

Correctable noise of quantum-error-correcting codes under adaptive concatenation

Jesse Fern

Berkeley Quantum Information Center, Department of Mathematics, University of California, Berkeley, California 94720, USA

(Received 28 March 2007; published 18 January 2008)

We examine the transformation of noise under a quantum-error-correcting code (QECC) concatenated repeatedly with itself, by analyzing the effects of a quantum channel after each level of concatenation using recovery operators that are optimally adapted to use error syndrome information from the previous levels of the code. We use the Shannon entropy of these channels to estimate the thresholds of correctable noise for QECCs and find considerable improvements under this adaptive concatenation. Similar methods could be used to increase quantum-fault-tolerant thresholds.

DOI: 10.1103/PhysRevA.77.010301

PACS number(s): 03.67.Lx

Decoherence is a significant problem for implementation of quantum computing. Quantum error correction [1,2] mitigates decoherence and gate errors by introducing redundancy. In particular, for many codes, one qubit is encoded into n qubits and each of those encoded into n qubits, so that after j levels of concatenation, n^j qubits are used. In previous work, concatenated quantum codes were analyzed using a channel map formalism to evaluate their performance and to determine the threshold error probability for perfect fidelity in the infinite-concatenation limit [3,4]. In order to optimize the recovery operators at each level of the code, we develop here an adaptive concatenation approach, which yields significantly larger values for the thresholds of correctable noise. This work extends previous work [5], by formulating it in the channel superoperator representation, which allows for arbitrary quantum channels. We introduce the idea of using the Shannon entropy of ensembles of channels to calculate the thresholds of correctable noise; we use this to find thresholds for two quantum-error-correcting codes under two different types of noise.

The n qubit Pauli operators are $\mathcal{P}_n = \{I, X, Y, Z\}^{\otimes n}$. We use the notation that $XY = X \otimes Y$. Each pair of Pauli operators σ, σ' either commutes or anticommutes, where $\sigma \in \mathcal{P}_n$. We let the quantity $\eta(\sigma, \sigma')$ be 1 if they commute, -1 if they anticommute.

We write a density matrix on n qubits as $\rho = (1/2^n) \sum_{\sigma \in \mathcal{P}_n} c_\sigma \sigma$. A channel is a map from density matrices to density matrices, and can be written as $\mathcal{N}(\rho) = \sum_i A_i \rho A_i^\dagger$, where the A_i are Kraus operators that act on the state. This channel has the matrix representation $\mathcal{N} = \sum_i \mathcal{O}(A_i)$, where $\mathcal{O}(M) = M \otimes M^{\dagger T}$, and \mathcal{N} is a matrix that acts on the vector form of ρ in the standard basis. By looking at how the 4^n Pauli operators $\sigma \in \mathcal{P}_n$ perform under the map $\mathcal{N}(\sigma)$, the channel can be written as a $4^n \times 4^n$ superoperator in the Pauli basis. Because channels are trace preserving and Hermitian preserving, this implies that one-qubit channels written in the Pauli basis have real entries and the first row is $\mathcal{N}_{I\sigma} = \delta_{I,\sigma}$. On one qubit, this is

$$\mathcal{N}^{(1)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ N_{XI} & N_{XX} & N_{XY} & N_{XZ} \\ N_{YI} & N_{YX} & N_{YY} & N_{YZ} \\ N_{ZI} & N_{ZX} & N_{ZY} & N_{ZZ} \end{bmatrix}.$$

A quasidensity matrix ρ is an unnormalized density matrix. It can be normalized to a density matrix $(1/p_s)\rho$, where $p_s = \text{tr} \rho = c_I$. Just as a density matrix can be written as a sum of quasidensity matrices, a channel can be written as a sum of quasichannels [4]. Quasichannels map density matrices to quasidensity matrices, and therefore do not have to preserve trace. If \mathcal{N} is a quasichannel and ρ is a density matrix, then the probability of the quasichannel occurring is $p_c = \text{tr}(\mathcal{N}\rho) = \sum_{\sigma} N_{I\sigma} c_\sigma$, which depends on the state ρ unless $\mathcal{N}_{I\sigma} = p_c \delta_{I,\sigma}$, in which case it gives the channel $(1/p_c)\mathcal{N} = (1/\mathcal{N}_{II})\mathcal{N}$ with probability p_c .

An $[[n, k, d]]$ quantum stabilizer code encodes k logical qubits into n physical qubits, and has a distance d . Typically, a quantum stabilizer code is given in terms of $n-k$ generators g_i and 4^k encoded Pauli operators $\bar{\sigma}$.

The g_i generate the stabilizer group S , which is isomorphic to \mathbb{Z}_2^{n-k} . The k logical qubits encode into the 2^k -dimensional code space C_S , which has the property that, if $|\psi\rangle \in C_S$ and $s \in S$, then $s|\psi\rangle = |\psi\rangle$. The elements of the set $\mathcal{C}(S) \subset \mathcal{P}_n$ are the 2^{n+k} Pauli operators which commute with S . They send states in C_S to states in C_S . Each of the 4^k different equivalence classes of $\mathcal{C}(S)/S$ corresponds to one of the 4^k encoded logical Pauli operators $\bar{\sigma}$ that act on the logical qubits in the code space.

To perform error detection, we measure each of the $n-k$ generators g_i , projecting into either the $+1$ or -1 eigenspace for each generator g_i . This gives us a syndrome bit β_i , which is 0 if the state is in the $+1$ eigenspace, and 1 if it is in the -1 eigenspace. These β_i form the 2^{n-k} possible error syndromes β . $\beta_0 = 0 = (0, 0, \dots, 0)$ is the syndrome corresponding to no error.

To recover from an error characterized by syndrome β , a recovery operator $R(\beta)$ must be chosen that returns the qubit to the code space. Since $g_i R(\beta) |\psi\rangle = R(\beta) |\psi\rangle$, then $\eta(R(\beta), g_i) = \beta_i$ for all i . The 2^{n+k} possible choices for $R(\beta)$ are all in the same equivalence class of $\mathcal{P}_n/\mathcal{C}(S)$.

We define a $4^n \times 4^k$ encoding operator \mathcal{E} that maps from a density matrix on k qubits to one on n qubits as

$$\rho \rightarrow \frac{1}{2^{n-k}} \mathcal{E} \rho.$$

This projects us into the code space of the code. In the Pauli basis, the 4^k columns of $(1/2^{n-k})\mathcal{E}$ act as the logical Pauli

operators in the code space, and 0 outside the code space. They are defined as

$$\mathcal{E}_I = \prod_i (I + g_i) \text{ and } \mathcal{E}_\sigma = \bar{\sigma} \mathcal{E}_I. \quad (1)$$

Now that the logical qubits are encoded into the code space, assume that some noise \mathcal{N} acts on the full 2^n -dimensional space of the n physical qubits of the code. The g_i operators are measured, giving an error syndrome β . Some unitary recovery operator $R(\beta)$ is then chosen. We can write the error-correction process as

$$\mathcal{T}_{\text{ideal}} = \sum_{\beta} \mathcal{O}(R(\beta)) \circ \mathcal{P}_{\beta} = \mathcal{P}_0 \circ \sum_{\beta} \mathcal{O}(R(\beta))$$

where $\mathcal{O}(R(\beta))$ is the error recovery superoperator, \mathcal{P}_{β} is a projection superoperator into the syndrome β space, and $\mathcal{P}_0 = (1/2^{n-k}) \mathcal{E} \circ \mathcal{E}^T$ (transpose of \mathcal{E} followed by \mathcal{E}) is a projection into the code space. \mathcal{E}^T decodes back into the original k qubit space. The whole process can then be written as

$$\mathcal{G} = \frac{1}{2^{n-k}} \mathcal{E}^t \circ \sum_{\beta} \mathcal{O}(R(\beta)) \circ \mathcal{N} \circ \mathcal{E}. \quad (2)$$

This represents the noise on the logical qubits \mathcal{G} in terms of the noise on the physical qubits \mathcal{N} . We can decompose \mathcal{G} as a sum of contributions from each error syndrome:

$$\mathcal{G} = \sum_{\beta} \mathcal{G}^{R(\beta)}, \quad \mathcal{G}^U = \frac{1}{2^{n-k}} \mathcal{E}^t \circ \mathcal{O}(U) \circ \mathcal{N} \circ \mathcal{E}. \quad (3)$$

Often it is assumed that each qubit experiences the same noise, so $\mathcal{N} = (\mathcal{N}^{(1)})^{\otimes n}$. Previous work [3,4] showed that the case of a code C where $k=1$ produces a map $\Omega^C: \mathcal{N}^{(1)} \rightarrow \sum_{\beta} (1/2^{n-k}) \mathcal{E}^t \circ \mathcal{O}(R(\beta)) \circ (\mathcal{N}^{(1)})^{\otimes n} \circ \mathcal{E}$ from the physical one-qubit noise on each qubit to the logical one-qubit noise on the encoded code space. This is useful for analyzing a code concatenated with itself an arbitrary number of times. A logical qubit on one level of the code is treated as a physical qubit at the next level of the code, allowing us to concatenate a code with itself many times. In particular, the code will correct noise if $\lim_{k \rightarrow \infty} \Omega^{C^k}(\mathcal{N}^{(1)}) = I$ [4]. However, this assumes that the recovery operators do not depend on syndrome information from the previous levels of error correction. Furthermore, it does not give as high thresholds as the optimal adaptive concatenation described below.

A Pauli matrix σ has the diagonal superoperator $\mathcal{O}(\sigma)$, where $\mathcal{O}(\sigma)_{\sigma'\sigma'} = \eta(\sigma, \sigma')$. Suppose the noise is diagonal in the Pauli basis. Then the \mathcal{G}^U from Eq. (3) become $\mathcal{G}_{\sigma'\sigma'}^{\sigma} = (1/2^{n-k}) \mathcal{E}_{\sigma'}^t \circ \mathcal{O}(\sigma) \circ \mathcal{N} \circ \mathcal{E}_{\sigma'}$. Since the nonzero entries of the column $\mathcal{E}_{\sigma'}$ are $\bar{\sigma}'_s$ where $s \in S$, it follows that

$$\mathcal{G}_{\sigma'\sigma'}^{\sigma} = \frac{1}{2^{n-k}} \sum_{s \in S} \eta(\sigma, \bar{\sigma}'_s) \mathcal{N}_{\bar{\sigma}'_s, \bar{\sigma}'_s}. \quad (4)$$

If a quasichannel is diagonal, it can be written as a sum of Pauli superoperators $\mathcal{N} = \sum_{\sigma} p_{\sigma} \mathcal{O}(\sigma)$, which corresponds to

having each Pauli error σ with probability p_{σ} . For a one-qubit quasichannel with probability $p = \sum_{\sigma} p_{\sigma}$, the diagonal parts are

$$\begin{aligned} [p, x, y, z] = & p_I [1, 1, 1, 1] + p_X [1, 1, -1, -1] + p_Y [1, -1, 1, -1], \\ & + p_Z [1, -1, -1, 1]. \end{aligned} \quad (5)$$

This yields the Pauli probabilities:

$$\begin{aligned} p_I = \frac{p+x+y+z}{4}, \quad p_X = \frac{p+x-y-z}{4}, \\ p_Y = \frac{p-x+y-z}{4}, \quad p_Z = \frac{p-x-y+z}{4}. \end{aligned} \quad (6)$$

Since the information entropy (Shannon) in terms of Pauli errors with probabilities q_i is $\sum_i h(q_i)$ where $h(x) = -x \log_2 x$, a quasichannel gives a contribution of $p \sum_{\sigma} h(p_{\sigma}/p) = -h(p) + \sum_{\sigma} h(p_{\sigma})$ to the entropy of a channel.

We now show how using syndrome information from the lowest levels of a code can be used to optimize the threshold of the code.

When we use the syndrome information, instead of having the logical encoded channel map written as a simple sum of channel contributions $\mathcal{G} = \sum_{\beta} \mathcal{G}^{R(\beta)}$ as in Eq. (3), we represent this additional information as $\sum_{\beta} \mathcal{G}^{R(\beta)} \otimes \beta$. If we perform error correction and ignore the syndrome information, we have simply \mathcal{G} . The probability of measuring a syndrome β depends on the density matrix state ρ , and is

$$p_{\beta} = \text{tr}(\mathcal{G}^{R(\beta)} \rho) = 2^n \sum_{\sigma} \mathcal{G}_{I\sigma}^{R(\beta)} c_{\sigma}.$$

If the first row of this quasichannel in the Pauli basis is all zero except for the first term, i.e., $\mathcal{G}_{I\sigma}^{R(\beta)} = p \delta_{I\sigma}$, then $p_{\beta} = \mathcal{G}_{II}^{R(\beta)} = p$. If the syndrome β is measured, the resulting noise is $(1/p_{\beta}) \mathcal{G}^{R(\beta)}$.

Given an $[[n, 1, d]]$ quantum code, assume that the syndrome β is measured. It has a corresponding equivalence class $\mathcal{P}_n/C(S)$. From this, we choose some representative element r_{β} to return to the code space. This element differs from the optimal recovery operator $R(\beta)$ by a logical Pauli operator $\bar{\sigma}$. The recovery operators for β are therefore equivalent to the four recovery operators $\bar{\sigma} r_{\beta}$, where $\bar{\sigma}$ is any one of the four logical encoded Pauli operators. We then have the quasichannel $\mathcal{G}^{r_{\beta}}$.

Now, for each of the n blocks of the code, we pass this quasichannel \mathcal{N}_i onto the next level of the code, and have the resulting noise $\mathcal{N} = \otimes_i \mathcal{N}_i$. To optimize the recovery operator based upon this noise, either we can find the optimal $\bar{\sigma} r_{\beta}$ at each level of the code, or we can apply a random recovery operator r_{β} at each level of the code, and optimize at the end. In the second case, since each block differed by a Pauli operator σ_i from optimal, the resulting noise differs by $\otimes_i \sigma_i$ from optimal. This is equivalent to having the recovery operator $\otimes_i \sigma_i \circ r_{\beta}$ instead of r_{β} . This is $\bar{\sigma} \circ r_{\alpha}$ for some logical Pauli operator $\bar{\sigma}$ and therefore some error syndrome α , and therefore differs by a logical Pauli operator from the optimal recovery operator at the top level of the code.

While the first method may seem more intuitive, the sec-

TABLE I. Encoded Pauli operators $\bar{\sigma}$ and encoding operator \mathcal{E}_σ for the two-qubit code.

σ	$\bar{\sigma}$	\mathcal{E}_σ
I	II	$II+ZZ$
X	XX	$XX-YY$
Y	XY	$XY+YX$
Z	IZ	$IZ+ZI$

ond may be more computationally efficient, since determining the optimal recovery operator at any given level of the code is nontrivial. At the last level of the code, we will choose the encoded Pauli operator σ that yields the channel closest to the identity channel, which corresponds to the maximal p_σ from Eq. (6). If the resulting noise is very close to the identity channel, then we have established that very little logical noise is introduced at each step of a quantum computation, allowing long quantum computations to be performed. In practice, the noise at the end of the concatenation will often be close to diagonal even though the initial physical qubit noise was not.

Suppose we wish to calculate the exact optimized channel map of an $[[n, 1, d]]$ quantum code after j levels of concatenation with itself. There are 2^{n-1} syndromes. Many of these syndromes may have the same channel components $\mathcal{G}^{R(\beta)}$, rendering it unnecessary to calculate all of them, reducing the complexity. However, for many codes it is still computationally too expensive to compute the optimized channel past the second level. Nevertheless, the overhead involved with implementing the optimization is low because the encoded channel from one level is simply passed on to the qubits of the next level of the code. This leads to efficient Monte Carlo simulation.

It is useful to look first at the two-qubit bit-flip code because of its simplicity; it is the classical two-bit repetition code. Its code space is spanned by the logical encoded states $|0\rangle=|00\rangle$, and $|1\rangle=|11\rangle$. The stabilizer group $S=\{II, ZZ\}$ is generated by the one generator $g_1=ZZ$. Table I lists the encoded Pauli operators and corresponding \mathcal{E}_σ [Eq. (1)]. Using Eq. (3), we can calculate various channel map components $\mathcal{G}^{R(\beta)}$, for example, $\mathcal{G}_{X,Z}^{IX}=\frac{1}{2}(-N_{XX,IZ}+N_{XX,ZI}+N_{YY,IZ}-N_{YY,ZI})$. In the case of diagonal noise, the channel map components can be found directly from Eq. (4). These are given in Table II, using the shorthand $\mathcal{N}_{\sigma,\sigma}=\sigma$.

Suppose now that we have the same diagonal noise $\mathcal{N}^{(1)}=[1, 1, x, x]$ on each qubit, with $x>0$. We see from Eq. (5) that this represents an X error with probability $p_x=(1-x)/2$. From Eq. (3),

$$\mathcal{G}^{II}=\left[\frac{1+x^2}{2}, \frac{1+x^2}{2}, x, x\right],$$

$$\mathcal{G}^{IX}=\mathcal{G}^{XI}=\left[\frac{1-x^2}{2}, \frac{1-x^2}{2}, 0, 0\right],$$

and the total map for this noise is

TABLE II. $\mathcal{G}^{R(\beta)}$ for diagonal noise, using the shorthand $\sigma=N_{\sigma,\sigma}$.

β	$R(\beta)$	$\mathcal{G}^{R(\beta)}$
0	II	$\frac{1}{2}[II+ZZ, XX+YY, XY+YX, IZ+ZI]$
1	XI	$\frac{1}{2}[II-ZZ, XX-YY, XY-YX, IZ-ZI]$
1	IX	$\frac{1}{2}[II-ZZ, XX-YY, YX-XY, ZI-IZ]$

$$\Omega^{bf_2}[1, 1, x, x]=\mathcal{G}=\mathcal{G}^{II}+\mathcal{G}^{IX}=\mathcal{G}^{II}+\mathcal{G}^{XI}=[1, 1, x, x].$$

It is evident that this code is not very useful for correcting bit-flip errors, since it leaves the bit-flip channel unchanged.

If the code is concatenated with itself, the resulting map on the bit-flip channel is $\Omega^{bf_2}\circ\Omega^{bf_2}[1, 1, x, x]=[1, 1, x, x]$, which is again ineffective at correcting bit-flip errors. The problem is that the recovery operators are also concatenated using this method. Let the recovery operators be $R=\{II, XI\}$. Then, for one of the syndromes, we apply the recovery operator $XXXI$ instead of the optimal $IIIX$. This syndrome has no error detected in the first block, an error detected in the second block, and an error detected at the top level of the code. The block noises are $\mathcal{N}_1=\mathcal{G}^{II}$ and $\mathcal{N}_2=\mathcal{G}^{XI}$. Applying the default XI recovery operator at the top level of the code (total recovery operator of $XXXI$) gives a logical noise of

$$\mathcal{G}^{XI}=\frac{1}{2}[II-ZZ, XX-YY, XY-YX, IZ-ZI]$$

$$=\frac{1}{2}\left[\frac{1+x^2}{2}\frac{1-x^2}{2}-x0, \frac{1+x^2}{2}\frac{1-x^2}{2}-x0, \frac{1+x^2}{2}0\right.$$

$$\left.-x\frac{1-x^2}{2}, \frac{1+x^2}{2}0-x\frac{1-x^2}{2}\right]$$

$$=\frac{1-x^4}{8}\left[1, 1, -\frac{2x}{1+x^2}, -\frac{2x}{1+x^2}\right].$$

For this quasicchannel, $p_X>p_I$ and $p_Y=p_Z=0$ [see Eq. (6)], and so the optimal recovery operator differs by encoded \bar{X} from this. Since $\bar{X}=XX$, this results in a recovery operator of IX instead of XI at the top level of the code (total recovery operator $IIIX$). This gives

$$\mathcal{G}^{IX}=[1, 1, -1, -1]\mathcal{G}^{XI}=\frac{1-x^4}{8}\left[1, 1, \frac{2x}{1+x^2}, \frac{2x}{1+x^2}\right].$$

The difference $\Delta=\mathcal{G}^{II}-\mathcal{G}^{XI}$ between these is

$$\Delta=[0, 0, YX-XY, ZI-IZ]=\left[0, 0, x\frac{1-x^2}{2}, x\frac{1-x^2}{2}\right].$$

Since this was the only suboptimal case, we optimize the total four-qubit code by adding Δ to the unoptimized channel map $[1, 1, x, x]$, yielding the optimized map

TABLE III. Critical values for depolarizing channel $(p_X, p_Y, p_Z) = (p, p, p)$.

Level	[[5,1,3]]	[[7,1,3]]
0	6.30965616%	6.30965616%
1	6.29873094%	6.25921455%
2	6.29795843%	6.26714580%
3	6.29850925%	6.268(8)%
4	6.299(0)%	6.269(6)%
5	6.299(3)%	6.270(0)%
6	6.299(5)%	6.270(3)%
7	6.299(6)%	6.270(3)%
∞	6.299(6)%	6.270(3)%
Unoptimized	4.58758548%	3.22981197%

$$\Omega^{bf4}[1, 1, x, x] = \left[1, 1, \frac{3}{2}x - \frac{1}{2}x^3, \frac{3}{2}x - \frac{1}{2}x^3 \right].$$

Since $x \rightarrow \frac{3}{2}x - \frac{1}{2}x^3$ converges to 1 for $x > 0$, iterating this map now yields a useful code for correcting X errors.

In general, the thresholds of correctable noise for a code under optimized recovery operators can be estimated by analyzing for which physical noise the entropy of the resulting logical noise is equal to some specific value near the threshold. Under repeated concatenation, these critical values converge to a threshold value. Generally, the entropy of the physical noise at the threshold is not 1. Noise below or above the threshold tends toward a logical entropy of 0 (no noise) or 2 (totally depolarizing). At the threshold, the entropy at each level of the code varies a lot less than the average encoded error. This allows for more efficient calculations of the thresholds than using the error rate [5]. We disprove the conjecture [5] that depolarizing noise can be corrected up to an entropy of 1 for the [[5, 1, 3]] code. The critical values are shown in Tables III and IV. These results indicate that the first two levels, which were calculated exactly, provide a good estimate of the error thresholds. Higher levels were evaluated with Monte Carlo calculations; for more details see [8]. In Table III, it can be seen that, for the depolarizing noise, the [[5, 1, 3]] code has a threshold (level ∞) of $p = 6.299(6)\%$, which is less than the physical noise with an entropy of 1 (level 0) of $p = 6.309\ 656\ 16\%$, disproving the conjecture of [5] that the threshold rate is equal to the latter.

TABLE IV. Critical values for a channel with independent probabilities p of bit flip and phase flips $(p-p^2, p^2, p-p^2)$.

Level	[[5,1,3]]	[[7,1,3]]
0	11.00278644%	11.00278644%
1	10.94668310%	10.94286393%
2	10.94728109%	10.95683308%
3	10.949(1)%	10.960(0)%
4	10.949(9)%	10.961(5)%
5	10.950(4)%	10.962(3)%
6	10.950(7)%	10.962(7)%
7	10.950(8)%	10.962(9)%
∞	10.951%	10.963%
Unoptimized	7.14780025%	6.45962393%

The value of p for which the channel Shannon entropy is 1 was calculated for various levels of concatenation for two different types of noise: the depolarizing channel (Table III) and independent bit and phase flips (Table IV). Tables III and IV compare our thresholds to the unoptimized thresholds obtained from the channel maps in Refs. [4,3]. The p values converge to the channel threshold for correctable noise at the ∞ level. The entropy at the thresholds for the [[5, 1, 3]] code [6] ranges from 0.993 [for $(p_X, p_Y, p_Z) = (p, 0, p)$ noise] to over 1. For the [[7, 1, 3]] code [7], both the phase-flip noise $(0, 0, p)$ and the channel $(p-p^2, p^2, p-p^2)$ have a critical value of $p = 10.963\%$, yielding minimum and maximum entropy of the thresholds as 0.4988 and 0.9976, respectively.

The results indicate that this adaptive concatenation approach greatly outperforms the concatenated channel map method. Table III shows that, for depolarizing noise, the value of p for the [[7, 1, 3]] code is increased from $p = 3.2298\%$ to $p = 6.270(3)\%$. The [[7, 1, 3]] code is particularly useful for quantum fault tolerance. It is therefore of interest to apply this adaptive concatenation method and similar entropy-based methods of threshold calculations to evaluate fault-tolerance thresholds [9].

For a discussion of which noise is correctable under the best code for correcting that noise, see [8].

We thank K. B. Whaley and M. Sarovar for reading and valuable criticisms of the manuscript. We thank the NSF for financial support under ITR Grant No. EIA-0205641.

[1] P. W. Shor, Phys. Rev. A **52**, R2493 (1995).

[2] E. Knill and R. Laflamme, Phys. Rev. A **55**, 900 (1997).

[3] B. Rahn, A. C. Doherty, and H. Mabuchi, Phys. Rev. A **66**, 032304 (2002).

[4] J. Fern, J. Kempe, S. Simić, and S. Sastry, IEEE Trans. Autom. Control **51**, 448 (2006).

[5] D. Poulin, Phys. Rev. A **74**, 052333 (2006).

[6] R. Laflamme, C. Miquel, J. P. Paz, and W. H. Zurek, Phys. Rev. Lett. **77**, 198 (1996).

[7] A. R. Calderbank and P. W. Shor, Phys. Rev. A **54**, 1098 (1996).

[8] J. Fern and K. B. Whaley, e-print arXiv:0708.1597.

[9] J. Fern, e-print arXiv:0801.2608