Error suppression in Hamiltonian-based quantum computation using energy penalties

Adam D. Bookatz,[*] Edward Farhi,[†] and Leo Zhou[‡]
Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
(Received 24 February 2015; published 10 August 2015)

We consider the use of quantum error-detecting codes, together with energy penalties against leaving the code space, as a method for suppressing environmentally induced errors in Hamiltonian-based quantum computation. This method was introduced in Jordan *et al.* [Phys. Rev. A **74**, 052322 (2006)] in the context of quantum adiabatic computation, but we consider it more generally. Specifically, we consider a computational Hamiltonian, which has been encoded using the logical qubits of a single-qubit error-detecting code, coupled to an environment of qubits by interaction terms that act one-locally on the system. Additional energy penalty terms penalize states outside of the code space. We prove that in the limit of infinitely large penalties, one-local errors are completely suppressed, and we derive some bounds for the finite penalty case. Our proof technique involves exact integration of the Schrodinger equation, making no use of master equations or their assumptions. We perform long time numerical simulations on a small (one logical qubit) computational system coupled to an environment and the results suggest that the energy penalty method achieves even greater protection than our bounds indicate.

DOI: 10.1103/PhysRevA.92.022317 PACS number(s): 03.67.Pp, 03.67.Lx

I. INTRODUCTION

A major problem on the road to building scalable quantum computers is the difficult task of protecting the system from errors, such as those due to unwanted environmental interactions. In the usual circuit model of quantum computation, the theory of quantum error correction has been well developed, culminating in the threshold theorem [1–5], which proves that, provided the error rate in a quantum computing system can be reduced to below a certain threshold, errors can be suppressed arbitrarily well using quantum error-correcting codes. The situation for the Hamiltonian model of quantum computation as used in, for example, adiabatic quantum computing, continuous-time quantum walks, and Hamiltonian simulation problems is less understood and no fault-tolerant theorem is known. In this paper, we take steps towards establishing such a theorem.

In the Hamiltonian model, the computational system is described by a Hamiltonian, which is a (possibly time-dependent) Hermitian operator, H_{comp}, that governs the time evolution of the system according to

$$i\frac{d}{dt}|\phi(t)\rangle = H_{\text{comp}}(t)|\phi(t)\rangle,$$

where $|\phi(t)\rangle$ is the state of the computational system at time t. In this model, the goal is to evolve some initial state $|\phi(0)\rangle$ to a final state $|\phi(T)\rangle$, the measurement of which reveals some information about the problem to be solved. Note that no instantaneous unitary gates are applied, nor are any intermediate measurements performed. To consider the effects of unwanted environmental interaction, one must consider the Hamiltonian $H_{\text{comp}} + H_{\text{environment}} + H_{\text{interaction}}$ that governs the evolution of the entire system-environment supersystem. The goal of error suppression is to ensure that the state of the system at time T is approximately as though the evolution had been governed by just H_{comp} alone.

It is not clear how to adapt the successful error-correcting code techniques of the circuit model to the Hamiltonian model. In a conventional quantum error-correcting code [6], each qubit is encoded as a logical qubit, comprised of several physical qubits, so that the occurrence of any single-qubit error on any physical qubit can be detected. The use of such a code in the error-correcting circuit model essentially consists of four steps: the state is encoded, the state is allowed to evolve, a measurement is made to determine what error has occurred (if any), and gates are applied to correct that error. In our Hamiltonian model, we do not allow intermediate measurements or the application of instantaneous gates, and therefore rule out any active determination and correction of errors; thus, a different strategy is required.

The error suppression strategy used in this paper is that of *energy penalties*, first suggested in [7], in which the system Hamiltonian is modified according to a quantum error-detecting code and a constant (time-independent) term is added to the Hamiltonian. This extra term, the energy penalty, penalizes states that have been corrupted by, say, single-qubit errors. It is believed that such a penalty will suppress the occurrence of environmentally induced errors, as it imposes an energy barrier that must be surmounted for an error to occur. In this work, we prove that, in principle, this energy penalty method does indeed work; we show that it successfully suppresses errors arbitrarily well when the penalty is arbitrarily large. (Throughout the paper we concentrate on one-local errors and use a one-qubit error detecting code. In the Appendix, however, we show that this result can be generalized to k-local errors when using a k-qubit error-detecting code.) We also explore (in the one-local error case) how well the penalty terms work when the penalty is not infinite but of a reasonable size. We then show the results of small-system numerical simulations that suggest that the achieved protection is even better than our bounds can predict.

We note that since we will not be performing active error correction, we do not need an error-*correcting* code, which gives information about which error occurred; rather, it suffices to use an error-*detecting* code, which only detects whether any error has occurred.

[*]bookatz@mit.edu
[†]farhi@mit.edu
[‡]leozhou@mit.edu

An error suppression technique using energy penalties has recently [8] been applied to the quantum annealing paradigm of computation. However, that method differs from the one discussed in this paper, as its energy penalty does not suppress phase-flip errors, while the penalty used in this paper suppresses arbitrary single-qubit errors. Other previously suggested Hamiltonian model error suppression methods include exploiting the Zeno effect [9,10] and using dynamical decoupling [11–14]. Some of these techniques require intermediate measurements, which is outside of the Hamiltonian paradigm, or require the ability to add rapidly time-dependent control terms to the Hamiltonian. The energy penalty method used in this paper remains in the Hamiltonian model paradigm and requires only the addition of a constant term to the encoded Hamiltonian. It would therefore be useful even when intermediate measurements and fast, active control are not available. A discussion of the similarity between the energy penalty, Zeno, and dynamical decoupling methods can be found in [15,16].

II. QUANTUM ERROR-DETECTING CODES

We first review some basic facts about quantum error detecting codes. Suppose that we have a $[[\ell,1]]$ quantum error detecting code, meaning that by encoding a single qubit as a logical qubit comprised of ℓ physical qubits, we can detect arbitrary one-qubit errors. Throughout this paper, we use this code to protect our system of n qubits, meaning that each qubit of the original H_{comp} is encoded to be ℓ qubits, so that the full encoded system consists of $n_s = \ell n$ qubits.

Specifically, for each qubit register i, the original computational basis states $|0\rangle_i$ and $|1\rangle_i$ are encoded as the ℓ-qubit logical states $|0_L\rangle_i$ and $|1_L\rangle_i$. The code space of the ith logical qubit is then the span of the logical states, $\{a|0_L\rangle_i + b|1_L\rangle_i : |a|^2 + |b|^2 = 1\}$. Associated with this code space is the projection operator

$$P_i = |0_L\rangle\langle 0_L|_i + |1_L\rangle\langle 1_L|_i,$$

where P_i acts as the identity on all physical qubits other than those associated with the logical qubit i. Note that states in the code space are invariant under P_i, whereas P_i kills states that are orthogonal to the code space of the ith qubit.

Saying that the code can detect arbitrary one-qubit errors is equivalent to saying that the code detects all single-qubit Pauli errors, i.e., an error caused by the application of a Pauli operator (X, Y, or Z) to any single physical qubit. Thus, for any single Pauli operator σ acting on one of the ℓ physical qubits comprising logical qubit i, we have

$$P_i \sigma P_i = 0. \tag{1}$$

The full code space for the entire logical space (over all n logical qubits) corresponds to the projector

$$P = P_1 P_2 \cdots P_n. \tag{2}$$

The quantum code also allows us to "encode" the Pauli operators X, Y, and Z as *logical operators* X_L, Y_L, and Z_L. Logical operators are Hermitian operators that have the same effect on the logical basis states as their corresponding Pauli operators have on the corresponding basis states. Furthermore, the logical operators associated with qubit i commute with the code space projector P_i, i.e., $X_L P_i = P_i X_L$, and similarly for Y_L and Z_L.

As a concrete example, consider the four-qubit code of Jordan-Farhi-Shor [7], in which

$$\begin{aligned}
|0_L\rangle &= \tfrac{1}{2}(|0000\rangle + i|0011\rangle + i|1100\rangle + |1111\rangle),\\
|1_L\rangle &= \tfrac{1}{2}(-|1010\rangle + i|1001\rangle + i|0110\rangle - |0101\rangle),\\
X_L &= Y \otimes \mathbb{1} \otimes Y \otimes \mathbb{1},\\
Y_L &= -\mathbb{1} \otimes X \otimes X \otimes \mathbb{1},\\
Z_L &= Z \otimes Z \otimes \mathbb{1} \otimes \mathbb{1}.
\end{aligned}$$

Observe that the logical operators have the same effect on logical qubits as the operators to which they correspond have on unencoded qubits, e.g., $X_L|0_L\rangle = |1_L\rangle$.

Using the logical operators, we can encode the Hamiltonian that acts on the system. Suppose that H_{comp} is some Hermitian operator on the original (n-qubit) system. Because the Pauli matrices (along with the identity) form a basis for all 2×2 matrices, we may generically write

$$H_{\text{comp}}(t) = \sum_{\substack{\sigma_i \in \{\mathbb{1},X,Y,Z\}\\ i=1,\ldots,n}} c_{\sigma_1,\ldots,\sigma_n}(t)\, \sigma_1 \otimes \cdots \otimes \sigma_n,$$

where the sum is over all possible choices of $\sigma_i \in \{\mathbb{1},X,Y,Z\}$ for each i. We may therefore encode the Hamiltonian by replacing X,Y,Z with X_L,Y_L,Z_L in the sum above, to obtain

$$H^{\text{L}}_{\text{comp}}(t) = \sum_{\substack{\sigma_i \in \{\mathbb{1},X_L,Y_L,Z_L\}\\ i=1,\ldots,n}} c_{\sigma_1,\ldots,\sigma_n}(t)\, \sigma_1 \otimes \cdots \otimes \sigma_n,$$

which is a Hamiltonian on the n_s-qubit encoded space built entirely out of logical operators (and $\mathbb{1}$). Since each logical operator commutes with each P_i, $H^{\text{L}}_{\text{comp}}$ also commutes with each P_i and with P.

Observe that the logical operators in the Jordan-Farhi-Shor code are all two-local. The encoding in this case thus doubles the locality of the original Hamiltonian, so that if the original Hamiltonian is two-local, the encoded one is four-local. As [7] points out, such an encoding is optimal (in terms of locality) if the code is to protect against arbitrary one-qubit errors.

III. HAMILTONIAN MODEL AND ENERGY PENALTIES

In this paper we consider a system coupled to an environment. We do not attempt to modify the environment or the system-environment interaction. However, we assume that we can modify the Hamiltonian of the system, and do so in two ways. As just discussed, we encode the original computational Hamiltonian in a quantum code. Furthermore, we add extra terms (acting only on the system) that penalize system states that are outside of the code space.

The combined system-environment Hamiltonian H, after encoding and penalty modifications, consists of three parts, and can be written as

$$H = H_0 + \lambda V + E_P \tilde{Q}.$$

We discuss each of these parts in turn.

(1) The first term is

$$H_0 = H_{\text{comp}}^{\text{L}} \otimes \mathbb{1}_{\text{env}} + \mathbb{1}_{\text{sys}} \otimes H_{\text{env}},$$

which governs the evolution in the absence of any system-environment interaction. Both $H_{\text{comp}}^{\text{L}}$ and H_{env} are in general time dependent. Evolution under $H_{\text{comp}}^{\text{L}}$ alone is equivalent to evolution under H_{comp} and represents the desired evolution we wish to protect.

Because the system Hamiltonian is encoded, the system consists of $n_s = \ell n$ qubits. The size of the environment will play no role in our discussion, except when we do simulations, and can be thought of as much larger than the system size. Note that $H_{\text{comp}}^{\text{L}}$ is built up from only logical operators and therefore commutes with each P_i. Since H_{env} (which acts only on the environment) trivially commutes with each P_i, we have

$$[H_0, P_i] = 0 \quad \text{for } i = 1, \ldots, n.$$

(2) λV is the error Hamiltonian, reflecting the coupling of the system to the environment, with λ serving as a time-independent (presumably small) parameter indicating the strength of the interaction (with units of energy), and V is a Hermitian operator acting on the full system-environment space. Our code is designed to protect against one-qubit errors, so we assume a one-qubit error model, i.e., that V acts one-locally on the system. Thus, we can write V as a sum of terms

$$V = \sum_{s=1}^{n_s} \sum_{\mu=X,Y,Z} \sigma_\mu^s \otimes B_\mu^s, \tag{3}$$

where σ_μ^s is the μth Pauli matrix acting on physical system qubit s and each B_μ^s is some operator acting on a small set of environmental qubits. We also allow the possibility that $B_\mu^s = \mathbb{1}$, which could represent one-local system control errors independent of the environment.

For convenience, we group terms in V according to the logical system qubit on which they act, so that

$$V = \sum_{i=1}^{n} V_i, \tag{4}$$

where each V_i is an operator whose one-local action on the system is only on the ℓ system qubits that comprise logical qubit i. Observe that V_i causes one-local errors on the system, as per Eq. (3), and that we are using a code that can detect arbitrary one-qubit errors, as per Eq. (1). Thus, we have that

$$P_i V_i P_i = 0 \quad \text{for } i = 1, \ldots, n,$$

which is crucial to our later analysis.

(3) $E_P \tilde{Q}$ is our time-independent energy penalty, which penalizes states outside of the code space. Specifically, E_P is a real constant with units of energy and $\tilde{Q}$ is the sum of the projectors $Q_i = \mathbb{1} - P_i$, i.e.,

$$\tilde{Q} = \sum_{i=1}^{n} Q_i = \sum_{i=1}^{n} (\mathbb{1} - P_i), \tag{5}$$

so we have a separate energy penalty for each logical qubit. In this context, $\tilde{Q}$ is to be understood as $\tilde{Q} \otimes \mathbb{1}_{\text{env}}$, since only the system is encoded. Observe that a state $|\psi\rangle$ is in the code space if and only if $\tilde{Q}|\psi\rangle = 0$, so $E_P \tilde{Q}$ applies an energy penalty of magnitude at least $|E_P|$ to states outside of (i.e., orthogonal to) the code space.

We point out that $\tilde{Q}$ is the sum of code-space projectors, differing from the penalty used in [7] which is a sum of code-space stabilizer generators. Note that the locality of $\tilde{Q}$ is that of each P_i, which is at most ℓ (i.e., 4 in the case of the Jordan-Farhi-Shor code).

The key point in this model is that V acts precisely one-locally on the system and we are using a quantum code that can detect one-qubit system errors. This enables us to penalize the states that arise from the action of V, and therefore have hope of suppressing V's effect. We can similarly consider the case in which V acts k-locally on the system as long as the quantum code can detect k-qubit errors. However, we consider only the one-local case throughout the paper, leaving the more general case to the Appendix.

IV. ERROR SUPPRESSION THROUGH ENERGY PENALTIES

A. Infinite E_P case

We first address the question of whether adding an energy penalty works even in principle; that is, we want to show that if E_P is arbitrarily large, errors are suppressed arbitrarily well. Let $U_0(t)$ and $U(t)$ be the evolution operators corresponding to the desired Hamiltonian, $H_0 = H_{\text{comp}}^{\text{L}} + H_{\text{env}}$, and the actual Hamiltonian, $H = H_0 + \lambda V + E_P \tilde{Q}$, respectively. That is, $U_0(t)$ and $U(t)$ obey

$$\begin{aligned} i\frac{d}{dt}U_0(t) &= H_0(t)U_0(t), \qquad U_0(0) = \mathbb{1}, \\ i\frac{d}{dt}U(t) &= H(t)U(t), \qquad\quad U(0) = \mathbb{1}. \end{aligned} \tag{6}$$

We wish to show that in the code space, as $E_P \to \infty$, $U(t)$ acts as $U_0(t)$. Our approach will be to show that the error induced by V is modulated by a term oscillating with frequency E_P in such a way so that for large E_P such errors are suppressed.

Our first step is to view λV as a perturbation and to work in the interaction picture using

$$H_{0P}(t) = H_0(t) + E_P \tilde{Q}$$

as the reference Hamiltonian. This corresponds to the evolution operator $U_{0P}(t)$, which obeys

$$i\frac{d}{dt}U_{0P}(t) = H_{0P}(t)U_{0P}(t), \quad U_{0P}(0) = \mathbb{1}.$$

Because H_0 commutes with each P_i, and therefore with $\tilde{Q}$, we have that

$$U_{0P}(t) = U_0(t)U_P(t), \tag{7}$$

where the evolution operator due to the error penalty alone is

$$U_P(t) = e^{-iE_P\tilde{Q}t}.$$

Now, the interaction picture evolution operator

$$U_I \equiv U_{0P}^\dagger U$$

obeys

$$i\frac{d}{dt}U_I = \lambda V_I U_I,$$

where

$$\lambda V_I(t) = \lambda U_{0P}^\dagger(t) V(t) U_{0P}(t). \tag{8}$$

These are just the usual interaction picture equations with a reference Hamiltonian H_{0P} and a perturbation λV. Taking conjugates, we get

$$U_I^\dagger = U^\dagger U_{0P} = U^\dagger U_0 U_P \tag{9}$$

and

$$\frac{d}{dt} U_I^\dagger = i\lambda U_I^\dagger V_I, \tag{10}$$

which upon integration gives

$$U_I^\dagger(T) = \mathbb{1} + i\lambda \int_0^T U_I^\dagger V_I dt. \tag{11}$$

Note that $\tilde{Q}P = 0$, so

$$U_P(t)P = e^{-iE_P\tilde{Q}t}P = P \tag{12}$$

and therefore

$$U_I^\dagger P = U^\dagger U_0 P.$$

Now, we multiply Eq. (11) on the right by P and use this last relation to get

$$U^\dagger(T)U_0(T)P = P + i\lambda \int_0^T U_I^\dagger V_I P dt.$$

Multiplying this by $U(T)$ gives

$$U_0(T)P = U(T)P + i\lambda U(T) \int_0^T U_I^\dagger V_I P dt, \tag{13}$$

which we can use to track the difference between the evolutions (in the code space) with and without the coupling to the environment. Our goal is to show that as E_P goes to infinity, this difference goes to zero. To this end, let

$$F(t) = \int_0^t V_I(\tau) P d\tau. \tag{14}$$

Using integration by parts, we see that

$$\begin{aligned}\int_0^T U_I^\dagger V_I P dt &= \int_0^T U_I^\dagger \frac{dF}{dt} dt \\ &= U_I^\dagger(T)F(T) - \int_0^T \frac{dU_I^\dagger}{dt} F dt \\ &= U_I^\dagger(T)F(T) - i\lambda \int_0^T U_I^\dagger V_I F dt,\end{aligned}$$

where Eq. (10) was used to obtain the final equality. Applying Eqs. (9) and (8) we can write this as

$$\int_0^T U_I^\dagger V_I P dt = U^\dagger(T)U_{0P}(T)F(T) - i\lambda \int_0^T U^\dagger V U_{0P} F dt$$

and using this in Eq. (13) we find that

$$\begin{aligned}U(T)P = U_0(T)P - i\lambda \Big[& U_{0P}(T)F(T) \\ & - i\lambda U(T) \int_0^T U^\dagger V U_{0P} F dt \Big],\end{aligned} \tag{15}$$

which is an exact expression and not just an expansion in λ.

We now focus on the operator $F(t)$ defined in Eq. (14), which using Eq. (8) for V_I and Eq. (7) for U_{0P} is

$$F(t) = \int_0^t U_P^\dagger U_0^\dagger V U_0 U_P P d\tau.$$

P commutes with H_0, and therefore also with U_0. Because of this and Eq. (12) we have

$$F(t) = \int_0^t U_0^\dagger e^{iE_P\tilde{Q}\tau} V P U_0 d\tau.$$

Consider

$$e^{iE_P\tilde{Q}\tau} V P = e^{iE_P\tilde{Q}\tau}(V_1 + \cdots + V_n)P, \tag{16}$$

where we have written V as the sum over terms associated with each logical qubit, as in Eq. (4). From the definitions in Eqs. (2) and (5), the first term is

$$e^{iE_P\tilde{Q}\tau} V_1 P = e^{iE_P Q_1 \tau} e^{iE_P Q_2 \tau} \cdots e^{iE_P Q_n \tau} V_1 P_1 \cdots P_n.$$

But, $P_2 P_3 \cdots P_n$ commutes with V_1, and $P_i Q_i = 0$ for all i, so we get

$$e^{iE_P\tilde{Q}\tau} V_1 P = e^{iE_P Q_1 \tau} V_1 P_1 \cdots P_n. \tag{17}$$

Our code protects against single-qubit errors and we are assuming that the coupling to the environment involves only single-qubit terms, so again,

$$P_1 V_1 P_1 = 0,$$

which implies that

$$V_1 P_1 = Q_1 V_1 P_1. \tag{18}$$

Because Q_1 is a projector, we have that

$$e^{iE_P Q_1 \tau} Q_1 = e^{iE_P \tau} Q_1. \tag{19}$$

The previous equations combine to give

$$e^{iE_P\tilde{Q}\tau} V_1 P = e^{iE_P\tau} V_1 P$$

and accordingly,

$$e^{iE_P\tilde{Q}\tau} V P = e^{iE_P\tau} V P. \tag{20}$$

Returning to $F(t)$, we thus have

$$F(t) = \int_0^t e^{iE_P\tau} U_0^\dagger(\tau) V(\tau) U_0(\tau) P d\tau. \tag{21}$$

Observe that $U_0^\dagger(\tau)V(\tau)U_0(\tau)P$ is independent of E_P. Therefore, we see that when E_P is large, the integrand of F is a rapidly oscillating function of τ due to the $e^{iE_P\tau}$. We can apply the Riemann-Lebesgue lemma to conclude that F vanishes as E_P goes to infinity. To be explicit, we perform an integration by parts to see that

$$\begin{aligned}F(t) &= \int_0^t e^{iE_P\tau} U_0^\dagger V U_0 P d\tau \\ &= \frac{1}{iE_P} \Big[e^{iE_P t} U_0^\dagger(t) V(t) U_0(t) \\ &\quad - V(0) - \int_0^t e^{iE_P\tau} \frac{d}{d\tau}(U_0^\dagger V U_0) d\tau \Big] P.\end{aligned} \tag{22}$$

The first two terms in the brackets do not grow with E_P and the third is bounded by t times the maximum magnitude of $\frac{d}{d\tau}(U_0^\dagger V U_0)$ which is independent of E_P. So as E_P goes to infinity, $F(t)$ goes to zero. Since both terms in the brackets in Eq. (15) contain F and are otherwise bounded independent of E_P, we have our E_P goes to infinity result:

Theorem. Suppose that the Hamiltonian of a system coupled to an environment is

$$H = H_{\text{comp}}^{\text{L}} + H_{\text{env}} + \lambda V + E_P \tilde{Q},$$

where V acts one-locally on the system, $H_{\text{comp}}^{\text{L}}$ is encoded in a quantum code that can detect single-qubit errors, and $\tilde{Q}$ is the operator defined in Eq. (5). Then, in the limit of an infinitely large energy penalty (positive or negative), the actual evolution in the code space is as if there were no errors due to V; i.e., for any time T,

$$\lim_{E_P \to \pm\infty} U(T)P = U_0(T)P,$$

where U and U_0 are the actual and error-free evolution operators defined in Eq. (6) and P is the code-space projection operator of Eq. (2).

This result applies to the evolution of both the system and the environment, and is therefore stronger than what we need, which is only that the system evolution be protected. We view our infinite E_P result as the starting point for large but finite E_P investigations.

Although throughout this paper we have focused only on the simplest case, where V acts one-locally on the system and a one-qubit quantum error-detection code is used, this simplification is not necessary. The theorem still holds as long as the error-detecting code can detect the errors that V causes, i.e., as long as $PVP = 0$, and therefore includes cases where V acts k-locally as long as the code can detect k-local errors. We show a proof of this in the Appendix. The remainder of the paper addresses the case where V acts one-locally but in which we use a finite, rather than infinite, penalty E_P.

B. Finite E_P case

1. Frequency analysis

We have just seen that for infinitely large E_P, the evolution in the code space in the presence of noise is the same as the desired noise-free evolution. We now want to know how large E_P must be to assure us that $F(t)$ is very small, so that the actual evolution in the code space is close to the desired one. It is helpful to consider the "natural frequencies" present in the expression for $F(t)$, as given by Eq. (21), which we informally analyze now.

If $f(t)$ is a (suitably nice) complex function, and $\tilde{f}(\omega)$ is its Fourier transfer, then

$$\begin{aligned}\int_0^t d\tau e^{iE_P\tau} f(\tau) &= \int_0^t d\tau e^{iE_P\tau} \int_{-\infty}^{\infty} d\omega e^{-i\omega\tau} \tilde{f}(\omega) \\ &= \int_{-\infty}^{\infty} d\omega \frac{e^{i(E_P-\omega)t} - 1}{i(E_P - \omega)} \tilde{f}(\omega).\end{aligned}$$

Suppose that there exists an ω_c such that $\tilde{f}(\omega)$ is non-negligible only for $|\omega| < \omega_c$. Then

$$\int_0^t d\tau e^{iE_P\tau} f(\tau) \approx \int_{|\omega|\leqslant\omega_c} d\omega \frac{e^{i(E_P-\omega)t} - 1}{i(E_P - \omega)} \tilde{f}(\omega).$$

Now, if E_P is much larger than ω_c, we can replace $1/(E_P - \omega)$ by $1/E_P$ in this integral, and may therefore conclude that the integral is small (shrinking as $1/E_P$).

The question is, therefore, what are the natural frequencies of $U_0^\dagger(\tau)V(\tau)U_0(\tau)$? If they are not too large, then $F(t)$ should be small for reasonably large values of E_P. Consider first the time-independent case, in which H_0 and V are time independent, so $U_0(\tau) = e^{-iH_0\tau}$. Certainly U_0 will have extremely large frequencies, namely $e^{-iE\tau}$ where E are eigenenergies of H_0; since H_0 includes the environment, E can scale with the size of the environment and be extremely large. However, the frequencies of $U_0^\dagger V U_0$, are expected to be much smaller. Inserting two complete sets of H_0 energy eigenstates, $|E\rangle$, we see that in the time-independent case,

$$U_0^\dagger(\tau)VU_0(\tau) = \sum_{E,E'} e^{i(E-E')\tau} |E\rangle\langle E|V\big|E'\big\rangle\big\langle E'\big|,$$

indicating that the frequencies are the energy *differences* induced by V. If V acts locally, we expect it would be unable to change the energy of the system and environment by a large amount—for example, it is unlikely that flipping just two spins in a spin chain will change the energy of the entire chain by more than a small amount. Therefore, we expect that $\langle E|V|E'\rangle$ is very small when $|E - E'|$ is large. If we make E_P larger than the largest $|E - E'|$ corresponding to any non-negligible $\langle E|V|E'\rangle$, we can conclude that F is small. To be more precise would require a specific model for the system, environment, and interaction. Still, we can make some progress on bounding F, even in the general time-dependent case.

2. Bounding F

We now bound the norm of $F(t) = \int_0^t e^{iE_P\tau} U_0^\dagger(\tau)V(\tau)U_0(\tau)P d\tau$. Since the norm of V is expected to grow linearly in the size of the system, and therefore in n, one would naively expect the same of F. However, the fact that each logical qubit is independently encoded allows us to do slightly better. Recall from Eq. (4) that we can write V as a sum of terms, V_i, where each V_i acts only on the ith logical system qubit (as well as the environment). Let

$$F_i(t) = \int_0^t e^{iE_P\tau} U_0^\dagger(\tau)V_i(\tau)U_0(\tau)P d\tau$$

so

$$F = \sum_{i=1}^n F_i.$$

We now show that

$$\|F\| \leqslant \sqrt{n} \max_i \|F_i\|. \tag{23}$$

Proof. Observe that

$$F^\dagger F = \sum_{i,j=1}^{n} F_i^\dagger F_j$$
$$= \sum_{i,j=1}^{n} \int_0^t \int_0^t d\tau_1 d\tau_2 e^{iE_P(\tau_2-\tau_1)}$$
$$\times [P U_0^\dagger(\tau_1) V_i(\tau_1) U_0(\tau_1)][U_0^\dagger(\tau_2) V_j(\tau_2) U_0(\tau_2) P]$$

and consider the terms with $i \neq j$. In the first P there is a P_j (i.e., $P = PP_j$) and it commutes with $U_0^\dagger(\tau_1), V_i(\tau_1)$, $U_0(\tau_1)$, and $U_0^\dagger(\tau_2)$. But $P_j V_j P = 0$ so these terms are 0. Consequently, the sum is only over $i = j$, i.e.,

$$F^\dagger F = \sum_{i=1}^{n} F_i^\dagger F_i,$$

and the claim follows since $\|F\|^2 = \max_{|\psi\rangle} \langle\psi| F^\dagger F |\psi\rangle$. ■

We now consider how to bound F_i (for any logical qubit i). In deriving Eq. (22), we assumed that $\frac{d}{d\tau}(U_0^\dagger V U_0)$ is finite; we now explicitly bound this term. By Eq. (6), we have

$$\frac{d}{d\tau}(U_0^\dagger V_i U_0) = -i U_0^\dagger [V_i, H_0] U_0 + U_0^\dagger \frac{dV_i}{d\tau} U_0.$$

Using this, Eq. (22) becomes

$$F_i(t) = \frac{1}{iE_P}\Bigg[e^{iE_P t} U_0^\dagger(t) V_i(t) U_0(t) - V_i(0)$$
$$+ i\int_0^t e^{iE_P\tau} U_0^\dagger [V_i, H_0] U_0 d\tau$$
$$- \int_0^t e^{iE_P\tau} U_0^\dagger \frac{dV_i}{d\tau} U_0 d\tau\Bigg] P$$

and taking the norm, using that $\|A+B\| \leqslant \|A\| + \|B\|$, $\|AB\| \leqslant \|A\|\|B\|$, $\|U_0\| = 1$, and $\|P\| \leqslant 1$, we obtain

$$\|F_i(t)\| \leqslant \frac{1}{|E_P|}\Bigg(\|V_i(t)\| + \|V_i(0)\|$$
$$+ \max_\tau \|[V_i(\tau), H_0(\tau)]\| t + \max_\tau \left\|\frac{dV_i}{d\tau}\right\| t\Bigg). \quad (24)$$

The norm $\|\frac{dV_i}{d\tau}\|$ will be bounded for reasonable V. For example, if the system control operations do not greatly change the environment surrounding each qubit, one expects that each V_i will likely stay fairly constant. Accordingly, we will ignore this term and the time dependence of V_i, in which case

$$\|F_i(t)\| \leqslant \frac{1}{|E_P|}(2\|V_i\| + \max_\tau \|[V_i, H_0(\tau)]\| t). \quad (25)$$

The commutator $[V_i, H_0] = [V_i, H_{\text{comp}}^{\text{L}} + H_{\text{env}}]$ involves the environment Hamiltonian, which may be extremely large; however, we now show that by making some reasonable physical assumptions, $\|[V_i, H_0]\|$ is independent of the size of the system and environment.

First, we assume that $H_{\text{comp}}^{\text{L}}$, H_{env}, and V_i are local operators. They can therefore each be written as a sum of terms, each term involving only a few qubits. Second, we make the assumption that each qubit (of the system and environment) appears in at most a few of these local terms of $H_{\text{comp}}^{\text{L}}$, H_{env}, and V_i. For example, if a Hamiltonian is two-local and *geometrically* local, say on a cubic lattice, so that each qubit only interacts with its immediate neighbors, then this restricts the number of terms in which any qubit appears, say to six for the cubic lattice. In terms of operator norms, these assumptions translate as follows.

For V_i we have

$$V_i = \sum_{s=1}^{\ell} \sum_{\mu=X,Y,Z} \sigma_\mu^s \otimes B_\mu^s,$$

where the sum over s is only over the ℓ system qubits that comprise the ith logical qubit. B_μ^s is an environmental operator that couples to σ_μ^s and only consists of a few local terms (because system qubit s only appears in a few local terms of V_i), each acting on only a few environmental qubits (by locality). Therefore,

$$\left\|B_\mu^s\right\| = \mathcal{O}(1)$$

independent of the system and environment sizes. (Recall that the coupling λ has units of energy, so the B_μ^s are dimensionless.) We thus have that

$$\|V_i\| = \ell\mathcal{O}(1). \quad (26)$$

Now, $H_0 = H_{\text{comp}}^{\text{L}} + H_{\text{env}}$ and both terms contribute to the commutator $[V_i, H_0]$. Let h_{sys}^s be the sum of all terms in $H_{\text{comp}}^{\text{L}}$ involving system qubit s, where s is a part of logical qubit i. Since there are only a few such terms, each of which acts on only a few system qubits, we can assert that

$$\left\|h_{\text{sys}}^s\right\| = \mathcal{E}\mathcal{O}(1),$$

where $\mathcal{E}$ is an energy scale parameter whose size is on the order of the size of the individual terms in $H_{\text{comp}}^{\text{L}}$. Similarly, let h_{env}^s be the sum of all terms in H_{env} that contain the environmental qubits that appear in B_μ^s for $\mu = X,Y,Z$. Since B_μ^s involves only a few environment qubits, which each appear in H_{env} in only a few, local terms, we have that

$$\left\|h_{\text{env}}^s\right\| = \mathcal{E}\mathcal{O}(1).$$

Then, since $\|A+B\| \leqslant \|A\| + \|B\|$, $\|[A,B]\| \leqslant 2\|A\|\|B\|$, and $\|A\otimes B\| = \|A\|\|B\|$,

$$\|[V_i, H_0]\| \leqslant \sum_{s=1}^{\ell} \sum_{\mu=X,Y,Z} \left\|\left[\sigma_\mu^s \otimes B_\mu^s, H_{\text{comp}}^{\text{L}}\right]\right\|$$
$$+ \left\|\left[\sigma_\mu^s \otimes B_\mu^s, H_{\text{env}}\right]\right\|$$
$$= \sum_{s=1}^{\ell} \sum_{\mu=X,Y,Z} \left\|\left[\sigma_\mu^s \otimes B_\mu^s, h_{\text{sys}}^s\right]\right\| + \left\|\left[\sigma_\mu^s \otimes B_\mu^s, h_{\text{env}}^s\right]\right\|$$
$$\leqslant 2\sum_{s=1}^{\ell} \sum_{\mu=X,Y,Z} \left\|\sigma_\mu^s\right\| \left\|B_\mu^s\right\| \left(\left\|h_{\text{sys}}^s\right\| + \left\|h_{\text{env}}^s\right\|\right).$$

Thus,

$$\left\|[V_i, H_0]\right\| = \ell\mathcal{E}\mathcal{O}(1) \quad (27)$$

independent of n and the size of the environment.

Applying the bounds of Eqs. (26) and (27) to Eq. (25) gives

$$\|F_i(t)\| \leqslant \frac{1}{|E_P|}[\ell\mathcal{O}(1) + \ell\mathcal{E}t\mathcal{O}(1)]$$

and using this in Eq. (23), we obtain

$$\|F(t)\| \leqslant \frac{\sqrt{n}}{|E_P|}\ell[\mathcal{O}(1) + \mathcal{E}t\mathcal{O}(1)]. \tag{28}$$

The term that grows with t represents a very weak bound for large t. We see from Eq. (21) that $F(t)$ is an integral over $[0,t]$ of an oscillating integrand and such integrals typically do not grow with t. For example, bounding

$$\int_0^t \sin(\omega\tau)d\tau \leqslant t,$$

while true, is not very helpful for large t. However, this is the best that we have been able to do for the general problem at hand. In Sec. V we will look at the full t dependence of small systems using numerical simulation.

3. Fidelity calculation

Suppose the system plus environment is initially in the pure state $|\psi\rangle$, and it evolves under U for time T. We begin in the code space of the system, so $P|\psi\rangle = |\psi\rangle$. The fidelity squared, $\mathcal{F}^2$, between the desired final state $U_0|\psi\rangle$ and the actual final state $U|\psi\rangle$ is given by

$$\mathcal{F}^2 = |\langle\psi|U_0^\dagger U|\psi\rangle|^2 = |\langle\psi|PU_0^\dagger UP|\psi\rangle|^2.$$

To evaluate this, we left-multiply Eq. (15) by $PU_0^\dagger$, and use Eq. (12) to give

$$PU_0^\dagger UP = P - i\lambda PF - \lambda^2 PU_0^\dagger U\int_0^T U^\dagger V U_{0P} F dt.$$

Because P commutes with U_0 and $PVP = 0$, we see from Eq. (21) that $PF = 0$. Therefore,

$$PU_0^\dagger UP = P - \lambda^2 PU_0^\dagger U\int_0^T U^\dagger V U_{0P} F dt\,,$$

so

$$\langle\psi|U_0^\dagger U|\psi\rangle = 1 - \lambda^2\langle\psi|PU_0^\dagger U\int_0^T U^\dagger V U_{0P} F d\tau\,|\psi\rangle\,. \tag{29}$$

Making the physical assumptions discussed above, we can immediately derive a bound on the fidelity. From Eq. (26), $\|V\| = n\ell\mathcal{O}(1)$, and using Eq. (28), along with the norm properties of $\|AB\| \leqslant \|A\|\|B\|$, $\|U_0\| = \|U\| = 1$, and $\|P\| \leqslant 1$, we obtain

$$\begin{aligned}\left|\lambda^2\langle\psi|PU_0^\dagger U\int_0^T U^\dagger V U_{0P} F d\tau\,|\psi\rangle\right| &\leqslant \lambda^2\max_{0\leqslant t\leqslant T}\|V\|\,\|F(t)\|\,T\\ &\leqslant \frac{\lambda^2 n^{3/2}\ell^2}{|E_P|}[T\mathcal{O}(1) + \mathcal{E}T^2\mathcal{O}(1)].\end{aligned}$$

Therefore, by the reverse triangle inequality, the fidelity is bounded by

$$\begin{aligned}\mathcal{F} &= |\langle\psi|U_0^\dagger U|\psi\rangle|\\ &= \left|1 - \lambda^2\langle\psi|PU_0^\dagger U\int_0^T U^\dagger V U_{0P} F d\tau\,|\psi\rangle\right|\\ &\geqslant 1 - \frac{\lambda^2 n^{3/2}\ell^2}{|E_P|}[T\mathcal{O}(1) + \mathcal{E}T^2\mathcal{O}(1)],\end{aligned}$$

so we are guaranteed good fidelity if we have

$$E_P \gtrsim \lambda^2 n^{3/2}\ell^2[T\mathcal{O}(1) + \mathcal{E}T^2\mathcal{O}(1)], \tag{30}$$

where by $\mathcal{O}(1)$ we mean the constants from Eqs. (26) and (27). For any efficient algorithm, $T \leqslant \text{poly}(n)$, so since λ, ℓ, and $\mathcal{E}$ are independent of n, it suffices for E_P to grow polynomially in the number of logical qubits n.

We assume that λ, the system-environment coupling, can be engineered to be small compared to the magnitudes of the individual terms in H_0. Accordingly, let us consider $\langle\psi|U_0^\dagger U|\psi\rangle$ to order λ^2. Working to this order, we can set $U = U_{0P}$ (as would occur if λ were zero) on the right hand side of Eq. (29):

$$\begin{aligned}\langle\psi|U_0^\dagger U|\psi\rangle = 1 &- \lambda^2\langle\psi|PU_0^\dagger U_{0P}\\ &\times\int_0^T U_{0P}^\dagger V U_{0P} F dt\,|\psi\rangle + \mathcal{O}(\lambda^3).\end{aligned}$$

Recall from Eqs. (7) and (12) that $U_{0P} = U_0 U_P$ and $PU_P = P$, so that $PU_0^\dagger U_{0P} = P$. Recalling the notation of Eq. (8) from the interaction picture, i.e., of $V_I \equiv U_{0P}^\dagger V U_{0P}$, and the definition of F in Eq. (14), we therefore have

$$\begin{aligned}\langle\psi|U_0^\dagger U|\psi\rangle &= 1 - \lambda^2\langle\psi|P\int_0^T V_I(t)F(t)dt\,|\psi\rangle + \mathcal{O}(\lambda^3)\\ &= 1 - \lambda^2\langle\psi|P\int_0^T V_I(t)\int_0^t V_I(\tau)P\,d\tau dt\,|\psi\rangle\\ &\quad + \mathcal{O}(\lambda^3).\end{aligned}$$

With perfect error suppression, $\mathcal{F}^2 \to 1$, so $1 - \mathcal{F}^2$ is a measure of error suppression failure. We calculate

$$\begin{aligned}1 - \mathcal{F}^2 &= 1 - |\langle\psi|U_0^\dagger U|\psi\rangle|^2\\ &= \lambda^2\langle\psi|P\int_0^T dt\int_0^t d\tau\; V_I(t)V_I(\tau)P\,|\psi\rangle\\ &\quad + \text{H.c.} + \mathcal{O}(\lambda^3),\end{aligned}$$

where H.c. denotes the Hermitian conjugate. But this conjugate involves

$$\begin{aligned}\left(\int_0^T dt\int_0^t d\tau\; V_I(t)V_I(\tau)\right)^\dagger &= \int_0^T dt\int_0^t d\tau\; V_I^\dagger(\tau)V_I^\dagger(t)\\ &= \int_0^T d\tau\int_0^\tau dt\; V_I(t)V_I(\tau),\end{aligned}$$

where in the last step we used the fact that V_I is Hermitian and relabeled $t \leftrightarrow \tau$, showing that this term is identical to the term of which it is the conjugate, except for the integration region. The original integrates over a region with $\tau < t$, while

the conjugate integrates the same integrand over a region with $t < \tau$, so their sum integrates over all $0 \leqslant \tau, t \leqslant T$. Thus,

$$1 - \mathcal{F}^2 = \lambda^2 \langle\psi| P \int_0^T dt\, V_I(t) \int_0^T d\tau\, V_I(\tau) P |\psi\rangle + \mathcal{O}(\lambda^3),$$

i.e.,

$$1 - \mathcal{F}^2 = \lambda^2 \langle\psi| F^\dagger F |\psi\rangle + \mathcal{O}(\lambda^3)$$

so

$$1 - \mathcal{F}^2 \leqslant \lambda^2 \|F\|^2 + \mathcal{O}(\lambda^3). \tag{31}$$

We see that a small $\|F\|$ corresponds to good error suppression.

We can combine this expression with Eq. (28) to obtain, at time T,

$$1 - \mathcal{F}^2(T) \leqslant \frac{\lambda^2 n}{E_P^2} \ell^2 [\mathcal{O}(1) + \mathcal{O}(1)\mathcal{E}T]^2 + \mathcal{O}(\lambda^3).$$

It is possible to write an expression for the λ^3 contribution. We find that the leading term in $1/E_P$ in the λ^3 contribution goes like $\lambda^3 T/E_P^2$. Again, we do not believe that this gives a useful bound for large T, but it may be useful in the small T regime.

V. NUMERICAL SIMULATION FOR ONE LOGICAL QUBIT

In this section, we discuss the results of a numerical simulation of one logical qubit, encoded as four physical qubits using the Jordan-Farhi-Shor [7] code, coupled to an eight-qubit environment according to

$$H = H_{\text{comp}}^{\text{L}} + H_{\text{env}} + \lambda V + E_P \tilde{Q}.$$

Since we track the evolution over long times, we find it too computationally expensive to work with more than 12 qubits total; therefore, for this paper we analyze only one logical qubit coupled to a modest size environment.

We choose the environment and the couplings as follows. The environment qubits are arranged on a randomly chosen three-regular graph and have two-local interactions between nearest neighbors. Each physical system qubit couples to a single, unique, randomly selected environment qubit. For simplicity, the environment and coupling Hamiltonians, H_{env} and V, are time independent.

We choose the environment Hamiltonian to be

$$H_{\text{env}} = \sum_{a=1}^{8} \alpha_a (\hat{n}_a \cdot \vec{\sigma}^a) + \sum_{\langle b,c\rangle} \alpha_{bc} (\hat{m}_b \cdot \vec{\sigma}^b) \otimes (\hat{\ell}_c \cdot \vec{\sigma}^c)$$

where each $\hat{n}_a$, $\hat{m}_b$, and $\hat{\ell}_c$ is a randomly chosen unit vector, $\vec{\sigma}^a = (\sigma_X^a, \sigma_Y^a, \sigma_Z^a)$ are the Pauli operators acting on environment qubit a, each α_a and α_{bc} is a coefficient chosen at random in the range of $[0.9, 1.1]$, and $\sum_{\langle b,c\rangle}$ denotes a sum over neighboring environment qubits on the three-regular graph.

In this small simulation, with one logical qubit, the system size is 4. The system-environment coupling has the form of Eq. (3), with the environmental operators chosen to be simple single-qubit terms, and is given by

$$V = \sum_{s=1}^{4} \beta_s (\hat{n}_s \cdot \vec{\sigma}^s) + \sum_{s=1}^{4} \gamma_s (\hat{m}_s \cdot \vec{\sigma}^s) \otimes \left(\hat{\ell}_s \cdot \vec{\sigma}_{\text{env}}^s\right),$$

where each $\hat{n}_s$, $\hat{m}_s$, and $\hat{\ell}_s$ is a randomly chosen unit vector, $\vec{\sigma}^s$ are the Pauli operators acting on system qubit s, and $\vec{\sigma}_{\text{env}}^s$ are the Pauli operators acting on the environment qubit that is coupled to system qubit s. Note that we have included single-qubit error terms, $\hat{n}_s \cdot \vec{\sigma}^s$, that are not coupled to any environment qubits but may arise from pure system errors. The coefficients β_s and γ_s are each chosen at random in the range $[0.9, 1.1]$. By design, V acts one-locally on the system.

The initial state is taken to be a pure product state of the system and environment,

$$|\psi\rangle = |\psi^s\rangle \otimes |\psi^e\rangle,$$

where the initial environment state $|\psi^e\rangle$ is a random eight-qubit state. We will study different choices for the initial system state $|\psi^s\rangle$ and the computational Hamiltonian $H_{\text{comp}}^{\text{L}}$. In order to compare the actual and desired dynamics, we evolve with U and U_0 defined in Eq. (6) to obtain

$$|\phi(t)\rangle = U(t)|\psi\rangle, \qquad t \in [0,T],$$
$$|\phi_0(t)\rangle = U_0(t)|\psi\rangle, \qquad t \in [0,T].$$

Note that because the system and environment are not coupled by H_0, we can write

$$|\phi_0(t)\rangle = \left|\phi_0^s(t)\right\rangle \otimes \left|\phi_0^e(t)\right\rangle$$

so that the state of the system at time t is $|\phi_0^s(t)\rangle$, independent of the environment. In the coupled case, on the other hand, the state of the system at time $t > 0$ is described by a density matrix,

$$\rho(t) = \text{Tr}_{\text{env}} |\phi(t)\rangle\langle\phi(t)|,$$

where the environment qubits have been traced out.

At any time t, we compare the actual versus coupling-free evolutions using the following measures:

(1) The squared fidelity of the total evolution,

$$\mathcal{F}^2(t) = |\langle\phi_0(t)|\phi(t)\rangle|^2.$$

As a result of our theorem, this measure goes to 1 as $E_P \to \pm\infty$. This fidelity also contains the fidelity of the environment's evolution, and accordingly is a stronger measure than what we need to track how well the computation is protected.

(2) The squared fidelity of the *system* evolution,

$$\mathcal{F}_s^2(t) = \left\langle\phi_0^s(t)\right|\rho(t)\left|\phi_0^s(t)\right\rangle.$$

This measure determines if the quantum computation in the presence of the coupling to the environment is following the desired evolution. The irrelevant bath degrees of freedom are traced out.

We first perform simulations for the time-independent computational Hamiltonian $H_{\text{comp}}^{\text{L}} = X_L$. Figure 1 shows the results of a typical simulation with $\lambda = 0.1$ and for a variety of E_P values, for both fidelity measures defined above. The initial system state in this case is a random superposition of $|0_L\rangle$ and $|1_L\rangle$, which can be viewed as a random superposition of the code-space eigenstates of $H_{\text{comp}}^{\text{L}} = X_L$, i.e., of

$$|\pm_L\rangle = \frac{1}{\sqrt{2}}(|0_L\rangle \pm |1_L\rangle).$$

We make the following observations:

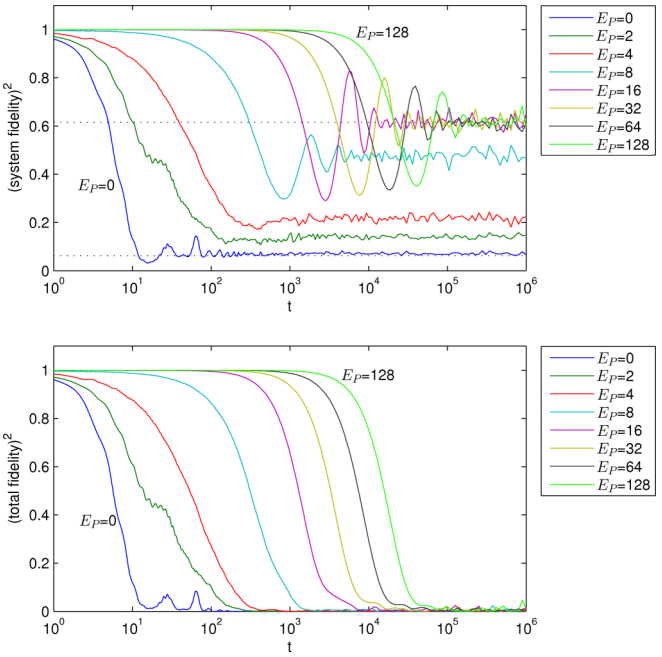

FIG. 1. (Color online) (Top) squared system fidelity $\mathcal{F}_s^2$ and (bottom) squared total fidelity $\mathcal{F}^2$ as functions of time t on a log scale for $\lambda = 0.1$ and initial system state $|\psi^s\rangle = \alpha|+_L\rangle + \beta|-_L\rangle$ with a random choice of α and β obeying $|\alpha|^2 + |\beta|^2 = 1$. Results are shown for increasing energy penalty strengths E_P. All data are for $H_{\text{comp}}^{\text{L}} = X_L$ on a four-qubit system, and for a particular random instance of H_{env}, V, $|\psi^s\rangle$, and $|\psi^e\rangle$ with eight environment qubits. The dashed line in the top panel is at $|\alpha|^4 + |\beta|^4$, which is 0.615 for this particular choice of $|\psi^s\rangle$; its significance will be explained later. The dashed line at $1/16$ is the expected long time system fidelity in the absence of protection.

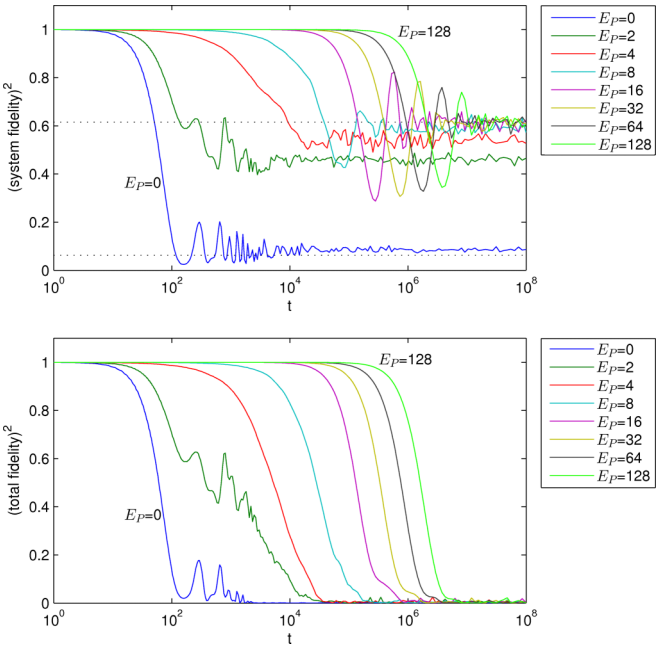

FIG. 2. (Color online) (Top) squared system fidelity $\mathcal{F}_s^2$ and (bottom) squared total fidelity $\mathcal{F}^2$ as functions of time t on a log scale for $\lambda = 0.01$. All other values are identical to those of Fig. 1, but the time scale has been increased because there is better protection for the smaller value of λ.

(1) In the absence of an energy penalty, i.e., when $E_P = 0$, the fidelities rapidly fall. We see that $\mathcal{F}_s^2$ falls to a value of about $1/16$, which is the expected fidelity between two random four-qubit system states. In other words, the state of the system is outside the code space and is uncorrelated with the state resulting from the desired evolution.

(2) For large E_P, near-perfect fidelity is maintained for a long time, both for the system ($\mathcal{F}_s$) and the system environment ($\mathcal{F}$). However, the fidelity eventually falls, and does so fairly abruptly (on a log scale). This kind of behavior would not be seen in a low-order power series expansion in time and is certainly not seen in expressions like Eq. (24) that have a linear term in t. Note that the larger the value of E_P, the longer near-perfect fidelity is maintained.

(3) For sufficiently large E_P, the general behavior is for the system fidelity $\mathcal{F}_s$ to approach an asymptotic value for large t, about which it has small fluctuations.

We have data for times greater than what we plot here that support this observation, but of course we cannot draw firm conclusions about what happens as $t \to \infty$. Still, we can say that the system fidelity stays fairly level away from zero at time scales much larger than any natural time scale involved in the simulation.

(4) The total fidelity $\mathcal{F}$ always falls to very close to 0 for very large t, indicating that the environment state is not as well protected as the system state is. This is unsurprising, as there is no preferred code space for the environment.

(5) In Fig. 2, we see qualitatively the same behavior for the same randomly chosen H_{env}, V, and $|\psi\rangle$, but with $\lambda = 0.01$ (rather than $\lambda = 0.1$). Note that for each E_P, the smaller λ value allows for good protection for longer times than the larger λ value allows.

It is interesting to compare the bounds of Eqs. (31) and (25) with our numerical observations. For the parameters used to generate Figs. 1 and 2, we have $\|V\| \approx 7$, $\|H_0\| \approx 12$, and $\|[V, H_0]\| \approx 17$ (significantly less than $\|V\| \cdot \|H_0\|$, in accordance with our previous discussion on locality). Equation (31) suggests that good fidelity squared, say of 0.9, can be achieved if $\lambda^2\|F\|^2 \lesssim 0.1$, so for $\lambda = 0.1$ we expect that we need $\|F\| \lesssim 3$. The bound in Eq. (25) indicates that for $E_P = 32$, $\|F\| \lesssim 3$ for $T \lesssim 5$, so that these two bounds together suggest that good fidelity can be maintained for time $T \lesssim 5$ if $E_P = 32$. However, in Fig. 1 we see that, in this case, we can maintain good $\mathcal{F}^2$ up to $T = 1000$. Similarly, for $\lambda = 0.01$ we expect that we need $\|F\| \lesssim 30$ [from Eq. (31)], which for $E_P = 32$ can be guaranteed for $T \lesssim 60$ [by Eq. (25)]; however, Fig. 2 indicates that in this case we can maintain good $\mathcal{F}^2$ up to $T = 100\,000$. We thus see that Eq. (25) is not really useful for large T, as our numerical results show good fidelity for far longer than our bounds can guarantee.

To address the question of how long good fidelity can be maintained, we note that for successful quantum computation it suffices to have high system fidelity $\mathcal{F}_s$; high total fidelity $\mathcal{F}$ is not required. Accordingly, we define the *protection time* t_{prot} to be the time at which the squared system fidelity $\mathcal{F}_s^2$ first

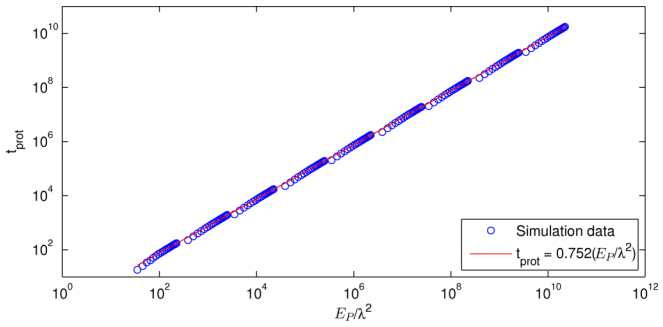

FIG. 3. (Color online) The protection time t_{prot} defined as the time at which the squared system fidelity $\mathcal{F}_s^2$ drops to 0.9, vs E_P/λ^2 for a range of E_P and λ values, specifically, $E_P \in \{35,45,\ldots,225\}$ and $\lambda \in \{10^{-4},3\times10^{-4},10^{-3},3\times10^{-3},10^{-2},3\times10^{-2},10^{-1},3\times10^{-1},1\}$ (each of the nine "clusters" of data in the figure corresponding to a different λ value). All Hamiltonian and initial-state values (other than E_P and λ) are kept identical to those of Fig. 1. The line shows that a linear relationship between t_{prot} and E_P/λ^2 fits the data well.

drops to 0.9. In Fig. 3 we plot t_{prot} for a variety of values of $E_P \in [35,225]$ and $\lambda \in [10^{-4},1]$ (with all other Hamiltonian and initial state values held fixed). Observe that, to a very good approximation, the data fit the relation

$$t_{\text{prot}} \propto E_P/\lambda^2$$

especially for larger values of E_P. We will later show a simple model that is consistent with this behavior.

We next address the question of what the system fidelity falls to at late times for large E_P. For the Hamiltonian $H_{\text{comp}}^{\text{L}} = X_L$, given $|\psi^s\rangle$ we can actually predict the long-term system fidelity. To help uncover this relationship, we plot in Fig. 4(a) the system fidelity as a function of time, with $|\psi^s\rangle$ taken to be $|0_L\rangle$. Note that the long term system fidelity is very near $1/2$ for $E_P \geqslant 16$. In Fig. 4(b) we show the same thing but with $|\psi^s\rangle = |+_L\rangle$, an eigenstate of X_L, and see that the long term system fidelity is very near 1 for $E_P \geqslant 16$. More generally, we observe that if we write

$$|\psi^s\rangle = \alpha|+_L\rangle + \beta|-_L\rangle, \tag{32}$$

with $|\pm_L\rangle$ being the code-space eigenstates of X_L and $|\alpha|^2+|\beta|^2=1$, then the long-term system fidelity is well approximated by $|\alpha|^4+|\beta|^4$. In Fig. 5 we show the long-term system fidelity versus $|\alpha|^4+(1-|\alpha|^2)^2$ for a set of randomly chosen $|\psi^s\rangle$ and the good fit is apparent.

We show in Fig. 6, for the three choices of $|\psi^s\rangle$ displayed in Figs. 1, 4(a), and 4(b), the probability to remain in the code space, $\langle\phi(t)|P|\phi(t)\rangle$. We see that it is close to 1 for all displayed times for $E_P \geqslant 16$, indicating that any loss of system fidelity is occurring because of errors inside the code space. With $H_{\text{comp}}^{\text{L}} = X_L$, the desired evolution, starting with the state in Eq. (32), is

$$\left|\phi_0^s(t)\right\rangle = \alpha e^{-it}|+_L\rangle + \beta e^{it}|-_L\rangle$$

since the code-space eigenvalues of X_L are ±1. Imagine that the only effect of the coupling to the environment is to induce dephasing in the $H_{\text{comp}}^{\text{L}}$ energy eigenbasis. Then the density

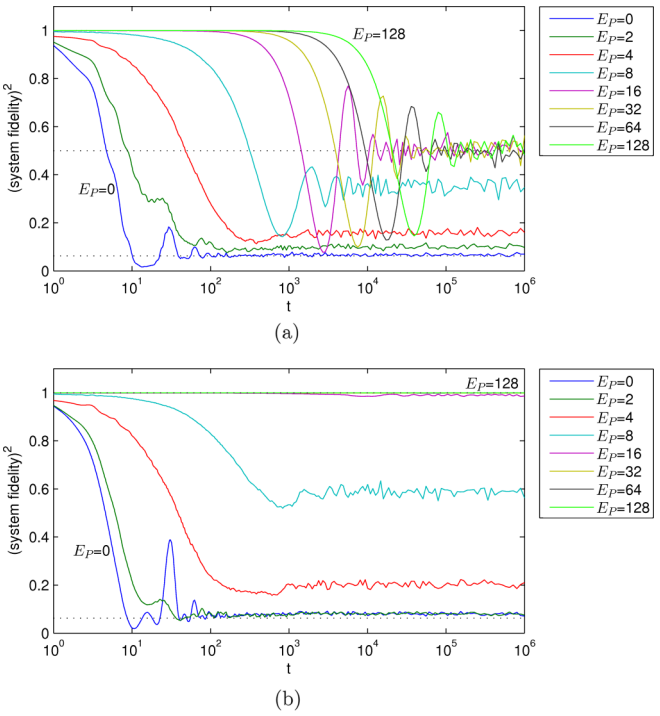

FIG. 4. (Color online) Squared system fidelity $\mathcal{F}_s^2$ as a function of time t on a log scale for initial states (a) $|\psi^s\rangle = |0_L\rangle$ and (b) $|\psi^s\rangle = |+_L\rangle$, with $H_{\text{comp}}^{\text{L}} = X_L$ and $\lambda = 0.1$. All Hamiltonian and initial environment state values are identical to those of Fig. 1. The dashed lines at $1/2$ (in the top figure) and $1/16$ (in both figures) serve as guides for the eye. Note that in the bottom figure, for $E_P \geqslant 16$, $\mathcal{F}_s^2$ remains close to 1 for the duration of the simulation.

matrix of the system will approach

$$\rho(t) = |\alpha|^2|+_L\rangle\langle+_L| + |\beta|^2|-_L\rangle\langle-_L|$$

and the squared system fidelity, $\langle\phi_0^s(t)|\rho(t)|\phi_0^s(t)\rangle$, is $|\alpha|^4+|\beta|^4$. That the data in Fig. 5 match this is good evidence that the

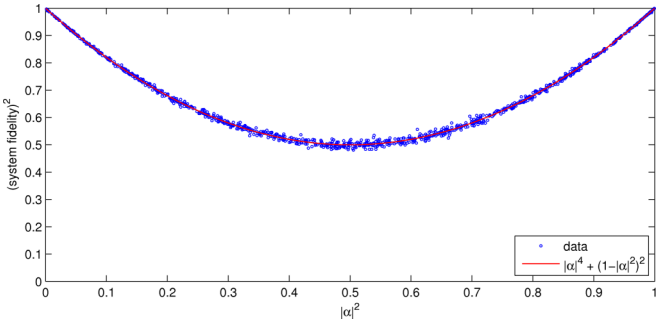

FIG. 5. (Color online) The long-term squared system fidelity $\mathcal{F}_s^2$ as a function of $|\alpha|^2$, where the initial system state is $|\psi^s\rangle = \alpha|+_L\rangle + \beta|-_L\rangle$ and $|\alpha|^2+|\beta|^2=1$. The curve $y=|\alpha|^4+(1-|\alpha|^2)^2$ is also shown, and the good fit is apparent. Each data point represents a random choice for α and β, as well as V, H_{env}, and the initial environment state $|\psi^e\rangle$. The computational Hamiltonian is $H_{\text{comp}}^{\text{L}} = X_L$ and $E_P = 128$. Each data point is the average $\mathcal{F}_s^2(T)$ over the times $T = \{1,2,\ldots,10\}\times10^8$ to account for fluctuations in time of $\mathcal{F}_s$ about the long-term system fidelity.

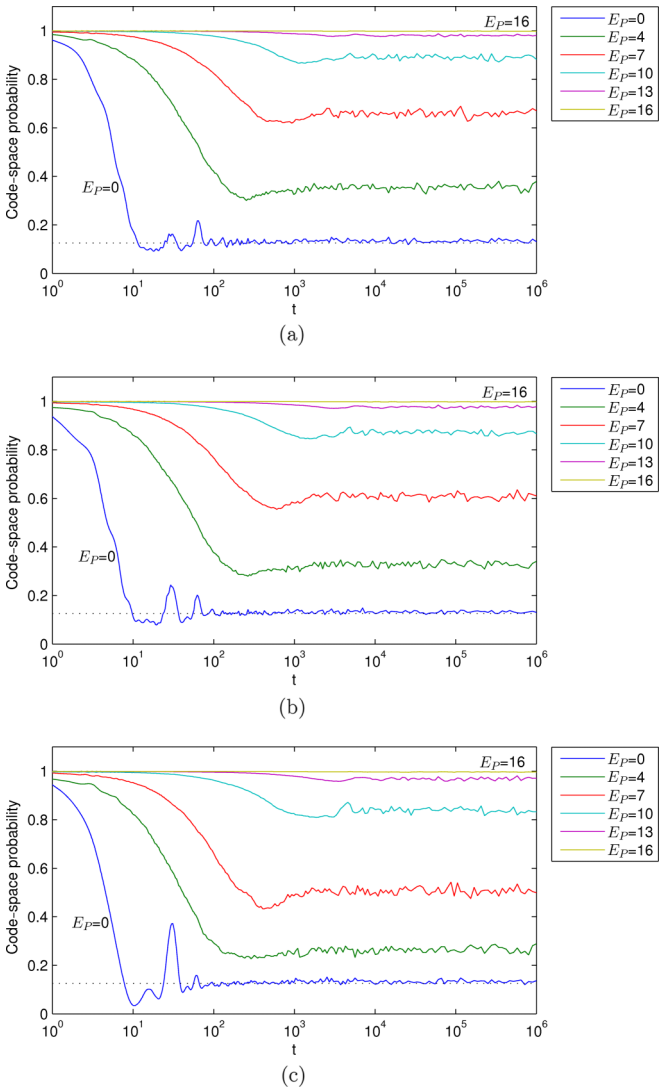

FIG. 6. (Color online) The probability $\langle\phi(t)|P|\phi(t)\rangle$ of the system being found in the code space for the three different initial states used in Figs. 1, 4(a), and 4(b), namely, (a) $|\psi^s\rangle = \alpha|+_L\rangle + \beta|-_L\rangle$, (b) $|\psi^s\rangle = |0_L\rangle$, and (c) $|\psi^s\rangle = |+_L\rangle$. The dashed line at 1/8 represents the expected probability for a maximally mixed four-qubit state to be found in the one-qubit code space. In all three cases, for $E_P \geqslant 16$ the code-space probability is very near 1 for all displayed times.

effect of the coupling to the environment is to cause dephasing in the energy basis of $H^{\mathrm{L}}_{\mathrm{comp}}$.

In our simulation we see that, for sufficiently large energy penalties, the system remains in the code space and decoheres inside the code space via dephasing of the energy eigenstates. We now present a simple phenomenological model that allows us to estimate t_{prot}, the time at which the effects of decoherence become appreciable. The model has three states. The first two states are the code-space eigenstates, $|+\rangle$ and $|-\rangle$, of the logically encoded two-level computational Hamiltonian with energies ω and $-\omega$. The third state is a penalty state, representing all the states orthogonal to the code space, and accordingly has energy $E_P \gg \omega$. The third state is coupled to the first two as a result of interactions with the environment, so that the effective Hamiltonian is

$$H_{\mathrm{eff}} = \begin{bmatrix} \omega & 0 & \lambda_+ \\ 0 & -\omega & \lambda_- \\ \lambda_+ & \lambda_- & E_P \end{bmatrix}.$$

Here, λ_+ and λ_- are the effective couplings of the first two states to the penalty state, and we assume that they are small compared to ω. We imagine that λ_+ and λ_- are proportional to some constant λ that represents the overall scale of the effective couplings. Expanding to lowest order in λ_+, λ_-, and $1/E_P$, we find that

$$\|\langle -|e^{-iH_{\mathrm{eff}}t}|+\rangle\|^2 \lesssim \left(\frac{\lambda_+\lambda_-}{\omega E_P}\right)^2,$$

so in this model the transition probability between states $|+\rangle$ and $|-\rangle$ is negligible for all time.

Treating the coupling as a perturbation, the effect of the coupling of the system states to the penalty state is to shift their energies. The perturbed energies are calculated to be

$$E_\pm = \pm\omega - \frac{\lambda_\pm^2}{E_P}$$

to lowest order in λ_+, λ_-, and $1/E_P$. Thus in this little model, at time t, the interaction-induced phase difference between $|+\rangle$ and $|-\rangle$ is

$$(E_+ - E_- - 2\omega)t = -\frac{\lambda_+^2 - \lambda_-^2}{E_P}\,t$$

so that the characteristic dephasing time is proportional to E_P/λ^2. Generalizing from the toy model to an encoded two-level logical system with a coupling to the environment of size λ and energy penalty term of size E_P, we guess that

$$t_{\mathrm{prot}} \propto \frac{E_P}{\lambda^2},$$

in agreement with the behavior seen in Fig. 3.

Returning to the simulation results, we have seen that a sufficiently large energy penalty keeps the system in the code space, even for large t. We also presented evidence that decoherence inside the code space occurs via dephasing in the energy basis. In particular, with a time-independent $H^{\mathrm{L}}_{\mathrm{comp}} = X_L$, starting in an energy eigenstate, say $|+_L\rangle$, we find that for sufficiently large E_P the system remains approximately in that eigenstate for the duration of the simulation. In adiabatic quantum computation [17], the state of the system is initially the ground state of a time-dependent computational Hamiltonian and, provided that the computational Hamiltonian is changed slowly enough, the evolving state is expected to remain near the instantaneous ground state. One might therefore expect good fidelity in the adiabatic computation case as well.

We now show the results of simulations for the one logical qubit adiabatic computation

$$H^{\mathrm{L}}_{\mathrm{comp}}(t) = \left(1 - \frac{t}{T}\right)X_L + \frac{t}{T}Z_L,$$

where the initial system state $|\psi^s\rangle = \frac{1}{\sqrt{2}}(|0_L\rangle - |1_L\rangle)$ is the ground state of $H^{\mathrm{L}}_{\mathrm{comp}}(0)$. The results are shown in Fig. 7 for

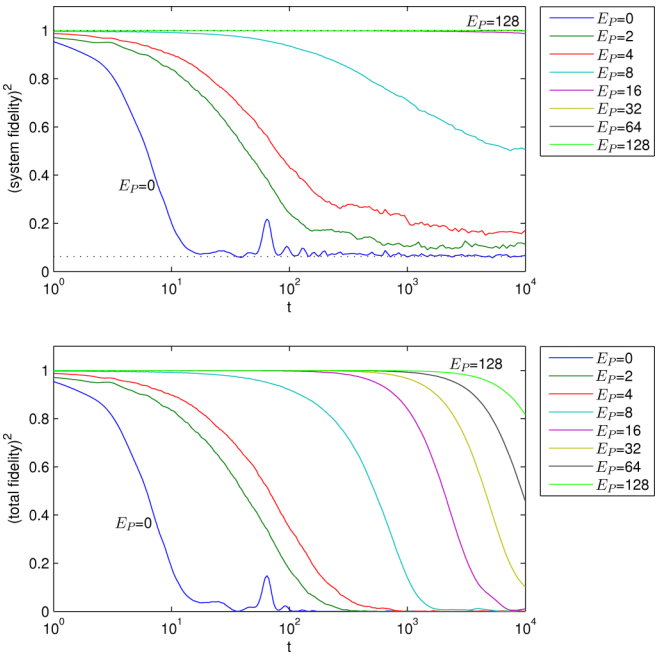

FIG. 7. (Color online) For the adiabatic computation, $H^{\mathrm{L}}_{\mathrm{comp}}(t) = (1-\frac{t}{T})X_L + \frac{t}{T}Z_L$ with $T = 10\,000$, the (top) squared system fidelity $\mathcal{F}_s^2$ and (bottom) squared total fidelity $\mathcal{F}^2$ as functions of time t for $\lambda = 0.1$. All data are for a particular random instance of H_{env}, V, and $|\psi^e\rangle$ with eight environment qubits, with the system initially in the ground state of $H^{\mathrm{L}}_{\mathrm{comp}}(0)$. Note that for $E_P \geqslant 16$, we have nearly perfect system fidelity throughout the evolution.

$T = 10\,000$. Observe that for $E_P \geqslant 16$, the system fidelity remains very high for the duration of the computation.

We emphasize that our numerical results are for a small system (one logical qubit made of four physical qubits) coupled to a small environment (of eight qubits). We do not know if the observations we have made for one logical qubit will hold in more complicated systems with many logical qubits. In particular, we would like to know if with a large number of qubits, modest energy penalties can keep the system in the code space and, if inside the code space, whether the decoherence is limited to dephasing in the energy basis. If so, this would be of great help in protecting adiabatic quantum computation. Furthermore, we are concerned that in our simulations, the size of the environment may be too small, especially given the large values of E_P that we are exploring. It would be disappointing if our encouraging small system simulation results are artifacts of having too small an environment or do not reflect what actually happens in large systems. Nevertheless, these numerical results, in conjunction with the proof that the energy penalty method works in the infinite E_P limit, suggest that the energy penalty method may be a useful approach towards the development of fault-tolerant Hamiltonian-based quantum computing.

VI. OUTLOOK

To use the energy penalty method in an actual device, some practical hurdles remain to be overcome. The logical operators used by the codes discussed in this paper need to be at least four-local (in order to detect arbitrary one-local errors), whereas physically implementable Hamiltonians are generally constrained to be two-local. The usual procedure to overcome such locality constraints is to use so-called perturbative gadgets (as introduced in [18]), which allow one to construct a two-local Hamiltonian whose low-energy subspace approximates a given desired Hamiltonian. Such techniques can perhaps be used here, too, to achieve error suppression using only two-local operations and energy penalties. Another technique that might work for certain situations is to use codes that do not correct arbitrary errors but have a smaller locality, similar to what was done in Ref. [8] in the context of quantum annealing. This would be useful in situations in which it is known that only certain types of errors are problematic. In addition to allowing for only two-local Hamiltonians, such codes may admit fewer physical qubits per logical qubit (i.e., a smaller value of ℓ), reducing the total error on the system and enabling numerical simulations for a larger number of logical qubits than we have been able to do in this paper.

Another potential hurdle is the scaling of the energy penalty. To maintain a given desired fidelity, it is possible that the magnitude of the required energy penalty E_P depends on the size of the system. Fortunately, Eq. (30) shows that under reasonable physical assumptions, such a scaling is at most polynomial in the number n of logical qubits. We hope that this inequality can be tightened further. For scalable implementation of Hamiltonian-based quantum computing with error suppression, it is likely that changing the energy penalty (even polynomially) to accommodate increases in the logical system size may be difficult to do. For a practical fault-tolerant theorem, it would be desirable for the error suppression to be ultimately achieved through the addition of extra qubits, as in the circuit-model case, rather than requiring hardware modifications (such as increasing the magnitude of energy penalty terms). Recently, it has been shown [19] how two-local perturbative gadgets can be used to achieve effectively large energy penalties using much weaker energy interactions, at the expensive of having additional qubits. It would be of great interest to see if such a technique could be applied here to develop a threshold theorem for scalable Hamiltonian-based quantum computing. Unfortunately, the technique in [19] requires a large overhead in the number of interaction terms per qubit, which is likely physically unrealistic and is in opposition to the physical assumptions we made in deriving our fidelity bounds. Nonetheless, it may be a fruitful avenue for future research.

VII. CONCLUSION

In this paper, we considered the energy penalty method of error suppression, i.e., the method of achieving error suppression by encoding a Hamiltonian using a quantum error detecting code and adding a constant term that penalizes states outside of the code space. We proved that this method does indeed work in principle. Specifically, we showed that, in the limit of an infinitely large energy penalty, the actual evolution of the system is precisely the evolution in the absence of unwanted control errors and environmental interactions, provided that the code can detect these errors. Moreover, we have provided some bounds governing the finite energy penalty scenario, allowing one to bound the energy penalty required to

attain the desired evolution with good fidelity. We believe that these bounds can be improved, as supported by our numerical evidence for a single logical qubit, and leave their tightening as an interesting open problem. We hope that progress in this area will eventually lead to a practical fault-tolerant paradigm for Hamiltonian-based quantum computation.

ACKNOWLEDGMENTS

The authors are grateful to R. Blume-Kohout, D. Lidar, S. Lloyd, and P. Shor for valuable discussions. E.F. thanks members of the Google Quantum Artificial Intelligence Lab for their input. We gratefully acknowledge support from the U.S. Department of Energy under Cooperative Research Agreement Contract No. DE-FG02-05ER41360, the Army Research Office under Grant No. W911NF-12-1-0486, and from the National Science Foundation under Award No. CCF-121-8176. L.Z. acknowledges support from the MIT Undergraduate Research Opportunities Program.

APPENDIX

In this paper we focused on the simplest case, where V acts one-locally on the system and the quantum error-detecting code can detect one-qubit errors. In this Appendix we show that this simplification is not necessary. As long as the error detecting code can detect the errors that V causes, our infinite energy penalty theorem still holds. This includes, for example, the case where V acts k-locally and the code can detect k-local errors. Specifically, the only requirement on V is that

$$PVP = 0. \tag{A1}$$

We now present a proof of this general case.

Define R_r (for $r = 0, \ldots, n$) to be

$$R_r = \sum \{A_1 \otimes \cdots \otimes A_n : A_i \in \{P_i, Q_i\} \text{ such that } |\{i : A_i = Q_i\}| = r\},$$

where as before, P_i is the code-space projector for the ith logical qubit and $Q_i = \mathbb{1} - P_i$. In other words, R_r is the sum of all terms, each of which is a tensor product of a total of n P_i's and Q_i's, one for each logical qubit, such that precisely r of these projectors are Q_i's. For example, $R_0 = P$, $R_n = Q_1 Q_2 \cdots Q_n$, and

$$R_1 = Q_1 P_2 \cdots P_n + \cdots + P_1 \cdots P_{n-1} Q_n.$$

Observe that the R_r are in fact a complete set of orthogonal projectors:

$$R_r^2 = R_r \qquad \text{for all } r,$$
$$R_r R_{r'} = 0 \qquad \text{for } r \neq r',$$
$$\sum_{r=0}^{n} R_r = \mathbb{1},$$

where the last equality can be obtained by expanding out $\mathbb{1} = \prod_i (P_i + Q_i)$.

Now, recall that $e^{iE_P \tau Q_i} P_i = P_i$ and that $e^{iE_P \tau Q_i} Q_i = e^{iE_P \tau} Q_i$. Therefore, using the definition of $\tilde{Q}$ in Eq. (5), we see that for any r,

$$U_P^\dagger(\tau) R_r = e^{iE_P \tau \tilde{Q}} R_r = \prod_{i=1}^{n} e^{iE_P \tau Q_i} R_r = e^{irE_P \tau} R_r$$

because each term in R_r consists of precisely r Q_i's. Applying $U_P^\dagger$ to $\mathbb{1} = \sum_{r=0}^{n} R_r$ therefore lets us write

$$U_P^\dagger(\tau) = \sum_{r=0}^{n} e^{irE_P \tau} R_r$$

so that applying $U_P^\dagger$ to VP gives

$$U_P^\dagger VP = \sum_{r=0}^{n} e^{irE_P \tau} R_r VP.$$

We now apply our key requirement of Eq. (A1) to see that the $r = 0$ term is $R_0 VP = PVP = 0$. Thus, we have

$$U_P^\dagger VP = \sum_{r=1}^{n} e^{irE_P \tau} R_r VP,$$

instead of the one-local version in Eq. (20), and our formula for F from Eq. (21) therefore generalizes to

$$F(t) = \sum_{r=1}^{n} \int_0^t e^{irE_P \tau} R_r U_0^\dagger(\tau) V(\tau) U_0(\tau) P d\tau.$$

Note that every term in $F(t)$ has a phase of $e^{irE_P \tau}$ for some $r > 0$. Applying the Riemann-Lebesgue lemma, we again conclude that in the infinite E_P limit $F(t) \to 0$ and our theorem follows. This form of F may be useful in deriving finite energy penalty bounds in the case that we have a code that can protect against more than one-local errors.

[1] D. Aharonov and M. Ben-Or, Fault-tolerant quantum computation with constant error, in *Proceedings of the Twenty-Ninth Annual ACM Symposium on the Theory of Computing* (ACM, New York, 1997), pp. 176–188.

[2] A. Y. Kitaev, Quantum error correction with imperfect gates, in *Quantum Communication, Computing, and Measurement*, edited by A. S. Holevo, O. Hirota, and C. M. Caves (Plenum, New York, 1997), pp. 181–188.

[3] A. Y. Kitaev, Quantum computations: Algorithms and error correction, Russ. Math. Surv. **52**, 1191 (1997).

[4] D. Gottesman, Stabilizer codes and quantum error correction, Ph.D. thesis, California Institute of Technology, Pasadena, CA, 1997.

[5] J. Preskill, Reliable quantum computers, Proc. R. Soc. London A **454**, 385 (1998).

[6] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2000).

[7] S. P. Jordan, E. Farhi, and P. W. Shor, Error-correcting codes for adiabatic quantum computation, Phys. Rev. A **74**, 052322 (2006).

[8] K. P. Pudenz, T. Albash, and D. A. Lidar, Error corrected quantum annealing with hundreds of qubits, Nat. Commun. **5**, 3243 (2014).

[9] B. Misra and E. C. G. Sudarshan, The Zeno's paradox in quantum theory, J. Math. Phys. **18**, 756 (1977).

[10] G. A. Paz-Silva, A. T. Rezakhani, J. M. Dominy, and D. A. Lidar, Zeno effect for quantum computation and control, Phys. Rev. Lett. **108**, 080501 (2012).

[11] L. Viola, E. Knill, and S. Lloyd, Dynamical decoupling of open quantum systems, Phys. Rev. Lett. **82**, 2417 (1999).

[12] D. Lidar, Towards fault tolerant adiabatic quantum computation, Phys. Rev. Lett. **100**, 160506 (2008).

[13] L. Viola and E. Knill, Robust dynamical decoupling of quantum systems with bounded controls, Phys. Rev. Lett. **90**, 037901 (2002).

[14] A. D. Bookatz, P. Wocjan, and L. Viola, Hamiltonian quantum simulation with bounded-strength controls, New J. Phys. **16**, 045021 (2014).

[15] K. C. Young, M. Sarovar, and R. Blume-Kohout, Error suppression and error correction in adiabatic quantum computation: Techniques and challenges, Phys. Rev. X **3**, 041013 (2013).

[16] P. Facchi, D. A. Lidar, and S. Pascazio, Unification of dynamical decoupling and the quantum Zeno effect, Phys. Rev. A **69**, 032314 (2004).

[17] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, Quantum computation by adiabatic evolution, arXiv:quant-ph/0001106..

[18] J. Kempe, A. Kitaev, and O. Regev, The Complexity of the Local Hamiltonian Problem, SIAM J. Comput. **35**, 1070 (2006).

[19] Y. Cao and D. Nagaj, Perturbative gadgets without strong interactions, arXiv:1408.5881 [quant-ph].