A study of the robustness of magic state distillation against Clifford gate faults

by

Tomas Jochym-O'Connor

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, 2012

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

Quantum error correction and fault-tolerance are at the heart of any scalable quantum computation architecture. Developing a set of tools that satisfy the requirements of fault-tolerant schemes is thus of prime importance for future quantum information processing implementations. The Clifford gate set has the desired fault-tolerant properties, preventing bad propagation of errors within encoded qubits, for many quantum error correcting codes, yet does not provide full universal quantum computation. Preparation of magic states can enable universal quantum computation in conjunction with Clifford operations, however preparing magic states experimentally will be imperfect due to implementation errors. Thankfully, there exists a scheme to distill pure magic states from prepared noisy magic states using only operations from the Clifford group and measurement in the Z-basis, such a scheme is called magic state distillation [1]. This work investigates the robustness of magic state distillation to faults in state preparation and the application of the Clifford gates in the protocol. We establish that the distillation scheme is robust to perturbations in the initial state preparation and characterize the set of states in the Bloch sphere that converge to the T-type magic state in different fidelity regimes. Additionally, we show that magic state distillation is robust to low levels of gate noise and that performing the distillation scheme using noisy Clifford gates is a more efficient than using encoded fault-tolerant gates due to the large overhead in fault-tolerant quantum computing architectures.

Acknowledgements

I would like to thank my supervisor, Dr. Raymond Laflamme, for being a great mentor, scientific advisor, and for fuelling my desire for learning through interesting questions and projects in physics and quantum information. I would also like to acknowledge all of my colleagues in Dr. Laflamme's group for stimulating interesting scientific discussions.

I would like to thank my collaborators, Yafei Yu, Bassam Helou, and Dr. Laflamme, for the hard work they provided in developing many of the results presented in the thesis. Their dedication to this project was inspirational and it was truly a joy to work with all of them. The work presented in Chapter 2 was begun by Bassam Helou, and the analytic derivation of the convergence of the noisy magic states were performed by Yafei Yu and myself, and all the results were discussed and reviewed with my supervisor Dr. Raymond Laflamme. I was the primary researcher on the results of Chapter 3 on magic state distillation with faulty Clifford gates, including deriving all the analytic results and providing the analysis on the resource counting and the comparative study between faulty and fault-tolerant magic state distillation. All results were discussed with my three collaborators on the project.

I would like to thank Dr. Joseph Emerson and Dr. Ben Reichardt for sitting on my advising committee. Their knowledge, insight, and guidance were very helpful in the direction and development of this project. I would additionally like to thank Dr. Kevin Resch for agreeing to sit on my examining committee.

My colleagues and I would like to acknowledge the support of Industry Canada, Quantum Works, and the National Science and Engineering Council of Canada for the funding of this project.

Finally, I would like to thank all of my family and friends for being ever so supportive as I pursue my scientific career, *je vous aime*.

Dedication

To my teachers, for providing knowledge and inspiration, and my parents, Anna and Roger, my aunt Betty, and my sister Gabrielle, the greatest teachers one could wish for.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction and Motivation

1.1 Introduction

Quantum information processing has the ability to efficiently address certain computational problems that have to date eluded classical algorithms. The promise of simulating quantum mechanical phenomena [2] or factoring large numbers [3] are among the most famous examples where a quantum computer could provide efficient scalable algorithms where no classical counterpart is known. So what is holding them back? Imperfect control of the quantum systems and unintended coupling to an external environment lead to errors in the implementation of the quantum computing device and are leading contributors to sources of noise. Techniques such as error correction have been developed to address classes of errors to recover quantum information that could otherwise be lost [4, 5, 6, 7, 8]. However, in order to be able to scale the devices of today into the processors of tomorrow, such faults will have to be corrected using methods that themselves do not introduce further errors in the device, motivating the construction of fault-tolerant quantum computing architectures [9, 10, 11, 12, 13].

The ability to fully exploit the power of quantum mechanics through quantum computation requires the use of universal control, that is to say, the ability to efficiently perform to arbitrarily high precision any unitary transformation in any Hilbert space that describes

the quantum state in question [14, 15]. Why this notion of universality? The notion of a quantum state evolution, or a quantum channel, is described by completely positive maps. Any such map can be described by a unitary transformation on a higher dimensional Hilbert space followed by tracing out the ancillary system, as established by the Stinespring dilation theorem [16]. As such, having the ability to perform arbitrary state evolution on a quantum computing device is equivalent to being able to perform an arbitrary unitary transformation of the extended Hilbert space describing the state of the computation. Developing techniques to establish such control is thus of great interest to all those pursuing the science of quantum information processing. The typical model for universal quantum computation involves preparation and measurement in the computational basis along with the ability to perform any single and two-qubit gates with arbitrarily high precision using a discrete set of gates available to the implementation in question, such a gate set is called a universal gate set. However, implementing a set of universal operations directly is not possible using transversal gates, a common fault-tolerant architecture. As a result, a different approaches to developing universal quantum computation have been taken such as gate teleportation through the use of magic states [1, 17, 18] and simulation of the Toffoli gate through preparation of "cat" states [9]. This work shall focus on the analysis of the former as it has more favourable resource requirements.

The Clifford gate set is an appealing gate set for many quantum computing architectures due to the ability to implement such gates transversally for many quantum error correcting codes [14]. However, the Clifford group is not universal for quantum computation and moreover computation based of Clifford group operations is efficiently classically simulatable [19]. Therefore, an additional resource is needed in the quest for universal quantum computation. An example of such a resource is a magic state. Magic states provide the ability to achieve universal quantum computation in conjunction with Clifford gate operations [1, 20, 21]. Moreover, multiple copies of noisy magic states can be distilled using a probabilistic method based of Clifford gate operations. It is this distillation process that is the central study of this work.

While the process of magic state distillation uses only a set of operations that can be implemented fault-tolerantly in order to reduce their error rate to arbitrarily low levels, performing such logical encodings is expensive in their number of required resources. Is

reducing the error rate of the Clifford gate operations to be arbitrarily small through a fault-tolerant encoding a necessary step to magic state distillation? This work shows that the magic state distillation protocol is robust to low levels of Clifford gate noise. Additionally, it is shown that, for certain noise levels, implementing the distillation protocol using faulty gates is more efficient than encoding the Clifford gates in a fault-tolerant architecture.

1.2 Mathematical Preliminaries

1.2.1 Descriptions of quantum mechanical systems

In this section, the definitions of fundamental quantum computation concepts are established and shall be used throughout the rest of this work.

A quantum mechanical state is described mathematically by using complex Hilbert spaces. The typical quantum state that we shall discuss in this work is a qubit, that is a state that lives in a 2-dimensional complex Hilbert space $\mathcal{H}$. A pure qubit state $|\psi\rangle$ can be described by a vector

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle. \tag{1.1}$$

The inner product of two vectors $|\psi_1\rangle = \alpha_1|0\rangle + \beta_1|1\rangle$ and $|\psi_2\rangle = \alpha_2|0\rangle + \beta_2|1\rangle$ in $\mathcal{H}$ is defined as,

$$\langle\psi_1|\psi_2\rangle = \alpha_1^*\alpha_2 + \beta_1^*\beta_2. \tag{1.2}$$

As such, the two vectors $|0\rangle$ and $|1\rangle$ are defined to be orthonormal and form the spanning basis set for the 2-dimensional complex Hilbert space.

The act of observing the state of a quantum mechanical system, or measurement of a quantum system, is given by the mathematical description of an Hermitian observable $\mathcal{O}$ whose measurement outcomes satisfy Born's rule [22]. The spectral decomposition theorem states that the observable $\mathcal{O}$ can be written as,

$$\mathcal{O} = \sum_i \lambda_i P_i, \tag{1.3}$$

where λ_i are real since $\mathcal{O}$ is Hermitian. P_i is a projection onto the subspace of states whose eigenvectors of the operator $\mathcal{O}$ have eigenvalue λ_i. Born's rule states that the probability of measuring a given eigenvalue λ_i for a pure quantum state $|\psi\rangle$ is given by $\langle\psi|P_i|\psi\rangle = \mathrm{Tr}\big(P_i|\psi\rangle\langle\psi|\big)$. Therefore, in order for the probabilities of all the different outcomes to sum to 1, the mathematical description of a pure quantum state $|\psi\rangle$ must be normalized, that is $\langle\psi|\psi\rangle = 1$. The post-measurement state is then given by the projection of the quantum state onto the subspace associated with the measured eigenvalue λ_i, $P_i|\psi\rangle$, and must be renormalized in order to satisfy the required normalization requirement for the description of a quantum mechanical state,

$$\frac{P_i|\psi\rangle}{\sqrt{\langle\psi|P_i|\psi\rangle}}. \tag{1.4}$$

Time evolution of a quantum state without measurement must therefore preserve the normalization condition of the quantum state. A time-evolution operator for a closed quantum system must be a unitary transformation U.

Definition 1 *A* unitary *operator U is an operator such that*

$$U^\dagger U = I, \tag{1.5}$$

where I is the identity operator on the input Hilbert space $\mathcal{H}$.

Thus far, we have described the postulates of quantum mechanics given pure quantum systems living in closed quantum systems, without noise introduced through the leakage of quantum information to the external universe, also known as open quantum system dynamics. In order to proceed to such a description of quantum state evolution, we must expand our definition of a quantum state.

Suppose one was given a quantum state with the following information, the state that has been given corresponds to the pure state $|\psi_i\rangle$ with probability p_i. Thus far we have discussed the description of the states in their mathematical form, but also in terms of their measurement observables. The mathematical description of the quantum system that is presented must therefore preserve the probabilities of measuring given outcomes for an

observable $\mathcal{O}\sum_i \lambda_i P_i$. The probability of measuring the eigenvalue λ_k must correspond to the weighted sum of each of the individual probabilistic outcomes,

$$\begin{aligned}\Pr(\lambda_k) = \sum_i p_i \mathrm{Tr}\big(P_k|\psi_i\rangle\langle\psi_i|\big) \quad &= \textstyle\sum_i \mathrm{Tr}\big(P_k p_i|\psi_i\rangle\langle\psi_i|\big) \\ &= \mathrm{Tr}\Big(P_k \textstyle\sum_i p_i|\psi_i\rangle\langle\psi_i|\Big), \end{aligned} \tag{1.6}$$

where the linearity of the trace function was used. Therefore, in order to satisfy the probability distribution of measurement results, the mixture of different pure quantum states can be described by the following mixed state *density matrix* ρ,

$$\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|. \tag{1.7}$$

Measuring the observable $\mathcal{O}$ will then record the measurement outcome λ_k according to an equation analogous to the Born's rule for pure quantum states,

$$\Pr(\lambda_k) = \mathrm{Tr}(P_k\rho). \tag{1.8}$$

In this new description of quantum states, state evolution must now map valid density matrices to valid density matrices. The density matrix is a positive operator, that is has nonnegative eigenvalues, so that the probability of obtaining measurement outcomes is always positive. The trace of the density matrix must also be equal to 1 (after renormalization in the case of measurement) in order to satisfy normalized probability distributions. A map Λ is said to be *positive* if $\Lambda(\rho)$ is positive, for all positive ρ. However, since quantum maps can also include ancillary systems, quantum state evolution is given by completely positive maps.

Definition 2 *A linear map $\mathcal{N}$ acting on a Hilbert space $\mathcal{H}_A$ is completely positive if and only if for any Hilbert space $\mathcal{H}_B$, $\mathcal{N}\otimes I_B$ is a positive map, where I_B is the identity operator on $\mathcal{H}_B$.*

Completely positive maps are not necessarily trace-preserving (TP), as measurement operators can leave the output state unnormalized and the mathematical description of the

state post-measurement will have to be normalized according to the measurement probability. While describing state evolution using completely positive maps seems complicated, it is analogous to state evolution by unitary transformation in larger Hilbert spaces, by the Stinespring dilation theorem.

Theorem 3 ([16]) *For any CPTP map $\mathcal{N}$ acting on a density matrix $\rho \in \mathcal{H}$, there exists an ancillary Hilbert space $\mathcal{H}_K$, a pure state $|\psi\rangle_K$, and a unitary transformation U such that,*

$$\mathcal{N}(\rho) = Tr_K\big(U(\rho \otimes |\psi\rangle\langle\psi|_K)U^\dagger\big). \tag{1.9}$$

Therefore simulating arbitrary quantum state evolution is equivalent to being able to perform arbitrary unitary rotations in a larger Hilbert space. As such, having access to a universal gate set that can simulate a unitary rotation of arbitrarily high precision and the ability to measure arbitrary subsystems of the larger Hilbert space provides the means to simulate the evolution of any quantum system. This is the underlying motivation behind developing a quantum device that has access to a universal gate set.

1.2.2 Pauli operators

The single-qubit Pauli operators are given by the matrices,

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{1.10}$$

$$Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad Z = -iXY = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.11}$$

The single-qubit Pauli matrices can be generated by X and Z, and generalizing the definition to n qubits, the n-qubit Pauli group is the set of operators generated by the tensor product of all possible combinations of the X and Z operators on each of the individual n qubits. Up to a global phase, the size of the Pauli group on n-qubits is 4^n. The n-fold tensor product of Pauli matrices form a basis for all Hermitian matrices of size 2^n by 2^n

and all density matrices can then be expressed by their coefficients in such a basis [23]. For example, an arbitrary single-qubit quantum state can be expressed as a density matrix of the following form,

$$\rho = \frac{1}{2}\Big(I + \alpha X + \beta Y + \gamma Z\Big). \tag{1.12}$$

The real coefficients α, β, and γ, correspond to the expected value of the observables X, Y, and Z, respectively. In order for ρ to be a valid normalized density matrix, the coefficients must lie within a sphere with radius 1, $\alpha^2 + \beta^2 + \gamma^2 \leq 1$. Therefore, a qubit can be thought of as a quantum system whose characterizing observables correspond to the Pauli matrices.

1.2.3 Quantum Gates

As described in Section 1.2.1, having the ability to perform arbitrary unitary rotations in a given Hilbert space is the ultimate goal for a quantum computing implementation. A group of quantum gates that can perform any unitary transformation with arbitrarily high precision efficiently, that is with polylogarithmic overhead, will be called a *universal gate set*. Before providing some examples of universal gate sets, we shall define another interesting group of gates for quantum computation.

Definition 4 *Let $\mathcal{P}_n$ be the set of Pauli operators on n-qubits. The Clifford group $\mathcal{C}_2$ on n-qubits is the set of unitary operators U such that,*

$$UPU^\dagger \in \mathcal{P}_n, \ \forall \ P \in \mathcal{P}_n. \tag{1.13}$$

The Clifford group maps Pauli matrices to Pauli matrices, and as such may not seem all that useful at first glance as this implies that evolution of a quantum state via Clifford group operations is classically efficiently simulatable by the Gottesman-Knill Theorem [10, 14, 19]. However, the Clifford gates augmented by a single other non-Clifford gate has shown to be universal for quantum computation [24, 25]. Yet Clifford gates exhibit many interesting properties for large scale quantum computations, and as such are of particular interest to this work.

A set of generators for the Clifford gates are the following,

$$S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, \quad H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad CNOT = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{1.14}$$

Additionally, examples of gates that are not members of the Clifford group and provide universality are,

$$\Lambda_{\pi/8} = \begin{pmatrix} e^{i\pi/8} & 0 \\ 0 & e^{-\pi/8} \end{pmatrix}, \quad \Lambda_{\pi/12} = \begin{pmatrix} e^{i\pi/12} & 0 \\ 0 & e^{-\pi/12} \end{pmatrix}. \tag{1.15}$$

1.3 Quantum Error correction

1.3.1 Motivation

Implementations of a computing device using physical systems are subject to errors in the ability to control the systems of interest coherently. While advances in the physical and computer sciences have increased the ability to control physical devices to behave and exhibit the properties that scientists seek when developing their computing algorithms, unintended faults do persist in even the most developed and studied physical systems, be it classical or quantum. In classical computation theory, error correcting algorithms are often based on the ability to copy classical information and use repetition codes [14, 26].

Consider the simplest example of a repetition code where the classical bits in question are copied to form a triplet,

$$\begin{aligned} 0 &\to 000 \\ 1 &\to 111. \end{aligned}$$

A classical code whose output is, say, the output to a function $f(c)$, where c is the classical input, would then be repeated three times,

$$f(c) \to f(c)f(c)f(c).$$

Say the classical computation had a probability of failure of p, then by comparing the output of the three computations and taking taking the majority result as the outcome of the computation, the probability of failure becomes $3p^2(1-p)+p^3$, where the errors are assumed to be independent. The probability is then reduced if $p < 1/2$, and the process can be iterated to reduce the error rate to be arbitrarily small.

In quantum information processing the procedure to correct error affecting the operations or states of the quantum computation is not so simple. Firstly, the no-cloning theorem prevents the copying an arbitrary quantum state, complicating the ability to implement a repetition code as in the classical case. Secondly, while classical computations are affected by a discrete set of errors, namely bit flip errors, quantum states can be subject to a range of errors that is infinite, from small rotations in the Bloch sphere to phase or bit flips. As such, the ability to detect and correct against such a range of errors seems to be at first glance a monumental task. However, the ability to perform non-classical operations through the use of entanglement will provide the necessary resource to achieve the desired goal.

1.3.2 Quantum Error Correcting Codes

In Section 1.3.1, an example of a classical repetition code was given, however what encompasses a quantum error correcting code (QECC)? This section begins with the definition of a QECC, and is followed by a set of motivating examples for why such a definition makes sense and has the desired properties for QEC. A quantum error is defined as a completely positive map that acts on a subspace of the overall Hilbert space that describes the quantum state of the computation. For example, for a multiple qubit system, if an error occurs on a single qubit, the error map will be a completely positive map on the 2-dimensional subspace corresponding to that qubit and will leave the remaining qubits unchanged (identity operation).

Definition 5 ([7, 8]) *For a given error $\mathcal{E} \in \mathfrak{B}(\mathcal{H})$, a quantum error correcting subspace code $\mathcal{C}$, is a subspace $\mathcal{C} \subset \mathcal{H}$ where there exists a recovery quantum operation $\mathcal{R} \in \mathfrak{B}(\mathcal{H})$*

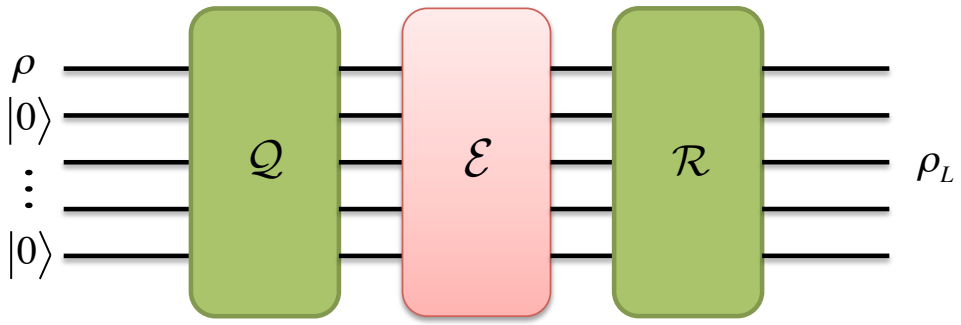


Figure 1.1: General architecture of a quantum error correcting scheme. A physical qubit ρ containing the quantum information to be stored and manipulated is coupled to a set of ancillary qubits to form a logical qubit $\rho_L \in \mathcal{C}$ using an encoding operation $\mathcal{Q}$. The logical qubit is then subject to some physical noise map $\mathcal{E}$. The logical quantum state can then be recovered by a chosen recovery operator $\mathcal{R}$ if the error map is quantum error correctable as in Definition 5. The physical state that is encoded is not restricted to being a single qubit, however for simplicity that is the case shown here.

such that

$$(\mathcal{R} \circ \mathcal{E})(\rho) = \rho \quad \forall\ \rho = P_{\mathcal{C}} \rho P_{\mathcal{C}}, \tag{1.16}$$

where $P_{\mathcal{C}}$ *is the projector of* $\mathcal{H}$ *onto* $\mathcal{C}$.

The idea behind the definition of a QEC subspace code is that for a given error map $\mathcal{E}$, there exists a code space $\mathcal{C}$ such that for all $\rho \in \mathcal{C}$, there is a recovery operator $\mathcal{R}$ that will bring the output state to its original state. What the recovery operator does to quantum states that are outside the codespace is irrelevant as quantum states used in the information processing are all chosen to reside in the codespace.

Quantum error correcting codes are typically developed using the structure pictured in Figure 1.1 where the physical unencoded qubit is coupled to a set of ancilla qubits through an encoding operation $\mathcal{Q}$, where the qubit lives in a logical qubit subspace $\mathcal{C}$. The logical qubit is then subject to some noise operation $\mathcal{E}$ that the quantum error correcting correcting code has been constructed to protect against, and as such, by performing a recovery

operation $\mathcal{R}$ the faulty state can be projected back onto the codespace by identifying the error and correcting it.

The simplest example that corrects for a Pauli error is the three-qubit bit flip code [5]. The encoding of the physical qubit is done through two $CNOT$ gates that map the input state basis $\{|0\rangle, |1\rangle\}$ to the codespace $\{|000\rangle, |111\rangle\}$, as shown in Figure 1.2. The noise that this quantum code is design to protect against is single qubit bit flips, which can flip a single physical state from $|0\rangle$ to $|1\rangle$. As such, if a single error occurs, the recovery operation maps the noisy state onto the codespace by in essence taking the majority rules argument through a set of $CNOT$ gates.

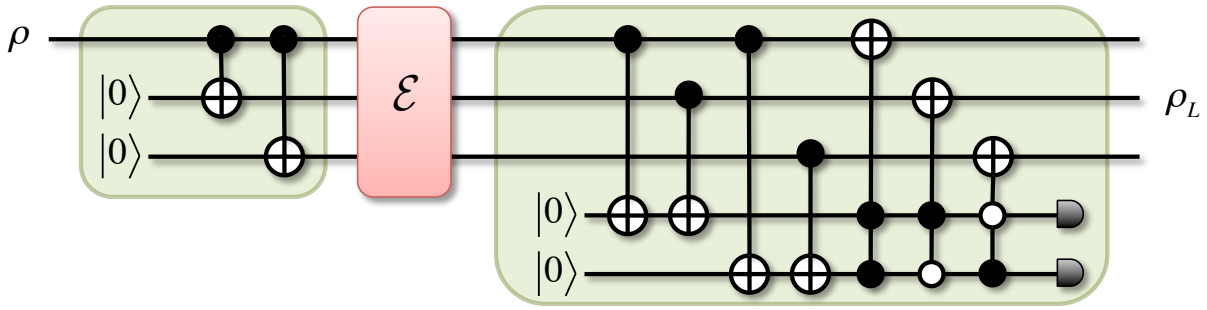


Figure 1.2: The 3-qubit code correcting against bit-flip errors. The physical input state ρ is encoded in a logical Hilbert space that is a subspace of a three-qubit Hilbert space. The encoding operation is performed using two ancilla qubits and coupling them to the input state through a pair of $CNOT$ gates. The logical qubit is then subject to some physical noise given by the red region in the quantum circuit. The subsequent recovery operation, given by the second green box, corrects against any single qubit bit-flip, and the error that occurred can be read out in the measurement results. The output state of the circuit is the corrected logical state, encoded in the three-qubit code.

Another way to characterize the above quantum code is by considering the operators that leave the codespace invariant. For the above 3-qubit code, the codespace is left invariant by the quantum operators ZZI and IZZ. The group of Pauli operators generated by the Pauli operators that leave the codespace invariant (in this case a group of just 4 elements), is called the *stabilizer operations* of the quantum error correcting code. As

such, it is equivalent to describe a quantum error correctable subspace by either the basis vectors that span the space, or the stabilizer generators in the larger Hilbert space. For a n qubit physical system, in order to encode k logical qubits there shall be $n-k$ independent stabilizer generators of the code.

In the stabilizer formalism [10], a logical quantum operation, that is an quantum operation that performs a transformation within the codespace, must leave the stabilizer group invariant. Let $\mathcal{C}$ be a quantum code, and $S = \langle S_k \rangle_{k=1}^{l}$ be the stabilizer group of the code, where S is the group generated by the set of generators S_k. Then for any logical state $|\psi_L\rangle$ in the codespace $\mathcal{C}$, the following holds by definition,

$$S_k|\psi_L\rangle = |\psi_L\rangle, \qquad \forall\ S_k \in S. \tag{1.17}$$

Consider now a unitary quantum gate that performs a transformation on a vector in the codespace $|\psi_L\rangle$,

$$U_L|\psi_L\rangle = U_L S_k|\psi_L\rangle = \big(U_L S_k U_L^\dagger\big)U_L|\psi_L\rangle, \tag{1.18}$$

thus the operator $U_L S_k U_L^\dagger$ stabilizes the transformed state, moreover since the identity must hold for all of the original stabilizer generators S_k, the group generated by $U_L S U_L^\dagger = \langle U_L S_k U_L^\dagger \rangle_{k=1}^{l}$ is the stabilizer group for the transformed codespace. In the case of a logical operation, the unitary must map vectors from the codespace the same codespace, as such the stabilizer group must remain invariant, $U_L S U_L^\dagger = S$. In the case of the three-qubit code, the operator XXX clearly maps the basis vectors $|000\rangle$ and $|111\rangle$ to one another, and as such is a logical operation. This is confirmed by the fact that the operator XXX leaves the stabilizer group invariant.

While the 3 qubit code is a simple example of a quantum code that corrects against bit-flips, developing a quantum error correcting code that corrects against arbitrary Pauli errors is of more practical use in typical implementations. It is worth noting that a single Z error on any of the qubits of the three qubit code will cause a logical operation. By performing a further encoding of the logical states that are protected against bit flip errors into a code that corrects against phase-flip errors, the resulting 9 qubit code will correct against an arbitrary error on single qubit.

While Shor's 9 qubit code is straightforward conceptually as the concatenation of a bit-flip and phase-flip code, it is not the smallest code to correct against arbitrary errors. The smallest such code is the 5 qubit quantum code [27, 7], whose stabilizers generators are

$$\begin{aligned} &XZZXI \\ &IXZZX \\ &XIXZZ \\ &ZXIXZ, \end{aligned} \tag{1.19}$$

and whose logical operators are given by

$$\begin{aligned} X_L &= XXXXX \\ Z_L &= ZZZZZ. \end{aligned} \tag{1.20}$$

While 5 qubits is the smallest number of qubits that can encode a single qubit and correct against arbitrary errors at a single location, the above code is has drawbacks for scalability. The 7 qubit Steane code [28] has more appealing fault-tolerant features, as shall be discussed in Section 1.4. The stabilizer generators for the 7 qubit code are

$$\begin{aligned} &XXXXIII \\ &XXIIXXI \\ &XIXIXIX \\ &ZZZZIII \\ &ZZIIZZI \\ &ZIZIZIZ, \end{aligned} \tag{1.21}$$

along with logical operators

$$\begin{aligned} X_L &= XXXXXXX \\ Z_L &= ZZZZZZZ. \end{aligned} \tag{1.22}$$

(a) Single qubit gates with no error propagation

(b) Two-qubit gate propagating errors

Figure 1.3: Difference in error propagation between one-qubit gates and two-qubit gates within an error correction codeblock.

1.4 Fault-tolerant quantum computation

1.4.1 Scaling quantum error correction

Motivated in Section 1.3.1, quantum error correcting codes shall play an essential role in any large scale quantum information processor as unavoidable errors will have to be addressed. While all the quantum information is stored in a higher dimensional logical Hilbert space, logical quantum operations will have to be applied to manipulate the logical quantum states in order to perform the desired algorithm.

Figure 1.3(a) provides an example of a set of quantum gates that cannot propagate a single error into multiple errors. However, as Figure 1.3(b) shows, a two-qubit gate such as a $CNOT$ gate can take a error that occurs on a single qubit, such as an X Pauli error on the control qubit, and turn it into a two-qubit physical error through the action of the gate. As a result, the application of logical gates using two-qubit gates within the codeblock is in general not robust to single qubit errors, which can lead to faults in the error correcting code, for this particular example if the quantum error correcting code would not correct against two arbitrary errors such propagation would lead to a logical fault. Therefore, in order to apply large scale quantum computations, quantum operations that propagate errors will have to be avoided, such operations will be called *fault-tolerant*.

In this work, a *fault-tolerant* quantum operation shall be defined as a quantum operation where if an error occurred at t physical locations in the implementation of the operation it shall produce at most t physical errors on the output codeblock.

1.4.2 A solution: Transversal Gates

The application of logical quantum gates that propagate correctable numbers of errors to errors that can no longer be tolerated by the quantum error correcting code are undesirable. There is, however, a simple solution to prevent such error propagation in the codespace: to not use two-qubit gates between qubits that encode a given logical quantum state. Such gates will be called *transversal*, and are formally defined below.

Definition 6 *A logical quantum gate $\mathcal{G}$ for a quantum code $\mathcal{C}$ is called a transversal gate if the individual quantum gates that construct the logical operation on the code $\mathcal{C}$ are located at at most* one *of physical qubits that form the quantum code $\mathcal{C}$.*

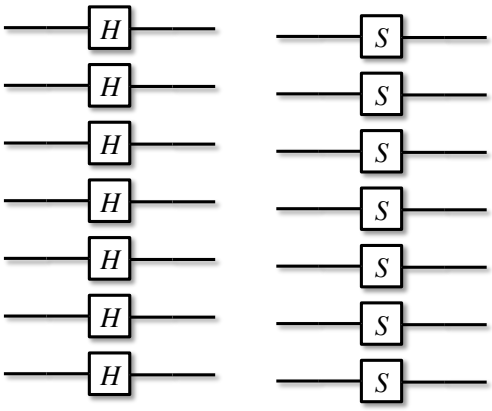
Figure 1.3(a) is an example of a logical gate that is transversal. Such gates are fault-tolerant since any set of errors that arise in the application of the physical gates that make up the logical gates will not propagate to other qubits in the quantum error correcting code as the propagation of errors can only occur through a two-qubit gate. The errors remain localized to the physical qubits where they occur, and since the quantum error correcting code has been developed with such errors in mind, there shall exist a recovery operation that will correct for these physical errors and the logical quantum gate can be applied while keeping the desired properties of the error correcting code.

The generators of the Clifford group are given by the set of gates,

$$S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, \quad H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad CNOT = \begin{pmatrix} I & 0 \\ 0 & X \end{pmatrix}, \tag{1.23}$$

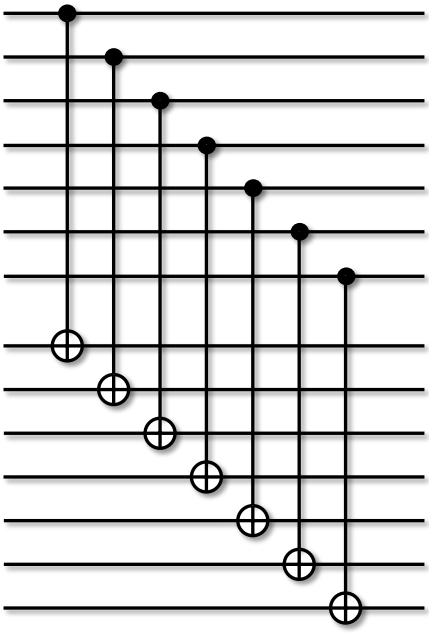
where I and X are the identity and Pauli-X operators on a single qubit. Figure 1.4 provides a set of transversal gates for the Clifford group generators in the 7-qubit Steane code. The two-qubit logical CNOT gate is indeed transversal since the two-qubit physical gates are between individual qubits in each logical code, not between multiple qubits in the same code.

The Gottesman-Knill theorem states that quantum operations based off elements of the Clifford group are efficiently simulatable on a classical computer [10], yet the addition



(a) Logical H

(b) Logical $S^{\dagger}$



(c) Logical $CNOT$

Figure 1.4: Transversal implementations of the Clifford group generators in the logical space of the 7-qubit Steane code. (a) The transversal Hadamard is implemented by performing Hadamard gates on each of the individual physical qubits forming the code. (b) The inverse of the phase gate S is applied by performing phase gates on each of the individual physical qubits. (c) Implementation of a $CNOT$ gate between two blocks of encoded qubits in the 7-qubit code.

of a single non-Clifford one-qubit gate is sufficient to provide universal quantum computation [24, 25]. As such, if there exists transversal gates for the generators of the Clifford group for quantum codes of as little as 7-qubits, perhaps there exists a set universal gates that are transversal for a more complex quantum code. However, the following theorem affirms that universal fault-tolerant quantum computation is not quite so simple.

Theorem 7 ([29, 30, 31]) *Let $\mathcal{C}$ be a quantum error correcting code, and $\mathcal{G}$ a set of logical unitary gates that are transversal for $\mathcal{C}$, then $\mathcal{G}$ cannot form a universal logical gate set for the code $\mathcal{C}$.*

While no universal set of transversal gates exist, there are other methods to develop fault-tolerant architectures. One such method is through the preparation of a certain class of quantum states that enable universal computation through gate teleportation [18], among such states are the magic states who are the main focus of this work and shall be discussed in great detail in Section 1.5.

1.4.3 Fault-tolerance threshold theorem

Before plunging into the realms of universal quantum computation using magic state distillation, it is worth further motivating the importance of fault-tolerant quantum computation architectures. Consider a physical device where the operations used to perform quantum information processing can be characterized by an independent error rate on each operation that is bounded from above by p. Given a quantum code $\mathcal{C}$ that can correct up to $(s-1)$ errors, and a quantum operation whose logical encoding requires A physical operations, if the logical gates are fault-tolerant, then in order for there to be a logical error there must be at least s independent faults. The probability of failure of the logical quantum operation at the first level of encoding p_1 can then be bounded as follows:

$$\begin{aligned} p_1 &\leq \textstyle\sum_{t=s}^{A} \binom{A}{t} p^t (1-p)^{A-t} \\ &\leq \textstyle\sum_{t=s}^{A} \binom{A}{t} p^t \\ &\leq \textstyle\sum_{t=s}^{A} \binom{A}{t} p^s \\ &\leq 2^A p^s = B p^s, \end{aligned} \tag{1.24}$$

where $B = 2^A$ is a constant.

If one then performs a second level of encoding, encoding the logical quantum states into an subsequent quantum code (for simplicity, assume it is the same quantum code), the subsequent failure of probability will be bounded from above by $p_2 \leq B(Bp^s)^s = B^{s+1}p^{s^2}$. Continuing the encoding procedure, the probability of failure at the k^{th} level can be bounded by,

$$p_k \leq B^{s^{k-1}+s^{k-2}+\ldots+1}p^{s^k} \leq (Bp)^{s^k}. \tag{1.25}$$

As such, as long as the initial error rate p is below some constant error *threshold* $p_{\text{th}} \leq B = 2^A$, the resulting encoded error rate can be made arbitrarily small.

Suppose each level of encoding requires r gates from the previous encoding level. Then the total number of gates required to encode to the k^{th} level is r^k. In order to simulate a logical gate that requires A encoded gates with error rate ϵ, the average error rate per gate must be ϵ/A. Thus the number of levels of encoding must satisfy,

$$(Bp)^{s^k} = \frac{\epsilon}{A} \rightarrow s^k = \frac{\log(A/\epsilon)}{\log(1/Bp)} \tag{1.26}$$

The total number of gates required to simulate A encoded gates at the desired error rate is given by Ar^k,

$$Ar^k = A(s^k)^{\log r/\log s} = A\Big(\frac{\log(A/\epsilon)}{\log(1/Bp)}\Big)^{\log r/\log s} = \mathcal{O}\big(A \cdot \text{poly}(\log(A/\epsilon))\big). \tag{1.27}$$

Therefore, A encoded gates can be simulated a polylogaritmic number of physical gates when the initial error rate of the physical gates p is below the fault-tolerance threshold p_{th}.

1.5 Magic State Distillation

As discussed in Section 1.4.2, there exists no set of universal quantum gates that are all transversal for a given quantum code. As such, another approach has to be used to augment the set of transversal Clifford gates to a universal gate set that is fault-tolerant. One such method is by preparing a quantum state that along with the Clifford group operations can

implement the necessary non-Clifford gate that provides universal control. Consider the following quantum state,

$$|T_0\rangle\langle T_0| = \frac{1}{2}\Big(I + \frac{1}{\sqrt{3}}(X + Y + Z)\Big), \tag{1.28}$$

which shall be referred to as the T *magic state*. Such a state can be used to implement the non-Clifford $\pi/12$-gate. The scheme to implement such a gate was outlined in the original work on magic state distillation (MSD) [1] and is given by the circuit below.

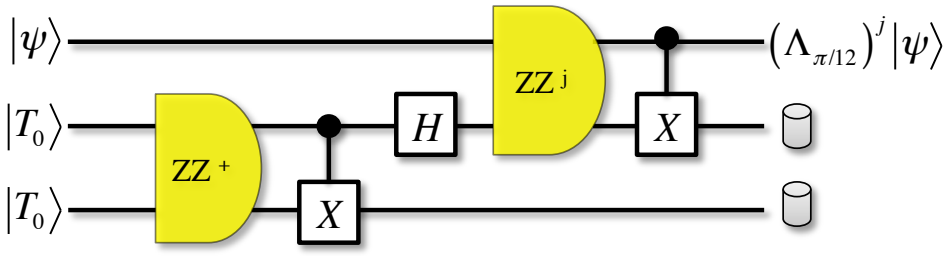


Figure 1.5: Implementation of the $\pi/12$-gate to an arbitrary state $|\psi\rangle$ using perfect magic states $|T\rangle$. The implementation of the gate is based off Clifford gates and $Z \otimes Z$ measurements, along with post-selection of the "+1" eigenspace of the first measurement. Depending on the outcome of the second measurement, either the $\pi/12$ or $-\pi/12$-gate is implemented, however the $\pi/12$ gate can be applied efficiently by repeating the process as this is equivalent to performing a classical random walk in the application of $\pi/12$ or $-\pi/12$ phases, which converges exponentially quickly to the phase desired.

Figure 1.5 applies the $\pi/12$ or $-\pi/12$-gate with equal probability, and the gate applied is given by the outcome of the second stabilizer measurement. Say one required specifically the $\pi/12$-gate to implement a given universal gate, then one can repeat the above quantum circuit with further magic states where the probability that one will need more that N repetitions of the algorithm decreases exponentially with N [1]. T type magic states are not the only states that can be used to implement a universal quantum gate, for example other magic states such as $|H\rangle = \cos\pi/8|0\rangle + \sin\pi/8|1\rangle$ can be used to implement the $\pi/8$-gate.

1.5.1 A scheme for magic state distillation

The magic state $|T_0\rangle$ can be used to implement the $\pi/12$-gate, which along with the Clifford gates, form a universal gate set. However, the preparation of such a state in any experimental implementation will have some faults associated with the faults in the physical device. As such, the implementation of a universal gate based on the perfect preparation of a particular state does not seem like a realistic scheme for a quantum information processor that will undergo some set of noisy operations. However, luckily perfect state preparation is not needed as there exists an quantum algorithm based on Clifford group operations that can take multiple noisy magic state and output a single magic state of higher fidelity. These properties, the ability to be distilled using Clifford group operations, and in turn to use the distilled state to perform a universal quantum gate, are the reason such states are called *magic*. The magic state distillation protocol [1] assumes the access to the following perfect quantum operations:

- Preparation of the quantum state $|0\rangle$
- Access to Clifford group operations
- Measurement in the Z-basis on all qubits.

The assumption of the ability to perform all such operations perfectly is based off the arguments presented in Section 1.4, that is all such operations can be performed fault-tolerantly for some quantum code. Such resources are not sufficient for universal quantum computation and a final universal resource is needed:

- Preparation of noisy magic states.

What are noisy magic states? This section shall explore the required initial fidelity that the noisy magic state preparation will have to have in order for the distillation protocol to be able to distill a magic state of arbitrarily high fidelity.

The magic state distillation protocol, first presented by Bravyi and Kitaev [1], begins with copies of quantum states of the form,

$$\rho_0 = (1-\epsilon)|T_0\rangle\langle T_0| + \epsilon|T_1\rangle\langle T_1|, \tag{1.29}$$

where $|T_1\rangle$ is the state orthogonal to $|T_0\rangle$. Such a state can be prepared from an arbitrary state ρ by using a dephasing transformation based on the Clifford gate $T = SH$,

$$\mathcal{T}(\rho) = \frac{1}{3}\Big(\rho + T\rho T^\dagger + T^\dagger \rho T\Big) = (1-\epsilon)|T_0\rangle\langle T_0| + \epsilon|T_1\rangle\langle T_1|, \tag{1.30}$$

where $\epsilon = \langle T_1|\rho|T_1\rangle$. The quantum operation $\mathcal{T}$ acts as a twirling operation in the Bloch sphere. The set of output states of the dephasing operation form an axis in the Bloch sphere for the range of possible values for $\epsilon \in [0, 1]$. Such a dephasing transformation can be implemented using classical equally distributed randomness between all three possible gate evolutions, T, $T^\dagger$, and the identity I. The importance of the dephasing protocol shall be addressed in great detail in Chapter 2, however for the description of protocol we shall include it here.

The quantum algorithm involves taking five copies of the noisy input magic states ρ_0, of the form found in Equation 1.29, and projecting this five-qubit state onto the codespace for the five-qubit error correcting code, whose stabilizers are $ZXIXZ$, $ZZXIX$, $XZZXI$, and $IXZZX$ [27, 7].

Projecting the five copies of the noisy magic state onto the codespace for the five-qubit code can be done by applying the decoding algorithm for the five qubit code, followed by 4 Z-basis measurements. The quantum circuit pictured in Figure 1.6 shows the decomposition of the decoding algorithm, in the orange box, in terms of the Clifford gate elements, that is in total 8 two-qubit gates and 5 one-qubit gates (by assuming that applying successive one-qubit Clifford gates is equivalent in experimental difficulty to applying a single different one-qubit Clifford that is made up of the composition of such gates). Various encoding/decoding circuits can be developed to encode the five qubit code, we chose to analyze the circuit presented in Figure 1.6 as once a qubit is used as a control qubit in a two-qubit gate, it is no longer used in any further two-qubit gates, thus minimizing the propagation of errors through the circuit [32].

Having projected onto the codespace by post-selecting on the measurement outcomes of the "+1" eigenstate in the Z-basis, the output state will have the form

$$\rho_{\text{out}} = \frac{1}{6}\Big((1-\epsilon)^5 + 5\epsilon^3(1-\epsilon)^2\Big)|T_0\rangle\langle T_0| + \frac{1}{6}\Big(\epsilon^5 + 5\epsilon^2(1-\epsilon)^3\Big)|T_1\rangle\langle T_1|, \tag{1.31}$$

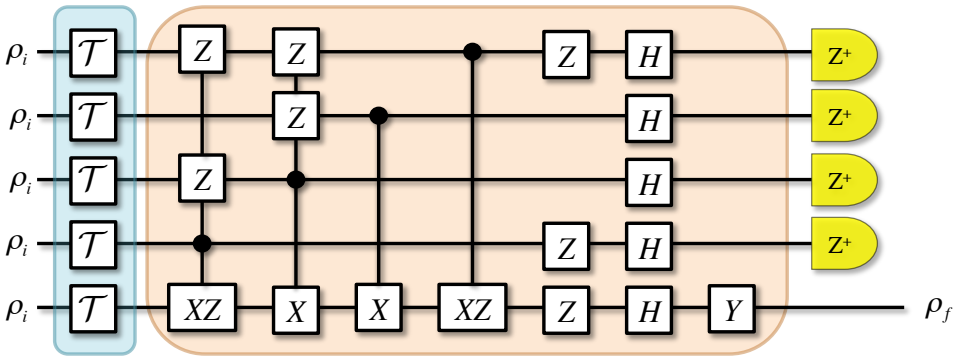


Figure 1.6: Analyzed circuit for the five-qubit magic state distillation protocol. The gate $\mathcal{T}$ represents the dephasing transformation. The gates X and Z are the standard Pauli gates, and the gate H is the Hadamard gate. The yellow measurement boxes represents the projection onto the trivial subspace upon post-selecting the measurement outcome "+1" for all the Z-basis measurements. All gates contained in the orange box represent the decoding circuit for the five-qubit quantum error correcting code.

where ϵ is the the error on the initial input state ρ_0. The normalization factor of the state corresponds to the probability of measuring the "+1" syndrome on all four measurement outcomes, and as such the success probability of the distillation iteration corresponds to

$$\mathrm{Tr}(\rho_{\mathrm{out}}) = \frac{1}{6}\Big((1-\epsilon)^5 + 5\epsilon^3(1-\epsilon)^2 + \epsilon^5 + 5\epsilon^2(1-\epsilon)^3\Big). \tag{1.32}$$

As such, the final normalized output state of the distillation protocol will have a modified error term,

$$\rho_1 = (1-\epsilon')|T_0\rangle\langle T_0| + \epsilon'|T_1\rangle\langle T_1|, \tag{1.33}$$

where the modified error term ϵ' has is related to the error on the input state ϵ by the following:

$$\epsilon' = \frac{\epsilon^5 + 5\epsilon^2(1-\epsilon)^3}{\epsilon^5 + 5\epsilon^3(1-\epsilon)^2 + 5\epsilon^2(1-\epsilon)^3 + (1-\epsilon)^5}. \tag{1.34}$$

By plotting the output fidelity of the magic state as a function of the input fidelity, as shown in Figure 1.7, one can notice a threshold value ϵ_{th} for the input state error such that

for all $\epsilon < \epsilon_{\text{th}}$, the error rate of the output state will be lower, $\epsilon' < \epsilon$. The threshold value for the input state error rate such that the magic state distillation protocol is beneficial is

$$\epsilon_{\text{th}} = \frac{1}{2}\left(1 - \sqrt{\frac{3}{7}}\right). \tag{1.35}$$

Throughout this work we shall typically refer to fidelity thresholds rather than error thresholds, that is the fidelity which the noisy magic states must be above in order to enable magic state distillation. We define the fidelity $F = 1 - \epsilon$, as such the fidelity threshold for perfect magic state distillation is

$$F_{\text{th}} = \frac{1}{2}\left(1 + \sqrt{\frac{3}{7}}\right) \approx 0.8273. \tag{1.36}$$

Therefore, when the initial state of the distillation protocol has high enough fidelity with respect to the magic state, 5 copies of such a state can be used to produce a single state of higher fidelity with respect to the magic state. This process can then be iterated by using 5 copies of distilled states and producing another state of higher fidelity, the number of initial states being consumed overall is then $5^2 = 25$ (not counting for the states that are thrown out when the measurement outcome does not produce a state in the codespace of the five-qubit code). The distillation can be repeated arbitrarily many times in this manner to produce a state with arbitrarily high fidelity. This idea if iterating the process to produce a magic state with arbitrarily high fidelity is described by the staircase in Figure 1.7.

1.5.2 Rate of convergence

A question that can be asked of magic state distillation is that seeing as each iteration of the protocol requires 5 copies of states produced from the previous iteration, the scaling of the number of magic states appears to be exponential. While this concern is valid, and the scaling of the number of noisy input magic states required scales poorly with the number of iterations that need to be run, the distillation protocol thankfully converges even faster than the divergence in the number of states being used. As such, in practice one would

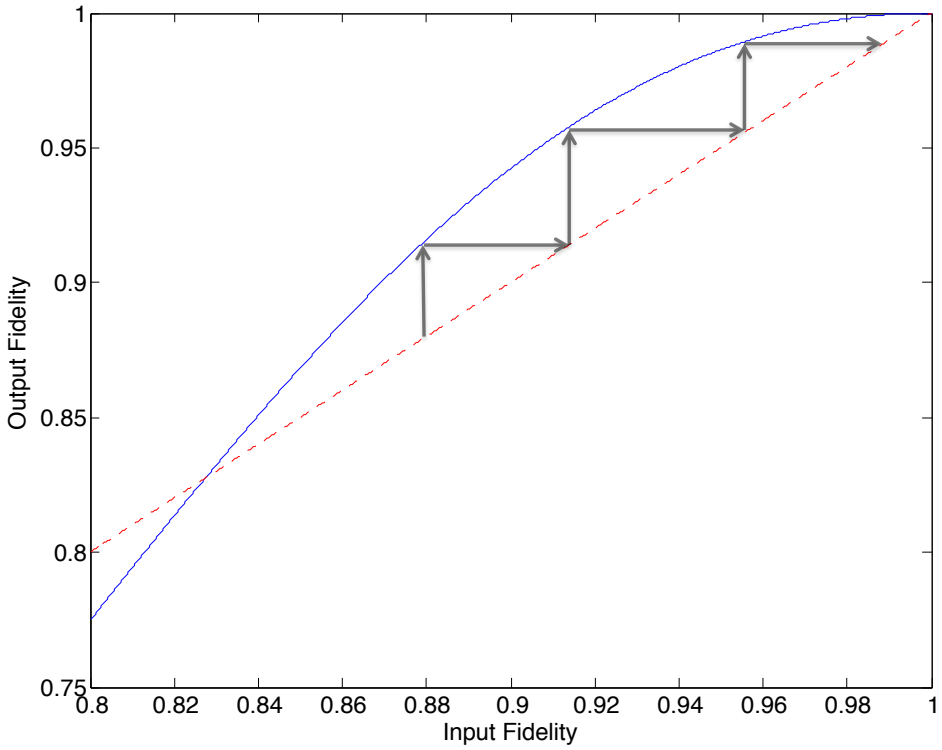


Figure 1.7: Plot of the output state fidelity of magic state distillation $\langle T_0|\rho_1|T_0\rangle$, as a function of the input state fidelity $\langle T_0|\rho_0|T_0\rangle$ in blue. The red dashed line is the plot corresponding to the input and output fidelities being equal. Magic state distillation is only useful in regimes where the blue curve is above the red dashed line. In such a region, multiple rounds of distillation can increase the output state fidelity of the magic state, one iteration at a time, as represented by the black arrow staircase.

rarely need more than a few iterations of the protocol to achieve a target reduced error rate for the magic state [1], typically below that set by the fault-tolerance threshold.

In the limit where the error rate in the input state is small, which occurs quite rapidly, the probability of success of the protocol is approximately 1/6, as given by Equation 1.32 where ϵ terms are assumed to be very small. As such, to distill a single copy of a higher fidelity magic state, it will require on average 30 copies of the magic states produced at the outcome of the previous scheme. Let k be the number of iterations of the protocol, then the total number of initial noisy magic states n will be

$$n = (30)^k \Rightarrow k = \log_{30} n = \frac{\log_2 n}{\log_2 30}. \tag{1.37}$$

Now calculating the error rate as a function of the number of iterations. In the limit of small error rate, the error on the output state to lowest order is $5\epsilon^2$. Each iteration reduces the error rate by a quadratic factor, and as such decreases exponentially,

$$\epsilon_{\text{out}}(k) = \frac{(5\epsilon)^{2^k}}{5} \tag{1.38}$$

$$\Rightarrow \epsilon_{\text{out}}(n) = \frac{(5\epsilon)^{n^\gamma}}{5}, \quad \gamma = \frac{1}{\log_2 30} \tag{1.39}$$

The error rate in the distilled magic states is reduced exponentially with respect to the number of noisy input magic states, as given in Equation 1.39 as long as the initial noisy magic states are of sufficiently high fidelity. Throughout this description of magic state distillation we have assumed that the distillation process can be done with perfect Clifford gates, however in practice a certain level of noise will always remain on the Clifford gates in the protocol. In Section 2 and 3, a study of the convergence rate of the distillation protocol is performed under certain classes of errors in the state preparation and gate application in the distillation scheme.

1.5.3 Other schemes

Many efforts have contributed to building protocols for magic state distillation and achieving tight noise thresholds for the noisy input ancillas to the distillation protocol in order

to understand the transition from classically simulatable quantum computation to genuine quantum computation. Bravyi and Kitaev provided another scheme that distilled a different magic state based of a 15-qubit code [1]. Reichardt then continued the derivation of many other distillation schemes based on CSS codes, such as the 7-qubit Steane code and 23-qubit Golay code [33]. Recently, Anwar *et al.* devised a distillation scheme for qutrits, again based off the 15-qubit code [34]. Such an array of different methods to derive states that provide universality along with the Clifford gates has led many to question exactly what states provide this notion of universality. Many efforts have been made to close the gap between the set of stabilizer states and those that provide universality [20, 21, 35, 36, 37], yet for certain regions in the Bloch sphere it remains an open question whether distillation is possible. Finally, an experimental demonstration of a single round of Bravyi and Kitaev's distillation protocol of T-type magic states has been performed in Nuclear Magnetic Resonance (NMR) [38].

Chapter 2

Robustness of magic state distillation

The material for Chapters 2 and 3 is taken from a paper I wrote with my colleagues Yafei Yu, Bassam Helou, and Raymond Laflamme at the Institute for Quantum Computing and the Physics Department at the University of Waterloo [39] (at the time of thesis submission the paper had been submitted for peer-review to the journal of Quantum Information and Computation, Rinton Press). The original manuscript can be found on the arXiv at: http://arxiv.org/abs/1205.6715.

The material presented in Chapters 2 and 3 has been edited slightly from its original form in order to comply with the structure of this work.

2.1 Analytic description of the convergence of distillation

As discussed in Section 1.5.1, the input states to Bravyi and Kitaev's magic state distillation protocol are assumed to be along the magic state axis. However, preparing such states may prove to be difficult experimentally. In this section we study the convergence of the distillation scheme under perturbations about the magic state axis. Moreover, we show that for low fidelity input states, the convergence to the magic state may be improved for

states away from the magic state axis. This suggests that while performing a dephasing operation to initialize the input states to be along the axis may be useful, it is not absolutely necessary in certain fidelity regimes.

The first step of Bravyi and Kitaev's distillation protocol [1] is to perform a dephasing operation $\mathcal{T}$ on five copies of the initial state of the quantum system,

$$\mathcal{T}(\rho) = \frac{1}{3}\left(\rho + T\rho T^\dagger + T^\dagger \rho T\right), \tag{2.1}$$

where $T = SH$ is a Clifford group gate. If the initial state of the system is expressed according to its Bloch sphere coordinates (x, y, z),

$$\rho = \frac{1}{2}\big(I + xX + yY + zZ\big), \tag{2.2}$$

the transfomation $\mathcal{T}$ is equivalent to projecting the state (x, y, z) onto the magic state axis connecting the states $|T_0\rangle$ and $|T_1\rangle$ in the Bloch sphere, where $|T_1\rangle$ is the state orthogonal to $|T_0\rangle$,

$$\mathcal{T}(\rho) = \frac{1}{2}\Big(I + \frac{x+y+z}{3}(X+Y+Z)\Big). \tag{2.3}$$

The dephasing operation leads to the ability to derive a clean threshold for the input fidelity of the initial states in order for the magic state distillation protocol to be beneficial. However, errors in the implementation of the quantum information processor could lead to a preparation of states away from the magic state axis. In this section, we provide an analysis of the effectiveness of the magic state distillation protocol for states prepared at an arbitrary location in the Bloch sphere and give modified target fidelities for state distillation under such noisy state preparation.

A circuit for magic state distillation is given in Figure 1.6. Upon following the steps outlined in Section 1.5.1, the initial noisy states must have a fidelity greater than $F_T = \frac{1}{2}(1 + \sqrt{\frac{3}{7}}) \approx 0.8273$ with respect to the magic state in order for the output state to be of higher fidelity, where the fidelity with respect to the magic state is defined as $F(\rho) = \langle T_0|\rho|T_0\rangle$. Repeating the protocol to obtain multiple copies of the state of increased fidelity ρ_m, the process can be iterated to obtain magic states with arbitrarily high fidelity.

We shall consider the scenario where the dephasing gates are omitted from the protocol, and only the gates in the orange box in Figure 1.6 are implemented, followed by post-selection upon obtaining outcomes of "+1" for Z-basis measurements on the top four qubits. The input state is now located at an arbitrary position in the Bloch sphere given by coordinates (x, y, z), then the output state coordinates are as follows:

$$\begin{aligned} x_{out} &= \frac{-z(z^4 - 5y^2 + 5x^2(y^2-1))}{(1 + 5z^2y^2 + 5x^2y^2 + 5z^2x^2)}. \\ y_{out} &= \frac{-y(y^4 - 5x^2 - 5z^2 + 5x^2z^2)}{(1 + 5z^2y^2 + 5x^2y^2 + 5z^2x^2)}, \\ z_{out} &= \frac{-x(x^4 - 5y^2 + 5z^2(y^2-1))}{(1 + 5z^2y^2 + 5x^2y^2 + 5z^2x^2)}, \end{aligned} \tag{2.4}$$

The plane of states that have fidelity F are the states in the Bloch sphere that satisfy,

$$x + y + z = \sqrt{3}(2F - 1). \tag{2.5}$$

For a given plane of constant input fidelity F_{in}, define a new coordinate system for that plane where r is the radial distance of the input state from the magic state axis, and θ as the angle between the distance vector and the modified x axis, as shown in Figure 2.1. The difference between the input fidelity and output fidelity can then be expressed according to these coordinates as

$$d = F_{out} - F_{in} = -\frac{2\Big(a(54 - 60a^2 + 14a^4 + 135r^4) + 15\sqrt{6}(-3 + a^2)r^3\cos 3\theta\Big)}{2\sqrt{3}\Big(108 + 20a^4 + 135r^4 + 60\sqrt{6}ar^3\cos 3\theta\Big)}, \tag{2.6}$$

where a is related to the input fidelity, $a = \sqrt{3}(2F_{in} - 1)$. As such, the ability of the distillation protocol to increase the fidelity of the input states depends on the distance of the initial states from the magic state axis and on the spatial angle with respect to the modified x axis as well as the input fidelity.

Figure 2.2 demonstrates the dependence on the distance from the magic state axis and angle for convergence to the magic state after many iterations for the set of states on the plane with fidelity $F = 0.886$ in the positive octant $(x, y, z > 0)$. States close to the magic

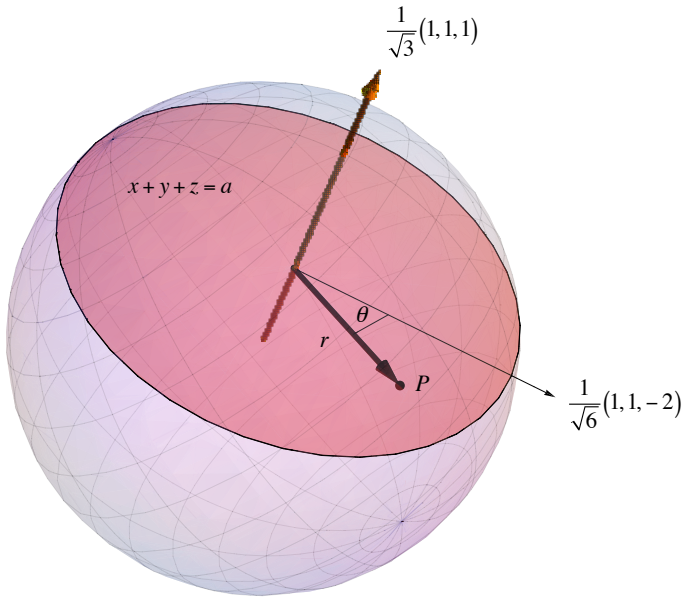


Figure 2.1: Labelling of the coordinates in the Bloch sphere used to derived the fidelity difference after a single run of the magic state distillation protocol, given in Eq. 2.6. The red plane is the plane of constant fidelity $x+y+z=a$, the orange arrow is the magic state axis, labeled by its normalized Bloch sphere coordinates $(1,1,1)/\sqrt{3}$, and the axis labelled by the vector $(1,1,-2)/\sqrt{6}$ is the modified x-axis. For a given point P, the coordinates a, r, and θ are defined as in the figure.

state axis converge to the magic state, while states far away from the axis converge to other undesirable states in the Bloch sphere. Perhaps more surprisingly, there are states below the distillation threshold set by Bravyi and Kitaev [1] that can converge to the magic state. As Figure 2.3 shows, states away from the magic state axis on the fidelity plane of $F = 0.8270$ converge to the magic state, while those on the axis converge to the maximally mixed state, as expected as F is below the fidelity threshold for states on the axis. Notice in Eq. 2.6 that the maximal increase in the fidelity of the state after the distillation procedure occurs for angles $\theta = 0,\ 2\pi/3,\ 4\pi/3$. Fixing $\theta = 0$, the difference in fidelity can show an increase as a function of the distance from the magic state axis r for small distances before

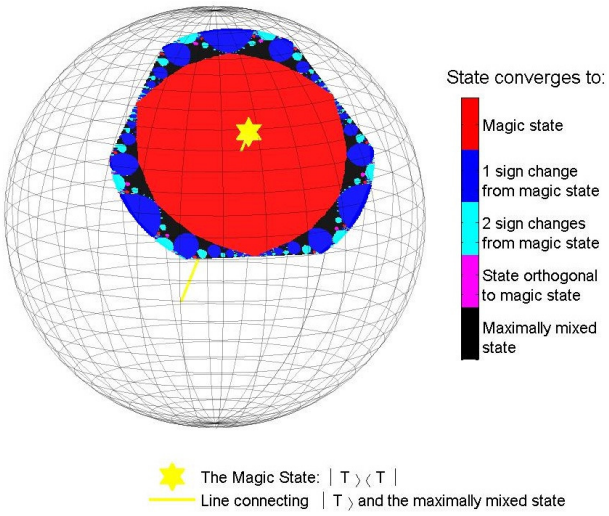


Figure 2.2: Final states of convergence after multiple rounds of distillation for states on the fidelity plane $F = 0.886$, that is a cut of the Bloch sphere with states of a fixed fidelity, as given by Equation 2.5. The states in red represent the states that converge to the magic state after multiple rounds of distillation. The states in light and dark blue converge to states with the coordinates $(\pm 1, \pm 1, \pm 1)/\sqrt{3}$ in the Bloch sphere, where the number of sign changes from the coordinates of the $|T\rangle$ magic state $(1, 1, 1)/\sqrt{3}$ is given by the shading of blue. The states in pink converge to the state orthogonal to $|T\rangle$, and black to the maximally mixed state.

dropping off rapidly as the distance from the axis grows. For an input fidelity just below the fidelity threshold set for on-axis states, the fidelity difference can be negative for states close to the axis and increase to be positive for certain distances away from the axis, as Figure 2.4(a) shows, but only for states whose angle is close to the angles of maximal increase, as plotted in Figure 2.4(b). Therefore at angles close to the angles of maximal increase, one can obtain states whose fidelity threshold can be below the on-axis threshold of $\frac{1}{2}(1+\sqrt{\frac{3}{7}}) \approx 0.8273$. As such, by performing the magic state distillation routine without the dephasing transformation, one can slightly lower the threshold for the fidelity of input states to be 0.8250 for states at the angle of maximal increase for the T-type distillation

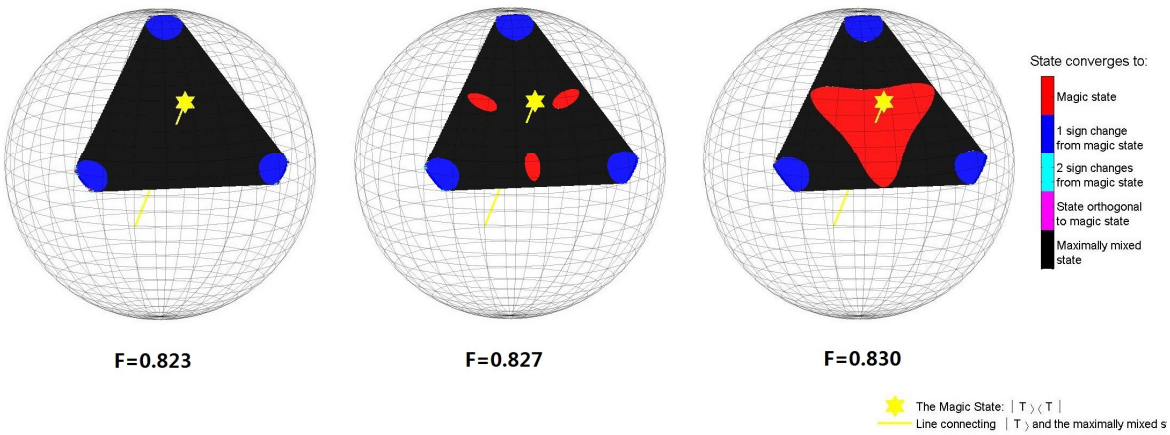


Figure 2.3: Convergence of the states around the perfect Clifford gate threshold 0.8273 for states along the magic state axis. All states much below the theoretical threshold ($F = 0.823$) converge away from the magic state axis, yet some states slightly below the threshold for on axis convergence converge to the magic state ($F = 0.8270$). This convergence only appears for states close to the three angles of maximum output fidelity on the fidelity plane. All states close to the magic state axis converge to the magic state for fidelities above the threshold ($F = 0.830$).

scheme.

We may thus conclude that the dephasing operation that projects the noisy magic states onto the magic state axis is not necessarily needed as there are regions of convergence around this axis. As such, noisy dephasing processes that would have slight perturbations off the magic state axis would not be detrimental to the convergence to the magic state. Furthermore, for certain states off the magic state axis and of fidelity just below that derived by Bravyi and Kitaev, omitting the dephasing operation would be beneficial as it would allow such states to remain useful for magic state distillation.

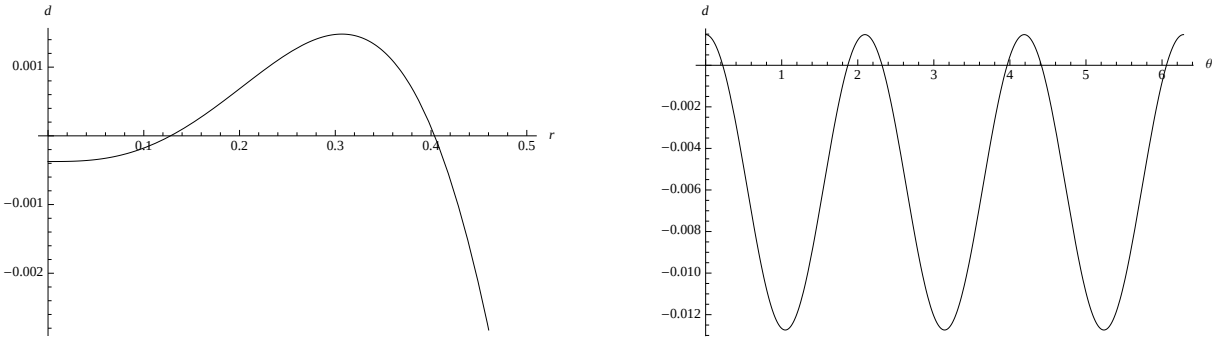


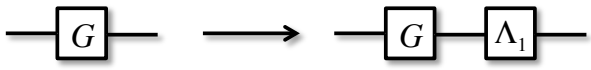
Figure 2.4: (a) Dependance of the difference in fidelity d, as defined in Equation 2.6, between the initial and final states as a function of distance from the magic state axis in the Bloch sphere, at fixed angle $\theta = 0$ and fidelity $F = 0.8269$. (b) Dependance of distance in fidelity between the initial and final states as a function of angle for fixed distance from the Bloch sphere $r = 0.3$ and fidelity $F = 0.8269$.

Chapter 3

Noisy magic state distillation

3.1 Effects of depolarizing noise on the convergence rate

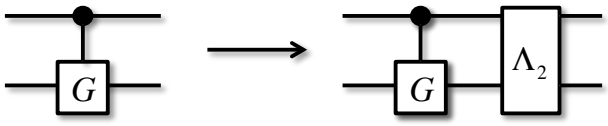
In an experimental realization of the implementation of magic state distillation, any quantum gate will introduce noise. As such, an interesting experimental question would be: to what level of noise is the distillation still beneficial? In order to address such a question, we consider the case where all the gates in the decoding circuit, shown by the boxed region in Figure 1.6, are subjected to depolarizing noise. Depolarizing noise is a common noise model for physical implementations of quantum information processing. A noisy one-qubit Clifford gate G will be modelled as follows:



where the gate Λ_1 performs the depolarizing transformation with a noise parameter p_1,

$$\Lambda_1(\rho, p_1) \to (1-p_1)\rho + p_1\frac{I}{2} = (1-\frac{3p_1}{4})\rho + \frac{p_1}{4}(X\rho X + Y\rho Y + Z\rho Z). \tag{3.1}$$

Similarly, we introduce noise to a two-qubit controlled Clifford gate by adding a two-qubit depolarizing gate, where the two-qubit depolarizing gate Λ_2 is defined by the transformation



with a two-qubit noise parameter p_2,

$$\begin{aligned}\Lambda_2(\rho, p_2) &\to (1-p_2)\rho + p_2\frac{I\otimes I}{4} \\ &= \left(1-\frac{15p_2}{16}\right)\rho + \frac{p_2}{16}\Big((I\otimes X)\rho(I\otimes X) + (I\otimes Y)\rho(I\otimes Y) + \dots \\ &\qquad + (Z\otimes Y)\rho(Z\otimes Y) + (Z\otimes Z)\rho(Z\otimes Z)\Big). \end{aligned} \tag{3.2}$$

We shall only consider errors affecting the gates in the decoding procedure. In order to compare the fidelity threshold for the noisy decoding procedure with the ideal magic state distillation protocol proposed by Bravyi and Kitaev [1], we assume the input states to the decoding procedure are along the magic state axis. This assumption is valid in the error strength regime we will be considering, as depolarizing errors in the dephasing transformation would lead to a negligible deviation of the input states from the magic state axis. We can relate the parameter p_i to the error per gate, which is defined as $E_i = 1 - F_i$, where F_i is the fidelity of the gate i. These values are related by the simple relationships: $p_1 = 2E_1$ and $p_2 = 4E_2/3$.

Assuming the error rate of the one and two-qubit gates are low (omitting quadratic and higher order terms in p_1 and p_2), the output (unnormalized) state will have the following matrix entries in the $|T_0\rangle$–$|T_1\rangle$ basis:

$$\begin{aligned}|T_0\rangle\langle T_0| : \frac{1-5p_1-8p_2}{6}\Big[(1-\epsilon)^5 + 5\epsilon^3(1-\epsilon)^2\Big] + \frac{p_1}{36}\Big[19 - 87\epsilon + 197\epsilon^2 - 164\epsilon^3 + 6\epsilon^4 + 32\epsilon^5\Big] \\ + \frac{p_2}{54}\Big[20 - 44\epsilon + 107\epsilon^2 - 106\epsilon^3 + 28\epsilon^4 + 8\epsilon^5\Big], \end{aligned} \tag{3.3}$$

$$\begin{aligned}|T_1\rangle\langle T_1| : \frac{1-5p_1-8p_2}{6}\Big[\epsilon^5 + 5\epsilon^2(1-\epsilon)^3\Big] + \frac{p_1}{36}\Big[3 + \epsilon + 61\epsilon^2 - 180\epsilon^3 + 166\epsilon^4 - 32\epsilon^5\Big] \\ + \frac{p_2}{54}\Big[13 - 4\epsilon + 37\epsilon^2 - 86\epsilon^3 + 68\epsilon^4 - 8\epsilon^5\Big], \end{aligned} \tag{3.4}$$

$$|T_0\rangle\langle T_1| : \frac{p_2(1+i)(-1+2\epsilon)}{432}\Big[(-2+3i)-(1+5i)\sqrt{3}+((-2-9i)+(3+10i)\sqrt{3})\epsilon + ((6+9i)-(3+6i)\sqrt{3})\epsilon^2+((-8-6i)+(2-8i)\sqrt{3})\epsilon^3+(4+4i\sqrt{3})\epsilon^4\Big], \tag{3.5}$$

where $\epsilon = 1 - \langle T_0|\rho_{\text{in}}|T_0\rangle$ is the error of the initial state. Note that depolarizing noise in the decoding procedure will introduce off diagonal terms in the $|T_0\rangle$–$|T_1\rangle$ basis, that is, will produce an output state that deviates away from the magic state axis. However, such matrix elements will be on the order of p_2, which for low levels of noise will be negligible compared to the strength of the coefficients along the diagonal terms. Therefore, we may assume that upon iterating the distillation protocol, the input states will always remain along the magic state axis. This assumption is also well motivated from the results of Chapter 2 that show that slight deviations off the magic state axis will not affect the convergence of the distillation scheme. For small values of ϵ, the output error ϵ_{out}, the normalized coefficient of the $|T_1\rangle\langle T_1|$ term, can be approximated to be $5\epsilon^2 + p_1/2 + 13p_2/9$. Thus in the limit of infinite iterations of the distillation protocol the error rate will be linear in the terms p_1 and p_2.

Motivated by the average one and two-qubit gate errors in a recent benchmarking experiment in NMR [40], we have chosen different values for the noise parameters p_1 and p_2. The graph in Figure 3.1 shows the output fidelity of the magic state distillation protocol for different error strengths of the one and two-qubit gates. The results for error strengths given by the benchmarking NMR results, $E_1 = 1.3 \times 10^{-4}$ and $E_2 = 4.7 \times 10^{-3}$, are given by the black curve. In order to obtain an appreciation of the decrease in output fidelity cause by the strength of both of these errors, we have also plotted the output fidelity corresponding to the case when the one and two-qubit errors have the same strength, namely 1.3×10^{-4} and 4.7×10^{-3}. Note that when both errors are on the order of 10^{-4} the decrease in fidelity is minimal, this leads one to conclude that the decrease in fidelity in the black curve is caused mostly by the larger error on the two-qubit gates. Finally, note that for the case of choosing an error model based on the results from Ref. [40], the new threshold for the minimal input fidelity is 0.842, which is larger than the theoretical noiseless threshold of 0.8273, and the maximum output fidelity that can be reached through repeated applications of the distillation protocol is 0.9895 with respect to the magic state.

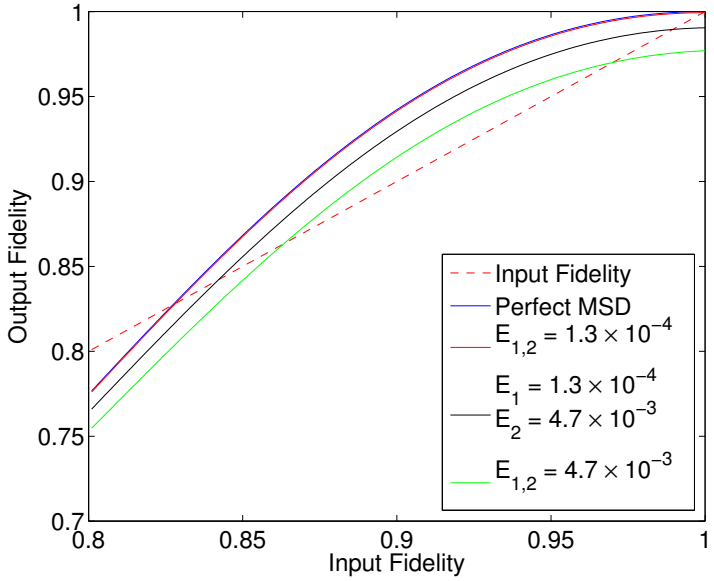


Figure 3.1: Output fidelities of the magic state distillation protocol as a function of the input fidelity of the five input states. The uppermost line indicates the noiseless magic state distillation, while the three other lines show a decrease of the output fidelity caused by an increase in the values of E_1 and E_2. As the error strength increases, the class of states for which the distillation protocol is beneficial decreases. Notice that once errors are introduced one can no longer increase the fidelity to be arbitrarily close to 1 for repeated applications of the protocol.

Table 3.1 shows how the threshold rate for magic state distillation and the maximal distillation fidelity is affected by different levels of one and two-qubit noise, p_1 and p_2. For levels of noise below 10^{-2} in either parameter, the distillation protocol performs very well. However, once the error rate increases above this level, the threshold for noisy magic state distillation begins to rise rapidely while the maximal fidelity of distillation begins to drop. The noise rate at which the distillation protocol no longer has any benefit (in the case where $p_1 = p_2$) is approximately 0.01581.

Due to the errors on the applied Clifford gates in the decoding procedure, distilling

p_1	p_2	$F_{\text{thres.}}$	$F_{\max}$
10^{-5}	0	0.8273 ± 10^{-4}	0.999995 ± 10^{-6}
10^{-4}	0	0.8274 ± 10^{-4}	0.99995 ± 10^{-5}
10^{-3}	0	0.8281 ± 10^{-4}	0.9995 ± 10^{-4}
10^{-2}	0	0.9648 ± 10^{-4}	0.9948 ± 10^{-4}
10^{-1}	0	n/a	n/a
0	10^{-5}	0.8274 ± 10^{-4}	0.999990 ± 10^{-6}
0	10^{-4}	0.8275 ± 10^{-4}	0.99986 ± 10^{-5}
0	10^{-3}	0.8295 ± 10^{-4}	0.9985 ± 10^{-4}
0	10^{-2}	0.8524 ± 10^{-4}	0.9820 ± 10^{-4}
0	10^{-1}	n/a	n/a
10^{-2}	10^{-2}	0.8642 ± 10^{-4}	0.9731 ± 10^{-4}
1.1×10^{-2}	1.1×10^{-2}	0.8694 ± 10^{-4}	0.9689 ± 10^{-4}
1.2×10^{-2}	1.2×10^{-2}	0.8752 ± 10^{-4}	0.9642 ± 10^{-4}
1.3×10^{-2}	1.3×10^{-2}	0.8819 ± 10^{-4}	0.9586 ± 10^{-4}
1.4×10^{-2}	1.4×10^{-2}	0.8899 ± 10^{-4}	0.9517 ± 10^{-4}
1.5×10^{-2}	1.5×10^{-2}	0.9006 ± 10^{-4}	0.9421 ± 10^{-4}
1.581×10^{-2}	1.581×10^{-2}	0.9214 ± 10^{-4}	0.9223 ± 10^{-4}

Table 3.1: Calculated values of the fidelity threshold F_{thres} for the noisy input magic states and the maximal fidelity of distillation $F_{\max}$ for noisy magic state distillation. The noise rates of the Clifford gates in the protocol are given by the parameters p_1 and p_2. The final entry of the table shows where the threshold fidelity and maximal fidelity of distillation become very close, signalling that distillation is not a useful process at levels of noise where $p_1 = p_2 = 0.01581$. The error bars on the results of the fidelity thresholds are a consequence of the finite precision of the numerical simulation that was performed to derive such bounds. The two sets of fidelity thresholds denoted "n/a" represent error regimes where magic state distillation is not beneficial for any input magic state fidelity.

states arbitrarily close to the magic state is no longer possible, thus the distilled output state will have the form,

$$\rho'_m = (1-\epsilon')|T_0\rangle\langle T_0| + \epsilon'|T_1\rangle\langle T_1|, \tag{3.6}$$

and the value of the error ϵ' will be fixed away from 0. However, the application of a universal non-Clifford gate is still possible with such a state, albeit with a certain level of noise. Following the procedure laid out in Ref. [1], shown in Figure 1.5, one can achieve the application of the gate $\Lambda_{\pi/12}$ to an arbitrary state $|\psi\rangle = a|0\rangle + b|1\rangle$ by using multiple copies of the noisy magic state ρ'_m instead of the perfect magic states that are given in Figure 1.5. The fidelity of the applied gate is

$$F = 1 - \frac{12\epsilon'|a|^2(1-|a|)^2}{3+(1-2\epsilon')^2} \geq 1-\epsilon', \tag{3.7}$$

where ϵ' is the error rate of the distilled magic state given in Equation 3.6 (see Appendix A for the calculation). As such, the lower bound for the fidelity of the applied universal gate is $1-\epsilon'$, which corresponds to the minimum gate fidelity. After multiple iterations of the distillation protocol, if the error of initial state was below the threshold for state distillation, the error rate ϵ' will be approximately $p_1/2 + 13p_2/9$. Thus, in order to apply the universal gate with fidelity above $1-\epsilon'$, one would typically have to apply a final fault-tolerant iteration in order to reduce the error rates of the original Clifford gates to a smaller logical error, as will be discussed in Section 3.2.

3.2 Comparison study between noisy MSD and fault-tolerant MSD

Fault-tolerant quantum computation is a method to reduce the error rate of a given quantum operation using states and gates encoded into a higher-dimensional Hilbert space. Such encodings, combined with the ability of projective measurement and post-selection, provide a means to increase the fidelity of our quantum gates at the expense of using additional qubits and quantum gates to perform the desired encodings of states and encoded operations [9, 10, 11, 12, 13, 41].

Consider a scenario where we are presented with a physical device that is not restricted in the number of qubits at its disposal but limited in the number of gates that can be applied coherently. We could then envision two different methods for applying the magic state distillation procedure, either by applying the protocol with the faulty gates at our disposal or first performing a fault-tolerant encoding of the qubits and then performing the distillation protocol with less noisy encoded logical quantum gates. The faulty distillation protocol, as described and analyzed in Section 3.1, suffers from a reduced convergence rate and a fundamental limit in the ability to distill a magic state with high fidelity. Thus using a fault-tolerance encoding of the Clifford gates appears to be advantageous. However the cost associated with encoding each of the Clifford gates in a fault-tolerant encoding outweighs the savings one would obtain in the number of Clifford gates one would need to apply at lower noise rates in the faulty magic state distillation scheme.

Figure 3.2 plots the total number of two-qubit gates that would be required for either scheme to achieve a desired final fidelity for the magic state. The red dots plotted in Figure 3.2 are the number of gates in the ideal setting with no noise, this can be thought of as a lower bound of the number of *logical* fault-tolerant gates that would need to be applied for any fault-tolerant encoding. As expected, the number of gates exhibit a step function behaviour, characteristic of the number of iterations of the distillation subroutine that would be required to achieve the final fidelity target. The smearing of the steps is due to the probability of measuring the trivial syndrome, which is reduced for initial states with lower fidelity.

One notices that the number of iterations for the fault-tolerant magic state distillation is lower than that of the faulty procedure, this is due to the fact that noise in faulty Clifford gates can decrease the convergent rate for states along the magic state axis and will lower the probability of measuring the trivial syndrome. Thus if one were to be presented with faulty Clifford gates and were to encode the states such that the error rate of the encoded operations were negligible, then the total number of encoded *logical* gates required would be lower than the number of faulty Clifford gates to increase the fidelity of the magic state to a desired level. However, in the regime of typical one and two-qubit gate errors, 10^{-3}–10^{-2}, the fault-tolerant encodings whose threshold rates are above these levels typically use on the order of 100–1000 two-qubit physical gates per encoded logical

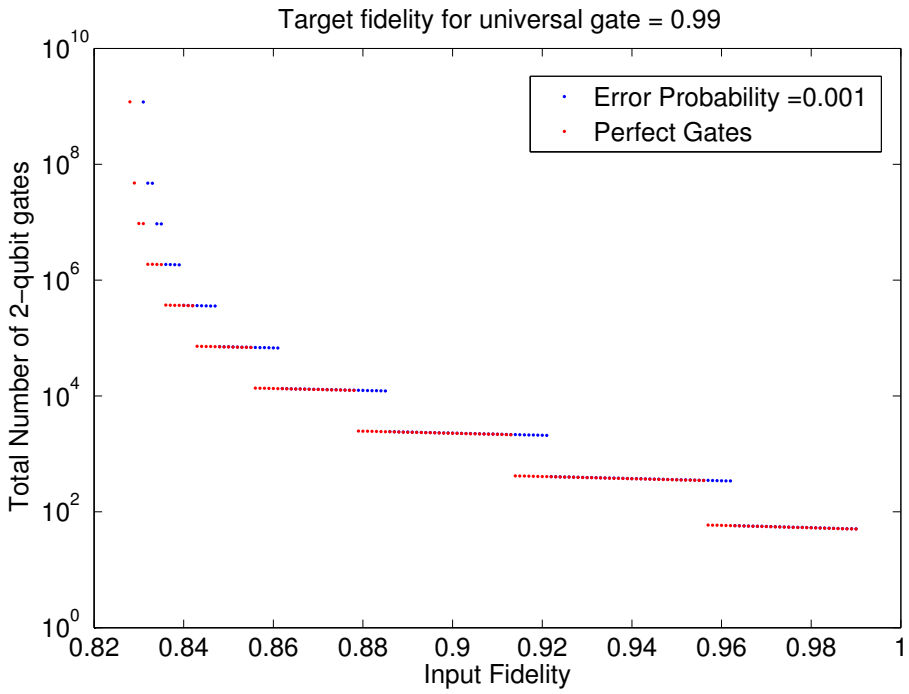


Figure 3.2: Total number of two-qubit gates as a function of input fidelity for the state distillation scheme to achieve an output fidelity of 0.99 for the magic state. The red dots denote the number of gates required for perfect gates and can be thought of as a lower bound in the number of fault-tolerant encoded logical gates that would be required. The blue dots represent the total number of gates for the faulty distillation protocol, where the gate error probability E_1, E_2 is 0.001. Note that the y-axis in this figure is plotted logarithmically in order to capture that the increase in the number of gates for each iteration of the distillation scheme scales approximately by a factor of 5.

gate [42]. As such, in order for a fault-tolerant encoded scheme to show an improvement in the total number of gates required for the full distillation protocol, the encoded scheme would have to use on the order of 100–1000 times less logical gates than that of the faulty distillation scheme. Consider the results from Figure 3.2, which compares the number of faulty gates, with error probability 10^{-3}, and the ideal case of encoded gates with no error. The reduction in the number of gates with no error compared to the faulty distillation scheme is greatest at the jumps in the number of gates that occur due to an increased number of iterations required. As such, the faulty distillation scheme may require an additional iteration of the distillation scheme, which is just repeating 5 times the previous level of distillation, leading to an overall gate cost on the order of 5 times more gates. However, as mentioned above, in order to reduce the error rate of the encoded states to be below that of the original error rate, each encoded *logical* gate would require on the order of 100–1000 *physical* gates. Therefore, while the number of logical gates is lower than the number of faulty gates required, the cost of each of the logical gates would drive the overall number of gates required in a fault-tolerant scheme to be much larger. We thus conclude that using faulty Clifford gates is more efficient for magic state distillation at noise levels comparable to those in current quantum information implementations. This analysis considers the number of two-qubit gates that would be required in the distillation protocol for either scheme as typical numbers of two-qubit gates have been studied extensively in past works [42], however such an analysis could be extended to one-qubit gates and we believe that the behaviour will be equivalent.

To summarize, this result would be important in a regime where Clifford gates have an error rate below that required for faulty magic state distillation, and applying a universal gate would be very faulty. Such a situation could arise if one was already at a certain level of fault-tolerant encoding, thus reducing the original physical Clifford gate error rates. In such a regime, performing faulty magic state distillation is more efficient than going to a further level of fault-tolerant encoding to reduce the error rates to be even smaller in order to perform close to perfect magic state distillation.

Finally, one should note that fault-tolerance would be required if the errors in the one and two-qubit gates were too high, preventing the distilled state to achieve the desired target fidelity for a prepared magic state after multiple rounds of the distillation. This is

due to the error of the distilled state always being approximately bounded from below by $p_1/2 + 13p_2/9$, where p_1 and p_2 are the error probabilities of the one and two-qubit gates, respectively. Therefore, if one requires to reduce the error of the distilled magic state, and subsequently the applied universal gate using the magic states, to be below some small target threshold, one would need to reduce the error rates of the one and two-qubit gates through a method such as fault-tolerance. However, this step would only have to be applied in the final iterations of the protocol in order to get the distilled state over the fidelity hump posed by the errors in the Clifford gates at one's disposal.

Chapter 4

Conclusion

In this work we have provided an extensive study on the robustness of magic state distillation to noisy input states and noise on the single and two-qubit Clifford gates that motivate the use of magic state distillation as a method to achieve universal quantum computation. Our findings have shown that magic state distillation is robust to noisy Clifford gates and that applying the magic state distillation scheme with noisy gates is an efficient process.

The first half of our analysis comprised of characterizing which input states to the distillation protocol would converge to the magic state after multiple rounds of T-type magic state distillation. Bravyi and Kitaev [1] provided an analysis of input states along the magic state axis based on the ability to perform a dephasing operation, however we extended the analysis to all input states in the Bloch sphere. The reason for such an analysis was two-fold: firstly, the dephasing operation that projects all states in the Bloch sphere to the magic state axis may be noisy and as such knowing if the states around the magic state axis converge is of prime importance. Secondly, there may be states off the magic state axis that converge more quickly than those on the magic state axis. We found that the magic state distillation process is indeed robust to such perturbations about the magic state axis as being slightly away from the axis does not greatly change convergence rate of the states. Additionally, we found that certain off-axis states of lower fidelity than the previously known threshold converge to the magic state. States of fidelity 0.8250 with respect to the magic state were shown to converge in a distillation scheme when the

previous minimal fidelity threshold derived for on-axis states was 0.8273. We conclude that not only is the dephasing transformation not always necessary, there are input states for which it can be detrimental their convergence to the magic state.

The second half of this project consisted of determining the resource cost for a noisy distillation scheme in terms of the number of quantum gates. The main question that was addressed was the trade-off between using noisy Clifford gates in the distillation scheme or very low noise Clifford gates that have been encoded fault-tolerantly. The noisy distillation scheme suffered from reduced convergence rates and an increase in the threshold for the input noisy magic states, yet distillation is still possible up to an error rate that is linear with respect to the error rates of the one- and two-qubit gates in a depolarizing error model. We showed that for levels of depolarizing noise that are under 10^{-2}, the magic state distillation scheme with noisy Clifford gates is very beneficial and almost as efficient as the perfect distillation scheme. As a consequence, we showed that if the fidelity of the required distilled magic state is not too high, a fault-tolerant encoding of the Clifford gates is not necessary and that the noisy implementation is the best approach for preparing states for universal quantum computation. However, in the case of wanting to prepare a high fidelity magic state for high fidelity universal gate applications, a fault-tolerant encoding of the Clifford gates can be applied in the final iterations of the protocol, but are not needed throughout the rest of the protocol as long as the initial noise rate of the Clifford gates is below 10^{-2}.

There are a few promising directions for this work and other works on magic state distillation. Firstly, it would be interesting to extend a similar analysis to the 15-qubit distillation code for H-type magic state distillation which can be used to implement a universal $\pi/8$-gate. Additionally, while studying the depolarizing error model provides a strong intuition on the effects of random errors in a quantum system, one could further extend the analysis to different error models. It could be that for different error models, the magic state distillation circuit chosen in this work is not as efficient as another choice of 5-qubit decoder, yet the circuit chosen would have to be on a case-by-case basis depending on the error model of quantum computation. Another path for possible research is to study the recycling of magic states, that is, would it be possible to use the garbage states that are thrown out at the end of the application of the universal gate in order to distill further

magic states. Work has been done showing that such output states are not believed to possibly be magic states [37], yet it could be possible that they could have some use in further distillation schemes, which could save some of the overhead in the overall universal gate scheme.

APPENDICES

Appendix A

Universal gate application with noisy magic states

In this Appendix we shall derive the bound on the error rate, given in Equation 3.7, for the application of the universal gate scheme, presented in Figure 1.5, with noisy magic states of the form of Equation 3.6. For the purpose of this calculation the Clifford gates are assumed to be perfect in order to characterize the error in the gate application due to the noise in the distilled magic state.

Consider the implementation of the two qubit operations on the last two qubits given in Figure 1.5, the input state has the form,

$$\begin{aligned}\rho'_m \otimes \rho'_m &= \Big((1-\epsilon')|T_0\rangle\langle T_0| + \epsilon'|T_1\rangle\langle T_1|\Big) \otimes \Big((1-\epsilon')|T_0\rangle\langle T_0| + \epsilon'|T_1\rangle\langle T_1|\Big) \\ &= \frac{1}{2}\Big(I + \frac{1-2\epsilon'}{\sqrt{3}}(X+Y+Z)\Big) \otimes \frac{1}{2}\Big(I + \frac{1-2\epsilon'}{\sqrt{3}}(X+Y+Z)\Big). \qquad \text{(A.1)}\end{aligned}$$

The state after post-selection of the "+1" eigenstate of the $Z \otimes Z$ observable is then

$$\begin{aligned}\frac{1}{\frac{1}{2}+\frac{1-2\epsilon'}{6}}\Big(\frac{1}{4}\big(1+\frac{1-2\epsilon'}{\sqrt{3}}\big)^2|00\rangle\langle 00| \quad &+\frac{2(1-2\epsilon')}{3}(-i|00\rangle\langle 11| + i|11\rangle\langle 00|) \\ &+\frac{1}{4}\big(1-\frac{1-2\epsilon'}{\sqrt{3}}\big)^2|11\rangle\langle 11|\Big). \qquad \text{(A.2)}\end{aligned}$$

By next performing a $CNOT$ gate, the state is projected back onto a single qubit state with an ancilla qubit $|0\rangle\langle 0|$, the qubit of information is characterized by the following Bloch sphere coordinates,

$$\frac{1}{2}\Big(I + \frac{2(1-2\epsilon')^2}{3+(1-2\epsilon')^2}Y + \frac{2\sqrt{3}(1-2\epsilon')}{3+(1-2\epsilon')^2}Z\Big). \tag{A.3}$$

The application of the Hadamard transformation maps the Bloch sphere coordinates, $Z \to X$ and $Y \to -Y$,

$$\frac{1}{2}\Big(I - \frac{2(1-2\epsilon')^2}{3+(1-2\epsilon')^2}Y + \frac{2\sqrt{3}(1-2\epsilon')}{3+(1-2\epsilon')^2}X\Big) = \begin{pmatrix} 1/2 & f(\epsilon')e^{i\theta} \\ f(\epsilon')e^{-i\theta} & 1/2 \end{pmatrix}, \tag{A.4}$$

where

$$f(\epsilon') = \frac{(1-2\epsilon')}{\sqrt{3+(1-2\epsilon')^2}}, \qquad \tan\theta = \frac{(1-2\epsilon')}{\sqrt{3}}. \tag{A.5}$$

Consider an arbitrary pure state $|\psi\rangle\langle\psi|$, $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, on which we would like to apply the $\pi/12$-gate. Performing the $Z \otimes Z$ measurement followed by the $CNOT$ on the pure state $|\psi\rangle\langle\psi|$ along with the prepared state will result in the following state on the first register, where the state is dependant on the measurement outcome,

$$\text{“}+1\text{” syndrome:} \qquad \begin{pmatrix} |\alpha|^2 & 2f(\epsilon')e^{i\theta}\alpha\beta^* \\ 2f(\epsilon')e^{-i\theta}\alpha^*\beta & |\beta|^2 \end{pmatrix} \tag{A.6}$$

$$\text{“}-1\text{” syndrome:} \qquad \begin{pmatrix} |\alpha|^2 & 2f(\epsilon')e^{-i\theta}\alpha\beta^* \\ 2f(\epsilon')e^{i\theta}\alpha^*\beta & |\beta|^2 \end{pmatrix}. \tag{A.7}$$

Comparing these outcomes to that of the application of the perfect $\pi/12$ or $-\pi/12$-gate, depending on measurement outcome,

$$\text{“}+1\text{” syndrome:} \qquad \begin{pmatrix} |\alpha|^2 & e^{i\pi/6}\alpha\beta^* \\ e^{-i\pi/6}\alpha^*\beta & |\beta|^2 \end{pmatrix} \tag{A.8}$$

$$\text{“}-1\text{” syndrome:} \qquad \begin{pmatrix} |\alpha|^2 & e^{-i\pi/6}\alpha\beta^* \\ e^{i\pi/6}\alpha^*\beta & |\beta|^2 \end{pmatrix}, \tag{A.9}$$

the fidelity of the imperfect gate application can be bounded from below by,

$$F = 1 - \frac{12\epsilon'|a|^2(1-|a|)^2}{3+(1-2\epsilon')^2} \geq 1-\epsilon', \tag{A.10}$$

which corresponds to the result in Equation 3.7.

References

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

[2] R. P. Feynman, "Simulating physics with computers," *International Journal of Theoretical Physics*, vol. 21, p. 467, 1982.

[3] P. Shor, "Algorithms for quantum computation: Discrete logarithms and factoring," *Proceedings., 35th Annual Symposium on Foundations of Computer Science*, pp. 124–134, 1994.

[4] R. Landauer, "Irreversibility and heat generation in the computing process," *IBM Journal of Research and Development*, vol. 5, pp. 183–191, 1961.

[5] P. W. Shor, "Scheme for reducing decoherence in quantum computer memory," *Phys. Rev. A*, vol. 52, p. 2493, 1995.

[6] A. W. Steane, "Error correcting codes in quantum theory," *Phys. Rev. Lett.*, vol. 77, p. 793, 1996.

[7] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, "Mixed state entanglement and quantum error correction," *Phys. Rev. A*, no. 54, p. 3824, 1996.

[8] E. Knill and R. Laflamme, "Theory of quantum error-correcting codes," *Phys. Rev. A*, vol. 55, pp. 900–911, 1997.

[9] P. Shor, "Fault-tolerant quantum computation," *Proceedings., 37th Annual Symposium on Foundations of Computer Science*, pp. 56–65, 1996.

[10] D. Gottesman, "Stabilizer Codes and Quantum Error Correction," *Ph.D. thesis, California Institute of Technology, Pasadena, CA*, 1997.

[11] E. Knill, R. Laflamme, and W. H. Zurek, "Resilient quantum computation," *Science*, vol. 279, pp. 342–345, 1998.

[12] J. Preskill, "Reliable quantum computers," *Proc. R. Soc. London A*, vol. 454, pp. 385–410, 1998.

[13] D. Aharonov, M. Ben-Or, R. Impagliazzo, and N. Nisan, "Limitations of noisy reversible computation," *arXiv:quant-ph/9611028*, 1996.

[14] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, England, 2000.

[15] C. M. Dawson and M. Nielsen, "The solovay-kitaev algorithm," *Quant. Inf. Comput.*, vol. 6, pp. 85–91, 2006.

[16] W. F. Stinespring, "Positive Functions on C*-algebras," *Proceedings of the American Mathematical Society*, pp. 211–216, 1955.

[17] E. Knill, "Fault-tolerant postselected quantum computation: schemes," *arXiv:quant-ph/0402171*, 2004.

[18] D. Gottesman and I. Chuang, "Quantum teleportation is a universal computational primitive," *Nature*, vol. 42, pp. 390–393, 1999.

[19] S. Aaronson and D. Gottesman, "Improved simulation of stabilizer circuits," *Phys. Rev. A*, vol. 70, p. 052328, 2004.

[20] B. W. Reichardt, "Quantum universality by state distillation," *Quant. Inf. Comput.*, vol. 9, pp. 1030–1052, 2005.

[21] E. Campbell and D. Browne, *On the Structure of Protocols for Magic State Distillation*, vol. 5906 of *Lecture Notes in Computer Science*. Springer Berlin / Heidelberg, 2009.

[22] M. Born, "Zur Quantenmechanik der Stoßvorgnge," *Zeitschrift fr Physik*, pp. 863–867, 1926.

[23] J. Emerson, "Open Quantum Systems (course notes)," 2011.

[24] D. DiVincenzo, "Two-qubit gates are universal for quantum computation," *Phys. Rev. A*, vol. 51, pp. 1015–1022, 1995.

[25] P. O. Boykin, T. Mor, M. Pulver, V. Roychowdhury, and F. Vatan, "On universal and fault-tolerant quantum computation," *Proceedings., 40th Annual Symposium on Foundations of Computer Science*, pp. 486–494, 1999.

[26] P. Kaye, R. Laflamme, and M. Mosca, *An Introduction to Quantum Computing*. Oxford University Press, Oxford, England, 2007.

[27] R. Laflamme, C. Miquel, J. P. Paz, and W. H. Zurek, "Bound states for magic state distillation in fault-tolerant quantum computation," *Phys. Rev. Lett.*, vol. 77, no. 198, 1996.

[28] A. W. Steane, "Multiple-Particle Interference and Quantum Error Correction," *Proc. Roy. Soc. Lond.*, vol. 452, pp. 2551–2577, 1996.

[29] X. Chen, H. Chung, A. W. Cross, B. Zeng, and I. L. Chuang, "Subsystem stabilizer codes cannot have a universal set of transversal gates for even one encoded qudit," *Phys. Rev. A*, vol. 78, p. 012353, 2008.

[30] B. Zeng, A. W. Cross, and I. L. Chuang, "Transversality Versus Universality for Additive Quantum Codes," *IEEE Transactions on Information Theory*, vol. 57, 2011.

[31] B. Eastin and E. Knill, "Restrictions on transversal encoded quantum gate sets," *Phys. Rev. Lett.*, vol. 102, p. 110502, 2009.

[32] M. Grassl, "Bounds on the minimum distance of linear codes and quantum codes." Online available at http://www.codetables.de, 2007.

[33] B. W. Reichardt, "Quantum universality from magic states distillation applied to CSS codes," *Quant. Info. Proc.*, vol. 4, p. 251, 2005.

[34] H. Anwar, E. T. Campbell, and D. E. Browne, "Qutrit magic state distillation," *arXiv:1202.2326*, 2012.

[35] E. T. Campbell and D. E. Browne, "Bound states for magic state distillation in fault-tolerant quantum computation," *Phys. Rev. Lett.*, vol. 104, p. 030503, 2010.

[36] V. Veitch, C. Ferrie, and J. Emerson, "Negative Quasi-Probability Representation is a Necessary Resource for Magic State Distillation," *arXiv:1201.1256*, 2012.

[37] J. T. Anderson, "On the Power of Reusable Magic States," *arXiv:1205.0289*, 2012.

[38] A. M. Souza, J. Zhang, C. A. Ryan, and R. Laflamme, "Experimental magic state distillation for fault-tolerant quantum computing," *Nature Comm.*, vol. 2, no. 169, 2011.

[39] T. Jochym-O'Connor, Y. Yu, B. Helou, and R. Laflamme, "The robustness of magic state distillation against errors in clifford gates," *arXiv:1205.6715 (submitted to Quant. Inf. Comput.)*, 2012.

[40] C. A. Ryan, M. Laforest, and R. Laflamme, "Randomized benchmarking of single and multi-qubit control in liquid-state NMR quantum information processing," *New J. Phys.*, vol. 11, p. 013034, 2009.

[41] P. Aliferis, D. Gottesman, and J. Preskill, "Quantum accuracy threshold for concatenated distance–3 codes," *Quant. Inf. Comput.*, vol. 6, pp. 97–165, 2006.

[42] A. Cross, D. P. DiVincenzo, and B. M. Terhal, "A comparative code study for quantum fault tolerance," *Quant. Inf. Comput.*, vol. 9, pp. 541–572, 2009.

[43] M. B. Plenio and S. Virmani, "Upper bounds on fault tolerance thresholds of noisy Clifford-based quantum computers," *New J. Phys.*, vol. 12, p. 033012, 2010.