19] T. P. Harty, D. T. C. Allcock, C. J. Ballance, L. Guidoni, H. A. Janacek, N. M. Linke, D. N. Stacey, and D. M. Lucas, Phys. Rev. Lett. **113**, 220501 (2014).
[20] A. D. Córcoles, J. M. Gambetta, J. M. Chow, J. A. Smolin, M. Ware, J. Strand, B. L. T. Plourde, and M. Steffen, Phys. Rev. A **87**, 030301 (2013).