Scalable and Robust Randomized Benchmarking of Quantum Processes

Easwar Magesan, J. M. Gambetta, and Joseph Emerson
Institute for Quantum Computing and Department of Applied Mathematics, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada
(Received 19 September 2010; revised manuscript received 27 January 2011; published 6 May 2011)

In this Letter we propose a fully scalable randomized benchmarking protocol for quantum information processors. We prove that the protocol provides an efficient and reliable estimate of the average error-rate for a set operations (gates) under a very general noise model that allows for both time and gate-dependent errors. In particular we obtain a sequence of fitting models for the observable fidelity decay as a function of a (convergent) perturbative expansion of the gate errors about the mean error. We illustrate the protocol through numerical examples.

DOI: 10.1103/PhysRevLett.106.180504 PACS numbers: 03.67.Pp, 03.65.Wj, 03.67.Lx

The challenge of characterizing the level of coherent control over a quantum system is a central problem in contemporary experimental physics and a fundamental task in the design of quantum information processing devices. Full characterization of any quantum process is possible though quantum process tomography (QPT) [1] which has been successfully applied to the measurement of up to three coupled qubits (two-level systems) in NMR [2,3], linear optics [4], atomic ions [5] and superconducting qubits [6,7]. However, QPT suffers from two shortcomings: the first is the often unrealistic assumption that measurements and state preparations admit much lower errors than the process which is being characterized; the second is that the number of experiments required grows exponentially with the number of qubits, and hence QPT becomes impractical for even moderately large systems.

Recently there has been significant interest in scalable methods for partial characterization of the noise affecting a quantum process [8–12], and, in particular, randomized benchmarking (RB) protocols [8,13]. For example, the RB protocol proposed in Ref. [13] is *conjectured* to provide an experimental means for estimating the average error rate for single-qubit Clifford gates. The simplicity of this protocol has motivated experimental implementations in atomic ions [13,14], NMR [15], superconducting qubits [6,16], and atoms in optical lattices [17].

However the above RB protocols leave open the question of how to implement a scalable multiqubit protocol in an unambigious manner. Another open issue is to determine sufficient conditions under which a RB protocol gives a reliable estimate; it is easy to construct examples where the decay rate estimated via RB protocols is not reliable. An unrealistic but simple example is when the error is gate-dependent and equal to the exact inverse of the target gate. In this example, implementing any gate results in the identity operation and so the error rate given by the protocol is always equal to zero, whereas in actuality there is substantial error on each gate.

In this Letter we propose a *scalable* randomized benchmarking protocol for Clifford gates on n-qubit quantum processors, requiring at most $O(n^2)$ quantum gates, $O(n^4)$ cost in classical preprocessing (to select each gate-sequence), and a number of single-shot repetitions that is independent of n. We give a rigorous analysis of the form of the observed fidelity decay for arbitrary time- and gate-dependent errors by developing a perturbative expansion of the errors about the average error. We prove that for time-independent and gate-independent errors the fidelity decay is exponential at a rate that determines the average error-rate of the noise model. By deriving conditions for when our perturbative expansion is convergent we are also able to prove that the protocol gives a reliable estimate of the average error-rate if the gate-dependent component of the error is sufficiently weak. In particular, we derive a fitting model for the observed fidelity decay which includes first-order correction terms due to gate dependence in the errors. This formula shows that weak gate dependence in the errors can lead to a deviation from the exponential decay, defining a partial test for such effects in the noise. Moreover, the model accounts for state preparation and measurement errors as they show up as independent fit parameters in the formula.

Our motivation for restricting to the Clifford group, as opposed to the full unitary group, is that our protocol is provably scalable in this case. In contrast, benchmarking the full unitary group is not a scalable task since generating a uniformly random unitary is an inefficient task. In spite of this limitation our method still provides significant information regarding the reliability of a quantum information processor for several independent reasons. First, the Clifford group is generated by a two-qubit entangling gate (such as the C-NOT) coupled with a set of single-qubit gates (such as the Hadamard and $\frac{\pi}{4}$ gates), and the unitary group can be generated by adding just one additional single-qubit gate not in the Clifford group (such as the $\frac{\pi}{8}$ gate). For many implementations the error-rate for a particular single-qubit rotation would not be expected to differ

 © 2011 American Physical Society

significantly from single-qubit rotations in the Clifford set. Furthermore, quantum computation is possible using only Clifford gates combined with ancilla states and measurements [18]. Lastly, most fault-tolerant schemes are based on stabilizer codes [19], which implies fault-tolerant quantum computation will likely be dominated by Clifford gates. Hence, scalably benchmarking Clifford gates is important for error estimation in universal quantum computation.

Our RB protocol consists of the following steps (repeated for several values of m):

Step 1—Generate a sequence of $m+1$ quantum operations with the first m chosen uniformly at random from the Clifford group and the final operation chosen so that the net sequence (if error-free) is the identity operation. Since the Cliffords form a group, this $m+1$th correction gate will also be a Clifford element. In practice each operation $\mathcal{C}_{i_j}$ will have some associated error and the entire sequence can be modeled by $\mathcal{S}_{\mathbf{i_m}} = \bigcirc_{j=1}^{m+1}(\Lambda_{i_j,j} \circ \mathcal{C}_{i_j})$, where $\mathbf{i_m}$ is the m-tuple $(i_1, \ldots, i_m)$ and i_{m+1} is uniquely determined by $\mathbf{i_m}$. In the above, $\Lambda_{i_j,j}$ is a quantum operation representing the error associated with the operation $\mathcal{C}_{i_j}$, and is allowed to depend independently on the time-step j. This is a very general noise model—the only assumption is that correlations in the noise are negligible on time scales longer than the time of the operation $\mathcal{C}_{i_j}$ (so that $\Lambda_{i_j,j}$ does not depend on earlier operations). This assumption becomes very well-motivated as n increases (see below).

Step 2—For each sequence measure the survival probability $\mathrm{Tr}[E_\psi \mathcal{S}_{\mathbf{i_m}}(\rho_\psi)]$. Here ρ_ψ is the initial state taking into account preparation errors and E_ψ is the POVM element that takes into account measurement errors. In the ideal noise-free case $\rho_\psi = E_\psi = |\psi\rangle\langle\psi|$.

Step 3—Average over random realizations of the sequence to find the averaged sequence fidelity,

$$F_{\mathrm{seq}}(m, \psi) = \mathrm{Tr}[E_\psi \mathcal{S}_m(\rho_\psi)], \tag{1}$$

where $\mathcal{S}_m$ is the average sequence operation

$$\mathcal{S}_m = \frac{1}{|\{\mathbf{i_m}\}|} \sum_{\mathbf{i_m}}^{|\{\mathbf{i_m}\}|} \mathcal{S}_{\mathbf{i_m}}. \tag{2}$$

Step 4—Fit the results for the averaged sequence fidelity (1) to the model

$$F_{\mathrm{seq}}^{(1)}(m, \psi) = A_1 p^m + B_1 + C_1(m-1)(q-p^2)p^{m-2} \tag{3}$$

derived later. A_1, B_1, and C_1 absorb state preparation and measurement errors as well as the error on the final gate. The average error rate r, defined below, is obtained from the relation $r = 1 - p - (1-p)/d$. For gate-independent errors the fitting results simplify to

$$F_{\mathrm{seq}}^{(0)}(m, \psi) = A_0 p^m + B_0, \tag{4}$$

where A_0 and B_0 absorb state preparation and measurement errors as well as the error on the final gate. Hence a nonzero fit value for q-p^2 is an indication of *weak gate dependence* in the errors.

The parameters r and p determined by the above protocol are defined as follows. Define Λ to be the average error for the set of error operators $\{\Lambda_{i_j,j}\}$,

$$\Lambda \equiv \frac{1}{|\{(i_j, j)\}|} \sum_{i_j,j}^{|\{(i_j,j)\}|} \Lambda_{i_j,j}. \tag{5}$$

The standard experimental figure of merit for a noise model Λ is the average gate fidelity $F_{\mathrm{ave}} = \int d\psi \langle\psi|\Lambda(|\psi\rangle\langle\psi|)|\psi\rangle$, which is the survival probability averaged over all pure input states. This naturally defines the average error rate as $r \equiv 1 - F_{\mathrm{ave}}$. Note that F_{ave} is equivalent to the average process fidelity and entanglement fidelity, up to normalization factors [18]. To define p, a key result in our analysis is that the Clifford group is a unitary 2 design. Therefore "twirling" Λ over the Clifford group gives $\frac{1}{K}\sum_l \mathcal{C}_l^\dagger \circ \Lambda \circ \mathcal{C}_l = \Lambda_{\mathrm{dep}}$, where $\Lambda_{\mathrm{dep}}(\rho) = p\rho + (1-p)\frac{\mathbb{1}}{d}$ is the unique depolarizing channel with the same average fidelity as Λ [20]. Hence, $F_{\mathrm{ave}} = p + (1-p)/d$, which relates the fidelity decay parameter p to the average error rate r as given above.

Derivation.—In the idealized case of gate and time independent errors we have $\Lambda_{i_j,j} = \Lambda$ for each i_j, j. Repeated application of the identity operation $\mathcal{C}_{i_j} \circ \mathcal{C}_{i_j}^\dagger$ in $\mathcal{S}_{\mathbf{i_m}}$ gives $\mathcal{S}_{\mathbf{i_m}} = \Lambda \circ \bigcirc_{j=1}^{m}(\mathcal{D}_{i_j}^\dagger \circ \Lambda \circ \mathcal{D}_{i_j})$ where we have used $\mathcal{C}_{i_{m+1}} \circ \cdots \circ \mathcal{C}_{i_1} = \mathbb{1}$ and for each j defined a new gate $\mathcal{D}_{i_j} = \mathcal{C}_{i_j} \circ \cdots \circ \mathcal{C}_{i_1}$ that is independent from the gates which where performed at time-steps earlier than j ($\mathcal{C}_{i_{j-1}}$, etc.). Substituting this into Eqs. (1) and (2) the average sequence fidelity is $F_{\mathrm{seq}}^{(0)}(m, \psi) = \mathrm{Tr}[E_\psi \Lambda \circ \Lambda_{\mathrm{twirl}}^{\circ m}(\rho_\psi)]$, where $\Lambda_{\mathrm{twirl}} = \sum_{i_j} \tilde{\Lambda}_{i_j}/K$ with $\tilde{\Lambda}_{i_j} = \mathcal{D}_{i_j}^\dagger \circ \Lambda \circ \mathcal{D}_{i_j}$. Thus we are left with an m-fold composition of gate-independent twirls over the Clifford group, and $F_{\mathrm{seq}}^{(0)}(m, \psi)$ reduces to Eq. (4) with $A_0 = \mathrm{Tr}[E_\psi \Lambda(\rho_\psi - \mathbb{1}/d)]$ and $B_0 = \mathrm{Tr}[E_\psi \Lambda(\mathbb{1}/d)]$.

More realistically, the noise operator can be both gate and time dependent $\Lambda \to \Lambda_{i_j,j}$. We can predict the behavior of $F_{\mathrm{seq}}(m, \psi)$ by considering a perturbative expansion of each $\Lambda_{i_j,j}$ about the mean error Λ. Defining $\delta\Lambda_{i_j,j} = \Lambda_{i_j,j} - \Lambda$ for each i_j, our perturbative approach will be valid provided each $\delta\Lambda_{i_j,j}$ is small in a sense made precise later. Using the same change of variables described above, i.e., $\mathcal{D}_{i_m} = \bigcirc_{j=1}^{m}\mathcal{C}_{i_j}$, we find that $\mathcal{S}_{\mathbf{i_m}} = \mathcal{S}_{\mathbf{i_m}}^{(0)} + \mathcal{S}_{\mathbf{i_m}}^{(1)} + \mathcal{S}_{\mathbf{i_m}}^{(2)} \ldots$ where $\mathcal{S}_{\mathbf{i_m}}^{(0)}$ corresponds to the gate-independent case, $\mathcal{S}_{\mathbf{i_m}}^{(1)}$ is the first-order correction and so on. The first-order correction consists of three terms defined by whether the small gate-dependent perturbation acts (1a) on the first gate, (1b) somewhere in the middle (there are $m-1$ of these terms), or (1c) on the final gate. Explicitly,

$$\mathcal{S}^{(1a)}_{\mathbf{i_m}} = \Lambda \circ \tilde{\Lambda}_{i_m} \circ \ldots \circ \tilde{\Lambda}_{i_2} \circ (\mathcal{D}^\dagger_{i_1} \circ \delta\Lambda_{i_1,1} \circ \mathcal{D}_{i_1})$$
$$\mathcal{S}^{(1b)}_{\mathbf{i_m}} = \Lambda \circ \tilde{\Lambda}_{i_m} \circ \ldots \circ (\mathcal{D}^\dagger_{i_j} \circ \delta\Lambda_{i_j,j} \circ \mathcal{D}_{i_j}) \circ \ldots \circ \tilde{\Lambda}_{i_1}$$
$$\mathcal{S}^{(1c)}_{\mathbf{i_m}} = \delta\Lambda_{i_{m+1},m+1} \circ \tilde{\Lambda}_{i_m} \circ \ldots \circ \tilde{\Lambda}_{i_1}.$$

Averaging $\mathcal{S}^{(1a)}_{\mathbf{i_m}}$ over $\mathbf{i_m}$ gives

$$\mathcal{S}^{(1a)}_m = \Lambda \circ \Lambda^{\circ m-1}_{\rm dep} \circ (\mathcal{Q}_1 - \Lambda_{\rm dep}), \tag{6}$$

where we define for each j, $\mathcal{Q}_j = \sum_{i_j} \mathcal{D}^\dagger_{i_j} \circ \Lambda_{i_j,j} \circ \mathcal{D}_{i_j}/K$. Note the correlations between the noise and the gate operations prevent this from being a depolarizing channel.

For the $m-1$ terms with $j \in \{2, \ldots, m\}$ (case b) averaging over $\mathbf{i_m}$ gives

$$\mathcal{S}^{(1b)}_m = \sum_{j=2}^{m} \Lambda \circ ((\mathcal{Q}_j \circ \Lambda)_{\rm dep} - \Lambda^{\circ 2}_{\rm dep}) \circ \Lambda^{\circ m-2}_{\rm dep}, \tag{7}$$

where "dep" represents the depolarization of the operator within brackets. The main trick used here is to reexpand $\mathcal{D}_{i_j} = \mathcal{C}_{i_j} \circ \mathcal{D}_{i_{j-1}}$ in order to depolarize $\mathcal{C}^\dagger_{i_j} \circ \delta\Lambda_{i_j,j} \circ \mathcal{C}_{i_j} \circ \Lambda$ under the twirling operation $\sum_{i_{j-1}} \mathcal{D}^\dagger_{i_{j-1}} \circ \cdot \circ \mathcal{D}_{i_{j-1}}/K$.

To find the expression for $\mathcal{S}^{(1c)}_m$ we use the fact that the Cliffords are a group. If $i_1, \ldots, i_{m-1}$ are fixed, averaging over the i_m index runs through every Clifford element with equal frequency in the $\mathcal{D}_{i_m}$ random variable. Since $\Lambda_{i_{m+1},m+1}$ is the error associated with the gate $\mathcal{D}^\dagger_{i_m}$, $\sum_{i_m} \delta\Lambda_{i_{m+1},m+1} \circ (\mathcal{D}^\dagger_{i_m} \circ \Lambda \circ \mathcal{D}_{i_m})/K$ is independent of the $i_1, \ldots, i_{m-1}$ indices and

$$\mathcal{S}^{(1c)}_m = (\mathcal{R}_{m+1} - \Lambda \circ \Lambda_{\rm dep}) \circ \Lambda^{\circ m-1}_{\rm dep}, \tag{8}$$

where $\mathcal{R}_{m+1} = \sum_{i_m} \Lambda_{i'_m,m+1} \circ (\mathcal{C}^\dagger_{i_m} \circ \Lambda \circ \mathcal{C}_{i_m})/K$. Here $\Lambda_{i'_m,m+1}$ denotes the error that arises when the Clifford operation $\mathcal{C}^\dagger_{i_m}$ is applied at final time step $m+1$.

Combining these three terms it can be shown that the average sequence fidelity is given by Eq. (3) with

$$A_1 = \mathrm{Tr}\Big[E_\psi \Lambda\Big(\frac{\mathcal{Q}_1(\rho_\psi)}{p} - \rho_\psi + \frac{(p-1)\mathbb{1}}{pd}\Big)\Big] + \mathrm{Tr}\Big[E_\psi \mathcal{R}_{m+1}\Big(\frac{\rho_\psi}{p} - \frac{\mathbb{1}}{pd}\Big)\Big], \tag{9}$$

$$B_1 = \mathrm{Tr}\Big[E_\psi \mathcal{R}_{m+1}\Big(\frac{\mathbb{1}}{d}\Big)\Big], \qquad C_1 = \mathrm{Tr}\Big[E_\psi \Lambda\Big(\rho_\psi - \frac{\mathbb{1}}{d}\Big)\Big], \tag{10}$$

where $q = \sum_{j=2}^m q_j/(m-1)$ and q_j is the depolarizing parameter defined by $(\mathcal{Q}_j \circ \Lambda)_{\rm dep}(\rho) = q_j\rho + (1-q_j)\mathbb{1}/d$. Note that A_1, B_1, and q can have a dependence on m if the errors are time dependent.

An important issue is determining when the zeroth or first-order expressions Eqs. (3) and (4), are valid approximations, and, more generally, when our perturbative expansion is convergent. To bound the error in fidelity at a given order we use the "$1 \to 1$" norm on linear superoperators maximized over Hermitian inputs [21], denoted $\|\cdot\|^H_{1\to1}$. We find that for each order k,

$$\|S^{(k)}_m\|^H_{1\to1} \le \sum_{j_k > \ldots > j_1} \gamma_{j_k} \ldots \gamma_{j_1}, \tag{11}$$

where $\gamma_j := \sum_i \|\Lambda_{i_j} - \Lambda\|^H_{1\to1}/K$ is a measure of the variation in noise at the jth time step. In the case where the noise is time independent Eq. (11) becomes

$$\|S^{(k)}_m\|^H_{1\to1} \le \binom{m+1}{k}\gamma^k,$$

and, hence,

$$|F^{(k+1)} - F^{(k)}| \le \|S^{(k)}_m\|^H_{1\to1} \le \binom{m+1}{k}\gamma^k.$$

Note that our choice of norm is motivated by the fact that it gives a tighter bound on the error in fidelity. Hence the $k+1$ order correction to the fidelity formula can be neglected provided that

$$(m+1-k)\gamma/(1+k) \ll 1.$$

Thus, we can ignore second order terms when the *variation* in error strengths satisfies $\gamma \ll 2/m$. Note that in practice one also needs $m \gg 1$ to have enough data points for a reasonable estimate of p in the fitting model. This issue is discussed further in Ref. [22].

At this point we emphasize our results hold for any number of qubits; however, there are two subtle remarks in the multiqubit case. First, our protocol provides an estimate of the average error rate associated with a Clifford operation, which for n qubits, consists of $O(n^2)$ generating gates. Hence our condition for weak variation in the errors corresponds to the error set associated with blocks of $O(n^2)$ gates. Second, we assume that non-Markovian noise effects are limited to time scales less than that of a typical Clifford operation. For large n this corresponds to time scales less than $O(n^2)\tau$, where τ is the time scale of a generating gate.

We now numerically illustrate the protocol by considering some *physically motivated* single-qubit noise models and provide an example of when the first-order correction term may be observed experimentally. First we consider time-independent unitary errors corresponding to over-under rotations. The unitary error was constructed by first finding the Hamiltonian for each $\mathcal{C}_j$ via $\mathcal{C}_j(\rho) = \exp(-iH_j)\rho\exp(iH_j)$. Next, $\exp(-iH_j)$ was diagonalized and to simulate the error one of the eigenvalues was multiplied by $e^{i\delta}$ and the other by $e^{-i\delta}$. Two cases for δ were analyzed: $\delta = 0.1$ (case A) and δ chosen uniformly at random in [0.075, 1.125] (case B). Numerical values for $F_{\rm seq}(m, \psi)$ are shown in Fig. 1 as blue points. The main point is that for both cases the first-order formula (green

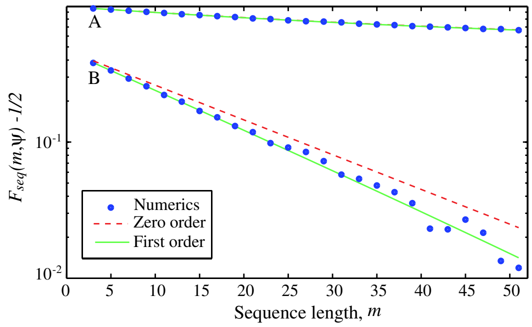

FIG. 1 (color online). Average sequence fidelity as a function of sequence length for unitary noise (see text for details). In these examples $B_1 = B_0 = 1/2$ since the noise model is unital and we assumed no state preparation or measurement errors. Hence, we have subtracted off these DC offsets so that pure exponentials will appear as straight lines (semilog plot). The nonexponential behavior for case B is evident.

line) models the data extremely well, whereas the zeroth-order formula (red dashed) only models the sequence fidelity in case A (when there is no variation in δ).

We also considered two other error models of practical relevance: unitary error with depolarizing noise and unitary error with amplitude damping. The depolarizing and damping parameters were chosen uniformly at random in 0.9875 ± 0.01 with the unitary error chosen as in case A. The results are summarized in Table I—in both cases the simulations are well approximated by the zeroth-order solution. These results further illustrate that the zeroth-order model gives a robust estimate of the error-rate for a variety of error models with small enough noise variation.

Lastly, we prove the efficiency of the protocol for arbitrary numbers of qubits. There are three main points to analyze: *Uniform sampling.*—Each Clifford element is uniquely determined by its action on the $2n$-generating elements of the Pauli group. Since randomly choosing $2n$ elements of the Pauli group that satisfy the required commutation relations is equivalent to inductively choosing random solutions to $2n$ sets of linear equations [which requires $O(n^3)$ operations], we can produce a random Clifford element in $O(n^4)$ (classical) operations. *Implementing Clifford operations.*—Any Clifford element can be decomposed into a sequence of $O(n^2)$ generators in $O(n^3)$ time [alternatively, there are slower methods for such a decomposition into $O(n^2/\log n)$ generators [23]]. *Averaging.*—The number of sequences of length m scales as $2^{mO(n^2)}$ and Hoeffding's inequality states that, with confidence δ and accuracy ϵ, the number of trials k needed for approximating the average sequence fidelity is no larger then $k = \ln(2/\delta)/2\epsilon^2$, which is independent of m and n.

In conclusion, we describe a scalable protocol for estimating average error-rates in noisy quantum information processors that consists of applying random sequences of Clifford gates and measuring the average sequence fidelity. The analysis admits both gate and time-dependent errors

TABLE I. Numerical results for the parameter p, error rate r, and our gate-dependence measure q-p^2 for the four cases of noise models considered. See text for details.

	Unitary A	Unitary B	Unitary and Dep.	Unitary and T_1
p	0.980	0.943	0.982	0.988
r	$1.05e$-2	$2.85e$-2	$8.75e$-3	$5.85e$-3
q-p^2	$-2.73e$-4	$-6.83e$-3	$-2.77e$-8	$-2.80e$-8

and is robust against state preparation and measurement errors. We derive zeroth and first-order fitting models for the experimental data and prove the validity of the models provided the variation in the errors is not too strong. We illustrate our results for some physically relevant noise models and provide an example of when a zeroth-order model fails to model the data.

We thank M. Silva, D. Gottesman, and D. Cory for valuable discussions. E. M. was supported by NSERC and CIFAR, J. M. G. was supported by CIFAR, Industry Canada, MITACS, MRI and NSERC and J. E. was supported by NSERC, CIFAR, and an Ontario ERA grant.

[1] I. Chuang and M. Nielsen, J. Mod. Opt. **44**, 2455 (1997).
[2] A. M. Childs, I. L. Chuang, and D. W. Leung, Phys. Rev. A **64**, 012314 (2001).
[3] Y. Weinstein *et al.*, J. Chem. Phys. **121**, 6117 (2004).
[4] J. L. O'Brien *et al.*, Phys. Rev. Lett. **93**, 080502 (2004).
[5] M. Riebe *et al.*, Phys. Rev. Lett. **97**, 220407 (2006).
[6] J. M. Chow *et al.*, Phys. Rev. Lett. **102**, 090502 (2009).
[7] R. C. Bialczak *et al.*, Nature Phys. **6**, 409 (2010).
[8] J. Emerson, R. Alicki, and K. Zyczkowski, J. Opt. B **7**, S347 (2005).
[9] B. Levi *et al.*, Phys. Rev. A **75**, 022314 (2007).
[10] J. Emerson *et al.*, Science **317**, 1893 (2007).
[11] M. Silva *et al.*, Phys. Rev. A **78**, 012347 (2008).
[12] A. Bendersky, F. Pastawski, and J. P. Paz, Phys. Rev. Lett. **100**, 190403 (2008).
[13] E. Knill *et al.*, Phys. Rev. A **77**, 012307 (2008).
[14] M. J. Biercuk *et al.*, Quantum Inf. Comput. **9**, 0920 (2009).
[15] C. Ryan, M. Laforest, and R. Laflamme, New J. Phys. **11**, 013034 (2009).
[16] J. M. Chow *et al.*, Phys. Rev. A **82**, 040305(R) (2010).
[17] S. Olmschenk *et al.*, New J. Phys. **12**, 113007 (2010).
[18] M. Nielsen and I. Chuang, *Quantum Computation and Information* (Cambridge University Press, Cambridge, UK, 2000).
[19] D. Gottesman, Phys. Canada **63**, 183 (2007).
[20] C. Dankert *et al.*, Phys. Rev. A **80**, 012304 (2009).
[21] J. Watrous, Quantum Inf. Comput. **5**, 058 (2005).
[22] E. Magesan, J. M. Gambetta, and J. Emerson (to be published)
[23] S. Aaronson and D. Gottesman, Phys. Rev. A **70**, 052328 (2004).