

# Quantum Error Correction with Biased Noise

Thesis by  
Peter Brooks

In Partial Fulfillment of the Requirements  
for the Degree of  
Doctor of Philosophy



California Institute of Technology  
Pasadena, California

2013  
(Defended May 24, 2013)

**Copyright notice.** Chapters 2 and 3 contain material from [1] and [2], respectively, which are copyrighted by the American Physical Society.

The remaining material is

© 2013

Peter Brooks

All Rights Reserved

To Kimi.

# Acknowledgements

I would first like to thank my advisor, John Preskill. John has been an exemplary mentor and role model, with a broad and deep knowledge of physics, an insightful mind, and a ready smile. I could not have asked for a better advisor, and I am immensely proud to have been his student.

I am been extremely lucky to have been blessed with a number of extraordinarily thoughtful mentors who have guided my journey through physics and life. I am especially grateful for guidance and support from Alexei Kitaev, Jim Harrington, Pat Burchat, Ian Fisher, Blas Cabrera and Fritz DeJongh. Most of all I would like to thank Bruce Keyzer, whose inspiring teaching instilled in me an incurable passion for physics and set this entire course in motion.

I owe much of what I have learned to conversations with collaborators and friends, especially Adam Paetznick, Cody Jones, Jeongwan Haah, Isaac Kim, and the entire IQI(M!) group at Caltech. And I could not have survived a Ph.D. without the help of Ann Harvey, to whom I am eternally grateful.

I could not have survived either without the support of my family and friends, especially my parents, Helen and Doug, my brothers, James and Will, and my loving girlfriend Kimi, the best thing ever to happen to me. Their loving support has made all of this possible.

# Abstract

Quantum computing offers powerful new techniques for speeding up the calculation of many classically intractable problems. Quantum algorithms can allow for the efficient simulation of physical systems, with applications to basic research, chemical modeling, and drug discovery; other algorithms have important implications for cryptography and internet security.

At the same time, building a quantum computer is a daunting task, requiring the coherent manipulation of systems with many quantum degrees of freedom while preventing environmental noise from interacting too strongly with the system. Fortunately, we know that, under reasonable assumptions, we can use the techniques of quantum error correction and fault tolerance to achieve an arbitrary reduction in the noise level.

In this thesis, we look at how additional information about the structure of noise, or “noise bias,” can improve or alter the performance of techniques in quantum error correction and fault tolerance. In Chapter 2, we explore the possibility of designing certain quantum gates to be extremely robust with respect to errors in their operation. This naturally leads to structured noise where certain gates can be implemented in a protected manner, allowing the user to focus their protection on the noisier unprotected operations.

In Chapter 3, we examine how to tailor error-correcting codes and fault-tolerant quantum circuits in the presence of dephasing biased noise, where dephasing errors are far more common than bit-flip errors. By using an appropriately asymmetric code, we demonstrate the ability to improve the amount of error reduction and decrease the physical resources required for error correction.

In Chapter 4, we analyze a variety of protocols for distilling magic states, which enable universal quantum computation, in the presence of faulty Clifford operations. Here again there is a hierarchy of noise levels, with a fixed error rate for faulty gates, and a second rate for errors in the distilled states which decreases as the states are distilled to better quality. The interplay of these different rates sets limits on the achievable distillation and how quickly states converge to that limit.

# Contents

<b>Acknowledgements</b>	<b>iv</b>
<b>Abstract</b>	<b>v</b>
<b>1 Background</b>	<b>1</b>
1.1 Overview . . . . .	1
1.2 Quantum computation . . . . .	3
1.2.1 The Pauli operators . . . . .	6
1.2.2 The Clifford group . . . . .	7
1.2.3 Universal computation . . . . .	9
1.3 Error correcting codes . . . . .	10
1.3.1 Classical error correction . . . . .	10
1.3.2 CSS codes . . . . .	12
1.3.3 Stabilizer codes and subsystem codes . . . . .	13
1.4 Fault tolerance . . . . .	14
1.4.1 Syndrome extraction techniques . . . . .	14
<b>2 Protected Gates for Superconducting Qubits</b>	<b>17</b>
2.1 Introduction . . . . .	17
2.2 The $0-\pi$ qubit . . . . .	19
2.2.1 Achieving superinductance . . . . .	23
2.2.2 Measurements . . . . .	24
2.3 Phase gate . . . . .	24
2.3.1 Ideal protected phase gate . . . . .	27

2.3.2	Effects of protected phase gate imperfections . . . . .	30
2.4	Sketch of the error estimate . . . . .	33
2.5	Encoding a qubit in an oscillator . . . . .	37
2.6	Imperfect grid states . . . . .	41
2.6.1	Bit-flip and phase errors . . . . .	42
2.6.2	Gate error estimate . . . . .	44
2.6.3	Gaussian case . . . . .	46
2.7	Diabatic error . . . . .	49
2.8	Squeezing error . . . . .	54
2.9	Simulations . . . . .	57
2.10	Nonzero temperature . . . . .	59
2.11	Perturbative stability . . . . .	64
2.12	Universality . . . . .	67
2.13	Conclusions . . . . .	70
Appendix 2.A	Adiabatic switch . . . . .	72
Appendix 2.B	Quantifying the gate error . . . . .	76
2.B.1	Fidelity . . . . .	78
2.B.2	Trace norm . . . . .	79
2.B.3	Kraus operators . . . . .	79
2.B.4	Diamond norm . . . . .	81
Appendix 2.C	Grid states . . . . .	82
2.C.1	Approximate codewords in $\varphi$ space and $Q$ space . . . . .	83
2.C.2	Phase gate overrotation . . . . .	85
2.C.3	Imaginary part of overrotation error . . . . .	86
Appendix 2.D	Poisson summation formula . . . . .	88
Appendix 2.E	Diabatic transitions in a two-level system . . . . .	89
<b>3</b>	<b>Asymmetric Bacon-Shor Gadgets for Biased Noise</b>	<b>93</b>
3.1	Introduction . . . . .	93
3.2	Dephasing biased noise . . . . .	95
3.3	Bias-compatible computation . . . . .	98

3.4	Bacon-Shor codes . . . . .	100
3.5	Fault-tolerant gadgets . . . . .	101
3.5.1	Logical $X$ measurement gadget . . . . .	102
3.5.2	Logical $Z$ measurement gadget . . . . .	102
3.5.3	Cat state preparation gadget . . . . .	105
3.5.4	$ +\rangle^L$ preparation gadget . . . . .	106
3.5.5	$ 0\rangle^L$ preparation gadget . . . . .	106
3.5.6	$\text{CNOT}^L$ gadget . . . . .	107
3.5.7	Scheduling . . . . .	109
3.5.8	Injection by teleportation . . . . .	111
3.6	Effective error strength . . . . .	111
3.6.1	Measurement failure . . . . .	112
3.6.2	Cat state preparation failure . . . . .	114
3.6.3	Shorter ancillas . . . . .	117
3.7	Geometrically local circuits . . . . .	118
3.8	Results for logical CSS gates . . . . .	122
3.8.1	Unrestricted gates . . . . .	122
3.8.2	Geometrically local gates . . . . .	124
3.8.3	Geometrically local gates and measurement bias . . . . .	126
3.9	State injection . . . . .	127
3.10	State distillation . . . . .	131
<b>4</b>	<b>Magic State Distillation with Noisy Clifford Gates</b>	<b>134</b>
4.1	Introduction . . . . .	134
4.2	Magic states . . . . .	136
4.3	$S$ -symmetric codes . . . . .	137
4.3.1	Doubly even codes . . . . .	138
4.3.2	$ i\rangle$ distillation . . . . .	140
4.4	$T$ -symmetric codes . . . . .	143
4.4.1	Triply even codes . . . . .	144
4.4.2	$ T\rangle$ distillation . . . . .	145

4.5	Triorthogonal codes . . . . .	146
4.5.1	Explicit triorthogonal codes . . . . .	149
4.5.2	$ T\rangle$ distillation from triorthogonal codes . . . . .	151
4.6	$H$ -symmetric codes . . . . .	151
4.6.1	$ H\rangle$ distillation . . . . .	152
4.7	Effects of noisy Clifford operations . . . . .	154
4.7.1	Error model . . . . .	155
4.7.2	Distillation circuit design . . . . .	156
4.7.3	Controlled- $H$ analysis . . . . .	156
4.7.4	Reed-Muller $ T\rangle$ distillation . . . . .	159
4.7.5	Bravyi-Haah triorthogonal $ T\rangle$ distillation . . . . .	165
4.7.6	Jones $ H\rangle$ distillation . . . . .	166
4.7.7	Comparison . . . . .	168
4.8	Discussion . . . . .	175
<b>5</b>	<b>Epilogue</b>	<b>176</b>
	<b>Bibliography</b>	<b>178</b>

# List of Figures

1.1	Transversal $H$ and CNOT gates. . . . .	15
1.2	Preparation of a verified 4-qubit cat state. . . . .	16
1.3	Steane error correction. . . . .	16
2.1	The energy $E(\theta)$ of the $0-\pi$ qubit. . . . .	19
2.2	Two-rung superconducting circuit underlying the $0-\pi$ qubit. . . . .	20
2.3	The circuit for the $0-\pi$ qubit. . . . .	23
2.4	An unprotected phase gate. . . . .	25
2.5	A protected phase gate. . . . .	26
2.6	An effective Josephson junction. . . . .	26
2.7	The profile of the tunable Josephson coupling $J(t)$ in the execution of the protected phase gate. . . . .	27
2.8	Coupling the qubit to the oscillator prepares a grid state in $\varphi$ space, a superposition of narrowly peaked functions governed by a broad envelope function. . . . .	28
2.9	Ideal codewords of the continuous variable code. . . . .	38
2.10	Approximate codewords of the continuous variable code. . . . .	40
2.11	The estimated gate error $ \eta_\varepsilon $ (on a log scale) as a function of the rotation error $\varepsilon$ , for $\kappa^{-2} = 40$ . . . . .	50
2.12	The estimated gate error $ \eta_\varepsilon $ (on a log scale) as a function of the rotation error $\varepsilon$ , for $\kappa^{-2} = 80$ . . . . .	51
2.13	The numerically computed and analytic prediction for the diamond-norm deviation from the ideal phase gate, for $\kappa^{-2} = \sqrt{L/C} = 80$ , $\Delta^{-2} = \sqrt{J_0 C} = 8$ , and $\tau_J/C = 80$ . . . . .	58
2.14	The minimum diamond-norm deviation of the protected phase gate from the ideal gate, as a function of the initial oscillator eigenstate $n$ , for $\sqrt{L/C} = 80$ and $\sqrt{J_0 C} = 8$ . . . . .	63

2.15	The minimum diamond-norm deviation of the protected phase gate from the ideal gate, as a function of the oscillator's anharmonicity parameter $\alpha = \lambda L \sqrt{L/C}$ . Here, as in Figure 2.13, $\sqrt{L/C} = 80$ and $\sqrt{J_0 C} = 8$ . . . . .	66
2.16	Logical CNOT gate acting on two blocks of the repetition code, shown here for code length $n = 3$ . . . . .	69
2.17	Adiabatic switch. For $\sqrt{L/C} \gg 1$ , the effective Josephson energy $J_{\text{eff}}$ of the switch is depends sensitively on circuit parameters. . . . .	72
2.18	The inverse effective capacitance of a Josephson junction as a function of $JC$ . . . . .	75
2.19	The effective Josephson parameter of the adiabatic switch as a function of $JC$ for $\sqrt{L/C} = 40$ . . . . .	77
3.1	Bacon-Shor qubits arranged in an $n \times m$ lattice. . . . .	100
3.2	Bacon-Shor stabilizers. . . . .	100
3.3	$Z$ measurement performed with $X$ measurement and an ancilla. . . . .	102
3.4	Gadget for a single row of the non-destructive $\mathcal{M}_{\frac{1}{2}}^L$ measurement . . . . .	104
3.5	Four versions of the CNOT gadget. . . . .	107
3.6	One-bit teleportation circuit. . . . .	107
3.7	Teleported CNOT gate. . . . .	108
3.8	Arrangement of Bacon-Shor qubits for geometrically local computation for $3 \times 5$ code block. . . . .	119
3.9	Optimal CNOT logical error rate versus physical error rate for various values of the bias. . . . .	123
3.10	Optimal CNOT logical error rate versus required number of physical gates for a physical error rate $\varepsilon = 10^{-4}$ at various values of the bias. . . . .	125
3.11	Optimal CNOT logical error rate versus physical error rate for various values of the bias, where two-qubit physical gates are assumed to be geometrically local. . . . .	126
3.12	Optimal CNOT logical error rate versus required number of physical gates for geometrically local asymmetric Bacon-Shor codes at various values of the bias, for a physical error rate $\varepsilon = 10^{-4}$ . . . . .	127

3.13	Bacon-Shor block size vs optimal (in number of gates) CNOT logical error rate for geometrically local asymmetric Bacon-Shor codes at various values of the bias, for a physical error rate $\varepsilon = 10^{-4}$ . . . . .	128
3.14	Optimal CNOT logical error rate versus required number of gates for geometrically local asymmetric Bacon-Shor codes, assuming a physical error rate $\varepsilon = 10^{-4}$ , bias $10^4$ , and various rates of the measurement error rate $\varepsilon_{\mathcal{M}}$ . . . . .	129
3.15	Injection error probability for the optimal geometrically local Bacon-Shor codes . . . .	130
3.16	Output error probability after $n$ levels of $ T\rangle$ distillation, with the values of $\varepsilon_{\text{in}}$ and $\varepsilon_{\text{CSS}}$ taken from the optimal values of Figure 3.15, for $\varepsilon = 10^{-4}$ and various biases. . .	132
4.1	Distillation circuit for the state $ i\rangle$ and equivalent circuit transformations. . . . .	141
4.2	Alternate distillation circuit for the state $ i\rangle$ . . . . .	141
4.3	Teleported circuits for the Clifford gates $Q$ and $S$ . . . . .	143
4.4	Distillation circuit for the state $ T\rangle$ and equivalent circuit transformations. . . . .	145
4.5	Teleported circuit for the non-Clifford gates $T$ . . . . .	146
4.6	Distillation circuit for the state $ H\rangle$ . . . . .	152
4.7	Circuit consuming two copies of the magic state $ H\rangle$ to perform a controlled- $H$ gate. .	153
4.8	Output error rate $\varepsilon_{\text{out}}$ from Equation (4.92), versus input error rate $\varepsilon_{\text{in}}$ for distillation based on the $[[15,1,3]]$ Reed-Muller code, for three different values of the Clifford operation failure rate $p$ . The dashed line represents the result of no distillation. . . .	163
4.9	Output error rate $\varepsilon_{\text{out}}$ from Equation (4.92) after $n$ rounds of distillation, for distillation based on the $[[15,1,3]]$ Reed-Muller code and various values of the Clifford operation failure rate $p$ , starting from a near-threshold input error rate of $\varepsilon^{(0)} = 0.15$ . . . .	164
4.10	Number of ways to combine one or two non-transverse faults and up to one additional non-transverse fault to create an encoded error on a particular output qubit, averaged over choice of output qubit, for various Bravyi-Haah code sizes $k$ . . . . .	167
4.11	Number of ways to combine one or two non-transverse faults and up to one additional non-transverse fault to create an encoded error on a particular output qubit, averaged over choice of output qubit, for various Jones transverse $H$ code sizes $k$ . . . . .	169

- 4.12 Output error rate  $\varepsilon_{\text{out}}$  from Equations (4.94) and (4.96) versus input error rate  $\varepsilon_{\text{in}}$  for distillation based on various Jones and Bravyi-Haah codes, for  $p = 10^{-12}$ . The dashed line represents the result of no distillation. . . . . 170
- 4.13 Output error rate  $\varepsilon_{\text{out}}$  from Equations (4.94) and (4.96) versus input error rate  $\varepsilon_{\text{in}}$  for distillation based on various Jones and Bravyi-Haah codes, for  $p = 10^{-4}$ . The dashed line represents the result of no distillation. . . . . 171
- 4.14 Output error probability  $\varepsilon_{\text{out}}$  as a function of number of rounds of distillation, for five different distillation codes, starting from  $\varepsilon_{\text{in}} = 0.01$  and with  $p = 10^{-10}$ . . . . . 172
- 4.15 Distillation limit as a function of the Clifford error rate, for various distillation procedures 173

# Chapter 1

# Background

*Plans are worthless, but planning is everything.*

–Dwight D. Eisenhower [3]

This chapter aims to give a brief introduction to the terminology and techniques of quantum computing, error correction and fault tolerance, with particular focus on material relevant for the rest of the thesis. For a fuller overview see [4–6].

## 1.1 Overview

It was first noted by Feynman [7, 8] and Deutsch [9, 10] that simulating quantum systems on a classical computer seems to require a calculation whose time scales exponentially with the size of the simulated system; whereas the nature is able to efficiently “simulate” such a system by the system itself. Therefore a computer designed to take full advantage of laws of nature, in particular the laws of quantum mechanics, might offer a fundamentally more powerful mode of computation than a classical computer, offering a counter-argument to the strong Church–Turing hypothesis [11, 12]. This intuition was further strengthened by the discovery of quantum algorithms with no classical analogue, especially by Shor’s discovery of polynomial-time algorithms for finding the factors of prime numbers and computing discrete logarithms [13]. Since then, more and more quantum algorithms have been found, with applications ranging from searches of unstructured databases [14] to simulations of quantum field theories [15, 16] and molecular properties [17–19].

A quantum computer is an interference device, a baroque elaboration of the classic double-slit experiment. The quantum bits of a quantum computer can be put into a superposition of

many different states, and, by a carefully designed pattern of interference, paths corresponding to the correct answer to the calculation interfere constructively while paths corresponding to wrong answers interfere destructively and cancel out to some degree. In many cases, the entanglement of different subsystems of the quantum computer with each other plays a key role in quantum algorithms.

A quantum state is a far more fragile state than its classical counterpart. It can be destroyed not only by disturbing the system directly, but also by measuring the system. Just as the interference fringes of the double-slit experiment are destroyed if it is possible to learn which path the particle has passed through, the interference necessary for a quantum calculation can be destroyed if information about the state of the computer leaks into its environment. Nevertheless, the theory of quantum error correcting codes and quantum fault tolerance have shown that, in principle, a sufficiently isolated quantum calculation can be protected against errors arising from its environment to an arbitrary degree. This is achieved by encoding the state of the calculation non-locally in a quantum error correcting code. Once encoded, the environment cannot access the computational state without measuring an extensive number of bare systems.

In this thesis, I look at how additional information about the structure of noise, or “noise bias,” can improve or alter the performance of techniques in quantum error correction and fault tolerance. In Chapter 2, I explore the possibility of designing qubits and quantum gates to be extremely robust with respect to errors in their operation. This naturally leads to structured noise where certain gates can be implemented in a protected manner, allowing the user to focus their protection on the noisier unprotected operations. In Chapter 3, I examine how to tailor error-correcting codes and fault-tolerant quantum circuits in the presence of dephasing biased noise, where dephasing errors are far more common than bit-flip errors. By using an appropriately asymmetric code, I demonstrate the ability to improve the amount of error reduction and decrease the physical resources required for error correction. In Chapter 4, I analyze a variety of protocols for distilling magic states, which enable universal quantum computation, in the presence of faulty Clifford operations. Here again there is a hierarchy of noise levels, with a fixed error rate for faulty gates, and a second rate for errors in the distilled states which decreases as the states are distilled to better quality. The interplay of these different rates sets limits on the achievable distillation and how quickly states converge to that limit.

## 1.2 Quantum computation

The state of a classical computer is a string of *bits*, each of which can take one of two values, 0 or 1. Equivalently, its state is a vector over the binary field  $\mathbb{Z}_2$ . A *computation* is a process which maps bit strings to other bit strings, either in a deterministic way, or with some probability. We can decompose any classical computation into a sequence of fundamental operations, called gates, and a connectivity diagram, a wiring, indicating how the outputs of one gate are mapped into the inputs of future gates.

The fundamental unit of quantum information is the quantum bit, or qubit. Unlike its classical analogue, the bit, which can take two values, the qubit can be described by a vector

$$|\psi\rangle = a|0\rangle + b|1\rangle, \quad (1.1)$$

where the components  $a$  and  $b$  are both complex numbers, called amplitudes. This state is called normalized if

$$\| |\psi\rangle \|^2 = \langle \psi | \psi \rangle = |a|^2 + |b|^2, \quad (1.2)$$

the norm of the qubit, is equal to 1.

We can perform a measurement on our qubit, which returns a classical bit, 0 or 1. The “logical basis” or “computational basis” measurement returns the value 0 with probability  $|a|^2$  and the value 1 with probability  $|b|^2$ . After the measurement, if we received the result 0, then we will find our qubit in the state  $|0\rangle$ , and if we received the result 1, then we will find our qubit in the state  $|1\rangle$ . In effect, the measurement projects the state of the qubit onto the two basis states,  $|0\rangle$  and  $|1\rangle$ . The normalization condition means that we will always find our qubit in one of the two states.

We can also perform a measurement in a different basis. One important basis is the “dual basis” measurement, which projects onto the two states

$$|+\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \quad (1.3)$$

$$|-\rangle = \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle), \quad (1.4)$$

where we will receive the first outcome with probability  $|a + b|^2$ , leaving the qubit in the state  $|+\rangle$ , and the second with probability  $|a - b|^2$ , leaving the qubit in the state  $|-\rangle$ . In general, we can

project onto any orthogonal set of basis states.

The state of the qubit can be reversibly manipulated by applying a unitary operation— a linear map from one state  $|\psi\rangle$  to a new state  $|\varphi\rangle$ . For a single qubit, such a map can be thought of as a  $2 \times 2$  complex matrix

$$U = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad (1.5)$$

satisfying  $U^\dagger U = I$ , where  $U^\dagger$  is the Hermetian adjoint

$$U^\dagger = \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix}, \quad (1.6)$$

and  $I$  is the identity matrix

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (1.7)$$

which maps any state  $|\psi\rangle$  to itself.

The state of  $n$  qubits requires  $2^n$  complex numbers to specify, one for each of the  $2^n$  classical length- $n$  bit strings, just as a probabilistic state of  $n$  bits would require  $2^n$  real numbers (probabilities) to describe. Some states, such as the state

$$\frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle) = |+\rangle \otimes |+\rangle, \quad (1.8)$$

can be decomposed into states on each individual qubit; such a state is called a product state. Other states, such as the “EPR pair” [20] state

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \quad (1.9)$$

cannot be decomposed in this way, and are called entangled states.

For any orthonormal basis  $\{|s_i\rangle\}$ , there is a natural complex inner product on states, defined

for states  $|\alpha\rangle = \sum_i a_i |s_i\rangle$  and  $|\beta\rangle = \sum_i b_i |s_i\rangle$  as

$$\langle\alpha|\beta\rangle = \sum_{i,j} a_i^* b_j \langle s_i | s_j \rangle = \sum_i c_i^* d_i. \quad (1.10)$$

The Hermetian adjoint, defined earlier, is the adjoint with respect to this inner product, so that the “ket”  $U|\psi\rangle$  corresponds to the “bra”  $\langle\psi|U^\dagger$ . In this notation, we can decompose any operator as

$$\sum_i |\alpha_i\rangle\langle\beta_i|, \quad (1.11)$$

which acts on the state  $|\psi\rangle$  as

$$\left( \sum_i |\alpha_i\rangle\langle\beta_i| \right) |\psi\rangle = \sum_i \langle\beta_i|\psi\rangle |\alpha_i\rangle. \quad (1.12)$$

When considering a quantum system with two subsystems  $\mathcal{H}_A \otimes \mathcal{H}_B$ , we cannot always describe the state of one subsystem  $\mathcal{H}_A$  as some “pure state”  $|\psi\rangle$ . For example, the EPR state

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle),$$

from the perspective of an observer of one half of the system, looks like a probabilistic (non-coherent) mixture of  $|0\rangle$  and  $|1\rangle$ , each with probability  $1/2$ . To describe such a state, we can introduce the density matrix description of a state. A pure state  $|\psi\rangle$  has density matrix  $\rho = |\psi\rangle\langle\psi|$ . For a state  $\sum_a |\psi_a\rangle \otimes |\phi_a\rangle$ , the “reduced state” on subsystem  $\mathcal{H}_A$  has density operator  $\rho = \sum_a |\psi_a\rangle\langle\psi_a|$ . A unitary operator acts on a density matrix  $\rho$  as

$$U : \rho \rightarrow U\rho U^\dagger. \quad (1.13)$$

### 1.2.1 The Pauli operators

It is helpful to understand a system of qubits by studying how various operators act on the states of the system. One useful set of states are the Pauli matrices

$$X = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix} \quad Y = \begin{pmatrix} & -i \\ i & \end{pmatrix} \quad Z = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}, \quad (1.14)$$

where by convention the empty entries are zeros. In bra-ket notation they can be written as

$$X = |0\rangle\langle 1| + |1\rangle\langle 0| \quad Y = -i|0\rangle\langle 1| + i|1\rangle\langle 0| \quad Z = |0\rangle\langle 0| - |1\rangle\langle 1|, \quad (1.15)$$

We can describe the computational basis measurement as a measurement of the eigenvalues of  $Z$ , and the dual basis measurement corresponds to measuring the eigenvalues of  $X$ ; we could similarly measure the eigenvalues of the  $Y$  operation. For a qubit consisting of a particle's spin, these three measurements correspond to measuring the spin along the  $x$ ,  $y$  and  $z$  axes.

Each Pauli operation describes a useful basis, defined by its eigenvectors. The eigenvectors of  $Z$  are  $|0\rangle$  and  $|1\rangle$ , while the eigenvalues of  $X$  are given by

$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \quad (1.16)$$

$$|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle), \quad (1.17)$$

with  $X|+\rangle = |+\rangle$  and  $X|-\rangle = -|-\rangle$ . The eigenvalues of  $Y$  are given by

$$|i\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) \quad (1.18)$$

$$|-i\rangle = \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle), \quad (1.19)$$

with  $Y|\pm i\rangle = \pm|\pm i\rangle$ .

The Pauli matrices anti commute with each other, obeying

$$\{\sigma_i, \sigma_j\} = \sigma_i\sigma_j + \sigma_j\sigma_i = 0 \quad (1.20)$$

(where  $\sigma_1 = X$ ,  $\sigma_2 = Y$ , and  $\sigma_3 = Z$ .) They form a group, in the sense that any product of Pauli operators is itself a Pauli operator, up to an irrelevant phase. The Pauli group on  $n$  qubits,  $\mathcal{P}_n$ , is any operation which can be written as the tensor product of single-qubit Pauli operators and identity operators.

More generally, we can measure the eigenvalues of any operator  $A$  whose eigenvalues are real; such an operator is called Hermetian and obeys  $A^\dagger = A$ . If two operators  $A$  and  $B$  commute, so that

$$[A, B] = AB - BA = 0, \quad (1.21)$$

then both of their eigenvalues can be measured simultaneously. Similarly, all of the eigenvalues of a set of mutually commuting operators may be measured simultaneously. In general, operators may neither commute nor anticommute.

### 1.2.2 The Clifford group

Another useful group of operations are the Clifford group of operations  $\mathcal{C}_1$ , which are defined to be the operations which transform Pauli operations into Pauli operations:

$$CPC^\dagger \in \mathcal{P}_n \text{ for } P \in \mathcal{P}_n, C \in \mathcal{C}_1. \quad (1.22)$$

Examples of Clifford operations include the Hadamard operation

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad (1.23)$$

which exchanges the logical and dual bases; the phase gate

$$S = \begin{pmatrix} 1 & \\ & i \end{pmatrix}; \quad (1.24)$$

and the two-qubit controlled-NOT or CNOT gate

$$\text{CNOT} = \begin{array}{c} \bullet \\ | \\ \oplus \end{array} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}, \quad (1.25)$$

which in the logical basis flips the value of the second qubit based on the value of the first. In fact, any operation in the Clifford group can be decomposed into these elementary operations. Another useful Pauli operation is the controlled- $Z$  operation

$$\text{CZ} = \begin{array}{c} \bullet \\ | \\ \bullet \end{array} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix}. \quad (1.26)$$

In terms of their effect on the Pauli operations, these Clifford operations can be described as

$$\begin{aligned} H : X &\rightarrow Z \\ Y &\rightarrow Y \\ Z &\rightarrow X, \end{aligned} \quad (1.27)$$

$$\begin{aligned} S : X &\rightarrow Y \\ Y &\rightarrow X \\ Z &\rightarrow Z, \end{aligned} \quad (1.28)$$

and

$$\begin{aligned} \text{CNOT} : IX &\rightarrow IX & IZ &\rightarrow ZZ \\ XI &\rightarrow XX & ZI &\rightarrow ZI, \end{aligned} \tag{1.29}$$

where the rest of the CNOT relations follow from linearity.

While the Clifford operations are insufficient for universal computation, they are useful for describing error correcting circuits.

### 1.2.3 Universal computation

We would like to be able to implement or approximate an arbitrary unitary operation  $U$ , to arbitrary precision  $\epsilon$ , by decomposing it into a sequence of fundamental operations or gates, just as an ordinary computation can be decomposed into logic gates from a finite set. The Clifford operations of Section 1.2.2 are insufficient for the task, because they generate only a finite group of unitaries.

It can be shown [4] that the Clifford operations together with any one non-Clifford single qubit gate is sufficient for universality. Moreover, an  $\epsilon$ -approximation to a unitary  $U$  can be constructed efficiently [6, 21–23]. One common choice for a universal set of gates is

$$\mathcal{C}_1 \cup \{T\}, \tag{1.30}$$

where  $T$  is the gate

$$T = \begin{pmatrix} 1 & \\ & e^{i\pi/4} \end{pmatrix}. \tag{1.31}$$

A more streamlined set is

$$\{H, T, \text{CNOT}\}, \tag{1.32}$$

which is universal because  $H$ ,  $S = T^2$  and CNOT complete the Clifford group, while  $T$  is non-Clifford.

## 1.3 Error correcting codes

Error correcting codes, both classical and quantum, are methods of storing information redundantly so that, even though a part of the information has been corrupted or lost, the stored information can be recovered with high probability.

### 1.3.1 Classical error correction

The simplest classical error correcting code is the repetition code, which encodes the message (0) as (000), and the message (1) as (111). If a single bit of the encoded message is flipped, the correct message can still be recovered by taking a majority vote of the bits. If each bit is flipped independently with probability  $p$ , then we will recover correctly with probability

$$P \geq 1 - 3p^2, \quad (1.33)$$

since at least two errors are required to fail, and there are  $\binom{3}{2} = 3$  ways to choose the two errors. This is an improvement over the unencoded case if  $3p^2 < p$ .

A set of classical error-correcting codes which generalize the concept of a repetition code are linear codes. These codes can be described by a binary matrix  $G$ , called the generator of the code. A code with a  $k \times n$  generator matrix will encode a  $k$ -bit message into an  $n$ -bit encoded message. The encoding of the  $k$ -bit message  $x$  is the  $n$ -bit message  $y = xG$ , where  $x$  is viewed as a row vector. The rows of  $G$  form a basis for the  $k$ -dimensional subspace of the  $n$ -dimensional binary vector space. The possible messages  $y = xG$  are called the codewords of the code,  $\mathcal{C}$ .

Equivalently, we can describe a binary code by a set of linear constraints that the codewords  $y$  must satisfy. These constraints can be written using an  $(n - k) \times n$  binary matrix  $H$ , called the parity check matrix or check matrix, as  $Hy = 0$ , which means that  $H$  and  $G$  must satisfy

$$HG^T = 0, \quad (1.34)$$

where  $G^T$  is the matrix transpose of  $G$ .

The effect of an error can be described as flipping some of the bits of the message string  $y$ , or taking

$$y \rightarrow \tilde{y} = y + e. \quad (1.35)$$

If  $y$  is a codeword, then

$$H(y + e) = Hy + He = He. \quad (1.36)$$

Typically, the error will not satisfy all of the constraints, so  $He \neq 0$ .  $He = s$  is called the syndrome of the error  $e$ . If the syndrome is nontrivial, we have detected the presence of the errors.

We would like to not only detect errors, but correct our corrupted codeword back to the original codeword. We cannot hope to correct all possible errors, because some errors will look exactly like codewords, but we can specify some smaller set of errors  $\{e_i\}$  that we would like to be able to correct. If all of the errors  $e_i$  have distinct syndromes, then we can determine which error has occurred by examining the syndrome. Upon seeing the syndrome  $s_i$ , we apply correction  $e_i$ , and restore the state correctly:

$$\tilde{y} \rightarrow \tilde{y} + e_i = (y + e_i) + e_i = y. \quad (1.37)$$

On the other hand, if two errors  $e_1$  and  $e_2$  have the same syndrome, then our correction will not succeed:

$$\tilde{y} \rightarrow \tilde{y} + e_1 = (y + e_1) + e_2 = y + e_1 + e_2 \neq y. \quad (1.38)$$

The recovered message is a codeword, but not the one that we originally stored, so the encoded message has been corrupted.

Typically, we are interested in errors that arise independently from each other on each bit of the message. A natural class of errors to protect against is all the errors of some small weight (where the weight of a bit string  $y$  is the number of 1's in  $y$ .) This is natural because higher weight errors require more independent events to occur, and so are less likely.

The distance of our code,  $d$ , is the minimum weight of any codeword  $y \in \mathcal{C}$ . If a linear code has distance  $d = 2t + 1$ , then each error of weight  $t$  or less has a distinct syndrome. This can be seen by noting that if  $He_1 = He_2 = s$ , then

$$H(e_1 + e_2) = He_1 + He_2 = s + s = 0, \quad (1.39)$$

so the string  $e_1 + e_2$  is a codeword, and therefore has weight at least  $d = 2t + 1$ . On the other hand, if each codeword has weight less than  $t$ , then  $\text{wt}(e_1 + e_2) \leq \text{wt}(e_1) + \text{wt}(e_2) \leq 2t$ ; therefore, two such codewords cannot be identical. A codeword which encodes  $k$  bits into  $n$  and has distance  $d$  can be denoted as an  $[n, k, d]$  code.

Every code  $\mathcal{C}$  has a *dual* code  $\mathcal{C}^\perp$ , whose generator is  $H^T$  and whose check matrix is  $G^T$ ; this is a well-defined code since

$$G^T(H^T)^T = G^T H = (HG^T)^T = 0. \quad (1.40)$$

The codewords  $\mathcal{C}^\perp$  are the set of all strings which are orthogonal to the codewords in  $\mathcal{C}$ . Since a binary string is self-orthogonal if its weight is even,  $\mathcal{C}$  and  $\mathcal{C}^\perp$  can have non-trivial intersection.

### 1.3.2 CSS codes

Classical binary codes, and their duals, are useful for defining a large class of quantum error-correcting codes, known as CSS codes for their inventors Calderbank, Shor, and Steane [24, 25].

If  $\mathcal{C}_1$  is a classical linear code with  $(n - k_1) \times n$  parity check matrix  $H_1$ , we can add additional constraints to  $H_1$  to form a subcode  $\mathcal{C}_2 \subset \mathcal{C}_1$  of  $\mathcal{C}_1$ , with  $(n - k_2) \times n$  check matrix  $H_2$ , where  $k_2 < k_1$ . The subcode  $\mathcal{C}_2$  defines an equivalence relation over  $\mathcal{C}_1$ , where  $u \equiv v$  if and only if there exists  $w \in \mathcal{C}_2$  such that  $u = v + w$ . Then the corresponding CSS code encodes  $k = k_1 - k_2$  qubits, associating a codeword with each equivalence class. The logical basis codewords  $|x\rangle$  are encoded as

$$|x\rangle_L = \frac{1}{\sqrt{2^{k_2}}} \sum_{v \in \mathcal{C}_2} |x + v\rangle, \quad (1.41)$$

a superposition of all the words in the coset  $w + \mathcal{C}_2$ .

Applying the Hadamard operation to each qubit,  $H^{\otimes n}$ , we transform to the dual basis state

$$\begin{aligned} H^{\otimes n} |x\rangle_L &= \frac{1}{\sqrt{2^n}} \sum_{u \in \mathbb{Z}_2^n} \frac{1}{\sqrt{2^{k_2}}} \sum_{v \in \mathcal{C}_2} (-1)^{u(x+v)} |u\rangle \\ &= \frac{1}{\sqrt{2^{n-k_2}}} \sum_{u \in \mathcal{C}_2^\perp} (-1)^{u \cdot x} |u\rangle, \end{aligned} \quad (1.42)$$

where we have used the identity [5]

$$\sum_{v \in \mathcal{C}} (-1)^{v \cdot u} = \begin{cases} 2^k & u \in \mathcal{C}^\perp \\ 0 & u \notin \mathcal{C}^\perp \end{cases}. \quad (1.43)$$

Therefore, the dual basis codewords consist of codewords in the dual code  $\mathcal{C}_2^\perp$ . If the code  $\mathcal{C}_1$  has distance  $d_1 = 2t_X + 1$  and the code  $\mathcal{C}_2^\perp$  has distance  $d_2^\perp = 2t_Z + 1$ , then the CSS code can correct  $t_X$

independent  $X$  errors and  $t_Z$  independent  $Z$  errors. Such a code would be denoted as an  $[[n, k, d]]$  code, where  $d = \min(d_1, d_2^\perp)$ .

### 1.3.3 Stabilizer codes and subsystem codes

A more general class of quantum error-correcting codes, containing the CSS codes, are the stabilizer codes [26, 27]. The codespace of an  $[[n, k, d]]$  stabilizer code is the simultaneous  $+1$  eigenspace of a set  $\mathcal{S}$  of stabilizer operations, an Abelian group of Pauli operations from  $\mathcal{P}_n$  with  $2^{n-k}$  elements. The stabilizer can be characterized by a set of  $n - k$  independent generators. The  $+1$  eigenspace of the stabilizers has dimension  $k$ .

The set of Pauli operators that commute with every element of  $\mathcal{S}$  is called the centralizer  $C(\mathcal{S})$  of  $\mathcal{S}$  in  $\mathcal{P}_n$ . Elements  $L$  in  $C(\mathcal{S})/\mathcal{S}$  act as logical operations, changing the encoded information without leaving the codespace. We can find  $k$  pairs of logical operators  $\{\bar{X}_i, \bar{Z}_i\}$  satisfying

$$[\bar{X}_i, \bar{X}_j] = 0 \tag{1.44}$$

$$[\bar{Z}_i, \bar{Z}_j] = 0 \tag{1.45}$$

$$[\bar{X}_i, \bar{Z}_j] = 0 \quad (i \neq j) \tag{1.46}$$

$$\{\bar{X}_i, \bar{Z}_i\} = 0 \tag{1.47}$$

An arbitrary operator  $E \in \mathcal{P}_n$  will generically anticommute with some stabilizer operations, and act as an error. We can detect this error by measuring a set of generators for the stabilizer group; because the stabilizers commute with the logical operations, doing so will not disturb the encoded information, and because the stabilizers commute with each other, we can measure them simultaneously. The set of measurement results will be called the syndrome. We can attempt to correct an error by applying the lowest weight operator  $E'$  that agrees with the syndrome. A code will have distance  $d$  if and only if  $C(\mathcal{S}) \setminus \mathcal{S}$  contains only elements of weight  $d$  and higher.

A subsystem code [28] is a stabilizer code where we reclassify some logical qubits as gauge degrees of freedom. Unlike the logical qubits, we do not attempt to store any information in the gauge degrees of freedom; they are allowed to vary arbitrarily. By moving low-distance logical operations into gauge operations, we can increase the distance of the code. Gauge degrees of freedom can often be useful for measuring stabilizers indirectly, as will be seen in Section 3.4.

## 1.4 Fault tolerance

Quantum fault-tolerance [5, 29, 30] is concerned not only with protecting information not only for storage, but also processing information in an encoded form. This can be achieved by replacing every physical gate in a quantum circuit with an encoded version of the gate, and also adding error-correction steps where syndromes are measured and corrections applied. If done correctly, encoded circuits can drastically reduce the error rates of circuits at the logical level. If the physical error rate is low enough, threshold theorems establish that an arbitrarily low error rate can be achieved [26, 30–36].

The goal of fault-tolerant circuit design is to ensure that errors at the physical level cannot build up to errors at the logical level over the course of a computation, even while the operations and measurements used to diagnose and correct errors are themselves faulty. Properly designed, fault-tolerant circuits can guarantee that each faulty location in the circuit can introduce no more than one error to the output (or more loosely, a constant number of errors.)

A key ingredient to making a computation fault tolerant is to avoid interacting a single physical qubit to too many other qubits. If a single qubit were to interact with every other qubit in the same code block, for example, then it is possible that an error on the single qubit could propagate to every other qubit in the code block, creating a correlated logical error from a single physical error.

One strategy for avoiding correlated errors is to use transversal encoded gates. A single-qubit encoded gate is called *transversal* if the only gates at the physical level are single-qubit gates — typically the same gate that is being implemented at the logical level. A transversal encoded gate on multiple logical qubits should only contain physical gates where no physical qubit is coupled to another physical qubit in the same code block, or two more than one physical qubit in any other code block. Equivalently, a transversal gate can be implemented by a depth 1 circuit with no two-qubit gates within a code block. Typically, this means that the  $i$ th qubit in the first block couples to the  $i$ th qubit in the second block, and so on, but a permutation of this arrangement would also be transversal. Examples of transversal single- and two-qubit gates are shown in Figure 1.1.

### 1.4.1 Syndrome extraction techniques

Realizing that syndrome measurements themselves can fail due to errors, a critical strategy for achieving fault-tolerance is to repeat syndrome measurements multiple times. There are a few

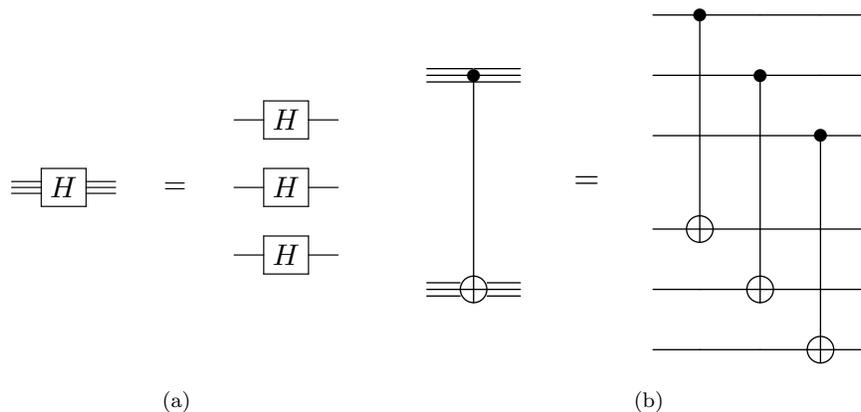


Figure 1.1: (a) Transversal  $H$  gate and (b) transversal CNOT gate, for a hypothetical 3-qubit code.

techniques for syndrome extraction.

In Shor error-correction [29], a cat state

$$|\text{cat}\rangle = \frac{1}{\sqrt{2}}(|00\dots 0\rangle + |11\dots 1\rangle) \quad (1.48)$$

is prepared for each syndrome operation, using a circuit similar to Figure 1.2 where the length of the cat state is equal to the weight of the stabilizer being measured. The cat state can be tested to verify that it has been prepared correctly; if not, we throw it away and try again. After preparing a verified cat state, we can apply either a CNOT (for  $Z$  stabilizers) or a CZ gate (for  $X$  stabilizers) from separate qubits of the the cat state to the nontrivial qubits of the stabilizer being measured. Then the stabilizer can be obtained by measuring the cat state in the  $X$  basis; the parity of the physical measurement outcomes is the stabilizer measurement outcome. This measurement should be repeated to ensure fault tolerance.

An alternative method of syndrome extraction for CSS codes is Steane style error correction [37], which extracts all of the  $X$  or  $Z$  syndrome information simultaneously. As depicted in Figure 1.3, a verified *logical*  $|+\rangle$  or  $|0\rangle$  state is prepared (and verified), and a transversal CNOT gate is performed between the ancilla block and the data block. This CNOT gate is chosen so that it has no logical effect, but will copy errors in either the  $X$  or  $Z$  basis onto the ancilla block. Finally, a measurement is performed and the syndrome information extracted.

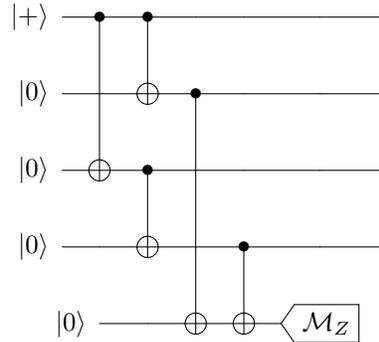


Figure 1.2: Preparation of a verified 4-qubit cat state.

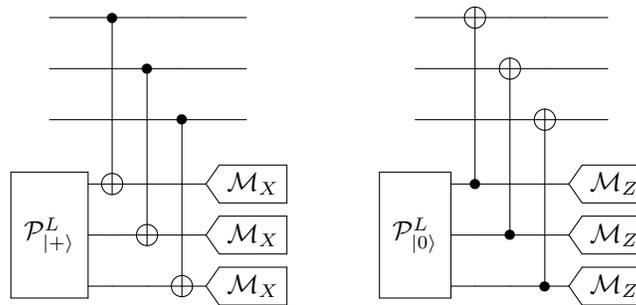


Figure 1.3: Steane error correction.

A third alternative is Knill style teleported error correction [38]. This approach uses a quantum teleportation circuit [39]. Effectively, errors are corrected by teleporting a state with errors onto a fresh ancilla qubit. Teleported error correction will be discussed in more detail in Section 3.5.6, where a variant of the approach is used extensively.

Depending on the code and situation, each of these techniques may have advantages and disadvantages. Shor style error correction can be applied to any code, while the other techniques will only be useful for codes with a transversal two-qubit operation such as the CNOT gate. When applicable, Steane and Knill error correction may offer smaller, more compact circuits for extracting all of the syndrome. Each of these methods will find use at different points in this thesis.

## Chapter 2

# Protected Gates for Superconducting Qubits

*Ecce ancilla Domini.*

–Luke 1:38

This chapter is based on work that was published in [1].

### 2.1 Introduction

Building a scalable quantum computer is a formidable challenge because quantum systems decohere readily and because their interactions are hard to control accurately, yet we hope to succeed someday by prudently applying the principles of quantum error correction and fault-tolerant quantum computing. In the standard “software” approach to quantum fault-tolerance [29], the deficiencies of noisy quantum hardware (if not *too* noisy) are overcome through clever circuit design, while in the alternative “topological” approach [21], the hardware itself is intrinsically resistant to decoherence. Both approaches exploit the idea that logical qubits can be stored and processed reliably when suitably encoded in a quantum system with many degrees of freedom; perhaps both approaches will be employed together in future quantum computing systems.

The best known version of the topological approach is based on nonabelian anyons, with quantum information stored in the fusion spaces of the anyons and processed by braiding the anyons, but it is important to search for other ways to realize quantum hardware such that intrinsic robustness results from how the information is physically encoded. One intriguing possibility is to

use superconducting circuits for this purpose. Specifically, several authors [40–42] have proposed designs for a superconducting “ $0-\pi$  qubit,” a circuit containing Josephson junctions. The circuit’s energy is a function of the superconducting phase difference  $\theta$  between the two leads of the circuit, and there are two nearly degenerate ground states, localized near  $\theta = 0$  and  $\theta = \pi$  respectively. The splitting of this degeneracy is exponentially small as a function of extensive system parameters, and stable with respect to weak local perturbations. Thus the  $0-\pi$  qubit should be highly resistant to decoherence arising from local noise.

For reliable quantum computing we need not just very stable qubits, but also the ability to apply very accurate nontrivial quantum gates to the qubits. A method for achieving protected single-qubit and two-qubit phase gates acting on  $0-\pi$  qubits, exploiting the error-correcting properties of a continuous-variable quantum code [43], was suggested in [42], and it was claimed that the gate errors can be exponentially small as a function of extensive system parameters. In this chapter we further develop and explore the ideas behind this protected gate.

Protected phase gates are executed by turning on and off a tunable Josephson coupling between an  $LC$  oscillator and a qubit or pair of qubits. Assuming the qubits are perfect, we show, using analytic arguments validated by numerical simulations, that the gate errors are exponentially small when the oscillator’s impedance  $\sqrt{L/C}$  is large compared to  $\hbar/4e^2 \approx 1\text{ k}\Omega$ , where  $L$  is the inductance and  $C$  is the capacitance of the oscillator. The gates are robust against small deformations of the device Hamiltonian and against small thermal fluctuations of the oscillator. The very large inductance in the superconducting oscillator, which is crucial for the high gate accuracy, may be quite challenging to achieve in practice, but the potential rewards are correspondingly substantial.

The internal structure of the  $0-\pi$  qubit is not relevant to our analysis, but for completeness we nevertheless explain in Section 2.2 the idea behind the qubit design proposed in [42], which also requires a large inductance in a superconducting circuit. We describe how a protected phase gate is executed in Section 2.3, and in Section 2.4 we outline our strategy for estimating the gate accuracy. We review the properties of continuous-variable quantum error-correcting codes in Section 2.5, and explain in Section 2.6 how the code provides protection against imperfect timing in the pulse that executes the gate. We analyze contributions to the gate error due to diabatic transitions and squeezing in Section 2.7 and Section 2.8, then compare our predictions with numerical simulations in Section 2.9. We discuss robustness with respect to thermal effects in Section 2.10 and with respect to Hamiltonian perturbations in Section 2.11. In Section 2.12 we explain how to obtain a complete

scheme for universal fault-tolerant quantum computation by augmenting the protected phase gates with measurements and unprotected noisy phase gates. Section 2.13 contains our conclusions, and some further details are contained in Appendices.

## 2.2 The $0-\pi$ qubit

For most of this chapter we will be concerned with the dynamics of gates built from the  $0-\pi$  qubit, and will treat the qubit itself as a black box. In this section we will outline the ideas behind the qubit itself, which was proposed in [42].

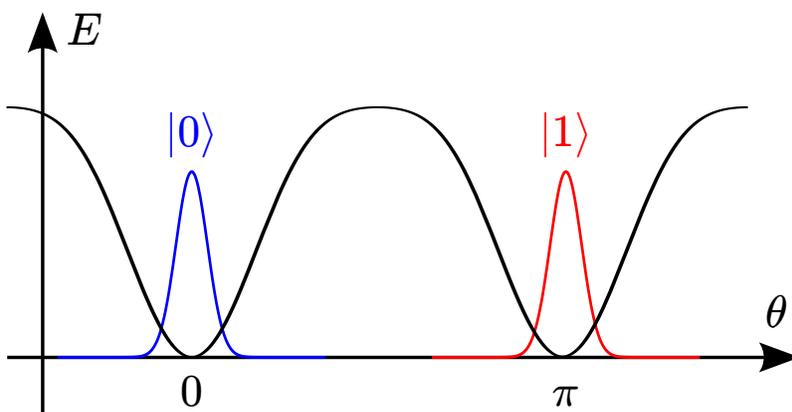


Figure 2.1: The energy  $E(\theta)$  of the  $0-\pi$  qubit. The energy is a periodic function with period  $\pi$  of the phase difference  $\theta$  between its two leads, aside from exponentially small corrections. The two basis states  $\{|0\rangle, |1\rangle\}$  of the qubit, localized near the minima of the energy at  $\theta = 0$  and  $\theta = \pi$  respectively, are nearly degenerate.

The goal in designing the qubit is to have a system which is naturally robust against both bit-flips and dephasing errors. The qubit itself will be a superconducting circuit with two leads, whose energy  $E(\theta)$  is a function of the phase difference  $\theta$  between the two leads. The circuit is designed so that this energy, which is naturally a periodic function of  $\theta$  with period  $2\pi$ , is in fact very close to a function with period  $\pi$ , having two minima at  $\theta = 0$  and  $\theta = \pi$ , separated by a large energy barrier. This energy barrier ensures that there are two well-localized states, one centered around each minima, which we will take to be the basis states  $|0\rangle$  (centered at  $\theta = 0$ ) and  $|1\rangle$  (centered

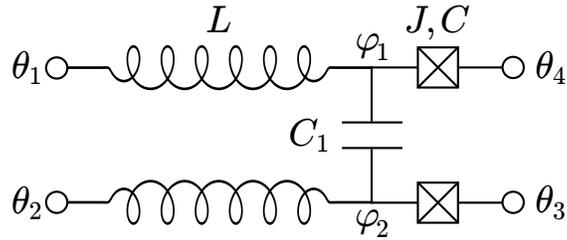


Figure 2.2: Two-rung superconducting circuit underlying the  $0-\pi$  qubit. If  $\sqrt{L/C}$  is large,  $C_1$  is large compared to  $C$ , and  $JC$  is not too large, then the circuit’s energy is a function of the combination of phases  $(\theta_2 + \theta_4) - (\theta_1 + \theta_3)$ , aside from corrections that are exponentially small in  $\sqrt{L/C}$ .

at  $\theta = \pi$ ) of an encoded qubit, as shown in Figure 2.1. The largeness of the barrier suppresses the odds of a bit flip error, and the *degeneracy* of the two states, enforced by the  $\pi$  periodicity, suppresses the odds of a dephasing error. The system is designed so that this degeneracy is robust against generic local perturbations, so that the error protection is not broken.

To get such a system, we start with the four-lead circuit shown in Figure 2.2. This circuit has two identical rungs, connected by a large capacitance  $C_1$ . Each rung consists of a Josephson junction, with Josephson energy  $J$  and intrinsic capacitance  $C$ , connected in series with an inductance  $L$ , chosen such that  $\sqrt{L/C}$  is large compared to the natural unit of impedance  $\hbar/(2e)^2 \approx 1.03 \text{ k}\Omega$ , and hence much larger than its “geometric” value  $4\pi/c \approx 377 \text{ }\Omega$  (where  $c$  is the speed of light), the impedance of free space. Achieving such a “superinductance” may be a daunting experimental challenge, but we take it for granted here that it is possible. The properties of a single rung, which can operate as an adiabatic switch when  $J$  varies, is discussed in more detail in Appendix 2.A.

We denote the value of the superconducting phase on the circuit’s four leads as  $\theta_1, \theta_2, \theta_3, \theta_4$  as shown, and the phase on either side of the capacitor connecting the rungs by  $\varphi_1, \varphi_2$ . Then the phase  $\varphi_+ = (\varphi_1 + \varphi_2)/2$  is insensitive to the value of the capacitance  $C_1$ , which we assume is much larger than  $C$ . Therefore the sum  $\varphi_+$  is a “light” variable with large fluctuations (assuming  $JC$  is not too large), while in contrast the difference  $\varphi_- = \varphi_1 - \varphi_2$ , which does feel the effect of the large capacitance  $C_1$ , is a well localized “heavy” variable. We assume that phase slips through the inductors are suppressed, so that we may regard  $\varphi_{\pm}$  as real variables rather than periodic phase variables with period  $2\pi$ .

A circuit with capacitance  $C_{\text{conv}}$  and inductance  $L_{\text{conv}}$  has Hamiltonian

$$H = \frac{q^2}{2C_{\text{conv}}} + \frac{\Phi^2}{2L_{\text{conv}}}, \quad (2.1)$$

where  $q$  is the charge on the capacitor and  $\Phi$  is the magnetic flux linking the circuit. We use the subscript “conv” to indicate that capacitance and inductance are expressed here in conventional units, while we will find it more convenient to use rationalized units such that

$$C = C_{\text{conv}}/(2e)^2, \quad L = L_{\text{conv}}/(\hbar/2e)^2, \quad (2.2)$$

so that

$$H = \frac{Q^2}{2C} + \frac{\varphi^2}{2L}, \quad (2.3)$$

where the charge  $Q = q/2e$  is expressed in units of the Cooper pair charge  $2e$ , and  $\varphi = (2e/\hbar)\Phi$  is the superconducting phase, such that  $\varphi = 2\pi$  corresponds to the quantum  $h/2e$  of magnetic flux.

Then  $[\varphi, Q] = i$ , and

$$\begin{aligned} \sqrt{L/C} &= \sqrt{L_{\text{conv}}/C_{\text{conv}}} / (\hbar/4e^2) \\ &\approx \sqrt{L_{\text{conv}}/C_{\text{conv}}} / (1.03 \text{ k}\Omega) \end{aligned} \quad (2.4)$$

is dimensionless. The ground state of the Hamiltonian Equation (2.3), with energy  $E_0 = 1/2\sqrt{LC}$ , has Gaussian wave function  $\psi(\varphi)$  such that

$$|\psi(\varphi)|^2 = \frac{1}{\sqrt{\pi\langle\varphi^2\rangle}} e^{-\varphi^2/2\langle\varphi^2\rangle}, \quad (2.5)$$

where

$$\langle\varphi^2\rangle = \frac{1}{2}\sqrt{\frac{L}{C}}. \quad (2.6)$$

There is also a Josephson energy term proportional to  $\cos \varphi$ , but for  $\sqrt{L/C} \gg 1$  the wiggles of the

cosine nearly average out, aside from an exponentially small correction:

$$\begin{aligned}
\langle \cos \varphi \rangle &= \frac{1}{\sqrt{\pi \langle \varphi^2 \rangle}} \int_{-\infty}^{\infty} \cos \varphi e^{-\varphi^2/2\langle \varphi^2 \rangle} d\varphi \\
&= \frac{1}{\sqrt{\pi \langle \varphi^2 \rangle}} \int_{-\infty}^{\infty} e^{-\varphi^2/2\langle \varphi^2 \rangle + i\varphi} d\varphi \\
&= e^{-\langle \varphi^2 \rangle/2} \\
&= \exp\left(-\frac{1}{4}\sqrt{\frac{L}{C}}\right).
\end{aligned} \tag{2.7}$$

The effective capacitance controlling the light phase  $\varphi_+$  is  $C_{\text{eff}} = 2C$ , and the effective inductance is  $L_{\text{eff}} = L/2$ . Therefore, in the circuit's ground state we have

$$\langle \varphi_+^2 \rangle = \frac{1}{2} \sqrt{\frac{L_{\text{eff}}}{C_{\text{eff}}}} = \frac{1}{4} \sqrt{\frac{L}{C}}. \tag{2.8}$$

The dependence of the Josephson energy on the strongly fluctuating light variable  $\varphi_+$  is proportional to

$$\langle \cos \varphi_+ \rangle = \exp\left(-\frac{1}{8}\sqrt{\frac{L}{C}}\right), \tag{2.9}$$

which is negligible when  $\sqrt{L/C}$  is large. We therefore need only consider the dynamics of the well localized heavy variable  $\varphi_-$ , which locks to the value

$$\phi_- = (\theta_4 - \theta_1) - (\theta_3 - \theta_2) = (\theta_2 + \theta_4) - (\theta_1 + \theta_3), \tag{2.10}$$

determined by the phases on the leads, so that the energy stored in the circuit is

$$E = f(\theta_2 + \theta_4 - \theta_1 - \theta_3) + O\left(\exp\left(-\frac{1}{8}\sqrt{\frac{L}{C}}\right)\right), \tag{2.11}$$

where  $f(\theta)$  is a periodic function with period  $2\pi$ .

Now, to devise a qubit, we twist the upper rung relative to the lower one and connect the leads as shown in Figure 2.3, thus identifying  $\theta_2$  with  $\theta_4$  and  $\theta_1$  with  $\theta_3$ . In addition, we add another large capacitance to ensure that tunneling events changing  $\theta_2 - \theta_1$  by  $\pi$  are heavily suppressed. The

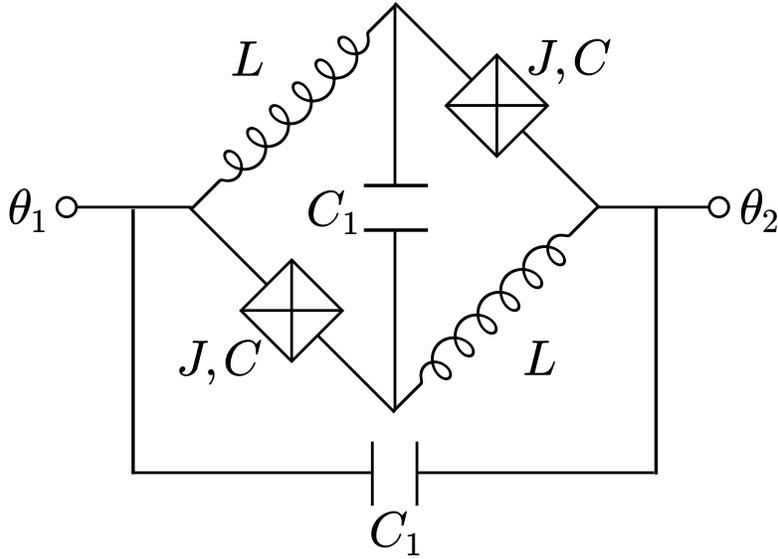


Figure 2.3: The circuit for the  $0-\pi$  qubit is obtained from the circuit in Figure 2.2 by twisting one of the rungs and connecting the leads, thus identifying  $\theta_2$  with  $\theta_4$  and  $\theta_1$  with  $\theta_3$ . In addition, another large capacitance is added to further suppress tunneling events that change  $\theta_2 - \theta_1$  by  $\pi$ .

energy of the resulting circuit is

$$E = f(2(\theta_2 - \theta_1)) + O\left(\exp\left(-\frac{1}{8}\sqrt{\frac{L}{C}}\right)\right), \quad (2.12)$$

where the ellipsis represents exponentially small corrections. Therefore, the energy is very nearly a periodic function with period  $\pi$  of the phase difference  $\theta_2 - \theta_1$ , with two nearly degenerate minima as in Figure 2.1.

### 2.2.1 Achieving superinductance

This robust degeneracy derives from the “superinducting” properties of each rung, *i.e.*, the large value of  $\sqrt{L/C}$ . One way to achieve a superinductor, suggested in [42], is to construct a long chain of  $N$  Josephson junctions, each with Josephson coupling  $\bar{J}$  and capacitance  $\bar{C}$ . Then the inductance of the chain is linear in  $N$ , and the capacitance is proportional to  $1/N$ , so  $\sqrt{L/C} \propto N$ , and the breaking of the degeneracy is exponentially small in the chain length. This suppression

arises because the correction terms in Equation (2.12) that break the  $\pi$ -periodicity are associated with quantum tunneling from one end to the other in the two-rung ladder. We also require  $\bar{J}\bar{C}$  to be large, to suppress phase slips due to tunneling across the chain, thus ensuring that  $\varphi_+$  can be regarded as a real variable rather than a periodic variable with period  $2\pi$ .

An impedance  $\sqrt{L/C} \approx 20$  has been achieved using long chains of devices [44–46]. Another possibility for achieving large  $\sqrt{L/C}$  is to use a long wire, thick enough to suppress phase slips, built from an amorphous superconductor with a large kinetic inductance. Whatever method is used, reaching, say,  $\sqrt{L/C}$  of order 100 may be quite challenging, but in this thesis we take it for granted that a robust  $0-\pi$  qubit can be realized. In fact, our scheme for implementing accurate quantum gates will also be based on superinducting circuits.

### 2.2.2 Measurements

We will need to be able to measure the qubit, in either the standard  $\{|0\rangle, |1\rangle\}$  basis (measurement of the Pauli operator  $Z$ ) or in the dual basis  $\{|+\rangle, |-\rangle\}$  (measurement of the Pauli operator  $X$ ). In principle, the  $Z$  measurement could be performed by connecting the two leads of the qubit with a Josephson junction, while inserting  $1/4$  of a flux quantum through the loop; then the current through the junction is proportional to  $\sin(\theta_2 - \theta_1 - \pi/2)$ , with sign dependent on whether  $\theta_2 - \theta_1$  is 0 or  $\pi$ .

For measuring  $X$ , we envision “breaking” the connection between  $\theta_1$  and  $\theta_3$  and then measuring the charge conjugate to the phase difference  $\theta_1 - \theta_3$ . The energy of the circuit is  $f(\theta_1 + \theta_3 - 2\theta_2)$ , so that if  $\theta_1$  advances adiabatically by  $2\pi$  with  $\theta_3$  fixed, then  $\theta_2$  advances by  $\pi$ ; if  $X = 1$  the wave function is invariant and if  $X = -1$  the wave function changes sign. Correspondingly, the dual charge is either an even or odd multiple of  $1/2$ . In practice, the  $X$  and  $Z$  measurements are bound to be noisy, but the limitations on measurement accuracy can be overcome by repeating the measurements or by using appropriate coding schemes, as we describe in Section 2.12.

## 2.3 Phase gate

Following [42], we will explain how to execute with high fidelity the single-qubit phase gate  $\exp(i\frac{\pi}{4}Z)$  and the two-qubit phase gate  $\exp(i\frac{\pi}{4}Z \otimes Z)$ . These gates are not sufficient by themselves for universal quantum computation, but we will discuss in Section 2.12 how they can be used as part of

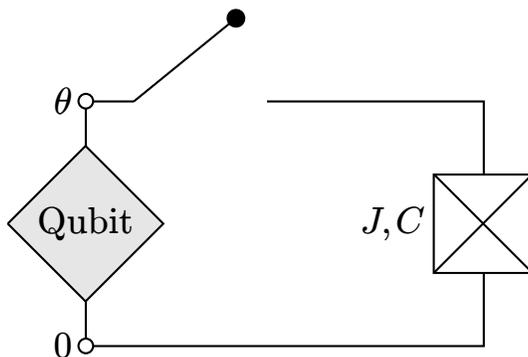


Figure 2.4: A phase gate can be applied to a qubit by coupling it to a Josephson junction, but the gate is not protected against pulse errors and other noise sources.

a universal fault-tolerant scheme.

First, for contrast, consider an example of an unprotected single-qubit gate implementation. As shown in Figure 2.4, we could close a switch that couples the qubit for time  $t$  to a Josephson junction with Josephson coupling  $J$ , in effect turning on a term  $J \cos \theta = JZ$  in the Hamiltonian, where  $\theta \in \{0, \pi\}$  is the phase difference across the qubit. After time  $t$  the unitary transformation  $\exp(-itJZ)$  has been applied. By choosing the time  $t$  appropriately, we can rotate the qubit about the  $z$  axis by any desired angle. However, this gate is sensitive to errors in the pulse that closes and opens the switch, and to other fluctuations in the circuit parameters. For example, if we leave the gate on for too long, the accumulated error is linear in the gate mistiming.

The protected single-qubit phase gate is executed as shown in Figure 2.5, by coupling the qubit to a “superinductive”  $LC$  circuit via a switch that pulses on and off. The switch is actually a tunable Josephson junction, which can be realized, as in Figure 2.6, by a loop containing two identical junctions, each with Josephson coupling  $J$ , linked by the magnetic flux  $(\eta/2\pi)\Phi_0$ , where  $\Phi_0 = h/2e$  is the flux quantum. The Josephson energy of this tunable junction is

$$\begin{aligned}
 E(\theta, \eta) &= -J \cos(\theta - \eta/2) - J \cos(\theta + \eta/2) \\
 &= -2J \cos(\eta/2) \cos \theta \\
 &= -J_{\text{eff}}(\eta) \cos \theta,
 \end{aligned} \tag{2.13}$$

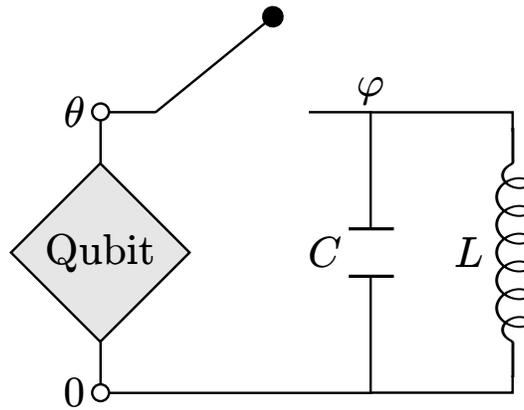


Figure 2.5: A protected phase gate is executed by coupling a qubit (or a pair of qubits connected in series) to a “super-quantum”  $LC$  circuit with  $\sqrt{L/C} \gg 1$ .

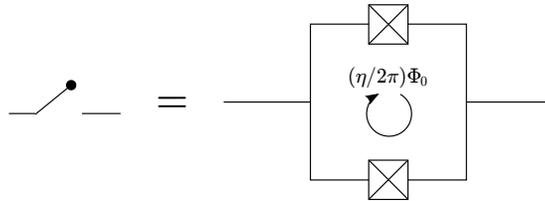


Figure 2.6: An effective Josephson junction can be tuned by adjusting the flux  $(\eta/2\pi)\Phi_0$  inserted in a circuit containing two identical junctions.

where  $\theta$  is the phase difference between the two leads on the loop. Thus the switch is “on” for  $\eta = 0$  and “off” for  $\eta = \pi$ . The “off” setting can be fairly soft — it is good enough for  $J_{\text{eff}}$  to be comparable to  $1/L$  rather than strictly zero — while in the “on” position we require  $J_{\text{eff}}C$  to be large. The inductance  $L$  and capacitance  $C$  of the circuit are unrelated to the inductance and capacitance for the  $0-\pi$  qubit discussed in Section 2.2, though we will again demand that  $\sqrt{L/C} \gg 1$ . From now on we will assume the  $0-\pi$  qubit is perfect, and will focus on realizing the robust phase gate under this assumption.

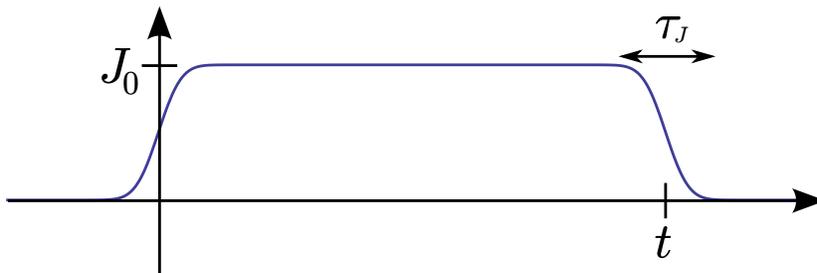


Figure 2.7: The profile of the tunable Josephson coupling  $J(t)$  in the execution of the protected phase gate.

### 2.3.1 Ideal protected phase gate

We will first describe how the protected phase gate works in the ideal case with no errors. In the following section we will discuss the effects of the imperfections and argue that their effects are minor.

Using the same normalization conventions as in Section 2.2, the Hamiltonian for the circuit can be expressed as

$$H(t) = \frac{Q^2}{2C} + \frac{\varphi^2}{2L} - J(t) \cos(\varphi - \theta), \quad (2.14)$$

where now  $J(t)$  is the time-dependent effective Josephson coupling of the tunable junction,  $\theta$  is the phase difference across the qubit, and  $\varphi$  is the phase difference across the inductor. We assume that phase slips through the circuit are strongly suppressed, so that  $\varphi$  can be regarded as a real variable rather than a periodic phase variable — when  $\varphi$  winds by  $2\pi$  the flux linking the  $LC$  circuit increases by one flux quantum. Depending on whether the state of the qubit is  $|0\rangle$  or  $|1\rangle$ , the phase  $\theta$  is either 0 or  $\pi$ ; hence, the Hamiltonian can be expressed as

$$H_{0,1}(t) = \frac{Q^2}{2C} + \frac{\varphi^2}{2L} \mp J(t) \cos \varphi, \quad (2.15)$$

with the  $\mp$  sign conditioned on the qubit's state.

Suppose the initial state  $|\psi^{\text{in}}\rangle$  of the oscillator is its ground state, a Gaussian wave function which is broad in  $\varphi$  space. We now begin to ramp on the tunable Josephson coupling  $J(t)$ , with the form shown in Figure 2.7: starting at zero, ramping on smoothly and then rounding off to

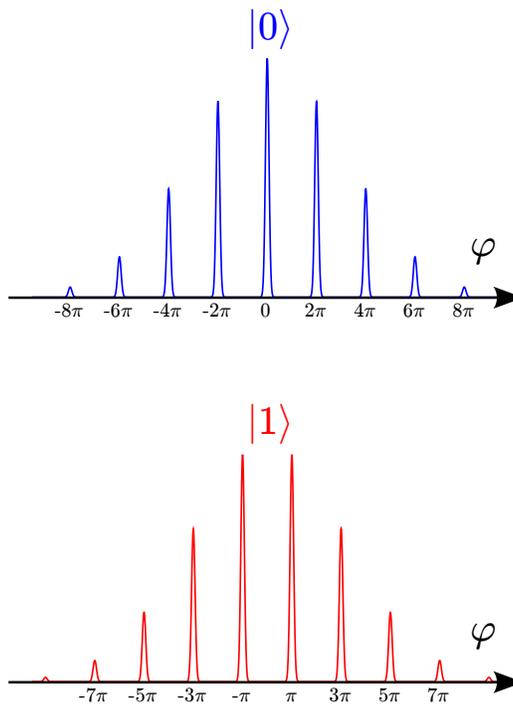


Figure 2.8: Coupling the qubit to the oscillator prepares a grid state in  $\varphi$  space, a superposition of narrowly peaked functions governed by a broad envelope function. The peaks occur where  $\varphi$  is an even multiple of  $\pi$  if the qubit's state is  $|0\rangle$ , and where  $\varphi$  is an odd multiple of  $\pi$  if the qubit's state is  $|1\rangle$ .

the constant value  $J_0$ . We assume that  $J(t)$  ramps up slowly enough to prepare adiabatically the ground state in each local minimum of the cosine potential, yet quickly enough to prevent the state from collapsing to just a few local minima with the smallest values of  $\varphi^2/2L$ . Thus, as  $J(t)$  turns on, the initial state of the oscillator evolves to become a “grid state,” as shown in Figure 2.8, a superposition of narrowly peaked functions governed by a broad envelope function. The width of the broad envelope is  $\langle\varphi^2\rangle \approx \frac{1}{2}\sqrt{L/C} \gg 1$ , as for the oscillator's initial state, while the width of each narrow peak is  $\langle(\varphi - \varphi_0)^2\rangle \approx \frac{1}{2}\sqrt{1/J_0C} \ll 1$ , the width of the ground state supported near the local minimum of the cosine potential.

If the state of the  $0-\pi$  qubit is  $|0\rangle$  and the coefficient of the cosine in Equation (2.15) is negative,

then the narrow peaks occur where  $\varphi$  is an even multiple of  $\pi$ . We denote this grid state of the oscillator as  $|0_C\rangle$ ; the subscript stands for “code,” since, as we will explain later, this state can be regarded as a basis state for a quantum error-correcting code. If the state of the qubit is  $|1\rangle$  and the coefficient of the cosine is positive, then the narrow peaks occur where  $\varphi$  is an odd multiple of  $\pi$ ; in that case we denote the grid state as  $|1_C\rangle$ . Thus, if the initial state of the  $0-\pi$  qubit is  $a|0\rangle + b|1\rangle$ , then when  $J(t)$  turns on, the joint state of the qubit and oscillator evolves according to

$$(a|0\rangle + b|1\rangle) |\psi^{\text{in}}\rangle \rightarrow a|0\rangle \otimes |0_C\rangle + b|1\rangle \otimes |1_C\rangle. \quad (2.16)$$

After the Gaussian grid state has been prepared, the Josephson coupling  $J(t)$  maintains its steady-state value  $J_0$  for a time  $t \equiv L\tilde{t}/\pi$ , where  $\tilde{t}$  is a rescaled time variable. While the coupling is on, each narrowly peaked function is stabilized by the strongly confining cosine potential, but the state is subjected to the Gaussian operation  $e^{-it\varphi^2/2L} = e^{-i\tilde{t}\varphi^2/2\pi}$ , due to the harmonic potential  $\varphi^2/2L$ , which alters the relative phases of the peaks. As  $\tilde{t}$  increases the oscillator states  $|0_C\rangle$  and  $|1_C\rangle$  evolve, but when  $\tilde{t}$  reaches 1, each returns to its initial value, apart from a state-dependent geometric phase. For the grid state  $|0_C\rangle$ , the peaks in  $\varphi$  space occur at  $\varphi = 2\pi n$ , where  $n$  is an integer, and the Gaussian operation

$$\begin{aligned} |\varphi = 2\pi n\rangle &\rightarrow e^{-2\pi i\tilde{t}n^2} |\varphi = 2\pi n\rangle \\ &= e^{-2\pi i n^2} |\varphi = 2\pi n\rangle \\ &= |\varphi = 2\pi n\rangle \end{aligned} \quad (2.17)$$

acts trivially. But for the grid state  $|1_C\rangle$ , the peaks occur at  $\varphi = 2\pi(n + \frac{1}{2})$ , and the operation

$$\begin{aligned} |\varphi = 2\pi n\rangle &\rightarrow e^{-2\pi i\tilde{t}(n+\frac{1}{2})^2} |\varphi = 2\pi(n + \frac{1}{2})\rangle \\ &= e^{-2\pi i(n^2+n+\frac{1}{4})} |\varphi = 2\pi(n + \frac{1}{2})\rangle \\ &= e^{-i\frac{\pi}{2}} |\varphi = 2\pi(n + \frac{1}{2})\rangle \end{aligned} \quad (2.18)$$

therefore modifies the phase of the state by the factor  $-i$ . Hence, the joint state of the qubit and

oscillator becomes

$$a|0\rangle \otimes |0_C\rangle + b|1\rangle \otimes |1_C\rangle \rightarrow a|0\rangle \otimes |0_C\rangle - ib|1\rangle \otimes |1_C\rangle. \quad (2.19)$$

To complete the execution of the phase gate, the tunable coupling  $J(t)$  ramps down from  $J_0$  to zero, again with a characteristic time scale  $\tau_J$  subject to the constraints specified above. As the coupling turns off, the state  $|0_C\rangle$  of the oscillator evolves to  $|\psi_0^{\text{fin}}\rangle$  and the state  $|1_C\rangle$  evolves to  $|\psi_1^{\text{fin}}\rangle$ ; the final joint state of the qubit and oscillator is

$$a|0\rangle \otimes |0_C\rangle - ib|1\rangle \otimes |1_C\rangle \rightarrow a|0\rangle \otimes |\psi_0^{\text{fin}}\rangle - ib|1\rangle \otimes |\psi_1^{\text{fin}}\rangle. \quad (2.20)$$

Thus, a perfect phase gate  $\exp(i\frac{\pi}{4}Z)$  has been applied to the qubit if  $|\psi_0^{\text{fin}}\rangle = |\psi_1^{\text{fin}}\rangle$ . If, on the other hand,  $|\langle\psi_1^{\text{fin}}|\psi_0^{\text{fin}}\rangle| < 1$ , then the qubit and oscillator are entangled in the final state, compromising the gate fidelity. Even if  $|\langle\psi_1^{\text{fin}}|\psi_0^{\text{fin}}\rangle| = 1$  so that there is no entanglement, the gate may be imperfect because the phase of  $\langle\psi_1^{\text{fin}}|\psi_0^{\text{fin}}\rangle$  deviates from zero.

The two-qubit phase gate  $\exp(i\frac{\pi}{4}Z \otimes Z)$  is executed using a similar procedure, but now two qubits connected in series are coupled to the  $LC$  oscillator. The total phase difference across the pair of qubits is either 0 for the states  $|0\rangle \otimes |0\rangle$  and  $|1\rangle \otimes |1\rangle$ , in which case the oscillator evolves to the final  $|\psi_0^{\text{fin}}\rangle$ , or  $\pi$  for the states  $|0\rangle \otimes |1\rangle$  and  $|1\rangle \otimes |0\rangle$ , in which case the oscillator evolves to the final state  $|\psi_1^{\text{fin}}\rangle$ . Again, the gate is executed perfectly if  $|\psi_0^{\text{fin}}\rangle = |\psi_1^{\text{fin}}\rangle$ .

### 2.3.2 Effects of protected phase gate imperfections

Suppose for now that the initial state  $|\psi^{\text{in}}\rangle$  of the oscillator is its ground state, a Gaussian wave function with  $\langle\varphi^2\rangle = \frac{1}{2}\sqrt{L/C}$  and  $\langle Q^2\rangle = \frac{1}{2}\sqrt{C/L}$ . (Other harmonic oscillator energy eigenstates will be considered in Section 2.10.) Because  $\sqrt{L/C} \gg 1$ , the wave function is broad in  $\varphi$  space and narrow in  $Q$  space. Hence, when the switch pulses on, the contribution to the expectation of the energy arising from the cosine potential is highly suppressed by the factor

$$\langle\cos\varphi\rangle = e^{-\langle\varphi^2\rangle/2} = \exp\left(-\frac{1}{4}\sqrt{\frac{L}{C}}\right), \quad (2.21)$$

as derived in Equation (2.7). Correspondingly, the energy is very insensitive to the state of the qubit, which determines the sign of the cosine potential. This suppression factor determines the characteristic scale of the error in the phase gate.

As we turn on the Josephson coupling  $J(t)$  with the form shown in Figure 2.7, the characteristic time  $\tau_J$  for the coupling to ramp on and off is subject to some constraints, which we will specify shortly. With  $J$  at its steady state value  $J_0$ , phase slips (tunneling events between successive minima of the cosine potential) are suppressed by the WKB factor

$$\exp\left(-\int_0^{2\pi} d\varphi \sqrt{2J_0C(1-\cos\varphi)}\right) = \exp(-8\sqrt{J_0C}). \quad (2.22)$$

We assume that  $\sqrt{J_0C}$  is large enough so that phase slips can be safely neglected.

A diabatic transition that excites the oscillator in the cosine well is most likely to occur while  $J(t)C$  is approximately one and the frequency of oscillations in the well is approximately  $1/C$ . We will pass through this regime once when ramping up the gate and once again ramping down. Landau-Zener theory indicates that the probability  $P_{\text{diab}}$  of such a transition scales like

$$P_{\text{diab}}(\tau_J) \sim \exp\left(-(\text{constant})\frac{\tau_J}{C}\right), \quad (2.23)$$

where  $\tau_J$  is the characteristic time for  $J(t)$  to ramp on. (We will discuss this error in more detail in Section 2.7.) Since diabatic effects also contribute to the error in the phase gate, we require  $\tau_J \gg C$ . Indeed, the diabatic error is comparable to the intrinsic error in Equation (2.21) for

$$\tau_J \sim \sqrt{LC}; \quad (2.24)$$

that is, when the ramping time is on the order of the period of the  $LC$  oscillator. During this ramping time, the envelope function of the Gaussian grid state is squeezed somewhat in  $\varphi$  space (and correspondingly spreads somewhat in  $Q$  space), but stays broad enough for the intrinsic error to remain heavily suppressed. In Section 2.8 we argue that the error arising from squeezing scales like

$$P_{\text{sq}}(\tau_J) \sim \exp\left(-(\text{constant})\frac{L}{\tau_J}\right); \quad (2.25)$$

hence, it too is comparable to the intrinsic error for  $\tau_J \sim \sqrt{LC}$ .

We will argue that under appropriate conditions  $\langle \psi_1^{\text{fin}} | \psi_0^{\text{fin}} \rangle \approx 1$  to extremely high accuracy so that the phase gate is nearly perfect. Note that we need not require the final state of the oscillator to match the initial state  $|\psi^{\text{in}}\rangle$ ; noise terms in the Hamiltonian may excite the oscillator, but the phase gate is still highly reliable as long as the oscillator's final state depends only very weakly on the state of the  $0-\pi$  qubit, *i.e.*, on whether the sign of  $J(t)$  is positive or negative. Indeed, the oscillator serves as a reservoir that absorbs the entropy introduced by noise. If not too badly damaged, the oscillator can be reused a few times for the execution of additional protected gates. Eventually, though, it will become too highly excited, and will need to be cooled before being employed again.

A gate error may arise if the coupling between qubit and oscillator remains on for too long or too short a time, *i.e.*, if  $\tilde{t} = 1 + \varepsilon$  rather than  $\tilde{t} = 1$ . But we will see that such timing errors do not much compromise the performance of the gate when  $\varepsilon$  is small; specifically, the gate error is  $\exp\left(-\frac{1}{4}\sqrt{L/C}\right) \times O(1)$  provided  $|\varepsilon| < 2\pi(L/C)^{-3/4}$ . Slightly overrotating or underrotating contributes to the damage suffered by the oscillator, but without much enhancing the sensitivity of the oscillator's final state to the qubit's state, and hence without much reducing the fidelity of the gate. We study the consequences of overrotation/underrotation in Section 2.6, and we confirm our findings using numerical simulations in Section 2.9. We also argue, in Sections 2.10 and 2.11, that the phase gate is robust against a sufficiently small nonzero temperature and against small perturbations in the Hamiltonian.

Let us summarize the sufficient conditions for the phase gate to be well protected. Just as for the realization of the  $0-\pi$  qubit itself, the execution of the protected phase gate relies on the construction of a “superinducting” circuit with  $\sqrt{L/C} \gg 1$ . This is a daunting experimental challenge, as we have already noted at the end of Section 2.2. To ensure high gate accuracy, we also assume that the steady state value  $J_0$  of the Josephson coupling between the  $0-\pi$  qubit and the oscillator satisfies  $\sqrt{J_0 C} \gg 1$ , and that the characteristic time scale  $\tau_J$  for the coupling to ramp on and off is  $O(\sqrt{LC})$ ; thus  $\tau_J$  is also small compared to the time  $L/\pi$  needed to execute the gate. Under these conditions, the error in the phase gate scales as  $\exp\left(-O\left(\sqrt{L/C}\right)\right)$ , and is stable with respect to small fluctuations in the implementation of the gate.

## 2.4 Sketch of the error estimate

A noisy quantum gate realizes a quantum operation  $\mathcal{N}_{\text{actual}}$ , and a useful way to quantify the error in the gate is to specify the deviation  $\|\mathcal{N}_{\text{actual}} - \mathcal{N}_{\text{ideal}}\|_{\diamond}$  from the ideal gate  $\mathcal{N}_{\text{ideal}}$  in the “diamond norm” [6]. As explained in Appendix 2.B, for the protected phase gate this diamond norm distance (assuming there are no bit flips) is

$$\|\mathcal{N}_{\text{actual}} - \mathcal{N}_{\text{ideal}}\|_{\diamond} = |1 - \langle \psi_1^{\text{fin}} |, \rangle \psi_0^{\text{fin}} | \quad (2.26)$$

where  $|\psi_{0,1}^{\text{fin}}\rangle$  denotes the final state of the oscillator when  $|0\rangle, |1\rangle$  is the state of the  $0-\pi$  qubit, as in Equation (2.20). Thus, we assess the gate accuracy by estimating the deviation of  $\langle \psi_0^{\text{fin}} | \psi_1^{\text{fin}} \rangle$  from 1.

To perform this estimate we track how the oscillator states  $|\psi_0(t)\rangle$  and  $|\psi_1(t)\rangle$  are related through three stages of evolution:

$$\begin{aligned} |\psi^{\text{in}}\rangle &\xrightarrow{J(t) \text{ turns on}} |\psi_0^{\text{begin}}\rangle \xrightarrow{J(t)=J_0} |\psi_0^{\text{end}}\rangle \xrightarrow{J(t) \text{ turns off}} |\psi_0^{\text{fin}}\rangle, \\ |\psi^{\text{in}}\rangle &\xrightarrow{J(t) \text{ turns on}} |\psi_1^{\text{begin}}\rangle \xrightarrow{J(t)=J_0} -i|\psi_1^{\text{end}}\rangle \xrightarrow{J(t) \text{ turns off}} -i|\psi_1^{\text{fin}}\rangle. \end{aligned} \quad (2.27)$$

In the first stage  $J(t)$  ramps on and the grid states are prepared – the initial state  $|\psi^{\text{in}}\rangle$  evolves to  $|\psi_0^{\text{begin}}\rangle$  if the  $0-\pi$  qubit’s state is  $|0\rangle$  and to  $|\psi_1^{\text{begin}}\rangle$  if the qubit’s state is  $|1\rangle$ . In the second stage  $J(t) = J_0$  and the grid state  $|\psi_0^{\text{begin}}\rangle$  evolves to  $|\psi_0^{\text{end}}\rangle$  while the grid state  $|\psi_1^{\text{begin}}\rangle$  evolves to  $-i|\psi_1^{\text{end}}\rangle$ , where ideally  $|\psi_{0,1}^{\text{end}}\rangle = |\psi_{0,1}^{\text{begin}}\rangle$ . In the third stage  $J(t)$  ramps off and the grid states  $|\psi_{0,1}^{\text{end}}\rangle$  evolve to the final oscillator states  $|\psi_{0,1}^{\text{fin}}\rangle$ , where ideally  $|\psi_1^{\text{fin}}\rangle = |\psi_0^{\text{fin}}\rangle$ .

Consider the first (or third) stage of the evolution, where the coupling  $J(t)$  ramps on (or off) in a time of order  $\tau_J$ . If  $\tau_J$  is not too large compared to the period  $2\pi\sqrt{LC}$  of the oscillator, then the harmonic potential term  $\varphi^2/2L$  may be treated perturbatively during this evolution stage. Hence, in first approximation the Hamiltonian is one of

$$H_0 = \frac{Q^2}{2C} - J(t) \cos \varphi, \quad H_1 = \frac{Q^2}{2C} + J(t) \cos \varphi, \quad (2.28)$$

depending on whether the state of the  $0-\pi$  qubit is  $|0\rangle$  or  $|1\rangle$ . This Hamiltonian commutes with the operator  $e^{-2\pi i Q}$ , which translates  $\varphi$  by  $2\pi$ ; therefore  $e^{-2\pi i Q}$  and the Hamiltonian can be

simultaneously diagonalized. We may express the eigenvalue of this translation operator as  $e^{-2\pi i q}$ , where  $q = Q - [Q] \in [-\frac{1}{2}, \frac{1}{2}]$  is the conserved Bloch momentum, and  $[Q]$  denotes the nearest integer to  $Q$ ; thus  $[Q]$  labels the distinct bands in the Hamiltonian's spectrum.

A diabatic transition between bands may be excited while  $J(t)$  varies, changing the value of  $[Q]$  by an integer, most likely  $\pm 1$ . If such transitions occur with nonnegligible probability, the final state of the oscillator will contain, in addition to a primary peak supported near  $Q = 0$ , also secondary peaks supported near  $Q = \pm 1$ ; the phases of the secondary peaks depend on whether the Hamiltonian is  $H_0$  or  $H_1$ , and therefore diabatic transitions contribute to the gate error. The probability of a diabatic transition cannot be computed precisely, but, as we will explain in Section 2.7, it can be analyzed semi-quantitatively, and is very small if  $\tau_J$  is sufficiently large.

For the purpose of discussing this diabatic error and other contributions to the deviation of  $\langle \psi_1^{\text{fin}} | \psi_0^{\text{fin}} \rangle$  from 1, we will find it useful to consider the operator

$$\begin{aligned} \bar{X} &\equiv (-1)^{[Q]} = \Pi_{\text{even}}^Q - \Pi_{\text{odd}}^Q \\ &= 2\Pi_{\text{even}}^Q - I = I - 2\Pi_{\text{odd}}^Q. \end{aligned} \tag{2.29}$$

Here  $\Pi_{\text{even}}^Q$  projects onto values of  $Q$  such that the nearest integer value  $[Q]$  is even and  $\Pi_{\text{odd}}^Q$  projects onto values of  $Q$  such that  $[Q]$  is odd. We denote this operator as  $\bar{X}$  because it can be regarded as the error-corrected Pauli operator  $\sigma^X$  acting on a qubit encoded in the Hilbert space of the oscillator, as we explain in Section 2.5. (Note that  $\bar{X}^2 = I$ .) Another (related) important property is that, since  $e^{\mp i\varphi}$  translates  $Q$  by  $\pm 1$ ,  $\bar{X}$  anticommutes with  $\cos \varphi$ :

$$\bar{X} \cos \varphi \bar{X} = -\cos \varphi. \tag{2.30}$$

Our argument showing that  $|\psi_1^{\text{fin}}\rangle \approx |\psi_0^{\text{fin}}\rangle$  has two main elements. On the one hand, we use approximate symmetries and properties of grid states to see that  $|\psi_1(t)\rangle \approx \bar{X}|\psi_0(t)\rangle$  at each stage of the oscillator's evolution, so that in particular  $|\psi_1^{\text{fin}}\rangle \approx \bar{X}|\psi_0^{\text{fin}}\rangle$ . On the other hand, we argue that if the time scale  $\tau_J$  for  $J(t)$  to turn on and off is suitably chosen, then the oscillator's final state is mostly supported near  $Q = 0$ , so that in particular  $\bar{X}|\psi_0^{\text{fin}}\rangle \approx |\psi_0^{\text{fin}}\rangle$ .

We note that the approximate Hamiltonians  $H_0$  and  $H_1$  in Equation (2.28) are related by

$$H_1 = \bar{X} H_0 \bar{X}. \tag{2.31}$$

By integrating the Schrödinger equation using the Hamiltonian  $H_0$  or  $H_1$  during the first stage of evolution while  $J(t)$  ramps on, we obtain the unitary time evolution operators  $U_0, U_1$ , which are related by

$$U_1 = \bar{X}U_0\bar{X}. \quad (2.32)$$

Thus, the initial oscillator state  $|\psi^{\text{in}}\rangle$  evolves to one of the states

$$\begin{aligned} |\psi_0^{\text{begin}}\rangle &= U_0|\psi^{\text{in}}\rangle \\ |\psi_1^{\text{begin}}\rangle &= U_1|\psi^{\text{in}}\rangle = \bar{X}U_0\bar{X}|\psi^{\text{in}}\rangle, \end{aligned} \quad (2.33)$$

and therefore

$$\begin{aligned} \langle \psi_1^{\text{begin}} | \bar{X} | \psi_0^{\text{begin}} \rangle &= \langle \psi^{\text{in}} | \bar{X} | \psi^{\text{in}} \rangle \\ &= \langle \psi^{\text{in}} | I - 2\Pi_{\text{odd}}^Q | \psi^{\text{in}} \rangle \\ &= 1 - 2\langle \psi^{\text{in}} | \Pi_{\text{odd}}^Q | \psi^{\text{in}} \rangle. \end{aligned} \quad (2.34)$$

We conclude that if the initial state is almost fully supported on even values of  $[Q]$  (for example, the oscillator ground state, a Gaussian in  $Q$ -space with width much less than  $1/2$ ), then  $|\psi_1^{\text{begin}}\rangle$  is very close to  $\bar{X}|\psi_0^{\text{begin}}\rangle$ .

So far, we have ignored the effects of the quadratic term  $\varphi^2/2L$  in the potential. This term can cause the wave function to broaden in  $Q$ -space and be squeezed in  $\varphi$  space, but we argue in Section 2.8 that this squeezing is a relatively small effect, so that the conclusion  $|\psi_1^{\text{begin}}\rangle \approx \bar{X}|\psi_0^{\text{begin}}\rangle$  still holds accurately. Specifically, the contribution to the gate error due to squeezing scales according to Equation (2.25), and hence becomes comparable to the other sources of error when we choose  $\tau_J \sim \sqrt{LC}$ .

During the second stage of the evolution, while  $J(t) = J_0$  is held constant, distinct peaks in the grid state acquire relative phases, and the condition  $|\psi_1(t)\rangle = \bar{X}|\psi_0(t)\rangle$  becomes badly violated. However, after a time  $t \approx L/\pi$ , the initial states  $|\psi_0^{\text{begin}}\rangle$  and  $|\psi_1^{\text{begin}}\rangle$  are restored, aside from the state dependent phase  $-i$ , and hence  $|\psi_1^{\text{end}}\rangle = \bar{X}|\psi_0^{\text{end}}\rangle$ , apart from a small error. Equivalently, the beginning states

$$|\psi_{\pm}^{\text{begin}}\rangle = \frac{1}{\sqrt{2}} \left( |\psi_0^{\text{begin}}\rangle \pm |\psi_1^{\text{begin}}\rangle \right) \quad (2.35)$$

are very nearly  $\bar{X}$  eigenstates with eigenvalues  $\pm 1$ , and this property is preserved by the ending states

$$|\psi_{\pm}^{\text{end}}\rangle = \frac{1}{\sqrt{2}} (|\psi_0^{\text{end}}\rangle \pm |\psi_1^{\text{end}}\rangle). \quad (2.36)$$

The  $\bar{X}$  eigenvalues of these states are highly stable with respect to timing errors in the gate, in which the coupling is left on for too short or too long a time, because these states are approximate codewords of a quantum code, well protected against logical phase errors. We study the errors resulting from imperfect timing in detail in Section 2.6, because they can be calculated explicitly and are the dominant errors in some parameter regimes.

In the third stage of the evolution, as in the first stage, it is a good first approximation to ignore the harmonic  $\varphi^2/2L$  term in the potential as the coupling  $J(t)$  ramps off. Using this approximation, the time evolution operators  $V_{0,1}$ , obtained by integrating the Schrödinger equation during the third stage when the state of the  $0-\pi$  qubit is  $|0\rangle, |1\rangle$ , are related by

$$V_1 = \bar{X}V_0\bar{X}; \quad (2.37)$$

hence, the final oscillator states are

$$|\psi_0^{\text{fin}}\rangle = V_0|\psi_0^{\text{end}}\rangle, \quad |\psi_1^{\text{fin}}\rangle = \bar{X}V_0\bar{X}|\psi_1^{\text{end}}\rangle, \quad (2.38)$$

and we conclude that

$$\langle\psi_1^{\text{fin}}|\bar{X}|\psi_0^{\text{fin}}\rangle = \langle\psi_1^{\text{end}}|\bar{X}|\psi_0^{\text{end}}\rangle. \quad (2.39)$$

Again, this conclusion is not modified much when the  $\varphi^2/2L$  term is properly taken into account, so we may infer that the condition  $|\psi_0(t)\rangle \approx \bar{X}|\psi_1(t)\rangle$  is well preserved during the final stage of evolution.

We have now seen that  $|\psi_1^{\text{fin}}\rangle \approx \bar{X}|\psi_0^{\text{fin}}\rangle$ , and it remains to show that  $\bar{X}|\psi_0^{\text{fin}}\rangle \approx |\psi_0^{\text{fin}}\rangle$ . This condition will be well satisfied, provided that the final state  $|\psi_0^{\text{fin}}\rangle$  of the oscillator, like the initial state  $|\psi^{\text{in}}\rangle$ , is almost fully supported in the interval  $Q \in [-\frac{1}{2}, \frac{1}{2}]$ . Logical errors may occur because of diabatic transitions between bands, which may change  $Q$  by an odd integer, or because of spreading in  $Q$  space, which may enhance the tails of the wave function outside  $[-\frac{1}{2}, \frac{1}{2}]$ . However, if diabatic transitions are rare and spreading is modest, as we expect if  $\tau_J$  lies in the appropriate range, then the gate will be highly accurate.

That our criterion for achieving  $|\psi_1^{\text{fin}}\rangle \approx |\psi_0^{\text{fin}}\rangle$  involves the operator  $\bar{X}$ , which has a sharp discontinuity at  $Q = \frac{1}{2} + \text{integer}$ , is really an artifact of an insufficiently careful treatment of diabatic transitions. The transitions occur with enhanced probability for  $Q$  close to  $\frac{1}{2} + \text{integer}$ , replacing the sharp edge in  $Q$  space by a rounded step with width of order  $C/\tau_J$ , as we will explain in Section 2.7.

## 2.5 Encoding a qubit in an oscillator

A continuous-variable quantum error-correcting code [43] underlies the robustness of the protected phase gate. The theory of quantum codes is not really essential for understanding our estimate of the gate accuracy, but this theory provides motivation for our construction of the protected gate, as well a convenient language for explaining how it works. Therefore, we will now review some of the relevant features of a code first described in [43].

In the version of the code we will use, a two-dimensional encoded qubit is embedded in the infinite-dimensional Hilbert space of a harmonic oscillator with position  $\varphi$  and conjugate momentum  $Q$ , satisfying  $[\varphi, Q] = i$ . The code space can be specified as the simultaneous eigenspace with eigenvalue 1 of the two commuting operators

$$M_Z = e^{2i\varphi}, \quad M_X = e^{-2\pi i Q}; \quad (2.40)$$

we say that  $M_Z$  and  $M_X$  are the generators of the code's "stabilizer group." Using the identity

$$e^A e^B = e^{[A,B]} e^B e^A \quad (2.41)$$

(where  $A$  and  $B$  commute with their commutator), we can easily verify that  $M_X$  and  $M_Z$  commute. The logical Pauli operators acting on the encoded qubit are

$$\bar{Z} = e^{i\varphi}, \quad \bar{X} = e^{-i\pi Q}. \quad (2.42)$$

One sees that  $\bar{X}$  and  $\bar{Z}$  commute with the stabilizer generators  $M_X$  and  $M_Z$ , and hence preserve the code space; furthermore they anticommute with one another, as the logical Pauli operators should.

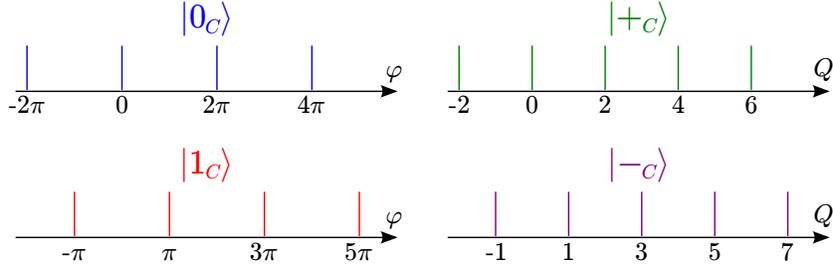


Figure 2.9: Ideal codewords of the continuous variable code. The  $\bar{Z} = \pm 1$  eigenstates  $|0_C\rangle, |1_C\rangle$ , expressed in  $\varphi$  space, are uniform superpositions of position eigenstates with  $\varphi$  an even or odd multiple of  $\pi$ , respectively. The  $\bar{X} = \pm 1$  eigenstates  $|+_C\rangle, |-_C\rangle$ , expressed in  $Q$  space, are uniform superpositions of momentum eigenstates with  $Q$  an even or odd integer, respectively.

The (unnormalizable) state  $|0_C\rangle^{\text{ideal}}$  is the unique  $\bar{Z}$  eigenstate with eigenvalue 1 in the code space. The condition  $\bar{Z} = 1$  requires the variable  $\varphi$  to be an integer multiple of  $2\pi$ , and the condition  $M_X = 1$  requires the codeword to be invariant under translation of  $\varphi$  by  $2\pi$ . Hence,  $|0_C\rangle^{\text{ideal}}$  is represented in  $\varphi$  space as the uniform superposition of delta functions

$$|0_C\rangle^{\text{ideal}} = \sum_{n=-\infty}^{\infty} |\varphi = 2\pi n\rangle; \quad (2.43)$$

the  $\bar{Z} = -1$  eigenstate  $|1_C\rangle^{\text{ideal}} = \bar{X}|0_C\rangle^{\text{ideal}}$ , obtained from  $|0_C\rangle^{\text{ideal}}$  by displacing  $\varphi$  by  $\pi$ , is

$$|1_C\rangle^{\text{ideal}} = \sum_{n=-\infty}^{\infty} |\varphi = 2\pi(n + \frac{1}{2})\rangle. \quad (2.44)$$

Similarly, the  $X = \pm 1$  eigenstates  $|\pm_C\rangle^{\text{ideal}}$ , invariant under translation of  $Q$  by 2, are represented in  $Q$  space as

$$\begin{aligned} |+_C\rangle^{\text{ideal}} &= \sum_{n=-\infty}^{\infty} |Q = 2n\rangle, \\ |-_C\rangle^{\text{ideal}} &= \sum_{n=-\infty}^{\infty} |Q = 2(n + \frac{1}{2})\rangle. \end{aligned} \quad (2.45)$$

This is depicted in Figure 2.9.

Weak noise may displace  $\varphi$  slightly, but the codewords  $|0_C\rangle^{\text{ideal}}$  and  $|1_C\rangle^{\text{ideal}}$  remain perfectly

distinguishable, and the error is correctable, as long as the value of  $\varphi$  shifts by less than  $\pi/2$  in either direction. Similarly, a shift in  $Q$  by less than  $1/2$  is correctable. In principle, we could diagnose the error by measuring  $M_Z$  to determine the value of  $\varphi$  modulo  $\pi$ , and  $M_X$  to determine the value of  $Q$  modulo 1, and then perform active error correction by applying the minimal shifts in  $\varphi$  and  $Q$  that return the damaged code state to the code space. (In our protected phase gate, however, the error correction will be passive rather than active.)

The unnormalizable ideal codewords, with infinite  $\langle\varphi\rangle$  and  $\langle Q\rangle$ , are unphysical. But if we coherently apply Gaussian distributed small shifts in  $\varphi$  and  $Q$  to the ideal codewords, we obtain the normalizable approximate codewords shown in Figure 2.10. The wave function in  $\varphi$  space is a superposition of narrow Gaussians, each of width  $\Delta \ll \pi/2$  (*i.e.*,  $\langle(\delta\varphi)^2\rangle = \frac{1}{2}\Delta^2$ , where  $\delta\varphi$  denotes the deviation from the center of the narrow Gaussian), governed by a broad Gaussian envelope with width  $\kappa^{-1} \gg 2$  (*i.e.*,  $\langle\varphi^2\rangle = \frac{1}{2}\kappa^{-2}$ ). The Fourier dual wave function in  $Q$  space is a superposition of narrow Gaussians, each of width  $\kappa$  (*i.e.*,  $\langle(\delta Q)^2\rangle = \frac{1}{2}\kappa^2$ , where  $\delta Q$  denotes the deviation from the center of the narrow Gaussian), governed by a broad Gaussian envelope with width  $\Delta^{-1}$  (*i.e.*,  $\langle Q^2\rangle = \frac{1}{2}\Delta^{-2}$ ). If  $\Delta$  and  $\kappa$  are sufficiently small, these approximate codewords retain good error correction properties. However, there is now an intrinsic error arising from the tails of the narrow Gaussians, with the probability of a logical  $\bar{Z}$  error (a shift in  $\varphi$  by more than  $\pi/2$ ) suppressed by  $e^{-\pi^2/4\Delta^2}$ , and the probability of a logical  $\bar{X}$  error (a shift in  $Q$  by more than  $1/2$ ) suppressed by  $e^{-1/4\kappa^2}$ .

Note that in Section 2.4 we used the notation  $\bar{X} = (-1)^{[Q]}$  for the logical  $X$  operator, where  $[Q]$  denotes the nearest integer to the real variable  $Q$ . The operator  $(-1)^{[Q]}$  acts in the same way as  $e^{-i\pi Q}$  on ideal codewords for which  $Q$  is an integer. By expressing the logical operator as  $\bar{X} = (-1)^{[Q]}$ , we are implicitly correcting phase errors that displace  $Q$ . That is, a  $Q$  eigenstate is decoded by shifting  $Q$  to the nearest integer value, and the eigenvalue of  $\bar{X}$  is determined by this ideal shifted value of  $Q$ , rather than the actual value of  $Q$  prior to the shift.

The first step in the execution of the protected phase gate described in Section 2.3 is the preparation of just such approximate codewords; compare Figure 2.8. If the state of the  $0\text{-}\pi$  qubit is  $|0\rangle$ , then the approximate  $\bar{Z} = 1$  eigenstate  $|0_C\rangle$  is prepared, and if the state of the  $0\text{-}\pi$  qubit is  $|1\rangle$ , then the approximate  $\bar{Z} = -1$  eigenstate  $|1_C\rangle$  is prepared. The narrowly peaked functions have width  $\Delta^2 = (J_0 C)^{-1/2}$  in  $\varphi$  space (though, because the potential is a cosine rather than harmonic, the tail of the peaked function decays more slowly than the tail of a Gaussian),

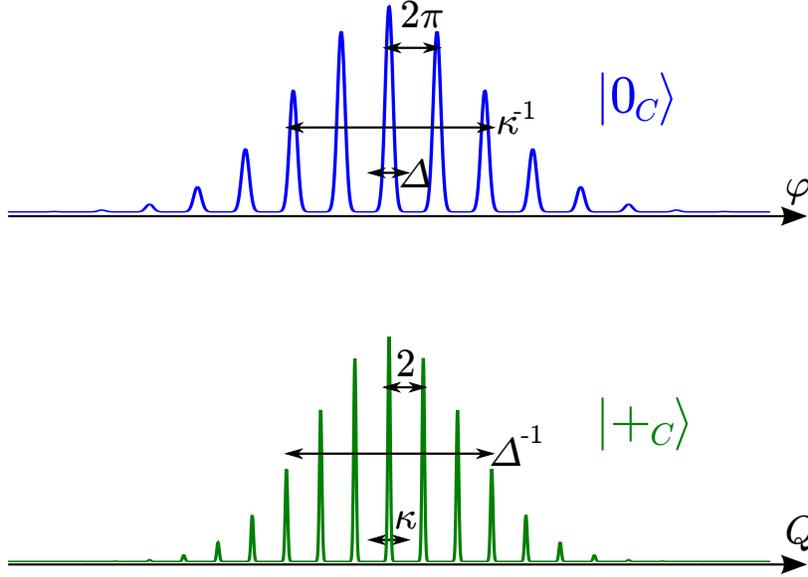


Figure 2.10: Approximate codewords of the continuous variable code. The codeword  $|0_C\rangle$ , expressed in  $\varphi$  space, is a superposition of Gaussian peaks, each of width  $\Delta$ , governed by a broad Gaussian envelope with width  $\kappa^{-1}$ . The codeword  $|+_C\rangle$ , expressed in  $Q$  space, is a superposition of Gaussian peaks, each of width  $\kappa$ , governed by a broad Gaussian envelope with width  $\Delta^{-1}$ .

and width  $\kappa^2 = (L/C)^{-1/2}$  in  $Q$  space. Hence, the intrinsic logical  $\bar{X}$  error of the approximate codewords, which is central to our estimate of the error in the phase gate, is suppressed by the factor  $\exp\left(-\frac{1}{4}\sqrt{L/C}\right)$ .

After the approximate codeword is prepared, the Gaussian unitary operator  $e^{-it\varphi^2/2L} = e^{-i\tilde{t}\varphi^2/2\pi}$  is applied (where  $\tilde{t} = \pi t/L$  is a rescaled time variable). This unitary operator rotates the code space, transforming the stabilizer generator  $M_X = e^{-2\pi i Q}$  according to

$$\begin{aligned} M_X &\rightarrow M'_X = e^{-i\tilde{t}\varphi^2/2\pi} e^{-2\pi i Q} e^{i\tilde{t}\varphi^2/2\pi} \\ &= e^{-2\pi i(Q + \varphi\tilde{t}/\pi)} = M_X e^{-2i\varphi\tilde{t}} e^{-2\pi i\tilde{t}}. \end{aligned} \quad (2.46)$$

Recalling that  $M_Z^{-1} = e^{-2i\varphi}$  is also a stabilizer generator, we see that the state returns to the code space at (rescaled) time  $\tilde{t} = 1$ , but during its excursion the codeword acquires a Berry phase, and thus a nontrivial logical operation is applied. Specifically, the logical operator  $\bar{X}$  is transformed

according to

$$\begin{aligned}
\bar{X} &\rightarrow \bar{X}' = e^{-i\varphi^2/2\pi} e^{-i\pi Q} e^{i\varphi^2/2\pi} \\
&= e^{-i\pi(Q+\varphi/\pi)} = \bar{X} e^{-i\varphi} e^{-i\pi/2} \\
&= -i\bar{X}\bar{Z} = \exp\left(i\frac{\pi}{4}\bar{Z}\right) \bar{X} \exp\left(-i\frac{\pi}{4}\bar{Z}\right), \tag{2.47}
\end{aligned}$$

while  $\bar{Z}$  remains invariant; hence, the logical operation acting on the code space is  $\exp\left(i\frac{\pi}{4}\bar{Z}\right)$ . An error in the logical gate arises if the coupling remains on for too long or too short a time (*i.e.*, if  $\tilde{t}$  is not precisely 1). However, this error is correctable with high probability if the timing error is small. We will study the consequences of overrotation/underrotation in Section 2.6.

## 2.6 Imperfect grid states

Now we will analyze the intrinsic phase errors in approximate codewords of the continuous variable code, and, in particular, how the phase error is affected by errors in the timing of the pulse that executes the phase gate.

In Section 2.5 we considered approximate codewords that can be described as ‘‘Gaussian grid states,’’ where the codeword is a superposition of narrow Gaussian peaks governed by a broad Gaussian envelope. The Fourier transform of such a wave function is also a Gaussian grid state, so that both logical bit flip errors and logical phase errors are suppressed.

But it actually suffices for the approximate codeword to be a superposition of narrow functions with a broad envelope; neither the peak nor the envelope needs to be Gaussian. Even a non-Gaussian grid is mapped to a conjugate non-Gaussian grid by the Fourier transform, so there is good protection against both  $\bar{X}$  and  $\bar{Z}$  errors. A Gaussian grid state could result from coherently applying Gaussian-distributed  $\varphi$  and  $Q$  shifts to an ideal codeword, assuming large shifts are suppressed. But we can also get a reasonable approximate codeword by applying more general small errors to the ideal codeword, with a distribution that is not necessarily Gaussian. What is important is that large shifts in both  $\varphi$  and  $Q$  are improbable, not the detailed form of the distribution.

This observation will be useful when we consider in Section 2.10 the execution of the protected phase gate in the case where the initial state of the harmonic oscillator is an excited state rather than the ground state. In that case, the envelope of the approximate codeword in  $\varphi$  space is not strictly

Gaussian, but rather Gaussian modulated by a Hermite polynomial, and its Fourier transform is also Gaussian modulated by a Hermite polynomial. Thus, in  $Q$  space, the narrow functions peaked at integer values of  $Q$  are also oscillator excited states. These functions have highly suppressed tails, ensuring that encoded phase errors are rare. In any event, considering more general kinds of grid states helps to clarify conceptually why the phase gate is robust.

### 2.6.1 Bit-flip and phase errors

Let  $f$  denote a narrow function in  $\varphi$  space, and  $\tilde{F}$  denote a broad envelope function in  $\varphi$  space. We express the approximate codewords as

$$\begin{aligned} |0_C\rangle &= \sqrt{2\pi} \sum_{n \text{ even}} \tilde{F}(\pi n) \int d\varphi f(\varphi - \pi n) |\varphi\rangle, \\ |1_C\rangle &= \sqrt{2\pi} \sum_{n \text{ odd}} \tilde{F}(\pi n) \int d\varphi f(\varphi - \pi n) |\varphi\rangle. \end{aligned} \quad (2.48)$$

The function  $f$  is normalized so that

$$\int |f(\varphi)|^2 d\varphi = 1, \quad (2.49)$$

and if the overlap between peaks centered at distinct integer multiples of  $\pi$  can be neglected, then  $|0_C\rangle$  and  $|1_C\rangle$  are normalized, provided

$$2\pi \sum_{n \text{ even}} |\tilde{F}(\pi n)|^2 \approx 1, \quad 2\pi \sum_{n \text{ odd}} |\tilde{F}(\pi n)|^2 \approx 1. \quad (2.50)$$

The intrinsic bit-flip error of the approximate codeword  $|0_C\rangle$  arises from the probability that  $\varphi$  lies closer to an odd multiple of  $\pi$  than to an even multiple, which can be estimated as

$$\begin{aligned} P_{\text{error}}^{|0_C\rangle} &\approx 2\pi \sum_{n \text{ even}} |\tilde{F}(\pi n)|^2 \times \left( \int_{-\infty}^{-\pi/2} d\varphi |f(\varphi)|^2 + \int_{\pi/2}^{\infty} d\varphi |f(\varphi)|^2 \right) \\ &\approx \int_{-\infty}^{-\pi/2} d\varphi |f(\varphi)|^2 + \int_{\pi/2}^{\infty} d\varphi |f(\varphi)|^2; \end{aligned} \quad (2.51)$$

the intrinsic error in  $|1_C\rangle$  can be estimated similarly. Thus, logical bit-flip errors are highly suppressed if  $f(\varphi)$  is a narrow, rapidly decaying function supported near zero.

The approximate codewords in the conjugate basis are

$$\begin{aligned} |+_C\rangle &= \frac{1}{\sqrt{2}} (|0_C\rangle + |1_C\rangle) \\ &= \sqrt{\pi} \sum_n \tilde{F}(\pi n) \int d\varphi f(\varphi - \pi n) |\varphi\rangle, \end{aligned} \quad (2.52)$$

$$\begin{aligned} |-_C\rangle &= \frac{1}{\sqrt{2}} (|0_C\rangle - |1_C\rangle) \\ &= \sqrt{\pi} \sum_n \tilde{F}(\pi n) (-1)^n \int d\varphi f(\varphi - \pi n) |\varphi\rangle, \end{aligned} \quad (2.53)$$

where

$$\pi \sum_n |\tilde{F}(\pi n)|^2 \approx 1. \quad (2.54)$$

We show in Appendix 2.C that these codewords can be expressed as

$$\begin{aligned} |+_C\rangle &= \sqrt{2} \int dQ \tilde{f}(Q) \sum_{m \text{ even}} F(Q - m) |Q\rangle, \\ &\approx \sqrt{2} \sum_{m \text{ even}} \tilde{f}(m) \int dQ F(Q - m) |Q\rangle, \end{aligned} \quad (2.55)$$

$$\begin{aligned} |-_C\rangle &= \sqrt{2} \int dQ \tilde{f}(Q) \sum_{m \text{ odd}} F(Q - m) |Q\rangle, \\ &\approx \sqrt{2} \sum_{m \text{ odd}} \tilde{f}(m) \int dQ F(Q - m) |Q\rangle, \end{aligned} \quad (2.56)$$

where

$$\begin{aligned} 2 \sum_{m \text{ even}} |\tilde{f}(m)|^2 &\approx \int dQ |\tilde{f}(Q)|^2 \approx 1, \\ 2 \sum_{m \text{ odd}} |\tilde{f}(m)|^2 &\approx \int dQ |\tilde{f}(Q)|^2 \approx 1. \end{aligned} \quad (2.57)$$

The intrinsic phase error of the approximate codeword  $|+_C\rangle$  arises from the probability that  $Q$  lies

closer to an odd integer than an even integer, which can be estimated as

$$\begin{aligned}
P_{\text{error}}^{|-c\rangle} &\approx 2 \sum_{m \text{ even}} |\tilde{f}(m)|^2 \times \left( \int_{-\infty}^{-1/2} dQ |F(Q)|^2 + \int_{1/2}^{\infty} dQ |F(Q)|^2 \right) \\
&\approx \int_{-\infty}^{-1/2} dQ |F(Q)|^2 + \int_{1/2}^{\infty} dQ |F(Q)|^2;
\end{aligned} \tag{2.58}$$

the intrinsic error in  $|-c\rangle$  is estimated similarly. Thus logical phase errors are highly suppressed if  $F(Q)$  is a narrow, rapidly decaying function supported near zero.

### 2.6.2 Gate error estimate

As explained in Section 2.4, the error in the protected phase gate can be expressed as

$$|1 - \langle \psi_1^{\text{fin}} | \psi_0^{\text{fin}} \rangle|, \tag{2.59}$$

where  $|\psi_{0,1}^{\text{fin}}\rangle$  denotes the final state of the oscillator (modulo the state dependent phase  $-i$  applied by the gate) when the state of the  $0-\pi$  qubit is  $|0\rangle, |1\rangle$ . Under conditions enumerated in Section 2.4, this quantity can be well approximated by the modulus of

$$\eta \equiv 1 - \langle \psi_1^{\text{end}} | \bar{X} | \psi_0^{\text{end}} \rangle, \tag{2.60}$$

where  $|\psi_{0,1}^{\text{end}}\rangle$  denotes the state of the oscillator as the coupling  $J(t)$  between oscillator and qubit starts to turn off, and  $\bar{X} = \Pi_{\text{even}}^Q - \Pi_{\text{odd}}^Q$  denotes the error-corrected logical operator.

Let us suppose that the states  $|\psi_{0,1}^{\text{begin}}\rangle$ , prepared when the coupling  $J(t)$  turns on, are the approximate codewords  $|0_C\rangle, |1_C\rangle$  depicted in Equation (2.48), where  $f(\varphi)$  is a narrow rapidly decreasing function and  $\tilde{F}(\varphi)$  is a broad envelope function. The coupling remains on for time  $t = (1 + \varepsilon)\frac{L}{\pi}$ , where  $\varepsilon$  is the fractional error in the timing of the gate. Then, as explained in Appendix 2.C, the states  $|\psi_{0,1}^{\text{end}}\rangle$  have the same form as  $|0_C\rangle, |1_C\rangle$ , but with  $\tilde{F}(\varphi)$  replaced by the function

$$\tilde{F}_\varepsilon(\varphi) = e^{-i\varepsilon\varphi^2/2\pi} \tilde{F}(\varphi). \tag{2.61}$$

We define states

$$|\psi_\pm^{\text{end}}\rangle = \frac{1}{\sqrt{2}} (|\psi_0^{\text{end}}\rangle \pm |\psi_1^{\text{end}}\rangle); \tag{2.62}$$

note that  $|\psi_{0,1}^{\text{end}}\rangle$  are normalized, since each is obtained by applying a unitary time evolution operator to the normalized state  $|\psi^{\text{in}}\rangle$  of the oscillator, but they are not necessarily orthogonal and hence the states  $|\psi_{\pm}^{\text{end}}\rangle$  are not necessarily normalized. We may write

$$\langle \psi_1^{\text{end}} | \bar{X} | \psi_0^{\text{end}} \rangle = \frac{1}{2} (\langle \psi_+^{\text{end}} | \bar{X} | \psi_+^{\text{end}} \rangle - \langle \psi_-^{\text{end}} | \bar{X} | \psi_-^{\text{end}} \rangle + \langle \psi_+^{\text{end}} | \bar{X} | \psi_-^{\text{end}} \rangle - \langle \psi_-^{\text{end}} | \bar{X} | \psi_+^{\text{end}} \rangle), \quad (2.63)$$

which, using Equation (2.29), has real part

$$\begin{aligned} \text{Re} \langle \psi_1^{\text{end}} | \bar{X} | \psi_0^{\text{end}} \rangle &= \frac{1}{2} \left( \langle \psi_+^{\text{end}} | I - 2\Pi_{\text{odd}}^Q | \psi_+^{\text{end}} \rangle - \langle \psi_-^{\text{end}} | 2\Pi_{\text{even}}^Q - I | \psi_-^{\text{end}} \rangle \right) \\ &= 1 - \langle \psi_+^{\text{end}} | \Pi_{\text{odd}}^Q | \psi_+^{\text{end}} \rangle - \langle \psi_-^{\text{end}} | \Pi_{\text{even}}^Q | \psi_-^{\text{end}} \rangle. \end{aligned} \quad (2.64)$$

Therefore, using Equation (2.60) we may estimate the real part of the gate error as in Equation (2.58):

$$\text{Re} \eta_\varepsilon \approx 2 \left( \int_{-\infty}^{-1/2} dQ |F_\varepsilon(Q)|^2 + \int_{1/2}^{\infty} dQ |F_\varepsilon(Q)|^2 \right). \quad (2.65)$$

To estimate the imaginary part we note that

$$\text{Im} \eta_\varepsilon = \frac{1}{2i} \left( \langle \psi_-^{\text{end}} | \bar{X} | \psi_+^{\text{end}} \rangle - \langle \psi_+^{\text{end}} | \bar{X} | \psi_-^{\text{end}} \rangle \right). \quad (2.66)$$

As in Equation (2.55), we express

$$\begin{aligned} |\psi_-^{\text{end}}\rangle &\approx \sqrt{2} \left( \int_{[Q] \text{ even}} dQ + \int_{[Q] \text{ odd}} dQ \right) \tilde{f}(Q) \sum_{m \text{ odd}} F_\varepsilon(Q-m)|Q\rangle, \\ \bar{X} |\psi_+^{\text{end}}\rangle &\approx \sqrt{2} \left( \int_{[Q] \text{ even}} dQ - \int_{[Q] \text{ odd}} dQ \right) \tilde{f}(Q) \sum_{m \text{ even}} F_\varepsilon(Q-m)|Q\rangle, \end{aligned} \quad (2.67)$$

and therefore obtain

$$\begin{aligned} \langle \psi_-^{\text{end}} | \bar{X} | \psi_+^{\text{end}} \rangle &\approx 2 \int_{[Q] \text{ even}} dQ |\tilde{f}(Q)|^2 \sum_{\substack{m \text{ even} \\ n \text{ odd}}} F_\varepsilon(Q-n)^* F_\varepsilon(Q-m) \\ &\quad - 2 \int_{[Q] \text{ odd}} dQ |\tilde{f}(Q)|^2 \sum_{\substack{m \text{ even} \\ n \text{ odd}}} F_\varepsilon(Q-n)^* F_\varepsilon(Q-m). \end{aligned} \quad (2.68)$$

Thus,  $\text{Im} \eta_\varepsilon$  is dominated by overlaps between sharply peaked functions  $\{F_\varepsilon(Q-m)\}$  centered at

neighboring values of  $m$ ; we show in Appendix 2.C that

$$\text{Im } \eta_\varepsilon \approx 4 \text{Im} \int_0^\infty dQ \text{Odd} [F_\varepsilon(Q + \frac{1}{2})^* F_\varepsilon(Q - \frac{1}{2})], \quad (2.69)$$

where  $\text{Odd}[G(Q)] \equiv \frac{1}{2} (G(Q) - G(-Q))$  denotes the odd part of the function  $G(Q)$ .

### 2.6.3 Gaussian case

The gate error estimate in Equations (2.65) and (2.69) is expressed in terms of the narrow function  $F_\varepsilon(Q)$ , whose Fourier transform  $\tilde{F}_\varepsilon(\varphi)$  is the broad envelope function that governs the grid state of the oscillator. To be concrete, let us now suppose that this function is Gaussian, the relevant case where the oscillator's initial state  $|\psi^{\text{in}}\rangle$  is the ground state.

The normalized ground state wave function is

$$\begin{aligned} \tilde{F}(\varphi) &= \left(\frac{\kappa^2}{\pi}\right)^{1/4} e^{-\kappa^2 \varphi^2 / 2}, \\ F(Q) &= \left(\frac{1}{\pi \kappa^2}\right)^{1/4} e^{-Q^2 / 2\kappa^2}, \end{aligned} \quad (2.70)$$

where

$$\kappa^{-2} = \sqrt{\frac{L}{C}}. \quad (2.71)$$

Therefore, for  $\varepsilon = 0$  the probability of an intrinsic phase error in the grid state is

$$P_{\text{error}}^{|\text{+c}\rangle} \approx 2 \int_{1/2}^\infty dQ |F(Q)|^2 \approx 2 \sqrt{\frac{\kappa^2}{\pi}} e^{-1/4\kappa^2}, \quad (2.72)$$

using the leading asymptotic approximation to the error function.

To find the probability of a phase error for  $\varepsilon \neq 0$ , we evaluate

$$\begin{aligned} F_\varepsilon(Q) &= \frac{1}{\sqrt{2\pi}} \int d\varphi e^{-iQ\varphi} \tilde{F}_\varepsilon(\varphi) \\ &= \frac{1}{\sqrt{2\pi}} \int d\varphi e^{-iQ\varphi} \left(\frac{\kappa^2}{\pi}\right)^{1/4} e^{-\kappa^2 \varphi^2 / 2} e^{-i\varepsilon \varphi^2 / 2\pi} \\ &= \left(\frac{\kappa^2}{\pi \kappa'^4}\right)^{1/4} \exp(-Q^2 / 2\kappa'^2), \end{aligned} \quad (2.73)$$

where

$$\kappa'^2 = \kappa^2 + \frac{i\varepsilon}{\pi}. \quad (2.74)$$

Thus,

$$\begin{aligned} \kappa'^{-2} &= \kappa^{-2} \left( 1 + \frac{i\varepsilon}{\pi\kappa^2} \right)^{-1} \\ &= \bar{\kappa}^{-2} \left( 1 - \frac{i\varepsilon}{\pi\kappa^2} \right), \end{aligned} \quad (2.75)$$

where

$$\bar{\kappa}^2 = \kappa^2 \left( 1 + \frac{\varepsilon^2}{\pi^2\kappa^4} \right); \quad (2.76)$$

therefore

$$\begin{aligned} |F_\varepsilon(Q)|^2 &= \left( \frac{\kappa^2}{\pi|\kappa'|^4} \right)^{1/2} \exp(-Q^2 \operatorname{Re}(\kappa'^{-2})) \\ &= \frac{1}{\sqrt{\pi\bar{\kappa}^2}} e^{-Q^2/\bar{\kappa}^2}. \end{aligned} \quad (2.77)$$

From Equation (2.65), our estimate of the real part of the gate error becomes

$$\operatorname{Re} \eta_\varepsilon \approx \frac{4}{\sqrt{\pi\bar{\kappa}^2}} \int_{1/2}^{\infty} dQ e^{-Q^2/\bar{\kappa}^2} \approx 4\sqrt{\frac{\bar{\kappa}^2}{\pi}} e^{-1/4\bar{\kappa}^2}. \quad (2.78)$$

For  $\varepsilon$  small compared to  $\pi\kappa^2$ , we may expand

$$\bar{\kappa}^{-2} = \kappa^{-2} - \frac{\varepsilon^2}{\pi^2\kappa^6} + \dots, \quad (2.79)$$

so that

$$\operatorname{Re} \eta_\varepsilon \approx \exp\left(\frac{\varepsilon^2}{4\pi^2\kappa^6}\right) \operatorname{Re} \eta_{\varepsilon=0}; \quad (2.80)$$

the overrotation of the gate has little effect on the real part of the gate error for  $\varepsilon \ll 2\pi\kappa^3$ . On the other hand, when  $\varepsilon$  is large compared to  $\pi\kappa^2$ , we have

$$\bar{\kappa}^{-2} \approx \frac{\pi^2\kappa^2}{\varepsilon^2}; \quad (2.81)$$

thus,  $\text{Re } \eta_\varepsilon = O(1)$  for  $\varepsilon \approx \pi\kappa$ .

To see that these results are reasonable, note that

$$e^{-i\varepsilon\varphi^2/2\pi} Q e^{i\varepsilon\varphi^2/2\pi} = Q + \varphi\varepsilon/\pi. \quad (2.82)$$

Thus, crudely speaking, overrotation by  $\varepsilon$  shifts  $Q$  by

$$\delta Q \approx (\varepsilon/\pi) \langle \varphi^2 \rangle^{1/2} = \varepsilon/(\pi\sqrt{2\kappa^2}). \quad (2.83)$$

We expect this shift to have a small effect if the amount of the shift is small compared to the width  $\langle Q^2 \rangle^{1/2} = \sqrt{\kappa^2/2}$  of the narrow peaks in  $Q$ -space, *i.e.*, for  $\varepsilon \ll \pi\kappa^2$ . On the other hand, for  $\varepsilon \approx \pi\kappa$ , the shift in  $Q$  is  $O(1)$ , and we expect the error probability to be large.

To estimate the imaginary part of the gate error, we note that

$$\begin{aligned} F_\varepsilon(Q + \tfrac{1}{2})^* F_\varepsilon(Q - \tfrac{1}{2}) &= \frac{1}{\sqrt{\pi\bar{\kappa}^2}} \exp \left[ -\frac{(Q - \tfrac{1}{2})^2}{2\kappa'^2} - \frac{(Q + \tfrac{1}{2})^2}{2\kappa'^*{}^2} \right] \\ &= \frac{1}{\sqrt{\pi\bar{\kappa}^2}} \exp \left[ -(Q^2 + \tfrac{1}{4}) \text{Re}(\kappa'^{-2}) + iQ \text{Im}(\kappa'^{-2}) \right] \\ &= \frac{1}{\sqrt{\pi\bar{\kappa}^2}} e^{-1/4\bar{\kappa}^2} e^{-Q^2/\bar{\kappa}^2} \exp \left( \frac{-i\varepsilon Q}{\pi\bar{\kappa}^2\kappa'^2} \right), \end{aligned} \quad (2.84)$$

and from Equation (2.69) we obtain

$$\begin{aligned} \text{Im } \eta_\varepsilon &= -\frac{4}{\sqrt{\pi\bar{\kappa}^2}} e^{-1/4\bar{\kappa}^2} \int_0^\infty dQ e^{-Q^2/\bar{\kappa}^2} \sin(\varepsilon Q/\pi\bar{\kappa}^2\kappa'^2) \\ &= -\frac{4}{\sqrt{\pi}} e^{-1/4\bar{\kappa}^2} I \left( \frac{\varepsilon}{\pi\bar{\kappa}\kappa'^2} \right), \end{aligned} \quad (2.85)$$

where

$$I(\alpha) \equiv \int_0^\infty dx e^{-x^2} \sin(\alpha x). \quad (2.86)$$

The integral  $I(\alpha)$  can be expressed in terms of Gamma functions with imaginary arguments, but, for our purposes, it will suffice to observe some of its properties. For small  $\alpha$  it has the power series expansion

$$I(\alpha) = \frac{\alpha}{2} - \frac{\alpha^3}{12} + \dots, \quad (2.87)$$

for large  $\alpha$  it has the asymptotic expansion

$$I(\alpha) = \frac{1}{\alpha} + \frac{2}{\alpha^3} + \dots, \quad (2.88)$$

and it attains its maximum value  $I = .5410\dots$  at  $\alpha = 1.8483\dots$

Combining the real part in Equation (2.78) with the imaginary part in Equation (2.85), our estimate of the gate error becomes

$$|\eta_\varepsilon| = \frac{4}{\sqrt{\pi}} e^{-1/4\bar{\kappa}^2} \sqrt{\bar{\kappa}^2 + \left[ I\left(\frac{\varepsilon}{\pi\bar{\kappa}\kappa^2}\right) \right]^2}. \quad (2.89)$$

The ratio of the imaginary and real parts is

$$\frac{\text{Im } \eta_\varepsilon}{\text{Re } \eta_\varepsilon} = -\bar{\kappa}^{-1} I\left(\frac{\varepsilon}{\pi\bar{\kappa}\kappa^2}\right) \approx -\frac{\varepsilon}{2\pi\bar{\kappa}^2\kappa^2}, \quad (2.90)$$

expanding to linear order in  $\alpha$ . Thus, the imaginary part of the error is smaller than the real part when  $\varepsilon$  is sufficiently small, but dominates by a factor of order  $\kappa^{-1}$  for  $\varepsilon \sim \pi\kappa^3$ . The error  $|\eta_\varepsilon|$  is bounded above by  $e^{-1/4\bar{\kappa}^2} \times O(1)$  for all  $\varepsilon$ , and hence by  $e^{-1/4\kappa^2} \times O(1)$  for  $\varepsilon < 2\pi\kappa^3$ .

In Figures 2.11 and 2.12, we plot the gate error estimate  $|\eta_\varepsilon|$  as a function of  $\varepsilon$  for  $\kappa^{-2} = 40$  and  $\kappa^{-2} = 80$ . Recall that, in the case where the *LC* circuit is initially in its ground state, we can identify  $\kappa^{-2}$  with  $\sqrt{L/C}$ .

## 2.7 Diabatic error

As explained in Section 2.4, the protected phase gate is very accurate if the final state vector of the oscillator depends only very weakly on the state of the  $0-\pi$  qubit:  $|\psi_1^{\text{fin}}\rangle \approx |\psi_0^{\text{fin}}\rangle$ . Our argument establishing high gate accuracy has two elements — we show that  $|\psi_1(t)\rangle \approx \bar{X}|\psi_0(t)\rangle$  at each stage of the oscillator's evolution, and also that  $\bar{X}|\psi_0^{\text{fin}}\rangle \approx |\psi_0^{\text{fin}}\rangle$ . In Section 2.6 we have seen that the condition  $|\psi_1(t)\rangle \approx \bar{X}|\psi_0(t)\rangle$  is stable with respect to imperfections in the timing of the pulse that executes the gate. Now we will consider rare diabatic transitions, occurring as the coupling  $J(t)$  ramps on and off, that contribute to the deviation of  $\bar{X}|\psi_0^{\text{fin}}\rangle$  from  $|\psi_0^{\text{fin}}\rangle$ .

While the coupling  $J(t)$  turns on or off, the harmonic  $\varphi^2/2L$  term in the potential can be treated perturbatively, where in first approximation the Hamiltonian is given by Equation (2.28);

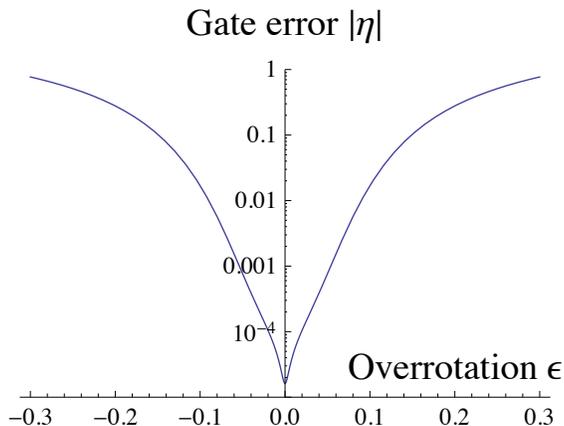


Figure 2.11: The estimated gate error  $|\eta_\epsilon|$  (on a log scale) as a function of the rotation error  $\epsilon$ , for  $\kappa^{-2} = 40$ .

we will consider the consequences of the harmonic term in Section 2.8. This Hamiltonian commutes with the operator  $e^{-2\pi i Q}$ , which translates  $\varphi$  by  $2\pi$ ; therefore,  $e^{-2\pi i Q}$  and the Hamiltonian can be simultaneously diagonalized. We express the eigenvalue of this translation operator as  $e^{-2\pi i q}$ , where  $q = Q - [Q] \in [-\frac{1}{2}, \frac{1}{2}]$  is the conserved Bloch momentum, and the integer  $[Q]$  labels the distinct bands in the Hamiltonian's spectrum.

A diabatic transition between bands may be excited while  $J(t)$  varies, changing the value of  $[Q]$  by an integer, typically  $\pm 1$ . If such transitions occur with nonnegligible probability, the final state of the oscillator will contain, in addition to a primary peak supported near  $Q = 0$  (where  $\bar{X} = 1$ ), also secondary peaks supported near  $Q = \pm 1$  (where  $\bar{X} = -1$ ). We will discuss the probability of a transition between bands while  $J(t)$  ramps on; a similar analysis applies to transitions occurring as  $J(t)$  ramps down.

The probability of a diabatic transition can be computed most reliably for  $q$  close to  $\pm \frac{1}{2}$ , since in that case the splitting between the lowest band and the first excited band is small when  $J(t)$  is small, and the continuous variable system can be well approximated by a two-level system. For example, when  $J = 0$ , the state in the lowest band with Bloch momentum  $q$  slightly less than  $\frac{1}{2}$  has

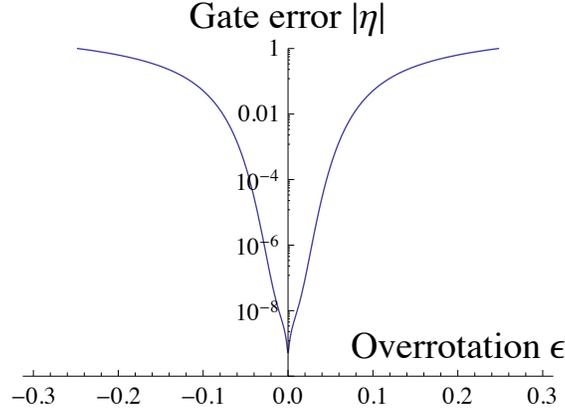


Figure 2.12: The estimated gate error  $|\eta_\epsilon|$  (on a log scale) as a function of the rotation error  $\epsilon$ , for  $\kappa^{-2} = 80$ .

charge  $Q = q$ , while the state in the first excited band has  $Q = q - 1$ . Hence, the splitting between bands is

$$\frac{1}{2C} ((q-1)^2 - q^2) = \frac{\frac{1}{2} - q}{C}. \quad (2.91)$$

Since  $e^{\mp i\varphi}$  translates  $Q$  by  $\pm 1$ , the perturbation  $J(t) \cos \varphi$  has matrix elements

$$\langle q-1 | J(t) \cos \varphi | q \rangle = \frac{J(t)}{2} = \langle q | J(t) \cos \varphi | q-1 \rangle, \quad (2.92)$$

and the effective two-level Hamiltonian is

$$H_{0,1}^{\text{eff}} = -\frac{\frac{1}{2} - q}{2C} \sigma^Z \mp \frac{J(t)}{2} \sigma^X, \quad (2.93)$$

where  $\sigma^{Z,X}$  are Pauli matrices. The energy eigenstates are  $\sigma^Z$  eigenstates for  $J(t) \rightarrow 0$  and  $\sigma^X$  eigenstates for  $J(t) \rightarrow \infty$ .

The time-dependent Schrödinger equation for this effective Hamiltonian can be solved exactly

if  $J(t)$  increases exponentially with time, as

$$J(t) = J_0 \exp(t/\tau_J^{\text{eff}}); \quad (2.94)$$

we show in Appendix 2.E that if the initial state as  $t \rightarrow -\infty$  is the ground state, then the probability that the final state is excited as  $t \rightarrow \infty$  is

$$\begin{aligned} P_{\text{diab}}(q, \tau_J^{\text{eff}}) &= \frac{1}{2} - \frac{1}{2} \tanh\left(\pi\left(\frac{1}{2} - q\right) \frac{\tau_J^{\text{eff}}}{2C}\right) \\ &\approx \exp\left(-\pi\left(\frac{1}{2} - q\right) \frac{\tau_J^{\text{eff}}}{C}\right), \end{aligned} \quad (2.95)$$

where the second equality holds when the argument of the tanh is large and positive.

We recall that if the initial state of the oscillator is the ground state or a low-lying excited state, then the probability distribution for  $q$  decays as

$$P(q) \sim \exp\left(-\sqrt{\frac{L}{C}} q^2\right); \quad (2.96)$$

hence, expanding in  $\delta = \frac{1}{2} - q$ , we find

$$\begin{aligned} P(q) P_{\text{diab}}(q, \tau_J^{\text{eff}}) &\sim \exp\left(\left(\delta - \frac{1}{4}\right) \sqrt{\frac{L}{C}}\right) \\ &\times \exp\left(-\pi\delta \frac{\tau_J^{\text{eff}}}{C}\right). \end{aligned} \quad (2.97)$$

Therefore, if

$$\tau_J^{\text{eff}} > \frac{1}{\pi} \sqrt{LC}, \quad (2.98)$$

the most likely diabatic transitions occur for  $q \approx \frac{1}{2}$ , where the two-level approximation is reasonable; we conclude in that case that the probability of a diabatic transition is suppressed by the factor  $\exp(-\frac{1}{4}\sqrt{L/C})$ . If, on the other hand,  $\tau_J^{\text{eff}} < \frac{1}{\pi}\sqrt{LC}$ , then the most likely diabatic transitions occur for  $q$  far from  $\pm\frac{1}{2}$ , and the two-level approximation cannot be justified.

If  $J(t)$  does not ramp on exponentially, then the exact solution in Appendix 2.E does not apply

directly. To estimate roughly the probability of a diabatic transition for more general pulse shapes, we note that the transition typically occurs when  $\sigma^Z$  and  $\sigma^X$  in the effective Hamiltonian have comparable coefficients, so that

$$\tau_J^{\text{eff}} \approx \left( \frac{J}{\dot{J}} \right)_{JC \approx 1-2q}. \quad (2.99)$$

If  $J(t)$  turns on like an error function with width  $\tau_J$ , then

$$J(t) = \frac{J_0}{\sqrt{\pi}} \int_{-\infty}^{t/\tau_J} dx e^{-x^2} \approx \frac{\tau_J J_0}{2|t|\sqrt{\pi}} e^{-t^2/\tau_J^2} \quad (2.100)$$

asymptotically for  $t/\tau_J \rightarrow -\infty$ , and we have

$$J/\dot{J} \approx \tau_J^2/2|t|; \quad (2.101)$$

combining Equations (2.99) to (2.101), we obtain

$$\frac{2\tau_J^{\text{eff}}}{\tau_J} = \left( \ln \left( \frac{J_0 C}{1-2q} \right) - O \left( \ln \ln \left( \frac{J_0 C}{1-2q} \right) \right) \right)^{-1/2}. \quad (2.102)$$

Although  $\tau_J^{\text{eff}}$  given by Equation (2.102) does not satisfy Equation (2.98) when  $q$  is very close to  $\frac{1}{2}$ , the dominant diabatic transitions may still occur for  $q \approx \frac{1}{2}$ , where the two-level approximation is applicable, provided  $\tau_J - \frac{1}{\pi}\sqrt{LC}$  is positive and sufficiently large. Otherwise, if the dominant value of  $q$  is far from  $\frac{1}{2}$ , we can anticipate that typical diabatic transitions occur for  $J(t)C = O(1)$ , where the band gap is  $O(1/C)$  and the transition probability is

$$P_{\text{diab}}(\tau_J) = \exp \left( -O \left( \frac{\tau_J}{C} \right) \right). \quad (2.103)$$

In the two-level approximation, which applies for  $|Q| \approx \frac{1}{2}$ , the probability of a jump from the lowest band ( $|Q| < \frac{1}{2}$ ) to the first excited band ( $|Q| > \frac{1}{2}$ ) matches the probability of a jump from the first excited band to the lowest band. Therefore, neglecting transitions to other bands and the higher-order probability of multiple transitions, and also ignoring other sources of error aside from diabatic jumps, we infer from Equation (2.95) that the probability of  $\bar{X} = 1$  (*i.e.*,  $|Q| < \frac{1}{2}$ ) in the

final state of the oscillator can be expressed as

$$P(|Q^{\text{fin}}| < \frac{1}{2}) \approx \int dQ^{\text{in}} P(Q^{\text{in}}) \tanh\left(\pi\left(\frac{1}{2} - |Q^{\text{in}}|\right)\frac{\tau_J^{\text{eff}}}{2C}\right); \quad (2.104)$$

a factor of two has been included to take into account that the transition could occur during either the ramping-up phase or the ramping-down phase. Because of the enhanced probability of a transition for  $|Q| \approx \frac{1}{2}$ , the  $Q^{\text{in}}$  integral has support extending beyond the range  $[-\frac{1}{2}, \frac{1}{2}]$ ; the tanh function smooths out the sharp edges at  $Q = \pm\frac{1}{2}$ , replacing them by rounded steps with width of order  $C/\tau_J^{\text{eff}}$ .

## 2.8 Squeezing error

In Section 2.4 we discussed how the state of the oscillator evolves as the coupling  $J(t)$  ramps on and off. There we used the idea that, because the oscillator's period is long compared to the time scale  $\tau_J$  for the coupling to turn on and off, we may as a first approximation ignore the  $\varphi^2/2L$  term in the potential as in Equation (2.28). Under that assumption we concluded that

$$\langle \psi_1^{\text{begin}} | \bar{X} | \psi_0^{\text{begin}} \rangle = \langle \psi^{\text{in}} | \bar{X} | \psi^{\text{in}} \rangle \approx 1, \quad (2.105)$$

where  $|\psi^{\text{in}}\rangle$  is the oscillator's initial state, and  $|\psi_{0,1}^{\text{begin}}\rangle$  denotes the state just after the coupling turns on, where  $|0\rangle, |1\rangle$  is the state of the  $0-\pi$  qubit. The second equality follows if the initial state of the oscillator has negligible support outside the interval  $Q \in [-\frac{1}{2}, \frac{1}{2}]$ .

How is this conclusion affected when the quadratic term  $\varphi^2/2L$  is included? If the coupling turns on slowly enough, this term can cause some squeezing of the wave function in  $\varphi$  space and, correspondingly, some spreading in  $Q$  space. To model crudely the effect of the spreading, consider first turning on  $J(t)$  using  $H_0$  or  $H_1$  in Equation (2.28), then applying the operator  $e^{-i\alpha\varphi}$  (which shifts  $Q$  by an amount  $\alpha$  that does not depend on the state of the  $0-\pi$  qubit). Denoting the time evolution operator as  $J(t)$ , turning on by  $U_0$  or  $U_1 = \bar{X}U_0\bar{X}$  as in Section 2.4, we then have

$$\begin{aligned} |\psi_0^{\text{begin}}\rangle &= e^{-i\alpha\varphi}U_0|\psi^{\text{in}}\rangle, \\ |\psi_1^{\text{begin}}\rangle &= e^{-i\alpha\varphi}\bar{X}U_0\bar{X}|\psi^{\text{in}}\rangle, \end{aligned} \quad (2.106)$$

and therefore

$$\langle \psi_1^{\text{begin}} | \bar{X} | \psi_0^{\text{begin}} \rangle = \langle \psi^{\text{in}} | \bar{X} U_0^{-1} \bar{X} e^{i\alpha\varphi} \bar{X} e^{-i\alpha\varphi} U_0 | \psi^{\text{in}} \rangle. \quad (2.107)$$

Now we note that

$$\begin{aligned} \bar{X} e^{i\alpha\varphi} \bar{X} e^{-i\alpha\varphi} &= (-1)^{-[Q]} e^{i\alpha\varphi} (-1)^{[Q]} e^{-i\alpha\varphi} \\ &= (-1)^{-[Q]} (-1)^{[Q-\alpha]} = (-1)^{[Q-\alpha]-[Q]}; \end{aligned} \quad (2.108)$$

furthermore,  $[Q - \alpha] - [Q]$  commutes with  $\cos \varphi$  and hence with  $U_0$ , because  $e^{-i\varphi}$ , acting by conjugation, increases both  $[Q - \alpha]$  and  $[Q]$  by 1 (while  $e^{i\varphi}$  decreases both by 1). Hence, we find

$$\begin{aligned} \langle \psi_1^{\text{begin}} | \bar{X} | \psi_0^{\text{begin}} \rangle &= \langle \psi^{\text{in}} | e^{i\alpha\varphi} \bar{X} e^{-i\alpha\varphi} | \psi^{\text{in}} \rangle \\ &= \langle \psi^{\text{in}} | (-1)^{[Q-\alpha]} | \psi^{\text{in}} \rangle; \end{aligned} \quad (2.109)$$

in particular, if  $|\psi^{\text{in}}\rangle$  is almost fully supported in the interval  $Q \in [-1/2 + |\alpha|, 1/2 - |\alpha|]$ , then  $|\psi_1^{\text{begin}}\rangle$  is very close to  $\bar{X} |\psi_0^{\text{begin}}\rangle$ . If  $|\psi^{\text{in}}\rangle$  is the Gaussian ground state with  $\langle Q^2 \rangle = \frac{1}{2} \sqrt{C/L}$ , the deviation of  $\langle \psi_1^{\text{begin}} | \bar{X} | \psi_0^{\text{begin}} \rangle$  from 1 is suppressed by the exponential factor

$$\begin{aligned} |1 - \langle \psi_1^{\text{begin}} | \bar{X} | \psi_0^{\text{begin}} \rangle| &\approx \exp\left(\left(\frac{1}{2} - |\alpha|\right)^2 \sqrt{\frac{L}{C}}\right) \\ &\approx \exp\left(|\alpha| \sqrt{\frac{L}{C}}\right) \exp\left(-\frac{1}{4} \sqrt{\frac{L}{C}}\right). \end{aligned} \quad (2.110)$$

How much spreading in  $Q$  space should be expected? To make a crude estimate of how the harmonic term affects the distribution in  $Q$  space, we note that

$$e^{-i\beta\varphi^2} Q e^{i\beta\varphi^2} = Q + 2\beta\varphi, \quad (2.111)$$

and choose  $\beta \approx \tau_J/2L$ , where  $\tau_J$  is the time scale for the coupling to turn on; using  $\langle \varphi^2 \rangle = \frac{1}{2} \sqrt{L/C}$  in the Gaussian ground state, we infer that  $Q$  is shifted by an amount of order

$$\alpha \sim \frac{\tau_J}{L} \left(\frac{L}{C}\right)^{1/4}. \quad (2.112)$$

Assuming  $\tau_J \sim \sqrt{LC}$  in order to suppress the diabatic error, we find that squeezing enhances the gate error by a factor

$$\exp\left(\left(\text{constant}\right)\left(\frac{L}{C}\right)^{1/4}\right); \quad (2.113)$$

that is, it contributes a subleading correction to the logarithm of the gate error.

More realistically, treating the harmonic term as a perturbative correction to the zeroth-order Hamiltonian, which has  $\varphi \rightarrow \varphi + 2\pi$  periodicity, the dynamics is governed by the effective Hamiltonian

$$H_{\text{eff}} = \epsilon_{J,C}(q) + \frac{\varphi^2}{2L}, \quad (2.114)$$

where  $q \in [-\frac{1}{2}, \frac{1}{2}]$  is the Bloch momentum and  $\epsilon_{J,C}(q)$  is the energy of the lowest band. Expanding this band energy to quadratic order, we have

$$\epsilon_{J,C}(q) \approx \frac{q^2}{2C_{\text{eff}}}. \quad (2.115)$$

The effective capacitance  $C_{\text{eff}}$  is approximately  $C$  for  $J$  small, but for  $JC \approx 1$ , the band curvature begins to flatten rapidly; correspondingly,  $C_{\text{eff}}$  increases sharply, as does the oscillator's period  $2\pi\sqrt{LC_{\text{eff}}}$ . The oscillator evolves adiabatically for  $J$  small, but its evolution freezes when its period becomes comparable to  $\tau_J$ , the characteristic time scale for the variation of the Hamiltonian. Therefore, the squeezing error is determined by the wave function's width in  $q$ -space at the time when the oscillator freezes; hence,

$$\begin{aligned} P_{\text{sq}} &\sim \exp\left(-\left(\text{constant}\right)\sqrt{\frac{L}{C_{\text{eff}}}}\right) \\ &\sim \exp\left(-\left(\text{constant}\right)\frac{L}{\tau_J}\right) \sim \exp\left(-\left(\text{constant}\right)\sqrt{\frac{L}{C}}\right), \end{aligned} \quad (2.116)$$

where we obtain the last equality by choosing  $\tau_J \sim \sqrt{LC}$  to suppress the diabatic error. Thus the contribution to the gate error due to squeezing is comparable to the other sources of error. We can use a similar argument to conclude that the squeezing error arising as the coupling  $J(t)$  turns off is also of the same order.

## 2.9 Simulations

We have compared the predictions from Sections 2.4 and 2.6 to numerical simulations of the single-qubit phase gate  $\exp(i\frac{\pi}{4}Z)$ . We solved the time-dependent Schrödinger equation for the Hamiltonians  $H_{0,1}$  in Equation (2.15), assuming the oscillator starts in the ground state (excited states will be considered in Section 2.10). These simulations were done in MATLAB using the fourth-order split-operator method, which is based on the identity

$$\begin{aligned} \exp(it(A+B)) &= \exp(i\frac{\gamma}{2}tA) \exp(i\gamma tB) \exp(i\frac{1-\gamma}{2}tA) \\ &\times \exp(i(1-2\gamma)tB) \exp(i\frac{1-\gamma}{2}tA) \exp(i\gamma tB) \exp(i\frac{\gamma}{2}tA) + \mathcal{O}(t^4), \end{aligned} \quad (2.117)$$

where

$$\gamma = \frac{1}{2 - \sqrt[3]{2}}, \quad (2.118)$$

and where  $A$  and  $B$  are the portions of the Hamiltonian that are diagonal in the position and momentum eigenbases, respectively. The full time evolution is broken up into many small steps with the Hamiltonian alternating between  $A$  and  $B$  and the Fourier transform or its inverse applied between successive steps.

We assume that the coupling  $\mp J(t) \cos \varphi$  between the oscillator and the  $0-\pi$  qubit turns on with an error-function profile,

$$J(t) = J_0 \left( \frac{1}{2} + \frac{1}{2} \operatorname{erf}(t/\tau_J) \right), \quad (2.119)$$

and that the coupling turns off after the time delay  $\tau \approx L/\pi$ , according to the time-reversed function  $J(\tau - t)$ ; the time scale  $\tau_J$  for turning the coupling on and off and the time delay  $\tau$  were chosen to optimize the gate accuracy. Then we varied the time delay  $\tau = \tau_0(1 + \varepsilon)$ , where  $\tau_0$  is the optimal value, to study the effect of overrotation/underrotation on the gate accuracy.

Figure 2.13 shows the results for a series of simulations with  $\sqrt{L/C} = 80$ ,  $\sqrt{J_0 C} = 8$  and  $\tau_J = \sqrt{LC}$ , plotted together with the analytic prediction Equation (2.89), arising from the overrotation error alone. For small  $\varepsilon$ , the gate error is

$$|\eta_\varepsilon| \approx 2 \times 10^{-8} \approx 10 \exp\left(-\frac{1}{4}\sqrt{\frac{L}{C}}\right), \quad (2.120)$$

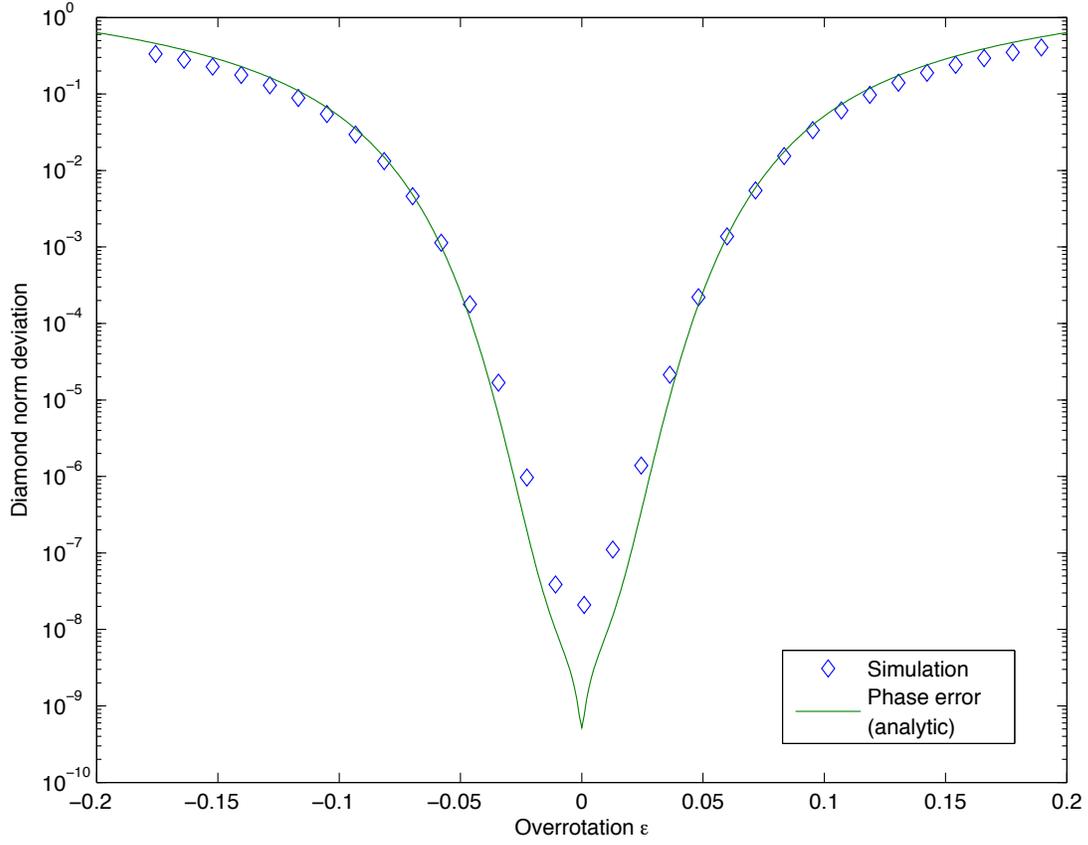


Figure 2.13: Blue diamonds show on a logarithmic scale the numerically computed diamond-norm deviation from the ideal phase gate, for  $\kappa^{-2} = \sqrt{L/C} = 80$ ,  $\Delta^{-2} = \sqrt{J_0 C} = 8$ , and  $\tau_J/C = 80$ , as a function of the overrotation parameter  $\varepsilon$ . The solid green line is the analytic prediction from Equation (2.89) for the overrotation error alone. The discrepancy for small  $\varepsilon$  arises from corrections due to diabatic transitions and squeezing.

substantially larger than the analytic prediction  $|\eta_\varepsilon| \approx 5 \times 10^{-10}$ , but roughly compatible with our expectations for the scale of the error due to diabatic transitions and squeezing, the dominant errors in this regime, when we choose  $\tau_J \approx \sqrt{LC}$ . For larger values of  $\varepsilon$  the overrotation error dominates and the numerical results agree well with the analytic prediction.

The simulations verify that the protected phase gate is much more robust than an unprotected gate. For an unprotected gate, executed by coupling the  $0-\pi$  qubit to a Josephson junction, as described in Section 2.3, achieving a gate error less  $10^{-4}$  in the diamond norm requires a pulse timing accuracy of order  $10^{-4}$ . In contrast, for a protected gate with the parameters specified above, 4% accuracy in the pulse timing suffices for achieving a gate error less than  $10^{-4}$ . When the pulse timing accuracy is better than 1%, the gate error is below  $4 \times 10^{-8}$ .

## 2.10 Nonzero temperature

Our estimate of the intrinsic gate error in Equations (2.65) and (2.69) applies to any grid-like state of the form Equation (2.55). In Equations (2.78) and (2.85) we treated the specific case of a Gaussian grid state, which is what arises during the execution of the protected phase gate in the case where the initial state  $|\psi_{\text{in}}\rangle$  of the harmonic oscillator is the ground state. Let us now consider the case where the initial state is excited, as occurs with nonzero probability at any nonzero temperature. To keep things simple, we will ignore the effects of overrotation, setting  $\varepsilon = 0$ . In that case,  $\text{Im} \eta_\varepsilon = 0$  because  $F_\varepsilon(Q)$  is real, so we only need to worry about the real part of the gate error. In this section we also neglect the corrections due to diabatic transitions and squeezing.

When the coupling to the  $0-\pi$  qubit is turned off, the harmonic oscillator Hamiltonian is

$$H = \frac{Q^2}{2C} + \frac{\varphi^2}{2L}, \quad (2.121)$$

and its ground state  $|\psi_0\rangle$  obeys

$$\begin{aligned} \langle \varphi | \psi_0 \rangle &= \left( \frac{1}{\pi} \sqrt{\frac{C}{L}} \right)^{1/4} e^{-\frac{1}{2} \sqrt{\frac{C}{L}} \varphi^2}, \\ \langle Q | \psi_0 \rangle &= \left( \frac{1}{\pi} \sqrt{\frac{L}{C}} \right)^{1/4} e^{-\frac{1}{2} \sqrt{\frac{L}{C}} Q^2}. \end{aligned} \quad (2.122)$$

Our estimate diamond norm deviation from the ideal gate becomes

$$\begin{aligned}
|\eta(0)| &\approx 4 \int_{1/2}^{\infty} dQ |\langle Q|\psi_0\rangle|^2 \\
&= 4 \left( \frac{1}{\pi} \sqrt{\frac{L}{C}} \right)^{1/2} \int_{1/2}^{\infty} dQ e^{-\sqrt{\frac{L}{C}} Q^2} \\
&\approx \frac{4}{\sqrt{\pi}} \left( \frac{L}{C} \right)^{-1/4} e^{-\frac{1}{4} \sqrt{\frac{L}{C}}}.
\end{aligned} \tag{2.123}$$

Numerically, we have, *e.g.*,  $|\eta(0)| \approx 1.6 \times 10^{-5}$  for  $\sqrt{L/C} = 40$  and  $|\eta(0)| \approx 5.2 \times 10^{-10}$  for  $\sqrt{L/C} = 80$ .

The oscillator's  $n$ th excited state is

$$|\psi_n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}} |\psi_0\rangle, \tag{2.124}$$

where

$$a^\dagger = \left( \frac{C}{4L} \right)^{1/4} \varphi - i \left( \frac{L}{4C} \right)^{1/4} Q, \tag{2.125}$$

which becomes

$$a^\dagger = i \left( \frac{C}{4L} \right)^{1/4} \frac{d}{dQ} - i \left( \frac{L}{4C} \right)^{1/4} Q \tag{2.126}$$

in the  $Q$  representation. Therefore, the  $n$ th harmonic oscillator excited state can be expressed in  $Q$  space as

$$\begin{aligned}
\langle Q|\psi_n\rangle &= \frac{1}{\sqrt{n!}} \left( - \left( \frac{C}{4L} \right)^{1/4} \frac{d}{dQ} + \left( \frac{L}{4C} \right)^{1/4} Q \right)^n \langle Q|\psi_0\rangle \\
&= \frac{2^{n/2} \pi^{-1/4}}{\sqrt{n!}} \left( \sqrt{\frac{L}{C}} \right)^{\frac{n}{2} + \frac{1}{4}} (Q^n + \dots) e^{-\frac{1}{2} \sqrt{\frac{L}{C}} Q^2},
\end{aligned} \tag{2.127}$$

where in the second line we have retained only the leading power of  $Q$  in the prefactor of the exponential. To estimate the probability of a logical phase error, we assume that this leading power dominates, and we also use the leading term in the asymptotic expansion

$$\int_x^\infty dt t^{2n} e^{-\alpha t^2} = \left( \frac{x^{2n-1}}{2\alpha} \right) e^{-\alpha x^2} (1 - O(1/x^2)) \tag{2.128}$$

to calculate

$$\begin{aligned}
|\eta(n)| &\approx 4 \int_{1/2}^{\infty} dQ |\langle Q|\psi_n\rangle|^2 \\
&\approx \frac{2^{n+2}}{n!\sqrt{\pi}} \left(\frac{L}{C}\right)^{\frac{n}{2}+\frac{1}{4}} \int_{1/2}^{\infty} dQ Q^{2n} e^{-\sqrt{\frac{L}{C}} Q^2} \\
&\approx \frac{1}{2^{n-2}n!\sqrt{\pi}} \left(\frac{L}{C}\right)^{\frac{n}{2}-\frac{1}{4}} e^{-\frac{1}{4}\sqrt{\frac{L}{C}}} \\
&= \frac{1}{2^n n!} \left(\frac{L}{C}\right)^{n/2} |\eta(0)|.
\end{aligned} \tag{2.129}$$

Thus, for example, the intrinsic gate error for the first excited ( $n = 1$ ) state is enhanced relative to the ground state by the factor  $\frac{1}{2}\sqrt{L/C}$ . This approximation is applicable when  $n$  is not too large, so that the leading power of  $Q$  dominates the prefactor of the exponential in the tail of the wave function at  $|Q| > 1/2$ ; in particular, we require that

$$\langle Q^2 \rangle_n = n \langle Q^2 \rangle_0 = \frac{n}{2} \left(\frac{L}{C}\right)^{-1/2} \ll \frac{1}{2}, \tag{2.130}$$

or

$$n \ll \sqrt{L/C}. \tag{2.131}$$

The energy of the  $n$ th oscillator state is  $E_n = n/\sqrt{LC}$ ; therefore, in the thermal ensemble with inverse temperature  $\beta$ , the probability that the oscillator is in the  $n$ th state is

$$P_n = \left(1 - e^{-\beta/\sqrt{LC}}\right) e^{-n\beta/\sqrt{LC}}. \tag{2.132}$$

Thus, if the oscillator is in a thermal state, while the intrinsic phase error probability for the  $n$ th state is enhanced by the factor  $(L/C)^{n/2}/2^n n!$ , it is also suppressed by the Boltzmann factor  $e^{-n\beta/\sqrt{LC}}$ . Summing up the error probabilities for all oscillator states, with the appropriate

Boltzmann weights, we find

$$\begin{aligned}
|\eta(\beta)| &= \left(1 - e^{-\beta/\sqrt{LC}}\right) \sum_{n=0}^{\infty} e^{-n\beta/\sqrt{LC}} |\eta(n)| \\
&\approx \left(1 - e^{-\beta/\sqrt{LC}}\right) |\eta(0)| \sum_{n=0}^{\infty} e^{-n\beta/\sqrt{LC}} \frac{1}{2^n n!} \left(\frac{L}{C}\right)^{n/2} \\
&= \left(1 - e^{-\beta/\sqrt{LC}}\right) \exp\left(\frac{1}{2} \sqrt{\frac{L}{C}} e^{-\beta/\sqrt{LC}}\right) |\eta(0)|.
\end{aligned} \tag{2.133}$$

The real part of  $\eta$  (and in the case we are considering there is no imaginary part) is essentially the probability of a logical phase error in the grid state, and hence to compute the gate error we need only compute this logical error probability for the thermal ensemble. Thus, the error at finite temperature is comparable to the zero-temperature error, provided that

$$\sqrt{\frac{L}{C}} e^{-\beta/\sqrt{LC}} \ll 1. \tag{2.134}$$

If, for example,  $\beta/\sqrt{LC} \approx 3$ , then compared to the zero-temperature case, thermal effects enhance the phase error probability by the factor 2.6 for  $\sqrt{L/C} = 40$  and 7.0 for  $\sqrt{L/C} = 80$ . We expect, then, that the protected phase gate remains reasonably robust, provided the temperature is smaller than or comparable to the frequency of the superinductive  $LC$  circuit.

Numerical results, plotted in Figure 2.14, show that the gate performance remains robust for excited eigenstates. In addition to the enhancement of the intrinsic phase error predicted by Equation (2.129), there is also a contribution to the gate error arising from diabatic transitions and squeezing, which becomes less important for more highly excited eigenstates.

Up until now, we have addressed how the accuracy of the phase gate is affected if, due to thermal fluctuations, the *initial* state of the  $LC$  oscillator is not the ground state, but we have not considered thermally excited transitions that might occur while the qubit and oscillator are coupled *during* the execution of the gate. A thermally activated transition between bands, like a diabatic transition, could flip the value of  $\bar{X} = (-1)^{[Q]}$  and cause the gate to fail. But, as in Section 2.7, the relevant

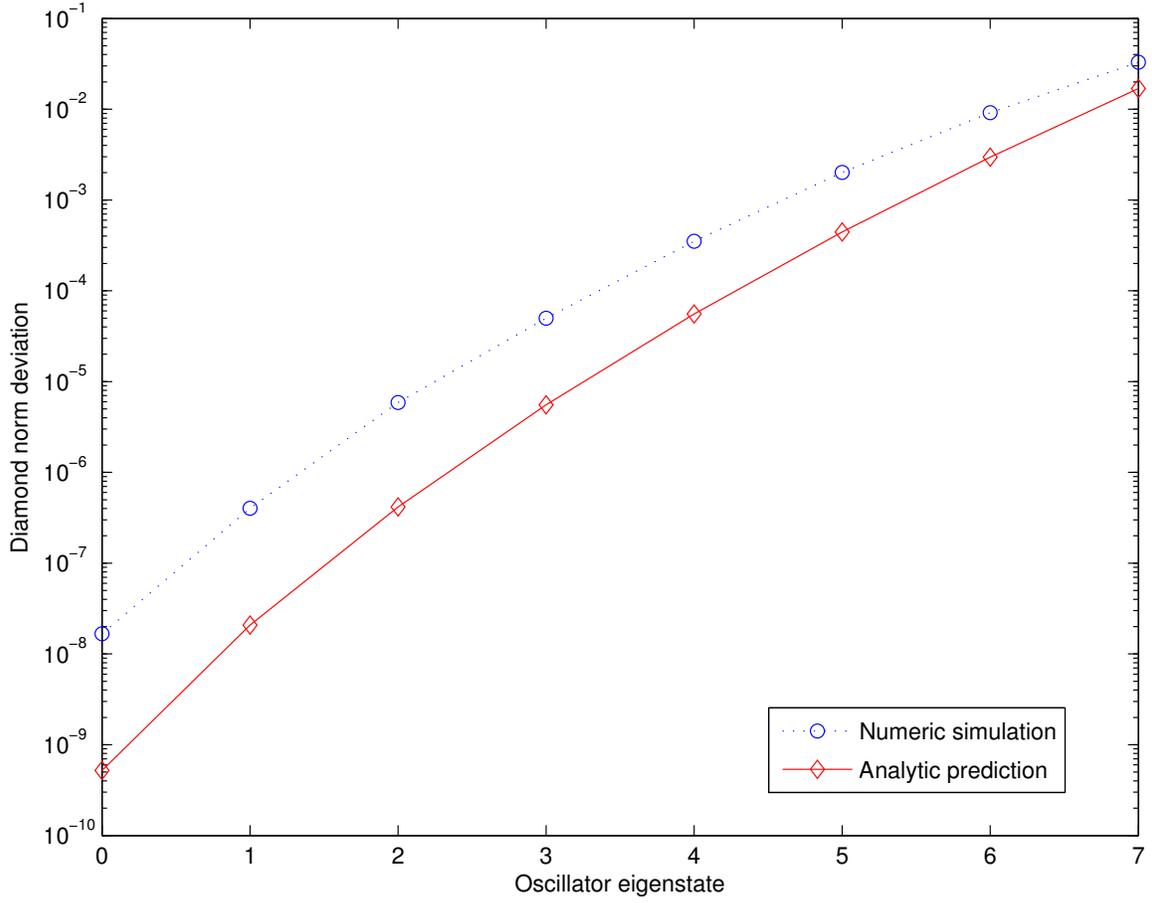


Figure 2.14: The minimum diamond-norm deviation of the protected phase gate from the ideal gate, as a function of the initial oscillator eigenstate  $n$ , for  $\sqrt{L/C} = 80$  and  $\sqrt{J_0 C} = 8$ . Results from numerical simulations are shown in blue, and the analytic prediction Equation (2.129) is shown in red. The discrepancy arises from corrections due to diabatic effects and squeezing, which are neglected in the derivation of Equation (2.129).

band gap is of order  $1/C$ , so such transitions are suppressed by a Boltzmann factor

$$\begin{aligned} P_{\text{thermal}}(\beta) &= \exp\left(-O\left(\frac{\beta}{C}\right)\right) \\ &= \exp\left(-O\left(\frac{\beta}{\sqrt{LC}} \cdot \sqrt{\frac{L}{C}}\right)\right), \end{aligned} \quad (2.135)$$

exponentially small in  $\sqrt{L/C}$  if  $\beta/\sqrt{LC} = O(1)$ . On the other hand, spontaneous decay of the oscillator during the execution of the gate is not likely to flip the value of  $\bar{X}$ , and hence has little impact on the gate accuracy.

## 2.11 Perturbative stability

Aside from its stability with respect to pulse timing errors and thermal effects, we also expect the protected phase gate to be robust against small deformations in the Hamiltonian of the  $LC$  oscillator and of the switch coupling the oscillator to the qubit. Suppose, for example, that the oscillator's potential energy  $V$  includes a small anharmonic term so that

$$V = \frac{\varphi^2}{2L} + \lambda\varphi^4. \quad (2.136)$$

Over time  $t$ , the effect of the anharmonic term on the charge  $Q$  is

$$e^{-i\lambda\varphi^4 t} Q e^{i\lambda\varphi^4 t} = Q + 4\lambda t \varphi^3; \quad (2.137)$$

thus, the charge spreads by an amount

$$\delta Q \approx 4\lambda t \langle \varphi^6 \rangle^{1/2} = 4\lambda t \sqrt{15} \langle \varphi^2 \rangle^{3/2}. \quad (2.138)$$

Comparing the contribution in Equation (2.83) to  $\delta Q$  arising from overrotating the gate, and choosing  $t \approx L/\pi$ , we see that the effect of the anharmonic term is roughly comparable to the effect of an overrotation error

$$\varepsilon \approx 4\sqrt{15} \lambda L \langle \varphi^2 \rangle = 2\sqrt{15} \lambda L \sqrt{L/C} \approx 7.75\alpha, \quad (2.139)$$

where  $\alpha \equiv \lambda L \sqrt{L/C}$  is a dimensionless parameter characterizing the strength of the anharmonic correction. For  $\sqrt{L/C} = 80$  and  $\sqrt{J_0 C} = 8$ , as in Section 2.9, numerical results plotted in Figure 2.15 confirm that the gate accuracy is not much affected by the anharmonic term for  $\alpha \lesssim 10^{-3}$ , as expected. In these simulations, we assumed that the initial state of the oscillator is the ground state of the unperturbed oscillator Hamiltonian, which for  $\alpha$  small has a large overlap with the ground state of the perturbed Hamiltonian.

The Josephson coupling between the  $0-\pi$  qubit and the oscillator is a periodic function of  $\varphi$  with period  $2\pi$ , but need not be a pure cosine potential. If we include a next-to-leading harmonic correction, the time-dependent Hamiltonian has the form

$$H_{0,1}(t) = \frac{Q^2}{2C} + \frac{\varphi^2}{2L} \mp f(t) J_0 [\cos \varphi \pm \beta \cos 2\varphi], \quad (2.140)$$

where  $f(t)$  varies between 0 and 1 as the coupling turns on or off, and  $\beta$  is a dimensionless parameter characterizing the strength of the perturbation. (Note that shifting  $\varphi$  by  $\pi$  changes the sign of  $\cos \varphi$  but not the sign of  $\cos 2\varphi$ .) For  $\sqrt{J_0 C} \gg 1$  and  $\beta \ll 1$ , we expect the wave function to be well localized near even or odd integer multiples of  $\pi$  in  $\varphi$  space (depending on whether the Hamiltonian is  $H_0$  or  $H_1$ ) while the coupling is turned on; hence, it should be a good approximation to expand the potential in a power series about these local minima, and therefore the perturbation is roughly equivalent to rescaling  $J_0$  according to

$$J_0 \rightarrow J_0 (1 + 4\beta). \quad (2.141)$$

The precise value of  $J_0$  does not strongly influence the phase error probability, as long as it is large enough to allow discrete peaks to form in  $\varphi$  space (*i.e.*, to strongly suppress phase slips); it instead determines the probability of an intrinsic bit flip error, as in Equation (2.51).

Numerical simulations confirm that the phase gate accuracy is insensitive to the perturbation in Equation (2.140) when  $\beta$  is small. For  $\sqrt{L/C} = 80$  and  $\sqrt{J_0 C} = 8$ , we find that the effect on the gate error is negligible for  $|\beta| \lesssim 0.05$ .

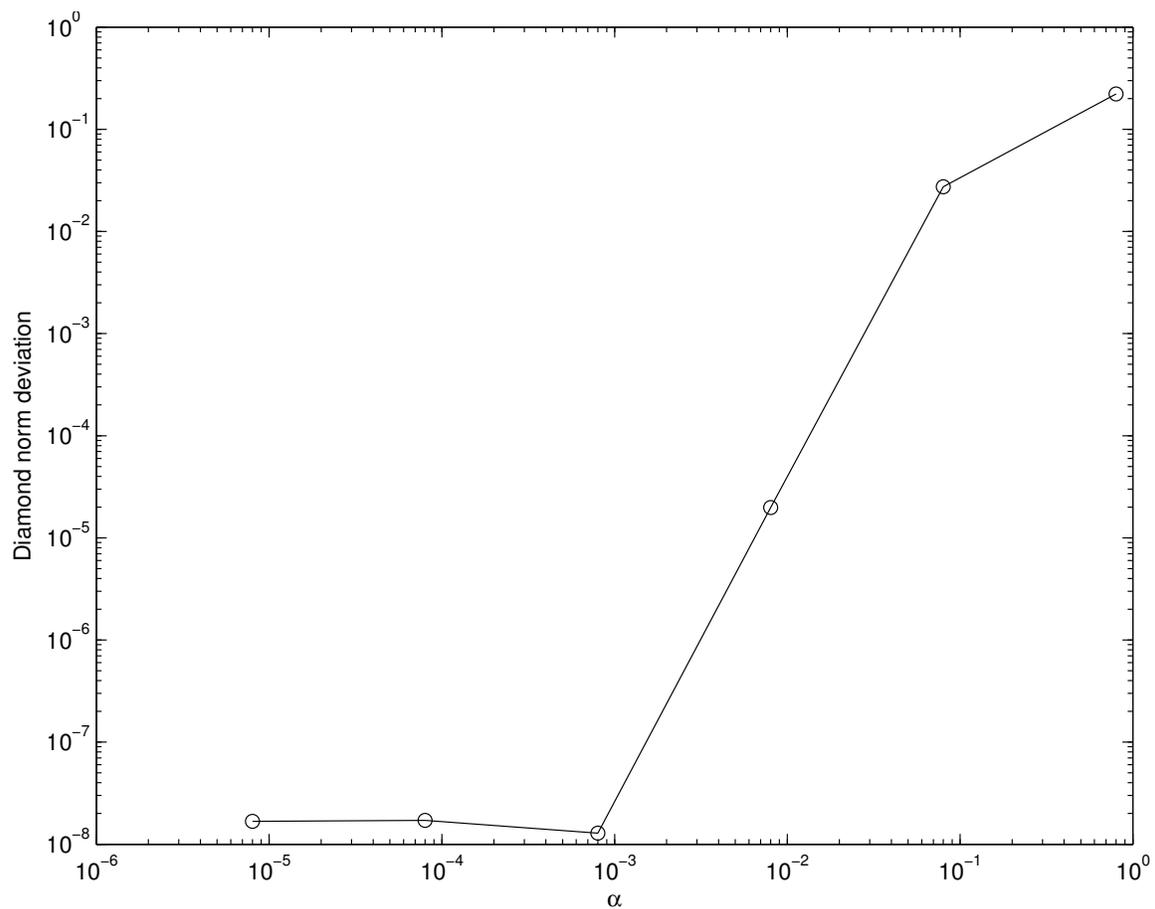


Figure 2.15: The minimum diamond-norm deviation of the protected phase gate from the ideal gate, as a function of the oscillator's anharmonicity parameter  $\alpha = \lambda L \sqrt{L/C}$ . Here, as in Figure 2.13,  $\sqrt{L/C} = 80$  and  $\sqrt{J_0 C} = 8$ .

## 2.12 Universality

We have focused so far on performing the single-qubit gate  $\exp(i\frac{\pi}{4}Z)$  and the two-qubit gate  $\exp(i\frac{\pi}{4}Z \otimes Z)$ ; in principle, these can be executed with very high fidelity by coupling a  $0-\pi$  qubit or a pair of such qubits to a superinductive  $LC$  circuit. Unfortunately, these gates are not adequate by themselves for universal quantum computing.

One way to obtain a universal gate set is to augment these gates by the following operations [42]:

1. Preparation of the single-qubit states  $|0\rangle$  and  $|+\rangle$  (the  $Z = 1$  and  $X = 1$  eigenstates).
2. Measurement of the single-qubit Pauli operators  $Z$  and  $X$ .
3. The single-qubit gate  $\exp(i\frac{\pi}{8}Z)$ .

(In fact (1) need not be regarded as independent of (2), since repeated noncommuting measurements can be used to achieve the state preparation, but we list these operations separately for clarity and completeness.)

The operation  $\exp(i\frac{\pi}{8}Z)$  could be executed by coupling the  $0-\pi$  qubit to a Josephson junction for a specified time, as in Figure 2.4. This unprotected gate might be fairly noisy. However, if all the other operations in the universal set were perfect, then scalable quantum computing would be possible, provided the noisy  $\exp(i\frac{\pi}{8}Z)$  gate meets the loose fidelity criterion  $F > .93$  [47]. Therefore, if the gates  $\exp(i\frac{\pi}{4}Z)$  and  $\exp(i\frac{\pi}{4}Z \otimes Z)$  are well protected, it follows that highly reliable universal quantum computation can be achieved, provided that the measurements, like these gates, have a very low error rate.

In practice, the measurements are likely to be noisy. However, if they can be performed non-destructively (with only a very low probability of changing the eigenvalue of measured operator), then they can be repeated multiple times to improve reliability.

We note that the CZ gate, the two-qubit gate, diagonal in the computational basis, with eigenvalues  $\{1, 1, 1, -1\}$ , can be constructed from protected gates using the decomposition

$$\begin{aligned} \text{CZ} &= \exp\left(i\frac{\pi}{4}(Z - I) \otimes (Z - I)\right) \\ &= \exp\left(i\frac{\pi}{4}Z \otimes Z\right) \exp\left(-i\frac{\pi}{4}Z \otimes I\right) \exp\left(-i\frac{\pi}{4}I \otimes Z\right) \end{aligned} \tag{2.142}$$

(up to an overall phase). One way to perform a nondestructive measurement of  $Z$  is to use the property

$$\begin{aligned} \text{CZ} : |0\rangle \otimes |+\rangle &\rightarrow |0\rangle \otimes |+\rangle, \\ |1\rangle \otimes |+\rangle &\rightarrow |1\rangle \otimes |-\rangle; \end{aligned} \tag{2.143}$$

we may apply CZ to the target qubit (the one to be measured) and an ancilla qubit prepared in the state  $|+\rangle$ , then perform  $X$  measurement on the ancilla qubit. If the CZ gate is not likely to induce a bit flip on the target qubit, this procedure can be repeated many times, and the measurement result can be determined by a majority vote of the outcomes.

If we are limited to using our protected gates, we cannot use the same trick to amplify an  $X$  measurement. Perhaps the charge measurement described in Section 2.2, though the outcome is noisy, can be done fault tolerantly, meaning that measurement procedure is not likely to flip the eigenvalue of  $X$ . In that case, amplification by repetition and majority voting will work. Otherwise, there are alternative ways to boost the measurement accuracy, using repetition coding.

For example, Chapter 3 and Ref. [2,48,49] describe a scheme for universal fault-tolerant quantum computing built from the CZ gate,  $|+\rangle$  preparation,  $X$  measurement, and, in addition, preparation of the single-qubit states

$$\begin{aligned} |-i\rangle &= \frac{1}{\sqrt{2}} (|0\rangle - i|1\rangle), \\ |T\rangle &= \frac{1}{\sqrt{2}} (|0\rangle + e^{-i\pi/4}|1\rangle), \end{aligned}$$

which can be achieved by applying  $\exp(i\frac{\pi}{4}Z)$  or  $\exp(i\frac{\pi}{8}Z)$  to  $|+\rangle$ . The main point of [2,48,49] is that this scheme works effectively when the noise in the CZ gate is highly biased, *i.e.*, when  $Z$  errors are much more common than  $X$  errors. The point we wish to emphasize here is that the scheme remains effective when the  $X$  measurement error rate is much higher than the CZ gate error rate.

The crucial element of the construction in [48,49] is a “teleported” encoded CNOT gate acting on blocks of a repetition code, shown in Figure 2.16 (and see related figure Figure 3.5.) In Chapter 3 this construction is extended to Bacon-Shor codes. The probability of a logical error in the CNOT

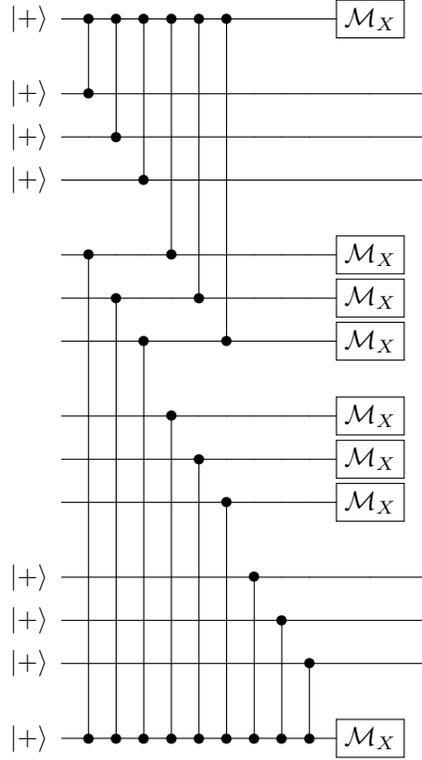


Figure 2.16: Logical CNOT gate acting on two blocks of the repetition code, shown here for code length  $n = 3$ . Ancilla qubits are prepared in the  $X = 1$  eigenstate  $|+\rangle$ , interact via CZ gates with the data qubits, then are measured in the  $X$  basis. (These preparations, gates, and measurements are repeated several times, and the result is determined by a majority vote; the measurement repetition is not shown.) Finally, each qubit in the two input blocks is measured in the  $X$  basis, the results are decoded by a majority vote in each block, and logical Pauli errors in the output blocks are inferred from the results.

gate can be bounded above as [48]

$$\begin{aligned} \varepsilon_{\text{CNOT}} &\leq 4 \binom{n}{\frac{n+1}{2}} (3n\varepsilon_g + 2\varepsilon_m)^{(n+1)/2} + 7n^2\varepsilon'_g \\ &\leq \sqrt{\frac{8}{\pi n}} (12n\varepsilon_g + 8\varepsilon_m)^{(n+1)/2} + 7n^2\varepsilon'_g, \end{aligned} \quad (2.144)$$

where  $\varepsilon_g$  is the probability of a dephasing error and  $\varepsilon'_g$  is the probability of a bit flip error in the CZ gate, and  $\varepsilon_m$  is the error probability in a  $|+\rangle$  preparation or an  $X$  measurement. Here  $n$  is the length of the repetition code, and we assume that each measurement is repeated  $n$  times; the

prefactor of 4 arises because the CNOT gadget contains four measurements decoded by majority vote, any of which might fail, and the factor  $3n\varepsilon_g + 2\varepsilon_m$  is an upper bound on the probability of error for each bit in each decoded block, neglecting the bit flip errors. (The second inequality is obtained from the Stirling approximation.) The term  $7n^2\varepsilon'_g$  is an upper bound on the probability that one or more CZ gates in the CNOT gadget have bit flip errors.

Just to illustrate the robustness of the CNOT gadget with respect to measurement and preparation errors (assuming bit flips are highly suppressed), note that for  $\varepsilon_m = .01$ ,  $\varepsilon_g = 10^{-5}$ , and  $\varepsilon'_g = 10^{-9}$ , by choosing  $n = 11$ , we find  $\varepsilon_{\text{CNOT}} < 10^{-6}$ . The analysis in [48] then shows that, by applying the state distillation ideas from [47], scalable quantum computing can be achieved by augmenting the CNOT gadget with  $| -i \rangle$  and  $| T \rangle$  preparations having the relatively high error rate  $\varepsilon_m$ . This example indicates one possible way in which noisy measurements can be tolerated if appropriate gates are highly reliable.

## 2.13 Conclusions

A “ $0-\pi$  qubit,” is a two-lead superconducting device whose energy is minimized when the superconducting phase difference between the leads is either 0 or  $\pi$ . This qubit, if properly designed, can be very robust with respect to weak local noise [40–42]. In this thesis, we have taken it for granted that near-perfect  $0-\pi$  qubits are attainable, and have asked whether quantum information encoded in such qubits can be processed fault tolerantly. Our conclusion, fleshing out a suggestion in [42], is that highly accurate nontrivial quantum gates can be executed by coupling one or two qubits to a superconducting  $LC$  oscillator with very large inductance. In principle, the gate error becomes exponentially small when  $\sqrt{L/C}$  is large compared to  $\hbar/4e^2 \approx 1.03 k\Omega$ , where  $L$  is the inductance and  $C$  is the capacitance of the oscillator.

We have estimated the gate accuracy using both analytic arguments and numerical simulations. The analytic arguments use approximations, in particular for the analysis of errors due to diabatic transitions and squeezing, which are validated by the simulations. The point of the simulations is not necessarily to capture fully the behavior of realistic devices, but rather to verify that the analytic arguments are on solid ground. The analysis applies to any sufficiently robust  $0-\pi$  qubit, regardless of its internal structure.

The protected gates are the single-qubit phase gate  $\exp(i\frac{\pi}{4}Z)$  and the two-qubit phase gate

$\exp(i\frac{\pi}{4}Z \otimes Z)$ . In both cases, the oscillator starts out in a low-lying energy state; then turning on a tunable Josephson coupling between the qubit(s) and the oscillator prepares a state of the oscillator protected by a continuous variable quantum error-correcting code. The coupling is kept on for a specified time, during which the code state acquires a nontrivial Berry phase, inducing a nontrivial encoded operation, with error probability exponentially small in  $\sqrt{L/C}$ . Then the coupling turns off, returning the oscillator to a low-lying state slightly different from the initial one.

The ramping of the coupling on and off takes place over a time scale  $\tau_J = O(\sqrt{LC})$ , chosen to achieve an optimal compromise between errors due to diabatic transitions (which favor large  $\tau_J$ ) and errors due to squeezing (which favor small  $\tau_J$ ). The gates are robust against generic small errors in the Hamiltonian and thermal effects, due to the good error-correcting properties of the quantum code. Entropy due to noise is mostly absorbed by the oscillator, inflicting little damage on the  $0-\pi$  qubit. These protected phase gates are not a universal gate set by themselves, but a universal fault-tolerant scheme can be built from the protected gates together with single-qubit measurements and noisy unprotected phase gates.

We do not know whether a very robust  $0-\pi$  qubit, and/or the superinductive  $LC$  circuit needed for protected phase gates, will prove to be feasible in superconducting devices. We hope that other settings will be found in which highly reliable quantum gates can be realized by using tunable couplings between qubits and oscillators.

## Appendix 2.A Adiabatic switch

Here we describe the static properties of a single rung of the two-rung circuit depicted in Figure 2.2. This single-rung circuit is of independent interest; because its effective Josephson energy depends sensitively on circuit parameters, it can serve as an “adiabatic switch,” which turns the coupling on and off between superconducting qubits.

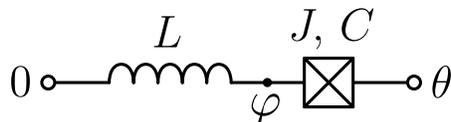


Figure 2.17: Adiabatic switch. For  $\sqrt{L/C} \gg 1$ , the effective Josephson energy  $J_{\text{eff}}$  of the switch is depends sensitively on circuit parameters.

The circuit shown in Figure 2.17 is described by the Hamiltonian

$$H = \frac{Q^2}{2C} + \frac{\varphi^2}{2L} - J \cos(\varphi - \theta), \quad \text{where} \quad Q = -i \frac{\partial}{\partial \varphi}. \quad (2.145)$$

The dynamical variable is the superconducting phase  $\varphi$  at the indicated point, and the phase difference  $\theta$  between the leads is fixed;  $Q$  is the electric charge operator in units of  $2e$ . While  $\theta$  is defined modulo  $2\pi$ , the variable  $\varphi$  is a real number because the states  $|\varphi\rangle$  and  $|\varphi + 2\pi\rangle$  differ by the phase winding in the inductor. We choose units such that  $\hbar = 1$  and express the capacitance and inductance in rationalized units such that  $C^{-1}$  and  $L^{-1}$  have dimensions of energy (or inverse time). Specifically,

$$C = \frac{C_{\text{conv}}}{(2e)^2}, \quad L = \frac{L_{\text{conv}}}{(\hbar/2e)^2}, \quad (2.146)$$

where the subscript “conv” refers to conventional units. We can also relate the parameters  $J, C$  to the commonly used charging energy  $E_C$  and Josephson energy  $E_J$ :

$$E_C = \frac{e^2}{2C_{\text{conv}}} = \frac{1}{8C}, \quad E_J = J_{\text{conv}} = J. \quad (2.147)$$

In most of the chapter we have studied this same Hamiltonian, where  $\theta = 0$  or  $\pi$  and  $J$  pulses on and off during the execution of a quantum gate. But here we focus on the case where  $J$  is time independent.

In the d.c. regime, the whole device behaves like an effective Josephson element — its ground state energy  $E(\theta)$  depends on the phase difference  $\theta$  between the leads. The effective Josephson energy can be characterized by

$$J_{\text{eff}} = E''(0) = \left. \frac{\partial^2 E}{\partial \theta^2} \right|_{\theta=0}. \quad (2.148)$$

We will see that this number varies by orders of magnitude when the circuit parameters change in a much narrower range.

Suppose that  $C$  and  $L$  are fixed, while  $J$  may vary from very small to very large values. In the limit  $J \rightarrow 0$ , we may treat the Josephson energy as a small perturbation of the  $LC$  oscillator, obtaining

$$E(\theta) = -J_{\text{eff}} \cos \theta + \text{const}, \quad (2.149)$$

where

$$J_{\text{eff}} = J \langle \cos \varphi \rangle, = J e^{-\langle \varphi^2 \rangle / 2} = J \exp\left(-\frac{1}{4} \sqrt{L/C}\right). \quad (2.150)$$

In the opposite limit  $J \rightarrow \infty$ , the dynamical phase is locked:  $\varphi \equiv \theta \pmod{2\pi}$ , and therefore

$$E(\theta) = \min_n \frac{(\theta + 2\pi n)^2}{2L} + \text{const}, \quad J_{\text{eff}} = L^{-1}. \quad (2.151)$$

If  $L$  is large, the effective Josephson coupling is suppressed in both limits, but the suppression is exponential only in the limit  $J \rightarrow 0$ .

We say that the circuit is superinducting if the dimensionless characteristic impedance is large, or equivalently if the impedance in conventional units is large compared to the superconducting impedance quantum  $R_Q = \hbar/(2e)^2 \approx 1.03 \text{ k}\Omega$ . Reaching this superinducting regime is a significant engineering challenge, quite hard to achieve using geometric inductance (except perhaps by constructing a coil with a very large number of turns). Indeed, the inductance of a loop of wire is accompanied by a parasitic capacitance such that  $\sqrt{L_{\text{conv}}/C_{\text{conv}}} \sim 4\pi/c \approx 377 \Omega$  (where  $c$  is the speed of light). This is the impedance of the free space, smaller than  $R_Q$  by the factor  $16\pi\alpha$ , where  $\alpha = e^2/\hbar c \approx 1/137$  is the fine structure constant. One possible way to realize a superinductor is to build a long chain of Josephson junctions [44–46]. Another is to use a long wire, thick enough to suppress phase slips, built from a material with large kinetic inductance.

In the case where  $\sqrt{L/C} \gg 1$ , we can compute the ground state energy of the Hamiltonian Equation (2.145) semi-analytically. First we note that in the limit  $L \rightarrow \infty$  the problem reduces to

a particle moving in a periodic potential:

$$H' = \frac{Q^2}{2C} - J \cos \tilde{\varphi}, \quad \text{where } \tilde{\varphi} = \varphi - \theta. \quad (2.152)$$

This approximate Hamiltonian  $H'$  preserves the quasimomentum  $q = Q \bmod 1$ , and therefore can be solved using Bloch wave functions. In the lowest Bloch band, the energy can be expressed as a function of the quasimomentum  $q \in [-\frac{1}{2}, \frac{1}{2}]$  in two different limits:

$$JC \ll 1 : \quad \varepsilon(q) = \frac{q^2}{2C}, \quad (2.153)$$

$$JC \gg 1 : \quad \varepsilon(q) = -J + \frac{\omega}{2} - 2\lambda \cos(2\pi q), \quad (2.154)$$

where

$$\omega = \sqrt{J/C}, \quad \lambda = \frac{4}{\sqrt{\pi}} J^{3/4} C^{-1/4} e^{-8\sqrt{JC}}. \quad (2.155)$$

In the case  $JC \gg 1$ , the system stays near a minimum of the cosine potential at  $\tilde{\varphi} = 2\pi n$  and occasionally tunnels to an adjacent minimum through the potential barrier. In Equation (2.155),  $\omega$  is the angular frequency for small oscillations about the potential minimum, and  $\lambda$  is the amplitude tunneling amplitude.

To compute  $\lambda$  we recall the semiclassical analysis [50] for a particle of mass  $m$  tunneling through a symmetric double-well potential  $V(x)$  with minima at  $x = a, b$ , which yields

$$\lambda = \frac{\omega}{\sqrt{\pi}} e^{-S + \omega\tau/2}, \quad (2.156)$$

where

$$S = \int_a^b dx \sqrt{2mV(x)}, \quad \tau = \int_{a+\Delta x}^{b-\Delta x} dx \sqrt{\frac{m}{2V(x)}}; \quad (2.157)$$

here  $\omega = \sqrt{V''(a)/m}$  and  $\Delta x = (V''(a)m)^{-1/4}$  is the width of the ground state localized around the potential minimum. We obtain Equation (2.154) using  $m = C$  and  $V(\tilde{\varphi}) = J(1 - \cos \tilde{\varphi})$ .

Defining the effective capacitance  $C_{\text{eff}}$  by

$$\frac{1}{C_{\text{eff}}} = \varepsilon''(0), \quad (2.158)$$

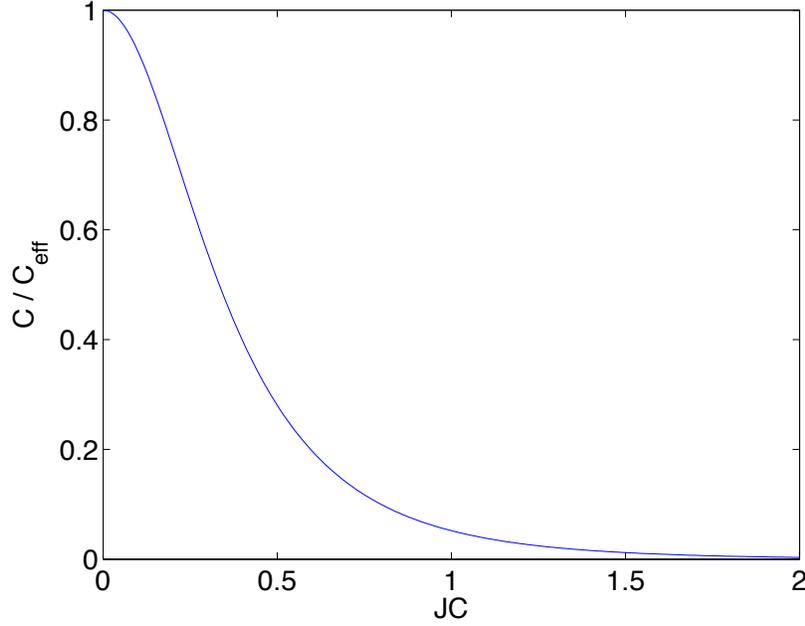


Figure 2.18: The inverse effective capacitance of a Josephson junction as a function of  $JC$ .

we obtain from Equations (2.154) and (2.155) the asymptotic values

$$\frac{C}{C_{\text{eff}}} = \begin{cases} 1, & \text{if } JC \ll 1, \\ 32\pi^{3/2}(JC)^{3/4}e^{-8\sqrt{JC}}, & \text{if } JC \gg 1. \end{cases} \quad (2.159)$$

The numerically computed value of  $C_{\text{eff}}$  as a function of  $JC$  is plotted in Figure 2.18.

Now we return to the original Hamiltonian  $H$  in Equation (2.145). For small but nonzero  $L^{-1}$ , the quasimomentum  $q$ , though not exactly conserved, is a slow variable which can be treated using the adiabatic approximation. We obtain the effective Hamiltonian

$$H_{\text{eff}} = \frac{\varphi^2}{2L} + \varepsilon(q), \quad \text{where } \varphi = i\frac{\partial}{\partial q} + \theta. \quad (2.160)$$

This problem is similar to the one we have just solved, Equation (2.152); now the variable  $q$  is periodic (defined modulo 1), and the parameter  $\theta$  plays the role of quasimomentum.

If  $\sqrt{L/C_{\text{eff}}} \gg 1$ , *i.e.*  $JC \ll \ln(L/C)$ , the ground state energy is given by the formula  $E(\theta) =$

$-J_{\text{eff}} \cos \theta + \text{const}$ , where  $J_{\text{eff}}$  is twice the tunneling amplitude. This expression is analogous to Equation (2.154), except that now we consider tunneling through the periodic “potential”  $\varepsilon(q)$  from an integer value of  $q$  to an adjacent integer value. From Equation (2.156) we find

$$J_{\text{eff}} = \nu C_{\text{eff}}^{-3/4} L^{-1/4} \exp\left(-\mu \sqrt{L/C_{\text{eff}}}\right), \quad (2.161)$$

where

$$\mu = 2 \int_0^{1/2} \sqrt{2C_{\text{eff}}(\varepsilon(q) - \varepsilon(0))} dq, \quad (2.162)$$

$$\nu = \frac{1}{\sqrt{\pi}} \exp\left(\int_0^{1/2} \left((2C_{\text{eff}}(\varepsilon(q) - \varepsilon(0)))^{-1/2} - q^{-1}\right) dq\right). \quad (2.163)$$

The parameters  $\mu$  and  $\nu$  are numbers of order 1: As  $JC$  increases from zero to infinity,  $\mu$  changes from  $1/4 = 0.25$  to  $2/\pi^2 \approx 0.2026$ , and  $\nu$  changes from  $1/\sqrt{\pi} \approx 0.564$  to  $4/\pi^{3/2} \approx 0.718$ .

For Equation (2.161) to apply, we also need that  $J$  not be too small:  $JC \gg (L/C)^{-1/4}$ . For smaller  $J$ , the adiabatic approximation breaks down near the point  $q = \pm 1/2$ , where two branches of the parabola  $\varepsilon(q) = q^2/(2C)$  meet each other. However, the exponential factor  $e^{-S}$  is still correct and coincides with  $\langle \cos \varphi \rangle$  from Equation (2.150).

To summarize, we have calculated the effective Josephson coupling in three different regimes, finding

$$JC \ll (L/C)^{-1/4} : \quad J_{\text{eff}} = J \exp\left(-\frac{1}{4} \sqrt{L/C}\right), \quad (2.164)$$

$$(L/C)^{-1/4} \ll JC \ll \ln(L/C) : \quad J_{\text{eff}} = \nu C_{\text{eff}}^{-3/4} L^{-1/4} \exp\left(-\mu \sqrt{L/C_{\text{eff}}}\right), \quad (2.165)$$

$$JC \gg \ln(L/C) : \quad J_{\text{eff}} = L^{-1}. \quad (2.166)$$

The numerical results for  $\sqrt{L/C} = 40$ , together with plots of Equations (2.164) to (2.166), are shown in Figure 2.19.

## Appendix 2.B Quantifying the gate error

The protected phase gate is executed by coupling the qubit to an oscillator for a prescribed time interval. We assume that the initial state of qubit and oscillator is a product state  $|\psi\rangle \otimes |\psi_{\text{in}}\rangle$ ,

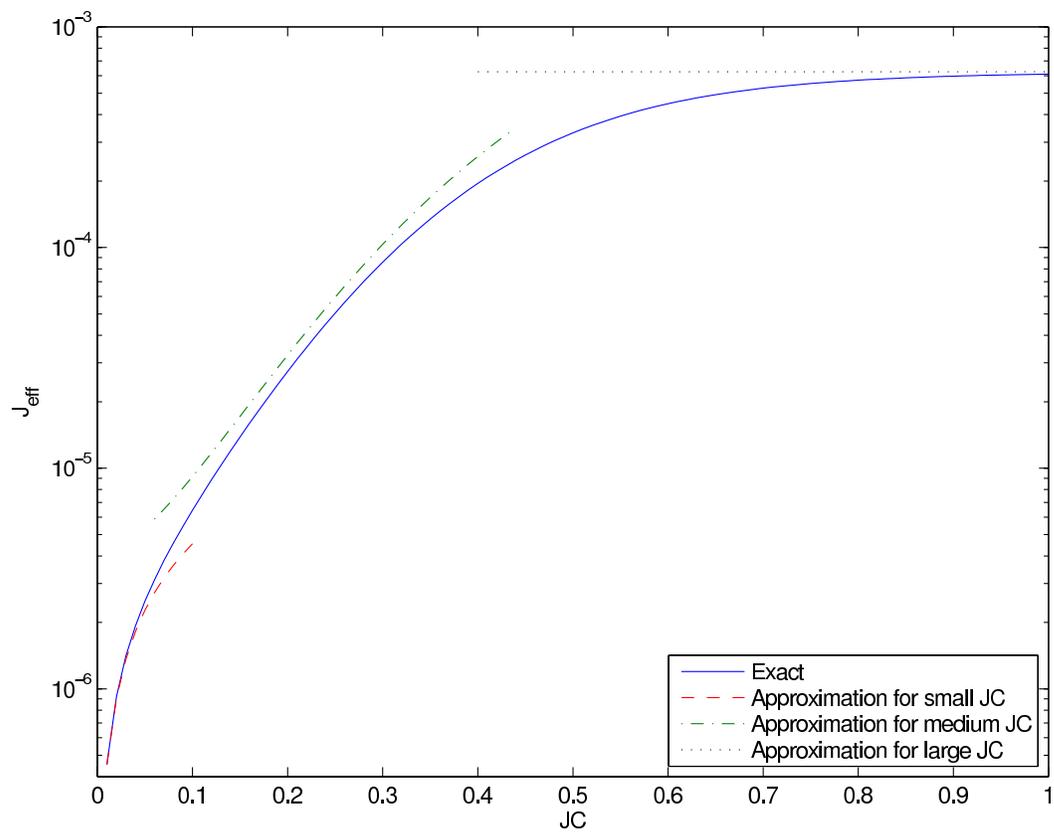


Figure 2.19: The effective Josephson parameter of the adiabatic switch as a function of  $JC$  for  $\sqrt{L/C} = 40$ .

where  $|\psi\rangle = a|0\rangle + ib|1\rangle$  is the initial (normalized) state of the qubit and  $|\psi_{\text{init}}\rangle$  is the initial state of the oscillator. After the coupling between qubit and oscillator is switched off, the joint state of qubit and oscillator becomes

$$|\psi'\rangle = a|0\rangle \otimes |\psi_0\rangle + b|1\rangle \otimes |\psi_1\rangle, \quad (2.167)$$

where  $\langle\psi_0|\psi_0\rangle = \langle\psi_1|\psi_1\rangle = 1$ . We assume that there are no bit flip errors, but there may be a phase error. If the gate is ideal, then  $\langle\psi_0|\psi_1\rangle = 1$ , and the gate rotates the phase of  $|1\rangle$  by the angle  $-\pi/2$  relative to the phase of  $|0\rangle$ . We wish to quantify the error, using some appropriate measure of the deviation of  $\langle\psi_0|\psi_1\rangle$  from 1.

### 2.B.1 Fidelity

One way to quantify the error is to use the fidelity of the actual state with the ideal state. Tracing out the oscillator, we obtain the final density operator for the qubit

$$\rho = \text{tr}_{\text{osc}}|\psi'\rangle\langle\psi'| = \begin{pmatrix} \langle\psi_0|\psi_0\rangle|a|^2 & \langle\psi_1|\psi_0\rangle ab^* \\ \langle\psi_0|\psi_1\rangle a^*b & \langle\psi_1|\psi_1\rangle|b|^2 \end{pmatrix} = \begin{pmatrix} |a|^2 & Ce^{-i\phi}ab^* \\ Ce^{i\phi}a^*b & |b|^2 \end{pmatrix}, \quad (2.168)$$

where

$$\langle\psi_0|\psi_1\rangle = Ce^{i\phi}, \quad (2.169)$$

and  $C = |\langle\psi_0|\psi_1\rangle|$  is real and nonnegative. The fidelity  $F$  with the ideal state  $|\psi_{\text{ideal}}\rangle = a|0\rangle + b|1\rangle$  is

$$\begin{aligned} F &= \langle\psi_{\text{ideal}}|\rho|\psi_{\text{ideal}}\rangle \\ &= |a|^4 + |a|^2|b|^2 (Ce^{i\phi} + Ce^{-i\phi}) + |b|^4 \\ &= \left(|a|^2 + |b|^2\right)^2 - 2(1 - C \cos \phi) |a|^2|b|^2. \end{aligned} \quad (2.170)$$

Thus the ‘‘infidelity’’ (the deviation of  $F$  from 1) is maximal when  $|a|^2 = |b|^2 = 1/2$ , and we conclude that

$$1 - F \leq \frac{1}{2} (1 - C \cos \phi). \quad (2.171)$$

Denoting  $\delta = 1 - C$ , we have

$$1 - F \approx \frac{1}{2}\delta + \frac{1}{4}\phi^2. \quad (2.172)$$

assuming  $\delta, \phi \ll 1$ .

## 2.B.2 Trace norm

Another useful measure is the deviation of  $\rho$  from the ideal density operator  $\rho_{\text{ideal}} = |\psi_{\text{ideal}}\rangle\langle\psi_{\text{ideal}}|$  in the trace norm. We see that

$$\rho_{\text{ideal}} - \rho = \begin{pmatrix} 0 & (1 - Ce^{-i\phi}) ab^* \\ (1 - Ce^{i\phi}) a^*b & 0 \end{pmatrix}, \quad (2.173)$$

whose eigenvalues

$$\pm |a||b|\sqrt{1 + C^2 - 2C \cos \phi} \quad (2.174)$$

have maximal absolute value for  $|a| = |b| = 1/\sqrt{2}$ . Hence, the trace norm satisfies

$$\begin{aligned} \|\rho_{\text{ideal}} - \rho\|_1 &\leq \sqrt{1 + C^2 - 2C \cos \phi} \\ &= |1 - \langle\psi_0|\psi_1\rangle|. \end{aligned} \quad (2.175)$$

Denoting  $\delta = 1 - C$ , we have

$$\begin{aligned} \|\rho_{\text{ideal}} - \rho\|_1 &\leq \sqrt{(1 - C)^2 + 2C(1 - \cos \phi)} \\ &\approx \sqrt{\delta^2 + \phi^2}, \end{aligned} \quad (2.176)$$

assuming  $\delta, \phi \ll 1$ .

## 2.B.3 Kraus operators

It is also useful to have a Kraus operator decomposition of the noise process. We may define states  $|\psi'_0\rangle$  and  $|\psi'_1\rangle$  such that

$$|\psi_0\rangle = e^{-i\phi/2}|\psi'_0\rangle, \quad |\psi_1\rangle = e^{i\phi/2}|\psi'_1\rangle, \quad (2.177)$$

and therefore

$$\langle \psi'_0 | \psi'_1 \rangle = C \geq 0. \quad (2.178)$$

In terms of these states,

$$\begin{aligned} |\psi'\rangle &= a|0\rangle \otimes |\psi_0\rangle + b|1\rangle \otimes |\psi_1\rangle \\ &= \left( e^{-i\phi/2}a|0\rangle + be^{i\phi/2}|1\rangle \right) \otimes \frac{1}{2}(|\psi'_0\rangle + |\psi'_1\rangle) \\ &\quad + \left( e^{-i\phi/2}a|0\rangle - be^{i\phi/2}|1\rangle \right) \otimes \frac{1}{2}(|\psi'_0\rangle - |\psi'_1\rangle), \end{aligned} \quad (2.179)$$

where

$$\begin{aligned} \left\| \frac{1}{2}(|\psi'_0\rangle \pm |\psi'_1\rangle) \right\| &= \sqrt{\frac{1}{2}(1 \pm C)}, \\ (\langle \psi'_0| + \langle \psi'_1|)(|\psi'_0\rangle - |\psi'_1\rangle) &= 0. \end{aligned} \quad (2.180)$$

Therefore,

$$\rho = \mathcal{N}(\rho_{\text{ideal}}) = M_0 \rho_{\text{ideal}} M_0^\dagger + M_1 \rho_{\text{ideal}} M_1^\dagger, \quad (2.181)$$

where

$$\begin{aligned} M_0 &= \sqrt{\frac{1}{2}(1+C)} \begin{pmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{i\phi/2} \end{pmatrix}, \\ M_1 &= \sqrt{\frac{1}{2}(1-C)} \begin{pmatrix} e^{-i\phi/2} & 0 \\ 0 & -e^{i\phi/2} \end{pmatrix}. \end{aligned} \quad (2.182)$$

Note that we have rotated away the ideal phase gate in our definition of the noise operation  $\mathcal{N}$ , so that  $\mathcal{N} = I$  corresponds to the ideal gate.

### 2.B.4 Diamond norm

In some versions of the quantum accuracy threshold theorem, the strength of Markovian noise is characterized by the deviation

$$\varepsilon = \|\mathcal{N} - U\|_{\diamond} \quad (2.183)$$

of a noisy gate  $\mathcal{N}$  from the corresponding ideal gate  $U$  in the “diamond norm” [6]. The advantage of the diamond norm is that it quantifies the damage inflicted by an operation that acts on a subsystem that might be entangled with a complementary subsystem, *e.g.*, a noisy gate acting on a qubit or pair of qubits that is entangled with the rest of the qubits in a quantum computer. The diamond norm  $\|\mathcal{E}\|_{\diamond}$  is defined as the  $L^1$  norm of the extended operator  $\mathcal{E} \otimes I$ ; that is,

$$\|\mathcal{E}\|_{\diamond} = \max_{\sigma} \|\mathcal{E} \otimes I(\sigma)\|_1. \quad (2.184)$$

If  $\mathcal{E}$  acts on a Hilbert space  $\mathcal{H}$  of dimension  $d$ , then  $I$  denotes the identity operator acting on another Hilbert space  $\mathcal{H}'$  of dimension  $d$ , and  $\sigma$  is a state on  $\mathcal{H} \otimes \mathcal{H}'$ .

For the operation defined by Equation (2.181), the two-qubit state  $\sigma$  that maximizes the  $L^1$  distance between  $(\mathcal{N} \otimes I)(\sigma)$  and  $\sigma$  is a maximally entangled pure state, which we may choose to be  $|\phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ . Letting  $\mathcal{N}$  act on the first qubit, we obtain an ensemble of two pure states,

$$\begin{aligned} \frac{1}{\sqrt{2}} \left( e^{-i\phi/2}|00\rangle + e^{i\phi/2}|11\rangle \right), & \quad \text{prob} = \frac{1}{2}(1+C), \\ \frac{1}{\sqrt{2}} \left( e^{-i\phi/2}|00\rangle - e^{i\phi/2}|11\rangle \right), & \quad \text{prob} = \frac{1}{2}(1-C), \end{aligned} \quad (2.185)$$

and the density operator can be expressed as a  $2 \times 2$  matrix acting on the span of  $|00\rangle$  and  $|11\rangle$ :

$$\begin{aligned} \mathcal{N} \otimes I(|\phi^+\rangle\langle\phi^+|) &= \begin{pmatrix} \frac{1}{2} & \frac{1}{2}Ce^{-i\phi} \\ \frac{1}{2}Ce^{i\phi} & \frac{1}{2} \end{pmatrix}, \\ (\mathcal{N} \otimes I - I \otimes I)(|\phi^+\rangle\langle\phi^+|) &= \begin{pmatrix} 0 & \frac{1}{2}(Ce^{-i\phi} - 1) \\ \frac{1}{2}(Ce^{i\phi} - 1) & 0 \end{pmatrix}. \end{aligned} \quad (2.186)$$

Comparing with Equation (2.173) in the case  $a = b = \frac{1}{\sqrt{2}}$ , we find

$$\|\mathcal{N} - I\|_{\diamond} = \max\|\rho_{\text{ideal}} - \rho\|_1 = \sqrt{1 + C^2 - 2C \cos \phi} = |1 - \langle \psi_0 | \psi_1 \rangle|. \quad (2.187)$$

Evidently, extending  $\mathcal{E}$  to  $\mathcal{E} \otimes I$  does not increase its maximal  $L^1$  norm; hence, the noisy gate's deviation from the ideal gate in the diamond norm coincides with the maximal trace distance deviation of the density operator  $\rho$  from the ideal density operator  $\rho_{\text{ideal}}$ .

The accuracy of the two-qubit phase gate can be analyzed in the same way. Now the final state of the oscillator depends on the total phase difference across a pair of  $0-\pi$  qubits; it is  $|\psi_0\rangle$  for the two-qubit states  $|00\rangle$  and  $|11\rangle$ , and  $|\psi_1\rangle$  for the two-qubit states  $|01\rangle$  and  $|10\rangle$ . The eigenvalues of  $\rho_{\text{ideal}} - \rho$  become doubly degenerate, and hence the diamond norm deviation from the ideal gate is twice as large as for the single-qubit phase gate.

Both the fidelity and the diamond norm are useful measures of the gate error. The significant difference is that the diamond norm deviation ( $\approx \sqrt{(1-C)^2 + \phi^2}$ ) is linear in  $\phi$  (for  $\phi$  small and  $C \approx 1$ ), while the infidelity is quadratic in  $\phi$ . The threshold theorem establishes a sufficient condition for scalable quantum computing, expressed as an upper bound on the diamond norm, and it applies under the pessimistic assumption that phase errors accumulate linearly with the circuit size. But if the phase errors are actually random, we might expect them to add in quadrature, and in that case the infidelity may be a more appropriate way to quantify the gate error.

## Appendix 2.C Grid states

Here we provide additional details concerning the properties of grid states, which were omitted from the discussion in Section 2.6.

### 2.C.1 Approximate codewords in $\varphi$ space and $Q$ space

Let  $f$  denote a narrow function in  $\varphi$  space, and  $\tilde{F}$  denote a broad envelope function in  $\varphi$  space. We express the approximate codewords of the continuous variable code as

$$\begin{aligned} |0_C\rangle &= \sqrt{2\pi} \sum_{n \text{ even}} \tilde{F}(\pi n) T(\pi n) \int d\varphi f(\varphi) |\varphi\rangle, \\ |1_C\rangle &= \sqrt{2\pi} \sum_{n \text{ odd}} \tilde{F}(\pi n) T(\pi n) \int d\varphi f(\varphi) |\varphi\rangle, \end{aligned} \quad (2.188)$$

where  $T(a)$  denotes the  $\varphi$  translation operator whose action is  $T(a)|\varphi\rangle = |\varphi + a\rangle$ . The function  $f$  is normalized so that

$$\int d\varphi |f(\varphi)|^2 = 1, \quad (2.189)$$

and if the overlap between peaks centered at distinct integer multiples of  $\pi$  can be neglected, then  $|0_C\rangle$  and  $|1_C\rangle$  are normalized provided

$$2\pi \sum_{n \text{ even}} |\tilde{F}(\pi n)|^2 \approx 1, \quad 2\pi \sum_{n \text{ odd}} |\tilde{F}(\pi n)|^2 \approx 1. \quad (2.190)$$

The approximate codewords in the conjugate basis are

$$\begin{aligned} |+_C\rangle &= \frac{1}{\sqrt{2}} (|0_C\rangle + |1_C\rangle) \\ &= \sqrt{\pi} \sum_n \tilde{F}(\pi n) T(\pi n) \int d\varphi f(\varphi) |\varphi\rangle, \\ |-_C\rangle &= \frac{1}{\sqrt{2}} (|0_C\rangle - |1_C\rangle) \\ &= \sqrt{\pi} \sum_n \tilde{F}(\pi n) (-1)^n T(\pi n) \int d\varphi f(\varphi) |\varphi\rangle, \end{aligned} \quad (2.191)$$

where

$$\pi \sum_n |\tilde{F}(\pi n)|^2 \approx 1. \quad (2.192)$$

To express these states in the  $Q$  basis, we use

$$\begin{aligned}\int d\varphi f(\varphi)|\varphi\rangle &= \int dQ \int d\varphi f(\varphi)|Q\rangle\langle Q|\varphi\rangle \\ &= \int dQ \tilde{f}(Q)|Q\rangle,\end{aligned}\tag{2.193}$$

where

$$\begin{aligned}\tilde{f}(Q) &= \frac{1}{\sqrt{2\pi}} \int d\varphi e^{-iQ\varphi} f(\varphi), \\ f(\varphi) &= \frac{1}{\sqrt{2\pi}} \int dQ e^{iQ\varphi} \tilde{f}(Q);\end{aligned}\tag{2.194}$$

since the function  $f$  is narrow in  $\varphi$  space, its Fourier transform  $\tilde{f}$  is broad in  $Q$  space. The translation operator is represented in  $Q$  space as  $T(a) = \exp(-iQa)$ , and therefore

$$\begin{aligned}|+c\rangle &= \sqrt{\pi} \sum_n \tilde{F}(\pi n) T(\pi n) \int d\varphi f(\varphi)|\varphi\rangle \\ &= \sqrt{\pi} \sum_n \tilde{F}(\pi n) \int dQ e^{-i\pi n Q} \tilde{f}(Q)|Q\rangle, \\ |-c\rangle &= \sqrt{\pi} \sum_n \tilde{F}(\pi n) (-1)^n T(\pi n) \int d\varphi f(\varphi)|\varphi\rangle \\ &= \sqrt{\pi} \sum_n \tilde{F}(\pi n) \int dQ e^{-i\pi n(Q-1)} \tilde{f}(Q)|Q\rangle.\end{aligned}\tag{2.195}$$

Now we reverse the order of the summation and integration, and use the Poisson summation formula, in the form (see Appendix 2.D):

$$\sqrt{\pi} \sum_n e^{-i\pi n Q} \tilde{F}(\pi n) = \sqrt{2} \sum_m F(Q - 2m),\tag{2.196}$$

where

$$\begin{aligned}F(Q) &= \frac{1}{\sqrt{2\pi}} \int d\varphi e^{-iQ\varphi} \tilde{F}(\varphi), \\ \tilde{F}(\varphi) &= \frac{1}{\sqrt{2\pi}} \int dQ e^{iQ\varphi} F(Q).\end{aligned}\tag{2.197}$$

The codewords  $|+C\rangle$  and  $|-C\rangle$  can now be expressed as

$$\begin{aligned} |+_C\rangle &= \sqrt{2} \int dQ \tilde{f}(Q) \sum_{m \text{ even}} F(Q-m)|Q\rangle, \\ |-_C\rangle &= \sqrt{2} \int dQ \tilde{f}(Q) \sum_{m \text{ odd}} F(Q-m)|Q\rangle. \end{aligned} \quad (2.198)$$

Here  $F$  is a narrow function centered at zero in  $Q$  space, so that  $|+_C\rangle$  has support near even integer values of  $Q$ , and  $|-_C\rangle$  has support near odd integer values of  $Q$ . Because  $\tilde{f}(Q)$  is slowly varying, it is nearly constant within each peak, and the codewords can be well approximated as

$$\begin{aligned} |+_C\rangle &\approx \sqrt{2} \sum_{m \text{ even}} \tilde{f}(m) \int dQ F(Q-m)|Q\rangle, \\ |-_C\rangle &\approx \sqrt{2} \sum_{m \text{ odd}} \tilde{f}(m) \int dQ F(Q-m)|Q\rangle, \end{aligned} \quad (2.199)$$

where

$$\begin{aligned} 2 \sum_{m \text{ even}} |\tilde{f}(m)|^2 &\approx \int dQ |\tilde{f}(Q)|^2 \approx 1, \\ 2 \sum_{m \text{ odd}} |\tilde{f}(m)|^2 &\approx \int dQ |\tilde{f}(Q)|^2 \approx 1. \end{aligned} \quad (2.200)$$

### 2.C.2 Phase gate overrotation

We wish to study the performance of the encoded phase gate, which rotates the relative phase of  $|0_C\rangle$  and  $|1_C\rangle$  by  $-\pi/2$ . For ideal codewords, this gate is achieved by applying the unitary operator  $e^{-i\varphi^2/2\pi}$ , which has the value 1 when  $\varphi$  is an even multiple of  $\pi$ , and has the value  $-i$  when  $\varphi$  is an odd multiple of  $\pi$ . Loosely speaking, this operation arises from the harmonic potential of the superinductor, turned on for a specified time interval. An error could occur because the timing is not precisely correct, so that  $e^{-i\varphi^2(1+\varepsilon)/2\pi}$  is applied instead, where  $\varepsilon \ll 1$ .

In  $\varphi$  space, each narrow peak in the approximate codeword is stabilized by the cosine potential, but when the phase gate is overrotated, the relative phases of the peaks are modified, with the peak localized near  $\varphi = n\pi$ , acquiring the phase  $e^{-i\varepsilon\pi n^2/2}$ . Thus, instead of Equation (2.195), the

approximate  $\bar{X}$  eigenstates become

$$|\pm_C\rangle_\varepsilon = \sqrt{\pi} \sum_n (\pm 1)^n e^{-i\varepsilon\pi n^2/2} \tilde{F}(n\pi) \times \int dQ e^{-i\pi n Q} \tilde{f}(Q)|Q\rangle. \quad (2.201)$$

We again use the Poisson summation formula

$$\sqrt{\pi} \sum_n e^{-i\pi n Q} \tilde{F}_\varepsilon(\pi n) = \sqrt{2} \sum_m F_\varepsilon(Q - 2m), \quad (2.202)$$

but now applied to the modified function

$$\tilde{F}_\varepsilon(\varphi) = e^{-i\varepsilon\varphi^2/2\pi} \tilde{F}(\varphi), \quad (2.203)$$

such that

$$\tilde{F}_\varepsilon(n\pi) = e^{-i\varepsilon\pi n^2/2} \tilde{F}(n\pi). \quad (2.204)$$

Therefore, as in Equation (2.55), the approximate codewords  $|\pm_C\rangle_\varepsilon$  can be expressed as

$$\begin{aligned} |+_C\rangle_\varepsilon &\approx \sqrt{2} \sum_{m \text{ even}} \tilde{f}(m) \int dQ F_\varepsilon(Q - m)|Q\rangle, \\ |-_C\rangle_\varepsilon &\approx \sqrt{2} \sum_{m \text{ odd}} \tilde{f}(m) \int dQ F_\varepsilon(Q - m)|Q\rangle, \end{aligned} \quad (2.205)$$

where

$$F_\varepsilon(Q) = \frac{1}{\sqrt{2\pi}} \int d\varphi e^{-iQ\varphi} \tilde{F}_\varepsilon(\varphi). \quad (2.206)$$

### 2.C.3 Imaginary part of overrotation error

As explained in Section 2.6.2, the imaginary part of the gate error due to overrotation is estimated as

$$\text{Im } \eta_\varepsilon = \text{Im } \langle \psi_-^{\text{end}} | \bar{X} | \psi_+^{\text{end}} \rangle, \quad (2.207)$$

where

$$\begin{aligned} \langle \psi_-^{\text{end}} | \bar{X} | \psi_+^{\text{end}} \rangle &\approx 2 \int_{[Q] \text{ even}} dQ |\tilde{f}(Q)|^2 \sum_{\substack{m \text{ even} \\ n \text{ odd}}} F(Q-n)^* F(Q-m) \\ &\quad - 2 \int_{[Q] \text{ odd}} dQ |\tilde{f}(Q)|^2 \sum_{\substack{m \text{ even} \\ n \text{ odd}}} F(Q-n)^* F(Q-m). \end{aligned} \quad (2.208)$$

We evaluate this expression as follows:

$$\begin{aligned} \langle \psi_-^{\text{end}} | \bar{X} | \psi_+^{\text{end}} \rangle &\approx 2 \sum_{m \text{ even}} |\tilde{f}(m)|^2 \int_{m-\frac{1}{2}}^{m+\frac{1}{2}} dQ (F(Q-m-1)^* + F(Q-m+1)^*) F(Q-m) \\ &\quad - 2 \sum_{n \text{ odd}} |\tilde{f}(n)|^2 \int_{n-\frac{1}{2}}^{n+\frac{1}{2}} dQ F(Q-n)^* (F(Q-n-1) + F(Q-n+1)) \\ &\approx \left( 2 \sum_{m \text{ even}} |\tilde{f}(m)|^2 \right) \left( \int_{-1}^0 dQ F(Q-\frac{1}{2})^* F(Q+\frac{1}{2}) + \int_0^1 dQ F(Q+\frac{1}{2})^* F(Q-\frac{1}{2}) \right) \\ &\quad - \left( 2 \sum_{n \text{ odd}} |\tilde{f}(n)|^2 \right) \left( \int_{-1}^0 dQ F(Q+\frac{1}{2})^* F(Q-\frac{1}{2}) + \int_0^1 dQ F(Q-\frac{1}{2})^* F(Q+\frac{1}{2}) \right) \\ &\approx 2 \int_0^1 dQ (\text{Odd} [F(Q+\frac{1}{2})^* F(Q-\frac{1}{2})] - c.c.) \\ &\approx 2 \int_0^\infty dQ (\text{Odd} [F(Q+\frac{1}{2})^* F(Q-\frac{1}{2})] - c.c.). \end{aligned} \quad (2.209)$$

To obtain the first equality, we suppose that the integral is dominated by the overlaps of peaked functions  $\{F(Q-m)\}$  centered at neighboring integer values of  $m$ , and approximate the slowly varying function  $\tilde{f}(Q)$  by a constant in each integral. We obtain the second equality by shifting the integration variable in each integral, and the third equality by using the normalization condition  $2 \sum_{m \text{ even}} |\tilde{f}(m)|^2 \approx 1 \approx 2 \sum_{n \text{ odd}} |\tilde{f}(n)|^2$ , while noting that the integral of the even part of  $F(Q+\frac{1}{2})^* F(Q-\frac{1}{2})$  cancels between the integrals over  $[-1, 0]$  and  $[0, 1]$ . ( $\text{Odd}[G(Q)] \equiv \frac{1}{2}(G(Q) - G(-Q))$  denotes the odd part of the function  $G(Q)$ , and  $c.c.$  denotes the complex conjugate.) Finally, because the integrand decays rapidly, we make a negligible error by extending the upper limit of integration from 1 to infinity. Thus we conclude that

$$\text{Im } \eta_\varepsilon \approx 4 \text{Im} \int_0^\infty dQ \text{Odd} [F(Q+\frac{1}{2})^* F(Q-\frac{1}{2})].$$

## Appendix 2.D Poisson summation formula

To derive Equation (2.196), we note that

$$G(Q) \equiv \sum_m F(Q - 2m) \quad (2.210)$$

is a periodic function of  $Q$  with period two, and therefore has a Fourier series expansion

$$G(Q) = \sum_n e^{-i\pi n Q} \tilde{G}_n, \quad (2.211)$$

where

$$\begin{aligned} \tilde{G}_n &= \frac{1}{2} \int_0^2 dQ e^{i\pi n Q} G(Q) \\ &= \frac{1}{2} \int_0^2 dQ \sum_m F(Q - 2m) e^{i\pi n Q} e^{-i2\pi m n} \end{aligned} \quad (2.212)$$

(in the last equality we have inserted  $e^{-i2\pi m n} = 1$ ). Now we can combine the integral over  $Q$  from 0 to 2 and the sum over  $m$ , obtaining an integral over  $Q$  from  $-\infty$  to  $\infty$ ; therefore,

$$\tilde{G}_n = \frac{1}{2} \int_{-\infty}^{\infty} dQ F(Q) e^{i\pi n Q} = \sqrt{\frac{\pi}{2}} \tilde{F}(\pi n), \quad (2.213)$$

and combining Equation (2.210) with Equation (2.211) yields

$$G(Q) = \sum_m F(Q - 2m) = \sqrt{\frac{\pi}{2}} \sum_n e^{-i\pi n Q} \tilde{F}(\pi n). \quad (2.214)$$

A more general formula is also sometimes useful:

$$\sum_m F(Q - 2m) e^{i\pi(Q-2m)\alpha} = \sqrt{\frac{\pi}{2}} \sum_n e^{-i\pi n Q} \tilde{F}(\pi(n + \alpha)). \quad (2.215)$$

We may obtain Equation (2.215) from Equation (2.214) by observing that multiplying  $F(Q)$  by  $e^{i\pi Q \alpha}$  is equivalent to shifting the argument of its Fourier transform  $\tilde{F}(\varphi)$  by  $\pi\alpha$ .

## Appendix 2.E Diabatic transitions in a two-level system

Here we derive a formula for the probability of a diabatic transition in a time dependent two-level system, used in Section 2.7 to estimate the probability of a transition in the oscillator as the coupling between the oscillator and the  $0-\pi$  qubit ramps on or off.

We consider the Schrödinger equation

$$\frac{d}{dt}|\psi(t)\rangle = -i H(t)|\psi(t)\rangle \quad (2.216)$$

(with  $\hbar = 1$ ) for the two-level time-dependent Hamiltonian

$$H(t) = -\Delta \sigma^Z - V_0 e^{t/\tau} \sigma^X, \quad (2.217)$$

where  $\sigma^Z, \sigma^X$  are the Pauli matrices. If we express the time  $t$  in units of  $\tau$ , and absorb  $V_0$  by shifting the time variable, the Hamiltonian becomes (assuming  $V_0 > 0$ )

$$H(t) = -u \sigma^Z - e^t \sigma^X, \quad (2.218)$$

where

$$u = \tau \Delta. \quad (2.219)$$

In the limit  $t \rightarrow -\infty$ , the second term is negligible, and the general solution becomes

$$|\psi(t)\rangle = c_0 e^{iut} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_1 e^{-iut} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad (2.220)$$

while in the limit  $t \rightarrow \infty$ , the first term is negligible and the solution is

$$|\psi(t)\rangle = c_+ e^{ie^t} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_- e^{-ie^t} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \quad (2.221)$$

Our goal is to find the S-matrix relating these two asymptotic solutions:

$$\begin{pmatrix} c_+ \\ c_- \end{pmatrix} = S \begin{pmatrix} c_0 \\ c_1 \end{pmatrix}. \quad (2.222)$$

Defining

$$|\psi(t)\rangle \equiv \begin{pmatrix} c_0(t) \\ c_1(t) \end{pmatrix} = \begin{pmatrix} e^{iut} \tilde{c}_0(t) \\ e^{-iut} \tilde{c}_1(t) \end{pmatrix}, \quad (2.223)$$

the Schrödinger equation becomes

$$\begin{aligned} \frac{d\tilde{c}_0}{dt} &= i e^{(1-2iu)t} \tilde{c}_1 \\ \frac{d\tilde{c}_1}{dt} &= i e^{(1+2iu)t} \tilde{c}_0. \end{aligned} \quad (2.224)$$

Assuming  $u > 0$ , the solution  $|\psi^{(0)}(t)\rangle$  that starts out in the ground state obeys the initial conditions  $\tilde{c}_0 \rightarrow 1$  and  $\tilde{c}_1 \rightarrow 0$  as  $t \rightarrow -\infty$ ; expanded as a power series in  $e^t$ , this solution is

$$\begin{aligned} \tilde{c}_0(t) &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \frac{\Gamma(\frac{1}{2} + iu)}{\Gamma(\frac{1}{2} + iu + n)} \left(\frac{e^t}{2}\right)^{2n} \\ \tilde{c}_1(t) &= \left(\frac{i}{2}\right) e^{(1+2iu)t} \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \frac{\Gamma(\frac{1}{2} + iu)}{\Gamma(\frac{3}{2} + iu + n)} \left(\frac{e^t}{2}\right)^{2n}. \end{aligned} \quad (2.225)$$

Matching this formula to the power series expansion for the Bessel function

$$J_\nu(x) = \sum_{n=0}^{\infty} \frac{1}{n! \Gamma(\nu + 1 + n)} \left(\frac{x}{2}\right)^{\nu+2n}, \quad (2.226)$$

we find

$$\begin{aligned} \tilde{c}_0(t) &= \Gamma(\frac{1}{2} + iu) \left(\frac{e^t}{2}\right)^{\frac{1}{2}-iu} J_{-\frac{1}{2}+iu}(e^t), \\ \tilde{c}_1(t) &= \frac{i}{2} e^{(1+2iu)t} \Gamma(\frac{1}{2} + iu) \left(\frac{e^t}{2}\right)^{-\frac{1}{2}-iu} J_{\frac{1}{2}+iu}(e^t), \end{aligned} \quad (2.227)$$

and therefore

$$|\psi^{(0)}(t)\rangle = 2^{-\frac{1}{2}+iu} e^{t/2} \Gamma(\frac{1}{2} + iu) \begin{pmatrix} J_{-\frac{1}{2}+iu}(e^t) \\ i J_{\frac{1}{2}+iu}(e^t) \end{pmatrix}.$$

To find the solution  $|\psi^{(1)}(t)\rangle$  that starts out in the excited state, it suffices to change the sign of  $u$  and interchange  $c_0, c_1$ ; hence,

$$|\psi^{(1)}(t)\rangle = 2^{-\frac{1}{2}-iu} e^{t/2} \Gamma(\frac{1}{2} - iu) \begin{pmatrix} i J_{\frac{1}{2}-iu}(e^t) \\ J_{-\frac{1}{2}-iu}(e^t) \end{pmatrix}.$$

From the asymptotic behavior

$$J_\nu(x) \approx \sqrt{\frac{2}{\pi x}} \cos\left(x - \left(\nu + \frac{1}{2}\right)\frac{\pi}{2}\right) \quad (2.228)$$

of the Bessel function as  $x \rightarrow \infty$ , we find how the solutions  $|\psi^{(0,1)}(t)\rangle$  behave for  $t \rightarrow \infty$ :

$$\begin{aligned} |\psi^{(0)}(t)\rangle &\approx \frac{2^{iu} \Gamma(\frac{1}{2} + iu)}{\sqrt{2\pi}} \left( e^{\frac{\pi}{2}u} e^{ie^t} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + e^{-\frac{\pi}{2}u} e^{-ie^t} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right), \\ |\psi^{(1)}(t)\rangle &\approx \frac{2^{-iu} \Gamma(\frac{1}{2} - iu)}{\sqrt{2\pi}} \left( e^{-\frac{\pi}{2}u} e^{ie^t} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + e^{\frac{\pi}{2}u} e^{-ie^t} \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right). \end{aligned} \quad (2.229)$$

Hence, we conclude that the S-matrix is

$$\begin{pmatrix} f(u) & f(-u) \\ f(-u)^* & -f(u)^* \end{pmatrix}, \quad (2.230)$$

where

$$f(u) = \frac{1}{\sqrt{2\pi}} 2^{iu} \Gamma(\frac{1}{2} + iu) e^{\frac{\pi}{2}u}. \quad (2.231)$$

The probability of a transition from the ground state  $|0\rangle$  to the excited state  $|-\rangle$ , or from the excited state to the ground state, is

$$P(0 \rightarrow -) = |f(-u)|^2 = \frac{1}{2\pi} \Gamma(\frac{1}{2} + iu)\Gamma(\frac{1}{2} - iu)e^{-\pi u}; \quad (2.232)$$

from the identity  $\Gamma(x)\Gamma(1-x) = \frac{\pi}{\sin \pi x}$ , we obtain

$$P(0 \rightarrow -) = P(1 \rightarrow +) = \frac{e^{-\pi u}}{2 \cos(i\pi u)} = \frac{e^{-\pi u}}{e^{\pi u} + e^{-\pi u}} = \frac{1}{2} (1 - \tanh \pi u). \quad (2.233)$$

(The probability that no transition occurs is given by the same formula, but with  $u$  replaced by  $-u$ .) For  $u$  large, *i.e.*, when the initial energy splitting  $2\Delta$  is large compared to the time scale  $\tau$  for the perturbation to turn on, the transition probability is exponentially suppressed:

$$P(0 \rightarrow -) = P(1 \rightarrow +) \approx e^{-2\pi u} = e^{-2\pi\tau\Delta}. \quad (2.234)$$

## Chapter 3

# Asymmetric Bacon-Shor Gadgets for Biased Noise

*go dreamless knaves on Shadows fed,  
your Harry's Tom, your Tom is Dick;  
while Gadgets murder squawk and add,  
the cult of Same is all the chic;*

—E.E. Cummings [51]

This chapter is based on work that was published in [2].

### 3.1 Introduction

The theory of fault-tolerant quantum computation [29, 52] has established that noisy quantum computers can operate reliably provided the noise is neither too strong nor too strongly correlated. In a fault-tolerant quantum circuit, carefully designed gadgets process logical qubits protected by quantum error-correcting codes.

Typical fault-tolerant gadgets are designed to work effectively against generic noise without any special structure. But in some physical settings, the noise is expected to be highly biased, with dephasing in the computational basis far more likely than bit flips. This chapter addresses how noise bias can be exploited to improve the reliability of fault-tolerant quantum circuits.

Specifically, we have in mind a setting in which, to an excellent approximation, the computational basis states  $\{|0\rangle, |1\rangle\}$  are the energy eigenstates for an unperturbed qubit, and single-qubit

gates which are diagonal in this basis can be performed by adiabatically adjusting the energy splittings between these states. Similarly, diagonal two-qubit gates are performed by adjusting the energy splittings of the four (approximate) energy eigenstates  $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ . Various noise sources may induce fluctuations in these energy spacings, and pulse imperfections may cause gates to be over-rotated or under-rotated; in either case the noisy gate deviates from the ideal gate, but remains diagonal. Other physical processes may induce transitions between energy eigenstates, but we will assume that this non-diagonal noise is far weaker than the dominant diagonal noise.

Our scheme for fighting biased noise is based on asymmetric Bacon-Shor codes [53–55]. These quantum codes combine together a length- $m$  repetition code protecting against dephasing and a length- $n$  repetition code protecting against bit flips; by “asymmetric” we mean  $m > n$ , so the code works more effectively against dephasing than against bit flips. A complete universal set of fault-tolerant gates can be constructed using only diagonal gates, plus single-qubit measurements in the  $X$  eigenstate basis  $\{|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)\}$  and preparation of single-qubit  $|+\rangle$  states.

In most previous studies of quantum fault-tolerance based on Bacon-Shor codes, these codes have been concatenated to build a hierarchy of codes within codes [56, 57]. Here we will mainly study the performance of a single large block code rather than a concatenated scheme. For the Bacon-Shor code family, used without concatenation, there is no accuracy threshold; rather, for a fixed value of the noise strength and noise bias, there is an optimal block size which achieves the best performance.

Our gadget constructions and analysis extend the results in [48, 49], where the case  $n = 1$  was regarded as the bottom layer of a concatenated code. In contrast to [48, 49], we consider the case where the gates are required to be geometrically local in a two-dimensional array. In the geometrically local case, logical blocks can be transported as needed via teleportation, with the fundamental operations still limited to diagonal gates,  $X$  measurements, and  $|+\rangle$  preparations.

We have obtained upper bounds on the optimal performance and overhead cost of logical gates by deriving analytic formulas, and our estimates are far from tight. We expect that significantly better results could be obtained using numerical simulations assuming independent biased noise, which we have not attempted. In particular, our analysis includes quite conservative estimates of the probability of failure for preparations of cat states that are used for measurements of logical Pauli operators.

Another particularly attractive approach to achieving quantum fault tolerance with geomet-

rically local gates is based on topological codes [58–60], which unlike Bacon-Shor codes have an accuracy threshold and hence can reach arbitrarily low error rates per logical gate. Topological codes can be adapted for optimized protection against biased noise, and they perform well against biased noise even without any such adaptation. Our current results do not conclusively identify noise parameter regimes for which Bacon-Shor codes are clearly superior to topological codes. However, our scheme has several appealing features — for example, the optimal error rate per logical gate is achieved with a relatively modest number of physical qubits per code block, and is not too adversely affected when qubit measurements are noisier than quantum gates. Furthermore, only relatively modest classical computational resources are needed to interpret error syndromes.

Typical fault-tolerant gadgets are designed to work effectively against generic noise without any special structure. But in some physical settings, the noise is expected to be highly biased, with dephasing in the computational basis far more likely than bit flips. This chapter addresses how noise bias can be exploited to improve the reliability of fault-tolerant quantum circuits.

We describe our biased noise model in Section 3.2, discuss appropriate gate sets in Section 3.3, and review Bacon-Shor codes in Section 3.4. In Section 3.5, we construct a set of fault-tolerant gadgets for logical CSS operations. We obtain upper bounds on the performance of our fault-tolerant logical CNOT gate in Section 3.6, repeat the analysis in Section 3.7 for the case where two-qubit gates are required to be geometrically local, and report numerical values in Section 3.8. In Sections 3.9 and 3.10 we discuss the state injection and state distillation procedures needed to complete our fault-tolerant universal gate set.

Fault-tolerant gadgets protecting against biased noise have also been discussed previously in [61–63].

## 3.2 Dephasing biased noise

Most quantum error correcting codes and fault tolerant circuits are designed without assuming any structure in the noise. For storing information, it is common to model the effects of noise as a depolarizing noise channel

$$\mathcal{N}_{\text{dep}}(\rho) = (1 - p)\rho + pI/2, \quad (3.1)$$

on each qubit, which replaces each qubit with the totally mixed state with some probability  $p$ . For error-correcting circuits, another useful class of error models are stochastic error models, where  $r$

locations are specified to be faulty, with probability  $\leq p^r$  and the resulting error at each location can be chosen arbitrarily— even adversarially— from some specified set, usually the Pauli errors  $\{X, Y, Z\}$ .

In the absence of more information about the structure of the noise, it is reasonable to take the attitude of assuming as little about the noise as possible, and these models are well justified. However, in practice, any physical implementation of a quantum computer will have additional noise structure, and it will be extremely beneficial to take into account as much of this structure as possible when designing and choosing codes and fault-tolerant gadgets. Details of the noise structure can be observed by the techniques of quantum process tomography [64–67].

Often knowledge about additional noise structure can make it harder to use error correction and fault tolerance effectively. For instance, the presence of faults which are correlated in time or space can reduce the effective distance of an error-correcting code, necessitating a larger code to achieve the same level of error correction. In other cases, this additional structure can lead to improvements in the effectiveness of error correction, as one can tailor the choice of code and gadget to the noise situation at hand.

One common such instance is where noise can be modeled by an amplitude damping channel

$$\mathcal{N}_{\text{damp}}(\rho) = E_0\rho E_0^\dagger + E_1\rho E_1^\dagger, \quad (3.2)$$

with

$$E_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix} \quad E_1 = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix}, \quad (3.3)$$

which describes an energy dissipating process, where there is a probability  $\gamma$  of relaxing from the excited state  $|1\rangle$  to the ground state  $|0\rangle$ . Under this noise model, it has been shown that an approximate recovery map can be easily constructed, allowing a simpler error correction procedure which compares favorably to other methods using exhaustive search [68]. Additionally, one can use “refrigeration” techniques to compute for an exponentially long time in the number of qubits, without an external supply of fresh ancillas [69]. Essentially this is because the amplitude damping channel itself provides the refrigeration capability to produce fresh qubits.

The noise structure we will be particularly interested in is the case of dephasing biased noise,

or simply “biased noise.” In this case, the probability  $\varepsilon$  of a dephasing error  $Z$ , which causes an unexpected phase difference between the  $|0\rangle$  and  $|1\rangle$  states, is larger than the probability  $\varepsilon'$  of other errors like the bit-flip error  $X$ , which flips the states  $|0\rangle$  and  $|1\rangle$  to each other. The ratio of these two rates will be called the bias  $\beta$ :

$$\beta = \varepsilon/\varepsilon'. \quad (3.4)$$

More precisely, when we consider faults that happens in circuits, we will consider a local stochastic biased noise model, which distinguishes between two groups of faults. The first group, dephasing faults or diagonal faults, can be modeled by inserting after the gate a trace-preserving completely positive map, whose Kraus operators are all diagonal in the computational basis. Because our fault-tolerant circuits consist of Clifford gates and measurements, we can always assume that the most harmful such fault would produce a Pauli  $Z$  error (or some product of Pauli  $Z$  errors on a multi-qubit gate) after the gate. These faults will be associated with the probability  $\varepsilon$ . The second group of faults, which we will call general or non-diagonal faults, are associated with the smaller probability  $\varepsilon'$ . They may consist of any trace-preserving completely positive map, with no restriction on the form of the Kraus operators. The most harmful of these faults will be something that produces a Pauli error. The argument also applies to more general non-diagonal faults—by expanding each Kraus operator in terms of Pauli operators, and noting that all syndrome bits are measured in the  $X$  basis, we can bound the probability for each syndrome measurement outcome in terms of  $\varepsilon$  and  $\varepsilon'$  as above. We will assume the local stochastic condition, which says that the sum of the probabilities for all fault paths with  $r$  dephasing faults and  $s$  general faults at specified circuit locations obeys

$$P \leq \varepsilon^r (\varepsilon')^s. \quad (3.5)$$

In our fault analysis, we will assume that, once the locations of the faults have been specified, the fault can be chosen adversarially to be most likely to cause an overall failure.

It is quite easy to envision situations in which there is a substantial dephasing noise bias when storing information for some period. In particular, it is very common to identify the computational basis  $\{|0\rangle, |1\rangle\}$  with the low-lying energy eigenstates of some physical system. In that case, there is an energetic penalty associated with flipping the state from  $|0\rangle$  to  $|1\rangle$ —namely, the energy gap  $\Delta$  between the two states. For thermal noise with temperature  $T$ , the probability of a bit flip will be suppressed by a factor like  $e^{-\Delta/kT}$ . By contrast, dephasing errors correspond to the *measuring* of

the qubit by its environment in the computational basis, which we can expect to be fairly common.

### 3.3 Bias-compatible computation

In the course of a computation, it is possible for much of the noise bias structure to be washed out by the complex action of a sequence of unitary evolutions. Gates such as the Hadamard operation  $H$  can flip  $Z$ -type errors into  $X$ -type errors, while other operations like the CNOT gate can copy one  $X$  error into many, according to the propagation rule

$$\text{CNOT} : XI \rightarrow XX. \quad (3.6)$$

Such gates can cause the resulting noise to more closely resemble an unbiased noise model. Furthermore, errors that occur in the middle of a particular operation can be transformed by the remaining portion into a different type of error. Nevertheless, by restricting to an appropriate gate set, one can maintain a noise bias throughout a computation. An appropriate set of gates,  $\mathcal{G}_{\text{fund}}$ , first noted in [48], consists of the following operations:

$$\mathcal{G}_{\text{fund}} = \{\mathcal{P}_{|+\rangle}, \mathcal{M}_Z, CZ\}. \quad (3.7)$$

The preparation  $\mathcal{P}_{|+\rangle}$  of a qubit in the  $+1$   $X$ -basis eigenstate

$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \quad (3.8)$$

preserves a bias because an error in the preparation is equivalent to some probability of producing  $|-\rangle = Z|+\rangle$ , and so behaves like a correct preparation followed by a  $Z$  error. Similarly, an erroneous result of an  $X$ -basis measurement,  $\mathcal{M}_X$ , is equivalent to a  $Z$  error just before a correct measurement.

The final operation, the two-qubit controlled  $Z$  operation

$$\text{CZ} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & -1 \end{pmatrix}, \quad (3.9)$$

is manifestly diagonal in the computational basis. It can be performed by a unitary evolution that never leaves the computational basis, so a dephasing  $Z$  error occurring partway through the evolution would remain a  $Z$ -type error.

The CZ operation also has helpful commutation properties that help preserve the dephasing bias:

$$\text{CZ} : XI \rightarrow XZ \quad (3.10)$$

$$IX \rightarrow ZX \quad (3.11)$$

$$ZI \rightarrow ZI \quad (3.12)$$

$$IZ \rightarrow IZ. \quad (3.13)$$

$Z$  errors commute with the CZ operation, while  $X$  errors propagate add a  $Z$  error on the partner qubit. This ensures while  $X$  errors propagate, they propagate into the more common  $Z$  errors.

The gate set  $\mathcal{G}_{\text{fund}}$  is highly restrictive: it does not even generate the entire Clifford group, let alone enough operations for universal quantum computation. Even the operations that are possible, such as the CNOT gate and  $Z$ -basis measurement, require extra ancillas and multiple fundamental operations to perform. Nevertheless, it is a sufficient set of operations to perform error correction on CSS type codes. With the appropriate use of state distillation and teleportation, discussed further in Chapter 4, we can complete a universal basis of gates.

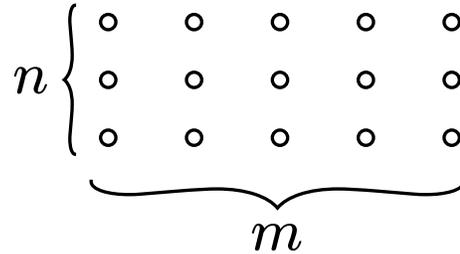


Figure 3.1: Bacon-Shor qubits arranged in an  $n \times m$  lattice.

### 3.4 Bacon-Shor codes

Bacon-Shor codes are a family of quantum subsystem codes. Conceptually the simplest possible quantum code, they can be constructed by concatenation of two classical repetition codes, one protecting against dephasing  $Z$  errors and one protecting against bit-flip  $X$  errors. By choosing the distances of these two codes separately, we can independently tailor the amount of error correction for each type of error. The code is called symmetric when these distances are chosen to be the same, and antisymmetric when they differ.

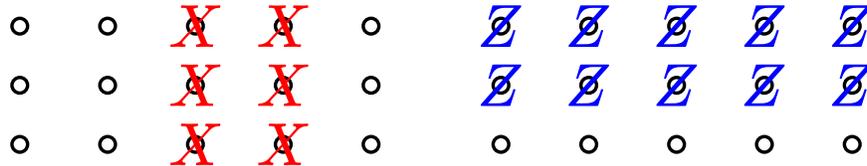


Figure 3.2:  $X$ -type stabilizer operator (left) and  $Z$ -type stabilizer operator (right).

A convenient way to think about this code is to picture the qubits as arranged in a rectangular lattice with  $n$  rows and  $m$  columns, as in Figure 3.1. A single qubit is encoded into this block. The  $X$ -type stabilizer generators apply an  $X$  operation to each qubit in a pair of adjacent columns, while the  $Z$ -type stabilizer generators do the same for  $Z$  operations over the rows. Examples are pictured

in Figure 3.2. Any  $X$ - and  $Z$ -type stabilizers intersect nontrivially on four qubits and therefore commute. The logical  $X$  and  $Z$  operators look similar to the stabilizers of the corresponding type, except that they act on a single row or single column. Each intersects nontrivially twice with any stabilizer generator and once with its logical partner, giving the appropriate commutation statistics.

As a subsystem code, the code block also contains extra gauge degrees of freedom that commute with the logical operator and the stabilizers, but not necessarily each other. These operations could be formed into logical pairs themselves, but they would have low distance and so are not useful to encode any information. The  $Z$ -type gauge operators are generated by weight-two operators, which apply a  $Z$  operation to a any pair of adjacent qubits in the same row, and the  $X$ -type operators are generated by similar operators, which apply an  $X$  operation to any pair of adjacent qubits in the same column.

It is very useful to note that each  $X$ -type stabilizer generator is a product of  $n$  weight-two gauge operators, and each  $Z$ -type stabilizer generator is a product of  $m$  weight-two gauge operators. This allows one to measure the error syndrome by measuring only the gauge operators, which are weight-two and, if our mental picture of the block also corresponds to a physical layout, geometrically neighboring. By taking the parity of the gauge measurement results for all the gauge operators in a particular pair of rows or columns, we determine the value of the corresponding stabilizer for that pair.

### 3.5 Fault-tolerant gadgets

This section will explain fault-tolerant circuits using the fundamental bias-compatible gate set

$$\mathcal{G}_{\text{fund}} = \{\mathcal{P}_{|+\rangle}, \mathcal{M}_Z, CZ\}. \quad (3.14)$$

to implement the logical gate set

$$\mathcal{G}_{\text{CSS}} = \{\mathcal{P}_{|0\rangle}^L, \mathcal{P}_{|+\rangle}^L, \text{CNOT}^L, \mathcal{M}_X^L, \mathcal{M}_Z^L\}. \quad (3.15)$$

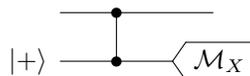


Figure 3.3:  $Z$  measurement performed with  $X$  measurement and an ancilla.

To achieve universal computation, we will supplement the logical gate set  $\mathcal{G}_{\text{CSS}}$  with the ability to prepare noisy copies of the physical states

$$|i\rangle = \frac{1}{\sqrt{2}} (|0\rangle + i|1\rangle) \quad (3.16)$$

and

$$|T\rangle = \frac{1}{\sqrt{2}} (|0\rangle + e^{i\pi/4}|1\rangle). \quad (3.17)$$

In Section 3.9 and Section 3.10, we will use the ideas of state injection and distillation to produce high-fidelity *logical* copies of these states. These states can be consumed to perform gates that complete a universal set.

### 3.5.1 Logical $X$ measurement gadget.

To perform a destructive logical  $X$  measurement  $\mathcal{M}_X^L$ , we first perform independent measurements  $\mathcal{M}_X$  on all qubits in the code block, and then compute the parity of the  $n$  measurement outcomes in each column. With no errors, each of these parity results would be the same, but some may be corrupted due to errors. Taking a majority vote of the  $m$  column parities yields the logical measurement outcome. The result agrees with an ideal measurement if the sum of the number of  $Z$  errors in the block and the number of faulty  $\mathcal{M}_X$  measurements is no larger than  $(m-1)/2$ , so that there are errors in less than half of the columns.

### 3.5.2 Logical $Z$ measurement gadget.

A destructive logical  $Z$  measurement  $\mathcal{M}_Z^L$  could be performed in much the same manner as the  $X$  measurement, with two modifications: using physical  $\mathcal{M}_Z$  measurements instead of  $\mathcal{M}_X$ , and taking parities of rows instead of columns. If we can only use  $\mathcal{M}_X$  as the fundamental measurement, we can use the circuit of Figure 3.3 to get a measurement in the  $Z$  basis indirectly, using an ancilla

qubit.

A non-destructive  $\mathcal{M}_Z^L$  measurement can be performed by measuring (non-destructively) the  $Z$ -basis parity of each length- $m$  row of the code block, and taking a majority vote of the  $n$  row outcomes. (We say the measurement is “non-destructive” if the ideal measurement procedure leaves encoded  $Z^L$  eigenstates intact.) In principle, the row parity measurement could be done by preparing a single ancilla qubit in the state  $|+\rangle$ , performing  $m$  CZ gates in succession acting on the ancilla qubit and the  $m$  qubits in the row, and finally measuring the ancilla qubit in the  $X$  basis. However, this procedure is not fault-tolerant for two reasons. First, a single  $X$  error acting on the ancilla can propagate multiple times, producing a high-weight  $Z$  error acting on the qubits in the row. Second, a single  $Z$  error acting on the ancilla qubit can flip the measurement outcome.

This first problem can be addressed by replacing the single-qubit ancilla with an  $m$ -qubit cat state

$$|+\text{cat}\rangle = \frac{1}{\sqrt{2}} \left( |0\rangle^{\otimes m} + |1\rangle^{\otimes m} \right), \quad (3.18)$$

so that each ancilla qubit interacts via a CZ gate with only one data qubit, limiting the error propagation. The fault-tolerant preparation of the cat state is discussed below. After the  $m$  CZ gates, the cat state is read out in the basis  $|\pm\text{cat}\rangle = \frac{1}{\sqrt{2}} \left( |0\rangle^{\otimes m} \pm |1\rangle^{\otimes m} \right)$  to determine the row parity. This measurement of  $X^{\otimes m}$  is performed destructively by measuring each of the  $m$  ancilla qubits in the  $X$  basis and computing the parity of the outcomes.

A cat state of this sort is essentially the most fragile possible state in the presence of dephasing noise: a single  $Z$  error on any of the cat state qubits can flip its state from  $|+\text{cat}\rangle$  to  $|-\text{cat}\rangle$ . At the same time, the problem it was meant to solve comes from the more rare  $X$  error events. Therefore it is not too surprising that, depending on the rates for diagonal and non-diagonal faults, it may be advantageous to use a shorter ancilla, prepared in a cat state of length  $p$ , where  $1 \leq p < m$ . In that case, some ancilla qubits interact with more than one data qubit, increasing the danger of error propagation, but, on the other hand, errors are less likely to occur during the preparation of the (shorter) cat state. However, using a length- $m$  cat state has the significant advantage that the  $\mathcal{M}_Z^L$  measurement can be conveniently executed using geometrically local gates on the two-dimensional lattice, as we will discuss in Section 3.7.

To address the second problem, that a single fault can flip the measured parity, the parity measurement is repeated  $r$  times for each row, each time with a fresh cat state, and the majority of

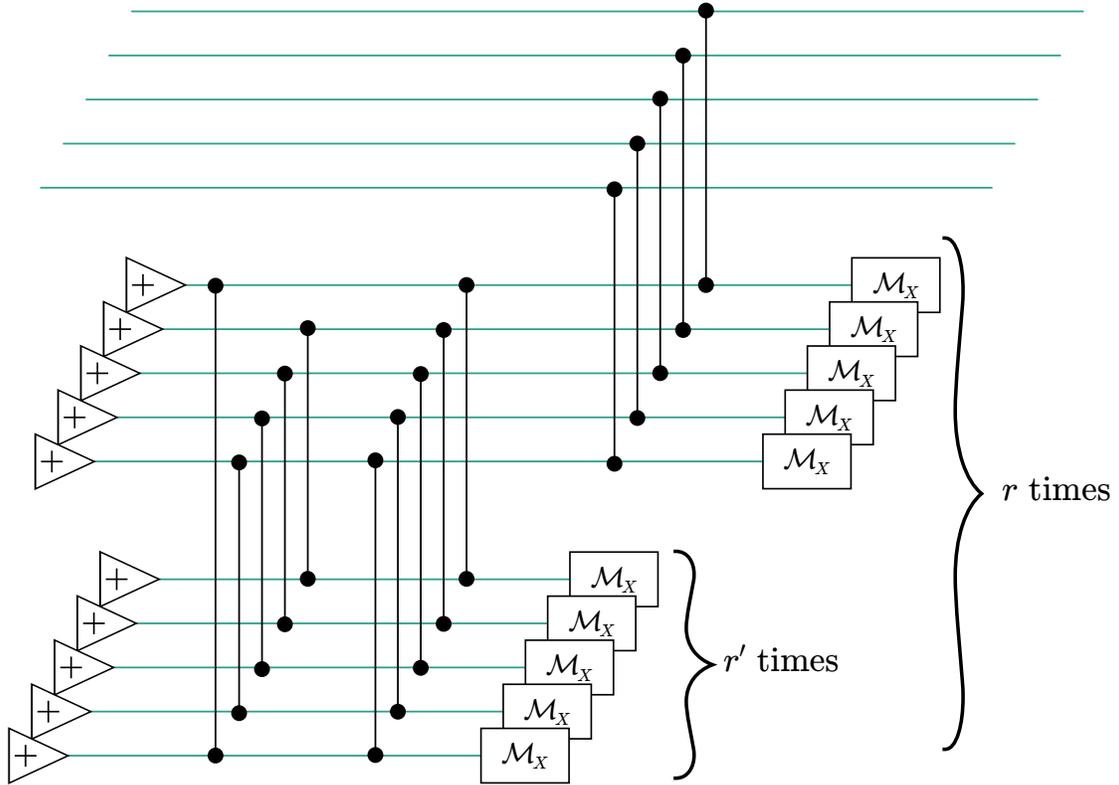


Figure 3.4: Gadget for a single row of the non-destructive  $\mathcal{M}_x^{\frac{1}{2}}$  measurement, including cat state preparation, where the cat state ancilla length  $p$  is equal to the row length  $m$ .

the results is computed. The repetition provides protection against  $Z$  errors in the ancilla, but, of course, there may also be (relatively rare)  $X$  errors acting on the data qubits due to faults during the  $Z^L$  measurement circuit. These faults cause a logical  $X^L$  error only if  $X$  errors occur in a majority of the  $n$  rows. The complete circuit is shown in Figure 3.4.

This procedure for measuring  $Z^L$  can easily be extended to build gadgets that perform the parity measurements  $\mathcal{M}_{ZZ}^L$ ,  $\mathcal{M}_{ZZZ}^L$  on multiple logical qubits.

### 3.5.3 Cat state preparation gadget

The length- $p$  cat state  $|+\text{cat}\rangle = \frac{1}{\sqrt{2}}(|0\rangle^{\otimes p} + |1\rangle^{\otimes p})$  may be characterized as the simultaneous eigenstate with eigenvalue 1 of  $X^{\otimes p}$  and of  $(p-1)$   $ZZ$  operators, each one acting on a pair of neighboring qubits. To prepare the cat state, we start by preparing the product state  $|+\rangle^{\otimes p}$  (an eigenstate of  $X^{\otimes p}$ ), and proceed to measure the  $(p-1)$   $ZZ$  operators. Each of these measurements is executed using yet another ancilla qubit. If every measurement has the outcome  $ZZ = +1$ , then the cat state has been prepared successfully, assuming the measurements are flawless. Otherwise, the measurement outcomes provide a syndrome pointing to  $X$  errors in the cat state: each  $ZZ = -1$  outcome identifies the endpoint of a chain of  $X$  errors. These  $X$  errors in the cat state need not be corrected; instead, we can keep track of their propagation as the computation proceeds. Specifically, the  $X$  errors in the cat state propagate to become  $Z$  errors acting on the measured row of the data block, which should be taken into account when we decode a later  $X^L$  measurement of that block. Following Knill [38], we say that the measured syndrome is used to update the “Pauli frame” of the computation.

To improve robustness, we add one more (redundant)  $ZZ$  measurement, acting on the first and last qubit in the cat state. Then an even number among the  $p$   $ZZ$  measurements should have the outcome  $-1$  if all measurements are correct; if, in fact, we find an odd number of  $-1$  outcomes, then we reject the syndrome. Adding further redundant checks would further improve the reliability of the syndrome extraction, but *this* extra check can be performed without increasing the time we have to wait to do measurements. We conduct  $r'$  rounds of syndrome measurement, including rejected rounds, so that the number of accepted rounds may be fewer than  $r'$ .

In principle, the cat state Pauli frame could be determined by performing a perfect matching on a two-dimensional graph representing the space-time history of the cat state syndrome measurement [58]. Instead, we will consider a much simpler and less effective method for decoding the syndrome

history. Though far from optimal, our procedure has two advantages over the perfect matching algorithm — it requires only very modest (classical) computational resources compared to perfect matching, and it is also much easier to study analytically.

In our unsophisticated scheme, we determine the Pauli frame using whichever syndrome occurs most frequently among the accepted rounds, which we call the “winning” syndrome. Our analysis does not depend on what rule is used when two or more syndromes tie for the distinction of being most frequent, since we pessimistically assume that the preparation gadget fails catastrophically in the event of such a tie. The state preparation protocol is deterministic in the sense that ancillas are never discarded, and the number of rounds of syndrome measurement is always  $r'$  irrespective of the measurement outcomes.

$X$  errors in the ancilla during the preparation can cause the cat state syndrome to evolve as it is measured repeatedly, and therefore the syndrome that occurs most often, even if valid when measured, may differ from the syndrome that is applicable when the cat state interacts with the data. The effect of this syndrome “drift” needs to be included in our error analysis.

### 3.5.4 $|+\rangle^L$ preparation gadget

The logical state  $|+\rangle^L$ , for a particular “choice of gauge,” is a product of  $m$  length- $n$  cat states, one for each column of the block. This state is an eigenstate with eigenvalue 1 of  $X^L = X^{\otimes n}$  acting on any column, and of the gauge operator  $ZZ$  acting on a pair of neighboring qubits in any column (and hence also a +1 eigenstate of all the  $Z$ -type check operators). We prepare these  $m$  length- $n$  cat states in parallel, using the same procedure for preparing length- $p$  cat states described above. In this case, we repeat the syndrome measurement for each column  $r_+$  times.

### 3.5.5 $|0\rangle^L$ preparation gadget.

The preparation of the logical state  $|0\rangle^L$  may be achieved by first preparing the state  $|+\rangle^L$  and performing the non-destructive measurement  $\mathcal{M}_{\frac{L}{2}}$ . If the measurement outcome is  $-1$ , we update the Pauli frame to account for the  $X^L$  difference.

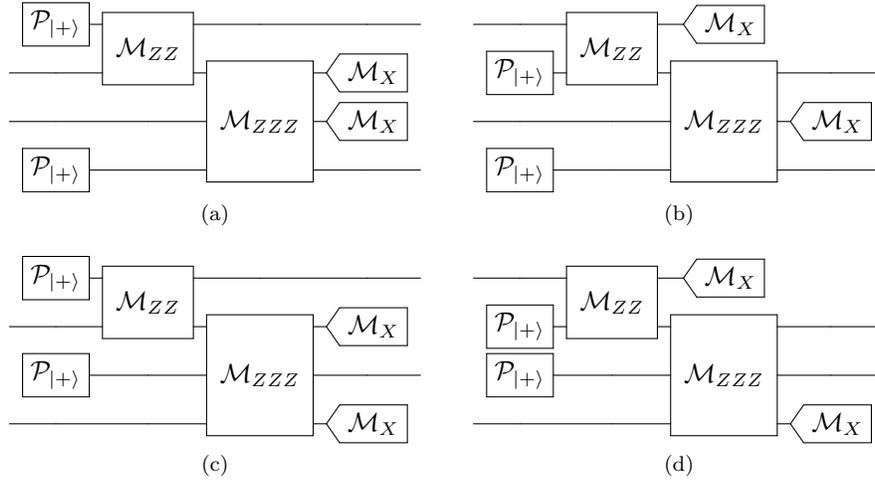


Figure 3.5: Four versions of the CNOT gadget.

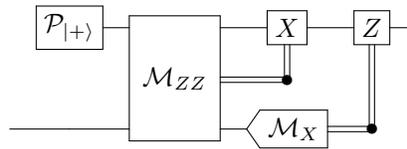


Figure 3.6: One-bit teleportation circuit.

### 3.5.6 CNOT<sup>L</sup> gadget.

The gadget shown in Figure 3.5 applies a fault-tolerant “teleported” CNOT<sup>L</sup> logical operation, up to a known Pauli error determined by the measurement outcomes [48]. Aside from executing the logical gate, the gadget also achieves error correction by measuring both incoming blocks and preparing fresh outgoing blocks.

To better understand how the gadget works, it is helpful to divide it into two parts. One part is the “one-bit teleportation” gadget shown in Figure 3.6, which performs error correction on a single block. This circuit transfers the single-qubit state  $|\psi\rangle$  from qubit 2 to qubit 1. We prepare qubit 1 in the state  $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ , perform the joint measurement  $\mathcal{M}_{Z^1Z^2}$  on the two qubits, and then perform the single-qubit measurement  $\mathcal{M}_{X^2}$ . After applying Pauli operators conditioned on the two measurement outcomes, the result is that the input state  $|\psi\rangle$  of qubit 2 is moved to qubit 1. The conditional  $X$  and  $Z$  gates in this teleportation gadget need not actually be executed; instead

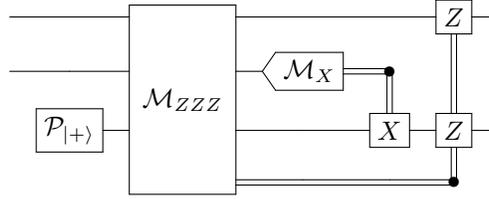


Figure 3.7: Teleported CNOT gate.

the Pauli frame is updated to reflect the measurement outcomes.

The other component of the CNOT gadget is shown in Figure 3.7, where again the Pauli frame updates are explicitly indicated. We may verify that this circuit realizes a CNOT gate by checking that it propagates Pauli operators according to the rule

$$\text{CNOT} : XI \rightarrow XX \quad (3.19)$$

$$IX \rightarrow IX$$

$$ZI \rightarrow ZI$$

$$IZ \rightarrow ZZ.$$

We denote the measurement outcomes of  $\mathcal{M}_X$  and  $\mathcal{M}_{ZZZ}$  as  $(-1)^a$  and  $(-1)^b$  respectively. Since  $|+\rangle$  is an eigenstate of  $X$  with eigenvalue  $+1$ ,  $XI$  acting on the input qubits is equivalent to  $XIX$  acting on the three qubits in the circuit, which commutes with  $\mathcal{M}_{ZZZ}$  and hence becomes  $XX$  acting on the output qubits. Likewise,  $IX$  acting on the input qubits is equivalent to  $IXX$ , which commutes with  $\mathcal{M}_{ZZZ}$  and becomes  $(-1)^a IX$  acting on the output qubits using the measurement outcome  $IXI = (-1)^a$ .  $ZI$  acting on the input qubits commutes with  $\mathcal{M}_{ZZZ}$  and trivially maps to  $ZI$  acting on the output qubits, while  $IZ$  acting on the input qubits becomes  $IZI$ , which maps to  $(-1)^b ZIZ$ , using the measurement outcome  $ZZZ = (-1)^b$ . Hence, we have found

$$\text{CNOT} : XI \rightarrow XX \quad (3.20)$$

$$IX \rightarrow (-1)^a IX,$$

$$ZI \rightarrow ZI$$

$$IZ \rightarrow (-1)^b ZZ,$$

which is corrected to Equation (3.19) by applying  $Z^a$  to both output qubits and  $X^b$  to the second qubit. When we combine the CNOT gate with one-bit teleportation of the first qubit, the measurements  $\mathcal{M}_{ZZ}$  and  $\mathcal{M}_{ZZZ}$  commute and hence can be done in either order.

Four different versions of the CNOT gadget are shown in Figure 3.5. Let us denote the blocks appearing in the figure as blocks 1, 2, 3, 4, counting from the top down. Version (a) of the CNOT gadget contains the preparations and measurements  $\mathcal{P}_{|+\rangle}[1]$ ,  $\mathcal{P}_{|+\rangle}[4]$ ,  $\mathcal{M}_{ZZ}[12]$ ,  $\mathcal{M}_{ZZZ}[234]$ ,  $\mathcal{M}_X[2]$ ,  $\mathcal{M}_X[3]$ . This gadget has some arbitrary features, which are exploited to construct the other versions.

Since  $ZZII \circ ZIZZ = IZZZ$ , we extract the same parity information if  $\mathcal{M}_{ZZZ}[234]$  is replaced by  $\mathcal{M}_{ZZZ}[134]$ ; also, since  $\mathcal{M}_{ZZ}$  acts symmetrically on its two measured blocks, we may permute the order of the blocks, in effect replacing  $\mathcal{P}_{|+\rangle}[1]$  and  $\mathcal{M}_X[2]$  with  $\mathcal{P}_{|+\rangle}[2]$  and  $\mathcal{M}_X[1]$ , thus obtaining version (b). Similarly, since  $\mathcal{M}_{ZZZ}$  acts symmetrically, we may replace  $\mathcal{P}_{|+\rangle}[4]$  and  $\mathcal{M}_X[3]$  with  $\mathcal{P}_{|+\rangle}[3]$  and  $\mathcal{M}_X[4]$ , thus obtaining versions (c) and (d). In all these realizations of the gadget, the upper incoming block is the control block and the lower incoming block is the target block; by replacing  $\mathcal{M}_{ZZ}[12]$  and  $\mathcal{M}_{ZZZ}[234]$  by  $\mathcal{M}_{ZZZ}[123]$  and  $\mathcal{M}_{ZZ}[34]$ , we can reverse the order of the CNOT.

This freedom to choose the incoming control block to be either block 1 or block 2, and to choose the incoming target block to be either block 3 or block 4, will be useful when we consider in Section 3.7 how to realize the CNOT gadget using geometrically local gates. No matter how we configure the gadget, one block participates in both  $\mathcal{M}_{ZZ}$  and  $\mathcal{M}_{ZZZ}$ . However, for versions (a) and (c) in Figure 3.5, the twice-measured block is an incoming (control) data block, while for versions (b) and (d), the twice-measured block is an outgoing (control) data block.

The  $\text{CNOT}^L$  gadget is characterized by four parameters: the length  $p$  of the cat states used in the  $Z$ -type measurements and the number of times measurements are repeated ( $r_+$  for the  $|+\rangle^L$  preparation,  $r'$  for the cat state preparation, and  $r$  for  $ZZ^L$  and  $ZZZ^L$  measurements). In principle, we could choose different values of these parameters in the measurements  $\mathcal{M}_{ZZ}^L$  and  $\mathcal{M}_{ZZZ}^L$ , but in practice we have found it optimal or near optimal to use the same parameters for both.

### 3.5.7 Scheduling

If gate locations are not restricted, and fresh ancilla qubits are always available on demand, then we can schedule the CZ gates so that no data qubit or ancilla qubit is idle during any time step. We assume that a preparation, measurement, or CZ gate can be performed in a single time step.

Suppose, for example, that we use the gadget design in Figure 3.5b, where the outgoing control data block participates in both  $\mathcal{M}_{ZZ}^L$  and  $\mathcal{M}_{ZZZ}^L$ , and let us also assume that the [12] cat states have length  $2m$  while the [234] cat states have length  $3m$ , matching the weight of the measured operators. We label the  $r$  ancilla registers used to measure a row operator  $r$  times by  $a \in \{1, 2, 3, \dots, r\}$  for the  $ZZ^L$  measurement and by  $b \in \{1, 2, 3, \dots, r\}$  for the  $ZZZ^L$  measurement. Suppose that  $\mathcal{P}_{|+\rangle}^L[2]$  is completed in time step  $-1$  and that  $\mathcal{P}_{|+\rangle}^L[3]$  is completed in time step  $0$ . Then the CZ gates are scheduled as follows: The [12] ancilla  $a$  interacts with block 2 in time step  $a - 1$  and with block 1 in time step  $a$ , while the [234] ancilla  $b$  interacts with blocks 2, 3, and 4 in time step  $b$ . Hence, processing of the incoming blocks (1 and 4) begins in time step 1, and processing of the outgoing blocks (2 and 3) ends in time step  $r$ .

The “preparation” of the [234] ancilla  $b$  begins in time step  $b + 1$ , but the “preparation” of the [12] ancilla  $a$  is “staggered.” The syndrome measurement for the first  $m$  qubits in ancilla  $a$  (those that interact with block 1) begins in time step  $a + 1$ , while the syndrome measurement for the last  $m$  qubits (those that interact with block 2) begins one step earlier (step  $a$ ). Despite this one time step delay, no qubits are ever idle — the ancilla qubit used to measure syndrome bit  $Z[m]Z[m + 1]$  in cat state  $a$  interacts with cat state qubit  $m + 1$  in time step  $a$ , and interacts with cat state qubit  $m$  in time step  $a + 1$ . Meanwhile, the ancilla qubit used to measure the syndrome bit  $Z[1]Z[2m]$  in cat state  $a$  interacts with cat state qubit  $2m$  in time step  $a + 1$ , and with cat state qubit 1 in time step  $a + 2$ , *etc.*

If, instead, we use the CNOT gadget design in Figure 3.5a, where the incoming control data block participates in both  $\mathcal{M}_{ZZ}^L$  and  $\mathcal{M}_{ZZZ}^L$ , the processing of the data blocks is not perfectly synchronized. If  $\mathcal{M}_{ZZ}^L$  is performed first, then the incoming and outgoing control blocks are one time step ahead of the incoming and outgoing target blocks, while, if  $\mathcal{M}_{ZZZ}^L$  is performed first, it is the other way around. Thus, if the incoming control block is either one step ahead or one step behind the incoming target block, we can perform the CNOT without leaving any qubits idle, while maintaining the one-step lag between the blocks. But if the two incoming blocks are synchronized, then, using the gadget in Figure 3.5a one block or the other would have to wait for one time step before the processing of the block begins.

### 3.5.8 Injection by teleportation

To complete a universal set of protected gates, we will use the magic state distillation procedure introduced by Bravyi and Kitaev [47]. First, we inject non-Clifford states into the code block using teleportation, and then we improve the fidelity of these encoded states via distillation. The distillation protocol is discussed further in Section 3.10; here we briefly describe the state injection procedure, with further details postponed until Section 3.9.

The state injection makes use of the “one-bit teleportation” circuit depicted in Figure 3.6, but with a physical qubit coming in and a logical qubit coming out. To inject a state into the Bacon-Shor block, first the state  $|\psi\rangle$  is prepared using a single physical qubit, while a logical block is prepared in the state  $|+\rangle^L$ . Then the weight- $(m+1)$  operator  $Z^L Z$  is measured on the single qubit and a single row of the block, and, finally, the single qubit is measured in the  $X$  basis. This procedure prepares the block in the state  $|\psi\rangle^L$ , up to a logical Pauli operator known from the measurement outcomes.

There are other possible injection procedures, but this one is particularly simple. A single fault in the circuit can cause a failure.

## 3.6 Effective error strength

We wish to derive an upper bound on the failure rate of the encoded gadgets of Section 3.5. Because the CNOT gadget is the most complicated and most likely gadget to fail, we can achieve an overall upper bound by bounding its failure rate. The CNOT gadget may produce a logical error if any one of its four measurements deviates from the ideal case. To estimate the failure probability for the CNOT gadget, we must enumerate the ways in which these measurements might fail.

We can separate the modes of failure into two types. In the first type, a series of faults, diagonal and non-diagonal, “directly” causes enough errors to flip the inferred measurement away from the ideal result. The second class consists of a series of failures that cause the inferred Pauli frame measurement of a cat state or  $|+\rangle^L$  state to be “catastrophically” bad, a term that will be defined more precisely in Section 3.6.2. Although the net effect of this second class of failure modes is also to flip the later measurement results, it is convenient to separate them out because a single such failure pattern can immediately cause a measurement to fail.

In estimating the failure probability for the logical measurements  $\mathcal{M}_X^L$ ,  $\mathcal{M}_{ZZ}^L$ , or  $\mathcal{M}_{ZZZ}^L$ , we

need to take into account possible faults in circuit locations preceding the CNOT gadget, which might cause these measurements to deviate from their ideal outcomes. In our fault-tolerant scheme, a logical CNOT gate is either preceded by teleportation into the code block or by another CNOT gate; in the worst case, each input block to the CNOT gadget is an output block from an immediately preceding CNOT gadget. Thus, in the preceding gadget, each block is prepared in the state  $|+\rangle^L$ , and then subjected to either  $\mathcal{M}_{ZZ}^L$  and  $\mathcal{M}_{ZZZ}^L$  (if it is the control block of the CNOT gadget) or just  $\mathcal{M}_{ZZZ}^L$  (if it is the target block of the CNOT gadget).

We will assume that failure of any of its preparation gadgets will cause the CNOT to fail, and we denote by  $P_{\text{err}}^*$  the failure probability for a logical measurement due to a cause other than failure of a preparation gadget. Thus, the failure probability for the CNOT can be bounded as

$$\begin{aligned} P_{\text{err}}(\text{CNOT}^L) &\leq P_{\text{err}}^*(\mathcal{M}_{ZZ}^L) + P_{\text{err}}^*(\mathcal{M}_{ZZZ}^L) + 2P_{\text{err}}^*(\mathcal{M}_X^L) \\ &\quad + 4P_{\text{err}}(\mathcal{P}_{|+\rangle}) + P_{\text{err}}(\mathcal{P}_{|ZZ\text{-cat}\rangle}) + P_{\text{err}}(\mathcal{P}_{|ZZZ\text{-cat}\rangle}) \\ &\quad + 2 \left( P_{\text{err}}(\mathcal{P}_{|ZZ\text{-cat}\rangle}) + P_{\text{err}}(\mathcal{P}_{|ZZZ\text{-cat}\rangle}) \right). \end{aligned} \quad (3.21)$$

The last term accounts for preparation errors that may have occurred in a  $\mathcal{M}_{ZZ}^L$  or  $\mathcal{M}_{ZZZ}^L$  immediately preceding the current CNOT gadget, acting on either one of the two input blocks to the current gadget. In the worst case, both incoming blocks were control blocks in the preceding CNOT gadgets, and were therefore subjected to both  $\mathcal{M}_{ZZ}^L$  and  $\mathcal{M}_{ZZZ}^L$ . The coefficient of  $P_{\text{err}}(\mathcal{P}_{|+\rangle})$  is 4 rather than 2 for a similar reason — the logical measurement outcomes in the current CNOT gadget may differ from ideal outcomes due to errors in the preparation of logical blocks either in the current gadget or in one of the two preceding gadgets.

### 3.6.1 Measurement failure

For an  $\mathcal{M}_X^L$  measurement to fail, more than half of the  $m$  length- $n$  columns must each have at least one  $Z$  error. In the worst case, each qubit in the block participates in  $3r + 2r_+ + 2$  circuit locations: a preparation, a measurement,  $2r_+$  CZ gates contained in the preparation of the logical  $|+\rangle^L$  state,  $r$  CZ gates in the logical  $Z$  measurement in the current gadget, and  $2r$  CZ gates in two logical  $Z$  measurements in the preceding gadget. Each of these locations could be faulty with probability  $\varepsilon + \varepsilon'$ , since either a diagonal or a non-diagonal fault could produce a phase error.

For the moment, let us assume for simplicity that the cat state has length  $2m$  for the  $ZZ^L$

measurement and length  $3m$  for the  $ZZZ^L$  measurement, so that no ancilla qubit interacts with more than one data qubit. (We will consider the case where the cat state is shorter in Section 3.6.3.) Because the preparation of the cat state actually occurs *after* the ancilla interacts with the data, we do *not* need to worry about  $X$  errors arising in the cat state preparation propagating directly to the data. We *do* need to worry that the Pauli frame inferred from the “winning” syndrome after  $r'$  rounds of syndrome measurement actually differs from the ideal Pauli frame. However, assuming that the syndrome is decoded correctly (as it will be if the preparation succeeds), a Pauli-frame  $X$  error afflicts a qubit only if at least one  $X$  error acted on that qubit during the multi-round syndrome measurement. The probability of such a Pauli-frame  $X$  error on any qubit in the cat state is therefore bounded above by  $2r'\varepsilon'$ , since two CZ gates act on each qubit in each round of syndrome measurement, and the syndrome measurement is repeated  $r'$  times in the cat state preparation. Because each row measurement is repeated  $r$  times, the total probability that a row measurement fails due to a Pauli-frame  $X$  error in a cat state is bounded above by  $2rr'\varepsilon'$ . Furthermore, in the worst case there are three logical  $Z^L$  measurements between the preparation and measurement of the block, in each of which a cat state interacts with the data.

Combining together the probability per qubit of a  $Z$  error caused directly by a fault with the probability of a  $Z$  error that propagates from the cat state, we conclude that the error probabilities for the  $X^L$  measurements can be bounded as

$$P_{\text{err}}^*(\mathcal{M}_X^L) \leq \binom{m}{\frac{m+1}{2}} [n(2r_+ + 3r + 2)(\varepsilon + \varepsilon') + 6nrr'\varepsilon']^{(m+1)/2}. \quad (3.22)$$

As explained in Section 3.5.7, there are no time steps in which “resting” qubits are subject to “storage errors.”

For the  $\mathcal{M}_{ZZ}^L$  measurement to fail, more than half of the  $n$  weight- $2m$  row parity measurements must have errors. An  $X$  error acting on any one of the  $2m$  qubits in the same row of the two blocks, either before or during the measurement, could flip the parity of the row. Before the  $\mathcal{M}_{ZZ}^L$  is completed, each qubit participates in  $2r_+$  CZ gates during the  $\mathcal{P}_{|+\rangle}$  operation,  $r$  CZ gates during the  $\mathcal{M}_{ZZ}^L$ , and, in the worst case, another  $2r$  CZ gates during the  $\mathcal{M}_{ZZZ}^L$  and  $\mathcal{M}_{ZZ}^L$  in the preceding CNOT gadget. Thus, the probability per qubit of a non-diagonal error is  $(2r_+ + 3r)\varepsilon'$ . There might also be an  $X$ -type Pauli frame error in the  $|+\rangle^L$  preparation, but we do not have to count that separately; if the preparation “succeeds,” then a Pauli-frame  $X$  error afflicts a qubit in the block

only if a non-diagonal fault damaged that same qubit during the preparation circuit.

In the absence of non-diagonal faults, the row parity measurement might still fail because at least one  $Z$  error acts on the cat state in each of  $(r+1)/2$  of the repeated measurements. If the cat state has length  $p$ , then in each repetition of the measurement there are  $2m + 2p + 2r'p$  circuit locations ( $2m$  CZ gates coupling the ancilla to the data, plus  $p$  qubit preparations,  $p$  qubit measurements and  $2pr'$  CZ gates for the cat state preparation). We conclude that the failure probability can be bounded as

$$P_{\text{err}}^*(\mathcal{M}_{ZZ}^L) \leq \binom{n}{\frac{n+1}{2}} \left[ 2m(2r_+ + 3r)\varepsilon' + \binom{r}{\frac{r+1}{2}} [(2m + 2p + 2pr')(\varepsilon + \varepsilon')]^{(r+1)/2} \right]^{(n+1)/2}. \quad (3.23)$$

Similarly,

$$P_{\text{err}}^*(\mathcal{M}_{ZZZ}^L) \leq \binom{n}{\frac{n+1}{2}} \left[ 3m(2r_+ + 4r)\varepsilon' + \binom{r}{\frac{r+1}{2}} [(3m + 2p + 2pr')(\varepsilon + \varepsilon')]^{(r+1)/2} \right]^{(n+1)/2}. \quad (3.24)$$

(The  $\mathcal{M}_{ZZZ}^L$  is preceded by a  $\mathcal{M}_{ZZ}^L$  acting on one of its input blocks inside the current CNOT gadget, plus additional  $Z$ -type logical measurements in the preceding CNOT gadgets acting on another one of its input blocks.)

### 3.6.2 Cat state preparation failure

Now it remains to estimate the probability of failure for the cat state and the logical state  $|+\rangle^L$ . Under our simplified procedure for choosing the “winning” syndrome, we first reject any syndrome that does not have a consistent interpretation— that is, an even number of nontrivial syndrome bits. Recall that the syndromes ideally detect the boundaries of error chains, so syndromes with an odd number of bits are inconsistent with this interpretation. Out of the remaining consistent syndromes, we choose the syndrome that occurs most often to be the “winning” syndrome.

If the winning syndrome leads us to exactly identify the correct Pauli frame, then we have succeeded. But we could instead identify some other syndrome which is slightly different than the correct Pauli frame. If the number of locations  $l$  where the inferred Pauli frame differs from the correct Pauli frame is smaller than the number of non-diagonal faults that occurred during the preparation, we can reinterpret this as being the same as a perfect syndrome decoding followed by non-diagonal faults which add the errors. These kinds of failures are already accounted for in

Section 3.6.1; we call them “valid” syndromes. We are therefore interested in the probability that the winning syndrome is “catastrophically” different, such that the previous condition fails.

Suppose that the winning syndrome occurs  $t$  times due to faults, and that there are  $u$  additional rounds that each contain at least one fault. Some of these additional rounds might be rejected, and some may not, but we can assume that they are all rejected because, by definition, they do not produce the winning syndrome. Suppose that non-diagonal faults occur in  $s$  of the  $r'$  rounds of syndrome measurement; these faults can alter the syndrome. There are  $r' - t - u$  rounds without any faults, and the number of distinct syndromes detected in these rounds is at most  $s + 1$ .

Now we can use the pigeonhole principle to obtain a lower bound on  $t$ , expressed in terms of  $r'$ ,  $u$ , and  $s$ . There are at most  $s + 1$  “valid” syndromes that can occur in syndrome measurement rounds that have no non-diagonal faults. Combining these with the winning syndrome, there are at most  $s + 2$  possible syndromes that can occur in the  $r' - u$  accepted rounds. Of these  $s + 2$  syndromes, the winning syndrome must occur at least as many times as any other syndrome; hence,

$$t \geq \left\lceil \frac{r' - u}{s + 2} \right\rceil, \quad (3.25)$$

where  $\lceil x \rceil$  denotes the smallest integer greater than or equal to  $x$ .

To bound  $P_{\text{err}}(\mathcal{P}_{|\text{cat}\rangle})$ , we sum over  $s$  and  $u$  in each cat state preparation step, estimating the number of possible fault histories using the upper bound Equation (3.25) on  $t$ . In the first of the  $t$  winning rounds, a particular winning syndrome is found, which differs in at least two bits from the actual syndrome in the beginning of that round. Then this same syndrome is found again in each of the remaining winning rounds. The sum over all possible winning syndromes, weighted by their probabilities, is bounded above by the probability that at least two measured syndrome bits are faulty in the first of the  $t$  winning rounds. Each  $ZZ$  measurement is performed using one  $|+\rangle$  preparation, two  $CZ$  gates, and one  $X$  measurement; therefore, the probability of error in the measurement of a single syndrome bit is bounded above by  $4\varepsilon + 2\varepsilon'$ , and the probability that at least two syndrome bits are in error is bounded above by  $\binom{p}{2}(4\varepsilon + 2\varepsilon')^2$ .

In each of the  $s$  rounds that contain  $X$  errors, we must sum over all the possible  $X$ -error patterns that can occur in that round. The sum over all  $X$ -error patterns, weighted by the probabilities, is bounded above by the probability that at least one nondiagonal fault occurs in that round. Since the round contains  $2p$   $CZ$  gates, this probability is in turn bounded above by  $2p\varepsilon'$ .

Once the  $X$ -error pattern has been chosen in each round that contains  $X$  errors, we know the actual syndrome at the beginning of each round. And once the winning syndrome is chosen in the first winning round, we know which syndrome bits must have errors in each of the remaining  $t - 1$  winning rounds. If the cat state preparation fails, then, by definition, at least two syndrome bits have errors in each of these rounds; hence, each winning round after the first has a probability weight bounded above by  $(4\varepsilon + 2\varepsilon')^2$ .

Taking into account that the rounds with non-diagonal faults can be chosen in at most  $\binom{r'}{s}$  ways, and enumerating the ways to choose which  $t$  rounds produce the winning syndrome and which  $u$  additional rounds have faults, we obtain

$$P_{\text{err}}(\mathcal{P}_{|\text{cat}\rangle}) \leq nr \sum_{s=0}^{r'} \sum_{u=0}^{r'} \binom{r'}{s} \binom{r'}{u+t} \binom{u+t}{u} \binom{p}{2} (4\varepsilon + 2\varepsilon')^{2t} (4p\varepsilon + 2p\varepsilon')^u (2p\varepsilon')^s, \quad (3.26)$$

and, similarly,

$$P_{\text{err}}(\mathcal{P}_{|+\rangle}) \leq m \sum_{s=0}^{r_+} \sum_{u=0}^{r_+} \binom{r_+}{s} \binom{r_+}{u+t} \binom{u+t}{u} \binom{n}{2} (4\varepsilon + 2\varepsilon')^{2t} (4n\varepsilon + 2n\varepsilon')^u (2n\varepsilon')^s, \quad (3.27)$$

where  $t$  denotes  $\lceil \frac{r'-u}{s+2} \rceil$ . The prefactor  $nr$  in Equation (3.26) arises because we use  $n$  length- $p$  cat states in each measurement, and each measurement is repeated  $r$  times. The prefactor  $m$  in Equation (3.27) arises because the encoded state  $|+\rangle^L$  is a product of  $m$  length- $n$  cat states.

We should add another contribution to the failure probability, because if half of the qubits (or more) have  $X$  errors, we might decode the cat-state syndrome incorrectly. A syndrome of the repetition code points to two possible  $X$ -error patterns, one low weight and one high weight. We always assume the low-weight interpretation is correct, so if the high-weight interpretation is actually correct, then a Pauli-frame error  $X^{\otimes p}$  is applied to the cat state (or worse if the cat state qubits are used multiple times). The additional contribution, then, is bounded by the probability that each of  $\lceil p/2 \rceil$  qubits in the cat state are each hit by  $X$  errors at least once in at least one of the cat state preparations. This upper bound is

$$P_{\text{high weight}} \leq nr \binom{p}{\lceil p/2 \rceil} (2r'\varepsilon')^{\lceil p/2 \rceil}. \quad (3.28)$$

### 3.6.3 Shorter ancillas

Cat states are used in measurements of the weight- $2m$  logical operator  $ZZ^L$  and the weight- $3m$  logical operator  $ZZZ^L$ . In the derivation of Equation (3.22) we assumed that the cat states have length  $2m$  and  $3m$  respectively, so that each of the cat state's qubits participates in just one CZ gate during the measurement. But we have found that the CNOT gadget may be more reliable if a shorter cat state is used instead.

If the length of the cat state is at least  $m$ , then each cat-state qubit interacts with only one qubit in each encoded block. But if the length is  $p < m$ , then a single cat-state qubit may interact with as many as

$$R = \lceil m/p \rceil \tag{3.29}$$

data qubits in a single block, and hence an  $X$  error on that cat-state qubit may propagate to produce  $Z$  errors in  $R$  distinct columns of one block. If  $k$  of the cat state's  $p$  qubits have  $X$  errors, these could generate  $Z$  errors in  $kR$  columns of the block; thus, an  $X^L$  measurement on the block could fail if there are

$$L(k) = (m + 1)/2 - Rk \tag{3.30}$$

additional columns that each contain at least one  $Z$  error.

In each measurement, each data block interacts with the ancilla for  $R$  consecutive time steps, where, in each step, at most  $p$  CZ gates act on the data block. The repeated measurements can be staggered as described in [48], to avoid time steps in which data qubits are idle.

We use the index  $i \in \{1, 2, 3, \dots, p\}$  to label the positions of the  $p$  qubits in the cat state, and design the measurement circuits so that the cat-state qubit  $i$  interacts with the same set of data qubits in each one of the measured code blocks. Since the  $ZZ^L$  measurement is performed on all  $n$  rows of each block, and also repeated  $r$  times, inside a  $\mathcal{M}_{ZZ}^L$  there are at most  $2nrR$  CZ gates that act on a cat-state qubit at position  $i$  (not counting CZ gates in the cat state preparation step), any of which could have a non-diagonal fault. Similarly, inside a  $\mathcal{M}_{ZZZ}^L$  there are at most  $3nrR$  CZ gates that act on a cat-state qubit at position  $i$ .

The  $\mathcal{M}_X^L$  inside a CNOT gadget is preceded by  $\mathcal{M}_{ZZZ}^L$  or  $\mathcal{M}_{ZZ}^L$  within the CNOT, and, in the worst case, by both  $\mathcal{M}_{ZZZ}^L$  and  $\mathcal{M}_{ZZ}^L$  in the preceding gadget, so, at most, there are  $8nrR$  CZ gates in these measurements where a non-diagonal fault could disturb cat-state qubit  $i$ .

An  $X$  error acting on cat-state qubit  $i$  might also occur because of an error in the winning

syndrome in one of the cat-state preparation steps, which can occur only if a faulty  $CZ$  gate acts on qubit  $i$  at least once during the  $r'$  repeated syndrome measurements. Noting that a damaged cat-state qubit in position  $i$  could cause  $Z$  errors in  $R$  columns of the encoded block, and that  $k$  damaged qubits in the length- $p$  cat state can be chosen in  $\binom{p}{k}$  ways, we conclude that if length- $p$  cat states are used, then Equation (3.22) should be replaced by

$$P_{\text{err}}^*(\mathcal{M}_X^L) \leq \sum_{k=0}^{k_{\text{max}}} \binom{p}{k} [8nrR\varepsilon' + 6nr'r'\varepsilon']^k \binom{m}{L(k)} [n(2r_+ + 3r + 2)(\varepsilon + \varepsilon')]^{L(k)}, \quad (3.31)$$

where

$$k_{\text{max}} = \lceil (m + 1)/2R \rceil. \quad (3.32)$$

### 3.7 Geometrically local circuits

In our analysis so far, we have assumed that the two-qubit  $CZ$  gates can be performed on any pair of qubits, with an error rate independent of the distance between the qubits. Now we will consider how the analysis is modified if  $CZ$  gates can be performed only on neighboring pairs of qubits.

There are a variety of possible architectures for fault-tolerant quantum computing using Bacon-Shor codes. To be concrete, we will consider an effectively one-dimensional arrangement, in which logical qubits are encoded in a ribbon of physical qubits with constant width in the vertical direction, and length in the horizontal direction proportional to the total number of logical qubits. Each logical qubit lives in a “bi-block” — a pair of Bacon-Shor code blocks, one storing the data processed by the computation, and the other used as an auxiliary block for teleporting gates. If desired, the data block can be shuffled back and forth between the left and right sides of the bi-block via a “one-bit-teleportation” circuit consisting of a  $|+\rangle^L$  preparation, a  $ZZ^L$  measurement, and an  $X^L$  measurement, as in Figure 3.6.

Each  $n \times m$  logical block is interlaced with an  $n \times (2m - 1)$  array of ancilla qubits, as in Figure 3.8. Of the  $2m - 1$  ancilla qubits in a row,  $m - 1$  are used to read out the outcome of a  $ZZ$  measurement performed on the qubit’s two horizontal neighbors, preparing an  $m$ -qubit cat state in the row. In addition, there is a column of  $n$  ancilla qubits at the boundary between two adjacent blocks, allowing the cat states to be extended to length  $2m$  or  $3m$  as needed for the measurements  $\mathcal{M}_{ZZ}^L$

and  $\mathcal{M}_{ZZZ}^L$  in the CNOT gadget. In the  $|+\rangle^L$  preparation gadget,  $n - 1$  ancilla qubits in a column are used to read out the outcome of a  $ZZ$  measurement on the qubit's two vertical neighbors, preparing an  $n$ -qubit cat state in a column.

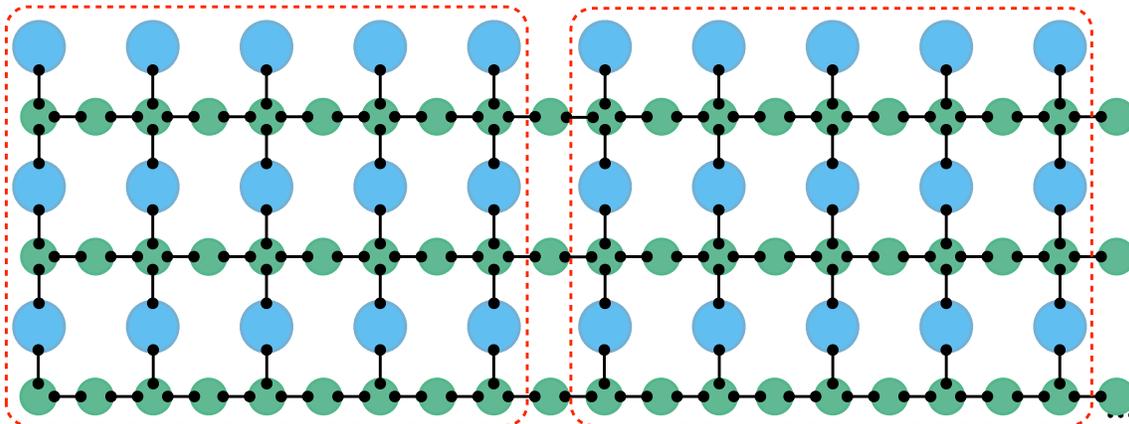


Figure 3.8: Arrangement of Bacon-Shor qubits for geometrically local computation for  $3 \times 5$  code block. Large blue circles indicate data qubits, small green circles ancilla qubits. Solid lines indicate locations between which CZ gates may be performed. Dotted lines separate qubits into code blocks. The ribbon may be extended arbitrarily far to the right and left with additional code blocks.

A CNOT gate is performed on a neighboring pair of logical bi-blocks using one of the variants of the CNOT gadget depicted in Figure 3.5. Let us denote the bi-block on the left as bi-block 1 and the bi-block on the right as bi-block 2. If the data in bi-block 1 is on the right side of the bi-block and the data in bi-block 2 is on the left side, then we may use the CNOT gadget design in Figure 3.5a to perform a CNOT gate in either direction (*i.e.*, with either logical qubit as the control and the other as the target) using cat states of length  $2m$  and  $3m$ , with the output data from the gate appearing on the left side of bi-block 1 and the right side of bi-block 2. Otherwise, we may use one of the alternative gadget designs discussed in Figure 3.5 to deal with the cases where the data in bi-block 1 is on the left side and/or the data in bi-block 2 is on the right side. Whichever design we choose, the CNOT gadget flips the data from one side to the other in both bi-blocks, and only the  $\mathcal{M}_{ZZZ}^L$  measurement reaches across the boundary between the two bi-blocks.

Note that with this method for performing a logical CNOT gate using geometrically local gates there is no need to swap the positions of pairs of physical qubits. The CZ gate is the only two-qubit gate used at the physical level, so that the assumption of highly biased noise remains physically

plausible. Bacon-Shor blocks can be swapped using circuits of logical CNOT gates.

To complete a universal gate set we also need to be able to inject noisy non-Clifford states into code blocks, which can then be purified by state distillation. This state injection can be performed using one-bit teleportation — the state  $|\psi\rangle$  is prepared using a single physical qubit, while a logical block is prepared in the state  $|+\rangle^L$ ; next, the weight- $(m+1)$  operator  $Z^L Z$  is measured on the single qubit and a single row of the block; finally, the single qubit is measured in the  $X$  basis. This procedure prepares the block in the state  $|\psi\rangle^L$ , up to a logical Pauli operator known from the measurement outcomes. To perform this task in a bi-block, we prepare  $|+\rangle^L$  on, say, the left side of the bi-block, and build a length- $(m+1)$  cat state using  $m$  of the ancilla qubits that accompany a row of the block on the left side, and one additional ancilla qubit on the right side. The single-qubit state  $|\psi\rangle$  is prepared using one physical qubit on the right side of the bi-block, which is adjacent to the  $(m+1)$ st qubit of the cat state, allowing the measurement of  $Z^L Z$  to be complete using local CZ gates.

If we wish to build a two-dimensional architecture, we may stack horizontal ribbons of logical qubits on top of one another. To perform CNOT gates on pairs of logical bi-blocks that are stacked vertically, we need to be able to perform  $\mathcal{M}_{ZZZ}^L$  using cat states shared by vertically stacked blocks. For this purpose we can use ancilla qubits in neighboring bi-blocks to establish a “channel,” linking the cat states neighboring corresponding rows of the vertically stacked blocks.

We note that in the case of unbiased noise, a two-dimensional architecture might be realized more simply, using symmetric Bacon-Shor codes. In that case, we can build a CNOT gate using either the gadget in Figure 3.5 or its Hadamard dual, with  $X$  and  $Z$  interchanged. The primal gadget uses horizontal cat states to execute CNOT gates on horizontally neighboring blocks, while the dual gadget uses vertical (dual) cat states to execute CNOT gates on vertically neighboring blocks.

For the geometrically local case, our analysis needs to be modified in several ways. As already indicated, we will not be able to choose the length  $p$  of the cat state to optimize the gadget. Rather, the locality constraint requires that the length of the cat state match the weight of the measured operator ( $p = 2m$  for  $\mathcal{M}_{ZZ}^L$  and  $p = 3m$  for  $\mathcal{M}_{ZZZ}^L$ ). Another change is that we must now consider the consequences of storage errors acting on idle qubits. Previously, we assumed that the off-line preparation of cat states can be scheduled so that the cat states are always available immediately when needed. But now there is only one set of  $m$  ancilla qubits to accompany each row of a data

block. Each time a cat state is measured, we must prepare a new one, and the preparation involves a syndrome measurement repeated  $r'$  times. In each round of syndrome measurement, all qubits except the ones at the ends of the cat state participate in two  $ZZ$  measurements, and because the same ancilla qubits are used for two different  $ZZ$  measurements, four time steps are required for each one (a preparation, two CZ gates, and an  $X$  measurement). Therefore, the data waits for  $8r'$  time steps before the cat state is ready to interact with the data. While the data qubits wait for the cat state to be ready, they may be subject to storage errors. We must estimate how these storage errors contribute to the failure probability of the logical measurements in the CNOT gadget.

We will denote by  $\varepsilon_s$  the probability per time step of a diagonal storage error and by  $\varepsilon'_s$  the probability per time step of a non-diagonal storage error. How many storage steps are included depends on the gadget design, which, in turn, depends on whether the data is on the left or right side of the incoming bi-blocks. In the worst case, the measurement  $\mathcal{M}_X^L$  is performed on a block that undergoes four logical  $Z$ -type measurements between its preparation and measurement—two in the current gadget plus another two in the immediately preceding gadget, and for each of these measurements the data waits for  $8r'$  steps as the cat state is prepared. If each  $Z$ -type logical measurement is repeated  $r$  times, then the total contribution to the  $Z$  error probability per qubit due to the identity gates in the circuit is  $32rr'(\varepsilon_s + \varepsilon'_s)$ . Therefore, Equation (3.22) is modified to become

$$P_{\text{err}}^*(\mathcal{M}_X^L) \leq \binom{m}{\frac{m+1}{2}} [n(2r_+ + 3r + 2)(\varepsilon + \varepsilon') + 32nrr'(\varepsilon_s + \varepsilon'_s) + 8nrr'\varepsilon']^{(m+1)/2}. \quad (3.33)$$

Furthermore, a non-diagonal storage fault on any of the  $2m$  data qubits can flip the outcome of  $\mathcal{M}_{ZZ}^L$ . This fault can occur in any of the  $8r'$  time steps while a cat state is prepared in either the  $\mathcal{M}_{ZZ}^L$  or one of the preceding measurements. Similarly, a non-diagonal storage fault on any of the  $3m$  data qubits can flip the outcome of  $\mathcal{M}_{ZZZ}^L$ , where this fault can occur in any of the  $8r'$  time steps while the cat state is prepared in either the  $\mathcal{M}_{ZZZ}^L$  or one of the preceding measurements. Thus, Equations (3.23) and (3.24) are replaced by

$$P_{\text{err}}^*(\mathcal{M}_{ZZ}^L) \leq \binom{n}{\frac{n+1}{2}} \left[ 2m(2r_+ + 3r)\varepsilon' + 2m(24rr'\varepsilon'_s) + \binom{r}{\frac{r+1}{2}} [(2m + 2p + 2pr')(\varepsilon + \varepsilon')]^{(r+1)/2} \right]^{(n+1)/2}, \quad (3.34)$$

$$P_{\text{err}}^*(\mathcal{M}_{ZZZ}^L) \leq \binom{n}{\frac{n+1}{2}} \left[ 3m(2r_+ + 4r)\varepsilon' + 3m(32rr'\varepsilon'_s) + \binom{r}{\frac{r+1}{2}} [(3m + 2p + 2pr')(\varepsilon + \varepsilon')]^{(r+1)/2} \right]^{(n+1)/2}. \quad (3.35)$$

Note that the number of storage locations in the circuit might be reduced by adding additional ancilla sites, so that more measurements can be performed in parallel, and/or by combining together measurement locations with immediately following preparation locations.

Geometric locality also requires that we remove the redundant cat state measurements that “wrap around” the code block. Thus, there is no syndrome parity check: any syndrome outcome is accepted, and a single diagonal fault can generate a syndrome error. There are also storage locations where the cat state qubits wait while ancilla qubits are prepared and measured. Because of these changes, Equations (3.26) and (3.27) are replaced by

$$P_{\text{err}}(\mathcal{P}_{|\text{cat}\rangle}) \leq nr \sum_{s=0}^{r'} \sum_{u=0}^{r'} \binom{r'}{s} \binom{r'}{u+t} \binom{u+t}{u} [4p\varepsilon + 2p\varepsilon' + 4p(\varepsilon_s + \varepsilon'_s)]^{t+u} (2p\varepsilon' + 4p\varepsilon'_s)^s \quad (3.36)$$

and

$$P_{\text{err}}(\mathcal{P}_{|+\rangle}) \leq m \sum_{s=0}^{r_+} \sum_{u=0}^{r_+} \binom{r_+}{s} \binom{r_+}{u+t} \binom{u+t}{u} [4n\varepsilon + 2n\varepsilon' + 4n(\varepsilon_s + \varepsilon'_s)]^{t+u} (2n\varepsilon' + 4n\varepsilon'_s)^s. \quad (3.37)$$

## 3.8 Results for logical CSS gates

### 3.8.1 Unrestricted gates

With our analytic upper bound on the effective error strength  $\varepsilon^{(1)}$  of the CNOT gadget, we can find the optimum choice of the code block size  $n \times m$ , as well as the four CNOT parameters  $r$ ,  $r'$ ,  $r_+$  and  $p$ , for any choice of the error strengths  $\varepsilon$  and  $\varepsilon'$ . We did this by brute force search over the parameter space. Figure 3.9 plots the results of this optimization for five choices of the bias  $\beta = \varepsilon/\varepsilon'$ .

These results were obtained by optimizing the size of a Bacon-Shor code block which is used by itself, rather than as part of a more complex concatenated code. Alternatively, we may consider using our Bacon-Shor gadgets at the bottom level of a concatenated coding scheme. In particular, we may estimate the accuracy threshold for biased noise achieved by such concatenated codes, as was done in Ref. [48] for the special case of the  $n = 1$  Bacon-Shor code (*i.e.*, the repetition code). We have extended the analysis of the accuracy threshold performed in [48] to more general Bacon-Shor codes. However, for bias above  $10^3$  we found no improvement over the threshold estimate

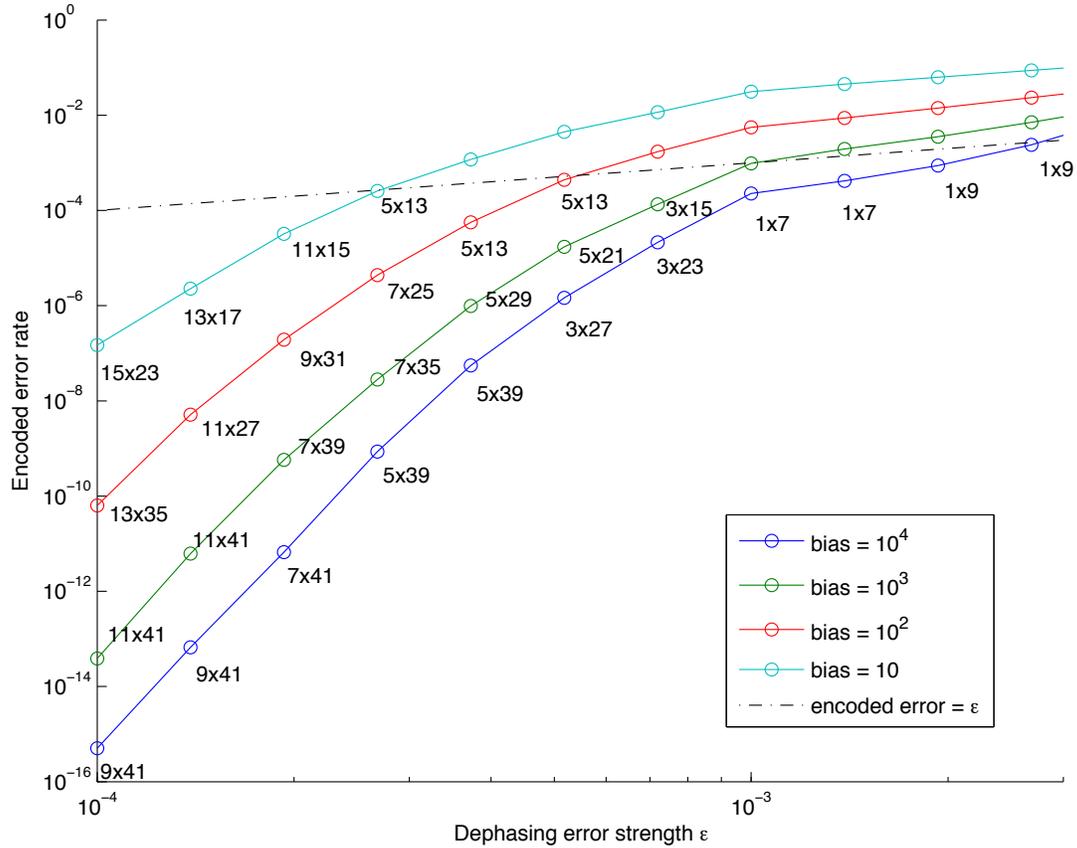


Figure 3.9: Optimal CNOT logical error rate versus physical error rate for various values of the bias. Labels represent the  $n \times m$  dimensions of the Bacon-Shor code block. For blocks with a single row, one can use the code studied in [48] which is a special case of the asymmetric Bacon-Shor code.

in [48], because the  $n = 1$  code turns out to provide the best threshold value. For values of the bias between 1 and  $10^3$ , the optimal value of  $n$  turns out to be greater than 1, and for that range of the bias we found modest improvements in the accuracy threshold estimate compared to [48].

Using asymmetric Bacon-Shor codes, we can exploit the noise bias to improve the number of physical gates needed to construct a fault-tolerant CNOT gate with a specified logical error rate. Figure 3.10 shows this overhead factor and also indicates the dimensions of the optimal code block. This plot includes a comparison with the performance of concatenated codes surveyed in [57] for the case of unbiased noise. For highly biased noise, asymmetric Bacon-Shor codes achieve a much lower logical error rate with a smaller number of physical two-qubit gates, compared to the performance of these previous constructions for unbiased noise.

### 3.8.2 Geometrically local gates

Optimizing the parameters for the case of geometrically local gates, assuming the storage error rate is negligible, we obtain the results displayed in Figures 3.11 and 3.12. Enforcing locality significantly weakens the performance of our constructions — at a bias of  $10^4$  and dephasing error strength of  $\varepsilon = 10^{-4}$ , the optimal effective error strength we can achieve increases from  $10^{-20}$  to  $10^{-9}$ , while requiring roughly 8 times as many gates. Therefore, in the geometrically local case our estimated logical error rate for asymmetric Bacon-Shor codes and highly biased noise is roughly similar to what can be achieved using surface codes with a similar number of logical gates, while disregarding the bias [59, 60]. The most important reason that nonlocal gates prove to be advantageous is that, if geometry is ignored, then the cat state used in (for example) the measurement  $\mathcal{M}_{ZZZ}^L$  can be chosen to be much shorter than the length- $3m$  cat state used in our geometrically local construction. The number of physical qubits per block required to achieve the optimal logical error rate stays relatively small, as indicated in Figure 3.13.

Our upper bound on the logical error rate is likely to be far from optimal, and we also expect that the performance could be substantially improved by using a more sophisticated (but harder to analyze) method like perfect matching for decoding the cat state syndrome history. Suitable lattice deformations may also make it more feasible to reduce the length of cat states substantially, yielding further enhancements in performance for the case of geometrically local gates. Such improvements might make asymmetric Bacon-Shor codes more competitive relative to surface codes.

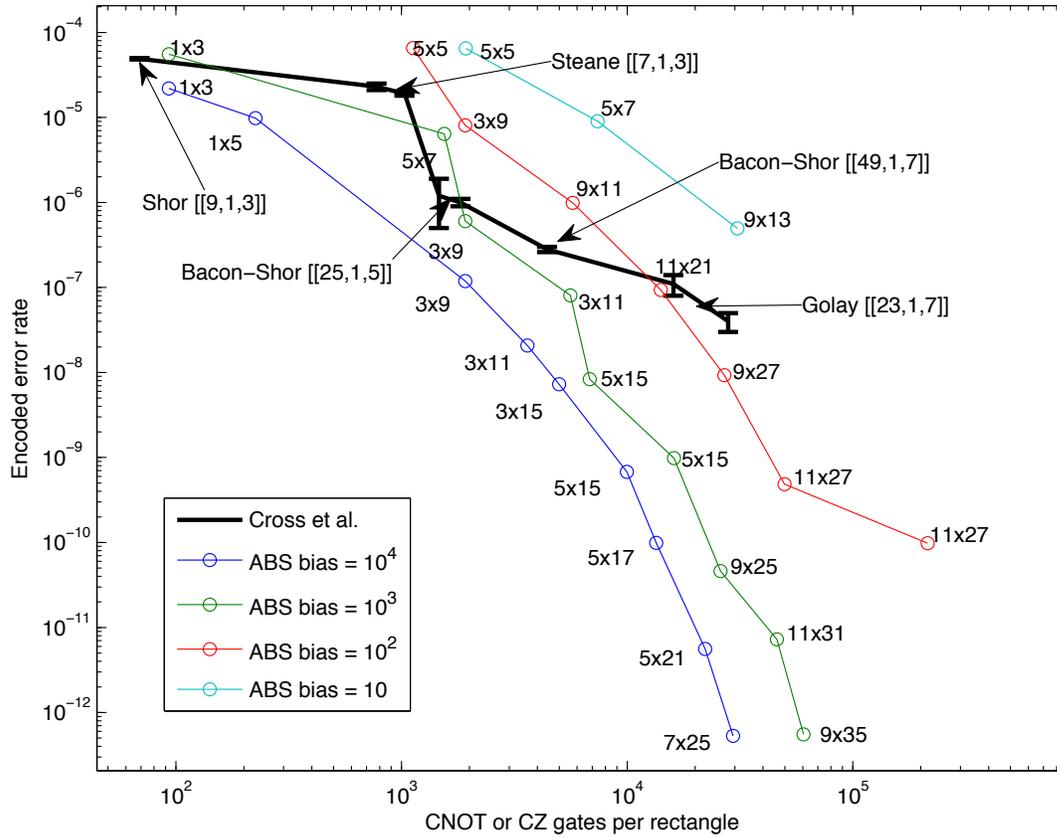


Figure 3.10: Optimal CNOT logical error rate versus required number of physical gates for  $n \times m$  asymmetric Bacon-Shor codes, for a physical error rate  $\varepsilon = 10^{-4}$  at various values of the bias, plotted alongside results for codes studied in [57] with no bias.

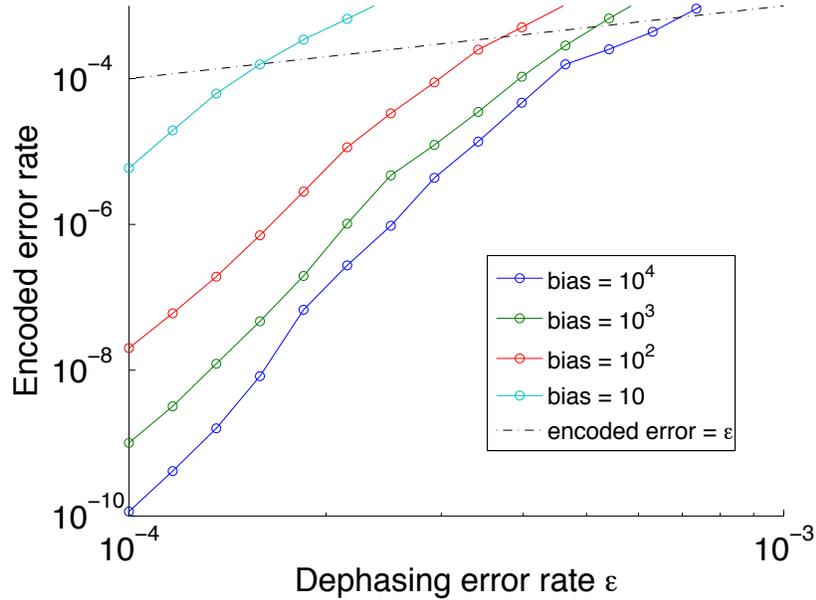


Figure 3.11: Optimal CNOT logical error rate versus physical error rate for various values of the bias, where two-qubit physical gates are assumed to be geometrically local.

### 3.8.3 Geometrically local gates and measurement bias

In some experimental settings, measurements are noisier than gates, in which case we say the noise has “measurement bias.” We may consider a noise model in which the dephasing error rate  $\epsilon$  in diagonal gates exceeds the rate  $\epsilon'$  for non-diagonal faults, while, in addition, the error rate  $\epsilon_{\mathcal{M}}$  for single-qubit measurements exceeds  $\epsilon$ . (We also continue to assume that  $\epsilon$  is the error rate for single-qubit preparations.) As shown in Figure 3.14, our geometrically local asymmetric Bacon-Shor code gadgets are somewhat robust against increasing measurement bias, because our gadgets contain considerably more gates than measurements (or preparations). Relative to the case  $\epsilon_{\mathcal{M}} = \epsilon$ , the performance of the logical CNOT gate is not much affected as the measurement bias  $\epsilon_{\mathcal{M}}/\epsilon$  rises to about 5. In contrast, surface-code gadgets, which contain a higher number of measurements relative to the number of gates, are more sensitive to measurement bias.

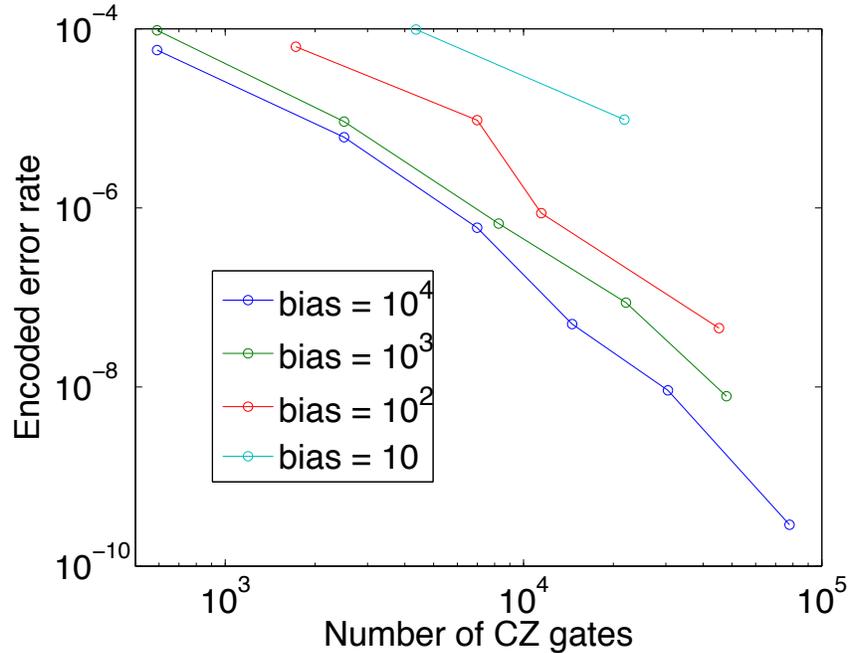


Figure 3.12: Optimal CNOT logical error rate versus required number of physical gates for geometrically local asymmetric Bacon-Shor codes at various values of the bias, for a physical error rate  $\varepsilon = 10^{-4}$ .

### 3.9 State injection

Having analyzed the performance of our fault-tolerant CSS gates, we now turn to the state injection and distillation protocols needed to complete a universal set of fault-tolerant gates. An arbitrary single-qubit state  $|\psi\rangle$  can be injected into the Bacon-Shor block using the “one-bit teleportation” circuit depicted in Figure 3.6. Here, a logical qubit is prepared in the state  $|+\rangle^L$ , the two-qubit measurement  $\mathcal{M}_{ZZ}^L$  is performed jointly on an unprotected qubit and the Bacon-Shor block using a length- $(m+1)$  cat state, and, finally, the  $X$ -basis measurement  $\mathcal{M}_X$  is performed on the unprotected qubit. To estimate the error in the state injection step, we should consider all the ways in which the outcomes of the measurements  $\mathcal{M}_{ZZ}^L$  and  $\mathcal{M}_X$  might deviate from their ideal values.

In the state distillation circuit, this injection step is directly followed by a CNOT gate. Hence, some sources of error in the teleportation circuit need not be attributed to the state injection step, as they are already included in our error estimate for the CNOT gate that follows. In particular,

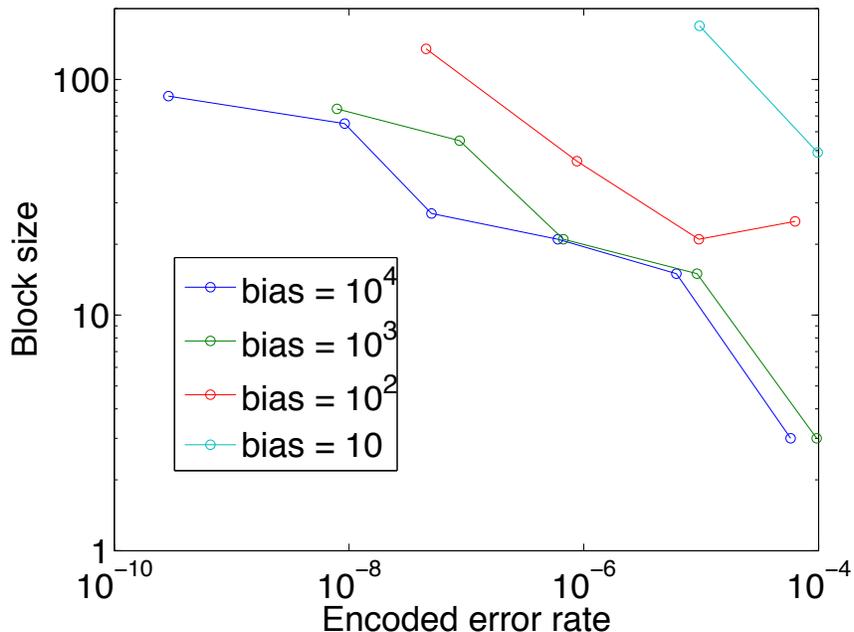


Figure 3.13: Bacon-Shor block size vs optimal (in number of gates) CNOT logical error rate for geometrically local asymmetric Bacon-Shor codes at various values of the bias, for a physical error rate  $\varepsilon = 10^{-4}$ .

the possibility of failure in the preparation of  $|+\rangle^L$  or a cat state is included in Equation (3.21), and conservatively at that, since we assumed there that the cat states had length  $2m$  or  $3m$  rather than  $m + 1$ . Here we assume that the number of repetitions  $r'$  of the cat state syndrome measurement, and the number of syndrome measurement repetitions  $r_+$  in the  $|+\rangle^L$  state preparation, match the number of repetitions in the following CNOT gate. If so, we may bound the injection error probability  $P_{\text{err}}(\mathcal{J})$  by

$$P_{\text{err}}(\mathcal{J}) \leq \varepsilon_\psi + P_{\text{err}}^*(\mathcal{M}_{ZZ}^L) + P_{\text{err}}^*(\mathcal{M}_X), \quad (3.38)$$

where  $\varepsilon_\psi$  bounds the probability of an error in the preparation of the unprotected state  $|\psi\rangle$  (we presume that this error is not necessarily diagonal in the  $Z$  basis).

We assume that  $\mathcal{M}_{ZZ}^L$  is repeated  $r$  times to improve its reliability. In the first measurement

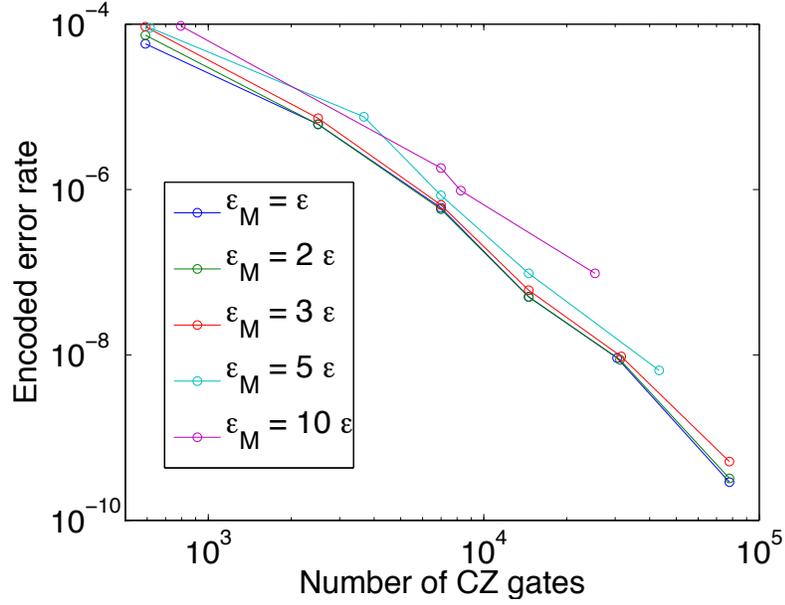


Figure 3.14: Optimal CNOT logical error rate versus required number of gates for geometrically local asymmetric Bacon-Shor codes, assuming a physical error rate  $\epsilon = 10^{-4}$ , bias  $10^4$ , and various rates of the measurement error rate  $\epsilon_{\mathcal{M}}$ .

we can time the preparation of the unprotected qubit and the cat state so that there are no storage errors prior to step in which the cat state interacts with the data. But in subsequent measurements, storage errors on the data may accumulate during the  $8r'$  time steps while the cat state is prepared for the next round. To estimate  $P_{\text{err}}^*(\mathcal{M}_X)$ , we note that each of the  $m + 1$  cat-state qubits is acted upon by two CZ gates during each of  $r'$  rounds of cat-state syndrome measurement, where a nondiagonal fault in any of these gates could result in a Pauli-frame  $X$  error. Furthermore, a storage fault acting on the unprotected qubit, a fault in final  $X$  measurement of the unprotected qubit, or a fault in the CZ gate that couples the unprotected qubit to the cat state could cause  $\mathcal{M}_X$  to fail. Therefore, we find

$$P_{\text{err}}^*(\mathcal{M}_X) \leq (r + 1)(\epsilon + \epsilon') + 8(r - 1)r'(\epsilon_s + \epsilon'_s) + 2(m + 1)rr'\epsilon'. \quad (3.39)$$

A nondiagonal error acting on any data qubit can flip the outcome of  $\mathcal{M}_{ZZ}^L$ . Otherwise, a diagonal fault in each of at least  $(r + 1)/2$  measurement rounds could cause the  $\mathcal{M}_{ZZ}^L$  to fail. Since in each

round there are  $m + 1$  gates coupling the qubits to the data,  $(m + 1)r'$  CZ gates in the cat state preparation, as well as  $m + 1$  single-qubit preparations and measurements, we find

$$P_{\text{err}}^*(\mathcal{M}_{ZZ}^L) \leq r\varepsilon' + 8r'(r - 1)\varepsilon'_s + m(2r_+ + r)\varepsilon' + \left(\frac{r}{r+1}\right)[(m + 1)(r' + 3)(\varepsilon + \varepsilon')]^{(r+1)/2}. \quad (3.40)$$

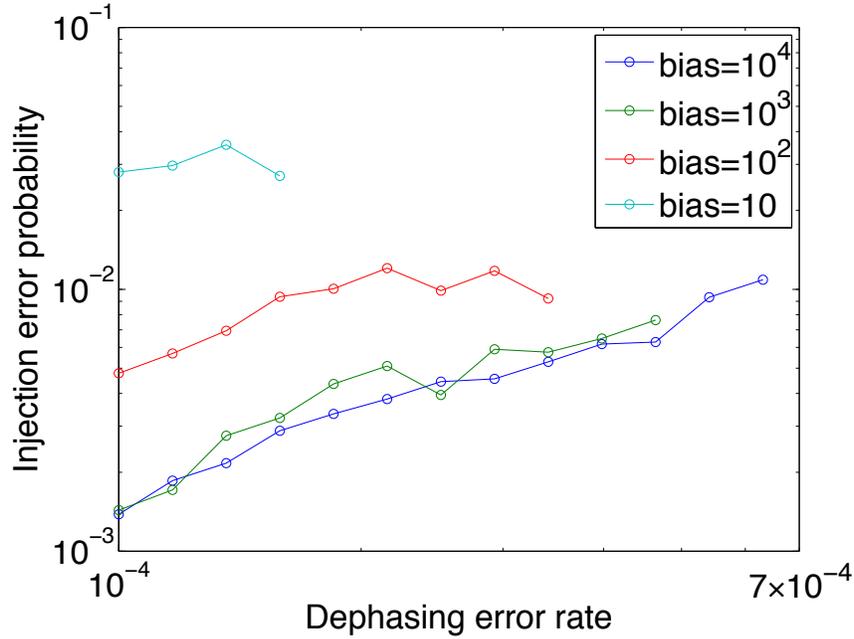


Figure 3.15: Injection error probability for the optimal geometrically local Bacon-Shor codes found in Figure 3.11, with the probability of an error on the unprotected state  $|\psi\rangle$ ,  $\varepsilon_\psi$ , set to be equal to the dephasing error probability  $\varepsilon$ .

Figure 3.15 bounds the injection error probability assuming the values of  $m$ ,  $r$  and  $r'$  chosen to optimize the geometrically local CNOT gadget. For large biases, the injection error probability is around 1%, well below the thresholds for successful distillation, as will be seen in Section 3.10.

By choosing  $r = 1$  (no repetition of the  $ZZ^L$  measurement), we can avoid storage locations, and the leading contribution to  $P_{\text{err}}(\mathcal{J})$  linear in  $\varepsilon$  (assuming  $\varepsilon' \ll \varepsilon$ ) is

$$P_{\text{err}}(\mathcal{J}) = 3\varepsilon + (m + 1)(r' + 3)\varepsilon + \dots. \quad (3.41)$$

Increasing to  $r = 3$  entails increasing the sensitivity to storage errors, while providing better

protection against diagonal errors in CZ gates:

$$P_{\text{err}}(\mathcal{J}) = 5\varepsilon + 16r'\varepsilon_s + \dots, \quad (3.42)$$

which might be a significant improvement if  $\varepsilon_s \ll \varepsilon$ .

The probability of error in state injection depends on  $m$  and also on  $r'$ , the number of times the cat-state syndrome measurement is repeated, which we assume matches the value of  $r'$  in the  $\text{CNOT}^L$  gate that follows the state injection step. We note, though, that the value of  $r'$  might increase gradually as state injection proceeds. If there are multiple rounds of state distillation, we might be willing to accept a larger  $\text{CNOT}^L$  error rate  $P_{\text{err}}(\text{CNOT}^L)$  in early rounds, where the error in the distilled state is higher, with  $P_{\text{err}}(\text{CNOT}^L)$  declining in later rounds as the state's purity improves. Adjustment of the Bacon-Shor block size, and hence of  $P_{\text{err}}(\text{CNOT}^L)$ , is easy to incorporate in our circuit constructions, as there is no need for the control and target blocks in the  $\text{CNOT}^L$  circuit to be of equal size. Using a smaller code in early rounds may save on overhead, but, more importantly, reducing the value of  $m$  and  $r'$  used in the first round improves the error  $P_{\text{err}}(\mathcal{J})$  of the initially injected state. However, for simplicity, we did not invoke this strategy of gradually increasing the block size in the analysis of state distillation reported in Section 3.10.

### 3.10 State distillation

So far we have seen how, using asymmetric Bacon-Shor codes, to perform fault-tolerant encoded versions of the operations in the ‘‘CSS set’’  $\mathcal{G}_{\text{CSS}}^L$ : the CNOT gate, preparations of the encoded states  $|0\rangle^L$  and  $|+\rangle^L$ , and the measurements  $\mathcal{M}_X^L$ ,  $\mathcal{M}_Z^L$ . In addition, to perform fault-tolerant universal quantum computation, we will need to prepare high-fidelity encoded versions of the states

$$\begin{aligned} |+i\rangle &= \frac{1}{\sqrt{2}} (|0\rangle + i|1\rangle), \\ |T\rangle &= \frac{1}{\sqrt{2}} (|0\rangle + e^{i\pi/4}|1\rangle). \end{aligned} \quad (3.43)$$

Using a  $|+i\rangle$  ancilla state and CSS operations, we can teleport the Clifford group gates  $Q = \exp(i\frac{\pi}{4}X)$  and  $S = \exp(-i\frac{\pi}{4}Z)$ , which suffice for generating the full Clifford group. Using a  $|T\rangle$  ancilla, the  $S$  gate, and CSS operations, we can teleport  $T = \exp(-i\frac{\pi}{8}Z)$ , completing a universal

gate set.

To prepare these encoded ancilla states, we first prepare noisy versions of the encoded states, and then use a distillation protocol to generate the needed high-fidelity versions of these states [47]. To distill the  $|+i\rangle$  ancillas, we need only CSS operations, and can use a circuit based on the  $[[7,1,3]]$  Steane code; it takes 7 noisy  $|+i\rangle$  ( $|-i\rangle$ ) ancillas and produces one clean  $|-i\rangle$  ( $|+i\rangle$ ) ancilla. The  $|T\rangle$  distillation protocol uses both CSS operations and  $|+i\rangle$  ancillas and is based on Reed-Muller codes [47]; it produces one clean  $|T\rangle$  ancilla from 15 noisy ones. Other, more efficient, distillation protocols have been proposed recently [70–72].

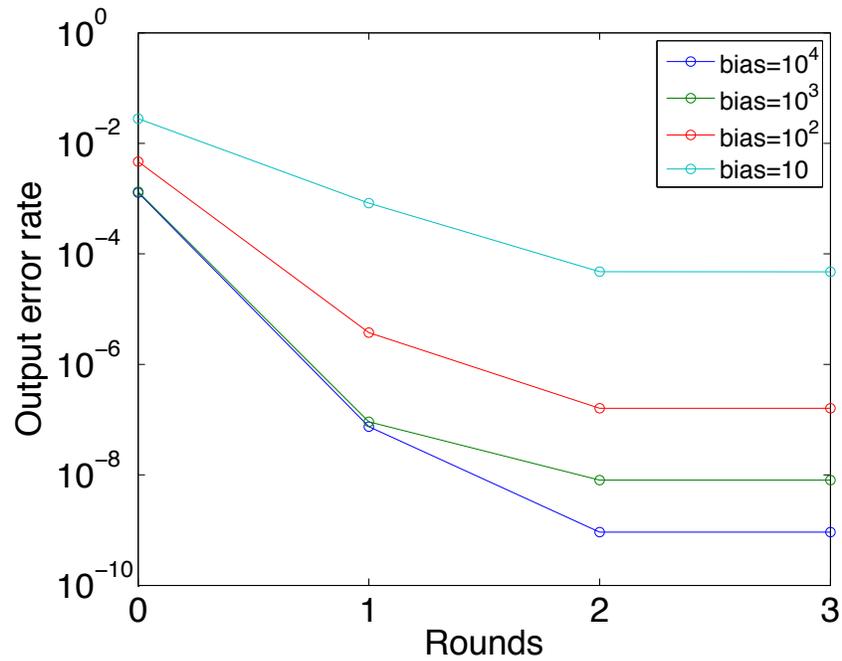


Figure 3.16: Output error probability after  $n$  levels of  $|T\rangle$  distillation, with the values of  $\varepsilon_{\text{in}}$  and  $\varepsilon_{\text{CSS}}$  taken from the optimal values of Figure 3.15, for  $\varepsilon = 10^{-4}$  and various biases. Only two rounds of distillation are required in each case to reach the floor.

In the original analysis [47] of the performance of these state distillation protocols, it was assumed that the CSS operations were perfect. In Chapter 4, we consider the case in which the CSS operations are themselves noisy. Here we consider the case where the noise level is set by the noisy operations protected by an optimally chosen Bacon-Shor code.

When the CSS operations are perfect, we can attain ancilla states with arbitrarily low error by performing a sufficient number of rounds of distillation, assuming that the fidelity of the initial ancilla with the desired ideal state exceeds a threshold value. However, when the CSS operations are noisy, the error rate  $\varepsilon_{\text{CSS}}$  for the CSS operations sets a nonzero floor on the error in the output ancilla states — once this floor is reached, further rounds of distillation produce no further improvement, because the error in the output is dominated by the errors in the distillation circuit rather than the errors in the input ancilla states.

For initial ancillas with errors that are very close to the threshold of the ideal distillation protocol, it can take many rounds of distillation in principle to reach the floor set by the CSS error rate. On the other hand, if the initial ancillas are not too noisy, then it will take just a few rounds of distillation to reach this floor. In Figure 3.16, the error rate of the output ancilla is plotted as a function of the number of rounds of distillation, where the input error rate is the injection error probability estimated in Figure 3.15. For the parameters shown, only two rounds of distillation suffice to reach the floor set by the logical CSS error rate. This relatively modest overhead cost for executing non-Clifford logical quantum gates is one of the advantages of our fault-tolerant scheme.

## Chapter 4

# Magic State Distillation with Noisy Clifford Gates

*In that Empire, the Art of Cartography attained such Perfection that the map of a single Province occupied the entirety of a City, and the map of the Empire, the entirety of a Province. In time, those Unconscionable Maps no longer satisfied, and the Cartographers Guilds struck a Map of the Empire whose size was that of the Empire, and which coincided point for point with it. The following Generations, who were not so fond of the Study of Cartography as their Forebears had been, saw that that vast Map was Useless, and not without some Pitilessness was it, that they delivered it up to the Inclemencies of Sun and Winters. In the Deserts of the West, still today, there are Tattered Ruins of that Map, inhabited by Animals and Beggars; in all the Land there is no other Relic of the Disciplines of Geography.*

—Jorge Louis Borges, “On Exactitude in Science” [73]

### 4.1 Introduction

If we wish to perform a large-scale quantum calculation, in the presence of a high rate of errors, we must ensure that we can detect and correct errors before they overwhelm the computation. Typically, this is achieved by encoding the system of interest into a subspace of a larger system, for example a quantum error-correcting code. If the error rate is not too big, typical errors will move the state outside of the code subspace and be correctable, and the rate of errors at the encoded level

will be small. If the error rate is small enough—below a threshold which depends on the choice of code—then we can achieve an arbitrary level of error correction by treating the encoded errors as physical errors at the *next* level of error correction. Depending on the code, other methods of reaching arbitrarily low error rates may also exist [58, 74, 75].

Using an error-correcting code during a computation entails developing encoded gadgets which replace the ideal circuit’s basic elements. It is important to design these gadgets fault-tolerantly, so that errors cannot spread out too far over the course of a computation. If errors are allowed to spread, then a single error could be transformed into a larger set of *correlated* errors, which would reduce the effective distance of the code. In the worst case, a small set of uncorrelated errors could evolve directly into an encoded error, at which point it would be undetectable and could have adverse effects on the higher-level computation.

One strategy for avoiding correlated errors is to use transversal encoded gates. Unfortunately, it is known that any local quantum error-correcting code cannot have a universal set of encoded gates [76]. Even if we replace transversal (depth 1) encoded gates with constant depth encoded gates, which correspond to finite spread of errors, the set of encoded gates in that form is non-universal for stabilizer codes in  $d = 1$  or 2 dimensions [77]. Codes with these dimensions are exactly the codes which are convenient to work with in actual physical implementations, where the gate arrangement is most naturally 2 dimensional. Therefore, for these code families, at least one encoded gate which is required for universality must be implemented in a non-transversal and non-constant-depth manner.

A particularly promising strategy for fault tolerance in the face of these no-go theorems is to supplement the available transversal gates with the ability to prepare high-quality copies of certain “magic states” [47]. Combining these states with appropriate teleported gates can enable operations which complete a universal basis, where the error in the gate depends on the quality of the magic state. Moreover, even though Clifford operations cannot produce these states, Clifford operations can *distill* them, creating some number of higher-fidelity copies from a greater number of low-fidelity copies. These distillation circuits are based on error-correcting codes with certain symmetries.

Most analyses of Clifford gates assume that the Clifford gates are perfect. In practice, the presence of noisy Clifford operations can limit the effectiveness of distillation protocols. The effects of noisy Clifford operations has been studied before for one choice of distillation routine [78]. At the same time, there has recently been tremendous growth in the discovery of new, more efficient

state distillation protocols [70–72, 79–82]. This work aims to extend the number of protocols that have been analyzed in the presence of noisy Clifford operations, and to analyze the comparative performance of these techniques.

Section 4.2 introduces magic states in more detail. Section 4.3 discusses codes with  $S$ -symmetry, and their use for distilling a non-magic state which is nonetheless useful for completing the Clifford group in circumstances where it is not readily available. Sections 4.4 and 4.5 introduce  $T$ -symmetric codes and related symmetries which allow for the distillation of the magic state  $|T\rangle$ , while Section 4.6 does the same for  $H$ -symmetry. An overview of the effects of noisy Clifford operation is provided in Section 4.7. We introduce in Section 4.7.1 our error model for noisy Clifford gates, and in Section 4.7.2 present our methods for completing the specification of distillation circuits. Section 4.7.3 provides some more details on the analysis of non-Clifford circuit elements used in some distillation procedures. Sections 4.7.4 to 4.7.6 provide the detailed accounting of upper bounds on various distillation protocols, and Section 4.7.7 provides a comparative analysis of the protocols. Section 4.8 discusses the results.

## 4.2 Magic states

Magic states are a set of pure states corresponding to symmetry points on the Bloch sphere. The single-qubit magic states can be divided into two types. The  $H$ -type magic states include the state

$$|H\rangle\langle H| = \frac{1}{2} \left[ I + \frac{1}{\sqrt{2}}(X + Z) \right], \quad (4.1)$$

an eigenstate of the Hadamard operation, and all states obtained from  $|H\rangle$  by the action of single-qubit Clifford operations. Another prominent  $H$ -type magic state is the state

$$|T\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\pi/4}|1\rangle). \quad (4.2)$$

There is one magic state of this type for each of the *edges* of the polyhedron whose vertices are the eigenstates of the Pauli operators  $X$ ,  $Y$  and  $Z$  (that is, the basis states  $|0\rangle$ ,  $|1\rangle$ ,  $|+\rangle$ ,  $|-\rangle$ ,  $|i\rangle$  and  $|-i\rangle$ .)

The other type of magic states correspond to the symmetry axes of the *faces* of the Pauli polyhedron, which represent symmetries under rotation by  $2\pi/3$ . The prototypical magic state of

this type is the state

$$|\Upsilon\rangle\langle\Upsilon| = \frac{1}{2} \left[ I + \frac{1}{\sqrt{3}}(X + Y + Z) \right], \quad (4.3)$$

which is an eigenstate of the rotation  $\Upsilon$  that takes  $X$  to  $Z$ ,  $Z$  to  $Y$ , and  $Y$  to  $X^1$ .

### 4.3 $S$ -symmetric codes

An  $[[n, k, d]]$  code is  $S$ -symmetric if the simultaneous application of the operator  $S$  to each physical qubit is equivalent to the application of the operator  $S_L$  to each *logical* qubit:

$$S^{\otimes n} = (S_L)^{\overline{\otimes} k}, \quad (4.4)$$

where

$$S = \exp(i\pi/4) \exp\left(-i\frac{\pi}{4}Z\right) = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \quad (4.5)$$

and  $S_L$  is the equivalent logical operator. Note that on the left-hand side of Equation (4.4), the tensor product  $\otimes$  is taken with respect to the physical basis of qubits, while on the right-hand side it is with respect to the logical basis, and written as  $\overline{\otimes}$ . It could be explicitly defined as as

$$(S_L)^{\overline{\otimes} k} = \prod_{a \in [k]} (S_L)_a. \quad (4.6)$$

The shorthand version will be used throughout without further comment.

A related symmetry is

$$S^{\otimes n} = (S_L^\dagger)^{\overline{\otimes} k}, \quad (4.7)$$

which I will call  $S$ -antisymmetric. There could also be cases where some logical qubits are symmetric and some antisymmetric:

$$\bigotimes_{i \in [n]} A_i = \prod_{a \in [k]} B_a, \quad (4.8)$$

---

<sup>1</sup>The state  $|\Upsilon\rangle$  and corresponding operator  $\Upsilon$  are elsewhere referred to as  $|T\rangle$  and  $T$ , where the state we call  $|T\rangle$  is called  $|A\rangle$ .

where

$$\begin{aligned} A_i &\in \{S, S^\dagger\} \\ B_a &\in \{S_L, S_L^\dagger\}, \end{aligned} \tag{4.9}$$

but this mixed symmetry does not occur in any common codes.

### 4.3.1 Doubly even codes

$S$ -symmetries occur in any CSS code where the  $X$ -basis stabilizer generators  $\mathcal{T}_X$  and logical operators  $\mathcal{L}_X$  obey the “doubly even” properties

$$\text{wt}(s \in \mathcal{T}_X) \equiv 0 \pmod{4} \tag{4.10}$$

$$\text{wt}(l \in \mathcal{L}_X) \equiv \pm 1 \pmod{4} \tag{4.11}$$

$$\text{wt}(s \wedge t \mid s, t \in \mathcal{L}_X \cup \mathcal{T}_X) \equiv 0 \pmod{2}, \tag{4.12}$$

where  $s \wedge t$  denotes the bitwise AND operation

$$(s \wedge t)_i = s_i t_i. \tag{4.13}$$

The case  $+1$  corresponds to  $S$ -symmetry, and  $-1$  to  $S$ -antisymmetry. Using the identity

$$\text{wt}(x \oplus y) = \text{wt}(x) + \text{wt}(y) - 2 \text{wt}(x \wedge y), \tag{4.14}$$

where  $\oplus$  denotes the bitwise XOR operation

$$(s \oplus t)_i = s_i + t_i \pmod{2}, \tag{4.15}$$

we can see that Equations (4.10) and (4.12) together imply that every stabilizer  $s \in \mathcal{S}_X$  (that is, not just the generators  $\mathcal{T}_X$ ) obeys

$$\text{wt}(s \in \mathcal{S}_X) \equiv 0 \pmod{4}, \tag{4.16}$$

and, furthermore, the combination of logical and stabilizer operations obey

$$\text{wt}(s \oplus t \mid s \in \mathcal{S}_X, t \in \mathcal{S}_X \cup \mathcal{L}_X) \equiv \pm 1 \pmod{4}, \quad (4.17)$$

where the sign  $\pm 1$  is the same sign for every term, depending only on the choice of symmetry.

Then we can express the operation of  $S^{\otimes n}$  on the all-zero codeword  $|0_L\rangle^{\otimes k}$  as

$$\begin{aligned} S^{\otimes n} |0_L\rangle^{\otimes k} &= S^{\otimes n} \sum_{s \in \mathcal{S}_X} X[s] |0\rangle^{\otimes n} \\ &= \sum_{s \in \mathcal{S}_X} X[s] (iSZ)[s] |0\rangle^{\otimes n} \\ &= \sum_{s \in \mathcal{S}_X} X[s] i^{\text{wt}(s)} |0\rangle^{\otimes n} \\ &= |0_L\rangle^{\otimes k}, \end{aligned} \quad (4.18)$$

where on the second line we have used the commutation property

$$SX = XS^\dagger = XSZ, \quad (4.19)$$

and on the final line we have used Equation (4.16). In other words, every term in the logical-basis expansion of  $|0_L\rangle^{\otimes k}$  has a number of  $|1\rangle$ 's which is a multiple of four, and so applying  $S$  to each of these terms leaves it unchanged. Applying any logical operator  $X_i^L = X[l]$  gives

$$\begin{aligned} S^{\otimes n} X_i^L |0_L\rangle^{\otimes k} &= S^{\otimes n} X[l] \sum_{s \in \mathcal{S}_X} X[s] |0\rangle^{\otimes n} \\ &= \sum_{s \in \mathcal{S}_X} X[s \oplus l] (iSZ)[s \oplus l] |0\rangle^{\otimes n} \\ &= \sum_{s \in \mathcal{S}_X} X[s \oplus l] i^{\text{wt}(s \oplus l)} |0\rangle^{\otimes n} \\ &= (i)^{\pm 1} X_i^L |0_L\rangle^{\otimes k}, \end{aligned} \quad (4.20)$$

where the final deduction makes use of Equation (4.17). Now each term in the logical-basis expansion has a number of  $|1\rangle$ 's which is 1 or  $-1$  modulo 4. These are exactly the properties that  $S_i^L$  or

$(S_i^L)^\dagger$  should obey:

$$\begin{aligned} S_i^L |0_L\rangle^{\otimes k} &= |0_L\rangle^{\otimes k} \\ S_i^L X_i^L |0_L\rangle^{\otimes k} &= i X_i^L |0_L\rangle^{\otimes k}. \end{aligned} \quad (4.21)$$

Therefore, any doubly even code has the  $S$ -symmetry property

$$S^{\otimes n} = (S_L)^{\otimes k}. \quad (4.22)$$

$S$ -symmetry occurs, for example, in the  $[[7,1,3]]$  Steane code [25], with stabilizer generators

$$\begin{pmatrix} X & \cdot & X & \cdot & X & \cdot & X \\ \cdot & X & X & \cdot & \cdot & X & X \\ \cdot & \cdot & \cdot & X & X & X & X \end{pmatrix} \quad (4.23)$$

and logical  $X$  operation

$$\begin{pmatrix} X & X & X & \cdot & \cdot & \cdot & \cdot \end{pmatrix}. \quad (4.24)$$

From the logical operation we can see that this case will have the antisymmetric form.

### 4.3.2 $|i\rangle$ distillation

Using an  $S$ -symmetric code allows us to distill the state

$$|i\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle) = S|+\rangle. \quad (4.25)$$

The basic distillation circuit is depicted in Figure 4.1. We can break up the initial  $|i\rangle$  states into  $S|+\rangle$ , and then use the  $S$ -symmetry property Equation (4.4) to push the  $S$  through the error correction and decoder. Since the logical  $S^L$  operations commute with the syndrome, they do not change the measured syndromes during the error correction; however, they do change the recovery operation, changing  $X$  operations to  $Y$  operations as they are conjugated by  $S$  gates.

From the transformed distillation circuit, we can see that in the case of perfect input ancillas,

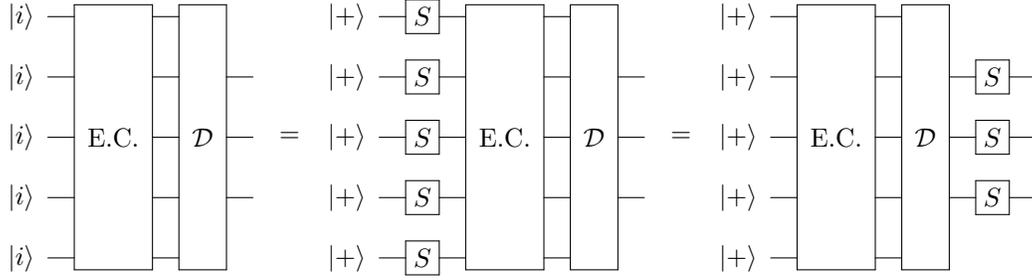


Figure 4.1: Distillation circuit for the state  $|i\rangle$  and equivalent circuit transformations.  $n$  copies of the input state  $|i\rangle$  are prepared, a cycle of error-correction is performed, and the state is decoded into  $k$  copies.

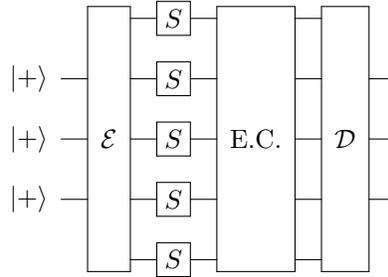


Figure 4.2: Alternate distillation circuit for the state  $|i\rangle$ .  $k$  copies of the input state  $|+\rangle$  are encoded, and noisy  $S$  gates are performed using  $n$  noisy  $|i\rangle$  states and the teleported  $S$  implementation of Figure 4.3. The circuit of a cycle of error-correction is performed, and the state is decoded into  $k$  copies.

the circuit is equivalent to performing error correction on the state  $|+\rangle^{\otimes n}$ . This state agrees with all of the  $X$ -type stabilizers, but not in general with the  $Z$ -type stabilizers. By measuring the  $Z$ -type stabilizers and determining an appropriate correction (which in this case is Pauli, and so can be tracked by updating the Pauli frame), we bring the state back to the codespace, and, in particular, to the logical state  $|+_{L}\rangle^{\otimes k}$ , since

$$\left(\langle +_{L} |^{\otimes k}\right) X_i^L \left(|+\rangle^{\otimes n}\right) = 0 \quad (4.26)$$

for each  $X_i^L$ . In the untransformed circuit, this is equivalent to bringing the state to  $|i_L\rangle^{\otimes k}$ . Then the final portion of the circuit, the decoder  $\mathcal{D}$ , transforms the logical state  $|i_L\rangle^{\otimes k}$  into the bare

state  $|i\rangle^{\otimes k}$ .

So far we have transformed  $n$  ideal copies of the state  $|i\rangle$  into  $k < n$  ideal copies, which we could just have easily done by throwing copies away. The usefulness of the distillation circuit comes into play when we now assume that the input states are noisy. It is useful to assume that the noisy state  $\rho$  takes the form

$$\rho = (1 - \varepsilon) |i\rangle\langle i| + \varepsilon |-i\rangle\langle -i| \quad (4.27)$$

for some probability  $\varepsilon > 0$ , where

$$|-i\rangle = \frac{1}{\sqrt{2}} (|0\rangle - i|1\rangle) = Z|i\rangle. \quad (4.28)$$

This can be achieved by applying the “twirling” map

$$\mathcal{Y}(\rho) = \frac{1}{2} (\rho + Y\rho Y), \quad (4.29)$$

which completely dephases an arbitrary state  $\rho$  in the  $\{|i\rangle, |-i\rangle\}$  basis. Note that this map is equivalent to applying the Pauli  $Y$  operation to each noisy state with probability  $1/2$ , which we can keep track of by updating the Pauli frame appropriately. Then all input errors are equivalent to independent local  $Z$  errors, since  $|-i\rangle = Z|i\rangle$ . If the error-correcting code has distance  $d$ , it can either correct up to  $(d-1)/2$  errors or *detect* up to  $d-1$  errors, leading in the latter case to output ancillas that have an error rate

$$\varepsilon_{\text{out}} = \mathcal{O}(\varepsilon^d). \quad (4.30)$$

If the pre-factor of  $\varepsilon^d$  is not too large, then we have reduced the effective error rate at the next level, and can recursively apply this procedure to reach arbitrarily clean input states.

An alternate distillation circuit is shown in Figure 4.2. In this case, the logical state  $|+L\rangle^{\otimes k}$  is prepared directly, rather than prepared through the syndrome measurement. The advantage is that now in the absence of errors both syndromes should be trivial, and so we can use both syndromes of code for error detection instead of error correction. With twirled inputs and negligible gate errors, we can skip the error correcting cycle for  $X$  errors. The disadvantage is the added circuit complexity of the encoding operation and teleported gates.

For  $S$ -antisymmetric codes, copies of the state  $|\pm i\rangle$  can be distilled from many copies of the state  $|\mp i\rangle$ .

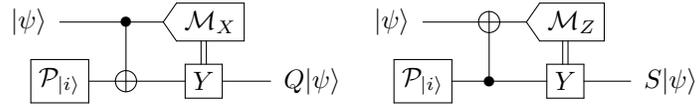


Figure 4.3: The Clifford gates  $Q = e^{i(\pi/4)X}$  and  $S = e^{i(\pi/4)Z}$  can be applied using teleportation and  $|i\rangle$  ancillas. The measurement determines a Pauli operator that is used to update the Pauli frame of the calculation, not actually performed.

The state  $|i\rangle$  is not a magic state, since the gates it enables are in the Clifford group, and in many settings it is taken for granted that we can perform them. Nevertheless, it can be useful to complete the Clifford group in situations like Chapter 3, where the physical gate set is restricted. Using the circuits of Figure 4.3, a single copy of the  $|i\rangle$  state may be consumed to perform either the gate  $S$  or the gate  $Q = e^{i(\pi/4)X}$ . Similarly, the state  $|-i\rangle$  may be used to perform the gates  $S^\dagger$  and  $Q^\dagger$ , or may be transformed to  $|i\rangle$  by application of the Pauli  $Z$  operator.

## 4.4 $T$ -symmetric codes

An  $[[n, k, d]]$  code which is  $T$ -symmetric obeys

$$T^{\otimes n} = (T_L)^{\otimes k}, \quad (4.31)$$

where

$$T = \exp\left(i\frac{\pi}{8}\right) \exp\left(-i\frac{\pi}{8}Z\right) = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{i} \end{pmatrix} \quad (4.32)$$

is the rotation about the  $Z$  axis by  $\pi/4$ . This can be slightly generalized to

$$T^{\otimes n} = (T_L^a)^{\otimes k}, \quad a \in \{1, 3, 5, 7\}, \quad (4.33)$$

as well as possibly mixed-symmetric varieties analogous to Equation (4.8).





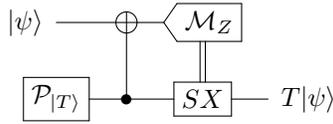


Figure 4.5: Teleported circuit for the non-Clifford gate  $T$ .

basis (where  $|T^\dagger\rangle = T^\dagger|+\rangle = Z|T\rangle$ .) This can be achieved by applying the dephasing map

$$\mathcal{T}(\rho) = \frac{1}{2} (\rho + SX\rho XS^\dagger), \quad (4.43)$$

which applies the operation  $SX$  to each copy of  $\rho$  with probability  $1/2$ . Unlike  $|i\rangle$ -distillation, this is a non-Clifford operation and must be physically applied if we wish to dephase, although the work of O'Connor *et al.* [78] has shown distillation protocols to be robust with regard to the removal of the dephasing step.

The state  $|T\rangle$  may be used to teleport the gate  $T$ , using the circuit of Figure 4.5.

## 4.5 Triorthogonal codes

Triorthogonal codes, introduced by Bravyi and Haah [71], generalize the idea of  $T$ -symmetric codes by noting that it is still useful for distillation to have codes that are not precisely symmetric but which are symmetric *up to Clifford operations*. This symmetry has the form

$$U_{\text{Clifford}} T^{\otimes n} = (T_L)^{\otimes k}, \quad (4.44)$$

where  $U_{\text{Clifford}}$  is a unitary which is inside the Clifford group.

As discovered in [71], such codes can be found from CSS codes with stabilizer generators  $\mathcal{T}_X$

and logical operators  $\mathcal{L}_X$  that obey the conditions

$$\text{wt}(s \in \mathcal{T}_X) \equiv 0 \pmod{2} \quad (4.45)$$

$$\text{wt}(l \in \mathcal{L}_X) \equiv 1 \pmod{2} \quad (4.46)$$

$$\text{wt}(s \wedge t \mid s \neq t \in \mathcal{L}_X \cup \mathcal{T}_X) \equiv 0 \pmod{2} \quad (4.47)$$

$$\text{wt}(s \wedge t \wedge u \mid s \neq t \neq u \in \mathcal{L}_X \cup \mathcal{T}_X) \equiv 0 \pmod{2}. \quad (4.48)$$

With the exception of the new condition Equation (4.48), these are relaxations of the equivalent conditions for triply even codes, Equations (4.34) to (4.36). In fact, every triply even code is necessarily also triorthogonal. Any matrix  $G$  with  $n$  columns and  $k$  odd-weight rows, whose rows obey Equation (4.36), generates a CSS-type stabilizer code, which encodes  $k$  qubits into  $n$ , where the even rows become  $\mathcal{T}_X$  and the odd rows become  $\mathcal{L}_X$ . (We will assume that the odd weight rows are the first  $k$  rows.) The final condition Equation (4.48) provides the symmetry property.

To see the symmetry property, note that

$$T^{\otimes n} |_{+L}\rangle^{\otimes k} = \sum_{f \in \mathcal{G}} e^{i\pi \text{wt}(g)/4} |g\rangle, \quad (4.49)$$

where  $\text{wt}(s)$  denotes the weight of the binary string  $s$ .  $\mathcal{G}$  is the set generated by  $\mathcal{L}_X \cup \mathcal{S}_X$ , or equivalently the set generated by the rows of the matrix  $G$ . We can write each  $f \in \mathcal{G}$  as

$$f = \sum_{a=1}^m x_a g^a \pmod{2}, \quad (4.50)$$

where  $g^a$  is the  $a$ th row of  $G$  and  $x_a$  is uniquely determined by  $f$ . Each component  $f_i$  of  $i$  is the parity of the bit string  $(x_1 g_i^1, x_2 g_i^2, \dots, x_m g_i^m)$ . We may write the parity  $\epsilon(s)$  of a bit string  $s$  as

$$\begin{aligned} \epsilon(s) &= \frac{1}{2} \left( 1 - (1 - 2)^{\text{wt}(s)} \right) = \sum_{p=1}^{\text{wt}(s)} \binom{\text{wt}(s)}{p} (-2)^{p-1} \\ &= \sum_{i=1}^m s_i - 2 \sum_{i < j} s_i s_j + 4 \sum_{i < j < k} s_i s_j s_k + \dots, \end{aligned} \quad (4.51)$$

where in the second line we have used the fact that  $\binom{\text{wt}(s)}{p}$  is the number of ways to choose  $p$

non-zero components of  $s$ . This leads to the identity

$$e^{i\pi\epsilon(s)/4} = \exp \left[ \frac{i\pi}{4} \sum_i y_i - \frac{i\pi}{2} \sum_{i<j} s_i s_j + i\pi \sum_{i<j<k} s_i s_j s_k \right], \quad (4.52)$$

where the higher order terms are eliminated because they are integer multiples of  $2\pi i$ . Applying this identity to the term  $e^{i\pi \text{wt}(f)/4}$  in Equation (4.49) leads to the expansion

$$\begin{aligned} e^{i\pi \text{wt}(f)/4} &= \exp \left[ \frac{i\pi}{4} \sum_{a=1}^m x_a \text{wt}(g^a) - \frac{i\pi}{2} \sum_{a<b} x_a x_b \text{wt}(g^a \wedge g^b) + i\pi \sum_{a<b<c} x_a x_b x_c \text{wt}(g^a \wedge g^b \wedge g^c) \right] \\ &= \exp \left[ \frac{i\pi}{4} \sum_{a=1}^m x_a \text{wt}(g^a) - \frac{i\pi}{2} \sum_{a<b} x_a x_b \text{wt}(g^a \wedge g^b) \right] \\ &= \exp \left[ \frac{i\pi}{4} \sum_{a=1}^k x_a + \frac{i\pi}{2} Q(x) \right], \end{aligned} \quad (4.53)$$

where in the second line we have used the triorthogonality condition Equation (4.48), and in the third line we have collected the desired effect, which is the same effect as the encoded  $(T_L)^{\otimes k}$ , and the undesired additional phase. We would like to show that we can get rid of this phase by the application of Clifford gates. Introducing the auxiliary tensors

$$\Gamma_a = \left\lfloor \frac{\text{wt}(g^a)}{2} \right\rfloor \quad (4.54)$$

and

$$\Delta_{ab} = \frac{\text{wt}(g^a \wedge g^b)}{2}, \quad (4.55)$$

we may rewrite the term  $Q(x)$  in Equation (4.53) as

$$Q(x) = \sum_{a=1}^m \Gamma_a x_a - 2 \sum_{a<b} \Gamma_{ab} x_a x_b. \quad (4.56)$$

To see this, it is useful to note that

$$\text{wt}(g^a) = \begin{cases} 2\Gamma_a + 1 & \text{if } 1 \leq a \leq k \\ 2\Gamma_a & \text{otherwise} \end{cases}. \quad (4.57)$$

Using  $B_{ap}$ , the (non-unique) one-sided inverse over  $\mathbb{F}_2$  of  $G_{ap}$ , defined by

$$\sum_a B_{ap} G_{aq} = \delta_{pq}, \quad (4.58)$$

we can write  $x_a$  and  $x_a x_b$  explicitly in terms of  $f$ . We have

$$\begin{aligned} x_a &= \sum_p B_{ap} f_p \pmod{2} \\ &= \sum_p B_{ap} f_p - 2 \sum_{p < q} B_{ap} B_{aq} f_p f_q \pmod{4} \end{aligned} \quad (4.59)$$

$$2x_a x_b = 2 \sum_{p,q} B_{ap} B_{bq} f_p f_q \pmod{4}. \quad (4.60)$$

This allows us to write  $Q(x)$  as

$$Q(x(f)) = \sum_{p=1}^n \Lambda_p f_p - 2 \sum_{p < q} \Omega_{pq} f_p f_q \pmod{4}, \quad (4.61)$$

where

$$\Lambda_p = \sum_a \Gamma_a B_{ap} - 2 \sum_{a < b} \Gamma_{ab} B_{ap} B_{bp} \pmod{4} \quad (4.62)$$

and

$$\Omega_{pq} = \sum_a \Gamma_a B_{ap} B_{aq} - \sum_{a < b} (B_{ap} B_{bq} + B_{bp} B_{aq}) \pmod{2}. \quad (4.63)$$

Then the unwanted phase factor  $e^{i\pi Q/2}$  may be cancelled by applying CZ gates to every pair of qubits  $p < q$  where  $\Omega_{pq} = 1$ , and the gate  $(S^\dagger)^{\Lambda_p}$  to each qubit  $p$ .

#### 4.5.1 Explicit triorthogonal codes

A family of triorthogonal codes with parameters  $[[3k + 8, k, 2]]$  may be explicitly constructed for any even  $k$ . These codes have stabilizers with the form

$\cdot$	$X$	$\cdot$	$X$	$\cdot$	$X$	$\cdot$	$X$	$X$	$\cdot$	$X$	$X$	$\cdot$	$X$	...	$X$	$\cdot$	$X$
$\cdot$	$\cdot$	$X$	$X$	$\cdot$	$\cdot$	$X$	$X$	$\cdot$	$X$	$X$	$\cdot$	$X$	$X$	...	$\cdot$	$X$	$X$
$X$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	...	$\cdot$	$\cdot$	$\cdot$							
$Z$	$Z$	$\cdot$	$Z$	$Z$	$\cdot$	$\cdot$	$\cdot$	...	$\cdot$	$\cdot$	$\cdot$						
$Z$	$Z$	$\cdot$	$Z$	$Z$	...	$\cdot$	$\cdot$	$\cdot$									
$\vdots$																	
$Z$	$Z$	$\cdot$	...	$\cdot$	$Z$	$Z$											
$Z$	$\cdot$	$Z$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$Z$	$\cdot$	$Z$	$\cdot$	$\cdot$	$\cdot$	...	$\cdot$	$\cdot$	$\cdot$
$Z$	$\cdot$	$Z$	$\cdot$	$Z$	$\cdot$	$Z$	...	$\cdot$	$\cdot$	$\cdot$							
$\vdots$																	
$Z$	$\cdot$	$Z$	$\cdot$	...	$Z$	$\cdot$	$Z$										
$Z$	$\cdot$	$\cdot$	$\cdot$	$Z$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$Z$	$\cdot$	$\cdot$	$Z$	...	$\cdot$	$\cdot$	$Z$
$\cdot$	$Z$	$\cdot$	$\cdot$	$\cdot$	$Z$	$\cdot$	$\cdot$	$\cdot$	$\cdot$	$Z$	$\cdot$	$\cdot$	$Z$	...	$\cdot$	$\cdot$	$Z$
$\cdot$	$\cdot$	$Z$	$\cdot$	$\cdot$	$\cdot$	$Z$	$\cdot$	$\cdot$	$\cdot$	$Z$	$\cdot$	$\cdot$	$Z$	...	$\cdot$	$\cdot$	$Z$
$\cdot$	$\cdot$	$\cdot$	$Z$	$\cdot$	$\cdot$	$\cdot$	$Z$	$\cdot$	$\cdot$	$Z$	$\cdot$	$\cdot$	$Z$	...	$\cdot$	$\cdot$	$Z$
$Z$	$Z$	$Z$	$Z$	$\cdot$	...	$\cdot$	$\cdot$	$\cdot$									

and logical operations with the form

$$\left( \begin{array}{cccc|cccc|cccc|cccc|cccc|cccc} \cdot & \cdot & \cdot & \cdot & X & X & X & X & X & X & X & \cdot & \cdot & \cdot & \dots & \cdot & \cdot & \cdot \\ \cdot & Z & Z & \cdot & Z & \cdot & \cdot & \cdot & \dots & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & X & X & X & X & \cdot & \cdot & \cdot & X & X & X & \dots & \cdot & \cdot & \cdot \\ \cdot & Z & Z & \cdot & Z & \dots & \cdot & \cdot & \cdot \\ \vdots & \vdots \\ \cdot & \cdot & \cdot & \cdot & X & X & X & X & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \dots & X & X & X \\ \cdot & Z & Z & \cdot & \dots & \cdot & \cdot & Z \end{array} \right) \cdot$$

Note that when  $k$  is even but not divisible by 4, the resulting code is not only triorthogonal but triply even, and so no extra Clifford operation is required to get  $T$ -symmetry.

#### 4.5.2 $|T\rangle$ distillation from triorthogonal codes

The circuit for distilling based on triorthogonal codes is identical to that of Figure 4.4, with the addition of the extra Clifford operation above, which is applied before error correction. Triorthogonal codes can also be used to distill Toffoli states [82].

### 4.6 $H$ -symmetric codes

An  $H$ -symmetric code obeys

$$H^{\otimes n} = (H_L)^{\overline{\otimes} k}, \quad (4.64)$$

where  $H$  is the Hadamard gate

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (4.65)$$

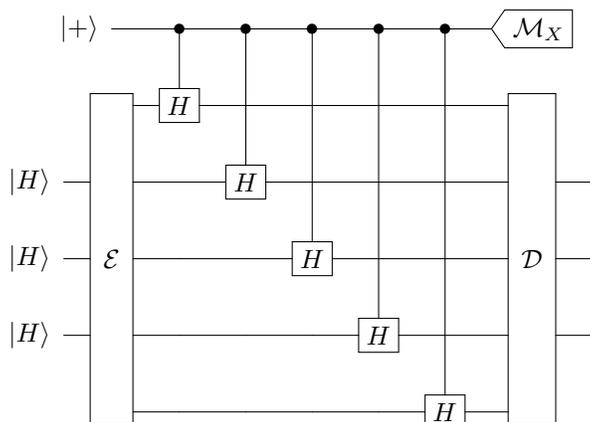


Figure 4.6: Distillation circuit for the state  $|H\rangle$ .  $k$  copies of the input state  $|H\rangle$  are prepared and encoded in an  $H$ -symmetric code. An additional  $2n$   $|H\rangle$  states are consumed to perform controlled- $H$  tests on an ancilla state, measuring the Hadamard parity of the states, using the circuits of Figure 4.7. The attempt is aborted if the ancilla state measurement outcome is  $-1$ , or if the syndrome detected during decoding is nontrivial.

Any self-dual classical code ( $\mathcal{C}^\perp = \mathcal{C}$ ) gives rise to an  $H$ -symmetric CSS code. In such a code, every  $X$ -type stabilizer has a corresponding  $Z$ -type stabilizer, found by replacing every  $X$  with a  $Z$  (or, equivalently, conjugating the stabilizer by Hadamards on each qubit), and, similarly, the partner logical  $Z$  of each logical  $X$  operation is found by bitwise Hadamard conjugation.

Codes with  $H$  symmetry appear to be easier to find than codes with  $T$  symmetry. Examples include the  $[[7, 1, 3]]$  Steane code [25] and a family of  $[[k+4, k, 2]]$  codes [72]. The  $[[4, 2, 2]]$  code [83] has a related symmetry, where  $H^{\otimes 4} = \text{SWAP}_L(H_L)^{\otimes 2}$ , and, similarly, the symmetric  $[[d^2, 1, d]]$  Bacon-Shor codes [29, 54] are  $H$ -symmetric up to a permutation of the qubits.

### 4.6.1 $|H\rangle$ distillation

$H$ -symmetric codes allow distillation of the magic state

$$|H\rangle = \cos(\pi/8)|0\rangle + \sin(\pi/8)|1\rangle, \quad (4.66)$$

the  $+1$  eigenstate of the Hadamard operation.

It is again convenient to assume that all the input magic states are dephased in the  $\{|H\rangle, |-H\rangle\}$

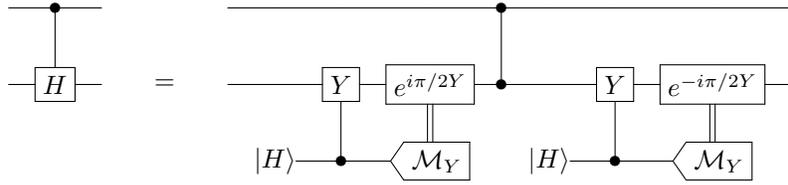


Figure 4.7: Circuit consuming two copies of the magic state  $|H\rangle$  to perform a controlled- $H$  gate.

basis, which can be achieved by applying the map

$$\mathcal{H}(\rho) = \frac{1}{2} (\rho + H\rho H^\dagger) = (1-p)|H\rangle\langle H| + p|-H\rangle\langle -H| \quad (4.67)$$

for some value of  $p$  which depends on  $\rho$ .

The prototypical distillation circuit is depicted in Figure 4.6. Instead of performing an error correcting cycle using the chosen code, a simple controlled- $H$  check is performed on the  $k$  encoded magic states  $|H\rangle$ . This is equivalent to a parity check in the  $\{|H\rangle, |-H\rangle\}$  basis, where

$$|-H\rangle = \sin(\pi/8)|0\rangle - \cos(\pi/8)|1\rangle = Y|H\rangle \quad (4.68)$$

is the  $-1$  eigenstate of the Hadamard operation. To perform the distillation,  $k$  copies of the input state  $|H\rangle$  are prepared and encoded. An additional  $2n$   $|H\rangle$  states are consumed to perform controlled- $H$  tests on an ancilla state, measuring the Hadamard parity of the encoded states, using the circuits of Figure 4.7. Because of the code's  $H$ -symmetry, a transversal controlled- $H$  test is equivalent to logical controlled- $H$  tests. Each error flips the result of the parity test, so the test can detect an odd number of errors. The distillation attempt is aborted and resulting states discarded if the ancilla state measurement outcome is  $-1$ , or if the syndrome detected during decoding is nontrivial. If the test and decoding are both successful, we have produced  $k$   $|H\rangle$  states from  $k + 2n$  states, while reducing the error probability to  $\varepsilon_{\text{out}} = \mathcal{O}(\varepsilon^d)$ , in the absence of noisy Clifford operations.

In some cases, the number of controlled- $H$  tests can be lowered by the prudent application of circuit identities [70].

Two copies of the state  $|H\rangle$  can be used to perform the non-Clifford operation controlled- $H$ ,

using the circuit of Figure 4.7, or a single  $|H\rangle$  may be transformed by Clifford operations into  $|T\rangle$  and used to perform a  $T$  gate using the circuit of Figure 4.5.

## 4.7 Effects of noisy Clifford operations

In the discussion above, and in most analyses of the distillation output error rates, we have assumed that the Clifford gates needed for distillation behave perfectly. This will never be exactly true, because in practice there will always be some probability for these gates to fail. However, we know from error-correcting threshold theorems [26, 30–36] that if the physical level of noise is sufficiently low, then we can reach an *arbitrarily low* level of noise for encoded Clifford operations. Because of this, the effects of noise on the Clifford operations can safely be ignored in the limit of arbitrarily high degree of error correction.

The effects of Clifford errors become more important, however, when we only have a finite degree of error correction available to us. Likewise, they become important when we wish to characterize the amount of resources required to reach a particular error rate— whether in terms of number of physical qubits, number of physical gates, depth of logical circuits, or other measure.

Including the effects of faulty Clifford gates can affect a distillation procedure in a number of ways. These faults can reduce the distillation threshold error rate, so that more perfect ancillas are required before distillation becomes viable. They can also set a floor on the achievable level of distillation, so that distillation protocols become ineffective below a certain error rate. Finally, they can reduce the rate of convergence to this error floor, so that more rounds of distillation are necessary to reach it. This reduction is achieved both by an increased rate in undetected errors on the output qubits, as well as more rounds of distillation which will be rejected due to detected errors.

The effects of noisy Clifford operations have been previously studied for distillation protocols based on the magic state

$$|\Upsilon\rangle = \cos\beta|0\rangle + e^{i\frac{\pi}{4}}\sin\beta|1\rangle, \quad \cos(2\beta) = \frac{1}{\sqrt{3}} \quad (4.69)$$

and distillation using the five qubit code [47, 78]. Here we analyze a set of distillation protocols for  $|T\rangle$  and  $|H\rangle$  states, including some more recent proposals. Specifically, we look at the original  $|T\rangle$

distillation protocol of Bravyi and Kitaev [47], the  $|T\rangle$  distillation protocols of Bravyi and Haah based on triorthogonal codes [71], and the  $|H\rangle$  distillation protocols of Jones [72].

### 4.7.1 Error model

We analyze each protocol under a local stochastic error model with error rate  $p$ , where the probability  $P$  of gates failing at  $r$  specified locations obeys

$$P \leq p^r. \tag{4.70}$$

The failure of a particular gate is modeled by replacing the gate by the combination of the ideal gate, followed by a Pauli error, which we assume can be chosen adversarially, from the set  $\{X, Y, Z\}$  for single-qubit gates and from the 15 possible nontrivial 2-qubit Paulis for two-qubit gates. A state preparation failure results in the preparation of the orthogonal state, and a measurement fault of a two-outcome measurement results in the measurement result being flipped. The probability  $p$  should be interpreted as the probability of failure for the Clifford operations at the highest encoded level, and can therefore be much smaller than the failure rate  $p_0$  of physical gates.

When analyzing  $|T\rangle$  distillation protocols, it is useful to focus on pure  $X$  and pure  $Z$  errors, because for those circuits any logical error produced with  $Y$  faults could also be produced with only  $X$  or only  $Z$  faults. For  $|H\rangle$  protocols, a more nuanced approach is required, as explained in Section 4.7.3.

To achieve an upper bound on the overall failure rate, we will systematically count (aided by computer) the single faults, pairs of faults, and in some cases triples of faults that can lead to an encoded error. For each tuple of faults that is not directly a logical error, we will also count the ways in which input errors can be combined with the faults to produce an encoded error. For larger tuples of faults, we will pessimistically assume that every such tuple leads to an encoded error.

We will similarly model the input magic states as having uniform input error probability  $\varepsilon$ , so that on each input the correct state is input with probability  $1 - \varepsilon$ , and the orthogonal state with probability  $\varepsilon$ . Where necessary, we will refer to  $\varepsilon_{\text{in}}$  as the input error rate, and  $\varepsilon_{\text{out}}$  as the (average) output error rate; or, for multiple rounds of distillation, we will refer to  $\varepsilon^{(j)}$  as the error rate after the  $j$ th round of distillation.

### 4.7.2 Distillation circuit design

The distillation protocols discussed in Sections 4.3 to 4.6 above are not entirely specified: even after choosing the code, there is freedom to choose particular encoding and decoding circuits, and freedom for  $|T\rangle$  distillation to specify the exact procedure for error correction.

For error correction, we will adopt Steane-type error correction, discussed in Section 1.4.1. We will choose encoding and decoding circuits based on the Latin rectangle scheduling method [57,84]; other methods exist which may be able to find more efficient and less error-prone circuits for special cases [85]. We have chosen the encoding and decoding circuits as best as possible to minimize the constant term in the output distilled error rate, averaged over the output qubits. Future improvements in choosing these circuits would result in better performance, so our analysis serves to upper-bound the achievable error rates using these codes.

### 4.7.3 Controlled- $H$ analysis

The distillation circuits for  $|T\rangle$  states involve only Clifford gates at the elementary level. These gates, by definition, propagate Pauli errors into Pauli errors:

$$CPC^\dagger \in \mathcal{P} \text{ for } P \in \mathcal{P}, C \in \mathcal{C}_1. \quad (4.71)$$

For that reason, such circuits are particularly amenable to a worst-case Pauli error model, where the inserted errors are chosen from the Pauli operations.

The distillation protocols for the  $|H\rangle$  state, by contrast, make use of non-Clifford circuit elements: the controlled  $H$  gates  $CH$ . A few Pauli operations propagate to Pauli operations:

$$\begin{aligned} CH_{12} \cdot Z_1 I_2 \cdot CH_{12}^\dagger &= Z_1 I_2 \\ CH_{12} \cdot I_1 Y_2 \cdot CH_{12}^\dagger &= Z_1 Y_2; \end{aligned} \quad (4.72)$$

however, in general the result is not a Pauli operation, for example a  $Z$  error on the target qubit propagates as

$$CH_{12} \cdot I_1 Z_2 \cdot CH_{12}^\dagger = CY_{12} \cdot S_1 Z_2, \quad (4.73)$$

becoming a non-Pauli operation that would then have to be propagated through the circuit.

Fortunately, we will not have to deal with arbitrary circuits of  $CH$  gates, and can take advantage

of the particular structure of the distillation circuits. In each case, we have a single ancilla qubit, prepared in the  $|+\rangle$  state, and a code block state that can be naturally expressed in the basis  $\{|H\rangle, |-H\rangle\}^{\otimes n}$ . A series of  $CH$  gates act from the ancilla to each of the physical qubits, which act trivially on a target  $|H\rangle$  and flip the control qubit from  $|+\rangle \leftrightarrow |-\rangle$  if the target is  $|-H\rangle$ . Finally, the ancilla qubit is measured in the  $X$  basis, which effectively measures the *parity* of the code block state in the  $\{|H\rangle, |-H\rangle\}$  basis, counting whether the number of  $|-H\rangle$ 's is odd or even. Ideally the state of the codeblock is  $|H\rangle^{\otimes n}$ ; conditioned on a successful measurement result, the state is projected onto even parity.

A  $Y$  error in the codeblock flips  $|H\rangle$  and  $|-H\rangle$ , flipping the ancilla state. An odd number of  $Y$  errors will be detected, while an even number can pass undetected. The effect of an  $X$  error can be determined as follows (ignoring normalization):

$$\begin{aligned}
(CH \cdot IX)|+\rangle|H\rangle &= CH(|0\rangle \otimes X|H\rangle + |1\rangle \otimes X|H\rangle) \\
&= (|0\rangle \otimes X|H\rangle + |1\rangle \otimes HX|H\rangle) \\
&= (|0\rangle \otimes X|H\rangle + |1\rangle \otimes Z|H\rangle) \\
&= |+\rangle \otimes (X + Z)|H\rangle + |-\rangle \otimes (X - Z)|H\rangle \\
&= |+\rangle|H\rangle + |-\rangle|-H\rangle
\end{aligned} \tag{4.74}$$

and similarly

$$(CH \cdot IX)|-\rangle|H\rangle = |-\rangle|H\rangle + |+\rangle|-H\rangle \tag{4.75}$$

$$(CH \cdot IX)|+\rangle|-H\rangle = |+\rangle|H\rangle - |-\rangle|-H\rangle \tag{4.76}$$

$$(CH \cdot IX)|-\rangle|-H\rangle = |-\rangle|H\rangle - |+\rangle|-H\rangle. \tag{4.77}$$

We can work out similar rules for  $Z$  by noting that  $Z = -iXY$ . The effect of these Pauli errors is to create a Bell pair state, in the basis  $\{|+\rangle, |-\rangle\}$  for the control qubit and  $\{|H\rangle, |-H\rangle\}$  for the target qubit.

The effect of  $X$  errors on an encoded input state  $|H_L\rangle^{\otimes n}$  is

$$(CH \cdot IX)|+\rangle|\psi\rangle = |+\rangle \otimes (X + Z)|\psi\rangle + |-\rangle \otimes (X - Z)|\psi\rangle, \tag{4.78}$$

and, for a multi-qubit target state with multiple errors, the effect is

$$(CH \cdot (I \otimes X[e]))|+\rangle|\psi\rangle = |+\rangle \otimes (X[e] + Z[e])|\psi\rangle + |-\rangle \otimes (X[e] - Z[e])|\psi\rangle. \quad (4.79)$$

Conditioned on measuring the control qubit to be in the  $|+\rangle$  state (which occurs with probability  $p = 1/4|\langle\psi|(X[e]+Z[e])|\psi\rangle|^2$ ), the operation  $X[e]+Z[e]$  is applied to the target state. This operator is not to be understood as a probabilistic mixture of  $X$  and  $Z$  errors, but rather a coherent mixture, and the resulting operator is in general a projection followed by a unitary rotation.

When  $|\psi\rangle$  is a codeword of an  $H$ -symmetric quantum error correcting code, we can use this additional structure to more closely determine the effect of this non-Pauli error. In this case (assuming no future errors), if  $X[e]$  would be a detected error, then the syndrome measurement during the decoding of the state will detect a syndrome consistent with either the error  $X[e]$  or the error  $Z[e]$ , each with probability  $1/2$ . If  $X[e]$  is equivalent to a stabilizer operation, then so is  $Z[e]$ , and the operation  $X[e] + Z[e]$  has no effect on the encoded information. Only when  $X[e]$  is equivalent to a logical error is  $X[e] + Z[e]$  a logical error as well (albeit a different one than  $X[e]$ .) This non-Pauli logical error is mapped onto a  $Y$ -type Pauli error by a subsequent twirling operator.



As the form of  $Z_L$  and  $X_L$  suggest, this code has distance 3 protection against  $Z$  errors but distance 7 protection against  $X$  errors.

Almost every single fault will cause a detected error and reject the round, so we can get a bound on this rate by counting the total elementary gates.

The distillation circuit, depicted schematically in Figure 4.4, utilizes a prepared logical  $|+\rangle_L$  state which uses 47 elementary operations. In addition, the use of Steane error correction requires a transversal CNOT, a transversal  $Z$ -basis measurement, and up to 7 single qubit  $X S^\dagger$  corrections (up to 14 elementary operations). The final decoding step requires 48 elementary operations. Altogether, there are 139 elementary operations, so we can bound the probability of a detected error as

$$P(\text{detected error}) \leq 139p + 15\varepsilon. \quad (4.82)$$

To bound the probability of an undetected logical error, we will examine all patterns of 3 or fewer faults and determine how close they are to a logical error. It is useful to distinguish faults that happen in the non-transversal part of the circuit with faults that occur in the transversal part, which contribute in the same way as input errors. The probability for any individual qubit of a non-transversal fault is  $p_{\text{trans}} = 3p$ , since there are three such locations per qubit (the transversal CNOT, the transversal  $\mathcal{M}_Z$  measurement, and possibly the correction operator  $X S^\dagger$ , of which only the  $S^\dagger$  is physically performed.)

Counting all patterns with 4 or more faults as leading to undetected failures, we can bound these higher order terms by

$$P \leq \binom{139}{4} p^4 \approx 1.49 \times 10^7 p^4, \quad (4.83)$$

which will be negligible for  $p \ll 0.01$ .

For the  $|+\rangle_L$  preparation, 20 of the 31 CNOT locations could fail in a way leading to an error which is one single-qubit Pauli error away from a logical error. The remaining 11 locations are two errors away, and there are 7 possible patterns of two errors that result in a logical error. The remaining errors would come from faults in the following transversal CNOT of the circuit, with probability  $p$  per qubit. This gives an overall contribution

$$P \leq 20p \cdot (\varepsilon + 3p) + 7 \cdot 11p \cdot (\varepsilon + 3p)^2 \quad (4.84)$$

to the probability of an undetected error. With two faults in the non-transversal section of the preparation, 57 of the  $\binom{32}{2} = 496$  pairs of faults can cause a logical error directly, and the remaining 412 pairs will be counted as one error away from a logical error, so these patterns contribute probability

$$P \leq 57p^2 + 412p^2 \cdot p. \quad (4.85)$$

There are 1,791 ways to choose 3 faults that lead directly to a logical error, adding probability

$$P \leq 1,791p^3. \quad (4.86)$$

Similarly, we can count faults that occur in the decoding  $\mathcal{D}$  itself. 8 of the 34 CNOT gates act directly on the outgoing qubit. If the circuit is properly arranged, 4 of the 34 CNOT gates can put an error only on the outgoing qubit, avoiding detection; 21 require one additional error, and 7 require two errors (each with in 7 possible ways to do so.) This contributes

$$P \leq 4p + 21p(\varepsilon + 3p) + 7 \cdot 7(\varepsilon + 3p)^2 \quad (4.87)$$

to the error probability.

Out of the  $\binom{34}{2} = 561$  pairs of faults in the decoder, 84 lead directly to a logical error; the remaining 477 will be counted as a single error away, which is true for nearly all of them. This contributes

$$P \leq 84p^2 + 477p^2(\varepsilon + 3p) \quad (4.88)$$

Similarly, we could pair a non-transversal fault in the decoder with a non-transversal fault in the logical  $|+\rangle_L$  preparation, such that the decoder fault completes the logical error. Summing up *all* of the pairs of faults involving either the encoder, decoder, or both, gives contribution

$$P \leq 284p^2 + 1,809p^2(\varepsilon + 3p) \quad (4.89)$$

to the probability of an undetected error.

Including faults in both the encoding and decoding, there are 14798 ways for triples of faults to

directly cause a logical error, giving a contribution

$$P \leq 14,798p^3 \quad (4.90)$$

Finally, there are 35 ways for 3 independent errors to produce a logical error. Each independent error could come from an input error, with probability  $\varepsilon$ , or from a transversal part of the circuit, with probability  $p_{\text{trans}} = 3p$ . This contributes

$$P \leq 35(\varepsilon + 3p)^3 \quad (4.91)$$

to the probability of an undetected error.

Summing up the contributions from Equations (4.83) to (4.91), we find the probability of an undetected logical error to be

$$\begin{aligned} \varepsilon_{\text{out}} = P(\text{undetected error}) &\leq 4p + 284p^2 + 14,798p^3 + 14,891,626p^4 \\ &\quad + 41p(\varepsilon + 3p) + 1,809p^2(\varepsilon + 3p) \\ &\quad + 126p(\varepsilon + 3p)^2 + 35(\varepsilon + 3p)^3. \end{aligned} \quad (4.92)$$

For sufficiently small values of  $p$ , the threshold  $\varepsilon_{\text{in}}$  for successful distillation is around 0.16. Figure 4.8 depicts the output error probability as a function of input error probability for various Clifford error rates  $p$ . The error rate  $p$  sets a floor on the achievable distillation, corresponding to the  $\varepsilon$ -independent term of Equation (4.92). Above  $p \gtrsim 0.003$ , the noise floor is too high for distillation to be successful. The slope of the curve, which determines the rate of convergence to the floor, is largely unaffected by the presence of noisy Clifford operations, except when close to the floor, because the noise floor term becomes significant before the slope terms do.

Figure 4.9 plots the output error probability after some number of rounds of distillation, for various values of  $p$ . The distillation rapidly approaches the minimum error rate. For  $10^{-20} \leq p \leq 10^{-4}$ —that is, for all practical purposes—and starting with  $\varepsilon_{\text{in}} = 0.01$ , this protocol requires no more than 3 rounds of distillation to reach the minimum error rate. We will require around  $15^r$  initial copies of the  $|T\rangle$  ancilla to distill a single clean copy through  $r$  rounds of distillation, neglecting the detected error probability. This translates to at most 3,375 noisy ancillas to produce a single clean copy.

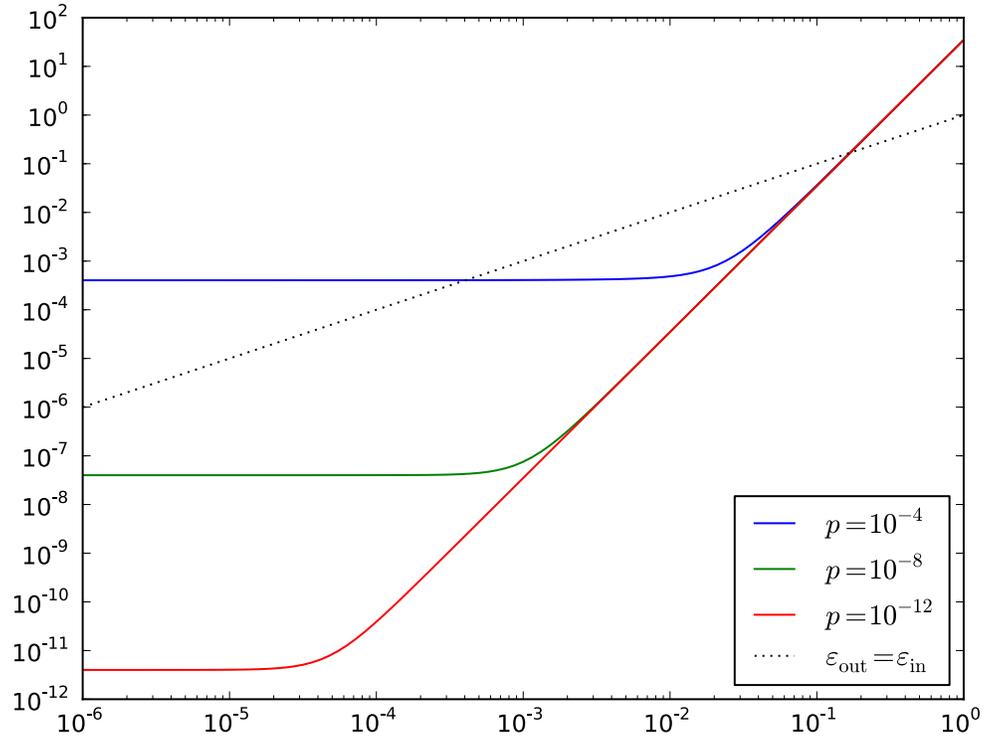


Figure 4.8: Output error rate  $\varepsilon_{\text{out}}$  from Equation (4.92), versus input error rate  $\varepsilon_{\text{in}}$  for distillation based on the  $[[15,1,3]]$  Reed-Muller code, for three different values of the Clifford operation failure rate  $p$ . The dashed line represents the result of no distillation.

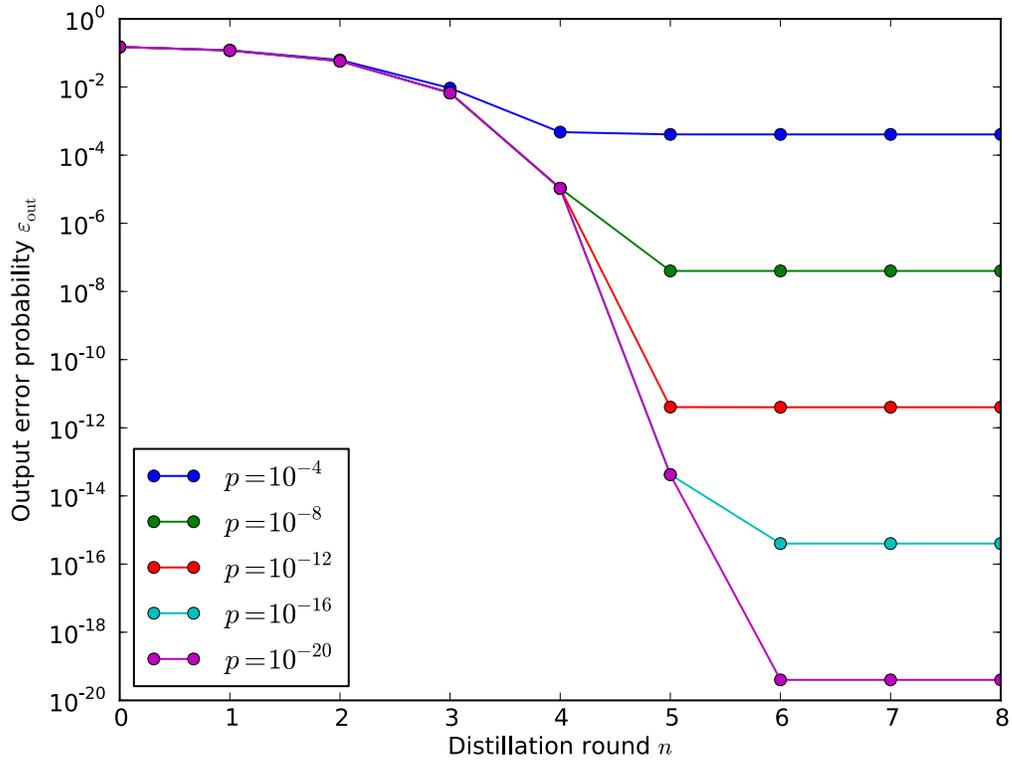


Figure 4.9: Output error rate  $\varepsilon_{\text{out}}$  from Equation (4.92) after  $n$  rounds of distillation, for distillation based on the  $[[15,1,3]]$  Reed-Muller code and various values of the Clifford operation failure rate  $p$ , starting from a near-threshold input error rate of  $\varepsilon^{(0)} = 0.15$ . Starting from  $\varepsilon = 0.01$  (round 3), at most 3 rounds of distillation are required to reach the floor.

### 4.7.5 Bravyi-Haah triorthogonal $|T\rangle$ distillation

The distillation protocol of Bravyi and Haah [71] can be used with any triorthogonal code; we will focus on the  $[[3k + 8, k, 2]]$  family of codes proposed in their paper and discussed in Section 4.5.1. In particular, we will focus on values of  $k$  satisfying  $k = 4a + 2$ ,  $a \in \mathbb{N}$ , since these values require no Clifford unitary to correct.

To bound the overall probability of an error, we again count the total number of circuit locations, assuming here that every isolated fault can be detected. In the preparation of the state  $|+\rangle_L$ , there are  $(3k + 8)$  preparation locations and  $(8k + 13)$  CNOTs, for  $(11k + 21)$  total locations. Injecting the  $|T\rangle$  states requires one CNOT, one measurement, and possibly one  $SX$  correction per physical qubit, for up to  $3 \cdot (3k + 8)$  total locations. To perform Steane error detection, we use the state  $|0\rangle_L$ , which has  $(3k + 8)$  preparations and  $(4k + 13)$  CNOTs, for a total  $(7k + 21)$  locations. In addition, for error detection there is a transversal CNOT and a transversal set of measurements, each with  $(3k + 8)$  locations. Finally, the decoding circuit contains  $(2k + 8)$  measurements and  $(10k + 13)$  CNOT gates, for a total  $(12k + 21)$  locations. In total we have  $45k + 103$  circuit locations, giving

$$P(\text{detected error}) \leq (45k + 103)p + (3k + 8)\varepsilon \quad (4.93)$$

for the probability of a detected error. In practical situations, we will expect  $k$  to be in the range 2-40, so in the extreme case we should require  $p \ll 5 \times 10^{-4}$  and  $\varepsilon_{\text{in}} \ll 0.007$  for detected errors to be negligible.

Instead of the probability of some undetected error, we are interested on the probability of an error on a *given* output qubit (averaged over the choice of output qubits) conditioned on having not detected any error. This quantity depends on the weights of the different outgoing errors. We have carefully chosen our preparation and decoding circuits so that a single fault cannot lead to a correlated error; however, a pair of faults can often do so.

As before, to bound the output error probability, we will count the number of ways in which a non-transversal fault or pair of faults can create a logical error directly. If a fault or pair of faults cannot create a logical error, we will look at whether one additional error can complete a logical error, and in how many ways this might happen. Any triple of fault locations will be assumed to be able to create a full-weight logical error.

From counting the different combinations of faults, aided by computer, we can arrive at an

$k$	$\varepsilon_{\text{out}}$ upper bound
2	$11p + 42p(\varepsilon + 3p) + 562p^2 + 1,046p^2(\varepsilon + 3p) + \binom{193}{3}p^3$
6	$12p + 96p(\varepsilon + 3p) + 1,656p^2 + 7,540p^2(\varepsilon + 3p) + \binom{373}{3}p^3$
10	$12p + 152p(\varepsilon + 3p) + 2,889p^2 + 21,228p^2(\varepsilon + 3p) + \binom{553}{3}p^3$
14	$12p + 208p(\varepsilon + 3p) + 4,273p^2 + 43,040p^2(\varepsilon + 3p) + \binom{733}{3}p^3$
18	$12p + 264p(\varepsilon + 3p) + 5,806p^2 + 73,994p^2(\varepsilon + 3p) + \binom{913}{3}p^3$
22	$12p + 320p(\varepsilon + 3p) + 7,488p^2 + 115,116p^2(\varepsilon + 3p) + \binom{1093}{3}p^3$

Table 4.1: Explicit upper bounds on output error rate  $\varepsilon_{\text{out}}$ , for Bravyi-Haah codes from  $k = 2$  to  $k = 22$ . Bounds for higher values of  $k$  become computationally expensive to determine.

expression for the probability of undetected error and the output error probability for different values of  $k$ , the number of encoded qubits. The output error probability is shown in Table 4.1 and the relative strengths of different terms shown in Figure 4.10. Because a single fault cannot lead to a correlated error, the first term is constant with respect to  $k$ ; however, higher order faults *can* cause correlated errors, leading to the expected linear and quadratic terms. Extracting the trends from Figure 4.10, we arrive at the approximate bound

$$\begin{aligned} \varepsilon_{\text{out}} \lesssim & 12p + 15kp(\varepsilon + 3p) + 323kp^2 + 233k^2p^2(\varepsilon + 3p) \\ & + (3k + 1)(\varepsilon + 3p)^2 + \binom{45k + 103}{3}p^3. \end{aligned} \quad (4.94)$$

#### 4.7.6 Jones $|H\rangle$ distillation

The distillation procedure of Jones [72] uses a  $[[k + 4, k, 2]]$  family of  $H$ -symmetric codes, possibly with multiple levels of concatenation. Here we will consider the first level distillation procedure.

The preparation and decoding circuits each use  $3k + 4$  CNOTs and 4 preparations or measurements, for a total  $3k + 8$  gates. The ancilla requires one preparation and one measurement. Finally, there are  $k + 4$  controlled-Hadamard operations, each with 7 elementary Clifford operations. This gives  $10k + 38$  total locations, for a detected failure probability of

$$P(\text{detected failure}) \leq (10k + 38)p + (k + 4)\varepsilon, \quad (4.95)$$

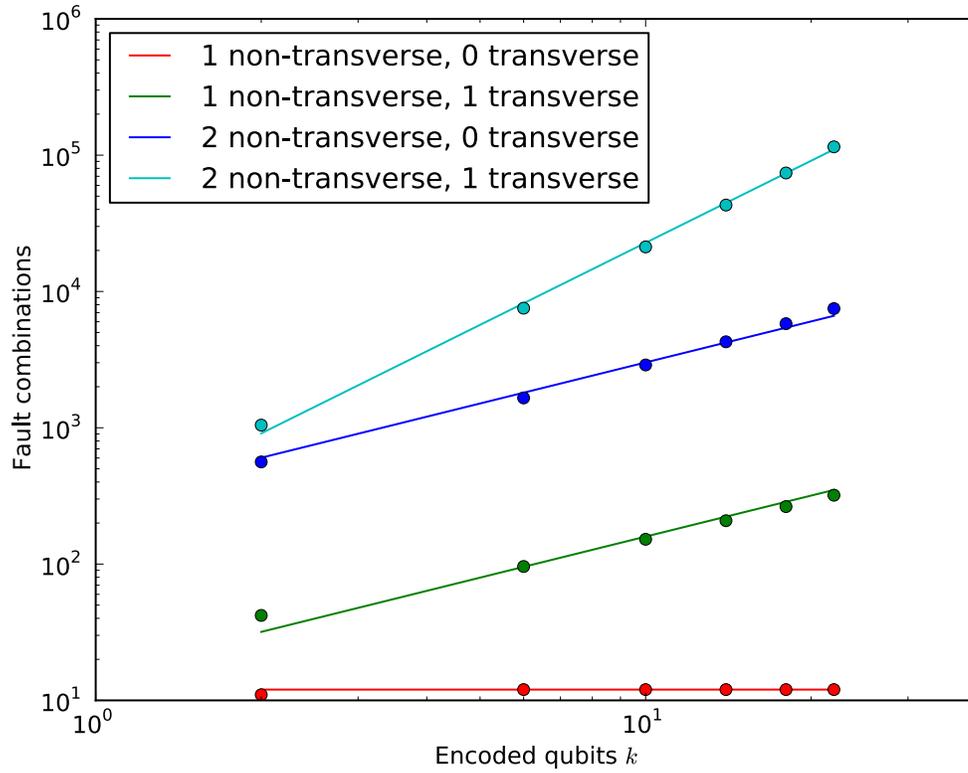


Figure 4.10: Number of ways to combine one or two non-transverse faults and up to one additional non-transverse to create an encoded error on a particular output qubit, averaged over choice of output qubit, for various Bravyi-Haah code sizes  $k$ . The lines are guides to the eye, obeying power laws which are constant, linear, linear, and quadratic from bottom to top.

which will be negligible for  $p \ll .002$  and  $\varepsilon \ll 0.02$  when  $k \leq 40$ .

In analyzing undetected errors, it is useful to separately consider the errors before and after the dephasing operation, as discussed in Section 4.7.3. Errors after the dephasing operation can produce an arbitrary error on the output. In the decoding step, *almost every* CNOT gate can fail in a way leading to an encoded error. However, due to the relatively simple circuit, there are not that many locations for failure compared to other codes, so this is not a major problem. The decoding circuit has been designed to ensure that these errors cannot be correlated with a single fault.

Errors resulting from the encoding side of the circuit pass through the controlled- $H$  test. If there are any  $X$  and  $Z$  errors, this is detected with probability  $1/2$  by the parity measurement. This helps to attenuate the effects of faults in the encoding circuit, as they have increased chances of being detected. The most effective errors on this side of the circuit are errors which produce an even-weight logical  $Y$  operation, which can be produced by as few as two physical  $Y$  errors. However, because of the way CNOTs propagate errors, any  $Y$  error is likely to be accompanied by  $X$  and  $Z$  errors on other qubits. These facts help ensure that the probability of an undetected error is small.

Again aided by a computer counting of combinations of faults, we can arrive at an expression for the probability of undetected error and the output error probability for different values of  $k$ , the number of encoded qubits. The output error probability is shown in Table 4.2 and the relative strengths of different terms shown in Figure 4.11. Extracting the trends from Figure 4.10, we arrive at the approximate bound

$$\begin{aligned} \varepsilon_{\text{out}} \lesssim & 5p + (3/2k + 3)p(\varepsilon + 8p) + (3k^2 + 21k + 15)p^2 \\ & + (3k + 1)(\varepsilon + 8p)^2 + \binom{10k + 38}{3} p^3. \end{aligned} \quad (4.96)$$

### 4.7.7 Comparison

The output error rate as a function of the input error rate is depicted for  $p = 10^{-12}$  in Figure 4.12 for representative Bravyi-Haah and Jones codes, as well as for Reed-Muller based distillation. Away from the noise floor, each Jones and Bravyi-Haah code's behavior for the same  $k$  is nearly identical, as is the case of perfect Clifford gates. Away from the noise floor, the slope of each curve is quadratic for Jones and Bravyi-Haah codes, and cubic for Reed-Muller. Due to the more compact circuits

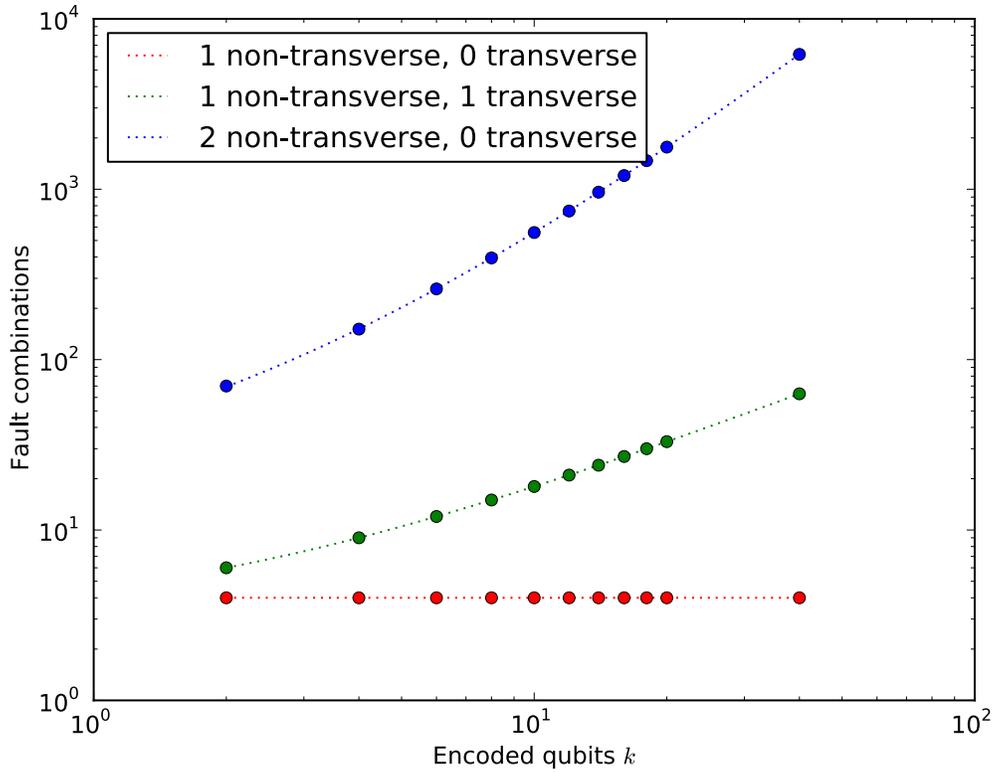


Figure 4.11: Number of ways to combine one or two non-transverse faults and up to one additional non-transverse fault to create an encoded error on a particular output qubit, averaged over choice of output qubit, for various Jones transverse  $H$  code sizes  $k$ . The dotted lines are fits to constant, linear, and quadratic polynomials, from bottom to top.

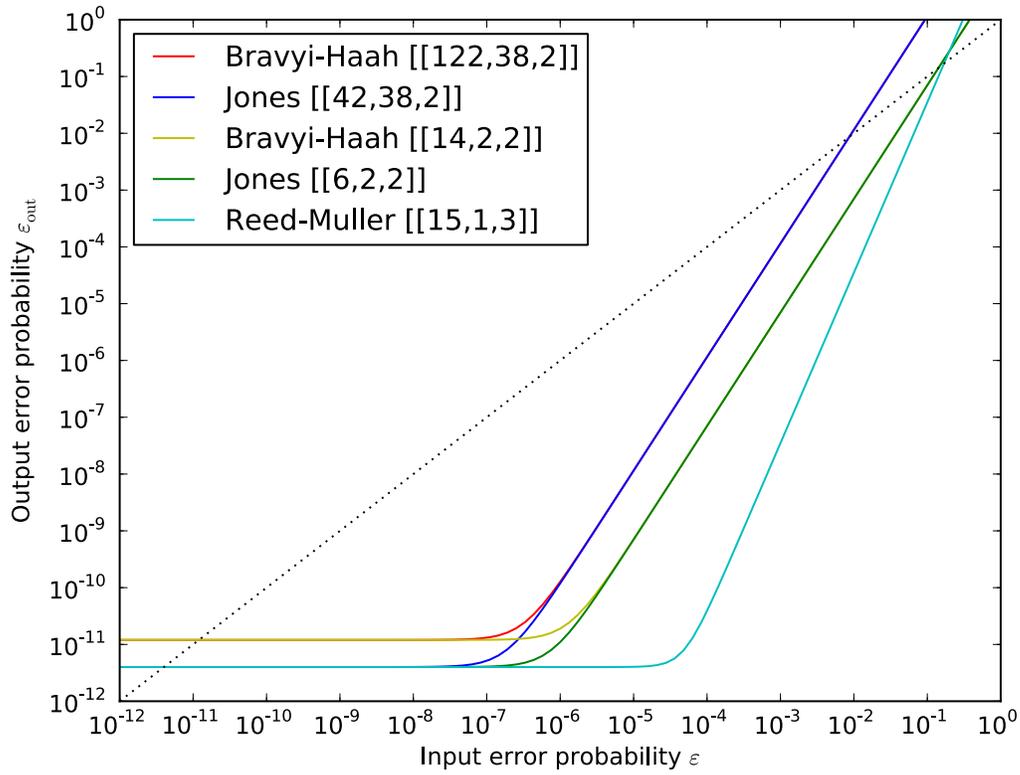


Figure 4.12: Output error rate  $\varepsilon_{\text{out}}$  from Equations (4.94) and (4.96) versus input error rate  $\varepsilon_{\text{in}}$  for distillation based on various Jones and Bravyi-Haah codes, for  $p = 10^{-12}$ . The dashed line represents the result of no distillation.

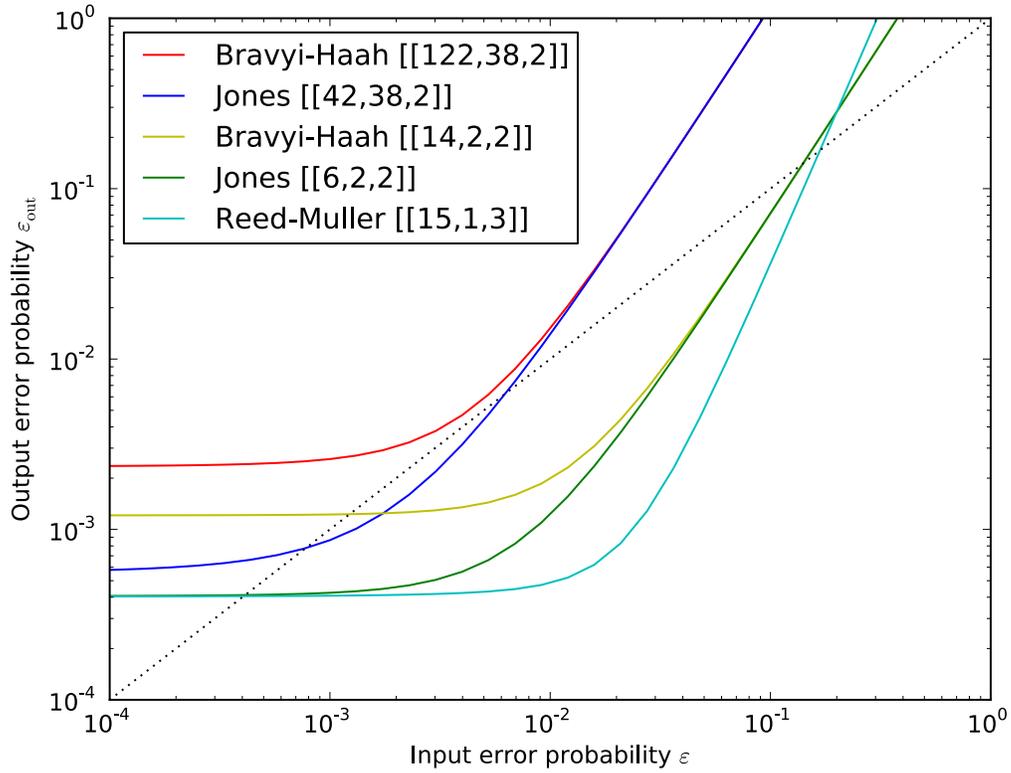


Figure 4.13: Output error rate  $\epsilon_{\text{out}}$  from Equations (4.94) and (4.96) versus input error rate  $\epsilon_{\text{in}}$  for distillation based on various Jones and Bravyi-Haah codes, for  $p = 10^{-4}$ . The dashed line represents the result of no distillation.

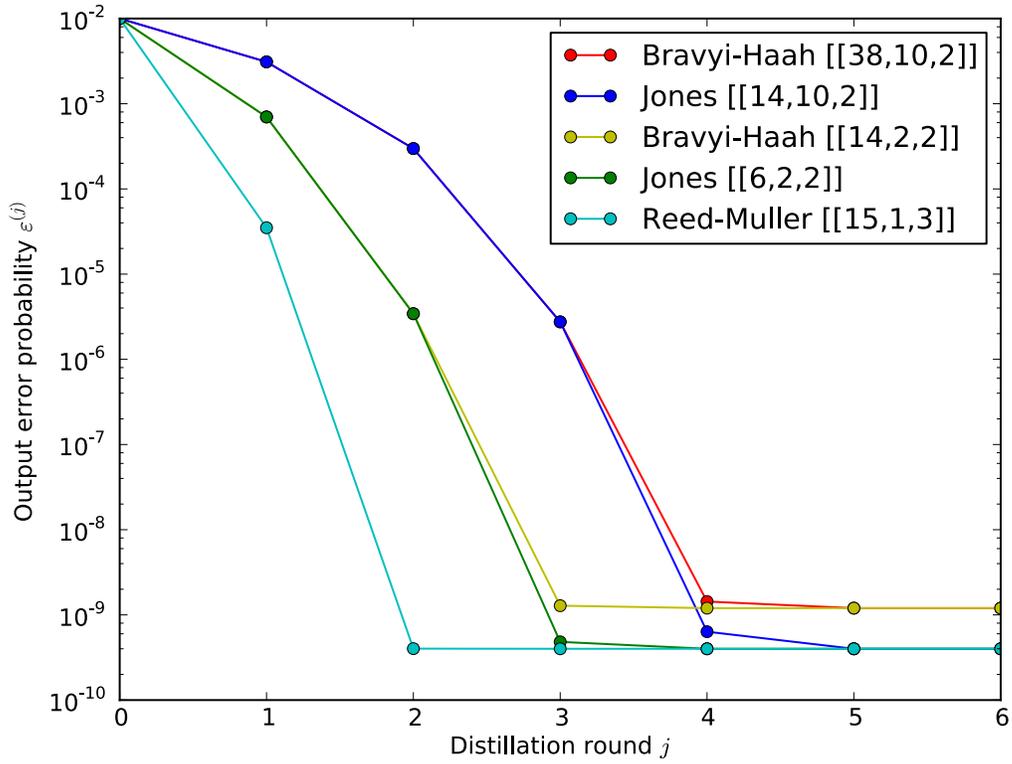


Figure 4.14: Output error probability  $\varepsilon_{\text{out}}$  as a function of number of rounds of distillation, for five different distillation codes, starting from  $\varepsilon_{\text{in}} = 0.01$  and with  $p = 10^{-10}$ .

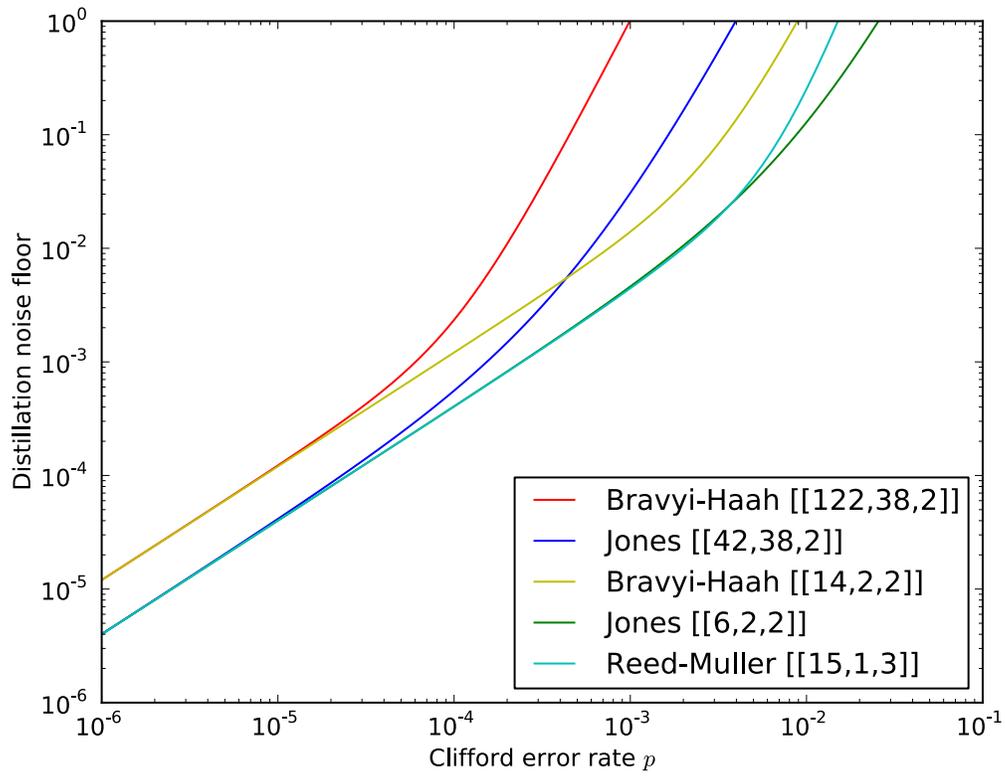


Figure 4.15: Distillation limit as a function of the Clifford error rate, for various distillation procedures

$k$	$\varepsilon_{\text{out}}$ upper bound
2	$5p + 6p(\varepsilon + 8p) + 70p^2 + \binom{58}{3}p^3 + 7(\varepsilon + 8p)^2$
4	$5p + 9p(\varepsilon + 8p) + 152p^2 + \binom{78}{3}p^3 + 13(\varepsilon + 8p)^2$
6	$5p + 12p(\varepsilon + 8p) + 260p^2 + \binom{98}{3}p^3 + 19(\varepsilon + 8p)^2$
8	$5p + 15p(\varepsilon + 8p) + 396p^2 + \binom{118}{3}p^3 + 25(\varepsilon + 8p)^2$
10	$5p + 18p(\varepsilon + 8p) + 558p^2 + \binom{138}{3}p^3 + 31(\varepsilon + 8p)^2$
12	$5p + 21p(\varepsilon + 8p) + 746p^2 + \binom{158}{3}p^3 + 37(\varepsilon + 8p)^2$
14	$5p + 24p(\varepsilon + 8p) + 962p^2 + \binom{178}{3}p^3 + 43(\varepsilon + 8p)^2$
16	$5p + 27p(\varepsilon + 8p) + 1,204p^2 + \binom{198}{3}p^3 + 49(\varepsilon + 8p)^2$
18	$5p + 30p(\varepsilon + 8p) + 1,472p^2 + \binom{218}{3}p^3 + 55(\varepsilon + 8p)^2$
20	$5p + 33p(\varepsilon + 8p) + 1,768p^2 + \binom{238}{3}p^3 + 61(\varepsilon + 8p)^2$
40	$5p + 63p(\varepsilon + 8p) + 6,188p^2 + \binom{438}{3}p^3 + 121(\varepsilon + 8p)^2$

Table 4.2: Explicit upper bounds on output error rate  $\varepsilon_{\text{out}}$ , for Jones codes from  $k = 2$  to  $k = 40$ .

for the Jones codes, their noise floor is slightly lower than the floor for Bravyi-Haah codes. The Jones floor is equivalent to the floor for Reed-Muller, to leading order. In Figure 4.13 the same comparison is plotted for a much higher Clifford error rate of  $p = 10^{-4}$ . This plot demonstrates the larger effect of correlated errors on large block-size codes, rendering the  $k = 38$  Bravyi-Haah code useless at this error rate, and making the  $k = 38$  Jones code nearly useless. The noise floors of the large codes are also significantly affected.

In Figure 4.14, the output error probability is given for five different choices of code, as a function of the number of rounds of distillation. While each code has a comparable distillation limit, the  $[[15, 1, 3]]$  distillation protocol, with its higher distance, is able to reach the limit in only two rounds, while each other protocol requires at least three or more. At the same time, the slower-converging codes can be more efficient in terms of number of resources. Two rounds of  $[[15, 1, 3]]$  distillation require  $15^2 = 225$  input states (ignoring rounds which are rejected due to detected errors), whereas four rounds of Bravyi-Haah  $[[38, 10, 2]]$  or Jones  $[[14, 10, 2]]$  distillation require  $(38/10)^4 \approx 208$  input states per output state. Three rounds of Bravyi-Haah  $[[14, 2, 2]]$  or Jones  $[[6, 2, 2]]$  distillation require  $7^3 = 343$  input states per output state. Depending on how many output distilled states are required, either the  $38 \rightarrow 10$  protocols or the  $15 \rightarrow 1$  protocol may be preferred.

Figure 4.15 shows the distillation limit for representative codes, as a function of the Clifford error

rate  $p$ . The lowest floors come from  $[[6, 2, 2]]$  Jones transverse H code and Reed-Muller  $[[15, 1, 3]]$  code. For error rates above  $\sim 10^{-4}$ , the  $k$ -dependence of the distillation floors of large block codes begins to limit the usefulness of these codes. This effect is more severe for the Bravyi-Haah codes than Jones codes. Despite its larger size, the original protocol based on the Reed-Muller  $[[15, 1, 3]]$  code has a lower floor than some larger codes. This suggests that higher-distance codes may continue to be useful, as long as the size tradeoff is not too severe compared to a distance-2 code.

## 4.8 Discussion

In general, the most useful sequence of protocols will depend on the particular Clifford error rate  $p$ , the ancilla input error rate  $\varepsilon_{\text{in}}$ , and the target output error rate  $\varepsilon_{\text{out}}$ . The presence of the Clifford error rate sets a floor on achievable distillation which depends on the choice of code, so it may be beneficial to switch to a code with the lowest noise floor in the final round of distillation, even if this is a less efficient code in terms of resources. As in the case of perfect Clifford operations, the less efficient codes also tend to have the highest thresholds, so the best sequence of distillation protocols will likely start with a less efficient code, switch to an efficient code to achieve the best resource rate, and switch back to an inefficient code in order to hit the lowest noise floor.

The bounds presented here depend on the choice of encoding and decoding circuits, as well as the error-correcting technique in some cases. Other choices for these parameters could result in better codes. Further development of the techniques developed here may allow for tighter bounds on the codes studied, as well as bounds on a wider range of codes, such as the protocol of Meier, Eastin and Knill [70] which is a relatively small circuit but with a more complicated structure of gates.

The results of this analysis show that, for the protocols studied, the effects of noisy Clifford operations do not have a major detrimental impact on their performance. The most prominent feature is the existence of a distillation noise floor, which for low enough Clifford error rates is a constant multiple, typically in the range 2 – 20 depending on the code and circuits, of the Clifford error rate. The convergence of ancilla quality to the noise floor is not significantly impacted by the presence of noisy Clifford operations. At higher error rates, the noise floor becomes more significant and, depending on the choice of distillation scheme, can prevent useful distillation.

## Chapter 5

# Epilogue

*Our imagination is stretched to the utmost, not, as in fiction, to imagine things which are not really there, but just to comprehend those things which are there.*

—Richard Feynman [86]

We have shown that the analysis of noise biases and hierarchies can be a fruitful approach to studying problems in quantum error correction and fault tolerance. Careful consideration of feature-rich noise landscapes, like the robust gates of Chapter 2 and the asymmetric circuits of Chapter 3, can lead to error correcting codes and fault-tolerant circuits with better performance than could be realized in flatter worlds with simpler, more generic noise models. Increased understanding of the combined effects of multiple error sources can improve our understanding of quantum circuits such as the distillation circuits of Chapter 4, and increase our certainty that they will behave as we hope, not just in ideal models but even in the presence of realistic defects in their operation.

The ability to carefully study and incorporate different competing sources of noise with different strengths will become increasingly important as quantum-mechanical circuits step closer to large-scale real-world implementations. In some cases, this will require us to be more cautious, taking care to ensure that our conclusions are unchanged in the face of new sources of errors. In other cases, we may find that additional knowledge about the structure of noise will allow us to *tailor* error-correcting circuits to the situation at hand, potentially using fewer resources than may be otherwise required.

In every case, designing fault-tolerant circuits in the quantum regime presents new, possibly unforeseen challenges, and correspondingly new opportunities to improve our understanding of

nature, and to innovate with new techniques, all with the hope of moving large-scale quantum computers closer to reality.

# Bibliography

- [1] P. Brooks, A. Kitaev, and J. Preskill, “Protected gates for superconducting qubits,” *Phys. Rev. A* **87** no. 5, (Feb., 2013) 052306, [arXiv:1302.4122](#) [quant-ph].
- [2] P. Brooks and J. Preskill, “Fault-tolerant quantum computation with asymmetric Bacon-Shor codes,” *Phys. Rev. A* **87** no. 3, (Mar., 2013) 032310, [arXiv:1211.1400](#) [quant-ph].
- [3] D. Eisenhower, *Public Papers of the Presidents of the United States, Dwight D. Eisenhower, 1957*. National Archives and Records Service, Government Printing Office, 1957.
- [4] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.
- [5] J. Preskill. Lecture notes.  
<http://www.theory.caltech.edu/people/preskill/ph229/#lecture>.
- [6] A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and Quantum Computation*, vol. 47 of *Graduate Studies in Mathematics*. American Mathematical Society, 2002.
- [7] R. P. Feynman, “Simulating physics with computers,” *Int. J. Theor. Phys.* **21** (1982) 467–488.
- [8] R. P. Feynman, “Quantum mechanical computers,” *Optics News* **11** no. 11, (1985) .
- [9] D. Deutsch, “Quantum Theory, the Church-Turing Principle and the Universal Quantum Computer,” *Proc. Roy. Soc. Lond. A* **400** no. 1818, (1985) 97–117.
- [10] D. Deutsch, “Quantum computational networks,” *Proc. Roy. Soc. Lond. A* **400** no. 1868, (1985) 73–90.

- [11] A. Church, “An unsolvable problem of elementary number theory,” *Amer. J. Math* **58** (1936) 345.
- [12] A. Turing, “On computable numbers, with an application to the Entscheidungsproblem,” *Proc. Lond. Math. Soc.* **42** (1936) 230.
- [13] P. W. Shor, “Algorithms for quantum computation: Discrete logarithms and factoring,” in *Proc. 35th Annual Symposium on Foundations of Computer Science*. Los Alamitos, CA: IEEE Computer Society Press, 1994. [arXiv:quant-ph/9508027](#).
- [14] L. Grover, “A fast quantum mechanical algorithm for database search,” in *Proc. 28th Annual ACM Symposium on the Theory of Computation*, pp. 212–219. New York: ACM Press, 1996.
- [15] M. H. Freedman, A. Kitaev, and Z. Wang, “Simulation of Topological Field Theories by Quantum Computers,” *Communications in Mathematical Physics* **227** (2002) 587–603, [arXiv:quant-ph/0001071](#).
- [16] S. P. Jordan, K. S. M. Lee, and J. Preskill, “Quantum Computation of Scattering in Scalar Quantum Field Theories,” *ArXiv e-prints* (Dec., 2011) , [arXiv:1112.4833 \[hep-th\]](#).
- [17] D. S. Abrams and S. Lloyd, “Simulation of Many-Body Fermi Systems on a Universal Quantum Computer,” *Physical Review Letters* **79** (Sept., 1997) 2586–2589, [arXiv:quant-ph/9703054](#).
- [18] D. S. Abrams and S. Lloyd, “Quantum Algorithm Providing Exponential Speed Increase for Finding Eigenvalues and Eigenvectors,” *Physical Review Letters* **83** (Dec., 1999) 5162–5165, [arXiv:quant-ph/9807070](#).
- [19] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, “Simulated Quantum Computation of Molecular Energies,” *Science* **309** (Sept., 2005) 1704–1707, [arXiv:quant-ph/0604193](#).
- [20] A. Einstein, B. Podolsky, and N. Rosen, “Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?,” *Phys. Rev.* **47** (May, 1935) 777–780.
- [21] A. Y. Kitaev, “Fault-tolerant quantum computation by anyons,” *Annals of Physics* **303** (Jan., 2003) 2–30, [arXiv:quant-ph/9707021](#).

- [22] A. W. Harrow, B. Recht, and I. L. Chuang, “Efficient discrete approximations of quantum gates,” *Journal of Mathematical Physics* **43** (Sept., 2002) 4445–4451, [arXiv:quant-ph/0111031](#).
- [23] C. M. Dawson and M. A. Nielsen, “The Solovay-Kitaev algorithm,” *eprint arXiv:quant-ph/0505030* (May, 2005) , [arXiv:quant-ph/0505030](#).
- [24] A. R. Calderbank and P. W. Shor, “Good quantum error-correcting codes exist,” *Phys. Rev. A* **54** (Aug., 1996) 1098–1105, [arXiv:quant-ph/9512032](#).
- [25] A. Steane, “Multiple-Particle Interference and Quantum Error Correction,” *Proc. Roy. Soc. Lond* **452** (1954) 2551–2557.
- [26] D. Gottesman, *Stabilizer codes and quantum error correction*. PhD thesis, California Institute of Technology, 1997. [arXiv:quant-ph/9705052](#).
- [27] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, “Quantum Error Correction via Codes over GF(4),” *ArXiv e-prints* (Aug., 1996) , [arXiv:quant-ph/9608006](#).
- [28] E. Knill, R. Laflamme, and L. Viola, “Theory of Quantum Error Correction for General Noise,” *ArXiv e-prints* (Aug., 1999) , [arXiv:quant-ph/9908066](#).
- [29] P. Shor, “Fault-tolerant quantum computation,” in *Foundations of Computer Science, 1996. Proceedings., 37th Annual Symposium on*, pp. 56–65, IEEE. 1996. [arXiv:quant-ph/9605011](#).
- [30] J. Preskill, “Reliable quantum computers,” *Royal Society of London Proceedings Series A* **454** (Jan., 1998) 385, [arXiv:quant-ph/9705031](#).
- [31] D. Aharonov and M. Ben-Or, “Fault Tolerant Quantum Computation with Constant Error,” *ArXiv e-prints* (Nov., 1996) , [arXiv:quant-ph/9611025](#).
- [32] D. Aharonov and M. Ben-Or, “Fault-Tolerant Quantum Computation With Constant Error Rate,” *ArXiv e-prints* (June, 1999) , [arXiv:quant-ph/9906129](#).
- [33] A. Y. Kitaev, “Quantum computations: algorithms and error correction,” *Russian Mathematical Surveys* **52** no. 6, (1997) 1191.

- [34] A. Y. Kitaev, “Quantum error correction with imperfect gates,” in *Quantum Communication, Computing, and Measurement*, A. S. Holevo, O. Hirota, and C. M. Caves, eds. New York: Plenum Press, 1997.
- [35] E. Knill, R. Laflamme, and W. Zurek, “Threshold Accuracy for Quantum Computation,” *ArXiv e-prints* (Oct., 1996) , [arXiv:quant-ph/9610011](#).
- [36] E. Knill, R. Laflamme, and W. H. Zurek, “Resilient quantum computation: error models and thresholds,” *Royal Society of London Proceedings Series A* **454** (Jan., 1998) 365, [arXiv:quant-ph/9702058](#).
- [37] A. M. Steane, “Active Stabilization, Quantum Computation, and Quantum State Synthesis,” *Physical Review Letters* **78** (Mar., 1997) 2252–2255, [arXiv:quant-ph/9611027](#).
- [38] E. Knill, “Quantum computing with realistically noisy devices,” *Nature* **434** (Mar., 2005) 39–44, [arXiv:quant-ph/0410199](#).
- [39] C. H. Bennett, G. Brassard, C. Crepeau, R. Jozsa, A. Peres, and W. K. Wootters, “Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels,” *Physical Review Letters* **70** (Mar., 1993) 1895–1899.
- [40] B. Douçot and J. Vidal, “Pairing of Cooper Pairs in a Fully Frustrated Josephson-Junction Chain,” *Physical Review Letters* **88** no. 22, (June, 2002) 227005, [arXiv:cond-mat/0202115](#).
- [41] L. B. Ioffe and M. V. Feigel’man, “Possible realization of an ideal quantum computer in Josephson junction array,” *Phys. Rev. B* **66** no. 22, (Dec., 2002) 224503, [arXiv:cond-mat/0205186](#).
- [42] A. Kitaev, “Protected qubit based on a superconducting current mirror,” *ArXiv e-prints* (Sept., 2006) , [arXiv:cond-mat/0609441](#).
- [43] D. Gottesman, A. Kitaev, and J. Preskill, “Encoding a qubit in an oscillator,” *Phys. Rev. A* **64** no. 1, (July, 2001) 012310, [arXiv:quant-ph/0008040](#).
- [44] V. E. Manucharyan, J. Koch, L. I. Glazman, and M. H. Devoret, “Fluxonium: Single Cooper-Pair Circuit Free of Charge Offsets,” *Science* **326** (Oct., 2009) 113, [arXiv:0906.0831 \[cond-mat.mes-hall\]](#).

- [45] N. A. Masluk, I. M. Pop, A. Kamal, Z. K. Mineev, and M. H. Devoret, “Implementation of low-loss superinductances for quantum circuits,” *Phys. Rev. Lett.* **109** (June, 2012) 137002, [arXiv:1206.2964](#) [cond-mat.supr-con].
- [46] M. T. Bell, I. A. Sadovskyy, L. B. Ioffe, A. Y. Kitaev, and M. E. Gershenson, “Quantum Superinductor with Tunable Nonlinearity,” *Physical Review Letters* **109** no. 13, (Sept., 2012) 137003, [arXiv:1206.0307](#) [cond-mat.supr-con].
- [47] S. Bravyi and A. Kitaev, “Universal quantum computation with ideal Clifford gates and noisy ancillas,” *Phys. Rev. A* **71** no. 2, (Feb, 2005) 022316.
- [48] P. Aliferis and J. Preskill, “Fault-tolerant quantum computation against biased noise,” *Phys. Rev. A* **78** no. 5, (Nov, 2008) 052331.
- [49] P. Aliferis, F. Brito, D. P. Di Vincenzo, J. Preskill, M. Steffen, and B. M. Terhal, “Fault-tolerant computing with biased-noise superconducting qubits: a case study,” *New Journal of Physics* **11** no. 1, (Jan., 2009) 013061, [arXiv:0806.0383](#) [quant-ph].
- [50] S. Coleman, “The double well done doubly well,” in *Aspects of Symmetry: Selected Erice Lectures*, pp. 341–344. Cambridge University Press, 1985.
- [51] E. E. Cummings, *No Thanks*. New York: Golden Eagle, 1935.
- [52] D. Gottesman, “An introduction to quantum error correction and fault-tolerant quantum computation,” in *Quantum Information Science and Its Contributions to Mathematics, Proceedings of Symposia in Applied Mathematics*, vol. 68, p. 13. 2009.
- [53] P. W. Shor, “Scheme for reducing decoherence in quantum computer memory,” *Phys. Rev. A* **52** no. 4, (Oct, 1995) R2493–R2496.
- [54] D. Bacon, “Operator quantum error-correcting subsystems for self-correcting quantum memories,” *Physical Review A* **73** no. 1, (Jan, 2006) 012340, [arXiv:quant-ph/0506023](#).
- [55] J. Napp and J. Preskill, “Optimal Bacon-Shor codes,” *ArXiv e-prints* (Sept., 2012) , [arXiv:1209.0794](#) [quant-ph].
- [56] P. Aliferis and A. W. Cross, “Subsystem Fault Tolerance with the Bacon-Shor Code,” *Physical Review Letters* **98** no. 22, (June, 2007) 220502, [arXiv:quant-ph/0610063](#).

- [57] A. W. Cross, D. P. DiVincenzo, and B. M. Terhal, “A comparative code study for quantum fault-tolerance,” *ArXiv e-prints* (Nov., 2007) , [arXiv:0711.1556 \[quant-ph\]](#) .
- [58] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, “Topological quantum memory,” *Journal of Mathematical Physics* **43** (Sept., 2002) 4452–4505, [arXiv:quant-ph/0110143](#).
- [59] R. Raussendorf, J. Harrington, and K. Goyal, “Topological fault-tolerance in cluster state quantum computation,” *New Journal of Physics* **9** (June, 2007) 199, [arXiv:quant-ph/0703143](#).
- [60] A. G. Fowler, A. M. Stephens, and P. Groszkowski, “High-threshold universal quantum computation on the surface code,” *Phys. Rev. A* **80** no. 5, (Nov., 2009) 052312, [arXiv:0803.0272 \[quant-ph\]](#) .
- [61] I. Gourlay and J. F. Snowdon, “Concatenated coding in the presence of dephasing,” *Phys. Rev. A* **62** no. 2, (Jul, 2000) 022308.
- [62] Z. W. E. Evans, A. M. Stephens, J. H. Cole, and L. C. L. Hollenberg, “Error correction optimisation in the presence of X/Z asymmetry,” *ArXiv e-prints* (Sept., 2007) , [arXiv:0709.3875 \[quant-ph\]](#) .
- [63] A. M. Stephens, Z. W. E. Evans, S. J. Devitt, and L. C. L. Hollenberg, “Asymmetric quantum error correction via code conversion,” *Phys. Rev. A* **77** no. 6, (June, 2008) 062335, [arXiv:0708.3969 \[quant-ph\]](#) .
- [64] I. L. Chuang and M. A. Nielsen, “Prescription for experimental determination of the dynamics of a quantum black box,” *Journal of Modern Optics* **44** (Nov., 1997) 2455–2467, [arXiv:quant-ph/9610001](#).
- [65] G. M. D’Ariano and P. Lo Presti, “Quantum Tomography for Measuring Experimentally the Matrix Elements of an Arbitrary Quantum Operation,” *Physical Review Letters* **86** (May, 2001) 4195–4198, [arXiv:quant-ph/0012071](#).
- [66] J. B. Altepeter, D. Branning, E. Jeffrey, T. C. Wei, P. G. Kwiat, R. T. Thew, J. L. O’Brien, M. A. Nielsen, and A. G. White, “Ancilla-Assisted Quantum Process Tomography,” *Physical Review Letters* **90** no. 19, (May, 2003) 193601, [arXiv:quant-ph/0303038](#).

- [67] M. Mohseni, A. T. Rezakhani, and D. A. Lidar, “Quantum-process tomography: Resource analysis of different strategies,” *Phys. Rev. A* **77** no. 3, (Mar., 2008) 032322, [arXiv:quant-ph/0702131](#).
- [68] H. Khoon Ng and P. Mandayam, “A simple approach to approximate quantum error correction based on the transpose channel,” *ArXiv e-prints* (Sept., 2009) , [arXiv:0909.0931 \[quant-ph\]](#).
- [69] M. Ben-Or, D. Gottesman, and A. Hassidim, “Quantum Refrigerator,” *ArXiv e-prints* (Jan., 2013) , [arXiv:1301.1995 \[quant-ph\]](#).
- [70] A. M. Meier, B. Eastin, and E. Knill, “Magic-state distillation with the four-qubit code,” *ArXiv e-prints* (Apr., 2012) , [arXiv:1204.4221 \[quant-ph\]](#).
- [71] S. Bravyi and J. Haah, “Magic-state distillation with low overhead,” *Phys. Rev. A* **86** no. 5, (Nov., 2012) 052329, [arXiv:1209.2426 \[quant-ph\]](#).
- [72] C. Jones, “Multilevel distillation of magic states for quantum computing,” *Phys. Rev. A* **87** no. 4, (Apr., 2013) 042305, [arXiv:1210.3388 \[quant-ph\]](#).
- [73] J. L. Borges, “On Exactitude in Science,” *Los Anales de Buenos Aires* **1** no. 3, (1946) .
- [74] P. Aliferis and J. Preskill, “Fibonacci scheme for fault-tolerant quantum computation,” *Phys. Rev. A* **79** no. 1, (Jan., 2009) 012332, [arXiv:0809.5063 \[quant-ph\]](#).
- [75] S. Bravyi, “Topological Quantum Order and Quantum Error Correcting Codes,” in *Simons Institute Workshop on Quantum Hamiltonian Complexity*. 2013.
- [76] B. Eastin and E. Knill, “Restrictions on Transversal Encoded Quantum Gate Sets,” *Physical Review Letters* **102** no. 11, (Mar., 2009) 110502, [arXiv:0811.4262 \[quant-ph\]](#).
- [77] S. Bravyi and R. König, “Classification of Topologically Protected Gates for Local Stabilizer Codes,” *Physical Review Letters* **110** no. 17, (Apr., 2013) 170503, [arXiv:1206.1609 \[quant-ph\]](#).
- [78] T. Jochym-O’Connor, Y. Yu, B. Helou, and R. Laflamme, “The robustness of magic state distillation against errors in Clifford gates,” *ArXiv e-prints* (May, 2012) , [arXiv:1205.6715 \[quant-ph\]](#).

- [79] C. Jones, “Low-overhead constructions for the fault-tolerant Toffoli gate,” *Phys. Rev. A* **87** no. 2, (Feb., 2013) 022328, [arXiv:1212.5069](#) [quant-ph].
- [80] C. Jones, “Distillation protocols for Fourier states in quantum computing,” *ArXiv e-prints* (Mar., 2013) , [arXiv:1303.3066](#) [quant-ph].
- [81] B. Eastin, “Distilling one-qubit magic states into Toffoli states,” *Phys. Rev. A* **87** no. 3, (Mar., 2013) 032321, [arXiv:1212.4872](#) [quant-ph].
- [82] A. Paetznic and B. W. Reichardt, “Universal fault-tolerant quantum computation with only transversal gates and error correction,” *ArXiv e-prints* (Apr., 2013) , [arXiv:1304.3709](#) [quant-ph].
- [83] E. Knill, “Fault-Tolerant Postselected Quantum Computation: Schemes,” *ArXiv e-prints* (Feb., 2004) , [arXiv:quant-ph/0402171](#).
- [84] A. M. Steane, “Fast fault-tolerant filtering of quantum codewords,” *ArXiv e-prints* (Feb., 2002) , [arXiv:quant-ph/0202036](#).
- [85] A. Paetznic and B. W. Reichardt, “Fault-tolerant ancilla preparation and noise threshold lower bounds for the 23-qubit Golay code,” *ArXiv e-prints* (June, 2011) , [arXiv:1106.2190](#) [quant-ph].
- [86] R. P. Feynman, *The Character of Physical Law: The 1964 Messenger Lectures*. MIT Press, 1967.