Approximation of realistic errors by Clifford channels and Pauli measurements

Mauricio Gutiérrez, Lukas Svec, Alexander Vargo, and Kenneth R. Brown

Phys. Rev. A **87**, 030302 — Published 22 March 2013

DOI: 10.1103/PhysRevA.87.030302

Approximation of realistic errors by Clifford channels and Pauli measurements

Mauricio Gutiérrez[1], Lukas Svec[2], Alexander Vargo[3], and Kenneth R. Brown[1*]
[1]*Schools of Chemistry and Biochemistry; Computational Science and Engineering; and Physics Georgia Institute of Technology, Atlanta, GA 30332-0400*
[2]*Department of Physics, University of Washington, Seattle, WA 98195-2350 and*
[3]*Department of Mathematics and Statistics, Haverford College, Haverford, PA 19041-1336*

(Dated: February 15, 2013)

The Gottesman-Knill theorem allows for the efficient simulation of stabilizer-based quantum error-correction circuits. Errors in these circuits are commonly modeled as depolarizing channels by using Monte Carlo methods to insert Pauli gates randomly throughout the circuit. Although convenient, these channels are poor approximations of common, realistic channels like amplitude damping. Here we analyze a larger set of efficiently simulable error channels by allowing the random insertion of any one-qubit gate or measurement that can be efficiently simulated within the stabilizer formalism. Our new error channels are shown to be a viable method for accurately approximating realistic error channels.

A system interacting with its environment will eventually reach thermal equilibrium. For finite temperatures, the equilibrium state will be a distribution of energy eigenstates weighted by Boltzmann factors. As a result the process of thermalization is generally non-unital. A unital process maps a completely mixed state to a completely mixed state and is compatible with thermalization only in the limit of infinite temperature or complete degeneracy of the energy eigenstates.

A quantum computer is a system that is often out of thermal equilibrium with its environment. Interactions with the environment can lead to errors in the computation. Fault-tolerant quantum error correction is one method for mitigating these errors with the advantage that provable arbitrary quantum computation is possible given constraints on the error rates and the error locality [1]. There are many possible error correcting codes ranging from concatenated [2–5] to subsystem [6] to topological codes [7–10]. It is typical to use simulation to determine the error correcting properties [11–13]. Although simulation of quantum systems is difficult, simulation of error correction can be done efficiently for stabilizer codes where the process of error correction only includes gates in the Clifford group [14, 15].

A standard error model is a depolarizing channel where a Pauli operator, chosen from a probability distribution, is applied at every possible error position [16]. The depolarizing channel efficiently simulates common laboratory processes such as dephasing. It also serves as a good approximation for unital channels and is appropriate for qubits with degenerate energy eigenstates. In nature, it is also common to encounter interactions with the environment where the process of thermalization leads to non-unital error channels. One example is spontaneous emission or amplitude damping where, given enough time, all density matrices map to a single pure state. If an error channel is far from unital then simulating it with Pauli errors gives large approximation errors, thus making it hard to extract useful results.

* ken.brown@chemistry.gatech.edu

In this paper, we go beyond simulating errors with the conventional Pauli depolarizing channel (PC). Rather than restricting to Pauli errors, we allow any subset of efficiently simulable gate errors to occur. In particular, we look at subsets generated by including all Clifford group operators and/or Pauli measurements to the PC. We show that adding Clifford errors and/or measurement errors always results in more accurate approximations and significant improvements for most error channels.

The stabilizer formalism allows for efficient classical simulation of quantum circuits when the initial state can be described by the measurement of a set of commuting Pauli operators and the gates in the circuit are Clifford operators, which map Pauli operators to Pauli operators by conjugation [15]. Single-qubit Clifford operators preserve the symmetry of the chiral octahedron [17]. This includes the identity, I; rotations about the vertices by π (Pauli operators), σ_j; rotations about the vertices by $\pi/2$ (S-like operators), S_v; rotations about the midpoint of each edge by π (Hadamard-like operators), H_e; and rotations about each face center by $2\pi/3$, R_f.

One can create an error process which is the weighted random application of these 24 unitary operators. We call this class of error models the Clifford Channel (CC) [18]. Most simulations of error correction circuits have used Pauli depolarizing channel (PC), which is a subset of CC consisting of only the random application of I or Pauli operators.

The stabilizer formalism also allows for efficient simulation of non-unital operations involving Pauli measurements and, optionally, conditional application of Clifford gates based on those measurements. In this paper, we limit ourselves to the set of operators that corresponds to measuring a Pauli operator and then conditionally applying a Pauli matrix such that all states map to the same state. We call these channels measurement induced translations. For each eigenstate, $|\lambda\rangle$, of a Pauli operator, we define the channel $\mathcal{E}_\lambda$ by two Kraus operators: $E_{\lambda 0} = |\lambda\rangle\langle\lambda|$ and $E_{\lambda 1} = |\lambda\rangle\langle\lambda^\perp|$. Notice that the effect of these two operators is to discard the state and replace it by $|\lambda\rangle$. We add these channels to our model with probability p_λ. The effect on a state, when represented on the Bloch sphere, is to translate it toward $|\lambda\rangle$. This allows us to generate non-unital

error channels that can be efficiently simulated. The extended models including measurement are labeled PMC and CMC. Table I describes the content of each channel class in terms of the underlying channel set and the number of free parameters.

TABLE I. Four error models compatible with the stabilizer formalism.

Channel Class	Channel Set	Parameters
PC	$\{I, \sigma_i\}$	3
PMC	$\{I, \sigma_i, \mathcal{E}_\lambda\}$	9
CC	$\{I, \sigma_i, S_v, H_e, R_f\}$	23
CMC	$\{I, \sigma_i, S_v, H_e, R_f, \mathcal{E}_\lambda\}$	29

In nature, there are many error channels that cannot be exactly represented in the stabilizer formalism. We examine two channels: the amplitude damping channel (ADC), and dephasing about an arbitrary axis in the x-y plane (Pol$_\phi$C). The ADC, represented in Equation 1, is the prototypical non-unital error channel [19]. It describes the energy dissipation of a two-level quantum system. However small, it is present in any non-degenerate physical system.

$$\text{ADC} = \mathcal{E}_A = \begin{cases} E_{A0} = |0\rangle\langle 0| \; + \; \sqrt{1-\gamma}\,|1\rangle\langle 1| \\ E_{A1} = \sqrt{\gamma}\,|0\rangle\langle 1| \end{cases} \tag{1}$$

The dimensionless parameter γ, which can take any real value between 0 and 1 can be related to the total energy lost to the environment. Numerous codes have have been developed specifically to combat ADC, but studying the effects of this error channel on a circuit has yielded only a handful of results [20–23]. All of the results assume γ to be small in order to expand the Kraus operators in a Taylor series expansion using the Pauli operator basis.

Unlike the ADC, the Pol$_\phi$C, which is represented in Equation 2, is unital. Yet unless the angle ϕ is a half-integer multiple of π, the depolarization occurs along a non-Pauli axis, and the quality of the PC approximation will vary with ϕ.

$$\text{Pol}_\phi\text{C} = \mathcal{E}_{xy} \begin{cases} E_{xy0} = \sqrt{1-p_\phi}\, I \\ E_{xy1} = \sqrt{p_\phi}\,[\cos(\phi)\, X \; + \; \sin(\phi)\, Y] \end{cases} \tag{2}$$

Here the parameter ϕ represents the angle of the polarization axis with respect to the X axis and p_ϕ the probability of error.

To study how closely our error models approximate target error channels, we compute the distance between the process matrix of our error model and the process matrix of the target error. As a distance measure on a single qubit, we employ the normalized Hilbert-Schmidt distance between the process matrices associated with each channel [24], $D^{\text{HS}}(\chi_1, \chi_2) = \frac{1}{8}\|\chi_1 - \chi_2\|^2_{\text{HS}}$ and $\|A\|_{\text{HS}} = \sqrt{tr(A^\dagger A)}$. For a multi-qubit study of non-unital errors, a more natural distance measure is the diamond norm [25]. Here we minimize the Hilbert-Schmidt distance over the parameter space of the model. We want our model to be an upper bound to the error induced on the system. Therefore, we perform the distance minimization with the constraint that for every initial pure state its trace distance to the resulting state after the target transformation is not greater than its trace distance to the resulting state after the model approximation, $D^{tr}(\rho, \text{Target}(\rho)) \leq D^{tr}(\rho, \text{Model}(\rho))$. The trace distance is calculated using the expression $D^{tr}(\rho, \sigma) = \frac{1}{2} tr|\rho - \sigma|$. We use the Hilbert-Schmidt distance for the analysis here due to ease of computation, but the method works for any distance measure or constraint [26].

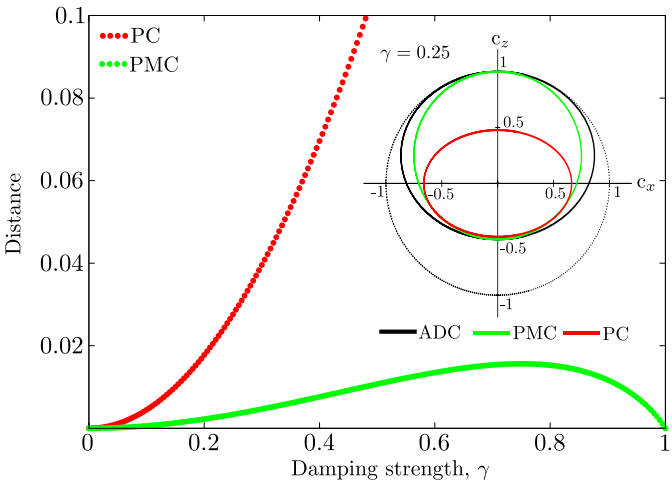

FIG. 1. (color online) Minimum Hilbert-Schmidt distance between two approximate error models and the ADC as a function of γ. For $\gamma > 0.5$, in order to satisfy the trace distance constraint, the PC results in highly inaccurate approximations. The inset figure shows the contraction of the Bloch sphere for the ADC, PC, and PMC for $\gamma = 0.25$ (see text).

Figure 1 shows the results of the approximation of the ADC with the PC and PMC error models. Numerical minimization revealed that the channels with additional Clifford operators, the CC and the CMC, did not improve the approximations achieved by the PC and the PMC, respectively. As the non-unital character of the ADC becomes more pronounced, the unital error models result in a less accurate approximation. When $\gamma > 0.5$ the trace distance constraint forces the unital error models to generate an approximation that results in a re-orientation of the mapped density matrices with respect to the initial Bloch sphere. At this point, the unital approximations become very inaccurate.

The addition of the measurement-induced translations considerably improves the approximation. In this case, the PMC and CMC yield valid approximations for the whole range of γ. The PMC and CMC can match the ADC perfectly only for $\gamma = 0$, which corresponds to I, and $\gamma = 1$, which corresponds to the measurement-induced translation $\mathcal{E}_0$. Interestingly, despite the large amount of operators in the CMC error model, the best approximation only employs I and $\mathcal{E}_0$. This allowed us to perform the minimization symbolically to obtain a simple analytical expression for the distance, $D^{\text{HS}}_{\text{PMC}} = \frac{1}{2}(\gamma - 1)(\gamma + 2\sqrt{1-\gamma} - 2)$, and for the Kraus operators in

the approximation, $\{\sqrt{1-\gamma}\,I\,,\ \sqrt{\gamma}\,|0\rangle\langle 0|\,,\ \sqrt{\gamma}\,|0\rangle\langle 1|\}$. In the limit of small γ the distance of the PC (red points) and the PMC (green points) both scale quadratically with γ. After Taylor-expanding the distance expression for the PMC and fitting the PC distance numerically in terms of γ, it was found that the ratio of the quadratic coefficients was 7.3. This number, which can be interpreted as the ratio of the two distances in the limit of small damping, $\lim_{\gamma\to 0}(D^{\rm HS}_{\rm PC}/D^{\rm HS}_{\rm PMC})$, shows the improved performance of the PMC over the PC. To further appreciate this improvement, notice that the distance between the ADC and the Identity channel also scales quadratically in the limit of small γ. In this case, $\lim_{\gamma\to 0}(D^{\rm HS}_{\rm I}/D^{\rm HS}_{\rm PMC}) = 5$. At small γ the distance between the ADC and the Identity is of comparable magnitude to the distance between the ADC and PC approximation. The fact that the quadratic coefficient is smaller for the Identity might lead one to believe that, for small γ, the Identity is a better approximation than the PC, but, of course, the Identity is not a valid option as it does not satisfy the trace distance constraint.

The results obtained by including or not including measurement operators are best illustrated in the inset of Figure 1. Here we examine, for $\gamma = 0.25$, the closest PC and PMC approximation under the constraint. The figure shows a cross section of the Bloch sphere (black dotted) and its transformation by the ADC (black solid) and the closest approximate PC (red) and PMC (green). The non-unital PMC preserves the asymmetry of the transformation along the z-axis. Notice that for these error channels the deformed Bloch sphere is still symmetric with respect to rotations around z.

As mentioned before, the distance constraint guarantees that for every initial pure state its distance to the resulting state after the target transformation is not greater than its distance to the resulting state after the model approximation. For both the ADC and its approximations the largest discrepancy between input and output occurs when the initial state is $|1\rangle$. The distance constraint has a nice geometric interpretation, provided that the error is sufficiently low. As shown in the inset of Figure 1, both the green and the red curves lie *inside* the black solid curve and further away from the initial states (black dotted curve). For the PMC approximation, this interpretation is satisfied for any value of γ. However, for the PC approximation this interpretation fails for $\gamma > 0.5$ when, as mentioned before, the best approximation results in a reorientation of the mapped density matrices with respect to the inital Bloch sphere. Furthermore, the PMC approximation satisfies the distance constraint for any input state, whether pure or mixed. It is impossible to satisfy this condition for the PC or any unital approximation: simply consider the maximally mixed state, which is mapped to itself by a unital channel, but mapped to a different state by a non-unital one. In fact, by this same argument, it is clear that the distance constraint is impossible to satisfy for every initial state, pure or mixed, when the approximation channel has a different fixed point from the target transformation.

Figure 2 shows the results of the approximation of the Pol$_\phi$C by the error models introduced earlier. Once again, each point corresponds to a numerical minimization. Because of the unital nature of this channel, it is the addition of the Clifford operators rather than the measurement operators that improve the approximation. For both the PC and the PMC, the distance between Pol$_\phi$C and the best approximation was found to be $D^{\rm HS}_{\rm PC} = \frac{1}{2}p^2\sin^2(2\phi)$ for $p \le 2/3$. When the Clifford operators are included in the approximate channel, the distance is reduced to $D^{\rm HS}_{\rm CC} = \frac{1}{6}p^2(\sin(2\phi)+\cos(2\phi)-1)^2$ for $p \le 6/7$ and $0 < \phi < \pi/4$. Not only does the distance decrease with the addition of the Clifford operators; the period of the distance function is reduced from $\frac{\pi}{2}$ to $\frac{\pi}{4}$, because between every two Pauli axes there is a Clifford axis. Notice the great improvement in the CC approximation with respect to the PC one. At the worst point of the CC (which in this interval occurs at $\phi = \pi/8, 3\pi/8$), the PC is 8.7 times worse. For the PC(CC) approximation when $p > 2/3(6/7)$, the trace distance constraint cannot be satisfied.

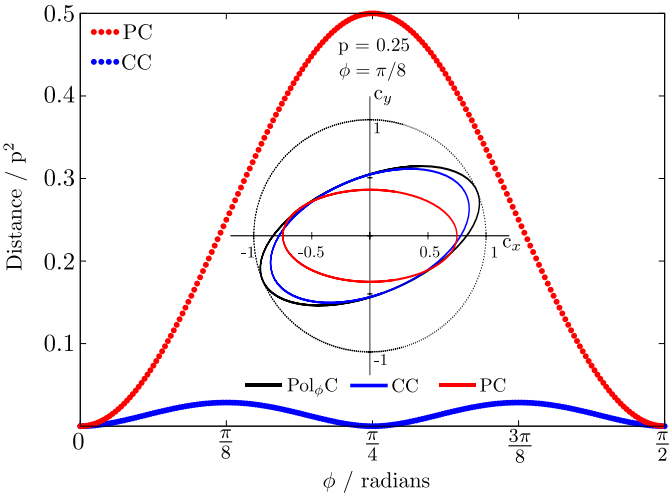

FIG. 2. (color online) Minimum Hilbert-Schmidt distance between several approximate error models and the polarization along an axis in the X-Y plane of the Bloch sphere as a function of the polarization angle. Although not shown, the results for PMC and CMC are the same as the results for PC and CC, respectively. The distances scale quadratically with p, so the results are normalized by p^2.

Despite the large number of operators in the CMC, the best approximation uses a small number of them: I, Z, and the two axes closest to the polarization axis. If we only employ Pauli axes, the best approximation in Kraus representation is $\{\sqrt{1-p_x-p_y-p_z}\,I\,,\ \sqrt{p_x}\,X\,,\ \sqrt{p_y}\,Y,\ \sqrt{p_z}\,Z\}$, where $p_x = p\cos^2(\phi)$, $p_y = p\sin^2(\phi)$, and $p_z = p\cos(\phi)\sin(\phi)$. If we employ the whole Clifford group, the best approximation in the interval $0 \le \phi \le \pi/4$ is given by $\{\sqrt{1-p_1-p_2-p_3}\,I\,,\ \sqrt{p_1}\,X\,,\ \sqrt{p_2}\,H_{XY}\,,\ \sqrt{p_3}Z\}$, where $H_{XY} = \frac{1}{\sqrt{2}}(X+Y)$, $p_1 = \frac{1}{3}p[2\cos(2\phi)-\sin(2\phi)+1]$, $p_2 = \frac{1}{3}p[2\sin(2\phi)-\cos(2\phi)+1]$, and $p_3 = \frac{\sqrt{2}-1}{6}p[\cos(2\phi)+\sin(2\phi)-1]$.

The inset in Figure 2 illustrates the closest PC (red) and CC (blue) approximations to the Pol$_\phi$C (black solid) with $p = 0.25$ and $\phi = \pi/8$. For the PC approximation, the great-

est discrepancy between the input and output states occurs when $\rho = \frac{1}{2}[I + \cos(\frac{3\pi}{4})\,X + \sin(\frac{3\pi}{4})\,Y]$. For the CC approximation, this occurs when $\rho = \frac{1}{2}[I + \cos(\frac{5\pi}{8})\,X + \sin(\frac{5\pi}{8})\,Y]$.

We have seen that the addition of the measurement-induced translations and the Clifford operators improves the approximation of two specific error channels. To determine how the method works for general errors, we generated 1000 random process matrices and computed the distance of the best approximation that each one of the 4 approximate channels could make. For the 1-qubit case, a process matrix is a 4×4 Hermitian positive matrix M with 4 constraints in the normalized Pauli basis: $tr(M)$ = 2, Re(M_{01}) = -Im(M_{23}) , Re(M_{02}) = Im(M_{13}) , and Re(M_{03}) = -Im(M_{12}). To generate this matrix we first create a 4×4 diagonal matrix D with real, positive diagonal entries that add to 2. We then create a 4×4 random unitary matrix U and apply this unitary transformation to D to obtain $M = UDU^\dagger$, which is positive with trace 2. We then enforce the last 3 constraints mentioned earlier and keep the random process if the matrix is still positive.

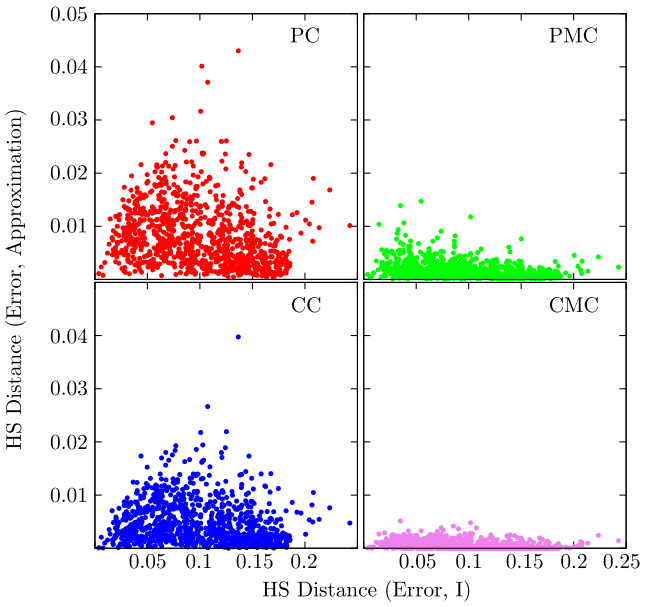

FIG. 3. (color online) Hilbert-Schmidt distance between the random error channels and the best approximations attained with each model as a function of the distance between the error channel and the error-free channel (I). The slope of a line joining the origin and a point represents the distance of the best approximation to that error relative to the magnitude of the error.

Figure 3 illustrates the distance between each random error channel and the best approximation as a function of the distance between the error channel and the identity channel. Notice that as the number of operators in the error models increases, both the mean and the median distance between each model and the random error decreases and the distributions become more compact, as summarized in Table II. For this data, the improvement of adding either Clifford gates, CC, or measurement-induced translation operators, PMC, over PC is comparable. The total set of operators leads to an order of magnitude improvement over the PC. In the case of the CMC,

TABLE II. Summary of the approximations obtained with each of the 4 error models.

Channel	Distance mean	Distance median	Distance variance
PC	1.7×10^{-2}	1.4×10^{-2}	1.4×10^{-4}
PMC	3.4×10^{-3}	2.4×10^{-3}	1.1×10^{-5}
CC	9.8×10^{-3}	7.5×10^{-3}	7.0×10^{-5}
CMC	1.1×10^{-3}	4.2×10^{-4}	2.2×10^{-6}

for the 1000 random channels tested, the number of non-zero parameters used in the approximations varied from 4 to 29 with a median of 12. This is in contrast to the ADC and the Pol$_\phi$C where only 1 and 3 parameters, respectively, are required due to the symmetry of the error channels.

We have presented an extension to the random Pauli error model which is still compatible with efficient simulation using the Gottesman-Knill theorem and leads to a computationally tractable description of realistic error models like amplitude damping. Our method can be extended to multi-qubit channels but the optimization becomes more difficult as the number of Clifford operators grows quickly with n. In many cases, symmetries of the underlying error channels will minimize the number of Clifford operators that must be considered. We also note that conditional measurements followed by Clifford gates can be used to generate a classical Toffoli (Z measurement followed by conditional CNOT) and to mimic thermalization processes for spin interactions (measurement of ZZ followed by a conditional spin flip). We plan to examine the performance of quantum error correcting codes under a wider range of non-unital error channels. Based on the relative distance of the PMC and PC approximations to the ADC for small errors, we expect that the CMC approximation of a non-unital channel will yield a nontrivial change in the code threshold of order unity relative to the PC approximation.

ACKNOWLEDGMENTS

The authors thank Aram Harrow for valuable comments. AV and KRB acknowledge support from the NSF (CHE-1037992). MG, LS, and KRB were also supported by ODNI-IARPA (ARO: W911NF-10-1-0231 and DOI:D11PC20167).

[1] D. Gottesman, in *Quantum Information Science and its Contributions To Mathematics* (2010), vol. 68 of *Proceedings of Symposia in Applied Mathematics*, pp. 13–58.

[2] P. W. Shor, Phys. Rev. A **52**, R2493 (1995).
[3] A. R. Calderbank and P. W. Shor, Phys. Rev. A **54**, 1098 (1996).
[4] D. P. DiVincenzo and P. W. Shor, Phys. Rev. Lett. **77**, 3260 (1996).
[5] E. Knill, Nature **434**, 39 (2005).
[6] D. Bacon, Phys. Rev. A **73**, 012340 (2006).
[7] A. Kitaev, Annals of Physics **303**, 2 (2003).
[8] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, J. Math. Phys. **43**, 4452 (2002).
[9] H. Bombin and M. A. Martin-Delgado, Phys. Rev. Lett. **97**, 180501 (2006).
[10] H. Bombin, Phys. Rev. A **81**, 032301 (2010).
[11] A. Cross, Ph.D. thesis, Massachusetts Institute of Technology (2009).
[12] P. Aliferis, D. Gottesman, and J. Preskill, Quantum Info. Comput. **6**, 97 (2006), ISSN 1533-7146.
[13] A. W. Cross, D. P. Divincenzo, and B. M. Terhal, Quantum Info. Comput. **9**, 541 (2009), ISSN 1533-7146.
[14] D. Gottesman, Ph.D. thesis, California Institute of Technology (1997). S. Aaronson and D. Gottesman, Phys. Rev. A **70**, 052328 (2004).
[15] S. Aaronson and D. Gottesman, Phys. Rev. A **70**, 052328 (2004).
[16] A. M. Steane, Phys. Rev. A **68**, 042322 (2003).
[17] W. van Dam and M. Howard, Phys. Rev. Lett. **103**, 170504 (2009).
[18] During the preparation of this manuscript, we learned about independent and related work on the Clifford channel, E. Magesan, D. Puzzuoli, C.E. Granade, and D.G. Cory, arXiv:1206.5407v1.
[19] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2001).
[20] D. W. Leung, M. A. Nielsen, I. L. Chuang, and Y. Yamamoto, Phys. Rev. A **56**, 2567 (1997).
[21] I. L. Chuang, D. W. Leung, and Y. Yamamoto, Phys. Rev. A **56**, 1114 (1997).
[22] R. Duan, M. Grassl, Z. Ji, and B. Zeng, in *2010 IEEE International Symposium On Information Theory* (2010), IEEE International Symposium on Information Theory, pp. 2672–2676.
[23] P. W. Shor, G. Smith, J. A. Smolin, and B. Zeng, IEEE Transactions On Information Theory **57**, 7180 (2011).
[24] M. D. Grace, J. Dominy, R. L. Kosut, C. Brif, and H. Rabitz, New Journal of Physics **12**, 015001 (2010).
[25] A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and Quantum Computation* (American Mathematical Society, Boston, MA, USA, 2002).
[26] A. Gilchrist, N. K. Langford, and M. A. Nielsen, Phys. Rev. A **71**, 062310 (2005).