Error-Suppression by Energy-Gap Protection for Quantum Computation in Open Systems

by

Xiang-Yu (Leo) Zhou

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

BACHELOR OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2014

Signature redacted

Author .....
Department of Physics
May 9, 2014

Signature redacted

Certified by.....
Edward Farhi
Professor of Physics
Director, Center for Theoretical Physics
Thesis Supervisor

Signature redacted

Accepted by .....
Nergis Mavalvala
Senior Thesis Coordinator, Department of Physics

Error-Suppression by Energy-Gap Protection for Quantum Computation in Open Systems

by

Xiang-Yu (Leo) Zhou

Submitted to the Department of Physics
on May 9, 2014, in partial fulfillment of the
requirements for the degree of
BACHELOR OF SCIENCE

Abstract

Adiabatic Quantum Computation, while attractive due to its "hands-off" approach and intrinsic tolerance of noise, has not been shown to be fully fault-tolerant in a satisfying manner. The protection of the evolution from noise and decoherence through the use of an energy penalty, recently proposed as a method to suppress error in adiabatic algorithms, is also appealing due to its passiveness. In this thesis, we first introduce the background on quantum computation, and discuss existing efforts towards fault-tolerant computation, specifically in the adiabatic model. Subsequently, we will prove a general result concerning the utility of energy-gap protection in generic (not necessarily adiabatic) quantum evolution in open system, and provides analytic bounds on the necessary energy penalty magnitude to achieve good protection. Evidence from numerical simulation is also given to demonstrate the practical usefulness of energy-gap protection for fault-tolerant quantum computation in open systems.

Thesis Supervisor: Edward Farhi
Title: Professor of Physics
Director, Center for Theoretical Physics

Acknowledgments

This thesis is formed from my collaboration with Adam Bookatz and Prof. Ed Farhi in the MIT Center for Theoretical Physics. I thank both of them for invaluable discussions that have always helped keep me *grounded* when my thinking starts going off course. In addition, I would like to thank the MIT Undergraduate Research Opportunities Program (UROP). Without the generous and consistent support of UROP over the last four years, my MIT research career would have been a lot less pleasant. Lastly, I would also like to thank those in the Physics Department office who have dependably provided me with resources whenever necessary.

Contents

List of Figures

Chapter 1

Introduction

Since computers become a pervasive and indispensable tool of our modern life, the thirst for greater computing power is now stronger than ever. As the transistors in ordinary classical computing chipsets get smaller and smaller per Moore's Law, it is conceivable that they will reach the quantum regime in the not-to-distant future. The idea of using quantum systems to do computation can be traced back to as early as Feynman[5], who discovered that quantum systems cannot be efficiently simulated by classical computers, due to the exponential growth in the classical description of a quantum state with the degrees of freedom. This opened up the door to the idea that computers with quantum components may possess computational power beyond classical machines, at least in certain problems e.g. simulating quantum systems.

Finding efficient algorithms that can solve problems with limited resources is an important problem in computer science and for real-world applications. In general, the criterion of efficiency that we are mostly interested in is the behavior of algorithms in the limit of large n, where n is the problem size. To aid our discussion, it's useful to introduce the big-O notation: given two functions $f(n)$ and $g(n)$, we say $f(n) = O(g(n))$ if and only if there are positive constants c, n_0 such that $|f(n)| \leq c|g(n)|$ for all $n > n_0$. Intuitively, it indicate that $g(n)$ is an upperbound on the asymptotic behavior of $f(n)$ for large n. Computer scientists consider a problem to be efficiently solvable if there is an algorithm that determines the solution in *polynomial time*, i.e. the amount of time necessary is $O(n^k)$ for some constant k. This polynomial definition of efficiency is a robust description of what can be computed deterministically in the real world, and the class of problem solvable in polynomial time is called the *complexity class* P, which stands for polynomial time. It has been a long-standing open problem in theoretical computer science to determine if $P \stackrel{?}{=} NP$, where NP (non-deterministic polynomial time) is the class of problems whose solutions can be verified in polynomial time with a classical computers. For quantum computer, the analogous complexity classes are BQP (bounded-error quantum polynomial time) and QMA (quantum Merlin Arthur).

While it was known for some time that quantum computers have capabilities beyond the Turing machines in classical computer science and provides speedup in certain situations due to so-called "quantum parallelism"[3], they largely remained obscure until Peter Shor's 1994 discovery of a quantum algorithm to factor integers in polynomial time[17]. The fact that quantum computers allow exponential speedup over any known classical algorithms for integer factorization, an important problem for cryptography and number theory, quickly made quantum information a subject fascinating to mathematicians, physicists, computer

scientists, and the information industry.

Despite their alleged usefulness, quantum computers face a serious practical obstacle: their fragile quantum states are highly sensitive to noise such as control errors and decoherence. Hence, the fault-tolerance of quantum computers is an important issue for its implementation that drew considerable attention[13]. Fortunately, the celebrated result of the *threshold theorem* has been proven for the standard circuit model of quantum computation that assures reliability of arbitrarily long quantum computation as long as the error per component is below a certain constant threshold. This means that the necessary accuracy of components do not need to scale as the problem size. If this isn't the case, then large computations would be physically intractable.

Farhi, Goldstone, Gutmann, and Sipser[4] recently found that quantum computation can also be done with adiabatic evolution, which in principle allows us to find the ground state in an arbitrary problem Hamiltonian. The adiabatic model of quantum computation seems appealing due to its simple description and natural suitability to solve optimization problems since we can encode the cost function of an optimization problem as the energy of the problem Hamiltonian. In addition, adiabatic quantum algorithms are enticing to experimentalists since it can be run passively without the need for measurements until the very end. This "hands-off" approach as well as the inherent resistance to noise[2] of adiabatic quantum computation made it one of the more promising paradigms of quantum computation for large-scale experimental implementation. Nevertheless, the quest for an analogous fault-tolerance threshold theorem in adiabatic quantum computation remains elusive. While some efforts have been made, such as the utilization of energy-gap protection[7] and dynamical decoupling[9] to suppress error, there has not been a satisfactory, rigorous, and generalizable solution.

In this thesis, we will review the basic ideas of quantum computation, specifically the standard circuit model, in Chapter 2. Then we will discuss the alternative model of adiabatic quantum computation and its motivation in Chapter 3. Finally, we will present in Chapter 4 a result more general than Ref. [7] that demonstrates and elucidates the practical utility of using energy-gap protection against decoherence and noise in not only adiabatic algorithms, but quantum evolution in open systems in general. Throughout the thesis, we will work in the units where $\hbar = k_B = 1$ for notation simplicity. The reader is assumed to know basic quantum mechanics, linear algebra, and elementary group theory.

Chapter 2

Standard Quantum Computation

2.1 Preliminary Quantum Information Theory

The fundamental building block of quantum computers is the *qubit*, which can be considered as a two-level quantum system. Physically, a qubit can be anything ranging from the spin state of a spin-$\frac{1}{2}$ particle, the polarization of a photon, the charge, flux, or phase of a Josephson junction, or the location of electron in quantum dot. Mathematically, the qubit is simply a vector in $\mathbb{C}^2$. Unlike a classical bit, which can be only be either 0 or 1, a qubit can be in a superposition such as $\alpha|0\rangle + \beta|1\rangle$.

Instead of qubits, it is also possible to perform quantum computation with three-level (qutrits) or even d-level (qudits) systems. While qudits may provide some speedup in certain computational tasks, the speedup is only of a constant factor since we can use multiple qubits to efficiently simulate a quoit, which is negligible in the grand scheme considering our polynomial definition of efficiency. Since qubits have the appeal of being the smallest quantum system, it is naturally considered as the fundamental block of quantum computers.

A quantum computer consists of a system of many qubits. An n-qubit system can be represented by a state vector in $\mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2 = \mathbb{C}^{2^n}$. As we can see, the dimension of the state space grows exponentially with the number of qubits, which is a reason that makes classical simulation of quantum systems difficult. The quantum state of a system can also be generalized to so-called *density matrix*, which are useful for describing statistical ensemble of quantum states, i.e. a quantum system whose state isn't completely known. Suppose a quantum system is known to be in state $|\psi_i\rangle$ with probability p_i, where $i = 1, 2, \ldots$ is an index, the density matrix of the system is defined as

$$\rho \equiv \sum_i p_i |\psi_i\rangle \langle\psi_i| \tag{2.1}$$

which in general are positive hermitian operators with trace 1. If $\rho = |\psi\rangle\langle\psi|$ for some $|\psi\rangle$, then the system is said to be in the *pure state* of $|\psi\rangle$. Otherwise, ρ is a *mixed state*. This density matrix formalism allows us to write

$$i\frac{d|\psi\rangle}{dt} = H|\psi\rangle \quad \rightarrow \quad \frac{d\rho}{dt} = -i[H, \rho] \tag{2.2}$$

$$|\psi(t)\rangle = U|\psi(0)\rangle \quad \rightarrow \quad \rho(t) = U(t)\rho(0)U^\dagger(t) \tag{2.3}$$

$$\langle A\rangle = \langle\psi|A|\psi\rangle \quad \rightarrow \quad \langle A\rangle = \mathrm{Tr}(\rho A) = \sum_i p_i \langle\psi_i|A|\psi_i\rangle \tag{2.4}$$

Since density matrix is a random ensemble of pure states, we can calculate its entropy, for which the von Niemann definition is

$$S(\rho) = -\mathrm{Tr}(\rho \log_2 \rho) = -\sum_{\lambda_i \neq 0} \lambda_i \log_2 \lambda_i \tag{2.5}$$

where λ_i are the eigenvalues of ρ. Observe that $S(\rho) = 0$ if and only if ρ is a pure state. A special case of mixed state is the *maximally mixed state*, in which ρ is a diagonal matrix with entries $1/d$, where d is the dimension of the Hilbert space. This has maximum entropy, $\log_2 d$. Note in this case, any pure state can be observed equal probability $1/d$.

Quantum computations can be considered as arbitrary unitary operations on qubits. The reason why the operation needs to be unitary is the fact that Schrödinger equation, which governs quantum evolution, is norm-preserving. In other words, a quantum state $|\psi\rangle$ should in principle always have norm $\langle\psi|\psi\rangle = 1$, irrespective of what is happening to it. Since any unitary operation U satisfy $U^\dagger U = I$, this is the necessary and sufficient condition for any operation on a quantum state $|\psi\rangle \to |\psi'\rangle = U|\psi\rangle$, because $\langle\psi'|\psi'\rangle = \langle\psi|U^\dagger U|\psi\rangle = \langle\psi|\psi\rangle$. The result of the computation is obtained by measuring certain observables on the final quantum state. Notably, any global phase on the state $|\psi\rangle$ is irrelevant for quantum computation, since even if we end up with $|\psi'\rangle = e^{i\theta}|\psi\rangle$, the value of an observable A is unchanged:

$$\langle\psi'|A|\psi'\rangle = \langle\psi|e^{-i\theta}Ae^{i\theta}|\psi\rangle = \langle\psi|A|\psi\rangle \tag{2.6}$$

The density matrix allows us to generalize quantum operations beyond the unitary operators to so-called *quantum channels*. This is necessary when we're examining a *reduced* density matrix, i.e. a density matrix that describes the state of only one part of an overall system. An example would be the density matrix of the system state when the system is coupled to a bath, obtained by partial tracing out the bath: $\rho_S = \mathrm{Tr}_B(\rho)$. It's common to use the *operator-sum representation* of a quantum channel, which is

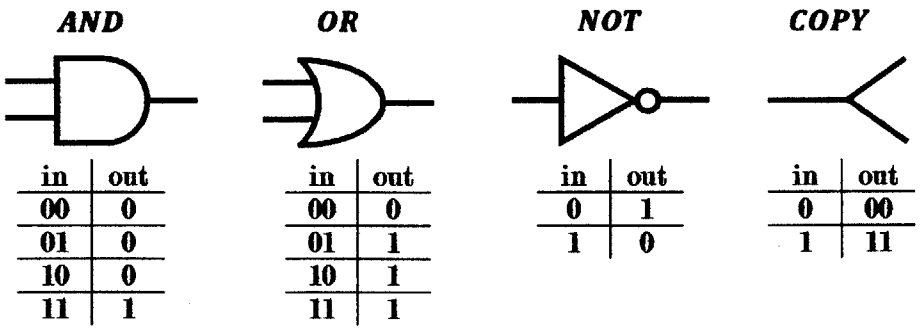
$$\rho \to \rho' = A(\rho) = \sum_k A_k \rho A_k^\dagger \tag{2.7}$$

where $\sum_k A_k^\dagger A_k = I$. The last condition ensures that the operations is trace-preserving, so that $A(\rho)$ is still a density matrix that satisfies $\mathrm{Tr}(A(\rho)) = 1$. The set of operators $\{A_k\}$ are also called Kraus operators. In the case where $|\{A_k\}| = 1$, we simply have a unitary operation $\rho \to A_1 \rho A_1^\dagger$ since $A_1^\dagger A_1 = I$.

2.2 The Circuit Model

In a classical digital computers, computations are performed via arbitrarily complex logical operations on bits. It turns out that any logical operations on any number of bits can be broken down to a sequence of logical operations drawn from a small set of elementary logic gates that operates on a few bits. Such a set of gates is *universal*, as they are capable of *universal classical computation*. An example of a universal set is $\{AND, OR, NOT, COPY\}$, shown in Fig. 2-1. Any classical computation can be envisioned as a circuit of these logic gates with a set of bits as input. However, this is not a minimal set, since $x\ OR\ y$ can be simulated with $NOT((NOT\ x) AND (NOT\ y))$, due to de Morgan's law. In fact, the set $\{NAND, COPY\}$, where $x\ NAND\ y = NOT(x\ AND\ y)$, is also universal.

Similarly, a quantum computation is essentially performing an arbitrary unitary opera-



in	out
00	0
01	0
10	0
11	1

in	out
00	0
01	1
10	1
11	1

in	out
0	1
1	0

in	out
0	00
1	11

Figure 2-1: A set of logic gates capable of universal classical computation, along with a table of their input/output.

tion on qubits as mentioned before. Analogous to the classical circuit of logic gates, quantum computations can be seen as feeding a set of qubits through a circuit of quantum logic gates, which individually can be considered as a unitary operation on a small number of qubits. For example, the Pauli matrices are quantum gates on a single qubit:

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \qquad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{2.8}$$

Two other examples of common quantum gates, the single-qubit Hadamard gate and the two-qubit controlled-NOT ($CNOT$) gate, are shown in Fig. 2-2. A set of quantum gates is *universal*, i.e. capable of *universal quantum computation*, if any unitary operations can be approximated to arbitrary accuracy by applying gates drawn from the set. For instance, the set of single-qubit gates and the $CNOT$ gate is universal.

While there are infinitely many possible unitary operations, it turns out that only a finite number of them is necessary for universality. A well-known universal set is the Hadamard, $\pi/8$, and $CNOT$ gates[12], where the $\pi/8$ gate is

$$T = e^{-i\frac{\pi}{8}Z} = \begin{bmatrix} e^{-i\pi/8} & 0 \\ 0 & e^{i\pi/8} \end{bmatrix} \tag{2.9}$$

It's not hard to see that we can easily obtain the Pauli gates (up to an irrelevant global phase) from these three gates, since $Z = iT^4$, $X = HZH$, and $Y = iXZ$. We can then

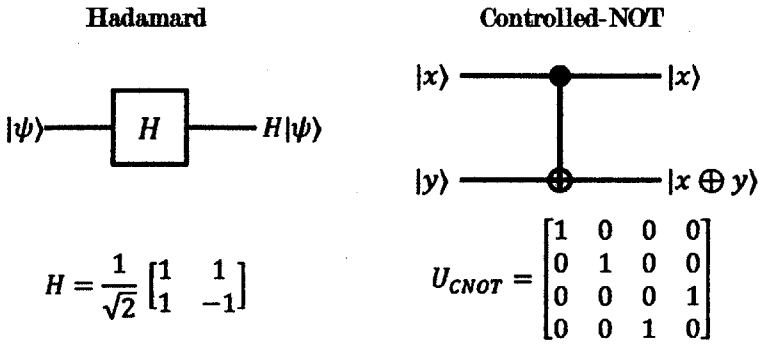


Figure 2-2: Examples of single-qubit and two-qubit quantum logic gates, along with their corresponding unitary operations. $\oplus$ is addition modulo 2.

use these gates to approximate arbitrary unitary operations. Although using approximate versions of gates in a circuit may seem problematic at first, note that the errors induced from a sequence of approximate gates is at most the sum of the errors in the individual gates. We can easily show this result: suppose U' and V' are to approximate gates U and V, we can use the metric of error $E(U') = \|U' - U\|$, where $\|A\|$ is the matrix norm of A defined as

$$\|A\| = \max_{|\psi\rangle:\langle\psi|\psi\rangle=1} \|A|\psi\rangle\| \tag{2.10}$$

Then, using the triangle inequality and unitarity, we can easily show

$$\begin{aligned} \|U'V' - UV\| &= \|U'V' - UV' + UV' - UV\| \\ &\leq \|U'V' - UV'\| + \|UV' - UV\| = \|U' - U\| + \|V' - V\| \\ E(U'V') &\leq E(U') + E(V') \end{aligned} \tag{2.11}$$

The general result for a long sequence of gates can easily be proven using induction, which demonstrate that the errors add at most linearly. Thus, a finite number of single-qubit gates can generate a subgroup that is dense in the set of all single-qubit unitary operations $SU(2)$. Here, a set S is said to be *dense* in $SU(2)$ if for any elements U of $SU(2)$ and a small number $\epsilon > 0$, we can find an element in $V \in S$ such that their distance $D(U, V) \leq \epsilon$, for some distance measure $D(x, y)$ such as $\|x - y\|$. Despite the fact that a finite set can generate a dense subgroup in the unitaries, we may still worry that an enormous number of gates can be necessary to approximate a general unitary operation. Fortunately, we are saved by the Solovay-Kitaev Theorem, which is rephrased from Ref. [12] below

Theorem 2.1 (Solovay-Kitaev theorem). *Let $\mathcal{G}$ be a finite set of elements in $SU(2)$ containing its own inverses. Suppose the set generates a dense subgroup in $SU(2)$. Let $\epsilon > 0$ be a precision parameter and $U \in SU(2)$ be any desired unitary operation, then there is a product V of $O(\log^c(1/\epsilon))$ elements from G, where $c \approx 4$, such that $\|U - V\| \leq \epsilon$.*

Hence, in order to approximate a circuit of m arbitrary single-qubit and $CNOT$ gates to a final accuracy of ϵ, we only need $O(m \log^c(m/\epsilon))$ gates from the finite set due to the Solovay-Kitaev theorem and the generalization of Eq. (2.11). This fits in our polynomial definition of efficiency, since the approximate circuit is only polylogarithmically larger.

The efficiency of a quantum algorithm can be measured as the number of gates necessary

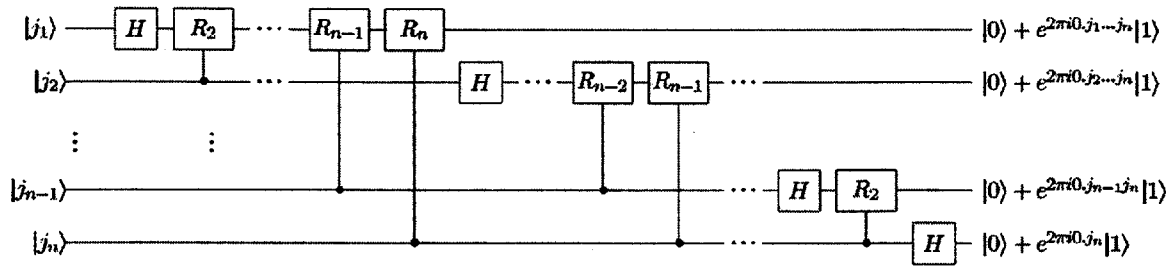


Figure 2-3: Quantum Fourier transform circuit on n qubits, adopted from Ref. [12]. Note the normalization factor of $1/\sqrt{2}$ on the final qubits, as well as the swap gates at the end of the circuit that reverse the order of the qubits, are omitted.

in the quantum circuit. For instance, quantum Fourier transform is the following operation

$$\sum_{j=0}^{N-1} x_j |j\rangle \rightarrow \sum_{k=0}^{N-1} y_k |k\rangle \quad \text{where} \quad y_k \equiv \frac{1}{N} \sum_{j=0}^{N-1} x_j e^{2\pi ijk/N} \tag{2.12}$$

where we use the shorthand $|j\rangle = |j_1 j_2 \cdots j_n\rangle$, so that $j_1 j_2 \cdots j_n$ is the binary representation of the integer j. For example, $|5\rangle \equiv |101\rangle = |1\rangle \otimes |0\rangle \otimes |1\rangle$. Without loss of generality, we assume $N = 2^n$, so that N can be represented as an n-bit number. Note that the vector of y_k's can be considered as the output of the classical discrete Fourier transform. The circuit for quantum Fourier transform is shown in Fig. 2-3, where R_k is the gate

$$R_k = \begin{bmatrix} 1 & 0 \\ 0 & e^{2\pi i/2^k} \end{bmatrix} \tag{2.13}$$

Each R_k is in fact part of a bigger, two-qubit controlled-R_k gate. In general, a controlled-U gate is defined as the operation $|c\rangle \otimes |t\rangle \rightarrow |c\rangle \otimes U^c |t\rangle \ \forall c = 0, 1$. Now, note the overall circuit requires $n + (n-1) + \cdots + 1 = n(n+1)/2 = O(n^2)$ gates. We can compare this with classical Fourier transform, where the fastest algorithm known (Fast Fourier Transform) requires $O(nN) = O(n2^n)$ gates. Therefore, by applying the quantum Fourier transform to problems such as integer factorization (Shor's algorithm) and estimating eigenvalues of unitary operator (phase estimation algorithm), we can achieve an exponential speedup by quantum computers over any known classical algorithm.

The clarity and success of using circuits of quantum gates to describe universal quantum computation is the reason why the circuit model is considered the standard method of quantum computation.

2.3 Fault-Tolerance

A huge problem with quantum computation is that the qubits are in delicate quantum states which are susceptible to environmental noise and decoherence. Since noise is practically unpreventable, it is important to have *fault-tolerant* quantum computation in the sense that a small error in one component may not cause the entire computation to fail. Hence, we have an important question for quantum computer scientists: is it possible to efficiently and reliably perform arbitrarily large quantum computations even when there is noise? Well, it turns out that this is possible as long as the noise in the individual quantum state is below some constant threshold, a result known as the *threshold theorem*.

2.3.1 Quantum Error-Correcting Codes

In order to protect against noise, a common and natural practice in computer science is to encode our messages by adding redundant information. This way, even if part of the encoded information is corrupted, the redundancy should allow us to decode the information and recover the message. For example, a naïve approach would be encoding the message by making many copies of the same message. If each message is corrupted with probability p where $p \ll 1$, then the probability of half or more of the copies are corrupted is $O(p^{n/2})$ where n is the number of copies; by perform majority voting with the n copies received, we can determine the correct message with only error probability $O(p^{n/2})$ which improves over the original p for $n \geq 3$.

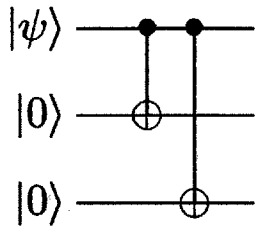


Figure 2-4: The encoding circuit for three-qubit flip code that takes one-qubit state $|\psi\rangle$ and, after applying two $CNOT$ gates, outputs the logical state $|\psi_L\rangle$ on three qubits. Adopted from Ref. [12].

A simple example of quantum error-correcting codes is the three-qubit bit flip code, which protect against errors in the form of random X operations occurring with probability p on any qubit, where X is the usual Pauli σ_x operator. It's called bit flip because $X|0\rangle = |1\rangle$, and $X|1\rangle = |0\rangle$. For instance, when there is a single qubit whose density matrix is ρ, the noise model can be presented as the following quantum channel:

$$\rho \to \rho' = \mathcal{E}(\rho) = (1-p)\rho + pX\rho X \tag{2.14}$$

In this code, we encode an arbitrary single-qubit state in the following way

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \to |\psi_L\rangle = \alpha|0_L\rangle + \beta|1_L\rangle = \alpha|000\rangle + \beta|111\rangle \tag{2.15}$$

where the subscripts L in $|0_L\rangle \equiv |000\rangle$ and $|1_L\rangle \equiv |111\rangle$ indicate that they are the ***logical*** $|0\rangle$ and $|1\rangle$ state. Each of the three qubits are called the ***physical*** qubits. The circuit that performs the encoding is shown in Fig. 2-4. Now, the possible states after sending the qubit through the noisy bit flip channel are

$$|\psi_0\rangle = \alpha|000\rangle + \beta|111\rangle \quad \text{(no error)} \tag{2.16}$$
$$|\psi_1\rangle = \alpha|100\rangle + \beta|011\rangle \quad \text{(bit flip on 1st qubit)} \tag{2.17}$$
$$|\psi_2\rangle = \alpha|010\rangle + \beta|101\rangle \quad \text{(bit flip on 2st qubit)} \tag{2.18}$$
$$|\psi_3\rangle = \alpha|001\rangle + \beta|110\rangle \quad \text{(bit flip on 3rd qubit)} \tag{2.19}$$

In general, quantum error-correction involves two stages:

1. *Syndrome diagnosis* In this stage, we perform measurements of some syndrome operators to determine what error, if any, occurred on the quantum state. For the bit flip code, the two syndrome operators necessary are $g_1 = Z \otimes Z \otimes I$ and $g_2 = I \otimes Z \otimes Z$. Note the syndrome measurements do not cause any change to the state since all possible states in Eq. (2.16)–(2.19) are simultaneous eigenstates of the syndrome operators. In general, syndrome measurements do not give any information about the logical state, i.e. amplitudes α and β, since otherwise it would destroy the superposition that we meant to preserve. The measurement outcomes in each of the four scenarios are summarized in the table below.

	$\|\psi_0\rangle$	$\|\psi_1\rangle$	$\|\psi_2\rangle$	$\|\psi_3\rangle$
g_1	$+1$	-1	-1	$+1$
g_2	$+1$	$+1$	-1	-1

2. *Recovery* In this stage, we correct the error by applying an operation on the system based on the measured value of the error syndrome. For the bit flip code, we will do nothing if $(g_1, g_2) = (1, 1)$, apply X on the first qubit if $(g_1, g_2) = (-1, 1)$, apply X on the second qubit if $(g_1, g_2) = (-1, -1)$, and apply X on the third qubit if $(g_1, g_2) = (1, -1)$. It's not hard to see that this restores the state to the initial state before any bit flip error occurred.

Thus, this three-qubit bit flip code is able to prevent arbitrary X error, as long as it occurs on one or fewer of the three physical qubits. This scheme fails, i.e. an error remain uncorrected, when two or more qubit experience the X error. This occurs with probability $3p^2(1-p)+p^3 = 3p^2 - 2p^3 = O(p^2)$, which is less than p provided $p < 1/2$.

Stabilizer codes are an important subset of quantum error-correcting codes, which are convenient group-theoretical representations of most quantum error-correcting codes. They will also prove useful for our discussion on energy-gap protection later. In the stabilizer formalism, we consider the Pauli group G_n on n-qubits, which are simply the set of all tensor products of n Pauli operators or the identity with coefficient ± 1 or $\pm i$. For instance, the single-qubit Pauli group is the set

$$G_1 \equiv \{\pm I, \pm iI, \pm X, \pm iX, \pm Y, \pm iY, \pm Z, \pm iZ\} \tag{2.20}$$

which forms a group under matrix multiplication. Note for any element $E \in G_n$, $E^{-1} = E^\dagger$.

An $[n, k]$ ***stabilizer code*** is the vector space $C \subset \mathbb{C}^{2^n}$ stabilized by a subgroup $S \subset G_n$, such that $-I \notin S$ and S has $n-k$ independent and commuting generators. The vector space C is also called the codespace, and the fact that it's ***stabilized*** by S means

$$C = \{|\psi\rangle : x\,|\psi\rangle = |\psi\rangle\ \forall x \in S\} \tag{2.21}$$

It is then obvious that the condition $-I \notin S$ is necessary or else C would the trivial vector space. Given that S has $n-k$ generators, the codespace dimension is 2^k, corresponding to k logical qubits. We can also define the orthogonal space $C^\perp$ as $C^\perp = \{|\phi\rangle : \langle\phi|\psi\rangle = 0\ \forall\, |\psi\rangle \in C\}$.

Denoting the generators as g_i's, we can write $S = \langle g_1, \ldots, g_{n-k}\rangle$. Then ***independence*** of generators means that $\nexists$ index i such that we can take out g_i from the set of generators and still have $S = \langle g_1, \ldots, g_{i-1}, g_{i+1}, \ldots, g_{n-k}\rangle$. The fact generators are ***commuting*** means $[g_i, g_j] = 0\ \forall i, j$.

We can construct an $[n, k]$ stabilizer code with the following procedure. First, find $\bar{Z}_1, \ldots, \bar{Z}_k \in G_n$ so that the set $\{g_1, \ldots, g_{n-k}, \bar{Z}_1, \ldots, \bar{Z}_k\}$ is independent and commuting. We can consider $\bar{Z}_j$ as the logical Pauli Z operator on the logical j^{th} qubit. Then, the logical computational basis state $|x_1 \cdots x_k\rangle_L$, where $x_j \in \{0, 1\}$, can be naturally defined as the state stabilized by $\langle g_1, \ldots, g_{n-k}, (-1)^{x_1}\bar{Z}_1, \ldots, (-1)^{x_k}\bar{Z}_k\rangle$.

For an error $E \in G_n$, note that for any element $g \in S$, E either anti-commutes ($Eg = -gE$) or commutes ($Eg = gE$) with g, because they are both tensor products of Pauli operators. In the anti-commuting case, the error can be detected by measuring the generators of S and perhaps be corrected afterwards. In the commuting case, we won't have to worry if $E \in S$ since the $E\,|\psi\rangle = |\psi\rangle\ \forall\, |\psi\rangle \in C$; if $E \notin S$ and commutes with all elements of S, then E may affect the computation and cannot be detected. With that in mind, it is natural to

look at the centralizer and normalizer group of S, which are

$$\text{centralizer: } Z(S) = \{E \in G_n : EgE^\dagger = g \ \forall g \in S\} \tag{2.22}$$

$$\text{normalizer: } N(S) = \{E \in G_n : EgE^\dagger \in S \ \forall g \in S\} \tag{2.23}$$

It turns out that $Z(S) = N(S)$ for any subgroup S of G_n that doesn't contain $-I$, since elements of G_n either commutes or anti-commutes with one another. Recall that the quantum error-correction condition stipulates that a set of errors $\{E_j\}$ is *correctable* if and only if

$$\textit{Error Correctability:} \qquad PE_j^\dagger E_k P = \alpha_{jk} P \tag{2.24}$$

where P is the projector onto the codespace C and $[\alpha_{jk}]$ is an Hermitian matrix. A simple theorem proven by Nielsen and Chuang[12] gives the condition on a set of errors that can be corrected by a stabilizer code, which we now rephrase below

Theorem 2.2 (Error-correction condition for stabilizer codes). *Let S be the stabilizer for a stabilizer code $C(S)$, and $\{E_j\} \subset G_n$ be a set of error operators such that $E_j^\dagger E_k \notin N(S) - S \ \forall j,k$. Then $\{E_j\}$ is a correctable set of errors for the code $C(S)$.*

The set $N(S) - S$ contains nontrivial operations on the codespace, which include both undetectable errors as well as operations on the logical qubit.

As an example, let us consider the stabilizer code version of the three-qubit bit flip code above, which is a [3,1] stabilizer code. The stabilizer group is generated by $g_1 = Z \otimes Z \otimes I$ and $g_2 = I \otimes Z \otimes Z$ as defined earlier. We can see that a Pauli X operator on any of the three physical qubit anti-commutes with either or both of g_1 and g_2, indicating that they are correctable. We can also define the logical Pauli operators as

$$\bar{X} = X \otimes X \otimes X, \quad \bar{Y} = Y \otimes X \otimes X, \quad \bar{Z} = Z \otimes I \otimes I \tag{2.25}$$

It isn't hard to check that

$$\bar{X}\,|0_L\rangle = |1_L\rangle, \quad \bar{Y}\,|0_L\rangle = i\,|1_L\rangle, \quad \bar{Z}\,|0_L\rangle = |0_L\rangle \tag{2.26}$$

$$\bar{X}\,|1_L\rangle = |0_L\rangle, \quad \bar{Y}\,|1_L\rangle = -i\,|0_L\rangle, \quad \bar{Z}\,|1_L\rangle = -\,|1_L\rangle \tag{2.27}$$

It is also easy to verify that $\bar{X}, \bar{Y}, \bar{Z}$ each commutes with both g_1 and g_2, and none of them can be generated by products of g_1, g_2. Unsurprisingly, they are in $N(S) - S$. In our future discussion, we will follow the convention where we use a bar over logical operator $\bar{A}$ to denote the encoded version of operator A.

2.3.2 Threshold Theorem for the Circuit Model

It turns out that we can apply error-correcting codes to achieve full fault-tolerance in the standard circuit model of quantum computation, which allows arbitrarily good quantum computation to be accomplished with faulty quantum gates as long as the error probability per gate is below a certain constant threshold. The discovery of the threshold theorem is an important achievement of quantum error-correction, since the threshold doesn't need to decrease with the problem size. This result is also nontrivial due to the fundamental issues associated with encoded quantum states and gates, because an error in any part of the encoded circuit may propagate due to long-range entanglement, and the procedure of error-correction itself may introduce error in the system.

The idea behind the threshold theorem is essentially performing quantum computation with encoded quantum states and gates and carefully designing the circuit in a way such that failures may not propagate beyond an encoded block of qubits. More specifically, we want a construction of the fault-tolerance circuit that allows us to reduce the net error probability from p in the original computation to $O(p^2)$ in the computation on the encoded states. The proof ultimately relies on concatenated codes to reduce the effective error rate to an arbitrarily low level. The main result and a sketch of proof per Nielsen and Chuang[12] is provided below.

Theorem 2.3 (Threshold theorem for quantum computation). *A quantum circuit containing $g(n)$ gates may be simulated with probability of error at most ϵ using*

$$O(\text{poly}(\log(g(n)/\epsilon))g(n)) \tag{2.28}$$

gates on hardware whose component fail with probability at most p, provided p is below some constant threshold, i.e. $p < p_{th}$, and given reasonable assumption about the noise in the underlying hardware.

Sketch of Proof. Suppose any single component (qubit) of a quantum computer has an error (fails) with probability p. The proof of the threshold theorem relies on a procedure of implementing fault-tolerant gates, so that the probability of introducing two or more error in the encoded state is $O(p^2)$. Then, it uses concatenation to construct a hierarchy of quantum circuits $C_0, C_1, C_2, \ldots$ where C_0 is the original circuit. For instance, at the first level, we replace all gates in C_0 with an encoded gate in C_1, implemented with a fault-tolerant procedure along with error-correction. At the second level, any gate in C_1 is replaced in C_2 by an encoded gate and error-correction through a fault-tolerant procedure. We can then repeat this procedure ad infinitum.

The probability of the entire code failing at each level is

Level	C_0	C_1	C_2	$\cdots$	C_k
Failure Probability	p	cp^2	$c(cp^2)^2$	$\cdots$	$(cp)^{2^k}/c$

where c is just some constant associated with the method of constructing concatenated code with a fault-tolerant procedure. Thus, if we want to achieve a final accuracy of ϵ, then we need each gate to be accurate to $\epsilon/g(n)$, so the necessary number of concatenations is

$$\frac{(cp)^{2^k}}{c} \leq \frac{\epsilon}{g(n)} \tag{2.29}$$

provided that $p < p_{\text{th}} = 1/c$. Suppose the size of the circuit grows exponentially in k, so the number of gates necessary in achieving an accuracy of ϵ is

$$\#\text{ gates} = d^k g(n) = \Big(\frac{\log(g(n)/c\epsilon)}{\log(1/cp)}\Big)^{\log d} g(n) = O\Big(\text{poly}\Big(\log\frac{g(n)}{\epsilon}\Big)g(n)\Big) \tag{2.30}$$

where d is just some constant associated with the concatenation procedure, and poly$(*)$ indicates a polynomial function of a fixed degree. Hence, we've shown that the fault-tolerant circuit that simulates the original circuit is only polylogarithmically larger, and thus completed the proof. □

For a 7-qubit Steane code, the threshold is roughly $p_{\text{th}} = 1/c \approx 10^{-4}$ from counting arguments. With more sophisticated calculation, typical thresholds are in the range of 10^{-5}–10^{-6}.

There are also some limitations to the threshold theorem[12]. Firstly, our proof of the threshold theorem requires many fault-tolerant operations be performed in parallel. Otherwise, sequential operations would take too long while errors accumulate in the circuit, causing the computation to fail. Secondly, the cost of classical computation and communication necessary for processing the syndrome of error-correction isn't taken into account. If these can't be done faster than the timescale over which errors occur, then error-correction would have no utility. But this is more of a fundamental problem for error-correcting scheme used than it is for the threshold theorem. Lastly, many fault-tolerant construction of gates and error-correcting procedures requires a significant amount of ancillary qubits, which impose another condition on implementation of fault-tolerant circuit-model quantum computation.

Chapter 3

Adiabatic Quantum Computation

3.1 The Adiabatic Model

The foundation of adiabatic quantum computation (AQC) is the well-known adiabatic theorem[10]. Roughly speaking, the theorem states that a quantum system evolving under a time-dependent Hamiltonian will remain essentially in its instantaneous eigenstate if there is a energy gap between its instantaneous eigenvalue and the rest of the spectrum, and if the evolution occurs slowly enough (i.e. adiabatically). Note that although the theorem can be proven for any instantaneous eigenstate, people are generally most interested in the ground state.

There are many mathematical statements of the adiabatic theorems. A rigorous version due to Jordan[6] is reproduced below

Theorem 3.1 (The Adiabatic Theorem). *Suppose $H(s)$ is finite-dimensional Hamiltonian that is twice-differentiable on $s \in [0,1]$, with an instantaneous nondegenerate ground state $|\phi_0(s)\rangle$. Also let $g(s) \equiv E_1(s) - E_0(s)$ be the instantaneous energy gap separating the ground state and the first excited state. If $|\psi(t)\rangle$ is the state obtained by evolving $|\phi_0(0)\rangle$ starting at $t = 0$ with $H(t/T)$ according to the Schrödinger equation, then with some choice of overall phase for $|\phi_0(t)\rangle$, we have*

$$\| \, |\psi(T)\rangle - |\phi_0(T)\rangle \, \| \leq \frac{1}{T}\Big[\frac{\|dH/ds\|_{s=0}}{g^2(0)} + \frac{\|dH/ds\|_{s=1}}{g^2(1)} + \int_0^1 ds\Big(\frac{5}{g^3(s)}\Big\|\frac{dH}{ds}\Big\|^2 + \frac{1}{g^2(s)}\Big\|\frac{d^2H}{ds^2}\Big\|\Big)\Big] \tag{3.1}$$

The above theorem essentially says that if we evolve the instantaneous nondegenerate ground state of a time-dependent Hamiltonian sufficiently slowly (i.e. adiabatically) over some period T, then the quantum state will remain in the corresponding instantaneous nondegenerate ground state with high probability, provided there is always a nonzero energy gap between it and nearby energy levels. Note that in the limit $T \to \infty$ of true adiabatic process, we reach the final ground state with probability one. Clearly, we can't wait forever in practice. For $H(s)$ with minimum gap $\gamma \equiv \min_s g(s)$ and maximum speed $\eta \equiv \max_s \|dH/ds\|$, Eq. (3.1) suggests that the minimum length T of the evolution period to ensure significant overlap with the final desired state satisfies is $T = O(\eta^2/\gamma^3)$ due to the first term in the integrand. However, there is a version of the adiabatic theorem due to Reichardt[14] that says if $H(s)$ is $(k+1)$-differentiable with $k \geq 1$, then the necessary time is $T = O(\eta^{1+\delta}/\gamma^{2+\delta})$, where $\delta = 1/k \leq 1$ depends on the differentiability of $H(s)$ and the definition of η is slightly

adjusted. In general, for a smooth adiabatic Hamiltonian, we expect $T \approx O(\eta/\gamma^2)$, but never beyond $T = O(\eta^2/\gamma^3)$.

Notice that the problem of finding the ground state of an arbitrary Hamiltonian is generally hard. This is true even if we restrict to k-local Hamiltonians, which are Hamiltonians that can be written as $H = \sum_i H_i$ where H_i acts on at most k qubits. In fact, Kempe *et al.*[8] proved that the problem of merely finding the ground state eigenvalue of a k-local Hamiltonian is QMA-complete for all $k \geq 2$. This means that if one can solve this problem efficiently, then all problems in QMA are efficiently solvable, where QMA is the quantum analog of the classical computation complexity class NP that contains problems considered very difficult. Fortunately, the adiabatic theorem affords us the possibility of tackling this difficult problem: we can start with a Hamiltonian whose ground state we know and can prepare, and apply adiabatic evolution towards a final Hamiltonian whose ground state we hope to know, we can in principle find the ground state of that final Hamiltonian. This adiabatic model of quantum computation, which was first studied by Farhi *et al.*[4], has the following standard procedure:

1. Encode the solution to a problem as the nondegenerate ground state of a local Hamiltonian H_f.

2. At $t = 0$, start with an initial local Hamiltonian H_0 such that the nondegenerate ground state is an easy-to-prepare state (e.g. $|0\cdots 0\rangle$). Prepare the quantum state $|\psi(0)\rangle$ in the ground state.

3. Slowly change the Hamiltonian over time from $t = 0$ to $t = T$, so that $H(0) = H_0$ and $H(T) = H_f$, assuming that there is a nonzero energy gap between the ground state and the first-excited state(s) at all time $t \in [0, T]$.

4. At $t = T$, measure the quantum state $|\psi(T)\rangle$ and obtain the solution with high probability due to the adiabatic theorem.

The requirement that our Hamiltonians be local (i.e. consisting of terms acting on only a few qubits) is reasonable, since the Hamiltonians corresponds to the energy of physical systems, which only appear in nature as few-body interactions. The locality constraint in our theoretical discussion allows experimentalists to physically set up the interactions between qubits in any feasible implementation of the adiabatic model.

The efficiency of the adiabatic model is measured by the necessary runtime T, instead of the number of gates as in the circuit model. A typical adiabatic algorithm utilizes constant-rate evolution, which is

$$H(s) = (1 - s)H_0 + sH_f \tag{3.2}$$

Note that it is possible to vary the speed of the adiabatic evolution in real time so as to increase efficiency, especially if one knows where the minimum gap is likely to occur. For instance, Roland and Cerf[15] found that the adiabatic Grover's algorithm for searching in an unsorted database of size n would've taken $O(n)$ time, which isn't an improvement over classical algorithms, but can be improved to a runtime of $O(\sqrt{n})$ if the rate is varied according to the instantaneous energy gap, which can be computed a priori due to the nature of the problem. Nonetheless, this approach cannot be generally taken due to the fact that the energy eigenvalue landscape for a generic Hamiltonian is unknown.

Despite a very different computational approach, the adiabatic model has been shown to polynomially equivalent to standard circuit-model quantum computation by Aharonov

et al.[1]. In their paper, Aharonov *et al.* demonstrated that any circuits of quantum gates can be efficiently simulated by 5-local Hamiltonian adiabatic algorithms, with the minimum energy gap $\gamma = O(1/m^2)$ where m is the number of gates, giving a runtime of $T = O(m^5)$. More efficient proofs can be found, such as the 3-local construction with $\gamma = O(1/m^2)$ discussed in Ref. [6, 11]. This suggests not only that the adiabatic model is capable of universal quantum computation, but also that it offers no computational advantage over the circuit-model.

Even though the adiabatic model is not more powerful in general than the circuit model, there are several reasons why the adiabatic model can be more appealing. One reason is its "hands-off" approach to quantum computation. Unlike in the circuit model, where we need to continuously monitor the quantum state and perform measurements whenever necessary, the adiabatic model only requires us to prepare the initial state and slowly evolve $H(t)$, and to only measure when the computation is complete. Since any interactions with the delicate state, especially measurements, can introduce error in the system, it seems ideal that adiabatic model only requires one measurement at the end. Hence, adiabatic algorithms may be easier to experimentally implement. In any case, a physically distinct version of quantum computation is always of interest to experimentalists who are looking for innovating schemes to harvest the power of quantum information processing. Additionally, the adiabatic model is a very natural construction of quantum computation in some sense, especially since the k-local Hamiltonian problem is QMA-complete[8]. Nevertheless, an oft-cited benefit of the adiabatic model relevant to this thesis is its intrinsic fault-tolerant properties, which we will discuss in the ensuing section.

3.2 Towards Fault-Tolerance in Adiabatic Quantum Computation

Even though adiabatic quantum computation is shown to be equivalent in computation power to standard circuit-model computation, which has been shown to be fault-tolerant, no analogous threshold theorem has been found for the adiabatic paradigm that demonstrates its fault-tolerance. While it may be tempting to translate the techniques for circuit-model fault-tolerance such as active error-correction to the adiabatic model, we're largely interested in fault-tolerance techniques that preserve the passive, hands-off approach that made adiabatic algorithms attractive in the first place. In the section, we will first demonstrate the intrinsic tolerance to noise and decoherence of adiabatic quantum computation, and then review recent efforts towards fault-tolerance in the adiabatic model.

3.2.1 Inherent Robustness of the Adiabatic Paradigm

It has been suggested that adiabatic quantum computation may have some inherent fault-tolerance, which makes it even more interesting and appealing. There are multiple reasons to believe so, which we will now enumerate.

Firstly, we have reason to believe that any control errors that occur during the adiabatic evolution are likely irrelevant as long as the initial and final Hamiltonian are correct. This is not true for quantum computation in the circuit model, where errors in intermediate steps have significant implications for the final result even if the last few operations are performed correctly. Since the answer (final state) in the adiabatic model only depends on the final Hamiltonian, any small control errors during the adiabatic evolution that causes

$H(t)$ to deviate from its intended path is irrelevant as long as it didn't lower the energy gap significantly, which seems unlikely in general. Childs *et al.*[2] provided numerical evidence that small unitary control errors do not strongly suppress the success probability of the computation, and may even increase it in certain cases. In particular, they found that low-frequency, off-resonance unitary control errors that vanish at the endpoints of the computational period may even improve success probability at large amplitude, since they can increase the minimum gap.

In addition, the adiabatic model is a continuous-time, smooth evolution of a quantum system in contrast to the discrete-time, pulse-like operations in the circuit model. While both may be useful depending on the preferences of and resources available to the experimenter, it can be argued that the pulse-sequences in the typical circuit-model implementation have a large bandwidth and can be more likely to introduce noise to the system. On the other hand, the slow and smooth adiabatic evolution may allow us to filter out most noises except the low frequency ones[6]. The conjecture that the adiabatic model can be resistant to high frequency control noise is corroborated by the numerical evidence of Childs *et al.*[2]. Moreover, Roland and Cerf[16] found analytically that the adiabatic model is resistant to noise from Gaussian random matrix, in that the error probability doesn't increase with problem size, provided that the cutoff frequency of the noise is either very high or very low with respect to the characteristic frequency of the system. Their results suggested that the adiabatic algorithm is fault-tolerant as long as the noise spectrum doesn't contain frequencies close to resonances of the system.

Lastly, the existence of a minimum energy gap in adiabatic algorithms suggests an inherent tolerance of thermal excitation. Physical intuition indicates that the excitation of a system from its ground state is unlikely if $\Theta \ll \gamma$, where Θ is the ambient temperature. Childs *et al.* examined the connection of γ and the maximum temperature Θ for efficient adiabatic algorithms with a noise model of photon bath, and found that the success probability increases as temperature decreases. They also found sufficiently low temperature may even improve performance compared to no decoherence, since decoherence at low temperature appears to drive transitions towards the ground state. Unfortunately, they didn't discover a connection between the energy gap and the maximum temperature necessary for reliable adiabatic quantum computation. Jordan *et al.*[7] furthered their study by introducing the energy-gap protection through error-detecting stabilizer codes, and demonstrated that the temperature only needs to shrink logarithmically with problem size for the photon bath noise model. We will now reproduce their result in the below section, and later prove a more general result concerning energy-gap protection in Chapter 4.

3.2.2 Stabilizer Code and Energy-Gap Protection

Jordan, Farhi and Shor[7] provided an example of a [4,1] stabilizer code that detects arbitrary 1-qubit noise. The code is defined by the generators and logical Pauli operators in Table 3.1, where we again use the notation where $\bar{A}$ denotes the logical, encoded version of operator A. The basis states for the logical qubit are

$$|0_L\rangle = \frac{1}{2}(|0000\rangle + i\,|0011\rangle + i\,|1100\rangle + |1111\rangle) \tag{3.3}$$

$$|1_L\rangle = \frac{1}{2}(-|0101\rangle + i\,|0110\rangle + i\,|1001\rangle + |1010\rangle) \tag{3.4}$$

Name	Operator
g_1	$X \otimes X \otimes X \otimes X$
g_2	$Z \otimes Z \otimes Z \otimes Z$
g_3	$X \otimes Y \otimes Z \otimes I$
$\bar{X}$	$Y \otimes I \otimes Y \otimes I$
$\bar{Y}$	$-I \otimes X \otimes X \otimes I$
$\bar{Z}$	$Z \otimes Z \otimes I \otimes I$

Table 3.1: Generators and logical Pauli operators for the Jordan-Farhi-Shor [4,1] stabilizer code.

The codespace C is spanned by the logical basis, which is equivalently defined as all wavefunctions fixed by the stabilizer group $S = \langle g_1, g_2, g_3 \rangle$, i.e.

$$C = \{\alpha |0_L\rangle + \beta |1_L\rangle\} = \{|\psi\rangle : x |\psi\rangle = |\psi\rangle \, \forall x \in S\} \tag{3.5}$$

Note the logical Pauli operators are 2-local. Also note that $\bar{X}\bar{Y} \neq i\bar{Z}$, but that's fine because their action on the codespace follows expected behavior, i.e. $\bar{X}\bar{Y} |\psi\rangle = i\bar{Z} |\psi\rangle \; \forall |\psi\rangle \in C$, since $\bar{X}\bar{Y}g_3 = i\bar{Z}$. It isn't hard to check that all 1-local operators anti-commute with at least one of the generators, which is the necessary condition for detection of arbitrary 1-local errors.

The fact that one can detect arbitrary 1-local errors with certain Hermitian operators indicates that it is also possible to impose an "energy penalty" on states with an error. Utilizing the fact that energy gap seems to provide protections against excitation, Jordan *et al.* proposed a method to suppress error in adiabatic quantum computation: encoding the original system Hamiltonian H_{S0} with the logical Pauli operators stabilizer code, and imposing an energy penalty Hamiltonian, so that the full system Hamiltonian becomes

$$H_S = \bar{H}_{S0} + H_P \tag{3.6}$$

$$\bar{H}_{S0} = H_{S0}(X \to \bar{X}, Y \to \bar{Y}, Z \to \bar{Z}) \tag{3.7}$$

$$H_P = -\sum_i E_p(g_{i1} + g_{i2} + g_{i3}) \tag{3.8}$$

where in the energy penalty Hamiltonian H_{SP}, we sum over all the logical qubits, each with generators g_{i1}, g_{i2}, g_{i3} from the Jordan-Farhi-Shor code. Note any state outside the codespace C incurs an energy increase relative to the ground state by at least E_p. In fact, the spectrum of H_{SP} on a single logical qubit consists of energy eigenvalues $\pm 3E_p$, $\pm E_p$, so the energy between the ground state and the first excited state with error is actually $2E_p$.

This energy-gap protection method of suppressing error is appealing because it doesn't require the experimenter to interfere with the system during its adiabatic evolution, which preserves "the hands-off" approach that the adiabatic algorithm takes. To investigate the efficacy of error-suppression via the energy penalty, Jordan *et al.* studied the following Hamiltonian

$$H = H_S + H_B + \lambda V \tag{3.9}$$

where H_S is the system Hamiltonian, H_B is the Hamiltonian of a photon bath (H_E in their

paper), and the noise model V is a coupling between spin qubits and the photon bath:

$$V = \sum_i \int_0^\infty d\omega [g(\omega) a_\omega \sigma_+^{(i)} + g^*(\omega) a_\omega^\dagger \sigma_-^{(i)}] \tag{3.10}$$

where $g(\omega)$ is the spectral density, a_ω is the annihilation operator for photon mode with frequency ω, and $\sigma_\pm^{(i)}$ are the raising and lowering operators for the i^{th} system spin qubit. This model was also studied by Childs *et al.* in Ref. [2]. They were able to derive the following master equation for the reduced density matrix ρ of the system

$$\frac{d\rho}{dt} = -i[H_S, \rho] - \sum_{a,b} M_{ab} \mathcal{E}_{ab}(\rho) \tag{3.11}$$

where the terms are defined as

$$\begin{aligned}
M_{ab} &= \sum_i \left[N_{ba} |g_{ba}|^2 \langle a|\sigma_-^{(i)}|b\rangle \langle b|\sigma_+^{(i)}|a\rangle + (N_{ab}+1)|g_{ab}|^2 \langle b|\sigma_-^{(i)}|a\rangle \langle a|\sigma_+^{(i)}|b\rangle \right] \\
\mathcal{E}_{ab}(\rho) &= |a\rangle\langle a|\rho + \rho|a\rangle\langle a| - 2|b\rangle\langle a|\rho|a\rangle\langle b| \\
N_{ba} &= \frac{1}{\exp[\beta(\omega_b - \omega_a)] - 1} \\
g_{ba} &= \begin{cases} \lambda g(\omega_b - \omega_a), & \omega_b > \omega_a \\ 0, & \omega_b \le \omega_a \end{cases}
\end{aligned}$$

where $|a\rangle$, $|b\rangle$ are instantaneous eigenstates of H_S with energy ω_a and ω_b, respectively. N_{ba} is the Bose-Einstein distribution at temperature $1/\beta$ assumed for the photon bath. Observe that M_{ab} is a scalar, while $\mathcal{E}_{ab}$ is an operator.

Assuming at $t = 0$, we start in the ground state $|0\rangle$ in the codespace. Then

$$\frac{d\rho}{dt}\Big|_{t=0} = -i[H_S, \rho] - \sum_{b \in C^\perp} M_{0b} \mathcal{E}_{0b}(\rho) \tag{3.12}$$

Since $|0\rangle$ is the ground state, we have $\omega_b \ge \omega_0$. Thus in the expression of M_{0b}, the term with $|g_{0b}|$ will vanish, and the term with $|g_{b0}|^2 = O(\lambda^2)$ is bounded by N_{b0}. Note for $|b\rangle \in C^\perp$, we have

$$\omega_b - \omega_0 = \langle b|H_{SL} + H_{SP}|b\rangle - \langle 0|H_{SL} + H_{SP}|0\rangle \ge \langle b|H_{SL}|b\rangle - \langle 0|H_{SL}|0\rangle + E_p \ge E_p$$

where we used the fact that $|0\rangle$ is the ground state of H_{SL} and $|b\rangle \in C^\perp$ incurs an energy penalty of at least E_p. Hence

$$N_{b0} = \frac{1}{\exp[\beta(\omega_b - \omega_0)] - 1} \le O(e^{-\beta E_p})$$

Since the $M_{0b}\mathcal{E}(\rho)$ is exponentially suppressed by E_p and only grows polynomially with the problem size n, we expect for large βE_p,

$$\frac{d\rho}{dt}\Big|_{t=0} = -i[H_S, \rho] + O(\lambda^2 e^{-\beta E_p}) \approx -i[H_S, \rho]$$

This is the primary result of Jordan *et al.*: for adiabatic algorithm on a system of spin

qubits coupled to a photon bath, it is sufficient for $\beta E_p = E_p/\Theta$ grow logarithmically with the problem size n. Nevertheless, this result is still unsatisfying since it relies on a narrow class of noise models. Also, the assumption of Bose-Einstein distribution for the photon bath is responsible for the good bound that Jordan *et al.* obtained.

We will take Jordan *et al.*'s idea of energy-gap protection further in Chapter 4 by proving a much more general result that applies to not only adiabatic quantum computation but also generic quantum evolution in a open system.

3.2.3 Dynamical Decoupling

Another interesting technique put forth for potentially helping adiabatic algorithms to achieve fault-tolerance is dynamical decoupling. First proposed by Viola *et al.*[18], dynamical decoupling was recently put to use for suppressing error in adiabatic quantum computation by Lidar[9]. The model of interest is the following Hamiltonian

$$H(t) = H_S(t) + H_B + \sum_j E_j \otimes B_j + H_C(t) \tag{3.13}$$

where $H_C(t)$ is a control Hamiltonian, E_j are single-qubit Pauli error operators. and B_j are corresponding bath operators. The system Hamiltonian $H_S(t)$ is assumed to be the time-dependent adiabatic algorithm Hamiltonian, while the bath Hamiltonian H_B as well as E_j and B_j are assumed to be time-independent.

As summarized by Young *et al.*[19], the typical dynamical decoupling scheme essentially involves encoding the system Hamiltonian in some stabilizer code, and applying the stabilizer group generators as unitary operations by manipulating $H_C(t)$, which in general is

$$H_C(t) = \begin{cases} 0, & t_k \le t < t_{k+1} - w \\ H_{DD}^{(k)}, & t_{k+1} - w \le t \le t_{k+1} \end{cases} \tag{3.14}$$

where $t_k = k(\tau + w)$, where τ and w are the pulse interval and width of the dynamical decoupling sequence. Here, the pulse Hamiltonians $H_{DD}^{(k)}$ are generators of the stabilizer group. The unitary evolution operator under $H_C(t) + H_B$ is

$$U_C(t) = e^{-i\int_0^t (H_C(\tau)+H_B)d\tau} = \prod_{j=1}^{K(t)} P_j \tag{3.15}$$

where $K(t) = \lfloor t/(\tau+w) \rfloor \in \mathbb{Z}$ gives the number of pulses up to time t, and P_j is the unitary operator corresponding to the j^{th} pulse. In the interaction picture of $H_C(t) + H_B$, we have

$$\tilde{H}_{DD}(t) = \tilde{H}_S(t) + \sum \tilde{E}_j(t) \otimes \tilde{B}_j(t) \tag{3.16}$$

where an operator A in the interaction picture is denoted with a tilde as $\tilde{A} \equiv U_C^\dagger A U_C$. Now, because E_j is a Pauli operator that either commutes or anti-commutes with the members of the stabilizer group, we have the effective Hamiltonian

$$\tilde{E}_j(t) = U_C^\dagger(t) E_j U_C(t) = (-1)^{p_j(t)} E_j \quad \text{where} \quad p_j(t) = \begin{cases} 0 & \text{if } [E_j, U_C(t)] = 0 \\ 1 & \text{if } \{E_j, U_C(t)\} = 0 \end{cases} \tag{3.17}$$

As we can see, a well-chosen dynamical decoupling sequence can cause $p_j(t)$ to oscillate rapidly between 0 and 1, which would cause the system-bath coupling to average to zero on timescales longer than the pulse period.

Despite its allure as an innovative approach for error-suppression in AQC, dynamical decoupling also requires a few caveats. First of all, it needs a large number of controlled pulses and thus constant albeit repetitive interference from the experimenter, which somewhat diminishes the "hands-off" appeal of adiabatic algorithms. Additionally, since dynamical decoupling calls for arbitrarily fast pulse sequences and short cycles to achieve arbitrary accuracy, one would worry that the it might cause arbitrarily large energy perturbations in the system due to the energy-time uncertainty principle. As a result, we decided to investigate the passive energy-gap protection in more depth for this thesis.

Chapter 4

Energy-Gap Protection in General Open-System Evolution

Energy-gap protection has the appeal that it preserves the "hands-off" approach of adiabatic quantum computation, since we only need to set up the energy penalty Hamiltonian in the beginning. In this chapter, we will analyze the use of energy-gap protection against decoherence in open-system quantum evolution. This result is applicable not only to the fault-tolerance of adiabatic quantum computation but also general quantum evolutions in an open system with a wide range of noise models.

Although density matrix formalism and master equations are generally necessary to study evolution in an open system, our analysis in this chapter does not require the use of mixed states and reduced density matrices of the system. In fact, we are able to prove results for the fully unitary evolution on the pure state of the overall system that includes the environment/bath, which is a more general result that what was done before in Jordan *et al*[7]. As we will see, the energy-gap protection effectively decouples the system and the bath, and provides protection for both the system evolution and the bath evolution.

4.1 The Hamiltonian Model

The model that we're considering is the quantum evolution under the following form of a Hamiltonian

$$H = H_S + H_B + \lambda V + E_p H_P \tag{4.1}$$

where H_S is the system Hamiltonian, H_B is the bath Hamiltonian, V is the error Hamiltonian, and H_P is the energy penalty Hamiltonian. Here it is understood that $H_S = H_S \otimes I_B$ acts trivially on the bath, and $H_B = I_S \otimes H_B$ acts trivially on the system. The parameters λ and E_p characterize the strength of the error/decoherence and the energy-gap protection, respectively, and both have units of energy. As a result, V and H_P are unitless in this convention. We will also denote

$$H_0 \equiv H_S + H_B \tag{4.2}$$

H_0 indicates the unperturbed, error-free Hamiltonian. The goal of this chapter is to analyze the fidelity of the quantum evolution under H when compared to H_0.

Our general noise model is the error Hamiltonian V of the following form:

General Noise Model:
$$V = \sum_j E_j \otimes B_j \tag{4.3}$$

The E_j's are the error operators acting on the system, and B_j's are corresponding bath operators. When $B_j = I$, we can consider $E_j \otimes B_j$ to a control error on the system Hamiltonian. When $B_j \neq I$, we have a coupling between the system and the bath that causes decoherence. Note in general E_j can be nonlocal, simultaneously operating on many physical qubits in the system. Thus, the index j is simply dummy variable the sums over all such possible error operators.

In the discussion that follows, we will sometimes make use of the simplifying assumption that we have a local noise model, i.e. the error Hamiltonian V can be written as a sum of Hamiltonians V_i acting on some subsystems (e.g. n different logical qubits). Mathematically, we write

Local Noise Model:
$$V = \sum_{i=1}^{n} V_i \quad \text{where} \quad V_i = \sum_s E_{i,s} \otimes B_{i,s} \tag{4.4}$$

where $E_{i,s}$'s only act on the i^{th} subsystem, with the index s summing over all the possibilities of errors within the subsystem. This doesn't allow the presence of an error E_j in V that operates on two different subsystem, for instance. This assumption, which we term **Noise Locality**, is a physically reasonable. For example, it is applicable to the fairly generic 1-local noise model where the noises E_j act 1-locally on the system qubits. The mathematical treatment necessary for generalization of our result beyond the local noise model is omitted in this chapter for succinctness, but can be found in Appendix A.

The experimenter has full control over H_S and H_P, while putting his/her best effort to suppress the magnitude of λV. In general, it is possible that H_S, H_B, and V can be time-dependent. In the following analysis, we won't make any assumptions regarding their time-dependence, and all of them are understood implicitly to be time-dependent unless otherwise noted.

4.2 Error-Detecting Code and Energy Penalty Hamiltonian

In order to impose energy-gap protection against a set of errors $\{E_j\}$, we need a quantum code that can detect these errors, since the idea is to impose an energy penalty on states with error. A set of errors $\{E_j\}$ is *detectable* by a quantum code if the projector P onto its codespace C satisfy

Error Detectability:
$$PE_jP = 0 \tag{4.5}$$

Note the contrast with the correctability condition in Eq. (2.24). As we can see, we don't consider the identity operation on the system as a detectable error, which can be treated as an overall phase or part of the bath Hamiltonian. We will thus assume without loss of generality that all terms in our error Hamiltonian V are nontrivial operations on the system.

Practically, we want an error-detecting scheme that is scalable in the quantum computing framework, which generally requires the Hamiltonians to be local. For a computation with n logical qubits, we propose an error-detecting stabilizer code C implementation that uses

a series of $[\ell, 1]$ stabilizer code C_i's, each encoding one logical qubit. This means that each logical qubit is encoded in ℓ physical qubits, and thus there are a total of $N_p = n \times \ell$ physical qubits for a computation on n logical qubits. The projector P onto the overall codespace C and the projector P_i onto the individual codespace C_i are

$$P = \prod_{i=1}^{n} P_i \tag{4.6}$$

$$P_i = |0_L\rangle\langle 0_L|_i + |1_L\rangle\langle 1_L|_i = \prod_{m=1}^{\ell-1} \frac{I + g_{i,m}}{2^{\ell-1}} \tag{4.7}$$

where the last equality can be obtained by noting the fact that the codespace C_i is the simultaneous $+1$ eigenvectors of all the generators $g_{i,m}$'s on the i^{th} logical subsystem. Note for an $[\ell, 1]$ stabilizer code, there are $\ell - 1$ generators that the index m is summed over here. We will also define $Q_i \equiv I - P_i$ as the projector onto $C_i^{\perp}$.

In general, we call a noise model V of Eq. (4.3) is detectable if and only if

$$PVP = 0 \tag{4.8}$$

which follows from Eq. (4.5). Now, for the local noise model in Eq. (4.4), we can see that this implies the errors E_{ij} are detectable by P_i, and the following condition must hold for all indices i, j iterated over the n logical qubits:

$$P_j V_i P = \begin{cases} 0 & i = j \\ V_i P & i \neq j \end{cases} \quad \text{and} \quad Q_j V_i P = \begin{cases} V_i P & i = j \\ 0 & i \neq j \end{cases} \tag{4.9}$$

This condition is a narrower version of Eq. (4.8), but it makes our calculation simpler. Furthermore, as we will see in Appendix A, we can generalize our results to the more general error Hamiltonians V that only need to satisfy Eq. (4.8).

To impose an energy penalty on states that are outside the codespace (i.e. states with detectable errors, in $C^{\perp}$), we consider the following energy penalty Hamiltonian

$$H_P = \mathcal{Q} \equiv \sum_{i=1}^{n} Q_i = \sum_{i=1}^{n} (I - P_i) \tag{4.10}$$

Note a state $|\Psi\rangle \in C$ if and only if $\mathcal{Q}|\psi\rangle = 0$, and any state in outside the codespace (in $C^{\perp}$) incurs an energy cost of at least E_p due to $E_p\mathcal{Q}$. This definition of H_P not only preserves the energy values of the states in the codespace, but also allows the energy penalty Hamiltonian to be local as it is a sum of terms that operates independently on each logical qubit. In contrast, $I - P$ can also serve as a candidate for H_P, but it is highly nonlocal and thus undesirable for the practically-minded.

An example of a quantum error-detecting code that detects the set of 1-local Pauli errors is the Jordan-Farhi-Shor [4,1] stabilizer code outlined in Table 3.1 of the previous section, and summarized here:

Name	Operator
g_1	$X \otimes X \otimes X \otimes X$
g_2	$Z \otimes Z \otimes Z \otimes Z$
g_3	$X \otimes Y \otimes Z \otimes I$
$\bar{X}$	$Y \otimes I \otimes Y \otimes I$
$\bar{Y}$	$-I \otimes X \otimes X \otimes I$
$\bar{Z}$	$Z \otimes Z \otimes I \otimes I$

$$|0_L\rangle = \frac{1}{2}(|0000\rangle + i|0011\rangle + i|1100\rangle + |1111\rangle) \tag{4.11}$$

$$|1_L\rangle = \frac{1}{2}(-|0101\rangle + i|0110\rangle + i|1001\rangle + |1010\rangle) \tag{4.12}$$

Per earlier convention, we denote logical operator $\bar{A}$ as the encoded version of operator A. Stabilizer formalism allows for easy translation of unencoded Hamiltonians to encoded ones, by replacing the Pauli operators in the unencoded Hamiltonian with the logical Pauli operators:

$$H_S(X, Y, Z) \to \bar{H}_S(\bar{X}, \bar{Y}, \bar{Z}) \tag{4.13}$$

This translation works because the Pauli operators (along with the identity operator) serve as a basis for arbitrary Hamiltonians on qubits.

In addition, the Jordan-Farhi-Shor code has the nice property that all its logical operators are 2-local. In fact, this encoding is optimal for protection against arbitrary 1-local errors, especially since the minimum size of a stabilizer code that detects arbitrary 1-local error is 4[7]. Hence, if the unencoded system Hamiltonian H_S is 2-local, then the encoded Hamiltonian H_{SL} will be 4-local. Coincidentally, the energy penalty Hamiltonian $\mathcal{Q}$ is also 4-local since it is a sum of 4-local terms, each involving some 4-local generators g_i. Therefore, using Jordan-Farhi-Shor encoding allows us to suppress arbitrary 1-local noise with an encoded Hamiltonian that needs only to be 4-local on the system.

4.3 Analysis of the Open-System Evolution

Now that we have our model of the Hamiltonian evolution in an open system with an energy penalty, given by Eqs. (4.1)(4.3)(4.10), we can proceed to analyze the utility of the error-suppression by energy-gap protection in the open-system quantum evolution. In places where we assume Noise Locality as in Eq. (4.4), we will make an explicit note, and point to the generalization in Appendix A when possible.

4.3.1 The infinite E_p limit

The goal of this section is to show that $U_0 P \approx UP$ in the limit of large E_p, i.e. that the evolution in the codespace remains coherent even in the open system provided E_p is large. To help us, let us define

$$H_{0P} \equiv H_0 + E_p H_P = H_0 + E_p \mathcal{Q} \tag{4.14}$$

which is the error-free Hamiltonian with the energy penalty added. Let $U_0(t)$, $U_{0P}(t)$ and $U(t)$ denote the unitary evolution operator under the Hamiltonians $H_0(t)$, $H_{0P}(t)$, and $H(t)$, respectively. Formally, U_x is the solution to the Schrödinger equation $i dU_x/dt = H_x U_x$ with the initial condition $U(0) = I$. Since $[H_0(t), \mathcal{Q}] = 0$ as H_0 is an encoded Hamiltonian, we

have

$$U_{0P}(t) = U_0(t)e^{-iE_p\mathcal{Q}t} = U_0(t)U_P(t) \tag{4.15}$$

where we denoted $U_P(t) \equiv e^{-iE_p\mathcal{Q}t}$. Since Q_i's act independently on the i^{th} logical subsystems, and Q_i's and P_i's are projectors with $Q_iP_i = 0$, we can obtain the following useful relation

$$U_PP = \prod_{i=1}^{n}(e^{-iE_pt}Q_i + P_i) \cdot \prod_{j=1}^{n} P_j = \prod_{i=1}^{n} P_i = P \tag{4.16}$$

which affirms that states in the codespace are unaffected by evolution under $H_P = \mathcal{Q}$. Note the products here are implicitly tensor products $\otimes$.

Let us now switch into the interaction picture where the standard evolution is under H_{0P}. Then we have $V_I = U_{0P}^\dagger V U_{0P}$ and $U_I = U_{0P}^\dagger U$. From these definitions, it is immediately clear that

$$UU_I^\dagger P = U_{0P}P = U_0U_PP = U_0P \tag{4.17}$$

Thus, we need to examine the behavior of $U_I^\dagger$ in order to understand the behavior of U and U_0. Note in the interaction picture, we have

$$i\frac{dU_I}{dt} = \lambda V_I U_I \implies \frac{dU_I^\dagger}{dt} = i\lambda U_I^\dagger V_I \implies U_I^\dagger(T) = I + i\lambda\int_0^T U_I^\dagger(t)V_I(t)dt$$

$$U_I^\dagger P = P + i\lambda\int_0^T U_I^\dagger V_I P dt \tag{4.18}$$

To ease the eye, we omitted the explicit time dependences in the operators in the last line and will continue to do so in the ensuing discussion unless otherwise required for clarity.

To help us analyze $U_I^\dagger P$, let us define

$$F(T) \equiv \int_0^T V_I P dt \tag{4.19}$$

Now we can perform integration by parts on the Eq. (4.18)

$$\begin{aligned} U_I^\dagger P &= P + i\lambda\int_0^T U_I^\dagger\frac{dF}{dt}dt = P + i\lambda\Big(U_I^\dagger F\Big|_0^T - \int_0^T \frac{dU_I^\dagger}{dt}Fdt\Big) \\ &= P + i\lambda\Big(U_I^\dagger F - i\lambda\int_0^T U_I^\dagger V_I F dt\Big) \end{aligned}$$

where we used the fact that $F(0) = 0$. By left-multiplying both sides with U, we have

$$\boxed{U_0P = UU_I^\dagger P = UP + i\lambda U_{0P}F + \lambda^2 U\int_0^T U_I^\dagger V_I F dt} \tag{4.20}$$

So it seems the operator F plays an important role in the difference of U_0P and UP.

Under the Noise Locality assumption of Eq. (4.4), observe the following

$$\begin{aligned} U_P^\dagger VP &= e^{iE_p Qt}\sum_i V_i P = \sum_{i=1}^n \prod_{j=1}^n (e^{iE_p t}Q_j + P_j)V_i P = \sum_{i=1}^n e^{iE_p t}V_i P \\ &= e^{iE_p t}VP \end{aligned} \tag{4.21}$$

where we used the fact that for $j \neq i$, $P_i V_i P = Q_j V_i P = 0$ and $P_j V_i P = Q_i V_i P = V_i P$ from Eq. (4.9). We can then simplify the expression for F through

$$\begin{aligned} F(T) &= \int_0^T V_I P dt = \int_0^T U_{0P}^\dagger V U_{0P} P dt = \int_0^T U_P^\dagger U_0^\dagger V U_0 U_P P \\ &= \int_0^T U_P^\dagger U_0^\dagger V U_0 P = \int_0^T U_0^\dagger U_P^\dagger V P U_0 dt = \int_0^T U_0^\dagger e^{iE_p t} V P U_0 dt \\ &= \int_0^T e^{iE_p t} U_0^\dagger(t) V(t) U_0(t) P dt \end{aligned} \tag{4.22}$$

where we made use of Eqs. (4.16)(4.21) and the fact that U_P and P commutes with U_0. The generalization of this formula with a nonlocal noise model is given in Appendix A, where the k-local terms in V would simply pull out a factor of $e^{ikE_p t}$ instead of $e^{iE_p t}$ in the integrand.

Immediately, we can see that since $U_0^\dagger V U_0 P$ is independent of E_p, we can apply the Riemann-Lebesgue Lemma and conclude that $F(T) \to 0$ when $|E_p| \to \infty$, demonstrating $U_0 P \approx UP$ in the infinite E_p limit. Explicitly, we can integrate by parts to get

$$\boxed{F(T) = \frac{1}{iE_p}\Big[e^{iE_p T}U_0^\dagger(T)V(T)U_0(T) - V(0) - \int_0^T e^{iE_p t}\frac{d}{dt}(U_0^\dagger V U_0)dt\Big]P} \tag{4.23}$$

This can be done as long as $U_0^\dagger V U_0$ is finite and differentiable in t, which is reasonable for any finite system and bath. Plugging this expression into Eq. (4.20), we can see that for finite λ

$$U_0 P = U\Big[I + O\Big(\frac{1}{E_p}\Big)\Big]P \tag{4.24}$$

Note that for a k-local noise model, the generalization in Appendix A shows that the effective protection on the terms in the noise model that act on k different logical qubits scales with $1/kE_p$ instead of just $1/E_p$.

Theorem 4.1. *Suppose we have an evolution in an open system with finite degrees of freedom, where the overall Hamiltonian is $H(t) = H_S(t) + H_B(t) + \lambda V(t) + E_p Q$ like we defined above. If the error Hamiltonian V satisfies the error-detection condition $PV(t)P = 0$ and λ is finite, then in the limit of infinitely large energy penalty E_p, the evolution in the codespace is the same as if there were no error at all. Mathematically, for all time t, we have*

$$\lim_{E_p\to\pm\infty} U(t)P = \lim_{E_p\to\pm\infty} U(t)\Big[I + O\Big(\frac{1}{E_p}\Big)\Big]P = \lim_{E_p\to\pm\infty} U_0(t)P = U_0(t)P \tag{4.25}$$

Remarkably, just as indicated by the Riemann-Lebesgue Lemma, the sign of E_p is irrelevant for this protection of coherence, so the energy gap can be positive or negative. Also interestingly, the evolution of the bath also retains high fidelity thanks to the energy-gap protection.

4.3.2 Behavior with finite E_p

Even though we proved that the evolution in the open system remains (from both system and bath perspectives) coherent for the codespace in the large E_p limit, we also want to answer the question: how large does E_p need to be for practically high enough fidelity, given the size of the computation n, the strength of decoherence λ, and the time of evolution T? To answer that question, we want to investigate the difference $U_0P - UP$, which primarily depends on $F(t)$. Hence, our first goal in this section is to bound $\|F\|$.

Bounding $\|F\|$ Naïvely, we can simply look at the expression of $F(T)$

$$\begin{aligned} F(T) &= \frac{1}{iE_p}\Big[e^{iE_pT}U_0^\dagger(T)V(T)U_0(T) - V(0) - \int_0^T e^{iE_pt}\frac{d}{dt}(U_0^\dagger V U_0)dt\Big]P \\ &= \frac{1}{iE_p}\Big[e^{iE_pT}U_0^\dagger(T)V(T)U_0(T) - V(0) - \int_0^T e^{iE_pt}U_0^\dagger\Big(i[H_0,V] + \frac{dV}{dt}\Big)U_0dt\Big]P \end{aligned}$$

Using the triangle and Cauchy-Schwarz inequalities, this give us a bound of

$$\textit{Bound 1:} \qquad \|F(T)\| \leq \frac{1}{|E_p|}(2\|V\| + \|[H_0,V]\|T + \|dV/dt\|T) \tag{4.26}$$

where each term is understood implicitly to be the maximum over $t \in [0,T]$. The linear dependence on T originates from the integral over time and is likely artificial, especially since the integrand oscillates with e^{iE_pt}. Note in this case $\|F(T)\| \propto n$, the computational size, because V normally grows linearly with n.

We can improve this bound by making assumptions on our system. We're interested in the following two assumptions:

1. **Noise Locality:** $V = \sum_{i=1}^n V_i$, where V_i operates independently on the i^{th} logical subsystem. This is the same assumption we made in Eq. (4.4).
2. **Geometric Locality:** All terms in $H_S(t)$, $H_B(t)$, and $V(t)$ are local, i.e. each acting on a bounded number of physical qubit.

Using the Noise Locality assumption, we can improve our Eq. (4.26) bound using the fact that we can define

$$F_i(T) = \int_0^T e^{iE_pt}U_0^\dagger V_i U_0 P dt \tag{4.27}$$

And noticing that since Q_i commutes with U_0 for all time and $Q_iV_jP = \delta_{ij}V_iP$ due to Eq. (4.9), then we have for $i \neq j$

$$\begin{aligned} PV_i(t_1)U_0(t_1)U_0^\dagger(t_2)V_j(t_2)P &= PV_i(t_1)Q_iU_0(t_1)U_0^\dagger(t_2)V_j(t_2)P \\ &= PV_i(t_1)U_0(t_1)U_0^\dagger(t_2)Q_iV_j(t_2)P = 0 \end{aligned}$$

Hence

$$\begin{aligned} F^\dagger F &= \sum_{i,j=1}^n F_i^\dagger F_j = \sum_{i,j=1}^n \int_0^T dt_1 \int_0^T dt_2 PU_0^\dagger(t_1)V_i(t_1)U_0(t_1)e^{-iE_pt_1}e^{iE_pt_2}U_0^\dagger(t_2)V_j(t_2)U_0(t_2)P \\ &= \sum_{i,j=1}^n \int_0^T dt_1 \int_0^T dt_2 e^{iE_p(t_2-t_1)}U_0^\dagger(t_1)PV_i(t_1)U_0(t_1)U_0^\dagger(t_2)V_j(t_2)PU_0(t_2) = \sum_{i=1}^n F_i^\dagger F_i \end{aligned}$$

And thus

$$\|F\|^2 = \max_{|\psi\rangle} \|F|\psi\rangle\|^2 = \max_{|\psi\rangle} \langle\psi|F^\dagger F|\psi\rangle = \sum_{i=1}^{n} \max_{|\psi\rangle} \langle\psi|F_i^\dagger F_i|\psi\rangle \leq n \max_i \|F_i\|^2$$

Then we can apply this argument to our bound Eq. (4.26) to obtain

Bound 2: $$\|F\| \leq \sqrt{n} \max_i \|F_i\| = \frac{\sqrt{n}}{|E_p|}(2\|V_i\| + \|[H_0, V_i]\|T + \|dV_i/dt\|T) \tag{4.28}$$

Note we essentially have the same bound as Bound 1, except now $\|F\| \propto \sqrt{n}$ instead of growling linearly with n.

Next, we can apply the Geometric Locality assumption to bound $[H_0, V_i]$, which in general can be arbitrarily large since H_0 also contains H_B, which is the bath Hamiltonian that may have extremely large energy levels. However, if H_S, H_B, and V are local, we note that V cannot take an energy eigenstate to something of drastically different. Mathematically, this means for energy eigenstate $|E\rangle$ of H_0, we have

$$\|[H_0, V]|E\rangle\| = \|H_0 V|E\rangle - EV|E\rangle\| = \|(H_0 - E)V|E\rangle\| \approx O(1) \tag{4.29}$$

This is true as long as $V|E\rangle$ doesn't give a superposition of energy eigenstates that has significant component in states with energy dramatically different from E. We also assume $\|dV/dt\|$ to be bounded like $\|V\|$. This argument, along with the Noise Locality assumption, indicates a new bound

Bound 3: $$\|F\| = \frac{\sqrt{n}}{|E_p|}(2\|V_i\| + O(1)T) \tag{4.30}$$

where $O(1)$ is a constant with the unit of energy.

Fidelity For a pure initial state $|\psi_0\rangle$ in the codespace, we denote

$$|\psi(T)\rangle = U(T)|\psi_0\rangle \qquad \text{and} \qquad |\phi_0(T)\rangle = U_0(T)|\psi_0\rangle \tag{4.31}$$

which are the states evolving with and without noise for time T. The fidelity between them is defined as

$$f(T) \equiv |\langle\phi_0(T)|\psi(T)\rangle| = |\langle\psi_0|U_0^\dagger U|\psi_0\rangle| = |\langle\psi_0|PU_0UP|\psi_0\rangle| \tag{4.32}$$

We can inserted the projector P because we assumed $|\psi_0\rangle \in C$. Therefore, in order to understand the behavior of fidelity, we simply need to examine the operator $PU_0^\dagger UP$. Since the projector P commutes with U_0 and we can get an expression for UP from Eq. (4.20), we have

$$\begin{aligned} PU_0^\dagger UP &= PU_0^\dagger PUP = PU_0^\dagger P\Big(U_0 P - i\lambda U_{0P} F - \lambda^2 U \int_0^T U_I^\dagger V_I F dt\Big)P \\ &= P - \lambda^2 PU_0^\dagger U \int_0^T U_I^\dagger V_I F dt = P - \lambda^2 PU_{0P}^\dagger U \int_0^T U_I^\dagger V_I F dt \\ &= P - \lambda^2 PU_I \int_0^T U_I^\dagger V_I F dt \end{aligned} \tag{4.33}$$

where we use the fact that $PU_{0P}F = U_{0P}PF = U_{0P}\int PV_IP = 0$ from line 1 to line 2, and $PU_0^\dagger = PU_P^\dagger U_0^\dagger = PU_{0P}^\dagger$ in line 2.

Note $U_I(\lambda = 0) = I$. If we expand U_I in orders of λ, we get $U_I = I + O(\lambda)$. Performing the expansion in λ up to $O(\lambda^2)$, we get

$$\begin{aligned} PU_0^\dagger UP &= P - \lambda^2 P(I + O(\lambda)) \int_0^T (I + O(\lambda)) V_I F dt = P - \lambda^2 P \int_0^T V_I F dt + O(\lambda^3) \\ &= P - \lambda^2 P \int_0^T V_I(t) \int_0^t V_I(t') P dt' dt + O(\lambda^3) \end{aligned}$$

Then

$$\langle\psi_0|PU_0^\dagger UP|\psi_0\rangle = 1 - \lambda^2 \langle\psi_0| \int_0^T dt \int_0^t dt' PV_I(t)V_I(t')P|\psi_0\rangle + O(\lambda^3)$$

If we look at squared fidelity, which has the nice interpretation as the probability of finding the real state $|\psi(T)\rangle$ to be in the noiseless state $|\phi_0(T)\rangle$, we can see that

$$\begin{aligned} f^2(T) &= |\langle\psi_0|PU_0^\dagger UP|\psi_0\rangle|^2 = \langle\psi_0|PU_0^\dagger UP|\psi_0\rangle \langle\psi_0|PU^\dagger U_0 P|\psi_0\rangle \\ &= 1 - \lambda^2 \langle\psi_0| \Big[\int_0^T dt \int_0^t dt' PV_I(t)V_I(t')P + \Big(\int_0^T dt \int_0^t dt' PV_I(t)V_I(t')P \Big)^\dagger \Big] |\psi_0\rangle + O(\lambda^3) \\ &= 1 - \lambda^2 \langle\psi_0| \Big[\int_0^T dt \int_0^t dt' PV_I(t)V_I(t')P + \int_0^T dt \int_0^t dt' PV_I(t')V_I(t)P \Big] |\psi_0\rangle + O(\lambda^3) \\ &= 1 - \lambda^2 \langle\psi_0| \Big[\int_0^T dt \int_0^t dt' PV_I(t)V_I(t')P + \int_0^T dt' \int_0^{t'} dt PV_I(t)V_I(t')P \Big] |\psi_0\rangle + O(\lambda^3) \\ &= 1 - \lambda^2 \langle\psi_0| \Big(\iint_{0\le t'\le t\le T} + \iint_{0\le t\le t'\le T} \Big) dt dt' PV_I(t)V_I(t')P|\psi_0\rangle + O(\lambda^3) \end{aligned}$$

Note that the two integrals have identical integrand and their integration regions sum to $[0, T]^2$, so

$$\begin{aligned} f^2(T) &= 1 - \lambda^2 \langle\psi_0| \Big[\int_0^T dt \int_0^T dt' PV_I(t)V_I(t')P \Big] |\psi_0\rangle + O(\lambda^3) \\ &= 1 - \lambda^2 \langle\psi_0| \Big[\int_0^T dt (V_I(t)P)^\dagger \int_0^T dt' V_I(t')P \Big] |\psi_0\rangle + O(\lambda^3) \\ &= 1 - \lambda^2 \langle\psi_0|F^\dagger F|\psi_0\rangle + O(\lambda^3) \end{aligned} \tag{4.34}$$

Assuming the λ-expansion in U_I is valid, Eq. (4.34) tells us that

$$1 - f^2(T) \le \frac{1}{2}\lambda^2 \|F^\dagger F\| + O(\lambda^3) \le \frac{1}{2}\lambda^2 \|F\|^2 + O(\lambda^3) \tag{4.35}$$

where we used the Cauchy-Schwarz Inequality: $\|F^\dagger F\| \le \|F\|^2$. If the assumptions of Noise and Geometric Locality are valid, so that we can utilize the bound on $\|F\|$ in Eq. (4.30), we can then show

$$1 - f^2(T) = O\Big(\frac{\lambda^2 n(1 + cT)^2}{E_p^2}\Big) \tag{4.36}$$

for some constant c that has the unit of time^{-1}, due to the ratio of $\|[H_0, V]\| + \|dV/dt\|$ to $2\|V\|$ in the bound of $\|F\|$. In other words, to ensure high enough fidelity, we need an energy

penalty that satisfies $E_p \gg c\lambda T\sqrt{n}$ for some constant c according to the above bound.

Nevertheless, we have reasons to believe that λ-expansion may not be accurate or even valid. It should be noted that an expansion in λ is essentially an expansion in T since with each λ comes an integral over time. Similar to how the expansion $\cos(t) = 1 - t^2/2! + t^4/4! + O(t^6)$ doesn't capture the long-term behavior, we don't expect the rough bound in Eq. (4.36) to necessarily capture the behavior of fidelity for large T.

4.4 Numerical Simulation

Since no fidelity bound found so far for a generic local noise model is better than Eq. (4.36), which is based on a λ-expansion, we turn to numerical simulation to determine and verify the asymptotic behavior of fidelity $f^2(T)$ with large E_p, small λ, and long T, assuming Noise and Geometric Locality.

The Model The system we simulated is the evolution of 1-logical qubit system coupled to n_b-qubit bath under

$$H = H_S + H_B + \lambda V + E_p H_P \tag{4.37}$$

Per usual, we will also denote $H_0 \equiv H_S + H_B$. The initial state of the system is a general state in the codespace, $|\psi_{0S}\rangle = \alpha|0_L\rangle + \beta|1_L\rangle$. The system Hamiltonian is the logical Pauli X operator:

$$H_S = \bar{X} \tag{4.38}$$

The system is encoded with the Jordan-Farhi-Shor code, where a logical qubit is represented by $\ell = 4$ physical qubits.

The bath Hamiltonian is

$$H_B = \sum_{a=1}^{n_b} c_a \sigma^{(i)} + \sum_{\{a,b\}\in G} c_{ab}\sigma^{(a)} \otimes \sigma^{(b)} \tag{4.39}$$

where $a, b \in \{1, 2, \ldots, n_b\}$ are indices of the bath qubits. Here G is a 3-regular graph, randomly generated using the pairing model. The reason why we choose to use 3-regular graphs is because they have the lowest locality while maintaining nontrivial configurations for a connected graph (the only connected 2-regular graph is a ring). The coefficients c_a and c_{ab} are numbers sampled from the uniform distribution on $[0.9, 1.1]$ to give more randomness to the bath. Each $\sigma = \vec{n} \cdot \vec{\sigma}$ is a random Pauli matrix, with $\vec{n}$ sampled uniformly from the unit sphere, and every σ that appears in the expression is randomly generated, independent of the others. This construction ensures a bounded and relatively constant norm of H_B and satisfies the Geometric Locality assumption of the qubits in the overall system. The initial state $|\psi_B\rangle$ of the bath is taken to be a random 2^{n_b} complex unit vector.

The noise model is taken to be (with $\ell = 4$)

$$V = \sum_{s=1}^{\ell} c_s \sigma^{(s)} + \sum_{s=1, b=b(s)}^{\ell} c_{sb}\sigma^{(s)} \otimes \sigma^{(b)} \tag{4.40}$$

where $b = b(s) \in \{1, 2, \ldots, n_b\}$ is the index of a bath qubit that is connected to the s^{th} physical qubit of the system. The coefficients c_s and c_{sb} are also sampled from the uniform

distribution on $[0.9, 1.1]$. In total, there are 2ℓ terms in V, each acting 1-locally on the system and producing an error detectable by the Jordan-Farhi-Shor code.

The energy penalty is simply

$$H_P = I - P = I - |0_L\rangle\langle 0_L| - |1_L\rangle\langle 1_L| = I - (I+g_1)(I+g_2)(I+g_3)/2^3 \tag{4.41}$$

Simulation Procedure Observe that the Hamiltonian H is fully time-independent. Hence, we can simulate the evolution exactly by diagonalizing H and computing the unitary evolution operator $U = e^{-iHt}$. For some given $(|\psi_{0S}\rangle, \lambda, E_p)$, the simulation procedure is then

1. Initialize a random instance of $(V, H_B, |\psi_B\rangle)$. Then $|\psi_0\rangle = |\psi_{0S}\rangle \otimes |\psi_B\rangle$.
2. Numerically diagonalize H and H_0 and find the exact unitary evolution operator $U(t) = e^{-iHt}$ and $U_0(t) = e^{-iH_0t}$.
3. At each time of interest, determine $|\psi(t)\rangle = U(t)|\psi_0\rangle$ and $|\phi_0(t)\rangle = U_0(t)|\psi_0\rangle$. Note $|\phi_0(t)\rangle = |\phi_{0S}(t)\rangle \otimes |\phi_{0B}(t)\rangle$ since the system and bath are separable in H_0. Then measure the following four metrics:
   (a) the squared fidelity of the total evolution:
   $$f^2(t) = |\langle\phi(t)|\psi(t)\rangle|^2 \tag{4.42}$$
   (b) the squared fidelity of the system evolution
   $$f_S^2(t) = \langle\phi_{0S}(t)|\rho_S(t)|\phi_{0S}(t)\rangle = \langle\psi(t)|(\rho_{0S}(t)\otimes I)|\psi(t)\rangle \tag{4.43}$$
   where $\rho_S(t) = \mathrm{Tr}_B|\psi(t)\rangle\langle\psi(t)|$, and $\rho_{0S}(t) = |\phi_{0S}(t)\rangle\langle\phi_{0S}(t)|$.
   (c) the probability of remaining in the codespace:
   $$p_C(t) = \langle\psi(t)|P|\psi(t)\rangle \tag{4.44}$$
   (d) the entanglement entropy:
   $$S(t) = -\mathrm{Tr}(\rho_S(t)\log_2\rho_S(t)) \tag{4.45}$$

Ideally, we want $f^2(t)$, $f_S^2(t)$ and $p_C(t)$ be close to 1 so the overall state $|\psi(t)\rangle$ as well as the state of the system $\rho_S(t)$ are evolving as close to the noiseless, error-free state $|\phi_0(t)\rangle$ as possible. In addition, we want there to be little entanglement between the system and the bath, so ideally $S(t) = 0$.

In order to check numerical accuracy throughout the simulations, we calculated $U(t)$ in two different ways in MATLAB: using either `expm()`, which uses the Padé approximation, or `eig()`, which allegedly uses LAPACK routines. Using `expm` on $-iHt$ straightforwardly gives us $U(t)$, while `eig` gives us the diagonal form of H and also allows easy computation of $U(t)$ by exponentiating the entries in the diagonal form. The norm of the difference in the state vector $|\psi(t)\rangle$ obtained through the two independent different algorithms is used as the estimated numerical precision after evolving for time t.

Results Fig. 4-1 show an example simulation of an instance of $(V, H_B, |\psi_B\rangle)$ with $|\psi_{0S}\rangle = |0_L\rangle$ and $\lambda = 0.01$ in a 6-qubit bath. As we can see, higher E_p provides better protection in

all four metrics, and allows $f^2(t), f_S^2(t), p_C(t)$ stay closer to 1 and $S(t)$ stay closer to 0 for longer period of time. It may be interesting to note that high E_p protects not only the system but also the bath evolution, evident in high values of $f^2(t)$ when E_p is sufficiently large. In addition, when $E_p = 1 = 10^2\lambda$, the protection is already pretty good up to $T = 1000$, so that the probability of measuring the correct state, both overall (f^2) and within the system (f_S^2), at the end remains above 0.8. When $E_p = 10 = 10^3\lambda$, the protection is so good that up to $t = 1000$, all probabilities look identical to 1.

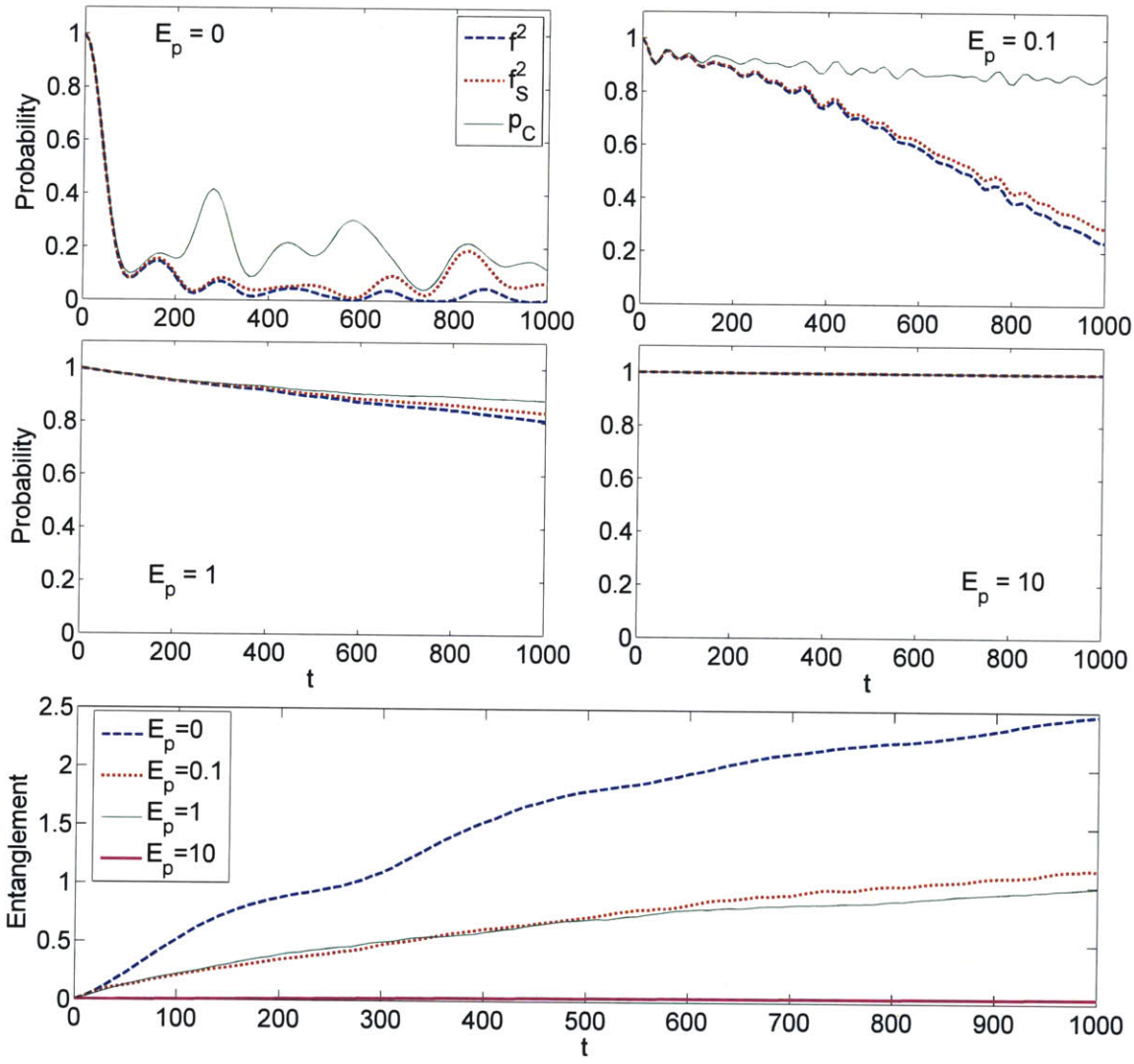


Figure 4-1: *Top 4:* For $\lambda = 0.01$ and $|\psi_{0S}\rangle = |0_L\rangle$, squared total fidelity $f^2(t)$, squared system fidelity $f_S^2(t)$, and probability of remaining in the codespace $p_C(t)$ up to $t = 1000$ for a particular instance of $(V, H_B, |\psi_B\rangle)$ with different E_p. *Bottom:* Entanglement entropy $S(t)$ between the system and the bath over time for the same instance with different E_p. The size of the bath is $n_b = 6$ qubits.

Asymptotic Behavior in Time (t)

Let us now examine a larger range of E_p, different initial system state $|\psi_{0S}\rangle$, and for evolution over much longer time. In Fig. 4-2 and 4-3, we show a typical result of the four metrics plotted over a log-time scale, with three different initial $|\psi_{0S}\rangle$ and various E_p. Both

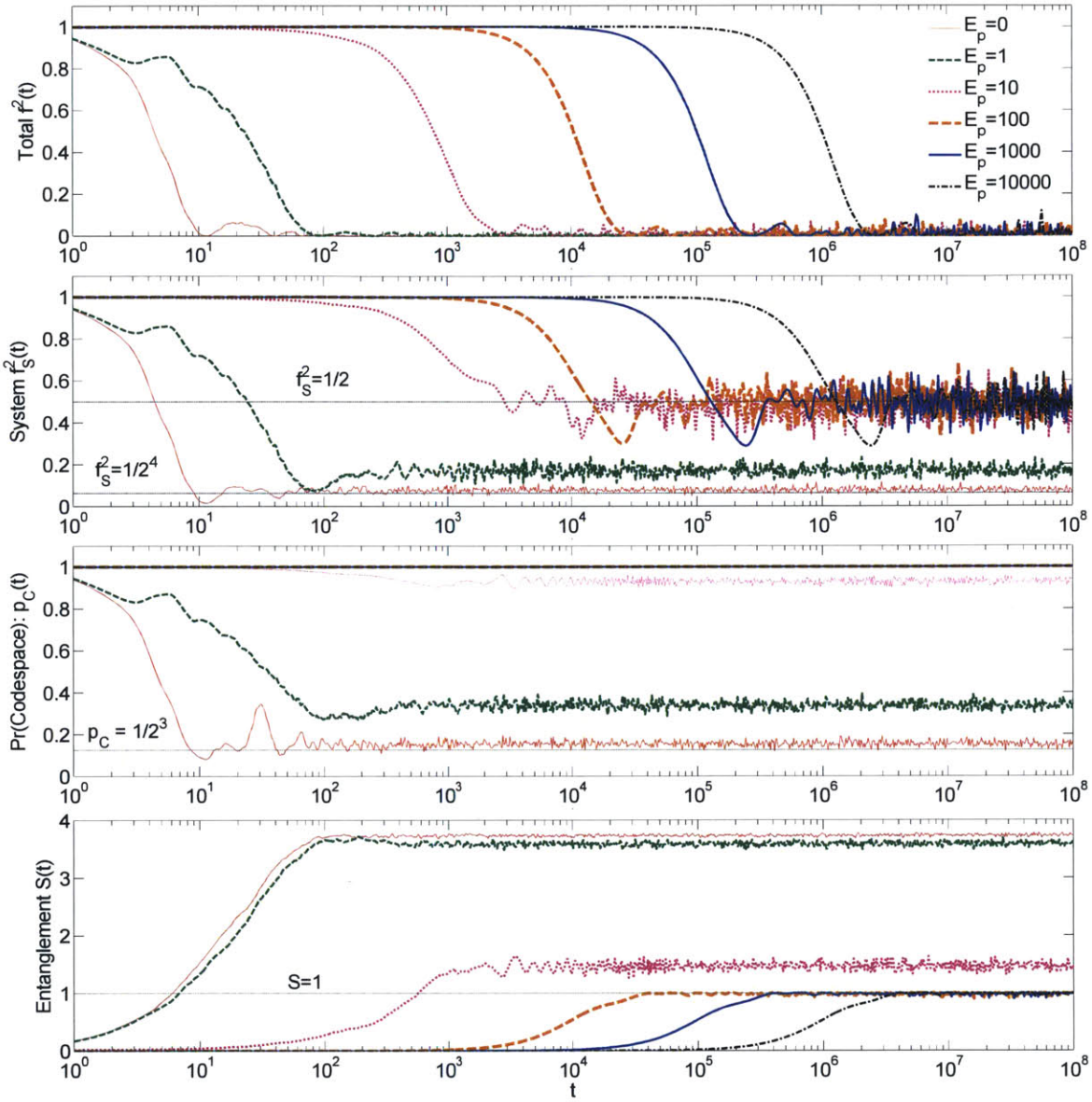


Figure 4-2: *From top to bottom:* For $\lambda = 0.1$ and $|\psi_{0S}\rangle = |0_L\rangle$, squared total fidelity $f^2(t)$, squared system fidelity $f_S^2(t)$, probability of remaining in codespace $p_C(t)$, and entanglement entropy $S(t)$ as a function of time up to $t = 10^8$ on a log-scale, with E_p of different orders of magnitude. All data are for a particular instance of $(V, H_B, |\psi_B\rangle)$. The bath size is $n_b = 6$. The worst-case numerical precision at $t = 10^8$ is approximately 10^{-4}.

For guiding the eye, we have plotted a few horizontal lines at interesting values. The two lines in the $f_S^2(t)$ plot are at values $1/2$ and $1/2^4$, which are expected probabilities of measuring a particular state when we have a maximally mixed 1- or 4-qubit state, respectively. The line in the $p_C(t)$ plot is at value $p_C = 1/2^3$, the expected probability of being in the codespace (dimension 2) when the system is in a maximally mixed 4-qubit state (dimension 2^4). The line in the bottom $S(t)$ plot is at value $S = 1$, which is the entropy of a maximally mixed state in a 1-qubit system.

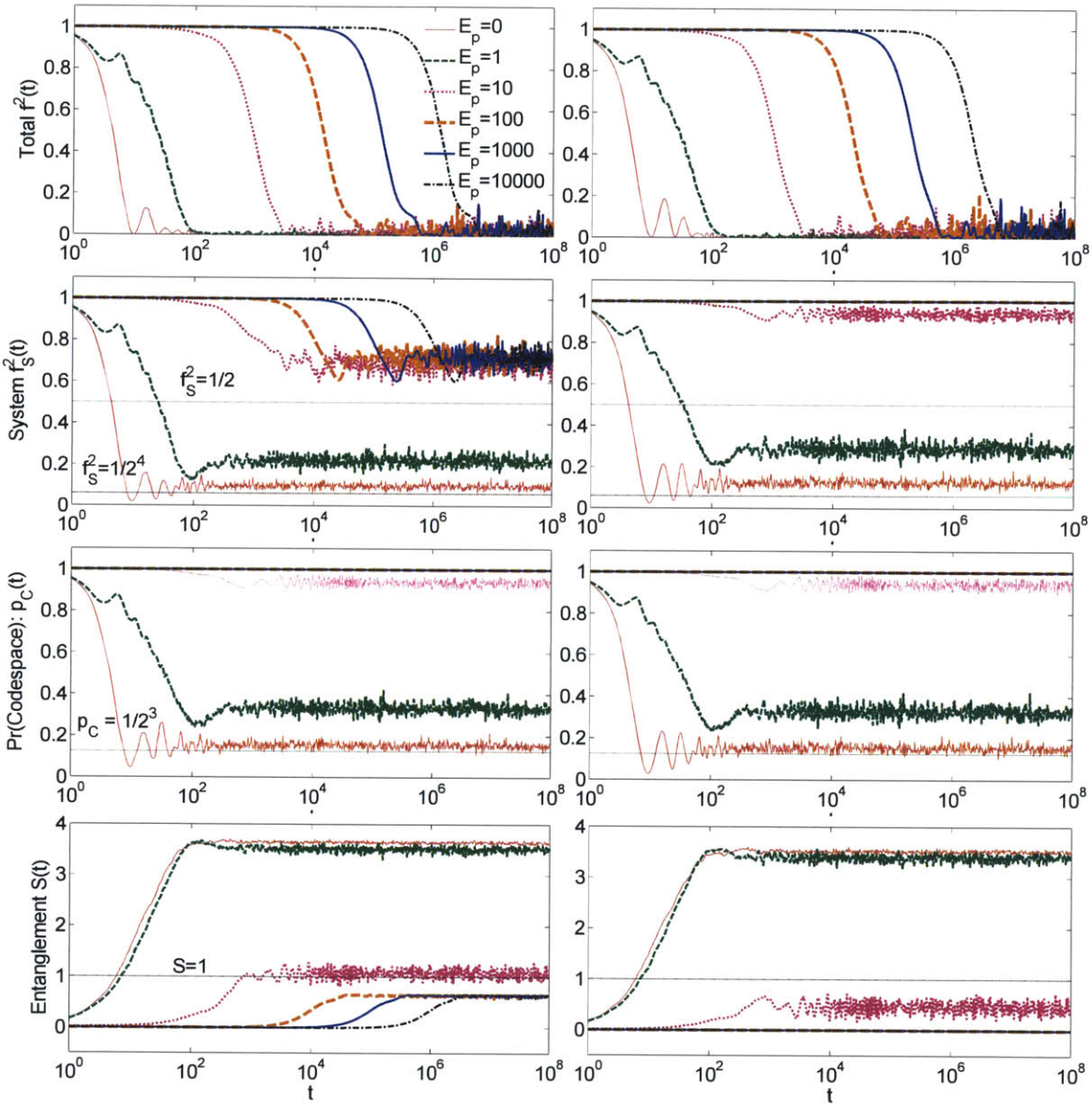


Figure 4-3: *Left:* $|\psi_{0S}\rangle$ is a random state in the codespace. *Right:* $|\psi_{0S}\rangle = (|0_L\rangle - |1_L\rangle)/\sqrt{2}$ is an eigenstate of $H_S = \bar{X}$. *From top to bottom:* For $\lambda = 0.1$, squared total fidelity $f^2(t)$, squared system fidelity $f_S^2(t)$, probability of remaining in codespace $p_C(t)$, and entanglement entropy $S(t)$ as a function of time up to $t = 10^8$ on a log-scale, with E_p of different orders of magnitude. All data are for the same instance of $(V, H_B, |\psi_B\rangle)$ as Fig. 4-2. The bath size is $n_b = 6$. The worst-case numerical precision at $t = 10^8$ is approximately 10^{-4}.

For comparison purposes, the same horizontal lines that guided the eyes in Fig. 4-2 are reproduced here. As we can see, $f_S^2(t)$ and $S(t)$ generally have different asymptotic behavior in time than the $|\psi_{0S}\rangle = |0_L\rangle$ case shown in the previous figure.

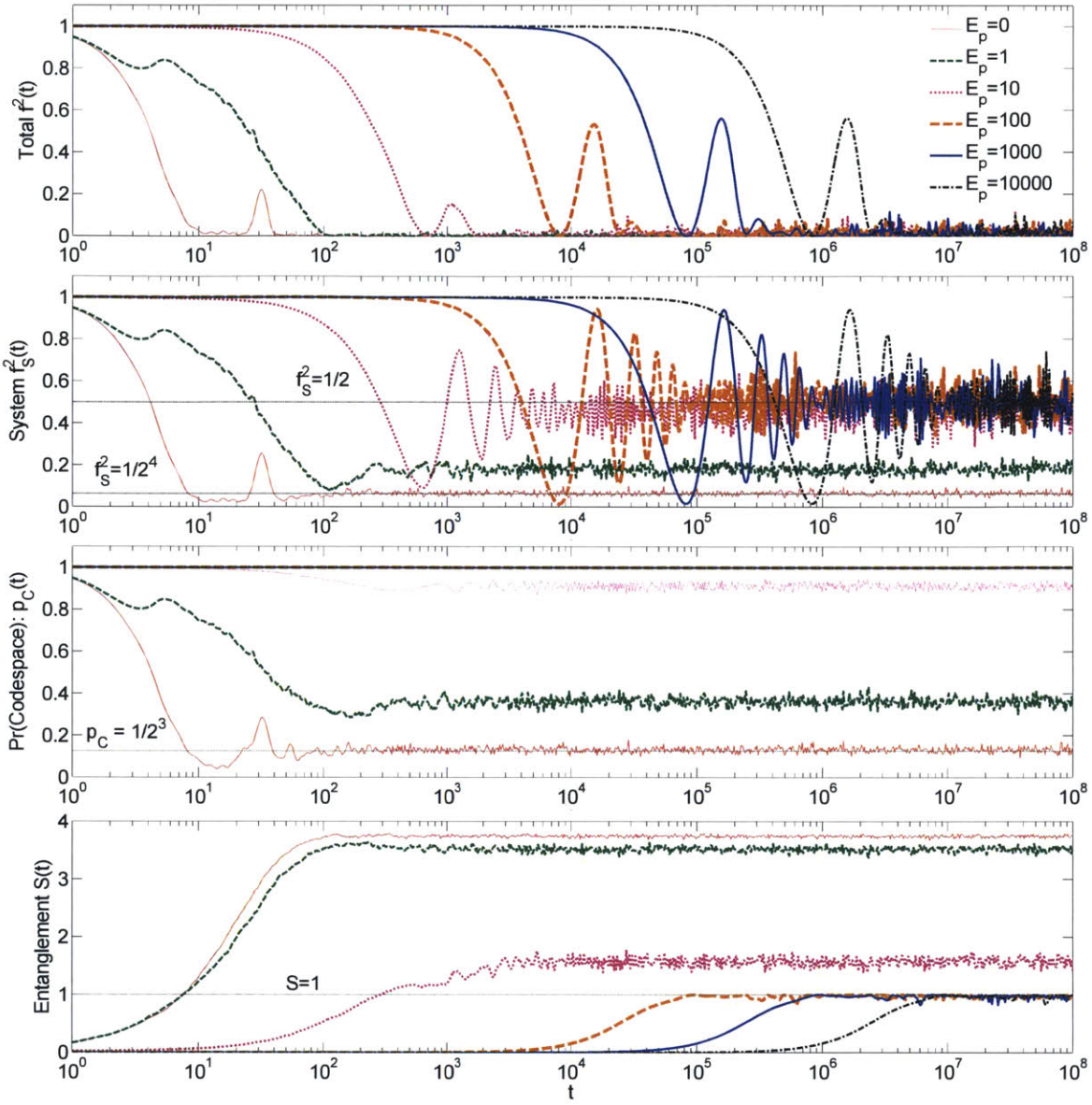


Figure 4-4: *From top to bottom:* For $\lambda = 0.1$ and $|\psi_{0S}\rangle = |0_L\rangle$, squared total fidelity $f^2(t)$, squared system fidelity $f_S^2(t)$, probability of remaining in codespace $p_C(t)$, and entanglement entropy $S(t)$ as a function of time up to $t = 10^8$ on a log-scale, with E_p of different orders of magnitude. All data are for a particular instance of $(V, H_B, |\psi_B\rangle$. The bath size is $n_b = 6$. The worst-case numerical precision at $t = 10^8$ is approximately 10^{-4}.

For guiding the eye, we have plotted a few horizontal lines at interesting values. The two lines in the $f_S^2(t)$ plot are at values $1/2$ and $1/2^4$, which are expected probabilities of measuring a particular state when we have a maximally mixed 1- or 4-qubit state, respectively. The line in the $p_C(t)$ plot is at value $p_C = 1/2^3$, the expected probability of being in the codespace (dimension 2) when the system is in a maximally mixed 4-qubit state (dimension 2^4). The line in the bottom $S(t)$ plot is at value $S = 1$, which is the entropy of a maximally mixed state in a 1-qubit system.

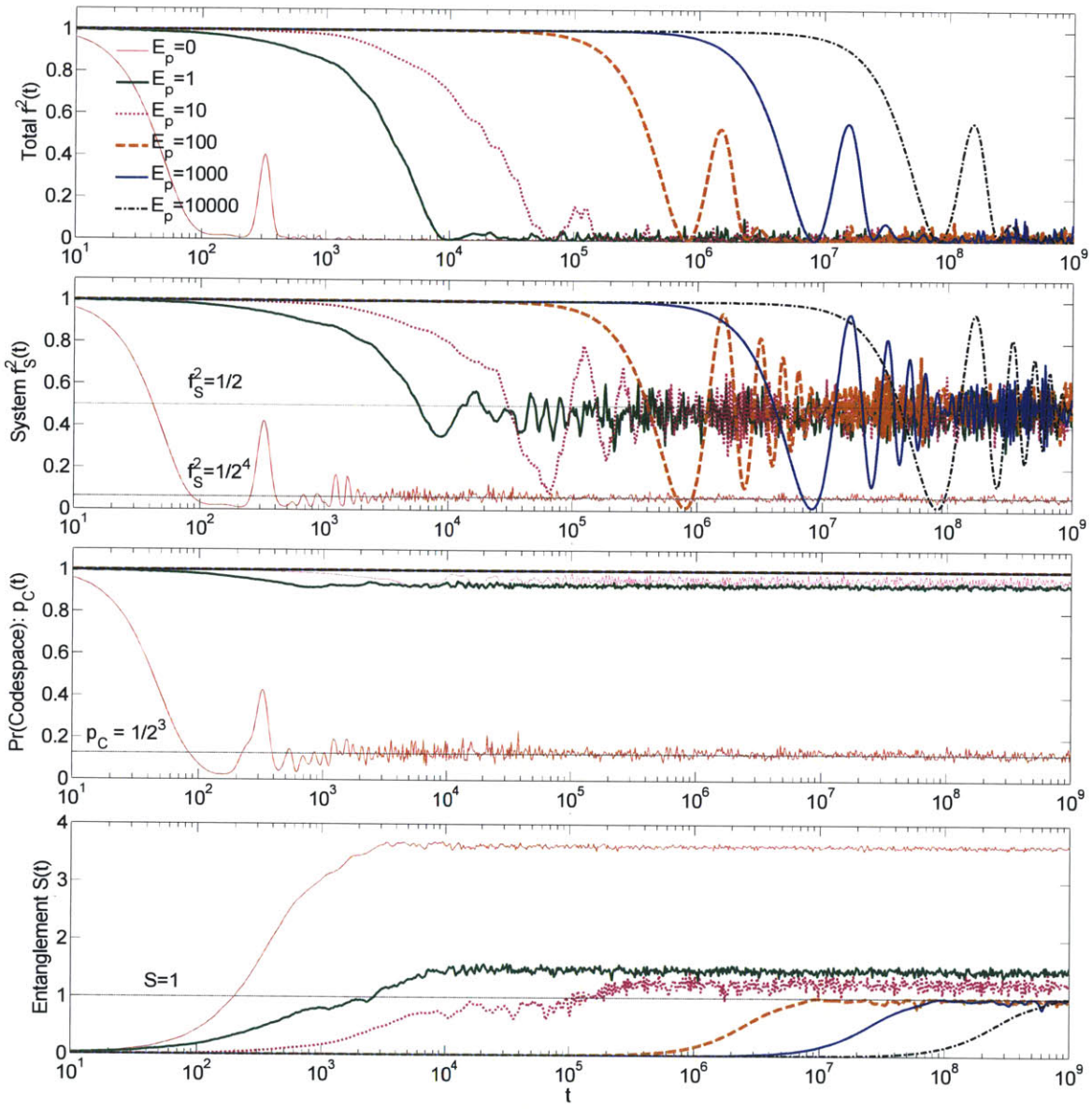


Figure 4-5: *From top to bottom:* For $\lambda = 0.01$ and $|\psi_{0S}\rangle = |0_L\rangle$, squared total fidelity $f^2(t)$, squared system fidelity $f_S^2(t)$, probability of remaining in codespace $p_C(t)$, and entanglement entropy $S(t)$ as a function of time up to $t = 10^9$ on a log-scale, with E_p of different orders of magnitude. All data are for the same instance of $(V, H_B, |\psi_B\rangle)$ as in Fig. 4-4, except with a different λ. The bath size is $n_b = 6$. The worst-case numerical precision at $t = 10^9$ is approximately 10^{-3}.

For guiding the eye, we have plotted a few horizontal lines at interesting values. The two lines in the $f_S^2(t)$ plot are at values $1/2$ and $1/2^4$, which are expected probabilities of measuring a particular state when we have a maximally mixed 1- or 4-qubit state, respectively. The line in the $p_C(t)$ plot is at value $p_C = 1/2^3$, the expected probability of being in the codespace (dimension 2) when the system is in a maximally mixed 4-qubit state (dimension 2^4). The line in the bottom $S(t)$ plot is at value $S = 1$, which is the entropy of a maximally mixed state in a 1-qubit system.

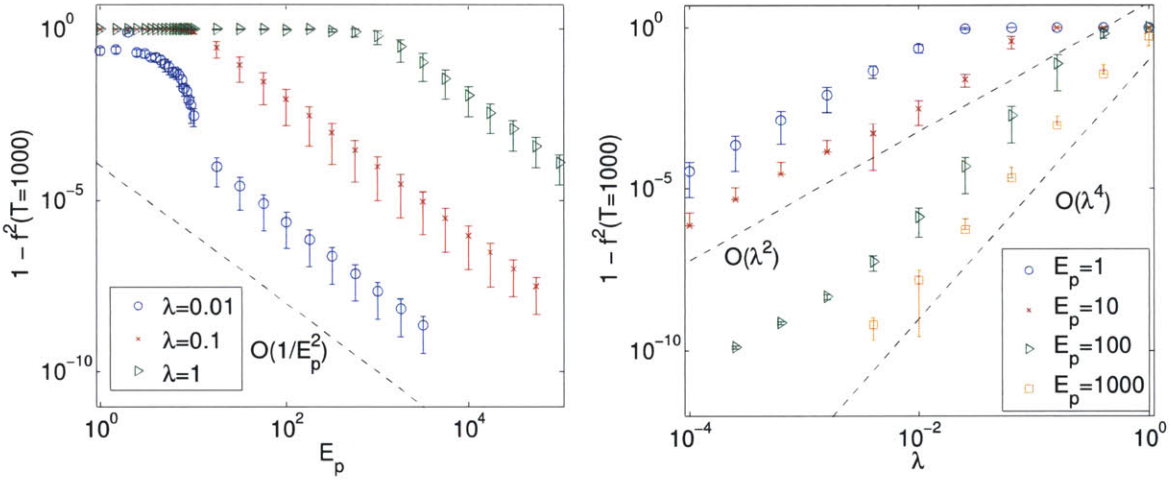


Figure 4-6: *Left:* $1 - f^2(T = 1000)$ vs. E_p for different fixed λ. *Right:* $1 - f^2(T = 1000)$ vs. λ for different fixed E_p. Size of the bath is $n_b = 6$ qubits. Each data point is the average over 10 random instances of $(V, H_B, |\psi_B\rangle)$, and the standard error is shown as error bars. We can see that in the large E_p and very small λ limit, $1 - f^2(T) = O(\lambda^2/E_p^2)$. However, for very large E_p and intermediate λ, we see $1 - f^2(T) = O(\lambda^4/E_p^2)$ instead. The dash line are guides for the eye for power-law dependences. All $1 - f^2(T)$ data points are larger than their estimated numerical precision by at least a factor of 10.

Chapter 5

Conclusion

As the interest in quantum computers grows, the ability to execute quantum algorithms fault-tolerantly, i.e. perform arbitrarily good quantum computation with noisy and limited resources, becomes an increasingly important problem. While the standard circuit model of quantum computation has been shown to be fault-tolerant with a constant threshold of error, we are also interested in the alternative model of adiabatic quantum computation due to its potential of easy, "hands-off" implementation and intrinsic robustness. We examined efforts towards fault-tolerant adiabatic quantum computation, and found the energy-gap protection through penalizing states outsize the codespace with energy E_p to be most promising for practical error-suppression in the adiabatic paradigm.

Building on the result of Jordan *et al.*[7], we were able to analytically demonstrate that the energy-gap protection works for generic Hamiltonian evolution in an open system with an arbitrary noise model, provided that the noise is detectable. Although we didn't find that the protection in general scales as $e^{-\beta E_p}$ as Jordan *et al.* illustrated for their specific photon bath noise model, we proved (Theorem 4.1) that the protection gives full fidelity of evolution in the codespace in the $E_p \to \pm\infty$ limit. Remarkably, our result is fully unitary and doesn't rely on the reduced density matrix of the system; this implies that the evolution of the bath is also protected by the energy penalty on the system. In addition, our result is applicable to not only adiabatic quantum computations but also quantum evolutions in general.

By assuming Noise and Geometric Locality, which are essentially assertions that all parts of the Hamiltonian act locally, we were able to deduce bounds on the fidelity of the evolution. For a system size of n and the magnitude of the error Hamiltonian λ, we found that by performing a λ-expansion up to $O(\lambda^2)$, the probability of error (i.e. probability of not finding exactly the desired noiseless state, or $1 - |\text{fidelity}|^2$) after evolving for time T is

$$1 - f^2(T) = O\Big(\frac{\lambda^2 n(1+cT)^2}{E_p^2}\Big) \tag{5.1}$$

for some bounded constant c that has the unit of time^{-1}. This suggests that the necessary energy-gap protection for high fidelity is $E_p \gg c\lambda T\sqrt{n}$.

Based on our numerical study of a 1-logical qubit evolving under the time-independent system Hamiltonian $H_S = \bar{X}$, encoded with the 4-qubit Jordan-Farhi-Shor stabilizer code and coupled locally to random baths of $n_b = 6$ qubits, we found the following empirical asymptotic behavior for the overall fidelity:

- Intermediate E_p: $1 - f^2(T) = O(\lambda^2 T^2 / E_p^2)$, which is essentially Eq. (5.1);
- Very large E_p: $1 - f^2(T) = O(\lambda^4 T^2 / E_p^2)$.

which indicates, along with other reasons, that the λ-expansion is not necessarily valid in general, and a better bound may be potentially obtained than Eq. (5.1).

In our numerical simulations, we also found that the energy-gap protection for eigenstates of the time-independent H_S is much better than expected. In fact, the fidelity of the evolution within the system remain essentially 1 even as the overall fidelity decays towards 0. This indicates that the energy-gap protection can be better than imagined for protecting adiabatic quantum computation, where the states are expected to be close to the instantaneous eigenstates of the time-dependent system Hamiltonian. Nevertheless, no conclusion cannot be comfortably drawn without further numerical study on energy-gap protected adiabatic algorithms, which are generally much harder to simulate due to the inability to solve their time-dependent Hamiltonians exactly.

Appendix A

Generalization Beyond Local Noise Model

In this appendix, we give mathematical reasoning that allows the generalization of our result beyond the local noise model of Eq. (4.4) to the more general noise model of Eq. (4.3), provided the error-detection condition holds

$$PVP = 0 \tag{A.1}$$

First of all, observe that in general, we can write V as

$$V = \sum_{k=1}^{n} W_i = \sum_i W_{1,i} + \sum_{i,j} W_{2,i}W_{2,j} + \ldots + W_{n,1}W_{n,2}\cdots W_{n,n} \tag{A.2}$$

$$\text{where} \quad W_k = \sum_{i_1,\ldots,i_k} W_{n,i_1}W_{n,i_2}\cdots W_{n,i_k} \tag{A.3}$$

where the index k indicates the locality in the logical qubits of the term, and i and j are the indices of the logical subsystem on which the corresponding term acts. Here, the multiplications are understood to be tensor products. We interpret W_k as the sum of the terms in V that acts on k logical subsystems (or k-local on the logical qubits), consisting of tensor products of k nontrivial terms $W_{k,i}$ each acting only on the i^{th} logical subsystem. Because of the generic error detection condition of Eq. (A.1) must hold, we have for all k

$$P_jW_{k,i}P = \begin{cases} 0 & i = j \\ W_{k,i}P & i \neq j \end{cases} \quad \text{and} \quad Q_jW_{k,i}P = \begin{cases} W_{k,i}P & i = j \\ 0 & i \neq j \end{cases} \tag{A.4}$$

Our earlier Noise Locality assumption is simply the belief that $V = W_1$. The place we used Noise Locality is Eq. (4.21), where we extracted a factor of e^{iE_pt} from $U_P^\dagger VP$ so we can applied Riemann-Lebesgue Lemma to the expression of $F(T)$ in Eq. (4.22). It turns out that we can reproduce the factor of e^{iE_pt} without assuming Noise Locality.

To that end, let us now define $Q \equiv I - P$. Note

$$QVP = VP \tag{A.5}$$

Since for each logical subsystem, its identity operation can be written as $I_i = P_i + Q_i$, we

can write out the explicit expression for Q as

$$\begin{aligned} Q &= I - P = \prod_{i=1}^{n}(P_i + Q_i) - \prod_{i=1}^{n} P_i \\ &= \prod_{i=1}^{n} P_i + \sum_{j=1}^{n} Q_j \prod_{i\neq j} P_i + \sum_{j<m}^{n} Q_j Q_m \prod_{i\neq j,m} P_i + \ldots + \prod_{i=1}^{n} Q_i - \prod_{i=1}^{n} P_i \\ &= \sum_{j=1}^{n} Q_j \prod_{i\neq j} P_i + \sum_{j<m}^{n} Q_j Q_m \prod_{i\neq j,m} P_i + \ldots + \prod_{i=1}^{n} Q_i \\ &= R_1 + R_2 + \ldots + R_n \end{aligned}$$

where R_h are the sum of all $\binom{n}{h}$ possible tensor products with h Q_i's and $(n-h)$ P_i's. Intuitively, R_h is the group of "h-local" terms in Q, with h Q_i's that act nontrivially on h of the logical subsystems. We can interpret Q_i and a nontrivial and P_i as a trivial operation on the codespace due to $Q_i P = 0$ and $P_i P = P$. Mathematically, R_h's are defined as

$$\begin{aligned} R_1 &= \sum_{j=1}^{n} Q_j \prod_{i\neq j} P_i = Q_1 P_2 P_3 \cdots P_n + \ldots + P_1 P_2 \cdots P_{n-1} Q_n \\ R_2 &= \sum_{j<m}^{n} Q_j Q_m \prod_{i\neq j,m} P_i = Q_1 Q_2 P_3 P_4 P_5 \cdots P_n + Q_1 P_2 Q_3 P_4 P_5 \cdots P_n + \ldots \\ &\cdots \\ R_n &= \prod_{i=1}^{n} Q_i \end{aligned}$$

Observe that

$$\begin{aligned} U_P^\dagger R_1 &= \prod_{i=1}^{n}(e^{iE_p t} Q_i + P_i) \cdot \sum_{j=1}^{n} Q_j \prod_{m\neq j} P_m = \sum_{j=1}^{n} e^{iE_p t} Q_j \prod_{m\neq j} P_m = e^{iE_p t} \sum_{j=1}^{n} Q_j \prod_{m\neq j} P_m \\ &= e^{iE_p t} R_1 \\ U_P^\dagger R_2 &= \prod_{i=1}^{n}(e^{iE_p t} Q_i + P_i) \cdot \sum_{j<m}^{n} Q_j Q_m \prod_{r\neq j,m} P_r = \sum_{j<m}^{n} e^{i2E_p t} Q_j Q_m \prod_{r\neq j,m} P_r \\ &= e^{i2E_p t} R_2 \\ &\cdots \\ U_P^\dagger R_n &= \prod_{i=1}^{n}(e^{iE_p t} Q_i + P_i) \cdot \prod_{j=1}^{n} Q_j = e^{inE_p t} R_n \end{aligned}$$

It's not hard to see that $U_P^\dagger R_h = e^{ihE_p t} R_h$ for all h. Now, using the conditions of Eq. (A.4), we can see that

$$R_h W_k P = \delta_{hk} W_k P \tag{A.6}$$

This conclusion can be reached by looking into the cases:

1. When $h > k$, there will always be an extra Q_i from R_h in any term of the expanded product expression of $R_h W_k P$ that doesn't match up with a $W_{k,i}$. As a result, Q_i annihilates the expression since Q_i commutes with all $W_{k,j}$ when $j \neq i$, and thus we

can move it to the rightmost part of the expression and use $Q_i P = 0$.

2. When $h < k$, there will always be an extra $W_{k,i}$ in any term of the expanded product expression of $R_h W_k P$ that doesn't match up with a Q_i from R_h, and will instead be paired with an P_i which annihilates the expression due to $P_i W_{k,i} P = 0$

3. When $h = k$, we recover $W_k P$ since we can pair each term in R_h with a term in W_k, so that the Q_i's in R_h act on the corresponding nontrivial $W_{k,i}$ in some term of W_k. Note the cross terms, where a term in R_h multiplies a term in W_k and their Q_i's and $W_{k,j}$'s don't match in terms of the locations of the logical subsystems, vanishes because any mismatch results in a direct multiplication of $Q_i P = 0$.

Hence, analogous to Eq. (4.21), we obtain the following expression

$$\begin{aligned} U_P^\dagger V P &= U_P^\dagger Q V P = U_P^\dagger \sum_{h=1}^{n} R_h \sum_{k=1}^{n} W_k P = \sum_{h,k=1}^{n} e^{ihE_p t} R_h W_k P \\ &= \sum_{h,k=1}^{n} e^{ihE_p t} \delta_{hk} W_k P = \sum_{k=1}^{n} e^{ikE_p t} W_k P \end{aligned} \tag{A.7}$$

Then, analogous to Eq. (4.22), we can simplify $F(T)$ as

$$F(T) = \int_0^T V_I(t) P dt = \sum_{k=1}^{n} \int_0^T e^{ikE_p t} U_0^\dagger(t) W_k(t) U_0(t) P dt \tag{A.8}$$

where for each term we can apply the Riemann-Lebesgue Lemma. Note if we integrate by parts, each term gets a protection by at least E_p, and the k-local term gets protection by kE_p.

Comment We didn't have to go through all the above trouble to generalize our result of the utility of the energy-gap protection if we had chosen a nonlocal energy penalty Hamiltonian

$$H_P = Q = I - P \tag{A.9}$$

Since Q is a projector (onto $C^\perp$) unlike $\mathcal{Q}$, this would give us the result

$$U_P^\dagger V P = U_P^\dagger Q V P = e^{iE_p Q t} Q V P = e^{iE_p t} Q V P = e^{iE_p t} V P \tag{A.10}$$

Although this would make our calculations far simpler, the requisite nonlocality in our Hamiltonian makes this unappealing for practical implementations.

Bibliography

[1] D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev. Adiabatic quantum computation is equivalent to standard quantum computation. *SIAM Journal of Computing*, 37:166–194, 2007.

[2] A. M. Childs, E. Farhi, and J. Preskill. Robustness of adiabatic quantum computation. *Phys. Rev. A*, 65:012322, Dec 2001.

[3] D. Deutsch. Quantum theory, the church-turing principle and the universal quantum computer. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 400(1818):97–117, 1985.

[4] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser. Quantum computation by adiabatic evolution. 2000. arXiv:quant-ph/0001106.

[5] R. P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6-7):467–488, 1982.

[6] S. P. Jordan. *Quantum Computation Beyond the Circuit Model.* PhD dissertation, Massachusetts Institute of Technology, Department of Physics, May 2008.

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

[8] J. Kempe, A. Kitaev, and O. Regev. The complexity of the local hamiltonian problem. *SIAM Journal on Computing*, 35(5):1070–1097, 2006.

[9] D. A. Lidar. Towards fault tolerant adiabatic quantum computation. *Phys. Rev. Lett.*, 100:160506, Apr 2008.

[10] A. Messiah. *Quantum Mechanics.* John Wiley & Sons, New York, 1958.

[11] D. Nagaj and S. Mozes. New construction for a qma complete three-local hamiltonian. *Journal of Mathematical Physics*, 48(7):–, 2007.

[12] M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press, Cambridge, UK, 2000.

[13] J. Preskill. Reliable quantum computers. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1969):385–410, 1998.

[14] B. W. Reichardt. The quantum adiabatic optimization algorithm and local minima. In *Proceedings of the Thirty-sixth Annual ACM Symposium on Theory of Computing*, STOC '04, pages 502–510, New York, NY, USA, 2004. ACM.

[15] J. Roland and N. J. Cerf. Quantum search by local adiabatic evolution. *Phys. Rev. A*, 65:042308, Mar 2002.

[16] J. Roland and N. J. Cerf. Noise resistance of adiabatic quantum computation using random matrix theory. *Phys. Rev. A*, 71:032330, Mar 2005.

[17] P. W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Foundations of Computer Science, 1994 Proceedings., 35th Annual Symposium on*, pages 124–134, Nov 1994.

[18] L. Viola, E. Knill, and S. Lloyd. Dynamical decoupling of open quantum systems. *Phys. Rev. Lett.*, 82:2417–2421, Mar 1999.

[19] 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, Nov 2013.