Topics in Fault-Tolerant Quantum Computation

by

Hillary Dawkins

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Science
in
Physics (Quantum Information)

Waterloo, Ontario, Canada, 2017

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

This thesis contains the results of research projects broadly related to the pursuit of universal fault-tolerant quantum computation. We are interested in questions such as which resources are required to achieve a quantum computational advantage, how to utilize such resources in practice, and how to characterize and control noise throughout a computation.

In part 1, we cover results pertaining to magic state distillation, one of the leading approaches to implementing universal fault-tolerant quantum computation. Early work in this area focused on determining the region of distillable states for qubit protocols, yet comparatively little is known about which states can be distilled and with what distillable region for $d > 2$. In the first project, we focus on $d = 3$ and present new four-qutrit distillation schemes that improve upon the known distillable region, and achieve distillation tight to the boundary of undistillable states for some classes of states. As a consequence of recent results, this implies that there is a family of quantum states that enable universality if and only if they exhibit contextuality with respect to stabilizer measurements. We also identify a new routine whose fixed point is a magic state with maximal sum-negativity i.e., it is maximally non-stabilizer in a specific sense. In the second project, we return to $d = 2$ and present a number of interesting new distillation routines based on small codes. Many of these distill noisy states right up to the boundary of the known undististillable region, while some distill toward non-stabilizer states that have not previously been considered.

In part 2, we address the issue of noise characterization. The characterization of noise in a quantum system serves as a foundation towards meeting two important goals. Firstly, we may wish to know about specific sources of noise in order to adapt or correct errors as much as possible. Secondly, we would like a way to verify that noise is below some threshold value in order to meet the criteria of threshold theorems. We consider the case where noise is approximately known to coincide with the generalized damping channel (encompassing the common intrinsic processes of amplitude damping and dephasing), but may contain additional unknown noise sources. We provide methods to accurately obtain the noise parameters, and compare with the results of a randomized benchmarking experiment. Using this information, we show how to make meaningful statements about fault-tolerance threshold theorems by considering the diamond distance.

Acknowledgements

First and foremost, I would like to thank my supervisor Joseph Emerson for his guidance and support throughout my time as a graduate student. I would also like to extend a thank you to my committee members, Raymond Laflamme and David Cory, for their valuable insights. Special thanks goes to Mark Howard and Joel Wallman for including me in their research. I am grateful for the patience and encouragement they have shown me throughout our collaborations. I would also like to thank all the members of the Emerson group for their company and feedback.

I thank the research groups of Rainer Blatt and David Poulin, for their generous hospitality while hosting my visits with them.

I thank my previous advisors, Alex Gezerlis and Paul Garrett, for their continued advice and encouragement.

Last but not least, I thank my friends and family for their continuous support. Thank you to my parents, Gary, Sophie, Lacy, Scott, Jenny and Joanna.

Dedicated to my grandmother, Hazel.

Table of Contents

List of Tables

List of Figures

Chapter 1

Outline and Summary of Results

A major obstacle in the path to building a scalable and robust quantum computer is overcoming the loss or corruption of quantum information due to unwanted interactions with the environment. We will present methods for overcoming such obstacles and achieving universal fault-tolerant quantum computation, as well as methods for characterizing the potential threats. We begin by reviewing representations of quantum states and channels (Chapter 2), as well as quantum error-correcting codes (Chapter 3).

In Chapter 4, we provide background information on one of the leading approaches to achieving universal fault-tolerant quantum computation: magic state distillation. The distillation protocol was first proposed by Knill [1] and Bravyi and Kitaev [2], where it was shown that impure input states may be consumed in an iterative procedure to produce a pure non-stabilizer state. Furthermore this non-stabilizer state can be used to implement a gate outside the Clifford gate set, and therefore universal fault-tolerant quantum computation is achieved using this scheme. Which types of quantum states are suitable as input states to this procedure then becomes a question of foundational interest, since any resource these states possess enables universality. For qubits, undistillability of a state is implied by it being a stabilizer state by the Gottesman-Knill theorem [3]. Therefore, showing that any non-stabilizer state can be distilled via some magic state distillation routine implies that this is the resource to promote quantum universality.

The idea of magic state distillation has since been extended to higher dimensional systems by Anwar, Campbell, and Browne [4, 5, 6]. Similar to qubits, there is a region of states known to be undistillable. For prime dimensions, the region of possibly distillable states was recently shown to coincide exactly with the region of states which exhibit state dependent contextuality with respect to stabilizer measurements [7]. It has been conjec-

tured that contextuality is the resource to enable universality in this regime [8]. However, proving such a conjecture would require showing that any state outside the undistillable region is indeed distillable by a magic state distillation routine.

In Chapter 5, we present our contribution towards this goal (based on [9], a collaboration with Mark Howard). We present novel magic state distillation protocols which use very few input states per round of distillation, and improve the region of suitable input states. Furthermore the improved distillable region saturates a bound in that we achieve distillation tight to the boundary of provably undistillable states for the first time in a three-dimensional system. This implies that there is a family of quantum states that allow universality if and only if they exhibit contextuality with respect to stabilizer measurements. The tightness result is highlighted geometrically through a certain parameterization of qutrit state space, which may be of independent interest. In addition to the main result, we present a family of a magic state distillation protocols and corresponding fixed points which can help fill out the landscape of known distillation procedures. Currently very little is known about what makes a good magic state distillation routine, or given some routine, which state will be distilled. Future work in this area can make use of the vast library of routines and corresponding fixed points we have collected. Our distilled states include a special fixed point of interest due to it being a maximally non-stabilizer state, which was not previously known to be directly distillable. In Chapter 6, we further increase the landscape of known distillation routines for qubit protocols (based on [10], a collaboration with Mark Howard). While transversality of a non-Clifford gate within a code often leads to efficient distillation routines, it appears to not be a necessary condition. We have examined a number of small stabilizer codes and highlight a handful which display interesting, albeit inefficient, distillation behaviour.

In the remaining chapters, we turn our attention to another fundamental component of universal fault-tolerant quantum computation: noise characterization. In Chapter 7, we review some of the most common noise processes associated with unwanted interactions with the environment and how they are measured in practice. Amplitude damping describes the thermal equilibration of ground and excited states, while phase damping describes the loss of coherences. The rates of amplitude damping and phase damping are measured via inversion recovery and Ramsey experiments respectively. We also introduce a generalized notion of these channels, which combines both processes and allows for relaxation at a finite temperature.

In Chapter 8 (based on a collaboration with Joel Wallman and Joseph Emerson), we show how the parameters of this generalized model can be estimated, accounting for state preparation and measurement errors, as well as small deviations from generalized damping. We then show how the resulting estimates can be compared with the model-independent

quantities from randomized benchmarking to rigorously establish how close the noise is to a generalized damping channel. Furthermore, relation to fault-tolerance threshold theorems often requires knowledge of the diamond distance of the noise channel from the identity. We derive bounds on the diamond distance for channels close to generalized damping in terms of the experimental parameters, which will allow experimentalists to easily check if threshold criteria are satisfied.

Chapter 2

Quantum Information Processing

Processing quantum information involves:

i) encoding of information into quantum *states*, described by ρ, a positive linear operator existing on a complex Hilbert space

ii) manipulation of information achieved by applying *gates*, ideally described by a unitary operator on the Hilbert space. We will often refer to the evolution of a state more generally as a *channel*, to encompass the application of ideal gates as well as noise processes. A quantum channel $\mathcal{E}(p)$ is said to be:

ii.a) *completely positive* (CP) if $\mathcal{E} \otimes \mathbb{1}(\rho)$ is positive whenever ρ is positive,

ii.b) *trace preserving* (TP) if $\mathrm{Tr}(\mathcal{E}(\rho)) = \mathrm{Tr}(\rho)$ for all ρ,

and ii.c) *unital* if $\mathcal{E}(\mathbb{1}/d) = \mathbb{1}/d$.

In this chapter, we outline the formalism we use to describe these operations. Both states and channels can be conveniently represented using the Pauli operators

$$D_{(x|z)} = \omega^{2^{-1}xz} X^x Z^z, x, z \in \mathbb{Z}_d \ , \tag{2.1}$$

for a d-dimensional system, where X and Z are defined by

$$X \left|j\right\rangle = \left|j+1\right\rangle , \qquad Z \left|j\right\rangle = \omega^j \left|j\right\rangle \tag{2.2}$$

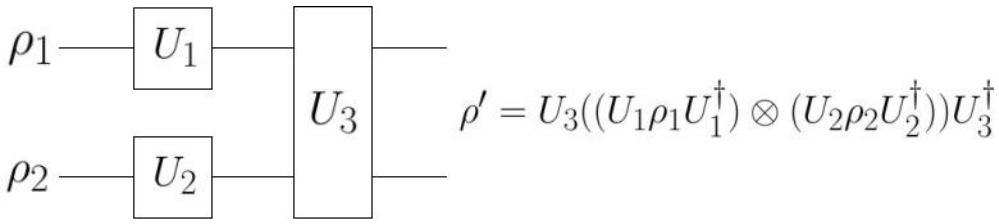


Figure 2.1: Example of a circuit diagram. Reading from left to right, we see the state is initially prepared as $\rho_1 \otimes \rho_2$, undergoes a transformation by the single qubit gates U_1, U_2 followed by the 2-qubit gate U_3, and is output in the state ρ'.

and $\omega = e^{2\pi i/d}$ is the d-th root of unity. For qubits, this reduces to the standard

$$\mathbb{1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \; X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \; Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \; Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{2.3}$$

Diagrammatically, we can describe a quantum computation by a circuit diagram as shown in Fig. 2.1.

2.1 Representations of states

Here we present two useful ways to represent states which will be referenced throughout this thesis.

2.1.1 The Bloch sphere

Any pure qubit state can be expressed as

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle \tag{2.4}$$

with respect to an orthonormal basis $\{|0\rangle, |1\rangle\}$ where $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi)$, so that the state is uniquely described by a point on the unit sphere. Mixed states, convex combinations of pure states, are described by points within the interior of the sphere. Both pure and mixed states can be described by a density matrix

$$\rho = \frac{1}{2}(\mathbb{1} + xX + yY + zZ) \tag{2.5}$$

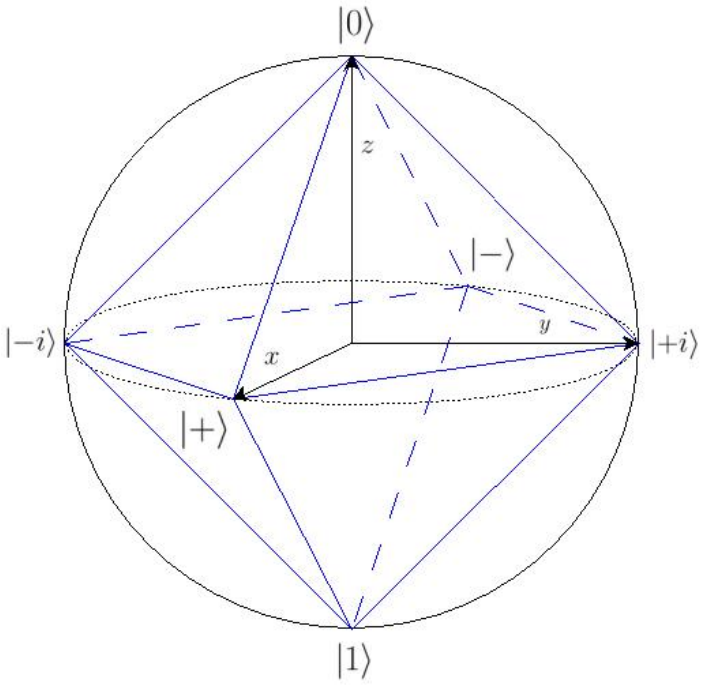


Figure 2.2: The Bloch sphere for qubits. The stabilizer octahedron is shown in blue.

where (x, y, z) are the Bloch coordinates and give the position of the state within the Bloch sphere. The single qubit stabilizer states are the eigenstates of the Pauli operators, and are usually denoted as $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle, |+i\rangle, |-i\rangle\}$. Any mixed state which is a convex combination of stabilizer states lies within the stabilizer octahedron, as shown in Fig. 2.2. When discussing qubits, we will often describe states using Bloch coordinates.

2.1.2 Discrete Wigner representation

The discrete Wigner representation [12, 13] of a single-qudit state ρ is a quasi-probability distribution over a $d \times d$ grid. The elements of the grid are given by

$$W_\rho(x, z) = \frac{1}{d} \operatorname{Tr}(A_{x,z}\rho) \tag{2.6}$$

where $A_{x,z}$ are the phase point operators

$$A_{x,z} = \frac{1}{d} D_{(x|z)} \left(\sum_{x',z'} D_{(x'|z')} \right) D^\dagger_{(x|z)}. \tag{2.7}$$

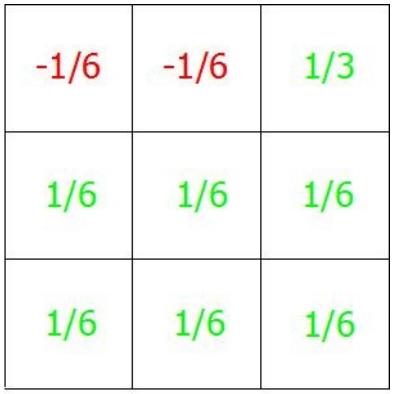


Figure 2.3: Discrete Wigner representation of the state $|\psi\rangle = \frac{1}{\sqrt{6}}(-|0\rangle + 2|1\rangle - |2\rangle)$.

For $d > 2$, the subset of quantum states that have nonnegative entries in all d^2 positions of the discrete Wigner function are known to be classically simulable [14]. Geometrically, this subset of states is prescribed by the Wigner polytope

$$\text{Wigner polytope} := \{\rho : \text{Tr}(\rho A_{x,z}) \geq 0, \quad x, z \in \mathbb{Z}_d\}.$$

It should be noted that the stabilizer states form a strict subset of the Wigner polytope. The sum-negativity [8] of a state,

$$sn(\rho) = \sum_{x,z:W_\rho(x,z)<0} |W_\rho(x,z)|, \tag{2.8}$$

is an operationally useful quantification of how far from the Wigner polytope a state is. As an example, consider the qutrit state $|\psi\rangle = \frac{1}{\sqrt{6}}(-|0\rangle + 2|1\rangle - |2\rangle)$. The Wigner representation is shown in Fig. 2.3, from which we can see $sn(|\psi\rangle) = \frac{1}{3}$, and therefore this state is not classically simulable. When discussing qudits of dimension greater than two, we will often refer to the Wigner representation.

2.2 Representations of channels

Here we present channel representations which will be used to describe quantum processes throughout this thesis. This discussion is based on [15, 16].

2.2.1 The Kraus-operator representation of quantum channels

The evolution of a quantum state may be described by the superoperator $\mathcal{E}(\rho)$ acting on ρ for a given time. A CPTP map admits an operator-sum decomposition

$$\mathcal{E}(\rho) = \sum_k A_k \rho A_k^\dagger \tag{2.9}$$

where $\{A_k\}$ is the set of Kraus operators. The Kraus operators must satisfy

$$\sum_k A_k^\dagger A_k = \mathbb{1} \tag{2.10}$$

in order for $\mathcal{E}(\rho)$ to be trace-preserving. Additionally, if $\mathcal{E}(\rho)$ is unital, we have the extra condition

$$\sum_k A_k A_k^\dagger = \mathbb{1}. \tag{2.11}$$

The maximum number of Kraus operators needed to describe a CP map acting on d-dimensional quantum states is d^2.

2.2.2 The Pauli-Liouville representation of quantum channels

A matrix representation of a superoperator, also called the process matrix, $\Lambda(\mathcal{E})$ is a $d^2 \times d^2$ matrix with elements

$$\Lambda_{ij}(\mathcal{E}) = \mathrm{Tr}(B_i^\dagger \mathcal{E}(B_j)) \tag{2.12}$$

where $\{B_i\}$ is a trace-orthonormal basis. The action of the channel is then given by

$$|\mathcal{E}(\rho)\big) = \Lambda |\rho\big) \tag{2.13}$$

where $|\rho\big)$ is the column vector representation of a state ρ with entries

$$|\rho\big)_i = \mathrm{Tr}(B_i^\dagger \rho). \tag{2.14}$$

The Pauli-Liouville representation of a channel is the process matrix given by the choice of basis

$$\{B_i\} = P_d/\sqrt{d} \tag{2.15}$$

which constrains all elements of $\Lambda(\mathcal{E})$ to be real. For trace-preserving channels, we can write $\Lambda(\mathcal{E})$ in block diagonal form

$$\Lambda(\mathcal{E}) = \begin{bmatrix} 1 & 0 \\ \Lambda_n & \Lambda_u \end{bmatrix} \tag{2.16}$$

where we have fixed $B_1 = \mathbb{1}/\sqrt{d}$. Λ_n and Λ_u are called the non-unital and unital parts of $\Lambda(\mathcal{E})$ respectively, and $\Lambda_n = 0$ for unital maps.

2.2.3 The Choi matrix

Another matrix representation of the channel $\mathcal{E}$ is given by the Choi matrix $J(\mathcal{E})$

$$J(\mathcal{E})_{ij} = \mathcal{E}(B_{ij}) \tag{2.17}$$

where each "element" $J(\mathcal{E})_{ij}$ is a $d \times d$ matrix, and $\{B_{ij}\}$ is the standard computational basis, $B_{ij} = |i\rangle\langle j|$. More explicitly,

$$J(\mathcal{E}) = \begin{bmatrix} \mathcal{E}(|0\rangle\langle 0|) & \mathcal{E}(|0\rangle\langle 1|) \\ \mathcal{E}(|1\rangle\langle 0|) & \mathcal{E}(|1\rangle\langle 1|) \end{bmatrix} \tag{2.18}$$

for $\mathcal{E}$ acting on a single qubit. The Choi matrix representation is unique and must be positive semi-definite if $\mathcal{E}$ is a CP map. Furthermore we can diagonalize the Choi matrix to find a minimal Kraus-operator representation

$$A_k = \sqrt{\lambda_k} V_k \tag{2.19}$$

where V_k are the eigenvectors of $J(\mathcal{E})$ with associated eigenvalues λ_k.

2.2.4 Master equations

The previous sections describe discrete time representations of channels, however it will sometimes be more convenient to give a continuous time description of evolution as governed by a master equation. The standard form of the evolution is given by

$$\frac{\partial \rho}{\partial t} = -i[H, \rho] + \sum_{i,j=2}^{4} A_{i,j}(B_i \rho B_j^\dagger - \frac{1}{2}\{B_j^\dagger B_i, \rho\}) \tag{2.20}$$

where

$$H = \frac{1}{2i}(F^\dagger - F), \quad F = \frac{1}{2}\sum_{i=2}^{4} A_{1,i}B_i \tag{2.21}$$

and $\{B_i\}$ is a trace-orthonormal basis with the choice $B_1 = \mathbb{1}/\sqrt{d}$. We can see that the evolution is completely specified by providing the coefficient matrix A, which is Hermitian and positive [16]. Since A is Hermitian, it may be diagonalized to give the Lindblad form of the evolution

$$\frac{\partial \rho}{\partial t} = -i[H, \rho] + \mathcal{D}(\rho) \tag{2.22}$$

where H describes the unitary evolution of ρ, and $\mathcal{D}(\rho)$ gives the dissipative term

$$\mathcal{D}(\rho) = \sum_{i=2}^{d^2} \lambda_i (V_i \rho V_i^\dagger - \frac{1}{2}\{V_i^\dagger V_i, \rho\}), \quad \lambda_i \geq 0. \tag{2.23}$$

2.3 Measures of noise

Finally, we need some way to measure the level of noise on a quantum system or the amount of error in our desired operations. In fact there are several ways to define noise, and we spend the remainder of this section describing the measures most relevant to our results.

The difference between two quantum states, ρ and σ, may be characterized by the state fidelity

$$F(\rho, \sigma) = \mathrm{Tr}\left(\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)^2. \tag{2.24}$$

In a similar fashion, the fidelity between channels $\mathcal{E}_1$ and $\mathcal{E}_2$ is defined as

$$\mathcal{F}_\psi(\mathcal{E}_1, \mathcal{E}_2) = F(\mathcal{E}_1(\psi), \mathcal{E}_2(\psi)) \tag{2.25}$$

which is dependent on the input state ψ. To obtain a state-independent measure of channel distance, we average over all pure states using the Fubini-Study measure

$$\mathcal{F}(\mathcal{E}_1, \mathcal{E}_2) = \int d\mu_{FS}(\psi) F(\mathcal{E}_1(\psi), \mathcal{E}_2(\psi)). \tag{2.26}$$

In our discussions, we wish to characterize the difference between an unwanted noise process $\mathcal{E}$ and the identity operation $\mathbb{1}$, $\mathcal{F}(\mathcal{E}) := \mathcal{F}(\mathcal{E}, \mathbb{1})$. Depending on the channel representation of $\mathcal{E}$, the computation of the average fidelity can take several forms. Given the Kraus operators $\{A_k\}$,

$$\mathcal{F}(\mathcal{E}) = \frac{\sum_k \mathrm{Tr}(A_k)\mathrm{Tr}(A_k^\dagger) + d}{d^2 + d}. \tag{2.27}$$

Equivalently, using the Pauli-Liouville representation,

$$\mathcal{F}(\mathcal{E}) = \frac{\mathrm{Tr}(\Lambda(\mathcal{E})) + d}{d^2 + d}. \tag{2.28}$$

Finally when given a continuous time description of the evolution, it may be more convenient to compute

$$\mathcal{F}(\mathcal{E}) = \frac{1}{N}\sum_{i=1}^{N} \langle\psi_i| \mathcal{E}(\psi_i) |\psi_i\rangle \tag{2.29}$$

where $|\psi_i\rangle$ are the pure stabilizer states, using the fact that these states form a state 2-design.

While the average fidelity is a useful quantity, many threshold theorems actually depend on the "worst-case error" as characterized by the diamond distance between $\mathcal{E}$ and the identity operation, $D(\mathcal{E})_\diamond = \frac{1}{2}||\mathcal{E} - \mathbb{1}||_\diamond$ where

$$||\Delta||_\diamond = \sup_\psi ||\mathbb{1} \otimes \Delta(\psi)||_1. \tag{2.30}$$

The diamond distance has the nice property that it is invariant under enlargement of the Hilbert space, and therefore the supremum in the above may be taken over pure states. The diamond norm may be computed using semi-definite programming techniques; for more details see [21]. For an analytic expression in the case where $\mathcal{E}$ is a Pauli channel, see [22].

Chapter 3

Quantum Error Correction and Fault Tolerance

In this section, we discuss the major themes that unify the two parts of this thesis: quantum error correction and fault tolerance. The strange properties of quantum mechanics that allow for an advantage over classical computation turn out to be a double-edged sword when trying to account for errors. For one, errors take on an infinite range of possibilities in contrast to the single bit-flip error in the classical setting. The detection and classification of errors is also problematic due to decoherence caused by measurement. Quantum error correction protocols tell us how to apply corrections given only certain information. However, it is always possible that our correction operations will be faulty. Fault-tolerant schemes implement error correction combined with other tools in such a way that errors do not propagate badly. In combination with error correction, the *fault-tolerance threshold theorem* [19, 20] ensures that meaningful quantum computation is achievable. Loosely stated, the threshold theorem guarantees that a certain amount of noise can be corrected or accounted for as long as we are promised that the noise is below some threshold value. We begin this section by introducing the Clifford group, an important class of operations in the context of fault tolerance. We then discuss the stabilizer formalism (adapted from [17]) which we frequently use to describe error-correcting codes. More specifically, we use code-word stabilized codes to describe the majority of the results presented in Part 1.

3.1 The Clifford group

The Clifford group [11] is the set of unitary operations that map Pauli operators to Pauli operators under conjugation. The elements (for the single-qudit case) can be written as

$$C = D_{(x|z)} U_F \tag{3.1}$$

$$U_F = \begin{cases} \frac{1}{\sqrt{d}} \sum_{j,k=0}^{d-1} \omega^{2^{-1}\beta^{-1}(\alpha k^2 - 2jk + \delta j^2)} |j\rangle \langle k|, \ \beta \neq 0 \\ \sum_{k=0}^{d-1} \omega^{2^{-1}\alpha\gamma k^2} |\alpha k\rangle \langle k|, \quad \beta = 0 \end{cases} \tag{3.2}$$

where $F = \left(\begin{smallmatrix} \alpha & \beta \\ \gamma & \delta \end{smallmatrix}\right)$ is a symplectic matrix and all variables are elements of $\mathbb{Z}_d = \{0, 1, \ldots, d-1\}$. We denote the d-level Pauli and Clifford groups acting on n qudits as $\mathcal{P}_d^n$ and $\mathcal{C}_d^n$ respectively. As we will see, the Clifford group plays an important role in the theory of fault-tolerant quantum computation. The Gottesman-Knill theorem [3] states that Clifford operations can be simulated efficiently using a classical computer. However the set of Clifford operations do not form a dense set in $\mathcal{U}(d^n)$ and therefore do not constitute a set of universal operations. Part 1 of this thesis is dedicated to the study of magic state distillation, one possible path to achieving a universal gate set.

3.2 The stabilizer formalism

The basic idea of quantum error correction is to encode information into a protected subspace of a larger Hilbert space. A quantum error-correcting code tells us how information should be encoded as well as which operations can be performed safely while leaving the subspace invariant. Within the stabilizer formalism, quantum error-correcting codes are described by the *stabilizer* of the code. The stabilizer $\mathcal{S}$ is comprised of Pauli operators which leave all codewords invariant (i.e. are $+1$ eigenstates). Given an Abelian group $\mathcal{S} \subseteq \mathcal{P}_d^n$ with $-\mathbb{1} \notin \mathcal{S}$, the codespace is defined as

$$\mathcal{T}(\mathcal{S}) = \{|\psi\rangle : M |\psi\rangle = |\psi\rangle \, \forall M \in \mathcal{S}\}. \tag{3.3}$$

Generally, a code will be described by its *generators*, a non-unique set of operators which generate the stabilizer group. In the context of error correction, the generators are Pauli measurements which are used to determine the error syndrome. As we will see in Chapter 4, the generators also play a functional role in the creation of magic states. The generators can be written in binary symplectic notation so that $G = (X^{x_1} \otimes X^{x_2} \otimes \ldots \otimes X^{x_n})(Z^{z_1} \otimes Z^{z_2} \otimes \ldots \otimes Z^{z_n})$ becomes the vector $(x_1, x_2, \ldots, x_n | z_1, z_2, \ldots, z_n)$. The complete set of generators can then be compactly described by the matrix $(X|Z)$.

Given the generators, we can look at properties of the code such as the size, distance and codespace projector. The *size* of a code k is the number of logical qudits, and is given by $k = n - r$ where r is the number of generators. The *distance* of a code tells us about the error correction capability. A code of distance $D = 2t + 1$ is able to correct t local errors. The distance is given by

$$D = \min\{\text{weight } M : M \in \mathcal{N}(\mathcal{S}) \backslash \mathcal{S}\} \tag{3.4}$$

where $\mathcal{N}(\mathcal{S})$ in the normalizer of $\mathcal{S}$ and global phase errors have been ignored (see [17] for the proof and further discussion). In a short-hand notation, it is common to describe a code as being an $[[n, k, D]]_d$ code. Finally, the *codespace projector* is an operator that projects arbitrary states in the enlarged Hilbert space onto the codespace (representing logical information). The projector onto the codespace of $\mathcal{S}$ in terms of generators $\{G_i\}$ is

$$\Pi_{\mathcal{S}} = \frac{1}{2^r} \prod_{i=1}^{r} (\mathbb{1} + G_i) \tag{3.5}$$

which will force states to be in the $+1$ eigenspace of all generators simultaneously. Consequently, if we add the projection onto the $+1$ eigenspace of the logical Z operator, we will arrive at the logical state $|0_L\rangle$:

$$|0_L\rangle\langle 0_L| = \frac{1}{2} \Pi_{\mathcal{S}} (\mathbb{1} + Z_L). \tag{3.6}$$

3.3 Code-word stabilized codes

We now describe an important class of stabilizer codes: code-word stabilized (CWS) codes (see [18]). In this representation, the code is given by a graph state written as an adjacency matrix, $\Gamma \in \mathbb{Z}_2^{n \times n}$, as well as a classical codeword, $w \in \mathbb{Z}_2^n$. We review the construction explicitly for qubits, which is the case we make use of throughout this thesis. An n-qubit stabilizer code with n generators has $k = 0$ and consequently specifies a 1-dimensional subspace of Hilbert space i.e, a stabilizer state. By applying local Cliffords, a generic stabilizer state $(X|Z)$ can be brought to the form $(\mathbb{1}_n|\Gamma)$ where Γ is the adjacency matrix of the corresponding graph. When stabilizer states take this particular form they are called graph states, sometimes denoted $|\Gamma\rangle$. We may choose the logical basis state $|0_L\rangle$ of our code to be the state

$$|0_L\rangle = |\Gamma\rangle = \sum_{x \in \mathbb{Z}_2^n} i^{x^T \Gamma x} |x\rangle . \tag{3.7}$$

Subsequently the logical $|1_L\rangle$ state is

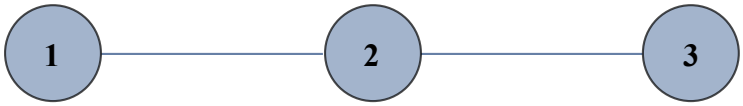
$$|1_L\rangle = Z^w |0_L\rangle = \sum_{x \in \mathbb{Z}_2^n} i^{x^T \Gamma x + 2w^T x} |x\rangle \,, \tag{3.8}$$

where w is the codeword and Z^w denotes the product of Z operators

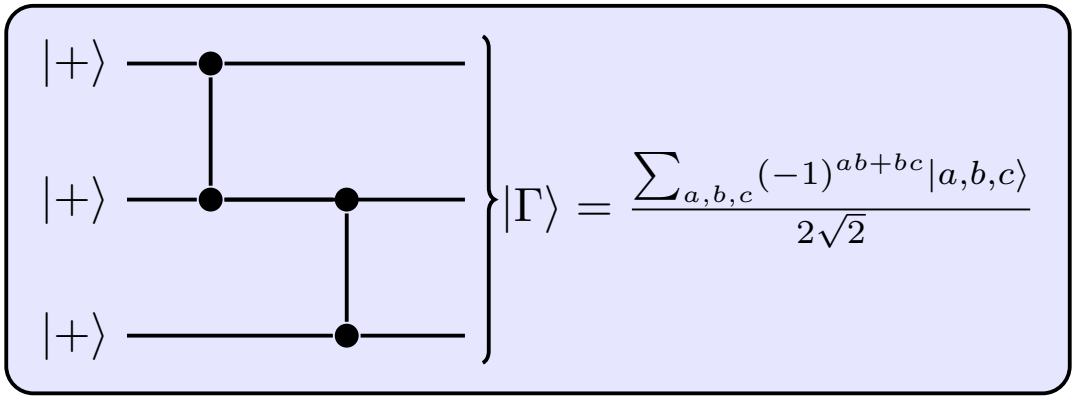
$$Z^w = Z^{w_1} \otimes ... \otimes Z^{w_n}. \tag{3.9}$$

The unitary encoding/decoding associated with these codes is quite straightforward as illustrated in Fig. 3.1b – it involves controlled-Z rotations applied to qubits whose vertices are connected by edges in the graph. A convenient way of visualizing CWS codes is to display the graph and highlight the subset of vertices $1 \leq i \leq n$ such that $w_i = 1$. For example the highlighted graph shown in Fig. 3.2 corresponds to

$$\Gamma = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}, \quad \omega = (1, 1, 0, 0). \tag{3.10}$$



(a) Graph with labelled vertices



(b) The graph state preparation circuit for the above graph

Figure 3.1: Creating the graph state associated with a given graph amounts to performing controlled-Z rotations between qubits whose vertices are connected by an edge in the graph.

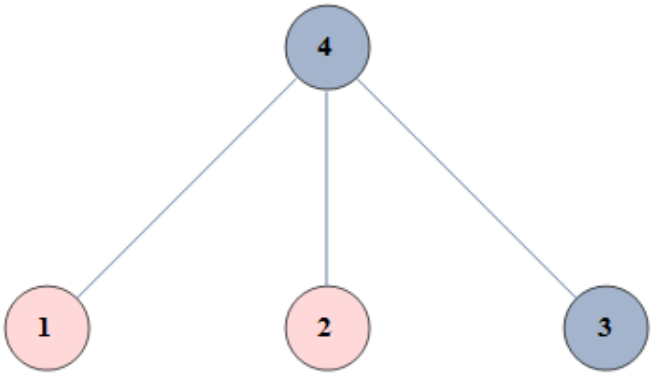


Figure 3.2: An example of a CWS code as described by a highlighted graph.

Chapter 4

Magic State Distillation Background

In the first part of this thesis, we study magic state distillation as a path to universal fault-tolerant quantum computation. We begin by reviewing the relevant background information.

4.1 The need for magic states

One of the main challenges facing the implementation of a large scale quantum computer is the ability to protect quantum information from decoherence, typically introduced by unwanted interactions with the environment. Encoding information into quantum error-correcting codes provides a partial answer [23, 24, 25]. However such encodings normally only allow a limited set of transversal or manifestly fault-tolerant operations, usually the stabilizer operations i.e., Clifford gates, preparation of stabilizer states and Pauli measurements. Stabilizer operations alone do not enable universality [3], which is achieved by adding access to any gate outside the Clifford group. A popular choice is the T-gate (or $\pi/8$-gate) and much work has been done on the subject of T + Clifford circuit synthesis [26, 27]. We define a magic state to be any state which can be used to implement a non-Clifford gate to a desired precision. For example, an equatorial state of the form

$$|\psi\rangle = |0\rangle + e^{i\theta}\,|1\rangle \tag{4.1}$$

implements a rotation

$$R(\theta) = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{bmatrix} \tag{4.2}$$

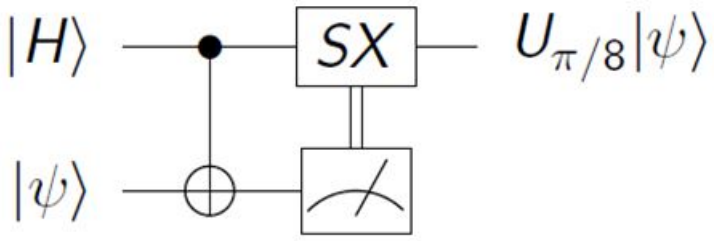


Figure 4.1: The equatorial state $|H\rangle = |0\rangle + e^{i\pi/4}|1\rangle$ is consumed by the state injection circuit to implement the $\pi/8$-gate on an arbitrary input.

via the state injection circuit shown in Fig. 4.1. Although alternatives exist [28, 29, 30], magic state distillation remains a leading implementation of universal fault-tolerant quantum computation. In the following sections we detail the distillation protocol, how to gauge the success of a routine, and types of magic states.

4.2 The distillation protocol

Having established that access to a magic state will allow for universal quantum computation, we turn to the task of obtaining such a state. By definition, a magic state is not a stabilizer state and therefore we cannot assume perfect preparation. Using only Clifford operations, the distillation protocol (first introduced by Knill [1] and Bravyi and Kitaev [2]) consumes noisy versions of the magic state to produce the final state of required purity.

Magic state distillation routines are described in terms of stabilizer error-correcting codes, which in turn are described by a set of generators $\{G_i\}$ consisting of Pauli operators (as described in Chapter 3). The protocol consists of iteratively applying the following steps:

1. Prepare n copies of the input state $\rho_{in}^{\otimes n}$, where n is the size of the code
2. Perform Pauli measurements corresponding to each of the $n-k$ generators G_i, and postselect on the desired outcome
3. Decode the logical qudit of the code onto k physical qudits.
4. Optionally perform a Clifford transformation based on the measurement outcome.

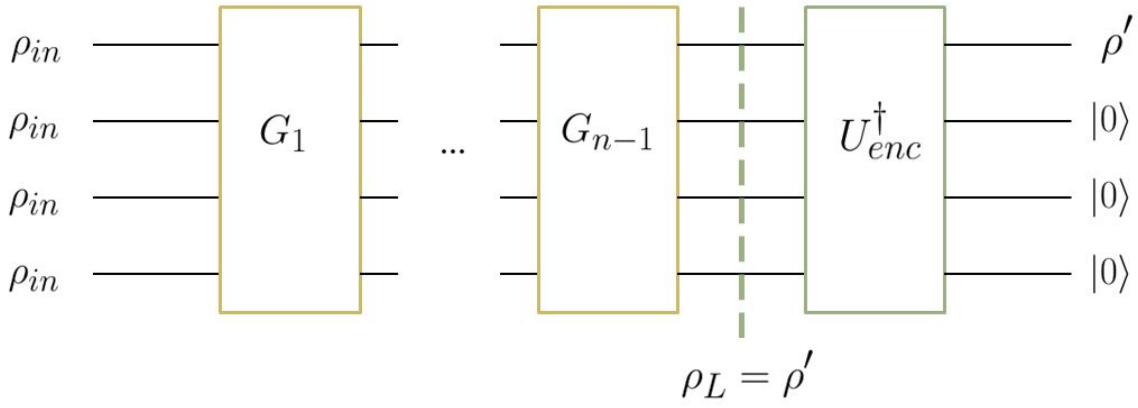


Figure 4.2: One round of distillation taking noisy versions of the magic state ρ_{in} to a higher fidelity output state ρ'.

One round of the protocol is shown in Fig. 4.2. When successful, the output state(s) will be purified in the direction of the target magic state. Typically $k = 1$ and consequently an $[[n, k = 1, D]]$ stabilizer code is used (see [31, 32, 33] for $k \geq 2$ however).

It was shown in [34] that any protocol restricted to Clifford operations cannot produce an output state of higher fidelity than a stabilizer reduction protocol. Therefore we lose no generality by restricting our search to distillation routines based on stabilizer codes. The success of any given routine can be measured in a number of ways, depending on the desired application. In the following section we describe these measures.

4.3 Useful quantities

The magic state distillation protocol describes an iterative process in which an output state is produced by consuming the output state at the previous step, when successful. When post-selecting on the trivial measurement outcome, the output state is given by

$$\rho' = \frac{1}{p_s} \Pi_c \rho^{\otimes n} \Pi_c \tag{4.3}$$

where Π_c is the codespace projector, and p_s is the *success probability*:

$$p_s = \mathrm{Tr}(\Pi_c \rho^{\otimes n}). \tag{4.4}$$

For qubits, one can assume that input states suffer from depolarizing noise and take the form

$$\rho_M = (1-p)\,|M\rangle\langle M| + p\frac{\mathbb{1}_2}{2}\ , \tag{4.5}$$

where $|M\rangle$ is the target magic state, and p is called the *error rate*. This is due to the fact that twirling to the depolarizing axis is a Clifford operation, and therefore may be implemented perfectly. We measure the efficiency of a routine based on the output error rate of a state as a function of the input error rate $p_{out} = f(p)$. The *error suppression* of a routine refers to the rate at which p_{out} tends to zero (i.e. if $p_{out} = c_1 p + \mathcal{O}(p^2)$ then we have linear error suppression).

The *threshold* of a routine refers to the maximum value of p for which distillation is successful (i.e. $p_{out} < p$), which we denote as p^*. For qubits, the relevant value of p^* always occurs along the depolarizing axis of the target magic state. For qudits, when a depolarizing noise model cannot be assumed, we may report a range of threshold values containing the depolarizing threshold (see [4]). The upper limit on any threshold is set by the boundary of undistillable states. A state which is classically simulable cannot possibly be used to achieve universal quantum computation, and therefore all classically simulable states are undistillable. As discussed in section 2.1, this region contains the stabilzer octahedron for qubits and the Wigner polytope for qudits. In both cases we denote the upper limit on the threshold as p_{poly}.

The *yield* [6] of a magic state distillation routine is defined as

$$Y(p, p_{out}) = \prod_{k=1\ldots N} \frac{p_s^{(k)}}{n} \tag{4.6}$$

where N is the total number of iterations needed to obtain p_{out} starting at initial error rate p, and $p_s^{(k)}$ is the probability of success on the k^{th} iteration. This quantity relates to the efficiency/resource overhead of a magic state distillation routine. Alternatively, the efficiency may sometimes be described by the *cost* of a routine, which is simply the inverse of yield, and equal to the total number of consumed noisy states.

4.4 Types of magic states

As previously stated, we define any state enabling the application of a non-Clifford gate to a desired precision to be a magic state. However there are several states that we consider

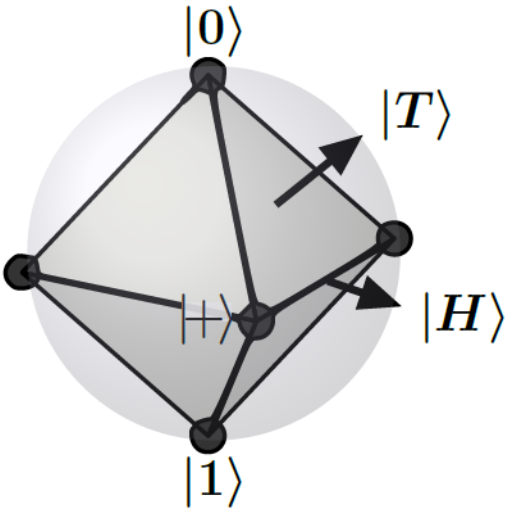


Figure 4.3: The stabilizer octahedron inscribed within the Bloch sphere: Six Pauli eigenstates form the vertices of an octahedron. States within the octahedron are provably undistillable so the best one can hope for is to distill states up to the boundary. The two pure non-stabilizer states singled out here, $|H\rangle$ and $|T\rangle$, were shown to be distillable by Bravyi and Kitaev [2].

to be special based on geometric considerations. In this section we review the common types of magic states that we might hope to distill.

For qubits, there are two special types of magic state as defined in the original proposal [2]. These are the "H-states" located outside the edges of the stabilizer polytope, and the "T-states" located outside the polytope faces as shown in Fig. 4.3. In Bloch coordinate representation, the magic states are given by

$$|H\rangle = 1/\sqrt{2}\{(\pm 1, \pm 1, 0), (\pm 1, 0, \pm 1), (0, \pm 1, \pm 1)\} \tag{4.7}$$

$$|T\rangle = 1/\sqrt{3}\{(\pm 1, \pm 1, \pm 1)\}. \tag{4.8}$$

Note that all states belonging to the same type of magic state are *Clifford equivalent*, meaning that we may rotate amongst them using only Clifford operations. The H-type states implement the $\pi/8$-gate (or Clifford equivalent gates), while the T-type states implement a $\pi/6$-gate [2].

For qudits, we define the generalized H-states to be states which will implement the

generalized $\pi/8$-gate [35]. There are $d^2(d^2-1)$ such states, defined as

$$|H\rangle = U_G\,|f_{abc}\rangle \quad a \in \mathbb{Z}_d^* \;\; b, c \in \mathbb{Z}_d \tag{4.9}$$

where U_G are the symplectic unitaries defined in 3.2 with

$$G = \left\{ \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, ..., \begin{bmatrix} 1 & d-1 \\ 0 & 1 \end{bmatrix} \right\} \tag{4.10}$$

and

$$|f_{abc}\rangle = \frac{1}{d} \sum_k \omega^{ak^3+bk^2+ck}\,|k\rangle\,. \tag{4.11}$$

For the special case of qutrits, we need to modify the above as

$$|f_{abc}\rangle = \frac{1}{3}(|0\rangle + \xi^{2a+6b+3c}\,|1\rangle + \xi^{a+6b+6c}\,|2\rangle) \tag{4.12}$$

where $\xi = e^{2\pi i/9}$ due to the irregular generalized $\pi/8$-gate in this dimension (see [35]). The generalized T-states (sometimes called the strange states [8]) are defined to be the states located the furtherest outside the Wigner polytope, in analogy to the geometrical interpretation for qubits. There are d^2 generalized T-states, corresponding to the facets of the Wigner polytope. In discrete Wigner representation, these would be the d^2 states with maximal sum-negativity of $sn = 1/d$ due to a single negative entry. For qutrits, we define a second class of magic state with maximal sum-negativity due to two negative entries of $-1/6$. There are 36 such states, defined in [8] as the Norrell states.

Chapter 5

Tight Qutrit Magic State Distillation

In the previous chapter, we introduce magic state distillation as one of the leading pathways to universal fault-tolerant quantum computation. Here, we present magic state distillation as a tool for studying resource theory. This chapter is based on [9], a collaboration with Mark Howard.

5.1 Motivation

A distillable state, that is a noisy state which can be brought to a magic state via distillation, can be thought of as a resource enabling universality. Therefore by studying the region of distillable states and their common properties, we can attempt to make statements about which properties are universality enabling resources. For example, we have seen that the stabilizer polytope represents the region of provably undistillable states in the qubit setting. Finding magic state distillation routines with tight thresholds to this polytope in both the H and T directions would imply that any non-stabilizer state enables universality. Reichardt [36] showed that the distillable region was tight to the stabilizer boundary along the octahedron edges in the H-direction (see Fig. 5.1), while Campbell and Browne [37, 34] showed that (for stabilizer codes of fixed length) there exists a region of undistillable non-stabilizer states outside the octahedron faces in the T-direction. We can ask similar questions in the qudit setting, where the provably undistillabe region of states is given by the Wigner polytope. Recently, the presence of negativity was shown to exactly coincide with the possibility of exhibiting state-dependent *contextuality* with respect to stabilizer measurements [7]. Therefore, it has been conjectured that contextuality is the resource that promotes universality in these dimensions [8]. Actually proving

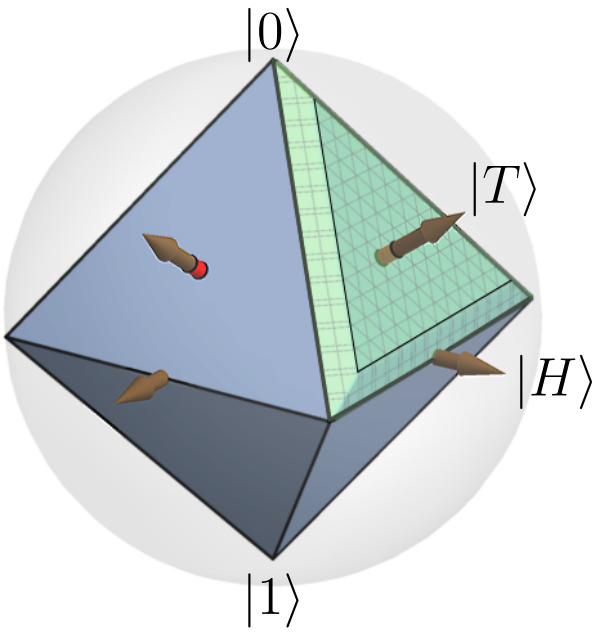


Figure 5.1: The qubit Bloch sphere and the stabilizer octahedron: The convex hull of stabilizer eigenstates – the stabilizer polytope – carves out a solid octahedron. H-type magic states lie outside the octahedron edges, and T-type magic states lie outside the octahedron faces. The green volume outside one face represents the region in which ancillas have not yet been shown to be distillable.

this conjecture requires showing that any state that is negatively represented is distillable by some magic state distillation routine. While magic state distillation protocols for all prime dimensions [4, 5, 6] have been found, no single protocol has been shown to distill states tight to the undistillable boundary. Here we focus on $d = 3$ and present distillation schemes that achieve tight distillation for some directions, as well as introduce a new class of magic state with maximal sum-negativity for qutrits.

5.2 Subspace of interest

We describe magic state distillation protocols that iteratively distill towards pure states within the $+1$-eigenspace of the $A_{0,0}$ phase point operator (2.7). This eigenspace is degen-

erate and has eigenvectors of the form $(a, b, b) \in \mathbb{C}^3$. Before describing the protocols, we give an overview of the geometrical interpretation of this eigenspace. We may parameterize pure states living within the (a, b, b) subspace using a polar angle θ and an azimuthal angle ϕ, via

$$|\psi\rangle = (a, b, b) = (\cos\theta, e^{i\phi}\sin\theta/\sqrt{2}, e^{i\phi}\sin\theta/\sqrt{2}) \tag{5.1}$$

where $\phi \in [0, 2\pi)$ and $\theta \in [0, \pi/2]$. The set of pure states corresponds to the surface of a sphere in analogy with the Bloch sphere for qubits. In our representation, points within the body of the sphere correspond to depolarized versions of the nearest surface state as in (4.5); states at distance $(1-p)$ from the centre of the sphere are depolarized by an amount p. In other words, a point in the interior of the sphere with spherical coordinates (r, θ, ϕ) corresponds to the state

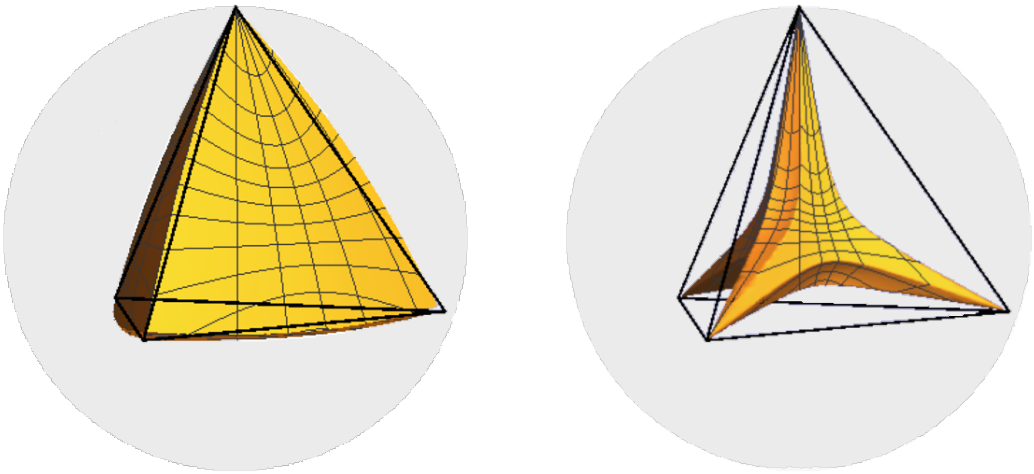
$$\rho = r\,|\psi(\theta,\phi)\rangle\,\langle\psi(\theta,\phi)| + (1-r)\frac{\mathbb{1}_3}{d} \tag{5.2}$$

where $|\psi(\theta, \phi)\rangle$ is given by (5.1) and $r = 1-p$. However unlike the qubit Bloch sphere, states in the interior no longer correspond to convex combinations of surface states in general. For example, in the qubit Bloch sphere picture, we expect an equal mixture of any two diametrically opposite points to correspond to the maximally mixed state. However in our representation, mixing the North pole $|0\rangle$ with the South pole $|N\rangle = (|1\rangle + |2\rangle)/\sqrt{2}$ gives

$$\frac{1}{2}\,|0\rangle\langle 0| + \frac{1}{2}\,|N\rangle\langle N| = \frac{1}{4}\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} \neq \frac{\mathbb{1}_3}{3}. \tag{5.3}$$

Despite the fact that our representation does not respect convexity, we feel it provides good intuition for the relevant geometry and symmetries. Mixtures of stabilizer states as well as states with positive Wigner function form a closed volume within the sphere (see Fig. 5.2). A convex 2-dimensional slice through qutrit state space can be found in [7], which depicts a similar hierarchy (i.e., the stabilizer polytope is a strict subset of the Wigner polytope).

In order to map the distillable region within the (a, b, b) subspace, it will suffice to partition the space into Clifford-equivalent sections and find the distillable region for only one such section. It is a well-known property of the Clifford group that $A_{0,0}$ is invariant under the symplectic part of the Clifford group i.e., $A_{0,0} = U_F A_{0,0} U_F^\dagger$ for all $F \in \mathrm{SL}(2, \mathbb{Z}_d)$. Consequently qutrit states of the form (a, b, b) remain in the $+1$ eigenspace of $A_{0,0}$ under all $|\mathrm{SL}(2, \mathbb{Z}_3)| = 24$ such transformations. Not only does the symplectic unitary $U_{\mathbb{1}_2} = \mathbb{1}_3$ fix every qutrit vector but the symplectic unitary $U_{-\mathbb{1}_2} = A_{0,0}$ also fixes every vector of



(a) (b)

Figure 5.2: Depolarized versions of qutrit states $(\cos\theta, e^{i\phi}\sin\theta/\sqrt{2}, e^{i\phi}\sin\theta/\sqrt{2})$: Points on surface of the sphere represent pure states, while every point in the interior is a depolarized version of the nearest surface state (see (4.5)). Figure (a) depicts the set of states that are non-negatively represented in the Wigner function and hence useless for MSD. The volume roughly corresponds to a "curvy tetrahedron" which hereafter will be referred to as the Wigner tetrahedron. Figure (b) depicts the set of states expressible as convex combinations of stabilizer states. A regular tetrahedron is shown as a visual aid and the pure states at its vertices are stabilizer states $\{|0\rangle, (1, \omega^k, \omega^k)/\sqrt{3}, k \in \mathbb{Z}_3\}$.

the form (a, b, b). Consequently, the set of non-trivial symplectic transformations acting on (a, b, b) states is $\mathrm{SL}(2, \mathbb{Z}_3)/\pm\mathbb{1}_2 = \mathrm{PSL}(2, \mathbb{Z}_3)$. This latter symmetry group is isomorphic to the rotation group of the tetrahedron (i.e. the alternating group A_4). Therefore we expect the entire space to partition into 12 Clifford-equivalent regions. These regions correspond to the 4 Wigner tetrahedron faces, further divided into 3 wedges each as shown in Fig. 5.3. We will consider distillation confined to the region defined as being contained by the three

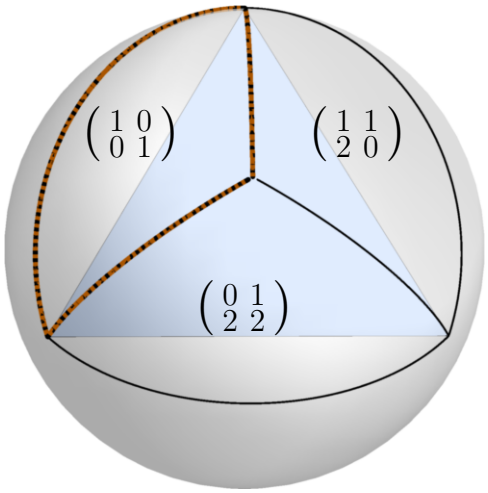


Figure 5.3: Clifford symmetries: The surface of the sphere lying outside one face of the Wigner tetrahedron is outlined in black and divided into 3 Clifford-equivalent wedges. Each wedge is labelled by the symplectic matrix F such that U_F maps states in that region back to the region of interest (highlighted).

lines

$$\text{Arc 1.} \quad (x, y, z) = (\sin 2\theta, 0, \cos 2\theta) \tag{5.4}$$

$$\begin{aligned}\text{Arc 2.} \quad (x, y, z) = \big((17 \sin 2\theta - 2\sqrt{2} \cos 2\theta)/18, \\ (\sin 2\theta + 2\sqrt{2} \cos 2\theta)/2\sqrt{3}, \\ (-\sqrt{2} \sin 2\theta + 5 \cos 2\theta)/9\big)\end{aligned} \tag{5.5}$$

$$\text{Arc 3.} \quad (x, y, z) = (\frac{1}{2} \sin 2\theta, \frac{\sqrt{3}}{2} \sin 2\theta, \cos 2\theta) \tag{5.6}$$

5.3 MSD routines

In this section we present two classes of magic states within the (a, b, b) subspace, give explicit distillation schemes for each, and sketch out the corresponding distillable region. Magic states within the (a, b, b) subspace can be generally split into two types, those lying outside the six Wigner tetrahedron edges, and those lying outside the four Wigner tetrahedron faces. By numerically searching over a large set of stabilizer codes, we find

that there exists many limiting states outside the Wigner tetrahedron edges, where limiting state is taken to mean the end point of the iterative procedure based on a given stabilizer code. In contrast, we find only one limiting state outside the Wigner tetrahedron faces, and this is the state geometrically located the furthest outside the Wigner tetrahedron in the middle of each face. These so-called Norrell states [8], which were not previously known to be directly distillable, have maximal sum-negativity $sn = \frac{1}{3}$ arising from a Wigner function of $-\frac{1}{6}$ in two phase space points. The set of numerically identified limiting states is shown in Fig. 5.4, and the corresponding stabilizer codes are listed in Appendix A. We note that a nice feature of all our distillation protocols is that they do not require twirling (i.e., dephasing in a particular basis) between rounds of MSD or cycling [4].

We now demonstrate distillation schemes based on stabilizer codes for both types of magic state. We begin with the edge states for which tight distillation thresholds were achieved, and then present the Norrell state which improves the distillation region outside the Wigner tetrahedron faces slightly. Out of the many edge-type magic state depicted in Fig. 5.4, the state $|E\rangle$ whose code has the best overall distillable region will be presented. The state

$$|E\rangle = \begin{pmatrix} 0.774149 \\ 0.447601 \\ 0.447601 \end{pmatrix}, \tag{5.7}$$

which lies on the x-z plane is distilled by a $[[4, 1, 2]]_3$ code given by generators:

$$\text{Edge}, |E\rangle : \quad \begin{array}{c|cccc|cccc} G_1 & 0 & 0 & 0 & 2 & 2 & 2 & 0 & 0 \\ G_2 & 1 & 1 & 0 & 1 & 1 & 1 & 2 & 2 \\ G_3 & 0 & 0 & 1 & 0 & 2 & 0 & 0 & 0 \\ Z_L & 2 & 0 & 0 & 2 & 2 & 2 & 1 & 2 \\ X_L & 0 & 0 & 0 & 0 & 1 & 2 & 0 & 0 \end{array} \tag{5.8}$$

in $(x|z)$ notation where each element is taken to be a generalized Pauli operator. This magic state $|E\rangle$ may be transformed into an equatorial state useful for state injection by first applying a symplectic unitary U_F with $F = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and then following the parity check and equatorialization procedures as outlined in [4]. For all states $|\theta\rangle := (\cos\theta, \sin\theta/\sqrt{2}, \sin\theta/\sqrt{2})$ along the Wigner tetrahedron edge, on the line $(x, y, z) = (\sin 2\theta, 0, \cos 2\theta)$ with $0 \leq \theta \leq \arccos 1/\sqrt{3}$, we find an error threshold given by

$$p^* = 1 - \frac{4}{1 + 3\cos 2\theta + 3\sqrt{2}\sin 2\theta}, \tag{5.9}$$

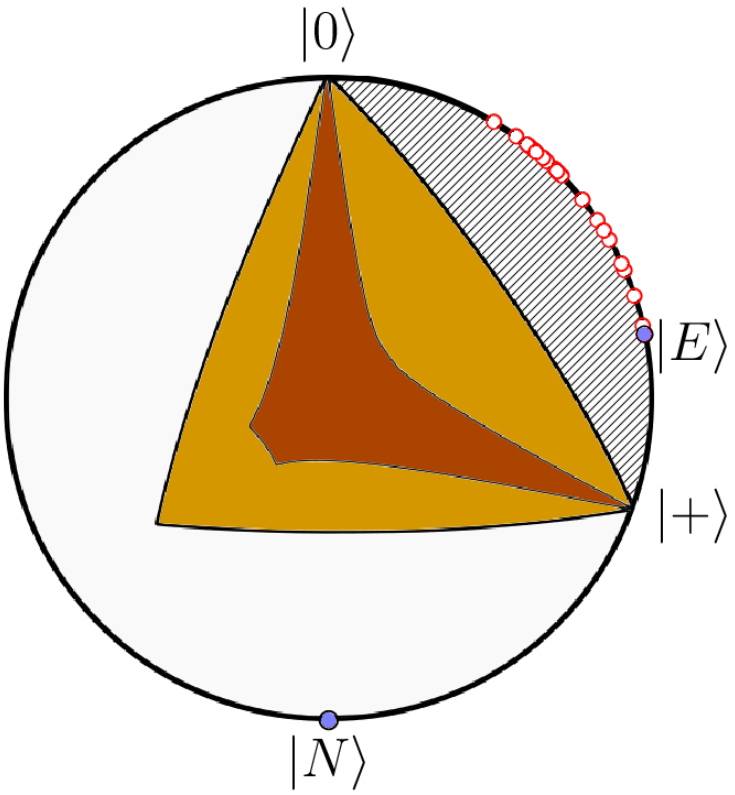


Figure 5.4: Cross section through the x-z plane showing limiting states of various MSD routines: Different circles correspond to different choices of 4-qutrit stabilizer code and every such code can distill all states in the hatched area. Two magic states of interest are denoted by filled circles; the $|E\rangle$ state lies on the arc joining $|0\rangle$ and $|+\rangle$, while the south pole $|N\rangle$ is Clifford equivalent to the Norrell state $|N'\rangle$ that we discuss in the text. The colored regions represent undistillable states inside the Wigner polytope (lighter, larger) and mixtures of stabilizer states (darker, smaller) i.e., this shows a 2-d slice through Fig. 5.2.

which corresponds to a state $\rho_\theta = (1-p^*)\left|\theta\right\rangle\left\langle\theta\right| + p^* \mathbb{1}/3$ with Wigner function

$$W_{\rho_\theta} = \begin{pmatrix} r & s & s \\ t & 0 & 0 \\ t & 0 & 0 \end{pmatrix} \quad \begin{aligned} r &= \tfrac{\cos 2\theta + \sqrt{2}\sin 2\theta + 3}{9\cos 2\theta + 9\sqrt{2}\sin 2\theta + 3} \geq 0 \\ s &= \tfrac{4\cos 2\theta + \sqrt{2}\sin 2\theta}{9\cos 2\theta + 9\sqrt{2}\sin 2\theta + 3} \geq 0 \\ t &= \tfrac{\sqrt{2}}{3\cot 2\theta + \csc 2\theta + 3\sqrt{2}} \geq 0 \end{aligned} \tag{5.10}$$

and therefore distillation is tight to the Wigner polytope boundary for all points along the Wigner tetrahedron edges. The maximally robust edge state occurs at $\theta = \frac{1}{2}\cos^{-1}(\frac{1}{\sqrt{3}})$ and this corresponds to the $+1$ eigenstate of the qutrit Fourier transform. Using our edge code, this state can tolerate depolarizing noise up to

$$p^* = 1 - \frac{4}{1+3\sqrt{3}} \approx 0.354438\ , \tag{5.11}$$

which is the best known depolarizing noise threshold for qutrits. The edge code distills points along the depolarizing axis of the Norrell state $\left|N'\right\rangle = (2,-1,-1)/\sqrt{6}$ with a threshold of

$$p^* = \frac{1}{3}\Big(5 - \frac{1}{q} - q\Big) \qquad \Big(q = \sqrt[3]{28 - 3\sqrt{87}}\Big) \tag{5.12}$$

$$\approx 0.304379. \tag{5.13}$$

The entire distillable region of this $[[4,1,2]]_3$ edge code confined to the Clifford-equivalent wedge of interest is shown in Fig. 5.5. The situation is analogous to the qubit picture wherein distillation is tight for all edges (H-type) and there is a pocket of undistillable states outside the Wigner tetrahedron faces (T-type).

The distillation region may be improved slightly by a second stabilizer code which has the Norrell state $\left|N'\right\rangle = (2,-1,-1)/\sqrt{6}$ as a limiting state. The Norrell state $\left|N'\right\rangle$ is Clifford equivalent to the south pole state $\left|N\right\rangle$ shown in Fig. 5.4. This $[[4,1,2]]_3$ code has generators:

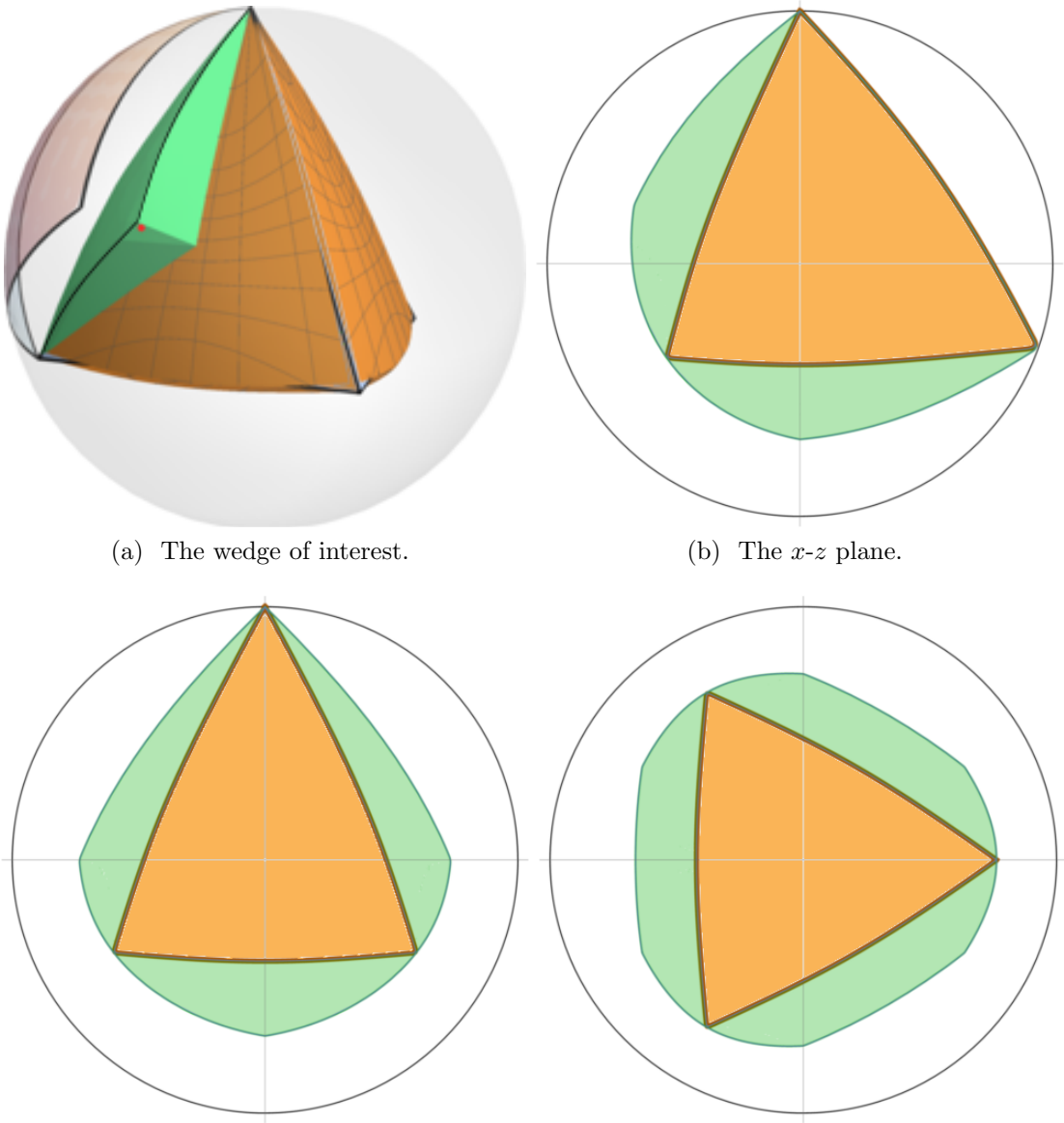
Face, $\left|N'\right\rangle$:

G_1	2	0	0	2	1	2	0	1
G_2	2	1	0	1	1	0	1	0
G_3	1	0	1	2	0	2	1	0
Z_L	1	0	0	2	1	0	1	2
X_L	0	0	0	0	1	2	1	2

(5.14)

We find that the Norrell state $\left|N'\right\rangle$ is distilled up to a threshold noise rate of

$$p^* \approx 0.32989\ , \tag{5.15}$$



(a) The wedge of interest.

(b) The x-z plane.

(c) The y-z plane.

(d) The x-y plane.

Figure 5.5: The remaining undistillable region highlighted in green, after using the edge code (5.8). The small red dot in (a) corresponds to the threshold point for distilling $|N'\rangle$ using the face code (5.14).

which increases the distillable region as shown in Fig. 5.5. The equatorialization procedure of [4] converts the Norrell state into something useful for state injection (two copies of $|N'\rangle$ enable implementation of the non-Clifford unitary $U = \mathrm{diag}(1, 1, -1)$ [4]).

The success probabilities $p_{\mathrm{succ}} = \mathrm{Tr}(\rho_{in}^{\otimes 4}\Pi_{\mathrm{C}})$ for both codes for points along the depolarizing axes of their respective limiting states begin at approximately $p_{\mathrm{succ}} = 0.12$ for $p = 0$ and drop to approximately $p_{\mathrm{succ}} = 0.07$ for $p = 0.3$, assuming post-selection on the trivial error syndrome. Similar to [4] error suppression appears to be linear, which means that despite the appreciable success probability, both codes are outperformed in terms of efficiency by a previously known qutrit code [5, 6]. Our focus was on maximizing the range of applicability of MSD schemes rather than their efficiency. Nevertheless, exploring the landscape of stabilizer codes with non-stabilizer limiting states may have practical consequences if these codes can be adapted or used in conjunction with others (as in [31]) to produce more efficient schemes.

5.4 Discussion

In summary, we have proven the existence of magic state distillation schemes with thresholds tight to the Wigner polytope boundary for $d = 3$. This represents a first step towards showing that all negatively represented states are distillable by some routine. Our result for the distillable region presents an interesting analogy to the qubit Bloch sphere distillable region. From the magic state distillation point of view, our results help to map out the set of codes which are useful for MSD and the set of limiting states for higher dimensional systems. In particular, finding stabilizer codes with distillation properties besides those known to be useful for magic state distillation by their transversality properties [2, 5] is useful. We have shown that a certain class of state with maximal sum negativity for qutrits, the Norrell state, is distillable by at least one stabilizer code. We have given a convenient parameterization of the $+1$-eigenspace of the $A_{0,0}$ phase space operator which allows us to easily visualize the distillable region and understand the geometric significance of the Norrell state. We identified a large set of limiting states within the (a, b, b) subspace, all located along the Wigner tetrahedron edges. These edge states allow for universality by following the parity check and equatorialization procedures outlined in [4] to produce equatorial states useful for state injection. Why these states in particular are limiting states, and what properties of the identified stabilizer codes make them useful for magic state distillation are interesting open questions.

Chapter 6

Small Codes for Qubit Magic State Distillation

In the previous chapter, we present the main result of our study of magic state distillation, namely the demonstration of tight thresholds for $d > 2$. In this chapter, we turn our attention to qubit routines, which may be of more practical interest. This chapter is based on [10], a collaboration with Mark Howard.

6.1 Motivation

Here we use small stabilizer codes, which means that the number of qubits, the number of measurements and the number of nonlocal operations involved in each round of distillation are all small. While these are attractive features, it seems that using a small code limits the amount of purification that can occur per round of distillation. Efficient codes for magic state distillation typically exhibit quadratic ($p \mapsto \mathcal{O}(p^2)$) or cubic ($p \mapsto \mathcal{O}(p^3)$) suppression of the error parameter p. The codes listed here all exhibit linear error suppression and consequently are not competitive with existing routines in terms of efficiency. Nevertheless we feel that exploring the landscape of codes that achieve distillation is still worthwhile. Reichardt [39] has previously summarized a handful of known distillation routines exhibiting the best thresholds. A number of the codes presented here achieve tight distillation right up to the boundary of the stabilizer octahedron. We also present codes that converge, upon iteration, to states that are not those that are typically considered i.e., the H-type and T-type magic states. None of our codes require twirling between rounds.

6.2 Results

All stabilizer codes are local Clifford equivalent to some CWS stabilizer code. Consequently CWS stabilizer codes form a subset of all stabilizer codes and so checking all combinations of graphs, Γ, and codewords, w, may still miss stabilizer codes that are useful for distillation. Nevertheless, for $n \leq 4$ qubits we iterated over all graphs on n vertices and for $n \in \{5, 6\}$ qubits we iterated over non-isomporphic graphs on n vertices. We were primarily focused on recording those CWS codes that achieve tight distillation but we also noted a number of non-tight codes that distilled to target states other than $|H\rangle$.

With the exception of one code, the codes that we present will be depicted graphically. This concise representation is possible because of the CWS formalism that we described in section 3.3. All the codes that we present converge, upon repeated iteration, to a pure state on the surface of the Bloch sphere. Codes that we describe below as being tight obey the following property: all non-stabilizer states in the same quadrant as $|M\rangle$ are distillable. For instance, if $|M\rangle$ has Bloch vector $(x, 0, z)$ with $x, z > 0$ then all states ρ satisfying $\mathrm{Tr}(X\rho) + \mathrm{Tr}(Z\rho) > 1$ converge to $|M\rangle$ under repeated iteration. (We confirmed this numerically by taking a random sample of 1000 points from the relevant region).

6.2.1 Codes achieving tight distillation

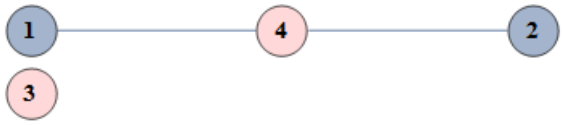
A 3-qubit code with generators

$$\begin{array}{cccc} G_1 & Z & I & Z \\ G_2 & X & Z & X \\ Z_L & X & X & Y \\ X_L & I & X & Z \end{array} \tag{6.1}$$

corresponding to the logical basis states

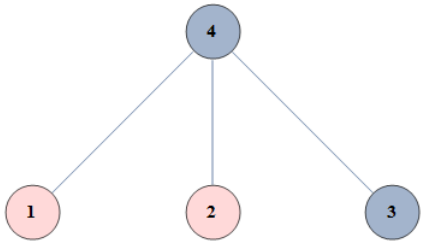
$$|0_L\rangle = \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ -i \\ 0 \\ 0 \\ 1 \\ 0 \\ i \end{pmatrix}, |1_L\rangle = \frac{1}{2}\begin{pmatrix} -i \\ 0 \\ 1 \\ 0 \\ 0 \\ -i \\ 0 \\ -1 \end{pmatrix} \tag{6.2}$$

(a) This code converges to $(x, y, z) = (\sin\theta, 0, \cos\theta)$ where $\theta = \arctan\sqrt{(\sqrt{5}-1)/2}$



(b) This code converges to $(x, y, z) = (0.66796, 0., 0.7442)$

(c) This code converges to $(x, y, z) = (0.81281, 0, 0.58252)$



(d) This code converges to $(x, y, z) = (0.64969, 0, 0.7602)$

(e) This code converges to $(x, y, z) = (1, 0, 1)/\sqrt{2}$

Figure 6.1: Codeword stabilized quantum codes with tight distillation thresholds. The classical codeword w associated with each graph is the binary vector with a "1" in lighter pink positions and "0" in darker blue positions.

distills an equatorial state in the y-z plane with Bloch coordinates $(0, -.83929, -.54369)$ up to a tight error threshold of $p_{oct} = .276921$.

Other codes achieving tight distillation are given in pictorial form in Fig. 6.1 and Fig. 6.2. It is a straightforward exercise to recover the generators and logical operators if necessary. The 5-qubit code in Fig. 6.1(e) distills to the $|H\rangle$-type magic state but is more efficient than the (also tight) 7-qubit Steane code. The efficiency of these codes is compared in Fig. 6.3.

6.2.2 Codes not achieving tight distillation

Here we highlight a small number of codes that, even though they do not achieve tight distillation, we still find to be noteworthy.

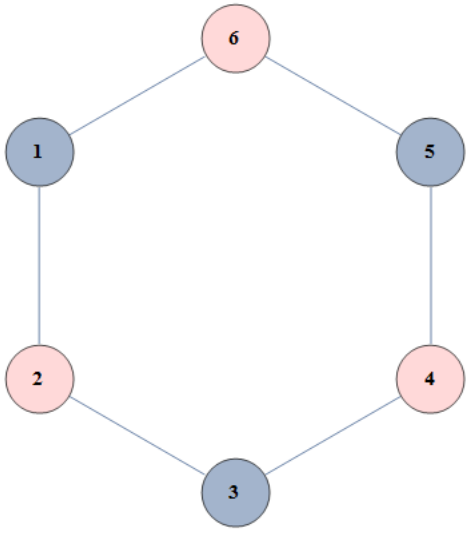
The $|T\rangle$-type states (depicted in Fig. 4.3) were shown to be distillable by Bravyi and Kitaev [2] using the perfect $[[5, 1, 3]]$ code. We are not aware of any additional routines for $|T\rangle$ states that have subsequently been developed. This is in marked contrast to $|H\rangle$-type distillation for which a number of codes have been found. For this reason we note the existence of a 4-qubit code in Fig. 6.4(b) that also converges to $|T\rangle$-type states. A visual comparison with the $[[5, 1, 3]]$ code in Fig. 6.4(a) indicates that they are closely related. Unfortunately, the threshold for the new code is worse than that of the $[[5, 1, 3]]$ code. Finding tight distillation routines for the $|T\rangle$ direction was already known to be more complex than the $|H\rangle$ case because of a no-go theorem in [37].

In Fig. 6.4(c) we depict a code that distills $|\pi/3\rangle = (|0\rangle + e^{i\pi/3}|1\rangle)/\sqrt{2}$. This is particulary interesting as the associated gate $U = \mathrm{diag}(1, e^{i\pi/3})$ is not transversal for any stabilizer code [42] which prevents it from being distillable by the most commonly used distillation techniques.

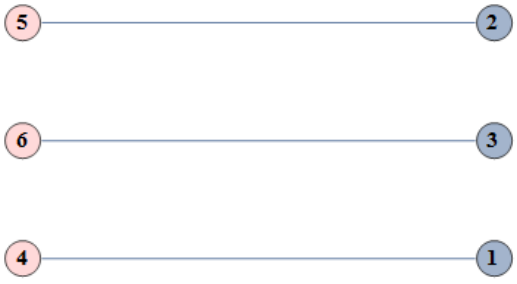
It is of interest to find distillation routines that distill $|V\rangle$-type magic states, which look like $(x, y, z) = (\frac{3}{5}, 0, \frac{4}{5})$ in the Bloch sphere picture. Supplementing Cliffords with such states leads to a set of gates – the V-basis – that is highly efficient for gate synthesis [40]. We did not find such a code but in Fig. 6.4(d) we depict a code that converges to a nearby state $(x, y, z) = (0.60965, 0., 0.79267)$.

6.3 Discussion

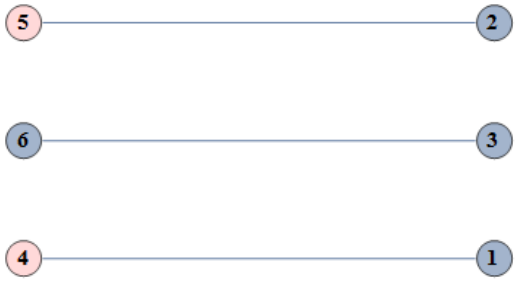
We have presented a collection of qubit magic state distillation schemes using small stabilizer codes, most of which achieve tight distillation up to the edge of the stabilizer octahedron. Two codes that are not tight were still noted because they converge to $|T\rangle$ and $|\pi/3\rangle$ respectively. Our distillation routines converge to equatorial states, which means that they can be used to implement rotations about a Pauli axis using standard half-teleportation techniques [2]. We have noted that inefficiency of our routines relates to linear error suppression and leave as an open question whether there is some way of boosting their efficiency by combining them with other codes or techniques. Another possible avenue is to use these codes to convert non-stabilizer states of one type to another, as was done in e.g., [41]. For example, if we have access to almost pure $|H\rangle$ states (after using the Reed-Muller code) then input these to the $|\pi/3\rangle$ routine, the output is an almost pure state somewhere between $|H\rangle$ and $|\pi/3\rangle$. Further analysis of the relative merits of methods such as this is left for future work.



(a) This code converges to $(x, y, z) = (0.84893, 0., 0.52851)$

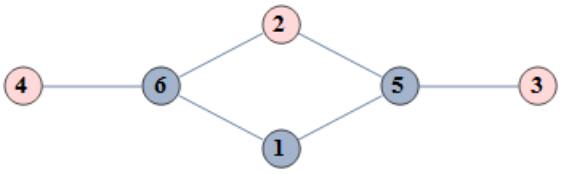


(b) This code converges to $(x, y, z) = (0.63544, 0., 0.77215)$



(c) This code converges to $(x, y, z) = (0.81281, 0., 0.58252)$

(d) This code converges to $(x, y, z) = (0.84534, 0., 0.53423)$



(e) This code converges to $(x, y, z) = (0.58252, 0., 0.81281)$

Figure 6.2: Codeword stabilized quantum codes with tight distillation thresholds. The classical codeword w associated with each graph is the binary vector with a "1" in lighter pink positions and "0" in darker blue positions.

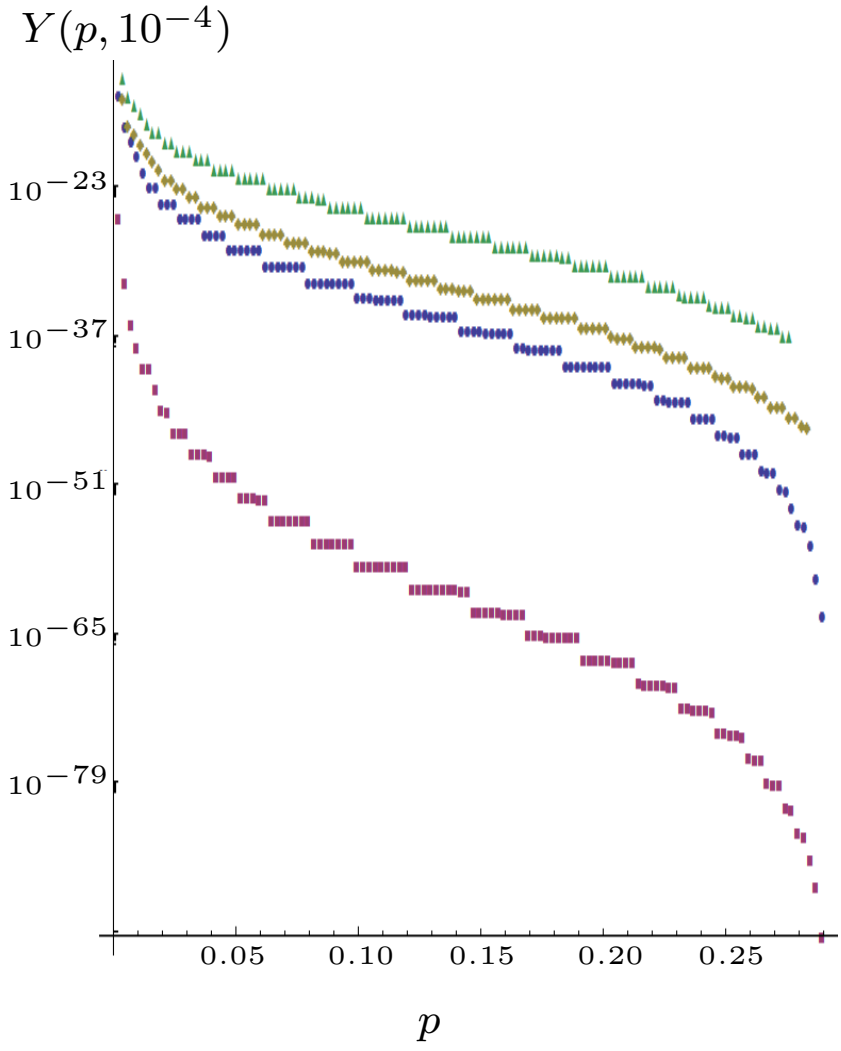
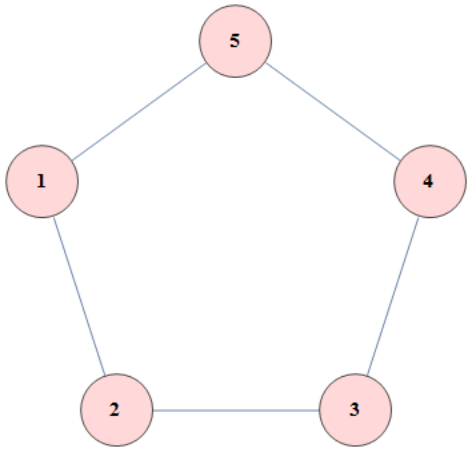
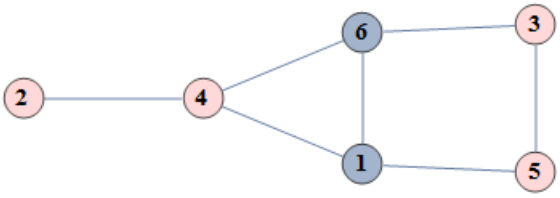


Figure 6.3: Curves showing the yield (efficiency) of various tight distillation routines as a function of input noise rate p. From bottom to top we have (*i*) the 7-qubit Steane code as applied by Reichardt [36] (*ii*) the 5-qubit code in Fig. 6.1(e), (*iii*) 3-qubit code from Fig. 6.1(a) and (*iv*) the 3-qubit code in Eq. (6.1).

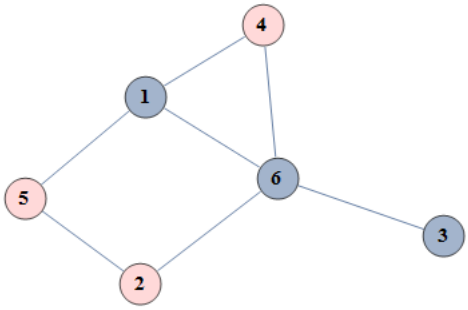


(a) This is the well-known 5-qubit code in CWS form, which Bravyi and Kitaev showed distilled $|T\rangle$

(b) This code also distills to $|T\rangle$ albeit less efficiently and with a worse threshold than the above code. Note that a trivial Z correction must be applied in between rounds



(c) This code converges to a state Clifford-equivalent to $|\pi/3\rangle = (|0\rangle + e^{i\pi/3}|1\rangle)/\sqrt{2}$. The limiting state has Bloch vector $(x, y, z) = (\frac{1}{2}, 0, \frac{\sqrt{3}}{2})$



(d) This code converges to $(x, y, z) = (0.60965, 0., 0.79267)$

Figure 6.4: Codeword stabilized quantum codes that do not achieve tight thresholds

Chapter 7

Noise Characterization Background

In the second part of this thesis, we turn our attention to the task of noise characterization. We begin with an overview of common noise processes as well as experimental characterization techniques.

7.1 Noise processes

7.1.1 Amplitude damping

Amplitude damping describes the relaxation of a quantum system from an excited state to the ground state of the system Hamiltonian. For example, let the energy eigenstates of a two-level system be $\{|0\rangle, |1\rangle\}$. If we prepare the population in the $|1\rangle$ state, we expect some of the population will relax to $|0\rangle$ to reach thermal equilibrium. *Pure amplitude damping* describes amplitude damping at zero temperature, where the steady state of the system is perfectly $|0\rangle$. In this section we describe the noise model for pure amplitude damping, and delay finite temperature considerations for section 7.1.3.

The Kraus-operator representation of the amplitude damping channel is $\mathcal{E}_{AD}(\rho) = \sum_k A_k \rho A_k^\dagger$ where

$$A_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{bmatrix}. \tag{7.1}$$

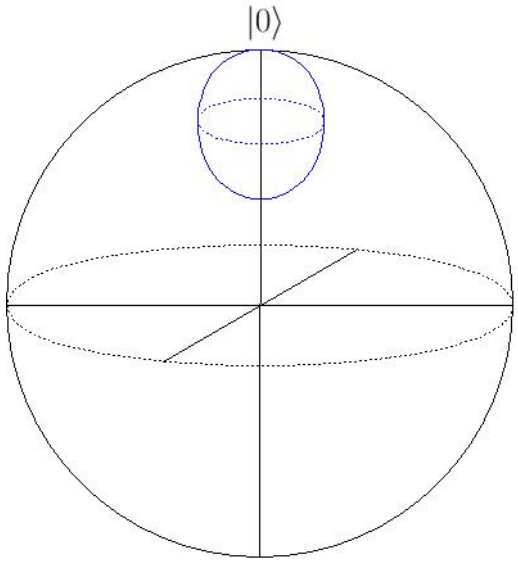


Figure 7.1: Bloch sphere before (black) and after (blue) application of the amplitude damping channel.

In the Pauli-Liouville representation,

$$\Lambda(\mathcal{E}_{AD}) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{1-\gamma} & 0 & 0 \\ 0 & 0 & \sqrt{1-\gamma} & 0 \\ \gamma & 0 & 0 & 1-\gamma \end{bmatrix}. \tag{7.2}$$

From either representation, we can work out the output state ρ' for a general input $\rho = \begin{bmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{bmatrix}$ to be

$$\rho' = \begin{bmatrix} \rho_{00} + \gamma\rho_{11} & \sqrt{1-\gamma}\rho_{01} \\ \sqrt{1-\gamma}\rho_{10} & (1-\gamma)\rho_{11} \end{bmatrix}. \tag{7.3}$$

The effect of amplitude damping is to increase the population of $|0\rangle$ while decreasing the population of $|1\rangle$ for $\gamma > 0$ as expected. We also observe the attenuation of the coherence terms, which is required to preserve positivity. Visually, we can understand the process of amplitude damping by observing its effect on the Bloch sphere, shown in Fig. 7.1. Taking a continuous time description, we can describe amplitude damping by the master equation

$$\frac{\partial\rho}{\partial t} = \Gamma\left[\sigma_+\rho\sigma_+^\dagger - \frac{1}{2}\{\sigma_+^\dagger\sigma_+, \rho\}\right] \tag{7.4}$$

where $\gamma = 1 - e^{-\Gamma \Delta t}$, from which we obtain

$$\rho' = \begin{bmatrix} 1 - e^{-\Gamma \Delta t} \rho_{11} & e^{-\Gamma \Delta t/2} \rho_{01} \\ e^{-\Gamma \Delta t/2} \rho_{10} & e^{-\Gamma \Delta t} \rho_{11} \end{bmatrix}. \tag{7.5}$$

The rate of equilibration is described by the characteristic time $T_1 = 1/\Gamma$, while the damping of off-diagonal terms is described by the time T_2^*. In this case, T_2^* takes its maximal value of $T_2^* = 2/\Gamma$ induced by amplitude damping. In the following section we see that T_2^* can be decreased by the process of phase damping.

7.1.2 Phase damping

Phase damping describes the decay of coherence terms caused by weak coupling to a bath. The Kraus operators are

$$A_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & 0 \\ 0 & \sqrt{\gamma} \end{bmatrix} \tag{7.6}$$

or alternatively,

$$\tilde{A}_0 = \sqrt{\alpha}\mathbb{1}, \quad \tilde{A}_1 = \sqrt{1-\alpha} Z \tag{7.7}$$

where $\alpha = 1/2(1 + 1/\gamma)$, from which we can see that phase damping is a Pauli channel. Consequently, the Pauli-Liouville representation is diagonal, with

$$\Lambda(\mathcal{E}_{PD}) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{1-\gamma} & 0 & 0 \\ 0 & 0 & \sqrt{1-\gamma} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{7.8}$$

The output state ρ' for a general input $\rho = \begin{bmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{bmatrix}$ is

$$\rho' = \begin{bmatrix} \rho_{00} & \sqrt{1-\gamma}\rho_{01} \\ \sqrt{1-\gamma}\rho_{10} & \rho_{11} \end{bmatrix}. \tag{7.9}$$

For $\gamma > 0$, the off-diagonal terms are attenuated as expected. In contrast to amplitude damping, there is no secondary effect. The effect of phase damping on the Bloch sphere

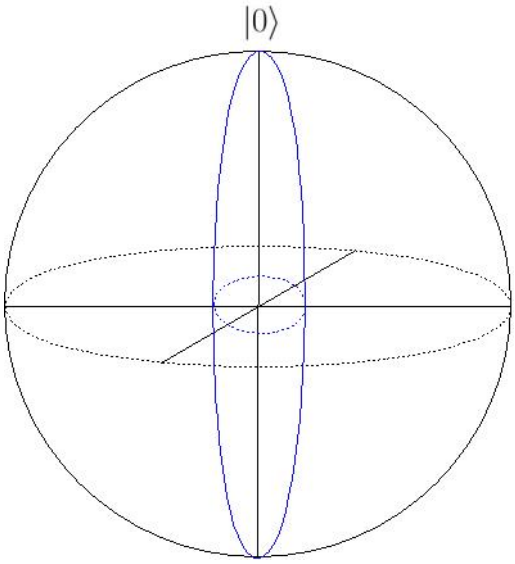


Figure 7.2: Bloch sphere before (black) and after (blue) application of the phase damping channel.

is shown in Fig. 7.2. Taking a continuous time description, phase damping is described by the master equation

$$\frac{\partial \rho}{\partial t} = \frac{\Gamma}{2}\Big[Z\rho Z - \frac{1}{2}\{ZZ, \rho\}\Big] \tag{7.10}$$

where $\gamma = 1 - e^{-2\Gamma\Delta t}$, from which we obtain

$$\rho' = \begin{bmatrix} \rho_{00} & e^{-\Gamma\Delta t}\rho_{01} \\ e^{-\Gamma\Delta t}\rho_{10} & \rho_{11} \end{bmatrix}. \tag{7.11}$$

In this case, the damping is described by the characteristic time $T_2^* = 1/\Gamma$. In the following section we describe how to combine amplitude damping with phase damping, while considering finite temperature, to arrive at the generalized damping channel.

7.1.3 Generalized damping

Experimental systems are typically subject to both amplitude damping at finite temperature and phase damping. These processes can be described by a generalized damping

master equation

$$\frac{\partial\rho}{\partial t} = \lambda\Gamma_1(\sigma_+\rho\sigma_+^\dagger - \frac{1}{2}\{\sigma_+^\dagger\sigma_+,\rho\}) + (1-\lambda)\Gamma_1(\sigma_-\rho\sigma_-^\dagger - \frac{1}{2}\{\sigma_-^\dagger\sigma_-,\rho\}) + \frac{\Gamma_2}{2}(Z\rho Z - \frac{1}{2}\{ZZ,\rho\}) \tag{7.12}$$

where $\sigma_\pm = (X \pm iY)/2$. Γ_1 and Γ_2 control the rate of amplitude and phase damping respectively and λ is a temperature-dependent parameter that governs the relative population of ground and excited states in equilibrium. The action of the generalized damping channel for a generic input state is given by

$$\rho(t) = \begin{bmatrix} e^{-\Gamma_1 t}(1-\rho_{11}-\lambda)+\lambda & \rho_{01}e^{-t(\Gamma_1/2+\Gamma_2)} \\ \rho_{10}e^{-t(\Gamma_1/2+\Gamma_2)} & e^{-\Gamma_1 t}(\rho_{11}+\lambda-1)-\lambda+1 \end{bmatrix}. \tag{7.13}$$

In a discrete-time representation, the Kraus operators are

$$\begin{aligned} A_0 &= \sqrt{\frac{1-\gamma_1/2-x}{a^2+b^2}}(a\left|0\right\rangle\left\langle 0\right| - b\left|1\right\rangle\left\langle 1\right|) \\ A_1 &= \sqrt{\lambda\gamma_1}\left|0\right\rangle\left\langle 1\right| \\ A_2 &= \sqrt{\frac{1-\gamma_1/2+x}{a^2+b^2}}(b\left|0\right\rangle\left\langle 0\right| + a\left|1\right\rangle\left\langle 1\right|) \\ A_3 &= \sqrt{(1-\lambda)\gamma_1}\left|1\right\rangle\left\langle 0\right| \end{aligned} \tag{7.14}$$

where $\gamma_1 = 1 - e^{-\Gamma_1 t}$, $\gamma_2 = 1 - e^{-2\Gamma_2 t}$, $x = [(1-\gamma_1)(1-\gamma_2) - \gamma_1^2\lambda(1-\lambda) + 1/4\gamma_1^2]^{1/2}$, $a = x + \gamma_1/2 - \gamma_1\lambda$, and $b = \sqrt{(1-\gamma_1)(1-\gamma_2)}$. These Kraus operators correspond to a process matrix in Pauli-Liouville representation given by

$$\Lambda(\mathcal{E}_{GD}) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{(1-\gamma_1)(1-\gamma_2)} & 0 & 0 \\ 0 & 0 & \sqrt{(1-\gamma_1)(1-\gamma_2)} & 0 \\ \gamma_1(2\lambda-1) & 0 & 0 & 1-\gamma_1 \end{bmatrix}. \tag{7.15}$$

The Pauli-Liouville matrix and master equation reduce to those given in the previous sections when $\lambda = 1$ (zero temperature) and either $\gamma_2 = 0$ (amplitude damping) or $\gamma_1 = 0$ (phase damping). The Kraus operators will not, but the Kraus-operator representation of a channel in not unique. The action of the generalized damping channel is shown in Fig. 7.3.

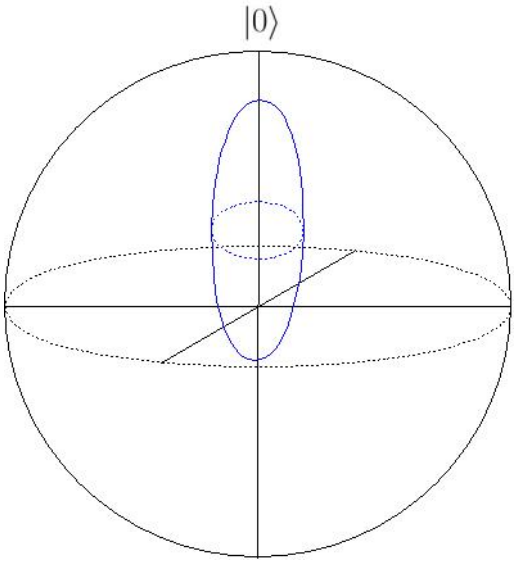


Figure 7.3: Bloch sphere before (black) and after (blue) application of the generalized damping channel.

7.2 Measuring noise

7.2.1 Inversion recovery

In an inversion recovery experiment, the state is initially prepared in the excited state $|1\rangle\langle 1|$, then allowed to evolve for a certain time before measuring the relative population levels of $|0\rangle\langle 0|$ vs. $|1\rangle\langle 1|$. The relevant experimentally measured quantity is $Q(t) = Tr(-Z\rho(t))$ which can then be fit to extract T_1. Under generalized damping, the expected $\rho(t)$ and $Q(t)$ would be

$$\rho(t) = \frac{1}{2}\begin{bmatrix}\lambda(1-e^{-\Gamma_1 t}) & 0 \\ 0 & 1-\lambda(1-e^{-\Gamma_1 t})\end{bmatrix} \tag{7.16}$$

$$Q(t) = 2\lambda e^{-\Gamma_1 t} - 2\lambda + 1 \tag{7.17}$$

where $T_1 = 1/\Gamma_1$.

7.2.2 Ramsey experiments

In a Ramsey experiment, the state is initially prepared in the $|+\rangle\langle+|$ state, and then allowed to precess freely around the z-axis for some time before being pulsed back to the z-axis so that the "length" may be measured as $Q(t) = Tr(\rho(t)Z)$. The resulting measure as a function of time will be a decaying sinusoidal function, and the decay may be fit to give T_2^*. Under generalized damping, the expected $\rho(t)$ and $Q(t)$ would be

$$\rho(t) = \frac{1}{2}\begin{bmatrix} 1 + e^{-t(\Gamma_2+\Gamma_1/2)} & (e^{-t\Gamma_1} - 1)(2\lambda - 1) \\ (e^{-t\Gamma_1} - 1)(2\lambda - 1) & 1 - e^{-t(\Gamma_2+\Gamma_1/2)} \end{bmatrix} \tag{7.18}$$

$$Q(t) = e^{-t(\Gamma_2+\Gamma_1/2)} \tag{7.19}$$

where $T_2^* = 1/(\Gamma_2 + \Gamma_1/2)$. Since Γ_1 is already known from T_1 estimation, accurate estimation of T_2^* becomes equivalent to accurate estimation of Γ_2. Note that if Hahn-echo [54] or CPMG [55, 56] pulse sequences have been applied to measure T_2, the resulting form of $Q(t)$ is the same as (7.19) with the redefinition $\Gamma_2 \to \Gamma_2^{H/CPMG}$ where $T_2 = 1/(\Gamma_2^{H/CPMG} + \Gamma_1/2)$.

7.2.3 Randomized benchmarking

Inversion recovery and Ramsey experiments are designed to measure a specific type of noise parameter when the noise model is known (or approximately known). When the noise model is unknown, we need to use other techniques. Ideally, we could completely characterize the noise process using quantum process tomography [43, 44]. However this method is inefficient in the number of qubits, and fails to account for state preparation and measurement (SPAM) errors [45]. Randomized benchmarking [47, 48, 22] provides an alternative which circumvents both of these issues. In traditional randomized benchmarking, we learn the average error rate over a group of operations which form a 2-design (taken to be the Clifford group in this discussion). The protocol may be adapted to learn the error rate on a particular gate [49], even for a gate outside the Clifford group [64]. In this section we briefly summarize the traditional protocol.

The basic idea of the protocol is to apply random sequences of Clifford operations of varying length. The sequence is followed by the inverse operation so that the output state would ideally be left unchanged. By observing the decrease in fidelity (between output and input states) as a function of sequence length, we learn the average error rate associated with the Clifford group $\mathcal{C}$. More explicitly, the protocol consists of:

1. For a given sequence length $m \in \mathbb{N}$, choose a sequence $\mathbf{k} = (k_1, ..., k_m)$ uniformly at random where $k_i \in \{1, ..., |\mathcal{C}|\}$.

2. Prepare an initial state ρ (usually $\rho \approx |0\rangle\langle 0|$).

3. Apply the random sequence of Clifford operations $C_{k_m} \circ ... \circ C_{k_1}$ to ρ.

4. Apply the inverse operation $C_{k_0} = \prod_{i=1}^{m} C_{k_i}^{\dagger}$.

5. Perform a projective measurement $\{E, \mathbb{1} - E\}$ (usually $E \approx \rho$).

6. Repeat steps $1 - 5$ enough times to estimate the survival probability to a desired precision.

7. Repeat steps $1 - 6$ for varying m.

The survival probability is expected to fit an exponential decay in the sequence length

$$Pr(m) = Ap^m + B \tag{7.20}$$

where SPAM errors have been absorbed into the coefficients A and B. Because the Clifford group forms a 2-design, randomly sampling over sequences approximates the discrete twirl over $\mathcal{C}$ which is equal to the full twirl over the Haar measure. The average error rate is invariant under the twirl and therefore the twirled noise channel is equal to the completely depolarizing channel with error rate

$$r = \frac{(d-1)(1-p)}{d}. \tag{7.21}$$

By fitting the survival probability to extract p, we learn the average error rate over the Clifford group r.

In a variant on this protocol where the application of the inversion gate is omitted, the decay parameter becomes the *unitarity* of the noise. The unitarity [63] provides a quantification of the coherence of noise and, taking a Pauli-Liouville representation, is defined as

$$\mathcal{U}(\mathcal{E}) = \frac{1}{d^2 - 1} \mathrm{Tr} \Lambda_u^{\dagger} \Lambda_u. \tag{7.22}$$

Together, the average error rate and unitarity give us insight into a noise process while making no assumptions about the specific model. In the following chapter, we see how randomized benchmarking may be combined with inversion recovery and Ramsey experiments to provide more information on the generalized damping channel.

Chapter 8

Characterization of the Generalized Damping Process

This chapter is based on a collaboration with Joel Wallman and Joseph Emerson.

8.1 Motivation

In order to achieve practical quantum computation, quantum systems must be precisely controlled despite being prone to intrinsic errors. Characterization of errors acting on a system is of the utmost importance in the quest to increase control. Complete characterization of a noise process is achievable via process tomography [43, 44], however this method requires exponentially growing resources and is susceptible to state preparation and measurement (SPAM) errors [45]. Consequently, there are two general approaches to error characterization in practice. In the first approach, we assume a specific noise model and estimate the corresponding parameters. For example, T_1 and T_2^* times are often quoted figures of merit used to characterize amplitude damping and dephasing errors [46], however their measurement may assume that they are the only noise processes present. In the second approach, we do not assume any noise model and estimate summary statistics. Randomized benchmarking [47, 48, 22] estimates the average error rate on a set of operations, however no information on the source of noise is gained. Lastly, we wish to connect experimental noise parameters obtained via either method to threshold theorems [50, 51]. The diamond distance [20] of a noise channel from the identity gives information on the worst case errors, which is often needed to make meaningful statements about threshold theorems, however it is not directly accessible by experiment.

Here we bridge the gap between these two methods and provide a link between T_1 and T_2^* estimation, randomized benchmarking, and the diamond distance. We consider cases where noise is close to the generalized damping channel, which allows for both amplitude damping and dephasing processes at finite temperature. The goal is to obtain accurate generalized damping parameters, relate these to the expected error rate and unitarity values, and thus validate the use of threshold theorems. Firstly, we consider how T_1 and T_2^* estimation can be affected by additional unknown noise. We show that T_1 estimation is mostly insensitive to additional noise sources, while T_2^* estimation can be affected greatly when measured via a traditional Ramsey experiment. We propose an alternative fidelity measurement, derive fit models, and provide a numerical comparison between the two methods. Secondly, we show how T_1 and T_2^* parameters are related to the error rate and unitarity values obtained from a randomized benchmarking experiment. Comparison of experimental measurements to these goal values provides a benchmark to check whether noise is limited by T_1 and T_2^* processes. Finally, we derive upper bounds on the diamond distance of the generalized damping channel to the identity. If the goal error rate and unitarity values are met, the application of these bounds and comparison to fault-tolerance threshold theorems are justified.

8.2 Perturbing the generalized damping model

Generalized damping is believed to account for the majority of the intrinsic noise in experimental systems, as shown by the prominence of T_1 and T_2^* values. However, it is likely that there are other unknown noise sources, which can be viewed as perturbations from a generalized damping process as follows. Referring to the general construction of a master equation (2.20), the generalized damping channel is described by the coefficient matrix

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \lambda\Gamma_1 & 0 & 0 \\ 0 & 0 & (1-\lambda)\Gamma_1 & 0 \\ 0 & 0 & 0 & \Gamma_2 \end{pmatrix} \tag{8.1}$$

in the basis $\{\mathbb{1}/\sqrt{2}, \sigma_+, \sigma_-, Z/\sqrt{2}\}$. The Hamiltonian can be estimated separately [52, 53], so we set $H = 0$ and so delete the first row and column of A. A general master equation with coefficient matrix

$$A = \begin{pmatrix} \lambda\Gamma_1 & \alpha_r - i\alpha_i & \beta^* \\ i\alpha_i + \alpha_r & (1-\lambda)\Gamma_1 & \delta^* \\ \beta & \delta & \Gamma_2 \end{pmatrix} \tag{8.2}$$

is then a perturbation about a generalized damping channel if the off-diagonal elements of A are small compared to the diagonal elements. We can also assume all variables in A are real without loss of generality by including any phase from β and δ into $\sigma_\pm$ respectively.

For a general state $\rho = (I + \boldsymbol{r}\cdot\boldsymbol{\sigma})/2$ where $\boldsymbol{r}$ is the Bloch vector and $\boldsymbol{\sigma} = (X, Y, Z)$, substituting (8.2) into (2.20) gives

$$\frac{\partial \boldsymbol{r}}{\partial t} = C\boldsymbol{r} + \boldsymbol{\lambda} \tag{8.3}$$

where

$$C = \begin{pmatrix} \alpha_r - \Gamma_2^* & \alpha_i & \beta \\ \alpha_i & -\alpha_r - \Gamma_2^* & 0 \\ \beta & 0 & -\Gamma_1 \end{pmatrix},$$
$$\boldsymbol{\lambda} = \begin{pmatrix} 2\sqrt{2}\delta - 2\beta \\ 0 \\ \Gamma_1(2\lambda - 1) \end{pmatrix}, \tag{8.4}$$

$\Gamma_2^* = \frac{\Gamma_1}{2} + \Gamma_2$, and we have redefined $\frac{\beta+\delta}{\sqrt{2}} \to \beta$. As C is Hermitian, it has an orthonormal set of eigenvectors $V = (v_1, v_2, v_3)$ with associated eigenvalues $\boldsymbol{\eta} = (\eta_1, \eta_2, \eta_3)$. Therefore, under a generic master equation, an initial state ρ with Bloch vector $\boldsymbol{r}(0) = \sum_j c_j(0)v_j$ evolves to the state with Bloch vector $\boldsymbol{r}(t) = \sum_j c_j(t)v_j$ at time t where

$$\frac{\partial c_j(t)}{\partial t} = \eta_j c_j(t) + \lambda_j, \tag{8.5}$$

which can be solved to give

$$c_j(t) = \frac{\lambda_j(e^{\eta_j t} - 1)}{\eta_j} + c(0)e^{\eta_j t} \tag{8.6}$$

The three different eigenvalues result in three characteristic decay time scales.

For example, when $\alpha_i = \beta = \delta = 0$ (more general than normally considered as α_r can be nonzero), $V = \mathbb{1}_3$ and so a state with Bloch vector $\boldsymbol{r} = (r_x, r_y, r_z)$ evolves to a state with Bloch vector components

$$\begin{aligned} r_x(t) &= r_x e^{t\left(\alpha_r - \Gamma_2^*\right)}, \\ r_y(t) &= r_y e^{t\left(-\left(\Gamma_2^* + \alpha_r\right)\right)}, \\ r_z(t) &= r_z e^{-\Gamma_1 t} - (2\lambda - 1)\left(e^{-\Gamma_1 t} - 1\right). \end{aligned} \tag{8.7}$$

Any non-zero value of α_r splits the degeneracy of the first two eigenvalues so that there is no longer a single T_2^* lifetime. This is reflected in the sensitivity of T_2^* experiments to perturbations as discussed in section 8.4.

Any experiment to estimate the generalized damping parameters will include SPAM errors. Throughout our analysis, we assume that SPAM errors result in the preparation and measurement

$$\begin{aligned} \rho &\to (1-k)\psi + k\psi^{\perp} \\ M &\to (1-n_1)M + n_2 I \end{aligned} \tag{8.8}$$

instead of the ideal versions ψ and M respectively, where $\psi^{\perp}$ is the state orthogonal to ψ and k and n_1 are small positive constants and n_2 is a small constant. This assumption can be made rigorous by performing a random operation that leaves ψ invariant immediately after the state is prepared and a random operation that leaves M invariant immediately before the measurement, where both of these random choices are made independently for each experimental run.

8.3 T_1 estimation

Under generalized damping, the output from an inversion recovery experiment (section 7.2.1) is

$$Q_0(t) = 2\lambda e^{-\Gamma_1 t} - 2\lambda + 1 \tag{8.9}$$

by (8.7) with $\boldsymbol{r} = (0, 0, 1)$. Therefore $T_1 = 1/\Gamma_1$ and λ can be estimated by estimating $Q(t)$ for multiple values of t and fitting to (8.9). In the presence of SPAM errors as in (8.8), (8.9) becomes

$$Q(t) = (1 - n_1)\, Q_0(t) - n_2, \tag{8.10}$$

which can be used to obtain Γ_1.

We now prove the robustness of the T_1 estimate to perturbations about generalized amplitude damping. For $\alpha_i = 0$, $Cv = \eta v$ where

$$\begin{aligned} v &= \left(\frac{-\sqrt{2}\beta}{\sqrt{4\beta^2 + g\left(\sqrt{4\beta^2+g^2}+g\right)}}, 0, \frac{\sqrt{\frac{g}{\sqrt{4\beta^2+g^2}}+1}}{\sqrt{2}} \right), \\ \eta &= \frac{1}{2}\left(g - \sqrt{4\beta^2+g^2}\right) - \Gamma_1, \end{aligned} \tag{8.11}$$

Table 8.1: Numerical estimation of Γ_1 obtained from simulation of a population inversion experiment.

	$Q(t)$ measurement (8.10)
Γ_1^{avg}	0.104657
Γ_1^{var}	3.670493E−5
R^2_{adj}	0.999997

and $g = \alpha_r + \Gamma_1 - \Gamma_2^*$. For $\alpha_i \neq 0$,

$$Cv = \eta v - \frac{\sqrt{2}\beta\alpha_i e_2}{\sqrt{4\beta^2 + g\left(\sqrt{4\beta^2+g^2}+g\right)}} \tag{8.12}$$

where $\{e_k\}$ is the canonical basis of $\mathbb{R}^3$, so there exists an eigenvector u with eigenvalue κ such that $u - v \in O(\alpha_i\beta)$ and $\eta - \kappa \in O(\alpha_i\beta)$.

The eigenvector v is the most relevant for T_1 estimation because

$$\begin{aligned} c_3(0) = e_3^\dagger v &= \frac{\sqrt{\frac{g}{\sqrt{4\beta^2+g^2}}+1}}{\sqrt{2}} + O(\alpha_i\beta) \\ &\approx 1 \end{aligned} \tag{8.13}$$

where we have used $g \gg \beta$, which holds near a generalized damping channel. Therefore the overlap with the other eigenvectors is negligible and so the decay rate of $Q(t)$ will be dominated by $\kappa = -\Gamma_1 + O(\beta^2)$. Therefore fitting $Q(t)$ to characterize a T_1 time is robust to perturbations around generalized damping channels. We demonstrate the robustness numerically in the following section.

8.3.1 Numerical simulation

For these simulations, we have used $\Gamma_1 = 0.1$, and all other noise components were chosen uniformly at random with intervals $\alpha_r, \alpha_i, \beta, \delta \in [0, \Gamma_1/2]$, $\Gamma_2 \in [0, 2\Gamma_1]$, and $\lambda \in [0.8, 1]$. Measurements were taken over a time interval on the order of T_1, using 100 data points. Values of Γ_1^{avg} and R^2_{adj} are averaged over 2000 noise matrices, while Γ_1^{var} provides the variance in the results. From the results of the numerical simulation (Tab. 8.1), we can see that (8.10) provides a good fit to $Q(t)$ under the noisy generalized damping channel, and extracts the value of Γ_1 accurately with a high probability. Fig. 8.1 shows a histogram of

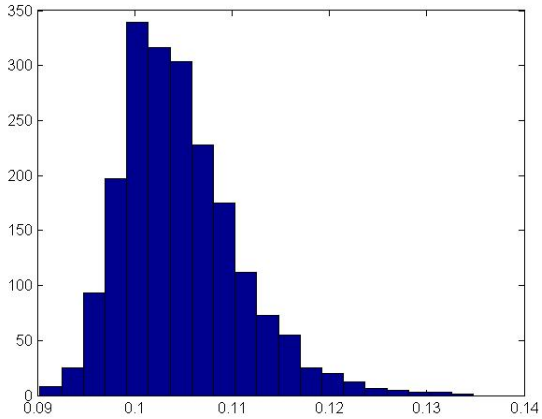


Figure 8.1: Histogram showing the estimated values of Γ_1 ($\Gamma_1^{act} = 0.1$) for 2000 random noise matrices.

the results. We find that the conclusions drawn from this simulation are unchanged when the noise distribution is varied or when finite sampling effects are taken into consideration.

8.4 T_2^* estimation

Under generalized damping, the expected $Q(t)$ obtained from a Ramsey experiment (section 7.2.2) is

$$Q(t) = e^{-t(\Gamma_2+\Gamma_1/2)} \tag{8.14}$$

where $T_2^* = 1/(\Gamma_2 + \Gamma_1/2)$.

We would like to check which additional components to the coefficient matrix (8.2) can affect the experimentally measured quantity $Q(t) = tr(\rho(t)Z)$. It was found that α enters $Q(t)$ to first order. Taking (8.2) with $\beta, \delta = 0, \alpha \neq 0$ we obtain

$$Q(t) = \frac{1}{2}e^{-t(\Gamma_1/2+\Gamma_2)}[e^{-|\alpha|t}(1 - \alpha_r/|\alpha|) + e^{|\alpha|t}(1 + \alpha_r/|\alpha|)]. \tag{8.15}$$

Thus, the use of $Q(t)$ to characterize a T_2^* time via measurement of Γ_2 is not very robust to additional noise. We propose obtaining Γ_2 using a fidelity measurement as follows.

8.4.1 Fidelity measurement

In order to minimize the effect of additional noise on Γ_2 estimation, we propose measuring the fidelity quantity

$$F = \langle +| \mathcal{E}(|+\rangle\langle +|) |+\rangle + \langle +i| \mathcal{E}(|+i\rangle\langle +i|) |+i\rangle \tag{8.16}$$

where measurement is understood to be in the rotating frame. In essence, this is equivalent to performing two Ramsey experiments with orthogonal preparations. To first order in the additional noise parameters, F is expected to be

$$F = e^{-\phi_0 t} + 1 \tag{8.17}$$

where $\phi_0 = \Gamma_2 + \Gamma_1/2$. Projective state preparation and measurement errors manifest themselves through the leading coefficients only so that the fit model becomes

$$F = Ae^{-\phi_0 t} + B. \tag{8.18}$$

In the case that Hahn-echo or CPMG pulse sequences have been applied, we need to modify (8.16) as

$$F = \langle +| S\mathcal{E}(|+\rangle\langle +|)S |+\rangle + \langle +i| S\mathcal{E}(|+i\rangle\langle +i|)S |+i\rangle \tag{8.19}$$

where S is the additional pulse sequence that has been applied.

8.4.2 Numerical comparison

In this section we compare the results of Γ_2 estimation from a Ramsey experiment with our proposed fidelity measurement. To allow for projective SPAM errors in the Ramsey experiment, we take a fit model of the form

$$Q(t) = Ae^{-t\phi_0}. \tag{8.20}$$

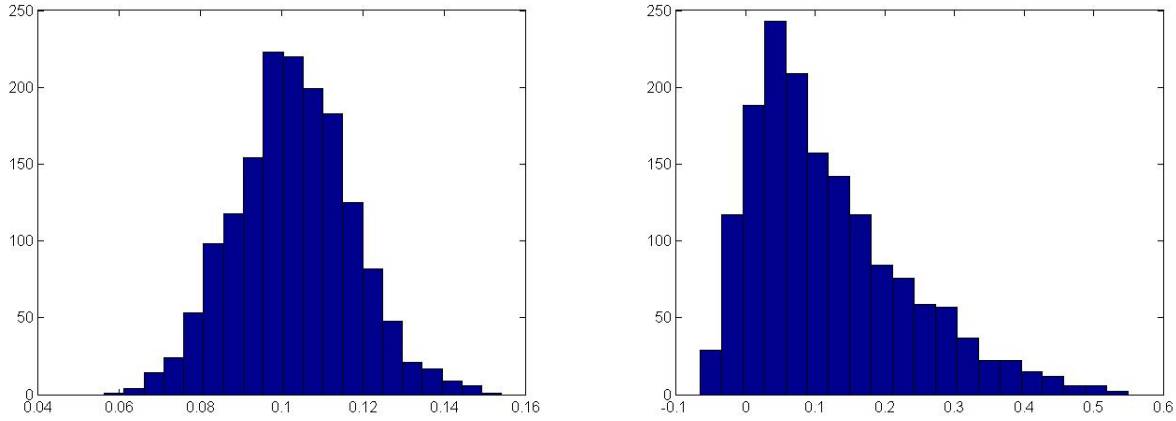
The fidelity measurement is fit using the form of (8.18). In both cases we obtain the estimate

$$\Gamma_2 = \phi_0 - \Gamma_1/2 \tag{8.21}$$

from the fit parameter ϕ_0 and previously estimated value of Γ_1.

Table 8.2: Numerical estimation of Γ_2 obtained from simulation of a Ramsey experiment compared to the proposed fidelity measurement.

	Fidelity meas. (8.18)	Ramsey exp. (8.20)
Γ_2^{avg}	0.102939	0.118594
Γ_2^{var}	2.099973E−4	0.012631
R^2_{adj}	0.999997	0.979670



(a) (b)

Figure 8.2: Histograms showing estimated values of Γ_2 ($\Gamma_2^{act} = 0.1$) for 1600 random noise matrices using a) fidelity measurement and b) a Ramsey experiment.

For these simulations we used $\Gamma_2 = 0.1$, Γ_1 was chosen uniformly at random in the interval $[0, 2\Gamma_2]$ and λ was chosen uniformly at random in the interval $[0.8, 1]$. All other noise components were chosen uniformly at random in the interval $[0, \Gamma_2/2]$. Measurements for both experiments were taken over a time interval on the order of T_2^*, using 100 data points. Values of Γ_2^{avg} and R^2_{adj} are averaged over 1600 noise matrices, while Γ_2^{var} provides the variance in the results.

Based on the R^2_{adj} values (Tab. 8.2), we can conclude that the fidelity measurement provides a much better fit to the observed data. Furthermore the fidelity measurement obtains a more accurate estimate of Γ_2, and obtains that value with a high probability based on the small variance. In contrast, the Ramsey experiment provides an inaccurate estimate, and is less reliable based on the large variance. The results are illustrated in Fig. 8.2.

8.4.3 Non-Markovian dephasing

The previous analysis of a Ramsey experiment and fidelity measurement assume Markovian noise, however we must also include the possibility of non-Markovian dephasing. To model the dephasing process, consider the Hamiltonian

$$H_{SB} = \hbar A(t)\sigma_z/2 \tag{8.22}$$

describing the interaction between the system and the bath. Here $A(t)$ is assumed to be a Gaussian random variable with zero mean $\langle A(t)\rangle = 0$ and covariance $S(t-t') = \langle A(t)A(t')\rangle$, where $\langle\cdot\rangle$ denotes averaging over the ensemble $\{A\}$. This Hamiltonian causes decay of the off-diagonal elements of $\rho(t)$ with corresponding Ramsey measurement

$$Q(t) = \exp\Big[-\int_{-\infty}^{+\infty} \frac{\sin^2(\omega t/2)}{\omega^2\pi}\tilde{S}(\omega)d\omega\Big] \tag{8.23}$$

where $\tilde{S}(\omega)$ is the Fourier transform of the covariance $S(t)$, also sometimes referred to as the spectral density or power spectrum of the noise [57, 58]. In the Markovian case where $\tilde{S}(\omega)$ can be approximated by a constant, we recover the exponential decay $Q(t) = e^{-t\phi_0}$ where $\tilde{S}(\omega) = \phi_0 = 1/T_2^*$. However in the non-Markovian case where memory effects are present, we obtain the Gaussian decay [58, 59]

$$Q(t) = Ae^{-(t\phi_0)^2} \tag{8.24}$$

which has been observed experimentally [60, 61, 62]. In the case that dephasing causes a Gaussian decay of off-diagonal elements, the fidelity measurement (8.16) is expected to take the form

$$F = Ae^{-(\phi_0 t)^2} + B. \tag{8.25}$$

We numerically simulate this case and again compare the performance of the two methods (Tab. 8.3).

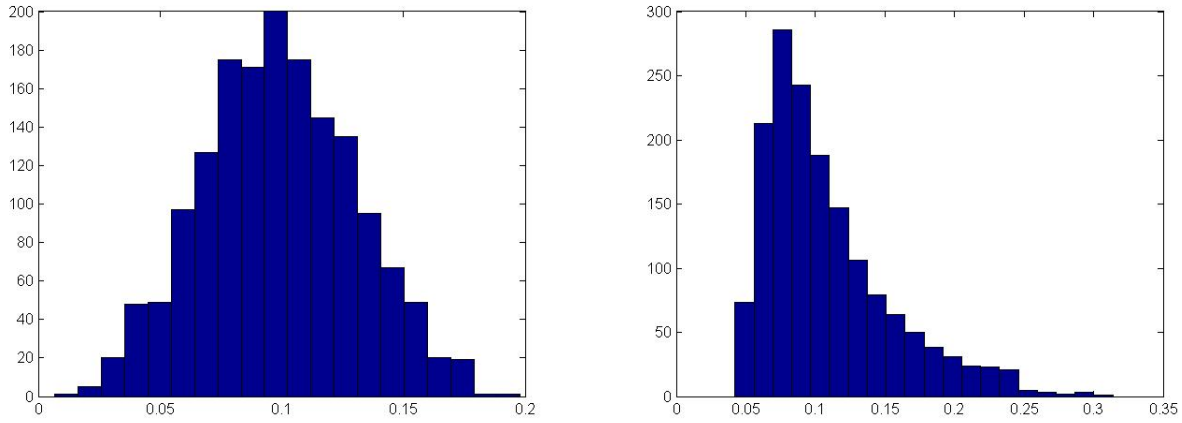
As in the Markovian case, we see that the fidelity measurement produces a better fit and provides a more accurate estimate of Γ_2 which is obtained with a higher probability. The results are illustrated in Fig. 8.3.

8.5 Linking T_1 and T_2^* estimations to RB

Following the analysis in the previous two sections, we now have accurate estimates for the generalized damping parameters $(\Gamma_1, \Gamma_2, \lambda)$. In this section we calculate the error rate

Table 8.3: Numerical estimation of Γ_2 obtained from simulation of a Ramsey experiment compared to the proposed fidelity measurement in the presence of non-Markovian dephasing.

	Fidelity meas. (8.25)	Ramsey exp. (8.24)
Γ_2^{avg}	0.098827	0.108287
Γ_2^{var}	9.694169E−4	0.002107
R^2_{adj}	0.999918	0.994462



(a) (b)

Figure 8.3: Histograms showing estimated values of Γ_2 ($\Gamma_2^{act} = 0.1$) for 1600 random noise matrices with non-Markovian dephasing using a) fidelity measurement and b) a Ramsey experiment.

and unitarity [63] expected in a randomized benchmarking experiment attributed to the generalized damping channel, which serve as goal values in the case when noise is limited by T_1 and T_2^* processes.

The error rate obtained from a randomized benchmarking experiment may be calculated as $r = 1 - \mathcal{F}$, where $\mathcal{F}$ is the average fidelity. For the generalized damping channel,

$$\begin{aligned} r_{\text{goal}} &= \frac{1}{2} - \frac{1}{6}e^{-\Gamma_1 \Delta t} - \frac{1}{3}e^{-(\Gamma_1/2+\Gamma_2)\Delta t} \\ &= \frac{1}{3} + \frac{\gamma_1}{6} - \frac{1}{3}\sqrt{(1-\gamma_1)(1-\gamma_2)}. \end{aligned} \tag{8.26}$$

The error rate obtained from an RB experiment may be compared to r_{goal} where Δt would be the average duration needed to apply one gate. Additionally, we can obtain more benchmark values if we perform the randomized benchmarking procedure with sequences of random Pauli gates and an additional projection step, as in [64]. The error rate r^σ represents the value obtained with the projection $\frac{1}{2}(\mathbb{1}+\sigma)$ applied at the end of the sequence where $\sigma \in \{X, Y, Z\}$:

$$r^X_{\text{goal}} = \frac{1}{2} - \frac{1}{6}\sqrt{(1-\gamma_1)(1-\gamma_2)} \tag{8.27}$$

$$r^Y_{\text{goal}} = \frac{1}{2} - \frac{1}{6}\sqrt{(1-\gamma_1)(1-\gamma_2)} \tag{8.28}$$

$$r^Z_{\text{goal}} = \frac{1}{2} - \frac{1}{6}(1-\gamma_1) \tag{8.29}$$

Similarly, the goal unitarity is given by

$$\mathcal{U}_{\text{goal}} = \frac{1}{3}\text{Tr}(\Lambda_u^\dagger \Lambda_u) = \frac{1}{3}(3 - 4\gamma_1 - 2\gamma_2 + 2\gamma_1\gamma_2 + \gamma_1^2). \tag{8.30}$$

Together these values provide a benchmark to check whether noise is limited by T_1 and T_2^* processes.

8.6 The diamond distance

In the previous section, we provide a way to test whether noise is limited by the generalized damping process. If that scenario holds, we would like some way to compare worst-case

gate errors to the fault-tolerance threshold theorem. In this section we provide upper bounds on the diamond distance between the generalized damping and identity channels, which accomplishes this goal.

Using two derivation techniques (see Appendix A) we obtain two upper bounds for $D_{GD} = \frac{1}{2}||\mathbb{1} - \Lambda_{GD}||_\diamond$

$$D_{GD} \leq \frac{1}{2}\Big[1 - \sqrt{(1-\gamma_1)(1-\gamma_2)} - \frac{1}{2}\gamma_1 + 2\lambda\gamma_1\Big] \tag{8.31}$$

$$D_{GD} \leq \max\begin{cases} 1 - \sqrt{(1-\gamma_1)(1-\gamma_2)} + \gamma_1(1-\lambda)) \\ 1 - \lambda(1-\gamma_1-\gamma_2) \end{cases}. \tag{8.32}$$

A lower bound may be obtained by evaluating $D_{GD} = \frac{1}{2}\sup_\rho||\mathbb{1} \otimes (I - \Lambda_{GD})(\rho)||_1$ for a given ρ. With $\rho = |01\rangle\langle 01|$, we obtain

$$D_{GD} \geq \lambda\gamma_1 \tag{8.33}$$

which we find to be numerically close to the actual diamond distance in the parameter regime of interest. Depending on the values of γ_1, γ_2 and λ, the minimum of (8.31) and (8.32) provides the upper bound on the diamond distance, assuming the noise is described by the generalized damping channel. In the more realistic case that the noise is not fully described by the generalized damping channel, we now provide an extension to the bounds (8.31 - 8.32) based on the actual measurements of the error rates r^σ and unitarity $\mathcal{U}$. Writing the total channel as

$$\Lambda = \Lambda_{GD} + \begin{pmatrix} 0 & 0 \\ \alpha & \mathcal{E} \end{pmatrix} \tag{8.34}$$

we can bound the diamond norm of Λ from the identity as

$$||\mathbb{1} - \Lambda||_\diamond \leq 2UB_{GD} + \sum_i |\alpha_i| + \sum_{i,j} |\mathcal{E}_{i,j}| \tag{8.35}$$

which is obtained by splitting Λ into Λ_{GD} plus a sum of one element matrices and applying the triangle inequality. The diamond norm of a matrix with a single non-zero entry in either the unital block or non-unital vector is given by the absolute value of that element as shown in Appendix B. Applying the Cauchy-Schwarz inequality

$$||\mathbb{1} - \Lambda||_\diamond \leq 2UB_{GD} + \sqrt{\sum_i |\alpha_i|^2} + \sqrt{\sum_{i,j} |\mathcal{E}_{i,j}|^2} \tag{8.36}$$

$$||\mathbb{1} - \Lambda||_\diamond \leq 2UB_{GD} + \sqrt{9r^2/2 - \gamma_1^2(2\lambda - 1)^2} + \sqrt{\sum_{i,j} |\mathcal{E}_{i,j}|^2} \tag{8.37}$$

since $\sqrt{\sum_i |\Lambda_i^n|^2} \leq 3r/\sqrt{2}$ by theorem 3 of [66], where Λ^n is the non-unital vector of Λ. Finally we obtain $\sqrt{\sum_{i,j} |\mathcal{E}_{i,j}|^2}$ from the measured values of r^σ and $\mathcal{U}$ compared to the goal values:

$$\begin{aligned}\sum_{i,j} |\mathcal{E}_{i,j}|^2 = \mathcal{U} - \mathcal{U}_{\text{goal}} - 12(1-\gamma_1)(r^Z_{\text{goal}} - r^Z) \\ - 12\sqrt{(1-\gamma_1)(1-\gamma_2)}(r^X_{\text{goal}} - r^X + r^Y_{\text{goal}} - r^Y).\end{aligned} \tag{8.38}$$

Also note that $||\mathbb{1} - \Lambda||_\diamond$ can be bounded in terms of the measured r and $\mathcal{U}$ directly as in [66]. A comparison of the bounds obtained from (8.37) and [66] is shown in Fig. 8.4. The upper bound given by (8.37) is only beneficial when the actual channel is very close to the generalized damping channel. Comparison of the bounds therefore provides another test of whether the noise is well described by generalized damping.

8.7 Summary

In summary, we have provided an in depth analysis of how commonly measured experimental quantities should behave for a channel close to generalized damping. Comparison to goal values provides a true benchmark for whether noise is limited by T_1 and T_2^* processes. Bounds on the diamond distance of the generalized damping channel from the identity provide a way to relate these noise parameters to fault-tolerance threshold theorems.

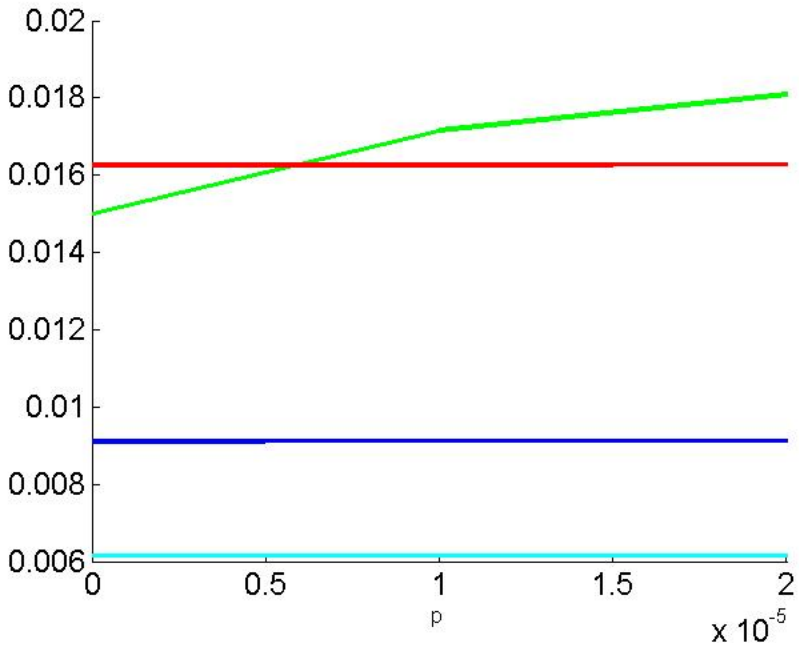


Figure 8.4: The diamond distance (blue) from the identity for random channels close to the generalized damping channel ($\Lambda = (1-p)\Lambda_{GD} + p\Lambda_{\text{rand}}$) as a function of p. Upper bounds are given by (8.37) (green) and [66] (red), and the lower bound is given by [66] (cyan). Each data point is averaged over 100 random channels.

Chapter 9

Conclusion

To summarize the results of this thesis, we have provided progress in two major areas related to fault-tolerant quantum computation. Firstly, we have studied magic state distillation as a method of achieving universality. One major goal is to distinguish a resource which is responsible for the quantum advantage over classical computation. We take the first step towards proving that contextuality is this resource for prime dimensional systems by demonstrating that tight distillation thresholds exist in at least some directions. Finding tight thresholds in the remaining directions will likely not be possible by brute force search, considering that the generalized T-state has not been shown to be distillable at all by any of the codes we searched over. Finding codes which distill the generalized T-state will likely require making use of transversality techniques, or generalizing qubit codes which distill qubit T-states. Another future direction would be to generalize the results of [37, 34], in an attempt to further limit the region of possibly distillable states. Secondly, we have provided many new qubit distillation routines based on small stabilizer codes. Although these exhibit linear error suppression, they may still be useful for some applications. These routines use fewer measurements and local operations per round of distillation, and therefore may be used in experimental systems where the size of the system is limited. The majority of magic states we found were unusual equatorial states, meaning they can implement rotations other than the usual $\pi/4$ rotation given by the H-state. It would be interesting to combine these routines and attempt to distill other equatorial states that are known to be useful, such as the V-state. Another interesting result was finding a 5-qubit code that distills the H-state with a tight threshold, making this the smallest known tight qubit code. Studying the properties of small codes exhibiting tight thresholds may illuminate the path to finding tight thresholds in higher dimensions. In the final part of this thesis, we study noise processes close to the generalized damping

channel and provide a link between parameter estimation, randomized benchmarking, and the diamond distance. Many experimentalists report T_1 and T_2^* times and can make use of the analysis presented here to compare these values to randomized benchmarking metrics. Such a comparison provides insight into how closely noise is limited by T_1 and T_2^* processes. The final ingredient is to make use of randomized benchmarking metrics to bound the diamond distance of the noise channel from the identity. By bounding the diamond distance, experimentalists may safely relate their results to fault-tolerance threshold theorems.

References

[1] E. Knill, "Quantum computing with realistically noisy devices" Nature **434**, 39 (2005).

[2] S. Bravyi and A. Kitaev, "Universal quantum computation with ideal Clifford gates and noisy ancillas" Phys. Rev. A **71**, 022316 (2005).

[3] D. Gottesman, "The Heisenberg Representation of Quantum Computers" Group22: Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics, eds. S. P. Corney, R. Delbourgo, and P. D. Jarvis, pp. 32-43 (Cambridge, MA, International Press, 1999). arXiv:quant-ph/9807006

[4] H. Anwar, E. T. Campbell, and D. E. Browne, "Qutrit magic state distillation " New J. Phys. **14**, 063006 (2012).

[5] E. T. Campbell, "Enhanced Fault-Tolerant Quantum Computing in d-Level Systems" Phys. Rev. Lett. **113**, 230501, (2014).

[6] E. T. Campbell, H. Anwar, and D. E. Browne, "Magic-State Distillation in All Prime Dimensions Using Quantum Reed-Muller Codes" Phys. Rev. X **2**, 041021 (2012).

[7] M. Howard, J. Wallman, V. Veitch, and J. Emerson, "Contextuality supplies the magic for quantum computation" Nature **510**, 351355 (2014).

[8] V. Veitch, S. A. H. Mousavian, D. Gottesman, and J. Emerson, "The resource theory of stabilizer quantum computation " New J. Phys. **16**, 013009 (2014).

[9] H. Dawkins and M. Howard, "Qutrit Magic State Distillation Tight in Some Directions" Phys. Rev. Lett **115**, 030501 (2015).

[10] M. Howard and H. Dawkins, "Small codes for magic state distillation" Eur. Phys. J. D **70**, 55 (2016).

[11] D. M. Appleby, I. Bengtsson, and S. Chaturvedi, "Spectra of phase point operators in odd prime dimensions and the extended Clifford group " J. Math. Phys. **49**, 012102 (2008).

[12] W. K. Wootters, "A Wigner-function formulation of finite-state quantum mechanics" Ann. Phys. **176**, 1 (1987).

[13] D. Gross, "Hudsons theorem for finite-dimensional quantum systems" J. Math. Phys. **12** 47 122107 (2006).

[14] V. Veitch, C. Ferrie, D. Gross, and J. Emerson, "Negative quasi-probability as a resource for quantum computation " New J. Phys. **14**, 113011 (2012).

[15] J. Emerson, "Theory of Open Quantum Systems with Applications" University course notes, 2015.

[16] H. -P. Breuer and F. Petruccione, "The Theory of Open Quantum Systems" Oxford University Press, 2002.

[17] D. Gottesman, "Surviving as a Quantum Computer in a Classical World" University course notes, 2013.

[18] A. Cross, G. Smith, J. A. Smolin, and B. Zeng, "Codeword Stabilized Quantum Codes" Information Theory, IEEE Transactions on **55**, 433-438 (2009).

[19] D. Aharonov and M. Ben-Or, "Fault-tolerant quantum computation with constant error" 29th ACM Symp. on Theory of Computing (STOC), 176-188 (1997).

[20] A. Kitaev, "Quantum computations: algorithms and error correction" Russ. Math. Surv. **52**, 1191-1249 (1997).

[21] J. Watrous, "Semidefinite programs for completely bounded norms" Theory of Comput. **5**, 217-238 (2009).

[22] E. Magesan, J. M. Gambetta, and J. Emerson, "Characterizing quantum gates via randomized benchmarking," Phys. Rev. A **85**, 042311 (2012).

[23] P. Shor, "Fault-tolerant quantum computation" in *Foundations of Computer Science, 1996. Proceedings., 37th Annual Symposium on* (IEEE, 1996), pp. 56-65.

[24] A. Kitaev, "Fault-tolerant quantum computation by anyons" Ann. Phys. **303**, 2 (2013).

[25] R. Raussendorf and J. Harrington, "Fault-Tolerant Quantum Computation with High Threshold in Two Dimensions" Phys. Rev. Lett. **98**, 190504 (2007).

[26] V. Kliuchnikov, D. Maslov, and M. Mosca, "Asymptotically Optimal Approximation of Single Qubit Unitaries by Clifford and T Circuits Using a Constant Number of Ancillary Qubits" Phys. Rev. Lett. **110**, 190502 (2013).

[27] V. Kliuchnikov, D. Maslov, and M. Mosca, "Practical Approximation of Single-Qubit Unitaries by Single-Qubit Quantum Clifford and T Circuits" IEEE Transactions on Computers **65**, 161-172 (2015).

[28] A. Paetznick and B. W. Reichardt, "Universal Fault-Tolerant Quantum Computation with Only Transversal Gates and Error Correction" Phys. Rev. Lett. **111**, 090505 (2013).

[29] T. Jochym-OConnor and R. Laflamme, "Using Concatenated Quantum Codes for Universal Fault-Tolerant Quantum Gates" Phys. Rev. Lett. **112**, 010505 (2013).

[30] C. Chamberland, T. Jochym-OConnor and R. Laflamme, "Thresholds for Universal Concatenated Quantum Codes" Phys. Rev. Lett. **117**, 010501 (2016).

[31] A. M. Meier, B. Eastin, and E. Knill, "Magic-state distillation with the four-qubit code" Quantum Information & Computation **13**, 195 (2013).

[32] S. Bravyi, and J. Haah, "Magic-state distillation with low overhead" Phys. Rev. A **86**, 052329 (2012).

[33] C. Jones, "Multilevel distillation of magic states for quantum computing" Phys. Rev. A **87**, 042305 (2013).

[34] E. T. Campbell and D. E. Browne, "On the Structure of Protocols for Magic State Distillation" Lecture Notes in Computer Science **5906**, 20 (2009).

[35] M. Howard and J. Vala "Qudit versions of the qubit "pi-over-eight" gate" Phys. Rev. A **86**, 022316 (2012).

[36] B. W. Reichardt, "Quantum Universality from Magic States Distillation Applied to CSS Codes" Quantum Information Processing **4**, 251 (2005).

[37] E. T. Campbell and D. E. Browne, "Bound States for Magic State Distillation in Fault-Tolerant Quantum Computation" Phys. Rev. Lett. **104**, 030503 (2010).

[38] W. van Dam and M. Howard, "Noise thresholds for higher-dimensional systems using the discrete Wigner function" Phys. Rev. A **83**, 032310 (2011).

[39] Ben W. Reichardt "Quantum universality by state distillation" Quantum Inf. Comput. 9:1030-1052, (2009)

[40] A. Bocharov, Y. Gurevich and K. M. Svore, "Efficient decomposition of single-qubit gates into V basis circuits" Phys. Rev. A **88**, 012313 (2013).

[41] G. Duclos-Cianci and K. Svore, "Distillation of nonstabilizer states for universal quantum computation" Phys. Rev. A **88**, 042325 (2013).

[42] Jonas T. Anderson and Tomas Jochym-O'Connor, "Classification of transversal gates in qubit stabilizer codes" Quantum Information & Computation **16**, 0771-0802 (2016).

[43] I. L. Chuang and M. A. Nielsen, "Prescription for experimental determination of the dynamics of a quantum black box" J. Mod. Opt. **44**, 2455 (1997).

[44] J. Poyatos, J. Cirac, and P. Zoller, "Complete Characterization of a Quantum Process: The Two-Bit Quantum Gate" Phys. Rev. Lett. **78**, 390 (1997).

[45] S. Merkel, J. Gambetta, J. Smolin, S. Poletto, A. Córcoles, B. Johnson, C. Ryan, and M. Steffen, "Self-consistent quantum process tomography" Phys. Rev. A **87**, 062119 (2013).

[46] T. Ladd, F. Jelezko, R. Laflamme, Y. Nakamura, C. Monroe, and J. O'Brien, "Quantum computers" Nature **464**, 45-53 (2010).

[47] J. Emerson, R. Alicki, and K. Zyczkowski, "Scalable noise estimation with random unitary operators," J. Opt. B Quantum Semiclassical Opt. **7**, S347 (2005).

[48] E. Magesan, J. M. Gambetta, and J. Emerson, "Scalable and Robust Randomized Benchmarking of Quantum Processes," Phys. Rev. Lett. **106**, 180504 (2011).

[49] E. Magesan, J. M. Gambetta, B. R. Johnson, C. A. Ryan, J. M. Chow, S. T. Merkel, M. P. da Silva, G. A. Keefe, M. B. Rothwell, T. A. Ohki, M. B. Ketchen, and M. Steffen, "Efficient Measurement of Quantum Gate Error by Interleaved Randomized Benchmarking," Phys. Rev. Lett. **109**, 080505 (2012).

[50] P. W. Shor, "Scheme for reducing decoherence in quantum computer memory" Phys. Rev. A **52**, R2493(R) (1995).

[51] S. Devitt, W. Munro, and K. Nemoto, "Quantum error correction for beginners" Rep. Prog. Phys. **76** 076001 (2013).

[52] S. G. Schirmer, A. Kolli, and D. K. L. Oi, "Experimental Hamiltonian identification for controlled two-level systems", Phys. Rev. A **69**, 050306(R) (2004).

[53] L. E. de Clercq, R. Oswald, C. Flühmann, B. Keitch, D. Kienzler, H. -Y. Lo, M. Marinelli, D. Nadlinger, V. Negnevitsky, and J. P. Home, "Estimation of a general time-dependent Hamiltonian for a single qubit", Nature Comm. **7**, 11218 (2016).

[54] E. L. Hahn, "Spin echoes" Phys. Rev. **80**, 580-594 (1950).

[55] H. Carr and E. Purcell, "Effects of diffusion on free precession in the nuclear magnetic resonance experiments" Phys. Rev. **94**, 630-638 (1954).

[56] S. Meiboom and D. Gill, "Modified spin-echo method for measuring nuclear relaxation times" Rev. Sci. Instrum. **29**, 688-691 (1958).

[57] J. Preskill, "Physics 219/Computer Science 219 Quantum Computation" University course notes, 2006.

[58] J. M. Taylor, "Dephasing" University course notes, 2008.

[59] P. Cappellaro, "22.51 Quantum Theory of Radiation Interactions Fall 2012", Massachusetts Institute of Technology: MIT OpenCourseWare, https://ocw.mit.edu. (2012).

[60] C. Barthel, J. Medford, C. Marcus, M. Hanson, and A. Gossard, "Interlaced dynamical decoupling and coherent operation of a singlet-triplet qubit" Phys. Rev. Lett. **105**, 266808 (2010).

[61] G. de Lange, Z. Wang, D. Riste, V. Dobrovitski, and R. Hanson, "Universal dynamical decoupling of a single solid-state spin from a spin bath" Science **330**, 60-63 (2010).

[62] J. Bylander, S. Gustavsson, F. Yan, F. Yoshihara, K. Harrabi, G. Fitch, D. Cory, Y. Nakamura, J. Tsai, and W. Oliver, "Dynamical decoupling and noise spectroscopy with a superconducting flux qubit" Nature Phys. **7** 565-570 (2011).

[63] J. Wallman, C. Granade, R. Harper, and S. T. Flammia, "Estimating the Coherence of Noise," New J. Phys. **17** 113020 (2015).

[64] A. Carignan-Dugas, J. Wallman, and J. Emerson, "Characterizing Universal Gate Sets via Dihedral Benchmarking," Phys. Rev. A **92**, 060302(R) (2015).

[65] J. J. Wallman and S. T. Flammia, "Randomized Benchmarking with Confidence" New J. Phys. **16**, 103032 (2014).

[66] J. J. Wallman, "Bounding experimental quantum error rates relative to fault-tolerant thresholds" arXiv:1511.00727v2 [quant-ph] (2016).

[67] R. Kueng, D. M. Long, A. C. Doherty and S. T. Flammia, "Comparing Experiments to the Fault-Tolerance Threshold" Phys. Rev. Lett. **117**, 170502 (2016).

APPENDICES

Appendix A

Stabilizer codes for edge-state distillation

Each set of generators describe a stabilizer code in $(x|z)$ notation which has edge state $|E\rangle = (a, b, b)$ as its limiting state. All limiting states lie outside the Wigner tetrahedron edge in the x-z plane as shown in Fig. 5.4.

a	b	x	z	Generators										
0.835983	0.388029	0.917501	0.397734	$G1$	1	0	0	0	0	0	1	2		
				$G2$	0	1	0	0	0	0	2	1		
				$G3$	0	0	1	2	2	1	0	0		
				Z_L	0	0	0	2	1	2	0	0		
				X_L	0	0	0	0	0	0	1	1		
0.941079	0.239134	0.636520	0.771260	$G1$	1	0	0	0	0	0	0	0	0	2
				$G2$	0	0	0	0	2	1	1	1	1	0
				$G3$	0	2	1	0	2	1	1	1	1	0
				$G4$	0	2	0	1	2	1	1	1	1	0
				Z_L	0	2	0	0	2	1	1	1	1	1
				X_L	0	0	0	0	0	0	1	1	1	0
0.952265	0.215861	0.581402	0.813616	$G1$	0	0	0	0	2	1	1	1	1	0
				$G2$	2	1	0	0	2	1	1	1	1	0
				$G3$	2	0	1	0	2	1	1	1	1	0
				$G4$	1	0	0	1	1	2	2	2	2	1
				Z_L	2	0	0	0	2	1	1	1	1	1
				X_L	0	0	0	0	0	1	1	1	2	0

0.842186	0.381263	0.908192	0.418554	$G1$	0	0	0	0	2	1	2	1	0	0
				$G2$	1	1	0	0	1	2	1	2	1	0
				$G3$	0	0	1	0	0	0	0	0	0	2
				$G4$	0	0	0	1	0	2	2	0	0	0
				Z_L	2	0	0	0	2	1	2	1	1	1
				X_L	0	0	0	0	0	1	2	0	0	0
0.881207	0.334271	0.833147	0.553052	$G1$	1	0	0	0	0	0	0	0	2	2
				$G2$	0	0	0	0	2	1	1	1	0	0
				$G3$	0	2	1	0	2	1	1	1	0	0
				$G4$	0	0	0	1	0	2	2	2	0	0
				Z_L	0	2	0	0	2	1	1	1	1	1
				X_L	0	0	0	0	0	0	1	1	0	0
0.811769	0.412936	0.948113	0.317936	$G1$	0	0	0	0	2	1	1	2	0	0
				$G2$	2	1	0	0	2	1	1	2	0	0
				$G3$	0	0	1	0	0	0	0	0	2	1
				$G4$	0	0	0	1	0	2	2	2	0	0
				Z_L	2	0	0	0	2	1	1	2	1	1
				X_L	0	0	0	0	0	1	1	0	0	0
0.964002	0.188015	0.512644	0.858601	$G1$	0	0	0	0	0	2	1	1	1	1	1	0
				$G2$	2	1	0	0	0	2	1	1	1	1	1	0
				$G3$	2	0	1	0	0	2	1	1	1	1	1	0
				$G4$	2	0	0	1	0	2	1	1	1	1	1	0
				$G5$	1	0	0	0	1	1	2	2	2	2	2	1
				Z_L	2	0	0	0	0	2	1	1	1	1	1	1
				X_L	0	0	0	0	0	0	1	1	1	1	2	0
0.921798	0.274124	0.714707	0.699424	$G1$	1	0	0	0	0	0	0	0	0	0	2	2
				$G2$	0	0	0	0	0	2	1	1	1	1	0	0
				$G3$	0	2	1	0	0	2	1	1	1	1	0	0
				$G4$	0	2	0	1	0	2	1	1	1	1	0	0
				$G5$	0	0	0	0	1	0	2	0	0	0	0	0
				Z_L	0	2	0	0	0	2	1	1	1	1	0	1
				X_L	0	0	0	0	0	0	0	1	1	1	0	0

0.943399	0.274124	0.622954	0.781983	$G1$	0	0	0	0	0	2	1	1	1	1	0	0
				$G2$	2	1	0	0	0	2	1	1	1	1	1	0
				$G3$	2	0	1	0	0	2	1	1	1	1	1	0
				$G4$	1	0	0	1	0	1	2	2	2	2	2	1
				$G5$	0	0	0	0	1	0	2	0	0	0	0	0
				Z_L	2	0	0	0	0	2	1	1	1	1	1	1
				X_L	0	0	0	0	0	0	1	1	1	2	0	0
0.863728	0.356353	0.622954	0.781983	$G1$	1	0	0	0	0	0	0	0	0	0	2	0
				$G2$	0	1	0	0	0	0	0	0	0	0	2	0
				$G3$	0	0	1	0	0	0	0	0	0	0	0	2
				$G4$	0	0	0	0	0	2	0	0	1	1	1	0
				$G5$	0	0	0	1	1	1	2	2	2	2	2	1
				Z_L	0	0	0	2	0	2	0	0	1	1	1	1
				X_L	0	0	0	0	0	0	0	0	0	1	2	0
0.872073	0.346041	0.853543	0.521023	$G1$	0	0	0	0	0	2	1	2	1	1	0	0
				$G2$	1	1	0	0	0	1	2	1	2	2	1	0
				$G3$	0	0	1	0	0	0	0	0	0	0	0	2
				$G4$	0	0	0	1	0	0	0	0	0	0	0	2
				$G5$	0	0	0	0	1	0	2	2	0	0	0	0
				Z_L	2	0	0	0	0	2	1	2	1	1	1	1
				X_L	0	0	0	0	0	0	1	2	0	0	0	0
0.899161	0.309443	0.786979	0.616980	$G1$	0	0	0	0	0	2	1	0	1	1	0	0
				$G2$	0	1	0	0	0	0	0	0	0	0	2	0
				$G3$	0	0	1	0	0	0	0	0	0	0	0	2
				$G4$	2	0	0	1	0	2	1	0	1	1	1	0
				$G5$	0	0	0	0	1	0	2	2	0	0	0	0
				Z_L	2	0	0	0	0	2	1	0	1	1	1	1
				X_L	0	0	0	0	0	0	1	0	0	1	0	0
0.782757	0.440052	0.974262	0.225417	$G1$	1	0	0	0	0	0	0	0	0	0	2	0
				$G2$	0	1	0	0	0	0	0	0	0	0	2	0
				$G3$	0	0	0	0	0	2	0	0	1	1	1	0
				$G4$	0	0	2	1	0	2	0	0	1	1	1	0
				$G5$	0	0	1	0	1	1	2	2	2	2	2	1
				Z_L	0	0	2	0	0	2	0	0	1	1	1	1
				X_L	0	0	0	0	0	0	0	0	1	1	2	0

				$G1$	0	0	0	0	0	2	1	0	1	1	0	0
				$G2$	2	1	0	0	0	2	1	0	1	1	0	1
				$G3$	0	0	1	0	0	0	0	0	0	0	0	2
0.925353	0.268068	0.701613	0.712558	$G4$	1	0	0	1	0	1	2	0	2	2	2	1
				$G5$	0	0	0	0	1	0	2	2	0	0	0	0
				Z_L	2	0	0	0	0	2	1	0	1	1	1	1
				X_L	0	0	0	0	0	0	1	1	0	2	0	0
				$G1$	0	0	0	0	0	2	1	0	1	1	0	0
				$G2$	2	1	0	0	0	2	1	0	1	1	0	1
				$G3$	2	0	1	0	0	2	1	0	1	1	1	0
0.945567	0.230113	0.615429	0.788192	$G4$	1	0	0	1	0	1	2	0	2	2	2	1
				$G5$	0	0	0	0	1	0	2	2	0	0	0	0
				Z_L	2	0	0	0	0	2	1	0	1	1	1	1
				X_L	0	0	0	0	0	0	1	1	1	2	0	0
				$G1$	0	0	0	0	0	2	1	0	1	2	0	0
				$G2$	2	1	0	0	0	2	1	0	1	2	0	1
				$G3$	2	0	1	0	0	2	1	0	1	2	1	0
0.931836	0.256594	0.676287	0.736638	$G4$	1	0	0	1	0	1	2	0	2	1	2	0
				$G5$	0	0	0	0	1	0	2	2	0	0	0	0
				Z_L	2	0	0	0	0	2	1	0	1	2	1	1
				X_L	0	0	0	0	0	0	1	1	1	2	0	0
				$G1$	1	0	0	0	0	0	0	0	0	0	2	2
				$G2$	0	1	0	0	0	0	0	0	0	0	2	2
				$G3$	0	0	0	0	0	2	1	1	1	1	0	0
0.919698	0.277629	0.722196	0.691689	$G4$	0	0	1	1	0	1	2	2	2	2	0	1
				$G5$	0	0	0	0	1	0	2	2	0	0	0	0
				Z_L	0	0	2	0	0	2	1	1	1	1	0	1
				X_L	0	0	0	0	0	0	0	0	1	2	0	0
				$G1$	1	0	0	0	0	0	0	0	0	0	2	2
				$G2$	0	0	0	0	0	2	1	1	1	1	0	0
				$G3$	0	1	1	0	0	1	2	2	2	2	2	1
0.934648	0.251429	0.664674	0.747134	$G4$	0	1	0	1	0	1	2	2	2	2	2	1
				$G5$	0	0	0	0	1	0	2	2	0	0	0	0
				Z_L	0	2	0	0	0	2	1	1	1	1	1	1
				X_L	0	0	0	0	0	0	0	1	2	2	0	0

0.923498	0.271248	0.708513	0.705698	$G1$	1	0	0	0	0	0	0	0	0	0	2	2
				$G2$	0	0	0	0	0	2	1	2	1	1	0	0
				$G3$	0	1	1	0	0	1	2	1	2	2	2	0
				$G4$	0	1	0	1	0	1	2	1	2	2	2	0
				$G5$	0	0	0	0	1	0	2	2	0	0	0	0
				Z_L	0	2	0	0	0	2	1	2	1	1	1	2
				X_L	0	0	0	0	0	0	0	1	2	2	0	0
0.939523	0.242173	0.643544	0.765409	$G1$	0	0	0	0	0	2	1	2	1	1	0	0
				$G2$	2	1	0	0	0	2	1	2	1	1	0	2
				$G3$	2	0	1	0	0	2	1	2	1	1	1	0
				$G4$	1	0	0	1	0	1	2	1	2	2	2	1
				$G5$	0	0	0	0	1	0	2	2	0	0	0	0
				Z_L	2	0	0	0	0	2	1	2	1	1	1	1
				X_L	0	0	0	0	0	0	1	1	1	2	0	0

Appendix B

Diamond distance bounds for the generalized damping channel

In Pauli-Liouville representation, the generalized damping channel is given by

$$\Lambda_{GD} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{(1-\gamma_1)(1-\gamma_2)} & 0 & 0 \\ 0 & 0 & \sqrt{(1-\gamma_1)(1-\gamma_2)} & 0 \\ \gamma_1(2\lambda-1) & 0 & 0 & 1-\gamma_1 \end{bmatrix} \tag{B.1}$$

where $\gamma_1 = 1 - e^{-\Gamma_1 \Delta t}$ and $\gamma_2 = 1 - e^{-2\Gamma_2 \Delta t}$. The diamond distance $D_{GD} = \frac{1}{2}||\mathbb{1} - \Lambda_{GD}||_\diamond$ is defined as

$$D_{GD} = \frac{1}{2}\text{sup}_\rho||\mathbb{1} \otimes (\mathbb{1} - \Lambda_{GD})(\rho)||_1. \tag{B.2}$$

B.1 Method 1 - Triangle inequality

We may split $\mathbb{1} - \Lambda_{GD}$ into its diagonal and non-diagonal part $\mathbb{1} - \Lambda_{GD} = \mathcal{E}_1 + \mathcal{E}_2$

$$\mathcal{E}_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1-\sqrt{(1-\gamma_1)(1-\gamma_2)} & 0 & 0 \\ 0 & 0 & 1-\sqrt{(1-\gamma_1)(1-\gamma_2)} & 0 \\ 0 & 0 & 0 & \gamma_1 \end{bmatrix} \quad \mathcal{E}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \gamma_1(2\lambda-1) & 0 & 0 & 0 \end{bmatrix} \tag{B.3}$$

and then apply the triangle inequality $||\mathbb{1} - \Lambda_{GD}||_\diamond \leq ||\mathcal{E}_1||_\diamond + ||\mathcal{E}_2||_\diamond$. Since $\mathcal{E}_1$ is a Pauli channel, the diamond norm is directly computable as $||\mathcal{E}_1||_\diamond = \frac{1}{2}\mathrm{Tr}(\mathcal{E}_1)$,

$$||\mathcal{E}_1||_\diamond = 1 - \sqrt{(1-\gamma_1)(1-\gamma_2)} + \frac{1}{2}\gamma_1. \tag{B.4}$$

Then $||\mathcal{E}_2||_\diamond = \sup_\rho ||\mathbb{1} \otimes \mathcal{E}_2(\rho)||_1$ where ρ can be assumed to be a pure state, $\rho = |\psi\rangle\langle\psi|$ and $|\psi\rangle = (a\ b\ c\ d)^T$. Working this out, we get:

$$\mathbb{1} \otimes \mathcal{E}_2(\rho) = \frac{\gamma_1(2\lambda-1)}{2}\begin{bmatrix} |a|^2+|b|^2 & 0 & ac^*+bd^* & 0 \\ 0 & -(|a|^2+|b|^2) & 0 & -(ac^*+bd^*) \\ ca^*+db^* & 0 & |c|^2+|d|^2 & 0 \\ 0 & -(ca^*+db^*) & 0 & -(|c|^2+|d|^2) \end{bmatrix} \tag{B.5}$$

which has eigenvalues $\lambda = \pm\frac{\gamma_1(2\lambda-1)}{4}[1 \pm \sqrt{1-4(ad-bc)(a^*d^*-b^*c^*)}]$ so that for any a, b, c, d, $\sum|\lambda_k| = \gamma_1(2\lambda-1)$ and therefore $||\mathcal{E}_2||_\diamond = \gamma_1(2\lambda-1)$. Putting things together, the upper bound is

$$||\mathbb{1} - \Lambda_{GD}||_\diamond \leq 1 - \sqrt{(1-\gamma_1)(1-\gamma_2)} - \frac{1}{2}\gamma_1 + 2\lambda\gamma_1. \tag{B.6}$$

B.2 Method 2 - Semidefinite programming

We use the semidefinite program defined in [67], which has the diamond distance as its optimal value. The diamond distance $D_{GD} = \frac{1}{2}||\mathbb{1} - \Lambda_{GD}||_\diamond$ will be bounded by the solution to the dual problem, $||\mathrm{Tr}_B(Z)||_\infty$ where Z is subject to the constraints $Z \geq J(\mathbb{1} - \Lambda_{GD})$ and $Z \geq 0$.

The Choi matrix of $\mathbb{1} - \Lambda_{GD}$ is given by

$$\begin{aligned} J(\mathbb{1} - \Lambda_{GD}) = {} & \gamma_1(\lambda-1)\,|00\rangle\langle 00| - \lambda\gamma_1\,|11\rangle\langle 11| + \gamma_1(1-\lambda)\,|01\rangle\langle 01| + \gamma_1\lambda\,|10\rangle\langle 10| \\ & + (\sqrt{(1-\gamma_1)(1-\gamma_2)} - 1)[|00\rangle\langle 11| + |11\rangle\langle 00|]. \end{aligned} \tag{B.7}$$

We pick a dual feasible point Z as

$$\begin{aligned} Z = {} & (1 - \sqrt{(1-\gamma_1)(1-\gamma_2)})\,|00\rangle\langle 00| + \gamma_1(1-\lambda)\,|01\rangle\langle 01| \\ & + \gamma_1\lambda\,|10\rangle\langle 10| + (1 - \lambda(1-\gamma_2))\,|11\rangle\langle 11|\,. \end{aligned} \tag{B.8}$$

Z is positive semidefinite since each coefficient in the above is non-negative, and also $Z \geq J(\mathbb{1} - \Lambda_{GD})$ is satisfied since

$$\begin{aligned} Z - J(\mathbb{1} - \Lambda_{GD}) = (1 - \sqrt{(1-\gamma_1)(1-\gamma_2)})[|00\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 00|] \\ + \gamma_1(1-\lambda)\,|00\rangle\langle 00| + (1 - \lambda(1-\gamma_1-\gamma_2))\,|11\rangle\langle 11| \end{aligned} \tag{B.9}$$

$$\begin{aligned} Z - J(\mathbb{1} - \Lambda_{GD}) = 2(1 - \sqrt{(1-\gamma_1)(1-\gamma_2)})\,|\psi_{Bell}\rangle\langle\psi_{Bell}| + \gamma_1(1-\lambda)\,|00\rangle\langle 00| \\ + (\sqrt{(1-\gamma_1)(1-\gamma_2)} - \lambda(1-\gamma_1-\gamma_2))\,|11\rangle\langle 11| \end{aligned} \tag{B.10}$$

where $|\psi_{Bell}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Note the $|11\rangle\langle 11|$ coefficient is always positive since $\sqrt{(1-\gamma_1)(1-\gamma_2)} \geq \lambda(1-\gamma_1-\gamma_2)$. Then we obtain an upper bound on the diamond distance as $\frac{1}{2}||\mathbb{1} - \Lambda_{GD}||_\diamond \leq ||\mathrm{Tr}_B(Z)||_\infty$

$$||\mathbb{1} - \Lambda_{GD}||_\diamond \leq \max\begin{cases} 2(1 - \sqrt{(1-\gamma_1)(1-\gamma_2)} + \gamma_1(1-\lambda))) \\ 2(1 - \lambda(1-\gamma_1-\gamma_2)) \end{cases}. \tag{B.11}$$

Appendix C

The diamond norm of single element matrices

Here we show that the diamond norm of a matrix $M = \begin{bmatrix} 0 & 0 \\ \alpha & \mathcal{E} \end{bmatrix}$ with a single non-zero element x in either the unital block $\mathcal{E}$ or non-unital vector α is equal to $|x|$. Due to the invariance of the diamond norm under unitary conjugation, we need to show this for only two cases: when x is some element of α, and when x is some element of $\mathcal{E}$.

C.1 x is an element of α

Let $\alpha = \begin{bmatrix} 0 \\ 0 \\ x \end{bmatrix}$, $\mathcal{E} = 0$, and $\rho = |\psi\rangle\langle\psi|$ where $|\psi\rangle = (a\ b\ c\ d)^T$. Then $||M||_\diamond = \sup_\rho ||\mathbb{1} \otimes M(\rho)||_1$.

$$\mathbb{1} \otimes M(\rho) = \frac{x}{2} \begin{bmatrix} |a|^2 + |b|^2 & 0 & ac^* + bd^* & 0 \\ 0 & -(|a|^2 + |b|^2) & 0 & -(ac^* + bd^*) \\ ca^* + db^* & 0 & |c|^2 + |d|^2 & 0 \\ 0 & -(ca^* + db^*) & 0 & -(|c|^2 + |d|^2) \end{bmatrix} \tag{C.1}$$

which has eigenvalues $\lambda = \pm\frac{x}{4}[1 \pm \sqrt{1 - 4(ad - bc)(a^*d^* - b^*c^*)}]$ so that for any a, b, c, d, $\sum |\lambda_k| = |x|$ and therefore $||M||_\diamond = |x|$.

C.2 x is an element of $\mathcal{E}$

Let $\mathcal{E} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & x \end{bmatrix}$ and $\alpha = 0$.

$$\mathbb{1} \otimes M(\rho) = \frac{x}{2} \begin{bmatrix} |a|^2 - |b|^2 & 0 & ac^* - bd^* & 0 \\ 0 & -|a|^2 + |b|^2 & 0 & -ac^* + bd^* \\ ca^* - db^* & 0 & |c|^2 - |d|^2 & 0 \\ 0 & -ca^* + db^* & 0 & -|c|^2 + |d|^2 \end{bmatrix} \tag{C.2}$$

which has eigenvalues $\lambda = \frac{x}{4}[\pm(|a|^2 - |b|^2 + |c|^2 - |d|^2) \pm \sqrt{1 - 4(ab^* + cd^*)(a^*b + c^*d)}]$ so that $\sum |\lambda_k| = |x||(|a|^2 - |b|^2 + |c|^2 - |d|^2)|$ with maximal value $|x|$, and therefore $||M||_\diamond = |x|$.