Degenerate Codes and Capacities of Quantum Channels

by

Tyler B. Jackson

A Thesis
presented to
The University of Guelph

In partial fulfilment of requirements
for the degree of
Doctor of Philosophy
in
Mathematics and Statistics

Guelph, Ontario, Canada

© Tyler B. Jackson, April, 2017

Abstract

Degenerate Codes and Capacities of Quantum Channels

Tyler B. Jackson
University of Guelph, 2017

Advisor:
Dr. Bei Zeng

No quantum system can be perfectly isolated from the environment and, as a result, no physical implementation of quantum information processing tasks can be completely free from noise. The best tool for combating such noise is the use of quantum error correcting codes (QECCs). The general requirements for QECCs have been documented for a while and yet construction of good codes and understanding their effect remains a difficult and active area of study. In this thesis I outline the work I have completed looking into both of these problems.

My general technique throughout this work is to first reduce the problem space as much as possible through use of group theory. Then to use numerical methods and bring computational power to bear on the problem. In chapter 2, I investigate the construction of good codes for the amplitude damping channel, using the codeword stabilized quantum code(CWS) framework. Through an exhaustive search method many new codes with better parameters than previously known are found. In chapter 3, I continue constructing good codes for the amplitude damping channel, this time using code concatenation techniques to find results that would be unfeasible to find via an exhaustive search.

Finally, chapter 4 broaches the difficult problem of determining if a quantum channel has capacity; the ultimate use of QECCs in a sense. Expanding on

the key works done on the problem, I develop the theory surrounding effective noise channels obtained from applying a QECC on multiple uses of a channel in order to determine if capacity exists. Using this framework and making use of computational power available today, these techniques allow us to find many very noisy non-Pauli channels that have positive capacity which previously had not been shown to have capacity.

Acknowledgements

There have been many people in my life, who have helped me get to where I am today: writing my dissertation, completing my doctorate degree. From teachers, to friends and colleagues, all have had a lasting impact in my life have helped shape the professional I have become and I am grateful to everyone of them.

Nevertheless, I would like to pick out a few of the people who have exercised a special influence in my life, taking here the opportunity to thank them for all their contributions to my success. First and foremost, I need to thank my parents. They have been there for me literally from day one, and never once has their support and enthusiasm for my goals wavered. I have no doubt that without them I would not be where I am today.

Second, I need to thank my supervisor Bei Zeng. I honestly could not have wished for a better supervisor. It is not only a joy to work on research problems with Bei, but with her constant encouragement, and untiring patience, she has guided and motivated me throughout my degree. It is my sincere hope to continue working with her on interesting problems even as my path diverges.

Next, I would like to acknowledge the great understanding shown by Casey and Preston, my partners in business. They have been incredibly patient and supportive of my goal of completing my PhD.

Finally, I need to thank Diana. Writing this thesis has been by my estimation the single hardest task I have ever undertaken. There were many times when I felt too tired to continue, or disappointed, when the narrative was not telling the story I wanted it to. During all those times she was there to help me push through. Every so often, the stress would be too much to handle, and Diana not only made sure I was fed, she made sure things kept working around me. Thank you.

Contents

Acknowledgements **iv**

List of Tables **vii**

List of Figures **viii**

1 Preliminaries **1**
1.1 Quantum States 2
1.1.1 Bell States 5
1.2 Operations 6
1.2.1 Pauli Operators 7
1.2.2 The Clifford Group 8
1.3 Measurement 9
1.4 Quantum Channels 9
1.4.1 Complementary Channels 12
1.5 Quantum Error Correcting Codes 13
1.5.1 Stabilizer codes 16

2 Search for CWS Codes Correcting Asymmetric Errors **18**
2.1 Introduction 18
2.2 The CWS Framework 21
2.2.1 The CWS standard form 24
2.2.2 Algorithm for constructing CWS codes 26
2.3 Error Model 28
2.4 The Exhaustive Search 32
2.5 Results 35
2.5.1 Codes correcting a single amplitude damping error 36
2.5.2 Codes correcting two amplitude damping errors 38
2.5.3 Codes detecting a single amplitude damping error and detecting multiple dephasing errors 39
2.6 Discussion 40

3 Concatenated Codes Correcting Asymmetric Errors **42**
3.1 Introduction 42
3.2 Distance of Codes Correcting Asymmetric Channels 45
3.2.1 Effective Distance 48
3.3 Concatenated Codes 52
3.4 Concatenated Codes for Asymmetric Channels 53

3.5 Results 57
3.6 Discussion 59

4 Capacities of Very Noisy Quantum Channels 64
4.1 Introduction 64
4.2 Existing Proof Techniques 69
4.2.1 Upper and Lower Bounds 70
4.2.2 Hashing code 73
4.2.3 Effective channels and code concatenation 75
4.3 Construction of Effective Channels from Codes 79
4.3.1 Using CWS codes to construct Effective Channels 80
4.3.2 Constructing the Effective Channel 83
4.4 Verification 94
4.5 Results 95
4.5.1 2- and 3-cat code on t_z-perturbed near-depolarizing channels 97
4.6 The Effectiveness of Degenerate Codes 99
4.7 Discussion 106

5 Conclusion 109

A Codeword Stabilized Quantum Code Search Algorithms 113

Bibliography 117

List of Tables

2.1 Dimension K of CWS codes $((n, K))$ of length n detecting the error set $\mathcal{E}^{\{1\}}$ for different length n. The column $d = 3$ lists the largest dimension of CWS codes that correct a single error for the depolarizing channel. The column $\mathcal{E}^{\{1\}}$ lists the largest dimension of CWS codes found detecting the error set $\mathcal{E}^{\{1\}}$. The column CSS lists the largest dimension of the known Calderbank-Shor-Steane (CSS) codes [7, 71] that can correct the error set $\mathcal{E}^{\{1\}}$, based on a construction proposed in [26]. The column GF(3) lists the largest dimension of codes correcting a single amplitude damping error based on a construction proposed in [65]. 36

2.2 Dimension K of CWS codes detecting the error set $\mathcal{E}^{\{3\}}$ (as defined in equation 2.21) for different length n and parameter r. For each value of r, the first column lists the largest dimension of stabilizer codes that detect all errors up to weight r as given in [7]; the second column lists the largest dimension of asymmetric codes detecting a single amplitude damping error and phase errors up to weight r as found in [23]; the third column lists the largest dimension of the CWS codes found by our search for codes detecting the error set $\mathcal{E}^{\{3\}}$. '–' means that no non-trivial codes exist based on the construction. The numbers labeled with * are the best parameters found by random search; otherwise the maximal dimension is obtained by exhaustive search. 39

3.1 The total effective weight of pairs of single qubit Pauli operators compared to the effective weight of their product. Notice the result weight column is always less than or equal to the input weight. 50

3.2 Concatenated codes $[\![n, k, d_e]\!]_2$ for the AD channel based on QMDS outer codes with qudit dimension 2, 4, 8, and 16. 61

3.3 Concatenated codes $[\![n, k, d_e]\!]_2$ for the AD channel based on non-QMDS outer qubit codes. 62

3.4 Concatenated codes $[\![n, k, d_e]\!]_2$ for the AD channel based on non-QMDS outer codes with qudit dimension 4 and 8. 63

4.1 Physical operators and their effective counter parts for 2-cat code . . 92

4.2 Shors 1996 results[66] 94

4.3 Capacity bounds on depolarizing channel using effective channels method 95

List of Figures

3.1 Outline of concatenated codes 42

4.1 Diagram showing the gap between channels with zero coherent information and those with known zero capacity (middle ring) 68
4.2 Diagram of effective channel and where it sits in larger error correcting framework. The effective channel is the action of the inner code acting on the physical channel. The outer code will be code achieving coherent information rate as from LSD[48, 64, 18]. 75
4.3 Comparison of capacity boundaries. The region below the "valid channels" line are all channels which satisfy the CPTP condition. The regions above any of the cat code lines are channels which have capacity under concatenation of the random stabilizer code with that code. 98
4.4 Zoomed in version of figure 4.3 to region of particular interest. . . . 99

Chapter 1

Preliminaries

Most of the world we observe is dominated by classical physics and systems that behave in a classical manner. As a result information theory was developed under some assumptions of the kinds of information processing tasks which were possible. However, as quantum mechanics was developed it became increasingly clear that under carefully controlled situations the effects of quantum mechanics and quantum systems could be developed and manipulated allowing for capabilities not originally available to information processing.

Information processing, while having been beautifully abstracted from the physical processes that implemented it by the study of information theory, suddenly has new physical processes which could be utilized including superposition of state, and entanglement. Thus, Quantum Information Processing and Quantum Information Theory were born to investigate the question of what this newly accessible resource could be used to do.

1.1 Quantum States

It is practically impossible to discuss information theory, either classical or quantum, without a notion of "state". In classical computing the fundamental unit of information is the bit, or binary digit. It can be in one of two states typically known as 0 and 1. There is an analogous quantity in quantum computing which is called a "qubit" or quantum bit. Instead of two distinct values, qubits are represented by unit vectors in a two dimensional complex Hilbert space [50].

Quantum mechanics tells us that quantum systems obey certain laws and equations such as Schrodinger's equation [61] or Born's rule [6]. In quantum mechanics "quantum states" are solutions to Schrodinger's partial differential equation. Transitions between states are effected through unitary operations.

However, working with raw wave functions will often obscure some of the fundamental behaviours occurring, much like how focusing on the electrons in a telegraph wire makes it hard to see the information it represents. This is not to say that an understanding of the underlying physical systems is unnecessary but rather that a deeper understanding of the qualities of information processing can often be easier if the physical system is abstracted. To this end we can introduce the idealized quantum state.

Physically a qubit is a quantum system with two basis states (or one that has more than two states but is constrained to only two). Typical examples include spin half particles such as the electron for which measuring its angular momentum yields either an "up" or "down". Another common system is a photon with either vertical or horizontal polarization. These distinct states form an orthonormal basis for the Hilbert space. The correspondence of any state to a particular basis vector

is arbitrary.

While it is important to understand the physical systems that implement a qubit, for the purposes of quantum information theory we deal with the abstraction of the qubit from the system which implements it. For us the states are vectors in the two dimensional complex Hilbert space, with some basic assumptions about behaviour, usually quoted as the axioms of quantum information [50] [57]. This is similar to the classical bit which can be representative of any number of systems; from magnetized strips to high and low voltage in a wire. The abstraction of the qubit from the physical system that implements it allows for consistency and broadly applicable results. The advantage of qubits over bits is captured in a property called superposition. A classical bit can be 0 or 1, but it cannot be some complex linear combination of 0 and 1. At best it can have some probability of being 0 or 1, which constrains bits to a line segment. The qubit, has states that can be described as $|\psi\rangle = a\,|0\rangle + b\,|1\rangle$ where a and b are complex numbers which satisfy $|a|^2 + |b|^2 = 1$ (the unit length constraint). This gives the qubit two complex degrees of freedom with one constraint.

The "standard basis" for this Hilbert space is typically represented as $|0\rangle = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ and $|1\rangle = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$. Here we are using Dirac notation for the states of the qubit/ vectors of the Hilbert space. $|a\rangle$ is "ket" a and is the vector $\overrightarrow{a}$ in the vector space. The Hermitian conjugate of $\overrightarrow{a}$, $(\overrightarrow{a})^\dagger$ is then written in Dirac notation as "bra" a, $\langle a|$. The inner product between two vectors $\overrightarrow{a}$ and $\overrightarrow{b}$ is represented as $\langle a|b\rangle$.

What is described above is a class of quantum states known as pure states. Pure states are quantum states that we have perfect information about. There is also another class of quantum states that represent probabilistic mixtures of pure states; called mixed states. In these cases there is imperfect information about what

pure state we are in. In order to represent mixed states we use density operators or density matrices.

Single qubit density matrices are 2-by-2 Hermitian matrices. The density matrix of a pure state $|\psi\rangle$ can be written as $|\psi\rangle\langle\psi|$ in outer product notation. A mixed state is then any state of the form:

$$\rho = \sum_{i=1}^{n} p_i |\psi_i\rangle\langle\psi_i|$$

where $\sum_{i=1}^{n} p_i = 1, \quad p_i \geq 0$.

A good example of a mixed state is the state produced by a source which outputs the states $|0\rangle$ and $|1\rangle$ with equal probability. The state produced by such a source is inherently different from the pure state that is an equal superposition of $|0\rangle$ and $|1\rangle$, normally written as $|+\rangle = \frac{1}{\sqrt{2}}[|0\rangle + |1\rangle]$.

To confirm that these states are in fact different one only needs to observe their density matrices which for the mixed state is:

$$\rho \quad = \tfrac{1}{2}[|0\rangle\langle 0| + |1\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 1|] = \tfrac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}. \tag{1.1}$$

and for the pure superposition is:

$$\rho \quad = \tfrac{1}{2}[|0\rangle\langle 0| + |1\rangle\langle 1|] = \tfrac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \tag{1.2}$$

In order for an arbitrary Hermitian matrix ρ to be a valid quantum state it must satisfy two properties. First $\mathrm{Tr}(\rho) = 1$; this is the normalization constraint of states. In the case of pure states this is equivalent to the unit length constraint

of state vectors.

$$\mathrm{Tr}(\rho) = \mathrm{Tr}(|\psi\rangle\langle\psi|) = \mathrm{Tr}(\langle\psi|\psi\rangle) = \langle\psi|\psi\rangle = 1 \tag{1.3}$$

The second property ρ must satisfy is being positive semi-definite. This represents the fact that when projecting onto a state, you cannot have negative probability of getting that result. Mathematically it is stated as $\langle x|\,\rho\,|x\rangle \geq 0$ for all vectors $|x\rangle$.

Larger systems, such as multi-qubit systems, can be represented by states in a higher dimensional Hilbert space. For n qubits this is the 2^n dimensional complex vector space. All of the same requirements hold from the single qubit case. Combining two systems together, say of two qubits in states ρ_1 and ρ_2 respectively is done by taking the tensor product of their states, $\rho_{12} = \rho_1 \otimes \rho_2$. This state now exists in the space $\mathbb{C}^2 \otimes \mathbb{C}^2 = \mathbb{C}^4$, though states which are a tensor product of states in subsystems (or a linear combination of them) are called separable states and are only a subset of all possible states in $\mathbb{C}^4$. When a state is not separable, it is called entangled.

1.1.1 Bell States

There is a particular set of states on two qubit systems so useful and pervasive that they warrant a subsection to themselves, the Bell states.

These are maximally entangled pairs of qubits which are named for Bell's theorem [2]. They were key to disproving the "local hidden variables" theory of quantum mechanics which was a proposed solution to the EPR paradox [22]. Bell states also achieve the maximal value of the Bell inequality [12].

Written in the standard basis the four Bell states are:

$$\left|\phi^+\right\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$
$$\left|\psi^+\right\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$$
$$\left|\psi^-\right\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$$
$$\left|\phi^-\right\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)$$

The Bell states also form a useful basis of two qubit Hilbert spaces, one that is used repeatedly throughout this work.

1.2 Operations

In order for a quantum state to be used in computation, it must be possible to apply operations to it. The description of the evolution of a quantum state in a closed system can be represented by a unitary operator.

An operator basis is a set of operators for which any unitary operator can be decomposed into a linear combination, much like basis vectors for a vector space. There are an infinite number of bases for the unitary operators, but the one used most often in this work are the Pauli operators.

1.2.1 Pauli Operators

The single qubit Pauli operators are 2-by-2 Hermitian matrices which form a basis for the space of 2-by-2 Hermitian matrices. In the standard basis they are written as:

$$\sigma_0 = I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \sigma_1 = X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix};$$
$$\sigma_2 = Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \sigma_3 = Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

For n-qubit systems, the n-qubit Pauli operators still form a basis for the space of 2^n-by-2^n Hermitian matrices. The n-qubit Pauli operators are defined as:

$$\mathcal{P} = \bigotimes_{j=1}^{n} s_j$$
$$s_j \in \{I, X, Y, Z\}$$

When there is no confusion in the meaning, the notation XYZ will be used for the operator $X \otimes Y \otimes Z$ for ease of reading. In some cases the notation X_i will be used to indicate the Pauli operator with X in the ith position and I's elsewhere, for example $X_2 = IXII$.

There are two groups of Pauli operators which we are normally concerned with. The first case has elements $a\sigma_j$ for $j \in \{0, 1, 2, 3\}$ and $a \in \{\pm 1, \pm i\}$. Note that $X \cdot Z = -iY$. Sometimes the phases produced are not of interest to us, this is normally the case when they would introduce unmeasurable global phases. In these cases we use the group of the Pauli operators mod phases. In this group $X \cdot Z = Y$.

In most cases this is the Pauli group we refer to, when phases need to be accounted for we make this clear.

Any operator that is not a Pauli operator is referred to as a *Non-Pauli* operator. Since the Paulis form a basis for all operators, any Non-Pauli operator can be represented as a linear combination of the Pauli operators.

1.2.2 The Clifford Group

The Pauli operators are of great interest for many reasons. They are Hermitian, unitary, traceless (orthogonal to the identity under the Hilbert-Schmidt inner product), they form a basis for Hermitian matrices, and in many experimental systems they are easy to measure.

The Clifford group is defined as the set of operators that map the set $\mathcal{P}$ of Paulis to itself under conjugation [27, 39]. That is to say that any element C of the Clifford group, a Clifford operator, is a bijective map $C : \mathcal{P} \to \mathcal{P}$. The Clifford operators are then the subgroup of the unitaries $U(2^n)$ which normalize the Pauli operators. The Clifford group can be generated by, or written as products of some combination of, the Hadamard (H), controlled-not($CNOT$) and phase (S) operators, typically written as:

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, S = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}, CNOT = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tag{1.4}$$

The Clifford group induces an very useful equivalence relation:

Definition 1.1. *Two sets of Pauli operators $P = \{P_i\}$ and $Q = \{Q_i\}$ are said to be*

Clifford equivalent if there exists some Clifford operator C such that $\forall_i \quad \exists_j \quad C^\dagger P_i C = Q_j$.

Clifford and Local Unitary equivalence are very useful in many circumstances as they allow you to either reduce to uniquely behaving case or to expand to all cases that have a similar behaviour.

1.3 Measurement

Whenever you perform computation or transmission of information there will normally come a time when you wish to acquire some result from it. In classical computing this can be done without perturbing the state of interest. In the quantum realm, it is impossible (except in special circumstances) to perform a measurement without disturbing the state. This is because when a measurement is made of a system, the system is no longer closed, and as a result does not evolve unitarily.

Instead measurements in general are described by a series of measurement operators M_m where the index m corresponds to the measurement result obtained by the measurement. If the initial quantum state was ρ then the probability of outcome m being obtained is $p(m) = \mathrm{Tr}(M_m \rho M_m^\dagger)$ and the post-measurement state is $\frac{M_m \rho M_m^\dagger}{p(m)}$ [50].

1.4 Quantum Channels

Quantum channels are the most general description of the evolution of an open quantum system. I have already mentioned unitary transformations which are a subset of all channels. A quantum channel is a completely-positive trace-preserving (CPTP) map between two Hilbert spaces. The complete positivity condition ensures

that if the input ρ is positive, then the output $\Lambda(\rho)$ is also positive, as well as positivity of a state where the channel is applied only to a subsystem.

Quantum channels can represent any number of physical processes, such as computation or communication. Quantum channels are often used to represent noise processes. The operator-sum (Kraus) representation of a channel says any channel Λ can be represented in the following form:

$$\Lambda(\rho) = \sum_i E_i \rho E_i^\dagger$$

with the constraint that $\sum_i E_i^\dagger E_i = I$ which ensures the trace preserving condition. The operators E_i are called the Kraus operators of the channel. When the Kraus operators of a channel are all from the Pauli operators, the channel is called a Pauli channel. In contrast, any channel which does not admit a representation with Pauli Kraus operators is a non-Pauli channel.

Since the Pauli operators form an operator basis, any operator E_i can be expanded as the linear combination of Pauli operators $E_i = \sum_j a_j P_j$.

Another important representation of a qubit quantum channel is as a Choi matrix. An equivalence between the action of a channel and a two qubit state found by acting the channel on one half of a Bell $|\phi^+\rangle$ state. Since the Pauli operators acting on the second qubit of the $|\phi^+\rangle$ state produce the other Bell states, any Pauli channel $\Lambda(\rho) = \sum_i p_i P_i \rho P_i^\dagger$ (p_i is the probability of operator P_i being applied) can be represented as a Choi matrix in the Bell basis as:

$$\begin{bmatrix} p_I & 0 & 0 & 0 \\ 0 & p_X & 0 & 0 \\ 0 & 0 & p_Y & 0 \\ 0 & 0 & 0 & p_Z \end{bmatrix}$$

Non-Pauli channels will have off diagonal-values. The probabilities associated with each sum component (Pauli or non) can be read off the Choi matrix representation when written in the Bell basis. Furthermore the probabilities for a tensor of channels can similarly be read off the tensor products of the Choi matrices. While we don't explicitly show this trick in this thesis, it is one of the more valuable ones for representing and manipulating a channel numerically. Importantly, the *Choi Rank*, the rank of the Choi matrix, tells us the minimum number of Kraus operators needed to represent the channel.

Finally, one of the most important results about channels comes from the characterization of all channels (up to unitary equivalence) as provided in [58]. Here it was shown that up to unitary equivalence, only 6 parameters are needed to specify any qubit channel. In this thesis we can therefore represent a qubit channel as a vector $[p_X, p_Y, p_Z, t_X, t_Y, t_Z]$.

Channels with these parameters have Choi matrix representations (in the Bell Basis) given by:

$$\begin{bmatrix} p_I & t_x & it_y & t_z \\ t_x & p_X & -t_z & -it_y \\ -it_y & -t_z & p_Y & t_x \\ t_z & it_y & t_x & p_Z \end{bmatrix}$$

where $p_I = 1 - p_X - p_Y - p_Z$. Here we see that when the last three components are

0 we have a Pauli channel.

1.4.1 Complementary Channels

There is a broader view of the action of a channel and that is to consider the action of a unitary operation acting on both the system of interest and some "environment". By including the environment the system is no longer open as is the case with the general channel as described before. The action of a channel Λ is given by:

$$\Lambda(\rho) = \mathrm{Tr}_E(U^\dagger \left|\Psi\right\rangle \left\langle\Psi\right|_{AE} U)$$

Here U is some unitary action on the combined AE Hilbert space and $\left|\Psi\right\rangle \left\langle\Psi\right|_{AE}$ is a purification of the state ρ. That is to say that if the environment is traced out the state ρ is re-obtained. This expresses the notion that noisy channels can be viewed as operations which transfer some information out of system A and into the environment.

The complementary channel expresses the information that is transferred to the environment and is described as:

$$\Lambda^c(\rho) = \mathrm{Tr}_A(U^\dagger \left|\Psi\right\rangle \left\langle\Psi\right|_{AE} U)$$

Notice that the main system A is now being traced out. Complementary channels play a key role in describing two important classes of channels.

Definition 1.2. *A degradable channel is a channel Λ for which there exists a "degrading channel" D such that $D \circ \Lambda = \Lambda^c$.*

Intuitively this means that a degradable channel is capable of simulating

its complementary channel. Similarly:

Definition 1.3. *An antidegradable channel is a channel* Λ *such that the complementary channel* Λ^c *has a degrading channel* D *allowing* $D \circ \Lambda^c = \Lambda$.

The intuition is similar here, in this case the information transferred to the environment is enough such that it could simulate the channel Λ.

1.5 Quantum Error Correcting Codes

When noise processes occur, they can have devastating effects on information content and computation. Errors can propagate turning small errors into much larger ones. Systems which experience less noise can sometimes be designed but in other cases the effect of the noise must be made inconsequential. This is where error correcting codes come in.

A common example from classical information theory is the bit flip code where you encode a single bit of information into three bits as $0 \to 000$ and $1 \to 111$. In this (and all) error correction scheme, we have used additional resources in order to create some redundancy in the message. If any single bit of the message were to be corrupted then it would be obvious not only what the error was but how to correct the message. In essence since there is only one bit of information being stored in three bits, the additional bits allow you to obtain information about the error that occurred.

Quantum error correcting codes work in much the same fashion, but are complicated by the fact that performing measurements will disturb (and potentially destroy) states [50] [6], and these arbitrary unknown quantum states can not be cloned [75].

Much as a classical error correcting code is a subset of bitstrings of some length, a QECC is a K dimensional subspace of a larger Hilbert space. I focus on the case where we encode into n qubits and so the total space is $(\mathbb{C}^2)^{\otimes n}$.

This subspace V_0 which can be referred to as the encoding or logical subspace has a projector onto P_0. When an error E occurs on the encoded state they might map V_0 to some other subspace V_E.

If we have a set of errors $\mathcal{E} = \{E_i\}$ that we want to protect our quantum state from we have some conditions for when this is possible. First let $|v_i\rangle$ be a basis for the encoded subspace V_0. There exist well known criteria for when a quantum code Q can detect and correct a set of errors $\mathcal{E}$ [50] which I state here.

Theorem 1.1 (Error Detection Criteria)**.** *A QECC Q is capable of detecting errors from an error-set $\mathcal{E}$ if and only if:*

$$\langle v_i| E |v_j\rangle = c_E \delta_{ij} \quad \forall E \in \mathcal{E}$$

Where c_E is some complex value dependent upon only E.

Theorem 1.2 (Error Correction Criteria)**.** *A QECC Q is capable of correcting errors from and error set $\mathcal{E}$ if and only if:*

$$\langle v_i| E_k^\dagger E_l |v_j\rangle = c_{k,l} \delta_{ij} \quad \forall E_k, E_l \in \mathcal{E}$$

Note that if a code Q detects all errors from an error-set $\mathcal{E}^\dagger \times \mathcal{E}$ then Q also corrects $\mathcal{E}$. The δ_{ij} components of these equations are to ensure that two basis states are not confused after an error. Errors which would produce this effect are called *Logical Errors*. These are useful in their own right as they can be used to

perform operations an a quantum state while still encoded.

When $i = j$ there are two options, either $E|v_i\rangle$ is orthogonal to $|v_i\rangle$ for all i in which case $c_E = 0$ or they are equal for all i. In the second case, the code is degenerate to the error E, that is to say degenerate errors have no effect on any state in V_0. Degenerate errors are an interesting phenomena which can be found in QECCs but not in classical error correcting codes. The fact that degenerate quantum codes can be immune to some errors suggests that performance better than that achieved by non-degenerate codes.

Quantum codes that use n qubits to encode a K dimensional space are commonly cited as $((n, K))$ codes. When K is a power of 2, $K = 2^k$, these codes can also be cited as $[[n, k]]$ codes. When, instead of qubits, qudits of dimension q are used these will be cited as $[[n, k]]_q$ codes.

The distance is usually defined in connection to the noise associated with the depolarizing channel, and acts as a metric to understand the error correcting capabilities of the code. In order to understand the distance we must first define the weight of an operator.

Definition 1.4. *The weight of an operator is the number of qubits which it acts non-trivially on. For example the operator $IXZZ$ is of weight 3.*

The distance of a code is then defined as follows:

Definition 1.5. *The distance of a code d is the lowest weight of an error for which the error detection criteria fails. The weight of the error is the number of non-identity operators in its product (assuming Pauli errors). A distance d code is capable of correcting $\lfloor \frac{d-1}{2} \rfloor$ arbitrary errors.*

The distance of any code can be equivalently thought of as the minimum

weight of a logical operator on the code, since the error detection criteria fails for any logical operator. I discuss the distance and how it can be adapted to a noise model other than the depolarizing channel further in chapter 3. Since the distance conveys useful information of the detectable error set of a code it will often be included in code notation as $[[n, k, d]]_q$ codes. This indicates a distance d code using n qudits of dimension q protecting a 2^k subspace from noise.

1.5.1 Stabilizer codes

One common formalism for quantum error correcting codes are the stabilizer codes [26]. The stabilizer formalism relies on stabilizer groups and stabilized subspaces which are defined as follows:

Definition 1.6. *A stabilizer S is a subgroup of the n-qubit Pauli operators S, not including $-I$ such that all members of S mutually commute. A vector space V_S is said to be stabilized by S if for any $|v\rangle \in V_S$ and $s \in S$ you have $s\,|v\rangle = |v\rangle$. In other words, V_S is the subspace of eigenvectors of S with eigenvalue $+1$.*

A stabilizer code is a quantum error correcting code where the $K = 2^k$ dimensional code subspace V_0 is a stabilizer subspace V_S for some stabilizer S. When S has 2^{n-k} mutually commuting operators, the dimension of V_S is 2^k.

When an error E occurs which anti-commutes with some element of S, V_S is mapped to an orthogonal subspace (and is therefore detectable). If the error is an element of S, then there is non effect on the code space V_S, as every state in V_S is a $+1$ eigenvector of S. If the error commutes with all elements of S but is not an element of S, that is E is an element of $Z(S) - S$, the center of S minus S itself, then the error maps $V_S \to V_S$ but the effect is non trivial. Operators which map the code subspace to itself in a non-trivial way are called *Logical Operators* and

represent encoded operations acting on the protected state.

The error detection criteria for stabilizer codes is then:

Theorem 1.3. *A stabilizer code with stabilizer S can detect an error-set $\mathcal{E}$ if and only if*

$$\forall E \in \mathcal{E} \quad E \notin Z(S) - S \tag{1.5}$$

Chapter 2

Search for CWS Codes Correcting Asymmetric Errors

2.1 Introduction

In order to achieve quantum behaviour, quantum systems must be isolated from the environment. If they are not sufficiently protected, the state of the system will become completely destroyed in a process known as decoherence. As our ability to manipulate systems at the quantum level improve, we are able to increasingly protect against this decoherence of state. Unfortunately no physical system is ever truly isolated from its environment. Small amounts of heat, stray radio waves or other imperfections in experimental setup can all corrupt the delicate quantum states.

This is not a new problem for computing, classical computers had to face much the same problems throughout their development. In order for computation to be feasible, the effect of any error introduced must be able to be made small within some bound. There are two primary approaches to accomplishing this. First

the technology can improve, i.e. the error rate of RAM modules today is so small that most non-enterprise grade RAM doesn't even include any mechanism for compensating for it. The probability that a bit in RAM suffers an error is so small that its likely to never happen in the life of the average machine.

The second method is to use error correcting/detecting codes to protect data against probably noise that might occur. Some examples of this include using checksums on important files to ensure that they have not been corrupted, or CDs and DVDs which use error correcting codes to protect against noise inducing situations such as scratches on the disk [13].

These two methods of protecting the information content stored in physical systems applies to the quantum regime as well. Producing more reliable quantum systems is a task better left to engineers and experimentalists, so this chapter and the next focus on producing quantum error correcting codes.

The field of quantum error correcting codes is very well studied and there exist frameworks for understanding and constructing quantum codes. One of these is the stabilizer formalism [7, 26, 8, 71] , though many other interesting codes have been developed which don't fall into the stabilizer formalism.

However most of the QECC constructed so far are designed with the assumption that the Pauli errors occur with equal likelihood. Essentially, most codes have been designed for the depolarizing channel :

$$\mathcal{E}_{\mathrm{DP}}(\rho) = (1-p)\rho + \frac{p}{3}(X\rho X + Y\rho Y + Z\rho Z), \tag{2.1}$$

where the Pauli X, Y, Z errors happen equally likely. Codes designed for this channel will correct or detect some number of arbitrary errors on the transmitted data.

While this case is interesting, quite often physical systems will have an asymmetry in the types and frequencies of errors. In these cases, it has been shown that codes which are tailored for the probable error sets can achieve better parameters than those which correct a wider set of errors than necessary[11, 46, 45, 25, 59, 60, 65, 21, 23]. Here by better parameters I am specifically referring to the dimension K of the encoded subspace for some number of physical qubits n.

One common asymmetric channel of interest in physical systems is the amplitude damping channel which describes situations such as spontaneous emission. There have been codes constructed for the amplitude damping channel [65, 25, 21] but there has not been a systematic study or search for codes correcting amplitude damping errors.

Codeword stabilized (CWS) quantum codes constitute the most general systematic framework for constructing quantum error-correcting codes (QECC) [14, 9, 10]. It encompasses all stabilizer codes [7, 26, 8, 71], as well as many non-additive codes with good parameters [55, 77, 45]. In recent years, it has been applied to various problems and is often able to produce very strong, promising results [30, 31, 49, 37, 33, 47, 1, 72, 40, 41].

The work of this chapter is in applying the CWS code construction techniques to the problem of finding quantum codes with good parameters which correct asymmetric noise models. In particular, we focus on the Pauli error models corresponding to single and multi-qubit amplitude, and phase dampening noise channels. Through use of local Clifford operations on graph states, we are able to exhaustively search for all possible codes up to length 9. With a similar method, we also look at codes for the Pauli error model that detect a single amplitude error and detect multiple phase damping errors. Many new codes with good parameters are found,

including non-additive codes and degenerate codes.

2.2 The CWS Framework

A QECC Q is a subspace of the space of n qubits $(\mathbb{C}^2)^{\otimes n}$ (here we focus on quantum systems of dimension $q = 2$, but the approach can be generalized to qudits of dimension $q > 2$). For a K-dimensional code space spanned by the orthonormal basis $\{|\psi_i\rangle : i = 1, \ldots, K\}$ and an error set $\mathcal{E}$, there is a physical operation detecting all the elements $E_\mu \in \mathcal{E}$ (as well as their linear combinations) if the error detection condition [4, 43]

$$\langle\psi_i| E_\mu |\psi_j\rangle = c_\mu \delta_{ij}, \quad c_\mu \in \mathbb{C}, \tag{2.2}$$

is satisfied. The notation $((n, K))$ is used to denote a QECC with length n and dimension K.

Definition 2.7. *An $((n, K))$ CWS code Q is described by two objects:*

1) *A stabilizer S that is an Abelian subgroup of the n-qubit Pauli group, has order 2^n, and does not contain $-I$; the group S is called the word stabilizer.*

2) *A set of K n-qubit Pauli operators $W = \{w_\ell \colon \ell = 1, \ldots, K\}$, which are called the word operators.*

Since S is of order 2^n, it stabilizes a unique quantum state $|S\rangle$, i.e., $s|S\rangle = |S\rangle$ for all $s \in S$. The code subspace V_0 is the span of the basis vectors given by $|w_\ell\rangle = w_\ell |S\rangle$.

It was shown that without a loss of generality w_1 can be chosen to be the identity I [14]. When w_1 is chosen in this manner the other $K - 1$ word operators also correspond to some of the logical operators on the encoded state.

According to Eq. (2.2), the code Q detects the error set $\mathcal{E}$ if and only if $\langle w_i| E |w_j\rangle = c_E \delta_{ij}$ for all $E \in \mathcal{E}$. When $\mathcal{E}$ consists of Pauli matrices, this error-detecting condition can be written in terms of S and w_i as below [14]:

For all $E \in \mathcal{E}$,

$$\forall i \neq j \colon w_i^\dagger E w_j \notin \pm S \tag{2.3}$$

and

$$\begin{aligned} &\left(\forall i \colon w_i^\dagger E w_i \notin \pm S\right) && \text{or} \\ &\left(\forall i \colon w_i^\dagger E w_i \in S\right) && \text{or} \\ &\left(\forall i \colon w_i^\dagger E w_i \in -S\right). \end{aligned}$$

Eq. (2.3) follows from if $i \neq j$ then $\langle S| w_i^\dagger E w_j |S\rangle = 0$. This situation only occurs when $w_i^\dagger E w_j |S\rangle$ is orthogonal to $|S\rangle$. Since $|S\rangle$ is a stabilizer state we can say that this will occur anytime $w_i E w_j \notin \pm S$.

Another way to look at this however is to say that E is not a logical error of the code as $w_i^\dagger E w_j \in S$ only when E rotates w_j to be the logical error w_i (or $s w_i$ for some $s \in S$). Errors which rotate basis states of the code to other basis states are exactly logical errors on the code space.

The second set of equations are the case for when $i = j$. In this case we have $\forall_i \langle S| w_i^\dagger E w_i |S\rangle = c_E$. Note that this constant is independent of the basis vector i. Equation (2.4) is the sub case of this when $c_E = 0$ which, like the previous case this means that $w_i^\dagger E w_i \notin \pm S$. Since condition (2.3) must also be satisfied this means that the error E must map the encoding subspace to some orthogonal

subspace.

This follows because all logical errors, which would map from the encoding subspace to itself, are ruled out by condition (2.3). Elements of S are also ruled out in this case as we are assuming the inner product is 0.

If we assume c_E is non-zero we have a more interesting case. Since c_E is constant for all i the only case that satisfies this is $w_i^\dagger E w_i \in S$ or $w_i^\dagger E w_i \in -S$ which can be shown to mean that $E \in \pm S$ since all w_i will either commute or anti-commute with any element of S. These are therefore degenerate errors which do not affect the code spaces.

If condition (2.4) holds for all $E \in \mathcal{E}$ different from identity, then the code Q is non-degenerate, otherwise it is degenerate.

For error correcting criteria we have [50]:

$$\langle c_i | E_l^\dagger E_k | c_j \rangle = C_{lk} \delta_{ij} \tag{2.4}$$

Rather than working through this more complicated statement, it is sufficient for our purposes to note that if in order to show that a code corrects and error-set $\mathcal{E} = \{E_i\}$ it is sufficient to show that the code will detect the error-set $\mathcal{E}' = \{E_i^\dagger E_j | E_i, E_j \in \mathcal{E}\}$. That is, if a code is capable of detecting all pairs of errors from an error set then it is capable of correcting that error set.

2.2.1 The CWS standard form

Every $((n, K))$ CWS code can be transformed, by local Clifford operations, into a standard form [14], where the word operators take the form $w_\ell = Z^{\mathbf{c}_\ell}$ and the word stabilizer is a graph state stabilizer with generators of the form $S_i = X_i Z^{\mathbf{r}_i}$, for some choices of classical n-bit strings $\mathbf{c}_\ell$ and $\mathbf{r}_i$. Here $Z^{\mathbf{c}_\ell} = Z^{c_{\ell,1}} \otimes \ldots \otimes Z^{c_{\ell,n}}$ and the r_is are the rows of the adjacency of some n-node graph $\mathcal{G}$.

In the standard form, any n-qubit Pauli error, which can be written in the form $E = \pm Z^{\mathbf{v}} X^{\mathbf{u}}$ for some classical n-bit strings $\mathbf{v}$ and $\mathbf{u}$ which represent the positions of Pauli Z and X errors, can be translated to classical errors via the map

$$\mathrm{Cl}_S(E = \pm Z^{\mathbf{v}} X^{\mathbf{u}}) = \mathbf{v} \oplus \bigoplus_{i=1}^{n} (\mathbf{u})_i \mathbf{r}_i. \tag{2.5}$$

In standard form the word operators are of the form $\{Z^{\mathbf{c}_\ell} : \mathbf{c}_\ell \in \mathcal{C}\}$. These word operators must still satisfy the error detection conditions for the the error-set $\mathcal{E}$. It was shown [14] that the error detection conditions for a CWS code in standard form are the following:

Theorem 2.4 (Error Detection Criteria for CWS codes in standard form)**.** *A CWS code in standard form specified by a graph state stabilizer $\mathcal{G}$ and classical code $\mathcal{C}$ will detect errors from an error set $\mathcal{E}$ if and only if:*

The classical binary code $\mathcal{C}$ detects all errors from $Cl_S(\mathcal{E})$, and that for each $E \in \mathcal{E}$

$$Cl_S(E) \neq \mathbf{0} \tag{2.6}$$

$$\textit{or} \qquad \forall \ell \colon Z^{\mathbf{c}_\ell} E = E Z^{\mathbf{c}_\ell}. \tag{2.7}$$

If Eq. (2.6) *holds for all* $E \in \mathcal{E}$*, the CWS code is non-degenerate, otherwise it is degenerate.*

Proof. To see that these conditions follow from the earlier conditions start with the condition that $\forall i, j w_i \dagger E w_j \notin \pm S$, which was the requiremented that basis states not be confused after an error. In standard form, the errors $E = Z^v X^u$ are equivalent (up to multiplication by elements of S) to the error of the form $E' = Z^{Cl_S(E)}$.

Substituting in we have, $\forall i \neq j$, $Z^{\mathbf{c}_i} Z^{Cl_S(E)} Z^{\mathbf{c}_j} \notin \pm S$. We can do this since we do not care about the global phase factor in this case. The only element of S which consists of only Z operators is the identity which corresponds to the bitstring $\mathbf{c}_i = 0^{\otimes n}$. As such, this condition simplifies to $\forall i, j \quad \mathbf{c}_i \oplus Cl_S(E) \oplus \mathbf{c}_j \neq 0$ which is the error detection condition for classical binary codes for detecting errors from $Cl_S(\mathcal{E})$.

In addition to this condition, the error detection criteria for CWS codes (2.3) required one of the following had to also be satisfied:

$$\forall i w_i \dagger E w_i \notin \pm S \qquad \text{or} \tag{2.8}$$

$$\forall i w_i \dagger E w_i \in S \qquad \text{or} \tag{2.9}$$

$$\forall i w_i \dagger E w_i \in -S \tag{2.10}$$

The first sub case is trivial since, by using the same logic as above, it follows that:

$$\forall i \quad Z^{c_i} Z^{Cl_S(E)} Z^{c_i} \neq Z^0 \tag{2.11}$$

$$\forall i \quad \mathbf{c}_i \oplus Cl_S(E) \oplus \mathbf{c}_i \neq 0 \tag{2.12}$$

$$Cl_S(E) \neq 0 \tag{2.13}$$

The second and third cases require us to consider the global phase. However, note that if $Z^{\mathbf{c}_i}$ anti-commutes with E for some i then we will have a contradiction since I is one of the codewords and will commute with any E we get the conflicting statements that E is both an element of S and an element of $-S$. Therefore these conditions are equivalent to requiring that $Z^{\mathbf{c}_i} E = E Z^{\mathbf{c}_i}$ for all i.

□

2.2.2 Algorithm for constructing CWS codes

Standard form is interesting because, for Pauli errors, the code conditions become essentially classical in nature. You can specify a code completely with a graph $\mathcal{G}$ and a classical code $\mathcal{C}$. You can also construct codes by choosing the graph $\mathcal{G}$, specifying the error-set $\mathcal{E}$ which you want the code to correct and using the error detection criteria (2.6) to search for suitable classical codes $\mathcal{C}$. This process first detailed in [14] will be outlined here. Pseudocode for these algorithms can be found in Appendix A.

The procedure takes as input a graph $\mathcal{G}$, along with an error set $\mathcal{E}$ to be corrected by the finished code. Using the error detection criteria for standard form CWS codes, one can construct a search for classical codes C that will satisfy them. In this way, we can maximize the dimension of the encoding space while ensuring

the error-set is still detectable.

First we take the graph G and error set $\mathcal{E}$ and construct the set $Cl_S(\mathcal{E})$, the classical errors that $\mathcal{C}$ must detect. During this process we also note any errors for which $Cl_S(E) = 0$. If this is the case, then any classical codeword $\mathbf{c}$ for which $Z^{\mathbf{c}}$ does not commute with E must be excluded from the possible codewords of C. This set of excluded codewords we call D.

After constructing these sets, the CWS clique graph G_C can be constructed. In this graph the vertices are all binary strings of length n which are potential codewords in the classical code C.

The algorithm loops through all pairs of vertices $\mathbf{v}_i$, $\mathbf{v}_j$ such that neither $\mathbf{v}_i$ and $\mathbf{v}_j$ are excluded from candidacy by being in D, and checks that for all classical errors $Cl_S(E)$ $\mathbf{v}_i \oplus Cl_S(E) \neq \mathbf{v}_j$. If satisfied, an edge is added between $\mathbf{v}_i$ and $\mathbf{v}_j$. In short, there exists an edge between any two vertices in G_C if a code including both vertices satisfy the error detection condition for a CWS code in standard form.

Definition 2.8. *A clique of a graph G is a subset of vertices such that every pair of vertices is connected by an edge in the graph G.*

At this point, the CWS clique graph G_C contains all information about valid classical codes that could be used such that the resultant CWS quantum code will detect the error set $\mathcal{E}$.

Two codewords in this graph can be in a code $\mathcal{C}$ together only if they are connected. In fact, any clique of this graph will be a valid code $\mathcal{C}$ that could be used to construct a CWS code Q. Such a code will have $((n, |\mathcal{C}|))$ where $|\mathcal{C}|$ is the number of codewords in the code $\mathcal{C}$. Since the number of codewords in the classical code directly determine the dimension of the encoded space of the CWS code, maximizing

the size of the clique is ideal. This is a classic graph theory problem known as the max clique problem which is a computationally hard problem, being both fixed parameter intractable [20] and hard to approximate[24]. In order to solve the max clique problem we used the algorithms and programs developed in [76].

2.3 Error Model

The most general quantum channels allowed by quantum mechanics are completely positive, trace-preserving linear maps that can be represented in the Kraus decomposition $\mathcal{E}(\rho) = \sum_k E_k \rho E_k^\dagger$ with $\sum_k E_k^\dagger E_k = I$ (see Sec:1.4).

One example generalizing the depolarizing channel $\mathcal{E}_{\mathrm{DP}}$ are the asymmetric Pauli channels which sends ρ to

$$(1 - p_x - p_y - p_z)\rho + p_x X\rho X + p_y Y\rho Y + p_z Z\rho Z, \tag{2.14}$$

where the Pauli X, Y, Z errors happen with probabilities p_x, p_y, p_z, respectively [38].

Depending on the noise model in different physical systems, we obtain different asymmetric quantum channels. We start with the amplitude damping channel $\mathcal{E}_{\mathrm{AD}}$ which models real physical processes such as spontaneous emission. The amplitude damping channel can be expressed in Kraus operator form as:

$$\mathcal{E}_{\mathrm{AD}}(\rho) = A_0 \rho A_0^\dagger + A_1 \rho A_1^\dagger, \tag{2.15}$$

where

$$A_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix}, \quad A_1 = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix}, \tag{2.16}$$

for some damping parameter γ.

When the system is at finite temperature, the noise model will not only contain the Kraus operator A_1 corresponding to emission, but also $A_1^\dagger$ corresponding to absorption [50]. Notice that

$$A_1 = \frac{\sqrt{\gamma}}{2}(X + iY), \quad A_1^\dagger = \frac{\sqrt{\gamma}}{2}(X - iY). \tag{2.17}$$

Hence, the linear span of the operators A_1 and $A_1^\dagger$ equals the linear span of X and Y. We can then equivalently formulate the error model by using the Pauli operators X and Y, which happen with equal probability. That is, if a code is capable of correcting t X- and t Y-errors, it can also correct t A_1- and t $A_1^\dagger$-errors.

Furthermore, notice that

$$A_0 = I - \frac{\gamma}{4}(I - Z) + O(\gamma^2), \tag{2.18}$$

while A_1 depends linearly on $\sqrt{\gamma}$. This then results in an asymmetry between the probabilities $p_x = p_y$ and p_z that the Pauli X, Y or the Pauli Z errors, happen respectively.

Apart from amplitude damping, another common noise in physical systems is dephasing, with Kraus operators given by $\sqrt{1-p}I$ and $\sqrt{p}(I \pm Z)/2$, or equivalently, in terms of I and Z with $p_z > 0$ and $p_x = p_z = 0$ [50]. In general, the system undergoes both amplitude damping and dephasing, resulting in a wide range for the parameters $p_x = p_y$ and p_z.

Therefore, in this work we consider the following asymmetric Pauli channel

$$\mathcal{E}_{\mathrm{AS}}(\rho) = (1 - (2p_{xy} + p_z))\rho + p_{xy}(X\rho X + Y\rho Y) + p_z Z\rho Z$$

where X and Y happen with equal probability $p_x = p_y = p_{xy}$.

In terms of Eq. (2.19), the asymmetric Pauli error model corresponding to amplitude damping is given by $p_{xy} \propto \gamma$ and $p_z \propto \gamma^2$. This is different from the amplitude damping error model in, e.g., [46, 45, 21, 65, 34], where the Kraus operators A_0 and A_1 are used. The main reason that we use the Pauli Kraus operators as our error sets is that this enables us to use the CWS framework to construct codes. Within the CWS framework, in order to transform the quantum error detection condition into a classical condition, it is more convenient to use Pauli errors, as was discussed in Sec. 2.2.2. In other words, since A_0 and A_1 are not Pauli operators, the CWS framework in standard form does not directly apply.

However, making codes for asymmetric Pauli error probabilities is still very fruitful. As a side effect, due to the structure of Eq. (2.17) and Eq. (2.18), our error model does not only correct the errors A_0 and A_1, but the resulting codes will be stronger in the sense that $A_1^\dagger$ can be corrected as well. Codes which do not also correct spontaneous absorption errors could (and in some cases do) have larger encoded subspaces K then we achieve. We will show that the results we obtain using this choice of asymmetric error model will outperform codes designed for the depolarizing channel in most cases making them still interesting and with physical grounding.

In this chapter, I consider three specific cases for asymmetric codes, as listed below. We use X_i, Y_i, Z_i to denote the Pauli X, Y, Z operators on the ith qubit. Notice that our method for generating the error sets is very general and can be straightforwardly generalized to deal with different relations between p_x, p_y, and p_z.

1. Codes correcting a single amplitude damping error: to improve the fidelity of

the transmitted state from $1-\gamma$ to $1-\gamma^2$, one only needs to correct a single A_1 error and detect a single A_0 error [34]. In terms of Pauli operators, the corresponding error set is given by

$$\mathcal{E}^{\{1\}} = \{I\} \cup \{X_i, Y_i, Z_i,\ X_iX_j, Y_iY_j, X_iY_j\}, \tag{2.19}$$

where $i, j = 1, \ldots, n$. Any code which detects this error-set will be able to correct a single X and Y error as well as detect a Z error. As such it can correct a single A_1 and $A_1^\dagger$ error by the linear span argument.

2. Codes correcting two amplitude damping errors: based on the analysis on the single error case above, the error set is given by

$$\mathcal{E}^{\{2\}} = \{E_\mu E_\nu \colon E_\mu, E_\nu \in \mathcal{E}^{\{1\}}\}. \tag{2.20}$$

A code that detects this error set in fact also corrects two $A_1^\dagger$ errors. One major benefit of using the Pauli error-set equivalents is the fact that the construction of the error-sets for multiple amplitude damping error is greatly simplified.

3. Codes detecting both a single amplitude damping error and multiple dephasing errors: detecting $\{X_i, Y_i, Z_i\}$ suffices to detect an arbitrary single qubit error (including a single amplitude damping error), and detecting all Z-errors up to weight r will allow to correct $\lfloor r/2 \rfloor$ Z-errors. The error set is

$$\mathcal{E}^{\{3\}} = \{I\} \cup \{X_i, Y_i \colon i = 1, \ldots, n\} \cup \mathcal{Z}_r, \tag{2.21}$$

where $\mathcal{Z}_r$ is the set of all Pauli Z operators up to weight r. A code that detects

this error set in fact detects both an arbitrary error and r phase errors.

2.4 The Exhaustive Search

Our goal is to find good codes detecting the error sets $\mathcal{E}^{\{j\}}$, for each of the three cases. For each code length n, we seek the largest dimension K of CWS codes for each error set $\mathcal{E}^{\{j\}}$, $j = 1, 2, 3$. This is done through a maximum clique search [10], by using the algorithms and programs developed in [76].

Furthermore we desire this search be exhaustive in the sense that the parameters we obtain are the best that any CWS code could produce for these error-sets. Without simplification this would require using every graph G with n nodes as an input. However It is possible to reduce the size of the problem somewhat through use of local equivalence. In fact we only need to use equivalence up to local Clifford operations as the error sets, the word operators and the word stabilizers are all Pauli.

To get to the standard form, one needs to apply local Clifford (LC) operations of the form $L = \bigotimes_i L_i$, where L_i are single-qubit Clifford operations [14]. This transforms the stabilizer S and word operators $\{w_\ell\}$ to the standard form, but at the same time also changes the error model.

For the depolarizing channel given in Eq. (2.14), the error set is invariant under LC operations, since in this model essentially all single-qubit errors happen equally likely. Another way to note this invariance is the symmetry in the elements of the error-set under any permutation of X, Y, and Z operators. Therefore, in order to search for a CWS code, one can simply use the standard form by starting from a stabilizer of the form $S_i = X_i Z^{\mathbf{r}_i}$, which corresponds to a graph state [36].

Furthermore, for a fixed length n, use of every possible graph state is not necessary. It is sufficient to consider all graph states up to LC equivalence as classified in [17]. This results in an exhaustive search for all possible CWS codes of length n.

Being able to restrict the search to graph states up to LC equivalence, instead of all stabilizer states of length n, has dramatically reduced the search space, and an exhaustive search for single-error-correcting codes for the depolarizing channel up to length $n = 10$ has been carried out. It turned out that the best CWS code with length $n = 9$ has dimension $K = 12$, beating the best stabilizer code of dimension $2^3 = 8$ [77]; for $n = 10$ the best CWS code has dimension $K = 24$, again beating the best stabilizer code of dimension $2^4 = 16$ [37].

For the asymmetric channels as given in Eq. (2.19), however, only considering all graphs states as classified in [17] and the error sets $\mathcal{E}^{\{j\}}$ $(j = 1, 2, 3)$ is not sufficient to exhaustively search for all possible CWS codes. This is due to the asymmetry between p_{xy} and p_z, which implies that the error sets are no longer invariant under LC operations. Therefore, in order to exhaustively search for all possible CWS codes by using the standard form, one will need to check all the possible error sets that are LC equivalent to a given $\mathcal{E}^{\{j\}}$.

Recall that the single-qubit Clifford group is generated by the Hadamard operator H and the phase operator P as given below [7, 26]

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}. \tag{2.22}$$

Since overall phase factors can be ignored, we only need to consider the action of the Clifford group on the Pauli matrices X, Y, Z modulo phase factors. The

Clifford group acts as the permutation group S_3 on three letters (we use $1, 2, 3$ to denote X, Y, Z, respectively). The group S_3 has order six, with the elements given by (in cycle notation) $\{\mathrm{id}, (123), (132), (12), (13), (23)\}$, where id denotes the identity permutation. All error sets $\mathcal{E}^{\{j\}}$, $j = 1, 2, 3$, are invariant with respect to interchanging X and Y. Hence it is sufficient to consider one representative from each of the three right cosets of (12), given by $\{\mathrm{id}, (12)\}$, $\{(13), (132)\}$, and $\{(23), (123)\}$. So effectively, we only need to test, e.g., the three permutations $\{\mathrm{id}, (13), (23)\}$.

Therefore, for each of the error sets $\mathcal{E}^{\{j\}}$, we have three cases for each qubit i: no permutation, permute Y and Z, or permute X and Z. To search for a length n code, this will reduce the total number of error sets from 6^n to 3^n for each graph state to be tested. Compared to codes for the depolarizing channel, the search space is enlarged by a factor of 3^n, due to the asymmetry between p_{xy} and p_z. Nevertheless, we can still handle the search for small n, in particular for the error sets $\mathcal{E}^{\{j\}}$, $j = 1, 2, 3$, up to length $n = 9$.

From here performing the exhaustive search is relatively straight forward. First take your error-set $\mathcal{E}^{\{j\}}$ and create the set of all LC equivalent error-sets $\mathcal{E}_{LC}$. This set will have cardinality 3^n as described before.

Using the set of all graphs $\mathcal{G}$ up to local Clifford equivalence obtained from [17], which for each n will have a different cardinality, and the set of all LC error-sets, an exhaustive search constitutes taking each pairing of elements in $\mathcal{G} \times \mathcal{E}_{LC}$. That is every pairing of a graph state and an error-set. The pseudocode for this process can be found in Appendix A.

This problem does grow rapidly in the number of qubits n for the code due to multiple factors. Amoung these we have to contend with the rapid growth in the

number of distinct graph states increase as well as the number of LC equivalent error-sets which grows exponentially in n. This is further compounded by the fact that for each pairing the CWS construction algorithm requires a max clique algorithm to be run on a graph with up to 2^n nodes. However it does have one saving grace which makes this problem feasible for brute force computing which is that each pairing of graph and error-set can be run in parallel and the results gathered after the fact.

Our search was run primarily on a 12 Core desktop server utilizing all of its resources for multiple months, however with enough resources such as a super computing cluster this time could be greatly reduced.

An interesting trade off which occurs, is that as the number of errors in the error-set grows, the more restrictions that are placed on the CWS clique graph and the average degree of the graph decreases. This makes the max clique algorithm run faster. What this means is that for the same n a problem such as searching for codes correcting two amplitude damping errors runs much faster than the same search for those correcting one amplitude damping error.

2.5 Results

As described above, in our search algorithm, we start from the CWS standard form and transform the error set $\mathcal{E}^{\{j\}}$ by LC operations. This has no effect on the code parameters $((n, K))$ found. However, to present the CWS codes found, we fix the error set $\mathcal{E}^{\{j\}}$ and equivalently transform the CWS standard form into a general CWS code.

2.5.1 Codes correcting a single amplitude damping error

We have conducted an exhaustive search for the error set $\mathcal{E}^{\{1\}}$ up to length $n = 9$, resulting in CWS codes correcting a single amplitude damping error. As already mentioned, the codes found not only correct a single error given by the Kraus operator A_1, at the same time they also correct a single error $A_1^\dagger$. In other words, the codes correct both single X errors and single Y errors (and detect single Z errors as well). We summarize our results in Table 2.1.

n	$d = 3$ [7, 77]	$\mathcal{E}^{\{1\}}$	CSS [26]	GF(3) [65]
5	2	2	2	2
6	2	4	2	5
7	2	8	8	8
8	8	10	8	16
9	12	20	16	24

Table 2.1: Dimension K of CWS codes $((n, K))$ of length n detecting the error set $\mathcal{E}^{\{1\}}$ for different length n. The column $d = 3$ lists the largest dimension of CWS codes that correct a single error for the depolarizing channel. The column $\mathcal{E}^{\{1\}}$ lists the largest dimension of CWS codes found detecting the error set $\mathcal{E}^{\{1\}}$. The column CSS lists the largest dimension of the known Calderbank-Shor-Steane (CSS) codes [7, 71] that can correct the error set $\mathcal{E}^{\{1\}}$, based on a construction proposed in [26]. The column GF(3) lists the largest dimension of codes correcting a single amplitude damping error based on a construction proposed in [65].

As we can see from the table, for the lengths $6, 7, 8, 9$, our codes outperform the best single-error-correcting codes for the depolarizing channel—which also correct the error set $\mathcal{E}^{\{1\}}$, i.e., a single amplitude damping error. In particular, for lengths 8 and 9, the best CWS codes we have found (of dimensions 10 and 20 respectively) are non-additive codes.

For lengths $6, 8, 9$, our codes also outperform the best known CSS codes that are specifically designed to detect the error set $\mathcal{E}^{\{1\}}$, based on a construction proposed in [26] (see also [65]). Therefore, for these lengths, we have found good codes that outperform all the previously known constructions for detecting the error

set $\mathcal{E}^{\{1\}}$.

Notice that the existence of a CWS code with dimension $K = 4$, and hence a subcode of dimension 3, implies the existence of a stabilizer code with the same parameters [10, Theorem 7]. Hence the $((6,4))$ codes we found, as listed in Table 2.1, include stabilizer codes encoding two qubits. As an example, one such code has stabilizer S generated by

$$\begin{array}{cccccc} X & X & I & I & Z & Z \\ X & Z & I & Z & I & X \\ Z & I & Y & Z & Y & Z \\ I & I & Z & X & I & Z \end{array}$$

It is straightforward to check that this code detects the error set $\mathcal{E}^{\{1\}}$, since no elements in $\mathcal{E}^{\{1\}}$ are also in $C(S) \setminus S$, where $C(S)$ is the centralizer of the stabilizer S.

However, with the exception of $n = 7$, the single-error-correcting codes constructed in [65] have larger dimensions than our codes. The codes constructed in [65] are specifically designed to correct the Kraus operators A_0 and A_1, these codes cannot detect the error set $\mathcal{E}^{\{1\}}$. As detection of the errors $\mathcal{E}^{\{1\}}$ implies that a single error $A_1^\dagger$ can be corrected as well. As such it is not a surprise that our codes have smaller dimensions.

Notice that the codes constructed in [65] are also CWS codes, but errors are handled in a different way than the Pauli error set $\mathcal{E}^{\{1\}}$. It remains open how to generalize the method of [65] to deal with more than one amplitude damping error, while the error set $\mathcal{E}^{\{1\}}$ can naturally be generalized, e.g., to $\mathcal{E}^{\{2\}}$ for correcting two amplitude damping errors.

2.5.2 Codes correcting two amplitude damping errors

We have performed an exhaustive search for codes correcting two amplitude damping errors, i.e., detecting the error set $\mathcal{E}^{\{2\}}$, up to length $n = 9$. In fact, the resulting codes correct any combination of X and Y errors up to weight two, as well as a single Z error.

No non-trivial CWS codes are found for length $n \leq 8$. For length $n = 9$, two LC-inequivalent codes encoding a single qubit have been found. These are both stabilizer codes encoding a single qubit, since the corresponding classical code $\mathcal{C}$ is trivially linear [14].

One code has the stabilizer S_1 generated by

$$\begin{array}{ccccccccc}
X & I & I & I & I & I & I & I & Z \\
Z & I & I & I & X & Z & I & Z & X \\
I & X & I & I & I & I & I & Z & I \\
I & Z & I & I & X & I & Z & X & Z \\
I & I & X & I & I & Z & I & I & I \\
I & I & Z & I & Y & Y & Z & Z & I \\
I & I & I & X & I & I & Z & I & I \\
I & I & I & Z & Y & Z & Y & I & Z
\end{array}$$

The other code has the stabilizer S_2 generated by

$$\begin{array}{ccccccccc}
X & I & I & I & I & I & I & I & Z \\
Z & I & I & I & Z & Z & Z & I & X \\
I & X & I & I & I & I & I & Z & I \\
I & Z & I & Z & Z & Y & I & Y & Z \\
I & I & X & I & I & I & Z & I & I \\
I & I & Z & Z & Z & X & X & Z & I \\
I & I & I & X & I & Z & I & I & I \\
I & I & I & I & X & I & Z & Z & Z
\end{array}$$

It is straightforward to check that these codes detect the error set $\mathcal{E}^{\{2\}}$, since no elements in $\mathcal{E}^{\{2\}}$ are also in $C(S_i) \setminus S_i$ (for $i = 1, 2$). Furthermore, both codes are degenerate since some of the elements in $\mathcal{E}^{\{2\}}$ are indeed in S_i, for instance $X_1 Z_9$.

These codes outperform the $((10, 2))$ code found in [21]. Recall that Shor's nine-qubit code, having the same parameters $((9, 2))$ as our codes, also corrects

two amplitude damping errors [26]. However, Shor's code only corrects the Kraus operators A_0 and A_1, but does not detect the error set $\mathcal{E}^{\{2\}}$. Therefore, for length $n = 9$, we have found good codes that outperform all the previous known constructions for detecting the error set $\mathcal{E}^{\{2\}}$.

2.5.3 Codes detecting a single amplitude damping error and detecting multiple dephasing errors

For the error set $\mathcal{E}^{\{3\}}$, we have performed an exhaustive search for different lengths n and Z-error detecting capabilities r up to $n = 8$, and a random search starting from randomly selected graph states for $n = 9$ and different r. Our results are listed in Table 2.2. We compare our results with the best stabilizer codes that detect all errors up to weight r as given in [7], and the codes detecting a single amplitude damping errors and Z errors up to weight r as found in [23].

n/r	1			2			3			4			5			6			7		
	stab.	[23]	CWS	stab.	[23]	CWS	stab.	[23]	CWS	stab.	[23]	CWS	stab.	[23]	CWS	stab.	[23]	CWS	stab.	[23]	CWS
5	4	5	6	2	-	4	-	-	2	-	-	2	-	-	-	-	-	-	-	-	-
6	16	16	16	2	2	8	-	-	4	-	-	2	-	-	2	-	-	-	-	-	-
7	16	22	24	2	8	16	-	-	8	-	-	2	-	-	2	-	-	2	-	-	-
8	64	64	64	8	8	20	-	8	16	-	-	4	-	-	2	-	-	2	-	-	2
9	64	93	96*	8	16	40*	-	8	20*	-	-	6*	-	-	4*	-	-	2*	-	-	2*

Table 2.2: Dimension K of CWS codes detecting the error set $\mathcal{E}^{\{3\}}$ (as defined in equation 2.21) for different length n and parameter r. For each value of r, the first column lists the largest dimension of stabilizer codes that detect all errors up to weight r as given in [7]; the second column lists the largest dimension of asymmetric codes detecting a single amplitude damping error and phase errors up to weight r as found in [23]; the third column lists the largest dimension of the CWS codes found by our search for codes detecting the error set $\mathcal{E}^{\{3\}}$. '$-$' means that no non-trivial codes exist based on the construction. The numbers labeled with * are the best parameters found by random search; otherwise the maximal dimension is obtained by exhaustive search.

As we can see from the table, for most lengths n and Z-error weight r, the CWS codes found outperform the known results. The entries for which we did not find improvements are $n = 6$, $r = 1$ and $n = 8$, $r = 1$. Codes with $r = 1$ detect

single Pauli errors, i.e., they are codes of minimum distance two. For even length, the corresponding stabilizer codes are known to have the largest possible dimension for single-error-detecting codes [54]. For odd length $n = 5, 7, 9$, we find codes with parameters matching those of the code family $((2m+1, 3 \times 2^{2m-3}, 2))$ given in [54]. Whenever the dimension is a power of two, the codes we found include stabilizer codes.

2.6 Discussion

Using the CWS framework we were able to construct a multitude of new quantum error correcting codes capable of protecting quantum states from amplitude damping or phase damping errors. Among these new codes we have found examples of codes which have parameters better than any previously found by other constructions for the precise error set we use. Better codes do exist but only because they restrict the error-set to only A_1 type errors for spontaneous emission and can't protect against spontaneous absorption, which for finite temperature should also be included in error set.

This process is very general and can be expanded to find codes correcting any error-set. Though this was an exhaustive search of CWS codes which include many types of codes, there may exist codes with better parameters which are not representable as a CWS code. So these new results are very interesting but not necessarily the last word on the subject.

The primary downside to this method of finding codes is the extreme computational costs associated with it. Performing an exhaustive search is an exponentially difficult problem in the size of the code. While the search is highly par-

allelizable and can achieve near linear speed-up in the number of processors used. Regardless for $n > 12$ I expect the problem would grow beyond what is feasible with super computer cluster time at your disposal.

It may be the case that, like many computationally hard problems, it is easier to approximate good solutions than it is to find the best solution. Unfortunately the max-clique problem does not have this property and there is not yet known to be any structure in the problem that could be exploited in an approximating code search. However if such structure could be found, that might make searches for larger codes feasible.

An alternative option is to abandon the search altogether and instead use the smaller codes we have found to construct larger codes protecting bigger spaces. This idea I explore further in the next chapter.

Chapter 3

Concatenated Codes Correcting Asymmetric Errors

3.1 Introduction

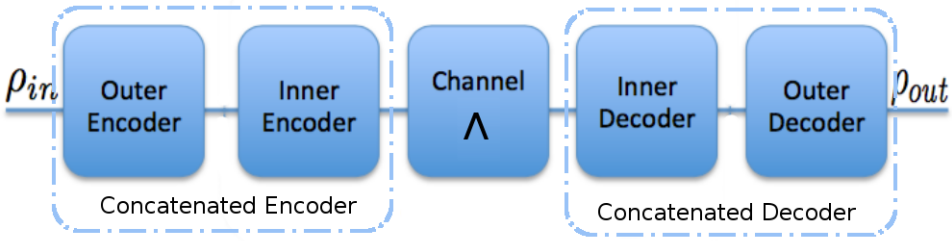

Figure 3.1: Outline of concatenated codes

In the previous chapter we illustrated the need for quantum error correcting codes for reliable quantum computing. We also noted that most attention for quantum error correcting codes has been focused on symmetric error models but there exist asymmetric error models which have not received as much attention. We focused on the amplitude and phase damping channels and through a brute force search found new codes with parameters better than any previously known for these error models.

The major draw back of the previous method was the exponentially large expense of computing resources. In this chapter, we extend the previous work using the method of concatenated codes in order to construct quantum error correcting codes for amplitude damping errors for numbers of qubits n far greater than is feasible with the search method of the previous chapter.

Quantum code concatenation involves encoding a state into a first (outer) QECC and then each of the physical qubits of this code are re-encoded into a second (inner) QECC. This process creates a new code with new parameters from two existing codes. In the picture above this new code combines the two encoding and decoding operations into a single one.

The relationship of the parameters of the concatenated code to the parameters of the inner and outer codes can be summed up in the following theorem.

Theorem 3.5. *Let the inner code $Q_i$be a qubit code with parameters $[n_1, k_1, d_1]_2$ and let the outer code Q_o be a qudit code with qudit dimension 2^{k_1} having parameters $[n_2, k_2, d_2]_{2^{k-1}}$, then the concatenated code Q has parameters $[n_1 n_2, k_1 k_2]_2$ with distance at least $d_1 d_2$*

Proof. To highlight how the process of concatenation works the outer code encodes a $2^{k_2^{k_1}}$ dimension state into n_2 qudits. Each of these qudits can then be re-encoded into the inner qubit code using n_1 qubits for each qudit of the outer code for a total of $n_1 n_2$ qubits. The dimension of the protected space is not changed by concatenation so it is $2^{k_1 k_2}$. Finally, the distance of the concatenated code is at least $d_1 d_2$ due to the fact that the logical operators of a concatenated code are constructed by taking each Pauli operator in the tensor product of the outer logical operator and replacing them with the n_1 qubit logical operator corresponding to it. As a result the

minimum weight logical operator is lower bounded by the product of the minimum weight logical operators of the composite codes. □

Example 3.1 (Shor's 9-qubit code)**.** *The classic example of code concatenation is Shor's 9-qubit code which can be seen as encoding a single qubit in the phase flip code followed by re-encoding each of the three qubits in the bit flip code.*

$$|0\rangle \quad \rightarrow |+++\rangle \quad \rightarrow \frac{1}{2\sqrt{2}}(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle) \tag{3.1}$$

$$|1\rangle \quad \rightarrow |---\rangle \quad \rightarrow \frac{1}{2\sqrt{2}}(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle) \tag{3.2}$$

Code concatenation can also be done in the stabilizer formalism. The process of constructing the stabilizers for the full code involve producing operators where the stabilizer of the inner code acts on each block of qubits unioned with the stabilizers of the outer code with each operator replaced with the corresponding logical operator of the inner code. The logical operators of the concatenated code are constructed similarly.

In this case the inner code has stabilizer generated by ZZI and IZZ and logical operators $\tilde{X} = XXX$ and $\tilde{Z} = ZZZ$. The outer code has stabilizer generated by XXI and IXX and logical operators $\tilde{X} = ZZZ$ and $\tilde{Z} = XXX$. Therefore the concatenated code's stabilizer is generated by :

$$ZZIIIIIII \qquad IZZIIIIII$$

$$IIIZZIIII \qquad IIIIZZIII$$

$$IIIIIIZZI \qquad IIIIIIIZZ$$

$$XXXXXXIII \quad IIIXXXXXX$$

and has as logical operators $\tilde{X} = ZZZZZZZZ$ *and* $\tilde{Z} = XXXXXXXXX$.

The process of using concatenation to find new codes with good parameters has already had many good results in particular multi-error-correcting AD codes are discussed in [21], based on a concatenation method. In particular, the inner code is chosen as the two-qubit code $\{|01\rangle, |10\rangle\}$ based on the classical dual-rail code, which results in a quantum erasure channel for the outer codes. Many good stabilizer AD codes are constructed by concatenating with the quantum erasure codes. However, due to the choice of the inner code, the rate of the constructed AD codes can never exceed 1/2.

For this work, we choose the inner codes as Calderbank-Shor-Steane (CSS) codes capable of correcting the asymmetric error-sets associated with the amplitude damping channel. By concatenating with outer codes correcting symmetric errors, many new codes with good parameters are found, which are better than the amplitude damping codes obtained by any previously known construction.

3.2 Distance of Codes Correcting Asymmetric Channels

In order to understand the performance of codes correcting a particular class of errors (most often the depolarizing channel) the distance of the code is the metric used.

Definition 3.9. *A code Q is of distance d if the error detection condition* (1.1) *is satisfied for all errors E that act non-trivially on at most $d-1$ qudits.*

The weight of an operator or error can be defined as [50]:

Definition 3.10. *The weight of a Pauli operator of the form* $P = \bigotimes_{i=1}^{n} P_i$ *where each* P_i *is one of single qubit Pauli operators* I*,* X*,* Y*, or* Z *is* $\sum_{i=1}^{n} \mathrm{wt}(P_i)$*. For this standard definition of weight, the weight of the single qubit Pauli operators are 0 for the* I *and 1 for* X*,* Y *and* Z*.*

The standard weight of a Pauli operator is equivalent to counting the number of non-identity single qubit operators in the tensor product. As such, the definition of distance can be reworded as:

Definition 3.11. *A code* Q *is of distance* d *if the error detection condition* (1.1) *is satisfied for all* $E \in \mathcal{E}$ *such that* $\mathrm{wt}(E) \leq d - 1$*.*

Alternatively, the definition of distance is sometimes given as the minimum weight of an operator E such that the error detection criteria is violated. These two alternative definitions are however equivalent.

The distance of a code is a useful measure for comparing codes because it gives information about the size of the error-set it is able to correct. Quantum error correcting codes are often cited as $((N, K, d))_q$. A code with distance d can not only detect any error up to weight $d - 1$ but can also correct any error of weight $\lfloor \frac{d-1}{2} \rfloor$ [50].

For this work, instead of the symmetric error model, we consider errors sets for the asymmetric amplitude damping (AD) channel. The amplitude damping channel has non-Pauli Kraus operators A_0, A_1 as in (2.16). Even though A_0 and A_1 are not Pauli operators, we can construct Pauli error models that lead to codes correcting AD errors.

It is important to notice in the process of "Pauli-ization" of the AD Kraus

operators:

$$A_1 = \frac{\sqrt{\gamma}}{2}(X + iY), \quad A_0 = I - \frac{\gamma}{4}(I - Z) + O(\gamma^2), \tag{3.3}$$

that A_0 is of a different order in γ than A_1. So the corresponding asymmetric error model as given in Eq. (2.19) has $p_{xy} \propto \gamma$ and $p_z \propto \gamma^2$.

In the case of asymmetric channels and the codes which correct asymmetric error models, such as $\mathcal{E}^{\{t\}}$, the standard definition of distance does not provide all that much useful information.

A t-error-correcting AD code (or t-code for short) improves the fidelity of the transmitted state from $1 - \gamma$ to $1 - \gamma^t$. For instance, for $t = 1$, we only need to correct a single A_1 error and detect a single A_0 error [34]. As before, this requires us to correct a single X and Y error, and detect a single Z error. In other words, a code that detects the error set $\mathcal{E}^{\{1\}}$ that is given by

$$\mathcal{E}^{\{1\}} = \{I\} \cup \{X_i, Y_i, Z_i,\ X_iX_j, X_iY_j, Y_iY_j\}, \tag{3.4}$$

with $i, j \in [1, n]$, is a 1-code that corrects a single AD error.

Pauli error models that lead to codes correcting t AD errors can be given similarly. For instance, codes detecting the Pauli error set given by $\mathcal{E}^{\{2\}} = \{A_\mu A_\nu \colon A_\mu, A_\nu \in \mathcal{E}^{\{1\}}\}$ are 2-codes that correct 2 AD errors. We will similarly denote by $\mathcal{E}^{\{t\}}$ the Pauli error set that results in t-codes. Once again a side effect of this process of finding a Pauli error set which will correct amplitude or phase damping errors results in codes which are slightly stronger than necessary in that they will also correct $A_1^\dagger$ errors.

3.2.1 Effective Distance

Returning to the discussion of distance, if we consider the error set $\mathcal{E}^1$ and any code correcting it, these codes would have distance 2 since it is missing some errors of weight 2 such as ZZ but has all errors of weight $1 = d - 1$ and therefore any code correcting $\mathcal{E}^1$ is a distance 2 code. This doesn't properly convey the information that the distance is supposed to in that using the standard definition these code correct no errors when in reality they correct one error of the type we are interested in. If we were to have a distance measure for asymmetric codes it should capture this sort of information. This is the motivation for the next few definitions.

When we examine the properties of the elements in $\mathcal{E}^{\{t\}}$, we note that from Eq. (2.17) it follows that each Z error contributes a factor of γ^2 to the noise, while X or Y errors contribute a factor of γ. In other words, when we consider each X or Y error as '1 error', then each Z error will be effectively '2 errors'. Motivated by this and the previous observations, we can formulate the following definition for 'effective weight'.

Definition 3.12. *For any tensor product E of Pauli errors, each tensor factor X or Y has effective weight* 1*, and each factor Z has effective weight* 2 *(as usual I has a weight contribution of* 0*). The effective weight of E is the sum of the effective weight of all factors X, Y, Z in E, and is denoted by* $\mathrm{wt}_e(E)$.

As an example of calculating the effective weight of an operator, take $E = XYIZ \in \mathcal{E}^{\{2\}}$, this has effective weight of 4. Simply calculated by adding 1 weight each for the X and Y and 2 more for the Z. The I contributes nothing.

This is nearly identical to the normal definition of weight, except that it has an asymmetry in the contributions of single qubit Pauli operators which is motivated

by the types of errors in our error-set.

This definition of effective weight leads to a definition of effective distance which restores the ability to properly describe the magnitude of detectable error-set.

Definition 3.13. *A code has effective distance d_e, if it detects Pauli errors of effective weight up to $d_e - 1$.*

If we examine the codes which detect the error-set $\mathcal{E}^{\{1\}}$, these are codes which correct a single amplitude damping error. The effective distance of this code would be the lowest effective weight an operator not in $\mathcal{E}^{\{1\}}$. In this case all errors with effective weight 2 or less are included in the set. Which is easy to see as the operators with effective weight 1 will be those with a single X or Y Pauli in any position. Those with effective weight two have two decompositions, either 2 operators from $\{X, Y\}$ in any pairs of positions, or a single Z operator in any position. which is precisely the definition of the error-set $\mathcal{E}^{\{1\}}$ which the code detects.

Therefore any code detecting the error-set $\mathcal{E}^{\{1\}}$ has distance 3. If the notion of effective distance is to make sense then such a code should be capable of correcting $\frac{3-1}{2} = 1$ errors which we have already shown it is.

In fact, we have the following results on the effective weight of the elements in $\mathcal{E}^{\{t\}}$.

Lemma 3.1. *The effective weight of the product of any two Pauli operators P_1, P_2 with effective weights w_1 and w_2 respectfully will have effective weight at most $w_1 + w_2$.*

Proof. This is a rather rudimentary fact which is also an extremely important property for our definitions to maintain their usefulness for larger error-sets. To show it, we can look at the result of multiplication position wise for each pairing of single-

qubit Paulis. For instance: If position i on P_1 and P_2 have are X and Y respectfully then their product will be Z. The total weight for this position both before an after is 2 and so this case is satisfied. All cases are outlined in table 3.1. □

P_1	P_2	result	input weight	result weight
I	X/Y	X/Y	1	1
I	Z	Z	2	2
X	X	I	2	0
X	Y	Z	2	2
X	Z	Y	3	1
Y	Y	I	2	0
Y	Z	X	3	1
Z	Z	I	4	0

Table 3.1: The total effective weight of pairs of single qubit Pauli operators compared to the effective weight of their product. Notice the result weight column is always less than or equal to the input weight.

If this lemma did not hold then our definition of effective weight and effective distance would lose the property that codes with distance d_e can correct $\frac{d_e-1}{2}$ errors of a particular type.

The previous lemma leads naturally to the following:

Lemma 3.2. *Any element $E \in \mathcal{E}^{\{t\}}$ has effective weight* $\mathrm{wt}_e(E) \leq 2t$.

Proof. Notice that any element $E \in \mathcal{E}^{\{t\}}$ will be a product of at most t elements from $\mathcal{E}^{\{1\}}$ as given in Eq. (3.4). Any element in $\mathcal{E}^{\{1\}}$ has at most effective weight 2. As such, following from lemma 3.1, E has at most effective weight $2t$. □

Obviously, the upper bound is achievable by some elements in $E \in \mathcal{E}^{\{t\}}$. Furthermore all elements of effect weight less than $2t$ are contained in this set.

Lemma 3.3. *The set $\mathcal{E}^{\{t\}}$ contains every element of effective weight less than or equal to $2t$.*

Proof. Assume a Pauli operator P with weight $w \leq 2t$. There are groups of options for how this weight can be distributed. Let n_Z be the number of Z Paulis present in the operator, then the number of Paulis from the set $\{X, Y\}$ is $n_{XY} = w - 2n_Z$. The number of Z Paulis is constrained by $2n_Z \leq w \leq 2t$ therefore $n_Z \leq t$. Any n_Z is possible then as the set $\mathcal{E}^{\{1\}}$ contains the single Z operators so simply choose the n_Z operators for the positions. As for the remaining X and Y operators, every pair of positions as well as single position operators containing X and Y are in $\mathcal{E}^{\{1\}}$ so only $\lceil \frac{n_{XY}}{2} \rceil = \lceil \frac{w}{2} \rceil - n_Z$ more operators from $\mathcal{E}^{\{1\}}$ are required to construct this operator P. The total number of operators required is then simply $\lceil \frac{w}{2} \rceil$. If $w = 2t$ then t operators need be multiplied together and as such $P \in \mathcal{E}^{\{t\}}$ for any $w < 2t$ this construction is still guaranteed to work. □

The previous two lemmas allow us to say the following:

Theorem 3.6. *If a code has effective distance $d_e = 2t + 1$, then it detects the error set $\mathcal{E}^{\{t\}}$, and is hence a t-code.*

Proof of which follows directly from the previous lemmas which allow us to say that $\mathcal{E}^{\{t\}}$ is the error-set of all elements with effective weight less than d_e. As such, this definition of effective distance restores the original purpose of the distance metric and captures the error correcting capabilities of codes.

3.3 Concatenated Codes

Now we are ready to present the concatenation method which is a specialization of the generalized concatenation procedure demonstrated in [72].

In this work, we focus on the construction of AD codes, which are qubit codes with $q = 2$. However, qudit codes with $q = 2^r$ are used as outer codes for the concatenation constructions in order to get qubit AD codes.

Theorem 3.7. *Starting from an inner $[\![n_1, k_1]\!]_2$ code $\mathcal{Q}_i$ with effective distance d_e, concatenation of an $[\![n_2, k_2, \delta]\!]_{2^{k_1}}$ qudit outer code $\mathcal{Q}_o$ with distance δ results in a concatenated code $[\![n_1 n_2, k_1 k_2]\!]_2$ with effective distance at least $d_e \delta$.*

Proof. The concatenated code $\mathcal{Q}$ is a stabilizer code with length $n_1 n_2$ as each of the n_2 qudits of dimension 2^{k_1} will be re-encoded into the code subspace of one instance of the inner code using n_1 qubits each for a total of $n_1 n_2$ qubits total.

The encoded dimension of the code will be the same as the dimension as the outer code as the inner code does not allow for additional information to be encoded. As such, the concatenated code protects k_2 qudits of dimension 2^{k_1} for a total protected dimension of $(2^{k_1})^{k_2}$, hence encoding $k_1 k_2$ qubits.

Denote the stabilizer of $\mathcal{Q}$ by $S_{\mathcal{Q}}$. It has two sets of generators. The first set is obtained by replacing each tensor factor in the generators of the outer code's stabilizer $S_{\mathcal{Q}_o}$ by the corresponding logical operator of the inner code. The second set is formed by the stabilizer S_i of the inner code acting on each block of n_1 qubits.

Example 3.2. *If the outer code had a stabilizer of XIZ and the inner code had logical operators $\tilde{X} = XX$ and $\tilde{Z} = ZI$ then the operator $(XX)(II)(ZI)$ would be an element of the first set of generators for the concatenated code.*

For the outer code $\mathcal{Q}_o$, any nontrivial logical operator in $C(S_{\mathcal{Q}_o}) \setminus S_{\mathcal{Q}_o}$ has

weight at least δ, where $C(S)$ is the centralizer of the stabilizer S. Likewise, the logical operators in $C(S_{\mathcal{Q}_i}) \setminus S_{\mathcal{Q}_i}$ of the inner code have effective weight at least d_e. The logical operators of the concatenated code are obtained by replacing each tensor factor in the logical operators of the outer code by the corresponding logical operator of the inner code precisely as with the first set of generators. Those operators have effective weight at least $d_e\delta$.

As is the case with standard concatenation of quantum codes [44], if minimum weight logical operators are used for both the inner and outer codes of the construction, then multiplying a logical operator of $\mathcal{Q}$ by an element of the stabilizer $S_{\mathcal{Q}}$ will not result in an effective weight less than $d_e\delta$.

Therefore the logical operators of the concatenated code will have effective weight no less than $d_e\delta$ and as such the effective distance of the concatenated code will be at minimum $d_e\delta$. $\square$

3.4 Concatenated Codes for Asymmetric Channels

To examine the power of the construction for AD codes given in Theorem 3.7, we will start with simple inner codes. We take classical linear binary codes of distance 2 with length r and dimension $r-1$ (hence cardinality 2^{r-1}). For any length r, such a distance 2 code will be formed by all bit strings of length r with even Hamming weight. For any such classical code $\mathcal{C}_r = [r, r-1, 2]_2$, there exists a corresponding quantum code $\mathcal{Q}_r = [\![r, r-1]\!]_2$ with code space spanned by the computational basis vectors $|\mathbf{c}_i\rangle$ for all $\mathbf{c}_i \in \mathcal{C}_r$. We first examine the effective distance of $\mathcal{Q}_r$.

Lemma 3.4. *The code $\mathcal{Q}_r$ defined above has effective distance $d_e = 2$.*

Proof. The stabilizer of the code $\mathcal{Q}_r$ consists of only two elements. The identity and the operator $Z^{\otimes r}$. It is easy to check that all basis states with an even number of 1's will be stabilized by this operator.

In order to determine the effective distance of the code we need only check the effective weights of the logical operators that are in $C(S_r) \setminus S_r$, where $C(S_r)$ is the centralizer of S_r. These are Pauli operators that commute with $Z^{\otimes r}$.

Obviously any operator which is a tensor product of only Z and I operators will commute, the smallest of these is a single Z (i.e., Z_i) operator having effective weight two. Operators with an odd number of X or Y will not appear in $C(S_r) \setminus S_r$ as they will always anti-commute with $Z^{\otimes r}$.

The tensor product of two X or Y operators (i.e., $X_iX_j, X_iY_j, Y_iX_j, Y_iY_j$) is in $C(S_r) \setminus S_r$. Therefore, every logical operator of $\mathcal{Q}_r$ has effective weight at least two, and hence the effective distance of $\mathcal{Q}_r$ is 2. □

Since the dimension of the quantum code $\mathcal{Q}_r$ is 2^{r-1}, it can be used as inner code for the concatenation with a qudit outer codes with single qudit dimension $q = 2^{r-1}$. For the construction of a t-code, we need effective distance $2t + 1$ for the concatenated code.

However note that if using the an inner code of $\mathcal{Q}_r$ and an outer code $\mathcal{Q}_o$ with properties $[\![n, k, \delta]\!]_{2^{r-1}}$ then theorem (3.9) states that the output code will be a $[\![nr, k(r-1), 2\delta]\!]_2$ code.

This code will have effective distance 2δ which means it is a t-code where $t = \lfloor \frac{2\delta-1}{2} \rfloor = \delta - 1$. Due to the rounding down however, this code has the same correction capability as a code with effective distance $2\delta - 1$.

As an example, assume a code has distance $6 = 2\delta$ where $\delta = 3$. A code with distance 6 will correct $t = \lfloor \frac{6-1}{2} \rfloor = 2$ AD errors. Consider another code with effective distance 5. This code will correct $t = \lfloor \frac{5-1}{2} \rfloor = 2$ AD errors, the same number of errors with a code with one less distance.

This observation motivates the next theorem in which we use this fact to obtain a slight improvement in the code parameters.

Theorem 3.8. *Given an $[\![n, k, \delta]\!]_{2^{r-1}}$ stabilizer code, a quantum code $\mathcal{Q}$ with parameters $[\![rn-1, (r-1)k]\!]_2$ and effective distance $d_e \geq 2\delta - 1$ can be constructed. This is a t-code with $t = \delta - 1$.*

Proof. We start from an $[\![n, k, \delta]\!]_{2^{r-1}}$ stabilizer code of length n, and each qudit has dimension 2^{r-1}. The first qudit is encoded into a trivial qubit code with parameters $[\![r-1, r-1, d_e = 1]\!]_2$. Essentially just encoding the data of the qudit with dimension $q = 2^{r-1}$ into $r-1$ qubits without any protection.

Each of the other qudits $j = 2, 3, \ldots, n$ is encoded into the code $\mathcal{Q}_r$ with parameters $[\![r, r-1, d_e = 2]\!]_2$.

The resulting concatenated code $\mathcal{Q}$ is a stabilizer code of length $(r-1) + (n-1)r = rn - 1$ and dimension $(2^{r-1})^k$, hence encoding $(r-1)k$ qubits.

Any logical operator of $[\![n, k, \delta]\!]_{2^{r-1}}$ has weight at least δ. Hence any logical operator of $\mathcal{Q}$ that acts trivially on the first qudit has effective weight at least 2δ.

Logical operators of $\mathcal{Q}$ that act non-trivially on the first qudit have effective weight at least $1 + 2(\delta - 1) = 2\delta - 1$. Therefore, the effective distance of $\mathcal{Q}$ is $d_e \geq 2\delta - 1$. $\square$

By choosing not to re-encode the first qubit of the outer code with the inner code we reduce the total number of qubits used by one whilst maintaining the

dimension of the protected space. We end up decreasing the distance of the code by one but this doesn't impact the error-correcting capability of the code due to the redundancy noted earlier. This process is clearer with an example:

Example 3.3. *Starting from the* $[\![5,1,3]\!]_2$ *code with stabilizer generated by*

$$\begin{array}{ccccc} X & Z & Z & X & I \\ I & X & Z & Z & X \\ X & I & X & Z & Z \\ Z & X & I & X & Z \end{array}$$

and encoding qubits $2,3,4,5$ *into the code* $\mathcal{Q}_2$ *stabilized by* ZZ*, we get a* $[\![9,1]\!]_2$ *code with effective distance* $d = 2\cdot 3 - 1 = 5$*, which corrects two AD errors. By choosing the logical operators for* $\mathcal{Q}_2$ *as* $\overline{X} = XX$ *and* $\overline{Z} = ZI$*, the stabilizer of the* $[\![9,1]\!]_2$ *code is generated by*

$$\begin{array}{ccccc} X & ZI & ZI & XX & II \\ I & XX & ZI & ZI & XX \\ X & II & XX & ZI & ZI \\ Z & XX & II & XX & ZI \\ I & ZZ & II & II & II \\ I & II & ZZ & II & II \\ I & II & II & ZZ & II \\ I & II & II & II & ZZ \end{array}$$

We remark that the $[\![9,1]\!]_2$ 2-code given above is in fact local Clifford equivalent to one of the $[\![9,1]\!]_2$ codes found in Chapter 2 via exhaustive numerical search for CWS codes detecting the error set $\mathcal{A}^{\{2\}}$. It is one of the best 2-codes known, which beats the $[\![10,1]\!]_2$ 2-code found in [21]. In fact, the construction

in [21] can be viewed as a special case of Theorem 3.8, by concatenating all qudits of an outer code with the inner code $\mathcal{Q}_2$. Notice that in [21], codes with effective distance 2δ are constructed in order to obtain t-codes with $t = \delta - 1$, which results in length $2n$ instead of $2n - 1$ as given by Theorem 3.8. In other words, by using Theorem 3.8, the length of any t-code constructed in [21] can be reduced by one.

For decoding, the inner code $[\![r, r-1]\!]_2$ will be used to detect single X- and Y-errors. This provides side-information on detected errors (erasures) for the outer code and allows to simultaneously correct e erasures and f erroneous blocks with r qubits each, as long as $e + 2f < \delta$.

3.5 Results

Using our theorem, 3.8 we can now construct new AD codes through our partial concatenation process with an inner code of $\mathcal{Q}_r$ with various outer codes. We will discuss the parameters of these new AD codes as well as compare the effective distance d_e of the new codes constructed via our concatenated method to the distance d_{lb} of the best known stabilizer codes. Note that this comparison is valid as the standard distance of a code provides information on how many arbitrary errors are capable to be corrected including amplitude damping errors.

However keep in mind at all times that the effective distance we use here only applies to amplitude damping errors. If these codes were to be used to consider different error models, the definition of effective distance as we have chosen here would be once again useless.

The best possible parameters for our concatenation technique are expected when the outer code is an optimal quantum code, and quantum MDS (QMDS) codes

in particular. QMDS codes are codes which attain the quantum Singleton bound $k + 2d = n + 2$ [43, 53] and have parameters $[\![n, n + 2 - 2d, d]\!]_q$. QMDS codes are known to exists for all $n \leq q + 1$, for $n = q^2 - 1, q^2, q^2 + 1$ and some $d \leq q + 1$, as well as for many parameters $n \leq q^2 + 1$, $d \leq q + 1$ [29]. In general it seems as if for a qudit QMDS code with qudit dimension q we have the bounds $d \leq q + 1$, and $n \leq q^2 + 1$, with the exception of codes $[\![4^m + 2, 2^m - 4, 4]\!]_{2^m}$ (see [32]).

In order to construct a t-code, we use QMDS codes $[\![n, n-2t, t+1]\!]_q$ where $q = 2^{r-1} \geq t$ as outer code and the code $\mathcal{Q}_r = [\![r, r - 1]\!]_2$ as inner code, yielding a t-code of length $rn - 1$ encoding $(r - 1)(n - 2t) = rn - n - 2rt + 2t$ qubits.

The parameters of our codes based on the concatenation of QMDS codes and the code $\mathcal{Q}_r$ are presented in Table 3.2. The last column labelled d_{lb} lists the largest known lower bound d_{lb} on the minimum distance of a stabilizer code for the depolarizing channel (see [28]). Here we consider only codes of length up to $n_{\max} = 128$. We only list the parameters $[\![n, k, d_e = 2t + 1]\!]_2$ of t-codes for which the effective distance d_e exceeds the lower bound d_{lb} (i.e., $d_e > d_{\text{lb}}$). Furthermore, we omit parameters for which we find even betters codes (smaller length, larger dimension, or larger effective distance).

In Tables 3.3 and 3.4 we list parameters of the best t-codes we found using outer codes that do not reach the quantum Singleton bound $k + 2d \leq n + 2$, but have the largest minimum distance among the known codes. The codes in Table 3.3 are based on qubit codes as outer codes and hence comparable to the codes in [21], but reducing the length by one as discussed above.

3.6 Discussion

In this work we formalized the concept of effective distance in the case of amplitude damping codes, allowing us to properly characterize the error tolerance of amplitude damping codes in a manner similar to what distance does for characterizing the effectiveness of codes correcting symmetric error models. Once defined, we were able to describe the characteristics of concatenated codes, which naturally led to developing the "partial concatenation" technique allowing us to use one less qubit in a code without sacrificing the error correction capabilities.

By using this process we construct many codes with parameters better than those formally known. We also rediscovered the 9-qubit codes capable of correcting two AD errors which we found in the previous chapter through exhaustive search. However the codes we construct are just a subset of the codes possible. The concatenation and partial concatenation techniques can be used in many other situations.

One example of which is using other asymmetric codes as inner codes to construct concatenated codes based on Theorem 3.7. Using a similar idea as in Theorem 3.8, one may also encode the first qudit of the outer $[\![n_2, k_2]\!]_{2^{k_2}}$ code into a trivial $[\![k_2, k_2]\!]_2$ code. This leads to the following corollary.

Corollary 3.9. *Concatenating an $[\![n_2, k_2, \delta]\!]_{2^{k_1}}$ qudit outer code $\mathcal{Q}_o$ with an inner asymmetric $[\![n_1, k_1]\!]_2$ code $\mathcal{Q}_i$ with effective distance d_e results in a code $[\![n_1(n_2 - 1) + k_2, k_1 k_2]\!]_2$ with effective distance at least $d_e(\delta - 1) + 1$, as well as a concatenated code $[\![n_1 n_2, k_1 k_2]\!]_2$ with effective distance at least $d_e \delta$.*

Example 3.4. *Choose the inner code to be the asymmetric $[\![8, 3, \{4, 2\}]\!]_2$ CSS code with X-distance $d_X = 4$ and Z-distance $d_Z = 2$, resulting in effective distance $d_e = 4$.*

It can be constructed from the first order Reed-Muller code and the repetition code. Its stabilizer is generated by

$$\begin{array}{cccccccc} Z & Z & Z & Z & I & I & I & I \\ Z & Z & I & I & Z & Z & I & I \\ Z & I & Z & I & Z & I & Z & I \\ Z & Z & Z & Z & Z & Z & Z & Z \\ X & X & X & X & X & X & X & X \end{array}$$

Based on Theorem 3.7, concatenating with a QMDS $[\![10, 2, 5]\!]_{2^3}$ *outer code results in a code* $[\![80, 6]\!]_2$ *with effective distance* $d_e = 20$*. This code is better than the best known stabilizer code* $[\![80, 6, 16]\!]_2$*. Using Corollary 3.9, we get a* $[\![75, 6]\!]_2$ *code with effective distance* $d_e = 17$*, correcting* $t = 8$ *AD errors. This again improves upon the best known stabilizer code* $[\![75, 6, 15]\!]_2$*. However, the* $t = 8$ *code with parameters* $[\![74, 6]\!]_2$ *listed in Table 3.4 has better parameters. Note that for both codes* $[\![8, 3, \{4, 2\}]\!]_2$ *and* $[\![2, 1, \{2, 1\}]\!]_2$ *(i.e., the code* $\mathcal{Q}_2$ *with the stabilizer generated by* ZZ*), the ratio between the* X*- and* Z*-distance is* 2*, resulting in an effective distance of* 4 *and* 2*, respectively. However, the* $[\![2, 1, \{2, 1\}]\!]_2$ *code has rate* $1/2$ *compared to rate* $3/8$ *for the* $[\![8, 3, \{4, 2\}]\!]_2$ *code, resulting in codes with better parameters.*

This example illustrates the flexibility of the method presented. It is also possible to use it for channels for which the asymmetry between p_{xy} and p_z is different than for the amplitude damping channel (see, e.g. [38]). In these cases we could redefine the effective distance as the situation suited and use concatenation once again.

Table 3.2: Concatenated codes $[\![n, k, d_e]\!]_2$ for the AD channel based on QMDS outer codes with qudit dimension 2, 4, 8, and 16.

t	concatenated code	outer code	d_{lb}
1	$[\![7, 2, d_e = 3]\!]_2$	$[\![4, 2, 2]\!]_2$	2
2	$[\![9, 1, d_e = 5]\!]_2$	$[\![5, 1, 3]\!]_2$	3
3	$[\![23, 4, d_e = 7]\!]_2$	$[\![8, 2, 4]\!]_{2^2}$	6
	$[\![26, 6, d_e = 7]\!]_2$	$[\![9, 3, 4]\!]_{2^2}$	6
	$[\![29, 8, d_e = 7]\!]_2$	$[\![10, 4, 4]\!]_{2^2}$	6
	$[\![41, 16, d_e = 7]\!]_2$	$[\![14, 8, 4]\!]_{2^2}$	6
4	$[\![26, 2, d_e = 9]\!]_2$	$[\![9, 1, 5]\!]_{2^2}$	8
	$[\![50, 18, d_e = 9]\!]_2$	$[\![17, 9, 5]\!]_{2^2}$	8
	$[\![39, 6, d_e = 9]\!]_2$	$[\![10, 2, 5]\!]_{2^3}$	8
	$[\![43, 9, d_e = 9]\!]_2$	$[\![11, 3, 5]\!]_{2^3}$	8
	$[\![47, 12, d_e = 9]\!]_2$	$[\![12, 4, 5]\!]_{2^3}$	8
	$[\![59, 21, d_e = 9]\!]_2$	$[\![15, 7, 5]\!]_{2^3}$	8
	$[\![75, 33, d_e = 9]\!]_2$	$[\![19, 11, 5]\!]_{2^3}$	8
5	$[\![47, 6, d_e = 11]\!]_2$	$[\![12, 2, 6]\!]_{2^3}$	10
	$[\![63, 18, d_e = 11]\!]_2$	$[\![16, 6, 6]\!]_{2^3}$	10
	$[\![71, 24, d_e = 11]\!]_2$	$[\![18, 8, 6]\!]_{2^3}$	10
	$[\![75, 27, d_e = 11]\!]_2$	$[\![19, 9, 6]\!]_{2^3}$	10
	$[\![79, 30, d_e = 11]\!]_2$	$[\![20, 10, 6]\!]_{2^3}$	9
	$[\![83, 33, d_e = 11]\!]_2$	$[\![21, 11, 6]\!]_{2^3}$	10
	$[\![91, 39, d_e = 11]\!]_2$	$[\![23, 13, 6]\!]_{2^3}$	10
	$[\![99, 45, d_e = 11]\!]_2$	$[\![25, 15, 6]\!]_{2^3}$	10
	$[\![103, 48, d_e = 11]\!]_2$	$[\![26, 16, 6]\!]_{2^3}$	10
	$[\![107, 51, d_e = 11]\!]_2$	$[\![27, 17, 6]\!]_{2^3}$	10
	$[\![111, 54, d_e = 11]\!]_2$	$[\![28, 18, 6]\!]_{2^3}$	10
6	$[\![95, 36, d_e = 13]\!]_2$	$[\![24, 12, 7]\!]_{2^3}$	12
	$[\![99, 39, d_e = 13]\!]_2$	$[\![25, 13, 7]\!]_{2^3}$	11
	$[\![103, 42, d_e = 13]\!]_2$	$[\![26, 14, 7]\!]_{2^3}$	11
	$[\![107, 45, d_e = 13]\!]_2$	$[\![27, 15, 7]\!]_{2^3}$	11
	$[\![111, 48, d_e = 13]\!]_2$	$[\![28, 16, 7]\!]_{2^3}$	11
	$[\![115, 51, d_e = 13]\!]_2$	$[\![29, 17, 7]\!]_{2^3}$	12
	$[\![119, 54, d_e = 13]\!]_2$	$[\![30, 18, 7]\!]_{2^3}$	12
	$[\![123, 57, d_e = 13]\!]_2$	$[\![31, 19, 7]\!]_{2^3}$	12
	$[\![127, 60, d_e = 13]\!]_2$	$[\![32, 20, 7]\!]_{2^3}$	11
7	$[\![127, 54, d_e = 15]\!]_2$	$[\![32, 18, 8]\!]_{2^3}$	13
6	$[\![79, 16, d_e = 13]\!]_2$	$[\![16, 4, 7]\!]_{2^4}$	12
7	$[\![119, 40, d_e = 15]\!]_2$	$[\![24, 10, 8]\!]_{2^4}$	14

Table 3.3: Concatenated codes $[\![n, k, d_e]\!]_2$ for the AD channel based on non-QMDS outer qubit codes.

t	concatenated code	outer code	d_{lb}
3	$[\![19, 2, d_e = 7]\!]_2$	$[\![10, 2, 4]\!]_2$	6
	$[\![23, 4, d_e = 7]\!]_2$	$[\![12, 4, 4]\!]_2$	6
4	$[\![21, 1, d_e = 9]\!]_2$	$[\![11, 1, 5]\!]_2$	7
	$[\![31, 4, d_e = 9]\!]_2$	$[\![16, 4, 5]\!]_2$	8
	$[\![35, 6, d_e = 9]\!]_2$	$[\![18, 6, 5]\!]_2$	8
5	$[\![31, 2, d_e = 11]\!]_2$	$[\![16, 2, 6]\!]_2$	10
	$[\![39, 4, d_e = 11]\!]_2$	$[\![20, 4, 6]\!]_2$	9
	$[\![41, 5, d_e = 11]\!]_2$	$[\![21, 5, 6]\!]_2$	9
	$[\![47, 6, d_e = 11]\!]_2$	$[\![24, 6, 6]\!]_2$	10
	$[\![55, 12, d_e = 11]\!]_2$	$[\![28, 12, 6]\!]_2$	10
6	$[\![33, 1, d_e = 13]\!]_2$	$[\![17, 1, 7]\!]_2$	11
	$[\![47, 3, d_e = 13]\!]_2$	$[\![24, 3, 7]\!]_2$	11
	$[\![49, 5, d_e = 13]\!]_2$	$[\![25, 5, 7]\!]_2$	11
	$[\![59, 8, d_e = 13]\!]_2$	$[\![30, 8, 7]\!]_2$	12
	$[\![63, 10, d_e = 13]\!]_2$	$[\![32, 10, 7]\!]_2$	12
7	$[\![47, 1, d_e = 15]\!]_2$	$[\![24, 1, 8]\!]_2$	13
	$[\![51, 4, d_e = 15]\!]_2$	$[\![26, 4, 8]\!]_2$	12
	$[\![59, 5, d_e = 15]\!]_2$	$[\![30, 5, 8]\!]_2$	13
	$[\![63, 6, d_e = 15]\!]_2$	$[\![32, 6, 8]\!]_2$	14
	$[\![65, 7, d_e = 15]\!]_2$	$[\![33, 7, 8]\!]_2$	13
	$[\![67, 8, d_e = 15]\!]_2$	$[\![34, 8, 8]\!]_2$	14
	$[\![71, 12, d_e = 15]\!]_2$	$[\![36, 12, 8]\!]_2$	14
8	$[\![49, 1, d_e = 17]\!]_2$	$[\![25, 1, 9]\!]_2$	13
	$[\![53, 3, d_e = 17]\!]_2$	$[\![27, 3, 9]\!]_2$	13
	$[\![69, 4, d_e = 17]\!]_2$	$[\![35, 4, 9]\!]_2$	15
	$[\![101, 19, d_e = 17]\!]_2$	$[\![51, 19, 9]\!]_2$	16
9	$[\![55, 2, d_e = 19]\!]_2$	$[\![28, 2, 10]\!]_2$	14
	$[\![71, 3, d_e = 19]\!]_2$	$[\![36, 3, 10]\!]_2$	15
	$[\![105, 17, d_e = 19]\!]_2$	$[\![53, 17, 10]\!]_2$	17
10	$[\![57, 1, d_e = 21]\!]_2$	$[\![29, 1, 11]\!]_2$	15
	$[\![81, 3, d_e = 21]\!]_2$	$[\![41, 3, 11]\!]_2$	18
	$[\![95, 4, d_e = 21]\!]_2$	$[\![48, 4, 11]\!]_2$	20
	$[\![97, 5, d_e = 21]\!]_2$	$[\![49, 5, 11]\!]_2$	19
11	$[\![83, 2, d_e = 23]\!]_2$	$[\![42, 2, 12]\!]_2$	19
	$[\![97, 3, d_e = 23]\!]_2$	$[\![49, 3, 12]\!]_2$	21
	$[\![99, 4, d_e = 23]\!]_2$	$[\![50, 4, 12]\!]_2$	20
	$[\![107, 8, d_e = 23]\!]_2$	$[\![54, 8, 12]\!]_2$	19
12	$[\![85, 1, d_e = 25]\!]_2$	$[\![43, 1, 13]\!]_2$	21
	$[\![101, 3, d_e = 25]\!]_2$	$[\![51, 3, 13]\!]_2$	21
	$[\![113, 5, d_e = 25]\!]_2$	$[\![57, 5, 13]\!]_2$	21
13	$[\![103, 2, d_e = 27]\!]_2$	$[\![52, 2, 14]\!]_2$	21
	$[\![115, 4, d_e = 27]\!]_2$	$[\![58, 4, 14]\!]_2$	22
	$[\![125, 6, d_e = 27]\!]_2$	$[\![63, 6, 14]\!]_2$	23
14	$[\![105, 1, d_e = 29]\!]_2$	$[\![53, 1, 15]\!]_2$	21
	$[\![117, 3, d_e = 29]\!]_2$	$[\![59, 3, 15]\!]_2$	23
15	$[\![119, 2, d_e = 31]\!]_2$	$[\![60, 2, 16]\!]_2$	23
16	$[\![121, 1, d_e = 33]\!]_2$	$[\![61, 1, 17]\!]_2$	25

Table 3.4: Concatenated codes $[\![n, k, d_e]\!]_2$ for the AD channel based on non-QMDS outer codes with qudit dimension 4 and 8.

t	concatenated code	outer code	d_{lb}
4	$[\![41, 8, d_e = 9]\!]_2$	$[\![14, 4, 5]\!]_{2^2}$	8
	$[\![44, 10, d_e = 9]\!]_2$	$[\![15, 5, 5]\!]_{2^2}$	8
	$[\![47, 12, d_e = 9]\!]_2$	$[\![16, 6, 5]\!]_{2^2}$	8
5	$[\![50, 10, d_e = 11]\!]_2$	$[\![17, 5, 6]\!]_{2^2}$	9
6	$[\![44, 2, d_e = 13]\!]_2$	$[\![15, 1, 7]\!]_{2^2}$	12
	$[\![56, 6, d_e = 13]\!]_2$	$[\![19, 3, 7]\!]_{2^2}$	12
	$[\![59, 8, d_e = 13]\!]_2$	$[\![20, 4, 7]\!]_{2^2}$	12
	$[\![74, 14, d_e = 13]\!]_2$	$[\![25, 7, 7]\!]_{2^2}$	12
	$[\![77, 16, d_e = 13]\!]_2$	$[\![26, 8, 7]\!]_{2^2}$	12
	$[\![80, 18, d_e = 13]\!]_2$	$[\![27, 9, 7]\!]_{2^2}$	12
7	$[\![104, 26, d_e = 15]\!]_2$	$[\![35, 13, 8]\!]_{2^2}$	14
	$[\![107, 28, d_e = 15]\!]_2$	$[\![36, 14, 8]\!]_{2^2}$	14
8	$[\![74, 6, d_e = 17]\!]_2$	$[\![25, 3, 9]\!]_{2^2}$	15
	$[\![92, 14, d_e = 17]\!]_2$	$[\![31, 7, 9]\!]_{2^2}$	16
	$[\![95, 16, d_e = 17]\!]_2$	$[\![32, 8, 9]\!]_{2^2}$	16
	$[\![110, 22, d_e = 17]\!]_2$	$[\![37, 11, 9]\!]_{2^2}$	16
9	$[\![77, 4, d_e = 19]\!]_2$	$[\![26, 2, 10]\!]_{2^2}$	16
	$[\![98, 10, d_e = 19]\!]_2$	$[\![33, 5, 10]\!]_{2^2}$	18
	$[\![101, 12, d_e = 19]\!]_2$	$[\![34, 6, 10]\!]_{2^2}$	17
	$[\![104, 14, d_e = 19]\!]_2$	$[\![35, 7, 10]\!]_{2^2}$	17
	$[\![110, 18, d_e = 19]\!]_2$	$[\![37, 9, 10]\!]_{2^2}$	17
	$[\![113, 20, d_e = 19]\!]_2$	$[\![38, 10, 10]\!]_{2^2}$	18
	$[\![116, 22, d_e = 19]\!]_2$	$[\![39, 11, 10]\!]_{2^2}$	18
10	$[\![95, 4, d_e = 21]\!]_2$	$[\![32, 2, 11]\!]_{2^2}$	20
	$[\![98, 6, d_e = 21]\!]_2$	$[\![33, 3, 11]\!]_{2^2}$	19
	$[\![101, 8, d_e = 21]\!]_2$	$[\![34, 4, 11]\!]_{2^2}$	19
	$[\![104, 10, d_e = 21]\!]_2$	$[\![35, 5, 11]\!]_{2^2}$	18
	$[\![107, 12, d_e = 21]\!]_2$	$[\![36, 6, 11]\!]_{2^2}$	18
	$[\![116, 14, d_e = 21]\!]_2$	$[\![39, 7, 11]\!]_{2^2}$	20
	$[\![119, 16, d_e = 21]\!]_2$	$[\![40, 8, 11]\!]_{2^2}$	20
	$[\![122, 18, d_e = 21]\!]_2$	$[\![41, 9, 11]\!]_{2^2}$	20
	$[\![125, 20, d_e = 21]\!]_2$	$[\![42, 10, 11]\!]_{2^2}$	20
	$[\![128, 22, d_e = 21]\!]_2$	$[\![43, 11, 11]\!]_{2^2}$	20
11	$[\![116, 10, d_e = 23]\!]_2$	$[\![39, 5, 12]\!]_{2^2}$	21
	$[\![119, 12, d_e = 23]\!]_2$	$[\![40, 6, 12]\!]_{2^2}$	21
	$[\![122, 14, d_e = 23]\!]_2$	$[\![41, 7, 12]\!]_{2^2}$	21
	$[\![125, 16, d_e = 23]\!]_2$	$[\![42, 8, 12]\!]_{2^2}$	21
12	$[\![116, 6, d_e = 25]\!]_2$	$[\![39, 3, 13]\!]_{2^2}$	22
	$[\![119, 8, d_e = 25]\!]_2$	$[\![40, 4, 13]\!]_{2^2}$	22
	$[\![128, 10, d_e = 25]\!]_2$	$[\![43, 5, 13]\!]_{2^2}$	23
8	$[\![107, 21, d_e = 17]\!]_2$	$[\![27, 7, 9]\!]_{2^3}$	16

Chapter 4

Capacities of Very Noisy Quantum Channels

4.1 Introduction

In the previous chapters the focus has been on constructing quantum codes which correct particular error models through two different methods; CWS construction and concatenation. In this chapter I apply these codes to the old and very hard problem of determining when a noisy channel has capacity.

If we have two parties, Alice and Bob, trying to communicate over a shared channel, a resource with which they are able to send messages, the fundamental question is always "are Alice and Bob capable of communicating". When the messages they wish to send are classical and the channel they are using is also classical we have a situation which was famously formalized and solved by Claude E. Shannon in his ground breaking work [63] and is considered the foundation of modern information theory.

The result I refer to is the channel coding theorem, which in brief states

that if X is the set of messages which Alice wishes to send, and Y are the set of messages that Bob receives under action of a noisy channel $\mathcal{N}$ then the maximum rate of communication, the capacity of the channel, is given by the maximum mutual information of the variables X and Y [63]. In the case of classical channel capacity, the mutual information completely characterizes the maximum rate at which information can be faithfully communicated.

An important question is what is the analogue case for quantum channels. However, as it stands, this question is not well defined since we haven't defined what kind of information Alice wishes to send to Bob. For instance Alice and Bob might want to send classical information using a quantum channel, in which case the capacity would be measured in bits of information. Additional constraints can also be considered such as requiring that Alice and Bob can generate a bit-string which no eavesdropper could reproduce, known as the private channel capacity. Finally various types of additional resources might be available such as a classical channel Alice and Bob can communicate with, or shared entangled pairs.

However, one situation is in a sense more fundamental and as such is what is usually referred to as the quantum capacity, and that is the maximum rate at which Alice can faithfully transmit qubits to Bob. As a side channel we allow forward classical communication which is known not to increase the capacity [4]. If a quantum channel can be used to faithfully trasmit quantum states, then this channel can also be used to generate shared maximally entangled pairs and the two problems are actually equivalent. To see this consider that if you can faithfully transmit one qubit, you can transmit one half of a maximally entangled Bell state producing one qubit of entanglement. In the other direction if you can produce one pure Bell state, it can be used with quantum teleportation to transmit a single

arbitrary qubit reliably. An intuitive definition of quantum channel capacity is as follows:

Definition 4.14. *The capacity of a quantum channel is the highest rate at which quantum states, measured in qubits, can be transmitted from sender to receiver with multiple independent uses of the noisy quantum channel*

This is the fundamental quantum channel behaviour because the ability to generate pure entangled pairs is sufficient for producing almost any other quantum behaviour including the other channel types described. For example classical information can be transmitted using either via straight forward encoding or super dense coding [5] and private bit-strings can be produced from shared Bell states using BB84 [3].

The question of importance is then: "when does a quantum channel Λ have capacity $Q(\Lambda) > 0$". One might hope that, like the mutual information for the classical channel, there might be a simple characterization of the quantum capacity. The quantum analogue of the mutual information is the coherent information which is defined as:

Definition 4.15. *The coherent information of a channel Λ is given by [62]*

$$I_c(\Lambda) = max_\rho S(\Lambda(\rho)) - S(I \otimes \Lambda(|\Psi\rangle\langle\Psi|))$$

where $|\Psi\rangle\langle\Psi|$ is some purification of ρ and S is the Von-Neumann Entropy given by $S(\rho) = -Tr(\rho \ln(\rho))$.

It has been shown that the coherent information of channel is the capacity of a single use of the channel $Q^{(1)}(\Lambda) = I_c(\Lambda)$[48, 64, 18]. In fact the capacity of

n uses of a channel Q^n is given by $Q^n = \frac{1}{n} I_c(\Lambda^{\otimes n})$. However $Q^{(n)}$ is only a lower bound on the channel capacity $Q \geq Q^{(n)}$.

Theorem 4.10 (LSD)**.** *The proper characterization of Q is given by[18]:*

$$Q(\Lambda) = \lim_{n\to\infty} Q^{(n)} = \lim_{n\to\infty} \frac{1}{n} I_c(\Lambda^{\otimes n}) \tag{4.1}$$

This result, the LSD theorem named after Lloyd, Shor, and Devetak who developed it, is the operational definition of quantum capacity, and states that in general, only in the limit of an infinite number of channel uses can we determine the true capacity of a channel. This means there exist channels for which n uses of the channel has a capacity greater than n times the capacity of a single channel use, a fundamentally quantum behaviour known as the super-additivity of the quantum channel capacity which is not found in classical capacity theory.

The capacity Q characterizes the rate at which qubits can be transmitted over the channel. However, sometimes the more interesting question is not what the precise value of Q is but rather determining if Q is non-zero since as long as Q is positive Alice and Bob can successfully communicate. If the rate is really low but not zero it might take a large number of channel uses to produce even one qubit but it can be done. On the other hand, if $Q = 0$ then no communication is possible with the available resources.

In determining whether a channel has capacity or not we can use some bounds. Since the coherent information is an achievable lower bound on the capacity, then any channel with non-zero coherent information must have non-zero capacity. This is actually the case for any $Q^{(n)}$ as they all act as lower bounds for Q. However of these $Q^{(1)}$ is the easiest to calculate.

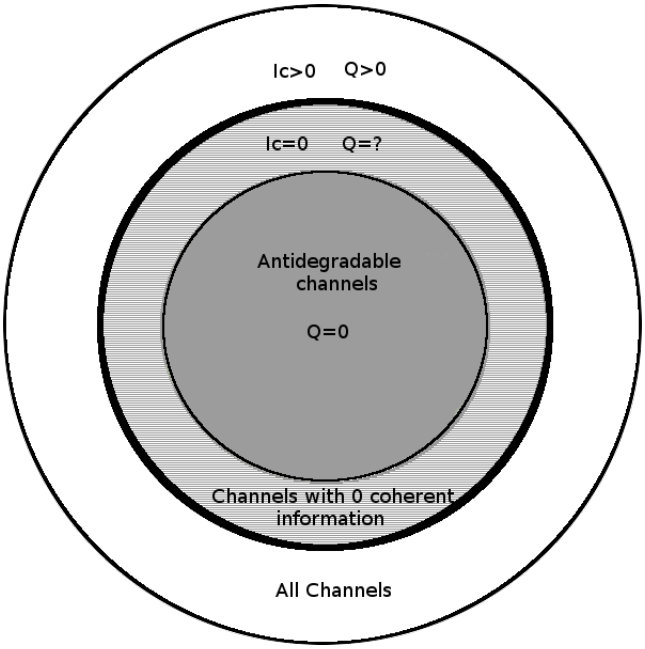

Figure 4.1: Diagram showing the gap between channels with zero coherent information and those with known zero capacity (middle ring)

On the other side, if a channel is anti-degradable, then it must have zero capacity. Its easy to see this because if an anti-degradable had capacity, then Alice and Bob are capable of producing pure, maximally entangled pairs. However the environment is able to simulate the channel to Bob, so Eve could also produce a pure, maximally entangled pair with Alice at the same time which breaks the monogamy of entanglement.

These results leave us with a picture like Fig: 4.1 where we see that there exists in general a gap. Any channel with $I_c > 0$ is known to have positive capacity, however there are channels with $I_c = 0$ which are not anti-degradable. The question of concern in this chapter is what can we say about channels in this gap.

There exist results showing that for certain classes of Pauli channels there do exist channels in the gap with non-zero capacity[66, 69, 42, 19]. Finding the same for non-Pauli channels is far more difficult to do, primarily due to the difficulty of calculating the coherent information on non-Pauli channels. The traditional way to

find results for non-Pauli channels is to "twirl" them to Pauli channels and use the corresponding results of Pauli channels. I will explore and expand this idea however I will show that there exist non-Pauli channels with zero coherent information but non-zero capacity using effective channels construction rather than relying on twirling. Furthermore I can find positive results for channels on which twirling can not prove capacity. These are new examples of channels in this gap that have not before been studied.

4.2 Existing Proof Techniques

As mentioned in the introduction, there is no general characterization or method to determine if a channel Λ with zero coherent information has non-zero capacity. However, there are several methods to prove that particular channels in this gap either have or don't have capacity.

An excellent example of this is the result that degradable channels have additive capacity [18]. So for degradable channels it is known that $Q(\Lambda) = Q^{(1)}(\Lambda)$. Any degradable channel in the gap has zero coherent information and as a result zero capacity.

The structure of degradable channels have been studied extensively and it has been shown that any qubit channel with 2 or less Kraus operators is either degradable or anti-degradable[74]. It was later shown that no qubit channel with greater than 2 Kraus operators is degradable [15]. For degradable channels there is no gap between where $I_c = 0$ and $Q = 0$. So when looking for examples of interesting qubit channels, any channel with Choi rank less than 3 can be excluded.

Outside the degradable channels, characterizations are far less general re-

lating only to specific channels or specific encodings. As a result several techniques have been developed to attempt to answer the core question for particular classes of channels. These techniques fall roughly into two categories; bounds and direct proofs through explicitly constructed codes.

4.2.1 Upper and Lower Bounds

Bound proofs typically rely on the fact that the capacity of a composition of channels is less than or equal to the capacity of either channel alone:

$$Q^{(n)}(\mathcal{N} \circ \Gamma) \leq Q^{(n)}(\Gamma) \tag{4.2}$$

This can be used to prove an upper bound on the coherent information in the following way. Assume your channel Λ can be represented as the composition of two other channels $\Lambda = \mathcal{N} \circ \Gamma$. Using eq.4.2 we can say that $I_c(\Lambda) \leq I_c(\Gamma)$. If you can find a pair of channels $\mathcal{N}$ and Γ that satisfies the above, and furthermore you know properties of Γ such as it having zero coherent information, this is a way to prove your channel Λ also has zero coherent information. If Γ is also degradable then it is called an additive extension of your channel and as a result implies your channel does not not have capacity [70].

Lower bounds can be obtained in the opposite situation when $\mathcal{N} \circ \Lambda = \Gamma$. In these situations $Q(\Lambda) \geq Q(\Gamma)$, so the capacity of Γ is a lower bound for the channel Λ. As a consequence, if the capacity of Γ is non-zero either by having non-zero coherent information or by other means then you can extend those results to Λ.

A closely related technique is that of “twirling” a channel.

Definition 4.16. *The $\mathcal{V}$-twirl $\mathcal{T}_{\mathcal{V}}$ of a channel $\mathcal{N}$ is described by:*

$$\mathcal{T}_V(\mathcal{N}) = \frac{1}{|\mathcal{V}|} \sum_{V \in \mathcal{V}} V^\dagger \mathcal{N}(V \rho V^\dagger) V.$$

Here $\mathcal{V}$ is some set of operators.

Even though this is not exactly a composition of channels, it is known that $Q^{(n)}(\mathcal{T}_{\mathcal{V}}(\Lambda)) \leq Q^{(n)}(\Lambda)$ [51]. Two of the more useful $\mathcal{V}$-twirls are the unitary and Pauli twirls. In the first, the set $\mathcal{V}$ is that of all unitary operators, the effect of which is to map any arbitrary channel to the depolarizing channel with same fidelity. In this case since the set of all unitary operators is infinite, the expression is actually an integral using the Haar measure [35] instead of a sum.

For example, if we use Mary Beth Ruskai's characterization of a general channel [58] it has six parameters and can be represented as a vector $[p_x, p_y, p_z, t_z, t_y, t_z]$. As described in chapter 1, the p_x, p_y, p_z correspond to Pauli component of the channel where as the t_x, t_y, t_z represent the non-Pauli shift component. The effect of the unitary twirl $\mathcal{T}_U$ is then to map a channel with parameters $[p_x, p_y, p_z, t_x, t_y, t_z]$ to one with parameters $[g, g, g, 0, 0, 0]$ where $g = \frac{p_x+p_y+p_z}{3}$. Notice the non-Pauli component has been wiped out by the twirl, this is a valuable behaviour as we can then use results on Pauli channels for non-Pauli channels. Unfortunately the unitary twirl is of limited practical usefulness as the depolarizing channel is the most noisy channel with a particular fidelity. So most channels where you can prove capacity from the unitary twirl technique already have non-zero coherent information and as a result can already be proved via LSD directly.

A more useful $\mathcal{V}$-Twirl is the Pauli Twirl. This is the twirling operation where $\mathcal{V} = \{I, X, Y, Z\}$.

Lemma 4.5. *The action of the Pauli twirl on the a qubit channel* Λ *with parameters* $[p_x, p_y, p_z, t_x, t_y, t_z]$ *is a twirled channel with parameters* $[p_x, p_y, p_z, 0, 0, 0]$.

Proof. A general qubit channel Λ can be expressed as:

$$\Lambda(\rho) = \sum_{i,j} \alpha_{ij} E_i \rho E_j^\dagger \tag{4.3}$$

where $E_i, E_j \in \{I, X, Y, Z\}$. Allow the numbering $E_0 = I$, $E_1 = X$, $E_2 = Y$ and $E_3 = Z$. We know [58] that $\alpha_{ii} = p_i$ except for $\alpha_{00} = 1 - (p_x + p_y + p_z)$.

We can separate these terms into two parts, the terms where $i = j$ and those where $i \neq j$. If we examine the action of the Pauli twirl on terms where $i = j$:

$$\frac{1}{|\mathcal{V}|} \sum_{V \in \{I,X,Y,Z\}} V^\dagger E_i V \rho V^\dagger E_i^\dagger V \tag{4.4}$$

$$= \frac{1}{4} \sum_{V \in \{I,X,Y,Z\}} E_i V^\dagger V \rho V^\dagger V E_i^\dagger \tag{4.5}$$

$$= E_i \rho E_i^\dagger \tag{4.6}$$

This shows that the Pauli channel components are left unchanged as for each V it will either commute or anti commute with both sides, cancelling out any phase.

The terms from the off diagonals will all be cancelled by similar reasoning. For any pair of E_i, E_j, where $i \neq j$ and i and j are not 0, two of the elements of $\mathcal{V}$ will anti-commute with one element of the pair and anti commute with the other element ($V = E_i$ and $V = E_j$) and two will commute with both or anti-commute with both. As a result these terms all get cancelled out. In the cases where one of the pair is I then all elements of $\mathcal{V}$ will commute with it but half will anti-commute

with the second element of the pair so the result is the same.

Therefore the parameters of the twirled channel are given by $[p_x, p_y, p_z, 0, 0, 0]$

□

The Pauli-twirl introduces less additional noise than the unitary twirl and therefore is more useful for finding lower bounds than the unitary twirl. As all channels can be represented as a Pauli part Λ_P and a non Pauli part Λ_{NP}, the Pauli twirl lets us say that if Λ_P has capacity then Λ also has capacity.

4.2.2 Hashing code

In the case of Pauli channels, the coherent information has a simple form $Q^{(1)}(\Lambda) = 1 - S(\sigma_\lambda)$ where S is the Von-Neumann entropy and σ_λ is the Choi matrix representation of the channel Λ. This form makes determining when a channel does not have coherent information exceedingly simple in comparison to the non-Pauli case.

An additional nice feature of Pauli channels is that there exists a quantum analogue to Shannon's random code [63], called the hashing code [4] which achieves the coherent information as a rate. It is heavily inspired by Shannon's random code and we review it here at a high level for the purpose of proving a couple theorems in the next section.

In the hashing code, Alice prepares a set of $|\phi^+\rangle \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ maximally entangled pure Bell states and sends one half of each to Bob through a noisy Pauli channel:

$$
\begin{aligned}
I \otimes \Lambda(|\phi^+\rangle\langle\phi^+|) &= p_I |\phi^+\rangle\langle\phi^+| + p_X I \otimes X |\phi^+\rangle\langle\phi^+| I \otimes X \\
&\quad + p_Y I \otimes Y |\phi^+\rangle\langle\phi^+| I \otimes Y + p_Z I \otimes Z |\phi^+\rangle\langle\phi^+| I \otimes Z \\
&= p_I |\phi^+\rangle\langle\phi^+| + p_X |\psi^+\rangle\langle\psi^+| + p_Y |\psi^-\rangle\langle\psi^-| + p_Z |\phi^-\rangle\langle\phi^-|
\end{aligned}
$$

The state shared between Alice and Bob is exactly the same as the Choi matrix representation of the channel Λ, σ_Λ.

At this point, to Alice and Bob their shared state a unknown sequence consisting of four characters for each of the shared Bell pairs. There are a minimum of 2^{2n} binary strings required to represent all possible sequences of shared bell pairs. However, using a typical strings argument [73], there are $2^n H(p) = 2^n S(\sigma_\Lambda)$ "typical sequences" which will contain almost all the probability of being generated.

Using only forward classical communication and local operations, Alice and Bob can collect the parity of any subset of the $2n$ bit string into one Bell pair. By measuring this Bell pair, Bob can rule out half of the remaining bit-strings. After k such steps, the remaining number of possible shared sequences is $\frac{2^n S(\sigma_\Lambda)}{2^k}$ which is equal to 1, or certainty about the sequence when $k = nS(\sigma_\Lambda)$. Since each of these k steps sacrifices one Bell pair, the total remaining shared Bell pairs is $n - nS(\sigma_\Lambda) = n(1 - S(\sigma_\Lambda))$. So when $1 - S(\sigma_\Lambda) > 0$ the hashing code gives capacity on a channel.

4.2.3 Effective channels and code concatenation

In 1996 Shor demonstrated that there exist channels which have capacity even when the hashing code does not. This was achieved by using a one-way entanglement purification protocol on the output of the physical channels. The outputs from the purification protocol were then used in the hashing code. By doing this, Shor showed it was possible to obtain capacity from the depolarizing channel for fidelities less than $f = 0.8107$, where the coherent information becomes zero.

Since one-way EPP is equivalent to error correction codes [4] Shor's original process can be reformulated as a concatenated code. The inner is a stabilizer code with encoding procedure:

$$|0\rangle \rightarrow |0\rangle^{\otimes n}$$

$$|1\rangle \rightarrow |1\rangle^{\otimes n}$$

This family of codes is commonly know as repetition codes or cat codes.

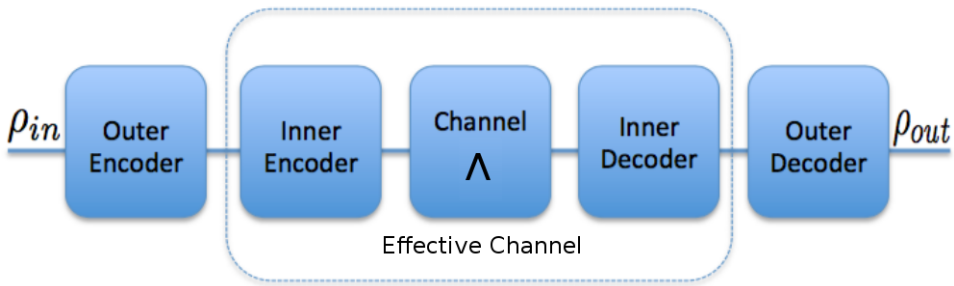

Figure 4.2: Diagram of effective channel and where it sits in larger error correcting framework. The effective channel is the action of the inner code acting on the physical channel. The outer code will be code achieving coherent information rate as from LSD[48, 64, 18].

As shown in figure 4.2, this inner code is then concatenated with the hash-

ing code as an outer code, for which the properties of achieving capacity are well understood. In this way quantum codes are used to protect the coherence of a state being sent through the physical channel. In some circumstances this concatenation is enough to achieve capacity on channels for which the hashing code alone not achieve capacity.

In order to analyze the effect of the inner quantum code, it is useful to introduce the "effective channel" sometimes also called the effective noise model [42]. The effective channel $\tilde{\Lambda}$ is produced by applying a quantum code Q to N uses of a physical channel Λ. Detection-correction QECCs will produce measurement results known as the syndrome i. For each i there is a syndrome-conditioned effective channel $\tilde{\Lambda}_i$.

Since our goal is to produce pure entangled pairs (for use in other interesting quantum processes) these effective channels can also be described by their Choi matrix representations σ_{Λ_i}. In fact since we are concatenating these inner codes with the hashing outer code, the effective channels will be used to send half of a Bell pair in which case the Choi matrices σ_{Λ_i} represent the actual quantum state output by the measurement conditioned effective channel. A formal definition of effective channel will be presented in the next section.

The question is then, given an effective channel $\tilde{\Lambda}$, what is the condition for it to have capacity? In his 1996 paper, Peter Shor uses:

$$\sum_i p_i(1 - S(\sigma_{\Lambda_i})) > 0 \tag{4.7}$$

as the condition to determine when an effective channel which produces output states σ_{Λ_i} with probability p_i.

Theorem 4.11. *If an effective Pauli channel produces k-qubit output states σ_{Λ_i} with probability p_i the rate of pure Bell pair production is $\sum_i p_i(k - S(\sigma_{\Lambda_i}))$*

Proof. When applying the hashing code to effective channels, instead of a single raw input state σ_Λ you have a set of Choi matrices σ_{Λ_i} representing the various effective channels conditioned on the measurement result i which occurs with probability p_i.

In order to determine the effectiveness of the hashing code on such effective channels we need to know how many likely strings there are. To determine this we can simplify the problem by noting that Bob has the measurement results i for each pair and can re-index the pairs such that every pair produced by the ith effective channel are together. This is not a physical change and since Alice's procedure does not (and can not) depend on the indices this is allowable.

The ith block will, in the large n limit, have a length of np_i. Within this block we have uniformity so we can use normal typical sequences for this sub sequence of pairs.

Since we are dealing with Pauli channels only at this point, for a k-qubit effective channel σ_{Λ_i} will be $4^k \times 4^k$ diagonal Choi matrices (in the Bell basis) containing the probability distribution of k-tuples of Bell-states being produced. This does not affect the formula for number of typical strings in form. Just consider each tuple as one "symbol". Therefore there are $2^{np_i S(\sigma_{\Lambda_i})}$ typical strings of these symbols.

The total number of typical strings for the complete sequence across all blocks is the product of all the sets of typical sequences for each block.

$$\prod_i 2^{np_i S(\sigma_{\Lambda_i})} = 2^{n \sum_i p_i S(\sigma_{\Lambda_i})}$$

$n\sum_i p_i S(\sigma_{\Lambda_i})$ is the number of Bell pairs which will need to be sacrificed in order to know with certainty which sequence is shared. However since each of the n symbols is actually a k-tuple of Bell states, there are kn Bell pairs shared. So after the hashing code is applied, $kn - n\sum_i p_i S(\sigma_{\Lambda_i}) = n(k - \sum_i p_i S(\sigma_{\Lambda_i}))$ shared Bell pairs remain. As a result, the hashing code guarantees capacity on effective channels where $\sum_i p_i(k - S(\sigma_{\Lambda_i})) > 0$.

□

An interesting by-product of the previous theorem is it gives a notion of negative information channels, which normally doesn't make sense since the maximization of coherent information always is at least 0. However in the context of effective channels some of the terms of the sum may actually be negative. What this indicates is that when those measurement values occur they produce effective channels which are so noisy they actually degrade the ability to produce pure Bell pairs from the rest of the less noisy effective channels. It does this by massively increasing the size of the set of typical sequences.

To be clear if $S(\sigma_{\Lambda_i})$ has a range of values then in some sense those with higher values will be degrading the effectiveness that could have been achieved if only the lowest entropy states could be used. However when it comes to the question of whether or not a channel has capacity, then any effective channel with $k - S(\sigma_{\Lambda_i}) > 0$ is contributing positively. Conversely, when this is not the case, the effective channel is so noisy that it actually can undo the capacity produced by the less noisy outputs. In this sense, they are providing a negative rate.

Using the same idea of re-indexing into blocks with the same measurement result you can actually prove a more general case for the coherent information and achievable rates on general channels.

Theorem 4.12. *An effective channel $\tilde{\Lambda}$, with measurement conditioned effective channels Λ_i occurring with probability p_i will have capacity when $max_\rho \sum_i p_i I_C(\Lambda_i, \rho) > 0$.*

Proof. Start by re-indexing into blocks of similar effective channels. Each block once again has np_i members in the large n limit. The coherent information is an achievable rate by LSD [48, 64, 18] with rate $I_C(\Lambda_i, \rho)$. Since Alice can't condition her actions on the measurement result i the ρ for each block must be the same. Therefore the number of qubits produced by the effective channels when maximized is $\max_\rho \sum_i np_i I_C(\Lambda_i, \rho) = n(\max_\rho \sum_i np_i I_C(\Lambda_i, \rho))$. Thus the generation rate of the effective channel is given by $\max_\rho \sum_i np_i I_C(\Lambda_i, \rho)$. □

4.3 Construction of Effective Channels from Codes

Now that we know when an effective channel has capacity, the next step is to be able to construct those effective channels from arbitrary quantum codes. Here I develop the theory around construction of the effective channels from Stabilizer codes.

Detection-Correction quantum error correcting schemes have many formulations from CSS codes and Stabilizer codes to CWS codes. Each has its own construction method but all of which satisfy the error correcting condition 1.2.

The error correcting condition is very useful when you have a specific set of errors you want to prove you can correct. But all detection-correction style codes have an alternative viewpoint in terms of subspaces and projectors into those subspaces. This view is the key to understanding the effect of a quantum code on the entirety of a channel which induces both correctable and uncorrectable errors.

Recall that a quantum error correcting code Q with parameters $((n, K))$ is a code where n qubits are used to encode K states. If K is a power of 2 such that $2^k = K$ then this code can encode k qubits in n qubits. It does this by encoding the states into a K dimensional subspace of the full 2^n dimensional complex Hilbert space. We denote this subspace as V_0 and there exists a projector into this subspace we can denote P_0.

The power of quantum error correcting codes comes when $2^n > K$. In particular it should be greater than some multiple of K. When this happens we can have multiple subspaces of equal dimension to V_0 which are all mutually orthogonal. We shall denote these as V_m and they have associated projectors P_m. For each of these subspaces, an error correcting code specifies a unique recovery operator F_m which will take any state in V_m to a state in V_0. A code can only define one such F_m per subspace V_m. If the recovery operator was not uniquely determined by the measurement result m then there would not be enough information to perform error correction.

The reason to look at the quantum code Q in this manner instead of focusing solely on the correctability of certain error sets is because it gives us a basis to examine the effect of the code not only on correctable errors, but also the action of the code on those errors which are not correctable.

4.3.1 Using CWS codes to construct Effective Channels

Moving out of the abstract discussion of code subspaces we use the CWS code framework to precisely describe how error correction produces the effective channel. The CWS formalism is used because it includes CSS codes, stabilizer codes, as well as many other interesting codes. Ultimately we will be restricting to

just stabilizer codes, but future work could potentially extend the process outlined here to the full class of CWS codes.

Recall, a CWS quantum error correcting code is specified by two objects. First the word stabilizer S, a subgroup of 2^n mutually commuting Pauli operators including the identity which stabilizes a unique n-qubit state $|S\rangle$. Second a set of K Pauli word operators $W = \{w_\ell\}_{\ell=0}^{K}$. The encoding subspace V_0 is given by $V_0 = \text{span}\{(w_\ell |S\rangle)\}$ and has an associated projector $P_0 = \sum_\ell w_\ell |S\rangle \langle S| w_\ell^\dagger$.

Finding the set of valid correction operators $F = \{F_m\}$ requires satisfying the error correction condition. Namely these operators must produce orthogonal error subspaces:

$$V_m = \text{span}\left\{(F_m^\dagger w_\ell |S\rangle)\right\}$$

which then can be mapped to V_0 by the correction operator F_m. Later we will discuss the choice of F further however it suffices to look at $F_m^\dagger$ as the set of correctable errors. Each V_m has an associated projector $P_m = F_m^\dagger P_0 F_m$.

The last piece of the picture is the encoding and decoding processes, however for CWS codes this is exceedingly simple to describe. If we assume that the protected space has a basis of $\{|i\rangle\}_{i=0}^{K}$ then the encoding process can be described by

$$C = \sum_\ell |w_\ell\rangle \langle \ell|$$

which will map from the protected state space into V_0. The decoding procedure of course need not be $C^\dagger$ however for simplicity we will assume it is. This assumption does not affect channel capacity calculations.

Given the above, we can describe the effective channel resultant from a CWS code acting upon the output of a physical channel.

Theorem 4.13. *Given a CWS code $\mathcal{Q}$ completely specified by W, S, F, V_m and C as described and a channel Λ which described in the Pauli operator basis has form:*

$$\Lambda(\rho) = \sum_{ij} \alpha_{ij} E_i \rho E_j^\dagger$$

The measurement conditioned effective channels Λ_m obtained after applying Q to Λ is given by:

$$\Lambda_m = \frac{1}{p(m)} \sum_{ij} \alpha_{ij} \tilde{E}_i \rho_{log} \tilde{E}_j^{\ \dagger}$$

Where $\tilde{E}_i = C^\dagger F_m E_i C$ and ρ_{log} is used to indicate the protected "logical" state so as to distinguish it from the general input state ρ or the physical state ρ_{phys}

Proof. In a detection-correction QECC such as the CWS codes, there are four primary steps, your state ρ_{log} is encoded into a state ρ_{phys} with support on the V_0 subspace:

$$\rho_{phys} = C \rho_{log} C^\dagger$$

This state is then passed through the channel Λ to produce:

$$\Lambda(\rho_{phys}) = \sum_{ij} \alpha_{ij} E_i C \rho_{log} C^\dagger E_j^\dagger$$

The detection-correction process involves projecting into the error subspaces. If measurement m is produced then after measurement the state will be in error subspace V_m which can be corrected by applying F_m. The output of which is given by:

$$\sum_{ij} \alpha_{ij} F_m P_m E_i C \rho_{log} C^\dagger E_j^\dagger P_m^\dagger F_m^\dagger$$

and finally, now that the state resides in the V_0 subspace the decoding procedure can be applied giving a description of the measurement conditioned effective channel (after normalization):

$$\begin{aligned}
\Lambda_m &= \frac{1}{p(m)} \sum_{ij} \alpha_{ij} C^\dagger F_m P_m E_i C \rho_{log} C^\dagger E_j^\dagger P_m^\dagger F_m^\dagger C && (4.8)\\
&= \frac{1}{p(m)} \sum_{ij} \alpha_{ij} C^\dagger P_0 F_m E_i C \rho_{log} C^\dagger E_j^\dagger F_m^\dagger P_0^\dagger C && (4.9)\\
&= \frac{1}{p(m)} \sum_{ij} \alpha_{ij} C^\dagger F_m E_i C \rho_{log} C^\dagger E_j^\dagger F_m^\dagger C && (4.10)
\end{aligned}$$

Here the second line follows because $P_m = F_m^\dagger P_0 F_m$ and the third line follows because the coding and decoding operations absorb the P_0

By defining the effective (logical) errors as:

$$\tilde{E}_i = C^\dagger F_m E_i C$$

and substituting in we have our result. □

4.3.2 Constructing the Effective Channel

Since this has been defined for arbitrary channels, the E_i can be any n-qubit Pauli operator. In essence we want to define an operator mapping from physical errors E_i to its logical effective error $\tilde{E}_i$.

Definition 4.17. *The operator map is given by* $M(E, m) = C^\dagger F_m E C$

One property we would like this mapping to have is $M(E, m) = 0$ for all but one measurement value m. This would be a beneficial feature to have as it means

the subspace $EV_0E^\dagger$ which E rotates to overlaps one and only one V_m and in fact $EV_0E^\dagger = V_m$. If this were not the case, then the measurement result is dependent not only on the error E but also on the input state ρ_{phys}. The following theorem states when we are in this situation.

Theorem 4.14. *For a given error E, $M(E,m) = 0$ for all but one m if and only if W is a group.*

Proof. If W is a group, then $WS = \{w \cdot s, W \in W, s \in S\}$ is also a group and a subgroup of the full Pauli group. The left cosets of WS in the Pauli group mod phases forms a partition with every E belonging to exactly one coset.

Each error E produces a subspace given by span $\{Ew\ell\,|S\rangle\}$. Furthermore if E_1 and E_2 are members of the same coset then $E_1 = E_2 \cdot w$ for some $w \in WS$ and therefore each coset corresponds to exactly one subspace. Any representative element can be chosen as the correction operation F_m which will rotate back to V_0. So $M(E,m)$ will only be non-zero for the value of m corresponding to the coset of $F_m^\dagger$.

If W is not a group, then it is not a group due to lack of closedness since associativity is inherited from the Pauli operators, WS can always be chosen to include the identity [14], and all Pauli operators are self inverse. If WS is not closed, then there is some element of $E \in WS$ such that $E \cdot WS \neq WS$ but at the same time $E \cdot WS \cap WS \neq \emptyset$. Similar to the first half, both $E \cdot W$ and WS correspond to subspaces, however now these subspaces are no longer orthogonal. As such when projected into the V_m subspaces, V_E with have non-zero overlap with V_0 and some other subspace resulting in $M(E,m)$ being non-zero for multiple values of m. $\square$

The case where W is not a group implies that the CWS code can only produce conditional phase operators on the logical space (through multiplication of elements in S). However when W is a group it is known that the CWS code is a stabilizer code [14, 10] which is still a very rich field to draw from and will be the focus moving forward.

Knowing which V_m an error E maps to is therefore a straight forward matter of determining which coset it belongs to. Determining what the action of the effective error is the second part to constructing the operator map. However since E belongs to the same coset as its correction operator, we know that $E = aF_m^\dagger R$ where $R \in WS$ is some remainder operator and a is some complex phase. It is clear to see that after correction:

$$F_m E = aF_m F_m^\dagger R = aR$$

This allows us to define the action of any arbitrary error by the actions of the operators in WS by $M(E, m) = aC^\dagger RC$. Written in a more useful form we have the lemma:

Lemma 4.6. *If $R \in WS$ and $F_m^\dagger \in F$ then $M(F_m^\dagger \cdot R, m) = aM(R, 0)$*

This allows us to reduce the problem of finding the mapping M of every Pauli error to finding the logical action of the elements of WS. We can further reduce the problem with the following observations of the behaviours of M:

Theorem 4.15. *The operator mapping M is a group homomorphism on WS that is if $R_i, R_j \in WS$ then $M(R_iR_j, 0) = M(R_i, 0)M(R_j, 0)$*

Proof. The mapping $M(R, 0) = C^\dagger RC$ where $C = \sum_\ell |w_\ell\rangle \langle \ell|$. Therefore:

$$\begin{aligned} M(R,0) &= C^\dagger RC \\ &= \sum_{k,\ell} |k\rangle \langle w_k| R |w_\ell\rangle \langle \ell| \\ &= \sum_{k,\ell} \langle S| w_k R w_\ell |S\rangle |k\rangle \langle \ell| \end{aligned}$$

We can prove the theorem directly as:

$$\begin{aligned} M(R_i,0)M(R_j,0) &= \sum_{k,m} \langle S| w_k R_i w_m |S\rangle |k\rangle \langle m| \sum_{n,\ell} \langle S| w_n R_j w_\ell |S\rangle |n\rangle \langle \ell| \\ &= \sum_{k,m,n,\ell} \langle S| w_k R_i w_m |S\rangle \langle S| w_n R_j w_\ell |S\rangle |k\rangle \langle m| |n\rangle \langle \ell| \\ &= \sum_{k,m,n,\ell} \delta_n^m \langle S| w_k R_i w_m |S\rangle \langle S| w_n R_j w_\ell |S\rangle |k\rangle \langle \ell| \\ &= \sum_{k,m,\ell} \langle S| w_k R_i w_m |S\rangle \langle S| w_m R_j w_\ell |S\rangle |k\rangle \langle \ell| \\ &= \sum_{k,\ell} \langle S| w_k R_i R_j w_\ell |S\rangle |k\rangle \langle \ell| \\ &= M(R_iR_j, 0) \end{aligned}$$

The second last line follows from the fact that $\sum_m w_m |S\rangle \langle S| w_m$ is just P_0 but all these states are in V_0 so it has no effect. □

Lemma 4.7. *If $c \in Z(WS)$, the center of the group WS, then $M(c,0) = I$*

Proof. Analyzing the action directly:

$$\begin{aligned} M(c,0) &= \sum_{k,\ell} \langle S| w_k c w_\ell |S\rangle |w_k\rangle \langle \ell| \\ &= \sum_{k,\ell} \langle S| w_k w_\ell c |S\rangle |w_k\rangle \langle \ell| \\ &= \sum_{k,\ell} \langle S| w_k w_\ell |S\rangle |w_k\rangle \langle \ell| \\ &= \sum_{k,\ell} \delta_\ell^k |w_k\rangle \langle \ell| \\ &= \sum_{\ell} |w_\ell\rangle \langle \ell| \end{aligned}$$

The second line follows because elements of $Z(WS)$ commute with all of WS. The third line is because $Z(WS) \subset S$. Therefore the only element of W which commutes with all of WS is the identity which is also in S. Since every element of WS is generated by elements of W and S, any center must be a subset of $I \cdot S = S$ □

The combination of the previous two results allows us to reduce the problem of finding the action of all elements of WS to that of finding the action of a generator of $WS/Z(WS)$. One such generator is the set $W \bigcup S/Z(WS)$, the only question remaining is what is the logical action of these operators.

Observation 1. *If $R \in W$, we have:*

$$M(R,0) = \sum_{k,\ell} \langle S| w_k R w_\ell |S\rangle |k\rangle \langle \ell|$$

Since W is a group, for each k there must exist some ℓ such that $w_k = Rw_\ell$ inducing a

switch between the $|\ell\rangle$ and $|k\rangle$ basis states. As such these are "X-type" logical errors.

Observation 2. *If $R \in S/Z(WS)$ then we have:*

$$M(R,0) = \sum_{k,\ell} \langle S| w_k R w_\ell |S\rangle |k\rangle \langle \ell| \tag{4.11}$$

$$= \sum_{k,\ell} (-1)^{f(R,w_\ell)} \langle S| w_k w_\ell R |S\rangle |k\rangle \langle \ell| \tag{4.12}$$

$$= \sum_{\ell} (-1)^{f(R,w_\ell)} |\ell\rangle \langle \ell| \tag{4.13}$$

Here we are using

$$f(A,B) = \begin{cases} 0 & \text{if } AB = BA \\ 1 & \text{if } AB = -BA \end{cases}$$

As is evident, the effective action of elements in this group is to introduce conditional phases. Therefore these are "Z-type" logical errors .

There is only one last consideration, which is, if Q encodes $k > 1$ qubits, there is an additional structure which needs to be determined, that of which qubit(s) the logical X and Z operators are acting on.

There is a degree of freedom here as the capacity of the effective channel, if not the effective channel description itself, will be identical under re-labelling of qubits. Care must be taken though as this choice can only be made once.

The simplest manner to do this, is to find a generator of the set W, $\langle W \rangle$. The elements of this generator R_{X_i} correspond to logical X_i operators acting on just one encoded qubit. In order to find the corresponding Z_i operators, one only needs to find an element of $S/Z(WS)$ which anti-commutes with R_{X_i} and no other element of $\langle WS \rangle$.

Thus, the general algorithm for finding the operator mapping can be summed up as:

Step 1 Find $Z(WS)$. For each $c \in Z(WS)$ set $M(c, 0) = I$.

Step 2 Find $\langle W \rangle$. For each $R_{X_i} \in \langle W \rangle$ set $M(R_{X_i}, 0) = X_i$.

Step 3 Find $S/Z(WS) - Z(WS)$. For each $R_{X_i} \in \langle W \rangle$ find $R_{Z_i} \in S/Z(WS)$ such that R_{Z_i} anti-commutes with only R_{X_i}. Set $M(R_{Z_i}, 0) = Z_i$.

Step 4 For all pairs of elements R_i R_j whose mappings are already known, $R_k = aR_iR_j$ set $M(R_k, 0) = aM(R_i, 0)M(R_j, 0)$ where a is some phase factor. After a new element R_k is added, all pairings with it should also be tested. After this step the mapping for all operators in WS is known.

Step 5 For all elements $F_m \in F$ and $R \in WS$, let $E = aF_m^\dagger R$ and set $M(E, m) = aM(R, 0)$. After this step, the mapping for all Pauli operators is known and the mapping is complete.

At this point an example might help clarify the process:

Example 4.5 (The 2-cat code)**.** *The 2-cat code is a stabilizer CWS code with* $W = \{II, XX\}$ *and* $S = \{II, IZ, ZI, ZZ\}$.

The cosets of WS *are* $WS = \{II, IZ, ZI, ZZ, XX, XY, YX, YY\}$ *and* $\{IX, IY, ZX, ZY, XI, XZ, YI, YZ\}$. *It is common to choose the set* F *such that it contains the lowest weight errors since those generally have the greatest probability of occurring. In this case we choose* $F_0 = II$ *and* $F_1 = IX$.

Step 1:

The center of WS is the subset of S which commutes with every element of W. In this case its fairly easy to manual find $Z(WS) = \{II, ZZ\}$. $M(II, 0) = M(ZZ, 0) = I$.

Step 2:

The generator of W, $\langle W \rangle = \{XX\}$. The 2-cat code encodes only one qubit so $M(XX, 0) = X$.

Step 3:

The set $S/Z(WS) - Z(WS) = \{IZ\}$. So set $M(IZ, 0) = Z$.

Step 4:

At this point we have the mappings for II, ZZ, XX, IZ for the other four note: $ZI = IZ \cdot ZZ$, $XY = iXX \cdot IZ$, $YX = iXX \cdot ZI$, $YY = -XX \cdot ZZ$. Using this information we have:

$$M(IZ, 0) = M(IZ, 0)M(ZZ, 0) \tag{4.14}$$

$$= Z \tag{4.15}$$

$$M(XY, 0) = iM(XX, 0)M(IZ, 0) \tag{4.16}$$

$$= iX \cdot Z = Y \tag{4.17}$$

$$M(YX, 0) = iM(XX, 0)M(ZI, 0) \tag{4.18}$$

$$= iX \cdot Z = Y \tag{4.19}$$

$$M(YY, 0) = -M(XX, 0)M(ZZ, 0) \tag{4.20}$$

$$= -X \tag{4.21}$$

Step 5:

To complete the mapping we need only calculate the values that for operators in the $F_1 = IX$ *coset. Similar to the last step we have:*

$$
\begin{aligned}
IX &= IX \cdot II \\
IY &= iIX \cdot IZ \\
ZX &= IX \cdot ZI \\
ZY &= iIX \cdot ZZ \\
XI &= IX \cdot XX \\
XZ &= -iIX \cdot XY \\
YI &= IX \cdot YX \\
YZ &= -iIX \cdot YY
\end{aligned}
$$

$$
\begin{aligned}
M(IX,1) &= M(II,0) = I \\
M(IY,1) &= iM(IZ,0) = iZ \\
M(ZX,1) &= M(ZI,0) = Z \\
M(ZY,1) &= iM(ZZ,0) = iI \\
M(XI,1) &= M(XX,0) = X \\
M(XZ,1) &= -iM(XY,0) = -iY \\
M(YI,1) &= M(YX,0) = Y \\
M(YZ,1) &= -iM(YY,0) = iX
\end{aligned}
$$

At this point the mapping is complete and is presented in table 4.1.

Table 4.1: Physical operators and their effective counter parts for 2-cat code

V_0		V_1	
II	$\tilde{I}$	IX	$\tilde{I}$
IZ	$\tilde{Z}$	IY	$i\tilde{Z}$
ZI	$\tilde{Z}$	ZX	$\tilde{Z}$
ZZ	$\tilde{I}$	ZY	$i\tilde{I}$
XX	$\tilde{X}$	XI	$\tilde{X}$
XY	$\tilde{Y}$	XZ	$-i\tilde{Y}$
YX	$\tilde{Y}$	YI	$\tilde{Y}$
YY	$-\tilde{X}$	YZ	$i\tilde{X}$

So, without belabouring the process any further, lets apply this code to the single qubit depolarizing channel with fidelity f given by:

$$\Lambda(\rho) = fI\rho I + gX\rho X + gY\rho Y + gZ\rho Z \tag{4.22}$$

Of course we are actually encoding our information into two qubits which individually are sent through the channel and have the same noise applied to each qubit. The resulting description of the total noise occurring is given by:

$$\begin{aligned}
\Lambda^{\otimes 2}(\rho) &= f^2 II\rho II + fgIX\rho IX + fgIY\rho IY + fgIZ\rho IZ + && (4.23)\\
& fgXI\rho XI + g^2 XX\rho XX + g^2 XY\rho XY + g^2 XZ\rho XZ + && (4.24)\\
& fgYI\rho YI + g^2 YX\rho YX + g^2 YY\rho YY + g^2 YZ\rho YZ + && (4.25)\\
& fgZI\rho ZI + g^2 ZX\rho ZX + g^2 ZY\rho ZY + g^2 ZZ\rho ZZ && (4.26)
\end{aligned}$$

The probabilities $p(m)$ are given by the sum of the components in each

component projected into V_m

$$p(0) = f^2 + 2fg + 5g^2 \tag{4.27}$$

$$p(1) = 4fg + 4g^2 \tag{4.28}$$

To describe the measurement conditioned effective channels $\Lambda_m(\rho)$ *simply replace the components of* $\Lambda^{\otimes 2}(\rho)$ *,* $E\rho E$ *with* $M(E,m)\rho M(E,m)^\dagger$*. Once done, the measurement conditioned effective channels are given by:*

$$\Lambda_0(\rho) = f^2 I\rho I + fgZ\rho Z + g^2 X\rho X + g^2 Y\rho Y + \tag{4.29}$$

$$g^2 Y\rho Y + g^2 X\rho X + fgZ\rho Z + g^2 I\rho I \tag{4.30}$$

$$\Lambda_1(\rho) = fgI\rho I + fgZ\rho Z + fgX\rho X + g^2 Y\rho Y + \tag{4.31}$$

$$fgY\rho Y + g^2 X\rho X + g^2 Z\rho Z + g^2 I\rho I \tag{4.32}$$

I should note for completeness, that the operator map is a super-operator [52] from the space of n-qubit channels to k-qubit channels and, in fact, is rather easily represented. The reason I chose not to represent the map in this manner is two-fold. First, I personally find looking at the super operators to be uninformative unless you have a solid grasp of the problem as I have presented above. Second, and more importantly the representation of the super operators are $\frac{2^n}{2^k}$ matrices, one for each measurement value, each of dimension $4^n \times 4^k$. Furthermore, even though they are sparse, in order to use them, you need to have a representation of your channel Λ as a $4^n \times 4^n$ matrix which is in generally not sparse.

As such we can save a lot of memory in the computation by representing it as a mapping of operators, and we can avoid constructing the full channel repre-

sentation by taking advantage of the fact that the full channel is a tensor product of n identical single-qubit channels. So every value of the representation of Λ can be calculated as a product of the elements of one 4×4 matrix. There is a trade off cost in time here, however since time on these scales is far cheaper than the memory that would be needed otherwise, I decided to tackle the problem this way.

4.4 Verification

Before employing the effective channels method to general channels, it is beneficial to verify our result by comparing them to known results on Pauli channels, in particular those of [66] and [69].

In Shor's original paper, he showed that using the n-cat code which can be described as :

$$|0\rangle \rightarrow |0\rangle^{\otimes n} \tag{4.33}$$

$$|1\rangle \rightarrow |1\rangle^{\otimes n} \tag{4.34}$$

concatenated with the hashing code as an outer code he could directly prove that any depolarizing channel with parameter $f \geq 0.8096$ had capacity. He produced table 4.2 showing the minimum fidelity possible while achieving capacity while using various length cat codes:

Table 4.2: Shors 1996 results[66]

Code	Min Fidelity
Hashing	0.8107
Cat 2	0.8114
Cat 3	0.8099
Cat 4	0.8101
Cat 5	0.8096
Cat 6	0.8100
Cat 7	0.8098

Using the effective channels procedure to calculate the coherent information, and the bisection method to determine minimum fidelity for which the coherent information was non-zero, the table 4.3 was produced.

Table 4.3: Capacity bounds on depolarizing channel using effective channels method

Code	Min Fidelity
Hashing	0.810708618164
Cat 2	0.0.811477661133
Cat 3	0.809872436523
Cat 4	0.810104370117
Cat 5	0.809646606445
Cat 6	0.809915161133
Cat 7	0.809768676758

For all codes up to the 7-cat code have perfect agreement. It is also possible to show the form of the 2-cat code effective channels coincides with the EPP output states that Shor calculates.

4.5 Results

The effective channel process is very general, requiring only a stabilizer CWS code Q and a channel Λ to be provided and the effective channel can be calculated along with the effective channel's coherent information.

Our goal is to find channels with $I_C(\Lambda) = 0$ for which $I_C(\tilde{\Lambda}) > 0$. The previous sections allow us to calculate these objects, but they don't tell us what codes to use or channels they will be effective on. Chief among the difficulties in tackling this problem is the fact that estimation of when the coherent information is zero on a general channel is computationally expensive as it requires showing that $I_C(\Lambda, \rho) = 0$ for all input states ρ. There does not exist a general characterization for when this is the case so the coherent information has to be estimated by sampling a large number of ρ such that you are fairly confident that the coherent information

is zero. When trying to find the precise boundary between zero and non-zero this can cause some errors.

Due to how difficult it is to estimate when the coherent information is zero we do not want to search all of the six parameter space for channels with zero coherent information. Luckily, the coherent information is a continuous function and reasonably well behaved, so if we start at a channel which we know has zero coherent information, it is likely that we can find other channels nearby that will also have zero coherent information. Here the concept of "near-Pauli" channels is beneficial which we define as:

Definition 4.18. *A near-Pauli channel is a channel for which the Pauli component of the channel p_x, p_y, p_z are the predominant values, the Kraus operators for these channels are only slightly perturbed from the Pauli operators.*

An example of a near-Pauli channel is to take some channel with parameters $[p_x, p_y, p_z, 0, 0, 0]$ and perturb it with some non-Pauli component like $[p_x, p_y, p_z, t_x, t_y, t_z]$ which normally will still represent a valid channel as long as t_x,t_y,t_z are small enough.

The primary benefit of investigating near-Pauli channels is the fact that, as long as the perturbation is small, the boundary between channels with and without coherent information is located close to the same boundary for the Pauli channel. Since finding the boundary for Pauli channels is much easier (and faster) than for non-Pauli channels, this makes finding candidates for interesting channels easier.

As for what codes to use, there is some speculation that degenerate codes are a valuable feature [66, 69, 68] though a precise understanding of why has not been found. Regardless, the cat-code family has already been shown to be capable

of achieving results on Pauli channels, so it is reasonable to attempt similar results on classes of non-Pauli channels.

4.5.1 2- and 3-cat code on t_z-perturbed near-depolarizing channels

We restrict the search for non-Pauli channels with zero coherent information to channels with the parameters $[g, g, g, 0, 0, t_z]$ where $g = \frac{1-f}{3}$ where f is the fidelity of the channel. Note that this is a near Pauli channel, where the Pauli channel that they are "near" is the depolarizing channel. Since the family of cat codes is capable of achieving capacity beyond the hashing bound on the depolarizing channel it seems reasonable that using the cat code on this family of near Pauli channels may have similar behaviour.

There are two free parameters to this family of channels, f and t_z. The depolarizing channel becomes anti-degradable when $f <= 0.75$ and it has coherent information when $f >= 0.8107$. So we investigate the range $0.75 < f < 0.8107$. For each value of f there are a couple of values of t_z we want to find. The first is the maximum value of t_z for which the parameters still constitute a valid quantum channel. The second, is the point the maximum t_z for which we can't find non-zero coherent information. Finally, we want to know the minimum t_z such that the effective channel for any input code has capacity.

All of these values are found by using the bisection method for binary functions where t_z is the varied parameter. The results are plotted in figure 4.3.

In figure 4.3, the cat 1 line is equivalent to the hashing bound, all channels above it have non-zero coherent information. Places where the cat2 or cat3 lines drop below the cat1 line are channels with zero coherent information and positive capacity.

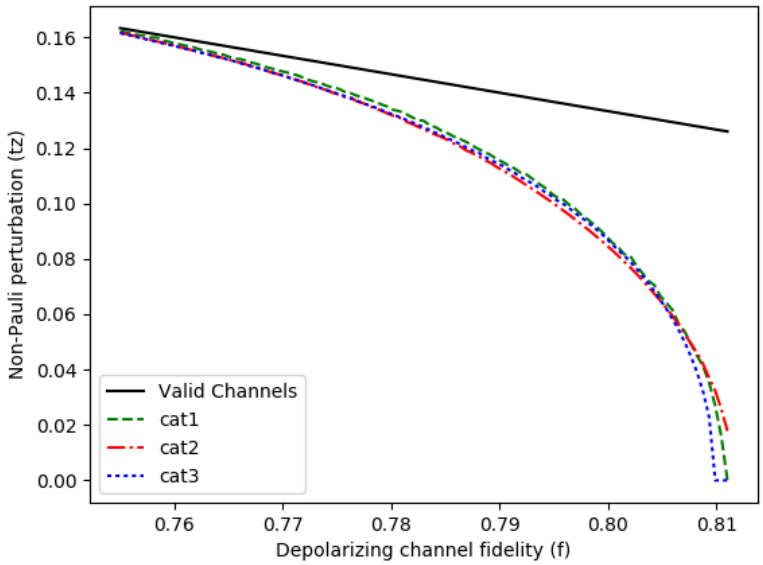

Figure 4.3: Comparison of capacity boundaries. The region below the "valid channels" line are all channels which satisfy the CPTP condition. The regions above any of the cat code lines are channels which have capacity under concatenation of the random stabilizer code with that code.

As we can see in the both figures, even though on the depolarizing channel the 2-cat code is not capable of achieving capacity beyond the hashing bound, for t_z perturbed channels the 2-cat code line drops below the hashing bound for most. The same can be seen for the 3-cat code, though interestingly, it does not perform better then the 2-cat code at low fidelities.

These are the first examples of a general non-Pauli channel with zero coherent information but non-zero capacity. it should be noted that these was nothing special about the choice of channels to look at, we could easily have picked any other class of families. Furthermore the near-Pauli approach is simply utilizing the properties of the well understood Pauli channels in order to simplify searching for the hashing bound. As such there is no reason to not expect to find these kinds of results to be rather common, much in the way it was shown that the cat code is capable of achieving capacity on most Pauli channels [69].

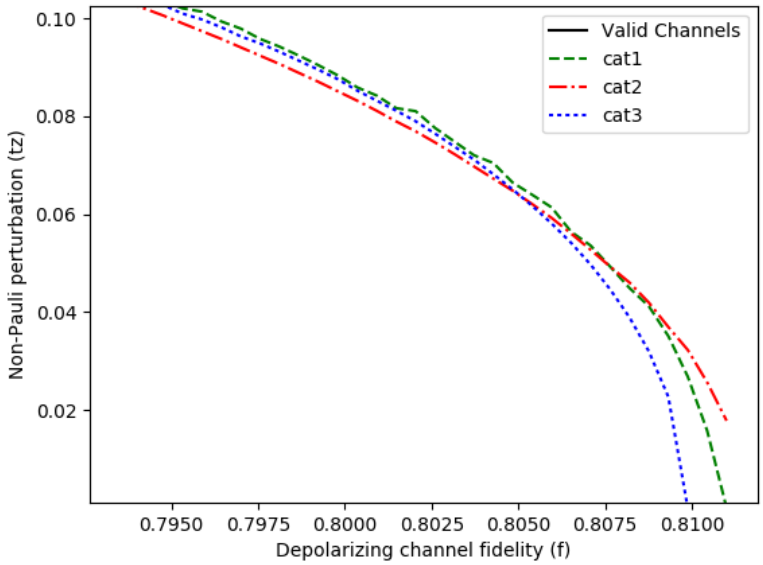

Figure 4.4: Zoomed in version of figure 4.3 to region of particular interest.

The procedure used to obtain these results is a very generally applicable tool which can be used to investigate many situations of interest. Indeed there are indications that results with wider margins are likely to be found in near-Pauli channels where the Pauli component is some ways off from the depolarizing channel.

4.6 The Effectiveness of Degenerate Codes

Even though we now have a framework for investigating if a channel has capacity under action of a particular code, I have not yet addressed the question of what kinds of codes are best for obtaining capacity beyond the hashing bound. Shor first illustrated that use of degenerate codes could achieve capacity beyond the hashing bound. Since then, a lot of the literature [19, 69] has attempted to answer what makes degenerate codes successful.

It is known that in order to achieve capacity beyond the hashing bound a code must be degenerate. The reason for this derives from how the coherent

information is proven to be an achievable capacity. Inspired by Shannon's random coding theorem [63], which in the classical case achieves the capacity of a channel, it was shown that the quantum analogue of using a random stabilizer codes achieves the coherent information of a channel. This proof is essentially classical in nature and does not account for fundamentally quantum behaviours, in particular code degeneracy. Quantum codes are degenerate when multiple errors in their error-set $\mathcal{E}$ have the same effect on the code. It is impossible for this to occur classically, but comes about naturally in the quantum world. In order for any code to perform better on a channel than the random stabilizer code, it must make use of degeneracy. Otherwise the codes being used acts no differently than a random stabilizer code.

More formally, a QECC Q which detects an error-set $\mathcal{E}$ is called degenerate if for some $E \in \mathcal{E}$ we have:

$$\forall_i \langle v_i | E | v_i \rangle = c_E \neq 0$$

This comes from the error detection criteria 1.1. When the $c_E = 0$ for all $E \in \mathcal{E}$ then the code is non-degenerate. It is clear from this, that code degeneracy is a feature which comes from both the code and the error-set being corrected/detected $\mathcal{E}$.

If we view what this means in the stabilizer CWS code framework we have the following equivalent condition:

$$\forall_i \langle S | w_i^\dagger E w_i | S \rangle = c_E \neq 0$$

It should be clear from discussion earlier that for this condition to be satisfied it is sufficient that $E \in \pm Z(WS)$ since these are elements of S which

commute with every $w_i \in W$ and as such we can say:

$$\begin{aligned} \langle S| w_i^\dagger E w_i |S\rangle &= \langle S| w_i^\dagger w_i E |S\rangle && (4.35) \\ &= \langle S| |S\rangle && (4.36) \\ &= 1 && (4.37) \end{aligned}$$

So in order for a code to be degenerate it is sufficient if the error-set and the center of WS have a non-trivial intersection. For error correction this is less straightforward, but by similar reasoning, a code is degenerate if for some $E \in \mathcal{E}$ which can be written in the form $F_m^\dagger R$ where $R \in Z(WS)$.

I note that any non-degenerate code with error-set $\mathcal{E}$ is can be "turned" into a degenerate code by expanding the error-set to $\mathcal{E} \bigcup Z(WS)$. This is partially why focusing the discussion on code degeneracy can be complex. The random stabilizer codes assume only one error can be corrected in each likely subspace, but any non-trivial stabilizer code will have more then one error per subspace which is correctable by a similar argument.

When the error-set of a channel corresponds to the entirety of errors which a noisy channel may produce, then the noise of the channel is perfectly correctable. Unfortunately we are rarely so lucky and more often the error-set $\mathcal{E}$ represents just the most likely subset of all errors. In a sense the choice of error-set for a code is a secondary factor, and that is why understanding the full behaviour of the channel, including uncorrectable errors, under action of a code has been the focus of this chapter and the effective channel techniques. I will now use it to attempt to explain why bad codes are sometimes good for achieving capacity, such as the cat code.

When designing a quantum code, which corrects some error-set the sole goal is to maximize the fidelity of the output state. In order for errors to be correctable they must map to the identity as an effective operator. That is $M(E, m) = I$ for all $E \in \mathcal{E}$. The effective channel construction shows that for some stabilizer CWS code specified by W, S and F then the operators which map to the identity are $F \cdot Z(WS)$. That is every element of F, every element of $Z(WS)$ and every element in their product.

In the effective channel framework, maximizing the output state fidelity is obtained by maximizing the weighted average of the fidelities of the measurement conditioned effective channels:

$$\sum_i p_i \left\langle \phi^+ \right| \sigma_{\Lambda_i} \left| \phi^+ \right\rangle \tag{4.38}$$

The error-set is chosen (for the symmetric channel) such that the weight of the error-set elements are as small as possible since those operators will have the greatest contribution to the noise. Assuming that P_I is the greatest component in the channel description, the elements of the error-sets with weight less than d will be the operators with the greatest contribution of noise and if corrected will maximize the value of (4.38). This is, in essence, the greedy algorithm trying to pack the highest value operators into $F \cdot Z(WS)$. This explains why under certain circumstances degenerate codes are beneficial as $|F \cdot Z(WS)| \geq |F|$.

On the other hand, the goal of finding codes which improve capacity of a channel has the goal of maximizing the average channel coherent information (see:4.12):

$$\max_\rho \sum_i p_i I_C(\Lambda_i, \rho) \tag{4.39}$$

or in the case of Pauli channels,

$$\sum_i p_i(1 - S(\sigma_{\Lambda_i})) \tag{4.40}$$

From this alone we can see that "good codes" which maximize (4.38) need not also maximize (4.39) or (4.40) since the former is concerned only with one value from each Λ_i and the latter is concerned with the complete behaviour of Λ_i.

Lets consider, as an example, 2-qubit stabilizer CWS codes acting on the depolarizing (Pauli) channel with fidelity f. The channel is given by:

$$\Lambda(\rho) = fI\rho I + gX\rho X + gY\rho Y + gZ\rho Z \tag{4.41}$$

The description of n uses of this channel is given by:

$$\Lambda^{\otimes n}(\rho) = \sum_{E \in \mathcal{P}_n} f^{n-\mathrm{wt}(E)} g^{\mathrm{wt}(E)} E\rho E^\dagger \tag{4.42}$$

Where E is an element of the n-qubit Pauli operators $\mathcal{P}_n$. When $f > g$ then the terms with the highest contribution $f^{n-\mathrm{wt}(E)} g^{\mathrm{wt}(E)}$ are those with the lowest weight $\mathrm{wt}(E)$. When constructing a QECC code, maximizing (4.38) corresponds to packing as many low weight elements into $F \cdot Z(WS)$ as possible.

We first need to choose a set S of 4 mutually commuting operators. S will contain II and due to symmetries of the depolarizing channel, choosing $S = \{II, IZ, ZI, ZZ\}$ allows us to explore all unique situations up to local Clifford equivalence. Now we have some freedom to choose either W or $Z(WS)$, though it it easier

to follow the discussion when choosing $Z(WS)$. In 2-qubit stabilizer CWS codes, $|Z(WS)| = 2^{2-k}$ where k is the number of encoded qubits. If $k = 0$, the code can not protect any state, and if $k = 2$ we have the trivial code. Therefore $k = 1$ and $|Z(WS)| = 2$. Up to uniqueness, there is only two options for $Z(WS)$ which we consider below:

Case 1: $S = \{II, IZ, ZI, ZZ\}$, $Z(WS) = \{II, IZ\} \rightarrow W = \{II, XI\}$

This code has the following table:

	Effective Operator			
m	I	X	Y	Z
0	II,IZ	XI,XZ	YI,YZ	ZI,ZZ
1	IX,IY	XX,XY	YX,YY	ZX,ZY

The choice of $F_1 = IX$ is arbitrary but is both maximizing for (4.38) and an unimportant choice for (4.40) since choosing any other error than F_1 constitutes a local unitary rotation on the effective channel which does not affect the capacity. On the whole this code does as good as is possible for packing in low weight errors into $F \cdot Z(WS)$. Being a 2 qubit code it is impossible to do better than this as no distance 1 two-qubit code exists.

We can also examine the measurement conditioned effective channels:

$$\begin{aligned}
\Lambda_0(\rho) &= \frac{1}{p_0}((f^2 + fg)I\rho I + (fg + g^2)X\rho X + (fg + g^2)Y\rho Y + (fg + g^2)Z\rho Z \\
\Lambda_1(\rho) &= \frac{1}{p_1}((2fg)I\rho I + (2g^2)X\rho X + (2g^2)Y\rho Y + (2g^2)Z\rho Z
\end{aligned}$$

both of these effective channels are depolarizing channels with fidelities $f_0' = \frac{f^2+fg}{f^2+4fg+3g^2}$ and $f_1' = \frac{2fg}{2fg+6g^2}$. Though it doesn't look like it, $f_0' = f_1' = f$ and as a result this code produces the same values for (4.40) as the hashing code.

Case 2: $S = \{II, IZ, ZI, ZZ\}$, $Z(WS) = \{II, ZZ\} \rightarrow W = \{II, XX\}$

This is the 2 cat code, as such, the table for the code is given by:

	Effective Operator			
m	I	X	Y	Z
0	II,ZZ	XX,YY	YX,XY	ZI,IZ
1	IX,ZY	XI,YZ	YI,XZ	XZ,IY

The thing to note about this code is that it is the worst possible code for (4.38) as $F \cdot Z(WS)$ contains only 2 low weight errors. Not that this comes as a surprise since the cat codes are really bad codes. As we know the minimum fidelity for the 2 cat code is actually worse than for the hashing code, but actually performs past the coherent information boundary on some non-Pauli channels.

The fact that what amounts to the "best" 2-qubit stabilizer code, at least in terms of error correction (4.38) actually results in a code which, while degenerate if considering the error-set $\mathcal{E} = \{II, IX, IY, IZ\}$, actually will perform identically to a random stabilizer code. On the other hand the bad 2-cat code we have shown on some non-Pauli channels can achieve capacity beyond the hashing bound.

This exemplifies the point that good codes which maximize (4.38), while still possibly degenerate still might not provide benefit beyond the hashing bound. However, this process could be expanded upon to find new good codes or possibly determine optimum codes for particular classes of channels. Ultimately the purpose of discussing this is to show that degeneracy itself is not likely the key factor in maximization of 4.39.

It has been suggested that the effectiveness of the cat code family comes from sacrificing all knowledge about particular classes of errors [19]. It is possible that this is the case, however I have seen that while the cat code achieves capacity beyond the hashing bound for a large group of Pauli channels [69], there are other "good" codes, such as the $[[6, 2]]$ code we found in the exhaustive search of chapter 2

which are capable of doing the same on various Pauli channels also, though notably not the depolarizing channel.

Unfortunately good answers don't exist at this point as to what features helps a code optimize (4.39) and as such general limits for capacity are probably still a way off. However in the hope of furthering that effort, one could construct quantum codes much like in the example above or like the code search of chapter 2 and search for any codes which might perform better than the cat code.

4.7 Discussion

This chapter outlined the work done in the attempt to answer a simple question. "Do there exist channels with capacity not covered by LSD theorem [48, 64, 18]". This is a very hard problem in general since, despite attempts, the quantum channel capacity does not seem to permit a simple form. In 1996 Shor showed it was possible on Pauli channels and since then there have been a couple techniques proposed that work on certain subsets of channels as well and a handful of well behaved families of channels for which analysis is easier. However even the majority of results have been on Pauli channels or channels which can be mapped to Pauli channels to provide bounds.

My approach as outlined in this chapter has been markedly different, most notably in attempting to tackle the problem directly by utilizing the large (frankly ridiculous) amount of computing power available today. Developing the effective channel framework to computationally determine the effect of quantum codes on arbitrary channels allows us to "try and see" if there are interesting channels.

The end result was that by use of the 2-cat and 3-cat codes I was able

to show that there are a collection of channels which do not have capacity from random stabilizer codes but can be proven to have non-zero capacity by using the cat codes concatenated with the random stabilizer as an outer code. The benefit of the approach presented is in its straight forwardness, relying on little other than coding theory. The downside is that aside from showing that the behaviour of the quantum capacity is interesting on non-Pauli channels, my work provides limited insight into the reasons why this is the case.

As for future directions of this work, I believe there to be many. This work was a blunt tool to crack the shell around working beyond the Pauli channels or special cases of the non-Pauli channels. There seems to be interesting results on most channels. The question of where the true boundary for quantum channel capacity is not answered by this work but this work might point to interesting behaviours which could be key to one day finding it.

The first most obvious extension to the work would be to attempt the same results on larger then 3-qubit cat codes. This has a high potential of producing good results since even the larger cat codes showed improvement of capacity on the depolarizing channel beyond the hashing point. These codes can also be directly applied to channels other than the near-depolarizing channels. The expectation is that there are a large number of near-depolarizing channels which have capacity beyond the hashing bound.

Other notable future works to be done are possibly extend the effective channels to non-stabilizer codes, though based on the structure of error mapping I personally feel that this might be less fruitful on all but subsets of very specific channels. A minor restriction this work had was in the fact that since the coherent information of multi-qubit channels is very hard to estimate. As a result any code

which had an encoded dimension greater than 2 was ignored. A possible solution is to concatenate the outer code with an inner one protecting a single qubit and is promising since there has been some success with concatenated codes [69].

Another direction for this work, and in fact my thesis in general, would be to extend the work of section 4.6. It is conceivable to attempt designing CWS codes specifically for the channel they are to be used on much like in chapter 2. However, using the insights in the effective channel construction its possible that we could optimize the choice of S, W and F to make the effective channels less noisy on the whole. This kind of search would likely be even more expensive than even the chapter 2 exhaustive search. It is possible to employ the same shortcuts that have been core to my thesis, such as Clifford equivalence and reducing the problem space through group theory, in order to make the problem somewhat tenable.

Finally, figuring out the "why" of these results would be a massive step forward towards being able to design codes on a per-channel basis. Ideally we would be able to take the channel parameters and design a code which will protect it better.

Chapter 5

Conclusion

Quantum computing is one of the most promising research areas for the future of computers but in order for them to come to fruition we have to deal with the fundamental problem that no system, quantum or classical, is completely immune to noise. In general, quantum systems are more sensitive to noise introduced by the environment due to the sensitivity of the physical implementations.

Noise in classical information theory was tackled in two ways, first the systems designed were improved in fidelity to the point where errors were highly improbable making classical computation highly reliable. For systems where this either wasn't sufficient or couldn't be implemented, a second technique, that of error correcting codes, has been used to great effect in protecting both data and computational operations.

If the implementations of quantum information theory are to be successful it is highly likely that our control and use of quantum systems will have to follow a similar path. Improving the fidelity of quantum systems as much as is possible and to employ error correction techniques for the rest.

The framework of general quantum error correcting codes has been well

documented, however constructing good codes and understanding their effects on the quantum systems is still a hard and important task. This thesis has outlined the work done on three separate but related projects all related to either the design of good quantum codes, or understanding the effect quantum codes have on quantum channels. All three projects are also related in the techniques brought to bear on the problems.

First of these techniques is to tackle a problem in a somewhat direct manner. Traditionally performing the calculations required to do this was beyond that which was available even 20 years ago. The explosion in computational power now available has resulted in being able to tackle computationally intense processes even in a mid-high end desktop. Without this previous works were more or less required to focus on situations with high degrees of symmetry or cases which could be simplified. The number of results obtained using these very creative techniques are impressive, But without the restriction to these sorts of situations we can perform searches and simulations which were previously completely impractical and as a result achieve results which have not previously been known.

The second technique prevalent throughout this work is in the use of group theory to help reduce the problem space. Even in light of the computational power available today as noted above, we should not do more work then is necessary. To that end the use of group theory is often very helpful in making statements which can restrict the problem in size to something more manageable. Of course, as with any optimization, you should not spend more time on the process of optimization than you will save by doing so. This is rarely a problem though as, in general, the problems you encounter when working with quantum systems will almost always be exponentially more expensive in the dimension of quantum system of concern.

Using these techniques, this work focused on three problems. First we investigated the construction of good quantum codes designed for asymmetric noise models, in particular the amplitude damping channel. Most codes designed to date have been for the depolarizing channel noise model, however by tailoring the codes to the noise model of interest better parameters can be found. In fact we performed an exhaustive search on CWS codes up to length 9 and found many codes which performed better than any previously known. If codes with better parameters exist, then they will require a different code framework.

Second we built new codes through concatenation for the same asymmetric channel. We noted that the effective channel of the amplitude damping channel when applying codes designed for amplitude damping noise was essentially symmetric in nature and as such doing concatenation of inner asymmetric codes with outer symmetric codes produced many new codes for the amplitude damping channel with lengths larger then was feasible through exhaustive search. Many of these also had better parameters then previously know.

Finally we addressed the problem of quantum channel capacity which involved essentially the same framework as before except now our channels were general and our outer code was fixed to the hashing or random stabilizer codes. This work was inspired by the large body of work trying to determine when a given quantum channel with zero coherent information has capacity. A problem which is possibly impossible to calculate an answer to directly. Previous efforts focused on Pauli channels due to the ease of computing the properties, our work was to formalize the framework of constructing the description of the effective channel and then using its properties to determine if the effective channel would have non-zero capacity. This process relied on both the CWS code framework as well as applications

of cosets of the Pauli operators in order to simplify the construction. In the end we were able to apply this work to a particular class of non-Pauli channels which were perturbed from the depolarizing channel, and by doing so, we found a range of channels with zero coherent information but positive capacity which had not been done in this direct manner before.

To conclude, this work demonstrated that although many of the problems in quantum error correcting codes and quantum channel capacities are computationally difficult, with clever use of algebra and sheer computational power many new and interesting results are possible. All of which have the potential to be useful for practical quantum systems in the future.

Appendix A

Codeword Stabilized Quantum Code Search Algorithms

In chapter 2 we relied on an algorithm for constructing CWS codes which took as input a graph $\mathcal{G}$ and an error-set $\mathcal{E}$ that we were designing the code to detect and returned as an output the largest dimension code which satisfied the error detection conditions. We outlined the procedure in 2.2.2 and include here for reference pseudocode for the algorithms.

Algorithms 1-3 first appeared in [10]. Algorithm 1 takes the input graph and error-set and returns the classical errors the code must correct $Cl_S(\mathcal{E})$ as well as the set of invalid classical codewords due to the degeneracy condition D. Algorithm 2 takes as input $Cl_S(\mathcal{E})$ and D and produces the graph $\mathcal{G}_C$ which contains all information of codewords which can appear in code $\mathcal{C}$ while satisfying the error detection conditions for the associated CWS code $\mathcal{Q}$. Algorithm 3 combines the first two algorithms and then runs a max clique solver to obtain the largest dimension code possible.

Algorithm 4 is the extension of this work to perform an exhaustive search

for all CWS codes detecting the error-set $\mathcal{E}$. In order to perform an exhaustive search, every possible input graph with n vertices must be used, however we optimized this search by using only local equivalent graphs which have been characterized in [17]. Restricting to this set of graphs as input is actually too restrictive to be exhaustive and as a result we are required to expand the number of error-sets under consideration, that is, to use a set of error-sets which are all local Clifford equivalent to the orginal error-set of interest $\mathcal{E}$. Up to global phase, the local Clifford operations act as the permutation group on three elements S_3 permuting the X,Y, and Z operators. By choosing our graphs and error-sets in this manner we guarantee searching for every possible CWS code thus making the search exhaustive.

So Algorithm 4 takes as input a set of graphs $\{\mathcal{G}\}$ and a set of error-sets $\{\mathcal{E}\}$ and returns the largest dimension code from every pairing of graph and error-set. It should be noted that the structure of algorithm 4 is highly parallelizable in that in theory every pair of $\mathcal{E}$ and $\mathcal{G}$ can be run independently of each other and the results collected at the end allowing for a near-linear speed up.

Algorithm 1 Setup$(\mathcal{E}, \Lambda)$: Compute $Cl_{\mathcal{G}}(\mathcal{E})$ and $D_{\mathcal{G}}(\mathcal{E})$, where $\mathcal{E}$ is a set of Pauli errors and Λ is the adjacency matrix associated with graph $\mathcal{G}$.

Require: $\Lambda^T = \Lambda$, $\Lambda_{ij} = \{0, 1\}$ and $\Lambda_{ii} = 0$
Ensure: $\mathrm{CL}[i] = \delta(\mathrm{String}(i) \in Cl_{\mathcal{G}}(\mathcal{E}))$ and $\mathrm{D}[i] = \delta(\mathrm{String}(i) \in D_{\mathcal{G}}(\mathcal{E}))$

```
for i ∈ {0,1}^n do
    CL[Integer(i)] ← 0
    D[Integer(i)] ← 0
end for
for error configuration E ∈ ℰ do
    ERR ← String(0)
    ERRX ← String(0)
    for (LOC, TYPE) in E do
        if TYPE is X or Y then
            ERR ← ERR ⊕ (row LOC of Λ)
            ERRX ← ERR ⊕ String(2^LOC)
        end if
        if TYPE is Z or Y then
            ERR ← ERR ⊕ String(2^LOC)
        end if
    end for
    CL[Integer(ERR)] ← 1
    if Integer(ERR) is 0 then
        for i ∈ {0,1}^n do
            if ERRX · i ≠ 0 then
                D[i] ← 1
            end if
        end for
    end if
end for
return (CL, D)
```

Algorithm 2 MakeCWSCliqueGraph$(\mathrm{CL}, \mathrm{D})$: Construct a graph whose vertices V are classical codewords and whose edges E connect codewords that can belong to the same classical code, according to the error model indicated by $Cl_{\mathcal{G}}(\mathcal{E})$ and $D_{\mathcal{G}}(\mathcal{E})$.

Require: CL and D are binary arrays of length 2^n
Ensure: $0^n \in V$, $0^n \neq v \in V \Rightarrow \mathrm{D}[v] = 0$ and $\mathrm{CL}[v] = 0$, $(v, w) \in E \Rightarrow \mathrm{CL}[v \oplus w] = 0$

```
V ← {0^n}
E ← ∅
for s ∈ {0,1}^n do
    if D[s] = 0 and CL[s] = 0 then
        V ← V ∪ {s}
        for v ∈ V \ {s} do
            if CL[v ⊕ s] = 0 then
                E ← E ∪ {(v, s)}
            end if
        end for
    end if
end for
return (V, E)
```

Algorithm 3 CWS-MAXCLIQUE$(\mathcal{E}, \Lambda)$: Find a quantum code $\mathcal{Q}$ detecting errors in $\mathcal{E}$, and providing the largest possible dimension K for the given input. The input Λ specifies the adjacency matrix of the graph $\mathcal{G}$. The output $\mathcal{C}$ is a classical code such that $\mathcal{Q} = (\mathcal{G}, \mathcal{C})$ is a CWS code detecting errors in $\mathcal{E}$.

Require: $\Lambda^T = \Lambda$, $\Lambda_{ij} = \{0, 1\}$ and $\Lambda_{ii} = 0\ \forall i$
Ensure: $K = |\mathcal{C}|$ is as large as possible for the given input, $0^n \in \mathcal{C}$, and $\mathcal{C}$ satisfies the error detection conditions for standard form CWS codes.

```
(CL, D) ← Setup(ℰ, Λ)
(V, E) ← MakeCWSCliqueGraph(CL, D)
𝒞 ← findMaxClique(V, E)
return 𝒞
```

Algorithm 4 ExhaustiveSearch$(\{\Lambda\}, \{\mathcal{E}\})$: Find the CWS code $\mathcal{Q}$ detecting errors in $\mathcal{E}$, and providing the largest possible dimension K. The input $\{\Lambda\}$ specifies the adjacency matrices of the set of graphs $\{\mathcal{G}\}$. The output $\mathcal{C}$ is the largest classical code such that $\mathcal{Q} = (\mathcal{G}, \mathcal{C})$ is a CWS code detecting errors in $\mathcal{E}$.

Require: $\Lambda^T = \Lambda$, $\Lambda_{ij} = \{0, 1\}$ and $\Lambda_{ii} = 0\ \forall i$
Ensure: words

```
M ← {}
for ℰ ∈ {ℰ} do
    for Λ ∈ {Λ} do
        𝒞 ← CWS-MAXCLIQUE(ℰ, *)
        M ← M ∪ 𝒞
    end for
end for
return findMax(M)
```

Bibliography

[1] S. Beigi, I. Chuang, M. Grassl, P. Shor, and B. Zeng. Graph concatenation for quantum codes. *Journal of Mathematical Physics*, 52(2):022201, February 2011.

[2] John S Bell. On the Einstein Podolsky Rosen paradox, 1964.

[3] Charles H Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. *Theoretical computer science*, 560:7–11, 2014.

[4] Charles H Bennett, David P DiVincenzo, John A Smolin, and William K Wootters. Mixed-state entanglement and quantum error correction. *Physical Review A*, 54(5):3824–3851, 1996.

[5] Charles H Bennett and Stephen J Wiesner. Communication via one-and two-particle operators on Einstein-Podolsky-Rosen states. *Physical review letters*, 69(20):2881, 1992.

[6] Max Born. Quantenmechanik der stoßvorgänge. *Zeitschrift für Physik*, 38(11-12):803–827, 1926.

[7] A. R. Calderbank, E. Rains, P. W. Shor, and N. J. A. Sloane. Quantum error correction via codes over $GF(4)$. *IEEE Transactions on Information Theory*, 44(4):1369–1387, 1998.

[8] A. R. Calderbank and P. W. Shor. Good quantum error-correcting codes exist. *Physcial Review A*, 54(2):1098–1105, August 1996.

[9] X. Chen, B. Zeng, and I. L. Chuang. Nonbinary codeword-stabilized quantum codes. *Physical Review A*, 78(6):062315, December 2008.

[10] I. Chuang, A. Cross, G. Smith, J. Smolin, and B. Zeng. Codeword stabilized quantum codes: Algorithm and structure. *Journal of Mathematical Physics*, 50(4):042109, April 2009.

[11] Isaac L Chuang, Debbie W Leung, and Yoshihisa Yamamoto. Bosonic quantum codes for amplitude damping. *Physical Review A*, 56(2):1114–1125, 1997.

[12] John F Clauser, Michael A Horne, Abner Shimony, and Richard A Holt. Proposed experiment to test local hidden-variable theories. *Physical review letters*, 23(15):880, 1969.

[13] Thomas M Cover and Joy A Thomas. *Elements of information theory*. John Wiley & Sons, 2012.

[14] A. Cross, G. Smith, J. A. Smolin, and B. Zeng. Codeword stabilized quantum codes. *IEEE Transactions on Information Theory*, 55(1):433–438, 2009.

[15] Toby S Cubitt, Mary Beth Ruskai, and Graeme Smith. The structure of degradable quantum channels. *Journal of Mathematical Physics*, 49(10):102104, 2008.

[16] Lars Eirik Danielsen. On self-dual quantum codes, graphs, and boolean functions. *arXiv preprint quant-ph/0503236*, 2005.

[17] Lars Eirik Danielsen and Matthew G Parker. On the classification of all self-

dual additive codes over GF(4) of length up to 12. *Journal of Combinatorial Theory, Series A*, 113(7):1351–1367, 2006.

[18] Igor Devetak and Peter W Shor. The capacity of a quantum channel for simultaneous transmission of classical and quantum information. *Communications in Mathematical Physics*, 256(2):287–303, 2005.

[19] David P DiVincenzo, Peter W Shor, and John A Smolin. Quantum-channel capacity of very noisy channels. *Physical Review A*, 57(2):830, 1998.

[20] Rod G Downey and Michael R Fellows. Fixed-parameter tractability and completeness ii: On completeness for w [1]. *Theoretical Computer Science*, 141(1-2):109–131, 1995.

[21] Runyao Duan, Markus Grassl, Zhengfeng Ji, and Bei Zeng. Multi-error-correcting amplitude damping codes. In *Information Theory Proceedings (ISIT), 2010 IEEE International Symposium on*, pages 2672–2676. IEEE, 2010.

[22] Albert Einstein, Boris Podolsky, and Nathan Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical review*, 47(10):777, 1935.

[23] Martianus Frederic Ezerman and Markus Grassl. Asymmetric quantum codes detecting a single amplitude error. In *Proceedings 2013 IEEE International Symposium on Information Theory (ISIT 2013)*, pages 922–926, 2013.

[24] Uriel Feige, Shafi Goldwasser, László Lovász, Shmuel Safra, and Mario Szegedy. Approximating clique is almost np-complete. In *Foundations of Computer Science, 1991. Proceedings., 32nd Annual Symposium on*, pages 2–12. IEEE, 1991.

[25] Andrew S Fletcher, Peter W Shor, and Moe Z Win. Channel-adapted quantum error correction for the amplitude damping channel. *IEEE Transactions on Information Theory*, 54(12):5705–5718, 2008.

[26] Daniel Gottesman. Stabilizer codes and quantum error correction. *arXiv preprint quant-ph/9705052*, 1997.

[27] Daniel Gottesman. The Heisenberg representation of quantum computers. *arXiv preprint quant-ph/9807006*, 1998.

[28] Markus Grassl. Bounds on the minimum distance of linear codes and quantum codes. Online available at http://www.codetables.de, 2007. Accessed on 2016-01-18.

[29] Markus Grassl, Thomas Beth, and Martin Rötteler. On optimal quantum codes. *International Journal of Quantum Information*, 2(1):55–64, 2004.

[30] Markus Grassl and Martin Rötteler. Non-additive quantum codes from Goethals and Preparata codes. In *Proceedings 2008 IEEE Information Theory Workshop (ITW 2008)*, pages 396–400, 2008.

[31] Markus Grassl and Martin Rötteler. Quantum Goethals-Preparata codes. In *Proceedings 2008 IEEE International Symposium on Information Theory (ISIT 2008)*, pages 300–304, 2008.

[32] Markus Grassl and Martin Rötteler. Quantum MDS codes over small fields. In *Proceedings 2015 IEEE International Symposium on Information Theory (ISIT 2015)*, pages 1104–1108, 2015.

[33] Markus Grassl, Peter Shor, Graeme Smith, John Smolin, and Bei Zeng. Generalized concatenated quantum codes. *Physical Review A*, 79(5):050306, 2009.

[34] Markus Grassl, Zhaohui Wei, Zhang-Qi Yin, and Bei Zeng. Quantum error-correcting codes for amplitude damping. In *Proceedings 2014 IEEE International Symposium on Information Theory (ISIT 2014)*, pages 906–910, 2014.

[35] Alfred Haar. Der massbegriff in der theorie der kontinuierlichen gruppen. *Annals of mathematics*, pages 147–169, 1933.

[36] M. Hein, J. Eisert, and H. J. Briegel. Multiparty entanglement in graph states. *Physical Review A*, 69:062311, 2004.

[37] Dan Hu, Weidong Tang, Meisheng Zhao, Qing Chen, Sixia Yu, and C. H. Oh. Graphical nonbinary quantum error-correcting codes. *Physical Review A*, 78(1):012306, 2008.

[38] Lev Ioffe and Marc Mézard. Asymmetric quantum error-correcting codes. *Physical Review A*, 75(3):032345, Mar 2007.

[39] Tyler Jackson. *Quantum Tomography with Pauli Operators*. PhD thesis, 2013.

[40] Tyler Jackson, Markus Grassl, and Bei Zeng. Codeword stabilized quantum codes for asymmetric channels. In *Information Theory (ISIT), 2016 IEEE International Symposium on*, pages 2264–2268. IEEE, 2016.

[41] Tyler Jackson, Markus Grassl, and Bei Zeng. Concatenated codes for amplitude damping. In *Information Theory (ISIT), 2016 IEEE International Symposium on*, pages 2269–2273. IEEE, 2016.

[42] Frederik Kesting, Florian Fröwis, and Wolfgang Dür. Effective noise channels for encoded quantum systems. *Physical Review A*, 88(4):042305, 2013.

[43] Emanuel Knill and Raymond Laflamme. Theory of quantum error-correcting codes. *Physical Review A*, 55(2):900–911, 1997.

[44] Emanuel Knill, Raymond Laflamme, and Wojciech H. Zurek. Resilient quantum computation: error models and thresholds. *Proceedings of the Royal Society of London Series A*, 454(1969):365–384, 1998.

[45] Ruitian Lang and Peter W. Shor. Nonadditive quantum error correcting codes adapted to the amplitude damping channel. arXiv:0712.2586 [quant-ph], 2007.

[46] Debbie W Leung, Michael A Nielsen, Isaac L Chuang, and Yoshihisa Yamamoto. Approximate quantum error correction can lead to better codes. *Physical Review A*, 56(4):2567–2573, 1997.

[47] Y. Li, I. Dumer, and L. P. Pryadko. Clustered Error Correction of Codeword-Stabilized Quantum Codes. *Physical Review Letters*, 104(19):190501, May 2010.

[48] Seth Lloyd. Capacity of the noisy quantum channel. *Physical Review A*, 55(3):1613, 1997.

[49] S. Y. Looi, L. Yu, V. Gheorghiu, and R. B. Griffiths. Quantum error correcting codes using qudit graph states. *Physical Review A*, 78(4):042303, 2008.

[50] Michael A. Nielsen and Isaac L. Chuang. *Quantum computation and quantum information*. Cambridge University Press, 2010.

[51] Yingkai Ouyang. Channel covariance, twirling, contraction, and some upper bounds on the quantum capacity. *Quantum Information & Computation*, 14(11-12):917–936, 2014.

[52] John Preskill. Lecture notes for physics 229: Quantum information and computation. *California Institute of Technology*, 16, 1998.

[53] Eric M. Rains. Nonbinary quantum codes. *IEEE Transactions on Information Theory*, 45(6):1827–1832, 1999.

[54] Eric M Rains. Quantum codes of minimum distance two. *IEEE Transactions on Information theory*, 45(1):266–271, 1999.

[55] Eric M Rains, RH Hardin, Peter W Shor, and NJA Sloane. A nonadditive quantum code. *Physical Review Letters*, 79(5):953, 1997.

[56] R. Raussendorf, S. Bravyi, and J. Harrington. Long-range quantum entanglement in noisy cluster states. *Physical Review A*, 71(6):062313, June 2005.

[57] Eleanor Rieffel and Wolfgang Polak. An introduction to quantum computing for non-physicists. *ACM Computing Surveys*, 32(3):300–335, 2000.

[58] Mary Beth Ruskai, Stanislaw Szarek, and Elisabeth Werner. An analysis of completely-positive trace-preserving maps on m2. *Linear Algebra and its Applications*, 347(1-3):159–187, 2002.

[59] Pradeep Kiran Sarvepalli, Andreas Klappenecker, and Martin Rötteler. Asymmetric quantum LDPC codes. In *Proceedings 2008 IEEE International Symposium on Information Theory (ISIT 2008)*, pages 305–309, 2008.

[60] Pradeep Kiran Sarvepalli, Andreas Klappenecker, and Martin Rötteler. Asymmetric quantum codes: constructions, bounds and performance. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 465, pages 1645–1672. The Royal Society, 2009.

[61] Erwin Schrödinger. An undulatory theory of the mechanics of atoms and molecules. *Physical Review*, 28(6):1049, 1926.

[62] Benjamin Schumacher and Michael A Nielsen. Quantum data processing and error correction. *Physical Review A*, 54(4):2629, 1996.

[63] C. E. Shannon. A mathematical theory of communication. *Bell Labs Technical Journal*, 27:379–423, 1948.

[64] Peter W Shor. The quantum channel capacity and coherent information. In *lecture notes, MSRI Workshop on Quantum Computation*, 2002.

[65] Peter W. Shor, Graeme Smith, John Smolin, and Bei Zeng. High performance single-error-correcting quantum codes for amplitude damping. *IEEE Transactions on Information Theory*, 57(10):7180–7188, 2011.

[66] Peter W Shor and John A Smolin. Quantum error-correcting codes need not completely reveal the error syndrome. *arXiv preprint quant-ph/9604006*, 1996.

[67] Graeme Smith. Quantum channel capacities. In *Information Theory Workshop (ITW), 2010 IEEE*, pages 1–5. IEEE, 2010.

[68] Graeme Smith, Joseph M Renes, and John A Smolin. Structured codes improve the Bennett-Brassard-84 quantum key rate. *Physical Review Letters*, 100(17):170502, 2008.

[69] Graeme Smith and John A Smolin. Degenerate quantum codes for Pauli channels. *Physical review letters*, 98(3):030501, 2007.

[70] Graeme Smith and John A Smolin. Additive extensions of a quantum channel.

In *Information Theory Workshop, 2008. ITW'08. IEEE*, pages 368–372. IEEE, 2008.

[71] A. Steane. Multiple particle interference and quantum error correction. *Proceedings of the Royal Society of London, Series A*, 452(1954):2551–2577, 1996.

[72] Yun-Jiang Wang, Bei Zeng, Markus Grassl, and Barry C Sanders. Stabilizer formalism for generalized concatenated quantum codes. In *Information Theory Proceedings (ISIT), 2013 IEEE International Symposium on*, pages 529–533. IEEE, 2013.

[73] Mark M Wilde. *Quantum information theory.* Cambridge University Press, 2013.

[74] Michael M Wolf and David Perez-Garcia. Quantum capacities of channels with small environment. *Physical Review A*, 75(1):012303, 2007.

[75] William K Wootters and Wojciech H Zurek. A single quantum cannot be cloned. *Nature*, 299(5886):802–803, 1982.

[76] Jingen Xiang, Cong Guo, and Ashraf Aboulnaga. Scalable maximum clique computation using MapReduce. In *Proceedings 2013 IEEE 29th International Conference on Data Engineering (ICDE)*, pages 74–85, 2013.

[77] Sixia Yu, Qing Chen, CH Lai, and CH Oh. Nonadditive quantum error-correcting code. *Physical Review Letters*, 101(9):090501, 2008.